



**HAL**  
open science

**Caractérisation des interactions  
glycosaminoglycannes/protéines dans le but de  
développer des molécules d'intérêt thérapeutique :  
exemples de l'Endocan et de l'IFNg**

Stéphane Sarrazin

► **To cite this version:**

Stéphane Sarrazin. Caractérisation des interactions glycosaminoglycannes/protéines dans le but de développer des molécules d'intérêt thérapeutique : exemples de l'Endocan et de l'IFNg. Sciences du Vivant [q-bio]. Université de Grenoble, 2007. Français. NNT : . tel-01063116

**HAL Id: tel-01063116**

**<https://theses.hal.science/tel-01063116>**

Submitted on 11 Sep 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## **THÈSE**

Pour l'obtention du grade de

**DOCTEUR DE L'UNIVERSITE JOSEPH FOURIER GRENOBLE I**

Discipline : Biologie Cellulaire et Moléculaire

Présentée et soutenue publiquement par

**Stéphane SARRAZIN**

Le 28 Juin 2007

**Caractérisation des interactions  
glycosaminoglycannes/protéines dans le but de  
développer des molécules d'intérêt thérapeutique :  
exemples de l'Endocan et de l'IFN $\gamma$**

Directeurs de thèse : Dr. Maryse DELEHEDDE et Dr. Hugues LORTAT-JACOB

**Membres du jury :**

Rapporteurs	Pr. Sylvie FOURNEL (Univ. Louis-Pasteur, Strasbourg) Pr. Fabrice ALLAIN (Univ. des Sciences et Technologies, Lille)
Président	Pr. Emmanuel DROUET (Univ. Joseph Fourier, Grenoble)
Examineurs	Dr. Philippe LASSALLE (Institut Pasteur, Lille) Dr. Maryse DELEHEDDE (Endotis Pharma, Loos) Dr. Hugues LORTAT-JACOB (Inst. Biologie Structurale, Grenoble)



# Remerciements

Je tiens en premier lieu à vous remercier, Hugues et Maryse, pour avoir co-dirigé ce travail de thèse. Partager la tâche n'était pas forcément évident, mais vous avez su l'accomplir à merveille, du début jusqu'à la fin.

Hugues, je dois te remercier tout d'abord pour la confiance que tu m'as accordée durant ces trois ans et demi (DEA inclus). Cela est un des éléments les plus importants, tu m'as certes guidé mais toujours laissé libre de mener à bien les projets comme je le souhaitais. A côté de cela, malgré ton emploi du temps de plus en plus chargé, tu as été constamment disponible, à l'écoute, tu as pris le temps de m'indiquer ton point de vue, de poser des critiques constructives sur les expériences, de me montrer que dans un laboratoire de recherche aussi "quand on veut on peut", quand on croit en quelque chose il faut se donner les moyens d'y arriver. A nouveau merci de me faire confiance dans l'année qui arrive, nous allons faire en sorte que les connaissances sur ta protéine fétiche s'enrichissent.

Maryse, je ne sais pas comment te remercier. Dès les premiers jours de l'aventure avec Endotis tu m'as fait entièrement confiance, tu as m'accordé la place d'un salarié Endotis et ce, même si 700 km me séparaient de Lille. Tu as su, par téléphone ou lors de quelques déplacements à l'IBS, encadrer les travaux qui ont été menés pour que ceux-ci satisfassent l'entreprise tout en m'offrant la marge de liberté qu'un laboratoire académique permet. Tu t'es démenée pour que je puisse participer à des meetings internationaux afin que ma formation doctorale soit la plus enrichissante possible. Tu m'as également permis d'aller travailler dans des laboratoires en Angleterre. Ta générosité, ta sincérité, ta gentillesse, et surtout la passion pour ce métier que tu véhicules me font te dire un énorme merci. Tu as été constamment disponible et à l'écoute et ce malgré toutes tes casquettes de chef de département, DRH, secrétaire, déléguée du personnel, assistante de R&D, responsable de la veille technologique, accueil des commerciaux, que tu sois à Lille, à Paris, dans le train, en RTT ou congés, à Helsinki, Cascais, en Suède ou à St Omer. Merci du fond du cœur.

Ensuite je souhaite remercier tous les membres du laboratoire "LEM" de l'IBS au sein duquel j'ai passé la plupart de mon temps :

Rabia, je te remercie pour ta gentillesse, ta serviabilité, tes conseils, ton humour, tes délicieux petits gâteaux, merci pour apporter tous les jours sans exception ton ravissant sourire au laboratoire.

Pascal, je te remercie pour ton humour léger, merci pour tes conseils, ta disponibilité, ta gentillesse et la fameuse tartiflette aussi divine que Collonges-la-Rouge peut être belle.

Mister Romain, qui à toi seul est une bande de jeunes, je te remercie pour arriver à la fois à partager l'humour léger de Pascal et conserver ton sérieux lorsqu'il est question de GAG. Merci pour apporter chaque jour au labo la chanson du loubard, capable à la fois de retranscrire le blues de la porte d'Orléans, les aventures de Jojo le démagog ou les derniers potins de la bande à Lucien.

Elodie, je suis vraiment très heureux d'avoir pu partager cette aventure de thésard avec toi. Je te remercie pour la bonne humeur et la joie de vivre que tu amènes avec toi chaque matin. Je

n'ai pas du te rendre la vie de labo simple tous les jours mais à aucun moment tu n'as fait preuve de susceptibilité, tu as toujours fait en sorte que la bombonne verte ne soit pas pleine... et t'es même allée jusqu'à me rendre service, notamment pour la mutagenèse. Je te dis un grand merci aussi pour ta disponibilité, ton soutien et ton aide. Tu as toutes les qualités pour être une leadeuse de boîte à outils.

Cédric, même si nous avons cohabité moins longtemps, dès ton arrivée tu t'es rendu disponible, tu as pris du temps pour réfléchir sur d'autres sujets que le tien, tu m'as rendu service et t'as réussi à adopter la GAG's attitude, pour un gars du 92 je trouve que ça fait déjà beaucoup et je t'en remercie !!

Aussi, j'aimerais dire un énorme merci à ma chère Pierrette, comme je te l'ai déjà dit tu es pour moi un exemple de vie à part entière, en te voyant, on se dit constamment "vivement que je vieillisse un peu que je retrouve une pêche d'enfer." Ton dynamisme et ta bonne humeur m'ont donné l'impression de passer la plupart de mon temps de labo dans un jardin de fleurs et d'oiseaux aux sifflements aigus !!

Sylvie, je te remercie pour ton dynamisme, ta force de travail, et pour ton professionnalisme, tu as été un exemple très motivant. Avant que tu rejoignes Lyon tu étais le métronome de notre équipe, je te remercie pour les conseils que tu m'as donnés en biacore, tu avais raison, l'immobilisation de PGs n'est pas forcément facile...

Pour ne pas faire trop long non plus, j'aimerais remercier d'une manière générale tout le LEM de Gérard. Merci Isabelle, Monique, Nicole, Pascale, Evelyne, Véronique, Florence, Hélène, Philippe, Marie-Claire, Hélène, Thomas, Benjamin et bien sûr Gérard. Vous avez tous été très chaleureux, très accueillants, toujours disponibles et arrangeants, c'était un réel plaisir de travailler à vos côtés, et je suis très heureux de pouvoir partager une année supplémentaire avec vous.

Merci également à Jean-Pierre en particulier pour toutes les analyses protéiques, tu as beau me répondre que tu fais juste ton boulot, je te dis quand même merci parce que ta rigueur de travail et les services que tu m'as rendus m'ont beaucoup aidés.

Je tiens également à remercier Bernard Dublet qui a toujours fait de son mieux pour gérer à la fois ses priorités et nos échantillons, ton aide en spectrométrie de masse a été très précieuse et très informative.

Au sein de l'IBS j'aimerais également remercier Franck Fieschi qui me suit depuis mes débuts à l'université et qui s'est toujours rendu disponible et de bons conseils.

Merci aussi à Laurence Serre et à Delphine Blot qui m'ont beaucoup aidé pour les expériences de cristallographie.

Toujours au sein de l'institut, j'aimerais remercier Alain, Sylvie et Linda, qui savent se faire discrets mais qui font un travail dont on mesure ne l'importance que lorsqu'ils ne sont pas là où lorsque l'on doit le faire nous même au sein d'autres laboratoires.

En tant qu'ingénieur de recherche au sein d'Endotis Pharma, j'aimerais remercier M. David Béchard pour qui j'ai beaucoup d'estime, son travail a été colossal, son dynamisme et sa force très motivante et sa confiance appréciable.

Merci aussi à Stéphane Degove, co-fondateur de l'entreprise qui a toujours été très sympathique à mon égard.

J'aimerais également remercier les personnes avec qui j'ai travaillé étroitement pour l'aboutissement de ce travail au sein du site d'Endotis Pharma de Lille. Estelle et Christelle merci pour tout, vous m'avez toujours accueilli avec beaucoup d'enthousiasme, avec malgré tout, à chaque fois, un sourire très communicatif. Même si vous n'êtes pas tout à fait du nord, vous m'avez permis de vérifier que "les gens du nord ont dans le cœur le soleil qu'ils n'ont pas dehors".

Merci également à Romain, Justine, Ouarda, Isabelle, Heidi, Guy et Loïc avec qui j'ai pu un peu mieux faire connaissance ces derniers mois.

Au sein d'Endotis (site de Romainville) j'aimerais remercier l'équipe des chimistes qui m'a très bien accueilli et dont le soutien a été fort appréciable en cette fin de rédaction. Merci aux membres d'EP4000, Cécile la sportive et Olivier from UK. Merci à François, Julien, Francisco, Erwan le roi des proverbes, Sandrine, Olivier Q, Sabrina et Fanny les reines des fous rires, Pascal dit Grabit, Audrey la parigote au rire génial, et bien sûr le plus passionné et le plus drôle de tous M. Seb Simpson.

Merci à Endotis Pharma pour avoir co-financé ce travail de thèse en partenariat avec l'ANRT. La convention Cifre est une façon très enrichissante de réaliser une thèse.

J'aimerais également remercier l'ANRT pour m'avoir permis de travailler au sein du PICR (Paterson Institute for Cancer Research) à Manchester. A cette occasion je remercie John Gallagher de m'avoir accueilli dans son laboratoire. Je tiens également à remercier Jon dit "the butcher" qui m'a guidé sur toutes les expériences durant ce stage d'un mois, qui a été constamment disponible et serviable. Et puis bien entendu, je voudrais adresser un énorme merci à Malcolm. Tu m'as encadré durant ce stage, tu as pris du temps pour faire en sorte que tout se passe bien au laboratoire, et surtout je tiens à t'adresser un énorme merci parce que tu as pris du temps en dehors du laboratoire pour qu'on puisse continuer à discuter dans un pub ou ailleurs, tu m'as fait visiter York, une ville qui te tient à cœur. Tu as été d'une grande générosité, d'une gentillesse débordante, avec à chaque instant un humour anglais qui te donne tout ton charme. J'aurais sincèrement aimé que tu puisses faire partie du jury mais pour certaines raisons cela n'a pas pu être possible.

Je tiens à remercier Dave Fernig pour m'avoir accueilli au sein de son laboratoire à l'université de Liverpool. Ton enthousiasme, ta disponibilité, les discussions scientifiques que nous avons pu avoir, ton extravagance et ta passion pour ton métier (qui ne te quitte même pas lorsque tu prépares de délicieux spaghettis) sont autant de qualités qui ont rendu mon séjour agréable.

Je souhaite remercier le Professeur David Bonnaffé du Laboratoire de Chimie Organique Multifonctionnelle de l'université de Paris Sud, qui a fait avec son équipe et en particulier avec M. André Lubineau un travail remarquable concernant la synthèse des mimétiques d'HS.

Je voudrais remercier le laboratoire de Benito Casu à Milan et plus particulièrement Marco Guerrini pour avoir accepté de réaliser l'analyse par RMN de la chaîne glycanique d'endocan.

Aussi, j'aimerais remercier Marie-Christine Slomianny du laboratoire "Unité de Glycobiologie Structurale et Fonctionnelle" de l'université des sciences et technologies à Lille pour sa sympathie et son enthousiasme vis-à-vis de la caractérisation biochimique d'endocan et de l'analyse des glycosylations.

Ensuite, je tiens à remercier l'ESH pour m'avoir accordé une bourse Marie-Curie me permettant d'assister à leur meeting à Sitges, et plus particulièrement Ghyslaine et Cécile pour leur sympathie, leur bonne humeur et leur sourire.

Enfin j'aimerais dire un grand merci au leader en tout genre, le number one, celui qui arrive au labo entre 7h00 et 7h02, avec qui j'ai pris énormément de plaisir à travailler (et il me comprend parce que lui aussi il aime travailler avec lui-même !!). On ne pourra pas citer toutes ses qualités mais la plus grande c'est...Marie...sa femme depuis peu. Merci à vous deux pour m'avoir accueilli et surtout supporté pratiquement une semaine sur deux en temps que colloc à Lille pendant quelques temps.

Je souhaite dire à tous les membres du jury que je suis très heureux et très honoré qu'ils aient accepté de juger ce travail.

Madame la Professeure Sylvie Fournel, vous m'avez fait l'honneur d'accepter de juger ce travail en qualité de Rapporteur. Soyez assurée de ma sincère gratitude.

Monsieur le Professeur Fabrice Allain, vous m'avez également fait l'honneur d'être rapporteur de cette thèse. Veuillez trouver ici l'expression de mes sincères remerciements pour l'intérêt que vous portez à ce travail et soyez assuré de ma profonde reconnaissance.

Monsieur le Professeur Emmanuel Drouet, vous m'avez fait l'honneur de juger ce travail en tant qu'examinateur et représentant de l'Université Joseph Fourier. Je vous en remercie. Veuillez trouver ici l'expression de ma plus vive reconnaissance.

Monsieur le Docteur Philippe Lassalle, c'est un réel plaisir de te compter parmi les membres du jury. A chacune de nos rencontres tu as montré un enthousiasme débordant, chacune de nos discussions ont été enrichissantes et au combien motivantes, j'ai hâte de pouvoir discuter à nouveau à l'occasion de la soutenance. Merci pour tout, trouve ici l'expression de mon respect le plus sincère.

De manière plus personnelle, j'aimerais remercier tout ceux avec qui il est bon d'oublier les activités professionnelles, ceux qui sont capables en un entraînement, ou un week-end de donner l'impression d'avoir passé au moins une semaine de vacances. Merci au Mantaille Sportif dans sa globalité, à toute mon équipe en particulier d'être aussi débile, merci aussi aux membres du front de libération de Mantaille !! Merci aussi à Stéphanie d'être toujours disponible pour papoter, d'avoir toujours la dernière blague à la mode à faire suivre.

Je pense bien entendu à ma famille qui est toujours là pour moi, merci du fond du cœur pour tout. Papy, cette thèse est certainement quelque peu inspirée du château de Barbe-Bleue et du roi Boson !

Je ne peux pas oublier ma ptite soeur qui est toujours là pour moi. Merci pour tout Soeurette adorée.

Je leur dois tout, ils m'ont toujours encouragé à faire ce qui me faisait plaisir, ils m'ont toujours supporté, et soutenu. Sans eux je n'aurais jamais pu envisager d'effectuer un tel travail. Ils sont exceptionnels et parfaits, je veux bien entendu parler de mes parents. Je vous remercie pour tout l'amour que vous me témoignez, il est réciproque.

Et finalement, la meilleure est pour la fin et elle s'appelle Ninette. Ma chérie je te remercie pour l'amour que tu me donnes au quotidien. Je ne pourrais pas citer toutes les qualités qui te caractérisent, il faudrait se relancer dans une thèse. Un simple merci ne suffirait donc pas alors juste un subtil : je t'aime.

# **TABLE DES MATIERES**

<b>INDEX DES FIGURES.....</b>	<b>14</b>
<i>FIGURES DE LA REVUE BIBLIOGRAPHIQUE.....</i>	<i>14</i>
<i>FIGURES DU TRAVAIL EXPERIMENTAL .....</i>	<i>15</i>
<b>ABREVIATIONS .....</b>	<b>17</b>
<b>INTRODUCTION .....</b>	<b>20</b>
<b>RAPPORT BIBLIOGRAPHIQUE .....</b>	<b>22</b>
<b>CHAPITRE I :</b>	
<b>LES PROTEOGLYCANNES (PGs) .....</b>	<b>23</b>
1. <b>GENERALITES .....</b>	<b>24</b>
2. <b>LES GLYCOSAMINOGLYCANNES (GAGs).....</b>	<b>24</b>
2.1. Généralités .....	24
2.2. Les familles de glycosaminoglycannes (GAGs).....	26
2.2.1. L'acide hyaluronique (AH).....	26
2.2.2. Les kératanes sulfates (KS).....	26
2.2.3. L'héparine et les héparanes sulfates (Hp et HS).....	28
2.2.4. Les chondroïtines sulfates (CS) .....	29
2.2.5. Le dermatanes sulfate (DS).....	30
2.3. Biosynthèse des glycosaminoglycannes .....	30
2.3.1. La biosynthèse des GAGs .....	30
2.3.1.1. La biosynthèse de l'AH, le seul GAG à n'être rattaché à aucun core protéique.....	31
2.3.1.2. La biosynthèse des GAGs liés à un core protéique.....	31
2.3.1.2.1. La biosynthèse des KS .....	33
2.3.1.2.2. La biosynthèse des Hp, HS, CS et DS.....	35
2.4. La dégradation des GAGs .....	41
2.4.1. Introduction .....	41
2.4.2. La dégradation des GAGs comme un moyen supplémentaire de régulation .....	41
2.4.3. Le recyclage ou turnover des GAGs .....	42
2.4.4. Les enzymes de dégradation utilisées au laboratoire .....	44
3. <b>LES CORES PROTEIQUES .....</b>	<b>46</b>
3.1. Introduction .....	46
3.2. Les protéoglycannes intracellulaires .....	46
3.3. Les protéoglycannes membranaires .....	48
3.3.1. Le CD44.....	48
3.3.2. Le bétaglycan .....	49
3.3.3. Les glypicans .....	51
3.3.4. Les syndécans.....	53
3.4. Les protéoglycannes extracellulaires .....	57
3.4.1. Les protéoglycannes de la membrane basale.....	57
3.4.1.1. Les collagènes XVIII et XV .....	57
3.4.1.2. L'agrine .....	57
3.4.1.3. Le bamacan .....	58
3.4.1.4. La famille du perlécan .....	60
3.4.2. La famille des hyalectans .....	62
3.4.2.1. Le versican .....	62
3.4.2.2. L'aggrécan .....	64
3.4.2.3. Le neurocan .....	66
3.4.2.4. Le brévican .....	66
3.4.3. La famille de protéoglycannes riches en leucine (SLRP).....	68
3.4.3.1. Introduction.....	68
3.4.3.2. Le biglycan .....	70
3.4.3.3. La décorine .....	71
3.4.3.4. Le lumican .....	71
3.4.3.5. La fibromoduline .....	72
3.4.4. Les protéoglycannes circulants.....	73
3.4.4.1. Le PG-100 ou MCSF1 .....	73
3.4.4.2. La bikunine .....	75

3.4.4.3. L'apolipoprotéine O .....	79
3.4.4.4. L'endocan .....	79

## **CHAPITRE II :**

### **L'ENDOCAN, UN PROTEOGLYCANNE ATYPIQUE ET MYSTERIEUX.. 80**

<b>1. INTRODUCTION .....</b>	<b>82</b>
<b>2. DU GENE AU PROTEOGLYCANNE .....</b>	<b>82</b>
2.1. Organisation génomique.....	82
2.2. Contrôle transcriptionnel d'endocan.....	84
2.3. Le core protéique d'endocan .....	86
2.3.1. Caractéristiques biochimiques .....	86
2.3.2. Expression de l'endocan.....	87
2.4. L'endocan est un protéoglycane à dermatane sulfate.....	88
2.5. L'endocan ou ESM-1 (pour endothelial cell specific molecule) est il réellement un DSPG spécifique des cellules endothéliales ? .....	89
<b>3. FONCTIONS D'ENDOCAN.....</b>	<b>90</b>
3.1. Rôle dans l'inflammation.....	90
3.2. Endocan dans la progression tumorale .....	93
<b>4. CONCLUSION.....</b>	<b>94</b>

## **CHAPITRE III :**

### **LES INTERACTIONS PROTEINES-GAGS ET LEUR POTENTIEL DANS LE DEVELOPPEMENT DE MOLECULES THERAPEUTIQUES..... 96**

<b>1. LES FONCTIONS DES GAGS VIA LEURS INTERACTIONS AVEC D'AUTRES MOLECULES.....</b>	<b>97</b>
1.1. Introduction .....	97
1.2. Les caractéristiques biochimiques et structurales d'une interaction GAG-protéine.....	97
1.3. Les fonctions associées aux interactions GAG-protéines.....	101
1.3.1. Les GAGs peuvent induire un changement conformationnel des protéines.....	103
1.3.2. Les GAGs participent aux interactions protéines-protéines .....	103
1.3.3. Les GAGs comme co-récepteurs .....	105
1.3.4. Les GAGs peuvent servir de récepteur d'internalisation.....	107
1.3.5. Les GAGs participent à l'organisation de la matrice et à l'adhésion cellulaire .....	107
1.3.6. Les GAGs permettent le stockage et le relargage de protéines.....	108
1.3.7. Les GAGs jouent le rôle de récepteurs d'attachement pour les pathogènes.....	109
1.3.8. Les GAGs localisent, concentrent et permettent les gradients de médiateurs.....	110
1.3.9. Les GAGs protègent certaines protéines vis-à-vis de la protéolyse .....	111
1.4. L'implication des GAGs dans l'inflammation .....	113
1.4.1. Les GAGs et le recrutement des cellules mobilisées lors de l'inflammation .....	113
1.4.2. Les GAGs et les médiateurs de l'inflammation .....	114
1.5. L'implication des GAGs dans les cancers .....	116
1.5.1. La croissance et la prolifération des cellules cancéreuses .....	116
1.5.2. Les phénomènes d'invasion et de métastase des cellules cancéreuses .....	117
1.5.3. Les GAGs et l'angiogenèse.....	118
1.6. Les fonctions des Dermatanes Sulfates (DS) en particulier.....	118
1.6.1. Les DS ont une activité anti-coagulante.....	119
1.6.2. Les DS et la matrice extracellulaire .....	119
1.6.3. Les DS et les facteurs de croissance .....	120
1.6.4. Les DS et les cytokines ou chimiokines .....	120
1.6.5. Les DS et les pathogènes .....	121
1.6.6. Les DS dans la cicatrisation.....	121
<b>2. APPLICATIONS THERAPEUTIQUES DES GAGS .....</b>	<b>122</b>
2.1. L'UTILISATION THERAPEUTIQUE DE GAGS NATURELS.....	122
2.1.1. Utilisation de l'acide hyaluronique .....	122
2.1.2. Les CS pour soulager l'ostéoarthrose (OA).....	128
2.2. LES MIMETIQUES DES GAGs .....	129
2.2.1. Le pouvoir anti-coagulant de l'héparine .....	129
2.2.1.1. L'héparine rencontre l'ATIII : une histoire fluide et "sang accroc" .....	129
2.2.1.2. Le premier mimétique de l'héparine est synthétisé.....	132

2.2.1.3.	La thrombine serait-elle jalouse du facteur Xa? .....	134
2.2.1.4.	La troisième génération d'anticoagulants dérivés d'héparine .....	136
2.2.2.	PI-88 : un mimétique spécifique des interactions peu spécifiques ? .....	138
2.2.3.	Les GAGs et l'amylose .....	140
2.3.	<b>LE POTENTIEL THERAPEUTIQUE DES MIMETIQUES DES GAGS : Les sucres peuvent-ils soigner tous les maux ? .....</b>	<b>142</b>
2.3.1.	Une "pseudo-longueur" d'avance dans le développement thérapeutique des mimétiques des GAGs .....	142
2.3.2.	Les mimétiques des GAGs ont un potentiel thérapeutique très vaste .....	144
2.3.2.1.	Les mimétiques et la malaria .....	144
2.3.2.2.	Les mimétiques et le VIH .....	144
2.3.2.3.	Les mimétiques comme traitement anti-cancéreux ? .....	146
2.3.2.4.	Les mimétiques comme traitement anti-inflammatoire ? .....	146
2.3.2.5.	Les mimétiques pour inhiber l'héparanase .....	147
2.3.2.6.	Des mimétiques pour la réparation tissulaire : Les RGTAs .....	148

## **CHAPITRE IV :**

### ***L'INTERFERON $\gamma$ , UN LIGAND DES GLYCOSAMINOGLYCANNES ... 150***

1.	<b><i>L'INTERFERON GAMMA (IFN<math>\gamma</math>) .....</i></b>	<b><i>151</i></b>
1.1.	<b>GENERALITES .....</b>	<b>151</b>
1.1.1.	Historique .....	151
1.1.2.	Du gène à la protéine .....	151
1.1.3.	Les sources cellulaires .....	153
1.1.4.	L'expression d'IFN $\gamma$ .....	154
1.1.4.1.	La voie de synthèse .....	154
1.1.4.2.	La régulation de l'expression .....	154
1.1.5.	Principales activités biologiques .....	156
1.1.5.1.	Principales activités observées <i>in vitro</i> .....	156
1.1.5.2.	Principales activités observées <i>in vivo</i> .....	157
1.2.	<b>LES LIGANDS DE L'IFN<math>\gamma</math> .....</b>	<b>160</b>
1.2.1.	Le récepteur membranaire de signalisation de l'IFN $\gamma$ .....	160
1.2.1.1.	Le récepteur à IFN $\gamma$ : l'IFN $\gamma$ R .....	160
1.2.1.2.	L'interaction IFN $\gamma$ /IFN $\gamma$ R .....	160
1.2.1.3.	La transduction du signal de l'IFN $\gamma$ .....	162
1.2.2.	L'interaction IFN $\gamma$ / HS .....	162
1.2.2.1.	Caractéristiques de l'interaction .....	162
1.2.2.2.	Rôle de l'interaction .....	164
2.	<b><i>LES TROUBLES PHYSIOLOGIQUES LIES À L'IFN<math>\gamma</math> .....</i></b>	<b><i>165</i></b>
2.1.	<b>EXEMPLES DE MALADIES ASSOCIEES A L'IFN<math>\gamma</math> .....</b>	<b>165</b>
2.2.	<b>EXEMPLES DE TRAITEMENTS POUR MODULER L'ACTIVITE D'IFN<math>\gamma</math> 165</b>	

### **OBJECTIFS DU TRAVAIL..... 167**

✓ ***ETUDE STRUCTURALE ET FONCTIONNELLE D'ENDOCAN ..... 168***

✓ ***ETUDE STRUCTURALE ET FONCTIONNELLE DE L'INTERACTION IFN $\gamma$ /HS : DEVELOPPEMENT D'UN MIME OLIGOSACCHARIDIQUE MODULANT L'ACTIVITE D'IFN $\gamma$  ..... 169***

### **ETUDE STRUCTURALE ET FONCTIONNELLE D'ENDOCAN ..... 172**

1.	<b><i>MISE EN PLACE D'UNE PLATEFORME DE PRODUCTION ET PURIFICATION D'ENDOCAN, DU CORE PROTEIQUE D'ENDOCAN ET DES CHAINES DS D'ENDOCAN .....</i></b>	<b><i>173</i></b>
1.1.	<b>PRODUCTION ET PURIFICATION D'ENDOCAN .....</b>	<b>174</b>
1.1.1.	Matériel et méthodes .....	174
1.1.2.	Résultats et discussion .....	177
1.2.	<b>PRODUCTION ET PURIFICATION DU CORE PROTEIQUE D'ENDOCAN ... .....</b>	<b>183</b>
1.2.1.	Matériel et méthodes .....	183

1.2.2.	Résultats et discussion .....	185
1.3.	<b>PURIFICATION DE LA CHAÎNE DS D'ENDOCAN.....</b>	<b>187</b>
1.3.1.	Matériel et méthodes .....	187
1.3.2.	Résultats et discussion .....	189
1.4.	<b>CONCLUSION.....</b>	<b>189</b>
2.	<b>CARACTERISATION STRUCTURALE ET BIOCHIMIQUE D'ENDOCAN. 191</b>	
2.1.	<b>CARACTERISATION DE LA CHAÎNE DE DERMATANE SULFATE .....</b>	<b>192</b>
2.1.1.	Introduction .....	192
2.1.2.	Etude comparative de la composition oligosaccharidique de la chaîne d'endocan provenant de cellules HUVEC et de cellules HEK293 .....	193
2.1.2.1.	Matériel et méthodes .....	193
2.1.2.2.	Résultats .....	196
2.1.3.	Analyse de la chaîne oligosaccharidique d'endocan par RMN .....	204
2.1.3.1.	Matériel et méthodes .....	204
2.1.3.2.	Résultats .....	204
2.1.4.	Discussion .....	206
2.1.5.	Perspectives .....	212
2.2.	<b>CARACTERISATION DU CORE PROTEIQUE D'ENDOCAN.....</b>	<b>215</b>
2.2.1.	Modélisation .....	215
2.2.2.	Recherche d'autres glycosylations que la chaîne de DS .....	219
2.2.2.1.	Matériel et méthodes .....	221
2.2.2.2.	Résultats .....	226
2.2.3.	Cristallographie .....	234
2.2.3.1.	Matériel et méthodes .....	234
2.2.3.2.	Résultats .....	235
2.2.4.	Discussion .....	238
2.2.5.	Perspectives .....	243
3.	<b>INTERACTIONS ENTRE ENDOCAN ET DES PROTEINES D'INTERETS 244</b>	
3.1.	<b>ETUDE D'INTERACTION ENTRE ENDOCAN ET DES ANTICORPS.....</b>	<b>245</b>
3.1.1.	Utilisation du système à résonance plasmonique de surface en cuvette : le biosensor IAsys.....	246
3.1.1.1.	Matériel et méthodes .....	246
3.1.1.2.	Résultats .....	248
3.1.1.3.	Discussion .....	250
3.1.2.	Mesure d'affinité entre des anticorps et l'endocan en temps réel et en flux continu : utilisation du Biacore .....	252
3.1.2.1.	Matériel et méthodes .....	252
3.1.2.2.	Résultats .....	254
3.1.2.3.	Discussion .....	255
3.2.	<b>ETUDE D'INTERACTIONS ENTRE ENDOCAN ET DES MÉDIATEURS BIOLOGIQUES .....</b>	<b>255</b>
3.2.1.	Introduction .....	255
3.2.2.	Matériel et méthodes .....	256
3.2.3.	Résultats .....	258
3.2.4.	Discussion .....	262
3.2.5.	Perspectives .....	266
4.	<b>ETUDE FONCTIONNELLE D'ENDOCAN .....</b>	<b>267</b>
4.1.	<b>INTRODUCTION.....</b>	<b>268</b>
4.2.	<b>CRIBLAGE DE LIGNEES CELLULAIRES EXPRIMANT OU NON ENDOCAN.....</b>	<b>269</b>
4.2.1.	Matériel et méthodes .....	269
4.2.2.	Résultats et discussion .....	271
4.3.	<b>EFFET <i>IN VITRO</i> D'ENDOCAN SUR LA CYTOKINE IFN<math>\gamma</math> .....</b>	<b>273</b>
4.3.1.	Matériel et méthodes .....	273
4.3.2.	Résultats .....	275
4.3.3.	Discussion .....	277
5.	<b>CONCLUSION ET PERSPECTIVES .....</b>	<b>280</b>

## **ETUDE STRUCTURALE ET FONCTIONNELLE DE L'INTERACTION IFN $\gamma$ /HS : DEVELOPPEMENT D'UN MIMÉ**

# OLIGOSACCHARIDIQUE MODULANT L'ACTIVITE D'IFN $\gamma$

.....290

1. <i>2O<sub>10</sub>: UN MIMETIQUE DES CHAINES HEPARANES SULFATES POUR MODULER L'ACTIVITE D'IFN<math>\gamma</math></i> .....	291
2. <i>ETUDE DE LA REGION C-TERMINALE D'IFN<math>\gamma</math></i> .....	299
2.1. <i>Introduction</i> .....	301
2.2. <i>Etude des deux domaines basiques C-terminaux d'IFN<math>\gamma</math></i> .....	303
2.2.1. <i>Matériel et méthodes</i> .....	303
2.2.2. <i>Premiers résultats et perspectives</i> .....	305
2.3. <i>Etude structurale de la région C-terminale d'IFN<math>\gamma</math></i> .....	307
2.3.1. <i>Résultats et discussion</i> .....	307
3. <i>CONCLUSION ET PERSPECTIVES</i> .....	311

## BIBLIOGRAPHIE .....315

## ANNEXES .....351

<i>PUBLICATIONS</i> .....	352
<i>TRAVAUX PRESENTES EN MEETINGS (Posters)</i> .....	353

## INDEX DES FIGURES

### Figures de la revue bibliographique

	Pages
<u>Figure 1: Motifs disaccharidiques des différents glycosaminoglycanes et leurs modifications</u>	18
<u>Figure 2: Diversité et organisation structurale des Héparanes Sulfates</u>	19
<u>Figure 3: Représentation du profil structural des KS I (A) et KS II (B)</u>	23
<u>Figure 4: Biosynthèse des chaînes oligosaccharidiques d'héparine / héparanes sulfates (Hp/HS) et chondroïtines / dermatanes sulfates (CS/DS)</u>	24
<u>Figure 5: Enzymes utilisées couramment en laboratoire pour couper les GAGs</u>	32
<u>Figure 6: Classification des protéoglycanes (PG) en fonction de leur localisation</u>	33
<u>Figure 7: Les transcripts et protéines issus de l'épissage alternatif du CD44</u>	34
<u>Figure 8: La famille des glypicans</u>	36
<u>Figure 9: Domaines fonctionnels des syndécans</u>	37
<u>Figure 9bis: Les syndécans et les glypicans sont les deux grandes familles d'HSPG de la surface cellulaire</u>	38
<u>Figure 10: Domaines structuraux du collagène XVIII et de son domaine NCI/endostatine</u>	39
<u>Figure 11: Domaines structuraux du perlécan humain</u>	41
<u>Figure 12: Représentation schématique des quatre membres de la famille des hyalectans</u>	42
<u>Figure 13: L'association de l'aggrécan et de l'acide hyaluronique</u>	43
<u>Figure 14: Tableau de classification des SLRPs</u>	44
<u>Figure 15a: Organisation moléculaire de certaines SLRPs</u>	45
<u>Figure 15b: Superposition des structures de la décorine et du biglycan</u>	46
<u>Figure 16 a: Représentation schématique de la structure de la bikunine</u>	50
<u>Figure 16 b: Le pré-<math>\alpha</math>-inhibiteur et l'inter-<math>\alpha</math>-inhibiteur</u>	51
<u>Figure 16 c: Biosynthèse de la bikunine et de la protéine pre-a-inhibitor contenant la bikunine</u>	52
<u>Figure 17: Représentation schématique des caractéristiques d'endocan</u>	54
<u>Figure 18: Expression d'endocan dans des tumeurs humaines</u>	64
<u>Figure 19: Conformation et flexibilité de l'acide uronique de l'Hp et des HS</u>	69
<u>Figure 20: Le glycoalyx</u>	70
<u>Figure 21: Les fonctions des GAGs</u>	71
<u>Figure 21bis: Changement de conformation de l'ATIII crée par un pentasaccharide issu de l'héparine</u>	72
<u>Figure 22: Importance des HS pour le recrutement leucocytaire au cours de l'inflammation</u>	79
<u>Figure 23: Les stades de la progression tumorale</u>	81
<u>Figure 24: Les différentes préparations commerciales d'acide hyaluronique existantes (données de 2000)</u>	88
<u>Figure 25: Principaux essais cliniques publiés comparant l'utilisation de l'acide hyaluronique contre placebo dans la gonarthrose (données de 2000)</u>	89
<u>Figure 26: Tableau récapitulatif des différents essais cliniques concernant l'utilisation de chondroïtine sulfate dans l'ostéorathrose (OA)</u>	90
<u>Figure 27a: Structure de la séquence pentasaccharidique d'héparine qui interagit avec l'antithrombine III</u>	93
<u>Figure 27b: Mimétiques du pentasaccharide liant l'ATIII conçus par synthèse chimique</u>	93
<u>Figure 28a: Modèle moléculaire du complexe ternaire ATIII/Hp/thrombine</u>	94
<u>Figure 28b: Représentation schématique de la structure que doit adopter un mime d'Hp pour lier à la fois l'ATIII et la thrombine</u>	94
<u>Figure 28c: Développement de mimétiques de l'héparine liant l'ATIII et la thrombine</u>	95
<u>Figure 29A: Exemples d'inhibiteurs potentiels de la thrombine</u>	96
<u>Figure 29B: Mécanisme d'action d'un inhibiteur bivalent ciblant aussi bien l'ATIII que la thrombine de façon indépendante</u>	96

<b><u>Figure 30:</u></b> Le PI-88 est un mélange d'oses modifiés, issus de levure qui miment les HS et présentent des effets anti-tumoraux intéressants	97
<b><u>Figure 31:</u></b> Inhibiteurs des interactions GAGs-Protéines amyloïdes développés par Neurochem	98
<b><u>Figure 32:</u></b> Structure de l'homodimère d'IFN $\gamma$	108
<b><u>Figure 33:</u></b> Structure cristallographique du complexe entre l'IFN $\gamma$ et la partie soluble de la chaîne $\alpha$ du récepteur à IFN $\gamma$	114
<b><u>Figure 34:</u></b> Schéma de la voie de signalisation JAK/STAT par l'IFN $\gamma$	115
<b><u>Figure 35:</u></b> Modèle de l'interaction IFN $\gamma$ / HS	116

## Figures du travail expérimental

	Pages
<b><u>Figure 1:</u></b> Système de production en bioréacteur CELLLine 1000	130
<b><u>Figure 2:</u></b> Tableau comparatif des différents modes de production d'endocan	130
<b><u>Figure 3A:</u></b> Première étape de la purification d'endocan par chromatographie échangeuse d'anions	131
<b><u>Figure 3B:</u></b> La deuxième étape de la purification d'endocan est une chromatographie d'immunoaffinité	132
<b><u>Figure 4A:</u></b> Première étape de la purification du core protéique d'endocan par chromatographie d'échange d'anions	133
<b><u>Figure 4B:</u></b> La deuxième étape de la purification du core protéique d'endocan est une chromatographie d'immunoaffinité	134
<b><u>Figure 5:</u></b> Nature de la chaîne glycanique d'endocan	142
<b><u>Figure 6a:</u></b> Taille de la chaîne glycanique d'endocan	143
<b><u>Figure 6b:</u></b> Courbe de Wasteson	143
<b><u>Figure 7a:</u></b> Analyse disaccharidique de la chaîne glycanique d'endocan	144
<b><u>Figure 7b :</u></b> Histogramme rapportant l'analyse disaccharidique de la chaîne glycanique d'endocan	144
<b><u>Figure 8a:</u></b> Calcul du ratio CS/DS par une digestion à la Ch ACI	145
<b><u>Figure 8b:</u></b> Tableau récapitulatif des résultats de la figure 8a	145
<b><u>Figure 9:</u></b> Calcul du ratio CS/DS par une digestion à la Ch B	146
<b><u>Figure 10:</u></b> Localisation des IdoA le long de la chaîne glycanique d'endocan	146
<b><u>Figure 11:</u></b> Spectre RMN 1H et HSQC (heteronuclear single quantum coherence) de la chaîne DS d'endocan issu d'HEK293	147
<b><u>Figure 12:</u></b> Régions a) et b) du spectre HSQC de la figure 11	148
<b><u>Figure 12bis:</u></b> Ratio obtenu entre les acides iduroniques et glucuroniques grâce à l'analyse des spectres de la figure 12	148
<b><u>Figure 13:</u></b> Fraction du spectre 2D-NOESY de la chaîne DS d'endocan issue de cellules HEK293	149
<b><u>Figure 14:</u></b> Modèle représentant la composition de la chaîne de dermatane sulfate d'endocan	150
<b><u>Figure 15A:</u></b> Séquence protéique d'endocan	152
<b><u>Figure 15B:</u></b> Représentation schématique de l'organisation d'endocan	152
<b><u>Figure 16:</u></b> Prédiction du désordre dans la structure d'endocan	153
<b><u>Figure 17:</u></b> Le motif EGF-like	154
<b><u>Figure 18A:</u></b> Spectre MALDI-TOF du core protéique d'endocan purifié	155
<b><u>Figure 18B:</u></b> Gel SDS-PAGE coloré au bleu de Coomassie du core protéique d'endocan	155
<b><u>Figure 19:</u></b> Principe du marquage des glycannes par oxydation au périodate suivi de l'ajout d'une biotine hydrazide	156
<b><u>Figure 20:</u></b> Révélation par chemiluminescence de la présence de biotine sur les potentiels O- ou N-glycosylations d'endocan	158

<b><u>Figure 21:</u></b> Spectrométrie de masse de l'endocan mutant traité ou non par la O-glycosydase ou la N-glycosydase F	159
<b><u>Figure 22:</u></b> Tableau récapitulatif des fragments polypeptidiques d'endocan, théoriques ou expérimentaux, qui doivent être retrouvés après une digestion à la trypsine	160
<b><u>Figure 23:</u></b> Tableau récapitulatif des fragments polypeptidiques d'endocan, théoriques ou expérimentaux, qui doivent être retrouvés après une digestion à l'endoprotéinase Lys-C	161
<b><u>Figure 24:</u></b> Tableau récapitulatif des fragments polypeptidiques d'endocan, théoriques ou expérimentaux, qui doivent être retrouvés après une digestion à l'endoprotéinase Glu-C	161
<b><u>Figure 25:</u></b> Composition monosaccharidique des O- et N-glycannes liés sur le core protéique d'endocan	162
<b><u>Figure 26:</u></b> Suivi de la purification du Fab de l'anticorps EP214 par gel SDS-PAGE coloré au bleu de Coomassie	165
<b><u>Figure 27:</u></b> Chromatographie d'exclusion du complexe core protéique d'endocan/Fab d'EP214	165
<b><u>Figure 28:</u></b> Modèle de l'organisation structurale du core protéique d'endocan	168
<b><u>Figure 29a:</u></b> Mesure de la constante d'affinité de l'interaction entre l'endocan et un anticorps (EP208-003) par le système IAsys en utilisant la valeur du Bmax	173
<b><u>Figure 29b:</u></b> Tableau récapitulatif des Kd calculés par la méthode de Scatchard en IAsys pour les anticorps anti-endocan EP208 et EP221	173
<b><u>Figure 30:</u></b> Calcul du Kd par les constantes cinétiques kasso et kdisso	174
<b><u>Figure 31:</u></b> Schéma représentant le système mis en place dans la cuvette IAsys pour mesurer des interactions entre l'endocan et les différents anticorps monoclonaux anti-endocan	175
<b><u>Figure 32:</u></b> Mesure d'interactions entre endocan et les différents anticorps anti-endocan par Biacore	176
<b><u>Figure 33:</u></b> Utilisation de la méthode de GMSA pour visualiser l'interaction entre le HGF/SF et la chaîne DS d'endocan	179
<b><u>Figure 34:</u></b> Utilisation de la méthode de GMSA pour visualiser l'interaction entre la chaîne DS d'endocan produit par les HUVECs et les facteurs FGF-2, SDF-1 $\alpha$ et IFN $\gamma$	180
<b><u>Figure 35A:</u></b> Tableau récapitulatif du criblage de médiateurs testés dans l'interaction avec endocan, son core protéique ou sa chaîne de DS par Biacore	181
<b><u>Figure 35B:</u></b> Mesure d'interactions entre HGF/SF, IFN $\gamma$ et endocan et/ou le core protéique d'endocan par Biacore	182
<b><u>Figure 36:</u></b> Gels d'agarose des RT-PCR réalisées pour détecter la présence d'ARNm d'endocan dans les différentes lignées cellulaires	188
<b><u>Figure 37:</u></b> Tableau récapitulatif des différentes lignées cellulaires criblées dans le but de détecter l'expression d'ARNm ou de protéines d'endocan	189
<b><u>Figure 38:</u></b> Étude de la signalisation induite par IFN $\gamma$ et effet de l'endocan	190
<b><u>Figure 39:</u></b> Histogramme récapitulatif des expériences de FACS permettant de détecter les molécules HLA-DR et ICAM-1 à la surface des cellules THCP1 soumises à différentes stimulations	191
<b><u>Figure 40:</u></b> Schéma des domaines D2 et D1 de la région C-terminale d'IFN $\gamma$	214
<b><u>Figure 41:</u></b> Représentation des mutants dans la région C-terminale d'IFN $\gamma$ qui seront utilisés pour étudier l'importance des domaines D1 et D2 dans les interactions avec les deux ligands d'IFN $\gamma$	215
<b><u>Figure 42:</u></b> Gel SDS-PAGE de l'expression de plusieurs mutants d'IFN $\gamma$ en système bactérien	216
<b><u>Figure 43:</u></b> Prédiction du désordre dans la structure d'IFN $\gamma$	217
<b><u>Figure 44:</u></b> Tableau récapitulatif de toutes les structures cristallines des complexes obtenues entre une protéine et des GAGs (HS, fragment d'Hp ou mimétiques)	219

# ABBREVIATIONS

aa : acide aminé	ECL: Electrogenerated ChemiLuminescence
ABD : Antithrombin Binding Domain	EDC: 1-Ethyl-3-(Dimethylaminopropyl)Carbodiimine
ADN : Acide DeoxyriboNucléique	EDTA : Acide Ethylène-Diamine-Tétraacétique
ADNc : Acide DésoxyriboNucléique complémentaire	EGF : Epidermal Growth Factor
AH : Acide Hyaluronique	EGFR : Epidermal Growth Factor Receptor
ApoO : Apolipoprotéine O	ELISA: Enzyme Linked ImmunoSorbent Assay
AREs : AU-Rich Elements	EMBL: European Molecular Biology Laboratory
ARN : Acide RiboNucléique	eNOS : endothelial Nitric Oxide Synthase
ARNm : Acide RiboNucléique messenger	ESM-1 : Endothelial Cell-Specific Molecule 1
Asn : Asparagine	ESRF: European Synchrotron Radiation Facility
ATIII : AntiThrombine III	FACS : Fluorescence-activated cell-sorting
BCA: Bicinchoninic Acid	FAK : Focal Adhesion Kinase
BET: Bromure d'ETHidium	FGF : Fibroblast Growth Factor
BLAST: Basic Local Alignment Search Tool	FGFR : Fibroblast Growth Factor Receptor
BSA: Sérum Albumine Bovine	GAG : GlycosAminoGlycanne
CE : Capillary Electrophoresis	GalN : Galactosamine
Ch ABC : Chondroïtinase ABC	GalNAc-T : N-acétyl Galactosamine Transférase
Ch ACI : Chondroïtinase ACI	Gal-T : GalactosylTransférase
Ch B: Chondroïtinase B	GlcA : Glucuronic Acid
CHO : Chinese Hamster Ovary Cell	GlcN : Glucosamine
ChPF : Chondroïtine Polymerizing Factor	GlcNAc : Glucosamine N-acétyl
ChSy : Chondroïtine Synthase	GlcNH <sup>3+</sup> : Glucosamine N-désacétyl
CKII : Caséine Kinase II	GlcNS : Glucosamine N-sulfaté
CMH : Complexe Majeur d'Histocompatibilité	Gly : Glycine
Cpm : Counts Per Minute	GMSA: Gel Mobility Shift Assay
CRE : cAMP Response Element	GPI : Glucosyl Phosphatidyl Inositol
CS : Chondroïtine Sulfate	GRO/MGSA: Growth Related Protein / Melanoma
CS/DS2ST : hexuronyl CS/DS 2-O-SulfoTransférase	Growth Stimulatory Activity
CSPG : Chondroïtine Sulphate ProteoGlycan	HAS : Hyaluronic Acid Synthase
CTAB: Cetyl TrimethylAmmonium Bromide	HBS : Hepes Buffer Saline
Cys : Cystéine	HCAEC : Human Coronary Artery Endothelial Cells
DEAE-Sephacel : DiEthylAminoEthyl-Sephacel	HCII : Heparin Cofactor II
dNTP : DésoxyriboNucléotide TriPhosphate	HCl: Acide Chlorhydrique
DS : Dermatane Sulfate	HDMVEC : Human Dermal Microvascular
DTT: DiThioThreitol	Endothelial Cells
E. coli : Escherichia coli	HEK : Human Embryonic Kidney

HGF/SF : Hepatocyte Growth Factor/Scatter Factor	KS : Keratan Sulfate
Hhex : Hematopoietically expressed homeobox	KSPG : Keratan Sulfate ProteoGlycan
HIO : Heparin-Induced Osteoporosis	LFA-1 : Lymphocyte Function-associated Antigen 1
HIT : Heparin-Induced Thrombocytopenia	LMWH : Low Molecular Weight Heparin
HLA-DR : Histocompatibility Leukocyte Antigens Class II	LPS : LipoPolySaccharides
Hp : Héparine	LRRs : Leucine Rich Repeats
Hpa : Héparanase	LT : Lymphocytes T
HPAEC : Human Pulmonary Artery Endothelial Cells	Lymphocytes TH : Lymphocytes T helper
HPC : Hematopoietic Progenitor Cell	MAF : Macrophage-Activating Factor
HPCE : High Performance Capillary Electrophoresis	MALDI-TOF: Matrix Assisted Laser Desorption Ionisation - Time Of Flight
HPLC : High Performance Liquid Chromatography	MCP1: Monocyte Chemotactic Protein-1
HRP: Horse Radish Peroxidase	MCP2: Monocyte Chemotactic Protein-2
HS : Héparane Sulfate	MCSF-1 : Macrophage Colony Stimulating Factor
HSPG : Heparan Sulfate Proteoglycan	MCSP : Melanoma CS Proteoglycan
HSQC : Heteronuclear Single Quantum Coherence	MgCl <sub>2</sub> : Chlorure de Magnésium
HSV : Herpes Simplex Virus	MHV : Mouse Hepatitis Virus
HUEC : Human Capillary Endothelial Cells	MIP : Macrophage Inflammatory Protein
HUVEC : Human Umbilical Vein Endothelial Cells	MMP : Matrix MétalloProtéinase
IBD : Inflammatory Bowel Disease	MPS : MucoPolySaccharidose
ICAM : Intercellular Adhesion Molecule	MuSK : Muscle-Specific receptor tyrosine Kinase
Ido A : Iduronic Acid	MWCO: Molecular Weight Cut Off
IFN $\gamma$ : Interféron gamma	NaOH: Hydroxyde de Sodium
Ig : Immunoglobuline	NDST : N-Déacetylase/N-SulfoTransférase
IGF : Insulin Growth Factor	NHS: N-Hydroxy-Sulfosuccinimide
IGFBP : Insulin Growth Factor Binding Protein	NK : Natural Killers Cells
IGFR : Insulin-like Growth Factor Receptor	NO : Nitric Oxid
IgG: Immunoglobuline G	OA : OstéoArthrose
IGIF : IFN $\gamma$ -Inducing-Factor	PAGE : Poly-Acrylamide Gel Electrophoresis
IJAK-STAT : Janus Kinase et Signal Transducers and Activators of Transcription	PAGEFS: Polyacrylamide Gel Electrophoresis of Fluorophore-labeled Saccharides
IL : InterLeukine	PBS: Phosphate Buffer Saline
INSERM : Institute National de la Santé et de la Recherche Médicale	PCR : Polymerase Chain Reaction
IP : Phosphatidylinositol	PDGF : Platelet Derived Growth Factor
IP10: Interferon Inducible Protein 10	PECAM-1 : Platelet Endothelial Cell Adhesion Molecule-1
I $\alpha$ I : Inter-alpha-Inhibitor complex	PF-4 : Platelet Factor 4
kDa : kilo Dalton	PG : ProtéoGlycane
KGF : Keratinocyte Growth Factor	PKC : Protéine Kinase C
KO : Knock Out	PM : Poids Moléculaire

RANTES : Regulated upon Activation, Normal T-cell Expressed and Secreted	Sulf : endosulfatases
RE : Réticulum Endoplasmique	SUMO: Small Ubiquitin-related Modifier
RMN: Résonance Magnétique Nucléaire	SVF : Sérum de Veau Fœtal
RT-PCR : Reverse Transcription-Polymerase Chain Reaction	TA : Température Ambiante
SDF : Stromal Derived Factor	TARC: Thymus and Activation-Regulated Chemokine
SDS: Sodium Dodecyl Sulfate	TBD : Thrombin Binding Domain
SDS-PAGE: Sodium Dodecyl Sulfate PolyAcrylamide Gel Electrophoresis	TBE: Tris-Borate-EDTA buffer
Ser : Sérine	TBS: Tris Buffer Saline
SFM : serum-free medium	TGF : Transforming Growth Factor
SIDA : Syndrome de l'immunodéficience acquise	Thr : Thréonine
SLC : Secondary Lymphoid Tissue Chemokine	TNF: Tumor Necrosis Factor
SLRP : Small Leucine-Rich Protein	VEGF : Vascular Endothelial Growth Factor
SPR : Surface Plasmon Resonance	VEGFR : Vascular Endothelial Growth Factor Receptor
STAT-1 : Signal Transduction and Activator of Transcription -1	VIH : Virus de l'Immunodéficience Humaine
	WB: Western Blot

# INTRODUCTION

Le travail présenté dans ce manuscrit est le fruit d'une collaboration étroite entre le laboratoire d'Hugues Lortat-Jacob de l'Institut de Biologie Structurale de Grenoble et la société de biotechnologie Endotis Pharma. Endotis Pharma est une société créée au début de l'année 2003. Elle avait à ce moment précis deux projets principaux. Le premier projet, portait sur le protéoglycane endocan (d'où le nom de la société), découvert en 1996 dans un laboratoire de l'Institut Pasteur de Lille dirigé par le Dr Philippe Lassalle. Endocan est rapidement apparu comme une molécule potentiellement intéressante comme marqueur biologique, mais aussi comme une cible thérapeutique. En effet, les premiers travaux de recherche ont montré que la présence de ce protéoglycane (PG) circulant était corrélée avec un mauvais pronostic dans plusieurs types de cancers. Dans ce cadre, le développement d'un kit de diagnostic pour le dosage de la molécule endocan s'avérait possible grâce aux anticorps monoclonaux spécifiques développés par Philippe Lassalle et son équipe. De plus, il avait été montré que certains de ces anticorps monoclonaux anti-endocan pouvaient bloquer la formation de tumeurs cancéreuses dans des modèles murins. Ces résultats plaçaient donc endocan comme une éventuelle cible thérapeutique et nécessitaient non seulement de développer des outils de détection mais aussi de mieux comprendre la nature et la structure de ce nouveau protéoglycane. C'est ce qui a constitué le premier thème de recherche qui sera développé dans ce manuscrit. Le second projet développé par la société Endotis Pharma (projet toujours en activité) était d'utiliser des mimétiques de GAGs pour moduler l'activité de protéines définies comme cibles thérapeutiques. Pour ce projet, la protéine interféron gamma ( $IFN\gamma$ ), dont l'interaction avec les héparanes sulfates (HS) avait largement été décrite par des travaux antérieurs d'Hugues Lortat-Jacob, constituait une première cible de choix, notamment dans certains désordres immunitaires où le contrôle des activités de l' $IFN\gamma$  pourrait s'avérer intéressant. Le second thème de recherche développé dans ce manuscrit a donc eu pour objectif d'enrichir les informations structurales et fonctionnelles sur l'interaction entre l' $IFN\gamma$  et les HS dans l'optique de développer un mime oligosaccharidique des HS capable de moduler l'activité de la cytokine.

Dans le but de faire avancer les deux projets d'Endotis Pharma, une convention CIFRE a été établie entre Endotis Pharma et le laboratoire d'accueil de l'Institut de Biologie Structurale. Cette convention a eu pour objectif de faire bénéficier à Endotis Pharma de toute

l'expérience, des méthodes de travail et les acquis détenus par l'équipe de recherche du Dr. Lortat-Jacob ainsi que des instruments de haute technologie qu'offrent le cadre grenoblois.

Ce manuscrit synthétise les travaux aussi bien expérimentaux que bibliographiques qui ont été effectués pour faire progresser ces deux thèmes de recherche. Tout d'abord, une revue bibliographique détaillée présente les différentes familles de PGs ainsi que leurs fonctions, en particulier leur capacité à interagir avec de nombreuses protéines. La partie suivante recouvre autant que faire se peut les molécules issues de ces travaux de caractérisation et actuellement développées comme ayant un intérêt potentiellement thérapeutique. Finalement, le dernier chapitre de cette revue bibliographique est consacré à décrire l'IFN $\gamma$ , les relations que cette cytokine possède avec les glycosaminoglycannes et l'intérêt de pouvoir moduler son activité.

Comme endocan est un PG dont les fonctions biologiques ne sont pour l'instant pas encore clairement identifiées, il semblait en effet intéressant dans cette partie bibliographique parfois dense, d'avoir une vue générale, sur les homologies ou les différences que pouvait avoir ce protéoglycane toujours mystérieux au sein de la gigantesque famille des PGs. Après la présentation des objectifs des travaux expérimentaux, les résultats exposés s'organisent selon les deux thèmes de recherche de ce travail de thèse. L'étude structurale et fonctionnelle d'endocan est tout d'abord exposée avec la présentation de la mise en place d'une plateforme technologique de production et de purification d'endocan, de sa chaîne glycanique et de son core protéique. Ensuite, les caractérisations structurales de la chaîne glycanique et du core protéique sont exposées de manière indépendante. Et enfin, certaines des interactions et effets que peut avoir l'endocan vis-à-vis de protéines cibles ont été évalués dans le but d'approfondir les fonctions de ce nouveau PG. Dans le chapitre suivant, les travaux sur la caractérisation structurale de l'interaction entre l'IFN $\gamma$  et les HS ont été exposés, notamment ceux qui ont mené à l'identification d'un mime oligosaccharidique capable de moduler l'activité de cette cytokine. Pour chacun de ces deux projets de recherche de ma thèse de doctorat, une conclusion et des perspectives seront énoncées.

# **RAPPORT BIBLIOGRAPHIQUE**

**CHAPITRE I :**  
**LES PROTEOGLYCANNES (PGs)**

## 1. GENERALITES

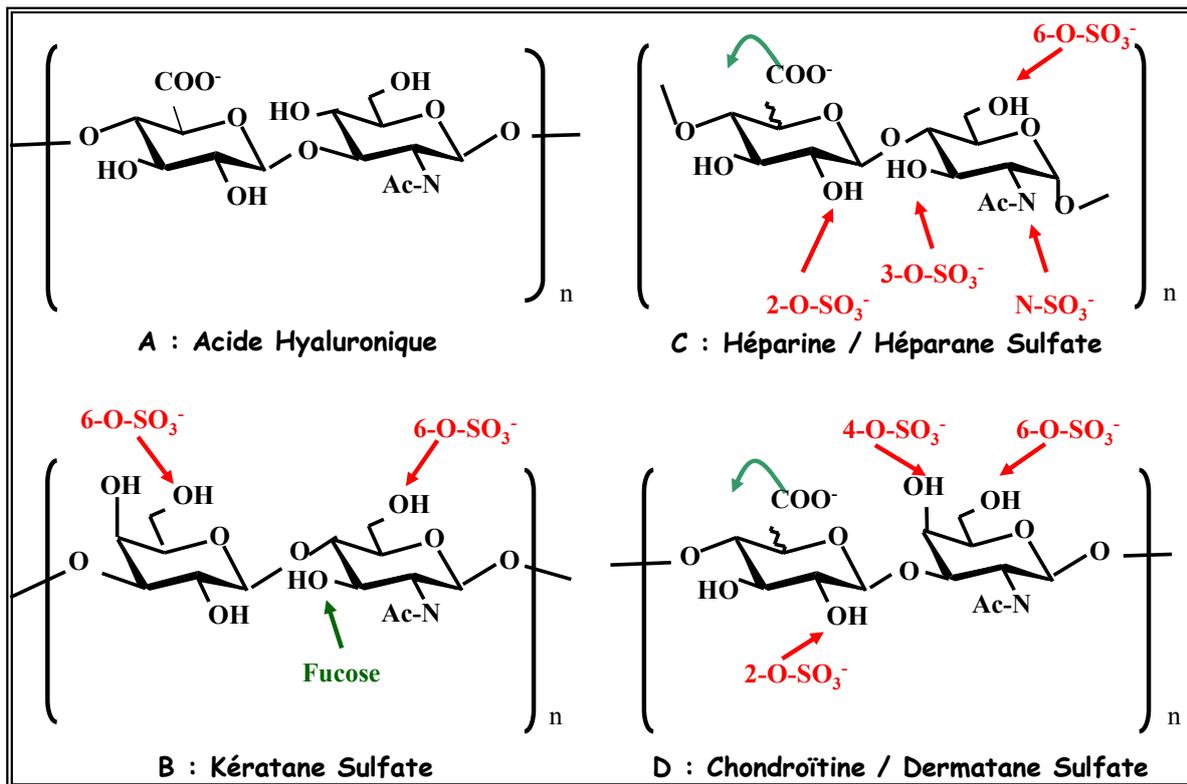
Les protéoglycannes (PGs) sont des macromolécules hydrophiles majoritairement situées à la surface des cellules, dans les matrices extracellulaires, et ils peuvent être dans certains cas intracellulaires voire circulants. **Ils sont constitués d'un "noyau" (core) protéique sur lequel sont greffés un ou plusieurs polysaccharides, principalement des glycosaminoglycannes (GAGs).** Le terme de core protéique va être utilisé tout au long du manuscrit pour désigner la partie polypeptidique d'un protéoglycanne. Leurs fonctions peuvent être exercées aussi bien par leur core protéique que par leurs chaînes oligosaccharidiques. Le core protéique va permettre principalement la localisation, alors que les glycosaminoglycannes agissent en particulier, en interagissant avec diverses protéines. Les protéines qui interagissent avec les GAGs comportent des domaines riches en acides aminés basiques, comme cela est le cas pour les chimiokines. La gamme de partenaires protéiques avec lesquels les GAGs sont amenés à interagir est si large (facteurs de croissance, cytokines, chimiokines, enzymes et autres protéines de la matrice extracellulaire) que les PGs sont impliqués dans de nombreux processus essentiels tels que la prolifération, le remodelage des matrices extracellulaires, la migration, l'angiogenèse, etc. Nombreux de ces processus se trouvent être déréglés lors de maladies et notamment au cours du développement tumoral (Iozzo et Cohen, 1993 ; Iozzo et San Antonio, 2001 ; Timar et al., 2002 ; Delehedde, 2002 ; Sanderson et al., 2004 ; Delehedde et al., 2005 ; Delehedde et al., 2006).

## 2. LES GLYCOSAMINOGLYCANNES (GAGs)

### 2.1. Généralités

Les GlycosAminoGlycannes (GAGs) sont de longs polysaccharides non ramifiés, constitués par polymérisation d'unités répétitives de disaccharides. Ces disaccharides, dont le terme biochimique est diholosides, sont composés de deux oses (c'est-à-dire de sucres simples non hydrolysables) qui sont reliés entre eux par une liaison osidique. La création d'une telle liaison bloque l'activité des fonctions qu'elle engage et notamment l'activité réductrice des fonctions aldéhydes des sucres ainsi que le phénomène de mutarotation.

Les disaccharides qui constituent les chaînes de GAGs, peuvent être plus ou moins sulfatés, et sont composés d'une manière générale (exception faite pour les kératanes sulfates) d'un acide hexuronique (soit un acide glucuronique (GlcA) ou un acide iduronique (IdoA)) et



**Figure 1: Motifs disaccharidiques des différents glycosaminoglycannes et leurs modifications**

**A :** L'acide hyaluronique est constitué de disaccharides reliés entre eux par des liaisons  $\beta$ 1-4. Les unités disaccharidiques sont composées d'un acide glucuronique et d'une N-acétylglucosamine, liés entre eux par une liaison  $\beta$ 1-3.

**B :** Les chaînes de kératane sulfate (KS) sont constituées d'unités disaccharidiques liées entre elles par une liaison  $\beta$ 1-3 et composées d'un galactose (à la place d'un acide hexuronique) et d'une N-acétylglucosamine. Les KS peuvent subir des 6-O-sulfatations sur les deux oses, et de façon unique chez les GAGs un fucose peut être ajouté en C3 de la N-acétylglucosamine.

**C :** L'héparine (Hp) ou les héparanes sulfates (HS) sont constitués d'unités disaccharidiques liées entre elles par une liaison  $\alpha$ 1-4 et composées d'un acide hexuronique et d'une N-acétylglucosamine liés par une liaison  $\alpha$ 1-4. L'Hp et les HS peuvent subir des 3, 6 ou N-sulfatations (après déacetylation) sur la N-acétylglucosamine ainsi qu'une 2-O-sulfatation sur l'acide hexuronique. L'épimérisation possible du carbone 5 transforme un acide glucuronique en un acide iduronique.

**D :** Les chondroïtines sulfates (CS) sont constitués d'unités disaccharidiques liées entre elles par une liaison  $\beta$ 1-4 et composées d'un acide glucuronique et d'une N-acétylgalactosamine liés par une liaison  $\beta$ 1-3. Les CS peuvent subir des 4- ou 6-sulfatations sur la N-acétylgalactosamine ainsi qu'une 2-O-sulfatation sur l'acide hexuronique. Des lors qu'un acide glucuronique subit une épimérisation en C5, la création d'un acide iduronique fait de la chaîne CS une chaîne de dermatane sulfate (DS).

d'une hexosamine (soit une glucosamine (GlcN) ou une galactosamine (GalN)). Selon la nature de l'hexosamine, les termes de glucosaminoglycannes ou de galactosaminoglycannes peuvent être utilisés.

Une des caractéristiques de ces chaînes oligosaccharides est leur importante hétérogénéité. En effet, la longueur des chaînes de GAGs pouvant être très variable, et les modifications structurales nombreuses (sulfatations, épimérisations), le nombre de combinaisons de structures possibles pour une chaîne oligosaccharidique apparaît infini.

Selon la nature des oses et la façon dont les disaccharides sont reliés entre eux, les GAGs peuvent être classés en différentes familles que nous allons détailler.

## **2.2. Les familles de glycosaminoglycannes (GAGs)**

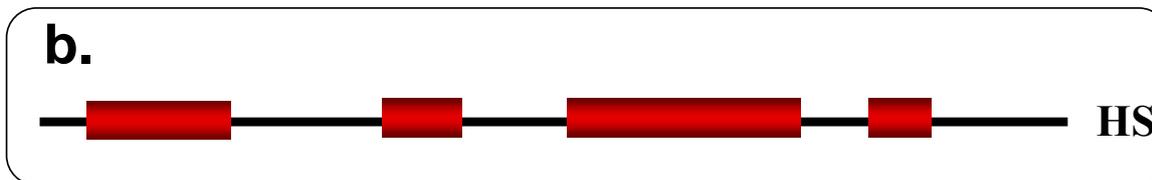
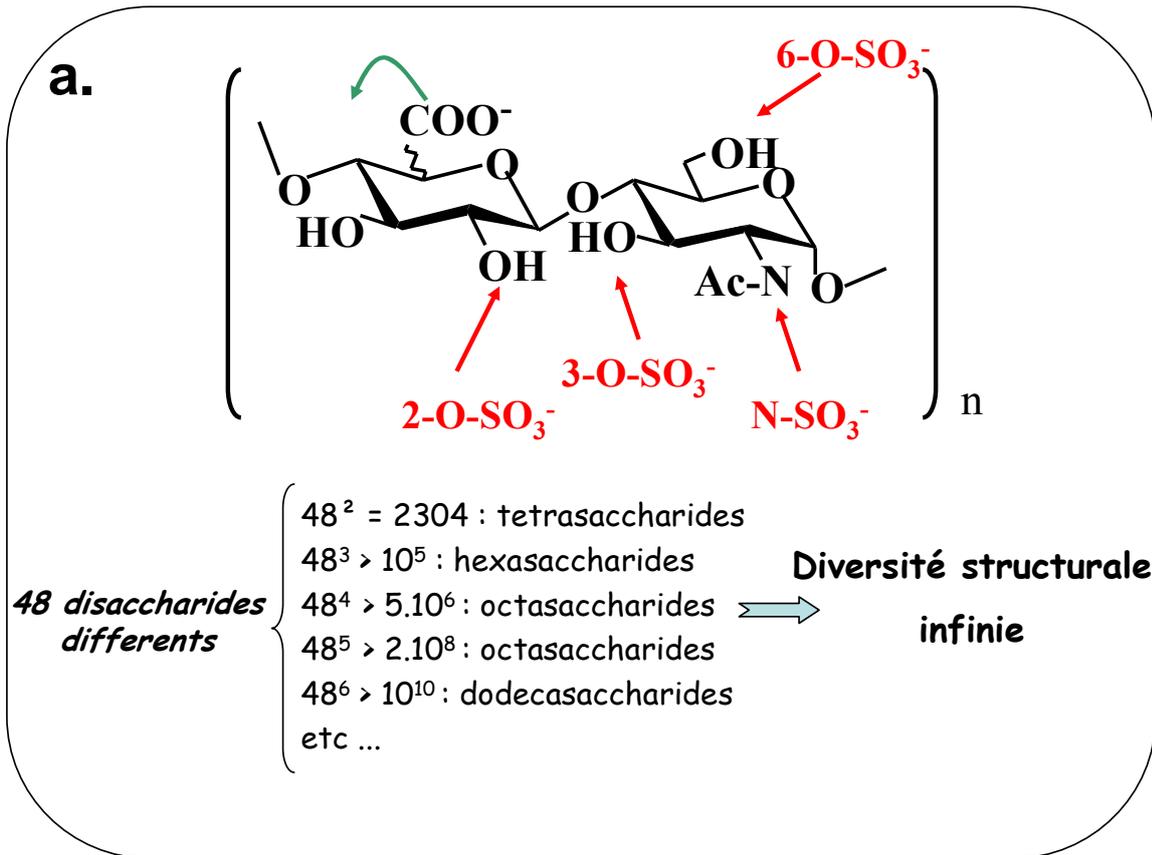
### **2.2.1. L'acide hyaluronique (AH)**

L'acide hyaluronique (du grec *hyalos* qui signifie vitreux, et d'uronique, parce que d'abord isolé de l'humeur vitrée, il possède un taux élevé d'acides uroniques) est un glycosaminoglycane largement réparti dans les tissus conjonctifs, épithéliaux et nerveux. On le trouve ainsi en abondance dans l'humeur vitrée, le liquide synovial et la peau. Il constitue également l'un des principaux composants de la matrice extracellulaire. Il a la particularité de n'être jamais relié à un corps protéique (il ne forme donc pas de protéoglycane en tant que tel), de n'être jamais sulfaté, et d'être parfois présent dans certaines bactéries (Fraser et al., 1997). L'AH est souvent défini comme un GAG qui permet le maintien de l'hydratation de la matrice extracellulaire cutanée.

La chaîne oligosaccharidique d'AH dont la composition est connue depuis les années 50 (Meyer et al., 1953), est constituée de disaccharides reliés entre eux par des liaisons  $\beta(1-4)$ . Les unités disaccharidiques sont composées de l'acide glucuronique et de la N-acétylglucosamine, liés entre eux par une liaison  $\beta(1-3)$  (Figure 1). Les polymères de cette unité récurrente peuvent atteindre une taille de  $10^2$  à  $10^4$  kDa (représentant ainsi environ 50000 unités disaccharidiques) soit une taille comparable à celle d'une bactérie !

### **2.2.2. Les kératanes sulfates (KS)**

Les kératanes sulfates (KS) ont été identifiés pour la première fois en 1939 (Suzuki et al., 1939) dans des extraits de cornée, mais réellement caractérisés bien plus tard (Meyer et al.,



**Figure 2 : Diversité et organisation structurale des Héparanes Sulfates**

a. Unité disaccharidique d'HS composée d'un acide hexuronique et d'une glucosamine N-acétylée. Ce motif disaccharidique 4GlcA $\alpha$ 1-4GlcNAc $\alpha$ 1 est ainsi répété n fois et peut subir jusqu'à 5 modifications telles qu'une déacétylation de la GlcNAc suivi d'une sulfatation; des sulfatations peuvent également avoir lieu en position 3 et 6 sur la GlcNAc ainsi qu'en position 2 sur l'hexosamine. Le carbone C5 de l'hexosamine peut subir une épimérisation qui convertit l'acide glucuronique (GlcA) en acide iduronique (IdoA).

b. Modèle de l'organisation d'une chaîne d'HS constituée d'une succession d'unités disaccharidiques. Les domaines les plus sulfatés appelés domaines NS sont représentés en rouge, ce sont des domaines dont la composition est semblable à celle de l'héparine. Les domaines peu ou non sulfatés sont représentés en noir (domaines NA). Ils sont composés principalement d'unités disaccharidiques qui ont subi très peu de modifications post-synthétiques. La succession de domaines NS et NA constitue une organisation dite de type SAS (Sulfaté Acétylé Sulfaté)

1953). Trois types de KS ont été ainsi identifiés : les KS de type I, II ou III (KSI ; KSII ; et KSIII). Cette distinction est liée au fait que les KSI se lient à un corps protéique par une asparagine (Asp) (N-glycosylation) alors que les KSII et III se lient aux sérines ou thréonines par une liaison GalNAc-O-Ser/Thr pour les KSII (Funderburgh, 2000) et Man-O-Ser pour les KSIII (Krusius et al., 1986).

Bien que les KS soient classés dans la famille des GAGs, ils présentent plusieurs particularités exceptionnelles. En effet, chaque unité disaccharidique de KS est constituée d'un galactose (à la place d'un acide hexuronique) et d'un glucose N-acétylé (Figure 1). Ces disaccharides (3Gal $\beta$ 1-4GlcNac $\beta$ 1) peuvent être sulfatés en C6 sur les deux oses. Un fucose peut également se brancher via son C1 sur le C3 de la glucosamine, en faisant de nouveau du KS un GAG à part (Greiling, 1994). Par ailleurs, les KS peuvent également dans certains cas être branchés et posséder des acides sialiques à leurs extrémités (Funderburgh, 2000).

### **2.2.3. L'héparine et les héparanes sulfates (Hp et HS)**

L'héparine a en fait été découverte en 1916 par Mc Lean qui, lors d'une expérience inattendue, a pu mettre en évidence l'effet anti-coagulant de ce polysaccharide (McLean, 1916). Dès 1935, l'Hp a été utilisée dans le traitement de l'embolie pulmonaire par voie veineuse et, en 1970, par la voie sous cutanée (Mc Lachelin, 1970).

L'héparine (Hp) ou les héparanes sulfates (HS) sont des GAGs dont les unités disaccharidiques sont composées d'un acide hexuronique et d'une glucosamine N-acétylée liées ainsi : 4GlcA $\alpha$ 1-4GlcNac $\alpha$ 1 (Figure 1). La glucosamine peut être soit N-acétylée (GlcNac), soit N-sulfatée (GlcNS). Une 6-O-sulfatation (GlcNac(6S) ou GlcNS(6S)) et même une 3-O-sulfatation sur les GlcNS (GlcNS(3S) ou GlcNS(3,6S)) peuvent également avoir lieu. Au cours de la synthèse sur laquelle nous reviendrons, l'acide glucuronique (GluA) peut être épimérisé en C5 pour donner l'acide iduronique (IdoA). Les deux formes épimérisées peuvent également subir ensuite une 2-O-sulfatation : IdoA(2S) et GlcA(2S). Il est rare, mais possible de trouver la glucosamine sous forme non glycosylée (GlcNH<sub>2</sub>) (Rabenstein, 2002) (Figure 2a).

Toutes ces modifications génèrent, le long d'une chaîne oligosaccharidique de ce type, une hétérogénéité et donc une variabilité de structure quasiment infinie. En effet, 48 disaccharides différents sont théoriquement possibles, et c'est leur combinaison qui va engendrer cette diversité structurale.

L'héparine, qui est le GAG le plus sulfaté, est composée majoritairement de disaccharides IdoA(2S)-(1-4)-GlcNS(6S) tout au long de sa chaîne.

Originellement identifiés comme étant des impuretés d'héparines (appelées heparitin sulfate) (Linker et al., 1958), les héparanes sulfates (HS) comportent en fait bien plus de variabilités dans leur composition et organisation structurale que l'Hp. En effet, les HS se composent de régions dites de type héparine et de régions bien moins sulfatées (Lindahl et al., 1998 ; Esko et Lindahl, 2001 ; Gallagher, 2001) (Figure 2b).

L'héparine et les héparanes sulfates n'ont pas le même lieu de synthèse et n'ont pas les mêmes fonctions. L'héparine est uniquement exprimée par les mastocytes situés dans les tissus conjonctifs. Elle est synthétisée sous forme de protéoglycane (PM : 750000 à 1000000 kDa) dont le core protéique est la serglycine. Plusieurs chaînes d'Hp (PM : 60000 à 100000 kDa) sont liées à ce core protéique. Une fois cette synthèse terminée, les chaînes d'Hp sont clivées aléatoirement pour donner des fragments d'Hp plus petits (PM : 5000 à 25000 kDa) qui vont être stockés dans des granules de sécrétion du cytoplasme des mastocytes. Les héparanes sulfates, quant à eux, sont exprimés par la plupart, voire toutes les cellules. Ils se trouvent rarement sous forme libre mais en liaison avec divers cores protéiques pour former les protéoglycannes à héparanes sulfates (HSPGs). Ces HSPGs se trouvent soit dans les matrices extracellulaires, soit à la surface des cellules (Rabenstein, 2002).

#### **2.2.4. Les chondroïtines sulfates (CS)**

Le chondroïtine sulfate (CS) ou sulfate de chondroïtine, est un GAG que l'on retrouve principalement dans les tissus conjonctifs. L'unité disaccharidique de base qui le compose est : -4GlcA $\beta$ 1-3GalNAc $\beta$ 1- (Figure 1). Comme les autres GAGs, il peut subir des modifications en fonction des cellules qui le synthétisent, de leur état d'activation ou du contexte physiologique (normal ou pathologique). Ainsi, sur des chaînes entières, la N-acétylgalactosamine peut être sulfatée en position C4 pour donner le chondroïtine sulfate de type A (**CS A**) ou en position C6 pour donner la **CS C**. Chez les calamars, les requins ou les crustacés, certains CS sont dits "très sulfatés" tels que le **CS D** (-4GlcA(2S)( $\beta$ 1-3)GalNAc(6S) $\beta$ 1) et le **CS E** (4GlcA( $\beta$ 1-3)GalNAc(4,6S) $\beta$ 1) et même le **CS iE** (4IdoA( $\beta$ 1-3)GalNAc(4,6S) $\beta$ 1) comportant un IdoA (ce qui en fait un DS).

### **2.2.5. Le dermatanes sulfate (DS)**

Les dermatane sulfate (DS) sont souvent classés comme étant une sous-classe des chondroïtines sulfates. On peut d'ailleurs les trouver sous le nom de **CS B**. En fait, on pourrait définir un DS comme une chaîne de CS qui comporte au minimum un acide iduronique issu de l'épimérisation en C5 d'un acide glucuronique d'un CS (Figure 1).

Ainsi, l'information structurale que peut contenir une chaîne de DS est plus diversifiée que celle d'un CS puisque l'acide glucuronique peut être épimérisé pour donner l'acide iduronique. Ce dernier peut alors être 2-O-sulfaté (la réaction d'épimérisation ayant préférentiellement lieu lorsque la galactosamine est 4-O-sulfatée).

L'épimérisation de l'acide uronique qui distingue les CS (A ou C) des DS n'est pas une modification anodine. En effet, la caractérisation d'interaction protéines-GAGs a permis de démontrer que cette épimérisation pouvait apporter une certaine spécificité aux DS. Les DS, par exemple, vont interagir avec l'héparine cofacteur II alors que les CS ne le peuvent pas (Mascellani et al., 1993). Il en est de même pour les interactions avec un facteur de croissance tel que le HGF/SF (Hepatocyte Growth Factor / Scatter Factor) (Lyon et al., 1998). Plus spécifiquement encore, seuls les DS parmi les GAGs sont capables d'activer le NF-kb et d'induire l'expression d'ICAM-1 circulant et endothélial (Penc, 1999).

## **2.3. Biosynthèse des glycosaminoglycannes**

### **2.3.1. La biosynthèse des GAGs**

Comme nous venons de le voir, il existe donc différentes familles de GAG, les Hp/HS, les CS/DS, les KS et l'AH. Chaque GAG a des caractéristiques différentes : ils diffèrent par leur structure, leur taille, leurs modifications post-synthétiques, leurs lieux de biosynthèse, leurs fonctions et leurs localisations. Il existe une large gamme d'enzymes spécifiques de chaque famille de GAGs qui agissent de manière très fine, en étant régulées par le microenvironnement biologique. Cependant, lorsque l'on parle de biosynthèse de GAGs, il est possible de définir deux groupes de GAG, à savoir ceux qui ne sont pas liés à un core protéique lors de leur élongation (l'AH) et ceux qui sont synthétisés depuis un point d'ancrage qui se trouve sur un core protéique (KS, Hp/HS et CS/DS) pour donner naissance aux protéoglycannes. Parmi les GAGs qui sont synthétisés depuis un core protéique, les Hp/HS et

CS/DS peuvent être distingués des KS puisque leur point d'ancrage sur le core protéique est clairement différent (Taylor et Gallo, 2006).

#### **2.3.1.1. La biosynthèse de l'AH, le seul GAG à n'être rattaché à aucun core protéique**

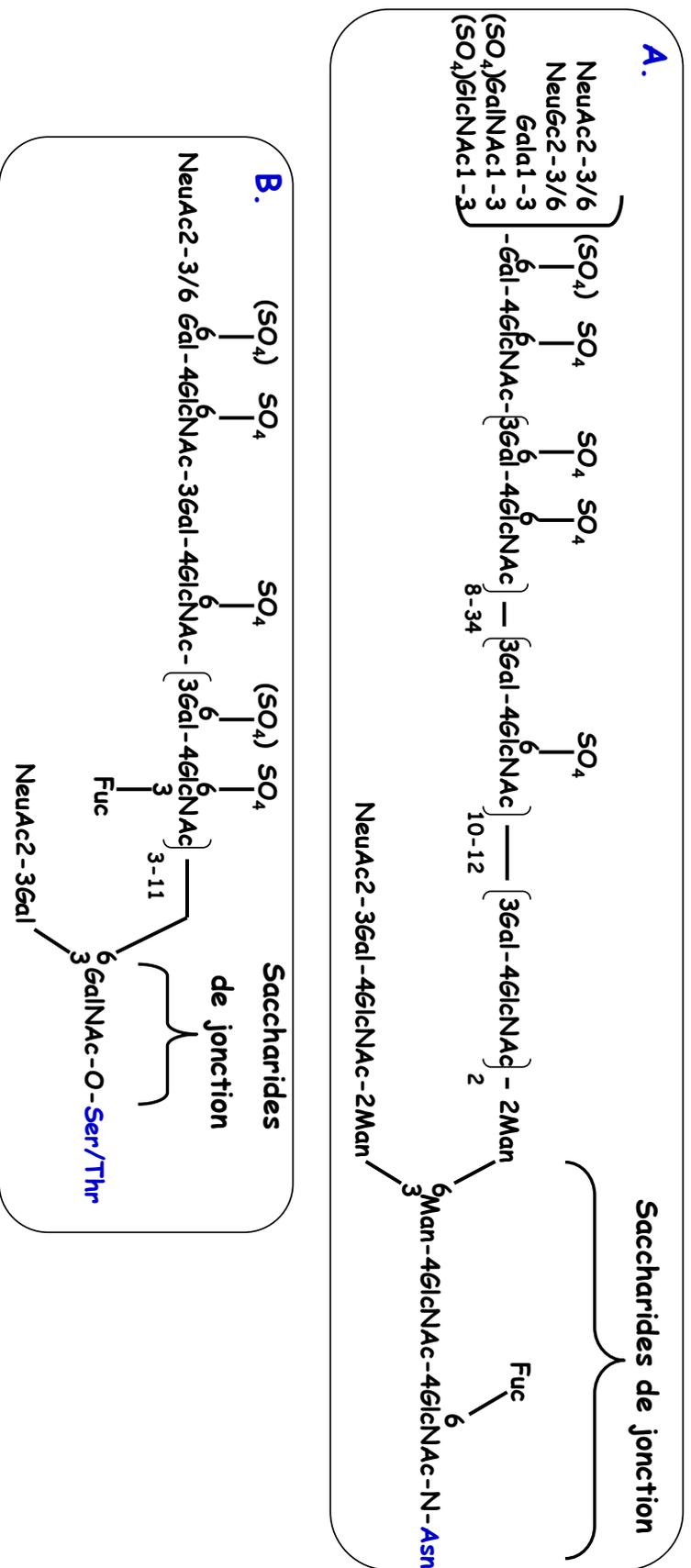
L'AH est assez particulier tout en étant l'un des GAGs les moins complexes. En effet, l'AH est directement synthétisé dans l'espace extracellulaire par les synthases d'AH (HAS) et ne subit aucune modification post-synthétique. Les HAS sont des protéines à sept domaines transmembranaires situées dans la membrane plasmique (Weigel et al., 1997). Chez les mammifères, trois HAS ont été identifiées (HAS I, II et III). Élément atypique chez les GAGs, certaines bactéries et virus possèdent également les gènes HAS, et peuvent donc synthétiser ou faire synthétiser (dans le cas des virus) leur propre AH (DeAngelis, 1999).

Les HAS agissent de manière unique. En effet, elles possèdent deux composantes enzymatiques, l'une qui permet l'élongation sur le GlcA et l'autre sur la GlcNAc. Contrairement à tous les autres GAGs, l'élongation de la chaîne oligosaccharidique par l'ajout de monosaccharides a lieu sur la partie réductrice de la chaîne (libre puisque non greffée sur un core protéique).

La taille des AH semble être un élément majeur de la régulation de leurs activités biologiques (Noble, 2002 ; McKee, 1996) et elle dépend de l'activité de chacune des enzymes HAS I, II et III. Les HAS I, II et III produisent des chaînes respectivement de  $2.10^5$  Da, de  $2.10^6$  Da et supérieure à  $2.10^6$  Da. La localisation bien distinguée des gènes HAS 1, 2 et 3 sur les chromosomes, pourrait expliquer une régulation pointue de la biosynthèse de l'HA (Itano et Kimata, 2002).

#### **2.3.1.2. La biosynthèse des GAGs liés à un core protéique**

A l'exception de l'AH, tous les autres GAGs sont synthétisés à partir d'un core protéique. Selon le core protéique et l'environnement biologique, la nature des GAGs synthétisés varie. Cependant, un même schéma général est observé. Sur une séquence protéique préférentielle reconnue, une séquence oligosaccharidique consensus d'ancrage (ou "stub") est ajoutée sur un acide aminé bien précis (la Ser ou la Thr pour l'Hp/HS, CS/DS et KSII et III, et l'Asn pour le KSI). L'élongation de la chaîne oligosaccharidique peut alors être initiée à partir de cette



**Figure 3: Représentation du profil structural des KS I (A) et KS II (B)**

Les saccharides de jonction avec le core protéique offre au KS un branchage qui n'est pas présent chez les autres GAGs et permet également une liaison avec un fucose dans le cas du KSII.

**A :** Oeben et al. (1987) proposent que les disaccharides proches de l'extrémité réductrice du KS I sont non-sulfatées. Se juxtaposent ensuite des unités disaccharidiques monosulfatées principalement avant que l'extrémité de la chaîne soit elle très sulfatée. Le second branchage issu du C3-Man des saccharides de jonction est représenté par une chaîne courte. Cependant une chaîne longue peut de temps en temps être observée.

**B :** Dans le cas du KSII les chaînes oligosaccharidiques sont plus courtes mais les unités disaccharidiques sont dans leur majorité disulfatées à l'exception de quelques unes qui peuvent être monosulfatées. De façon unique chez les GAGs, un fucose peut être ajouté tout au long de la chaîne sur le GlcNAc en position C3.

(d'après Funderburgh, 2000)

séquence saccharidique d'ancrage. Cependant, l'Hp/HS et les CS/DS possèdent une même séquence tétrasaccharidique consensus qui est bien différente de celles des KS I, II ou III.

#### 2.3.1.2.1. La biosynthèse des KS

##### La séquence saccharidique consensus de liaison au core protéique

Comme il a été décrit précédemment, il existe 3 classes de KS (KSI, II et III). La différence majeure entre chacun d'entre eux est due à leur liaison à un core protéique.

Les KSI, les plus étudiés jusqu'à maintenant, sont liés au core protéique par une séquence oligosaccharidique consensus de composition : Man-6Man-4GlcNAc-4GlcNAc-N-Asn. Le premier GlcNAc en partant de l'Asn peut avoir un fucose lié en position C6. Par ailleurs, le premier mannose peut avoir en plus d'un mannose lié en C6, un autre mannose lié en C3, cela créant en quelque sorte un branchage (Oeben et al., 1987) (Figure 3a).

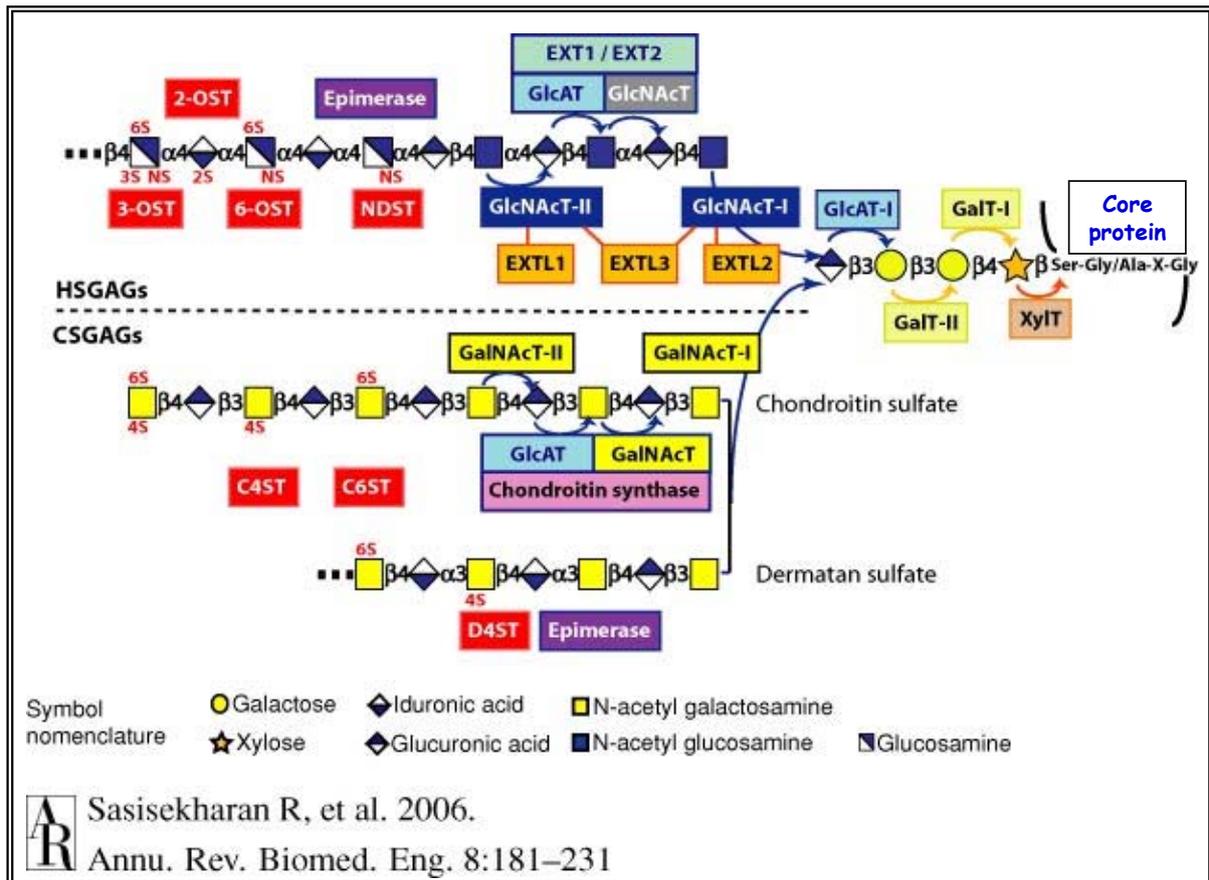
Les KSII se lient quant à eux par une séquence consensus très simple de type : 6GalNAc-O-Ser/Thr. A noter qu'en C3 du GalNAc, un branchage peut également avoir lieu (Figure 3b).

Les KSIII, qui restent peu étudiés, sont liés par une liaison Man-O-Ser (Krusius et al., 1986).

##### La synthèse des chaînes de KS I, II et III

Les chaînes oligosaccharidiques de KS sont synthétisées par des glycosyltransférases : les **galactosyltransférases** ( $\beta$ 4Gal-T) et les **N-acétylglucosaminyltransférases** (Gn-T) qui chacune à leur tour assemblent les monomères de Gal et de GluNAc. Pas moins de sept gènes codant pour les  $\beta$ 4Gal-T ont été identifiés ; cependant les fonctions de chacune de ces enzymes sont encore mal connues (Amado et al., 1999). En ce qui concerne les Gn-T, un certain nombre de ces enzymes sont connues mais aucune n'a encore été mise en évidence comme étant celle utilisée lors de la synthèse de KS (Funderburgh, 2000). Des travaux récents ont cependant montré que la  $\beta$ 3Gn-T7 pourrait être utilisée lors de la polymérisation des KS (Seko et Yamashita, 2003).

L'élongation d'un KS se déroule de la manière suivante : la  $\beta$ 3Gn-T7 ajoute la GlcNAc en position non réductrice de la chaîne. Cette dernière peut être 6-O-sulfatée par une sulfotransférase par exemple de type GlcNAc6ST-5 (Akama et al., 2002). Selon le lieu de biosynthèse des KS, on peut également trouver les GlcNAc6ST-1 et 2 (Kusche-Gullberg et Kjellen, 2003). Le Gal peut ensuite être ajouté en bout de chaîne par une  $\beta$ 4Gal-T et la  $\beta$ 3Gn-



**Figure 4: Biosynthèse des chaînes oligosaccharidiques d'héparine / héparanes sulfates (Hp/HS) et chondroïtines / dermatanes sulfates (CS/DS)**

La synthèse des GAGs de type Hp/HS et CS/DS nécessite un support protéique (core protéique). La sérine d'une séquence protéique préférentielle va permettre d'initier la synthèse d'un tétrasaccharide d'ancrage commun aux Hp, HS, CS et DS de séquence -4GlcA(β1-3)Gal(β1-3)Gal(β1-4)Xylβ1-O-Ser. Cette région de liaison entre le core protéique et les chaînes oligosaccharidiques proprement dites est créée par l'ajout d'un xylose sur la sérine, suivi par l'ajout de deux résidus galactosamines ensuite par un acide glucuronique. Chaque réaction est catalysée par des glycosyltransférases spécifiques qui sont respectivement la xylosyltransférase (XylT), la galactosyltransférase I (GalT-I) puis la galactosyltransférase II (GalT-II) et enfin la glucuronyltransférase I (GlcAT-I). Les enzymes GlcNAc transférase I et GalNAc transférase I ou II permettent respectivement l'ajout sur le tétrasaccharide d'ancrage d'une GlcNAc ou d'une GalNAc. Cette étape est décisive pour le choix de biosynthèse entre les Hp/HS ou CS/DS. Les GAGs peuvent alors être synthétisés par l'ajout alternatif et répété d'un acide glucuronique et d'une hexosamine N-acétylée. Les Hp/HS sont synthétisés par l'HS polymérase composée de deux enzymes EXT1 et 2 qui possèdent les activités de GlcNAc transférase et GlcA transférase. La synthèse des CS/DS est possible grâce à la Chondroïtine Synthase (ChSy). Cette enzyme récemment découverte possède les deux activités enzymatiques qui sont les GalNAc transférase II et GlcA transférase II. Elle a une bien meilleure activité en présence du ChPF (chondroïtine polymerizing factor) non représenté sur la figure (Sugahara et al., 2000 et 2003). Les enzymes (rectangles rouges) correspondent aux sulfotransférases responsables en partie de la diversité structurale. L'épimérase en C5 est quant à elle représentée dans un rectangle violet.

(D'après Sasisekharan et al., 2006)

T7 reprend son travail. Les 6-O sulfatations sur les Gal ont lieu lorsque plusieurs disaccharides sont en place et sont réalisées par la sulfotransférase KSGal6ST (Fukuta et al., 1997).

Les KS ont été identifiés dans de nombreux tissus, ce qui signifie que malgré quelques spécificités qui leur sont propres, les enzymes de la biosynthèse des KS sont exprimées dans la plupart des cellules. En revanche, un nombre limité de cores protéiques sont connus comme pouvant lier des KS. Funderburgh (2000) propose d'ailleurs que l'un des facteurs guidant la biosynthèse des KS pourrait être l'expression de ces cores protéiques, et qu'il faudrait donc, pour caractériser la biosynthèse des KS dans son intégralité, identifier les régulations de l'expression de ces cores protéiques.

#### 2.3.1.2.2. La biosynthèse des Hp, HS, CS et DS

##### La séquence saccharidique consensus de liaison au core protéique

Pour les Hp, HS, et CS/DS, la séquence consensus de liaison au core protéique est un tétrasaccharide de type : **4GlcA(β1-3)Gal(β1-3)Gal(β1-4)Xylβ1-O-Ser**. Sa formation a lieu par le transfert d'un UDP-Xylose via la **xylosyltransférase** (Xyl-T) sur le groupement hydroxyle d'une sérine du core protéique pour donner une liaison Xylβ1-O-Ser. A lieu ensuite l'ajout des deux galactoses et enfin l'ajout de l'acide glucuronique. Ces réactions sont réalisées respectivement par les **galactosyltransférases I et II** (GalT-1 et 2) puis par la **glucuronyltransférase I** (GlcAT-1) (Sugahara et Kitagawa, 2000) (Figure 4).

##### Le choix de l'élongation : une chaîne Hp/HS ou CS/DS ?

Les deux "familles" de GAGs, les Hp/HS et les CS/DS (dans le sens où la composition du disaccharide de base est différente) sont synthétisées à partir du même tétrasaccharide consensus. Le mécanisme qui définit le choix de la synthèse de l'une ou l'autre des deux familles a été mal compris pendant de nombreuses années, jusqu'à la découverte de deux enzymes clés : **la GlcNAc Transférase I** (GlcNAcT-1) et **la chondroïtine GalNAc Transférase I ou II** (chondroitin GalNAcT-I ou II selon la localisation de la biosynthèse). Ces enzymes initient la synthèse des chaînes d'Hp/HS et de CS/DS, respectivement par l'ajout d'une GlcNAcα1-4 ou d'une GalNAcβ1-4 au tétrasaccharide d'ancrage, par un mécanisme qui semble être dépendant de séquences protéiques particulières autour de la sérine (Esko et Zhang, 1996). La GlcNAcT-I possède un homologue, l'EXTL2, capable

d'ajouter à la séquence consensus un GalNAc mais avec une liaison  $\alpha(1-4)$  alors que pour initier un CS une liaison  $\beta(1-4)$  est requise. Cette possibilité n'a encore été reliée à aucune fonction biologique *in vivo*. Le choix de l'élongation d'une chaîne polysaccharidique d'Hp/HS ou de CS/DS serait dans certains cas, non pas dépendant des enzymes présentes dans l'environnement mais plutôt de la séquence protéique et donc du core protéique sur lequel a lieu l'élongation oligosaccharidique. L'activité de la GalNAcT-1, par exemple, serait dépendante du core protéique. En effet, Nadanaka et al. (1999) ont montré que l'élongation d'une chaîne de CS à partir du tétrasaccharide d'ancrage n'est possible que si ce dernier est lié au core protéique.

Certaines études ont montré que le tétrasaccharide consensus pouvait également subir des modifications ayant leur importance dans le choix de synthèse des Hp/HS ou des CS/DS. En effet, pour certaines chaînes de CS/DS, les tétrasaccharides d'ancrage peuvent avoir les compositions suivantes :  $\text{GlcA}\beta/\text{IdoA}(\alpha 1-3)\text{Gal}(\mathbf{4-O-sulfate})(\beta 1-3)\text{Gal}(+/-\mathbf{6-O-sulfate})(\beta 1-4)\text{Xyl}$  ; ou alors :  $\text{GlcA}(\beta 1-3)\text{Gal}(+/-\mathbf{6-O-sulfate})(\beta 1-3)\text{Gal}(\mathbf{6-O-sulfate})(\beta 1-4)\text{Xyl}$  et même :  $\text{GlcA}(\beta 1-3)\text{Gal}(\beta 1-3)\text{Gal}(\beta 1-4)\text{Xyl}(\mathbf{2-O-phosphate})$  (Sugahara et al., 1988 ; 1992 ; DeWaard et al., 1992). Le Xyl(2-O-phosphate) a été trouvée sur des chaînes d'HS ou de CS. En revanche, la 4-O-sulfatation sur un Gal a été trouvé uniquement sur des chaînes de CS (Yamada et al., 1995). L'enzyme GlcNAc transférase I, qui est essentielle pour l'initiation de la polymérisation des Hp/HS, semble en fait inhibée par le résidu Gal(4-O-sulfaté) (Kitagawa et al., 1995).

### La polymérisation des Hp/HS

Lorsque la synthèse des Hp/HS a débuté par l'ajout d'une GlcNAc sur le tétrasaccharide d'ancrage, les séquences répétées de  $4\text{GlcA}(\alpha 1-4)\text{GlcNAc}\beta 1$  constituant l'unité disaccharidique de base des Hp/HS, peuvent alors être ajoutées. Les enzymes **EXT1** et **EXT2** sont responsables de l'élongation des chaînes par l'ajout de ces disaccharides (Figure 4). Ces enzymes sont inactives dans le réticulum endoplasmique (RE) et sont indépendantes l'une de l'autre. En revanche, dans l'appareil de Golgi, elles forment un complexe qui est appelé l'HS polymérase. Ce complexe possède ainsi les deux activités transférases du type GlcA-T et GlcNAc-T, lui permettant de polymériser les chaînes oligosaccharidiques (McCormick et al., 2000).

### La polymérisation des CS/DS

Le mécanisme de polymérisation des chaînes de CS a été pendant longtemps mal connu. En théorie, deux activités enzymatiques sont nécessaires, à savoir une GlcA transférase et une GalNAc transférase. Bien qu'après les années 1960 ces activités enzymatiques aient été supposées, il a fallu attendre 1985 pour que les travaux de Rohrmann puissent révéler l'existence de deux GalNAc transférases (GalNAcT-I et II) (Rohrmann et al., 1985). Plus de dix ans plus tard, la GlcA transférase II et la GalNAc transférase II ont été co-purifiées à partir de sérum de bœuf (Tsuchida et al., 1999). Jusque là, seule la comparaison avec une chondroïtine synthase identifiée chez une bactérie (et capable de produire des chondroïtines) permettait de supposer qu'une seule et même protéine comportait les deux activités enzymatiques (DeAngelis et Padgett-McCue, 2000). Grâce aux progrès réalisés dans le séquençage du génome humain et, consécutivement, grâce à l'enrichissement des banques de données, le groupe de Sugahara a révélé l'existence d'enzymes spécifiques et a réussi à décortiquer le mécanisme de la polymérisation des chaînes de CS/DS. **La chondroïtine synthase (ChSy)** est ainsi la seule responsable de la polymérisation des chaînes de CS grâce à ses différentes activités enzymatiques de  $\beta$ 1-3-GlcA transférase et de  $\beta$ 1-4-GalNAc transférase (Kitagawa et al., 2001). Des études ont également démontré que la chondroïtine synthase avait une meilleure activité en présence du ChPF (**chondroitin polymerizing factor**) (Kitagawa et al., 2003).

### Les modifications de post-polymérisation (déacétylations, épimérisations et sulfatations)

Toutes les modifications des GAGs (dont les sulfatations) ont lieu dans l'appareil de Golgi grâce à des enzymes spécifiques qui agissent directement sur les polysaccharides en cours d'élongation ou déjà polymérisés. Ces modifications sont essentielles pour offrir aux oligosaccharides toute l'information dont ils doivent être porteurs pour être fonctionnels. Par exemple, si les HS ne sont pas sulfatés durant le développement d'une souris, la mort néonatale de cette dernière est inévitable (Bullock et al., 1998). L'environnement biologique a une très grande importance sur les différentes modifications qui peuvent orner les GAGs. Chaque modification est effectuée sous le contrôle d'enzymes régulées de manière fine par un réseau complexe afin de conduire aux "structures" de GAGs induits par le microenvironnement biologique et donc de répondre aux besoins de ce microenvironnement à un instant donné.

## Modifications des Hp/HS

Comme il a déjà été mentionné, l'Hp est un GAG qui est fortement sulfaté tout au long de sa chaîne. Ainsi, contrairement aux HS qui subissent des modifications de façon plus dispersée, chaque unité de chaque disaccharide d'Hp est le substrat de différentes enzymes que nous allons décrire ci-après.

Toutes les modifications sur le polymère d'Hp se déroulent de façon ordonnée. Tout d'abord, certaines des GlcNAc vont être N-déacétylées puis N-sulfatées. Ces deux réactions sont catalysées par une famille d'enzymes comportant les deux activités spécifiques : elles sont appelées les **N-déacétylases/N-sulfotransférases** (NDST). Il existe ainsi chez les mammifères, quatre NDST (NDST1, 2, 3 et 4). La glucosamine peut également rester N-déacétylée (GlcNH<sup>3+</sup>) (Toida et al., 1997). Bien que rare, cette modification ne semble pas anodine puisque des fonctions lui ont été associées (Norgard-Sumnicht et Varki, 1995).

Ensuite, les épimérisations en C5 des acides glucuroniques par la **C5-épimérase** ont lieu pour donner des acides iduroniques (IdoA). La présence d'IdoA favorise alors la 2-O-sulfatation de l'IdoA par la **2-O-sulfotransférase** qui reconnaît de préférence ce motif. Cette 2-O-sulfatation est d'autant plus favorable à l'IdoA qu'elle bloque l'épimérisation. Il est d'ailleurs possible de trouver la 2-O-sulfotransférase sous forme de complexe avec la C5-épimérase (Esko et Selleck, 2002). Cette sulfatation peut cependant avoir lieu parfois sur les GlcA (Rong et al., 2001). La famille des **6-O-sulfotransférases** (6-OST) exerce ensuite son activité en permettant l'ajout d'un sulfate en position 6 de la GlcNS, voire de la GlcNAc. Quatre 6-OST ont été dénombrées : les 6-OST1, 2 et 3, plus une variante de la 2, appelée la 6-OST2S. Dans quelques rares cas, un sulfate en position 3 de la GlcNS peut être ajouté par l'une des six **3-O-sulfotransférases** (3-OST1, 2, 3A, 3B, 4, ou 5) (Kusche-Gullberg et Kjellen, 2003). L'HS-3-OST 3A semble avoir une activité accrue en présence de GalNH<sup>3+</sup> (Liu et al., 1999).

Toutes ces modifications permettent d'offrir à cette succession de disaccharides de composition très simple une grande hétérogénéité structurale. Cette hétérogénéité structurale va permettre à l'Hp, aux HS et aux GAGs de manière quasi-générale d'engager des interactions très spécifiques avec certains de leurs ligands. Les différentes 3-O-sulfotransférases, même si leur activité est au fond similaire, sont par exemple capables d'offrir des profils de sulfatations spécifiques. En effet, la 3-OST-1 crée des séquences oligosaccharidiques plutôt favorables à une interaction avec l'antithrombine III alors que la 3-OST-3 sera favorable à la création de séquences permettant l'interaction avec la protéine gD de l'enveloppe du virus de l'Herpès (Yabe et al., 2001).

Les modifications apportées aux HS se déroulent de la même manière que pour l'Hp. En revanche, les enzymes de modification n'agissent pas uniformément sur le polymère d'HS dans sa totalité. Des régions subiront donc plus de modifications que d'autres, cela va permettre de créer au sein des oligosaccharides d'HS une diversité structurale plus importante que chez l'Hp. Il est ainsi possible de trouver sur la même chaîne d'HS, des régions très sulfatées qui sont de type héparine et sont appelées les domaines NS (NS-domain) et ensuite des régions non modifiées, non sulfatées, appelées domaines NA (NA-domain). La succession de domaines NS et NA constitue des séquences appelées SAS (pour Sulfaté Acétylé Sulfaté) (Figure 2b). Du fait que les différentes O-sulfatations ne peuvent en général avoir lieu que sur des disaccharides N-sulfatés, ce sont en quelque sorte les NDST qui dessinent l'ébauche de l'organisation des domaines NS et NA sur les HS.

### Modifications des CS/DS

Les CS/DS ne subissent pas de N-déacétylation contrairement aux Hp/HS. En revanche, comme pour l'héparine, l'épimérisation en C5 d'un GlcA par une **C5-épimérase** spécifique donne un IdoA. Dès lors que sur une chaîne de CS un GlcA est épimérisé en IdoA, il sera question d'un GAG de type DS ou CS B. Trois familles différentes de sulfotransférases sont responsables des modifications des CS/DS : celles qui modifient les GalNAc en C4 ou en C6 et celles qui modifient l'acide hexuronique en C2.

Dans la famille des **chondroïtine 4-O-sulfotransférases**, trois enzymes, les C4ST1, 2 et 3 agissent sur des disaccharides de type "CS" : GlcA-GalNAc et seule la C4ST1 est spécifique des GAGs de type "DS" : IdoA-GalNAc. Cette dernière agit immédiatement après l'épimérisation du GlcA, et la sulfatation qu'elle crée empêche l'épimérisation inverse (Silbert et Sugumaran, 2002). A l'image de ces deux enzymes, la C5-épimérase et la 2-O-sulfotransférase qui se complexent pour permettre la modification des HS, il est peut être juste de penser que la C5-épimérase et les 4-O-sulfotransférases modifient les CS également en se complexant.

Deux **chondroïtine 6-O-sulfotransférases** ont été pour l'instant identifiées : la C6ST et la C6ST-2. Elles agissent principalement dans des régions riches en GlcA pour ajouter un sulfate en position C6 des GalNAc. Une 6-O-sulfotransférase spécifique du dermatane sulfate (D6ST) a cependant été identifiée chez le bœuf et l'existence d'un homologue humain pourrait expliquer les rares séquences de IdoA-GalNAc(6S) retrouvées dans des cellules issues de cordons ombilicaux humains (Nadanaka et al., 1999).

Enfin, une seule **hexuronyl CS/DS 2-O-sulfotransférase** (CS/DS2ST) a été mise à jour pour l'instant. Elle est capable de sulfater aussi bien les GlcA que les IdoA. Cependant l'ajout de 2-O-sulfate sur les IdoA est bien plus fréquent lorsque la GalNAc est 4-O-sulfatée, et les quelques rares GlcA(2S) peuvent être créés dans le seul cas où la GalNAc adjacente est sulfatée en position 6 (Silbert et Sugumaran, 2002).

### Une nouvelle forme de modification des GAGs

Un autre moyen d'affiner la composition structurale des GAGs a été récemment mis à jour. En effet, d'autres enzymes, les **endosulfatases**, qui agissent sur les cellules comme dans le milieu extracellulaire ont été découvertes (Ai et al., 2006). Elles devraient permettre de définir de manière encore plus fine le profil de sulfatation qui avait été défini lors de la synthèse dans l'appareil de Golgi. Ainsi, ces enzymes agissent sur les GAGs en supprimant certains sulfates et permettent de ce fait, de fonctionnaliser encore plus spécifiquement les chaînes oligosaccharidiques.

Les endosulfatases sont, pour l'instant, toutes rassemblées dans la famille des Sulfs. La première d'entre elles à avoir été étudiée est la QSulf1 (Dhoot et al., 2001). La QSulf1 est une 6-O-endosulfatase spécifique des HS qui agit sur des unités disaccharidiques du type IdoA2S-GlcANS6S ou GlcA-GlcNS6S, mais pas sur des IdoA-GlcNS6S (Ai et al., 2003). Une seconde enzyme de la même famille (QSulf2) a récemment été identifiée et semble avoir les mêmes spécificités que son homologue QSulf1 (Ai et al., 2006). La dénomination QSulf de ces enzymes vient du fait qu'elles furent clonées à partir de cDNA de caille (Quail en anglais). Leurs homologues chez l'Homme sont les HSulf-1 et 2 (Ai et al., 2006 ; Uchimura et al., 2006).

Récemment, Uchimura et ses collaborateurs ont montré qu'*in vitro*, les protéines VEGF, FGF-1, SDF-1 et SLC interagissaient bien moins avec de l'Hp préalablement traitée avec l'enzyme HSulf-2 alors que le FGF-2 se fixait de la même manière (Uchimura et al., 2006). La découverte de ces enzymes a été un élément important pour comprendre la spécificité de certaines interactions protéines-GAGs et par conséquent leurs fonctions.

## 2.4. La dégradation des GAGs

### 2.4.1. Introduction

Les GAGs sont présents à la surface de la plupart des cellules. Ils présentent une grande diversité de taille, de composition et de structure, ce qui leur permet d'exercer un éventail de fonctions diverses et variées. Leur structure est contrôlée et peut subir des modifications. Cela peut se traduire par des fragmentations partielles ou par une dégradation totale qui nécessite dans ce cas, un processus complet de recyclage des saccharides libérés. De ce fait, lorsque l'on parle de dégradation de GAGs, il est important de distinguer, d'un côté le remodelage fonctionnel des chaînes oligosaccharidiques (pour permettre une activité biologique précise), et de l'autre ce que l'on peut appeler recyclage ou "turnover" des GAGs.

### 2.4.2. La dégradation des GAGs comme un moyen supplémentaire de régulation

Comme nous venons de le voir, la biosynthèse des GAGs est complexe au regard du nombre important d'enzymes qui rentrent en jeu. Cela a l'avantage de réguler au mieux la structure de ces oligosaccharides et ainsi de multiplier les activités biologiques (portées par les GAGs) nécessaires à l'organisme dans un contexte normal ou pathologique.

Un autre moyen de réguler les interactions entre les GAGs et les protéines de la matrice extracellulaire est la dégradation de ces mêmes chaînes oligosaccharidiques. La dégradation des GAGs est possible grâce à l'action d'**hydrolases** qui peuvent être soit des **exoglycosidases**, soit des **endoglycosidases**.

Les endoglycosidases sont capables de couper les chaînes oligosaccharidiques tout au long de celles-ci, en sachant tout de même que la plupart de ces enzymes ont des séquences préférentielles de reconnaissance et de dégradation. Parmi les endoglycosidases, on distingue les kérotanases (I et II), les hyaluronidases (I et II) et les héparanases.

La famille des enzymes les plus étudiées à ce jour est sans doute celle des héparanases qui permettent la dégradation des HS. Trois héparanases ont été identifiées. La CTAP-III (connective tissue activating peptide III) qui est issue de la protéolyse d'une chimiokine exprimée lors d'un phénomène d'inflammation ou de cicatrisation (Proudfoot et al., 1997). L'héparanase Hpa 1 qui est exprimée dans le sous-endothélium (Dempsey et al., 2000) et surexprimée dans les cellules tumorales ou métastatiques d'un grand nombre de tissus (Vlodavsky et al., 1999 ; Hulett et al., 1999 ; Kussie et al., 1999). Et enfin l'héparanase Hpa 2

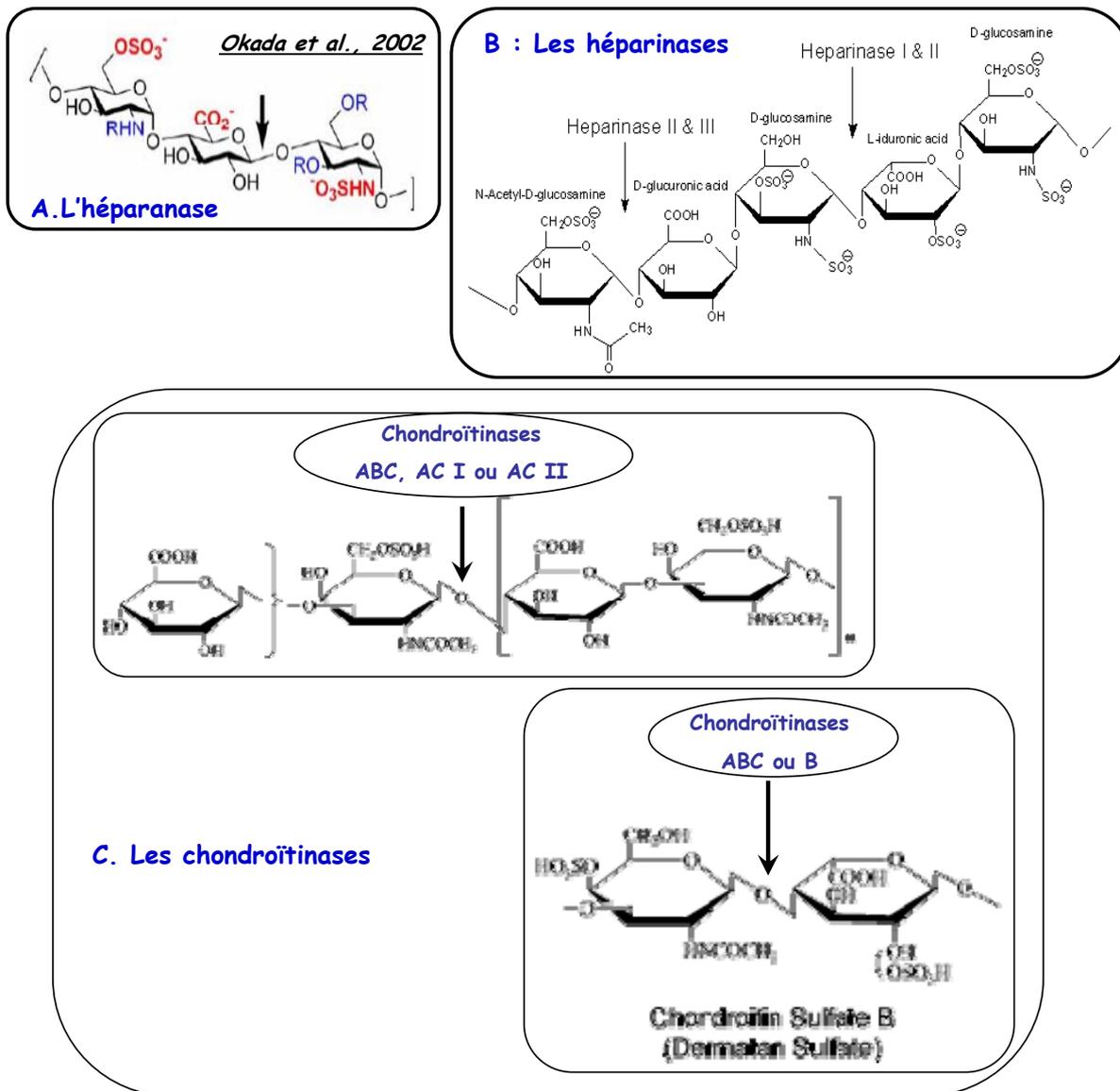
(McKenzie et al., 2000) qui ne serait pas exprimée dans les mêmes tissus que l'Hpa 1 mais dont la preuve de l'existence n'a pas pu être confirmée à ce jour, laissant donc perplexes de nombreux spécialistes.

Un des premiers rôles attribués aux héparanases localisées à l'extérieur des cellules est le remodelage des membranes basales après une blessure, par exemple, ou aux sites d'inflammation, pour favoriser la migration leucocytaire (Hoogwerf et al., 1995 ; Parish et al., 1998 ; Ircke et al., 1998 ; Dempsey et al., 2000). La dégradation des HS semblerait également jouer un rôle pour éviter l'agrégation de molécules de type  $\beta$ -amyloïde, avec les HSPGs, telle qu'elle est observée, en particulier chez les patients atteints de la maladie d'Alzheimer (Li et al., 1997). La fragmentation des HS dans les matrices extracellulaires, voire même dans les endosomes à l'intérieur des cellules, pourrait également faciliter le transport des protéines liées aux HS vers leurs sites d'action (Sperinde et Nugent, 1998 ; Tumova et al., 1999). Les HSPGs constituant une large "réserve" de facteurs de croissance, la dégradation des HS par l'Hpa peut permettre par un simple clivage, leur relargage, et par conséquent leur action à distance (Vlodavsky et al., 2006).

Deux pistes de dégradation non enzymatique des GAGs ont été récemment mises à jour. La première est due à la myéloperoxidase (MPO), qui est une enzyme exprimée par les phagocytes activés et qui catalyse la production d'une espèce chimique réactive : l'acide hypochloreux (HOCl). Cet HOCl va générer des dérivés de GAGs N-chlorés, qui au cours d'une homolyse réductive seront fragmentés (Rees et Davies, 2006). La seconde dégradation du même type est induite par l'éosinophile peroxydase (EPO) et la MPO qui catalysent la production d'acide hypobromique (HOBr). Cet HOBr va générer des dérivées de GAGs N-bromés qui vont se décomposer soit naturellement, soit en présence de  $Fe^{2+}$  ou  $Cu^+$  (Rees et al., 2007).

### **2.4.3. Le recyclage ou turnover des GAGs**

Les PGs ou les chaînes oligosaccharidiques qui doivent être recyclés, sont internalisés dans les cellules au sein d'endosomes. Les endoglycosidases vont alors fragmenter les chaînes oligosaccharidiques ce qui va permettre notamment de dissocier les protéines encore liées aux GAGs, de faciliter la dépolymérisation complète et, également, de supprimer les GAGs des cores protéiques. Ceux-ci sont alors recyclés dans l'appareil de Golgi afin d'être ornés de nouveaux profils oligosaccharidiques (Edgren et al., 1997). Les endosomes peuvent ensuite fusionner avec les lysosomes au sein desquels la dépolymérisation complète des GAGs sera



**Figure 5: Enzymes utilisées couramment en laboratoire pour couper les GAGs**

**A :** Structure minimale d'HS pour que l'héparanase (produite chez l'homme) puisse couper entre une GlcA et une GlcNS. Les groupes nécessaires pour la reconnaissance du trisaccharide sont en rouge. L'affinité de l'enzyme pour cette séquence est meilleure lorsque  $R=SO_3^-$

**B :** Les sites de coupures spécifiques des enzymes du type héparinase (non produites chez l'homme) vis-à-vis de l'héparine ou des héparanes sulfates. Les héparinases coupent les liaisons glycosidiques  $\alpha(1-4)$  entre la GlcA et un acide uronique (IdoA ou GluA). Plus spécifiquement, l'héparinase I coupe l'Hp et les HS sur la liaison entre une hexosamine et un acide uronique O-sulfaté. Cette enzyme coupe également le pentasaccharide spécifique de l'ATIII. L'héparinase II coupe les HS, et moins spécifiquement l'Hp sur la liaison entre une hexosamine et un acide uronique. L'héparinase III coupe la liaison entre l'hexosamine et l'acide glucuronique des HS.

**C :** La chondroïtinase ABC coupent les chondroïtines sulfates A, B, ou C sur la liaison entre l'hexosamine et un acide uronique (iduronique ou glucuronique). La chondroïtinase AC I catalyse l'élimination de la liaison entre une N-acétyl hexosamine et un acide glucuronique et non un acide iduronique. Elle est fonctionnelle sur les CS-B (DS). La chondroïtinase AC II catalyse le clivage de la liaison entre une N-acétyl galactosamine et un acide glucuronique, mais en revanche ne fonctionne pas sur le CS-B. La chondroïtinase B coupe par élimination les chaînes CS-B (DS) sur la liaison 1-4 entre les hexosamines et les acides iduroniques. Le 2-O-sulfate sur l'acide iduronique est favorable à l'action de cette enzyme.

achevée grâce à une variété d'exoglycosidases agissant par l'extrémité non-réductrice des fragments oligosaccharidiques. Parmi les exoglycosidases, peuvent être nommées la  $\beta$ -glucuronidase, l' $\alpha$ -L-iduronidase, la  $\beta$ -galactosidase, l' $\alpha$ -N-acétylglucosaminidase, et la  $\beta$ -N-acétylhexosaminidase qui clivent de manière très spécifique les polysaccharides.

Le recyclage des GAGs est un mécanisme essentiel dans la physiologie normale de la cellule. Ceci s'illustre notamment et malheureusement chez l'Homme lors de la mucopolysaccharidose (MPS). La MPS est une maladie génétique rare, autosomique et récessive qui se traduit par des manifestations pathologiques progressives dans la plupart des organes. Cette maladie est due à une anomalie du gène codant pour certaines enzymes lysosomiales. Les cellules sont alors incapables de produire l'enzyme ou la produisent en trop faible quantité et ne sont donc pas en mesure de cataboliser les chaînes de GAGs, qui s'accumulent pour provoquer finalement des lésions cellulaires, tissulaires et organiques irréversibles. Il existe au moins douze formes identifiées de la maladie selon la nature de l'enzyme déficiente. La plus connue est la MPS I, liée à l'absence de la  $\alpha$ -L-Iduronidase (Ashworth et al., 2006).

#### **2.4.4. Les enzymes de dégradation utilisées au laboratoire**

Mises à part les enzymes naturellement exprimées chez les eucaryotes, d'autres enzymes capables de dégrader les GAGs ont été découvertes chez des bactéries. Celles-ci sont pour la plupart commercialisées puisque très utiles dans les expériences de caractérisation des GAGs et sont classées suivant leur spécificité d'action. Les hyaluronidases permettent ainsi de dégrader l'AH. Les héparinases I, II et III, ainsi que la K5 lyase sont utilisées pour dégrader les Hp et HS. Et finalement, les chondroïtinases ABC, ACI, ACII, B sont nécessaires pour la digestion des CS et DS.

Ces enzymes reconnaissent des séquences oligosaccharidiques particulières (Figure 5). Cela permet, dans le cadre du laboratoire, de les utiliser pour isoler des sous-domaines structuraux dans les chaînes de GAG.

	<b>NOM</b>	<b>Type de GAG</b>
<b>PG membranaires</b>	Syndécans Glypicans Bétaglycan Les isoformes de CD44	Principalement HS et CS Principalement HS Principalement HS HS ou CS
<b>PG de la matrice extracellulaire</b>	Collagènes XVIII et XV Agrine Bamacan Famille du Perlecan Versican (265-370 kDa) Aggrécan (env. 220 kDa) Neurocan (env 136 kDa) Brévican (env. 100 kDa) Perlécan Decorine Biglycan Lumican Fibromoduline	Principalement CS (10-30) CS et KS (env. 100) CS (3-7) CS (1-3) Principalement HS DS DS ou CS KS KS
<b>PG intracellulaires</b>	Serglycine	HS ou CS
<b>PG circulants</b>	Bikunine PG100 ou MCSF1 ApoO Endocan	CS Unique chaîne de CS CS Unique chaîne de DS

**Figure 6: Classification des protéoglycannes (PGs) en fonction de leur localisation**

Le tableau classifie de manière non exhaustive les PG qui peuvent être trouvés dans l'organisme. Leur localisation est indiqué ainsi que leur nom et la nature de leur(s) chaîne(s) glycanique(s). Dans certains cas des informations sur le poids moléculaire et le nombre de chaînes glycanique sont indiquées.

### 3. LES CORES PROTEIQUES

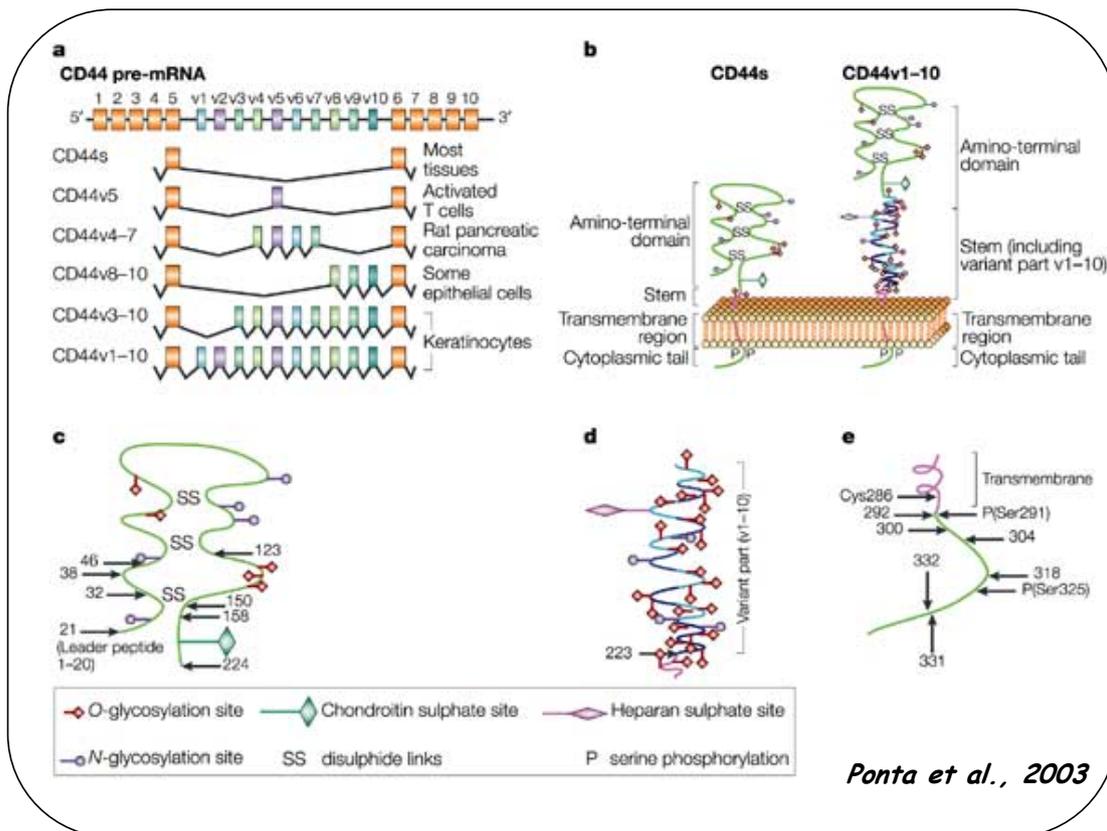
#### 3.1. Introduction

La partie protéique ou core protéique des protéoglycannes constitue la structure de base sur laquelle sont ajoutés les GAGs. Comme pour la majorité des protéines, l'expression des cores protéiques est régulée en fonction des besoins des cellules ou de leur environnement. Ils permettent, de par leur composition polypeptidique, l'initiation de la synthèse de GAGs qui leur sont spécifiques. D'une manière générale, ils guident la localisation spatio-temporelle des PGs. Les PGs sont généralement classés en quatre grandes familles : les PGs intracellulaires, les PGs membranaires (ou péri-cellulaires), extracellulaires et, depuis peu, les PGs circulants (Figure 6).

#### 3.2. Les protéoglycannes intracellulaires

Le protéoglycanne intracellulaire le plus connu est **la serglycine**. La serglycine est dite intracellulaire car elle est majoritairement localisée dans des vésicules de type granules de sécrétion. Elle possède le plus petit core protéique (10 à 15 kDa) de tous les PGs identifiés. Il comporte ainsi huit séquences peptidiques de type Ser/Gly consécutives (chez l'Homme) sur lesquelles se lient en fonction du type cellulaire, des chaînes de CS, d'HS ou d'héparine. La serglycine a été localisée dans différents types de cellules telles que les cellules hématopoïétiques, les cellules endothéliales ou encore les cellules embryonnaires. Selon le type cellulaire, son expression est soit constitutive, soit induite. La fonction de la serglycine reste encore bien mal connue. Cependant grâce à ses chaînes de GAGs, elle pourrait être utilisée pour constituer une sorte de "réserve" de protéines telles que des cytokines, des facteurs de croissance ou même des protéases. Ces protéines sont ainsi protégées de toute protéolyse jusqu'au moment où leur sécrétion extracellulaire par les vésicules devient nécessaire (Schick et al., 2001 ; Schick et al., 2003 ; Humphries et al., 1999).

La serglycine est le seul PG qui possède des chaînes d'Hp et ce, lorsqu'elle est exprimée dans les granules de sécrétions de mastocytes (Yurt et al., 1977).



**Figure 7: Les transcrits et protéines issus de l'épissage alternatif du CD44**

- Le CD44 est codé par un gène qui possède 20 exons dont 10 (appelés variant ou "v") sont sensibles à l'épissage alternatif. Le plus petit isoforme de CD44 ou CD44 standard (CD44s) est exprimé de façon ubiquitaire alors que le variant le plus long est exprimé uniquement dans quelques tissus épithéliaux.
- La structure du CD44s est comparée avec celle du plus long des isoformes. La différence entre les deux se fait uniquement au niveau de la "tige" (stem) de la protéine.
- Représentation du domaine N-terminal des isoformes de CD44. Ce domaine permet l'interaction avec l'acide hyaluronique (grâce aux a.a. 32-123). Pour la stabilité de ce module, 6 cystéines formant 3 ponts disulfures semblent être primordiales pour le bon repliement du domaine et ce quelque soit les isoformes. A noter que chez tous les isoformes, une chaîne de chondroïtine sulfate peut être ancrée via ce domaine.
- Le CD44s est composé de 46 a.a. Selon les isoformes il peut atteindre 381 a.a. comme dans le cas du CD44v1-10. L'unique site d'attachement d'une chaîne d'HS se trouve être sur un fragment protéique codé par l'exon V3 et ne sera donc pas présent sur tous les isoformes et notamment le CD44s.
- Domaine cytoplasmique correspondant à la région C-terminale de la protéine. Ce domaine permet la fixation de protéines aux fonctions importantes dans l'organisation du cytosquelette et dans la signalisation. L'erzin, radixin moesin (ERM) se fixent par exemple aux a.a. 292-300. Les a.a. 331 et 332 sont eux nécessaires à la distribution basolatérale dans les épithéliums polarisés.

(D'après Ponta et al., 2003)

### 3.3. Les protéoglycannes membranaires

Il existe quatre grandes familles de PGs membranaires, le CD44, le bétaglycan, les glypicans et les syndécans.

#### 3.3.1. Le CD44

Le core protéique du CD44 est une protéine transmembranaire de classe I. Le CD44 est impliqué dans de nombreux processus cellulaires tels que la différenciation, la mobilité, la croissance et la survie. Même s'il existe un seul gène codant pour le CD44, l'épissage alternatif des 20 exons qui le composent, permet de parler de la famille du CD44. Ainsi du plus petit isoforme à celui qui a subi le moins d'épissage, nous trouverons le CD44s, les CD44v5, v4-7, v8-10, v3-10 et v1-10 (Figure 7a) (Ponta et al., 2003).

L'isoforme CD44 standard (CD44s) qui est celui exprimé dans la plupart des tissus, est composé de trois domaines. Le domaine C-terminal cytoplasmique de 70 acides aminés, le domaine transmembranaire (23 aa.) et la partie N-terminale qui est extracellulaire. Le domaine N-terminal peut être divisé en deux régions. Une première région située à l'extrémité (exons 1 à 5) qui est largement conservée entre les espèces, et une seconde (exons 6 à 10) proche de la membrane, bien moins conservée, où sont liées les chaînes de CS. Entre ces deux régions extracellulaires, les dix exons (v1 à v10) qui subissent un épissage différentiel, peuvent s'insérer entre les exons 5 et 6 (Figure 7a) (Martin et al., 2003). Certains de ces exons peuvent permettre le greffage de chaînes HS.

La régulation et la localisation des CD44 ont été largement étudiées. Il est connu que les cytokines ont beaucoup d'influence sur le taux et le lieu d'expression du CD44. L'interféron gamma par exemple est capable de diminuer le taux d'expression dans les cellules tumorales de pancréas ou au contraire d'augmenter celui de CD44v10 dans les carcinomes du côlon (Ringel et al., 2001 ; Wimmenauer et al., 1999).

Les CD44 sont exprimés dans une large gamme de cellules telles que les cellules hématopoïétiques (lymphocytes T, lymphocytes B, monocytes, granulocytes, érythrocytes), certaines cellules épithéliales (kératinocytes, chondrocytes, cellules mésothéliales), certaines cellules endothéliales et neurales. Les deux isoformes les plus représentés sont le CD44s exprimé principalement sur les cellules hématopoïétiques, les leucocytes et même les fibroblastes, et ensuite le CD44H, qui est lui, plutôt associé aux cellules épithéliales notamment en cours de division (Brown et al., 1991). Ces deux isoformes ont clairement des

fonctions différentes. Par exemple, le CD44s peut se fixer à l'AH et permettre l'infiltration lymphocytaire, alors que le CD44H ne le peut pas (Picker et al., 1989 ; Liao et al., 1993).

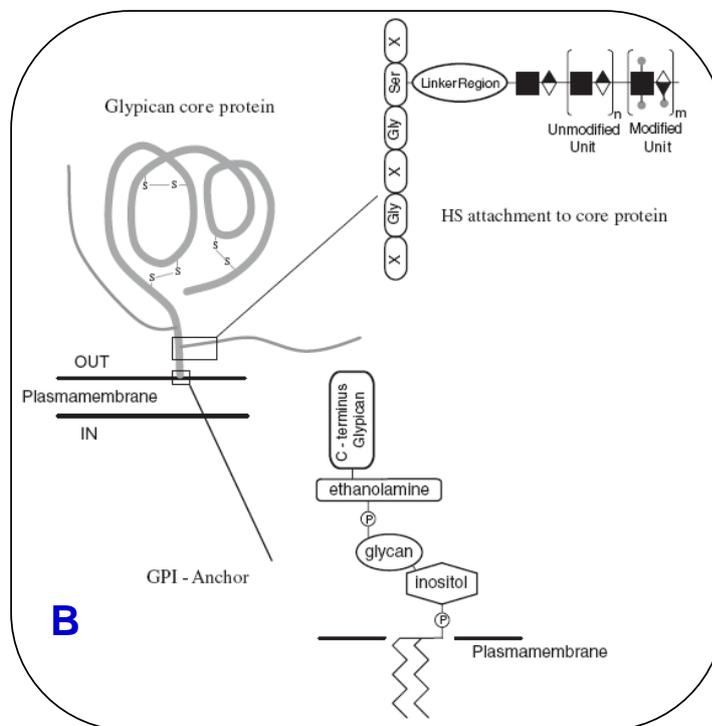
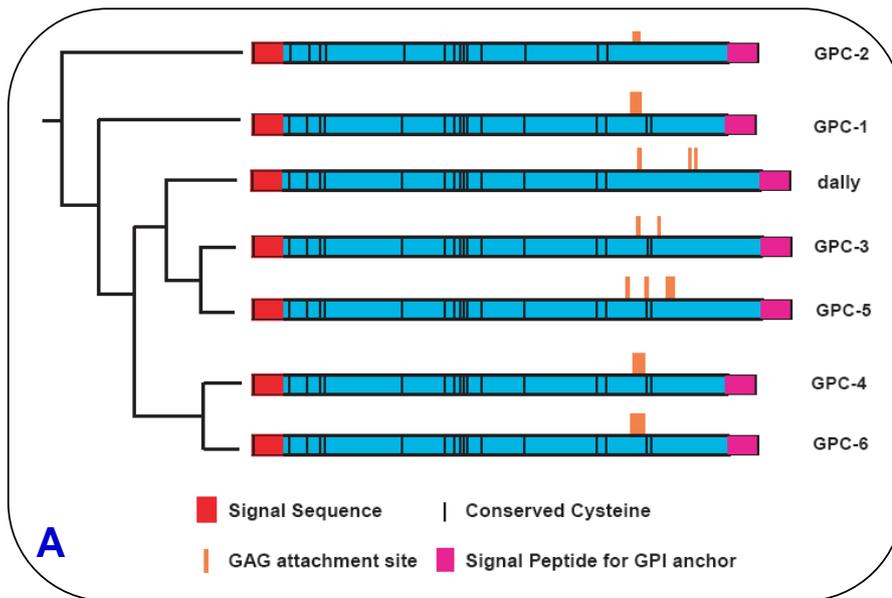
Les CD44 ont de nombreuses fonctions, dont beaucoup sont encore connues de façon approximative. La première fonction qui lui a été attribuée et qui est bien sûr la plus évidente, est celle de récepteur de l'acide hyaluronique, justifiant ainsi l'appartenance des CD44 à la famille des hyaladérines (Toole, 1990).

En tant que protéoglycane, les CD44 qui possèdent les chaînes de CS sont capables de fixer des composants de la matrice extracellulaire comme le collagène ou la fibronectine. En revanche, le CD44v3 qui porte une chaîne d'HS peut fixer des facteurs de croissance de type FGF-2, HB-EGF (Bennett et al., 1995 ; Jackson et al., 1995) ou encore l'HGF/SF (Van der Voort et al., 1999).

Les fonctions du CD44, trop nombreuses pour être toutes détaillées ici (récepteurs de molécules d'adhésion, régulateurs des signaux de transduction, etc.) peuvent être exercées aussi bien par la partie protéique que par les chaînes de GAGs (de type CS et parfois HS) voire par la coopération des deux, et ont été clairement détaillées dans les revues suivantes Ponta et al., (2003) et Martin et al., (2003).

### **3.3.2. Le bétaglycan**

Le bétaglycan est un protéoglycane transmembranaire de type I. Il a été découvert tout d'abord comme un récepteur du Transforming Growth Factor bêta (TGF- $\beta$ ) (Massagué, 1985) et a été appelé, récepteur TGF- $\beta$ -type III au regard de la forte affinité de son core protéique pour le TGF- $\beta$ . La partie extracellulaire du bétaglycan possède un domaine de type endogline (E-domaine) et un domaine de type uromoduline (U-domaine). L'inhibine A qui est une hormone folliculo-stimulante, est capable de se fixer avec une forte affinité sur le U-domaine alors que le TGF- $\beta$  va lui interagir avec les deux domaines. Le bétaglycan est considéré dans le cas de l'inhibine A comme un co-récepteur (Bernard et al., 2002). Sur la majorité des cellules, le bétaglycan est le récepteur principal du TGF- $\beta$ . L'une des fonctions associées à ce PG est de présenter le TGF- $\beta$  au récepteur de type II par lequel il transmet son signal (Wang et al., 1991). Les fonctions associées au GAGs (CS et HS) qui peuvent être présents sur le core protéique ne sont pas encore bien définies.



**Figure 8: La famille des glypicans**

**A** : Diagrammes montrant l'organisation des différents domaines des cores protéiques du glypican. La localisation des sites putatifs d'attachement des GAGs sont en orange. Les glypicans 1 à 6 sont ceux retrouvés chez l'homme, le dally étant l'homologue chez la drosophile. Le schéma montre la conservation des 14 cystéines réparties de la même façon pour tous les glypicans. Le dendrogramme suggère que le glypican-3 /-5 et le glypican-4 /-6 forment des sous-familles (d'après *Bernfield et al., 1999*).

**B** : Représentation schématique des glypicans. Le core protéique est localisé totalement dans l'espace extracellulaire. Il possède plusieurs ponts disulfures qui permettent l'organisation de la partie protéique et en font un domaine globulaire compact. Les chaînes d'HS sont covalamment liées à des résidus sérines dans des régions consensus se trouvant non loin de la surface cellulaire. Le core protéique est lié à la membrane plasmique par le carboxyl c-terminal via une ancre GPI. Le composant "glycan" de l'ancre GPI est une trimannose-glucosamine (d'après *De Cat et David 2001*).

### 3.3.3. Les glypicans

Les glypicans, initialement découverts chez la drosophile, (GPC1 à 6) sont une famille de six cores protéiques connus pour l'instant chez l'Homme, issus de gènes distincts. Ils sont membranaires grâce à leur liaison à une ancre lipidique de type glycosyl phosphatidyl inositol (GPI). Tous ont en commun i) une séquence signal à l'extrémité N-terminale qui va permettre la sécrétion, ii) une région d'environ 50 kDa qui contient un profil caractéristique de 14 cystéines très conservées, iii) une région proche de la membrane plasmique composée de séquences d'attachement de GAGs de type Ser-Gly (2 à 4), iv) une extrémité C-terminale hydrophobe impliquée dans la liaison au GPI (Figure 8A et 9bis). La région comportant les 14 résidus cystéines fortement conservée laisse penser qu'une partie de la protéine adopte une conformation qui doit être spécifique à la famille des glypicans, (importante par exemple pour des interactions de type protéine-protéine), riche en ponts disulfures et donc plutôt globulaire (Figure 8B et 9bis). Le domaine moins bien conservé proche de la membrane plasmique apporte sans doute la spécificité de chaque glypican et cela notamment pour la composition en chaînes oligosaccharidiques qui sont exclusivement des HS (De Cat et David, 2001), à l'exception du glypican-5 qui peut posséder des chondroïtines sulfates (Saunders et al., 1997). Il a été démontré par Chen et Lander (2001) que la sélectivité de la biosynthèse d'HS plutôt que de CS sur la partie protéique du glypican-1, est due à une séquence de la région globulaire du glypican, située à 70 aa des sites d'attachement des GAGs.

La caractéristique essentielle qui définit ce type de protéoglycane est l'ancre GPI. Cette dernière pourrait permettre la localisation préférentielle du glypican dans des microdomaines membranaires appelés rafts, et permettre par conséquent une fonction locale (Simmons et Toomre, 2000). A l'inverse, la localisation des glypicans en dehors de ces rafts serait défavorisée (Simmons et Toomre, 2000). Cependant, les rafts sont la plupart du temps localisés sur la surface apicale des cellules polarisées, alors que par exemple, le glypican-1 est décrit comme majoritairement présent à la surface basolatérale de cellules épithéliales. De manière intéressante, Mertens et al. (1996) ont montré que l'absence d'HS sur le glypican-1 permettrait une localisation apicale de ce PG, alors que la présence de ces GAGs favoriserait sa localisation du côté basolatéral. Cette donnée est tout à fait importante pour la compréhension des fonctions des glypicans. En effet, comme l'ancre GPI favorise les phénomènes d'internalisation et de re-externalisation, l'action d'enzymes telle que l'héparanase pourrait influencer sur la localisation du core protéique du glypican et ainsi



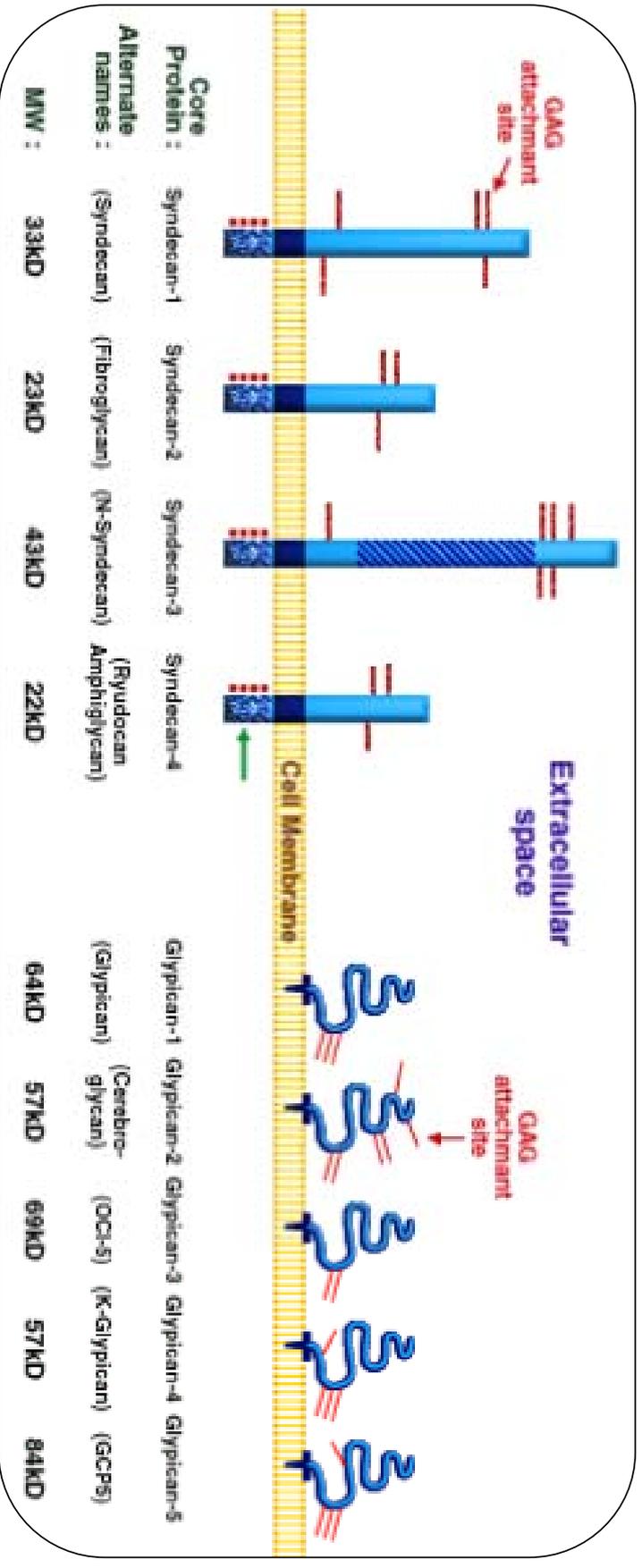
participer aux phénomènes de transcytose (phénomène permettant à une molécule de traverser une cellule) (Mertens et al., 1996).

L'ancre GPI se révèle également très importante pour le catabolisme ou "turnover" des glypicans. Grâce à elle, l'endocytose, le transport jusqu'aux lysosomes et la dégradation complète du PG ne prennent au total pas plus de trente minutes. Les glypicans peuvent être internalisés avec leurs ligands, pour en être ensuite séparés. Cela est un phénomène important dans l'internalisation ou le recyclage de protéines liant les HS. Par ailleurs, lorsque les glypicans sont internalisés, leur core protéique peut arborer de nouveaux types d'HS ayant une composition structurale différente. De cette manière, la cellule peut dans un contexte particulier internaliser plus spécifiquement et plus facilement certains types de protéines (revue de Bernfield et al., 1999). De Cat et David (2001) émettent l'hypothèse que les glypicans pourraient être sélectifs quant à la taille des ligands avec lesquels ils interagissent. Ils expliquent que la "tête globulaire" de large taille pourrait constituer une sorte de filtre ne laissant passer que les plus petites molécules, libres d'interagir avec les chaînes d'HS.

D'une manière générale, les glypicans sont exprimés préférentiellement durant les phases de développement embryonnaire, et leurs niveaux d'expression varient en fonction du stade de développement et du type cellulaire. Chez l'Homme, une mutation sur le gène codant pour le glypican-3 est responsable du syndrome de Simpson-Golabi-Behmel (SGBS) qui se manifeste par une croissance excessive dès la naissance (grosse langue, faciès grossier, anomalies rénales) (DeBaun et al., 1991).

#### **3.3.4. Les syndécans**

Les syndécans sont une famille de PG composée de quatre protéines qui sont nommées de syndécans 1 à 4 et qui sont issues de gènes distincts (Figure 9 et 9bis). Chaque syndécan est de nature transmembranaire, dite de type I et comporte : i) une région C-terminale intracellulaire fortement conservée, divisée en trois sous-régions qui sont les domaines C1 et C2 identiques chez tous les syndécans et le domaine V, variable selon le type de syndécan, qui se trouve donc entre les C1 et C2 ; ii) un domaine transmembranaire conservé ; iii) un domaine extracellulaire (encore appelé ectodomaine) qui est clairement la région la moins conservée entre tous les syndécans. Il comporte cependant des séquences consensus telles que les acides aminés Ser-Gly qui permettent l'attachement des GAGs, ou encore des domaines sensibles à la protéolyse ou d'autres permettant l'oligomérisation ; iv) une séquence signal en position N-terminale (Bernfield et al., 1999).



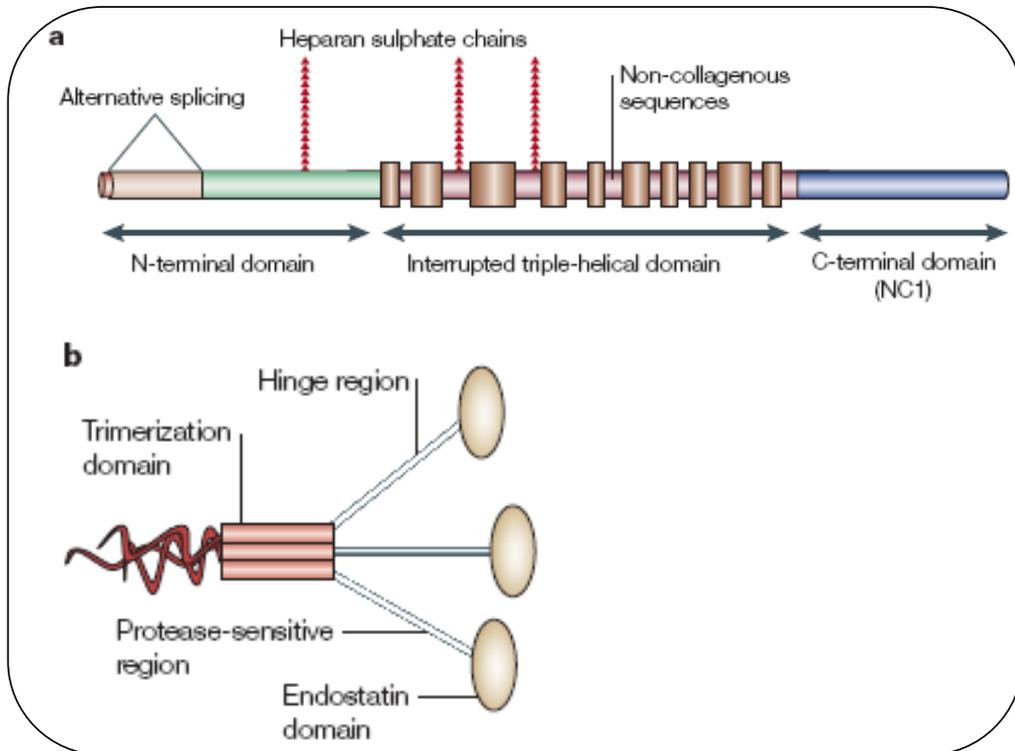
**Figure 9bis : Les syndécans et les glypicans sont les deux grandes familles d'HSPG de la surface cellulaire**  
 Représentation des différents membres de la famille des syndécans et de celle des PG membranaires à ancre GPI, les glypicans. En rouge sont indiqués les sites d'attachement de GAGs potentiels ou identifiés. Pour les syndécans, sont représentés : le domaine transmembranaire (bleu foncé), le domaine intracellulaire (bleu à relief) avec les tyrosines conservées (points rouges). Pour les glypicans, sont représentés le core protéique (bleu) et l'ancre GPI (bleu foncé). Le nom de chacun des membres des deux familles est indiqué ainsi que le nom alternatif et le poids moléculaire du core protéique.  
 D'après Tetsuhito Kajima (Nagoya University, School of Medicine)

Les différences structurales entre les syndécans sont le résultat de variations dans le nombre, le type et la composition des chaînes de GAGs (Bernfield et al., 1999). Même si le nombre et le type de GAG sont en général spécifiques de chaque syndécan, la structure fine de la chaîne oligosaccharidique sera plutôt dépendante de la cellule et de son état d'activation plutôt que de la nature du core protéique du syndécan (Carey, 1997).

Chaque syndécan possède des chaînes d'HS liées sur l'ectodomaine. Il est cependant assez fréquent que des chaînes de type CS soient greffées également sur le core protéique (Uéno et al., 2001).

Selon le stade de développement embryonnaire ou encore les conditions physiologiques, les taux et les lieux d'expression des syndécans sont très variables. Cependant, chaque syndécan apparaît être préférentiellement exprimé dans certains types cellulaires. Le syndécan-1 sera par exemple trouvé sur les cellules épithéliales et les cellules circulantes malignes. Le syndécan-2 est présent sur les cellules endothéliales et les fibroblastes, alors que le syndécan-3 serait plutôt exprimé dans les cellules du système nerveux. Enfin le syndécan-4, moins sélectif, est exprimé sur la plupart des cellules (Götte, 2003).

Les fonctions des syndécans semblent par ailleurs assez diversifiées. Ils représentent la forme majeure d'HSPG à la surface des cellules, et agissent comme co-récepteurs, afin de localiser, réguler, et activer bon nombre facteurs de croissance, de cytokines et de chimiokines. Cependant, les syndécans semblent impliqués dans d'autres processus biologiques. Dans certains cas, les syndécans s'oligomérisent, offrant une couverture en HS de la surface cellulaire plus importante (Bernfield et al., 1999). Ils peuvent également jouer le rôle d'un récepteur "classique" capable de transmettre un signal. Par exemple, la région cytoplasmique du syndécan-4 est capable de fixer le phosphatidylinositol 4,5-bisphosphate (PIP2) et d'activer une enzyme : la protéine kinase C alpha (PKC $\alpha$ ). Cette région peut également interagir avec des protéines de signalisation ou de structures telles que CASK (Calcium/calmodulin dependent serine protein kinase), FAK (Focal Adhesion Kinase), la syndesmose et la paxilline. Ces interactions sont certainement liées (d'un point de vue de la co-localisation et de la signalisation) à la formation de complexes macromoléculaires incluant le syndécan-4, la fibronectine et les intégrines  $\alpha 5\beta 1$ . Ce complexe semble nécessaire à la formation des plaques focales d'adhésion et au réarrangement de l'actine en fibres de stress. Les syndécans semblent donc étroitement impliqués dans les processus d'adhérence cellulaire (Beauvais et Rapraeger, 2004 ; Saoncella et al., 1998) et en partie dans l'organisation du cytosquelette (Yoneda et Couchman, 2003 ; Woods, 2001).



**Figure 10: Domaines structuraux du collagène XVIII et de son domaine NC1/endostatine**

**a:** Schéma de la chaîne  $\alpha 1$  du collagène XVIII (La chaîne  $\alpha 1$  du collagène XV étant structurellement homologue). Ils appartiennent à la sous-famille de collagène dite des multiplexines du fait de leurs domaines en triple hélice (orange) interrompus par des séquences non-collagéniques. Ils contiennent un domaine N-terminal non-collagénique qui peut subir un épissage alternatif chez le collagène XVIII, et un domaine C-terminal (NC1) non collagénique.

**b:** Le domaine NC1 du collagène XVIII contient trois sous-domaines distincts : un domaine de trimérisation impliqué dans l'assemblage non covalent de l'homotrimère (rouge); un domaine intermédiaire (vert) sensible à la protéolyse, et un domaine C-terminal aux propriétés angiogéniques : l'endostatine (jaune).

(d'après Iozzo 2005)

L'ensemble de ces interactions protéine-protéine, la formation de complexes macromoléculaires, et de ces oligomérisations avec les molécules "charpentes" de la matrice extracellulaire et du milieu intracellulaire sont cependant encore mal comprises. De plus en plus d'indications viennent néanmoins suggérer que la formation de ces complexes est utile pour localiser les GAGs et par conséquent pour permettre l'action de protéines dont l'activité biologique est dépendante des GAGs.

### **3.4. Les protéoglycannes extracellulaires**

#### **3.4.1. Les protéoglycannes de la membrane basale**

##### **3.4.1.1. Les collagènes XVIII et XV**

Sur l'ensemble des collagènes identifiés, cinq se trouvent être des protéoglycannes. Les collagènes de type IX, XII et XIV appartiennent à la sous-famille de collagène dite FACIT (Fibril Associated Collagen with Interrupted Triples helix) localisés principalement dans les tissus conjonctifs et ornés pour la plupart de chaînes CS (Li et al., 2000).

Les collagènes XVIII et XV quant à eux, sont également des molécules hybrides : à la fois collagène et protéoglycane. Ils font partie de la sous-famille de collagènes des multiplexines. Leur structure est décrite sur la Figure 10. Le collagène XVIII porte préférentiellement des chaînes HS même si dans certains cas, *in vitro*, des chaînes de CS ont été identifiées (Dong et al., 2003). Le collagène XV se différencie de son homologue en portant des chaînes CS. Le collagène XVIII aurait des activités associées à certaines fonctions oculaires. En revanche, la déficience en collagène XV est associée plus facilement à des désordres cardiaques ou des muscles squelettiques (Iozzo, 2005). D'une manière générale, peu de choses sont connues sur l'utilité des chaînes de GAG des collagènes même si la régulation de l'activité de facteurs de croissance ou des phénomènes d'adhérence cellulaire peuvent être supposés, comme notamment dans la progression de la maladie d'Alzheimer (Van Horssen et al., 2002).

##### **3.4.1.2. L'agrine**

L'agrine est une protéine de plus de 2000 acides aminés qui comporte environ 400 kDa de glycosylation, principalement du côté N-terminal, dont trois chaînes d'HS ce qui la classe dans la famille des HSPGs (Tsen et al., 1995). Un épissage alternatif permet de donner

naissance à deux isoformes d'agrine dont l'un est une protéine transmembranaire de type II (TM-agrin) et l'autre est extracellulaire. Cette différence de localisation leur donne à chacune des fonctions très distinctes (Burgess et al., 2000).

Dans les jonctions neuromusculaires, l'agrine permet le rassemblement des récepteurs à l'acétylcholine, (AChRs) grâce à la région C-terminale du core protéique et elle participe ainsi à la formation de l'appareil post-synaptique (Bezakova et al., 2003). Dans le cerveau et plus précisément dans le ganglion cervical supérieur, l'agrine est importante pour réguler la formation des synapses cholinergiques interneurales. Dans les synapses immunologiques, sur les lymphocytes T activés, l'agrine semble également être une des molécules responsables de l'agrégation de rafts contenant par exemple, des complexes CD3/T-cell récepteurs, complexes qui ne se forment que si l'agrine n'est pas glycosylée.

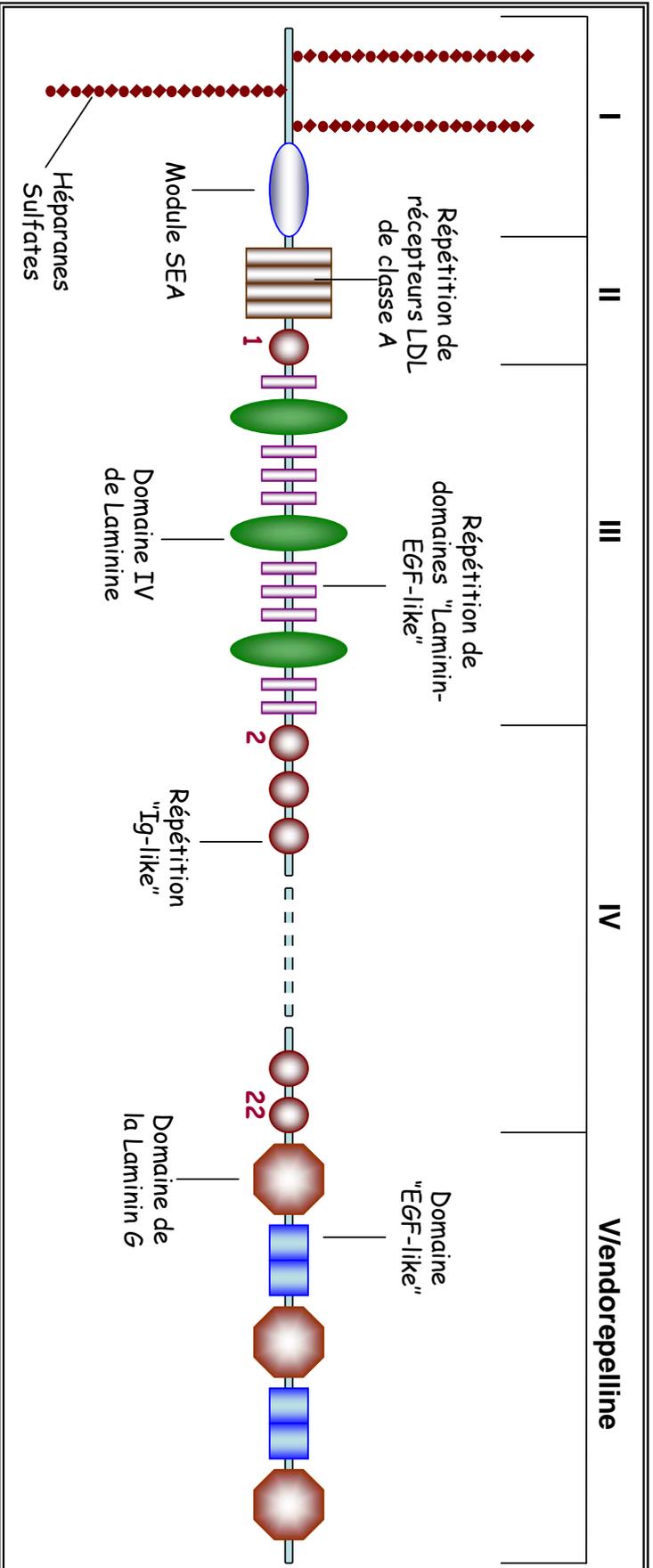
Bezakova et ses collaborateurs proposent également que les glycosylations permettent de réguler les interactions protéine-protéine sur le core protéique, et pourraient ainsi influencer l'orientation des protéines impliquées dans la reconnaissance de type cellule-cellule comme dans le cas de l'agrine déglycosylée qui est impliquée dans l'activité des synapses immunologiques. Les glycosylations ont sans aucun doute une influence sur l'agrégation des AChRs (Bezakova et al., 2003).

#### **3.4.1.3. Le bamacan**

Historiquement, le bamacan (BASment Membrane-Associated chondroitin protéoglyCAN) est le troisième PG des membranes basales à avoir été découvert, après le perlécan et l'agrine. A l'exception du collagène XV, le bamacan est le seul PG des membranes basales, connu jusqu'à présent pour posséder des chaînes CS. Son core protéique est constitué de cinq domaines où les GAGs peuvent être ajoutés sur les domaines II et V (Iozzo, 1998).

Le bamacan est exprimé tardivement durant l'organogénèse. Il jouerait un rôle dans la stabilisation de la membrane basale. Sa distribution est fortement perturbée lorsque la stabilité de la membrane basale est compromise (Ehara et al., 1994).

De façon surprenante, le bamacan est également une protéine intracellulaire de la famille des SMC3 (Structural Maintenance of Chromosome 3) (Ghiselli et al., 1999). Sa dérégulation semble être également étroitement liée à l'oncogénèse (Ghiselli et Iozzo, 2000).



**Figure 11: Domaines structuraux du perlecan humain**

Les cinq domaines de la partie protéique de perlecan (indiqués en chiffre romain) ont été schématisés. Ils contiennent des modules et répétitions qui sont régulièrement retrouvés chez les protéines de la matrice extracellulaire ou de la surface cellulaire impliqués dans l'adhésion ou la croissance cellulaire. Le domaine I contient trois séquences consensus d'attachement des HS : Ser-Gly-Asp et un module SEA qui favorise la glycosylation. Un autre site potentiel d'attachement des HS dans le domaine V n'a pas été schématisé. Dans le domaine II se trouvent quatre répétitions semblables au récepteur LDL (Low density lipoprotein) de classe A suivi par une répétition immunoglobuline (Ig)-like. Le domaine III est lui composé de trois domaines globulaires homologues au domaine IV de la laminine entre lesquels neuf domaines "laminin-EGF (Epidermal Growth Factor)-like" se répètent. Le domaine IV s'organise autour de 21 domaines Ig-like de façon très similaires à ce qui est observé chez la protéine N-CAM (neural cell adhesion molecule). Enfin le domaine V ou endorepelline contient trois domaines globulaires homologues à la laminine G séparés par deux répétitions d'EGF-like.

(d'après Tozzo, 2005)

#### 3.4.1.4. La famille du perlécan

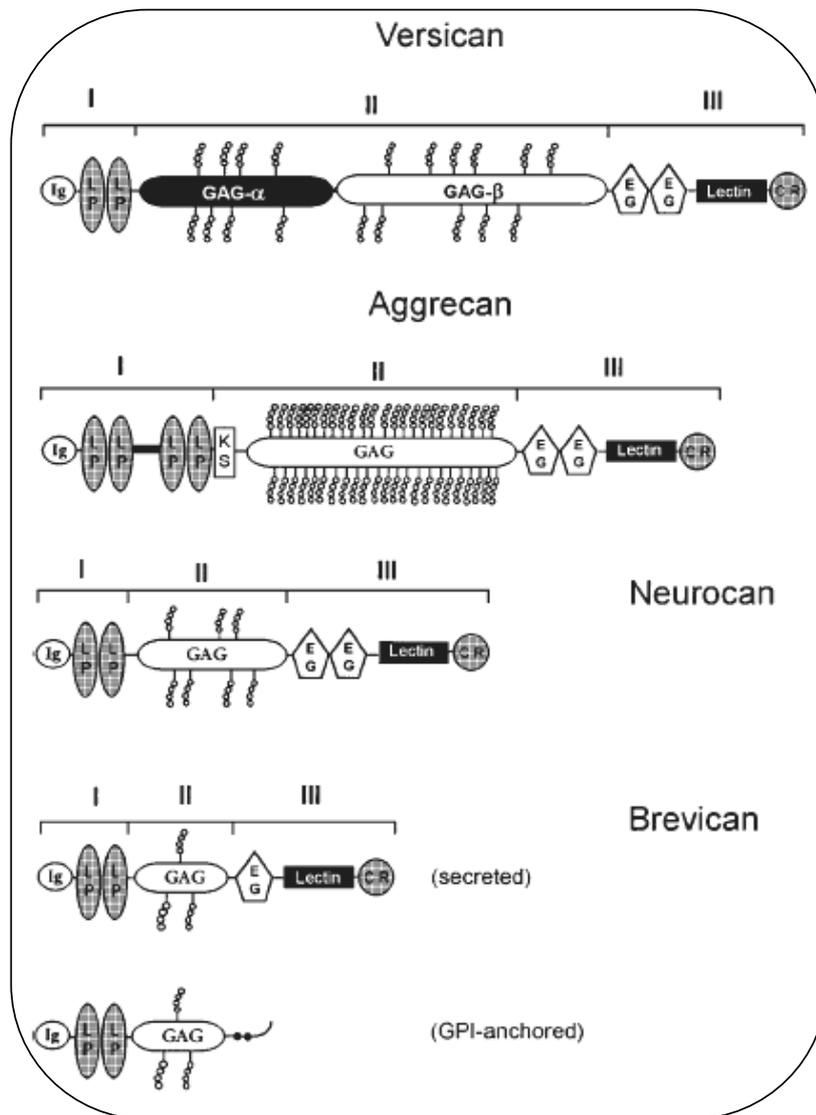
Le perlécan tient son nom de son apparence semblable à un collier de perles lorsqu'il est observé en microscopie électronique. Cette protéine modulaire est l'une des plus grandes chez les vertébrés (800 kDa) et les invertébrés (470 kDa) (Iozzo, 2005). Le perlécan possède cinq domaines (Figure 11). Il fait partie de la famille des HSPGs car il comporte trois chaînes HS liées sur le domaine I et une autre potentiellement liée sur le domaine V. Sa composition modulaire lui permet d'interagir avec de nombreux autres partenaires protéiques. Par exemple, les domaines III et IV peuvent interagir avec un facteur de croissance tel que le FGF-7 (Mongiat et al., 2000). De nombreux autres exemples ont été décrits (revue dans Iozzo, 2005).

Le perlécan est un PG présent dans la membrane basale de la plupart des cellules endothéliales et épithéliales. Il peut également être associé à la surface cellulaire grâce à ses interactions avec les intégrines, et notamment l' $\alpha 2\beta 1$  (Iozzo, 1984 ; Bix et Iozzo, 2005). Il est également exprimé dans des tissus non vascularisés tels que le cartilage et les tissus conjonctifs (SunarRaj et al., 1995 ; Melrose et al., 2004).

Ainsi, chez les souris déficientes en perlécan (KO), la moitié des embryons ne survit pas alors que l'autre moitié possède de sévères anomalies au niveau du cartilage et du cerveau et décède de troubles respiratoires assez rapidement après leur naissance.

En revanche, la suppression d'une région du gène de perlécan qui supprime l'attachement de deux des trois chaînes HS n'est pas fatale. Les souris ont cependant des yeux de taille réduite et une dégénération de leur système oculaire a été observée.

Chez l'Homme, des mutations dans le gène *HSPG2* codant pour le perlécan causent des désordres physiologiques. Par exemple, l'absence du perlécan déstabilise la matrice du cartilage à cause notamment de la mauvaise régulation des FGFs. Une délétion du domaine V du perlécan se traduit par le syndrome Schwartz-Jampel affectant les muscles et qui est lié à la non-dégradation de l'acétylcholine, normalement effectuée par le complexe composé de l'acétylcholinestérase, du perlécan, de l'agrine, de l'acétylcholine récepteur et du MuSK (Muscle-specific receptor tyrosine kinase) (Iozzo, 2005). Pour en apprendre davantage sur les fonctions du perlécan il y a la revue récente de Knox et Whitelock (2006).



**Figure 12: Représentation schématique des quatre membres de la famille des hyalectans**

A l'intérieur d'un même protéoglycane, sont mis en évidence, les différents domaines par les chiffres I, II et III. Ces domaines sont composés de modules : Ig pour les répétitions de type immunoglobuline, LP pour les modules de type "link-protein", GAG pour les domaines constituant les sites d'ancrage pour les GAGs ( $\alpha$  et  $\beta$  font valoir l'épissage alternatif), EG pour les modules "EGF-like", Lectin pour les modules "C-type lectin-like", CR pour le module protéique de "Complement Regulatory", KS pour le module d'attachement des KS.

(d'après Iozzo 1998)

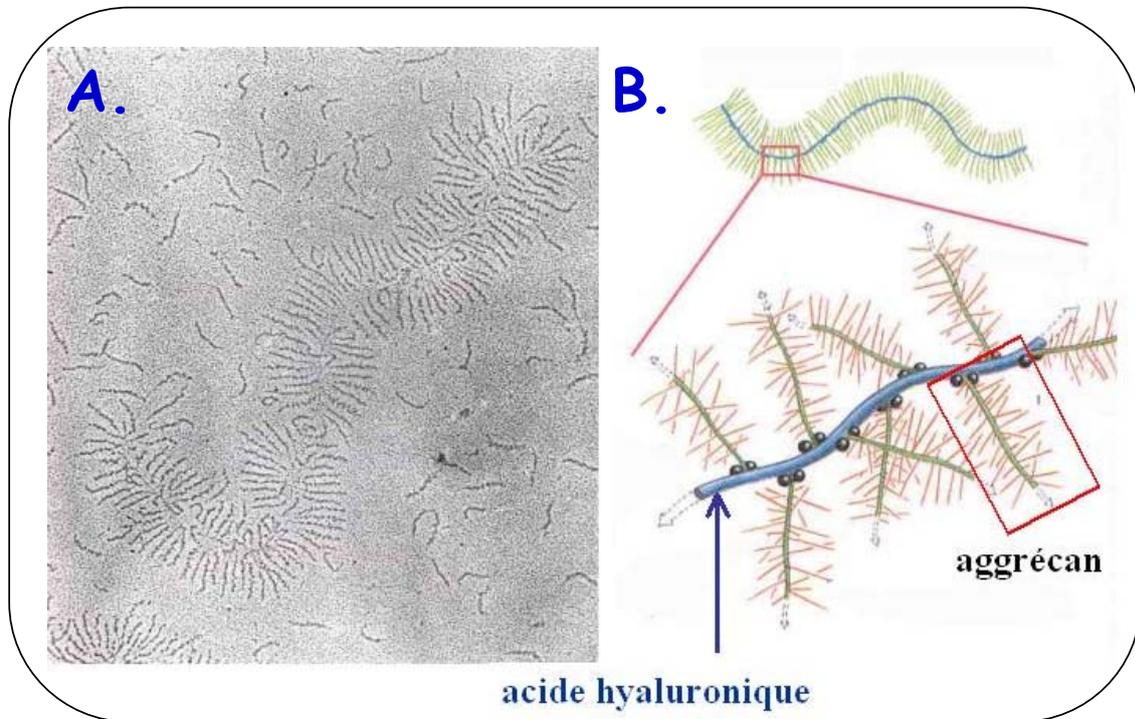
### 3.4.2. La famille des hyalectans

Les protéoglycannes qui composent la famille des hyalectans (également nommée famille des lecticans) sont similaires par leur séquence peptidique et par les fonctions qui leur sont attribuées (Iozzo, 1998). Quatre protéines issues de gènes différents sont membres de cette famille, à savoir le versican, l'aggrécan, le neurocan et le brévican (Figure 12). Le terme "hyalectan" vient du fait que ce sont des protéoglycannes qui se lient à l'acide hyaluronique et aux lectines. Ils comportent trois domaines conservés : la région N-terminale (domaine I) qui va permettre les interactions avec l'AH ; un domaine central (domaine II) sur lequel les chaînes de GAGS sont ancrées ; et enfin la partie C-terminale (domaine III) qui interagit avec les lectines (Iozzo, 1998). Les hyalectans sont des CSPG. L'aggrécan possède également des GAGs de type KS dans un sous-domaine composé de séquences consensus hexapeptidiques (Heinegard et al., 1985). Le neurocan possède également des chaînes KS. En revanche, aucune chaîne d'HS n'a été décrite dans la littérature. Le nombre de sites d'ancrage des GAGs sur le domaine central est respectivement pour l'aggrécan, le versican, le neurocan et le brévican, d'environ 120, 20, 7 et 3 (Yamaguchi, 2000).

#### 3.4.2.1. Le versican

Le versican est le membre le plus représenté de la famille des hyalectans. Il est composé de modules peptidiques. Son domaine I (N-terminal) contient un module Immunoglobuline suivi de deux modules de liaison ("link-protein") qui sont impliqués dans la liaison à l'acide hyaluronique. Le domaine II contient deux régions de fixation des GAGs : les domaines GAG- $\alpha$  et GAG- $\beta$  codés par deux exons différents. Ils contiennent jusqu'à 30 séquences consensus permettant l'attachement des GAGs ainsi que d'autres sites de N- et O-glycosylations (Iozzo, 1998). L'épissage alternatif de ce domaine II crée quatre variants possibles chez les versicans. Le plus répandu, V0, contient les deux domaines GAG- $\alpha$  et GAG- $\beta$ , les trois autres contiennent soit le domaine GAG- $\beta$  (V1), soit le domaine GAG- $\alpha$  (V2), soit aucun des deux (V3). Le domaine III, quant à lui, est composé de deux modules de type EGF suivi par un "C-type lectin domain" et enfin un module de type "complement regulatory protein" (Figure 12) (Iozzo, 1998).

L'épissage alternatif semble avoir un rôle fonctionnel et permet une expression différentielle de chaque isoforme en fonction du type cellulaire. Les isoformes V0 et V1 par



**Figure 13: L'association de l'aggrécan et de l'acide hyaluronique**

**A :** Photo de microscopie électronique de l'agrégat formé entre une chaîne d'acide hyaluronique et plusieurs molécules d'aggrécan.

**B :** Dessin représentant une chaîne d'acide hyaluronique (bleu) sur laquelle, grâce à des protéines de liaison (boule noire), sont fixées de manière régulière, tout au long de la chaîne, des molécules d'aggrécan.

(d'après [http://anatomy.yonsei.ac.kr/slide/histo/C&B\\_2003.files/slide0050.htm](http://anatomy.yonsei.ac.kr/slide/histo/C&B_2003.files/slide0050.htm))

exemple sont exprimés dans les fibroblastes, les chondrocytes, les hépatocytes, alors que les kératinocytes expriment uniquement l'isoforme V1 (Iozzo, 1998).

Les fonctions des versicans sont diverses. Agrégés avec l'acide hyaluronique, ils permettent la rétention d'eau lors des processus de remodelage cellulaire (réparation après blessure ou inflammation) (Delehedde et al., 2001), ceci même si le nombre de chaînes de GAGS du versican est moins important que l'aggrécan. Une forte localisation de versicans a été observée dans les tissus en prolifération intense, notamment dans les zones riches en AH associées à l'inflammation (Bensadoun et al, 1997 ; Nara et al., 1997 ; Brown et al., 1999 ; Venkatesan et al., 2000 ; Johnson, 2001). D'une manière générale, la grande diversité fonctionnelle du versican est liée au fait qu'il interagit avec de nombreuses protéines de la matrice extracellulaire et de la surface cellulaire (Wu et al., 2005). Récemment, la surexpression du versican et de la décorine, qui sont tout deux des CSPG, a été associée à un mauvais pronostic dans le cancer du pancréas (Skandalis et al., 2006). L'absence d'expression du versican chez le rat (par la technique des RNAi) altère les cellules musculaires lisses (prolifération cellulaire et migration réduites) et permet sur ces cellules une intensification du phénotype lié à l'élastogénèse (augmentation du taux d'expression des protéines de l'élastogénèse) (Huang et al., 2006).

#### **3.4.2.2. L'aggrécan**

L'aggrécan est le plus volumineux de tous les PGs (Iozzo, 1998). La composition du corps protéique reste semblable à celle du versican à l'exception d'un domaine favorable à la fixation des chaînes KS (Figure 12). La région centrale (domaine II) peut comporter jusqu'à 120 chaînes de CS qui ont une taille variable de 20 à 60 kDa (Iozzo, 1998). L'aggrécan est le PG majeur des cartilages, notamment du cartilage hyalin (essentiellement articulations). Cela n'exclut pas qu'il est également présent dans la matrice extracellulaire d'autres types de tissus. Dans les cartilages, son expression massive et sa forte affinité pour l'AH permettent de créer des complexes macromoléculaires anioniques de tailles très importantes, atteignant parfois 200 000 kDa et pouvant donc être perçus comme une agrégation, d'où le terme "aggrécan" qui a été donné à ce PG (Figure 13). Ces complexes macromoléculaires sont souvent stabilisés par des protéines qui font le lien entre deux molécules qu'on appelle les "link protein" ou protéines de liaison. On trouve parmi ces protéines, le TSG-6 et le CD44. La fonction de ces complexes macromoléculaires est principalement de conférer au cartilage les propriétés qui lui sont associées, à savoir une grande résistance au stress mécanique et une

	NOM (Taille du core protéique)	GAG (nombre de chaînes)	Particularité (nb d'exons, nb de LRR, région N-terminale riche en cystéine)
<b>Classe I</b>	Décorine (40 kDa)	DS (1 chaîne)	CX <sub>3</sub> CXCX <sub>6</sub> C
	Biglycan (40 kDa)	DS (2 chaînes)	10 LRR
	Asporine		8 exons
<b>Classe II</b>	Fibromoduline (42 kDa)	KS (2-3 chaînes)	3 exons
	Lumican (38 kDa)	KS (3-4 chaînes)	10 LRR
	Kératocan (38 kDa)	KS (3-5 chaînes)	CX <sub>3</sub> CXCX <sub>9</sub> C
	PRELP (44 kDa)	KS (2-3 chaînes)	
	Ostéoadhérine/ostéomoduline (42 kDa)	KS (2-3 chaînes)	
<b>Classe III</b>	Epiphican/PG-Lb/DSPG3 (35 kDa)	DS (2-3 chaînes)	7 exons
	Mimican/ostéoglycine (35 kDa)	KS (2-3 chaînes)	6 LRR
	Opticine/oculoglycan		CX <sub>2</sub> CXCX <sub>6</sub> C

**Figure 14: Tableau de classification des SLRPs**

Les SLRPs sont classées selon l'organisation de leur gène et l'homologie des membres à l'intérieur de cette famille. Dans la première colonne sont listées les trois classes de SLRP. Dans la seconde colonne sont présentés les PGs appartenant aux différentes classes avec le poids moléculaire de leur core protéique lorsqu'il est connu. La troisième colonne rend compte de la nature et du nombre de chaînes de GAGs qui sont liées au core protéique. Enfin la dernière colonne donnent des informations sur le nombre d'exons, le nombre de LRR (leucine-rich repeats) et les séquences à l'extrémité N-terminale riches en cystéines (C est une cystéine et X un autre acide aminé).

(d'après Iozzo, 1998 et Young et al., 2003)

importante élasticité. Ces complexes polyanioniques sont en effet capables de retenir l'eau dans une proportion très importante et l'aggrécan peut ainsi créer un volume hydrodynamique mille fois supérieur à son propre volume (Roughley et al., 2006 ; Dudhia et al., 2005 ; Caterson et al., 2000 ; Iozzo, 1998).

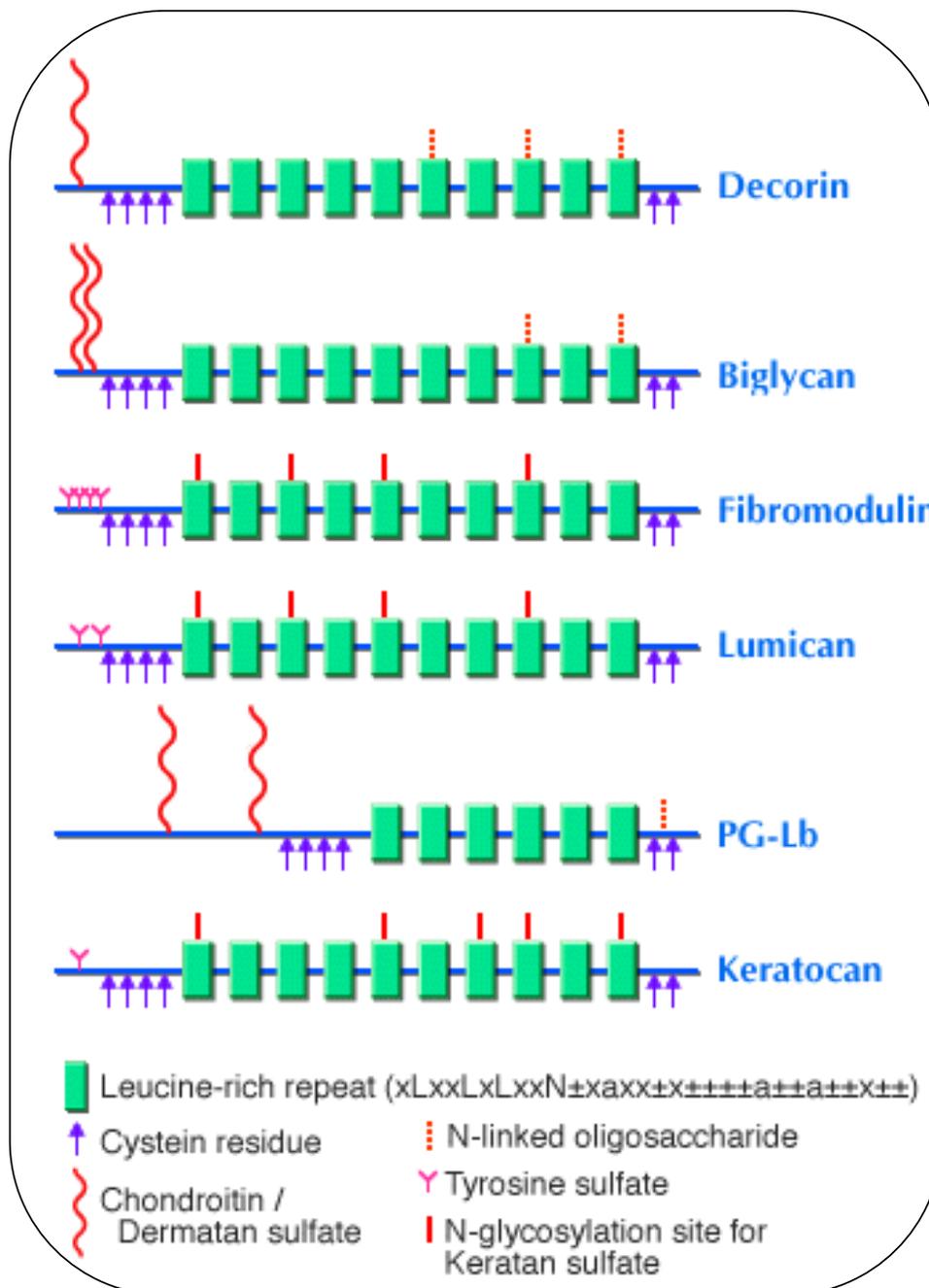
#### **3.4.2.3. Le neurocan**

Le neurocan est le troisième membre de la famille des hyalectans à avoir été cloné (Rauch et al., 1992). Sa composition protéique est similaire à celle du versican à l'exception de son domaine II, qui est plus restreint et comporte en moyenne trois chaînes de CS. Cependant sept séquences consensus pour l'attachement des CS sont présentes (Figure 12). La taille moyenne de ces chaînes CS varie de 20 à 35 kDa. Ces chaînes de CS sont principalement sulfatées en position C4 sur la GalNAc. Cependant, la composition de ces chaînes peut être amenée à évoluer d'une manière très significative. Par exemple, chez le rat de sept jours, les CS du neurocan seront pour 20 % des CS 6-O-sulfatées (CS C) alors que chez le rat adulte, moins de 3 % de CS C sont retrouvés (Rauch et al., 2001).

De manière intéressante, le neurocan comme le brévican sont des PG retrouvés principalement dans le cerveau (Rauch et al., 2001). Le neurocan est capable, avec son domaine C-terminal, d'interagir avec des molécules d'adhésion neurale telles que les Ng-CAM, N-CAM permettant une inhibition de la croissance neuritique et des interactions homophiliques créées par ces dernières. Sa localisation semble, dans certains cas, dictée par des interactions avec des molécules telles que la tenascine (-C ou -R) notamment via sa région N-terminale (Rauch et al., 2001).

#### **3.4.2.4. Le brévican**

Le brévican, du latin *brevis* (court) est le quatrième CSPG de la famille des hyalectans à avoir été découvert (Iozzo, 1998). Son nom est dû au fait que la région de fixation des GAGs sur le core protéique est courte (Figure 12). De ce fait, le brévican n'est capable de supporter que seulement une à trois chaînes oligosaccharidiques de type CS. Dans le cerveau humain, le brévican est le hyalectan le plus représenté. Le brévican existe soit sous une forme entière (145 kDa) laquelle possède des GAGs, soit sous une forme dégradée de façon protéolytique (80 kDa) qui ne possède plus les sites d'ancrage des CS. Ce phénomène de dégradation qui offre un core protéique ayant des fonctions différentes, se retrouve également chez l'aggrécan



**Figure 15a: Organisation moléculaire de certaines SLRPs**

Sur cette figure sont représentés de façon schématique les cores protéiques des SLRPs les plus étudiées. Bien que l'échelle ne soit pas respectée, cette figure permet de visualiser la localisation des sites de glycosylation, des LRR ainsi que des regroupements de cystéines le long de la partie protéique.

(D'après Yukiko Mochizuki, Tamayuki Shinomura, Koji Kimata (Aichi Medical University, Institute for Molecular Science of Medicine))

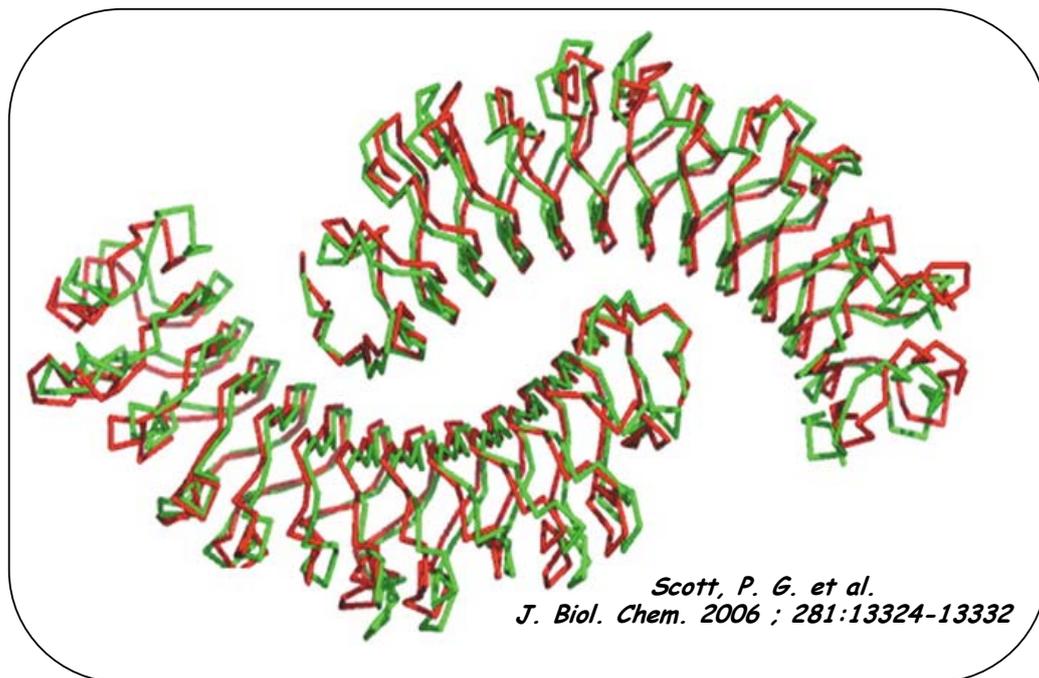
et cette coupure est dépendante des MMP (Matrix Metalloproteinase) (Yamaguchi, 1999). La forme tronquée du brévican devient plus abondante dans le stade tardif du développement embryonnaire. Chez le rat, en plus du brévican sécrété, un brévican lié aux membranes plasmiques via une ancre GPI, a été caractérisé (Figure 12) (Iozzo, 1998).

### **3.4.3. La famille de protéoglycannes riches en leucine (SLRP)**

#### **3.4.3.1. Introduction**

La famille des protéoglycannes riches en leucine appartient à la grande famille des SLRPs (Small Leucine-Rich Protein). Elle englobe les PGs de la matrice extracellulaire dont le core protéique, d'environ 40 kDa, est composé de séquences LRRs (Leucine Rich Repeats), délimitées de chaque côté par des regroupements de cystéines (Figure 14 a). Sur la base de leur organisation génomique et protéique, les SLRPs ont été classés en trois sous-classes dénommées tout simplement classe I, classe II et classe III (Figure 14). Dans la classe I, on trouvera pour les plus connus le biglycan, la décorine et l'asporine. La classe II est, elle, composée notamment du lumican, de la fibromoduline, du PRELP, du kératocan et de l'ostéoadherine/ostéomoduline. Enfin, la troisième classe inclut l'opticine, l'épiphycan et l'ostéoglycine (McEwan et al., 2006).

Les homologies dans la composition générale de ces SLRPs sont importantes, cependant quelques différences sont observées. Dans la classe I, chez la décorine et le biglycan par exemple, un signal pro-peptidique juxtaposé à un signal peptidique a été identifié dans le domaine I (Figure 15 a). Il semble nécessaire à l'activité de la xylosyltransférase, laquelle permet l'initiation de la biosynthèse des GAGs (Sawhney et al. 1991 ; Oldberg et al., 1996 ; Hocking et al., 1996). Les SLRP de classe II ne possèdent pas ce signal pro-peptidique, mais semblent en revanche subir une dégradation protéolytique après la perte de leur signal peptidique, cela afin de devenir actifs. Cette différence est en corrélation avec la composition en GAGs de chacune des classes de SLRP. En effet, alors que le biglycan et la décorine possèdent des chaînes de CS, les PG de la classe II vont, d'une manière générale, être ornés de chaînes de KS ancrés sur des asparagines (Iozzo, 1998). Le domaine II (Figure 15 a) composé principalement d'une séquence consensus comportant des cystéines (Cx<sub>2-3</sub>CxCx<sub>6-9</sub>C) permet aussi de faire la distinction entre les 3 classes. En effet, les PGs de la classe I ont leurs Cys espacées par 3, 1, 9 résidus, ceux de la classe II par 3, 1, 9 et ceux de la classe III par 2, 1, 6 résidus (McEwan et al., 2006). Les acides aminés présents entre les cystéines sont fortement



***Figure 15b: Superposition des structures de la décorine et du biglycan***

Sur cette figure sont superposés les atomes C- $\alpha$  du biglycan (vert) et ceux de la décorine (rouge). Une structure dimérique quasi-identique qui reflète l'homologie des deux molécules.

*(D'après Scott et al., 2006)*

conservés par classe. Les éléments les plus conservés entre les classes sont bien sûr les séquences LRR de compositions LxxLxLxxNxLSxL, où L est une leucine, isoleucine ou valine, S est une sérine ou thréonine. Dix répétitions sont couramment observées, bien que chez les SLRP de classe III seulement six soient présentes (Iozzo, 1998). La région C-terminale (domaine IV) des SLRP est caractérisée par au moins deux cystéines. Chez le biglycan, un pont disulfure a été identifié, créant une boucle qui permet l'interaction du biglycan avec le collagène. Cette interaction semble cruciale pour l'activité du biglycan dans le contrôle de la fibrillogénèse (Rada et al., 1996).

La famille des SLRP ne cesse de s'agrandir. Aussi, les principales caractéristiques du biglycan, de la décorine, du lumican et de la fibromoduline seront uniquement détaillées ci-après car ils sont les plus étudiés pour l'instant. Ils ont par ailleurs en commun la capacité de fixer le TGF- $\beta$  (Transforming Growth Factor beta) et de jouer un rôle dans la régulation de la fibrillogénèse du collagène (Wadhwa et al., 2005).

#### **3.4.3.2. Le biglycan**

Le biglycan appartient à la classe I des SLRP. Il possède deux chaînes de CS ancrées du côté N-terminal. Comme cela a été observé pour la décorine, il se dimérise en solution grâce à ses LRR. Cette dimérisation permettrait de stabiliser la protéine (Scott et al., 2006) (Figure 15b).

La suppression, chez la souris, du gène codant pour le biglycan, a permis d'évaluer sa fonction. Young et al. ont ainsi montré (2003) que la perte de l'expression du biglycan chez la souris dont le gène a été inactivé permet d'observer une longueur d'os réduite ainsi qu'une masse et densité minérale plus faibles. Un processus d'ostéoarthrose encore plus flagrant a été observé chez une souris modifiée génétiquement pour ne pas exprimer à la fois le biglycan et la fibromoduline (Wadhwa et al., 2005a ; 2005b).

Il est également connu que le biglycan aide à la formation des réseaux de collagène de type IV, préférentiellement grâce à son interaction avec ces derniers (Wiberg et al. 2002). Toujours dans le contexte de la matrice extracellulaire, le biglycan (au même titre que la décorine) se complexerait aux matrilines de type I, II ou III pour créer un lien entre les collagènes de type VI et l'aggrécan (Wiberg et al., 2003). Ce lien étroit avec l'aggrécan permet de rappeler que les aggrécanases 1 et 2, qui sont les enzymes responsables notamment de coupures protéolytiques dans la partie protéique de l'aggrécan, sont également capables de

couper le biglycan (et non pas la décorine) dans son domaine riche en LRR (Melching et al., 2006).

#### **3.4.3.3. La décorine**

A l'image du biglycan son homologue, la décorine appartient aux PGs de la famille des SLRPs de classe I (Iozzo, 1998). C'est l'un des rares PGs à n'avoir qu'une seule chaîne de DS liée à son core protéique (Roughley, 2006). La décorine possède également trois sites de N-glycosylation qui semblent nécessaires à la sécrétion extracellulaire (Seo et al., 2005). Stable sous forme dimérique, ce PG, membre le plus étudié des SLRP, exerce son activité aussi bien avec son corps protéique qu'avec sa chaîne de DS. Il permet de réguler la formation et la stabilisation des fibrilles de collagènes (Iozzo, 1998). Il est par ailleurs capable de constituer une sorte de réservoir pour le TGF- $\beta$  (Kinsella et al., 2001 ; Kresse et al., 2001) et de moduler l'activité de ce facteur de croissance (Iozzo, 1998 ; Droguett et al., 2006). La décorine permet de contrôler le devenir de nombreux types cellulaires en interagissant par exemple avec des récepteurs de la famille des ErbB dans les cellules tumorales. Cette liaison serait en relation avec une réduction des métastases tumorales dans le cancer du sein (Santra et al., 2000 ; Reed et al. 2005). L'interaction de la décorine avec d'autres récepteurs du type tyrosine kinase a également été montrée, notamment avec l'EGFR (Epidermal Growth Factor Receptor) (Santra et al., 2002) ou avec l'IGFR (Insulin-like Growth Factor Receptor) sur les cellules endothéliales (Schönherr et al., 2005). De plus, la décorine semble jouer un rôle dans le contrôle de l'angiogenèse. Grant et coll. (2002) ont par exemple remarqué que l'ajout du PG pouvait *in vivo* affecter la croissance tumorale par l'inhibition de la croissance vasculaire. L'absence du gène codant pour la décorine chez la souris semble responsable d'une peau fragile et relâchée ainsi que d'un changement dans la morphologie des fibrilles de collagène (Corsi et al., 2002).

#### **3.4.3.4. Le lumican**

Le lumican appartient à la famille des SLRP de classe II (Iozzo, 1998). Il a la particularité d'être retrouvé soit sous la forme d'une glycoprotéine non sulfatée, soit sous la forme d'un protéoglycane. Alors qu'une à deux chaînes de polylactosamines peuvent décorer la glycoprotéine de lumican, ce sont des KS (5 sites potentiels) qui ornent plus spécifiquement la forme "protéoglycane" du lumican, le rangeant dans la catégorie des KSPG (Naito 2005). Ce

changement de modification post-traductionnelle semble jouer notamment un rôle dans la transparence de la cornée. Dans la cornée adulte, la forme "PG" est principalement présente alors que dans celle de l'embryon la forme "glycoprotéine" est toujours majoritaire (Cornuet et al. 1994). De plus, la suppression du gène codant pour le lumican chez la souris crée une opacification de la cornée et un relâchement de la peau, pouvant s'expliquer par une répartition et un assemblage non homogènes du collagène (Chakravarti et al., 1998).

Le lumican est le PG majeur de la plupart des stromas (Iozzo, 1998). Il est parallèlement surexprimé principalement dans les régions de fibroses pathologiques par les cellules stromales dont les VSMCs (Vascular Smooth Muscle Cells). Sa faible expression dans les tumeurs du sein est souvent associée à un mauvais pronostic. Outre ses capacités à organiser les fibrilles de collagène et à rendre la cornée transparente, le lumican permet également de réguler l'expression d'un autre KSPG qui est le kératocan (Carlson et al., 2005). Ce mécanisme semble être nécessaire pour assurer un niveau suffisant d'expression de petits protéoglycannes utiles à l'organisation du collagène dans le stroma.

#### **3.4.3.5. La fibromoduline**

La fibromoduline est également un PG de la classe II dans la famille des SLRP (Iozzo, 1998). Son domaine LRR peut comporter jusqu'à cinq chaînes de KS. On le retrouve principalement dans les tissus interstitiels, le cartilage, les tendons et aussi dans l'os en formation. L'inactivation du gène codant pour la fibromoduline n'entraîne, chez la souris, aucun trouble apparent majeur et la viabilité comme la fertilité de ces animaux mutants sont conservées (Svensson et al., 1999). L'organisation tissulaire du tendon d'Achille chez ces souris est cependant anormale, probablement à cause d'une modification de l'architecture des collagènes caractérisée par des fibrilles de diamètre nettement plus petit (Chakravarti, 2003). Cette nécessité de fibromoduline dans les tendons est confirmée par les travaux d'Ameye et al. (2002) qui observent une ossification des tendons ainsi que le développement d'ostéoarthrite chez des souris déficientes en fibromoduline.

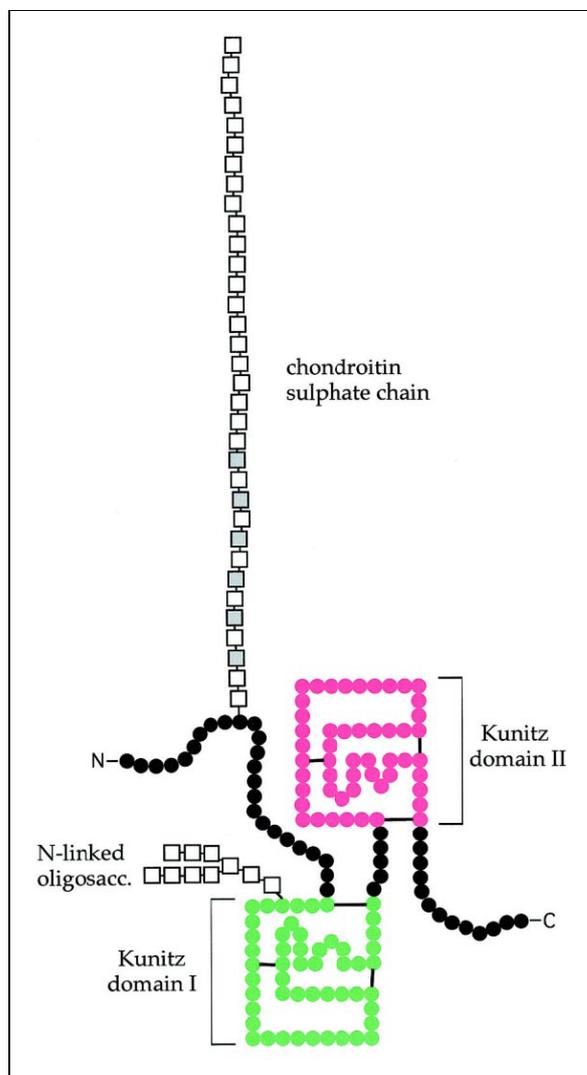
*In vitro*, la fibromoduline et le lumican ont le même site de fixation sur les fibres de collagènes de type I (Svensson et al., 2000). Ezura et al. (2000) suggèrent de ce fait que la fibromoduline occuperait, durant les phases avancées du développement, des fonctions comparables à celles du lumican dans le domaine de la fibrillogénèse lors des premières phases du développement chez la souris.

D'une manière générale pour le lumican et la fibromoduline, peu de rôles sont encore connus concernant la partie GAG de ces PGs. Les KS permettent probablement l'hydratation de la région cornéenne (même si la capacité d'hydratation de KS est plus faible que celle des DS). Les KS ont également sans aucun doute des fonctions sur le stroma puisque leur absence semble être en cause dans les désordres tels que la dystrophie cornéenne maculaire. Par ailleurs, il a été observé que les macrophages murins ne peuvent pas lier le lumican lorsqu'il possède des KS, alors qu'en leur absence, l'interaction devient possible. D'autres expériences ont alors suggéré des propriétés antiadhésives à ce type de GAGs (Funderburgh, 2000).

### **3.4.4. Les protéoglycannes circulants**

#### **3.4.4.1. Le PG-100 ou MCSF1**

La partie protéique (75 kDa) du PG-100 fait environ 106 kDa après l'ajout d'oligosaccharides sur deux asparagines et quelques O-glycosylations, d'où son nom de proteoglycan-100. A cela s'ajoute une unique chaîne de CS (Schwartz et al., 1990). Il est sécrété principalement par les macrophages mais a également été retrouvé sécrété par des lignées cellulaires d'ostéosarcome ou endothéliales. *In vitro*, son expression depuis une lignée endothéliale est régulée par différents facteurs. Par exemple, en réponse à du FGF-2 ou FGF-7 et du TGF- $\beta$ , la sécrétion du PG-100 est accrue. Cette observation n'est pas valable pour d'autres "petits" CSPG, tels que le biglycan et la décorine (Nelimarkka et al., 1997). Dans le cas du PG-100, la fonction attribuée à l'entité "protéoglycane" dans sa totalité n'est pas claire dans la littérature. En effet, grâce à une recherche d'homologues par alignement de séquences d'ADNc, il a été démontré que le PG-100 a la même séquence protéique que la protéine connue sous le nom de MCSF-1 (Macrophage Colony Stimulating Factor). Le MCSF-1 est le principal régulateur de la survie, de la prolifération, et de la différenciation des phagocytes mononucléaires. Il joue un rôle important dans l'immunité innée, contre le cancer et les maladies liées à l'inflammation. Ses fonctions dans l'immunité et l'inflammation ont récemment fait l'objet d'une revue (Chitu et Stanley, 2006). Partenheimer et coll. (1995) ont cependant mis en avant une fonction potentielle de la chaîne CS particulièrement intéressante. En effet, leurs travaux suggèrent que la partie CS du PG-100 aurait pour rôle de permettre le stockage de la protéine dans un environnement extracellulaire en inhibant son activité (activité CSF-1) lorsqu'elle n'est pas nécessaire. L'activité CSF-1 est, en effet, à priori issue



**Figure 16 a: Représentation schématique de la structure de la bikunine**

La chaîne polypeptidique est constituée par deux domaines de type Kunitz, reliés par un court peptide. Chaque acide aminé est représenté par un cercle et les sucres par un carré. On trouve lié au polypeptide une chaîne de chondroïtine sulfate ainsi qu'un N-glycane.  
(D'après Fries et Blom, 2000)

d'un fragment de dégradation du PG-100 de 50 kDa ne comportant pas la chaîne de GAG (Partenheimer et al., 1995).

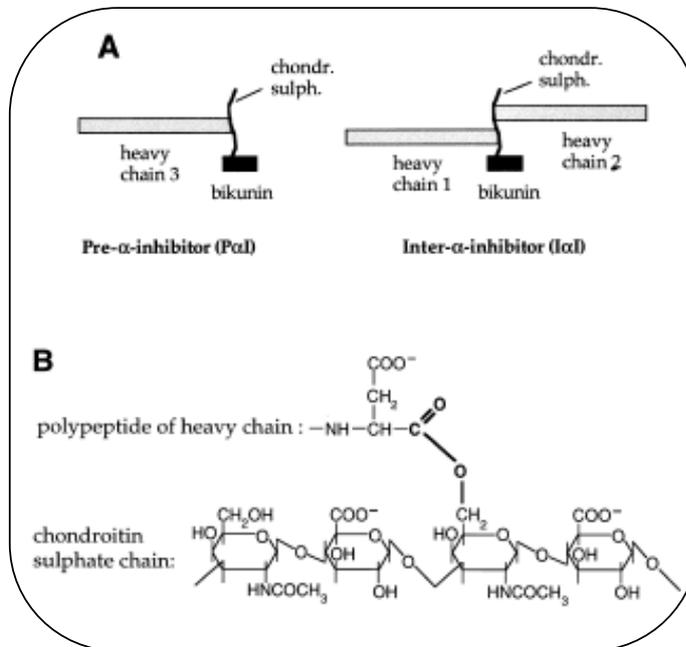
#### 3.4.4.2. La bikunine

L'UTI (Urinary Trypsin Inhibitor), l'ASTI (Acid Stable Trypsin Inhibitor), l'HI30, la migraine, l'urastine ou même l'ulinastatine se sont révélés n'être qu'un seul et même protéoglycane : la bikunine. Ce PG tient son nom des deux domaines Kunitz qui constituent la majeure partie de la partie protéique. Le domaine de type Kunitz est structuralement bien caractérisé : c'est un polypeptide de 7 kDa qui possède trois ponts disulfures. Ce motif structural est défini comme un domaine inhibiteur de protéases. Le core protéique de la bikunine est composé, outre ces deux domaines, d'une séquence de 5 acides aminés faisant le lien entre eux, et de courtes séquences N- et C-terminales (Figure 16a). La structure cristallographique de la bikunine, dont les extrémités N- et C-terminales sont manquantes a pu être résolue (Xu et al., 1998 ; code pdb : 1BIK).

L'unique chaîne de GAG de type CS qui fait de la bikunine un PG, se trouve être ancrée dans la région N-terminale de la protéine. Cette région du core protéique est sensible à la protéolyse. La chaîne de CS a été caractérisée comme plutôt courte (7 à 8 kDa) lorsque la bikunine est sécrétée par des hépatocytes de rat. D'autres études sur le PG "humain" ont montré que cette chaîne de GAG se trouve être uniquement sulfatée sur les cinq GalNAc les plus proches du core protéique (Fries et Blom, 2000).

Une N-glycosylation d'environ 2 kDa s'ajoute à cet ensemble protéoglycannique. Ainsi avec un core protéique de 16 kDa, la bikunine dans son intégralité fait environ 25 kDa (pour un volume hydrodynamique de protéine globulaire de référence de 67 kDa).

Depuis son mode d'expression jusqu'à sa présence dans le plasma, la bikunine est un PG complètement atypique. En effet, le core protéique est exprimé dans le réticulum endoplasmique sous forme d'un précurseur polypeptidique contenant à la fois le core protéique de la bikunine et l' $\alpha_1$ -microglobuline. C'est sur ce précurseur que, dans l'appareil de Golgi, la chaîne CS est ajoutée sur la partie constituant la bikunine. Ensuite, une liaison covalente va avoir lieu entre la chaîne CS et un ou deux polypeptides appelés respectivement chaîne lourde 3 ou 1 et 2 ("heavy chain"). Cette liaison covalente a lieu par estérification entre le carbone- $\alpha$  de l'acide aminé C-terminal de la chaîne lourde et un résidu non sulfaté de GalNAc (Figure 16b). Cette liaison peu courante chez les GAGs, permettrait, dans des vésicules de sécrétion, la protéolyse du précurseur créant ainsi la séparation et la libération



**Figure 16 b: Le pré- $\alpha$ -inhibiteur et l'inter- $\alpha$ -inhibiteur**

**A** : Représentation schématique des deux protéines duplasma liées à la bikunine : pré- et inter- $\alpha$ -inhibiteur. Dans ces protéines, un ou deux polypeptides de 75 ou 155 kDa, appelés "heavy chain 3" ou "heavy chain 1 et 2" (en gris), sont liés covalamment à la chaîne de chondroïtine sulfate de la bikunine (en noir). Plus de 90% de la bikunine présente dans le plasma est ainsi complexée.

**B** : Structure chimique de la liaison entre les "heavy chains" et la chaîne de chondroïtine sulfate de la bikunine. Le carbone alpha de l'acide aminé C-terminale de la "heavy chain" est covalamment attaché via une liaison ester à un résidu GalNAc non sulfaté de la chaîne de CS.

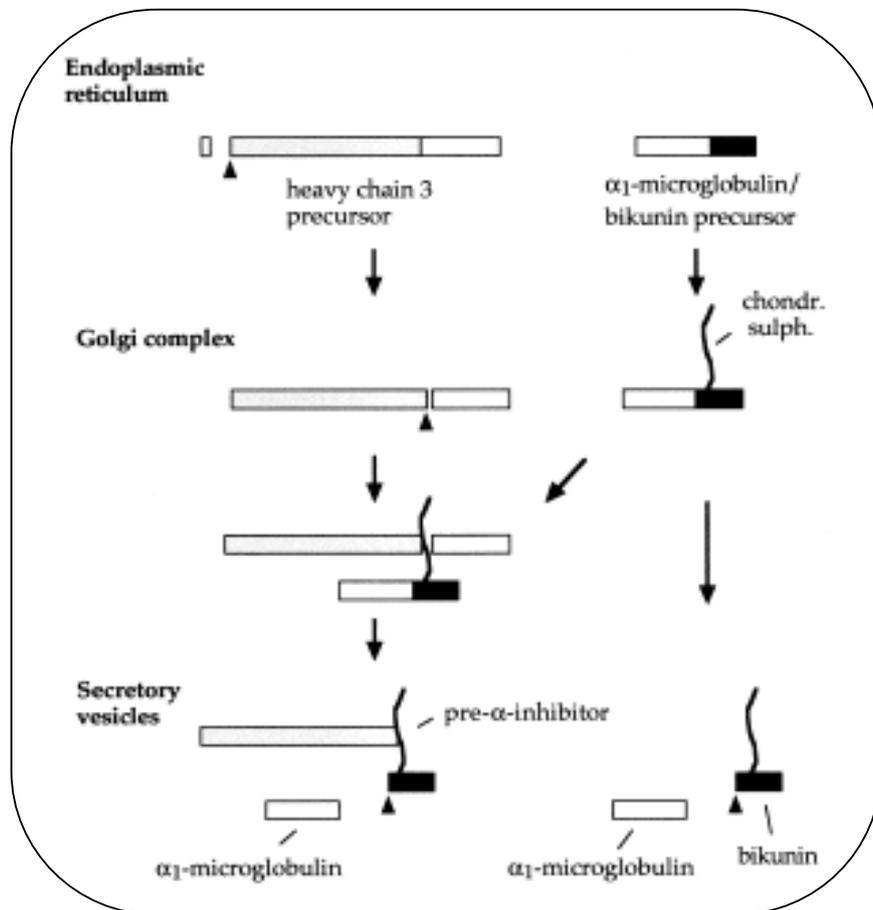
(D'après Fries et Blom, 2000)

dans le sang, de l'  $\alpha_1$ -microglobulin et de la bikunine (Fries et Blom, 2000). Cette dernière semble être dans 90 % des cas, liée avec la ou les chaînes lourdes pour former un complexe respectivement dénommé : pre- $\alpha$ -inhibitor (complexe avec la "heavy chain" 3) ou inter- $\alpha$ -inhibitor (complexe avec les "heavy chain" 1 et 2). La forme inter- $\alpha$ -inhibitor est la forme la plus représentée chez l'Homme (Figure 16c).

Des expériences ont montré, par RT-PCR, que la bikunine était exprimée dans les reins, l'intestin, l'estomac, le pancréas et le foie. De même, par immunohistochimie, la bikunine a été observée dans le foie, les reins, la peau, la vésicule biliaire, le côlon, l'appendice et aussi le cerveau. Dans le plasma, ce PG circulant est à 90 % sous forme de complexe protéique, à une concentration de l'ordre de 4-7  $\mu$ M. Dans l'urine, il est détecté sous forme non complexée à environ 0,005  $\mu$ M. Sa demi-vie sous forme non complexée dans le sang est de l'ordre de 4 à 30 minutes et sa présence dans l'urine semble due à une réabsorption non complète. En revanche, le taux de bikunine dans les urines apparaît bien plus élevé dans des pathologies telles que les pneumonies, l'arthrite rhumatoïde ou encore dans le cas de cancers. De plus, il a été révélé que le taux de bikunine non complexée dans le plasma augmente dans des conditions inflammatoires (Fries et Blom, 2000).

Les fonctions de la bikunine décrites pour l'instant sont certes nombreuses mais restent floues. Il est connu qu'elle est capable d'inhiber des protéases telles que la trypsine, la chymotrypsine, l'élastase granulocytaire, la plasmine, la cathepsine G et l'acrosine. De plus, *in vivo* et *in vitro*, la bikunine peut inhiber le processus de métastase de cellules tumorales, certainement en inhibant la plasmine membranaire. Aussi, la mise en présence de TSG-6 et d'inter- $\alpha$ -inhibitor permettait le relargage d'une des chaînes lourdes pour laisser place à une liaison covalente entre les TSG-6 et la chaîne CS de la bikunine. Cela augmente l'inhibition de la plasmine par la bikunine et pourrait constituer une explication de l'effet anti-inflammatoire de l'inter- $\alpha$ -inhibitor lors d'injections intraveineuses. La bikunine est également capable d'inhiber la stimulation des neutrophiles par les LPS (Fries et Blom, 2000).

Récemment, le complexe inter- $\alpha$ -inhibitor (I $\alpha$ I) semble avoir révélé quelques secrets sur la bikunine. En effet, présent dans le sang, l'I $\alpha$ I va atteindre les matrices riches en AH ; un phénomène de transestérification va permettre la libération des chaînes lourdes (jusque là liées à la bikunine) pour les ancrer de façon covalente sur l'AH créant ainsi le complexe SHAP-HA (Serum Derived Hyaluronan Associated Protein - Hyaluronic Acid). Ce complexe permet la stabilisation des matrices extracellulaires riches en AH. Il semble également jouer



**Figure 16 c: Biosynthèse de la bikunine et de la protéine pre-α-inhibitor contenant la bikunine**

La bikunine est issu d'un précurseur qui est lié à l'α1-microglobuline. Ce précurseur acquiert une chaîne de CS dans l'appareil de Golgi. La liaison entre la chaîne de CS et les "heavy chains" peut alors avoir lieu. Cette réaction semble médiée par le pro-peptide C-terminal de la "heavy chain". Lorsque cette entité atteint les vésicules de sécrétion, l'α1-microglobuline mature est relarguée et laisse seul le pre-α-inhibiteur. Finalement, les deux protéines sont libérées dans le courant circulatoire, la bikunine étant la plupart du temps complexée à la "heavy chain" (sous forme de pre-α-inhibitor).

(D'après Fries et Blom, 2000)

un rôle dans l'arthrite rhumatoïde ou la maladie de Crohn, notamment dans le recrutement des leucocytes (Zhuo et al., 2004).

Dans le même temps, il a été confirmé que la bikunine est capable de bloquer l'induction du TNF $\alpha$  par les LPS dans les macrophages de façon dose dépendante et ainsi de bloquer l'expression de l'interleukine-1 $\beta$  et -6 (IL-1 $\beta$  et IL-6) (Matsuzaki et al., 2004). De plus, la bikunine permet *in vivo* la survie de souris pourtant traitées avec une dose létale de LPS. Tout cela permet de penser que ce PG est un candidat potentiel pour un traitement anti-inflammatoire notamment, par exemple, dans le cas du choc septique.

La bikunine inhibe l'invasion tumorale et le phénomène de métastase. Son efficacité *in vivo* comme agent thérapeutique a été testée chez la souris. Pour les doses les plus importantes administrées, les tumeurs décroissent de 40 % (Kobayashi et al., 2004). La même équipe a également développé des anticorps anti-bikunine qui pourraient être des outils de diagnostic notamment dans le cas du cancer des ovaires. Ils ont en effet montré que le taux élevé de bikunine dans le plasma révélait de bons pronostics dans ce type de cancer (Matsuzaki et al., 2005).

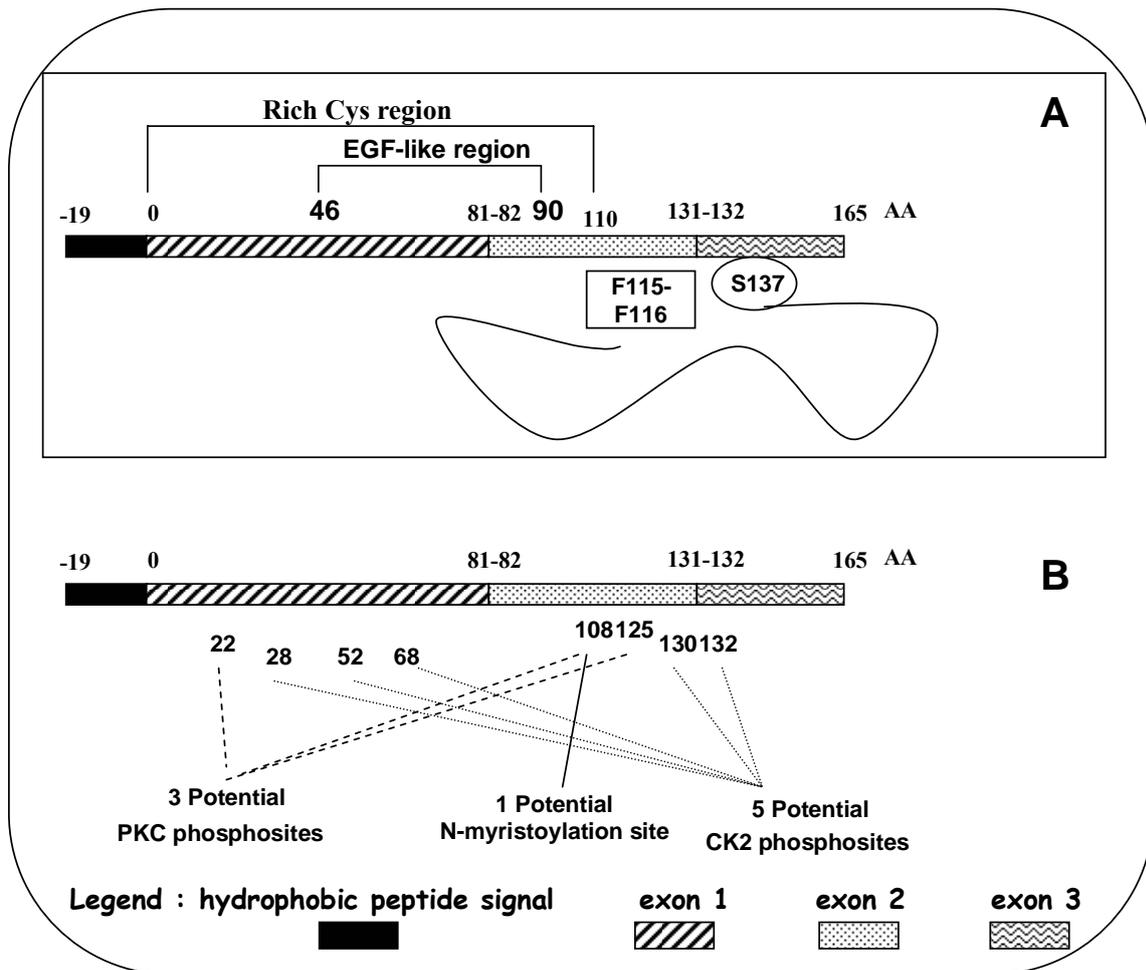
#### **3.4.4.3. L'apolipoprotéine O**

Récemment, dans des modèles d'obésités, une apolipoprotéine surexprimée dans le cœur a été identifiée. Pour la première fois, une apolipoprotéine a été caractérisée comme appartenant à la famille des protéoglycannes. En effet, un traitement à la chondroïtinase a permis de montrer qu'elle possède une ou plusieurs chaînes de CS. Dans la continuité du classement (par ordre alphabétique) des membres de cette famille de protéines, elle a été nommée apolipoprotéine O ou apoO. L'apoO possède un core protéique de 198 acides aminés dans lesquels est inclu un "peptide signal" de 23 acides aminés. Un traitement à la Naringenine ou au CP-346086 indique que la sécrétion d'apoO requiert l'activité de la protéine responsable du transfert des triglycérides microsomaux. Le gène d'apoO est surexprimé dans le cœur des diabétiques. De plus, l'apoO induit l'efflux de cholestérol depuis les macrophages, et il est possible que ce PG tout à fait récent soit impliqué dans des mécanismes de protection du myocarde contre l'accumulation de lipides (Lamant et al., 2006).

#### **3.4.4.4. L'endocan**

Les caractéristiques et fonctions d'endocan ne seront pas détaillées dans cette partie puisque le chapitre suivant lui est entièrement consacré.

**CHAPITRE II :**  
**L'ENDOCAN, UN PROTEOGLYCANNE**  
**ATYPIQUE ET MYSTERIEUX**



**Figure 17: Représentation schématique des caractéristiques d'endocan**

L'endocan est représenté avec son peptide signal hydrophobique (noir) et les exons 1, 2 et 3.

**A** : Sur cette figure, sont mis en évidence, la région N-terminale riche en cystéines (18 cystéines pour 110 acides aminés), la région potentielle EGF-like (a.a. 46 à 90), la région riche en phénylalanine (F115-116) et le site de O-glycosylation sur la sérine 137 sur laquelle l'unique chaîne de DS peut se lier.

**B** : Cette figure met en valeur différents sites potentiels de modifications post-traductionnelles tels que 3 sites de type PKC, un unique site de N-myristoylation et les 5 sites de CK2 (Casein Kinase II)

(d'après Sarrazin et al., 2006)

## 1. INTRODUCTION

L'endocan ou ESM-1 (pour **Endothelial Cell-Specific Molecule 1**) fait partie des protéoglycannes à dermatane sulfate. Il ne contient qu'une seule chaîne de DS (Lassalle et al., 1996 ; Béchard et al., 2001a ; Béchard et al., 2001b). Sa découverte s'est faite dans le laboratoire de Philippe Lassalle, à l'Institut Pasteur de Lille, au cours d'une étude dont le but était initialement d'isoler, par criblage d'une banque d'ADNc d'HUVEC, un auto antigène de 55 kDa potentiellement impliqué dans une forme d'asthme sévère (Lassalle et al., 1996).

Par ailleurs, l'endocan a également la particularité d'être un PG circulant. Il a été retrouvé dans les sites d'inflammation ainsi que dans un environnement tumoral. Comme il a été décrit dans le chapitre précédent, les PGs que ce soit via leur core protéique ou via leurs chaînes de GAGs sont souvent d'importants modulateurs des protéines, capables de contrôler les événements inflammatoires ou tumoraux. Des études récentes ont montré que le taux d'endocan circulant augmente significativement chez les patients atteints de cancer ainsi que ceux présentant un sepsis (Béchard et al., 2000 ; Scherpereel et al., 2003 ; 2006). Plusieurs des indications laissent donc penser que l'endocan pourrait avoir un rôle important dans ces pathologies.

Ce "jeune" protéoglycane d'une dizaine d'années est encore mal connu. Nous allons faire le point sur l'ensemble des données qui le concernent et, ainsi, offrir un maximum d'éléments nécessaires pour appréhender au mieux ses fonctions potentielles.

## 2. DU GENE AU PROTEOGLYCANNE

### 2.1. Organisation génomique

L'endocan ou ESM-1 a été cloné à partir d'une banque d'ADNc de cellules endothéliales humaines (HUVECs) en 1996 par l'équipe de Philippe Lassalle (U416, INSERM, Institut Pasteur de Lille) (Lassalle et al., 1996). La partie protéique de l'endocan est le produit d'un seul gène *esm1*, localisé sur la partie proximale du bras long du chromosome 5 (5q11.2). Ce gène *esm1*, qui s'étale sur 12 kb, est composé de trois exons séparés par deux introns (Béchard et al., 2001a). Les exons 1, 2 et 3 comptent respectivement 362, 150, et 1560 paires de bases. L'exon 1 et une partie de l'exon 2 codent pour la région N-terminale de l'endocan (110 premiers acides aminés) riche en cystéines (18 Cys) (Figure 17). L'exon 2 code

également pour un domaine riche en phénylalanines ( $^{113}\text{FPFFQY}^{118}$ ) qui semble important pour l'activité de l'endocan (Scherpereel et al., 2003). Enfin, le plus grand de tous, l'exon 3, code pour un court domaine C-terminal de 33 paires de bases (à cause d'un codon stop dans l'exon). Cette région C-terminale porte sur la sérine 137, la O-glycosylation responsable de l'ancrage d'une chaîne unique de dermatane sulfate (Figure 17) (Bécharde et al., 2001a).

L'ADNc d'endocan de 2006 pb contient un cadre fermé de lecture de 552 nucléotides débutant par un codon d'initiation ATG au sein d'une séquence consensus Kozak ( $\text{AAC}^{+1}\text{ATGA}$ ). L'analyse de l'ADNc d'endocan a mis en évidence outre les 552 pb qui permettent de dicter la traduction des 184 acides aminés d'endocan, une large séquence 3'-UTR (untranslated region) qui contient une série de AREs (AU-rich elements). Ce genre d'élément est assez fréquent chez les ARNm de cytokines et d'oncogènes. En général, les AREs sont reconnus comme des éléments *cis* qui affectent la stabilité des ARNm et qui peuvent être régulés par la MAP Kinase p38 (Mitogen Activated Protein Kinase p38) (Chen et Shyu, 1995 ; Frevel et al., 2003 ; Shaw et Kamen, 1986) ce qui signifie que l'ARNm d'endocan pourrait être instable. La région 3'-UTR contient également cinq sites potentiels de polyadénylation, seul le dernier est considéré comme fonctionnel (Lassalle et al., 1996).

Le promoteur de l'endocan humain a été cloné et séquencé par Tsai et al. (2002). Il est composé de 3888 pb (côté 5' du gène *esm1* bien sur) et contient les boîtes CAAT, GC et TATA ( $^{-39}\text{TATAAAA}$ ) caractéristiques (Tsai et al., 2002). Certains HSPGs tels que le perlécan ou certains CS/DSPGs comme le biglycan, ne possèdent pas de TATA box (Bernfield et al., 1999). Le promoteur de l'endocan contient également plusieurs sites potentiels de fixation pour les facteurs de transcription tels que trois motifs Ets (-77 à -74, -63 à -60 et +16 à +19), un élément CRE-like (cAMP Response Element) (-68 à -62) (qui possède 6 nucléotides sur les 8 que compte une séquence consensus CRE) et une séquence palindromique de 6 pb ( $^{-58}\text{GCATGC}^{-53}$ ). Une étude de ces séquences par mutagenèse a montré que les motifs Ets-binding (-77 à -74) et CRE-like (-68 à -62) doivent être notamment responsables du taux élevé d'expression d'endocan dans les cellules endothéliales. Par la méthode de l'empreinte ("footprinting") sur l'ADN du promoteur dans son intégralité, il a également été montré que le second site Ets (-63 à -60) et la séquence palindromique peuvent inhiber l'expression de l'endocan et ainsi être utilisés comme des répresseurs transcriptionnels (Tsai et al., 2002). De manière intéressante, les facteurs Ets-1 sont capables de réguler l'expression de plusieurs facteurs angiogéniques et autres facteurs responsables du remodelage de la matrice extracellulaire, notamment impliqués dans l'invasion tissulaire (Oikawa et Yamada, 2003). Les sites de fixation de type Ets pour les facteurs de transcription

sont retrouvés, par exemple, dans les promoteurs des protéines telles que les metalloprotéinases, enzymes responsables de la dégradation de certaines protéines de la matrice. De plus, ces sites de fixation Ets sont nécessaires pour l'activité maximale de molécules impliquées dans l'angiogenèse telles que Flt-1, Tie-1, Tie-2 et la cadhérine de l'endothélium vasculaire (Wakiya et al., 1996 ; Gory et al., 1998). Récemment, il a d'ailleurs été démontré que la protéine Hhex, un répresseur de l'expression de plusieurs gènes impliqués dans l'angiogenèse (VEGFR-1/Flt-1, VEGFR-2/KDR/Flk-1, Tie-1, Tie-2, neurophiline-1), est aussi un répresseur de l'endocan 8828 (Cong et al., 2006). Le promoteur du gène ESM-1 contient 10 séquences putatives de liaison de Hhex, dont la plus importante dans le contrôle de l'expression d'endocan semble située en 5' entre -8835 et -8828 (Cong et al., 2006). Parmi les autres séquences consensus dans le promoteur de l'endocan, trois motifs GATA, trois motifs AP4, deux motifs AP1, et un site NF1 ont été identifiés (Tsai et al., 2002).

## **2.2. Contrôle transcriptionnel d'endocan**

L'ARNm d'endocan est exprimé préférentiellement par l'endothélium (Lassalle et al., 1996) et son expression est constitutive (Nasu et al., 1999). La région -81 à +58 du promoteur et, plus particulièrement, le premier motif Ets (-77) et le motif CRE-like (-68), portent cette spécificité d'expression (Tsai et al., 2002). Ces deux motifs sont également retrouvés dans les promoteurs de gènes spécifiques de l'endothélium tels que le eNOS, l'ICAM-2 ou le VEGFR-1. Cependant, l'expression endothéliale d'endocan est plus forte et semble plus spécifique d'endocan que pour eNOS et l'ICAM-2 (Tsai et al., 2002).

L'ARNm d'endocan a été originellement décrit comme étant spécifique du poumon et du rein (Lassalle et al., 1996). Aucune expression n'avait pu être détectée dans le cœur, le cerveau, le placenta et le pancréas, qui sont pourtant des organes hautement vascularisés (Lassalle et al., 1996). Récemment, l'expression d'endocan à des niveaux beaucoup plus faibles que dans le poumon et le rein, a été mise en évidence dans des adipocytes isolés, des extraits de testicules, d'intestins grêles et de rates, ainsi que dans des cellules issues de glioblastome (Aitkenhead et al., 2002 ; Wellner et al., 2003 ; Janke et al., 2006 ; Adam et al., en préparation). Alors que par Northern Blot, un seul ARNm d'endocan avait été observé chez des cellules HUVEC, les RT-PCR réalisées par Aitkenhead et al. (2003) révèlent l'existence d'un épissage alternatif d'endocan. Des observations similaires ont été également faites par le groupe de Lassalle, notamment par Western Blot et par RT-PCR (Grigoriu et al., 2006). Ce phénomène de régulation génétique chez les eucaryotes supérieurs est couramment

utilisé et concerne bon nombre de PGs tels que les hyalectans (versican, agrécan, brévican) ou encore les PGs de type SLRPs (tels que la décorine et le biglycan) (Timar et al., 2002 ; Iozzo 1998 ; Prydz et Dalen, 2000). L'épissage alternatif peut être considéré comme l'un des modes principal de régulation de l'activité biologique des protéines (Brinkman, 2004), et encore plus significativement chez les PGs en permettant de moduler leurs profils de glycosylation. L'exemple des différents isoformes du CD44 qui a été décrit dans le chapitre sur les PGs en est ainsi une belle illustration.

La transcription du gène d'endocan est également sous le contrôle de nombreux facteurs. En effet, *in vitro*, il a été montré que des cytokines pro-inflammatoires telles que le TNF $\alpha$  (Tumor Necrosis Factor) et l'interleukine 1 (IL-1) sont capables de façon dose dépendante, d'augmenter la quantité d'ARNm d'endocan dans les HUVECs (Lassalle et al., 1996). L'IFN $\gamma$ , un autre médiateur inflammatoire, ainsi que l'IL-4 n'ont pas d'effet sur le taux d'ARNm d'endocan seul. En revanche, l'IFN $\gamma$  est capable de rendre nulle l'augmentation de la quantité d'ARNm induite par le TNF $\alpha$  sur des cellules endothéliales (Lassalle et al., 1996). Le TGF- $\beta$ 1 (et non le TGF- $\beta$ 2) a également été caractérisé comme inducteur de l'expression de l'ARNm d'endocan (Kirwan et al., 2005 ; Zhao et al., 2004). Cette information coïncide avec le motif Smad3/Runx2 identifié dans le promoteur qui répond au TGF- $\beta$ 1. L'augmentation du taux d'ARNm d'endocan dans des adipocytes a également été observée après un traitement avec l'acide rétinoïque ou le PMA (ester de phorbol). Cette stimulation semble emprunter la voie PKC, activant ainsi le facteur de transcription AP-1 qui a deux sites de liaison dans le promoteur d'endocan (Wellner et al., 2003). Un autre exemple avec l'enzyme lipoprotéine lipase (LPL) (responsable de l'hydrolyse des lipoprotéines circulantes riches en triglycérides), qui, par l'activation de HuR (protéine qui se fixe sur les séquences ARE et stabilise les ARNm), permet, entre autre, une augmentation de la quantité d'ARNm d'endocan dans des cellules endothéliales (Tschernatsch et al., 2006). Récemment, une étude a montré que deux hormones, l'hydrocortisone et l'insuline, diminuent l'expression de l'ARNm d'endocan dans les adipocytes. En revanche, de manière intéressante, l'IL-6 et le TNF- $\alpha$  ne modifient pas l'expression de l'ARNm d'endocan dans ce type cellulaire (Janke et al., 2006). D'autres molécules impliquées dans la croissance et surtout l'angiogenèse semblent également avoir une influence sur l'expression du gène d'endocan. En effet, le HGF/SF, le VEGF ou la combinaison de ces deux molécules augmentent très significativement le taux d'ARNm d'endocan dans les HUVECs et, cela, dès 24 heures après le traitement (Gerritsen et al., 2003 ; Rennel et al., 2007).

Chez l'Homme, de nombreuses études par biopuces ont montré que l'ARNm d'endocan est augmenté dans divers cancers (Sarrazin et al., 2006). Une étude récente sur les profils d'expression de plusieurs gènes dans des mélanomes plus ou moins agressifs révèle que dans le cas de mélanomes très agressifs, le taux d'ARNm d'endocan est multiplié par 44. Dans cette étude, l'endocan apparaît être un des gènes dont l'expression est la plus significativement augmentée, avec l'expression de Tie-1, gène également impliqué dans le cancer (augmentation de 100 fois) (Hendrix et al., 2003). De la même manière, dans un modèle d'angiogenèse, l'ajout de FGF-2 et de VEGF entraîne une multiplication par quatre des taux d'ARNm d'endocan (Aitkenhead et al., 2002). Ainsi, en plus d'être dépendant de certains médiateurs inflammatoires, le gène de l'endocan est sous le contrôle de médiateurs angiogéniques et de certains oncogènes.

### **2.3. Le core protéique d'endocan**

#### **2.3.1. Caractéristiques biochimiques**

Aucune évidence structurale ne permet aujourd'hui de classer l'endocan dans une famille de PGs décrite au chapitre précédent. En effet, chez l'endocan, aucun site de fixation aux lectines n'a pu être observé comme chez les hyalectans, ou aucun domaine riche en répétition de leucines (LRR) comme chez les SLRPs (Iozzo, 1998).

L'endocan peut cependant être caractérisé par son taux élevé de cystéines, à savoir 18 cystéines dans les 110 premiers acides aminés (Figure 17), ce qui représente 10,9 % de l'ensemble du core protéique. Cette région présente, de plus, des homologies avec d'autres familles de protéines. Par exemple, l'endocan présente 15 à 28 % d'homologies de séquences avec certains membres des IGFBP (Insulin Growth Factor Binding Protein). Ces protéines sont caractérisées par leur affinité pour les IGF et par leur séquence riche en cystéines (Lassalle et al., 1996 ; Hwa et al., 1999). L'endocan a, de ce fait, été classé comme correspondant à l'IGFBP-rP6 (Insulin Growth Factor Binding Protein rP6) (Hwa et al., 1999), même si aucune interaction de l'endocan avec un IGF n'a été caractérisée à ce jour. L'endocan a aussi des homologies de structure de l'ordre de 13 à 20 % avec d'autres protéines également riches en cystéines dans leur région N-terminale, les protéines CNN. Les protéines CNN incluent le CTGF (Connective tissue growth factor ou IGFBP-rP2), la Cyr-61 (Cystéine-rich 61 ou IGFBP-rP4) ou la protéine NOV (nephroblastoma overexpressed ou IGFBP-rP3). Ces protéines CNN sont principalement sécrétées et impliquées dans de nombreuses activités

biologiques comme la prolifération, l'adhésion, l'invasion ou l'angiogenèse (Desnoyers, 2004). Par ailleurs, l'endocan partage 74 % d'homologies de séquences en acides aminés avec la PG25, une protéine identifiée dans la glande pinéale du rat (Wang et al., 1997). La PG25 a donc été classifiée comme étant l'orthologue d'endocan chez le rat.

Biochimiquement, les 552 paires de bases du cadre de lecture codent pour une protéine de 184 acides aminés. A l'extrémité N-terminale, l'endocan possède une séquence de 19 acides aminés riches en résidus hydrophobes caractéristiques d'une séquence de type signal peptidique (Lassalle et al., 1996). Ainsi, le core protéique d'endocan mature, une fois le signal peptidique clivé, est composé de 165 acides aminés ce qui correspond à 18122 Da.

L'endocan mature apparaît cependant sur gel SDS-PAGE comme une protéine de 50 kDa et en gel filtration, comme une protéine globulaire de 400 kDa (Bécharde et al., 2000 ; Bécharde et al., 2001a ; Bécharde et al., 2001b). Ce constat s'explique par la liaison sur la sérine 137 d'une chaîne oligosaccharidique identifiée comme étant du dermatane sulfate (Figure 17). Par recherche bioinformatique, trois sites putatifs de O-glycosylation ont été identifiés (S16, T120, S137) (Lassalle et al., 1996). La mutation de chacun d'entre eux a affecté le poids moléculaire d'endocan dans le seul cas de la mutation sur la S137, indiquant que la glycosylation avait lieu sur ce site pour former le DSPG (Lassalle et al., 1996 ; Bécharde et al., 2001a).

### **2.3.2. Expression de l'endocan**

*In vivo*, chez l'Homme, l'endocan a été détecté sous forme de protéoglycane dans le courant circulatoire. Le taux sérique d'endocan a été dosé à 1 ng/mL chez les sujets sains (Bécharde et al., 2000). Le taux d'endocan circulant est largement augmenté dans les pathologies où il existe une stimulation endothéliale, comme chez les patients atteints de choc septique aiguë ou sévère et chez des patients atteints de cancer (Grigoriu et al., 2006).

En écho à la régulation de l'expression des ARNm d'endocan, les médiateurs inflammatoires ont également des effets sur l'expression du core protéique d'endocan. Ainsi, le TNF- $\alpha$  ou l'IL-1 $\beta$  permettent d'accroître la synthèse et la sécrétion d'endocan par des cellules endothéliales HUVECs *in vitro*. L'IFN $\gamma$ , en revanche, régule négativement la production d'endocan alors que, par exemple, l'IL-4 (interleukine-4) n'a pas d'effet (Bécharde et al., 2000).

## 2.4. L'endocan est un protéoglycane à dermatane sulfate

En 2001, l'ESM-1 (Endothelial Cell Specific Molecule-1), produit naturellement par les cellules HUVECs, a été caractérisé comme étant une protéine supportant une unique chaîne oligosaccharidique de DS. De cette observation, l'ESM-1 rejoint la famille des protéoglycannes à dermatane sulfate et le nom d'endocan lui est alors attribué (Bécharad et al., 2001a). La dégradation par la chondroïtinase ABC (spécifique des CS/DS et non des HS) de l'endocan permet, en gel SDS PAGE, d'observer une migration à 20 kDa alors que l'endocan non traité migre à 50 kDa. Ce résultat confirme l'appartenance de la chaîne oligosaccharidique d'endocan à la sous-famille glycosaminoglycannique des CS/DS. Il a également été observé que l'endocan n'est jamais sensible à la dégradation par l'héparinase, et n'est donc jamais produit comme un HSPG. De façon plus détaillée, la chaîne GAG d'endocan se trouve être sensible à la chondroïtinase B (spécifique des DS) ce qui permet d'affirmer que la chaîne oligosaccharidique d'endocan est du type chondroïtine sulfate B (CS-B), appelée également dermatane sulfate (DS) (Bécharad et al., 2001a).

Comme il a été décrit dans le premier chapitre, les fonctions des PGs sont dépendantes à la fois des cores protéiques et des chaînes oligosaccharidiques. Ainsi, le travail de Bécharad et de ses collaborateurs a permis de caractériser en détail la chaîne oligosaccharidique d'endocan (Bécharad et al., 2001a). L'analyse des disaccharides obtenus après digestion totale de la chaîne et séparation par HPLC révèle que la chaîne DS d'endocan est composée pour 66,5 % de GalNAc 4-O-S et pour 19,4 % de GalNAc 6-O-S. Les 14,1 % qui restent correspondent à des disaccharides di- ou non-sulfatés. De manière intéressante le taux d'IdoA ne dépasse pas les 6 % et ces IdoA semblent être regroupés. Cette épimérisation en C5 de l'acide glucuronique semble avoir un rôle très important dans les interactions avec les protéines et donc dans la régulation des activités des PGs (Delehedde, 2002 ; Lyon et al., 2002 ; Trowbridge et Gallo, 2002). De façon intéressante, une des premières fonctions qui a pu être attribuées à l'endocan, est celle de corécepteur pour le HGF/SF (Bécharad et al., 2001a). L'HGF/SF est principalement produit par les fibroblastes et les cellules vasculaires des muscles lisses pour agir en majeure partie sur les cellules épithéliales et endothéliales, ainsi que sur certaines cellules stromales (Delehedde et al., 2001 ; Bhowmick et al., 2004). L'endocan, via sa chaîne glycanique, augmenterait l'effet du HGF/SF sur la prolifération des cellules HEK293 (Bécharad et al., 2001a) ou des kératinocytes HaCat (Delehedde, Fernig, Lassalle, communication personnelle). Les paramètres cinétiques de l'interaction entre le HGF/SF et le DS d'endocan ont été analysés par résonance plasmonique de surface et un Kd

de 3 à 6 nM a été mesuré. Les interactions de l'HGF/SF avec l'Hp/HS ou les DS ont déjà été décrites (Lyon et al., 2002 ; Lyon et al., 1994 ; Lyon et al., 1998 ; Delehedde et al., 2002 ; Catlow et al., 2003 ; Lyon et al., 2004) et nous y reviendrons dans le chapitre III.

## **2.5. L'endocan ou ESM-1 (pour endothelial cell specific molecule) est il réellement un DSPG spécifique des cellules endothéliales ?**

Depuis sa découverte en 1996, dans les cellules HUVECs (Human Umbilical Vein Endothelial Cells), l'expression d'endocan avait été restreinte aux seules cellules endothéliales, (Lassalle et al., 1996 ; Bécharad et al., 2000 ; Gerritsen et al., 2003 ; Tsai et al., 2002). En effet, des transcripts d'endocan ont également été détectés dans d'autres cultures de cellules endothéliales de diverses origines telles que les HCAEC (Human Coronary Artery EC), HPAEC (Human Pulmonary Artery EC), HDMVEC (Human Dermal Microvascular EC) et HUEC (Human Capillary EC purifiée de tissus adipeux) (Tsai et al., 2002 ; Aitkenhead et al., 2002 ; Abid et al., 2006). Par comparaison, d'autres PGs ont été décrits chez les cellules endothéliales. Par exemple, le perlécan est abondant dans les matrices extracellulaires et dans la membrane basale de tissus normaux et de vaisseaux sanguins (Iozzo, 1998). D'autres PGs à CS/DS, tels que le biglycan ou le PG-100, sont produits par les cellules endothéliales (Nelimarkka et al., 1997). En revanche, contrairement à ce qui a été observé pour l'endocan, le TNF- $\alpha$  diminue l'expression du biglycan *in vitro* dans les cellules endothéliales (Nelimarkka et al., 1997). Parmi les PGs transmembranaires, les quatre membres de la famille des syndécans ainsi que certains isoformes de CD44 ont été décrits comme étant largement exprimés par les cellules endothéliales (Murakami et al., 2002 ; Henke et al., 1996 ; Kainulainen et al., 1996 ; Kulseth et al., 1999 ; Foster-Horvath et al., 2004). De même, l'expression du syndécan-1 est largement réduite en présence de TNF $\alpha$  et cela de manière dose-dépendante (Kainulainen et al., 1996).

Dans les cellules endothéliales telles que les HUVECs, l'endocan est également présent dans les lysats cellulaires et migre en gel SDS-PAGE avec un poids moléculaire apparent de 20 kDa, correspondant ainsi au polypeptide de 165 acides aminés ne possédant pas de chaîne oligosaccharidique (Bécharad et al., 2001a et b ; Bécharad et al., 2000). Aucun rôle n'a à ce jour encore été attribué à ce polypeptide.

Depuis sa découverte en 1996, l'expression d'endocan avait été quasi exclusivement observée dans des cellules endothéliales, d'où son nom d'ESM-1 pour Endothelial cell Specific Molecule 1 (Lassalle et al., 1996 ; Tsai et al., 2002 ; Aitkenhead et al., 2002 ;

Gerritsen et al., 2003 ; Rennel et al., 2007). Cependant, l'expression d'endocan a été récemment observée *in vitro* dans d'autres types cellulaires tels que les adipocytes, les cellules de mélanomes, des cellules rénales et également des cellules issues de glioblastomes (Wellner et al., 2003 ; Hendrix et al., 2003 ; Seftor et al., 2002 ; Ross et al., 2000 ; Abid et al., 2006 ; Adam et al., en préparation).

Grâce à des anticorps spécifiques, des analyses par immunohistochimie ont permis d'obtenir de nouvelles informations. Ainsi, dans les cellules HUVECs, la détection d'endocan se traduit par un marquage diffus dans le cytoplasme et aucun marquage membranaire n'a pu être visualisé. L'expression semble constitutive au niveau du poumon, au niveau des veines, des artères et des capillaires alvéolaires, ainsi qu'au niveau des cellules épithéliales des bronches et des glandes sous-muqueuses. Sur des coupes histologiques de reins, l'expression d'endocan est observée au niveau des cellules épithéliales tubulaires. Enfin, dans l'intestin, l'expression d'endocan a lieu au niveau des capillaires et des veines de la "*lamina propria*". L'expression d'endocan est ainsi détectée aussi bien dans les endothélias que dans les épithéliaux (Bécharde et al., 2000). Lorsque la lignée cellulaire HEK293 est transfectée avec de l'ADNc de l'endocan, un marquage diffus cytoplasmique est également observé (aucun marquage sur la membrane) (Scherpereel et al., 2003). De même, au niveau des adipocytes, l'expression d'endocan est localisée dans le compartiment intracellulaire (Wellner et al., 2003).

### **3. FONCTIONS D'ENDOCAN**

#### **3.1. Rôle dans l'inflammation**

L'implication de l'endocan dans les phénomènes inflammatoires a été une des premières fonctions identifiées. En effet, *in vitro*, il a été observé que l'ajout de LPS (LipoPolySaccharides) conduit à une augmentation de l'expression d'endocan dans les cellules HUVECs. De plus, une étude récente révèle qu'au cours du choc septique, le taux sérique d'endocan serait directement relié au stade d'évolution de la maladie. En effet, le taux d'endocan circulant augmente progressivement à mesure que l'on observe les patients atteints de sepsis, de sepsis sévère ou de choc septique. Les taux les plus élevés ont ainsi été retrouvés chez les patients qui décèdent dans les 10 jours (Scherpereel et al., 2006). Par ailleurs, comme il a déjà été mentionné précédemment, l'expression d'endocan est régulée par des cytokines inflammatoires telles que le TNF- $\alpha$  et l'IL-1. L'IFN $\gamma$  est capable de supprimer l'effet du TNF- $\alpha$  (Lassalle et al., 1996).

De plus, à l'image d'autres PGs, le core protéique d'endocan semble interagir directement avec d'autres protéines. *In vitro*, l'endocan, via son corps protéique, interagit avec les leucocytes (à la fois les monocytes et les lymphocytes) du sang chez l'Homme (Bécharde et al., 2001b). Parmi les autres PGs, il y a de nombreux exemples pour lesquels les interactions via le core protéique et via la chaîne oligosaccharidique coopèrent indépendamment pour exercer les fonctions biologiques du PG. Ainsi, dans la matrice extracellulaire, la décorine se fixe au collagène via son core protéique et à la tenascine via sa chaîne de DS (Iozzo, 1998 ; Elefteriou et al., 2001). Dans le cas de l'endocan, il a été mis en évidence que, via son core protéique, il est capable *in vitro* de se fixer aux intégrines CD11a/CD18, également appelées LFA-1 (Lymphocyte Function-associated Antigen 1). Un Kd non-négligeable de 18,7 nM a été déterminé par résonance plasmonique de surface (Bécharde et al., 2001b). Le principal ligand des LFA-1 est la molécule ICAM-1 (InterCellular Adhesion Molecule 1) (Anderson et Siahaan, 2003). Or, l'endocan *in vitro* est capable de se fixer aux lymphocytes et d'inhiber l'interaction avec de l'ICAM-1 soluble de manière dose-dépendante. Ces résultats suggèrent que l'endocan pourrait inhiber les phénomènes d'adhésion leucocytaire qui sont dépendants des LFA-1 (Bécharde et al., 2001b). Par ailleurs, l'augmentation du taux sérique d'endocan permettrait *in vivo* de limiter l'adhésion des leucocytes à l'endothélium dans un modèle d'inflammation chez le Rat (Tissier et al., 2004).

Pour enrichir les comparaisons entre l'endocan et les PGs issus des endothéliums, il convient de remarquer que les PGs de l'endothélium vasculaire ont, depuis longtemps déjà, été décrits comme étant des ligands de molécules impliquées dans le recrutement leucocytaire, l'un des premiers phénomènes à pouvoir être observé lors d'une réponse inflammatoire (Delehedde et al., 2005 ; Steeber et al., 2005 ; Parish, 2006 ; Taylor et Gallo, 2006). Le recrutement depuis le flux sanguin jusqu'aux tissus est régulé par une succession d'interactions entre les leucocytes et l'endothélium, grâce à des récepteurs incluant les intégrines, les sélectines et bien entendu les PGs de la surface cellulaire (Vaday et Lider, 2000 ; Schor et al., 2000). Durant le recrutement leucocytaire, les PGs endothéliaux, via leurs GAGs, permettent d'immobiliser et de présenter des messagers inflammatoires tels que MIP-1b (Macrophage Inflammatory Protein 1b), l'IL-8 (Interleukin 8) ou le PF-4 (Platelet Factor 4). Ces messagers peuvent induire l'activation d'intégrines qui vont permettre l'adhésion et la transmigration des leucocytes à travers les parois des vaisseaux (Steeber et al., 2005 ; Tanaka et al., 1993 ; Parish et al., 2006 ; Gallo et al., 2006). Dans ce contexte, il y a de nombreux exemples où les GAGs solubles, au même titre que les PGs solubles, sont capables d'inhiber le recrutement leucocytaire (Schor et al., 2000 ; Nelson et al., 1993 ; Lever et al., 2000).

Tumeurs	Nombre de patients pour l'étude	Marqueurs associés	Références
sein	78	Risque important de métastases	Van't Veer et al., 2002
rein	8	VEGF, c-met	Amatschek et al., 2004
rein	9	Augmentation de la vascularisation des tumeurs et de l'inflammation	Lenburg et al., 2003
utérus, estomac, rectum, sein, rein	14	VEGF, VEGF-R	Aitkenhead et al., 2002
poumon	23	Risque important de métastases	Borczuk et al., 2004
poumon	54	VEGF, Corrélié à la stimulation angiogénique des tumeurs	Grigoriu et al., 2006

**Figure 18: Expression d'endocan dans des tumeurs humaines**

(d'après Sarrazin et al., 2006)

Ainsi, même si à l'heure actuelle l'implication d'endocan dans le recrutement leucocytaire semble dépendant de son core protéique, il n'est pas exclu que sa chaîne glycanique joue également un rôle dans ce genre de processus.

### **3.2. Endocan dans la progression tumorale**

Les PGs sont en partie caractérisés par leurs capacités à interagir avec une multitude de partenaires tels que les facteurs de croissance, les chimiokines, les cytokines ou les enzymes. Ils sont donc consécutivement impliqués dans la régulation de processus tels que la prolifération, la différenciation, le remodelage tissulaire, l'angiogenèse, l'invasion tumorale et les métastases observées pendant la progression tumorale (Iozzo et Cohen, 1996 ; Iozzo et San Antonio, 2001 ; Iozzo, 2001 ; Timar et al., 2002 ; Grant et al., 2002 ; Delehedde, 2002 ; Sanderson et al., 2004 ; Reed et al., 2005).

L'endocan pourrait agir de la même façon au regard des effets de la chaîne glycanique observés vis-à-vis d'un facteur angiogénique, le HGF/SF (Bécharde et al., 2001). En effet, il a été montré que, via sa chaîne de DS, l'endocan est capable de réguler *in vitro* l'activité de l'HGF/SF suggérant un rôle potentiel de l'endocan dans le contrôle de la croissance et du développement tumoral, au même titre que dans l'angiogenèse lors de la progression tumorale.

Récemment, dans un modèle *in vivo*, les conséquences de la surexpression d'endocan dans la progression tumorale ont été décrites. Ainsi, l'injection sous-cutanée chez des souris SCID de cellules HEK293 modifiées génétiquement pour exprimer endocan induit la formation de tumeurs. Ce phénotype tumoral est dû, d'une part à la présence de la chaîne glycanique d'endocan (ce qui va dans le sens des interactions observées avec HGF/SF) et, d'autre part à la phénylalanine 116 du core protéique (Scherpereel et al., 2003). De plus, dans des tissus humains, l'augmentation du taux d'ARNm d'endocan est corrélée avec l'apparition de métastases et avec un mauvais pronostic dans différents types de cancers tels que les cancers du sein, du rein et des poumons (Figure 18). Dans le but de définir des marqueurs dans le cancer du sein, une étude de l'expression de 70 gènes dont celui d'endocan a été réalisée sur 78 patientes. L'expression d'endocan a, dans cette étude, été reliée étroitement au fort risque de métastase et à un décès des patientes dans les 5 ans (Van't Veer et al., 2002). De plus, parmi les 1234 gènes identifiés comme ayant une expression qui varie dans les cellules de carcinomes rénaux, le taux d'expression d'ARNm d'endocan est augmenté jusqu'à 3 fois et apparaît en corrélation, à la fois avec la vascularisation tumorale et l'inflammation dans le

cancer du rein (Lenburg et al., 2003). De même, dans une étude récente, le gène d'endocan est un de ceux dont l'expression est la plus augmentée (x 2) parmi les 8 échantillons de cellules de cancer de rein (comparées à des cellules normales). Il a également été mis en évidence que l'expression du VEGF et de c-Met (le récepteur du HGF/SF), deux protéines fortement impliquées dans l'angiogenèse, était également augmentée et, ce, de façon non négligeable (Amatschek et al., 2004). Dans une étude comparable, les expressions des gènes d'endocan, du VEGF et du VEGFR se trouvent être également augmentées dans des cellules rénales issues de biopsies de carcinomes (dans 5 cas sur 14) (Aitkenhead et al., 2002). Lors d'une étude de l'expression génique d'endocan sur 23 patients atteints du cancer du poumon, l'endocan fait partie des 42 gènes permettant d'évoquer un mauvais pronostic ainsi qu'un risque important de décès dans les 12 mois (Borczuk et al., 2004). Finalement, lors d'une autre étude, il a été également montré que l'augmentation du taux d'ARNm d'endocan était associée à des événements de l'angiogenèse (Figure 18), en particulier, au regard de la corrélation qui a été faite avec différents marqueurs spécifiques de l'angiogenèse comme le VEGF ou le VEGFR dans des échantillons de tumeurs de sein, de vessie et du côlon chez l'Homme (Aitkenhead et al., 2002).

Une étude récente (Grigoriu et al., 2006) démontre que dans les cancers broncho-pulmonaires, l'expression d'endocan est clairement augmentée au niveau des structures vasculaires tumorales. L'augmentation de l'expression de l'ARNm d'endocan a également été corrélée à celle du VEGF. Dans le sérum des patients, le taux d'endocan circulant se trouve être également augmenté avec un taux maximum pour ceux ayant le plus mauvais pronostic de la maladie (Grigoriu et al., 2006 ; Abid et al., 2006).

Comme on vient de le voir, dans les tumeurs humaines comme dans des modèles murins de xénogreffes, le taux d'ARNm d'endocan augmente donc très clairement. Il faut ajouter à cela les résultats des quantifications du taux d'endocan circulant réalisées par dosage ELISA. Ainsi, dans le sérum de patients atteints à un stade tardif du cancer du poumon, le taux d'endocan circulant se trouve être clairement supérieur à celui obtenu chez des individus sains (Scherpereel et al., 2003). Toutes ces informations font d'endocan un marqueur tumoral potentiel (pour revue, voir Sarrazin et al., 2006).

#### **4. CONCLUSION**

Depuis sa découverte en 1996, l'endocan ou ESM-1 apparaît être un protéoglycane à part entière, capable d'exercer potentiellement ses fonctions via son core protéique et via sa chaîne

glycannique. De plus en plus de données, suite à l'analyse de son promoteur jusqu'à la quantification de son expression dans des environnements sains ou pathologiques, laissent clairement penser qu'endocan aurait un rôle tant dans les phénomènes inflammatoires que dans la progression tumorale. A l'heure actuelle, ses caractéristiques de PG circulant et l'ensemble des connaissances acquises depuis 10 ans permettent de positionner l'endocan comme un marqueur potentiel dans le diagnostic de certaines pathologies, comme le choc septique et aussi de certains types de cancers. Il est également possible d'envisager qu'il puisse être une cible thérapeutique dans certains types de cancers (Sarrazin et al., 2006).

L'ensemble des connaissances disponibles sur l'endocan a fait l'objet d'une revue ajoutée en annexe de ce mémoire :

Sarrazin, S., E. Adam, M. Lyon, F. Depontieu, V. Motte, C. Landolfi, H. Lortat-Jacob, D. Béchar, P. Lassalle, M. Delehedde. (2006). "Endocan or endothelial cell specific molecule-1 (ESM-1): A potential novel endothelial cell marker and a new target for cancer therapy." Biochim Biophys Acta. **1765**(1): 25-37

**CHAPITRE III :**  
**LES INTERACTIONS PROTEINES-GAGS**  
**ET LEUR POTENTIEL DANS LE**  
**DEVELOPPEMENT DE MOLECULES**  
**THERAPEUTIQUES**

# 1. LES FONCTIONS DES GAGS VIA LEURS INTERACTIONS AVEC D'AUTRES MOLECULES

## 1.1. Introduction

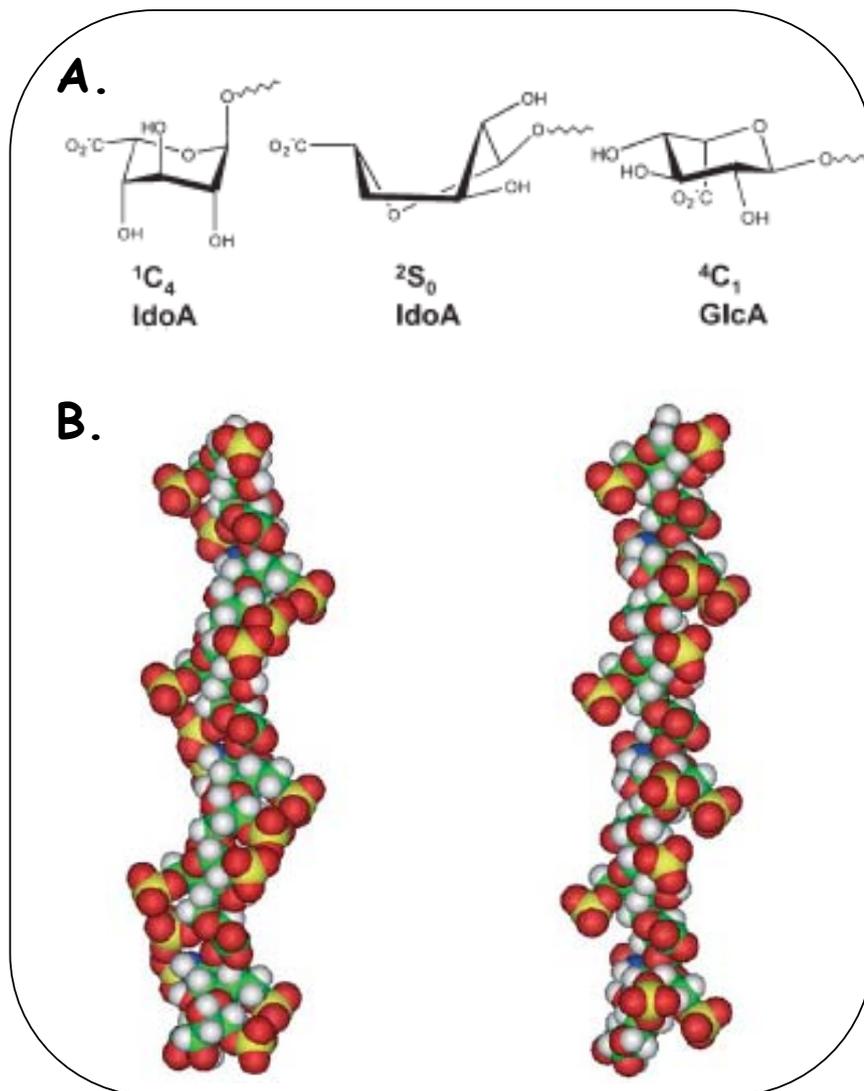
Comme il a été présenté dans le premier chapitre, les GAGs sont présents essentiellement dans les matrices extracellulaires ainsi qu'à la surface des cellules et peuvent également se trouver dans le courant circulatoire.

Dans ce chapitre, nous allons nous intéresser de manière préférentielle aux fonctions spécifiques des GAGs et à l'hétérogénéité de leurs structures qui leurs permettent d'interagir avec un très large panel de protéines. Un des objectifs de cet exposé est ainsi de mettre en avant les applications thérapeutiques qui sont, ou qui pourraient être, issues de l'étude et de la caractérisation des interactions GAG/protéines.

## 1.2. Les caractéristiques biochimiques et structurales d'une interaction GAG-protéine

**L'interaction d'un GAG avec une protéine est principalement de nature électrostatique.** Elle a lieu d'une manière générale entre certaines régions anioniques des chaînes oligosaccharidiques et des sites peptidiques ou protéiques de nature cationique tels que les groupements ammonium des lysines ou des arginines ( $\text{NH}_3^+$ ), le guanidinium de l'arginine ( $\text{pKa} = 13,6$ ) ou l'imidazolium de l'histidine (plus rare). Des centaines de protéines ont été identifiées comme interagissant avec l'héparine qui est le GAG le plus sulfaté et donc le plus anionique de tous. Cette propriété est d'ailleurs utilisée pour la purification de protéines par chromatographie d'affinité sur colonne d'héparine. Dans la plupart des cas, les interactions GAG-protéines pourraient paraître n'être que peu spécifiques (Lindahl et al., 1998 ; Salmivirta et al., 1996 ; Kreuger et al., 2006). Cependant, de nombreuses données permettent de penser que certaines séquences glycaniques ou du moins certaines régions glycaniques (domaines NS, NA/NS, NA) sont organisées pour interagir spécifiquement avec des protéines et contrôler ainsi leurs activités (Lindahl et al., 1998 ; Turnbull et al., 2001 ; Salmivirta et al., 1996).

**Les domaines de fixation à l'héparine** (Heparin Binding Domains) ont été déterminés dans de nombreuses protéines (Hileman et al., 1998). Ils **sont composés d'un nombre important de résidus basiques (lysines et arginines) et, dans certains cas, d'histidines.** La



**Figure 19: Conformation et flexibilité de l'acide uronique de l'Hp et des HS**

**A :** Représentation des différentes configurations de l'acide uronique dans l'héparine ou les HS. Tout au long d'une chaîne d'HS ou d'Hp, les résidus GlcN et GlcA sont stables dans la conformation  ${}^4C_1$ , alors que le résidu IdoA peut osciller entre les conformations d'énergie équivalente qui sont la  ${}^1C_4$  (chaise) et la  ${}^2S_0$  (bateau).

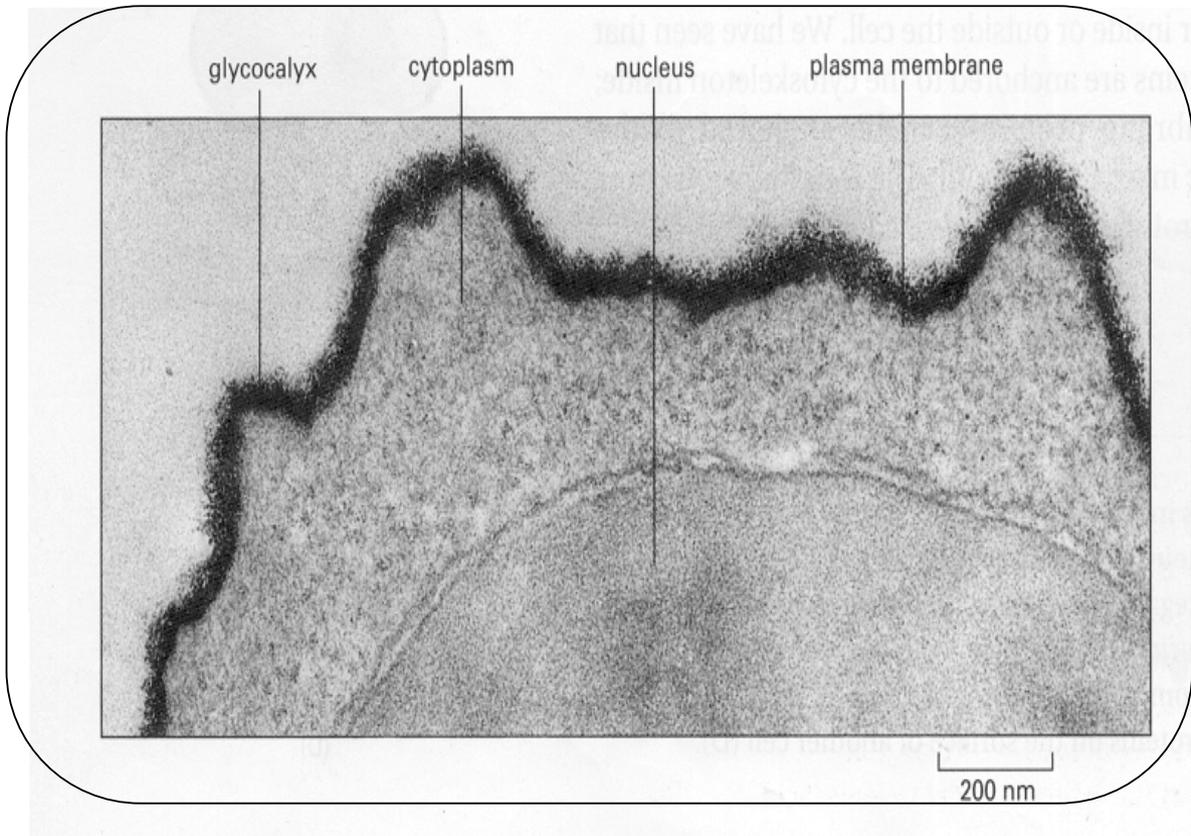
**B :** Structure de l'héparine obtenue par RMN. A gauche les IdoA sont dans une conformation  ${}^1C_4$ , et à droite dans une conformation  ${}^2S_0$ . Les sulfates sont représentés en rouge et jaune.

(D'après Coombe et Kett 2005)

structure du complexe FGF-2 avec un hexasaccharide par diffraction au rayon X a également révélé que des groupements fonctionnels de chaînes latérales d'autres acides aminés peuvent également interagir avec l'héparine, comme la fonction amine des chaînes latérales de l'arginine et de la glutamine (Faham et al., 1996). Les séquences consensus riches en acides aminés basiques sont du type : XBBBXXBX, XBBXBX et XBBBXXBBBXXBBX avec B comme acide aminé basique et X pour un acide aminé hydrophatique (Rabenstein, 2002). Ces séquences consensus peuvent ensuite adopter une structure secondaire qui permet le rassemblement des acides aminés basiques à l'extérieur et l'orientation des "X" vers l'intérieur de la protéine (Hileman et al., 1998).

Les propriétés électrostatiques de l'interaction GAG-protéine ont suscité la prise en considération de **la présence ou non de cations (en particulier pour le cas du Zinc et du Cuivre) dans les interactions.** En effet, de nombreuses protéines se fixent mieux sur l'Hp/HS en présence d'ions positifs. C'est le cas de la protéine du précurseur bêta-amyloïde, de l'interleukine 5 (IL-5), de la protéine prion, de "l'héparine cofactor II" ou encore de l'endostatine (Bush et al., 1994 ; Olsen et al., 1996 ; Lipscombe et al., 1998 ; Gonzales-Iglesias et al., 2002 ; Eckert et al., 2003 ; Ricard-Blum et al., 2004). Les cations peuvent avoir tout d'abord un effet sur les protéines. Ils peuvent favoriser l'oligomérisation, ce qui dans certains cas, facilite l'interaction des protéines avec l'héparine (Gonzalez-Iglesias et al., 2002). Ils peuvent également induire un changement conformationnel permettant l'interaction de la protéine avec l'héparine, comme dans le cas de l'héparine co-facteur II (Eckert et al., 2003). L'interaction de cations avec les GAGs ne dépend pas uniquement des forces électrostatiques puisque le  $Zn^{2+}$ , par exemple, va interagir sélectivement avec l'Hp et pas avec les autres GAGs (Parrish et Fair, 1981). Des données de RMN indiquent que l'acide iduronique est l'ose préférentiel pour la fixation des cations. Ces derniers seraient même capables de favoriser la conformation  ${}^1C_4$  par rapport à la  ${}^2S_0$  (Whitfield et Sarkar, 1991 ; Whitfield et al., 1992 ; Whitfield et Sarkar, 1992). **Le changement de conformation d'un acide iduronique dans une chaîne de GAG offre une flexibilité qui semble être importante et même nécessaire dans certaines interactions.** De plus, ce changement conformationnel permet de contrôler la manière dont le groupement 2-O-sulfate de l'IdoA (dans le cas où il est présent) peut s'agencer pour permettre les interactions avec les protéines (Coombe et Kett, 2005) (Figure 19).

De ce fait, si les cations sont capables de favoriser une conformation plutôt qu'une autre, il est évident que cela aura une répercussion au niveau des interactions avec les protéines. *In vivo*, l'effet de ces ions peut être uniquement local et ainsi apporter une réelle spécificité



***Figure 20: Le glycocalyx***

Le glycocalyx est une couche de polyholosides présente sur la surface membranaire externe des cellules animales ainsi que sur les parois externes de certaines bactéries. Sur cette image de microscopie électronique, le glycocalyx constitue une large bande sombre à la surface de la cellule. Dans certains épithéliums, l'épaisseur de cette couche peut atteindre  $0,5 \mu\text{m}$ .

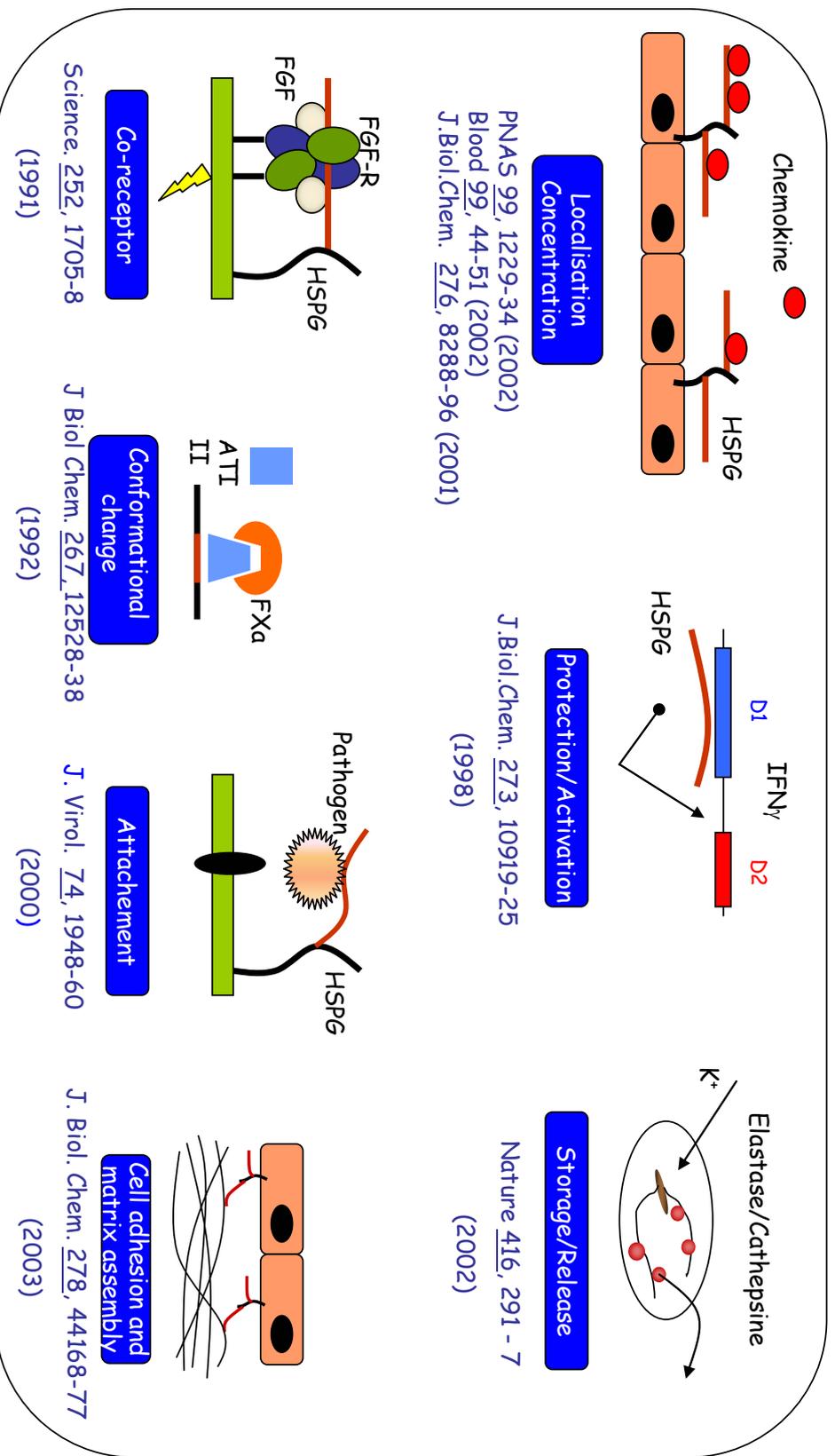
*(D'après : molecular biology of the cell, 4th edition)*

spatio-temporelle (Coombe et Kett, 2005). Par exemple, le taux circulant de  $Zn^{2+}$  dans le sang est très faible. En revanche, il est 30 à 60 fois plus élevé dans les plaquettes. Celles-ci contiennent également les PGs de type décorine et biglycan, connus pour avoir des cores protéiques qui stockent le zinc. Dans les sites où les plaquettes sont activées, une augmentation de la concentration en zinc pourrait alors avoir un rôle dans les interactions qui concernent les GAGs (Coombe et Kett, 2005).

Enfin, bien que la majorité des études d'interactions *in vitro* soient réalisées à pH neutre, il ne faut pas oublier qu'*in vivo*, le pH peut localement varier, et **la nature électrostatique de l'interaction GAG-protéine peut être dépendante du pH**. Par exemple, des lésions inflammatoires sont responsables d'une acidification locale créée par la production d'acide lactique par les neutrophiles (Lardner, 2001). Le pH des fluides issus de sites d'inflammation est de l'ordre de 6.1 (Trevani et al., 1999). Le microenvironnement tumoral est également plus acide, avec un pH d'environ 5.5 (Helminger et al., 1997). Les interactions GAG-protéines peuvent être réellement dépendantes du pH et notamment lorsque celles-ci mettent en jeu les histidines. Par exemple, "l'histidine-proline-rich glycoprotein" et la protéine prion, toutes deux dépendantes de cations, n'interagiront pas de la même façon via leur histidine lorsque le pH sera supérieur ou non au pKa de celle-ci (Borza et Morgan, 1998 ; Gonzalez-Iglesias et al., 2002). D'autres protéines sont sensibles au pH dans l'interaction avec les GAGs. C'est le cas de la forme non fibrillaire du peptide bêta-amyloïde, du GM-CSF (granulocyte macrophage colony stimulating factor) et même du VEGF (Gupta-Bansal et al., 1995 ; Wettreich et al., 1999). Dans le cas du VEGF, l'isoforme (VEGF<sub>121</sub>) qui ne possède pas la région d'interaction avec les GAGs, est capable, à un pH acide, d'interagir tout de même avec les HS (Goerges et Nugent, 2003). Le VEGF<sub>165</sub> qui possède la région d'interaction avec les GAGs a une affinité pour les HS augmentée lorsque le pH est abaissé. L'acidification permet en fait l'exposition d'un nouveau site de fixation aux HS dans les régions partagées par les deux isoformes du VEGF (Goerges et Nugent, 2003).

### 1.3. Les fonctions associées aux interactions GAG-protéines

A la surface des cellules, l'enveloppe créée par les protéoglycannes porteurs de différents GAGs est appelée **le glycocalyx**. Nous pouvons voir sur la Figure 20, une image du glycocalyx réalisée par microscopie électronique qui rend compte de la densité de cette "couche" de molécules oligosaccharidiques. Cette image traduit combien les chaînes polysaccharidiques sont importantes et, intuitivement, nous devinons qu'elles **doivent avoir**



**Figure 21 : Les fonctions des GAGs**

Ce schéma rend compte, grâce à des exemples, des différentes fonctions (liste non exhaustive) que les GAGs peuvent jouer dans divers processus. Ils permettent en particulier la création de gradient de chimiokines en localisant et concentrant celles-ci, ils jouent un rôle dans la protection des protéines contre la protéolyse. Leur bio-disponibilité permet également de stocker et relarguer certains facteurs. Les GAGs sont également bien connus comme étant des récepteurs d'attachement, notamment pour les pathogènes, et même des co-récepteurs sans lesquels certains signaux ne pourraient pas être transmis. Ils permettent bien sûr l'adhésion cellulaire et l'assemblage de la matrice en interagissant avec des molécules clés telles que les intégrines. Sans oublier les interactions avec les protéines pouvant créer des changements de conformation comme cela est le cas avec l'ATIII.

**des fonctions évidentes**, en particulier dans les grands processus biologiques nécessitant une communication entre les cellules. Ces processus peuvent être la prolifération, la migration, la différenciation ainsi que les interactions cellules-cellules ou cellules-matrice extracellulaire.

Parmi les fonctions principales des GAGs, peuvent être citées : **l'induction du changement conformationnel de certaines protéines** (cas de l'antithrombine III), **la participation dans les interactions protéine-protéine** (cas de l'antithrombine III et de la thrombine), **la fonction de co-récepteur** (exemple du FGF), ainsi que celle de récepteur ou plutôt **de site d'ancrage pour divers pathogènes** tels que le VIH. Les GAGs ont également **une utilité dans le stockage et le relargage de protéines** (pour l'élastase et la cathepsine par exemple), **ainsi que dans l'adhésion cellulaire et l'assemblage matriciel**. Les fonctions des GAGs dans **la localisation et la concentration** des chimiokines (exemple du SDF-1 $\alpha$ ), des cytokines (exemple de l'IFN $\gamma$ ), de facteurs de croissance ou autres médiateurs ainsi que dans **la protection de protéines vis-à-vis d'agents de la protéolyse** (exemple de l'IFN $\gamma$ , et du SDF-1 $\alpha$ ) peuvent être également citées (Figure 21).

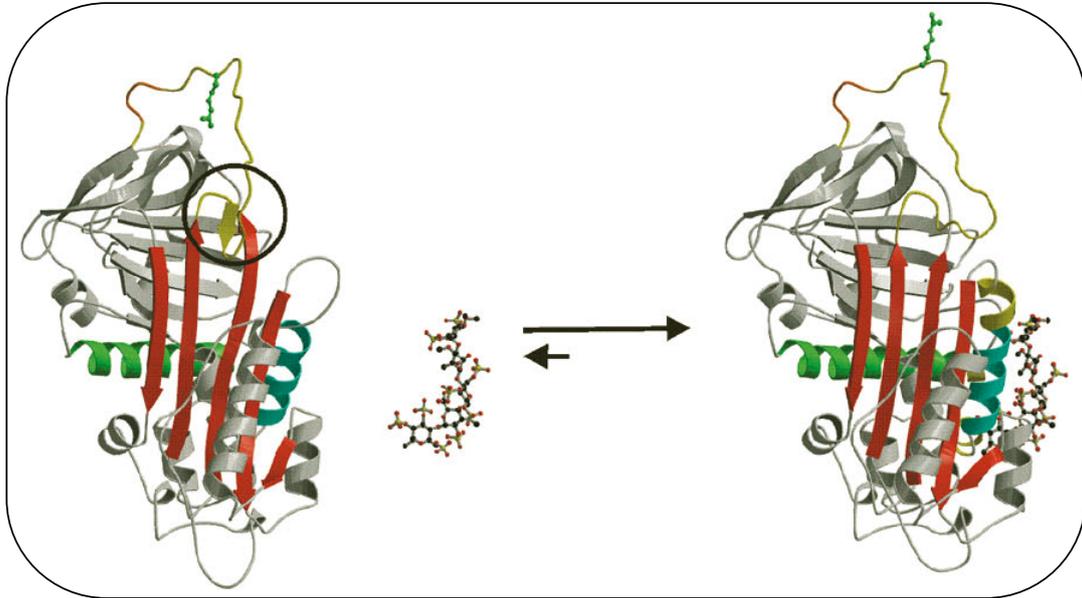
### 1.3.1. Les GAGs peuvent induire un changement conformationnel des protéines

L'interaction la plus étudiée ou, du moins, la plus caractérisée est celle de l'Hp et de l'inhibiteur de protéase à sérine : l'antithrombine III (ATIII) (Rosenberg et al., 1976 ; Andersson et al., 1976 ; Höök et al., 1976). Cette interaction induit un changement de conformation de l'ATIII (Figure 21bis) qui peut alors exercer son activité d'inhibiteur vis-à-vis de la thrombine ou du Facteur Xa qui sont les deux protéases à sérine impliquées dans la voie de coagulation (Olson et al., 1992). Une inhibition de l'une ou de ces deux protéases ralentit ou annule, en effet, le processus de coagulation.

Les études de l'interaction de l'Hp avec l'ATIII ont débouché sur le développement de mimétiques de l'Hp actuellement utilisés comme médicament ou déjà en cours d'essais cliniques, comme nous le détaillerons dans le paragraphe *2.1.Le pouvoir anticoagulant de l'héparine*.

### 1.3.2. Les GAGs participent aux interactions protéines-protéines

L'exemple le plus connu est celui de la formation du complexe ternaire : ATIII-thrombine-héparine. La fixation de l'Hp à l'ATIII induit un changement conformationnel de



***Figure 21bis: Changement de conformation de l'ATIII crée par un pentasaccharide issu de l'héparine***

L'Héparine catalyse l'inhibition de la thrombine par l'antithrombine. La figure montre que la fixation d'un pentasaccharide spécifique de type héparine induit un changement conformationnel de l'ATIII. Ce changement conformationnel se traduit par l'expulsion d'une région (entourée par un cercle noir sur la représentation de gauche) de la RCL (Reactive Center Loop représentée en jaune) en dehors de la partie globulaire de la protéine. L'expulsion de cette région augmente la flexibilité du RCL (Jaune) et libère une arginine (représentée en vert sur la boucle). Cette arginine et la flexibilité de la boucle vont permettre d'engager l'interaction avec la thrombine et son inhibition.

*(d'après Li et al., 2004)*

l'ATIII (Olson et al., 1992). Ce changement va lui permettre alors d'interagir avec la thrombine dans la mesure où le fragment d'Hp forme également un complexe avec cette dernière. Le modèle d'interaction entre les trois molécules est présenté dans la Figure 28a. Cette interaction et ses applications thérapeutiques sont détaillées dans le paragraphe : *2.1.Le pouvoir anticoagulant de l'héparine.*

### **1.3.3. Les GAGs comme co-récepteurs**

Pendant longtemps, en opposition aux récepteurs dits de "haute affinité" de type tyrosine kinase, les HSPGs et, particulièrement les GAGs, ont été définis comme des récepteurs de basse affinité (Israeli et al., 1989). Les termes de récepteur ou ligand peuvent être utilisés pour les GAGs lorsqu'ils permettent une interaction avec les protéines sans pour autant transduire le signal de ces dernières. Le terme de co-récepteur est, lui, utilisé lorsque le GAG est directement impliqué dans la transduction du signal d'un messenger en coordination avec son récepteur.

L'interaction la plus étudiée entre un facteur de croissance et des GAGs est, sans nul doute, celle des FGFs (Fibroblasts Growth Factors) avec les héparanes sulfates. Dans ce cas précis, les HS jouent un rôle de co-récepteur. Il existe aujourd'hui, au moins 23 FGFs différents répertoriés (Yamashita et al., 2005 ; Itoh et Ornitz, 2004) dont les plus étudiés pour leurs interactions avec les GAGs sont principalement le FGF-1 (FGF-a ou acide), le FGF-2 (FGF-b ou basique) et le FGF-7 (ou encore appelé KGF pour Keratinocyte Growth Factor).

En utilisant des cellules dépourvues d'HS, Yayon et al. (1991) ont initialement démontré que l'héparine est nécessaire pour que le FGF-2 se fixe à son récepteur FGFR1. Ils suggèrent que le polysaccharide agit en induisant un changement conformationnel du facteur de croissance nécessaire à la reconnaissance du récepteur (Yayon et al., 1991). A la même époque, Rapraeger et al. ont démontré que l'Hp/HS était nécessaire à l'activité du FGF-2 mais ont proposé que cela passe par la formation d'un complexe ternaire FGF2-Hp/HS-FGFR1 (Rapraeger et al., 1991). Dans ce modèle de complexe ternaire, la transduction de signaux via des récepteurs à tyrosine kinase (RTK) nécessite la dimérisation de ces derniers, permettant alors leur autophosphorylation.

Les controverses dans la formation du "complexe FGFR-FGF-HS-FGF-FGFR" ont été et sont toujours nombreuses. Ornitz et al. ont tout d'abord proposé que les HS permettent la dimérisation de deux FGFs (un octasaccharide d'HS serait la taille minimale pour dimériser ces FGFs). Ce complexe FGF-HS-FGF permet alors la dimérisation des récepteurs FGFR. Ce

modèle exclut l'interaction des HS avec le FGFR (Ornitz et al., 1992). En 1994, Kroizman et ses collaborateurs parviennent à la même conclusion en ajoutant que les FGFs se dimérisent de part et d'autre (*en trans*) de l'HS ce qui ne permet pas une interaction forte entre les deux FGFs.

La question sur l'orientation en *cis* ou en *trans* des FGFs sur les HS s'est alors posée. L'orientation *cis* des FGFs peut laisser penser que deux molécules se dimérisent puis que les HS viendraient stabiliser le dimère. C'est ce que proposaient Venkataraman et ses collaborateurs (1996). Les données structurales ont petit à petit apporté des informations supplémentaires mais ne permettent toujours pas de trancher sur l'organisation dimérique des deux FGFs le long d'un HS. La RMN serait en faveur d'un dimère FGF *cis* stabilisé par les HS (Moy et al., 1997) et certaines données laissent même penser que les HS peuvent, dans le même temps, interagir avec le FGFR (Safran et al., 2000). L'orientation *trans* suggère que deux FGFs se fixent à l'opposé l'un de l'autre sur la chaîne hélicoïdale d'Hp selon les résultats de cristallographie aux rayons X lorsque FGF1 et des décasaccharides d'Hp sont mis en présence (Di Gabriele et al., 1998).

La formation de ce complexe multimérique n'est cependant pas très claire. Les données apportées par Kan et al. (1993) rejoignent le modèle de Rapraeger dans lequel l'Hp interagit non seulement avec deux FGFs mais également avec deux FGFRs. Cela serait un élément essentiel pour la transduction du signal.

Les progrès réalisés grâce à la production de l'ectodomaine du FGFR recombinant par les bactéries *E.coli* ont permis d'apporter des données structurales sur la formation du complexe FGF-FGFR-Hp. Dans les structures obtenues, il apparaît clairement que le FGF se lie au niveau des régions D2, D3 et de l'inter-région D2-D3 du FGFR (Pellegrini et al., 2000). Les structures de FGF-FGFR seules révèlent une stoechiométrie 2:2 avec l'apparition d'un "canyon" de charges basiques qui correspondent avec la présence de l'Hp dans les conditions où elle a été ajoutée. Le désaccord sur l'agencement de l'Hp dans ce canyon ainsi que sur l'ordre des interactions entre les trois molécules est discuté dans deux revues de Mohammadi et al. (2005a ; 2005b).

D'une manière générale, il est désormais admis que le FGF transmet son signal via son récepteur FGFR et en présence du co-récepteur de type Hp/HS (Mohammadi et al., 2005a ; 2005b). Les FGFs et FGFR forment un complexe 2:2 qui doit être stabilisé par l'Hp/HS. La présence du GAG doit avoir un effet synergique sur les interactions en renforçant l'interaction FGF-FGFR et en induisant le rapprochement FGFR-FGFR. La dimérisation du récepteur

permet alors la transduction du signal par autophosphorylation (Mohammadi et al., 2005 a et b).

Les GAGs ont dans cet exemple, une fonction de co-récepteur et montrent à nouveau qu'ils permettent la stabilisation d'interactions protéine-protéine.

#### **1.3.4. Les GAGs peuvent servir de récepteur d'internalisation**

Comme il a été dit précédemment, le recyclage des PGs s'effectue par internalisation, transportant les PGs de la membrane jusqu'aux lysosomes situés dans le cytoplasme. Il n'est pas exclu que, lors de l'internalisation, des protéines liées aux GAGs des PGs soient également co-internalisées. Ceci est notamment le cas pour les protéines telles que le complexe follistatine/activine, la vitronectine, la thrombospondine, le FGF-2 ainsi que l'ATIII et la lipoprotéine lipase dont la clairance plasmatique dépend en grande partie des PGs (Bernfield et al., 1999). La lipoprotéine lipase (LPL) est, par exemple, une enzyme produite par les adipocytes qui est essentielle dans le métabolisme lipidique. Sous forme monomérique, elle n'est pas active et interagit peu avec les GAGs. En revanche, sa dimérisation lui permet de devenir fonctionnelle et d'interagir très fortement avec les HS ( $K_d = 0,3 \text{ nM}$ ) (Lookene et al., 1997). Cela lui permet d'être localisée à la surface cellulaire où elle peut fixer et agir sur les lipoprotéines. Ces interactions facilitent l'endocytose et le catabolisme des lipoprotéines et de la lipase (Rabenstein, 2002). Des études ont montré que les syndécans et les perlécans étaient des récepteurs de la LPL, permettant son internalisation (Fuki et al., 2000 a ; Fuki et al., 2000 b).

#### **1.3.5. Les GAGs participent à l'organisation de la matrice et à l'adhésion cellulaire**

La localisation des GAGs et leurs caractéristiques physico-chimiques leur permettent également d'exercer des fonctions dans l'organisation de la matrice et dans l'adhésion cellulaire.

D'une manière générale, les GAGs, tels que les CS/DS ou encore l'AH, jouent un rôle très important dans tous les tissus cartilagineux. En effet, ils permettent d'une part l'hydratation de ces tissus et, d'autre part, offrent une résistance mécanique face aux forces de compression, grâce à leur organisation et aux interactions qu'ils créent avec de nombreux partenaires de la matrice extracellulaire (Roughley et Lee., 1994).

D'un point de vue plus moléculaire, l'interaction entre les chaînes d'HS portées par des PG de la surface cellulaire et la fibronectine est, par exemple, très importante dans les adhésions cellules-matrices (Sekiguchi et al., 1983). Dans les interactions du même type, les GAGs peuvent interagir avec certains collagènes. Le collagène V interagit ainsi avec les HS pour renforcer l'adhésion cellulaire (LeBaron et al., 1989). Le collagène IV peut lui interagir via sa partie non collagénique avec les DS, alors que dans sa région composée de triples hélices, l'interaction sera spécifique des Hp/HS (Koliakos et al., 1989). Les collagènes I à VI ainsi que le collagène XIV sont connus pour posséder des sites de fixation à l'Hp/HS ou aux DS (Delehedde et al., 2001). Toutes ces interactions font partie intégrante de l'organisation de la matrice et ont un rôle majeur dans l'adhésion cellulaire via les PG de surfaces. La laminine, une protéine de la matrice extracellulaire essentiellement présente dans les membranes basales, interagit également avec les GAGs de type HS ou CS (Sung et al., 1997 ; Talts et al., 1999). Cette interaction semble importante pour l'adhésion cellulaire mais également lors de la polarisation cellulaire des épithélia (Delehedde et al., 2001).

Les HS sont également connus pour fixer des molécules d'adhésion telles que PECAM-1 (Platelet Endothelial Cell Adhesion Molecule-1) ou N-CAM (Neural Cell Adhesion Molecule) (Bernfield et al., 1999).

*In vitro*, des cellules CHO (Chinese Hamster Ovary Cell) dépourvues d'HS ne sont plus capables de s'attacher sur une surface recouverte de fibronectine et ne forment plus d'adhésion focale et de "F-actin stress fibers" (LeBaron et al., 1988).

Les CS se trouvent également être impliqués dans des interactions avec des molécules de la matrice extracellulaire ou de la surface des cellules. Le versican, par exemple, interagit avec des molécules telles que les L-sélectines, les P-sélectines (des molécules d'adhésion de la surface cellulaire des leucocytes ou des cellules endothéliales) et encore les CD44. Ces interactions peuvent être inhibées par les CS E, par les DS ainsi que par les HS, uniquement dans les interactions concernant les L- et P-sélectines. Outre les HS, tous les CS et l'AH sont impliqués dans les interactions concernant le CD44 (Kawashima et al., 2000).

### **1.3.6. Les GAGs permettent le stockage et le relargage de protéines**

Dans les mastocytes, lieu de production et de relargage de l'Hp, il a été démontré que l'Hp était essentielle pour le stockage de protéases telles que les MCP-4 ou MCP-5 (Mast Cell Protease), ou la carboxypeptidase A (MC-CPA) (Humphries et al., 1999 ; Forsberg et al., 1999).

Dans le même registre, il a également été montré que les neutrophiles étaient capables d'agir face aux micro-organismes sous l'influx d'un signal potassique dépendant du pH, grâce à un relargage de protéases telles que l'élastase et la cathepsine G qui sont stockées associées à des GAGs (Reeves et al., 2002).

Par ailleurs, les facteurs de croissance tels que les FGFs et les VEGF sont stockés, stabilisés, et protégés grâce aux interactions avec des chaînes de GAGs dans les membranes basales. Or, dans les phénomènes d'angiogenèse par exemple, sous l'action d'un stimulus tel que l'hypoxie et par le biais de l'héparanase (connues pour couper les HS), ces facteurs sont libérés et exercent à leur tour leurs fonctions angiogéniques (Vlodavsky et al., 2002).

### **1.3.7. Les GAGs jouent le rôle de récepteurs d'attachement pour les pathogènes**

Outre les fonctions des GAGs essentielles à l'organisme qui les génère, les caractéristiques des glycosaminoglycannes peuvent également être utilisées par les pathogènes tels que certains parasites, bactéries ou virus.

Parmi les pathogènes se servant des GAGs comme récepteurs d'ancrage, il y a les parasites tel que *Plasmodium falciparum* (responsable de la forme létale de la malaria), (Barragan et al., 1999) qui se fixe en particulier sur les CS A (Fried et al., 2006). Les GAGs peuvent également servir de site d'attachement pour des bactéries telles que *Streptococcus pyrogenes*, *Streptococcus mutans*, *Staphylococcus*, *Helicobacter pylori* (Bernfield et al., 1999 ; Ruiz-Bustos et al., 2001).

De nombreux virus sont également très affins des GAGs qu'ils utilisent pour s'adsorber à la surface des cellules avant de fusionner avec une cellule cible. Parmi les nombreux virus connus pour se fixer aux GAGs peuvent être cités : le VIH (Virus de l'Immunodéficience Humaine), le virus Herpès et aussi le virus de la Dengue (Pour revue, voir Lortat-Jacob et al., 2005). Dans le phénomène d'infection, la fixation des virus aux cellules est une étape critique. Une fois le point d'ancrage trouvé et donc l'interaction avec les GAGs ou autres molécules d'ancrage faite, l'infection des cellules cibles sera plus aisée. Par exemple, il a été montré que les syndécans à la surface des macrophages étaient une classe abondante de "récepteurs" du VIH-1 (Saphire et al., 2001). D'autres travaux montrent que les syndécans sont capables de capturer et de protéger le virus sur des cellules endothéliales. Le virus garde alors toutes ses propriétés infectieuses, dans l'attente d'être présenté aux lymphocytes T (Bobardt et al., 2003). Les revues de Spillmann (2001) et de Lortat-Jacob et al. (2005) expliquent plus en détail l'importance de ces interactions et le rôle qu'elles jouent dans l'infection virale.

Certaines bactéries ont également une manière tout à fait originale d'utiliser les GAGs. En effet, *Pseudomonas aeruginosa* est, par exemple, capable via un de ses facteurs de virulence, le LasA, d'induire le relargage de fragments de syndécans après l'activation de métalloprotéinases. Les fragments d'HS des syndécans ainsi relargués se fixent alors à la bactérie et lui permettent de détourner les processus de défense de l'organisme hôte (Park et al., 2001).

### **1.3.8. Les GAGs localisent, concentrent et permettent les gradients de médiateurs**

Les GAGs sont largement impliqués dans les processus de signalisation cellulaire en modulant la biodisponibilité de messagers tels que les facteurs de croissance, les cytokines ou les chimiokines. La plupart des messages transmis par ce genre de protéines sont locaux et très souvent temporaires.

Les chimiokines, de petites cytokines responsables du chimiotactisme (**chemotactic cytokines** a donné "chemokines") sont, pour la majorité, des protéines avec un pI d'environ 9, possédant des régions riches en acides aminés basiques, qui leur permettent d'interagir avec les GAGs (Lortat-Jacob et al., 2002). Dès leur sécrétion, les chimiokines sont capturées par les GAGs ce qui évite leur diffusion et donc permet d'avoir une concentration maximum du médiateur à l'endroit de leur sécrétion. Grâce aux propriétés cinétiques de ces interactions chimiokines-GAGs, les chimiokines peuvent se déplacer de GAG en GAG et ainsi constituer un gradient qui correspond au chimiotactisme. Les cellules circulantes peuvent alors, par exemple remonter jusqu'à un site d'inflammation en suivant le gradient chimio-attractants de facteurs ainsi formés sur l'endothélium et dans la matrice extracellulaire (Handel et al., 2005 ; Tanaka et al., 1993 ; Gilat et al., 1994 ; Bernfield et al., 1999). Nous y reviendons dans : *1.4. Les GAGs dans l'inflammation.*

D'autres exemples peuvent être cités avec le MIP-1 $\beta$  (Macrophage Inflammatory Protein) et RANTES (Regulated upon Activation, Normal T-cell Expressed and Secreted) pour lesquels les HS de surface sont capables de les localiser, de créer un gradient proche de leurs lieux de production et également de les présenter à leur récepteur spécifique. Cela leur confère la capacité d'induire l'adhésion des lymphocytes T, une fonction qui est abolie lorsque les GAGs de surface ont été préalablement supprimés par voie enzymatique (Hoogewerf et al., 1997 ; Kuschert et al., 1999 ; Proudfoot et al., 2003).

De même, il a été montré que les cellules progénitrices hématopoïétiques (HPC) sont guidées par des gradients de SDF-1 créés grâce aux PGs (Netelenbos et al., 2002 ; pour revue, voir Adam et al., 2006). De la même manière, l'injection de polysaccharides sulfatés (fucoïdan) à des souris ou à des singes, permet de déplacer le SDF-1 $\alpha$  de la moelle osseuse au plasma, entraînant dans le même temps, un relargage dans le sang des HPC qui étaient confinées dans la moelle, là où le SDF-1 $\alpha$  était présenté par les PGs (Sweeney et al., 2002). Ce travail illustre parfaitement le rôle des GAGs dans la localisation de médiateurs locaux.

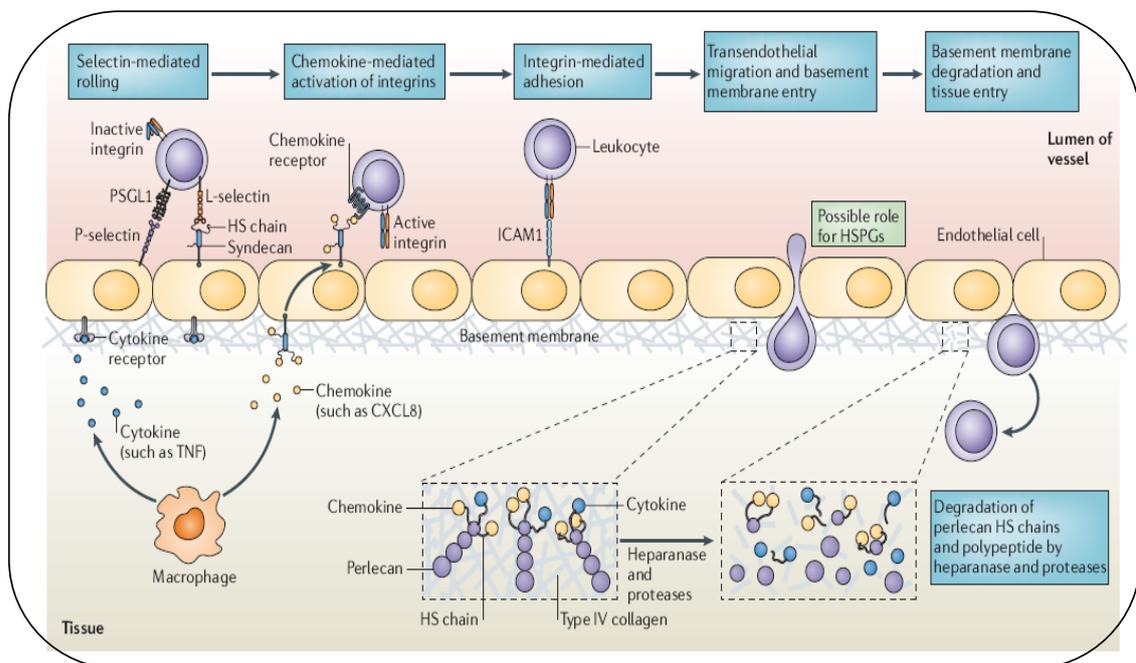
### **1.3.9. Les GAGs protègent certaines protéines vis-à-vis de la protéolyse**

Les médiateurs déjà cités tels que les chimiokines, cytokines et facteurs de croissance, sont utilisés par l'organisme de façon localisée et souvent très temporaire. Leur élimination se doit d'être très efficace lorsque les besoins n'existent plus. Ils sont, d'une manière générale, extrêmement sensible à la protéolyse, et donc le temps de demi-vie de ces molécules est en général très court (Overall et al., 2002 ; Flannery, 2006).

Le SDF-1 $\alpha$  est, par exemple, sensible à la protéase à sérine CD26/dipeptidyl peptidase (DPP IV). Cette protéase coupe la chimiokine à l'extrémité N-terminale ce qui lui fait perdre ses propriétés chimiotactiques. Or, il a été montré que les HS étaient capables de protéger SDF-1 $\alpha$  vis-à-vis de cette protéase (Sadir et al., 2004).

Les HS permettent également de protéger la protéolyse de facteurs tels que FGF-2, qui est lui sensible à la plasmine (Saksela et al., 1988) ou le TGF $\beta$  (Transforming Growth Factor  $\beta$ ) (McCaffrey et al., 1994) , l'IL-7 (interleukine-7) (Clarke et al., 1995), et l'IFN $\gamma$  (interféron gamma) dont les interactions avec les HS seront plus longuement détaillées dans le chapitre IV (Lortat-Jacob et Grimaud, 1991 ; Lortat-Jacob et al., 1996).

Dans un autre registre, les HS permettent de conserver l'intégrité de gp120, une glycoprotéine de l'enveloppe du VIH sensible à l'action de la PDI (pour protein disulfide-isomerase). Cette PDI clive deux des neuf ponts disulfures de gp120 pour permettre la fusion du virus avec la cellule cible. Or, il a été montré que l'interaction de gp120 avec les HS ne permettait plus à la PDI d'exercer son activité réductrice, exemple supplémentaire de la fonction protectrice des GAGs (Barbouche et al., 2005).



**Figure 22: Importance des HS pour le recrutement leucocytaire au cours de l'inflammation** Suite à un stimulus inflammatoire, les macrophages présent dans les tissus produisent des cytokines tel que le TNF (tumeur-necrosis factor) qui vont induire rapidement l'expression de molécules d'adhésions (P-selectine (Platelet) ou la E-selectine (Endothéliale)). L'interaction entre la P-selectine et des ligands comme le PSGL1 (P-selectin glycoprotein ligand 1) initie probablement la première phase d'adhésion des leucocytes sur la paroi endothéliale (**Rolling**). Ce phénomène de rolling est également consolidé par l'interaction entre les L-selectines des lymphocytes et les HS des cellules endothéliales. Ces HS sont localisés à la surface cellulaire sous la forme d'HSPG qui peuvent être les syndécans. Par ailleurs, le phénomène de **trancytose** des chimiokines au travers la paroi endothéliale est dépendant des HS. Par exemple, les héparanes sulfates sont capables de présenter aux leucocytes dotés de récepteurs à chimiokine, les médiateurs tel que CXCL8 (CXC-chémokine ligand 8) produit par les macrophages avoisinant. Ce processus de reconnaissance active les intégrines leucocytaires et facilite ainsi l'adhésion. Les HS des cellules endothéliales peuvent assister la migration trans-endothéliale en interagissant avec des protéines telles que le kinyogène (qui régule la perméabilité vasculaire). Une fois la migration des leucocytes à travers l'endothélium, ils vont être guidées par les chimiokines dans la membrane basale qui sont présentés par des HSPG tels que le perlecan, l'agrine ou le collagène de type VIII. La migration des leucocytes dans la membrane basale sub-endothéliale nécessite le déploiement d'enzymes et protéases utiles pour solubiliser la matrice. L'héparanase est une enzyme sécrétée par les leucocytes, elle permet de dégrader les HS, et va dans le même temps libérer les facteurs de croissance qui étaient associés aux HSPGs. Cela peut conduire à une réponse inflammatoire chronique en induisant les processus d'angiogénèse et de remodelage tissulaire.

(d'après Parish 2006)

## **1.4. L'implication des GAGs dans l'inflammation**

De par leurs interactions avec une large gamme de molécules via leurs chaînes de GAGs, les PGs se trouvent impliqués dans les différentes étapes observées lors des phénomènes inflammatoires tels que le recrutement de cellules de l'inflammation, le relargage, la diffusion et la régulation de médiateurs inflammatoires, ainsi que le réarrangement des matrices.

### **1.4.1. Les GAGs et le recrutement des cellules mobilisées lors de l'inflammation**

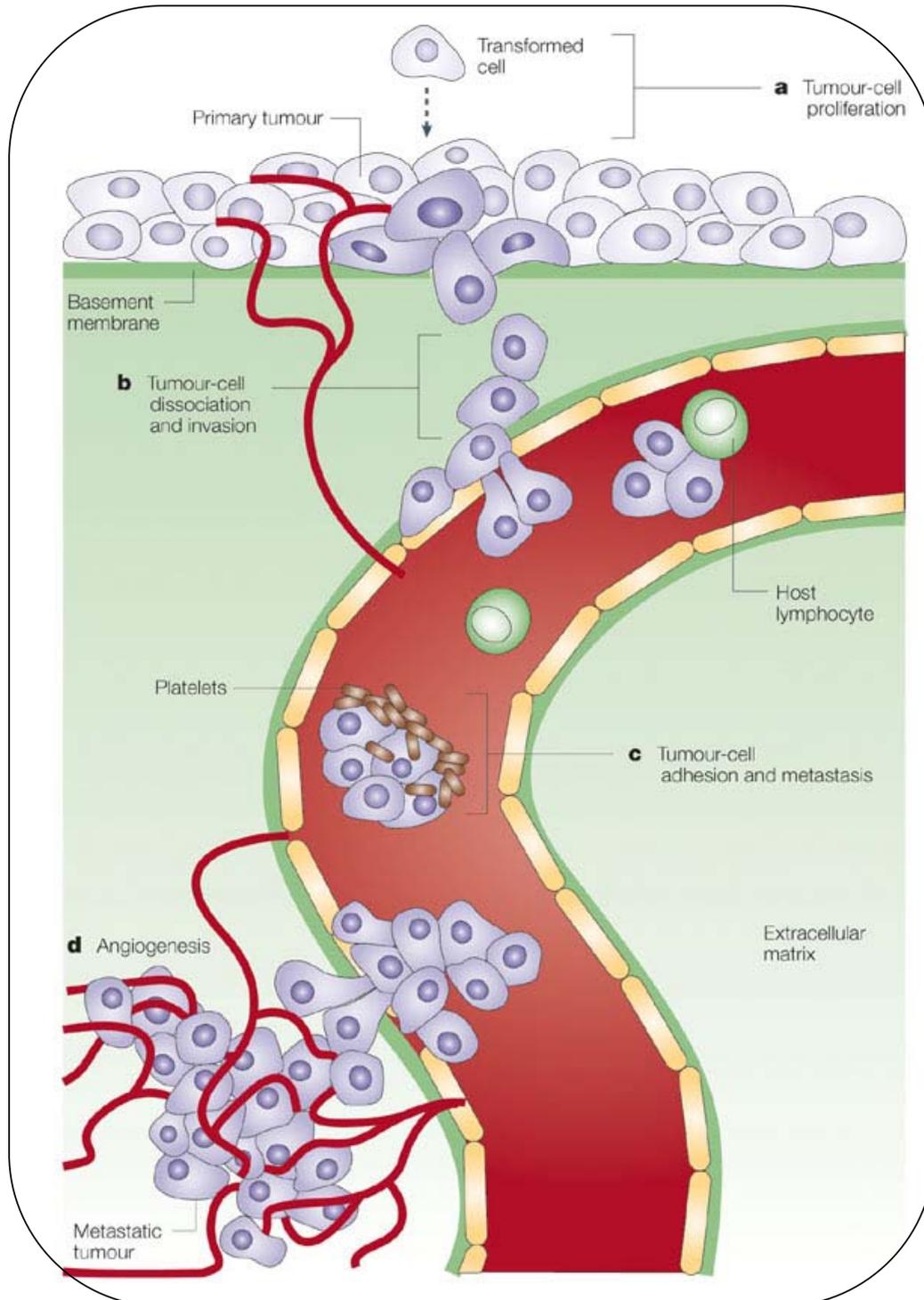
Lors d'une réponse inflammatoire, après l'activation des médiateurs de l'inflammation, l'un des premiers événements est le recrutement de cellules inflammatoires. Le recrutement de leucocytes dans les tissus est régulé par une succession d'interactions entre le leucocyte et l'endothélium rendues possible par des récepteurs d'adhésion tels que les intégrines, les sélectines et les PGs de surface (Figure 22). L'adhésion des leucocytes est en partie due à l'interaction de Mac-1 (Diamond et al., 1995) ou de la L-sélectine avec les HSPGs des cellules endothéliales (El Habbal et al., 1995, Wang et al., 2005). Dans des expériences vérifiant ce phénomène d'adhésion via les HS, l'injection de LMWH (Low Molecular Weight Hp) permet d'inhiber les phénomènes de "rolling", d'adhésion et d'extravasation de leucocytes à travers un endothélium stimulé au TNF $\alpha$  (Tumor Necrosis Factor) (Wan et al., 2001). Outre les interactions directes qui permettent l'adhésion des leucocytes, les PGs guident l'adhésion leucocytaire en immobilisant et présentant des chimiokines à la surface des endothélia, telles que MIP1 $\beta$  (Macrophage Inflammatory Protein 1 $\beta$ ), l'IL-8 ou le PF-4 (Platelet Factor 4). Ces chimiokines induisent l'activation des intégrines des leucocytes et sont également les médiateurs de l'adhésion et de la transmigration à travers la paroi vasculaire (Tanaka et al., 1993 ; Rot, 1992). Par ailleurs, l'expression de CD44 permet la fixation aux acides hyaluroniques. Cette interaction permet notamment l'adhésion des lymphocytes sur les parois endothéliales (Nandi et al., 2000). De plus, lors de phénomènes inflammatoires, l'expression de PGs peut généralement être induite. Leur présence peut alors modifier la fixation des cytokines et réguler la transduction de leurs signaux vers les cellules de l'inflammation. Par exemple, l'expression de syndécan-2 est induite sur les macrophages traités avec des médiateurs inflammatoires tels que le TNF- $\alpha$  et l'IL-8 $\beta$  (Interleukine 8 $\beta$ ) (Clasper et al., 1999).

### 1.4.2. Les GAGs et les médiateurs de l'inflammation

La réponse inflammatoire est orchestrée par de nombreux facteurs tels que les cytokines, les facteurs de croissance, les cyclophilines, les protéases, les antiprotéases et les endoglycosidases provenant des cellules du site inflammatoire ou encore du plasma (Vaday et Lider, 2000 ; Carpentier et al., 1999 ; Allain et al., 2002 ; 2003 ; Vanpouille et al., 2004 ; Melchior et al., 2006). Leurs interactions avec les chaînes de GAGs confèrent aux PGs des propriétés essentielles de la réponse inflammatoire. Comme nous l'avons déjà détaillé, ces interactions permettent en particulier de localiser, concentrer, créer des gradients chimioattractants de protéines ou de protéger les cytokines ou facteurs de croissance des dégradations protéolytiques. En plus de cela, il convient de prendre en compte toute la machinerie de synthèse et d'expression des PGs et les nombreuses enzymes régulées par des facteurs pro-inflammatoires, enzymes qui permettent de dégrader les GAGs et de remodeler les matrices extracellulaires facilitant ainsi les migrations et l'infiltration des cellules de l'inflammation (Delehedde et al., 2002).

Par exemple, vis-à-vis des chimiokines, qui sont le fil d'Ariane des leucocytes leur permettant de rejoindre le site d'inflammation, les HS sont connus pour avoir trois fonctions principales. La première est relative à l'interaction proprement dite puisqu'elle concerne la protection et leurs capacités à participer à l'oligomérisation des chimiokines (Sadir et al., 2004 ; Vives et al., 2002 ; Hoogewerf et al., 1997). La seconde fonction est d'immobiliser la chimiokine sur la surface laminaire des cellules endothéliales et de permettre la création d'un gradient chimioattractant (Middleton et al., 2002). Enfin la troisième fonction, concerne le processus de transcytose c'est-à-dire le processus selon lequel les chimiokines vont pouvoir traverser la barrière endothéliale. Ce cheminement est clairement rendu possible par les HSPGs, comme cela a été particulièrement bien décrit pour l'IL-8 (Middleton et al., 1997).

Une fois la première barrière endothéliale passée, les leucocytes doivent pouvoir gagner le site de l'inflammation et pour cela, franchir la membrane basale sub-endothéliale pour pouvoir continuer leur migration. Le perlécan, le collagène XVIII et l'agrine sont des constituants majeurs de la membrane basale formant un réseau très dense qui rend la migration cellulaire difficile. Les leucocytes doivent donc déployer une large gamme d'enzymes de dégradation. Ainsi, pour faciliter la migration, le désassemblage de la matrice est rendu possible par des enzymes, les protéases bien sûr, mais aussi par des enzymes plus particulièrement spécifiques des GAGs telle que l'héparanase qui coupe les HS. L'effet de l'héparanase est double : elle va à la fois faciliter la migration en désassemblant la matrice, et



**Figure 23: Les stades de la progression tumorale**

La prolifération tumorale (a) est la première étape, à la base de la progression tumorale. b | Pendant l'invasion, les cellules tumorales dégradent et migrent à travers les membranes basales et les matrices extracellulaires. c | Les cellules tumorales disséminées dans le courant circulaire s'agrègent avec des cellules hôtes tels que les plaquettes et les lymphocytes et éventuellement vont se loger dans les petits vaisseaux des organes à distance. d | Pour une croissance pathologique des tumeurs primaires ou des métastases, les processus d'angiogénèse sont nécessaires. Dans chacun de ces processus, les GAGs ou la dégradation des GAGs rentrent en jeu.

(d'après Fuster et Esko, 2005)

dans le même temps libérer les cytokines ou autres facteurs pro-inflammatoires stockés par les HS des PGs et ainsi accroître la réponse inflammatoire (Parish, 2006).

## **1.5. L'implication des GAGs dans les cancers**

Les GAGs et les PGs ont une importance majeure dans les processus liés au développement et la progression des cancers (Figure 23). Des changements dans leur niveau d'expression, ou dans celui des enzymes responsables de leur biosynthèse ou de leur dégradation (Narita et al., 2006) ont souvent été corrélés aux différents stades de la progression tumorale. De nombreuses données mettent en évidence l'implication des GAGs dans les cancers (Pour revue Sanderson et al., 2004 ; Toole, 2004 ; Sasisekharan et al., 2002 ; Iozzo, 2005 ; Yip et al., 2006 ; Filmus, 2001).

### **1.5.1. La croissance et la prolifération des cellules cancéreuses**

Une des caractéristiques des cellules cancéreuses est leur capacité à proliférer rapidement. Ce phénomène est en partie contrôlé par les HSPGs de surface qui servent, soit de co-récepteurs, soit tout simplement à localiser, concentrer et potentialiser l'activité de différents facteurs de croissance tels que le FGF-2, le FGF-1, ou le VEGF (Bernfield et al., 1999). C'est alors l'action de l'héparanase qui va les libérer de leur lieu de "stockage" en digérant les chaînes d'HS pour permettre d'exercer leurs activités (Vlodavsky et al., 2006). Les PGs à chondroïtine/dermatane sulfate (CS/DSPGs) semblent également jouer un rôle dans le contrôle de la transduction de signaux. Par exemple, le MCSP (Melanoma CS Proteoglycan) facilite la progression de mélanomes en induisant les voies ERK 1 et 2 (Extracellular signal-Regulated Kinase) et FAK (Focal Adhesion Kinase) qui sont deux voies importantes pour la croissance et l'invasion tumorale (Yang et al., 2004). La décorine, quant à elle, est capable, en modulant l'activité du récepteur EGF (Epidermal Growth Factor), de contrôler la prolifération cellulaire (Reed et al., 2005). Récemment, il a également été démontré que l'enzyme "hyaluronan synthase 2", en partie responsable de la synthèse de l'AH, était capable de réguler la signalisation via ErbB-2 sur des cellules cancéreuses du sein (Ghatak et al., 2005) et que sa suppression empêchait la prolifération de ce type cellulaire (Udabage et al., 2005).

### 1.5.2. Les phénomènes d'invasion et de métastase des cellules cancéreuses

Les capacités des cellules cancéreuses à migrer et à venir se loger dans les tissus avoisinants impliquent des changements dans l'expression des molécules de la surface cellulaire, ainsi que dans l'expression des enzymes responsables de la dégradation de la matrice extracellulaire.

Comme nous l'avons dit, les PGs et les GAGs sont des constituants majeurs de la matrice extracellulaire et l'expression des PGs à la surface des cellules permet les interactions cellules/matrice extracellulaire. Ainsi, des changements de leur expression peuvent réduire les interactions des cellules et promouvoir l'invasion des cellules cancéreuses. Par exemple, le versican produit par des cellules cancéreuses de prostate, inhibe l'adhésion cellulaire à la fibronectine (Sakko et al., 2003). De même, les syndécans en relation avec les intégrines (Bernfield et al., 1999) et l'acide hyaluronique via le CD44 (Toole 2004) contribuent à la mobilité des cellules cancéreuses, notamment grâce à l'activation du cytosquelette. De plus, les cellules cancéreuses ont la particularité de sécréter des enzymes telles que les metalloprotéinases, les héparanases ou les hyaluronidases pour pénétrer dans les membranes basales et dans les matrices extracellulaires afin d'atteindre les tissus avoisinants (Sanderson et al., 2004 ; Cohen et al., 2006).

Les métastases résultent de la dissémination de cellules cancéreuses dans la circulation et notamment des interactions que ces dernières peuvent créer avec les cellules endothéliales, les leucocytes ou encore les plaquettes. Toutes ces interactions guident la colonisation des organes ou tissus par les cellules tumorales. L'héparanase, par exemple, facilite les métastases en dégradant les HS (Maxhimer et al., 2002 ; Murry et al., 2006 ; Marchetti et al., 2001 ; Vlodavky et al., 2006). La dégradation partielle du syndécan-1 rend ce PG soluble et circulant. Cet ectodomaine ou forme non-glycosylée lui permet de réguler l'adhésion de cellules cancéreuses aux endothéliales des vaisseaux sanguins ou lymphatiques et, ainsi, de promouvoir leur association avec différentes cellules hôtes initiant les métastases (Götte et al., 2002 ; Yang et al., 2005 ; Ma et al., 2006). En revanche, *in vivo*, l'effet antimétastatique de l'héparine a été attribué à l'interférence qu'elle générerait avec les P-sélectines endothéliales (Ludwig et al., 2004).

### **1.5.3. Les GAGs et l'angiogenèse**

Au delà de 2 mm les tumeurs cancéreuses ou métastases ont besoin d'un support nutritif que seul le système vasculaire peut combler. C'est pourquoi, très rapidement autour d'une tumeur, le phénomène d'angiogenèse prend place. Les principaux facteurs guidant le processus d'angiogenèse sont les VEGFs (vascular endothelial growth factors), les FGFs (Fibroblasts Growth Factors) et les angiopoïétines. Cependant, les PGs et les GAGs sont également impliqués, puisque chacun des facteurs cités se lie à l'héparine ou aux HS.

Récemment, il a été montré que chez des souris déficientes en syndécan-1, l'angiogenèse était accrue (Götte et al., 2002). Chez les souris transgéniques qui à l'inverse surexpriment ce PG, les vaisseaux sanguins étaient tous anormalement dilatés (Elenius et al., 2004). Ce constat place le syndécan-1 dans une position où sa régulation apparaît essentielle au contrôle de l'angiogenèse et de la vascularisation.

Comme il a déjà été mentionné, l'héparanase permet l'induction de l'angiogenèse en libérant les facteurs angiogéniques tels que le FGF-2 et le VEGF (Zetser et al., 2006).

L'inhibition de l'expression du perlécan par un anti-sens ou l'observation de souris déficientes en perlécan révèle une diminution de l'angiogenèse autour des tumeurs et une baisse de la croissance tumorale (Sharma et al., 1998 ; Zhou et al., 2004).

Les CS peuvent, eux aussi, avoir un effet anti-angiogénique en inhibant la migration transendothéliale de monocytes (Liu et al., 2005). La décorine est également capable de supprimer l'angiogenèse tumorale en inhibant la production de VEGF par les cellules cancéreuses (Grant et al., 2002).

### **1.6. Les fonctions des Dermatanes Sulfates (DS) en particulier**

La plupart des fonctions de GAGs illustrées jusqu'à ce jour font souvent référence à des études qui prennent exclusivement en compte les HS ou l'héparine. Cependant, de plus en plus de données viennent nous éclairer sur l'importance des chondroïtines et des dermatanes sulfates (CS et DS). A l'image de l'Hp ou des HS, les CS/DS sont également impliqués dans le développement embryonnaire, l'homéostasie, le système cardiovasculaire, la tumorigenèse, ainsi que dans la réponse inflammatoire, la cicatrisation, les fibroses ou les infections. Ils semblent donc essentiels au contrôle de la destinée cellulaire. Au même titre que l'Hp ou les HS, les CS (dont le DS n'est qu'une variante résultant de l'épimérisation en C5 de l'acide glucuronique) ont une influence importante sur les différentes molécules présentes dans le

même micro-environnement. Les DS ont, en plus des HS, la particularité d'être les GAGs principalement exprimés dans des tissus particuliers comme la peau ou les cartilages. Par ailleurs, ils sont fortement relargués dans la circulation lors de la cicatrisation.

### **1.6.1. Les DS ont une activité anti-coagulante**

Une des interactions impliquant les DS qui est la plus étudiée aujourd'hui est celle qu'ils ont avec l'héparine cofacteur II (HCII). L'HCII est en fait un homologue de l'antithrombine III. Il inhibe donc les protéases à sérine telle que la thrombine, d'où la propriété anti-coagulante des DS. Cet effet peut être 1000 fois supérieur en présence soit de DS, soit d'Hp. Dans le cas de l'Hp, un complexe ternaire héparine/HCII/fibrine peut se former alors que dans le cas du DS, seule une interaction DS/HCII semble nécessaire. Les DS sont plus efficaces que l'héparine lorsque la thrombine est liée à la fibrine. Les DS pourraient être un bon moyen de traitement ou de prévention des thromboses (Liaw et al., 2001). Cette interaction semble relativement spécifique des sulfates et non de la charge globale du DS. Ainsi un DS de composition disaccharidique IdoA(2S)-Gal(4S) peut activer l'HCII alors qu'un DS de composition IdoA(2S)-Gal(6S) ne le fait pas (Mulloy, 2005).

Les DS peuvent également influencer la coagulation en potentialisant l'effet de l'APC (Activated Protein C) qui est un autre inhibiteur de la cascade de coagulation (Fernandez et al., 1999).

### **1.6.2. Les DS et la matrice extracellulaire**

Les DS sont abondants dans la peau et les cartilages, deux régions où les matrices extracellulaires ont une importance majeure (Towbridge et Gallo, 2004). Dans ces matrices extracellulaires, la chaîne de DS de la décorine est par exemple capable de se fixer à la ténascine-X, protéine colocalisée avec les fibres de collagènes (Elefteriou et al., 2001). Or, les patients déficients en ténascine-X ou les souris dépourvues de décorine présentent une peau d'une extrême fragilité (Burch et al., 1997 ; Danielson et al., 1997). Cela suggère que l'association entre la ténascine-X et les collagènes est rendue possible grâce à la chaîne de DS de la décorine et que cette fonction de "liant" est nécessaire pour l'intégrité de la peau.

### **1.6.3. Les DS et les facteurs de croissance**

Les facteurs qui lient les HS ou l'Hp les plus étudiés jusqu'à ce jour sont, sans aucun doute, les membres de la famille des FGFs (Harmer, 2006). Certains membres de cette famille interagissent avec les DS qui sont également capables de promouvoir l'activité des facteurs de croissance comme peuvent le faire les HS (dès lors que les fragments oligosaccharidiques sont assez longs) (Taylor et al., 2005). Les DS peuvent même avoir un effet plus important que les HS sur la potentialisation de l'effet du FGF-2 dans la prolifération cellulaire (Penc et al., 1998). Ainsi, lors de l'ajout de chondroïtinase B, les fibroblastes du derme deviennent moins sensibles à l'effet de prolifération induite par le FGF-2, en raison de la perte de leurs DS (Denholm et al., 2000 ; 2001).

Les DS se lient également au HGF/SF (Hepatocyte Growth Factor / Scatter Factor). Le HGF/SF est impliqué dans l'embryogenèse, la morphogenèse, la différenciation, la mobilité cellulaire, la régénération tissulaire et aussi dans l'angiogenèse (Kemp et al., 2006). Un dérèglement dans la signalisation via son récepteur tyrosine kinase (c-met) induit un processus de tumorigenèse ou de métastases dans de nombreux cancers. Les DS interagissent avec le HGF/SF mais avec une affinité 10 à 100 fois plus importante que les HS (Lyon et al., 1998 ; Delehedde et al., 2002 ; Catlow et al., 2003). Dans le cas des HS, cette interaction avec le HGF/SF nécessite au minimum un octasaccharide qui possède des IdoA non sulfatés avec une GalNAc (4S) (Lyon et al., 1998).

### **1.6.4. Les DS et les cytokines ou chimiokines**

Parmi la cinquantaine de chimiokines recensées, nombreuses sont celles qui ont la capacité de lier, via des rassemblements d'acides aminés basiques, les CS/DS. C'est le cas notamment de l'IL-8, des MIP-1 $\alpha$  et  $\beta$  (Macrophage Inflammatory Peptide), de RANTES (Regulated on Activation Normal T cell Expressed and Secreted) et de MCP-1 (Monocyte Chemoattractant Protein) (Kuschert et al., 1999).

L'exemple de la chimiokine RANTES peut être cité. Cette chimiokine est capable de lier les GAGs avec des affinités différentes. Ainsi Proudfoot et al. (2001) ont classé les affinités de RANTES vis-à-vis des GAGs de la manière suivante : Hp>DS>HS=CS. L'IdoA des DS semble donc être important puisque l'affinité pour les DS est 500 fois supérieure à celle des CS (Kuschert et al., 1999).

L'interféron gamma (IFN $\gamma$ ), qui fera l'objet du quatrième chapitre, est impliqué dans la réponse immunitaire ou inflammatoire. En plus de ses interactions bien caractérisées avec les HS, il interagit également avec les DS. Cette interaction avec les DS de mastocytes permet, indépendamment du récepteur à IFN $\gamma$  (IFN $\gamma$ R), de présenter la cytokine aux macrophages et d'induire, via l'IFN $\gamma$ R, la production d'oxide nitrique (NO) chez les macrophages (Brooks et al., 2000).

#### **1.6.5. Les DS et les pathogènes**

Outre les fonctions de récepteur d'attachement pour les pathogènes, les DS et DSPGs sont en partie utilisés par des pathogènes pour modifier la résistance de leur hôte. En effet, les pathogènes comme *Pseudomonas aeruginosa*, *Enterococcus faecalis* ou *Streptococcus pyogenes* relarguent des protéinases qui clivent la partie protéique des DSPGs tels que la décorine, libérant ainsi les chaînes de DS (Schmidtchen et al., 2001). Les défensines et les cathélicidines sont de courts peptides cationiques essentiels pour la résistance aux infections (Nizet et al., 2001). Les DS ainsi libérés peuvent se fixer et inhiber les peptides cationiques anti-microbiens tel que l' $\alpha$ -défensine. De ce fait, l'inactivation de ces peptides causée par les interactions avec les DS est un moyen efficace utilisé par les pathogènes pour augmenter leur potentiel infectieux.

#### **1.6.6. Les DS dans la cicatrisation**

Les DS sont un constituant majeur de la peau et représentent environ 0,3 % de son poids sec. Les deux couches supérieures de la peau sont le derme et l'épiderme. Le terme de dermatane sulfate dérive en fait de "derme", ce constituant de la peau directement en contact avec le monde extérieur. La localisation préférentielle de ces GAGs et leurs fonctions dans la coagulation, la croissance cellulaire et les défenses immunitaires sont des éléments importants qui favorisent leur implication dans les processus de cicatrisation.

Lors d'une blessure, parmi tous les PGs solubles sécrétés, la majorité sont des DSPGs (8 à 23  $\mu$ g/mL) (Penc et al., 1998). Ce mélange de PGs est capable de potentialiser l'effet du FGF-2 (facteur important pour la réparation des tissus) dans la prolifération cellulaire. Il a été mis en évidence que cet effet était dû majoritairement aux DSPGs et non aux HSPGs (Penc et al., 1998).

De même, il se trouve que seuls les DS (Hp, HS ou autres CS n'ont aucun effet) sont capables d'induire l'expression d'ICAM-1 (Intercellular Adhesion Molecule-1) par le NF- $\kappa$ b (Nuclear Factor-kappa b) et donc de permettre une meilleure adhésion leucocytaire sur les parois endothéliales lors de la cicatrisation, et d'une manière étroitement liée lors de processus inflammatoires (Penc et al., 1999).

## **2. APPLICATIONS THERAPEUTIQUES DES GAGS**

### **2.1. L'UTILISATION THERAPEUTIQUE DE GAGS NATURELS**

#### **2.1.1. Utilisation de l'acide hyaluronique**

L'acide hyaluronique est présent naturellement dans divers tissus du corps, comme la peau, le cartilage, et l'humeur vitrée de l'oeil. Il est donc adapté aux applications biomédicales visant ces tissus.

Le premier produit biomédical à base d'AH, Healon®, qui est toujours utilisé à l'heure actuelle, a été développé dans les années 1970 et approuvé pour des usages en chirurgie ophtalmique (i.e. transplantation de cornée, cataracte, glaucome, etc.). L'AH est également utilisé pour traiter l'ostéoarthrite du genou (Bellamy et al. 2006). De tels traitements sont administrés sous forme d'injections dans l'articulation du genou. On leur prête le pouvoir d'augmenter la viscosité du liquide synovial et de lubrifier ainsi l'articulation.

Dans certains cancers, les taux d'AH sont en corrélation avec la malignité et avec un pronostic défavorable. Dès lors, l'AH est souvent utilisé comme marqueur tumoral comme dans le cas du cancer de la prostate (Lokeshwar et al., 2005) et celui du sein (Corte et al., 2006). Il peut également être utilisé pour suivre l'évolution de la maladie.

L'AH est également utilisé en médecine esthétique depuis quelques temps.

#### *Dans le traitement de la gonarthrose (arthrose du genou)*

On sait depuis plus d'un siècle que la viscosité du liquide articulaire est due à la présence d'un polysaccharide. En 1934, Meyer et Palmer réussirent à l'isoler et à définir chimiquement sa structure primaire. Ils le baptisent alors "acide hyaluronique". Les années 50 ont permis de préciser sa structure macromoléculaire sous forme sodique libre dans le liquide synovial et de caractériser ses propriétés et ses fonctions rhéologiques dans l'écologie articulaire.

Nom commercial (année commercialisation)	Laboratoire	Poids Moléculaire moyen (Da)	N. d'injections / Dose unitaire
Artz (1987 - Japon)	Seikagaku (Japon)	600 - 1 200 000	5 / 2ml
Healonnon commercialisé	Biotrics Inc. (USA)	2 000 000	2 / 2 ml
Hyalgan (1988 - Italie)	Fidia (Italie)	500 - 750 000	3 à 5 / 2 ml
Synvisc (1992 - Canada)	Biomatrix (USA)	7 000 000 + gel	3 / 2 ml
NRD 101	Roussel (Japon)	1,8 - 2 000 000	essais avec 5
Orthovisc (1999 - Canada)	Anika (USA)	2 000 000	3 / 2 ml
Arthrum (France)	LCA (France)	annoncé pour 4 000 000	3 / 2 ml

**Figure 24: Les différentes préparations commerciales d'acide hyaluronique existantes (données de 2000)**

(<http://www.maitrise-orthop.com/corpusmaitri/orthopaedic/95/maheu/maheu.shtml>)

L'arthrose est une dégénérescence du cartilage des articulations qui conduit à la destruction plus ou moins rapide du cartilage qui enrobe l'extrémité des os. Le nombre de patients arthrosiques consultant chaque année en France est estimé à 3,5 millions. La gonarthrose ou arthrose du genou évolue au début le plus souvent sur un mode intermittent (poussées douloureuses, entrecoupées d'intervalles calmes), puis continue. Les symptômes principaux sont la douleur et la gêne fonctionnelle. Lors de poussées douloureuses inflammatoires, une hydarthrose nécessitant une ponction évacuatrice, peut avoir lieu.

L'AH est, à lui seul, responsable de la viscosité et de l'élasticité du liquide synovial des articulations. La diminution de sa concentration et de son poids moléculaire dans une articulation pathologique concourt à l'augmentation des phénomènes inflammatoires synoviaux et donc à l'aggravation de l'agression du cartilage. L'ensemble des propriétés que possède l'AH a conduit au développement du concept de «viscosupplémentation» (Peyron et al., 1993). Ce concept se fonde sur l'hypothèse, toujours discutée, que les injections intra-articulaires d'AH pourraient : i) restaurer les constantes rhéologiques du liquide synovial, ii) promouvoir la synthèse d'un AH de PM plus élevé et donc plus fonctionnel, iii) améliorer ainsi la mobilité et la fonction articulaire, iv) présenter en plus, une action antalgique.

En 1966, les premières injections intra-articulaires d'AH ont été réalisées chez des chevaux de course atteints d'arthropathies post-traumatiques. Ce traitement leur a permis de reprendre un entraînement normal (Nizolek et White, 1981 ; Butler et al., 1970). De même, Rydell et Balasz (1972), Peyron et Balasz (1974), puis Weiss (1980) ont injecté une fraction d'un sel sodique d'AH dans des genoux arthrosiques humains obtenant des résultats pertinents sur la symptomatologie.

Aussi, depuis une quinzaine d'années, les essais cliniques visant à apprécier l'efficacité de l'administration de l'AH dans l'arthrose, principalement au niveau du genou, se sont multipliés (Maheu et al., 1994). Le poids moléculaire des AH testés varie beaucoup, allant de 500 000-750 000 Da pour le Hyalgan<sup>®</sup> à 7 millions pour le Synvisc<sup>®</sup> (Figure 24).

A partir de ces différentes préparations dont la liste n'est certainement pas exhaustive, près de 40 essais cliniques réalisés dans la gonarthrose peuvent être recensés.

La Figure 25 présente les résultats des principales études menées avec une méthodologie conforme aux exigences actuelles des essais cliniques dans l'arthrose.

Pour synthétiser les résultats des études réalisées, la plupart des essais cliniques conduits avec l'AH mettent en évidence une supériorité sur le placebo dans la gonarthrose, quelque soit le produit utilisé. Trois essais se sont révélés négatifs. Le bénéfice clinique débute, selon les auteurs entre deux et quatre semaines après la première injection et peut se prolonger jusqu'à

Auteur (année)	Produit / contrôle	Schéma	N patients	N injections	durée suivi	Résultats
Dixon (1988)	Hyalgan / tampon Ph	DA, GP, R	63	11 en 23 semaines	11 M	AH > placebo de S5 à M6
Dougados (1993)	Hyalgan/excipient	SA, GP, R	110	4	12 M	AH > placebo à S7 jusqu'à 1 an
Moreland (1993)	Synvisc /arthrocent.	DA, GP, R	104	3	6 M	AH > pbo chez pts en poussée à JO
Adams (1993)	Synvisc /Sol saline	DA, GP, R	118	3	6 M	AH > placebo jusqu'à M6
Puhl (1993)	Artz /Sol saline	DA, GP, R	95	5	3 M	AH > placebo
Martin (1994)	Hyalgan /placebo	DA, GP, R	37	3	2 M	AH > placebo mais limite
Henderson (1994)	Hyalgan /Sol saline	DA, GP, R	92	5	5 M	AH = placebo
Scale (1994)	Synvisc /Sol saline	DA, GP, R	80	2 vs 3	12 S	3 inj > 2 > placebo
Carrabba (1995)	Hyalgan /Sol saline /arthrocent.	DA, GP, R	100	5 vs 3 vs 1	2 M	5 inj > 3 > 1, placebo, et arthrocent
Lohmander (1996)	Artz /Sol saline	DA, GP, R	240	5	4,5 M	AH = placebo; AH > pbo c/o > 65 ans
Wobig (1998)	Synvisc /Sol saline	DA, GP, R	110	3	3 M	AH > placebo
Altman (1998)	Hyalgan / AINS / pbo	DA, GP, R	495	5	6 M	AH > pbo et >= AINS à M6
Huskisson (1999)	Hyalgan / excipient	DA, GP, R	100	5	6 M	AH > placebo
Brandt (2000)	Orthovisc /Sol saline	DA, GP, R	226	3	6 M	AH > placebo

N = nombre ; DA = double aveugle ; GP = groupes parallèles ; R = randomisée ; M = mois ; AH = acide hyaluronique ; S = semaine ; SA = simple aveugle ; AINS = anti-inflammatoires non stéroïdien.

**Figure 25: Principaux essais cliniques publiés comparant l'utilisation de l'acide hyaluronique contre placebo dans la gonarthrose (données de 2000)**

(<http://www.maitrise-orthop.com/corpusmaitri/orthopaedic/95/maheu/maheu.shtml>)

six mois ou un an après le début du traitement. L'action est symptomatique et se marque par une diminution de la douleur et de la gêne fonctionnelle. La consommation d'antalgiques ou d'AINS (Anti-Inflammatoire Non Stéroïdiens) a été diminuée chez les patients traités. Les traitements à répétition semblent avoir encore plus d'effets bénéfiques (Kolarz et al., 2003).

Actuellement sur le marché français, plusieurs préparations d'AH sont disponibles, toutes commercialisées en qualité de dispositif médical avec : le Synvisc<sup>®</sup>, l'Arthrum<sup>®</sup>, l'Orthovisc<sup>®</sup> (ou Viscorneal Ortho<sup>®</sup>) et le Hyalgan<sup>®</sup>. Trois sont d'origine aviaire: le Hyalgan<sup>®</sup> et l'Orthovisc<sup>®</sup>, issus de crêtes de coq, ainsi que le Synvisc<sup>®</sup>, polymère de très haut poids moléculaire. L'Arthrum<sup>®</sup> est obtenu par fermentation de streptocoques. Ils sont indiqués dans la gonarthrose symptomatique, c'est-à-dire douloureuse. Ils apportent le bénéfice d'un traitement anti-arthrosique d'action lente mais prolongée.

#### L'acide hyaluronique en ophtalmologie

En ophtalmologie, en 1968-70, l'utilisation de l'acide hyaluronique (Healon), introduit sous l'endothélium cornéen comme facteur de « visco-supplémentation » en complément de la chirurgie de la cataracte est utilisé pour la première fois et connaît un succès rapide. De nos jours, des collyres à base d'AH sont accessibles pour une utilisation quasi-quotidienne lorsque l'oeil est irrité (Oxyl<sup>®</sup>).

#### L'acide hyaluronique pour les gencives

L'AH se trouve être également utilisé par les dentistes. En effet, des produits tels que le Hyalugel<sup>®</sup> (délivré en spray) sont utilisés pour favoriser la cicatrisation des gencives.

#### L'utilisation "semi médicale" de l'acide hyaluronique

L'utilisation "non thérapeutique" de l'AH est en plein essor. En effet, ses propriétés hydratantes (l'AH peut retenir jusqu'à mille fois son propre poids en eau) en font un ingrédient de base dans la confection de produits cosmétologiques, servant alors d'humectant de la peau et remplaçant peu à peu la glycérine. Outre ses capacités hydratantes, les laboratoires Pierre Fabre ont récemment découvert que des fragments courts d'AH sont capables de pénétrer dans la peau et permettent d'activer le mécanisme des rétinoïdes qui stimulent la synthèse de collagène et d'élastine. Ces fragments d'AH permettent également d'induire la synthèse naturelle d'AH par la peau. Cette activité régénératrice est par exemple utilisée dans la gamme Eluage<sup>®</sup> d'Avène.

Author(s)	Year	Nbr of Patients	Follow-up Period	Complications	Results
Rovetta	1991	40	25 wk	Well tolerated	Higher therapeutic effect on all symptoms of OA
Olivero et al.	1991	200	6 mo	3% mild adverse effect	Considerable improvement in both pain and mobility
Mazieres et al.	1992	120	3 mo	Well tolerated	Slowly effective against symptoms of OA and reduction of the need for NSAIDs
Morreale et al.	1996	146	3 mo	Minor	Benefits of CS appeared later but lasted for up to 3 months after treatment.
Fleish et al.	1997	56	1 yr	Well tolerated	Improvement in mobility, joint effusion and swelling
Bucsi & Poor	1998	80	6 mo	Minor	43% reduction in joint pain
Bourgeois et al.	1998	127	3 mo	No adverse events	Significant reduction in joint pain
Uebelhart et al.	1998	42	1 yr	None	Decreased joint pain and improved mobility
Verbruggen et al.	1998	119	3 yr	Not documented	Radiographic demonstration of decrease in number of patients with "new" erosive finger-joint OA
Uebelhart et al.	2004	120	1 yr	Minor adverse events	Decreased pain and improved knee function. Reduction of the progression of medial femorotibial joint space
Rovetta et al.	2004	24	2 yr	Not documented	Efficacy of oral CS in improving some aspects of erosive OA of the hands

**Figure 26: Tableau récapitulatif des différents essais cliniques concernant l'utilisation de chondroïtine sulfate dans l'ostéorathrose (OA)**

(d'après Volpi 2006)

Depuis quelques années, l'AH est également utilisé en médecine esthétique comme produit de comblement des rides. Des produits à base d'AH tels que Restylane®, Juvederm®, Hydracell®, Téosyal®, Hyaluderm®, Surgiderm®, sont utilisés par les médecins en remplacement du collagène, car les risques d'allergies sont moindres et l'effet dure plus longtemps (6 à 18 mois). Par ailleurs, sa résorbabilité empêche tout accident esthétique définitif.

### 2.1.2. Les CS pour soulager l'ostéoarthrose (OA)

L'ostéoarthrose (OA) et les douleurs musculo-squelettiques qui lui sont associées sont l'une des causes majeures d'invalidité aux Etats-Unis et dans de nombreux pays développés. Les causes de l'ostéoarthrose sont multifactorielles. Bien que l'âge soit un des facteurs principal, les facteurs mécaniques, hormonaux, environnementaux (sport, surpoids) et génétiques ne peuvent pas être exclus. L'OA qui touche les articulations, devient symptomatique lorsque les jonctions articulaires sont tellement endommagées que cela occasionne des gênes et des douleurs. La détérioration de ces jonctions articulaires est principalement due à une dégradation, une perte et/ou un non renouvellement du cartilage, constituant majeur des articulations. En plus des dommages occasionnés par l'absence de ces cartilages, les fragments de dégradation de ces derniers peuvent être antigéniques et ainsi provoquer une réaction inflammatoire qui vient s'ajouter aux douleurs dues aux frottements articulaires. Le cartilage des jonctions articulaires est composé à 98 % de matrices extracellulaires (dont principalement des collagènes fibrillaires et des PGs tel que le CSPG agrégé) et seulement par 2 % de cellules. Les fonctions principales conférées aux CSPGs sont l'hydratation et les propriétés de résistance aux compressions mécaniques du cartilage.

L'idée de "remplacer" les CS manquant dans le cartilage pour soulager l'OA a donc été testée (Figure 26) et différentes études cliniques se révèlent plutôt satisfaisantes même si ce n'est pas un traitement radical (McAlinton, 2000). Tous ces essais permettent aux laboratoires Genevrier de commercialiser : **Chondrosulf®** (ou Condrosulf) comme anti-arthrosique. Les CS sont également utilisés en association avec la glucosamine pour soigner l'ostéoarthrite. Les laboratoires SFB commercialisent par exemple le mélange sous forme de "gélules végétales".

En plus de reconstituer le cartilage, l'administration de CS peut avoir un effet secondaire tout à fait intéressant. L'administration orale de CS ou de Chondrosulf®, semble avoir un

important effet anti-inflammatoire (au moins chez le rat) ce qui va dans le sens du traitement nécessaire au soulagement de l'OA (Ronca et al., 1998 ; Distler et Anguelouch, 2006).

De manière anecdotique, les CS sont d'autre part utilisés en collyre pour des larmes artificielles (Lacrypos®).

## **2.2. LES MIMETIQUES DES GAGs**

### **2.2.1. Le pouvoir anti-coagulant de l'héparine**

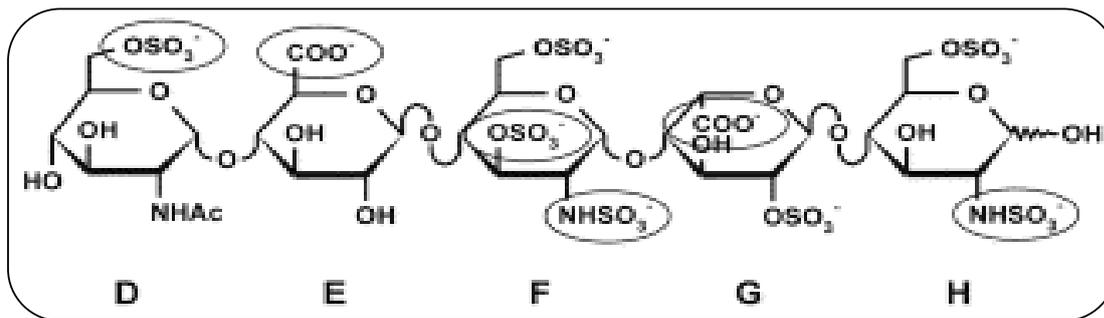
Lorsque l'on parle d'interactions entre les GAGs et les protéines, l'interaction de l'héparine avec l'antithrombine III (ATIII) est l'exemple de référence. Cette interaction est la mieux caractérisée et a permis le développement du premier mimétique de l'héparine, actuellement sur le marché, comme médicament anti-thrombotique. Parce que cette "histoire" est pour l'instant unique et inévitable lorsque l'on parle d'interactions GAGs-protéines, les étapes qui ont menées de la découverte de l'interaction au développement d'un médicament seront détaillées ci-après.

#### **2.2.1.1. L'héparine rencontre l'ATIII : une histoire fluide et "sang accroc"**

Le fait que la thrombine soit inhibée (et donc qu'elle ne puisse plus exercer ses fonctions d'agent coagulant) lorsqu'elle est ajoutée à un plasma sans fibrine (la fibrine étant le produit de l'action de la thrombine sur le fibrinogène permettant la coagulation) a été observé dès le 19<sup>ème</sup> siècle (Contejean, 1895). Cela a été associé à la présence d'un inhibiteur spécifique de l'enzyme appelé l'antithrombine ou ATIII. Quelque temps après, **en 1916, Mc Lean découvre par hasard l'existence de l'héparine ainsi que ses propriétés anticoagulantes** (Mc Lean, 1916). Ce n'est qu'en 1939 que l'Hp apparaît être un anti-coagulant efficace uniquement en présence d'un composé plasmatique, appelé initialement l'héparine cofacteur (Brinkhous et al., 1939). L'utilisation clinique de l'héparine comme agent anticoagulant se fait dès 1937 (Messmore, 2004). Le lien entre l'héparine cofacteur et l'ATIII n'a été fait que dans les années 1950. A ce moment là, **il était suggéré que l'activité de l'ATIII était accéléré par l'héparine** (Waugh et Fitzgerald, 1956 ; Monkhouse et al., 1955). Cette supposition a été validée en 1968 par les travaux de Abildgaard, qui purifia pour la première fois de l'ATIII (Abildgaard, 1968), puis confirmée par les travaux de l'équipe de Rosenberg en 1973 (Rosenberg et Damus, 1973 ; Damus et al., 1973).

La notion de spécificité de l'interaction entre l'Hp et l'ATIII sera en fait initiée en 1976 par le groupe de Rosenberg. En centrifugeant, un mélange d'Hp et d'ATIII sur un gradient de sucrose, il observa que 1/3 seulement de l'héparine avait migré avec l'ATIII, alors que les 2/3 restant n'avaient pas changés de place bien qu'ayant utilisé un excès d'ATIII. Cela signifiait donc que **seulement certains fragments "actifs" d'héparine étaient capables de former un complexe avec l'ATIII**. Il a également démontré que seuls ces fragments actifs d'héparine avaient un effet anticoagulant. Au même moment, l'équipe de Lindahl utilisa les techniques de chromatographie d'échanges d'ions et de filtration sur gel pour séparer ces différentes "espèces" d'héparine ou différents fragments d'héparine issus de dégradations enzymatiques. Ils révélèrent que seule une fraction d'héparine de charge très négative fixait l'ATIII et avait des effets anticoagulants (Höök et al., 1976). Ces fragments possédant constitutivement ces propriétés avaient au minimum la taille d'un hexasaccharide (Hopwood et al., 1976). La même année, une autre équipe s'est intéressée à l'effet de la taille de ces fragments d'héparine liant l'ATIII sur différents facteurs de la coagulation. L'étude révéla que **l'inhibition de la thrombine par l'ATIII nécessitait un fragment d'héparine relativement long alors que l'inhibition du facteur Xa était indépendante de la taille du fragment d'héparine capable de fixer l'ATIII** (Andersson et al., 1976).

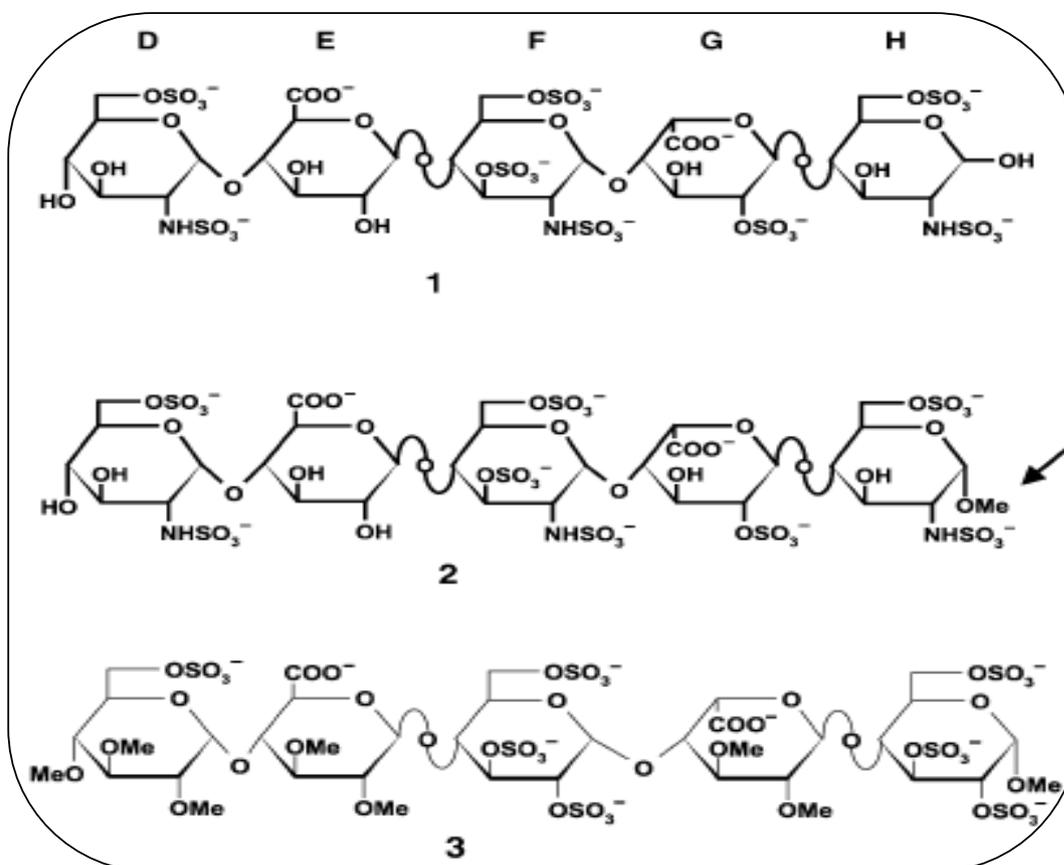
Les études successives menées ensuite par les équipes de Rosenberg et de Lindahl ont permis d'obtenir plus d'informations sur la structure du fragment d'Hp pouvant lier l'ATIII. En particulier, Lindahl et ses collaborateurs proposèrent qu'un octasaccharide contenant une séquence tetrasaccharidique semblait importante pour l'interaction avec l'ATIII et pour l'activité anticoagulante de l'Hp (Thunberg et al., 1980). Le groupe de Choay, de son côté, ayant souhaité identifier la séquence oligosaccharidique nécessaire et suffisante pour inhiber le facteur Xa dans le but de la produire par synthèse chimique. Comme Lindahl venait de le montrer, ils purent identifier un octasaccharide anticoagulant qui avait une partie de sa séquence commune avec l'octasaccharide de Lindahl (Choay et al., 1980 ; Choay et al., 1981a). La collaboration de l'équipe de Choay avec le groupe de Casu pour la RMN de cet octasaccharide a permis l'identification d'un élément rare dans le spectre RMN d'un tel polysaccharide. Dans le même temps, la collaboration de Lindahl avec le groupe de Leder dans une étude utilisant une désulfatase spécifique, révéla la présence très importante d'une 3-O-sulfatation sur une glucosamine N-sulfatée. Cette 3-O-sulfatation correspondait au pic inhabituel observé en RMN par Casu et ses collaborateurs (Lindahl et al., 1980 ; Casu et al., 1981). Les investigations menées par la suite proposèrent qu'un hexasaccharide (Choay et al., 1981b) ou qu'un heptasaccharide (Tunberg et al., 1982) était suffisant pour fixer l'ATIII et



**Figure 27a: Structure de la séquence pentasaccharidique d'héparine qui interagit avec l'antithrombine III**

Les groupements entourés sont essentiels à l'interaction avec l'ATIII. La suppression de l'un d'entre eux diminue énormément l'affinité.

(d'après Petitou et al., 2003)



**Figure 27b: Mimétiques du pentasaccharide liant l'ATIII conçus par synthèse chimique**

Le développement d'un protocole multi-étapes de synthèse chimique a permis d'obtenir le pentasaccharide 1 qui a été le premier pentasaccharide synthétique mimant l'Hp dans l'interaction avec l'ATIII. De cette synthèse, ont ensuite été conçus l'agent antithrombotique 2 mieux connu sous le nom de fondaparinux ou Arixtra®, puis Idraparinux (3). Ce dernier, est composé de groupements -OMe mais possède tout de même les groupements nécessaires à l'interaction. Ce profil de substitution lui a permis d'augmenter son temps de demi-vie d'un facteur dix. Il est actuellement testé en phase III dans la prévention et le traitement de thromboses dans les fibrillations atrial.

(d'après Petitou et Van Boeckel, 2004)

inhiber le facteur Xa. Ces hexa et heptasaccharide avaient en commun une séquence pentasaccharidique suggérant qu'un simple pentasaccharide de structure bien définie (Figure 27a) devait suffir à activer l'ATIII contre le facteur Xa.

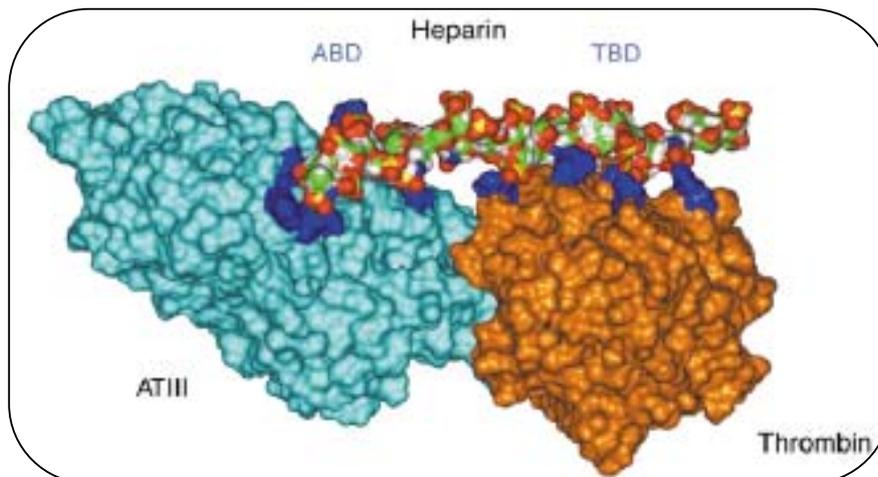
### **2.2.1.2. Le premier mimétique de l'héparine est synthétisé**

#### Premiers essais de synthèse, le pari

Un pentasaccharide était-il donc suffisant? Le développement de ce pentasaccharide synthétique (Figure 27b) fut possible grâce à une collaboration entre les deux groupes français de Choay et de Sinaÿ. Ce pentasaccharide s'est révélé être complètement actif vis-à-vis de l'ATIII et du facteur Xa (Petitou et al., 1986 ; Choay et al., 1983). Les essais de produits parallèles, issus de la synthèse chimique du type tétrasaccharide ou pentasaccharide sans la 3-O-sulfatation, ont permis de confirmer que le 3-O-sulfate du saccharide F (Figure 27a) est une condition *sine qua non* pour que l'ATIII se lie à l'Hp et que le site de fixation à l'Hp de l'ATIII nécessite au minimum un pentasaccharide (Petitou et al., 1988).

#### Les druides chimistes modifient la recette, la naissance de fondaparinux

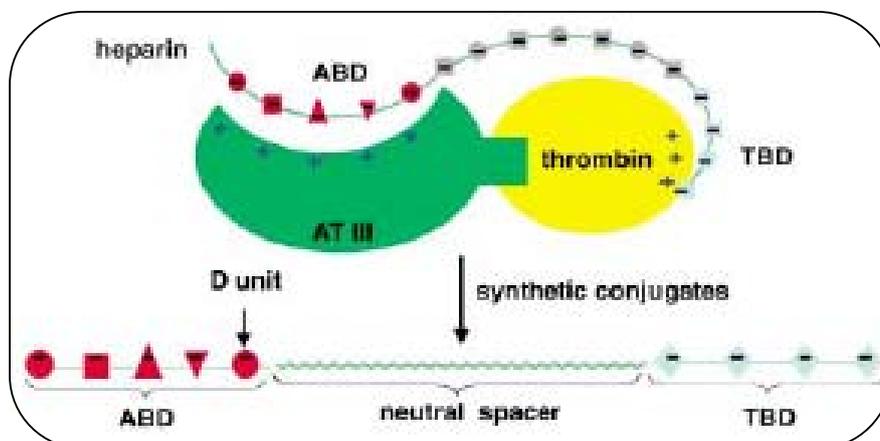
Le premier pentasaccharide, qui allait mimer l'effet de l'Hp, préparé par synthèse chimique possédait une légère modification par rapport à la structure naturelle de l'héparine impliquée dans l'interaction. Un groupement N-sulfate avait été positionné à la place d'un groupement N-acétyl sur l'ose D (Figure 27b), mais n'entraînait pas de modification de l'activité biologique (Petitou et al., 1986). Ce pentasaccharide nécessitait cependant une synthèse composée de plus de 60 étapes. Par ailleurs, son action ciblait uniquement le Facteur Xa et non pas la thrombine. De plus, la demi-vie du pentasaccharide était annoncée inférieure à celle de l'héparine (1 heure environ) (Petitou et Van Boeckel, 2004). Dès cette époque, quelques modifications chimiques ont été envisagées pour permettre une amélioration de cet anticoagulant potentiellement commercial. Sanofi et Organon ont poursuivi leurs travaux en préparant un nouveau pentasaccharide : SR90107 ou Org 31540 (Figure 27b). La nouveauté ? Un simple groupement méthyl sur le centre anomérique de l'ose H qui modifiait les propriétés de cette extrémité réductrice en bloquant la réactivité du groupement aldéhyde (Petitou et Van Boeckel, 2004). En 2001, après des essais cliniques concluants, ce pentasaccharide dénommé alors fondaparinux (ou SR90107 ou Org 31540) est enregistré en Europe et aux Etats-Unis comme un nouveau antithrombotique sous le nom d'Arixtra (Bauer et al., 2002). Son temps



**Figure 28a: Modèle moléculaire du complexe ternaire ATIII/Hp/thrombine**

Ce modèle montre clairement que l'Hp forme un pont entre les deux protéines. Six à huit unités monosaccharidique entre les deux protéines n'engagent aucune interaction avec celles-ci. L'ABD correspond à l'ATIII Binding Domain et le TBD au Thrombin Binding Domain.

*(d'après Petitou et Van Boeckel, 2004)*



**Figure 28b: Représentation schématique de la structure que doit adopter un mime d'Hp pour lier à la fois l'ATIII et la thrombine**

Ce schéma est issu du modèle moléculaire de la figure 25a. Il permet de décrire les caractéristiques requises par le mimétique synthétique développé pour mimer l'Héparine dans la liaison avec les deux protéines. Ainsi, ce glycoconjugué synthétique est composé du pentasaccharide liant l'ATIII (ABD), lié à un bras espaceur neutre d'une longueur suffisante, lui-même connecté à un domaine chargé négativement et se fixant à la thrombine (TBD).

*(d'après Petitou et Van Boeckel., 2004)*

de demi-vie chez l'Homme est d'environ 17 heures et nécessite donc uniquement une seule administration journalière à la place des deux ou trois effectuées pour l'Hp.

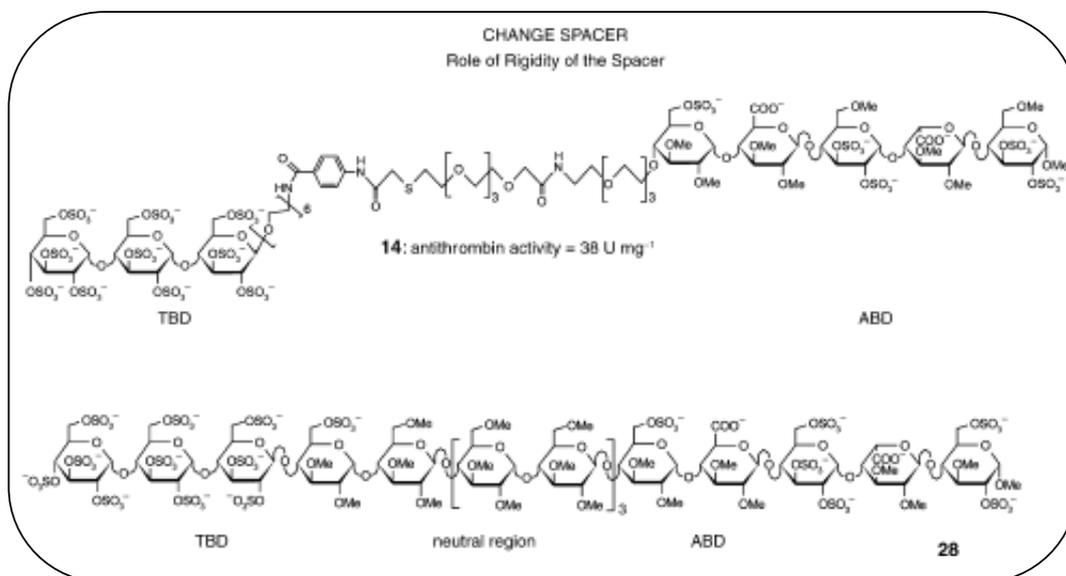
### Le perfectionnisme gaulois : de Fondaparinux à Idraparinux...

Pour tenter d'améliorer les propriétés pharmacologiques de ce médicament qui doit faire face aux nombreux antithrombotiques commercialisés de même nature mais non synthétiques, (à savoir des fragments d'héparine de bas poids moléculaires tels que le Lovenox® ou Fraxiparine® fabriqués par dégradation chimique ou Innohep® obtenu par clivage enzymatique) des analogues de fondaparinux ont très rapidement été envisagés. Par exemple, un pentasaccharide complètement O-sulfaté et O-méthylé, dénommé Idraparinux, a été synthétisé (Figure 27b). Plus facile à synthétiser (puisque cela peut être fait à partir de glucose et selon 25 étapes seulement), il est surtout plus actif (activité anti-Xa de 2700 U/ $\mu$ mol contre 1200 U/ $\mu$ mol pour fondaparinux) et possède un temps de demi-vie bien supérieur au fondaparinux (120 heures) (Herbert et al., 1998). Les essais cliniques de phase II montrent qu'une injection hebdomadaire est suffisante. Il est actuellement testé en phase III dans la prévention et le traitement de thromboses dans les fibrillations atriales (trouble du rythme supra-ventriculaire complexe, de mécanisme encore imparfaitement élucidé) (Kubitza et Haas, 2006).

#### **2.2.1.3. La thrombine serait-elle jalouse du facteur Xa?**

L'inhibition de la thrombine par l'héparine nécessite un fragment oligosaccharidique bien plus grand qu'un pentasaccharide. Ainsi fondaparinux et idraparinux sont de bons inhibiteurs du facteur Xa mais en aucun cas de la thrombine, elle-même pourtant impliquée dans la coagulation. Bien que le changement conformationnel de l'ATIII induit par le pentasaccharide soit nécessaire, l'inhibition de la thrombine requiert en plus la stabilisation du complexe ATIII-thrombine par un fragment d'héparine d'au moins 18 saccharides (Danielsson et al., 1986). Un modèle d'interaction entre les trois molécules a été proposé dès 1995 (Grootenhuis et al., 1995). Il a été confirmé cinq ans plus tard par des études de cristallographie (Figure 28a).

De ce modèle, naît le concept **d'une deuxième classe d'anticoagulants, capables de cibler à la fois le facteur Xa et la thrombine**. Ces nouveaux mimétiques sont composés du domaine de fixation à l'ATIII (ABD) et d'un domaine liant la thrombine (TBD). Ces deux domaines sont reliés par une région neutre qui peut être soit un bras espaceur, soit des



***Figure 28c: Développement de mimétiques de l'héparine liant l'ATIII et la thrombine***

Basé sur le raisonnement de la figure 25b, plusieurs mimétiques de l'héparine ayant pour but de lier l'ATIII et la thrombine en même temps ont été développés. Chacun d'entre eux possède à leurs extrémités, d'un côté un domaine ABD et de l'autre un domaine TBD. Les deux molécules présentées dans cette figure diffèrent principalement par la région reliant les deux domaines d'interaction. De cette façon, le rôle de la rigidité de cette région a pu être évalué. La molécule 14, qui contient un bras espaceur flexible a une activité anti-thrombotique bien inférieure à la molécule 28 qui possède une région centrale également neutre mais bien plus rigide. La molécule 28 a elle-même une activité anti-thrombotique dix fois supérieure à celle de l'Hp. Cela fait valoir l'importance de la composition du domaine centrale pour le développement de tels mimétiques.

*(d'après Petitou et Van Boeckel, 2004)*

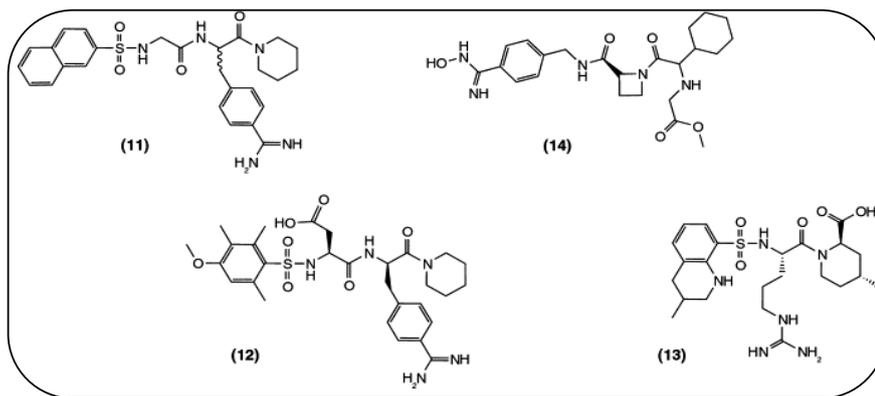
disaccharides non sulfatés (Figure 28b). Dans la revue de Petitou et Van Boeckel (2004), les différents essais réalisés sur la modification de la taille des bras espaceurs, ou de la charge des groupements de la région TBD, sont détaillés. Il s'est avéré que la séquence intermédiaire entre l'ABD et le TBD malgré sa non-interaction avec les deux protéines, avait une importance au moins au niveau de la rigidité, une propriété que n'offre pas un simple bras espaceur. En revanche, un hexadecasaccharide dont la région entre l'ABD et le TBD reste neutre (aucun -O-sulfate seulement des -O-Me) conserve cette rigidité et offre ainsi une activité antithrombique supérieure à celle de l'Hp ( $1200 \text{ U.mg}^{-1}$  contre  $160 \text{ U.mg}^{-1}$ ) (Figure 28c).

Cet hexadecasaccharide (**SanOrg 123781**) a l'avantage de ne pas fixer le PF-4 (Platelet Factor 4) (Savi et al., 2003). L'interaction avec le PF-4 lors de traitements à l'Hp ou avec des fragments d'Hp de bas poids moléculaire (LMWH) est la cause première d'effets secondaires indésirables (Napolitano et al., 2006). Des premiers essais *in vivo* de SanOrg123781 dans un modèle porcin de thrombose artérielle ont été prometteurs (Bal Dit Sollier et al., 2004). Dans le même registre, la société Momenta lance actuellement (octobre 2006) un essai clinique (phase I) du M118, une molécule faite pour fonctionner de la même manière (mais qui est hétérogène puisque obtenue enzymatiquement) dans les syndromes coronaires aigus ([www.momentapharma.com](http://www.momentapharma.com)).

#### 2.2.1.4. La troisième génération d'anticoagulants dérivés d'héparine

Les effets secondaires du type HIT (heparin-induced thrombocytopenia), dûs à l'héparine qui complexe le PF-4, ont fait stopper l'utilisation de l'héparine dans les traitements de thromboses. Ainsi pour traiter certaines thromboses, sont utilisés les LMWH (héparines de bas poids moléculaires qui agissent toujours en activant l'ATIII avec des effets secondaires moindres), fondaparinux, ainsi que les inhibiteurs de thrombine proprement dit qui ciblent directement le site actif de cette enzyme. C'est le cas notamment du Refludan® et l'Argatroban® qui ne sont pas des mimétiques de l'héparine (Figure 29A).

De ces inhibiteurs de thrombine, et des pentasaccharides développés pour activer l'ATIII, est née une molécule hybride qui possède en théorie deux actions. La première est l'activation de l'ATIII pour inhiber le facteur Xa et de manière indépendante, la seconde action est l'inhibition de la thrombine par le blocage de son site actif (Figure 29B). L'optimisation d'une première molécule (Org39913) a donné Org42675 (Figure 29B). La molécule Org42675 a une activité anti-Xa similaire à celle de fondaparinux ainsi qu'une capacité à inhiber la thrombine

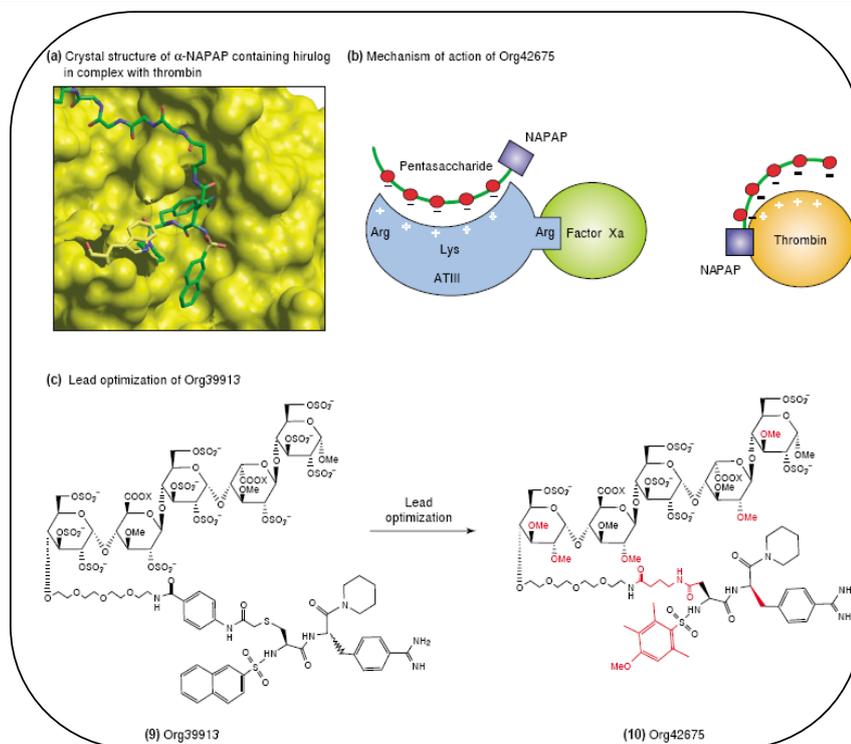


**Figure 29A: Exemples d'inhibiteurs potentiels de la thrombine**

$\alpha$ -NAPAP (11); son analogue  $\alpha$ -NAPAP (12); Argatroban (13); et Ximelagatran (14) sont les quatre inhibiteurs présentés sur cette figure. Peuvent être ajoutés les peptides d'hirudine de séquence en acide aminés:

XTYTDCTESGQNLCCEG5SNVCGQG6NKCILG5SDG6KNQCVTG6GTPKPQSHNDGDFEEIPEEYLQ où X=L dans le cas du Refludan® et X=V dans le cas du Revasc® (deux inhibiteurs commercialisés).

(d'après Martin de Kort et al., 2005)



**Figure 29B: Mécanisme d'action d'un inhibiteur bivalent ciblant aussi bien l'ATIII que la thrombine de façon indépendante**

A - Représentation de la structure cristalline (1QUR) du  $\alpha$ -NAPAP en complexe avec la thrombine.

B - Mécanisme d'action d'Org42675, l'inhibiteur bivalent.

i) L'inhibition sélective de la cascade de coagulation induite par la protéase fXa résulte de la fixation du pentasaccharide à l'ATIII.

ii) De plus, la partie NAPAP d'Org42675 est capable de bloquer directement, de façon indépendante, le site actif de la thrombine.

C - Optimisation de Org39913 (9) qui a permis le développement d'Org42675 (10). Les optimisations sont distinguées en rouge.

(Martin de Kort et al., 2005)

comparable à celle d'Argatroban. Son temps de demi-vie, en revanche, est bien meilleur que les inhibiteurs de thrombine ou encore fondaparinux ( $t_{1/2}$  multiplié par 10 vis-à-vis d'Argatroban chez le rat) et aucune réactivité avec le PF-4 n'a été observée. La molécule Org42675 a une activité inhibitrice double : elle est puissante et sélective. Elle peut ainsi être utilisée dans les thromboses à faible risque hémorragique (Martin de Kort et al., 2005).

### **2.2.2. PI-88 : un mimétique spécifique des interactions peu spécifiques ?**

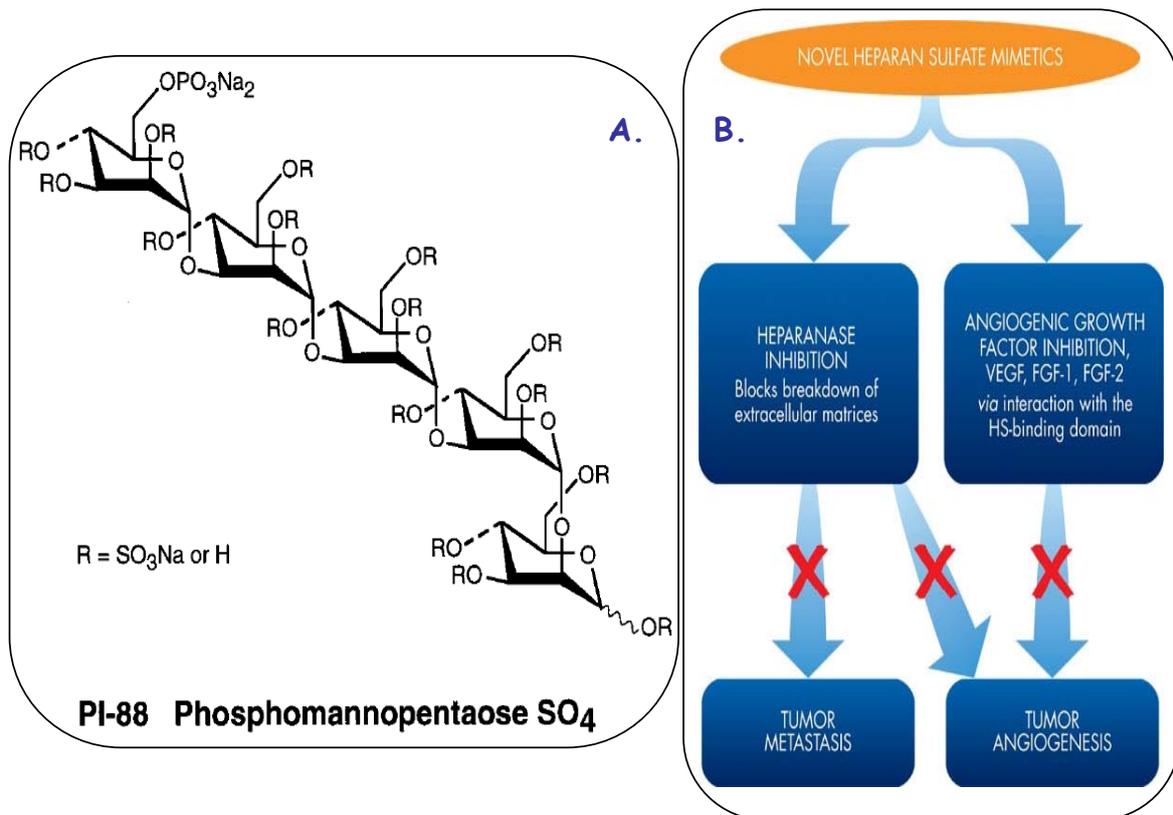
Comme il a déjà été décrit, les GAGs et les PGs jouent un rôle important dans le processus de progression tumorale. En effet, la croissance tumorale est liée en partie au remodelage des membranes basales et des matrices extracellulaires (composées notamment d'HSPGs) facilitant ainsi les phénomènes d'angiogenèse, d'invasion et de métastase. La régulation de l'activité et de la biodisponibilité de facteurs de croissances tels que le VEGF, les FGFs, ou le HGF/SF est également à prendre en considération.

Bloquer le remodelage des matrices autour des tumeurs et réguler l'effet des facteurs de croissance permettrait de lutter contre la progression tumorale. Les cibles directement en relation avec les HSPGs sont alors les facteurs pro-oncogènes (comme le VEGF et les FGFs), et les enzymes de dégradation telles que l'héparanase.

C'est sur cette idée de mimer les HS pour inhiber l'héparanase, ou pour délocaliser certains facteurs de croissance, qu'en 1999, a été développée, la molécule PI-88 dans les laboratoires australiens de Christopher Parish. La molécule PI-88, ou plutôt les molécules composant le PI-88, sont issues de la production d'exopolysaccharides par les levures *Pichi holsti*. De l'exopolysaccharide, les oses (principalement les phosphomannopentaoses et phosphomannotetraoses) sont purifiés (séparés du core phosphomannane) et sulfatés chimiquement pour donner le PI-88 qui est donc un mélange hétérogène dont le poids moléculaire varie de 1200 à 3100 Da (Ferro et al., 2002). La composition de base pour le phosphomannopentaose est le P-6-Man- $\alpha$ -(1-3)-Man- $\alpha$ -(1-3)-Man- $\alpha$ -(1-3)-Man- $\alpha$ -(1-2)-Man (Parish et al., 1999) (Figure 30).

#### *Le PI-88 en laboratoire :*

*In vitro*, le PI-88 est capable d'inhiber la croissance des cellules tumorales et *in vivo* de bloquer les métastases en limitant l'angiogenèse et en inactivant l'héparanase (Parish et al., 1999). En 2003, sur des rats, le PI-88 semble capable de fixer directement FGF-2, de bloquer la signalisation cellulaire et, en particulier, la voie ERK(1/2) (extracellular signal-regulated kinase 1/2), ainsi que d'inhiber la prolifération de cellules vasculaires de muscles lisses



**Figure 30: Le PI-88 est un mélange d'oses modifiés, issus de levure qui miment les HS et présentent des effets anti-tumoraux intéressants**

**A :** Représentation d'un des oses (modifié chimiquement) séparé du core phosphomannanes d'exopolysaccharide issu de la culture de levures *Pichi holstii*.

La molécule PI-88, ou plutôt les molécules composant PI-88, sont issues de la production d'exopolysaccharides par les levures *Pichi holstii*. De l'exopolysaccharide, les oses, principalement les phosphomannopentaoses et phosphomannotetraose sont purifiés (séparées du core phosphomannane) et sulfatés chimiquement pour donner le PI-88 qui est donc un mélange hétérogène dont le poids moléculaire varie de 1200 à 3100 Da (Ferro et al., 2002). La composition de base pour le phosphomannopentaose est le P-6-Man- $\alpha$ -(1-3)-Man- $\alpha$ -(1-3)-Man- $\alpha$ -(1-3)-Man- $\alpha$ -(1-2)-Man. (d'après Parish et al., 1999).

**B :** Schéma utilisé par Progen pour résumer les effets que peut avoir PI-88.

La société Progen qui développe PI-88, propose que ce "mélange oligosaccharidique", principalement composé de phosphomannopentaose sulfaté, peut être capable, sur le principe de mimétisme des HS, d'inhiber la progression tumorale de deux manières. Premièrement, PI-88 est capable d'inhiber l'héparanase ce qui leur permettrait de contrôler par ce biais les processus d'angiogénèse ou de métastases. Deuxièmement, en inhibant les interactions des facteurs pro angiogéniques (VEGF, FGF-1, FGF-2) avec les HS, l'angiogénèse pourrait, ici aussi, être plus facilement régulée. (d'après [www.progen.com](http://www.progen.com))

(Francis et al., 2003). En 2005, le lien entre la capacité du PI-88 à inhiber l'angiogenèse et sa capacité à inhiber VEGF est établie, l'inhibition du VEGF étant en partie une conséquence de l'inhibition de l'héparanase qui n'est, de ce fait, plus apte à relarguer les facteurs de croissance présents dans la matrice extracellulaire. Ces données sont encourageantes pour traiter les cancers dans lesquels l'héparanase est surexprimée et les facteurs de croissance mobilisés (Joyce et al., 2005 ; Miao et al., 2006).

Comme il a été décrit dans le chapitre sur les fonctions des GAGs, ces derniers peuvent servir de récepteurs pour certains pathogènes tels que *Plasmodium falciparum* (malaria) et l'HSV (Herpes Simplex Virus). La capacité du PI-88 à inhiber l'adhésion ou à éliminer ces pathogènes déjà liés aux GAGs de l'organisme a récemment été testée *in vitro*. Aussi, Nyberg et al. (2004) ont montré que le PI-88 inhibait l'adhésion du HSV-1 et HSV-2 à des cellules épithéliales. De même, Adams et al. (2006) ont, eux, montré que le PI-88 inhibait la fixation de *Plasmodium falciparum* sur son récepteur placental, à savoir les CS A (Chondroïtine-4-sulfate) avec une IC<sub>50</sub> de 7,4 µM.

#### Le PI-88 comme anti-cancéreux en phase clinique

Le PI-88 est actuellement l'anti-cancéreux le plus avancé de la société australienne Progen. En mimant les HS, il a pour but de traiter les tumeurs primaires ainsi que les métastases. En effet, le PI-88 est capable, *in vitro*, d'inhiber l'Hpa et de délocaliser les facteurs de croissance en inhibant les interactions entre le VEGF, les FGF-1 et FGF-2 avec les HS (Figure 30). Il est actuellement en essais cliniques de phase II dans différents types de cancers et notamment les cancers du foie.

### **2.2.3. Les GAGs et l'amylose**

Les amyloses regroupent un ensemble de maladies tissulaires caractérisées par la présence de protéines anormalement insolubles dans différents tissus humains. Ces protéines forment des agrégats moléculaires appelés "substance amyloïde". Ces dépôts envahissent progressivement les tissus et empêchent ainsi leur bon fonctionnement, entraînant le développement des différents symptômes et maladies. Leur évolution est le plus souvent sévère avec la destruction progressive et irréversible des organes atteints. Il existe différentes formes d'amyloses dont la classification repose sur la clinique et la nature biochimique de la protéine amyloïde impliquée dans la formation des "agrégats". Les formes les plus fréquentes sont les amyloses AL (immunoglobuliniques), les amyloses AA (inflammatoires) ou amyloses ATTR (héréditaires). Outre ces formes systémiques, il existe les amyloses localisées, limitées

Nom courant	Nom commercial	Formule	Champs thérapeutiques	Stade de développement
Eprodissate	Kiacta® Précédemment appelé Fibrillex	1,3-propane disulfonate (1,3PDS)	l'amylose amyloïde A	phase II/III terminée en 2005 Dernière décision en Avril 2007
Tramiprosate	Alzhemed®	3-amino-1-propane sulfonique (3APS)	maladie d'Alzheimer	Phase III aux USA OK Phase III en EU qui vient d'être validée
Tramiprosate	Cerebril®	3APS	Prévention des accidents cérébrovasculaires (ACV) hémorragiques causés par angiopathie amyloïde cérébrale.	Essais clinique phase II au U.S.

**Figure 31 : Inhibiteurs des interactions GAGs-Proteines amyloïdes développés par Neurochem**

à un seul type de tissu ou organe, ainsi que les amyloses spécifiques telles que la maladie d'Alzheimer, la maladie de Parkinson, le diabète de type 2, les maladies dues aux prions telles que l'encéphalopathie spongiforme bovine.

Dans ce genre de maladies, la formation des protéines anormalement fibrillaires (amyloïdes) qui sont responsables des symptômes est à priori facilitée par la présence des GAGs (Ben-Zaken et al., 2003 ; Horonchik et al., 2005). Les GAGs permettent également, une fois que les protéines sont sous forme amyloïdes, de stabiliser cet état (Gillmore et Hawkins, 2006). Des expériences *in vivo* ont montré, par l'utilisation de l'héparanase et donc par dégradation des HS, que ces derniers étaient nécessaires à l'amyloïdogénèse (Li et al., 2005). Récemment, des mimétiques d'HS dérivés de dextran et synthétisés dans le laboratoire de Denis Barritault se sont révélés capables, *in vitro*, de bloquer la formation de la forme amyloïde du prion (Schonberger et al., 2003). Certains de ces mimétiques se sont également révélés capables de bloquer l'interaction du prion avec un de ses récepteurs cellulaires (Gauczynski et al., 2006). En 1995, Kisilevsky et coll. ont montré, *in vivo*, que l'administration, par voie orale, de composés sulfatés ou de sulfonates anioniques de faible poids moléculaire (135 à 1000 Da) permettait de réduire la progression de la forme amyloïde AA. Sur la base de ces composés et sur cette idée d'inhiber les interactions entre les protéines potentiellement amyloïdes et les GAGs responsables de l'amyloïdogénèse et de la stabilisation des agrégats amyloïdes, la société canadienne Neurochem est en train de développer deux mimétiques des GAGs qui ciblent les amyloses systémiques de type AA : la maladie d'Alzheimer, et les accidents cérébrovasculaires (ACV) hémorragiques (Figure 31). Leur but est de prévenir la formation de fibrilles AA ainsi que le dépôt de ces fibrilles dans les organes, en inhibant l'interaction avec les GAGs présents naturellement. Les données récentes sur l'efficacité du Tramiprosate (Alzhemed®) dans la maladie d'Alzheimer apparaissent très prometteuses et soulèvent beaucoup d'espoir (Gervais et al., 2006).

### **2.3. LE POTENTIEL THERAPEUTIQUE DES MIMETIQUES DES GAGS : Les sucres peuvent-ils soigner tous les maux ?**

#### **2.3.1. Une "pseudo-longueur" d'avance dans le développement thérapeutique des mimétiques des GAGs**

Lors du développement de molécules à visée thérapeutique, de nombreuses interrogations restent en suspens lors de l'initiation des phases cliniques chez l'Homme. **La molécule va-t-**

## **elle engendrer une réaction immunitaire ? Quels vont être les effets secondaires liés au traitement ?**

Dans le cas des mimétiques de GAGs ayant un potentiel effet thérapeutique, l'idée est toujours d'administrer un "pseudo-GAG" pour moduler l'activité de protéines dont les fonctions sont connues pour être GAGs dépendantes.

Dans le cadre où les interactions GAGs-protéines visées sont dépendantes d'un profil de sulfatation plutôt que d'une structure glycosaminoglycannique bien définie (comme cela a pu l'être pour l'ATIII), **l'héparine** peut être considérée comme **la molécule** "thérapeutique" qui, potentiellement, a le plus d'effets thérapeutiques. En revanche, l'héparine est **très peu spécifique** et donc peut **engendrer plusieurs effets secondaires**.

Or, comme il a été mentionné, **l'héparine est utilisée depuis 1937** comme agent anti-coagulant. Depuis cette période il a été possible de mesurer et d'apprendre à connaître les effets secondaires liés à son utilisation. **Les effets indésirables ne sont pas très nombreux**. L'ostéoporose induite par un traitement prolongé à l'héparine (HIO : Heparin-Induced Osteoporosis) peut être citée (Griffith et al., 1965 ; Jaffe et Willis, 1965 ; Miller et DeWolfe, 1966 ; Rupp et al., 1982 ; Dahlman, 1993 ; Dahlman et al., 1994 ; Barbour et al., 1994). L'effet secondaire majoritaire lors de traitements à l'héparine est bien sûr l'effet HIT ou Heparin-Induced Thrombocytopenia. L'héparine, lors de son administration par voie intraveineuse et du fait de sa faible spécificité, est capable de complexer le PF-4. Cela induit chez cette chimiokine un changement de configuration qui crée une réaction immunitaire contre le complexe ainsi formé. Le complexe immun (Hp-PF-4-IgG) peut alors activer les plaquettes et dérégler le processus de coagulation générant des thromboses artérielles ou veineuses (Visentin et al., 2000 ; Amiral et Meyer, 1998).

**L'utilisation de LMWH** (Low Molecular Fractionated Heparin) comme anti-coagulant qui n'est autre que de l'Hp fractionnée et purifiée par chromatographie, **s'est révélée engendrer moins d'effets indésirables** de type HIT (Warkentin et al., 1995 ; Menajovsky 2005). Des composés encore plus petits et bien plus spécifiques tels que le pentasaccharide liant l'ATIII (Amiral et al., 1997) ou des fragments d'Hp reliés par des oligosaccharides non sulfatés se sont révélés être des composés également mieux tolérés (Petitou et al., 1999).

Ainsi, **l'expérience** acquise **depuis ces décennies** dans **l'utilisation de l'héparine ou des LMWH chez l'Homme** permet d'anticiper, de manière théorique, les effets indésirables que pourraient avoir les futurs mimétiques de l'Hp, ce qui constitue un avantage considérable.

Dans le développement thérapeutique de molécules mimant l'Hp, **la question de la demi-vie** peut également être posée à juste titre, puisque celle de l'Hp n'est que d'une heure.

L'exemple de fondaparinux, développé précédemment, permet d'être optimiste, puisqu'une simple méthylation a permis d'obtenir une demi-vie augmentée de l'ordre de 17 heures.

### **2.3.2. Les mimétiques des GAGs ont un potentiel thérapeutique très vaste**

#### **2.3.2.1. Les mimétiques et la malaria**

L'infection des érythrocytes par le parasite *Plasmodium falciparum* est plus connue sous le nom de malaria. Chaque année, plus d'un million de personnes décèdent de cette maladie. Le désordre causé par ce parasite est lié à une fixation excessive des érythrocytes infectés à l'endothélium vasculaire (cytoadhérence) et de l'autoagglutination des érythrocytes infectés et non infectés, qui entraîne consécutivement des occlusions vasculaires responsables de la pathologie (Alonzo, 2006).

Lors de l'infection, les GAGs jouent le rôle de récepteurs d'adhésion. Afin de réduire considérablement la fixation du parasite aux érythrocytes, l'Hp a tout d'abord été utilisée pour traiter des cas sévères de malaria (Munir et al., 1980 ; Rampangan., 1991). Les effets secondaires observés n'ont malheureusement pas permis d'utiliser l'Hp en routine et de façon chronique même si elle avait un rôle bénéfique sur la maladie. Anna M. Vogt et al. (2006) ont obtenu des résultats *in vitro* et *in vivo* prometteurs sur le relargage du parasite en utilisant de l'Hp traitée au périodate (dGAG) (et donc dépourvue de ses activités anti-coagulantes) dans deux nouveaux modèles, chez le rat et le singe. La conclusion de ces expériences était que ces dGAGs possédaient le même effet que l'Hp ou que les HS. Ces résultats sont actuellement exploités par la société suédoise Dilafor.

Une étude récente propose également que le PI-88, dont nous avons déjà parlé, pourrait également être un bon moyen de mimer l'effet thérapeutique de l'Hp sans en avoir les effets anti-coagulants (Adams et al., 2006).

#### **2.3.2.2. Les mimétiques et le VIH**

Le VIH, pour virus de l'immunodéficience humaine, est la cause de l'une des maladies les plus meurtrières du monde avec environ 3 millions de décès chaque année.

Brièvement, l'infection par le VIH est rendue possible par la fixation de la glycoprotéine d'enveloppe du virus, gp120, sur la protéine CD4 et sur les récepteurs à chimiokines CXCR4 et CCR5 présents sur des lymphocytes T. Cette interaction permet la fusion entre le virus et la

membrane cellulaire et rend ainsi possible l'internalisation du génome viral dans les lymphocytes (Moore et Stevenson, 2000). Les GAGs ont, par ailleurs, un rôle plus en amont de récepteur d'ancrage afin de concentrer le virus à la surface des cellules et de le présenter aux lymphocytes (Bobardt et al., 2003).

Plusieurs composés polyanioniques peuvent permettre d'inhiber l'attachement aux cellules, l'entrée et la réplication du virus (Rider, 1997 ; De Clercq, 1989).

Récemment, deux dérivés des polysaccharides K5 d'E.coli (K5-OS(H) et K5-N,OS(H)) se sont révélés capables d'inhiber la réplication du virus dans des cellules T et des macrophages. Ils permettent également d'empêcher la formation de syncytia induite par le virus. Ces molécules pourraient donc être intéressantes dans la prévention de l'attachement du virus et/ou son entrée dans les cellules (Vicenzi et al., 2003).

De plus, gp120 se fixe aux HSPGs via un domaine C-terminal basique de sa boucle V3 (Variable) (Moulard et al., 2000). Rusnati et coll. se sont rendus compte que les polysaccharides K5-OS(H) et K5-N,OS(H) sont capables d'inhiber cette interaction avec une IC50 comprise entre 1 à 10 nM (Rusnati et al., 2005).

Enfin, Vives et al. (2005) ont montré que les sites d'interactions de gp120 pour les corécepteurs du VIH pouvait également être ciblés par des molécules de type HS (Vives et al., 2005).

Par ailleurs, un facteur de transactivation du VIH, appelé Tat, permet la transcription du génome viral et l'induction de l'expression des co-récepteurs du VIH. Il représente de ce fait une cible potentielle dans la lutte contre l'infection virale. Or, Tat possède un domaine basique qui lui permet d'interagir avec les HSPGs (Rusnati et Presta, 2002). Des mimétiques de GAGs ont donc été testés pour moduler l'effet de ce facteur de transactivation. Des résultats positifs *in vitro* et *in vivo* dans l'inhibition de l'activité de Tat ont été reportés (Rusnati et Presta 2002). Les molécules K5-OS(H) et K5-N,OS(H) se fixent à Tat avec une haute affinité (3 à 20 nM) et inhibent l'interaction avec les HSPGs, au même titre que son internalisation cellulaire et donc que ses activités de transactivateur (Rusnati et al., 2005).

Dans le but de réduire les propriétés virulentes soit de gp120, soit de Tat, des études cliniques utilisant des polyanions tels que les dextrans ou des pentosanes polysulfates ont déjà été publiées. Cependant, ces mimétiques ont pour l'instant été que peu efficaces (Abrams et al., 1989 ; Pluda et al., 1993 ; Marshall et al., 1997 ; Swain et al., 1995 ; Schwartzmann et al., 1996 ; Flexner et al., 1991).

### **2.3.2.3. Les mimétiques comme traitement anti-cancéreux ?**

Comme il a été décrit dans ce chapitre (*1.5. L'implication des GAGs dans le cancer*), les GAGs ont un rôle dans beaucoup de processus liés au cancer aussi bien dans la croissance tumorale et l'angiogenèse qui lui sont associées, que dans la migration de cellules tumorales, l'adhésion et l'extravasation de ces cellules lors de la métastase. A tout moment, ces cellules "immortelles" sont en lien étroit avec les polysaccharides linéaires que sont les GAGs. Il semble donc important d'arriver à réguler les facteurs oncogènes tels que les facteurs de croissance (FGFs, VEGF, PDGF-BB...), les enzymes de remodelage de la matrice (héparanase, protéases) ou autres molécules d'adhésion (sélectines) pour espérer contrôler le développement tumoral. Or, il se trouve que bon nombre des facteurs oncogènes se lient ou sont dépendants des interactions avec les GAGs. L'idée d'aller moduler l'action de ces facteurs via des mimétiques de GAGs a progressivement fait son chemin dans les approches thérapeutiques envisageables pour contrôler les événements tumoraux.

Parmi les patients atteints de cancer, nombre d'entre eux ont de grands risques de présenter des complications de types VTE (venous thromboembolism) qui nécessitent un traitement à base d'Hp ou de LMWH (Smorenburg et al., 1999). Ainsi, l'effet direct de l'Hp ou de fragments d'Hp chez l'Homme atteint de cancer a déjà pu être en partie appréhendé et peut donner une idée de la pertinence de l'utilisation de mimétiques des GAGs en oncologie. Dans une revue, Smorenburg et Van Noorden (2001) ont fait un inventaire des effets répertoriés de l'Hp sur la progression tumorale et métastatique. Les effets de l'Hp et des LMWH dans le cancer ont également fait l'objet de plusieurs autres revues (Zacharski et Ornstein, 1998 ; Zacharsky et al., 2000 ; Zacharsky et Loynes, 2002 ; 2003).

### **2.3.2.4. Les mimétiques comme traitement anti-inflammatoire ?**

Comme nous l'avons déjà détaillé dans les fonctions des GAGs (*1.4. L'implication des GAGs dans l'inflammation*), ces polysaccharides, qu'ils soient à la surface des cellules ou dans les matrices extracellulaires, jouent un rôle important tout au long de la réponse inflammatoire, que ce soit dans le processus de recrutement des cellules de l'inflammation (rolling, triggering, adhesion) ou de transmigration (Delehedde et al., 2002 ; Lassalle et al., 2006, Melchior et al., 2006).

Il a été démontré que l'héparine (soluble) interagissait avec de nombreuses molécules d'adhésion, permettant le recrutement des cellules de l'inflammation comme les L-sélectines

(Koenig et al., 1998), MAC-1 (Diamond et al., 1995), les P-sélectines (Skinner et al., 1991), ou encore PECAM-1 (Platelet/endothelial cell adhesion molecule 1) (Watt et al., 1993).

Ces interactions permettent à l'Hp *in vitro*, d'inhiber l'adhésion des cellules de l'inflammation aux cellules endothéliales des parois des vaisseaux (Lever et Page 2002). *In vivo* l'Hp peut inhiber le recrutement leucocytaire (Lever et al., 2000 ; Xie et al., 1997 ; Ley et al., 1991 ; Tangelder et Arfors, 1991).

En plus de ces effets expérimentaux, dès 1960, Dolowitz et Dougherty (1960 ; 1965) ont montré que l'Hp peut avoir un réel effet dans le traitement des maladies inflammatoires. L'héparine a été utilisée pour pouvoir traiter l'asthme par inhalation (Diamant et al., 1996 ; Ahmed et al., 1993). L'héparine a été utilisée également dans les traitements d'arthrites rhumatoïde (Gaffney et Gaffney, 1996), d'IBD (Inflammatory bowel disease) (Gaffney et al., 1991 ; 1995 ; Evans et al., 1997), d'infarctus du myocarde (Theroux et al., 1993) ou de rhinites allergiques (Vancheri et al., 2001). Aucun cas d'hémorragie n'a été observé dans ces études cliniques.

Ces exemples confortent l'idée que l'utilisation de mimétiques des GAGs peut avoir un avenir sérieux dans les stratégies thérapeutiques anti-inflammatoires. Cela a d'ailleurs été également initié avec des héparines dé-O-sulfatées, qui en perdant leur activité anti-coagulante conservent leurs propriétés anti-inflammatoires (Fryer et al., 1997), ou encore avec l'utilisation du pentosane polysulfate comme anti-inflammatoire dans l'arthrose du genou (Innes et al., 2000). La société Paringenix développe ainsi un dérivé d'Hp qui ne possède plus d'activité anti-coagulante (dérivé d'héparine 2- et 3-O-désulfatée) mais conserve une activité anti-inflammatoire intéressante notamment dans les problèmes d'ischémie cardiaque, suite à un infarctus du myocarde.

### **2.3.2.5. Les mimétiques pour inhiber l'héparanase**

Comme cela a été précisé dans le chapitre I, l'héparanase est une endo- $\beta$ -glucuronidase qui dégrade les HS à la surface cellulaire ou présents dans les matrices extracellulaires. Cette enzyme est utile dans de nombreux processus tels que le recyclage des HSPGs ou le remodelage tissulaire essentiel lors du développement, de l'implantation embryonnaire, de la transmigration inflammatoire, de la cicatrisation et de l'angiogenèse (Vlodavsky et al., 2002).

Dans les processus pathologiques, tels que les maladies inflammatoires ou le cancer, de par ses capacités à remodeler la matrice extracellulaire, l'Hpa joue également un rôle

essentiel. Elle favorise l'infiltration des leucocytes ou des cellules tumorales dans le cas de l'inflammation ou du développement tumoral (Sanderson et al., 2004).

Depuis le clonage de son gène en 1999, l'héparanase est une des cibles thérapeutiques les plus étudiées en oncologie. La plupart des cellules cancéreuses surexpriment cette enzyme pour se frayer un chemin dans la matrice extracellulaire. De nombreuses molécules inhibitrices ont été développées (Ferro et al., 2004 ; 2007). L'Hpa, dont l'activité enzymatique est dirigée contre les HS ou l'Hp, possède un site de coupure préférentiel au sein des GAGs. Certains autres sites de l'Hp permettent en revanche d'inhiber l'activité de l'enzyme. De ces observations, est née l'idée que certains mimétiques de l'Hp pourraient inhiber l'Hpa, et ainsi réduire ses effets pathologiques non souhaités (Ferro et al., 2004 ; Ma et al., 2006).

Parmi ces potentiels inhibiteurs d'héparanase héparinoïdes (mimant les héparines), deux classes peuvent être distinguées. La première est celle des polysaccharides sulfatés modifiés naturellement ou chimiquement. La seconde est celle des polymères polyanioniques qui ne sont pas des carbohydrates. Les inhibiteurs de type polysaccharides sulfatés peuvent être l' $\lambda$ -carrageenan, le fucoïdane, des "pentosan polysulfate", les oligomannuronates sulfates, des dextrans de sulfate, la 6-O-carboxyméthyl chitine III, le laminarine sulfate ou le calcium spirulan. Dans la classe des polymères polyanioniques sont représentés le RG-13577 (polymère de (4-hydroxyphenoxy)acetic acid), les poly(N-acryl amino acids) tels que la poly(N-acrylleucine), les phosphorothioate oligodéoxynucleotides et les molécules développées par Lipinski dénommées les "Lipinski's Rule of Five" (Lipinski et al., 2001). La plupart de ces inhibiteurs d'héparanase ne peuvent toutefois pas être utilisés de façon thérapeutique en raison de leur faible spécificité et des effets secondaires qu'ils pourraient engendrer (effets anti-coagulants par exemple) (Ferro et al., 2004).

Aujourd'hui, seul le PI88 (phosphomannose) est en essai clinique de phase II pour différents types de cancers. Les candidats thérapeutiques potentiels en tant que mimétiques des GAGs pour inhiber l'héparanase peuvent être l'oligomannuronate sulfate JG3 (Zhao et al., 2006) ou une des molécules conçues par Lipinsky : le KI-105 (Ishida et al., 2004).

### **2.3.2.6. Des mimétiques pour la réparation tissulaire : Les RGTAs**

Depuis quelques années déjà, dans le but de développer un agent de cicatrisation et de réparation de tissus, Caruelle et Barritault développent des mimétiques des GAGs sur la base de polymères de dextran modifiés chimiquement. Ces molécules, rassemblées sous le nom de RGTA (Regenerating Agents), ont pour but d'augmenter la biodisponibilité des facteurs de

croissance liant l'héparine, à proximité d'un site de blessure (Barritault et Caruelle, 2006). Comme peuvent le faire des HS, le but de ces mimétiques est de protéger les facteurs de croissance des dégradations protéolytiques intervenant dans ces conditions.

Cette façon d'utiliser les mimétiques des GAGs est tout à fait originale. En effet, dans les cas où des médiateurs tels que des facteurs de croissance sont ciblés (anti-cancéreux, anti-inflammatoire), le but est en général de réduire leur activité en les délocalisant, or, dans ce cas précis, l'effet inverse est souhaité.

Dans cette famille de RGTA, le RG1192 a par exemple été utilisé dans plusieurs essais *in vivo* de régénérations tissulaires variées (Meddahi et al., 1996a ; Albo et al., 1996 ; Fredj-Reygrobellet et al., 1994 ; Yamauchi et al., 2000 ; Desgranges et al., 1999). Il n'a pas d'effet anticoagulant et permet de potentialiser l'activité des FGF-1 et 2 en les protégeant des dégradations (Meddahi et al., 1995 ; 1996b) et en servant de co-récepteur comme peut le faire l'Hp (Rouet et al., 2005a). Le RG1192 se révèle également être un bon inhibiteur de l'Hpa avec une IC50 de 2,5 nM (Rouet et al., 2006).

Un autre membre de la famille, l'OTR4120, est un mimétique capable d'induire l'angiogenèse en potentialisant l'affinité du VEGF (facteur pour lequel il a une forte affinité (2,2 nM)) pour ses récepteurs VEGFR-1 et 2 et pour la neuropiline-1 (Rouet et al., 2005b).

**CHAPITRE IV :**  
**L'INTERFERON  $\gamma$ , UN LIGAND DES**  
**GLYCOSAMINOGLYCANNES**

# 1. L'INTERFERON GAMMA (IFN $\gamma$ )

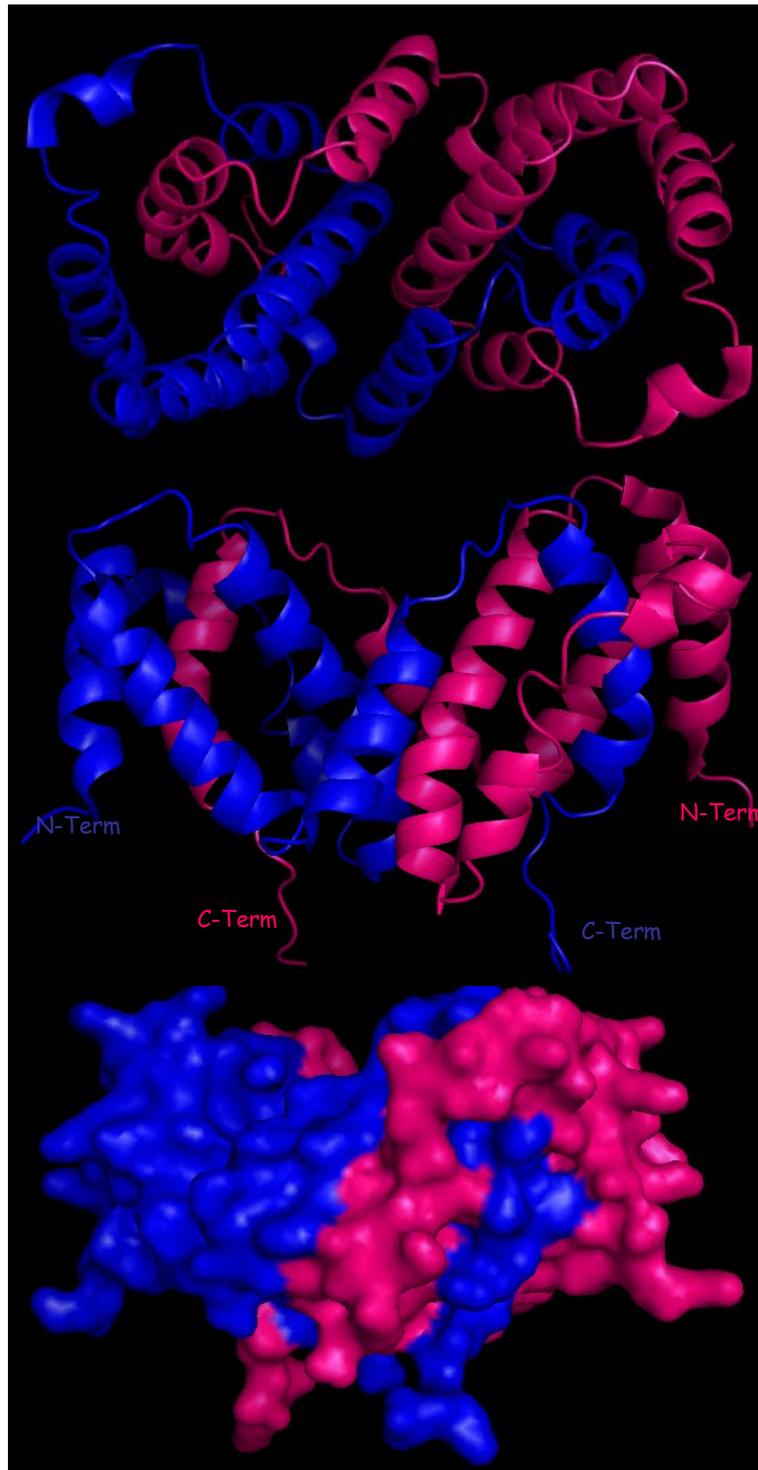
## 1.1. GENERALITES

### 1.1.1. Historique

L'interféron fût découvert en 1957 par Isaacs et Lindenmann en observant que certaines cellules en culture, infectées par un virus, produisaient une protéine les rendant résistantes aux infections virales. Historiquement, le nom "interféron" est relatif à ce phénomène d'interférence (Nagano et Kojima, 1954). Les interférons de type I, (l'interféron  $\alpha$  et  $\beta$ , IFN- $\epsilon$ , IFN- $\kappa$ , IFN- $\omega$  et les 13 variants d'IFN- $\alpha$ ) ont ainsi une activité antivirale importante. Deux équipes de recherche indépendantes ont plus tard identifié une protéine ayant le même type d'activité et l'ont nommé différemment : l'interféron immun ou l'interféron de type II et le MAF pour Macrophage-Activating Factor (Nathan et al., 1971 ; 1973 ; 1983, Billiau et Vandebroek, 2000). L'interféron  $\gamma$  (le seul à être de type II), bien qu'ayant aussi une activité antivirale a surtout des activités immunomodulatrices. Les premiers travaux qui ont permis de définir le facteur activateur de macrophage (MAF) comme un interféron sont ceux de Wheelock et al. (1965), même si cela avait été initié par Gresser et Nacify (1964). Dans les années 70, les termes d'interféron immun (Falcoff, 1972) ou d'interféron de type II (Youngner et Salvin, 1973) ont été utilisés. En 1980, le comité international de nomenclature officialise les noms utilisés pour les interférons de type I ou II. L'interféron immun ou de type II est devenu l'interféron gamma (IFN $\gamma$ ), les interférons de type I tels que l'interféron leucocytaire ou l'interféron fibroblastique sont respectivement devenus l'interféron alpha (IFN $\alpha$ ) et l'interféron bêta (IFN $\beta$ ).

### 1.1.2. Du gène à la protéine

Bien qu'assigné au départ à la bande chromosomique 12q24.1 par hybridation *in situ* (Trent et al., 1982), Bureau et al. (1995) ont plus précisément localisé le gène d'IFN $\gamma$  humain sur la bande 12q25. Ce gène de 6 kb contient trois introns et quatre exons codant pour un polypeptide de 166 acides aminés dont les 23 premiers constituent la séquence signal (Devos et al., 1982 ; Naylor et al., 1983). L'ADN complémentaire du gène humain d'IFN $\gamma$  a été cloné pour la première fois en 1982 (Gray et al., 1982). Ainsi, une fois sa séquence signal clivée,



**Figure 32: Structure de l'homodimère d'IFN $\gamma$**

IFN $\gamma$  est actif sous la forme d'un homodimère. Sur cette figure est représentée la structure cristallographique de ce dimère avec une couleur par monomère. En haut, une "vue de dessus", au centre une "vue de coté" avec les extrémités N- et C-terminale, et en bas une visualisation dite de surface. Le dimère a été extrait de la structure du complexe IFN $\gamma$ /3IFN $\gamma$ R résolue par Thiel et al., 2000. Les 10 derniers acides aminés de chaque monomère ne sont pas présent dans cette structure.

(D'après Thiel et al., 2000)

l'IFN $\gamma$  humain mature est une glycoprotéine homodimérique composée de deux polypeptides de 143 acides aminés, d'un poids moléculaire de 17,1 kDa. Chaque monomère possède deux sites de glycosylation sur les résidus asparagines en position 25 et 97 (Sareneva et al., 1996), d'où sa sécrétion sous deux formes : 20 et 25 kDa (Yip et al., 1982 ; Kelker et al., 1984 ; Rinderknecht et al., 1984 ; Sareneva et al., 1996). L'IFN $\gamma$  humain ne possède pas de cystéine, ce qui exclut toute conformation tertiaire liée à des ponts disulfures. Sous forme cristalline et non glycosylée, l'IFN $\gamma$  est un homodimère de 34 kDa stabilisé par des liaisons non covalentes et par l'imbrication entre les hélices  $\alpha$  des deux monomères. Ces deux polypeptides sont agencés de façon tête-bêche de telle sorte que l'extrémité N-terminale d'un monomère est située à proximité de l'extrémité C-terminale de l'autre monomère. Chacun d'entre eux est composé par sept hélices  $\alpha$  (A à F) (constituées chacune de 9 à 21 acides aminés) liées par des boucles et une séquence C-terminale (acides aminés 124 à 143) qui semble être très flexible et adopter de multiples conformations (Ealick et al., 1991) (Figure 32). Cette région C-terminale contient deux domaines d'acides aminés basiques (D1 : 125-131 et D2 : 137-140) qui confèrent à la cytokine une grande sensibilité aux attaques protéolytiques (Trinchieri et Perussia, 1985 ; Döbeli et al., 1988 ; Curling et al., 1990 ; Jarpe et Johnson, 1990) et qui semblent jouer un rôle important dans l'activité de la protéine (Nacheva et al., 2003).

### **1.1.3. Les sources cellulaires**

L'IFN $\gamma$  appartient à la famille des cytokines et peut être typiquement classé comme une lymphokine. Il est produit exclusivement par les cellules NK (Natural Killers Cells) et certaines populations de lymphocytes T telle que la sous-classe Th1 des lymphocytes CD4+ et certains lymphocytes cytotoxiques CD8+ (Fong et al., 1990). D'une manière générale, chez l'Homme, les lymphocytes T qui expriment l'antigène membranaire CD30 ont été identifiés comme les principaux producteurs d'IFN $\gamma$  (Alzona et al., 1994). De façon marginale, la production d'IFN $\gamma$  a été observée, *in vitro*, dans des cultures de phagocytes mononucléaires (Fultz et al., 1993 ; Gessani et Belardelli, 1998), de granulocytes neutrophiles (Yeaman et al., 1998), de neurones (Neumann et al., 1997), ou de lignées cellulaires d'origine neurale (Watanabe et al., 1989).

#### 1.1.4. L'expression d'IFN $\gamma$

##### 1.1.4.1. La voie de synthèse

La production d'IFN $\gamma$ , suite à une stimulation, est induite via la voie de la phospholipase C (Williams et al., 1993). Celle-ci va hydrolyser le phosphatidylinositol (IP) conduisant à la libération de diacylglycérol (DAG) et d'inositol triphosphate (IP3). Ce dernier entraîne une augmentation cytosolique d'ions Ca<sup>2+</sup> qui, en association avec le DAG, active la protéine kinase C (PKC). Une cascade de phosphorylation est alors induite, aboutissant à l'activation de la transcription et de la traduction de l'IFN $\gamma$ . La transcription du gène codant pour l'IFN $\gamma$  est régulée au niveau du promoteur et des séquences introniques, notamment par méthylation de l'ADN (Young et al., 1995).

##### 1.1.4.2. La régulation de l'expression

Les cellules NK et les lymphocytes T ne produisent pas d'IFN $\gamma$  lorsqu'elles sont dans un état stationnaire. Cependant, plusieurs agents peuvent les activer et induire la production de la cytokine.

###### Les stimuli exogènes

Les stimuli exogènes pouvant induire la production d'IFN $\gamma$  sont variés. Généralement ce sont des ligands de récepteurs situés sur les cellules productrices. Certains se manifestent *in vivo* dans l'environnement cellulaire en réponse à des situations telles que le cancer, les allergies, les infections, les traumatismes ou les réactions auto-immunes. D'autres sont plus connus puisqu'ils sont expérimentalement utilisés de manière artificielle pour induire la production (Billiau et Vandebroek, 2000).

De façon non exhaustive, les stimuli exogènes peuvent être i) toutes les substances reconnues par les lymphocytes T comme du non-soi, ii) certains micro-organismes tel que le *Staphylococcal enterotoxins* produisant des superantigènes, iii) les lipopolysaccharides (LPS) (également désignés comme des endotoxines) des membranes externes de bactéries gram négatives, iv) des lectines végétales telle que la phytohemagglutinine A (PHA) et la concanavaleine A (Con A). Expérimentalement, les anticorps dirigés contre des composés de la membrane des lymphocytes sont couramment utilisés pour induire l'expression d'IFN $\gamma$  (Billiau et Vandebroek, 2000).

### Les stimuli endogènes

Les deux principaux facteurs endogènes pouvant induire la production d'IFN $\gamma$  sont les interleukines 12 et 18 (IL-12 et IL-18). L'IL-12, également connue sous le nom de facteur stimulant les cellules NK ou de facteur de maturation cytotoxique des lymphocytes, est une cytokine produite principalement par les macrophages activés et les lymphocytes B qui exercent ses effets sur les cellules T ou les cellules NK. De plus, l'IL-12 joue un rôle majeur dans le développement et la différenciation de lymphocytes Th1 producteurs d'IFN $\gamma$  (Brunda, 1994 ; Trinchieri, 1995). L'IL-18 est une protéine initialement décrite comme IGIF pour IFN $\gamma$ -Inducing-Factor. Le précurseur de l'IL-18 est produit par divers types cellulaires tels que les phagocytes mononucléaires ou les kératinocytes. Ce précurseur (Pro-IL-18) est clivé en IL-18 activé par la caspase 1 alors que la caspase 3 va désactiver la cytokine. L'IL-18 induit ou augmente le taux de production d'IFN $\gamma$  en présence d'autres cytokines telles que l'IL-2 ou l'IL-12. Le rôle de l'IL-12 est en partie lié à une expression accrue du récepteur de l'IL-18 sur les cellules T. *In vivo*, de nombreuses études ont démontré le rôle crucial de ces deux cytokines dans l'induction de l'IFN $\gamma$  par des agents exogènes (Dinarello et al., 1998 ; Billiau et Vandebroek, 2000).

### Les modulateurs de l'IFN $\gamma$

Si certaines cytokines permettent d'induire la production d'IFN $\gamma$ , l'IL-10 en particulier va être un inhibiteur fort de sa production (Moore et al., 1993). L'IL-10, décrit comme un facteur d'inhibition de la synthèse de cytokine, est produit par les cellules Th2, les monocytes, les lymphocytes B ou les phagocytes mononucléaires. L'effet antagoniste de l'IL-10 sur l'IFN $\gamma$  est réciproque et la balance entre les deux cytokines apparaît être un élément important dans les premières phases de la réponse immunitaire.

L'IL-4 est également un antagoniste de l'IFN $\gamma$  et réciproquement. Sans rentrer dans les détails, selon la réponse immunitaire vers laquelle l'organisme s'oriente (Th1 ou Th2), l'une ou l'autre des cytokines prendra le dessus.

Le TGF $\beta$ 1 (Transforming Growth Factor  $\beta$ 1), bien qu'étant capable d'induire la transcription du gène de certaines cytokines, est un facteur qui a la capacité d'inhiber la production d'IFN $\gamma$  induite par des facteurs exogènes tels que les LPS ou la PHA (Billiau et Vandebroek, 2000).

La prostaglandine est également un élément de régulation de l'IFN $\gamma$ . Elle dirige la réponse immunitaire dans la direction des cellules Th2. En effet, elle inhibe la production de cytokines telles que l'IFN $\gamma$  ou l'IL-2 par les cellules Th1 mais n'affecte pas les cellules Th2 et la production d'IL-4 qui leur est associée.

Outre ces antagonistes d'IFN $\gamma$ , le TNF $\alpha$  va principalement avoir un effet synergique vis-à-vis de l'IFN $\gamma$ , et cela a été observé aussi bien *in vitro* qu'*in vivo* (Billiau et Vandebroek, 2000).

### **1.1.5. Principales activités biologiques**

L'IFN $\gamma$ , comme beaucoup d'autres cytokines, exerce de multiples fonctions et est capable de réguler et moduler plus de 200 gènes distincts (Boehm et al., 1997). En particulier, il est un facteur clé de la régulation des différents stades de la réponse immune car il possède des propriétés à la fois antivirales, antiprolifératives et immunomodulatrices (Ijzermans et Marquet, 1989 ; Williams et al., 1993). Il participe également à la régulation de diverses protéines de la matrice extracellulaire. D'abord connu pour son activité antivirale, l'IFN $\gamma$  constitue avant tout un agent immunomodulateur important (De Maeyer et De Maeyer-Guignard, 1992). Pour permettre la protection antivirale, l'IFN $\gamma$  induit de nombreuses protéines antivirales (la 2',5'-oligoadénylate synthétase par exemple) (Billiau et Vandebroek, 2000).

#### **1.1.5.1. Principales activités observées *in vitro***

*In vitro*, l'IFN $\gamma$  est bien connu pour ses capacités à potentialiser la chaîne respiratoire des macrophages. Cela permet, entre autre, d'augmenter la production d'ion superoxyde ( $O_2^-$ ) ou d'espèces réactives de l'oxygène ( $H_2O_2$ ) (Nathan et al., 1983). Ceci semble étroitement lié aux capacités de l'IFN $\gamma$  à réguler la transcription du gène codant pour une enzyme du complexe de la NADPH oxydase (Cassatella et al., 1990). Dans le même registre, l'IFN $\gamma$  joue un rôle important dans la production d'oxyde nitrique (NO). En particulier, l'IFN $\gamma$  est un inducteur important du tétrahydrobioptérine, facteur limitant de la production de NO. A noter que le NO réagit avec l' $H_2O_2$  pour générer des espèces réactives de l'oxygène ; ce phénomène est toxique pour les cellules, notamment pour celles dont la glycolyse est bloquée (Dijkmans et Billiau, 1991). Ainsi le NO produit par les macrophages peut avoir un effet cytotoxique pour les cellules voisines ce qui peut s'avérer utile envers des cellules cancéreuses ou contre des infections par exemple de types bactériennes ou protozoaires. En revanche, cela peut causer des dommages vis-à-vis de cellules ou tissus sains.

Bien sûr, l'une des activités les plus référencées de l'IFN $\gamma$  est l'induction des antigènes du complexe majeur d'histocompatibilité (CMH) de classe II sur de nombreux types cellulaires. L'expression d'antigènes du CMH de classe I peut dans certains cas, mais plus rarement, être

induite par la cytokine. Les cellules qui expriment ces antigènes vont être des cibles plus facilement identifiables pour les lymphocytes T cytotoxiques.

L'IFN $\gamma$  exerce également ses activités de cytokine pro-inflammatoire en augmentant sur les phagocytes ou les neutrophiles, l'expression de Fc $\gamma$ RI (Fc $\gamma$  Receptor I), de même avec l'augmentation de l'expression des molécules ICAM-1 (Intercellular Adhesion Molecule 1) dont les interactions avec les LFA-1 (Leukocyte Function-Associated antigen) sont en partie responsables du recrutement leucocytaire.

De plus, l'expression d'IFN $\gamma$  dans le microenvironnement cellulaire joue un rôle important aussi bien dans la différenciation que dans la maturation des phagocytes mononucléaires (MPCs), à savoir notamment les monocytes, les précurseurs situés dans la moelle et les différents types de macrophages tissulaires.

Même si l'IFN $\gamma$  possède, dans de nombreux types cellulaires, une action inhibitrice de la prolifération, l'effet de la cytokine serait plutôt inverse sur les cellules T. L'ajout d'IFN $\gamma$  permet d'augmenter l'activité des lymphocytes T cytotoxiques. Cette observation est une résultante d'un effet direct d'IFN $\gamma$  sur l'activation des précurseurs T en lymphocytes T CD8+ cytotoxiques. Sur les lymphocytes T CD4+ auxiliaires (de type Th2), en revanche, l'IFN $\gamma$  aurait plutôt un effet inhibiteur (Siegel, 1988).

Enfin, pour en revenir à l'activité qui a fait connaître l'IFN $\gamma$ , l'activité antivirale, elle provient des capacités de la cytokine à induire chez les cellules un état de résistance aux infections virales. Le mécanisme adopté est similaire à celui des IFN $\alpha$  et  $\beta$  et son activité spécifique souvent 10 à 100 fois inférieure à ces derniers. Ainsi comme l'IFN $\alpha$  ou  $\beta$ , l'IFN $\gamma$  agit en levant, par exemple, l'inactivation des gènes codant pour la 2'-5'-oligoadénylsynthase et la PKR (dsRNA-dependent protein kinase) (Billiau et Vandebroek, 2000).

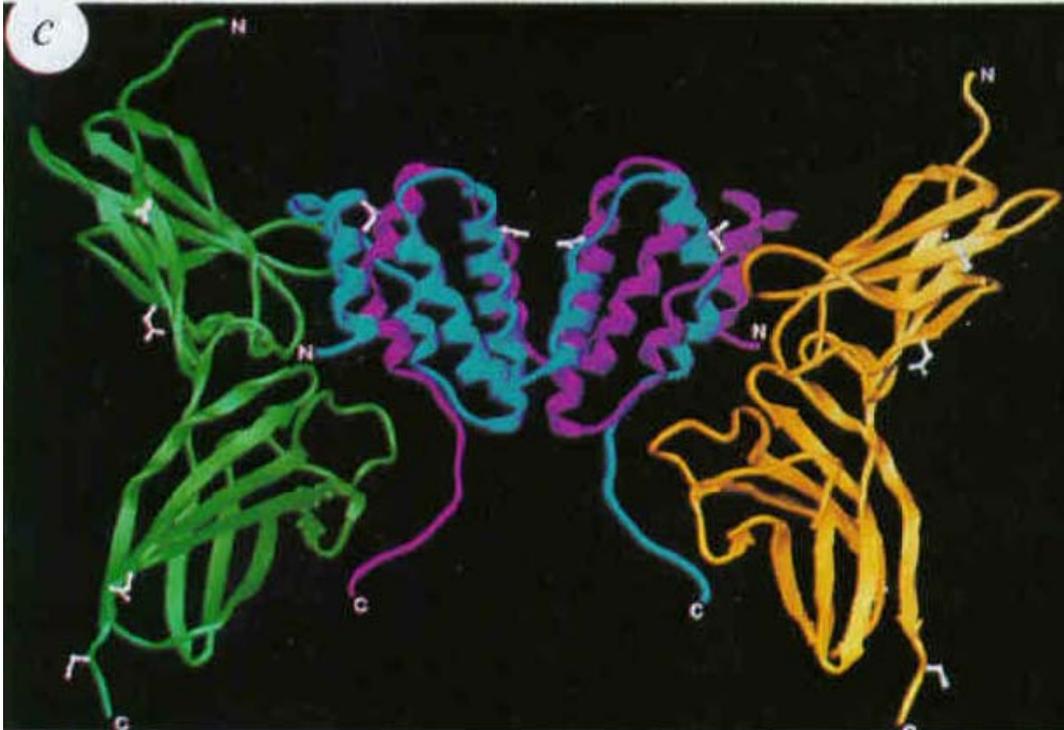
#### **1.1.5.2. Principales activités observées *in vivo***

Vis-à-vis des maladies infectieuses, il a été montré que les souris déficientes en IFN $\gamma$  ne survivaient pas à une infection par *Mycobacterium tuberculosis*. De même, des études ont révélé que les macrophages induits à l'IFN $\gamma$  protégeaient l'organisme contre des infections bactériennes telles que *Listeria* (Peck, 1989) ou *Salmonella* (Kagaya et al., 1989). Il a été démontré que la production de NO induite par l'IFN $\gamma$  était aussi utile par exemple pour défendre l'organisme contre *Trypanosoma cruzi* (Mayer et al., 1993), ou *Histoplasma capsulatum* (Nakamura et al., 1994).

Comme il a été mentionné plusieurs fois, l'IFN $\gamma$  joue un rôle important pour protéger l'organisme vis-à-vis d'éventuelles infections virales. Ainsi, l'apport en IFN $\gamma$  va par exemple permettre de protéger les souris ou les rats d'une infection expérimentale par le cytomegavirus ou CMV (Fennie et al., 1988 ; Haagmans et al., 1994), par le MHV (Mouse Hepatitis Virus) (Lucchiari et al., 1992), ou par le virus de l'Herpès humain (Stanton et al., 1995). Dans ce genre d'expérimentation, comme Billiau et Vandebroek (2000) le soulèvent, il n'est pas facile de savoir si l'effet antiviral observé est réellement dû à une activité antivirale ou à une activité immunomodulatrice, et, par conséquent, plutôt à l'activation des cellules NK ou plutôt à la maturation des LT vers les lymphocytes T cytotoxiques (Billiau et Vandebroek, 2000).

Certains modèles expérimentaux ont permis de placer l'IFN $\gamma$  comme un médiateur à prendre en considération dans les processus tumoraux. En effet, il a été observé que le rejet de tumeurs expérimentales chez la souris était lié à la présence de la cytokine dans le tissu tumoral. Ce phénomène est alors inhibé par l'utilisation d'anti-IFN $\gamma$  (Prat et al., 1987 ; Jarpe et al., 1989). De même, l'utilisation de cellules à fort pouvoir métastatique modifiées pour sur-exprimer IFN $\gamma$  se révèle largement moins tumorigène (Gansbacher et al., 1990). Dans le même sens, des souris n'exprimant pas le récepteur d'IFN $\gamma$  ou déficientes en STAT-1 (impliqué dans la voie de signalisation de la cytokine) développent plus rapidement et fréquemment des tumeurs induites par le méthylcholanthrène que les souris normales (Billiau et Vandebroek, 2000).

Les réponses des leucocytes et des cellules endothéliales à une stimulation par l'IFN $\gamma$  relèvent souvent de l'immunostimulation. Ceci étant, nombreuses expérimentations ont également mis en avant des propriétés immunosuppressives de la cytokine. Billiau et Vandebroek (2000) évoquent parfaitement ces propriétés opposées selon les pathologies immunitaires dont il est question. Par exemple, l'IFN $\gamma$  a des effets immunosuppresseurs mis en évidence par Konieczny et al. (1998) dans des modèles de greffes chez des souris privées d'IFN $\gamma$ . Les effets immunostimulants de l'IFN $\gamma$  ont, quant à eux, été observés dans des modèles de colites (Rennick et al., 1997 ; Guy-Grand et al., 1998).



**Figure 33: Structure cristallographique du complexe entre l'IFN $\gamma$  et la partie soluble de la chaîne  $\alpha$  du récepteur à IFN $\gamma$**

Le dimère d'IFN $\gamma$  interagit avec deux IFN $\gamma$ R $\alpha$ . Chaque chaînes polypeptidique composant l'IFN $\gamma$  (bleu/rose) ou son récepteur (vert/jaune) sont distinguées par une couleur différente. Les résidus 125 à 132 de l'IFN $\gamma$  ont été disposées par modélisation après résolution de la structure.

(Figure extraite de Walter et al., 1995)

## **1.2. LES LIGANDS DE L'IFN $\gamma$**

### **1.2.1. Le récepteur membranaire de signalisation de l'IFN $\gamma$**

#### **1.2.1.1. Le récepteur à IFN $\gamma$ : l'IFN $\gamma$ R**

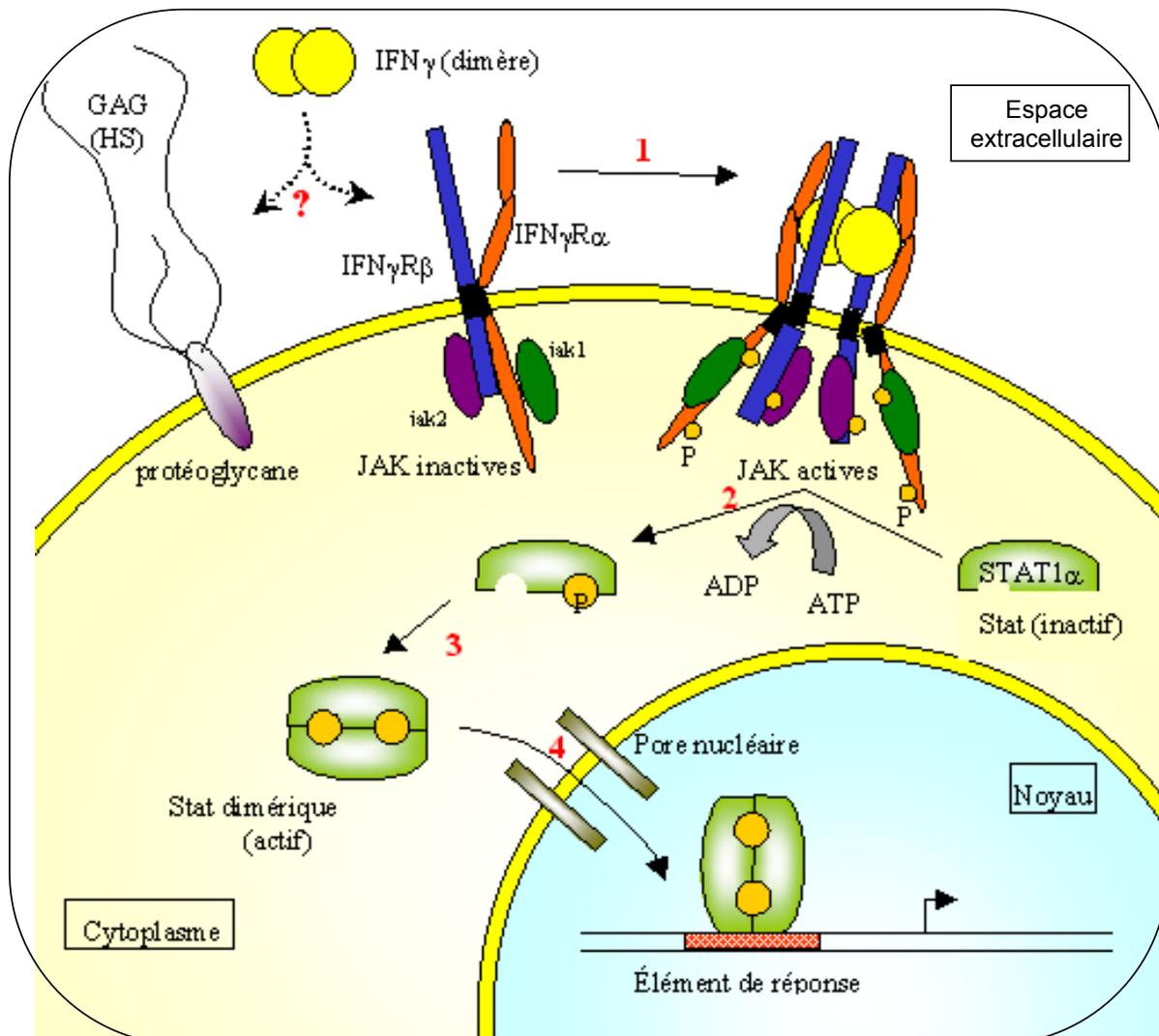
L'IFN $\gamma$  transmet son signal aux cellules cibles via un récepteur spécifique. Ce récepteur, l'IFN $\gamma$ R, est exprimé de manière ubiquitaire sur la grande majorité voire toutes les cellules (Schroder et al., 2004). Selon les cellules et les conditions microenvironnementales, 500 à 20000 récepteurs à IFN $\gamma$  par cellule peuvent être présents. L'IFN $\gamma$ R est composé de deux sous-unités qui sont les chaînes polypeptidiques IFN $\gamma$ R $\alpha$  et IFN $\gamma$ R $\beta$ .

L'IFN $\gamma$ R $\alpha$  permet l'interaction avec la cytokine avec une affinité de l'ordre du nanomolaire ( $K_d = 0,1 - 1 \text{ nM}$ ) (Anderson et al., 1982). Cette sous-unité transmembranaire (472 acides aminés) est composée d'un domaine extracellulaire de 228 aa, d'une région transmembranaire de 24 aa et d'un domaine intracellulaire de 220 aa. Le domaine intracellulaire comporte un site de fixation à Jak1, riche en résidus sérines et thréonines qui vont pour la plupart être phosphorylés durant la signalisation.

L'IFN $\gamma$ R $\alpha$  est constitutivement associé avec une autre protéine transmembranaire, l'IFN $\gamma$ R $\beta$ . Cette dernière joue un rôle moins important pour l'interaction, mais est primordiale pour la transduction du signal. Comme l'IFN $\gamma$ R $\alpha$ , elle est composée d'un domaine extracellulaire (226 aa), d'une région transmembranaire (25 aa) et d'un domaine intracellulaire de 64 aa. Ce dernier domaine possède un site de liaison pour Jak2 (Greenlund et al., 1994).

#### **1.2.1.2. L'interaction IFN $\gamma$ /IFN $\gamma$ R**

Comme il a été mentionné dans le paragraphe précédent, l'affinité de la cytokine pour son récepteur membranaire est de l'ordre du nanomolaire. Cette interaction met en jeu un IFN $\gamma$  (dimérique) et deux hétérodimères d'IFN $\gamma$ R( $\alpha/\beta$ ) (Fountoulakis et al., 1992). Les éléments structuraux de l'IFN $\gamma$  nécessaire à l'interaction ont été déterminés (Lundell et Narula 1994). Trois régions distinctes sont importantes, à savoir la boucle entre les deux premières hélices alpha (résidus 18-26), une séquence de l'hélice F (108-124) et une courte section de la région C-terminale (résidus 128-131) (Sadir et al., 1998 ; Walter et al., 1995) (Figure 33). Comme



**Figure 34: Schéma de la voie de signalisation JAK/STAT par l'IFN $\gamma$**

**1** : Le dimère IFN $\gamma$  interagit avec son récepteur (IFN $\gamma$ R) qui se dimérise, et induit le rapprochement des Janus kinases (jak) et leur transphosphorylation (●), ainsi que la phosphorylation de la Tyr 701 sur l'IFN $\gamma$ R $\alpha$ .

**2** : cette configuration permet l'activation des facteurs de transcription STAT1 $\alpha$  par phosphorylation.

**3** : les STAT phosphorylés s'homodimérisent et sont transloqués dans le noyau où ils activent les éléments de réponse (**4**).

**?** : L'IFN $\gamma$  a deux ligands : les protéoglycannes à HS et l'IFN $\gamma$ R. Les affinités sont comparables, mais l'abondance des protéoglycannes à la surface des cellules pourrait constituer un " passage presque obligé" pour l'IFN $\gamma$ .

Le mécanisme qui fait passer l'IFN $\gamma$  des structures HS à son récepteur reste encore à déterminer.

*D'après (Sakatsume et al., 1995; Stark et al., 1998; Levy, 1995; Briscoe et al., 1996 ; Greenlund et al., 1995; Heim et al., 1995 ; Pelligrini et Schindler, 1993)*

pour la structure cristallographique de l'IFN $\gamma$  seul, la structure du complexe n'a pas permis de résoudre la structure de la région C-terminale de la cytokine, trop flexible pour donner un signal interprétable.

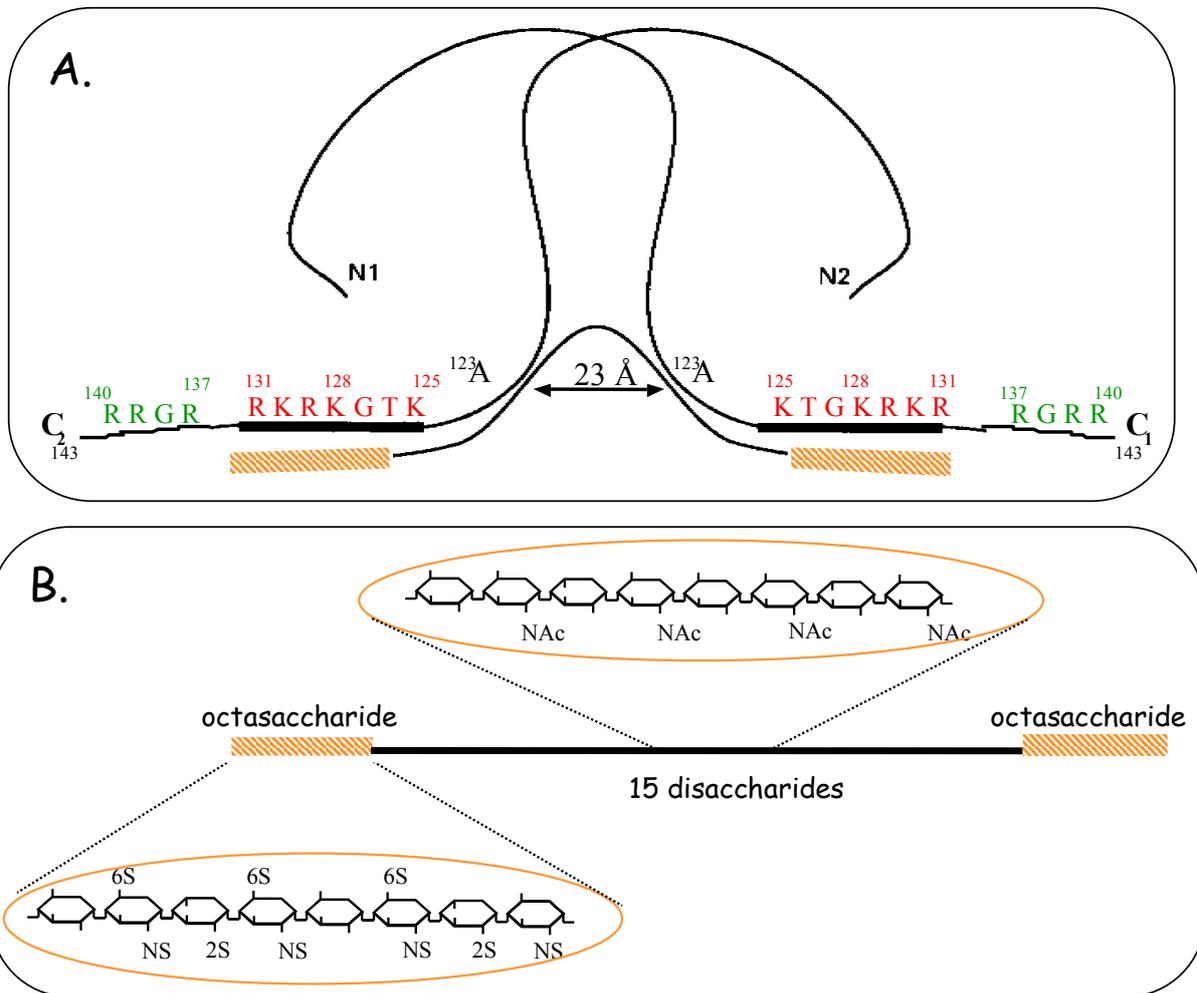
### **1.2.1.3. La transduction du signal de l'IFN $\gamma$**

L'IFN $\gamma$  utilise, pour transmettre son message aux cellules cibles, la voie de signalisation JAK-STAT (Janus Kinase et Signal Transducers and Activators of Transcription). Pour ce faire, les deux sous-unités de l'IFN $\gamma$ R, l'IFN $\gamma$ R $\alpha$  et l'IFN $\gamma$ R $\beta$  interagissent de manière constitutive au niveau de leur partie cytoplasmique avec les protéines kinases qui sont respectivement Jak1 et Jak2 (Sakatsume et al., 1995). Ces deux kinases interagissent avec des régions identifiées, proches de la membrane plasmique (Stark et al., 1998). La fixation de l'IFN $\gamma$  (dimérique) induit la dimérisation des deux sous-unités alpha du récepteur, ce qui permet le rapprochement des kinases (Levy, 1995). Les Jaks s'activent par phosphorylation croisée et permettent ainsi la phosphorylation des tyrosines du récepteur. L'activation de Jak2 se produit en premier et est indispensable à l'activation des Jak1 (Briscoe et al., 1996). Après phosphorylation, les deux chaînes alpha du récepteur offrent des sites proches de la partie C-terminale, reconnus par les domaines SH2 des facteurs STAT1 $\alpha$  (Greenlund et al., 1995; Heim et al., 1995). Ces derniers sont à leur tour phosphorylés par les Jaks (Schindler et al., 1992). Ceci va leur donner la capacité de se dimériser pour former le facteur GAF (Gamma Activation Factor) qui, une fois dans le noyau, va se fixer sur des séquences cibles de type GAS (Gamma Activation Site) en amont des gènes régulés par l'IFN $\gamma$  (Pelligrini et Schindler, 1993) (Figure 34).

## **1.2.2. L'interaction IFN $\gamma$ / HS**

### **1.2.2.1. Caractéristiques de l'interaction**

A l'origine de la découverte de cette interaction, des travaux mettent en évidence que l'IFN $\gamma$  interagit de manière importante avec des membranes basales. Le ligand principal alors identifié est l'héparane sulfate. L'affinité de cette interaction, de l'ordre de 1,5 nM, justifie l'intérêt porté à sa caractérisation (Lortat-Jacob et al., 1991). Des investigations plus poussées ont permis de démontrer que deux domaines riches en acides aminés basiques de la région C-



**Figure 35: Modèle de l'interaction IFN $\gamma$  / HS**

**A :** D'après ce modèle, l'IFN $\gamma$  homodimérique interagit par l'intermédiaire de ses deux domaines D1 (acides aminés en rouge) avec un segment d'HS caractéristique (deux courts domaines octasaccharidiques N-sulfatés (orangé) reliés par un long domaine N-acétylé d'environ 14 - 15 disaccharides). Le domaine D2 (deuxième domaine basique de la région C-terminale de l'IFN $\gamma$ ) est représenté en vert.

**B :** Séquence oligosaccharidique d'HS nécessaire pour l'interaction avec l'IFN $\gamma$ . Cette séquence est composée de deux domaines octasaccharidiques fortement sulfatés (important pour l'interaction) reliés entre eux par une quinzaine de disaccharides non sulfatés.

(d'après Lortat-Jacob et col., 1995).

terminale de la cytokine sont impliqués dans l'interaction avec les HS (Lortat-Jacob et Grimaud, 1991 ; Vanhaverbeke et al., 2004). Ces domaines C-terminaux, KTGKRKR et RGRR, dénommés respectivement D1 et D2, interagissent avec des domaines structuraux d'HS de type héparine (Lortat-Jacob et Grimaud, 1992). Grâce à une approche biochimique de type "empreinte moléculaire", la séquence oligosaccharidique interagissant avec la cytokine a pu être partiellement caractérisée. Il s'agit d'un domaine central N-acétylé riche en Glc-A, d'environ 7 kDa, encadré par deux domaines N-sulfatés de 6 à 8 saccharides. Ces deux domaines anioniques semblent alors interagir avec les régions C-terminales basiques de la protéine (Figure 35). Dans la caractérisation d'interactions protéines-GAGs, ce fut la première fois que ce type d'organisation oligosaccharidique (domaine SAS) fut mis à jour (Lortat-Jacob et al., 1995). L'organisation des GAGs en domaine SAS est schématiquement la succession de régions totalement sulfatées (S) entre lesquelles s'intercalent des régions pas ou peu sulfatées (A pour acétylé).

#### 1.2.2.2. Rôle de l'interaction

L'interaction de l'IFN $\gamma$  avec les HS met principalement en jeu le domaine D1 de la région C-terminale de la cytokine (Lortat-Jacob et Grimaud, 1991). Les deux domaines de la région C-terminale sont particulièrement sensibles aux coupures protéolytiques. De plus, l'activité de l'IFN $\gamma$  est potentialisée si les 10 derniers acides aminés ont été clivés (coupures en aval du domaine D1). A l'inverse, cette activité est largement diminuée si les 14 (ou plus) derniers acides aminés sont supprimés (coupures dans ou en amont de D1) (Lortat-Jacob et al., 1996 ; Döbeli et al., 1988). La fixation de l'IFN $\gamma$  aux HS, protégeant spécifiquement le domaine D1, mais pas le domaine D2, des dégradations protéolytiques représente un mécanisme de régulation de l'activité de la cytokine.

Des études *in vivo* ont montré que la fixation de l'IFN $\gamma$  aux HS régulaient l'élimination plasmatique, la distribution tissulaire et l'accumulation locale de la protéine (Lortat-Jacob et al., 1996b).

Ainsi l'adhésion aux HS permet de **localiser, protéger et moduler les effets de l'IFN $\gamma$**  (Lortat-Jacob et al., 1994 ; Lortat-Jacob et al., 1996a ; Lortat-Jacob et al., 1996b ; Sadir et al., 1998).

## 2. LES TROUBLES PHYSIOLOGIQUES LIES À L'IFN $\gamma$

### 2.1. EXEMPLES DE MALADIES ASSOCIEES A L'IFN $\gamma$

L'IFN $\gamma$  est capable de réguler au moins 200 gènes différents. Il est impliqué de façon majeure dans la régulation de la réponse immune, et a une activité pro-inflammatoire importante (Boehm et al., 1997).

Dans les troubles tel que l'**athérosclérose**, une *maladie inflammatoire* chronique des vaisseaux avec des lésions dans les parois artérielles, l'IFN $\gamma$  semble être un médiateur central (Leon et Zuckerman, 2005). Cependant, les travaux réalisés ne permettent pas, pour l'instant, de dire si, dans ce désordre là, la cytokine serait un pro- ou anti-athérogène. De plus, l'effet d'IFN $\gamma$  dans la balance athérogénique semble potentiellement lié à l'évolution des troubles (Harvey et Ramji, 2005).

De même, dans les maladies inflammatoires ou auto-immunes du côlon (colites) telles que **la maladie de Crohn** et **la colite ulcéraive**, les taux de cytokines pro-inflammatoires (IFN $\gamma$ , TNF- $\alpha$ ) détectés dans le sang ou les tissus, sont supérieurs à la normale. Ces protéines ont été définies comme des cibles thérapeutiques potentielles (Ghosh et al., 2006).

Lorsque la production d'IFN $\gamma$  est dérégulée, de *nombreuses maladies auto-immunes* peuvent être développées. Il se peut donc que le taux d'IFN $\gamma$  soit trop important, ce qui est le cas en particulier **pour l'arthrite rhumatoïde, les différentes scléroses, le diabète de type I, la schizophrénie, ainsi que différentes maladies auto-immunes de la peau** (Skurkovich et Skurkovich, 2003).

Dans le cas de la **leucémie**, les médecins ont souvent recours à la greffe de la moelle osseuse. Une greffe acceptée par l'organisme peut être signe de guérison. Malheureusement, il arrive parfois que les lymphocytes T matures génèrent « la maladie du greffon contre l'hôte » (GVH), une maladie mortelle. Il s'est avéré que dans le cas de la GVH, l'apport d'IFN $\gamma$  semble bénéfique dans la lutte contre cet effet secondaire grave. Cependant, l'IFN $\gamma$  semble avoir un effet néfaste sur la greffe en elle-même (Yang et al., 2005).

### 2.2. EXEMPLES DE TRAITEMENTS POUR MODULER L'ACTIVITE D'IFN $\gamma$

Dans les pathologies liées à une surexpression d'IFN $\gamma$ , la modulation de l'IFN $\gamma$  endogène et, plus particulièrement, son inhibition est importante. Dans le cas des maladies

inflammatoires ou auto-immunes telle que la maladie de Crohn, le but est de réduire significativement l'action d'IFN $\gamma$ . Ainsi, **Fontolizumab**, un anticorps monoclonal humanisé anti-IFN $\gamma$  est actuellement en essais cliniques et dans certaines conditions se trouve avoir un effet positif sur la maladie (Hommes et al., 2006 ; Reinisch et al., 2006 ; Ghosh et al., 2006).

**Innogenetics** utiliserait quant à elle un anticorps monoclonal anti-IFN $\gamma$  (**INNO-202**) pour traiter la septicémie (Lainee et al., 2005).

Pour information, Skurkovich et Skurkovich (2003) citent, dans leur revue, les maladies autoimmunes pour lesquelles des anticorps anti-IFN $\gamma$  peuvent s'avérer efficaces. Cela est par exemple le cas pour l'arthrite rhumatoïde, différentes scléroses, et plusieurs maladies autoimmunes de la peau. De même dans le rejet de greffe (comme par exemple la transplantation de cornée), l'inhibition de l'IFN $\gamma$  a été envisagée (Skurkovich et Skurkovich, 2003).

Outre les désordres liés à une surexpression d'IFN $\gamma$ , il existe certaines pathologies liées à un manque évident de la cytokine. Ainsi dans certains cas, de l'IFN $\gamma$  exogène et recombinant, peut être utilisé comme molécule thérapeutique.

Par exemple, la société **InterMune** a déjà commercialisé, aux USA, Actimmune® (IFN $\gamma$ ) via Genentech pour traiter les granulomatoses chroniques et l'ostéoporose, et souhaite le tester pour soigner les infections par le virus de l'hépatite C. En Europe, **Boehringer Ingelheim** le commercialise sous le nom d'Immukine®. InterMune réalise également des essais cliniques de phase III avec de l'IFN $\gamma$  pour le traitement de fibroses pulmonaires idiopathiques.

Aussi, dans le cas du cancer où l'immunosurveillance est souvent insuffisante, les propriétés anti-tumorales et immunostimulantes de l'IFN $\gamma$  peuvent être utiles. (Brandacher et al., 2006). En complément du TNF $\alpha$ , l'apport de l'IFN $\gamma$  apparaît intéressant pour lutter contre les mélanomes de la peau (Lejeune et al., 1998). Il en est de même avec l'IL-12 dans le traitement de mélanomes métastatiques (Kim et al., 1996) ou avec le 5-fluorouracil dans le cancer colorectal (Pavlidis et al., 1996). La société **Transgène** développe (phase clinique II) le **TG-1042**, un adénovirus non-répliquatif contenant le gène codant pour l'IFN $\gamma$ , connu également sous le nom d'adéno-interféron-gamma (Ad-IFN $\gamma$ ). Il est indiqué pour les cas de lymphomes cutanés à cellules B et potentiellement dans certaines tumeurs solides primaires et métastatiques.

# **OBJECTIFS DU TRAVAIL**

Le fil conducteur de ce travail, tant bibliographique qu'expérimental, repose sur le fait que les PGs, d'une manière générale, sont des macromolécules essentielles dans la plupart des processus qui permettent le bon fonctionnement de l'organisme. Ainsi, les travaux présentés dans ce manuscrit et résultant de ces trois dernières années d'études ont été organisés en deux grands exposés afin de mettre en avant comment ces macromolécules peuvent être caractérisées, et surtout comment de telles caractérisations peuvent permettre le développement de molécules présentant un intérêt thérapeutique.

La première partie porte ainsi sur l'étude structurale et fonctionnelle d'un protéoglycane récemment découvert: l'endocan. La seconde partie expose les travaux de caractérisation structurale et fonctionnelle de l'interaction entre IFN $\gamma$  et les HS qui ont permis le développement d'un mime oligosaccharidique permettant de moduler les activités biologiques de cette cytokine.

### ✓ ***ETUDE STRUCTURALE ET FONCTIONNELLE D'ENDOCAN***

Les mécanismes d'action d'endocan sont encore mal connus mais, comme cela a été dit, dans les processus tels que le développement tumoral ou l'inflammation, ce PG récemment découvert semble avoir des fonctions non négligeables qu'il reste encore à approfondir. Sa caractérisation structurale ainsi que l'étude des interactions qu'il est capable d'engager avec différents médiateurs peuvent ouvrir des voies permettant de mieux comprendre son fonctionnement et ses utilités.

Dans un premier temps, il a donc été nécessaire de mettre au point une plateforme de production et purification, à la fois, d'endocan, de son core protéique et de sa chaîne glycanique. Dans le cadre d'Endotis Pharma, cette plateforme était nécessaire pour le développement d'un kit de dosage d'endocan et surtout pour la réalisation d'essais *in vivo*, très consommateurs. Dans le cadre de ce travail de thèse, l'obtention de ces trois molécules a permis de réaliser l'ensemble des autres travaux. Cette plateforme a été optimisée pour obtenir chacune de ces molécules à l'échelle du milligramme. Le mode de production original mise en place fait l'objet d'un article en fin d'écriture : Adam et al., en préparation.

La caractérisation structurale de la chaîne d'endocan et du core protéique d'endocan a été réalisée de façon indépendante.

-A l'issue de ce travail, un modèle de l'organisation structurale de ces chaînes oligosaccharidiques a été proposé (résultats non publiés).

- L'approche structurale du core protéique d'endocan par cristallographie a été engagée. Elle n'a pas permis pour l'instant d'aboutir à l'obtention de cristaux nécessaires à la résolution d'une structure pour endocan. Malgré tout, plusieurs aspects structuraux ont pu être mis à jour. Le profil de l'organisation des ponts disulfures a été tracé, un motif EGF-like a pu être dégagé. La présence sur le core protéique d'endocan de glycosylations de type O- et N-glycannes de nature non GAG a été remarquée.

Une étude de l'interaction entre endocan et différents ligands a été réalisée. Elle a permis, premièrement, d'obtenir l'affinité d'anticorps monoclonaux vis-à-vis d'endocan, développés dans le cadre d'une trousse de dosage de type ELISA, ou comme agents thérapeutiques. Deuxièmement, un criblage de l'interaction entre plusieurs médiateurs pro-inflammatoires ou pro-tumoraux et endocan, son core protéique ou sa chaîne glycanique a été fait, et les affinités de l'HGF/SF et de l'IFN $\gamma$  vis-à-vis de la chaîne glycanique d'endocan ont pu être mesurées.

Finalement, d'un point de vue plus fonctionnel, l'effet d'endocan sur la protéine IFN $\gamma$  a pu être mesuré. Via sa chaîne de DS, endocan apparaît être dans ces tests un inhibiteur de la cytokine. Ensuite, un criblage pour détecter la présence de l'ARNm d'endocan ou de l'endocan sécrété a été effectué sur plusieurs lignées cellulaires. Les résultats confirment qu'endocan ne peut plus être restreint à ses qualités de molécule spécifique des cellules endothéliales.

**✓ ETUDE STRUCTURALE ET FONCTIONNELLE DE L'INTERACTION IFN $\gamma$ /HS : DEVELOPPEMENT D'UN MIMÉ OLIGOSACCHARIDIQUE MODULANT L'ACTIVITE D'IFN $\gamma$ .**

La région C-terminale de l'IFN $\gamma$  est importante pour l'activité de la cytokine. Or, l'IFN $\gamma$  interagit par le biais de cette région C-terminale avec les HS. Un modèle de cette interaction a été précédemment proposé (Lortat-Jacob et al., 1995).

Sur ce modèle d'interaction, des mimes oligosaccharides ont été pensés et obtenus par synthèse chimique : Lubineau et al., 2004.

L'objectif a alors été de tester si ces molécules étaient capables de mimer les HS dans l'interaction avec l'IFN $\gamma$ , et surtout d'évaluer, *in vitro*, si certaines étaient capables de moduler l'activité de la cytokine. Une molécule (appelée 2O<sub>10</sub>) s'est révélée avoir les mêmes propriétés que l'héparine *in vitro* vis-à-vis d'IFN $\gamma$ . Elle constitue donc un bon mime de l'Hp et des HS et se place comme une molécule intéressante pour inhiber l'activité de l'IFN $\gamma$ . Cela a fait l'objet d'une publication en 2005 : Sarrazin et al., 2005.

L'interaction de l'IFN $\gamma$  avec les HS est rendue possible par des domaines basiques (D1 et D2) de la région C-terminale de la protéine. Afin d'étudier plus en détail l'importance que peut avoir chacun des acides aminés basiques de ces domaines dans l'interaction avec les HS ou pour le mécanisme d'action de la cytokine, un travail de mutagenèse sur ces domaines a été entrepris.

Un travail a été mené pour résoudre par cristallographie la structure de la région C-terminale d'IFN $\gamma$  jusqu'à présent inconnue. Pour cela, des expériences de cristallogénèses ont été réalisées sur des complexes entre différents GAGs naturels ou de synthèse afin de stabiliser la région C-terminale de la protéine jugée trop labile. Ce travail n'a pour l'instant abouti qu'à quelques pistes qu'il reste encore à explorer.

Au total, les travaux réalisés au cours de ces études ont donné lieu à la publication de deux manuscrits originaux et quatre articles de revue.

Lubineau, A., H. Lortat-Jacob, O. Gavard, S. Sarrazin, D. Bonnaffé (2004). "Synthesis of tailor-made glycoconjugate mimetics of heparan sulfate that bind IFN-gamma in the nanomolar range." Chemistry **10**(17): 4265-82.

Sarrazin, S., D. Bonnaffé, A. Lubineau, H. Lortat-Jacob. (2005). "Heparan sulfate mimicry: a synthetic glycoconjugate that recognizes the heparin binding domain of interferon-gamma inhibits the cytokine activity." J Biol Chem **280**(45): 37558-64.

Sarrazin, S., E. Adam, M. Lyon, F. Depontieu, V. Motte, C. Landolfi, H. Lortat-Jacob, D. Béchar, P. Lassalle, M. Delehedde. (2006). "Endocan or endothelial cell specific molecule-1 (ESM-1): A potential novel endothelial cell marker and a new target for cancer therapy." Biochim Biophys Acta. **1765**(1): 25-37 Epub 2005 Aug 26.

Delehedde M., S. Sarrazin, E. Adam, V. Motte, C. Vanpouille. (2006). "Proteoglycans and glycosaminoglycans: Complex Molecules with modulating activity" In "New developments in Therapeutic Glycomics". Eds. Maryse Delehedde and Hugues Lortat-Jacob. Chapter 1.

Lassalle P., E. Adam, S. Sarrazin, F. Depontieu, V. Motte, C. Villedieu, S. Lequien, B. Gregoriu, N. de Freitas-Carreras, C. Vanpouille, M. Delehedde. (2006). "'Proteoglycans and glycosaminoglycans: key players in inflammation" In "New developments in Therapeutic Glycomics". Eds. Maryse Delehedde and Hugues Lortat-Jacob. Chapter 2.

Adam E, V. Motte, S. Sarrazin and M. Delehedde. (2006). "Stromal Derived Growth Factor-1: a target for GAG-based therapies?" In "New developments in Therapeutic Glycomics". Eds. Maryse Delehedde and Hugues Lortat-Jacob. pp 283-302. Chapter 12.

Deux articles sont également en cours de preparation :

Adam E., J-F. Minéo, S. Sarrazin, M. Debunne, V. Motte, C. Villedieu, R-M. Siminski, M. Baroncini, P. Lassalle, S. Blond, C-A. Maurage, and M. Delehedde. (Prêt). " Endocan expression in human glioblastoma cells: regulation by inflammatory and pro-angiogenic factors and localization in palisading area in tumor sections"

Adam E., S. Sarrazin, C. Landolfi, V. Motte, P. Lassalle and M. Delehedde. (Fin de rédaction) "Efficient long-term and high yielded production of a recombinant proteoglycan in eukaryotic HEK293 cells using a membrane-based bioreactor."

**ETUDE STRUCTURALE ET  
FONCTIONNELLE D'ENDOCAN**

**1. MISE EN PLACE D'UNE PLATEFORME  
DE PRODUCTION ET PURIFICATION  
D'ENDOCAN, DU CORE PROTEIQUE  
D'ENDOCAN ET DES CHAINES DS  
D'ENDOCAN**

## 1.1. PRODUCTION ET PURIFICATION D'ENDOCHAN

### 1.1.1. Matériel et méthodes

#### *Culture cellulaire de la lignée HEK293 surexprimant endocan ou le core protéique d'endocan*

Les lignées stables de cellules HEK293 sur-exprimant le gène d'endocan ou du mutant d'endocan (S137A) ont été obtenues de Philippe Lassalle (Bécharde et al., 2001). Ces cellules sont maintenues en culture adhérente en milieu DMEM contenant 10 % de sérum de veau fœtal (SVF, Cambrex), 2 mM de L-glutamine (Gibco Invitrogen), 100 U/ml de pénicilline (Gibco Invitrogen), et 100 µg/ml de streptomycine (Gibco Invitrogen). La quantité nécessaire de cellules HEK293 permettant la production d'endocan recombinant a été obtenue par amplification en condition adhérente. Toutes les cellules sont maintenues à 37°C en incubateur à 5 % de CO<sub>2</sub>, 95 % d'humidité.

#### *Production d'endocan recombinant en flasks*

La culture des cellules HEK293 pour la production d'endocan en flasks se fait en condition de semi-adhérence, dans du milieu SFMII qui est un milieu spécifique ne contenant pas de SVF, mais qui est supplémenté en glutamine (2 mM), en pénicilline (100 U/ml) et en streptomycine (100 µg/ml). Après amplification, 50 millions de cellules pour 100 mL de milieu de culture sontensemencées dans des flasks de 175 cm<sup>2</sup>. Après une période de quatre jours, 70 millions de cellules sont récupérées, rincées et réensemencées dans les flasks. Le milieu de culture contenant endocan est alors récupéré et congelé à -20°C. Les deux premières récoltes ainsi faites en milieu SFMII ne contiennent que très peu d'endocan.

#### *Production d'endocan recombinant en spinner*

Les conditions de culture sont les mêmes qu'en flasks à l'exception que chaque spinner contient un volume de 800 mL de milieu dans lequel est libéré l'endocan recombinant. La rotation appliquée dans le spinner doit être la plus douce possible afin que les cellules puissent pousser en amas sans sédimenter. Une rotation trop brusque ou trop lente n'est pas favorable.

#### *Production d'endocan recombinant en bioréacteur CL1000*

La culture en bioréacteur des cellules HEK293 a bénéficié de l'expertise d'Estelle Adam concernant ce type de culture, initialement développée chez Endotis Pharma pour la culture

des hybridomes et la production d'anticorps (Adam et al., en préparation). Pour une culture en bioréacteur, les cellules HEK293 surexprimant endocan doivent être préalablement adaptées au milieu SFMII dans les flasks. La concentration en cellules mortes ainsi que la concentration en endocan produit sont déterminées à chaque passage des cellules comme pour les cultures en flask ou spinner. Une fois les cellules adaptées au milieu SFMII et lorsqu'un nombre suffisant de cellules est atteint, elles sont passées dans le bioréacteur CL1000 (Integra Biosciences). Les cellules HEK293 cultivées en bioréacteur pour la production d'endocan sont toujours en suspension dans du milieu défini SFMII ne contenant pas de SVF (Gibco Invitrogen), mais supplémenté en glutamine (2 mM), pénicilline (100 U/ml) et streptomycine (100 µg/ml). Les cellules HEK293 en phase exponentielle de croissance sont donc lavées puis reprises dans 20 ml de milieu SFMII. Les cellules sont alors introduites dans le compartiment cellulaire du bioréacteur CL1000 à raison de  $400 \times 10^6$  cellules / CL1000. Le compartiment nutritif du bioréacteur CL1000, quant à lui, est rempli de milieu SFMII (800 mL). Le bioréacteur ainsi préparé est placé dans un incubateur à 5 % de CO<sub>2</sub>, 95 % d'humidité et 37°C. Après une semaine de culture, le compartiment nutritif est vidé et la totalité des cellules est prélevée. Après centrifugation à 1600 trs/min, 5 minutes et à 4°C, le surnageant contenant l'endocan recombinant est récolté et conservé à -80°C jusqu'à purification. Un aliquot est également conservé à -20°C pour le suivi de la production par dosage ELISA. Le culot cellulaire est remis en suspension en milieu SFMII, et la quantité nécessaire de cellules est réensemencée en milieu SFMII dans le compartiment cellulaire. Les récoltes suivantes ont ensuite lieu tous les 3 ou 4 jours, suivant toujours le même procédé. A chaque récolte, une numération cellulaire est effectuée à l'aide d'hématocytomètre (cellule de Thoma), avec exclusion des cellules mortes en utilisant la coloration au bleu trypan. La culture en CL1000 peut être ainsi maintenue pendant plusieurs semaines.

#### *Purification d'endocan*

Le pH du milieu de culture récolté est tout d'abord ajusté à pH 8 avant l'application sur une colonne de DEAE-Sephacel (Sigma). La colonne est lavée avec 6 volumes de tampon 0,2 M NaCl, 50 mM Tris, pH 8. L'endocan ainsi fixé est élué avec un tampon 1 M NaCl, 50 mM Tris, pH 8. La salinité des fractions récoltées est abaissée à 0,2 M par dialyse en utilisant une membrane Spectra Por7 (1000 MWCO). L'échantillon contenant l'endocan est alors appliqué sur une colonne d'affinité de type HiTrapNHS (Amersham) sur laquelle l'anticorps monoclonal EP204 avait préalablement été immobilisé à raison de 8 mg par mL de résine. La colonne d'immunoaffinité est rincée avec un tampon Tris 50 mM, NaCl 0,2 M, pH 8 puis

l'endocan est élué avec une solution de  $MgCl_2$  4 M. Les 6 premiers volumes sont récupérés puis dessalés sur une colonne HiPrep 26/10 Desalting Sephadex G25 ; 53 mL (Amersham).

#### *ELISA Endocan*

Le dosage ELISA Endocan est un ELISA sandwich. Un premier anticorps monoclonal anti-endocan, clone EP214 (2  $\mu$ g/ml, tampon carbonate/bicarbonate 0,1 M pH 9.6), est adsorbé en plaque 96 puits (Maxisorp, Nunc), pendant la nuit à 4°C. Après 3 lavages en tampon PBS contenant 0,05 % Tween 20, les plaques sont saturées pendant 1 heure en PBS 10 % SVF. Les échantillons à tester, les surnageants de culture, ou les sérums sont alors déposés, dilués ou non en tampon PBS 10 % SVF, pour obtenir un volume final de 100  $\mu$ l/puits. Une gamme étalon réalisée à partir d'endocan recombinant est également déposée dans les puits de la même manière (de 20 ng/ml à 0,3 ng/ml). Après deux heures d'incubation à température ambiante, et cinq lavages en tampon PBS contenant 0,05 % Tween 20, la présence d'endocan dans les puits est révélée par une première incubation avec un anticorps monoclonal anti-endocan, clone EP215 (2  $\mu$ g/ml en tampon PBS 10 % SVF) pendant une heure, puis avec un anticorps polyclonal anti-IgG1 de souris couplé à de la biotine (clone A85-1, BD Pharmingen) pendant une heure également. Une solution de streptavidine-peroxydase est ensuite ajoutée dans les puits, puis incubée pendant 20 minutes à l'obscurité. Chacune de ces trois incubations est suivie de cinq lavages en tampon PBS contenant 0,05 % Tween 20. Le complexe ainsi formé est alors détecté grâce à l'ajout du substrat de l'enzyme peroxydase, le substrat TMB (3,3',5,5'-Tetramethylbenzidine, Sigma), pendant 30 minutes à l'obscurité. Après neutralisation de la réaction colorimétrique par addition d'acide chlorhydrique 2N, l'absorbance est mesurée à 450 nm et 570 nm (comme filtre de référence) grâce au lecteur de microplaques Multiskan ascent (Thermolabsystems).

#### *Gel SDS-PAGE*

Pour visualiser l'endocan, mais surtout le core protéique d'endocan des gels SDS-PAGE de 12 % sont utilisés régulièrement (gel de concentration 4 %, et gel de séparation 12 %). Les échantillons sont préparés dans une solution de Tris/HCl 30 mM, SDS 1 %, glycérol 5 %, bleu de bromophénol 0,1 %, pH 6.8 et chauffés à 90°C pendant 5 minutes. La migration s'effectue à 15 mA dans le gel de concentration et à 30 mA dans le gel de séparation. Les protéines sont colorées au bleu de Coomassie (40 % méthanol, 7 % acide acétique, 0,25 % (w/v) Bleu Coomassie, 53 % H<sub>2</sub>O) le gel est ensuite décoloré (méthanol 10 %, acide acétique glacial 7 %, H<sub>2</sub>O 83 %).

### *Western Blot*

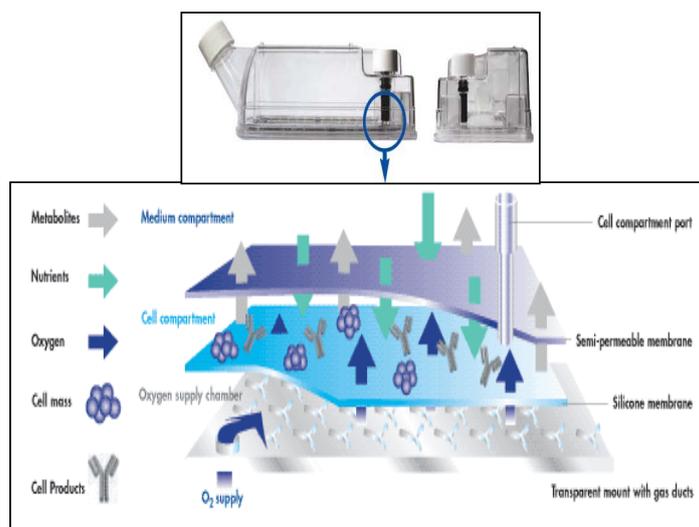
Afin de détecter de manière plus spécifique et sur de plus faibles quantités la présence d'endocan ou certains fragments d'endocan, le Western Blot peut être utilisé. Pour cela, les protéines d'un gel SDS-PAGE sont transférées sur une membrane de PVDF (Bio-Rad) (transfert 1 heure 100V 4°C). La membrane est saturée en protéines dans un tampon TBS (Tris / HCl 50 mM, NaCl 0,15 M, pH 7.4) contenant 5 % de lait en poudre, 1 heure à température ambiante. Un anticorps de souris monoclonal anti-endocan (2 µL à 2 mg/mL dans 20 mL) est ajouté sur la membrane dans du TBS/Tween/lait (TBS contenant 0,05 % de Tween 20 et 5 % de lait en poudre), pendant 1 heure. Les anticorps non fixés sont supprimés par trois rinçages de TBS contenant du Tween 20. L'anticorps polyclonal anti-souris couplé à la HRP (Amersham) est ajouté (1/10000°) dans du TBS Tween Lait, pendant 1 heure, le surplus est rincé trois fois avec du TBS Tween. On révèle ensuite la présence de la HRP, donc d'endocan, par chemiluminescence (Kit ECL Amersham Biosciences), et autoradiographie (Hyperfilm ECL, Amersham).

### *Dot Blot*

Dans certains cas, pour le suivi en routine d'endocan, pendant une chromatographie par exemple, le Dot Blot est une méthode plus rapide que le Western Blot. Les anticorps utilisés sont spécifiques d'endocan et ne créent pas de bruit de fond. En utilisant un appareil à dot-blot, l'échantillon est directement appliqué dans un volume de 200 µL de PBS sur une membrane de nitrocellulose. La procédure de révélation de la protéine sur la membrane de nitrocellulose est ensuite identique à celle utilisée dans le cas du Western Blot.

## **1.1.2. Résultats et discussion**

La production et la purification d'endocan recombinant ont été entreprises pour mener à bien plusieurs projets. Dans le cadre des travaux de thèse, l'endocan ainsi obtenu allait pouvoir être utilisé pour étudier les interactions avec d'autres molécules ou pour développer des essais visant à caractériser ses fonctions. Dans le même temps, l'endocan purifié pouvait constituer une source de chaînes de DS d'endocan, amenées à être caractérisées ou fractionnées afin d'obtenir une librairie oligosaccharidique de type DS. Dans le cadre du développement d'un kit de dosage d'endocan, le PG utilisé comme standard de référence était également nécessaire. De même, une production conséquente d'endocan était nécessaire pour



**Figure 1: Système de production en bioréacteur CELLine 1000**

Le système de production CL1000 (CELLine 1000) développé par Integra Biosciences a été conçu principalement pour la production d'anticorps monoclonaux. Le principe du bioréacteur est très simple. Le système se compose de deux compartiments distincts, séparés par une membrane semi-perméable (10 kDa). Les biomolécules produites sont ainsi retenues dans le compartiment inférieur (qui contient également la masse cellulaire). En revanche les métabolites produits par les cellules diffusent dans le compartiment supérieur, compartiment qui sert également de réserve nutritive (milieu de culture). L'apport en oxygène se fait par la surface inférieure qui est directement en contact avec les cellules pour une meilleure oxygénation des cellules vouées à croître de manière très dense.

	Production d'endocan en CL1000	Production d'endocan en flasks ou spinner	Optimisation liée à l'utilisation des CL1000
Méthode de production	CL1000	Suspension en flasks 175 cm <sup>2</sup> ou spinners	
Concentration estimée en Endocan	environ <b>200 µg/ml</b>	environ <b>5 µg/ml</b>	→ Concentration en endocan x 40
Volume de surnageant récolté / semaine	<b>40 ml / CL1000</b> (2 passages)	équivaldrait à 1,6 L de surnageant soit 2 passages de 8 flasks 175 cm <sup>2</sup>	→ Réduction importante du volume de récolte à purifier → Gain de place dans les incubateurs
Temps de manipulation / semaine	<b>20 à 30 min / CL1000</b> (pour 2 passages)	<b>1.5 heures minimum</b> (pour 2 passages)	→ Gain de temps → Réduction des risques de contamination
Consommables	Très faible	Flasks, Pipettes, Tubes à centrifuger, etc...	→ Économie énorme en consommables

**Figure 2: Tableau comparatif des différents modes de production d'endocan**

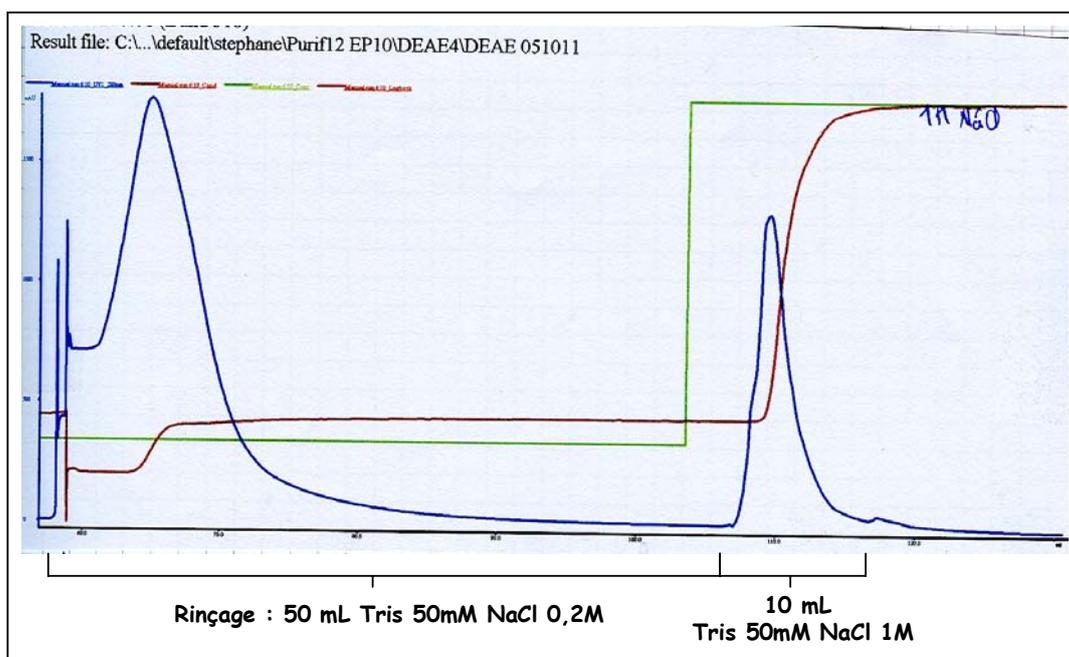
Dans le but d'optimisations, la production d'endocan a été mise au point en flasks, spinner, et bioréacteur CL1000. La production en protéines, les volumes à traiter, le temps de manipulation, l'utilisation de consommables ont été estimés et retranscrits dans ce tableau.

réaliser des essais *in vivo* chez l'animal très consommateurs de protéines recombinantes. Ainsi, en utilisant le protocole de production mis en place pour les anticorps et initialement développé chez Endotis Pharma par Estelle Adam (Adam et al. En préparation) et en adaptant la purification d'endocan mise en place par Bécharde et al. (2001), nous sommes passés d'une production de quelques centaines de microgrammes à plusieurs milligrammes d'endocan recombinant (Figure 1).

Pour chaque bioréacteur, les conditions optimales de production (ensemencement, temps entre chaque passage, viabilité) ont été évaluées (données non montrées). Si la quantité de cellules par millilitre de milieu consommé est uniquement prise en compte, les conditions d'ensemencement sont identiques pour les trois moyens de cultures. En revanche, pour les CL1000, si l'on prend en considération uniquement le compartiment cellulaire qui compte pour 20 mL alors que l'ensemble du bioréacteur utilise 800 mL, alors dans ce cas, l'ensemencement peut être considéré comme quarante fois plus important pour une CL1000 que pour un flask ou un spinner. Une comparaison sur la rentabilité de ces trois moyens de production a été faite (Figure 2). De façon identique à ce qui vient d'être expliqué, la production d'endocan apparaît quarante fois plus importante pour une CL1000 si seul le compartiment cellulaire est pris en compte. Cependant, rapporté au nombre de cellules totales et au volume de milieu consommé, le taux de production est identique. L'énorme avantage du bioréacteur CL1000 est lié au fait que les volumes à traiter (purifier) sont quarante fois moins importants. Cela est un avantage pour le temps de purification et par conséquent pour l'intégrité du PG. Concernant les avantages techniques, la CL1000 est également favorable à un gain de place dans les incubateurs, un gain de temps, une réduction des risques de contamination et une économie très importante sur du long terme (une CL1000 peut être utilisée pendant 6 mois). L'ensemble de ces travaux fait l'objet d'un manuscrit en préparation :

Adam E., S. Sarrazin, C. Landolfi, V. Motte, P. Lassalle and M. Delehedde. (Fin de rédaction)  
"Efficient long-term and high yielded production of a recombinant proteoglycan in eukaryotic HEK293 cells using a membrane-based bioreactor."

Concernant la purification, le protocole (Bécharde et al., 2001) a principalement été conservé. Il est constitué de deux principales étapes de chromatographie. Tout d'abord, une étape d'échange d'anions est réalisée. Figure 3A le chromatogramme d'une élution de colonne DEAE-sephacel est présenté. Pour avoir également travaillé avec des gradients de NaCl, l'élution de l'endocan débute à environ 0,4 M NaCl (données non montrées). L'élution par gradient n'a pas été conservée pour éviter une dilution trop importante de notre

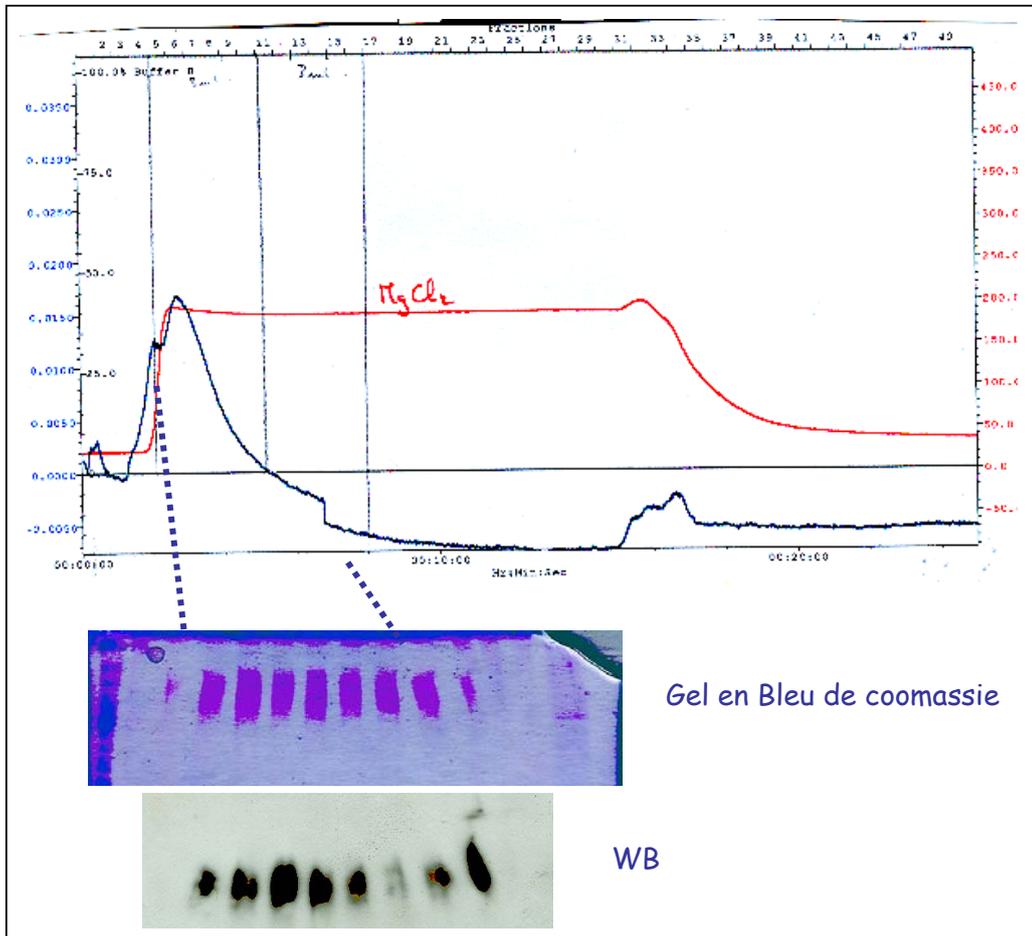


**Figure 3A: Première étape de la purification d'endocan par chromatographie échangeuse d'anions**

Sur une colonne C10/10 (5 mL de résine DEAE-Sephacel (Sigma), le milieu de culture récolté est ajusté à pH 8 et chargé à 0,25 mL/min. Le chromatogramme présenté, obtenu sur un Akta basic (Amersham), montre la phase de rinçage de la colonne (50 mL) puis l'éluion de l'endocan par l'ajout d'une solution à 1M NaCl. Le tracé bleu permet de suivre l'absorbance à 280 nm, le tracé rouge la conductivité et le tracé vert le pourcentage théorique du tampon B contenant 1M NaCl. Durant la phase de rinçage beaucoup de protéines ne sont pas retenues sur la colonne. Le gradient à 1 M NaCl permet d'éluer les protéines (dont l' endocan) qui étaient retenues sur la colonne.

échantillon. Selon les modes de production utilisés, la taille des colonnes a dû être adaptée pour que le chargement soit effectué en moins de 24 heures. Par exemple, pour la purification de dix récoltes, le volume à traiter pour les surnageants issus des CL1000 (180 mL) est différent du volume à traiter pour une culture en flask ou en spinner (8 Litres). Dans le premier cas, une colonne C10/10 (Amersham) avec 5 mL de résine suffit à avoir un débit de 0,25 mL/min, dans le second une colonne XK50/50 (avec 30 mL de résine) permettant un débit de 5 mL/min est plus adapté. Les volumes d'élution à traiter sont bien entendu à prendre en considération dans les étapes suivantes. Après abaissement de la salinité par dialyse, la seconde étape de purification est réalisée par immunoaffinité. Les colonnes utilisées pour l'ensemble des purifications ont été des HiTrapNHS (Amersham). La Figure 3B montre le profil du chromatogramme obtenu lors de l'élution d'endocan sur une colonne "HiTrapNHS-EP204" par une solution de MgCl<sub>2</sub> 4M. D'autres tampons plus couramment utilisés en laboratoire pour effectuer l'élution d'une colonne d'immunoaffinité ont été testés comme, par exemple, une solution de glycine 10 mM pH 2, mais aucune d'entre elles n'ont donné de résultats positifs. Afin de supprimer la solution "visqueuse" de MgCl<sub>2</sub>, la mise en place d'un protocole de désalage utilisant une colonne HiPrep 26/10 Desalting a donné des résultats plus intéressants que l'utilisation de la dialyse qui avait du mal à faire disparaître toutes les traces de MgCl<sub>2</sub> (élément très important pour l'utilisation de l'endocan *in vivo* puisque le Mg peut être toxique).

Dans un souci de dosage et du contrôle de la qualité des protéines recombinantes ainsi produites, plusieurs mesures ont été définies. Tout d'abord, concernant la quantification d'endocan, le test ELISA originellement développé par Philippe Lassalle et repris par nos laboratoires (Endotis Pharma et IBS), fonctionne très bien à condition de posséder un endocan étalon. Une autre méthode de dosage a pu être utilisée après purification d'endocan. C'est le dosage BCA (Pierce ou Sigma), avec toujours de l'endocan recombinant comme étalon et non de la BSA. Afin de définir la pureté et la concentration exacte d'un échantillon d'endocan purifié (pour définir un étalon par exemple), le meilleur moyen a été l'analyse protéique (analyse du ratio de chaque acide aminé dans un échantillon). Il a été remarqué que dans le cas d'endocan, le coefficient d'extinction molaire d'endocan variait énormément en fonction du tampon pour un même lot d'endocan. La chaîne de DS semble avoir une influence dans la détermination du coefficient d'extinction molaire. Ce moyen de dosage de la quantité d'endocan n'a donc pas été conservé. Concernant l'intégrité de l'endocan, un Western Blot a été réalisé sur chaque lot d'endocan purifié, avec l'anticorps EP221 capable de détecter endocan mais aussi les fragments peptidiques issus de la dégradation de ce PG.



**Figure 3B: La deuxième étape de la purification d'endocan est une chromatographie d'immunoaffinité**

Sur une colonne HiTrapNHS (amersham) sur laquelle l'anticorps monoclonal EP204 a préalablement été immobilisé, l'endocan récupéré par élution de la chromatographie d'échange d'ions est chargé en boucle (5 passages à 0,1 mL/min). Le chromatogramme présenté, obtenu sur un Akta basic (Amersham), montre la phase d'élution de la colonne avec une solution de 4 M MgCl<sub>2</sub>. Le tracé bleu permet de suivre l'absorbance à 280 nm, et le tracé rouge la conductivité. L'ajout de MgCl<sub>2</sub>, qui fait augmenter la conductivité, permet d'éluer l'endocan (tracé bleu). Un gel SDS-PAGE ainsi qu'un WB sont réalisés sur chacune des fractions de l'élution, cela permet de vérifier la présence d'endocan (WB) et permet d'avoir une idée sur la présence de potentielles protéines contaminantes (coloration au bleu de coomassie).

Grâce à l'intégration de ces différentes optimisations des systèmes de production cellulaire (augmentation de la biomasse et diminution des volumes), il a ainsi été possible de produire par semaine environ 7 à 8 mg d'endocan par CL1000. La purification de cette protéine recombinante dure ensuite globalement deux jours et le contrôle de l'intégrité et de la pureté (contrôle de la qualité) dure une autre journée. Les méthodes ainsi développées permettent d'obtenir avec un rendement de 50 %, environ 3 à 4 mg de protéoglycane avec un très bon degré de pureté, ce qui est très satisfaisant pour un système d'expression eucaryote.

## **1.2. PRODUCTION ET PURIFICATION DU CORE PROTEIQUE D'ENDOCAN**

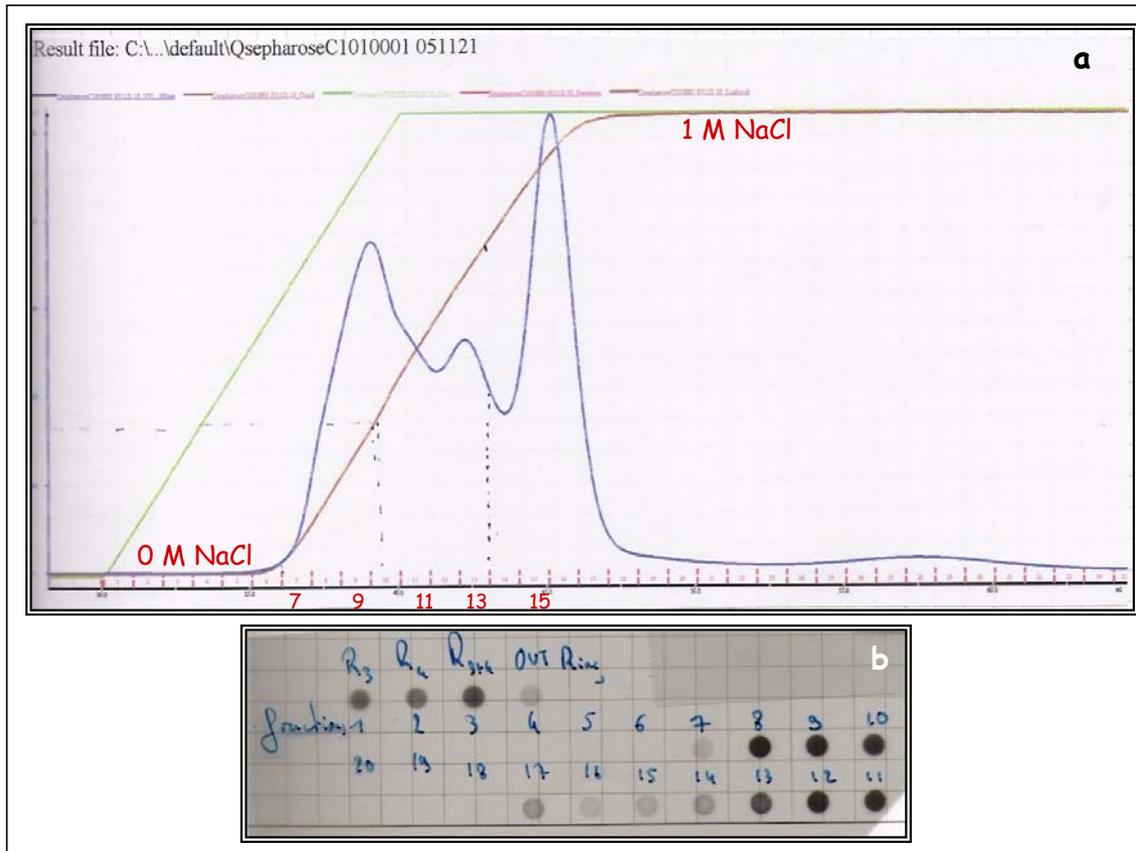
### **1.2.1. Matériel et méthodes**

#### *Production du core protéique d'endocan*

La production du mutant d'endocan (S137A) ne possédant pas la chaîne DS est réalisée en bioréacteur CL1000 dans les mêmes conditions que celles décrites pour l'endocan sauvage à l'exception que l'ensemencement optimal doit se faire avec 160 millions de cellules / CL1000.

#### *Purification du core protéique d'endocan*

Le pH du milieu de culture récolté est ajusté à pH 9 avant l'application sur une colonne de Q-Sepharose (Amersham) (5 mL de résine dans une colonne C10/10). La colonne est lavée avec 6 volumes de tampon 50 mM Tris, pH 9. L'endocan ainsi fixé est élué par un gradient de 0 à 1 M NaCl, 50 mM Tris, pH 9. Les fractions contenant le core protéique d'endocan (visualisation par SDS-PAGE ou WB) sont rassemblées. Le core protéique est alors appliqué sur une colonne d'affinité de type HiTrapNHS (Amersham) sur laquelle l'anticorps monoclonal EP236 avait préalablement été immobilisé à raison de 8 mg par mL de résine. La colonne d'immunoaffinité est rincée avec un tampon Tris 50 mM NaCl 0,2 M, pH 8 puis l'endocan est élué avec une solution de glycine 10 mM pH 2. Des fractions de 900 µL sont récupérées, et instantanément neutralisées par 100 µL de Tris 250 mM. Les fractions contenant le core protéique d'endocan sont rassemblées. Le changement de tampon et la concentration des échantillons se font par filtration à travers une membrane de porosité inférieure à 10 000 Da (Amicon Ultra Centrifugal Filter Devices 10000 MWCO de Millipore).



**Figure 4A: Première étape de la purification du core protéique d'endocan par chromatographie d'échange d'anions**

Sur une colonne C10/10 (5 mL de résine Q-Sepharose) (Amersham), le milieu de culture récolté est ajusté à pH9 et chargé à 0,25 mL/min.

a - Le chromatogramme présenté, obtenu sur un Akta basic (Amersham) montre l'éluion de l'endocan par un gradient de 0 à 1 M de NaCl. Le tracé bleu permet de suivre l'absorbance à 280 nm, le tracé rouge la conductivité et le tracé vert le gradient théorique de NaCl. Le gradient permet d'éluier les protéines (dont le core protéique d'endocan) selon trois "pics" principaux.

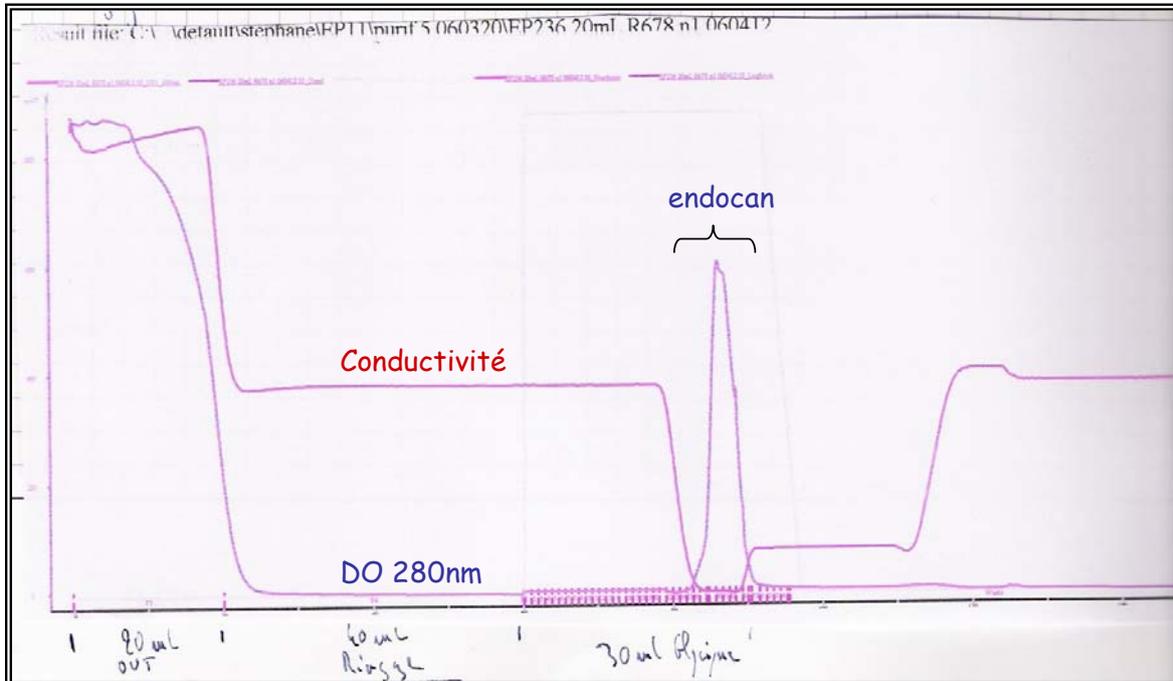
b - Le suivi des chromatographies par Dot-Blot permet de dire que la partie protéique d'endocan est éluée dans le premier pic. Sur le Dot Blot sont déposés : le milieu de culture chargé sur la colonne (R3 = récolte 3, R4 = récolte 4, R3+4 = mélange des deux récoltes qui a été chargé sur la colonne), les protéines non retenues sur la colonne = OUT, les protéines décrochées pendant le rinçage = Rinç, un échantillon de chacune des fractions = fractions 1 à 20.

### 1.2.2. Résultats et discussion

A l'image de l'endocan sauvage, la production et la purification du core protéique d'endocan ont été également mises en place pour mener à bien plusieurs projets. Dans le cadre des travaux de thèse, la partie protéique d'endocan ainsi obtenue devait pouvoir être utilisée pour étudier les interactions avec d'autres molécules, pour développer des essais visant à caractériser les fonctions d'endocan et également pour permettre d'entreprendre des travaux de cristallisation. Dans le cadre des projets d'Endotis Pharma, le core protéique d'endocan devait également être utilisé pour des essais *in vivo* chez la souris. Comme pour l'endocan, l'objectif était de mettre en place un système de production et purification permettant l'obtention de plusieurs milligrammes de protéines.

Le système d'expression utilisé reste le même que pour la production d'endocan sauvage, à la différence que sur l'ADNc d'endocan dans le vecteur d'expression, une mutation sur le codon de la sérine 137 a été introduite. La sérine 137 est alors remplacée par une alanine, ce qui ne permet plus l'ancrage de la chaîne de DS sur l'endocan (Bécharde et al., 2001). La lignée stable d'HEK293 transfectée avec le vecteur pcDNA3 contenant l'ADNc muté conçue par l'équipe de Philippe Lassalle (Institut Pasteur de Lille) a également été transférée au sein d'Endotis Pharma. La lignée exprimant le core protéique d'endocan permet, de la même façon que pour endocan, l'expression dans le milieu de culture. Les cellules sont pourtant moins adhérentes que les cellules HEK293 exprimant endocan. Bien qu'une production en flask soit possible, la culture en spinner n'a, en revanche, pas pu être adaptée. Pour des raisons évidentes d'efficacité, la production du core protéique d'endocan a été réalisée en CL1000. Les conditions optimales d'utilisation ont été définies et nécessitent un ensemencement de 160 millions de cellules par CL1000 avec un passage tous les trois jours et demi au cours duquel la viabilité est estimée à environ 50 %.

Concernant la purification du core protéique, plusieurs méthodes de chromatographie d'échange d'ions ont été tentées. En effet, des colonnes telles que la DEAE-sephacel et la SP-sepharose ont par exemple été testées avec différents tampons et à différents pH. Nos résultats ont montré que la Q-sépharose était la colonne qui donnait les meilleurs résultats à pH 9. La Figure 4A montre que le core protéique d'endocan sur une telle colonne est élué de 0 M NaCl à 0,5 M NaCl. Comme cela a été établi pour la purification d'endocan, la seconde étape de purification est une chromatographie d'immunoaffinité. Le core protéique d'endocan ne réagit pas de la même façon que l'endocan vis-à-vis des anticorps monoclonaux développés. Ainsi, en utilisant la résonance plasmonique de surface (SPR ; Biacore), des conditions de



**Figure 4B:** La deuxième étape de la purification du core protéique d'endocan est une chromatographie d'immunoaffinité

Sur une colonne HiTrapNHS (amersham) sur laquelle l'anticorps monoclonal EP236 a préalablement été immobilisé, l'endocan récupéré lors de la chromatographie d'échange d'ions est chargé en boucle (5 passages à 0,1 mL/min). Le chromatogramme présenté, obtenu sur un Akta basic (Amersham), montre la phase de fin de chargement (20 mL OUT), la phase de rinçage (40 mL Rinçage) et la phase d'éluion du core protéique d'endocan avec une solution de glycine 10 mM pH2. Le tracé bleu permet de suivre l'absorbance à 280 nm, le tracé rouge la conductivité.

purification via l'anticorps EP236 ont été étudiées, et plus particulièrement les conditions nécessaires à l'élution. Par comparaison avec l'endocan, le fait de ne pas devoir effectuer une élution avec une forte concentration de  $MgCl_2$  est réellement intéressant (volume d'élution, déssalage). Le chromatogramme d'immunoaffinité correspondant à l'utilisation de l'anticorps EP236 est présenté Figure 4B.

En appliquant le protocole mis en place, il est possible d'obtenir pour la production d'une CL1000 pendant une semaine, et après 1 jour et demi de purification environ 5 mg de core protéique d'endocan. Contrairement à ce qui avait été observé pour l'endocan, la concentration du core protéique d'endocan purifié peut être estimée grâce à la lecture d'absorbance à 280 nm. Le coefficient d'extinction molaire mesuré est de  $1,06 \text{ mg}^{-1} \cdot \text{mL} \cdot \text{cm}$ .

La production du core protéique d'endocan en utilisant un système bactérien avait été mise en place chez E.coli de type Origami (facilitant la formation des ponts disulfures). Les 11 cystéines des 110 premiers acides aminés suggèrent que plusieurs ponts disulfures prennent place pour structurer la protéine. Comme aucun test d'activité spécifique d'endocan n'est pour l'instant disponible, la renaturation correcte de la protéine apparaît difficile à appréhender et donc ce mode de production n'a pas été retenu pour l'instant.

### **1.3. PURIFICATION DE LA CHAÎNE DS D'ENDOCAN**

#### **1.3.1. Matériel et méthodes**

##### *Purification de la chaîne de dermatane sulfate d'endocan*

##### Par digestion enzymatique du core protéique

200  $\mu\text{L}$  d'endocan (200  $\mu\text{g}/\text{mL}$ ) sont mis en présence de 22  $\mu\text{L}$  de papain 10X (10  $\text{mg}/\text{mL}$  de papain dans un tampon acétate de sodium 1 M, EDTA 50 mM et Cystéine 50 mM pH 7) pendant 24 heures à  $65^\circ\text{C}$ . 24  $\mu\text{L}$  de papain 10X sont rajoutés, la réaction est incubée pendant 5 autres heures. Les protéines sont précipitées au TCA. Pour cela, 25  $\mu\text{L}$  de TCA 100 % sont ajoutés pendant 30 minutes, à  $4^\circ\text{C}$ , sous agitation. Après centrifugation, le surnageant est conservé (chaîne DS). Le culot est rincé avec 50  $\mu\text{L}$  de TCA 10 %, pendant 30 minutes, à  $4^\circ\text{C}$ , sous agitation. Après centrifugation, le surnageant est conservé et ajouté avec le premier surnageant pour être dialysé contre un tampon phosphate 20 mM pH 6. Les chaînes DS sont ensuite purifiées en utilisant 500  $\mu\text{L}$  de résine DEAE-Sephacel (Sigma) dans un biospin (Biorad). Pour cela, les chaînes oligosaccharidiques sont déposées sur la résine préalablement équilibrée. Après 30 minutes de mise en contact, la résine est rincée avec 10 volumes d'un

tampon Tris 50 mM, NaCl 0,15 M, puis les chaînes sont éluées avec une solution de Tris 50 mM, NaCl 1 M pour être ensuite dialysées contre de l'eau, puis finalement lyophilisées.

#### Par coupure chimique ( $\beta$ -élimination) de la liaison entre le core protéique et la chaîne DS

Un volume d'échantillon et un volume d'une solution de NaOH 100 mM, NaBH<sub>4</sub> 2 M ainsi qu'1  $\mu$ L de rouge de phénol (colorant indicateur de pH) et de l'acide acétique glacial (jusqu'à ce que le rouge de phénol vire au jaune) sont co-incubés dans un tube « haut » (en raison de la production de gaz), pendant 24 heures, à 37°C. La réaction est stoppée en ajoutant du NaOH concentré jusqu'à ce que le rouge de phénol vire au pourpre (pH 7). Les chaînes de DS sont alors purifiées de la même façon qu'au paragraphe précédent, en utilisant une résine DEAE-Sephacel.

#### *Gel d'acrylamide pour oligosaccharides (PAGE)*

Afin de visualiser les chaînes de DS d'endocan, après purification nous avons réalisé des électrophorèses en gels d'acrylamide pour oligosaccharides. Ils se composent d'un gel de concentration de 4 % (3,8 mL d'H<sub>2</sub>O ; 625  $\mu$ L Tris 1M pH 6,8 ; 830  $\mu$ L acrylamide (49 :1) 50 % (Sigma) ; 5  $\mu$ L TEMED ; 50  $\mu$ L persulfate d'ammonium 10 %) et d'un gel de séparation de 30 % (1,8 mL H<sub>2</sub>O, 3 mL Tris 1,5 M pH 8,8, 7,2 mL acrylamide (19 :1) 50 % (Sigma), 15  $\mu$ L TEMED, 18  $\mu$ L persulfate d'ammonium 10 %). Les échantillons sont déposés avec 20 % de glycérol. Un marqueur de migration contenant du bromophénol, du xylène cyanol et du rouge de phénol dans du glycérol 20 % peut être utilisé. Le tampon de migration (20 mA) utilisé est composé de Tris 25 mM / Glycine 200 mM. Le gel est ensuite coloré avec du bleu Azure A (0,08 % dans H<sub>2</sub>O) pendant 10 à 15 minutes, puis rincé à l'eau.

#### *Quantification des chaînes de DS par dosage colorimétrique des acides uroniques (dosage carbazole selon la méthode de Bitter et Muir, 1962)*

La quantité de DS est estimée par référence à une gamme de concentration de DS de source commerciale (50, 25, 12,5, 6,25, 0  $\mu$ g/mL). 150  $\mu$ L de tétraborate de sodium (0,025 M dans H<sub>2</sub>SO<sub>4</sub>) sont placés dans un tube en verre refroidi à 4°C. Lentement, 30  $\mu$ L d'échantillon (à doser ou de la gamme) sont déposés sur la surface. Les tubes sont fermés avec un bouchon en téflon, agités puis incubés pendant 10 minutes, à 100°C. Après refroidissement à température ambiante, 6  $\mu$ L de carbazole (0,125 % dans de l'éthanol) sont ajoutés. Les tubes sont alors à nouveau agités, puis incubés 15 minutes, à 100°C. Après refroidissement à température ambiante, l'absorbance est mesurée à 530 nm. Pour la gamme de concentration, le graphique de l'absorbance en fonction de la concentration est tracé et permet de définir la

courbe étalon. La concentration en acide uronique d'un échantillon est obtenue en se rapportant à la courbe étalon.

### **1.3.2. Résultats et discussion**

L'intérêt de savoir purifier la chaîne DS d'endocan était multiple. Pour la caractérisation de la chaîne, cela allait être bien évidemment nécessaire, autant à l'échelle du microgramme pour la caractérisation de la chaîne radiomarquée qu'à l'échelle du milligramme pour l'étude par RMN. La purification par  $\beta$ -élimination était adaptée à ce type d'étude car il était important de dissocier complètement la chaîne DS du core protéique afin de ne laisser aucune trace "protéique" pouvant interférer lors de l'analyse RMN. Pour évaluer la nature des interactions entre différents médiateurs et endocan, il était également très intéressant de posséder suffisamment de DS purifié afin d'être en mesure de dissocier la nature d'éventuelles interactions (entre le core protéique ou la chaîne DS). Pour l'étude des interactions avec le DS d'endocan, il était utile de posséder à l'extrémité de la chaîne, une amine qui allait pouvoir permettre l'immobilisation du DS sur une surface Biacore. Dans ce cas, la purification du DS par la papaïne était la méthode la plus favorable. Pour les médiateurs amenés à interagir avec le DS d'endocan, une banque oligosaccharidique serait en fait nécessaire afin de caractériser la structure (taille, sulfatation) du DS impliqué dans l'interaction. Finalement, toujours dans l'objectif d'analyser de façon dissociée les fonctions à la fois du core protéique et de la chaîne d'endocan, dans des essais visant à mettre en évidence les fonctions d'endocan, il était intéressant de posséder le DS d'endocan.

## **1.4. CONCLUSION**

Pour la mise en place de plateformes de purification d'endocan, de son core protéique et de sa chaîne de DS nous avons suivi des protocoles classiquement utilisés pour les PGs et pour endocan en particulier (Iozzo : proteoglycan protocols, 2001 ; Béchard et al., 2001). Cependant, concernant la production, l'adaptation des cellules HEK293 à la culture en bioréacteur CL1000 a été une optimisation remarquable. Ce genre de bioréacteur, encore peu utilisé, avait été initialement développé pour la production d'anticorps pour remplacer notamment la production en ascites. A notre connaissance, aucune donnée bibliographique n'a encore été référencée concernant ce procédé de production pour les cellules eucaryotes HEK293. Nos résultats montrent que ce système de culture cellulaire a été hautement

bénéfique pour augmenter le rendement de production de l'endocan recombinant, pour diminuer les volumes à traiter en purification et, consécutivement, pour maintenir l'intégrité de ce PG fortement sensible à la dégradation.

Pour résumer, la plateforme technologique mise en place a donc permis d'obtenir des molécules recombinantes à l'échelle du milligramme par semaine, à partir d'un système d'expression eucaryote, avec un très bon degré de pureté, que ce soit pour l'endocan, son core protéique ou sa chaîne de DS.

## **2. CARACTERISATION STRUCTURALE ET BIOCHIMIQUE D'ENDOCAN**

## 2.1. CARACTERISATION DE LA CHAÎNE DE DERMATANE SULFATE

### 2.1.1. Introduction

Comme mentionné dans le rapport bibliographique, l'endocan est un PG qui possède une seule chaîne de DS liée sur la sérine 137 de son core protéique (Bécharde et al., 2001a) initialement détecté spécifiquement dans les cellules endothéliales telles que les cellules HUVEC (Lassalle et al., 1996). Différents travaux ont montré depuis que son expression peut s'étendre aux adipocytes (Wellner et al., 2003), aux glioblastomes (Adam et al., en préparation), aux mélanomes et aux cellules rénales tumorales (Hendrix et al., 2003 ; Seftor et al., 2002 ; Ross et al., 2000 ; Abid et al., 2006).

Comme précédemment exposé, la synthèse et les modifications qui définissent la nature et la structure des GAGs sont exécutées par une large gamme d'enzymes. Ces enzymes sont elles-mêmes régulées par différents facteurs externes tels que des médiateurs (facteurs de croissance, cytokine, etc) mais également par les types cellulaires eux-mêmes (Viola et al., 2006). Autant dire que l'hétérogénéité des GAGs au sein d'une cellule, et donc encore plus d'un type cellulaire à l'autre, est dépendante d'une multitude de facteurs.

Le premier objectif de caractérisation d'endocan était de caractériser et de comparer la nature de la chaîne glycanique d'endocan produit naturellement par les cellules endothéliales HUVEC à celle issue d'endocan synthétisé par des cellules HEK293 génétiquement modifiées pour le produire de manière recombinante. En effet, l'endocan produit par le système artificiel d'expression utilisant les cellules HEK293 qui est une lignée cellulaire immortalisée issue de rein embryonnaire, pourrait être différent de l'endocan "naturel" des cellules endothéliales.

Pour effectuer la caractérisation de la chaîne d'endocan et, plus particulièrement, pour comparer celle issue de l'endocan de cellules HUVEC et de cellules HEK293, la méthode classique du radiomarquage des DS suivie de digestions enzymatiques spécifiques et de chromatographies a été utilisée. Cette méthode avait déjà été employée par Bécharde et coll. en 2001. Une partie de cette étude a été réalisée dans le laboratoire du Pr John Gallagher et grâce à la collaboration amicale du Dr Malcolm Lyon lors d'un séjour pour lequel j'ai bénéficié d'une bourse de voyage internationale de la part de l'ANRT. Dans un deuxième temps, afin d'apporter des résultats complémentaires à cette approche classique, des premières expériences de caractérisation par RMN ont été entreprises sur le DS d'endocan issu des surnageants de cellules HEK293. Ces premiers essais ont été réalisés par le Dr Marco Guerini après envoi de ces DS purifiés à l'institut Ronzoni (Milan, Italie).

## **2.1.2. Etude comparative de la composition oligosaccharidique de la chaîne d'endocan provenant de cellules HUVEC et de cellules HEK293**

### **2.1.2.1. Matériel et méthodes**

#### *Culture des deux lignées cellulaires*

La culture des cellules HEK293 s'effectue de la même manière que cela a été décrit page 121 (1.1.1 Matériel et méthodes). Lorsque les cellules HEK293 ont une croissance régulière en milieu SFMII, la glucosamine tritiée ( $^3\text{H}$ ) (Amersham) est ajoutée. Ainsi deux mCi diluées dans 250 mL de milieu SFMII sont répartis dans cinq flasks de 175 cm<sup>2</sup> contenant chacune au final 700 000 cellules / mL. Après quatre jours de culture en présence de glucosamine tritiée, le milieu est collecté et pourra être utilisé pour la purification.

Les cellules HUVECs (Human Umbilical Vein Endothelial Cells) (Tebu-Bio) sont cultivées dans du milieu Medium 200 (Tebu-Bio) supplémenté avec 10 % de Low Serum Growth Supplement (Tebu-Bio). La glucosamine tritiée ( $^3\text{H}$ ) est ajoutée directement dans le milieu de culture après le troisième passage de ces cellules non immortalisées. Ainsi trois mCi sont mélangées dans 400 mL de milieu qui est réparti dans 16 flasks de 175 cm<sup>2</sup> dont les cellules sont à confluence. Après quatre jours de culture en présence de glucosamine tritiée, le milieu est collecté et utilisé pour la purification.

#### *Quantification de l'endocan purifié et radiomarquage*

Un dosage protéique de type Micro BCA (Pierce) a été utilisé pour déterminer la quantité d'endocan purifié depuis les cellules HEK293 et les cellules HUVECs.

Tout au long de la purification, la radioactivité de chaque échantillon a été suivie par comptage de la scintillation. Pour cela, 10  $\mu\text{L}$  d'échantillon sont ajoutés à 200  $\mu\text{L}$  d'eau et 2 mL de tampon de scintillation (Beckman), correctement mélangés et lus dans un détecteur de scintillation.

#### *Analyse structurale de la chaîne de DS d'endocan issue des cellules HEK293 et des cellules HUVECs*

##### *Nature de la chaîne*

Pour déterminer la nature de la chaîne oligosaccharidique d'endocan, l'endocan (10 000 cpm) a été digéré avec la chondroïtinase ABC (0,5 IU/mL dans un tampon de

digestion Tris 20 mM, pH 8.0 ajusté avec l'acide acétique, à 37°C, pendant 2 heures). L'endocan digéré avec la ChABC a ensuite été déposé sur une colonne de gel filtration de type CL6B (20 mL,  $V_0 = 6$  mL,  $V_t = 21$  mL) équilibrée en PBS. La radioactivité de chaque fraction est ensuite mesurée par comptage de la scintillation.

#### Taille de la chaîne

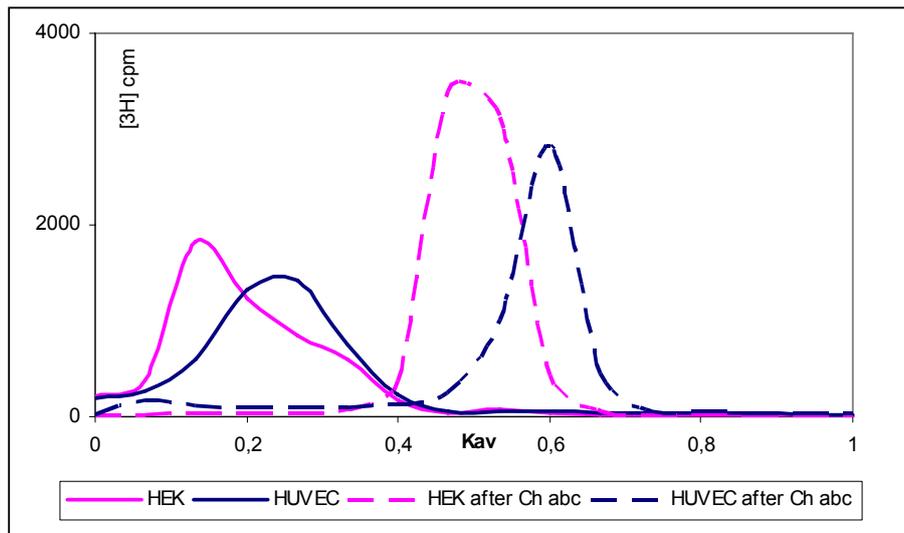
Pour comparer la longueur des chaînes d'endocan de cellules HUVECs et celles des cellules HEK293, les chaînes purifiées d'endocan (10 000 cpm) ont été appliquées sur une colonne CL-6B dans un tampon PBS. La colonne est étalonnée selon la courbe de calibration de CL-6B comme définie par Wasteson et le débit est fixé à 8 mL/h.

#### Analyse disaccharidique

40 000 cpm de DS radiomarqué d'endocan sont digérés totalement jusqu'à l'obtention de disaccharides grâce à l'addition de 100 mIU de chondroïtinase ABC (Sigma) dans un tampon de Tris 20 mM, pH 8.0 (acide acétique) à 37 °C toute la nuit. Le pH de la digestion (5 000 cpm) est corrigé avec de l'eau distillée/HCl (pH 3.5). Les échantillons (1 mL) sont chargés sur une colonne HPLC échangeuse forte d'anions de type Hypersil (5  $\mu$ m ; 0.46 x 25 cm; ThermoQuest ; Runcorn UK) préalablement équilibrée en eau (pH 3.5). Après un lavage à l'eau (pH 3.5), les disaccharides sont élués par un gradient linéaire de 0 à 0,3 M NaCl, pH 3.5 à un débit de 1 ml/min. Des fractions de 0.5 ml sont collectées et pour chacune d'entre elles, le tritium contenu est évalué par scintillation. Les différents disaccharides élués sont identifiés par comparaison avec les profils d'éluion de standards disaccharidiques de CS/DS dont la structure est connue, et qui sont détectés par une lecture de l'absorbance à 232 nm. Les analyses ont été réalisées sur deux digestions enzymatiques différentes.

#### Analyse du ratio acide glucuronique / acide iduronique

50 000 cpm de chaîne glycanique d'endocan radiomarquée sont digérés par la chondroïtinase ACI (100 mIU / mL dans un tampon 75 mM NaCl, 20 mM phosphate, pH 7.2) ou la chondroïtinase B (100 mIU/mL dans un tampon 20 mM Tris, pH 8.0 (acide acétique)). La digestion est initiée par l'ajout de la moitié de l'enzyme pendant 16 heures à 37°C puis le reste de l'enzyme est ajouté pour une nouvelle incubation de 2 heures. Les chaînes traitées aux chondroïtinases ACI ou B sont ensuite déposées sur une colonne de gel filtration de type Bio-Gel P10 (1x170 cm), équilibrée en PBS, à un débit de 4 ml/h. Des fractions de 1 mL sont collectées et leur radioactivité mesurée. Le  $V_0$  et le  $V_t$  de la colonne ont préalablement été



**Figure 5: Nature de la chaîne glycanique d'endocan**

Tracé rapportant la scintillation détectée à la sortie d'une colonne de filtration sur gel de type CL-6B lorsque de l'endocan radiomarqué issu de cellules HEK293 (rose) ou HUVECs (bleu), et traité (pointillés) ou non (trait plein) à la chondroitinase ABC a été déposé. Sur l'axe des abscisses, le 0 correspond au  $V_0$  (6 mL) et 1 correspond au  $V_t$  (21 mL) de la colonne.

déterminés avec, respectivement, de l'immunoglobuline et du dichromate de sodium. Le pourcentage de liaisons spécifiquement dégradées par les enzymes est calculé en utilisant un algorithme standard qui rapporte la distribution de radioactivité détectée dans chaque pic d'oligosaccharide de taille identifiée, à la radioactivité totale.

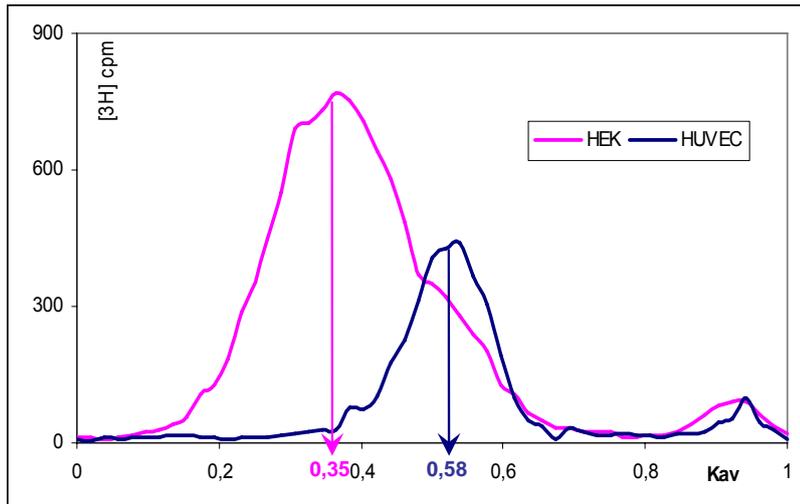
#### *Localisation des IdoA le long de la chaîne d'endocan*

10 000 cpm de chaîne glycanique d'endocan radiomarquée sont digérés par la chondroïtinase B (100 mIU/mL dans un tampon 20 mM Tris, pH 8.0 (acide acétique)). La digestion est initiée par l'ajout de la moitié de l'enzyme, pendant 16 heures, à 37°C, puis le reste de l'enzyme est ajouté pour une nouvelle incubation de 2 heures. Les chaînes traitées à la chondroïtinase B sont ensuite déposées sur une colonne de gel filtration de type CL-6B équilibrée en PBS, à un débit de 8 ml/h et étalonnée selon la courbe de calibration CL-6B définie par Wasteson (1971). Des fractions de 1 mL sont collectées et leur radioactivité mesurée.

### **2.1.2.2. Résultats**

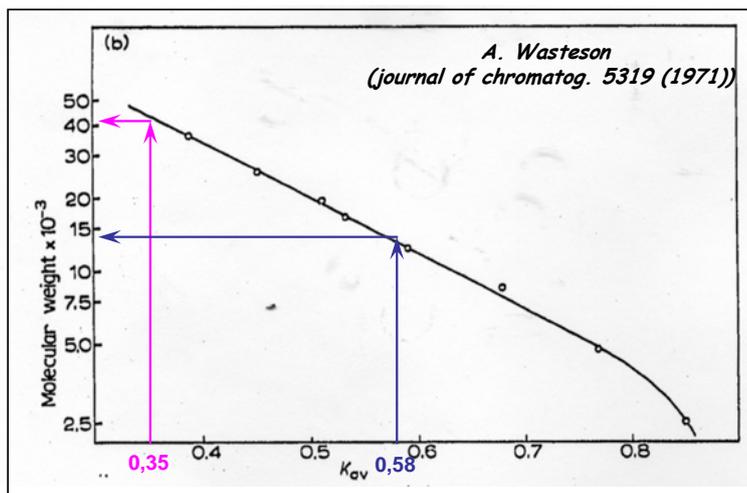
#### Nature de la chaîne de GAG d'endocan issu de cellules HEK293 et de cellules HUVECs

Afin de vérifier la nature de l'unique chaîne oligosaccharidique liée à endocan sur la sérine 137, la chondroïtinase ABC (enzyme dégradant spécifiquement les CS et DS) est utilisée. Sur une courte colonne de chromatographie d'exclusion de type CL-6B (20 mL), de l'endocan radiomarqué entier ou ayant été préalablement traité à la chondroïtinase ABC est déposé. Les résultats sont présentés dans la Figure 5. Il apparaît clairement que le traitement à la chondroïtinase ABC (courbe pointillée) dégrade la chaîne oligosaccharidique d'endocan, et cela pour l'endocan de cellules HEK293 (rose) ou l'endocan naturellement exprimé par les cellules HUVECs (bleu). Cette dégradation est visualisée par une migration retardée (création de fragments plus courts) de l'échantillon lorsqu'il est traité à la Ch ABC. Cependant, on peut remarquer que les pics d'élution des échantillons traités à la Ch ABC n'ont pas le même coefficient d'élution. Cela laisse penser que dans le cas de l'endocan issu des cellules HEK293 la dégradation n'est pas complète. La spécificité de l'enzyme utilisée, vis-à-vis des chaînes de DS et CS, permet ainsi de certifier que l'endocan produit naturellement par les cellules HUVECs ou artificiellement par les cellules HEK293 est composé de chaînes oligosaccharidiques de type CS/DS.



**Figure 6a: Taille de la chaîne glycanique d'endocan**

Tracé rapportant la scintillation détectée à la sortie d'une colonne de filtration sur gel de type CL-6B (calibrée selon les critères de Wasteson (1971)) lorsque des chaînes de CS/DS d'endocan radiomarqué issues respectivement de cellules HEK293 (rose) et HUVECs (bleu), ont été déposées. Sur l'axe des abscisses le 0 correspond au  $V_0$  (32 mL) et 1 correspond au  $V_t$  (83 mL). Les valeurs moyennes (0,35 et 0,58) des temps d'éluion pour les chaînes d'endocan d'HEK293 ou d'HUVECs respectivement, ont été indiqués sur l'axe des abscisses.



**Figure 6b: Courbe de Wasteson**

Courbe de calibration de la colonne CL-6B utilisée pour définir la taille des chaînes de CS/DS d'endocan. Cette courbe de référence a été initiée par A. Wasteson. Les tailles moyennes des GAGs d'endocan d'HUVEC ou HEK293 ont été reportées sur l'axe des abscisses.

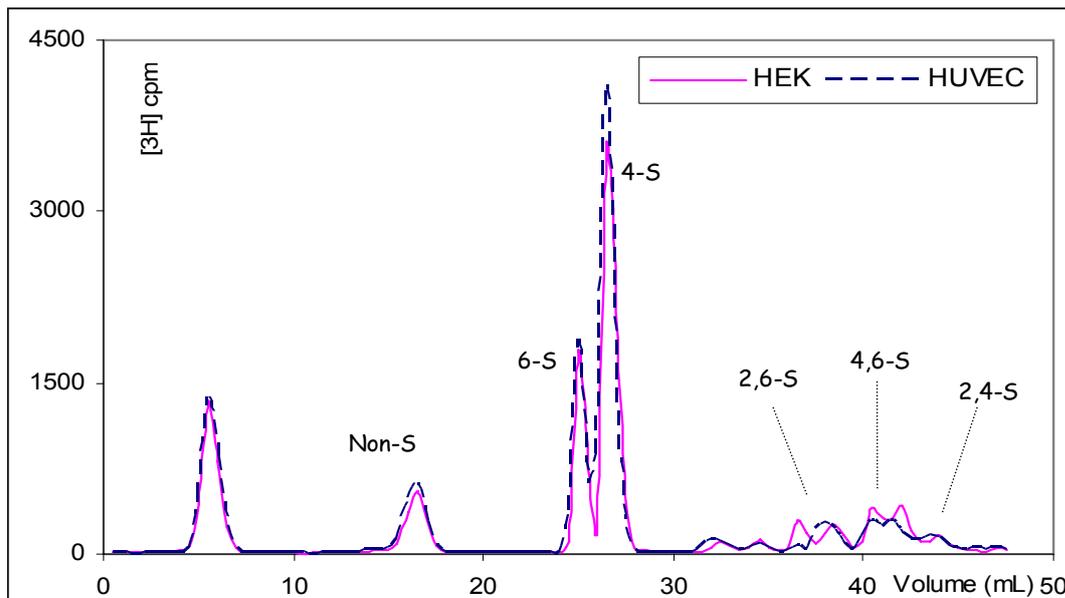
### Taille de la chaîne CS/DS d'endocan issue de cellules HEK293 et de cellules HUVEC

La chaîne oligosaccharidique d'endocan est séparée du core protéique par  $\beta$ -élimination. L'estimation de sa taille a pu être réalisée en utilisant une colonne d'exclusion stérique de type CL-6B préalablement calibrée. Ainsi, la courbe standard de migration de CS/DS en résine CL-6B, définie par Wasteson (1971) (Figure 6b), est utilisée. Elle permet de définir, pour un volume d'élution donné, un poids moléculaire théorique de la chaîne. Comme le montre la Figure 6a, la taille de la chaîne d'endocan est très hétérogène. En effet, pour la chaîne issue des cultures d'HEK293, les chaînes glycaniques sont détectées depuis le facteur d'élution 0,2 jusqu'à 0,6. La chaîne issue des cultures d'HUVECs, en revanche, est éluee entre les facteurs d'élution 0,4 et 0,6 ce qui révèle sans doute une plus grande homogénéité. Une valeur moyenne du poids moléculaire de la chaîne CS/DS d'endocan pour chaque lignée cellulaire a été définie. Dans le cas des HEK293, le coefficient d'élution moyen de la chaîne d'endocan est de 0,35 alors que, pour les cellules HUVECs, il est de 0,58 (Figure 6a). Lorsque ces valeurs sont reportées sur la courbe standard des migrations (Figure 6b), la taille moyenne de la chaîne glycanique d'endocan produite par les cellules HEK293 est de 40 kDa, alors qu'elle est simplement de 15 kDa sur de l'endocan produit naturellement par les cellules HUVECs. Sur la Figure 6a, plus précisément sur le profil d'élution de la chaîne d'endocan issue d'HEK293, il est intéressant de remarquer qu'un épaulement se distingue du large pic d'élution majoritaire, et surtout que celui-ci se trouve à l'endroit même où la chaîne d'endocan des cellules HUVECs est éluee.

Sur les deux tracés de cette même figure, un léger pic apparaît aux environs du coefficient d'élution 0,9. Le poids moléculaire du contenu des fractions, si toutefois cela était du CS/DS, serait inférieur à 2500 Da. Sa nature reste encore inconnue. Il apparaît cependant clairement qu'il peut être séparé des chaînes CS/DS proprement dites par une simple colonne DEAE échangeuse d'anion sur laquelle une concentration de 0,25 M NaCl ne lui permet pas d'interagir avec la phase solide contrairement à la chaîne d'endocan (données non montrées).

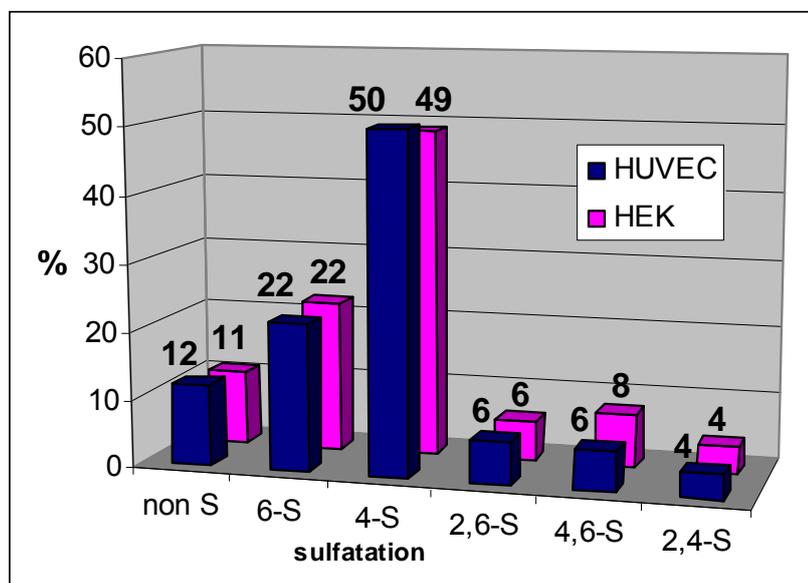
### Analyse disaccharidique de la chaîne CS/DS d'endocan

Le séquençage des GAGs n'a pas encore été rendu possible car leur hétérogénéité rend la tâche vraiment difficile. Cependant, depuis longtemps, il est possible de définir le profil des unités disaccharidiques qui composent les chaînes de GAGs, c'est-à-dire connaître dans la globalité, mais non de manière séquentielle, la proportion des différentes structures d'unités disaccharidiques, variable selon les modifications (sulfatations) qui composent une chaîne de GAG. Ce travail a donc été entrepris sur le DS d'endocan issu de cellules HEK293 et



**Figure 7a: Analyse disaccharidique de la chaîne glycanique d'endocan**

Tracé rapportant la scintillation détectée à la sortie d'une colonne HPLC d'échange d'ion forte de type hypersil. 5000 cpm d'échantillon (disaccharides de CS/DS, d'endocan d'HUVEC (bleu) ou d'HEK293 (rose)) dans 1 mL ont été chargés, rincés, puis élués par un gradient linéaire de NaCl de 0 à 0,3 M (2 à 47 mL). Les différents disaccharides ont pu être identifiés par une chromatographie identique pour laquelle des standards ont été utilisés et dont l'éluion pouvait être suivie à 232nm.



**Figure 7b: Histogramme rapportant l'analyse disaccharidique de la chaîne glycanique d'endocan**

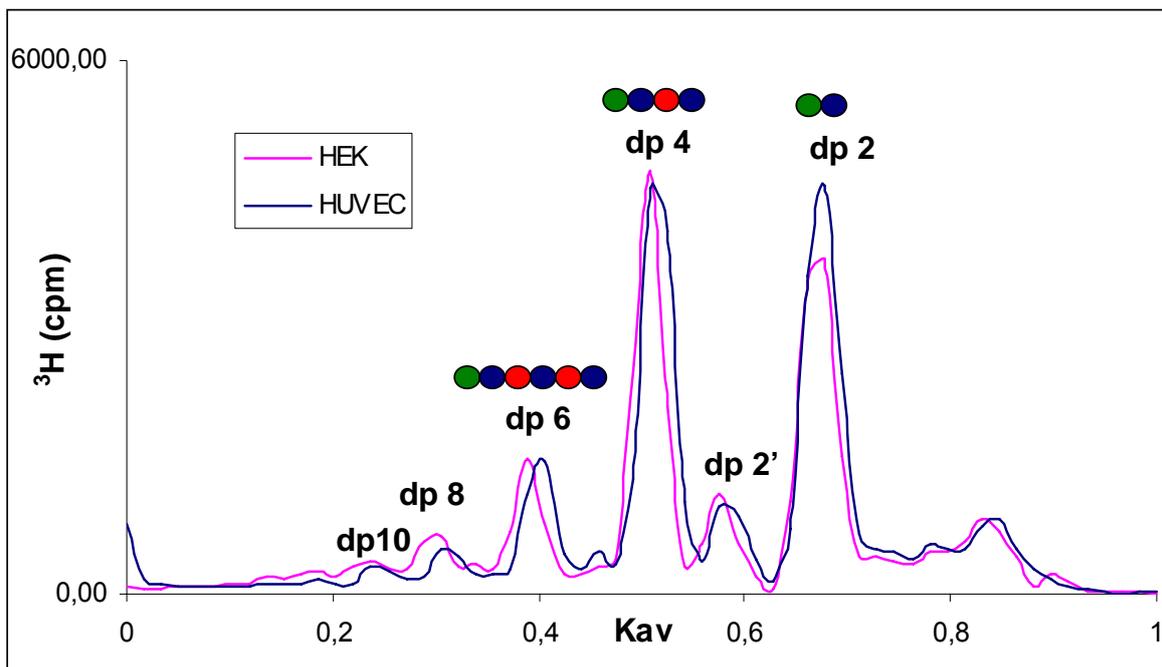
Retranscription (en pourcentage de la scintillation totale) des tracés obtenus par HPLC pour l'analyse disaccharidique (figure 3b). Pour chaque disaccharide (non-sulfaté, 6-S; 4-S; 2,6-S; 4,6-S; 2,4-S) la proportion qu'il représente est indiquée en haut de chaque histogramme pour les disaccharides venant d'HUVEC (bleu) ou d'HEK (rose).

HUVECs. Pour cela, la chaîne d'endocan a été réduite en unités disaccharidiques de manière enzymatique grâce à la chondroïtinase ABC. Ces unités disaccharidiques ont ensuite été séparées par chromatographie d'échange d'ions, selon leur structure ou, plus exactement, selon leur charge et donc leur degré de sulfatation. La Figure 7a représente le profil d'élution des différents disaccharides. Chacun d'entre eux a pu être identifié par une calibration préalable de la colonne avec des disaccharides standards de "structure" connue. Ainsi, il apparaît très clairement que les disaccharides majoritaires sont ceux qui possèdent des GalNAc(4-OSO<sub>3</sub>) ou GalNAc(6-OSO<sub>3</sub>). La proportion de chacun des pics de chromatographie a été calculée (Figure 7b). L'information principale de cette analyse disaccharidique provient de la comparaison entre les deux sources de cellules. En effet, le profil de sulfatation des disaccharides obtenus pour les deux lignées est, d'une manière surprenante, quasiment identique. Ensuite, d'un point de vue de la composition proprement dite, les disaccharides 4-sulfatés représentent plus de la moitié de la composition de la chaîne (49+8+4 % pour les cellules HEK293 Figure 7b), et les disaccharides 6-sulfatés comptent pour environ un tiers de la chaîne (22+6+8 % pour les cellules HEK293). Les 2-O-sulfatations sur les acides uroniques sont présentes sur environ 10 % de la chaîne. Parmi les unités 2-O-sulfatées, 6 % comportent également une 6-O-sulfatation, et 4 % une 4-O-sulfatation. Par ailleurs, environ 10 % de la chaîne glycanique d'endocan n'a pas subi de modification.

#### Proportion d'acide glucuronique versus acide iduronique

Les chaînes de DS sont des chaînes de CS qui possèdent au moins une épimérisation sur le carbone 5 d'un de leurs acides glucuroniques pour donner un acide iduronique. Pour connaître la proportion globale d'acides iduroniques de la chaîne d'endocan, les enzymes chondroïtinases ACI et B ont été utilisées. L'enzyme Ch ACI digère toutes les liaisons de type GalNAc–GlcA alors que l'enzyme Ch B préférera les liaisons GalNAc–IdoA. Les fragments obtenus pour chaque digestion ont été séparés par chromatographie de filtration sur gel. En fonction de la taille des fragments obtenus, le taux de GlcA ou IdoA a pu être évalué. En effet, chaque liaison résistante à l'enzyme (par exemple un tétrasaccharide correspond à une liaison non clivée) correspond à la présence d'un IdoA dans le cas de la Ch ACI ou à d'un GlcA dans le cas de la Ch B.

La Figure 8a représente le profil d'élution du DS d'endocan digéré par la Ch ACI à la sortie de la colonne de filtration sur gel. Comme pour l'analyse disaccharidique, la première observation est la grande similitude des profils d'élution obtenus pour les deux lignées cellulaires. La proportion de chaque fragment est présentée dans le tableau de la Figure 8b.



**Figure 8a: Calcul du ratio CS/DS par une digestion à la ChACI**

Tracé rapportant la scintillation détectée à la sortie de la colonne de filtration sur gel de type P10. 50000 cpm d'échantillon (chaîne d'endocan d'HEK293 (rose) ou d'HUVEC (bleu) digérée avec la ChACI) ont été déposés et donc élués en fonction de leur taille à un débit de 4 mL/h. Les différents saccharides détectés ont pour les plus abondants été représentés. (Cercle bleu : GalNAc; vert : GlcA; rouge : IdoA)

	dp2	dp2'	dp4	dp6	dp8	dp10	dp12	dp14	CS	DS
HEK	34%	8,68%	32,64%	12,22%	5,75%	3,58%	2,07%	0,97%	65,80%	34,20%
HUVEC	37,94%	7,79%	36%	11,25%	3,86%	2,33%	0,52%	0,31%	69%	31%

**Figure 8b: Tableau récapitulatif des résultats de la figure 8a**

Tableau qui retranscrit (en pourcentage de la scintillation totale) les profils d'éluion obtenus par la filtration sur gel de la figure 8a. La proportion des fragments est indiquée pour chaque lignée cellulaire. Dans les deux dernières colonnes est indiquée la proportion que représente le nombre d'IdoA (DS) ou de GlcA (CS).

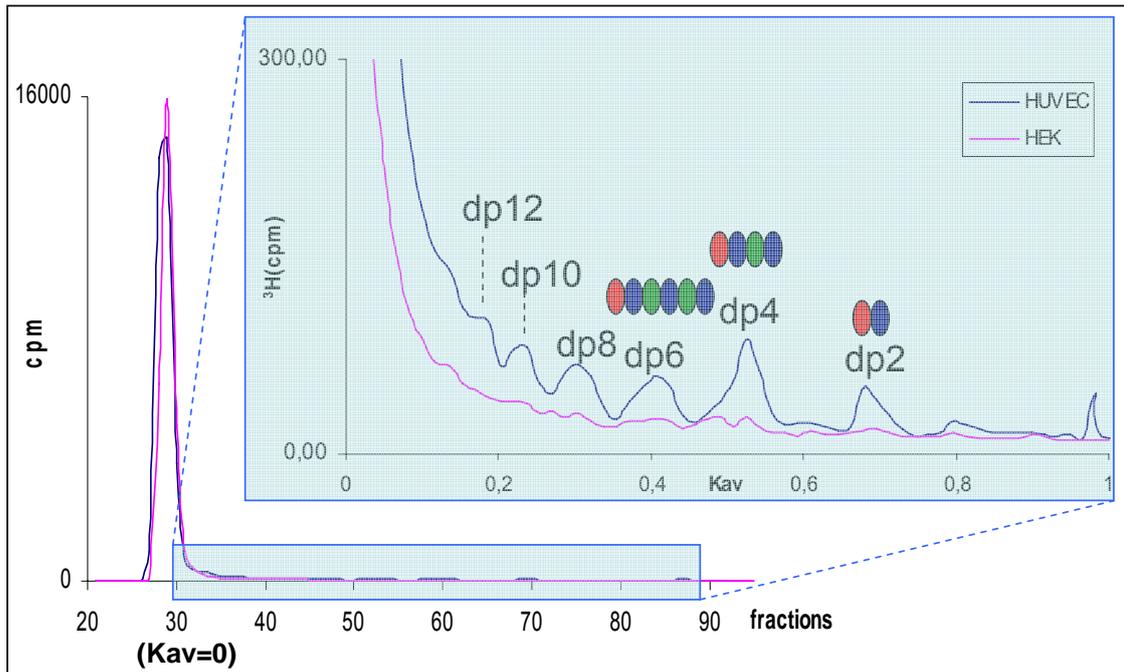
Comme il apparaît sur la Figure 8a, les dp2 et dp4 sont les espèces majoritaires. Les dp6 ou saccharides plus longs, quant à eux, sont les fragments qui traduisent une succession d'acides iduroniques sur une séquence de la chaîne. Ils représentent environ 18 % pour les cellules HEK293 et 13 % pour les cellules HUVEC, la proportion totale d'IdoA étant de 34 % pour le DS d'endocan des cellules HEK293 et de 31 % pour celui des cellules HUVEC. Cela signifie qu'environ la moitié des IdoA est disposé de manière regroupée et que l'autre moitié est un peu plus répartie sur la chaîne.

La digestion de la chaîne DS d'endocan par la Ch B n'a pas permis d'obtenir des fragments qui auraient pu confirmer que 30 % de la chaîne d'endocan est composée de DS. En effet, comme le montre la Figure 9, très peu de fragments inférieurs à 12 disaccharides ont pu être détectés. La fonctionnalité de l'enzyme a bien entendu été testée sur un DS témoin au préalable. Néanmoins, il faut noter que la chondroïtinase B est connue pour être une enzyme qui nécessite certaines structures particulières, telles qu'un regroupement des IdoA et leur 2-O-sulfatation, pour fonctionner. Ces exigences de l'enzyme nous fournissent donc des informations sur l'organisation structurale de la chaîne du GAG d'endocan comme cela est présenté dans la Figure 14.

#### Localisation des IdoA le long de la chaîne glycanique d'endocan

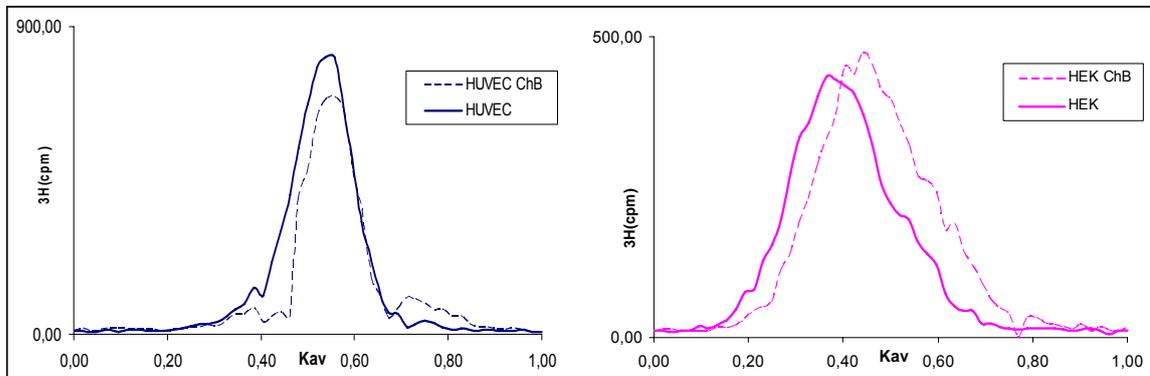
La filtration sur gel de type P10 a permis de séparer les fragments les plus courts issus des digestions par les enzymes Ch ACI et B. Pour observer les fragments les plus longs ou non digérés après traitement à la Ch B, une filtration sur gel de type CL-6B a été réalisée. Ainsi, la comparaison entre un échantillon traité ou non à la Ch B a permis de se faire une idée de l'emplacement des IdoA le long de la chaîne.

Sur la Figure 10, les chromatogrammes de ces chromatographies d'exclusion ont été représentés. Dans le cas de la chaîne glycanique d'endocan issue des cellules HUVEC, le traitement à la Ch B ne change guère le profil d'élution vis-à-vis de la chaîne non traitée. Une légère digestion des chaînes les plus longues peut cependant être constatée. En contrepartie, cela se traduit par l'apparition de fragments aux alentours du facteur d'élution 0,7 - 0,8. Dans le cas de la chaîne glycanique d'endocan issue d'HEK293, le profil d'élution est différent. Lorsque la chaîne est traitée à la Ch B, le large pic d'élution est déplacé dans sa totalité d'un facteur d'élution d'environ 0,1. En se reportant à la courbe standard de Wasteson (Figure 6b), un déplacement du facteur d'élution de 0,1 pour une chaîne de 40 kDa correspond à une diminution de sa taille de 15 kDa.



**Figure 9: Calcul du ratio CS/DS par une digestion à la ChB**

Tracé rapportant la scintillation détectée à la sortie de la colonne de filtration sur gel de type P10. 50 000 cpm d'échantillon (chaîne d'endocan d'HEK293 (rose) ou d'HUVEC (bleu) digérée avec la ChB) ont été déposés et donc élués en fonction de leur taille à un débit de 4 mL/h. Le profil d'éluion dans le cadre bleu est un zoom. Les différents saccharides détectés ont pour les plus abondants été représentés. (Cercle bleu : GalNAc; vert : GlcA; rouge : IdoA)



**Figure 10: Localisation des IdoA le long de la chaîne glycanique d'endocan**

Tracés rapportant la scintillation détectée à la sortie d'une colonne de filtration sur gel de type CL-6B (calibrée selon les critères de Wasteson (1971)) lorsque des chaînes de CS/DS d'endocan radiomarquées issues de cellules HEK293 (rose) ou HUVECs (bleu), préalablement digérées (pointillés) ou non (trait plein) à la ChB, ont été déposées. Sur l'axe des abscisses, le 0 correspond au  $V_0$  (32 mL) et 1 correspond au  $V_t$  (83 mL).

### 2.1.3. Analyse de la chaîne oligosaccharidique d'endocan par RMN

#### 2.1.3.1. Matériel et méthodes

##### *Purification de la chaîne de DS*

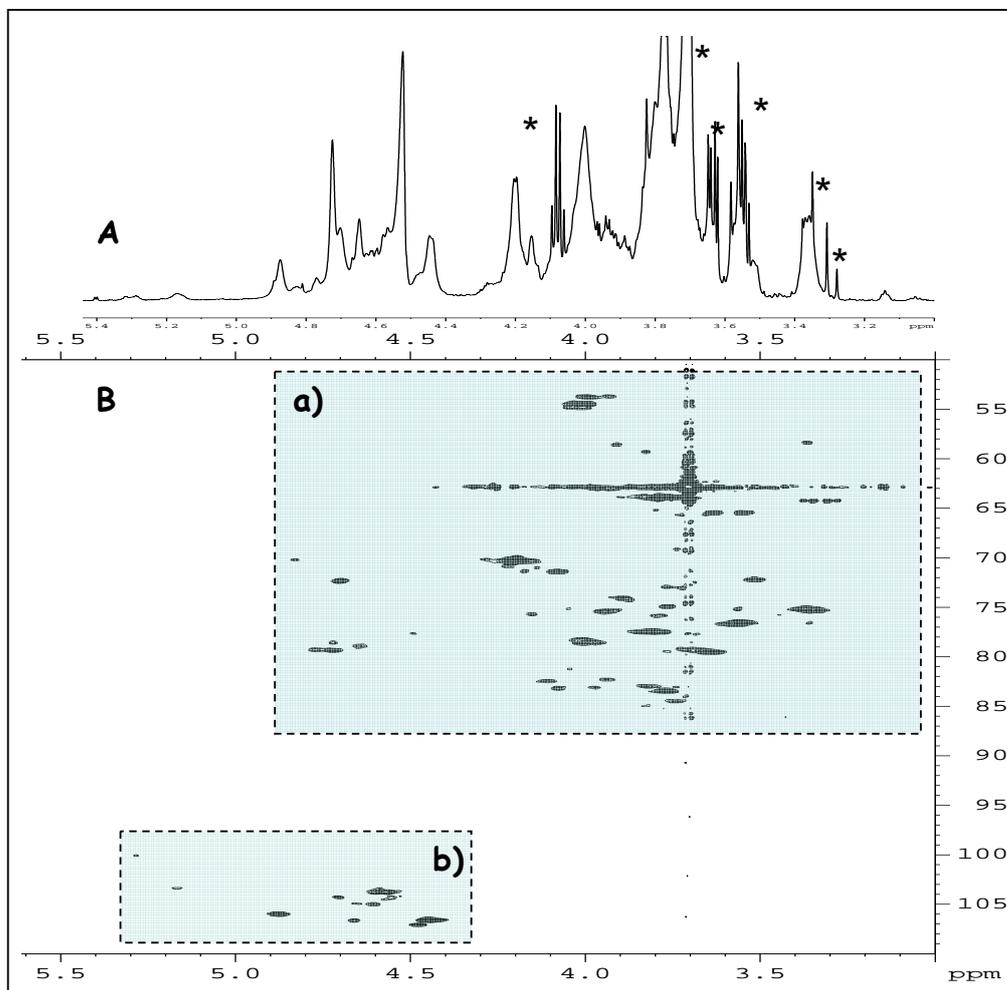
La chaîne de DS d'endocan a été purifiée par  $\beta$ -élimination selon le protocole décrit dans la section "*matériel et méthodes*" de la partie 1.3. *Purification de la chaîne DS d'endocan.*

##### *Préparation des échantillons pour la RMN*

Les spectres RMN mono- et bi-dimensionnels ont été obtenus avec un spectromètre Bruker Avance à 600 MHz équipé d'une cryo-sonde TCI (5mm) de haute sensibilité. Pour la détection des protons, 1 mg de chaîne DS d'endocan a été repris dans 500  $\mu$ L de D<sub>2</sub>O (99,9 %) lyophilisé puis resuspendu trois fois afin de supprimer toute trace d'eau résiduelle. L'échantillon a été ensuite dissout dans 0,7 mL d'un tampon phosphate 10 mM pH 7.4 contenant 0,6 M NaCl et 3 mM d'EDTA et analysé par le Dr Marco Guerrini dans le laboratoire du Pr Benito Casu (Ronzoni Institute, Milan).

#### 2.1.3.2. Résultats

L'analyse structurale de la composition de la chaîne d'endocan par RMN est réalisée, en collaboration, avec le Dr Marco Guerrini qui travaille dans le groupe du professeur Benito Casu, à l'institut Ronzoni à Milan (Italie). Deux lots de DS lui avaient été confiés, soit au total 5 mg de DS d'endocan. Les expériences ont été réalisées grâce à un spectromètre de 600 MHz, à une température de 337 degrés Kelvin, soit environ 62°C. Le spectre <sup>1</sup>H réalisé a révélé la présence d'impuretés dans les échantillons (Figure 11A). Un seul des deux échantillons a de ce fait pu être exploité en partie. Des expériences 2D de type HSQC (heteronuclear single quantum coherence) ont été entreprises (Figure 11B et Figure 12) dans le but d'obtenir la composition en oses de la chaîne. Grâce à cette méthode, la plupart des signaux anomériques ont pu être obtenus. En fait, le spectre 1D contenait trop de signaux tels que les protons anomériques, le H4 du GalNAc,4SO<sub>3</sub> et le H5 des IdoA qui s'entrecroisaient et rendaient le spectre ininterprétable. C'est pour cette raison que le spectre HSQC et plus particulièrement les pics relevant des interactions avec les protons anomères ont pu être intégrés selon une méthode développée par le laboratoire du Dr Marco Guerrini (Guerrini et al., 2005) (Figure 12). Ainsi, l'analyse de ce spectre a permis d'estimer la proportion de



**Figure 11: Spectre RMN <sup>1</sup>H et HSQC (heteronuclear single quantum coherence) de la chaîne DS d'endocan issue d'HEK293**

Ces spectres ont été obtenus avec un spectrophotomètre Bruker Avance 600 MHz à une température T=337K (62°C).

**A** : spectre <sup>1</sup>H de 0 à 8,5 ppm. Les étoiles désignent des impuretés présentes dans l'échantillon et qui ne sont pas des GAGs.

**B** : spectre HSQC. Les cadres en pointillés a) et b) contiennent les signaux utilisés pour calculer la substitution des monosaccharides par la méthode HSQC (Guerrini et al., 2005).

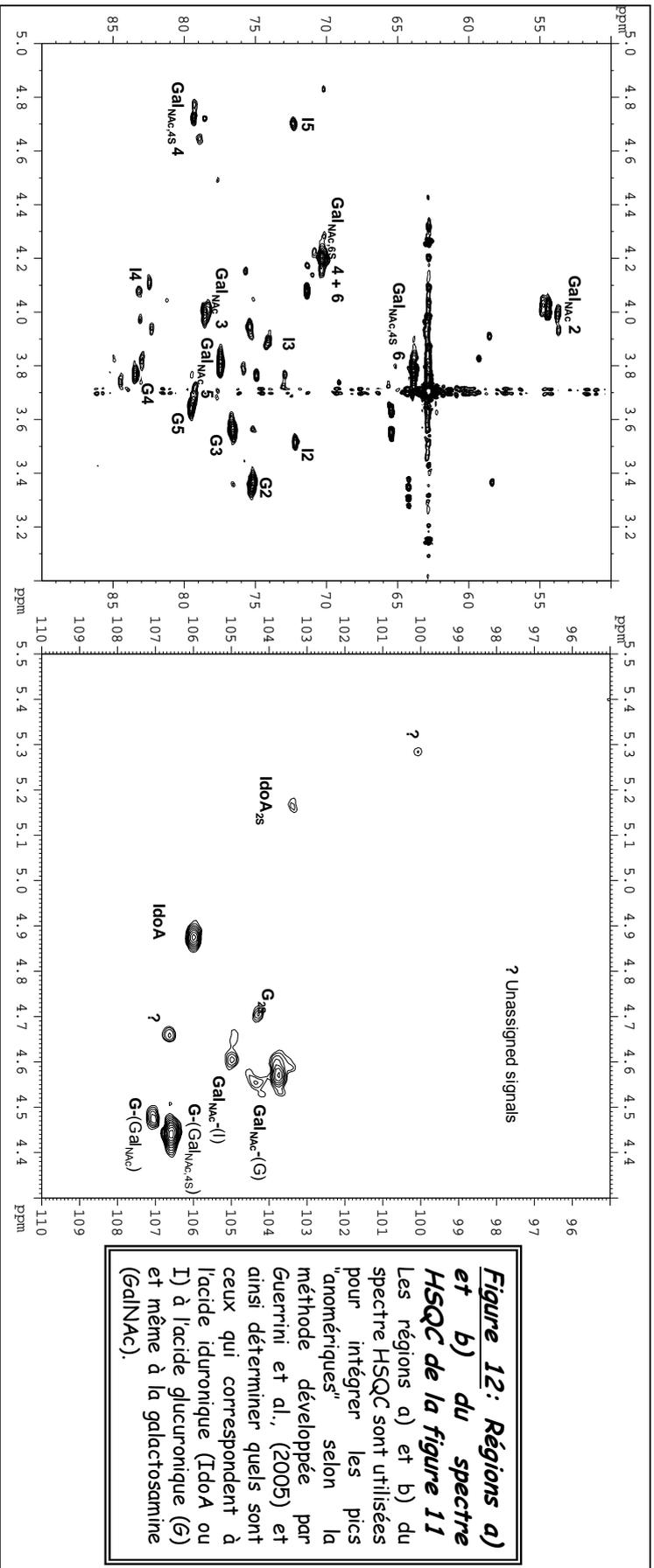
chaque acide uronique au sein du DS d'endocan (Figure 12bis). Le taux d'acides iduroniques serait donc de 29 % (dont 5 % de IdoA,2S) pour 71 % d'acides glucuroniques (dont 7 % serait 2-sulfaté). Ce spectre n'a malheureusement pas pu donner d'informations supplémentaires sur d'éventuelles séquences caractéristiques.

Cependant, un spectre NOESY 2D (Nuclear Overhauser enhancement spectroscopy) a pu être réalisé sur les échantillons fournis de DS purifié. Une partie de ce spectre est représentée sur la Figure 13. Il met en évidence les contacts dipolaires entre le H1 de l'IdoA et le H4 de la GalNAc,4SO<sub>3</sub>. En revanche, aucun contact n'a été détecté entre le H1 de l'IdoA et le H4 de la GalNAc,4OH, révélant donc l'absence de séquence du type IdoA-GalNAc,4OH au sein de la chaîne de DS d'endocan.

#### 2.1.4. Discussion

Un PG par définition comporte un core protéique sur lequel sont liés un ou plusieurs polysaccharides de type GAG. Généralement, plusieurs chaînes de GAGs sont covalamment attachées au polypeptide. Cependant il est possible, pour la décorine et la bikunine par exemple, qu'une seule chaîne de CS soit présente (cf. rapport bibliographique). Après la découverte d'endocan (Lassalle et al., 1996), la caractérisation de sa partie glycanique a été entreprise sur un endocan produit de façon recombinante par des cellules HEK293 (Bécharde et al., 2001a). De cette manière, il a été établi qu'endocan était un PG qui possède une seule chaîne de GAG, et plus précisément une chaîne de dermatane sulfate (DS), liée sur la sérine 137 du core protéique. L'utilisation d'enzymes telles que les chondroïtinases ACI et B leur avait permis d'évaluer le taux d'IdoA à 10 % contre environ 90 % de GlcA. Les résultats de l'analyse disaccharidique ont donné 2,8 % pour les disaccharides non sulfatés, 19,4 % pour les mono-6-*O*-sulfatés, 66,5 % de mono-4-*O*-sulfatés, 5,1 % de di-2,6-*O*-sulfatés, 4,6 % de di-4,6-*O*-sulfatés, et 1,5 % pour les disaccharides di-2,4-*O*-sulfatés.

L'intérêt de ce nouveau travail de caractérisation du DS d'endocan issu de deux lignées cellulaires que l'on a entrepris dès 2004 était multiple. Tout d'abord, il était intéressant de pouvoir confirmer et compléter le travail initié par l'équipe du Dr. Lassalle. Ensuite, il était important de pouvoir comparer la chaîne d'endocan produit de façon naturelle (par les cellules HUVEC dans cette étude) à une chaîne d'endocan produit de manière recombinante dans un système d'expression que l'on pouvait penser artefactuel utilisant des cellules de production intensive telles que les HEK293 (cellules immortalisées de rein embryonnaire). Cette comparaison est importante pour évaluer le biais que peut constituer l'utilisation d'endocan



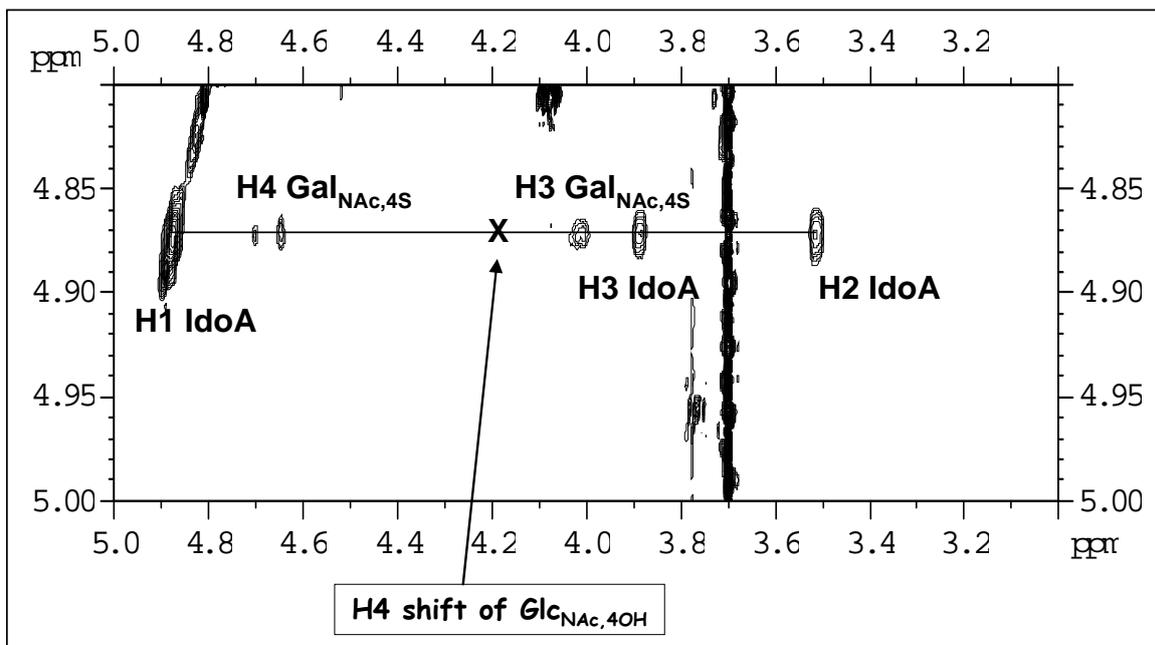
**Figure 12: Régions a) et b) du spectre HSQC de la figure 11**  
 Les régions a) et b) du spectre HSQC sont utilisées pour intégrer les pics "anomériques" selon la méthode développée par Guerrini et al., (2005) et ainsi déterminer quels sont ceux qui correspondent à l'acide iduronique (IdoA ou I) à l'acide glucuronique (G) et même à la galactosamine (GalNAc).

**Figure 12bis: Ratio obtenu entre les acides iduroniques et glucuroniques grâce à l'analyse des spectres de la figure 12**

recombinant dans l'étude des fonctions et activités d'un endocan "naturel", notamment dans un but thérapeutique (présence d'une chaîne glycanique, nature, taille etc.).

Pour caractériser la chaîne oligosaccharidique d'endocan, plusieurs méthodes pouvaient être utilisées. De plus en plus d'équipes utilisent la spectroscopie de masse pour caractériser les GAGs. Cette méthode est un réel progrès et certains travaux permettent aujourd'hui de donner la structure de plusieurs unités disaccharidiques successives (Zamfir et al., 2003). Certains laboratoires ont aussi développé des méthodes permettant de s'affranchir du radiomarquage des oligosaccharides mais sont, d'un point de vue quantitatif, encore à améliorer. C'est le cas, par exemple, de la méthode par PAGEFS (Polyacrylamide Gel Electrophoresis of Fluorophore-labeled Saccharides) (Viola et al., 2006) ou de l'électrophorèse capillaire (HPCE ou CE) (Mitropoulou et al., 2001). Malgré le développement de ces diverses méthodologies, les protocoles les plus utilisés jusqu'à maintenant, combinent le radiomarquage des chaînes oligosaccharidiques lors de leur biosynthèse à l'analyse par chromatographie des fragments ou des unités disaccharidiques obtenus suite à une dégradation enzymatique et/ou chimique. L'avantage du radiomarquage est lié à la faible quantité de GAGs nécessaire aux analyses biochimiques et structurales. Cet argument est important dans le cas de l'endocan qui a été purifié des cellules HUVECs. Par ailleurs, La RMN est un bon moyen d'obtenir des informations sur l'enchaînement des séquences disaccharidiques, mais elle est toutefois gourmande en échantillon (Guerrini et al., 2001 ; 2005).

La caractérisation de la chaîne DS d'endocan s'est donc effectuée en deux temps. Un premier travail a été réalisé à Manchester dans le laboratoire du professeur John Gallagher au "Paterson Institute for Cancer Research" (PICR), en collaboration avec le Dr Malcolm Lyon. Il a permis de caractériser et comparer les DS issus de deux sources cellulaires différentes. Ainsi, il a été vérifié que le GAG de l'endocan était, dans ces deux types de cellules, un CS/DS (Figure 5). Ensuite, le poids moléculaire et, par conséquent, la longueur de la chaîne ont également été évalués (Figure 6a et b). Nos résultats montrent que la taille des chaînes est en moyenne de 40 kDa pour la chaîne portée par de l'endocan issu des cellules HEK293 et de 15 kDa pour celle de l'endocan issu des cellules HUVEC. Cette différence constitue la différence majeure entre ces deux chaînes. Une chaîne de 40 kDa, (si le poids de l'unité disaccharidique est estimé à 480 Da), compte pour environ 80 disaccharides et la chaîne de 15 kDa issue d'HUVEC compterait en moyenne 30 disaccharides. Sur le profil d'éluion de la chaîne DS d'endocan d'HEK293 (Figure 6a) un épaulement aux alentours de 15 kDa peut être observé. Nous nous sommes alors interrogés pour savoir si cet effet était propre à l'expérience



**Figure 13: Fraction du spectre 2D-NOESY de la chaîne DS d'endocan issue de cellules HEK293**

Le spectre montre les contacts dipolaires entre le H1 de l'IdoA et le H4 de la GalNac,4SO<sub>3</sub> mais non du contact avec le H4 de la GalNac,4OH, ce qui indique l'absence des séquences IdoA-GalNac,4OH.

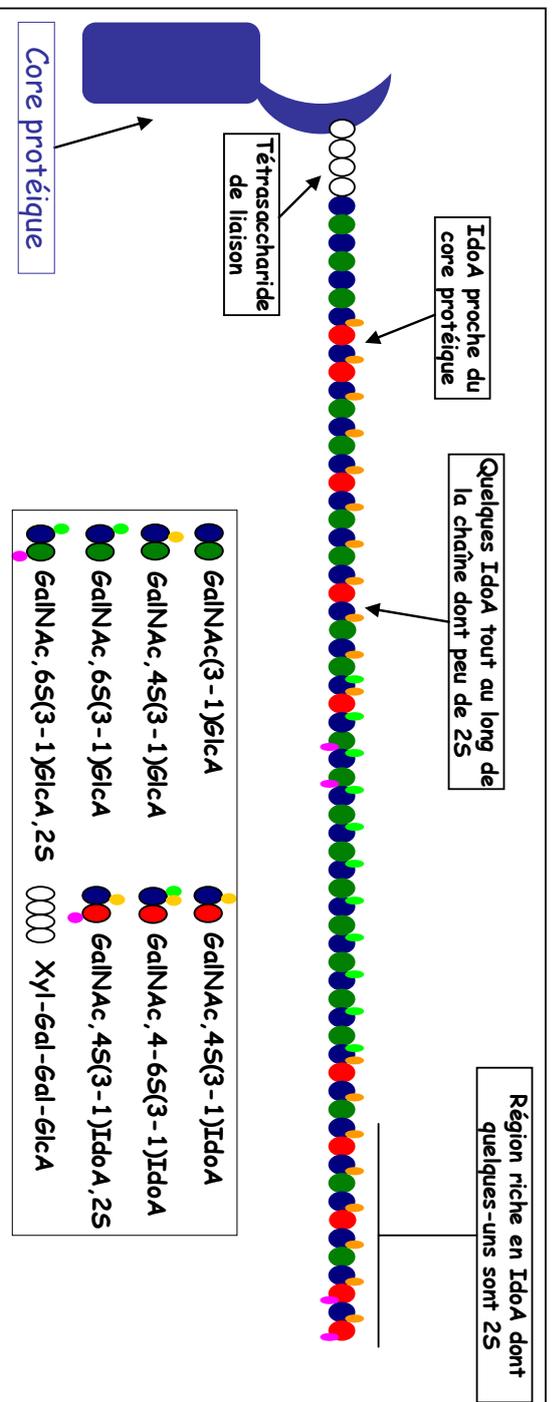
ou, si cet épaulement qui prend place au même endroit que l'élution de la majorité des chaînes DS d'endocan d'HUVEC avait une réelle signification. De plus, sur cette figure, on détecte du matériel radiomarké qui a été élué à une taille bien inférieure à celle des chaînes de DS. La détection par scintillation laisse penser que ce matériel correspond à un ose. Deux hypothèses ont été retenues. Soit ce signal correspond au tétrasaccharide d'ancrage du GAG sur le core protéique, soit cela correspond à un autre type de glycosylation. Aucune de ces deux hypothèses ne peut être validée avec la technologie utilisée à ce moment là. D'autres techniques nous permettront de vérifier la dernière hypothèse lors des études de la caractérisation du core protéique.

La structure des disaccharides de la chaîne d'endocan a ensuite été analysée par une méthode classique après digestion enzymatique à la ChABC et passage en HPLC (Figure 7). Les résultats obtenus dans la comparaison de la composition de la chaîne DS issue des cellules HUVECs à celle issue des cellules HEK293 sont étonnants. En effet, les compositions sont réellement et proprement identiques. Leurs analyses donnent 10 % pour les disaccharides non sulfatés, 22 % pour les mono-6-*O*-sulfatés, 50 % de mono-4-*O*-sulfatés, 6 % de di-2,6-*O*-sulfatés, 6 à 8 % de di-4,6-*O*-sulfatés, et 4 % pour les disaccharides di-2,4-*O*-sulfatés. Par rapport aux résultats obtenus dans l'étude de Béchard et al. (2001a), une diminution du nombre de mono-4-*O*-sulfatation peut être observée au profit probablement des di-4,6-*O*- ou des di-2,4-*O*-sulfatés. Par ailleurs, 7 % supplémentaires sont attribués dans notre étude aux disaccharides non sulfatés.

Toujours de manière enzymatique, les utilisations des Ch ACI et Ch B ont permis d'avoir une idée sur la proportion d'acide glucuronique et d'acide iduronique. Après digestion et chromatographie d'exclusion (Figure 8a), il s'est avéré que la chaîne oligosaccharidique d'endocan issue de cellules HEK293 ou de cellules HUVEC, était composée d'environ 30 % d'IdoA pour 70 % de GluA. Cette valeur ne rejoint pas le ratio IdoA/GlcA qui avait été établi en 2001 (Béchard et al.) et qui était de 10 % / 90 %.

Les résultats de l'analyse disaccharidique et du calcul de la proportion IdoA/GlcA sont intéressants. En effet, mise à part la longueur des chaînes oligosaccharidiques, la composition de la chaîne de DS d'endocan issues des cellules HEK293 ou HUVECs s'est révélée identique et ce, malgré le type cellulaire. Ces résultats ouvrent plusieurs perspectives qui seront discutés dans le paragraphe suivant.

La seconde partie du travail de caractérisation de la chaîne oligosaccharidique d'endocan a été réalisée en collaboration avec le Dr Marco Guerrini dans le laboratoire du Pr Benito Casu de l'institut Ronzoni (Milan, Italie). Ce groupe développe, depuis quelques années



**Figure 14: Modèle représentant la composition de la chaîne de dermatane sulfate d'endocan**

Ce modèle de la chaîne d'endocan prend en compte toutes les informations qui ont pu être obtenues par le travail réalisé avec le radio-marquage suivi des chromatographies et de l'analyse RMN. C'est un modèle de 30 disaccharides qui peut convenir pour représenter objectivement la chaîne DS de l'endocan d'HUVEC.

Pour ce faire, le ratio IdoA/GlcA de 30 % / 70 % a été respecté (chromatographie et RMN). Aucun IdoA-GalNAc, 4OH n'a été positionné (RMN). Les 5 % d'IdoA<sub>2S</sub> et 7 % de GlcA ont été respectés (RMN). Les proportions dans la structure des unités disaccharidiques ont été respectées, à savoir 11 % de non-sulfaté, 22 % de 6-5, 50 % de 4-5, 6 % de 2,6-5, 7 % de 4,6-5, 4 % de 2,4-5. La taille et la proportion des fragments qui peuvent être obtenus par la digestion à la ChACI ont également été prises en compte, à savoir 36 % de dp4 et 11 % de dp6. La mauvaise digestion par la ChB nous a conduit à placer les IdoA de manière assez dispersée, surtout pour les IdoA étant connus pour être non 2-S, cette configuration étant peu favorable à l'enzyme. Un regroupement d'IdoA à tout de même été placé à l'extrémité de la chaîne, cette disposition favorable à la ChB (d'autant plus si les IdoA sont 2-S) permettrait d'expliquer le léger déplacement observé sur la colonne CL-6B correspondant à la légère diminution de la taille de la chaîne obtenue avec l'endocan d'HUVEC. De même, deux IdoA successifs ont été placés près du core protéique afin d'expliquer la disparition du DS sur le Western Blot d'endocan traité ou non avec la ChB présentés dans les travaux de Becharh et al., (2001). La disposition des sulfates (4-5, 6-5 ou 2-5) a été faite selon le raisonnement que les sulfotransférases agissent sur des disaccharides successifs.

maintenant, l'analyse structurale des GAGs par RMN. Le but était qu'ils puissent nous donner un maximum d'informations sur la structure et l'organisation de la chaîne d'endocan issue de cellules HEK293. En effet, les quantités d'oligosaccharide exigées par ce type d'expérience ne permettaient pas d'envisager une analyse de la chaîne oligosaccharidique d'endocan issue d'HUVEC. Malgré la présence d'impuretés dans les échantillons fournis, des premiers résultats intéressants ont été obtenus. Tout d'abord, grâce à l'analyse du spectre HSQC et particulièrement à l'intégration des pics de protons anomériques, le ratio IdoA/GlcA a été évalué à 29 % / 71 %, avec un pourcentage de 2-O-sulfatation de 5 % pour les IdoA et 7 % pour les GlcA. Ces résultats confirment ceux obtenus par filtration sur gel après une digestion enzymatique. Il convient de signaler que dans le cas des travaux réalisés à Manchester, les GAGs avaient été radiomarqués sur une production d'endocan réalisée en flasks. L'endocan utilisé pour fournir les DS permettant la RMN a été produit ultérieurement en CL1000 qui permet d'augmenter la concentration cellulaire (augmentation de la biomasse). Ces résultats similaires montrent la constance de la structure de la chaîne DS d'endocan.

La réalisation d'un spectre NOESY 2D, permettant de se focaliser sur les interactions dipolaires, a permis de mettre en évidence l'absence de séquences IdoA-GalNAc,4OH et la présence de séquence IdoA-GalNAc,4SO<sub>3</sub> au sein du DS d'endocan.

A partir des résultats obtenus, une représentation globale de l'organisation de la chaîne glycanique d'endocan a pu être réalisée telle que représentée en Figure 14. Ce modèle a été réalisé pour une chaîne de 15 kDa d'endocan issu des cellules endothéliales HUVEC. Si nous voulions représenter la chaîne DS d'endocan issue des cellules HEK293, il conviendrait en fait d'aligner deux chaînes d'endocan d'HUVEC à la suite l'une de l'autre. Cette disposition offrirait au milieu de la chaîne un regroupement d'IdoA qui serait favorable à la digestion par une Ch B. Cela justifierait le décalage observé lors d'une telle digestion à la sortie de la colonne CL-6B, interprété comme une coupure au milieu de la chaîne de 40 kDa pour donner deux longs fragments.

### **2.1.5. Perspectives**

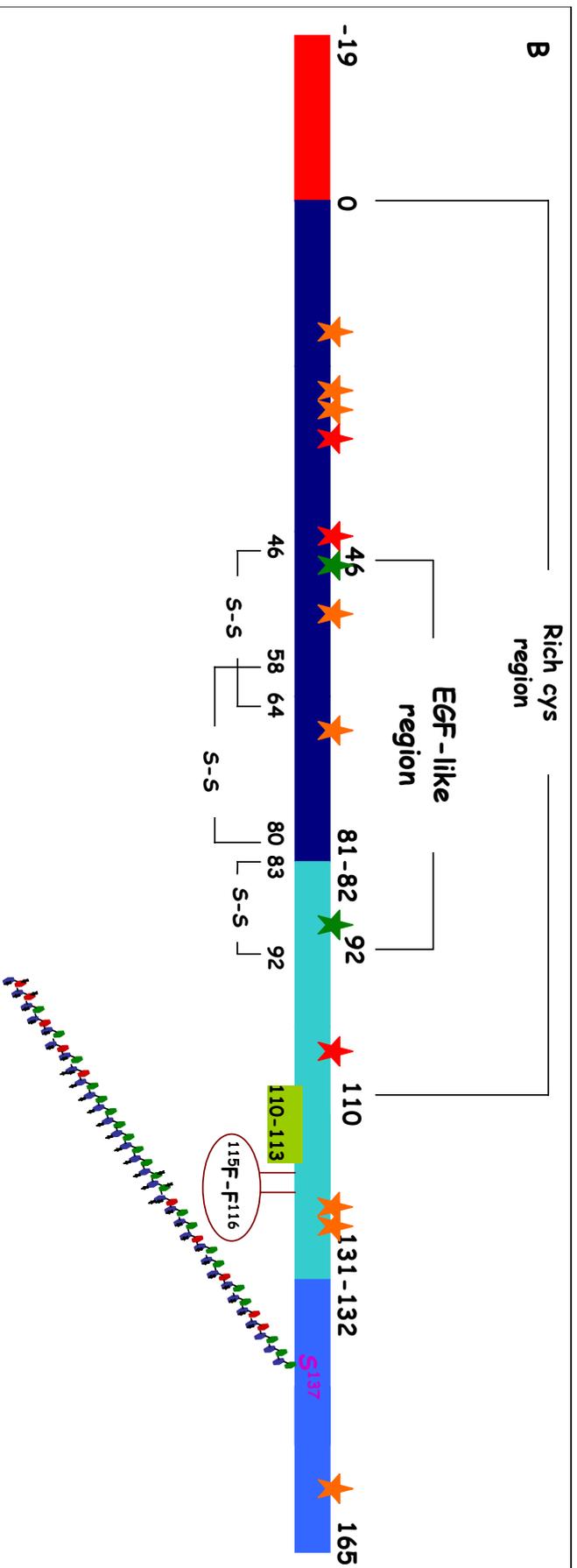
Au regard de ces résultats, les perspectives de ce travail sont nombreuses. Tout d'abord, concernant la comparaison des chaînes issues de lignées cellulaires différentes, de nouvelles expériences permettraient d'approfondir nos résultats. Nous pouvons, par exemple, imaginer utiliser un traitement à la Ch ACI, pour que les fragments obtenus soient individuellement réduits en unités disaccharidiques et analysés par HPLC. Ce genre d'expériences offrirait des

informations sur la structure des disaccharides inclus dans des séquences comportant des IdoA successifs. Par ailleurs, la RMN pourrait offrir plus de résultats que ceux obtenus jusqu'ici, notamment sur un nouvel échantillon de DS contenant moins d'impuretés, afin d'obtenir d'autres informations sur l'enchaînement des oses.

D'un point de vue plus général, les similitudes entre la chaîne produite par les cellules HEK293 et celle produite par les cellules HUVEC sont assez intrigantes. Plusieurs expériences complémentaires pourraient être mises en place pour approfondir et élargir ces observations. En effet, il serait intéressant d'établir la comparaison des chaînes d'endocan produits de façon naturelle, avec un plus grand nombre de lignées cellulaires l'exprimant. Depuis nos travaux débutés en 2003, il a été en effet mis en évidence que d'autres types de cellules que les cellules endothéliales exprimaient endocan. Ainsi les cellules tumorales de mélanomes ou de glioblastomes (Sefkor et al 2003 ; Adam et al, soumis à publication), qui peuvent être cultivées en laboratoire, probablement plus facilement que des adipocytes (Welner et al 2003). Plus récemment, il a été observé que des cellules phénotypiquement complètement différentes, les cellules HT1080 issues de fibrosarcome, produisent de l'endocan (Motte, Adam, Villedieu et Delehedde, communication personnelle). Ces cellules ne sont ni des cellules endothéliales ni des cellules épithéliales, mais des cellules de type fibroblastique. Il serait intéressant donc de connaître la nature et la composition de la chaîne de GAG de l'endocan synthétisé par ces autres types de cellules. Si, quelque soit la lignée cellulaire, la chaîne de DS d'endocan était la même, l'hypothèse que le core puisse être un élément régulateur pourrait être considérée. Cela a été par exemple exploré pour la décorine (Seidler et al., 2007 ; Cheng et al., 1994). Pour vérifier cette hypothèse, une seconde série d'expériences peut être également proposée. Elle consisterait à comparer, pour une même lignée cellulaire, la composition oligosaccharidique des chaînes de deux CS/DSPGs différents, dont endocan. Si les deux PG ont des chaînes identiques, la composition structurale de celles-ci sera plus facilement attribuée à la lignée cellulaire. En revanche, si la composition varie entre deux PGs d'une même lignée cellulaire, cela serait un argument en faveur du rôle "régulateur" du core protéique.

Dans le même genre d'expérience, il a été montré que l'expression d'endocan pouvait être modulée par différents facteurs de croissance ou cytokines tels que le  $TNF\alpha$ , l'IL1, l'IFN $\gamma$ , le VEGF ou encore le HGF/SF, probablement dans un but fonctionnel (revue dans Sarrazin et al., 2006). Par conséquent, il serait intéressant d'analyser l'effet de ces différents facteurs sur la structure de la chaîne d'endocan issue, par exemple, des HUVEC. Des travaux ont montré que certains facteurs avaient une influence sur la composition des chaînes de type GAGS

A			
-19	MKSVLLLTTL	LYPAHLVAAW	SNNVAVDPCQ
42	RGETCYRTVS	GNDGMKCGPG	LRCOPNSGED
102	ICDRGTGKCL	KRPPEQYSVT	KSSNRVSLT
162	LNPR		
			HCDSSSECKSS
			PRCKRTVLLDD
			CGCCRYCAAG
			PEGEFGICK
			DCPYGTGMD
			GRFTCNCOSS
			EHDMA <sup>S</sup> GDGN
			IVREEVYKEN
			AAGSPVMRKW



**Figure 15A : Séquence protéique d'endocan**

Le polypeptide d'endocan est issu de la traduction de 3 exons, l'exon 1 (rouge et bleu foncé), l'exon 2 (bleu clair) et l'exon 3 (bleu intermédiaire). Ce polypeptide est caractérisé par une séquence "signal" (rouge) en N-terminal et une séquence qui correspond à la protéine mature (Bleu). En rose a été indiqué la sérine 137 sur laquelle est lié covalamment le GAG (GS/DS) d'endocan. Les deux phénylalanines présentées comme importantes pour l'activité d'endocan (Bechard et al., 2001) ont été soulignées (exon 2).

**Figure 15B : Représentation schématique de l'organisation d'endocan**

Le code couleur de ce schéma est le même que figure 1A. Sont mentionnés sur ce schéma, la région riche en cystéines (18/110a.a.), la région du probablémotif EGF-like avec les ponts disulfures associés, la sérine 137 où la chaîne de DS est covalamment liée, la région riche en phénylalanine (F115 et 116), le site potentiel de sumoylation (110-113), et les sites putatifs de phosphorylation sur les sérines 15, 20, 21, 51, 67, 124, 129, 155 (étoiles oranges), les thréonines 27, 45, 107 (étoiles rouges) et les tyrosines 47, 87 (étoiles vertes).

(Tiedemann et al., 2005 ; Carter et al., 2003). Par exemple, dans certains cancers, les chaînes de DS s'enrichissent en IdoA. Or, comme endocan est surexprimé dans certains types de cancers, on peut donc s'attendre, lors d'une stimulation par des facteurs augmentant son taux d'expression, à ce que le taux d'IdoA de sa chaîne augmente également.

## **2.2. CARACTERISATION DU CORE PROTEIQUE D'ENDOCAN**

### **2.2.1. Modélisation**

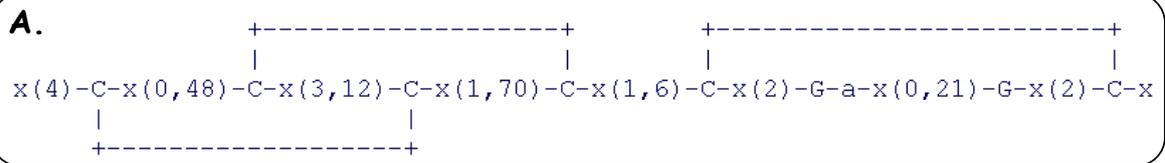
#### **Caractéristiques biochimiques**

L'endocan est un polypeptide de 184 acides aminés (dont la séquence est représentée à la Figure 15A) issu de la traduction de 3 exons (Figure 15B). Il comporte une séquence signal en N-terminal de 19 acides aminés (Figure 15B) ce qui en fait une protéine mature de 165 acides aminés. Le core protéique d'endocan mature a un poids moléculaire de 18 122,4 Da et un point isoélectrique théorique de 6,83 (logiciel ProtParam sur [www.expasy.org](http://www.expasy.org)). Ces valeurs ne tiennent pas compte des modifications post-traductionnelles, dont la chaîne glycanique de DS qui est covalamment liée à la sérine 137 (Figure 15B).

#### **Caractérisation structurale**

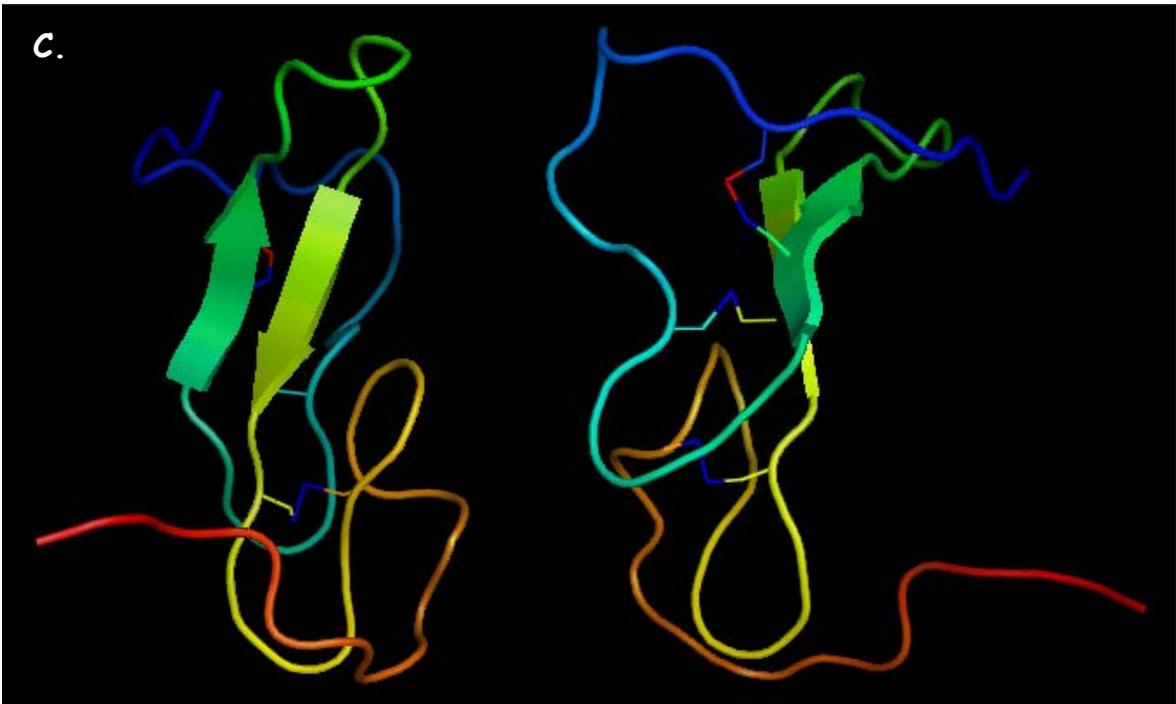
##### **Un probable motif de type EGF**

Une recherche d'homologues protéiques d'endocan par comparaison de séquences peptidiques (BLAST sur [www.expasy.org](http://www.expasy.org)) a été effectuée. Aucune protéine ne présentait plus de 50 % d'homologie avec l'endocan. Cependant, une séquence peptidique particulière a été retrouvée dans de nombreuses protéines. Cette séquence "consensus" commune à ces protéines est un motif de type EGF-like (Epidermal Growth Factor – Like), décrit dans la Figure 16A. Ce motif, d'environ trente à quarante acides aminés, identifié pour la première fois dans l'EGF (Epidermal Growth Factor), est retrouvé de manière plus ou moins conservée dans de nombreuses autres protéines (Davis, 1990 ; Appella et al., 1988 ). Les PGs qui contiennent ce type de motif pouvant être cités de manière non exhaustive sont : l'agrine, le perlécan (2 motifs), l'aggrécan (1 motif), le versican (2 motifs), le neurocan (2) et le brévican (1). Les protéines qui possèdent ce motif EGF-like ont en commun d'être extracellulaires, liées à la membrane ou sécrétées. Ce motif est composé de six cystéines impliquées dans des ponts disulfures, comme cela est représenté dans la Figure 16A. La structure de plusieurs domaines EGF-like a été résolue. D'une manière générale, elle est caractérisée par deux



**B.**

-19 MKSVLLLTTL	LVPAHLVAAW	SNNYAVDCPQ	HCDSECKSS	PRCKRTLDD	CGCCRVCAAG
42 RGETCYRTVS	GMDGMKCGPG	LRCQPSNGED	PFGEFEGICK	DCPYGTFGMD	CRETCNCQSG
102 ICDRGTGKCL	KFPFFQYSVT	KSSNRVSLT	EHDMSGDGN	IVREEVVKEN	AAGSPVMRKW
162 LNPR					



**Figure 16: Le motif EGF-like**

**A** : la séquence indiquée sur cette figure représente les conditions nécessaires pour définir une séquence EGF-like. A savoir, 'C' sont les cystéines impliquées dans les ponts disulfures, 'G' sont les glycines très souvent conservées, 'a' sont des acides aminés aromatiques également souvent conservés et 'x' peut être n'importe quel acide aminé.

**B** : Est représentée sur cette figure, la séquence en acide aminé d'endocan avec en rouge les acides aminés qui respectent la règle d'une structure EGF-like indiquée figure 15.

**C** : Cette figure représente la structure EGF (réf PDB : 1EGF) obtenue par RMN. Classiquement, le motif EGF se compose de deux feuillets bêta anti-parallèles (représentés par des flèches) suivis par une boucle mise en forme par les ponts disulfures. La structure de droite est le produit d'une rotation de 60° environ de la structure de gauche selon l'axe vertical et dans le sens des aiguilles d'une montre si l'on regarde par en dessous. Les ponts disulfures sont indiqués sur les structures.

feuillet bêta anti-parallèles suivis, du côté C-terminal, par une boucle formée par un pont disulfure (Figure 16C). La séquence protéique d'endocan de l'acide aminé 40 à 91 contient tous les éléments requis pour être définie comme une séquence EGF-like (Figure 16B), à la seule condition que les ponts disulfures puissent se former.

#### Une région riche en cystéines du côté N-terminal

Dans les 110 premiers acides aminés du core protéique d'endocan, sont réparties 18 cystéines. Cette forte présence de cystéines laisse pressentir que la région N-terminale de la protéine est structurée, car stabilisée par des ponts disulfures. Le logiciel IUPred (<http://iupred.enzim.hu>) a été développé pour prédire les régions plus ou moins structurées d'une protéine. Nous l'avons utilisé pour évaluer "le désordre" du core protéique d'endocan. En effet, l'analyse théorique qui a été faite à partir de la séquence protéique de l'endocan, indique que la région N-terminale, jusqu'à l'acide aminé 132, est assez ordonnée (Figure 17) alors que la région du côté C-terminal (à partir de l'acide aminé 133) semble peu structurée, moins globulaire et donc plus labile. Un désordre de plus petite envergure peut être remarqué sur la séquence EGF-like qui vient d'être définie.

#### Modifications post-traductionnelles

L'utilisation de logiciels spécialisés a permis d'évaluer, de manière purement bioinformatique, les modifications post-traductionnelles que pourrait subir la partie protéique d'endocan. Ainsi, aucun site potentiel de N-glycosylation n'a été détecté (utilisation du logiciel NetNGlyc sur [www.expasy.org](http://www.expasy.org)). Cependant, quatre d'entre eux ont obtenu un score non négligeable. Ce sont les aspargines N4, N68, N125, et N141. En utilisant les logiciels YinOYang, NetOGlyc, et récemment OGPET (expasy), aucun site potentiel de O-glycosylation n'a pu clairement être proposé. Par ailleurs, l'utilisation de "Prosit Motif Search" a permis de détecter la séquence consensus de liaison de la chaîne DS d'endocan dont l'existence est bien sûr connue et prouvée par mutagenèse (Bécharde et al., 2001).

Ensuite, grâce au logiciel NetPhos (expasy), plusieurs sites de phosphorylation ont pu être proposés, dont sept sur des sérines (Ser 15, 20, 21, 51, 67, 129, 155), deux sur des thréonines (Thr 45, 107), et deux sur des tyrosines (Tyr 47,87). Le logiciel "Prosit Motif Search", quant à lui, classe les sites de phosphorylation différemment. Il propose trois sites de phosphorylation par la protéine kinase C (PKC), qui sont la sérine 21, la thréonine 107 et le dernier, qui n'avait pas été proposé par NetPhos, la sérine 124. De même, quatre sites de phosphorylation par la caséine kinase II (CKII) ont été proposés dont la thréonine 27, non donné par NetPhos, et également les sérines 51, 67 et 129.



**Figure 17: Prédiction du désordre dans la structure d'endocan**

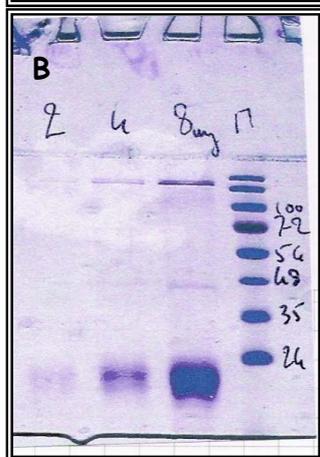
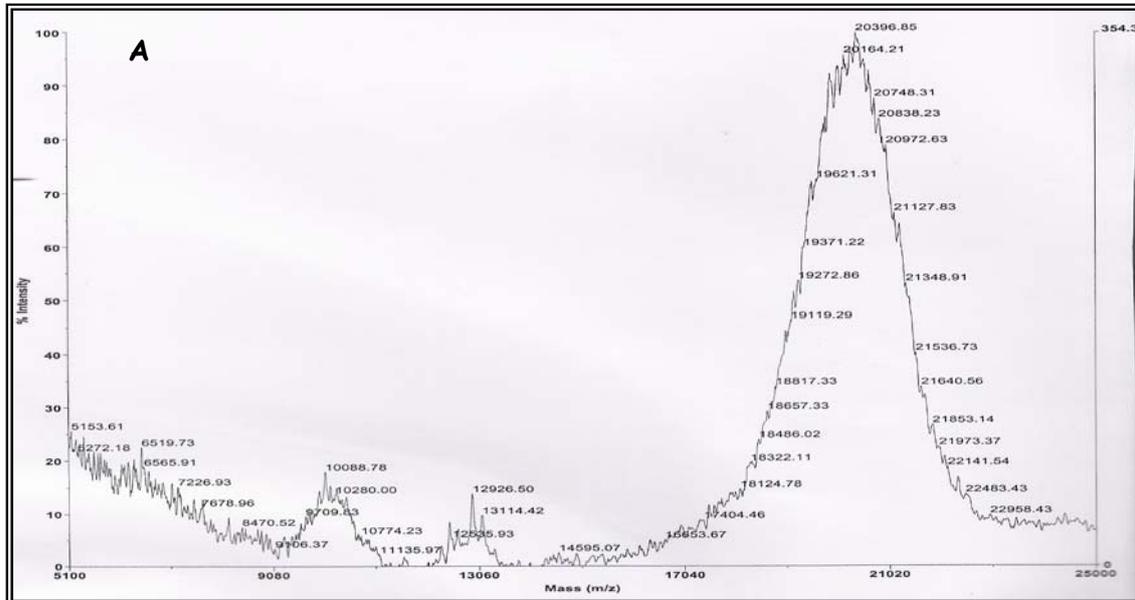
Le graphique représente le désordre probable de la protéine (en ordonnée) en fonction des acides aminés (abscisses) (obtenu par le logiciel IUpred). La bande en bleu foncé indique quels sont les résidus d'endocan qui seraient préférentiellement situés dans un domaine globulaire. Ceci est reporté en dessous du graphique, avec les a.a. appartenant à une région stable en majuscule bleu et ceux qui appartiendraient plutôt à une région désordonnée en minuscule (bleu ciel). Ainsi, la région supposée globulaire serait entre les acides aminés 1 et 133, ensuite, coté C-terminal la protéine serait bien plus désordonnée.

Nous avons suggéré en début d'année 2006, qu'un site de N-myristoylation serait potentiellement présent sur endocan (prédiction de "Prosite Motif Search") (Sarrazin et al., 2006). Cependant, l'utilisation d'un nouveau logiciel MYR, créé, disponible depuis peu et surtout spécialisé dans la prédiction des sites de N-myristoylation, ne nous a pas permis de détecter sur l'endocan de site favorable à ce genre de modification post-traductionnelle.

Récemment, le logiciel SUMOplot, développé par la société Abgent, a également été mis à disposition sur expasy. Il permet de prédire les sites potentiels de sumoylation. La sumoylation est l'ajout de la protéine SUMO-1 (small ubiquitin-related modifier, connue également sous le nom de PIC1, UBL1, Sentrin, GMP1, et Smt3) sur une protéine, par une liaison covalente sur une lysine incluse dans une séquence de type : B/K-X-D-E où B est un résidu hydrophobique, K est la lysine d'ancrage, X un acide aminé quelconque et D et E des résidus acides. Ce logiciel a donc permis sur endocan de mettre en évidence deux sites potentiels de sumoylation, qui sont respectivement <sup>55</sup>MKCG<sup>58</sup> et <sup>110</sup>LKFP<sup>113</sup>. La première séquence présente une probabilité d'existence de 63 % ce qui classe le site comme faible selon le logiciel alors que la seconde séquence est classée comme un site à forte probabilité avec un score de 80 %.

### **2.2.2. Recherche d'autres glycosylations que la chaîne de DS**

Comme cela est décrit précédemment, de l'endocan muté sur la position 137 (S137A), et donc dépourvu de sa chaîne oligosaccharidique, avait été produit. Cet endocan mutant a plusieurs utilités dont celle de pouvoir être étudié directement en cristallographie (i.e. sans devoir enlever le GAG) comme nous le détaillerons dans le prochain paragraphe. Après de nombreuses tentatives en variant les conditions cristallographiques, les résultats toujours négatifs obtenus nous ont conduits à nous interroger plus précisément sur le core protéique d'endocan. En effet, la cristallographie nécessite une protéine pure (ce qui était le cas), et homogène (ce qui n'était peut-être pas le cas). Or, bien que les glycosylations soient capables dans certains cas de favoriser la cristallogénèse, particulièrement en rendant les protéines plus solubles, dans bien d'autres cas, elles peuvent rendre l'échantillon non-homogène et donc non cristallisable. En plus de ces problèmes récurrents de cristallisation, plusieurs indices ont orienté notre recherche vers d'autres glycosylations ou d'autres éléments pouvant être liés au core protéique et faisant varier sa taille. En effet, lorsque la protéine est analysée par MALDI-TOF, le pic correspondant n'est pas étroit, au contraire le signal est étalé de la taille du core protéique (18106,4 Da) jusqu'à une taille de 4000 daltons supérieurs (environ 22000 Da) à la



**Figure 18A:** Spectre MALDI-TOF du core protéique d'endocan purifié

Le core protéique d'endocan a une masse théorique de 18106 Da. En spectrométrie de masse, le profil obtenu montre un pic large de 4000 Da qui débute à la masse attendue et dont le sommet se trouve aux alentours des 20200 Da.

**Figure 18B:** Gel SDS-PAGE coloré au bleu de Coomassie du core protéique d'endocan

Le core protéique d'endocan migre en gel SDS-PAGE non pas à 18 kDa mais aux alentours de 20 kDa en une bande plutôt large. Sur les gels, 2, 4 et 8  $\mu$ g de core protéique d'endocan ont été déposés à gauche du marqueur de poids moléculaire.

taille de la protéine d'endocan (Figure 18A). Ceci est confirmé de façon visuelle sur SDS-PAGE, puisque la bande correspondant au core protéique d'endocan purifié en condition non réductrice n'est pas fine mais peut varier d'environ 4000 Da au dessus de la masse théorique (Figure 18B). Par ailleurs, lors de la caractérisation de la chaîne glycanique d'endocan, lorsque cette dernière avait été radiomarquée, séparée du core protéique, puis analysée par chromatographie d'exclusion stérique (Figure 6A de la caractérisation de la chaîne d'endocan), des fragments de faibles tailles avaient été observés, mais leur nature n'a jamais été identifiée. Il n'est pas exclu que cela puisse être des glycosylations non glycosaminoglycanniques.

Pour ces raisons, un travail a été mis en place pour évaluer si le core protéique d'endocan dépourvu de sa chaîne de DS possède d'autres glycosylations ou d'autres modifications qui pourraient ainsi créer un endocan non-homogène. Une méthode de marquage des oses a été utilisée afin d'avoir une première idée sur la présence d'éventuelles glycosylations, puis, différentes digestions enzymatiques suivies d'analyses par spectroscopie de masse ont été effectuées.

#### **2.2.2.1. Matériel et méthodes**

##### *Détection en Western Blot des glycosylations par un marquage à la biotine*

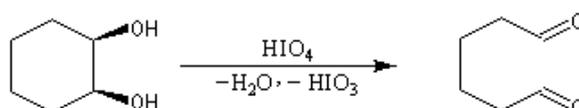
###### Principe du marquage (Figure 19)

Afin de marquer les glycosylations potentielles d'endocan à la biotine, le protocole que nous avons utilisé est couramment utilisé pour marquer les anticorps, par exemple, et les glycoprotéines d'une manière générale. Il a lieu en deux étapes qui sont tout d'abord, une oxydation des alcools vicinaux d'un ose par le périodate de sodium (ou l'acide périodique), créant ainsi une coupure "oxydante". En fait, la liaison entre les deux fonctions alcools est coupée formant, de cette manière, deux composés carbonylés (aldéhydes) sur lesquels, dans un deuxième temps, l'amine primaire de l'hydrazide (lié à la biotine par un bras espaceur) peut réagir. La biotine est ainsi liée de manière covalente à l'ose décyclisé.

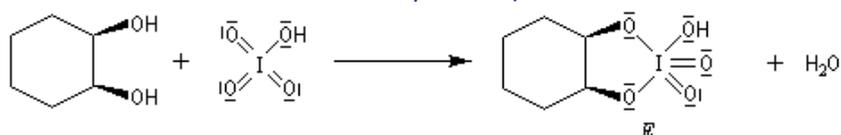
Deux façons d'effectuer le marquage à la biotine ont été testées. Soit l'oxydation au périodate suivie de la biotinylation se fait directement sur la membrane de PVDF sur laquelle la protéine a été transférée, soit le marquage se fait en solution et les échantillons sont ensuite analysés par Western Blot.

**Création d'aldéhydes par coupure oxydative sur les oses :**

Les  $\alpha$ -glycols vicinaux subissent une réaction de coupure en présence d'acide périodique  $\text{HIO}_4$  pour donner des composés carbonylés. Il s'agit de la réaction de Malaprade.

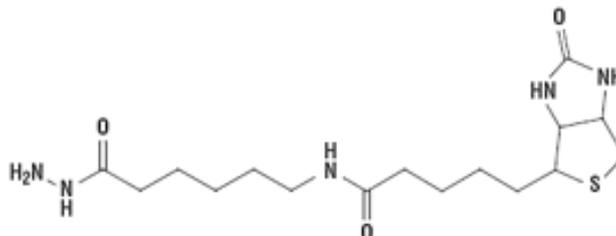


Elle fait intervenir un ester de l'acide périodique comme intermédiaire.



**Ajout de la biotine :**

La biotine hydrazide peut alors se coupler via son amine aux aldéhydes ainsi formés.



**Figure 19: Principe du marquage des glycanes par oxydation au périodate suivi de l'ajout d'une biotine hydrazide**

### Protocole

Marquage en solution : la protéine est dissoute dans un tampon d'acétate de sodium à 100 mM, pH 5.5, à une concentration de 5 mg/mL. La solution de périodate de sodium ( $\text{NaIO}_4$ , Sigma) à 100 mM (21,4 mg/mL) est préparée dans l'eau juste avant utilisation. A l'obscurité, le périodate (10 mM final) est ajouté (ou non dans le cas de l'échantillon témoin) à la protéine par aliquote toutes les deux minutes (5 x 20  $\mu\text{L}$  pour 1 mL de protéine). Il faut ensuite attendre 20 minutes après l'ajout du dernier aliquote. La protéine est déssalée par filtration sur gel (PD10) contre un tampon d'acétate de sodium à 100 mM, pH 5.5. La biotine(LC)hydrazide (Vector) est ensuite dissoute dans du DMSO à 50 mg/mL et 40  $\mu\text{L}$  de cette solution est ajoutée (ou non dans le cas de l'échantillon témoin) à 1 mL de protéine. L'ensemble est laissé à incuber de 2 heures à toute une nuit, à température ambiante. La protéine est récupérée par PD10 et sa biotinylation est visualisée par Western Blot à l'aide de l'extravidine-HRP (1/5000) (Sigma). Un Western Blot similaire est réalisé avec, cette fois-ci, un anticorps primaire monoclonal anti-endocan puis un anticorps secondaire anti-souris-HRP (Amersham). Ensuite, la détection se fait par chimioluminescence (ECL) et autoradiographie (ECL-hyperfilms Amersham).

Marquage directement sur une membrane de PVDF : L'endocan mutant, l'endocan et un anticorps (glycoprotéine témoin) sont déposés (2  $\mu\text{g}$ ) sur un gel SDS-PAGE 12 %. Après migration des échantillons, un transfert d'une heure à 100V sur une membrane de PVDF (Bio-Rad) est effectué. La membrane est rincée trois fois en PBS, puis elle est incubée dans le périodate de sodium (10 mM), ou non (témoin négatif) à l'obscurité, pendant 20 minutes. Trois rinçages au PBS sont effectués puis la membrane est incubée dans une solution de biotine-hydrazide (Vector) à 10 mM pendant une heure à température ambiante (TA). Les groupements qui n'ont pas réagi sont bloqués avec une solution de Tris 20 mM pendant 20 minutes. La membrane est ensuite rincée en PBS, puis saturée avec une solution de TBS contenant 5 % de lait, pendant une heure. La membrane est rincée au TBS-Tween20 (0,2 %) puis incubée avec l'extravidine-HRP (1/5000<sup>ème</sup>) (Sigma), pendant une heure, à température ambiante. La détection de la biotine se fait par chimioluminescence en utilisant les réactifs ECL et les "ECL-hyperfilm" (Amersham).

### *Déglycosylations enzymatiques et analyse par spectrométrie de masse*

*Digestion à la O-Glycosydase* : La protéine (10  $\mu\text{g}$ ) est incubée dans un tampon phosphate de sodium 200 mM à pH 7.2 contenant 50 mM d'EDTA. 3 mU d'enzyme (Roche) sont ajoutées et la réaction a lieu 26 heures, à 37°C.

*Digestion à la N-Glycosydase F* : La protéine (10 µg) est incubée dans un tampon phosphate de sodium 200 mM à pH 7.2 contenant 50 mM d'EDTA. 5 U d'enzyme (Roche) sont ajoutées. La réaction a lieu 26 heures, à 37°C.

*Spectroscopie de masse* : Les échantillons sont dilués au 1/10<sup>ème</sup> dans du TFA 0,1 %. Ensuite 1 µL de cette dilution est mélangé à 1 µL d'acide sinapinique saturé dans 50 % d'acétonitrile, TFA 0,3 % directement sur la plaque du MALDI-TOF (Voyager Elite, Perspective). Pour les fragments de plus petite taille (550 Da), la matrice utilisée est l'acide  $\alpha$ -cyano-4-hydroxycinnamique.

#### *Digestions peptidiques et spectrométrie de masse*

*Digestion à la trypsine* : La protéine (5 µg) est incubée (100 µL final) dans un tampon phosphate de sodium 20 mM à pH 7.2, contenant 5 mM d'EDTA. 2,5 µg de trypsine sont ajoutés au volume réactionnel. La réaction a lieu 24 heures, à 37°C. Ensuite la O-glycosydase est ajoutée (3 mU) et la réaction se poursuit pendant 48 heures supplémentaires.

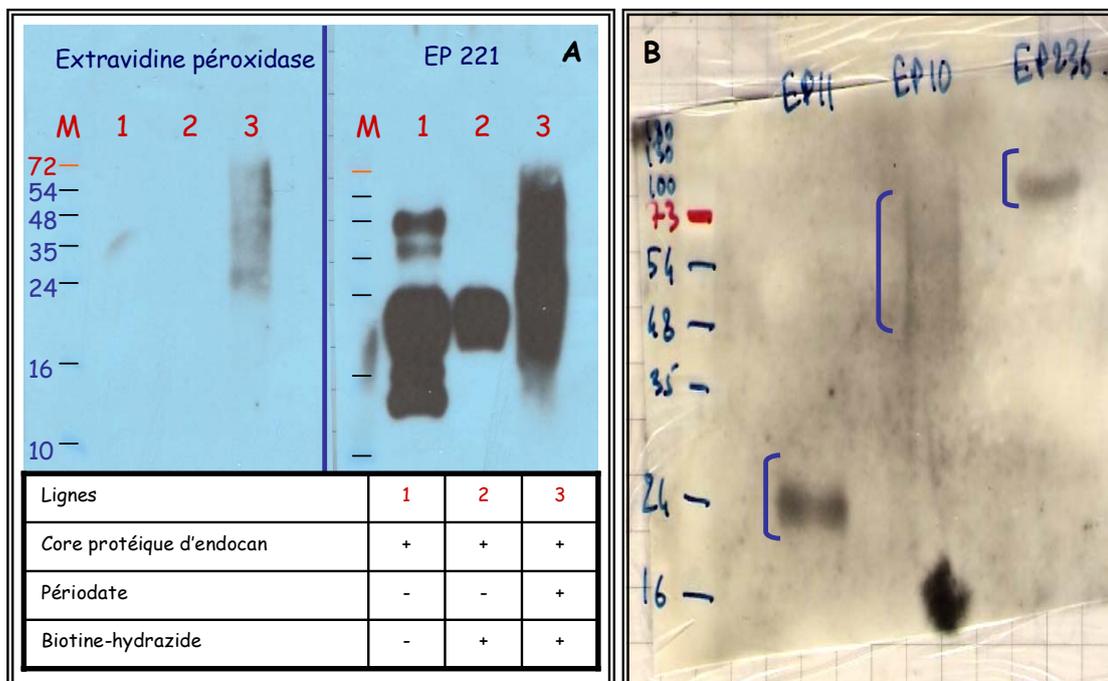
*Digestion à l'endoprotéinase Lys-C* : La protéine (5 µg) est incubée dans un tampon de digestion de type Tris-HCl 25 mM, pH 8.5 contenant 1 mM d'EDTA. 5 % (w/w) d'endoprotéinase Lys-C est ajouté à la protéine. La réaction a lieu toute la nuit, à 37°C.

*Digestion à l'endoprotéinase Glu-C* : La protéine est incubée dans un tampon de digestion de type carbonate d'ammonium 25 mM, pH 7.8. 5 % (w/w) d'endoprotéinase Glu-C est ajouté à la protéine. La réaction a lieu toute la nuit, à 25°C.

*Spectroscopie de masse* : Les échantillons sont dilués au 1/10<sup>ème</sup> dans du TFA 0,1 %. Ensuite 1 µL de cette dilution est mélangé à 1 µL d'acide sinapinique saturé dans 50 % d'acétonitrile, TFA 0,3 % directement sur la plaque du MALDI-TOF (Voyager Elite, Perspective). Pour les fragments de plus petite taille (550 Da), la matrice utilisée est l'acide  $\alpha$ -cyano-4-hydroxycinnamique.

#### *Méthanolyse du blot protéique d'endocan, HPLC et spectrométrie de masse : méthode H.F.B.A.*

Sur 300 µg de core protéique d'endocan, les glycanes potentiels sont hydrolysés par un traitement au méthanol/HCl à 80°C. Les monosaccharides ainsi libérés sont dérivés en heptafluorobutyrate et analysés en chromatographie en phase gazeuse (GC), couplée ou non à la spectrométrie de masse (GC-MS), selon le protocole de Zanetta et coll. (1999). La lysine (0,5 µg) est introduite comme témoin de référence avec les monosaccharides ainsi libérés.



**Figure 20: Révélation par chemiluminescence de la présence de biotine sur les potentiels O- ou N-glycosylations potentielles d'endocan**

A : Le marquage à la biotine est effectué en solution et les échantillons sont ensuite soumis à un western blot révélé soit avec l'extravidine peroxidase (révélation de la biotine) soit avec l'anticorps primaire EP221 et un anti souris HRP secondaire (révélation du core protéique). Pour chacune des révélations, la ligne M correspond au marqueur de poids moléculaire, et les lignes 1, 2 et 3 au core protéique d'endocan traité ou non au périodate et à la biotine hydrazide.

B : Le marquage à la biotine est effectué directement sur la membrane de PVDF issue d'un Western Blot. Sur le Western Blot ont été déposés de gauche à droite : le marqueur de poids moléculaire, le core protéique d'endocan (EP11), l'endocan entier (EP10 = témoin positif), un anticorps (EP236 = autre témoin positif). La détection aux bons poids moléculaires des protéines est indiquée par des crochets bleus.

L'analyse en GC permet de quantifier et d'évaluer les rapports molaires entre les différents monosaccharides. La GC-MS confirme l'identification des monosaccharides.

#### 2.2.2.2. Résultats

Dans le but d'identifier les modifications responsables des variations de poids moléculaire (poids observé en spectrométrie de masse versus poids théoriques, i.e. 18106 Da à 22500 g/mol), plusieurs expériences ont été réalisées sur la partie protéique d'endocan

##### Marquage du mutant d'endocan à la biotine par ses glycanes putatifs

Tout d'abord, l'idée était, comme cela est couramment réalisé sur des glycoprotéines, de marquer les glycosylations du core protéique d'endocan à la biotine (où un anticorps jouant le rôle de glycoprotéine témoin). L'utilisation de périodate de sodium permet une oxydation des alcools vicinaux des oses constituant la glycosylation. Les aldéhydes ainsi formés constituent alors de bonnes cibles pour qu'une molécule d'hydrazide liée par un bras espaceur à une biotine (utilisée comme marqueur) vienne réagir et se lier de manière covalente.

Dans un premier temps, l'expérience a été réalisée en solution, le core protéique d'endocan ainsi marqué à la biotine (ou non) a été analysé par Western Blot et révélé soit avec des anticorps spécifiques de la partie protéique, soit avec l'extravidine peroxydase (Horse Radish Peroxydase) qui révèle la biotine uniquement (Figure 20A). L'immunorévélation permet d'observer que l'oxydation au périodate crée un changement dans le profil de migration du core protéique. En effet, lorsque celui-ci a subi l'oxydation, une migration sous forme de traînée (ou "smear") apparaît avec une bande principale légèrement plus haute que le core protéique qui n'a subi que le traitement à la biotine-hydrazide. Lors de la révélation à l'extravidine, seulement la "traînée" est détectée. Cela peut laisser penser que le marquage des glycosylations à la biotine a fonctionné et donc qu'il y a bien des glycosylations présentes sur le core protéique mutant d'endocan. Cependant, le marquage n'est pas assez net pour le certifier et valider.

C'est pourquoi, afin d'éviter tout problème de migration en SDS-PAGE, le marquage a été effectué directement sur la membrane de PVDF sur laquelle les protéines ont été transférées (Figure 20B). Pour ce marquage, un anticorps utilisé en tant que glycoprotéine témoin, l'endocan et le corps protéique d'endocan ont été traités. La révélation a été faite par l'extravidine-HRP uniquement. La membrane traitée à la biotine hydrazide mais non oxydée par le périodate au préalable ne présente aucun signal. En revanche, celle traitée de manière à marquer les glycosylations (Figure 20B) révèle un signal clair pour le témoin, c'est-à-dire



l'anticorps. Un signal est observé sur la traînée correspondant à la migration classique du PG endocan. Par ailleurs, un signal de 24 kDa sur la piste du core protéique d'endocan est présent, ce qui laisse donc bien penser à des glycosylations éventuelles sur ce core protéique.

Ces résultats ont conduit à suivre davantage la piste des glycosylations supposées sur le mutant d'endocan. Pour cela, un travail de traitement de l'endocan mutant par des enzymes de déglycosylation telles que la O-glycosydase et la N-glycosydase F a été mis en œuvre.

#### Déglycosylation des glycanes putatifs de l'endocan mutant

Comme il a été souligné précédemment, le mutant d'endocan en spectrométrie de masse fait valoir un poids moléculaire de 20400 +/- 2200 Da pour un poids moléculaire théorique de 18106 Da. L'objectif de l'expérience était d'utiliser les enzymes de déglycosylation sur le mutant d'endocan et d'analyser par spectrométrie de masse, les variations potentielles de son poids moléculaire. Comme le montrent les spectres de masse des Figures 21A et B, l'utilisation de la O-glycosydase permet de faire se déplacer légèrement le large pic vers une taille moyenne plus proche de celle du core protéique. L'utilisation de la N-glycosidase, en revanche, ne crée aucune modification du temps de vol de l'endocan muté (Figure 21C). L'utilisation combinée des deux enzymes de déglycosylation donne un résultat semblable à celui obtenu avec la O-glycosydase seule.

#### Identification des sites putatifs de glycosylation à l'aide d'endoprotéinases

Afin de compléter l'analyse sur les éventuels sites de glycosylation d'endocan, une dernière série d'expériences a été entreprise. Le core protéique d'endocan a alors été fragmenté de manière enzymatique (trypsine, endoprotéinase Glu-C ou Lys-C) et les fragments peptidiques obtenus ont été comparés avec le profil de la digestion théorique. Ce travail peut permettre d'identifier quels sont les résidus, la séquence peptidique ou le résidu subissant les modifications post-traductionnelles (glycosylations ou autres). En effet, si certains résidus ne sont pas retrouvés après digestion, et si une déglycosylation permet de les retrouver à leur poids moléculaire théorique, la séquence de glycosylation sera ainsi identifiée.

Par ailleurs, cette procédure en conditions non réductrices, permet d'analyser les ponts disulfures qui peuvent se former entre les différents fragments (par l'obtention d'un pic d'une taille de deux fragments contenant des cystéines).

Masse	a. a.	Séquence peptidique	Digestion à la trypsine	Digestion à la trypsine et traitement à la O-glycosydase.
2185.8372	1-19	W <b>S</b> NNYAVD <b>C</b> PQH <b>C</b> D <b>S</b> SE <b>C</b> K	NON	<b>OUI</b>
446.2358	20-23	<b>S</b> SPR	Trop petit	Trop petit
250.1220	24-25	<b>C</b> K	Trop petit	Trop petit
175.1189	26-26	R	Trop petit	Trop petit
1084.4220	27-36	<b>T</b> VLDD <b>C</b> G <b>C</b> CR	NON	NON
576.2922	37-42	<b>V</b> CAAGR	Faible	NON
728.3032	43-48	<b>G</b> ET <b>C</b> YR	OUI	OUI
925.4117	49-57	<b>T</b> V <b>S</b> GMDGMK	NON	NON
602.3079	58-63	<b>C</b> GPGLR	OUI (ou 145-159)	OUI (ou 145-159)
1956.8102	64-81	<b>C</b> Q <b>P</b> S <b>N</b> GEDPFGE <b>E</b> FG <b>I</b> CK	Faible	OUI
1364.5068	82-93	<b>D</b> CPY <b>G</b> TFGMDCR	OUI	OUI
1328.5028	94-105	<b>E</b> TC <b>N</b> C <b>Q</b> S <b>G</b> ICDR	OUI	OUI
362.2034	106-109	<b>G</b> TK	Trop petit	Trop petit
363.2060	110-112	<b>C</b> LK	Trop petit	Trop petit
1263.6408	113-122	F <b>P</b> FFQ <b>S</b> VTK	NON	Très légère apparition
463.2259	123-126	<b>S</b> SNR	Trop petit	Trop petit
1931.9279	127-144	F <b>V</b> SLTEHDMAAGDGNIVR	Faible	<b>OUI</b>
603.3348	145-149	EEVVK	OUI (ou 58-63)	OUI (ou 58-63)
1031.4938	150-159	ENAAG <b>S</b> PVMR	OUI	OUI
147.1128	160-160	K	Trop petit	Trop petit
685.3780	161-165	WLNPR	OUI	OUI

**Figure 22: Tableau récapitulatif des fragments polypeptidiques d'endocan, théoriques ou expérimentaux, qui doivent être retrouvés après une digestion à la trypsine**

La trypsine coupe les chaînes peptidiques après les lysines et arginines. La troisième colonne du tableau montre les séquences peptidiques qui doivent être obtenues après une telle digestion. La seconde colonne donne les positions des premiers et derniers acides aminés de ces séquences et la première colonne donne leur masse. La quatrième colonne indique si oui ou non, sur le spectre de masse les fragments correspondants ont été retrouvés, de même pour la cinquième colonne où le spectre de masse a été obtenu après un traitement à la trypsine et à la O-glycosydase.

### 1<sup>ère</sup> étape : digestion à la trypsine

Le core protéique d'endocan est tout d'abord soumis à une digestion par la trypsine, enzyme qui coupe la protéine après toutes les lysines et toutes les arginines. Les fragments théoriques qui doivent être obtenus sont indiqués dans le tableau de la Figure 22. Dans la colonne 4, sont indiqués les fragments observés en spectrométrie de masse. Les fragments d'une taille inférieure à 550 Da n'ont cependant pas pu être étudiés. Plusieurs pics sur le spectre correspondent ainsi directement aux fragments attendus. Clairement des glycosylations ne peuvent pas intervenir sur des fragments dont la taille théorique correspond à celle obtenue sur le spectre expérimental. Or, les séquences peptidiques 58-81, 82-109, 110-112 et 123-149, malgré une faible apparition, n'ont pu être identifiées clairement en l'absence de traitement par la glycosydase. Ainsi, les sérines 67, 123, 124 et 129, et les thréonines 107 et 131 peuvent donc potentiellement être glycosylées. Le spectre du mutant d'endocan traité à la trypsine est ensuite comparé à celui du même échantillon traité en plus à la O-glycosydase (Figure 22 colonne 5). La variation la plus évidente est l'apparition du pic correspondant à la séquence 1-19. L'augmentation du pic de la séquence 127-145 peut également être pris en compte, en revanche la faible apparition de la séquence 113-122 n'est sans doute pas significative.

D'autre part, des pics ne correspondant pas à un fragment unique ont été obtenus. Par exemple, un signal est obtenu à 1687,95 Da. Il peut correspondre à la liaison entre les fragments 27-36 et 58-63 voire même celle entre les fragments 94-105 et 110-112. De même à 2684 Da, pour les fragments 43-48 et 64-81, ou à 3265 Da pour les fragments 1-19 et 27-36. Ces liaisons sont dues aux ponts disulfures entre deux cystéines.

### 2<sup>ème</sup> étape : digestion à l'endoprotéinase Lys-C

La digestion tryptique crée de nombreux fragments de petite taille. L'endoprotéinase Lys-C induit uniquement des coupures au niveau des résidus de lysine et peut théoriquement conduire à la formation de dix fragments. La digestion en l'absence et en présence d'agent réducteur (DTT) a été réalisée. Sur le même principe d'analyse que pour la digestion tryptique, le tableau présenté en Figure 23 indique les masses attendues pour chacun de ces dix fragments théoriques, les conditions réductrices utilisées et bien sûr la présence des fragments expérimentaux. En présence de l'agent réducteur, seuls les fragments 58-81, 82-109, 110-112, et 123-149 n'ont pas donné de signal évident. Ce résultat implique que les sérines 67, 123, 124 et 129, et les thréonines 107 et 131 peuvent potentiellement être glycosylées.

masse	a.a.	Séquence peptidique	Lys-C	Lys-C + DTT
2185.8372	1-19	W <b>S</b> N <b>N</b> Y <b>A</b> V <b>D</b> C <b>P</b> Q <b>H</b> C <b>D</b> <b>S</b> <b>S</b> E <b>C</b> K	NON	OUI
677.3399	20-25	<b>S</b> S <b>P</b> R <b>C</b> K	OUI	OUI
3413.4768	26-57	R <b>T</b> V <b>L</b> D <b>D</b> C <b>G</b> C <b>C</b> R <b>V</b> C <b>A</b> A <b>G</b> R <b>G</b> E <b>T</b> <b>C</b> <b>Y</b> R <b>T</b> V <b>S</b> G <b>M</b> D <b>G</b> M <b>K</b>	NON	OUI
2540.1002	58-81	<b>C</b> G <b>P</b> G <b>L</b> R <b>C</b> Q <b>P</b> S <b>N</b> G <b>E</b> D <b>P</b> F <b>G</b> E <b>E</b> F <b>G</b> I <b>C</b> K	NON	Faible
3017.1773	82-109	D <b>C</b> P <b>Y</b> G <b>T</b> F <b>G</b> M <b>D</b> C <b>R</b> E <b>T</b> C <b>N</b> C <b>Q</b> S <b>G</b> <b>I</b> C <b>D</b> R <b>G</b> T <b>G</b> K	NON	Faible
363.2060	110-112	<b>C</b> L <b>K</b>	Trop petit	Trop petit
1263.6408	113-122	F <b>P</b> F <b>F</b> Q <b>Y</b> S <b>V</b> T <b>K</b>	OUI	OUI
2960.4530	123-149	<b>S</b> S <b>N</b> R <b>F</b> V <b>S</b> L <b>T</b> E <b>H</b> D <b>M</b> A <b>A</b> G <b>D</b> G <b>N</b> I <b>V</b> R <b>E</b> E <b>V</b> V <b>K</b>	OUI	Faible
1159.5888	150-160	E <b>N</b> A <b>A</b> G <b>S</b> P <b>V</b> M <b>R</b> K	OUI	OUI
685.3780	161-165	W <b>L</b> N <b>P</b> R	OUI	OUI

**Figure 23: Tableau récapitulatif des fragments polypeptidiques d'endocan, théoriques ou expérimentaux, qui doivent être retrouvés après une digestion à l'endoprotéinase Lys-C**

L'endoprotéinase Lys-C coupe les chaînes peptidiques après les lysines. La troisième colonne du tableau montre les séquences peptidiques qui doivent être obtenues après une telle digestion. La seconde colonne donne les positions des premiers et derniers acides aminés de ces séquences et la première colonne donne leur masse. La quatrième colonne indique si oui ou non, sur le spectre de masse les fragments correspondants ont été retrouvés. De même pour la cinquième colonne où le spectre de masse a été obtenu après un traitement à la Lys-C et au DTT (agent réducteur ayant pour but de rompre les liaisons disulfures).

Masse	a.a.	Séquence peptidique	Glu-C	Glu-C + DTT
1954.7330	1-17	W <b>S</b> N <b>N</b> Y <b>A</b> V <b>D</b> C <b>P</b> Q <b>H</b> C <b>D</b> <b>S</b> <b>S</b> E	NON	OUI
2873.2878	18-44	<b>C</b> K <b>S</b> S <b>P</b> R <b>C</b> K <b>R</b> T <b>V</b> L <b>D</b> D <b>C</b> G <b>C</b> C <b>R</b> V <b>C</b> A <b>A</b> G <b>R</b> G <b>E</b>	NON	OUI
2747.1826	45-70	<b>T</b> C <b>Y</b> R <b>T</b> V <b>S</b> G <b>M</b> D <b>G</b> M <b>K</b> <b>C</b> G <b>P</b> G <b>L</b> R <b>C</b> <b>Q</b> P <b>S</b> N <b>G</b> E	NON	NON
693.2726	71-76	<b>D</b> P <b>F</b> G <b>E</b> E	Trop petit	Trop petit
2041.8275	77-94	<b>F</b> G <b>I</b> C <b>K</b> D <b>C</b> P <b>Y</b> G <b>T</b> F <b>G</b> M <b>D</b> C <b>R</b> E	NON	OUI
4252.0082	95-132	<b>T</b> C <b>N</b> C <b>Q</b> S <b>G</b> <b>I</b> C <b>D</b> R <b>G</b> T <b>G</b> K <b>L</b> K <b>F</b> P <b>F</b> F <b>Q</b> Y <b>S</b> V <b>T</b> K <b>S</b> S <b>N</b> R <b>F</b> V <b>S</b> L <b>T</b> E	NON	NON
1513.6699	133-146	<b>H</b> D <b>M</b> A <b>A</b> G <b>D</b> G <b>N</b> I <b>V</b> R <b>E</b> E	OUI	OUI
474.2922	147-150	<b>V</b> V <b>K</b> E	Trop petit	Trop petit
1696.9064	151-165	<b>N</b> A <b>A</b> G <b>S</b> P <b>V</b> M <b>R</b> K <b>W</b> L <b>N</b> P <b>R</b>	OUI	OUI

**Figure 24: Tableau récapitulatif des fragments polypeptidiques d'endocan, théoriques ou expérimentaux, qui doivent être retrouvés après une digestion à l'endoprotéinase Glu-C**

L'endoprotéinase Glu-C coupe les chaînes peptidiques après les acides glutamiques. La troisième colonne du tableau montre les séquences peptidiques qui doivent être obtenues après une telle digestion. La seconde colonne donne les positions des premiers et derniers acides aminés de ces séquences et la première colonne donne leur masse. La quatrième colonne indique si oui ou non, sur le spectre de masse les fragments correspondants ont été retrouvés. De même pour la cinquième colonne où le spectre de masse a été obtenu après un traitement à la Glu-C et au DTT (agent réducteur ayant pour but de rompre les liaisons disulfures).

En comparant l'échantillon traité au DTT et celui permettant la formation des ponts disulfures, il est donc possible de supposer que les fragments 1-19 et 26-57 sont reliés par ce type de liaison, d'autant plus qu'un pic correspondant à la somme de ces deux fragments a été obtenu sur l'échantillon non réduit. La même analyse peut être faite concernant une interaction par un pont disulfure entre les fragments 26-57 et 58-81, et de manière encore plus nette entre les fragments 58-81 et 82-109.

### 3<sup>ème</sup> étape : digestion à l'endoprotéinase Glu-C

Afin d'augmenter le nombre de combinaisons possibles, une digestion à l'endoprotéinase Glu-C a été réalisée. Cette enzyme coupe la liaison peptidique au niveau des acides glutamiques. Comme pour la Lys-C, l'échantillon a été traité en présence ou en absence de DTT. Le tableau de la Figure 24 compare les fragments théoriques et les fragments obtenus expérimentalement. Ainsi, en conditions réductrices, les séquences 45-70, 71-76, 95-132 et 147-10 manquent sur le spectre. Ce résultat n'apporte pas d'information supplémentaire concernant les glycosylations potentielles. La comparaison avec les conditions non réductrices confirme que les fragments 1-17, 18-44 et 77-94 sont impliqués dans des ponts disulfures. Pour le fragment 1-17 cela signifie que le pont disulfure 9-S-S-13 est impossible puisque ce fragment ne contient que deux cystéines et que l'une est impliquée dans un pont disulfure.

### Méthanolyse du core protéique d'endocan et identification des monosaccharides libérés

Les glycannes putatifs ont été séparés du core protéique d'endocan par méthanolyse. Après un marquage des monosaccharides qui permet leur suivi, une analyse HPLC a été réalisée. La lysine est ajoutée comme une référence pour la quantification. Le tableau de la Figure 25 détaille la quantité et la nature des monosaccharides détectés. Ainsi, il apparaît que le core protéique d'endocan porte environ 1 % de glycannes (3 µg pour 300 µg). La présence de galactose (Gal), de N-acétyl galactosamine (GalNAc), de N-acétyl glucosamine (GlcNAc) et d'acide N-acétyl neuraminique (NeuAc) confirmée en GC-MS (résultats non montrés) révèle une O-glycosylation. Aussi, la présence de mannose laisse supposer la présence de N-glycane. Pour les O-glycannes, le rapport molaire se calcule à partir de la GalNAc en position réductrice. Ainsi, à une mole de GalNAc correspond 0,857 mole de Gal, 0,439 mole de GlcNAc et 0,671 mole de NeuAc.

Monosaccharide	Fuc	Gal	Man	GalNAc	GlcNAc	NeuAc	Masse totale
Masse moléculaire	164	180	180	221	221	309	
Coefficient/lysine	0.99	0.894	1.730	0.790	1.560	2.5	
Surface des pics	0	44574	10232	45917	39809	97580	
S.pic/S. Lys.	0	1,2601	0,28925	1,29804	1,12537	2,7585	
Diviser par le coeff.		1,4095	0,16720	1,64309	0,72139	1,1034	
Nbre de nanomoles		3.8761	0.45979	4.51850	1.98383	3.0343	
Rapport molaire		0.857		1	0.439	0.671	
Masse en $\mu\text{g}$		0.6976	0.0827	0.9985	0.4384	0.9376	3.155

**Figure 25: Composition monosaccharidique des O- et N-glycannes liés sur le core protéique d'endocan**

Les glycannes présents sur le core protéique d'endocan (300  $\mu\text{g}$ ) sont hydrolysés par méthanolyse (méthanol/HCl à 80°C). Les monosaccharides obtenus sont dérivés en heptafluorobutyrate et analysés en chromatographie en phase gazeuse (GC), couplée ou non à la spectrométrie de masse (GC-MS). Les résultats de la chromatographie permettent d'évaluer les rapports molaires entre les différents monosaccharides. Pour estimer la quantité de monosaccharide libéré, la lysine est utilisée comme référence. Dans cette expérience 0,5  $\mu\text{g}$  de lysine ont été utilisés. Il existe un coefficient entre chaque monosaccharide est la lysine (coefficient/lysine). Il permet par la comparaison des aires obtenues sur le chromatogramme, de connaître les rapports massiques entre chacun d'entre eux. L'aire obtenue pour la lysine dans cette expérience est de 35374. Il est possible que l'analyse sous-estime la quantité de glycanne détecté mais le core protéique d'endocan porterait au moins 1 % de glycanne (O- ou N-glycanne) (3  $\mu\text{g}$  pour 300  $\mu\text{g}$ ).

### 2.2.3. Cristallographie

Dans le but d'obtenir des informations structurales sur le core protéique d'endocan, des expériences de cristallographie ont été entreprises. D'une manière générale, l'objectif était de trouver les conditions pour permettre la formation de cristaux qui, par diffraction aux rayons X, auraient pu donner des informations sur l'organisation structurale d'endocan. Pour cela, un premier criblage des conditions doit être réalisé. Ce premier criblage fait intervenir des conditions de cristallisation (composition des tampons) très variées. Ainsi, si des premiers cristaux se forment dans l'une des conditions, un second criblage autour de cette condition peut être réalisé (variation du pH, ou de la concentration des précipitants par exemple).

#### 2.2.3.1. Matériel et méthodes

##### Production et purification du core protéique d'endocan

Le core protéique d'endocan est obtenu selon les méthodes déjà détaillées dans la première partie des résultats.

##### Production et purification des Fab d'anticorps monoclonaux dirigés contre endocan

L'anticorps monoclonal EP214 a été produit et purifié par une collaboratrice, Estelle Adam, au sein de l'entreprise Endotis Pharma (Loos). La purification du Fab de cet anticorps a été réalisée selon le protocole fourni avec le kit de purification développé par Pierce : "ImmunoPure Fab Preparation Kit". Cette purification est faite en incubant l'anticorps avec de la papaïne afin de détacher le Fab et la partie Fc de l'anticorps. Cette dernière est retenue grâce à une colonne de Protéine A, et le Fab non retenu sur ce type de colonne est récupéré. Le protocole utilisé est rigoureusement identique à celui fourni par Pierce avec le kit, sans aucune modification.

##### Création du complexe core protéique d'endocan / Fab

L'endocan mutant (300  $\mu$ L à 22 mg/mL) et le Fab de l'anticorps EP214 (300  $\mu$ L soit 6 mg) sont indépendamment passés en filtration sur gel (Superdex 75, Amersham, tampon PBS  $\frac{1}{4}$ ) à 1 mL/min. Les protéines sont collectées dans un volume de 1,5 mL. Le complexe est formé entre le Fab et l'endocan mutant (3 mL final) durant la nuit à 4°C. Il est ensuite concentré sur Amicon Ultra-4 (10 MWCO) jusqu'à 350  $\mu$ L et est injecté sur une colonne de filtration sur gel. Un gel SDS-PAGE révèle que le pic majoritaire correspond au complexe endocan/Fab. Les fractions correspondant à ce pic sont rassemblées et concentrées puis

déssalées contre un tampon Tris-HCl pH 7.8, 150 mM NaCl. Au final, 200 µL de complexe endocan/ Fab à 5 mg/mL ont été obtenus.

#### Criblage des conditions de cristallisation

##### Core protéique d'endocan seul

Plusieurs criblages ont été effectués à différentes concentrations d'endocan, selon la disponibilité du matériel.

*Premier criblage* : Il a été effectué par le robot de cristallisation de l'EMBL (Cartesian) capable de faire des nanogouttes de cristallisation d'environ 200 nL. De l'endocan à 3,45 mg/mL a été utilisé pour ces expériences. Il a été placé, de manière robotisée, en présence des agents précipitants des kits de cristallisation : Crystal Screen I et II, Crystal Screen Lite, Peg-ion, Memb Fac, Natrix, Ammonium sulfate, Screen Malonate et Formate, Quick Screen, PEG 6k, MPD, PEG Li-Cl, Mme 5000 et 96 autres conditions (Hampton Research), à 20°C, soit au total 600 conditions testées.

*Second criblage* : Le mutant d'endocan a été préparé à 22 mg/mL et le criblage a été effectué par le robot de cristallisation de l'IBS (TECAN) sur les kits de cristallisation Crystal Screen kits I and II (Hampton Research, CA), aux températures de 4°C et 20°C.

*Troisième criblage* : Une nouvelle production du mutant d'endocan a fourni un échantillon protéique à 26,7 mg/mL. Les conditions de criblage des kits Crystal Screen I et II à 4°C et 20°C ont été utilisées grâce au robot disponible à l'IBS.

##### Complexe endocan / Fab

Un seul criblage a été réalisé. Celui-ci a été effectué avec le robot de l'EMBL (Grenoble) en utilisant les kits Crystal Screen kits I and II (Hampton Research, CA), à 4°C ou 20°C et sur un échantillon de complexe protéique à la concentration d'environ 5 mg/mL.

### **2.2.3.2. Résultats**

Afin de pouvoir structurellement caractériser le core protéique d'endocan, une étude par cristallographie a été engagée. Le mutant d'endocan a été produit et purifié à partir d'un système eucaryote (les cellules HEK293) avec une mutation sur la sérine 137, afin que la chaîne de DS ne puisse plus être ajoutée au core protéique. Ce mutant avait l'avantage de fournir un échantillon protéique de composition théoriquement homogène (qu'un PG entier ou qu'un PG auquel on aurait supprimé les GAGs par dégradation chimique ou enzymatique). La première phase de l'étude cristallographique a consisté à identifier quels pouvaient être les précipitants les plus favorables. Durant cette première phase, différents facteurs ont pu être

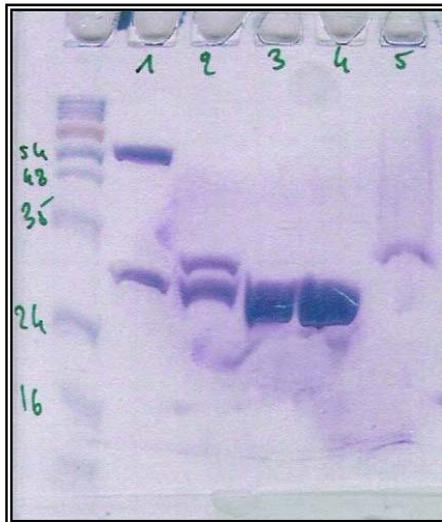
évalués, malheureusement, aucun des cristaux obtenus n'a permis de générer la moindre carte de diffraction.

#### *Cristallisation du core protéique d'endocan*

Trois essais ont pu être effectués. Le premier essai est celui qui a permis le plus large criblage de conditions de cristallisation de manière automatisée. Malheureusement, l'échantillon était à 3,5 mg/mL, une concentration certainement trop faible puisque, sur les 600 conditions testées, très peu ont fait précipiter la protéine, signe que la concentration en protéine utilisée n'était certainement pas assez élevée. Des cristaux se sont cependant formés dans six conditions contenant du phosphate comme agent précipitant, mais aucun n'était de nature protéique.

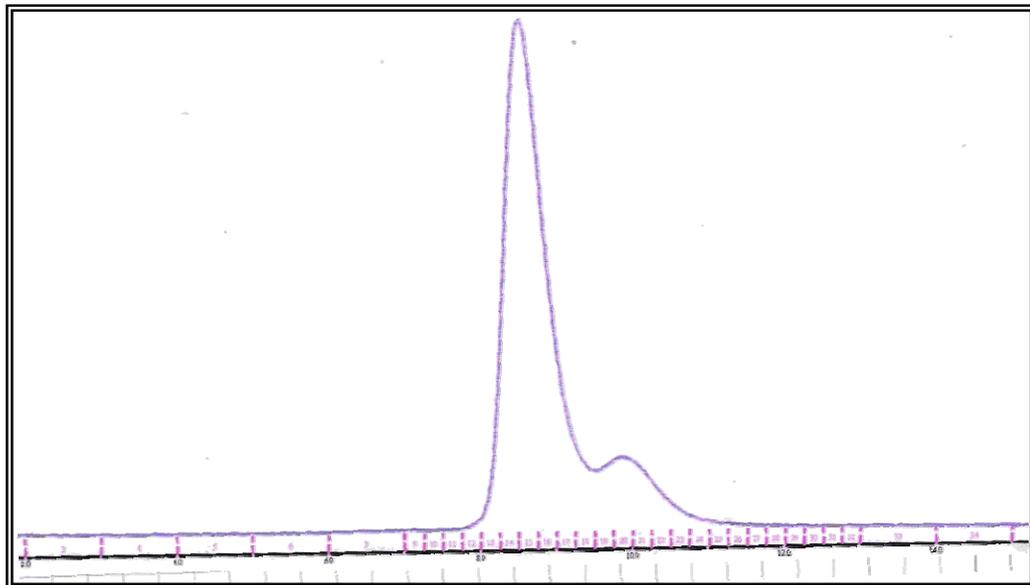
A la suite de ces premiers résultats, le second essai a été réalisé à une concentration en protéine plus élevée (22 mg/mL), sur une gamme de conditions cristallographiques moins importante, mais variée. Les échantillons ont été placés à des températures différentes : soit à 4°C, soit à 20°C, puisque la température est une variable importante pour certaines protéines, notamment celles qui se dégradent à température ambiante. Les gouttes de cristallisation ont été réalisées par le robot de l'IBS qui fait des gouttes de taille plus importante (1 µL) que celui de l'EMBL (200 nL), ce qui peut faciliter, dans certains cas, la formation des cristaux. Ce second essai a fait précipiter la protéine dans environ 50 % des puits. Dans certaines conditions et à 20°C uniquement, des cristaux se sont formés. Ceux-ci ont été isolés, plongés dans l'azote liquide afin d'être analysés sur une ligne de rayonnement synchrotron de l'ESRF. Pour la plupart, les signaux observés ont révélé que les cristaux seraient plutôt de nature saline. En revanche, dans la condition suivante : 0,2 M d'acétate d'ammonium et 0,1 M d'acétate de sodium trihydrate, pH 4.6, 30 % PEG 4000, les cristaux avaient une apparence protéique, mais malheureusement étant trop fragiles, ils n'ont pas pu être isolés et analysés au synchrotron. Cette condition avait déjà donné lieu à la formation de petits cristaux lors du premier essai. Elle était, dans ce premier essai, la seule condition expérimentale à ne pas contenir de phosphate, se détachant ainsi des autres conditions où des cristaux avaient pu être observés. A 4°C, aucune condition n'a permis d'obtenir de cristal, même si le taux de précipitation de la protéine semblait identique.

Un troisième essai, dans des conditions identiques à celle du second essai mais à une concentration d'endocan mutant encore plus élevée (26,7 mg/mL) est en toujours en cours. Pour l'instant, aucun cristal n'a été obtenu que ce soit à 4°C ou à 20°C.



**Figure 26: Suivi de la purification du Fab de l'anticorps EP214 par gel SDS-PAGE coloré au bleu de Coomassie**

Le Fab est séparé de la partie Fc par une coupure à la papaine. La partie Fc est ensuite retenue sur une colonne Protéine A, ce qui permet de récupérer le Fab qui lui n'est pas retenu. Pour suivre la purification, le gel SDS-PAGE présenté est réalisé. A droite du marqueur de poids moléculaire ont été déposés : l'anticorps EP214 (10 µg) (1), EP214 après digestion à la papain (10 µg) (2), le Fab récupéré (20 µg = 3 et 60 µg = 4), et la partie Fc éluée de la colonne protéine A (10 µg) (5).



**Figure 27: Chromatographie d'exclusion du complexe core protéique d'endocan/Fab d'EP214**

Pour obtenir un complexe pur entre la partie protéique d'endocan et le Fab d'EP214, les deux polypeptides sont mis en présence, avec un léger excès d'endocan. Pour supprimer l'excès d'endocan, le mélange est soumis à une filtration sur gel de type superdex 75 dans du PBS. Le chromatogramme obtenu montre un pic majoritaire (complexe = 70 kDa) et un épaulement de poids moléculaire inférieur (partie protéique d'endocan = 20 kDa).

### Cristallisation du complexe protéique endocan mutant / Fab

L'endocan est une protéine soluble d'environ 18,1 kDa. D'après l'analyse protéique réalisée après purification, le mutant apparaissait quasiment pur. Ces caractéristiques intéressantes auraient pu laisser penser que la cristallisation allait être facilitée. Cependant, les essais réalisés sur le mutant d'endocan n'ont été que très peu fructueux. Plusieurs facteurs peuvent rendre une cristallisation difficile. Tout d'abord, il faut un échantillon très homogène. Or comme on vient de le voir, l'analyse par spectrométrie de masse laisse penser que ce corps protéique est porteur de modifications post-traductionnelles qui font varier sa taille de 4000 Da. De plus, les régions non structurées d'une protéine peuvent la rendre flexible à certains endroits et ainsi l'empêcher de s'organiser en un réseau régulier nécessaire pour la cristallisation. Les prédictions bioinformatiques que nous avons réalisées concernant le désordre au sein de la protéine ont révélé que les 33 derniers acides aminés d'endocan sont organisés de façon peu structurée. C'est pourquoi, il a été entrepris des travaux de co-cristallisation de l'endocan mutant avec le Fab de l'anticorps monoclonal EP214 dont l'épitope se trouve précisément dans la région C-terminale que l'on souhaitait stabiliser. La Figure 26 présente le gel SDS-PAGE qui a permis de vérifier la pureté du Fab d'EP214 produit. Le complexe qui a été formé entre le core protéique d'endocan et le Fab a été purifié par filtration sur gel dont le chromatogramme est présenté en Figure 27. Le complexe ainsi purifié a été utilisé à 5 mg/mL pour cribler 200 conditions à 4°C et à 20°C. Aucun cristal ne s'est pour l'instant formé.

#### **2.2.4. Discussion**

L'endocan est un PG encore bien mal connu et le but essentiel de ce projet de recherche était de tenter d'en apprendre d'avantage sur la structure de cette molécule.

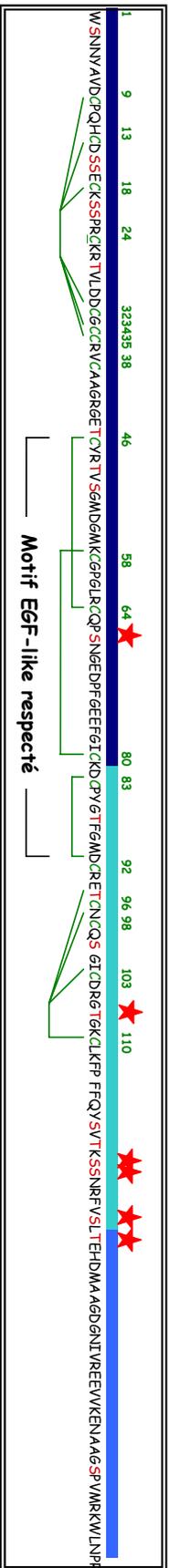
Comme nous l'avons vérifié, l'endocan est composé d'un corps protéique et d'une unique chaîne de DS liée à une sérine en position 137 (Bécharde et al., 2001a). Un mutant de ce PG circulant avait été précédemment développé, ne possédant plus de chaîne de GAG puisque la sérine 137 a été remplacée par mutagenèse par une alanine. L'obtention de quantité raisonnable de ce mutant par les techniques de production et de purification développées, a permis d'initier des expériences de cristallographie de manière à essayer de donner une structure protéique à ce PG circulant. Malgré le large criblage cristallographique effectué (1800 conditions au total), plusieurs éléments peuvent expliquer pourquoi cette petite protéine pourtant dépourvue de sa chaîne de GAG n'a pas permis l'organisation d'un cristal jusqu'à

présent. Dans le cas de l'endocan mutant, la purification par passage sur une colonne d'immunoaffinité a permis d'avoir une protéine pure. D'un point de vue de l'homogénéité de l'échantillon, la spectrométrie de masse ainsi que la visualisation de ce mutant sur gel SDS-PAGE en conditions non réductrices, révèle un poids moléculaire moyen de 20 000 Da environ (+/- 2000 Da) pour 18106 attendu. Cette variation de poids moléculaire, de nature inconnue, est un signe d'hétérogénéité, souvent peu favorable à la cristallisation. Ces éléments nous ont donc poussé à caractériser le core protéique d'endocan d'une façon différente que par diffraction aux rayons X.

Dans un premier temps, une approche bioinformatique a permis de se familiariser avec l'endocan, ses caractéristiques biochimiques, l'organisation structurale qu'il pourrait adopter et ses potentielles modifications post-traductionnelles. De cette manière, comme cela a déjà été analysé par des travaux antérieurs (Lassalle et al., 1996), aucune séquence consensus pour un éventuel site de N-glycosylation n'apparaît clairement même si quatre sites potentiels ne sont pas à écarter. La sérine 137 a bien été définie comme étant un site préférentiel pour l'attachement d'un GAG. Par contre, aucun logiciel ne permet d'évaluer réellement les sites potentiels de O-glycosylations chez l'Homme. Parallèlement, plusieurs sites potentiels de phosphorylations ont pu être observés (Figure 1B), certains étant théoriquement plus spécifiques à la caséine kinase II (CKII) d'autres plutôt à la protéine kinase C (PKC) (Sarrazin et al., 2006). Toujours selon cette analyse bioinformatique, une séquence consensus de sumoylation a été jugée probable sur la séquence 110-113 (80 % de probabilité). De plus, la protéine est composée de 18 cystéines dans les 110 premiers acides aminés (165 a.a. au total), ce qui laisse fort à penser que ces 110 acides aminés N-terminaux forment une région assez dense, structurée et plutôt globulaire à l'inverse de la région C-terminale qui apparaît certainement plus labile. Cette analyse de l'organisation de la protéine est confortée par les résultats d'une évaluation du désordre au sein de la protéine. En effet, un logiciel de la prédiction des régions non-structurées des protéines a permis d'évaluer qu'endocan était une protéine globulaire jusqu'à l'acide aminé 132 puis que les 33 derniers acides aminés formaient une région très peu structurée. Une recherche d'homologues protéiques par comparaison de séquence (BLAST) a également été effectuée. Aucune protéine connue jusqu'à maintenant ne possède un profil peptidique identique à celui d'endocan. Cependant, les protéines qui comportent des similarités avec endocan ont toutes en commun un motif de type EGF-like. Ce motif, décrit Figure 16, est principalement caractérisé par la formation de ponts disulfures et semble cohérent au regard du nombre de cystéines dans cette région (a.a. 46-92) de l'endocan.

Expérimentalement, la caractérisation de l'organisation de la protéine (pont disulfures principalement) et l'étude d'éventuelles modifications post-traductionnelles ont été réalisées principalement par spectrométrie de masse. La variation du poids moléculaire de 4000 Da nous a orienté préférentiellement vers les glycosylations qui induisent, en règle générale, un ajout de poids moléculaire d'au moins 800 Da. L'ubiquitination constitue également un ajout d'au moins 800 Da, et malgré la détection d'un site potentiel de sumoylation cette piste n'a pas été suivie. Les phosphorylations en revanche ne rajoutent que 80 Da, la variation de poids moléculaire observée ne peut donc pas leur être attribuée mais cela n'exclut cependant pas que l'endocan ne soit pas sujet à ce genre de modification.

Une première expérience de marquage des glycanes putatifs à la biotine (facilitant ainsi la détection par chimioluminescence) a révélé que le core protéique d'endocan pouvait également comporter d'autres oses. Un premier traitement à la N-glycosidase, suivi d'une analyse en spectrométrie de masse, semble rendre peu probable la présence de N-glycosylation sur endocan. La même expérience avec la O-glycosydase a permis d'observer un rétrécissement du pic d'environ 800 à 1000 Da (Figure 21), cependant le signal spectrométrique reste toujours assez large. Des cartes peptidiques ont alors été construites après digestion par les endoprotéinases (trypsine, Lys-C, et Glu-C), de manière à pouvoir identifier des sites potentiels de modifications post-traductionnelles (ces dernières peuvent en effet modifier le temps de vol des peptides analysés en spectrométrie de masse). La comparaison des spectres obtenus pour l'endocan mutant traité à la trypsine seule ou à la trypsine suivi d'un traitement à la O-glycosydase, a permis de révéler clairement l'apparition du fragment 1-19, ce qui laisse supposer que les sérines 2, 15 et 16 pourraient être O-glycosylées. Concernant les sérines 100 et 155 ainsi que les thréonines 45, 87, et 95, il n'y a pas de chance qu'elles soient O-glycosylées puisque les fragments les contenant ont tous été retrouvés sans l'ajout de glycosydase. Pour les autres sérines et thréonines, la création de fragments de trop petite taille, ou la non apparition des pics sur les spectres ne nous permettent pas de conclure. La digestion par la Lys-C en présence ou non de DTT, créant des fragments de plus grande taille que le traitement avec la trypsine, a permis de remettre en question la présence de glycosylations sur le fragment 1-19. Avec ce traitement, les séquences 58-81, 82-109, 110-112 et 123-149 n'ont pu être clairement identifiées en l'absence de glycosydase. Ainsi, peuvent potentiellement être glycosylées les sérines 67, 123, 124 et 129, et les thréonines 107 et 131. La même expérience réalisée avec la Glu-C a permis d'identifier, en conditions réductrices, tous les fragments sauf les peptides 45-70 et 95-132 ce qui laisse toujours les sérines 67, 123, 124, et 129 et thréonines 107 et 131 comme des sites potentiels



**Figure 28 : Modèle de l'organisation structurale du core protéique d'endocan**

D'après les spectres de masse obtenus après les différentes digestions, en présence ou non de O-glycosydase ou DTT, les sérines 67, 123, 124, et 129 et thréonines 107 et 131 sont de probables sites de O-glycosylation (étoile rouge). Également, les liaisons disulfures potentielles ont été indiquées en vert. Il est probable que les liaisons disulfures d'endocan respectent le motif EGF-like qui avait bio-informatiquement été défini. En compilant toutes les informations, les ponts disulfures les plus probables sont C9-C32 ; C18-C35 ; C46-C64 ; C58-C80 ; C83-C92 ; C96-C103 ; C98-C110.

de glycosylation. De plus, l'absence de pic à 4252 Da pour le fragment 95-132 mais la présence d'un pic large non identifié aux alentours de 4808 et 4887 Da est un signe supplémentaire de glycosylation et même potentiellement de glycosylation et de phosphorylation (+ 80 Da).

Les résultats ainsi obtenus n'étaient cependant pas en mesure d'établir clairement que le core protéique d'endocan est soumis à des N- ou des O-glycosylations. C'est pourquoi, en collaboration avec Marie-Christine Slomianny (laboratoire du Pr Jean-Claude Michalski, Université des Sciences et Technologies de Lille) une étude plus poussée de détermination de la présence de glycosylation a été effectuée. De cette étude, il apparaît que le core protéique d'endocan comporte environ 1 % de glycannes. Parmi ces glycannes, une ou plusieurs O-glycosylations sont présentes, et dans certains cas, il apparaît qu'au moins une N-glycosylation pourrait avoir lieu puisque des mannoses ont été trouvés.

La constitution de la carte peptidique, en l'absence ou en présence de DTT a également permis d'obtenir des informations sur les liaisons disulfures potentielles. En effet, l'utilisation de la trypsine a fait apparaître un fragment à 1687,95 Da qui correspond soit à la liaison entre les fragments 27-36 et 58-63 soit à une liaison des fragments 94-105 et 110-112. Il en est de même à 2558 Da pour les fragments 58-63 et 64-81, à 2684 Da pour les fragments 43-48 et 64-81 ou à 3265 Da pour les fragments 1-19 et 27-36. Le fragment 82-93 n'a pas été retrouvé associé à d'autres fragments peptidiques et a été observé à la bonne taille, laissant donc supposer que ses cystéines 83 et 92 ne peuvent former un pont disulfure uniquement qu'entre elles. Ce constat respecte d'ailleurs le motif EGF-like précédemment identifié lors de l'analyse bioinformatique (Figure 17B). En utilisant l'enzyme Lys-C, en présence ou non d'agent réducteur, un nouveau profil de digestion a été obtenu. Il a permis de vérifier que les fragments 1-19 et 26-57 sont donc bien reliés entre eux par des ponts disulfures, au même titre que les fragments 26-57 et 58-81 ou encore les fragments 58-81 et 82-109. La cystéine 24, en revanche, ne semble pas impliquée dans une liaison de type pont disulfure. L'utilisation de la Glu-C, en conditions réductrices ou non, a également permis de fournir des détails supplémentaires sur l'organisation structurale d'endocan. Par exemple, pour le fragment 1-17, il apparaît évident que le pont disulfure C9-C13 est totalement impossible.

Grâce à tous ces indices, un modèle de l'organisation d'endocan (ponts disulfures et glycosylations) respectant les données qui viennent d'être exposées a été proposé (Figure 28). Ainsi, un ou plusieurs ponts disulfures entre les cystéines 9, 13, 18 et 32, 34, 35 sont à prendre en considération. Sur la séquence protéique 46-92, l'organisation des ponts disulfures respecterait le motif EGF-like qui avait été détecté par homologie de séquence. Un pont

disulfure entre les cystéines 96, 98 et 103 n'est pas à exclure, de même un pont disulfure entre ces mêmes acides aminés et la cystéine 110 reste également possible. Par bioinformatique, certains logiciels (Geno3D, Job CPH, phyre) proposent, par homologies avec des protéines dont la structure a été résolue, de modéliser une protéine à partir de sa séquence peptidique. Cela a été réalisé pour endocan. En respectant les résultats obtenus et les modèles obtenus, les ponts disulfures les plus probables pourraient donc être : C9-C32 ; C18-C35 ; C46-C64 ; C58-C80 ; C83-C92 ; C96-C103 ; C98-C110.

### **2.2.5. Perspectives**

L'organisation structurale du core protéique d'endocan est encore loin d'être résolue. Les résultats obtenus permettent cependant d'ouvrir quelques pistes. Concernant les modifications post-traductionnelles, endocan, outre sa chaîne de DS liée en S137 semble également posséder des glycosylations de type O- et N-glycosylation, l'objectif serait maintenant de les localiser. Parallèlement, il serait intéressant d'approfondir la présence d'éventuelles phosphorylations. D'un point de vue plus structural, l'utilisation d'autres protéases en conditions réductrices ou non pourrait valider les ponts disulfures suggérés. Concernant la cristallisation du core protéique, une seconde phase de criblage autour de la condition : 0,2 M d'acétate d'ammonium et 0,1 M d'acétate de sodium trihydrate, pH 4.6, 30 % PEG 4000 pourrait être entreprise. Pour cela, l'obtention d'endocan nu de toute glycosylation s'avère importante, et la stabilisation avec un Fab une voie à poursuivre.

### **3.INTERACTIONS ENTRE ENDOCAN ET DES PROTEINES D'INTERETS**

Depuis sa découverte, le protéoglycane endocan a suscité de nombreux intérêts au regard de son implication potentielle, notamment dans la progression tumorale ou dans les phénomènes inflammatoires comme, par exemple, la septicémie. Le développement par le laboratoire du Dr. Philippe Lassalle d'anticorps monoclonaux anti-endocan a permis d'engager plusieurs pistes de recherche pour mieux connaître ce nouveau PG. Tout d'abord, grâce à ces anticorps spécifiques, une méthode de dosage par ELISA a été mise au point par ce même laboratoire. Après transfert de ces outils chez Endotis Pharma, ce dosage ELISA devait être perfectionné pour devenir un kit standardisé de dosage d'endocan. Dans le cadre d'une éventuelle mise sur le marché de ce kit, il était important de caractériser plus précisément ces anticorps et notamment de connaître leur affinité, non déterminée au commencement de mon travail de thèse.

Par ailleurs, certains anticorps apparaissaient comme potentiellement thérapeutiques. En effet, chez des souris ayant reçu des cellules tumorales modifiées pour surexprimer endocan, ces anticorps étaient capables de réduire les tumeurs formées (Scherpereel et al., 2003). Dans ce cadre d'applications, il était donc également très important d'évaluer l'affinité pour l'endocan de ces anticorps potentiellement thérapeutiques.

En plus du fait qu'endocan est clairement surexprimé dans certains types de cancers (Pour revue, voir Sarrazin et al., 2006), il a été montré *in vitro* que sa chaîne oligosaccharidique a une influence sur l'activité de facteurs tumoraux comme le HGF/SF (Bécharde et al., 2001). Cela nous a donc logiquement conduit à évaluer les interactions potentielles qu'endocan pouvait avoir avec d'autres molécules impliquées dans les événements cancéreux ou au cours de l'inflammation.

### **3.1. ETUDE D'INTERACTION ENTRE ENDOCAN ET DES ANTICORPS**

L'étude de l'affinité de certains anticorps pour l'endocan a été réalisée par résonance plasmonique de surface (SPR) soit en cuvette avec un biosenseur IAsys, soit en flux continu avec un Biacore. J'ai eu en effet la chance de pouvoir tester et, dans une certaine mesure, comparer, les deux systèmes de biosenseurs développés pour les études d'interactions les plus performants sur le marché. Les différents instruments Biacore (manuel, automatisé) sont utilisés depuis plusieurs années maintenant par le Dr Hugues Lortat-Jacob et son équipe (Grenoble). J'ai bénéficié d'une bourse de voyage internationale de l'ANRT pour effectuer un stage dans le laboratoire de Pr David Fernig (Liverpool) qui possède de son côté les différents instruments manuels comme automatisés mais développés par Iasys/Farfield.

### 3.1.1. Utilisation du système à résonance plasmonique de surface en cuvette : le biosensor IAsys

#### 3.1.1.1. Matériel et méthodes

##### Biotinylation d'endocan par sa chaîne de DS : oxydation au périodate puis ajout de la biotine par le groupement hydrazide

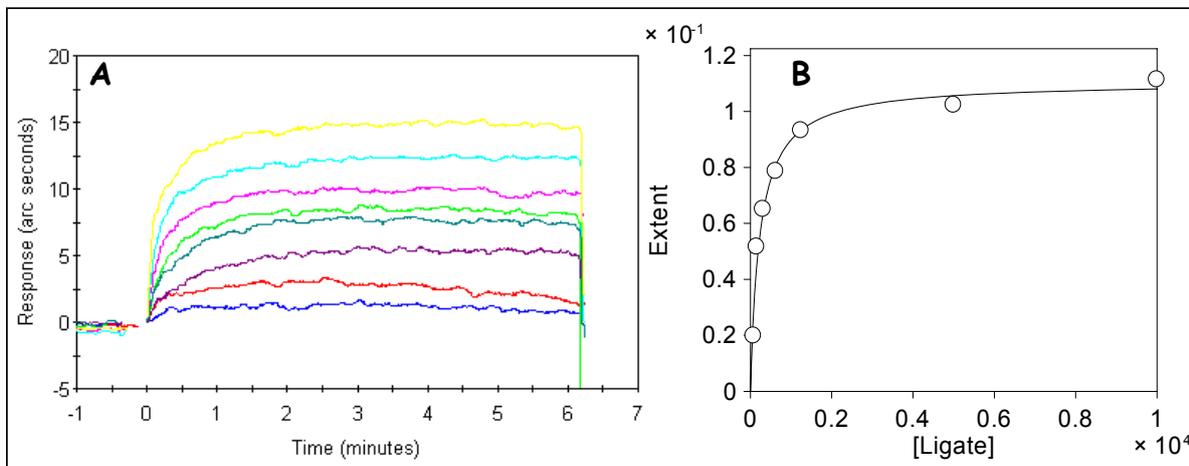
Pour ces expériences, 100  $\mu\text{L}$  d'endocan à 187  $\mu\text{g}/\text{mL}$  ont été utilisés. Le périodate de sodium ( $\text{NaIO}_4$ ) est ajouté (200  $\mu\text{L}$  à 1 M) au PG pour une concentration finale de 660 mM. La réaction se fait dans le noir, à 37°C, pendant une heure. Le périodate est séparé de l'endocan à l'aide d'une filtration sur gel de type G25 équilibrée en PBS. Les fractions contenant l'endocan sont rassemblées (2,4 mL). 30  $\mu\text{L}$  d'une solution de biotin(LongArm)hydrazide (Vector) dissoute à 25 mM dans du DMSO sont ajoutés. La réaction est concentrée sous vide (speed-vac) jusqu'à obtenir un volume de 200  $\mu\text{L}$ . 10  $\mu\text{L}$  de la solution de biotine sont de nouveau ajoutés pour une incubation toute la nuit à température ambiante. L'endocan biotinylé est dessalé par gel filtration sur une G25. Les fractions contenant l'endocan biotinylé sont concentrées sous vide (speedvac) jusqu'à un volume de 100  $\mu\text{L}$ .

##### Préparation d'une surface streptavidine et immobilisation de l'endocan biotinylé

Les cuvettes IAsys sont composées de deux canaux. La surface de chaque canal est couverte d'aminosilane qui est activé par 3 fois 50  $\mu\text{L}$  de cross-linker BS3 (BisSulfosuccinimidyl Suberate) à 1 mM. La streptavidine (2,5 mg/mL finale dans 40  $\mu\text{L}$  de tampon NaPi pH 7.7) est alors ajoutée pour être capturée par ses groupements amines. 600 arcs de streptavidine sont alors immobilisés. L'aminosilane activé est désactivé par ajout d'une solution de 2 M Tris. Après avoir vérifié que les anticorps ne se fixent pas de manière non spécifique à la surface de streptavidine, l'endocan biotinylé est ajouté (2 fois 50  $\mu\text{L}$ ).

##### Interaction avec les anticorps et mesure d'affinités

Les anticorps sont préparés à une concentration 5X dans un tampon PBS avec du Tween à 0,02 % (PBST). Pour chaque mesure d'interaction, 10  $\mu\text{L}$  d'anticorps sont ajoutés à 40  $\mu\text{L}$  de PBST. Une fois l'état d'équilibre atteint (après 4 minutes environ), la cuvette est vidée puis rincée 3 fois avec 50  $\mu\text{L}$  de PBST. La dissociation est ensuite observée durant environ 4 minutes. La surface est régénérée avec 3 fois 50  $\mu\text{L}$  de NaCl 1,5 M / NaOH 20 mM. La valeur des affinités est estimée selon deux méthodes : soit par la valeur du Bmax à l'équilibre pour



**Figure 29a: Mesure de la constante d'affinité de l'interaction entre l'endocan et un anticorps (EP208-003) par le système IAsys en utilisant la valeur du Bmax**

**A** : Le graphique montre la phase d'association jusqu'à l'équilibre, d'une gamme de concentration d'anticorps EP208-003 (100; 50; 25; 12,5; 6,25; 3,125; 1,56; 0,78  $\mu\text{M}$ ) avec l'endocan immobilisé via sa chaîne de DS biotinylée sur une surface de streptavidine.

**B** : Le graphique représente la valeur du Bmax (réponse de fixation maximale pour une concentration donnée) pour chaque concentration en anticorps mis en présence de l'endocan dans la cuvette. De cette manière, en utilisant la méthode de Scatchard, le logiciel FastFit peut donner une valeur du Kd. Dans cet exemple il est de  $2.3 \cdot 10^{-6} \pm 3.3 \cdot 10^{-7}$  M.

Anticorps	Kd par Bmax
EP208-003	$2.3063e-006 \pm 3.33794e-007$
EP208-004	$1.89058e-006 \pm 7.29107e-007$
EP208-005	$1.9029e-006 \pm 3.07725e-007$
EP221	$1.43602e-006 \pm 2.93903e-007$
EP236	Interactions non intégrable.

**Figure 29b: Tableau récapitulatif des Kd calculés par la méthode de Scatchard en IAsys pour les anticorps anti-endocan EP208 et EP221**

Les sensorgrammes obtenus pour l'anticorps EP236 n'ont pas donné de résultats intégrables. Pour les trois lots d'anticorps EP208 et pour l'EP221, les Kd calculés par cette méthode sont de l'ordre de  $2 \mu\text{M}$ .

plusieurs concentrations en anticorps (Scatchard), soit en intégrant les phases d'association et de dissociation pour obtenir les valeurs cinétiques de l'interaction ( $k_{\text{asso}}$  et  $k_{\text{disso}}$ ), et ainsi donner une constante d'affinité sachant que  $K_d = (k_{\text{disso}}/k_{\text{asso}})$ . L'interprétation des signaux d'interactions a été réalisée par le logiciel FastFit (Affinity Sensors).

### 3.1.1.2. Résultats

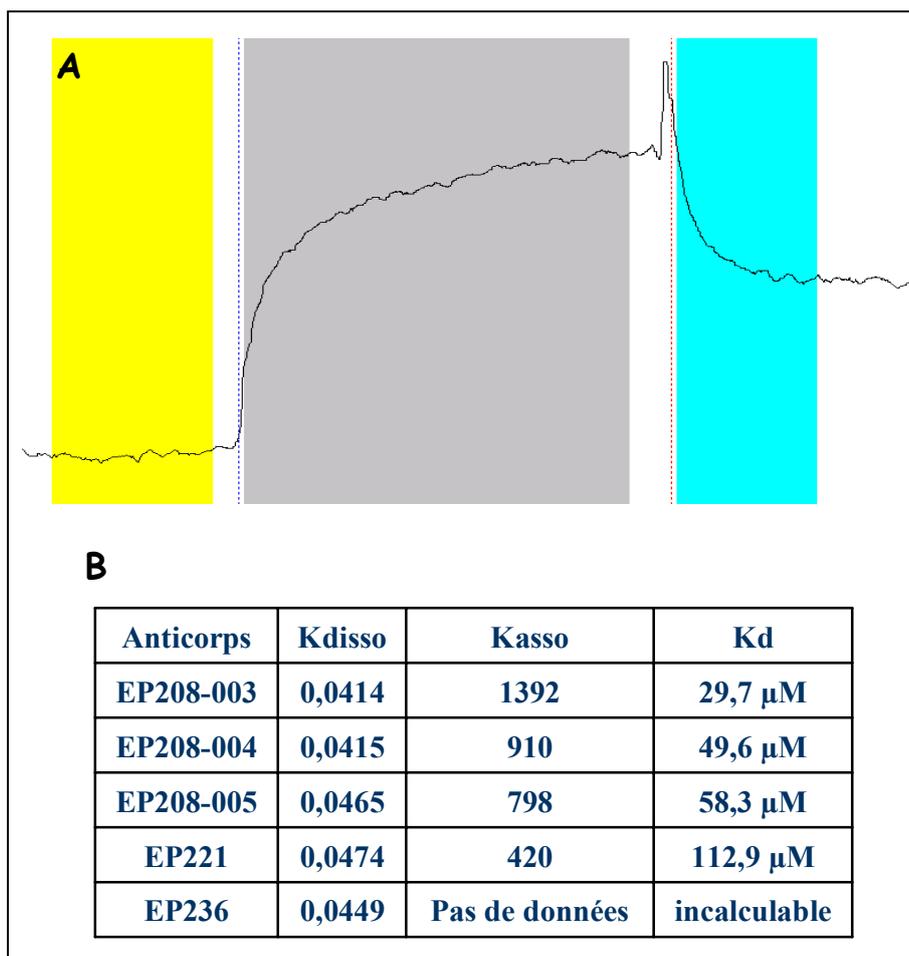
Dans ce travail de mesure des interactions entre endocan et les anticorps via le biosensor IA Sys, cinq anticorps ont été étudiés : les anticorps EP208, EP214, EP215, EP221, et EP236. Pour l'anticorps EP208, trois lots étaient disponibles. L'immobilisation d'endocan sur la surface d'amilosilane a été réalisée via le DS de l'endocan laissant ainsi libre la partie protéique. Afin de valider les données obtenues, deux méthodes pour mesurer le  $K_d$  de l'interaction de chaque anticorps avec l'endocan ont été utilisées. La première méthode consiste à mesurer le  $B_{\text{max}}$  de l'interaction (soit la valeur maximale du signal d'interaction pour une concentration donnée en anticorps). En travaillant avec plusieurs concentrations d'anticorps par la méthode de Scatchard, utilisant le  $B_{\text{max}}$  obtenu pour chacune des concentrations en anticorps, le  $K_d$  peut alors être calculé. La seconde méthode consiste à mesurer les constantes cinétiques d'association et de dissociation et de réaliser le rapport permettant d'obtenir une constante d'affinité. D'une manière générale, les anticorps EP214 et EP215 n'ont donné aucun signal d'interaction, cela sera discuté plus loin.

#### Mesure du $K_d$ par la méthode de Scatchard

Sur la Figure 29a sont montrés les signaux d'interaction ainsi que l'intégration des  $B_{\text{max}}$  en fonction des concentrations en anticorps du lot EP208-003. Le tableau de la Figure 29b rassemble les  $K_d$  calculés selon les anticorps ou encore selon les lots d'anticorps EP-208. D'une manière générale, pour les anticorps EP208 (3 lots) et EP221, les affinités mesurées sont de l'ordre du micromolaire ce qui ne correspond pas à ce que l'on peut attendre d'un anticorps de haute affinité. Les interactions observées pour l'anticorps EP236 n'ont donné aucun signal interprétable.

#### Mesure du $K_d$ par les constantes cinétiques d'association et de dissociation

Afin de valider les affinités obtenues par la méthode de Scatchard, le  $K_d$  des interactions a été évalué en mesurant les constantes d'affinités cinétiques que sont la constante d'association ( $k_{\text{asso}}$ ) et la constante de dissociation ( $k_{\text{disso}}$ ). La valeur du  $K_d$  est, de cette manière, calculée



**Figure 30: Calcul du Kd par les constantes cinétiques kasso et kdisso**

**A:** Sensorgramme type d'une interaction entre un anticorps et endocan obtenu par le système SPR IAsys.

Ce sensorgramme permet de calculer à l'aide du logiciel FastFit les constantes cinétiques d'une interaction. La phase d'association utilisée pour le calcul de la constante d'association est contenue dans la partie grise, de même pour la constante de dissociation calculée à partir du signal de dissociation contenu dans la partie bleue.

**B:** Tableau récapitulatif des valeurs de constante d'association et dissociation mesurées (moyenne de deux gammes de huit concentrations) pour chaque anticorps, et permettant de calculer le Kd de l'interaction.

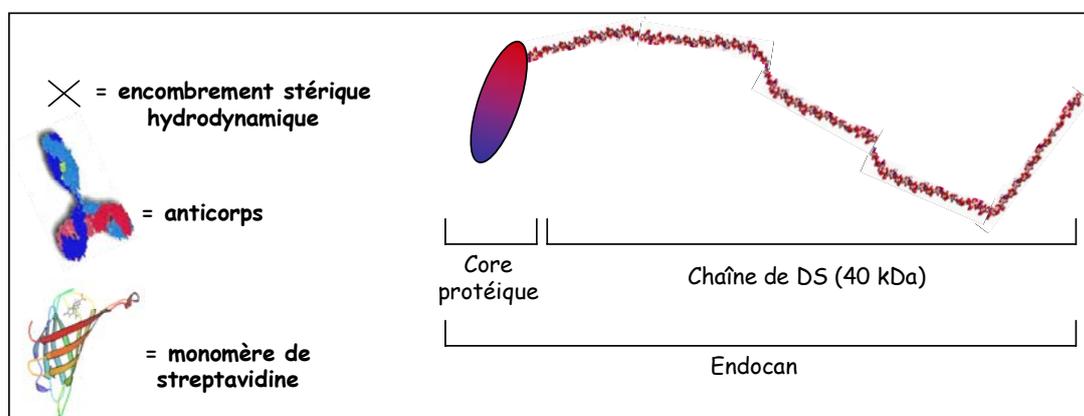
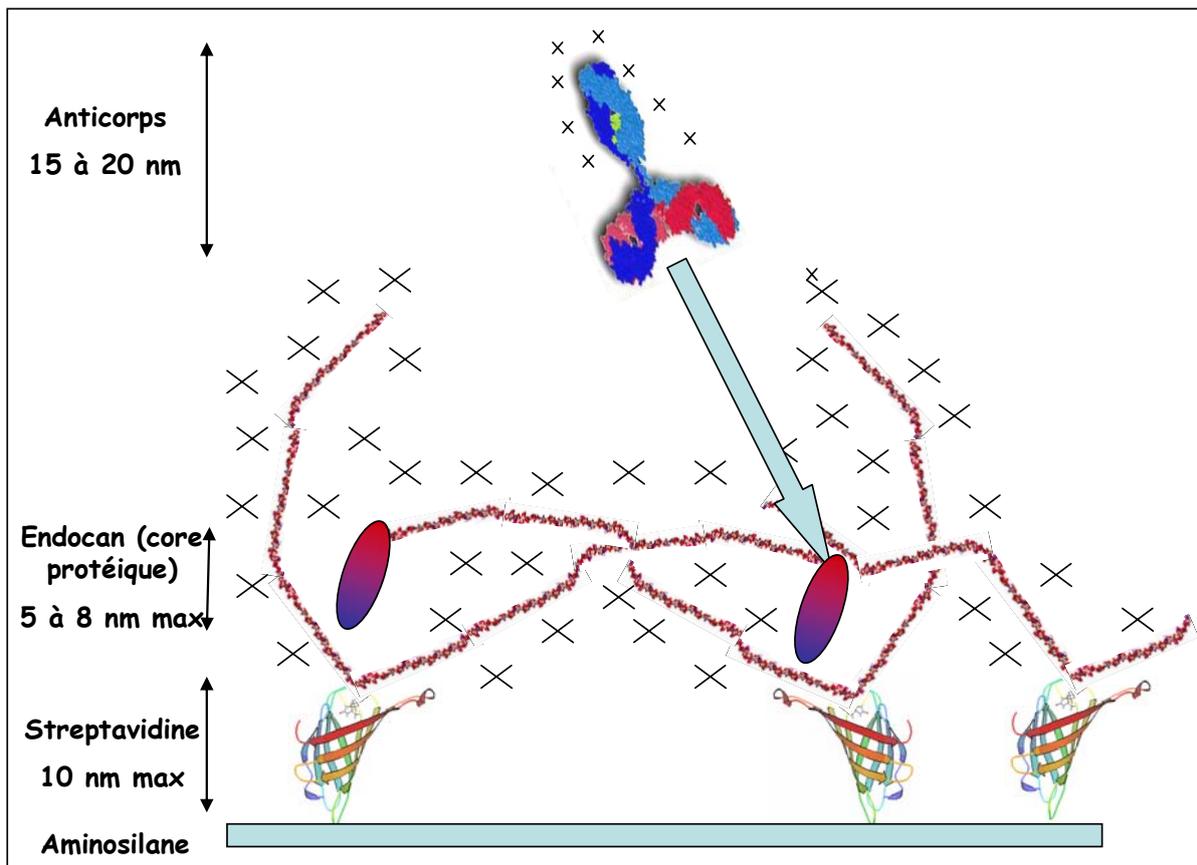
Comme lors de l'utilisation de la méthode de Scatchard, les sensorgrammes obtenus pour l'anticorps EP236 ne permettent pas de calculer un kasso et donc un Kd.

comme étant le rapport de  $k_{\text{diss}}/k_{\text{asso}}$ . La Figure 30A montre les différentes phases de l'interaction, et le tableau de la Figure 30B récapitule les valeurs obtenues pour chaque anticorps. Avec cette méthode, les  $K_d$  obtenus pour les anticorps EP208 et EP221 sont de l'ordre de quelques dizaines de micromolaire.

### 3.1.1.3. Discussion

Comme on vient de le voir, caractériser une interaction entre un anticorps et un protéoglycane n'est pas aussi simple que cela pourrait l'être avec une simple protéine sans aucune chaîne de GAG. Les résultats obtenus en utilisant le système SPR IAsys n'ont pas été satisfaisants. En effet, pour l'anticorps EP208 (potentiellement thérapeutique selon les données publiées chez la souris), un  $K_d$  de l'ordre de  $2 \mu\text{M}$  a été mesurée par la méthode de Scatchard. En se servant de la méthode des constantes cinétiques, un  $K_d$  de l'ordre de  $30$  à  $60 \mu\text{M}$  a pu être mesuré. Les trois lots d'anticorps EP208 testés ont donné des résultats similaires et ce quelque soit la méthode de mesure du  $K_d$  utilisée. Pour l'anticorps EP221, la méthode de Scatchard donne un  $K_d$  de l'ordre de  $1,5 \mu\text{M}$  alors que l'interprétation des constantes cinétiques permet d'estimer le  $K_d$  à environ  $110 \mu\text{M}$ . Pour l'anticorps EP236, seule la constante de dissociation a pu être mesurée. Elle correspond à celle observée pour les deux autres anticorps. En revanche, aucune constante d'association n'a été obtenue. Les expériences de Scatchard n'ont également pas donné de résultats interprétables. Ainsi, aucun  $K_d$  pour cet anticorps n'a pu être attribué. Également, pour les anticorps EP214 et EP215, (utilisés pour le développement du kit de dosage d'endocan circulant) aucun signal d'interaction avec l'endocan n'a pu être observé par cette méthode. Or, en règle générale, les anticorps monoclonaux de forte affinité ont des affinités pour leur antigène de l'ordre du nanomolaire.

D'une manière générale, les résultats sont homogènes et clairement reproductibles puisque les trois lots d'anticorps EP-208 ont donné des résultats similaires. L'anticorps EP221 donne également des résultats proches de ceux obtenus pour EP208 et la constante de dissociation obtenue pour l'anticorps EP236 est semblable à celles obtenues pour les autres anticorps. Malgré tout, ces résultats semblent traduire le côté artefactuel induit par le système mis en place et par l'outil de mesure qu'est le système SPR IAsys dont les interactions sont réalisées en cuvette fermée et sous agitation. Afin d'appréhender visuellement le système qui a été utilisé pour ces expériences, la Figure 31 a pour but de mettre à l'échelle les différentes molécules utilisées, et d'évaluer ainsi les artefacts qui ont pu être créés. Une telle



**Figure 31:** Schéma représentant le système mis en place dans la cuvette IAsys pour mesurer des interactions entre l'endocan et les différents anticorps monoclonaux anti-endocan

Sur une surface d'aminosilane, la streptavidine est immobilisée. L'endocan biotinylé par sa chaîne oligosaccharidique est capturé par la streptavidine. Le système représenté prend en compte que le core protéique d'endocan est plutôt globulaire, que les chaînes d'endocan font 40 kDa, qu'un disaccharide fait 450 Da et que la longueur d'un disaccharide est de 0,8 nm.

représentation indique qu'il est fort probable que la mauvaise affinité mesurée des anticorps anti-endocan soit directement liée à la mauvaise accessibilité du PG dans le système expérimental que nous avons créé. En effet, il pourrait exister un encombrement stérique hydrodynamique dû aux chaînes de DS assez important, rendant aux anticorps, l'accès au core protéique difficile.

Par ailleurs nous savons que la partie N-terminale de l'endocan a un pI assez élevé. A pH neutre, cette région est donc susceptible d'interagir directement avec la chaîne DS de l'endocan. La fixation du corps protéique d'endocan sur l'endocan immobilisé a d'ailleurs été observée (résultats non montrés). Or, l'anticorps monoclonal EP215 est spécifique de la région N-terminale de l'endocan. Le fait que cet anticorps ne puisse pas se fixer à son épitope est peut-être le signe d'une inaccessibilité causée par une interaction "inter-endocan". Concernant l'anticorps monoclonal EP214 spécifique de la région C-terminale, il est probable que la séquence reconnue par cet anticorps soit perdue lors de la biotinylation de l'endocan. En effet, lors des expériences de spectrométrie de masse réalisées pour caractériser le core protéique d'endocan, il s'est avéré que du côté C-terminal, le fragment 160-165 était extrêmement sensible à la dégradation et qu'une incubation de 48 heures à 37°C pouvait altérer l'intégrité du core protéique dans cette région. Or, ce fragment fait partie de l'épitope d'EP214 qui se trouve du côté C-terminal, sa dégradation expliquerait l'absence d'interaction de cet anticorps pour l'endocan immobilisé.

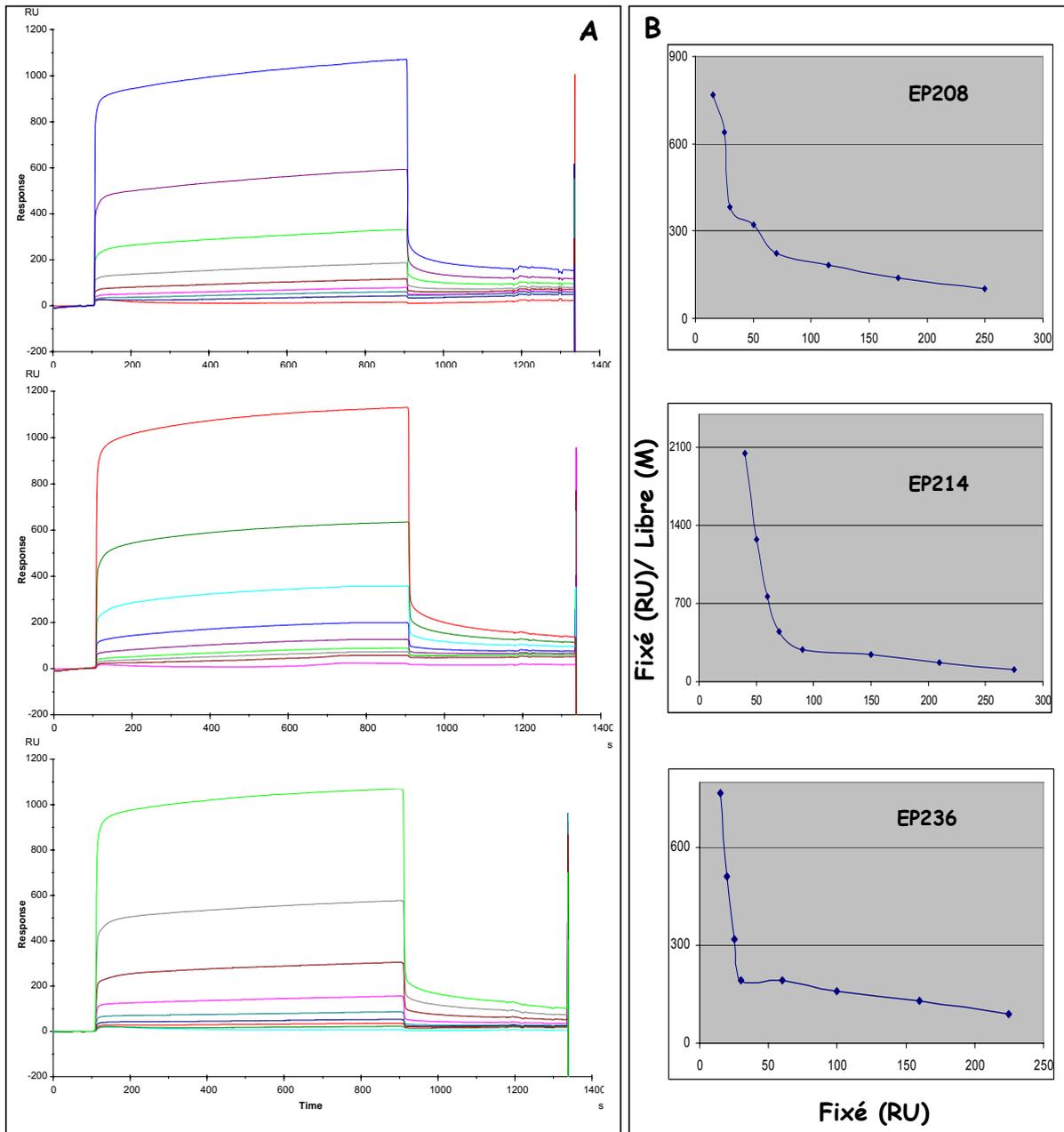
### **3.1.2. Mesure d'affinité entre des anticorps et l'endocan en temps réel et en flux continu : utilisation du Biacore**

Comme le système IAsys, le Biacore est un appareil capable de mesurer en temps réel l'interaction entre des molécules grâce à la technologie de résonance plasmonique de surface ([www.Biacore.com](http://www.Biacore.com)). A la différence du IAsys, pour le Biacore, l'interaction d'un analyte en solution avec un ligand immobilisé sur une surface se fait en flux continu. De cette manière, l'affinité de trois anticorps monoclonaux (EP208, EP214 et EP236) pour la molécule d'endocan a été mesurée.

#### **3.1.2.1. Matériel et méthodes**

##### *Immobilisation des anticorps*

Les expériences ont été réalisées sur un Biacore 3000 en utilisant les 4 canaux d'une sensorchip CM4 (Biacore). Les anticorps EP208, EP214 et EP236 ont été immobilisés sur



**Figure 32: Mesures d'interactions entre endocan et les anticorps anti-endocan par Biacore**

**A** : Sensorgramme obtenu lors du passage de différentes concentrations (2,5 1,25 0,625 0,3125 0,15625 0,0781 0,039 0,01950  $\mu\text{M}$ ) d'endocan sur les anticorps EP208 (haut), EP214 (milieu) ou EP236 (en bas) immobilisés (20000 RU) sur une sensorchip biacore de type CM3.

**B** : Représentation de Scatchard qui pour chaque concentration en endocan passé sur les anticorps immobilisés, retranscrit le taux d'endocan fixé par rapport à la concentration utilisée en fonction du taux d'endocan fixé (fixé/libre =  $f(\text{fixé})$ ).

trois des canaux, le dernier canal étant réservé au témoin négatif, où seul un traitement EDC/NHS est effectué. Le tampon HBS-EP (10 mM HEPES, 150 mM NaCl, 3,4 mM EDTA, 0,05 % de détergent P20, pH 7.4) est utilisé pour l'ensemble de ces expériences. A un débit de 5  $\mu\text{L}/\text{min}$ , et une température fixe de 25°C, la surface de la sensorchip est activée avec 50  $\mu\text{L}$  d'un mélange contenant 0,2 M 1-ethyl-3-(diméthylaminopropyl)carbodiimine (EDC) et 0,05 M de N-hydroxy-sulfosuccinimide (NHS). 50  $\mu\text{L}$  d'anticorps (50  $\mu\text{g}/\text{mL}$  dans 10 mM de tampon acétate, pH 5) sont injectés à un débit de 10  $\mu\text{L}/\text{min}$ . De cette façon, environ 20 000 RU d'anticorps ont été immobilisés. Les groupements NHS qui n'ont pas réagi sont bloqués par 50  $\mu\text{L}$  d'éthanolamine 1 M (pH 8.5) à 5  $\mu\text{L}/\text{min}$ .

#### *Mesure de l'interaction entre endocan et les anticorps immobilisés, et calcul de l'affinité*

Dans le but d'obtenir une mesure de l'affinité entre endocan et les anticorps immobilisés sur la surface CM4, une gamme de concentration d'endocan est réalisée. Ainsi, 200  $\mu\text{L}$  d'endocan (100 à 0,39  $\mu\text{g}/\text{mL}$ , dilution de 2 en 2), sont injectés à un débit de 15  $\mu\text{L}/\text{min}$ . Le temps de dissociation est fixé à 240 secondes, ensuite, 150  $\mu\text{L}$  de SDS 0,1 % suivis de 50  $\mu\text{L}$  de glycine-HCl 10 mM pH 1.85 sont utilisés à un débit de 50  $\mu\text{L}/\text{min}$  pour régénérer la surface.

A l'équilibre, le taux d'endocan fixé (bound) est mesuré pour chaque concentration utilisée d'endocan. Grâce à la méthode de Scatchard, le  $K_d$  de l'interaction à l'équilibre est défini comme suit :  $(\text{Bound}/\text{Free} = -\text{Bound}/K_d + N_o/K_d$  ; "Bound" correspond à la réponse de résonance à l'équilibre (RU), "Free" correspond à la quantité d'endocan passée sur la surface (nM) et  $N_o$  correspond au nombre de site). En traçant "endocan fixé/endocan libre" en fonction de la quantité fixée, pour plusieurs concentrations en endocan, on obtient une droite, dont la valeur absolue de l'inverse de la pente nous donne le  $K_d$  à l'équilibre.

#### **3.1.2.2. Résultats**

Après avoir immobilisé les anticorps sur une surface Biacore (20000 RU), différentes concentrations d'endocan (2,5 1,25 0,625 0,3125 0,15625 0,0781 0,039 0,0195  $\mu\text{M}$ ) ont été injectées sur la surface. La Figure 32A montre les sensorgrammes d'interaction obtenus. En prenant le taux d'endocan fixé aux anticorps à l'état d'équilibre pour chaque concentration d'endocan, par la méthode de Scatchard, un  $K_d$  peut être calculé. La Figure 32B montre pour chaque anticorps, les représentations de Scatchard, à savoir le taux d'endocan fixé, par rapport au taux d'endocan libre en fonction du taux d'endocan fixé. Les tracés obtenus se décomposent pour chaque anticorps en deux parties linéaires en fonction des concentrations

d'endocan utilisées. Pour les concentrations en endocan les plus faibles (0,0195 à 0,15625  $\mu\text{M}$ ), la valeur absolue de la pente du segment obtenu est largement plus importante que celle du segment obtenu pour les concentrations les plus fortes. Cela doit être dû à des interactions non spécifiques qui viennent s'ajouter pour les trop fortes concentrations. Le  $K_d$  correspond alors à la valeur absolue de l'inverse de la pente ( $K_d = -1/\text{pente}$ ) pour le segment des plus petites concentrations. Les  $K_d$  calculés pour chaque anticorps sont de 78 nM pour EP208, 19 nM pour EP214 et 26 nM pour EP236.

### **3.1.2.3. Discussion**

La mesure d'affinité entre endocan et des anticorps monoclonaux en cuvette (IASys) a donné des valeurs qui semblaient être dépendantes du système de mesure mis en place. Le Biacore à l'avantage d'être un système moins statique puisque l'interaction se fait en flux continu. Ainsi son utilisation a permis de donner des valeurs certainement plus proches de la réalité. De plus, contrairement à ce qui avait été mis en place pour le IASys, ce sont les anticorps qui ont été immobilisés (20000 RU) et l'endocan a été ajouté comme analyte. Différentes concentrations d'endocan ont été passées sur la surface possédant les trois anticorps. En utilisant la méthode de Scatchard, les affinités des anticorps EP208, EP214 et EP236 pour l'endocan mesurées sont respectivement 78, 19 et 26 nM. La fixation d'endocan sur la surface puis le passage des anticorps en solution, à différentes concentrations a également été testée par la technique du Biacore. Il s'est avéré que, dans ce cas de figure, aucune dissociation n'a pu être observée.

Ainsi, pour mesurer par Biacore l'affinité d'un anticorps pour endocan et, peut-être pour d'autres PGs, la méthode la plus appropriée semble être celle où les anticorps sont immobilisés sur la surface.

Cette méthode de fixation d'endocan à la surface n'a pas été testée avec le système IASys, la comparaison avec les résultats obtenus en Biacore ne peut donc pas être faite.

## **3.2. ETUDE D'INTERACTIONS ENTRE ENDOCAN ET DES MÉDIATEURS BIOLOGIQUES**

### **3.2.1. Introduction**

Du fait de la surexpression d'endocan dans certains types de cancers et lors des processus inflammatoires (Scherpereel et al., 2003 ; 2006 ; Lassalle et al., 2006 ; Sarrazin et al., 2006), il apparaissait intéressant d'évaluer l'effet que peut avoir ce PG sur les principaux médiateurs

impliqués dans de tels processus. Ce travail avait été en quelque sorte initié par l'étude de l'effet de l'endocan sur le HGF/SF (Bécharde et al., 2001a).

### 3.2.2. Matériel et méthodes

#### GMSA : Gel mobility shift assay ou "retard sur gel d'agarose"

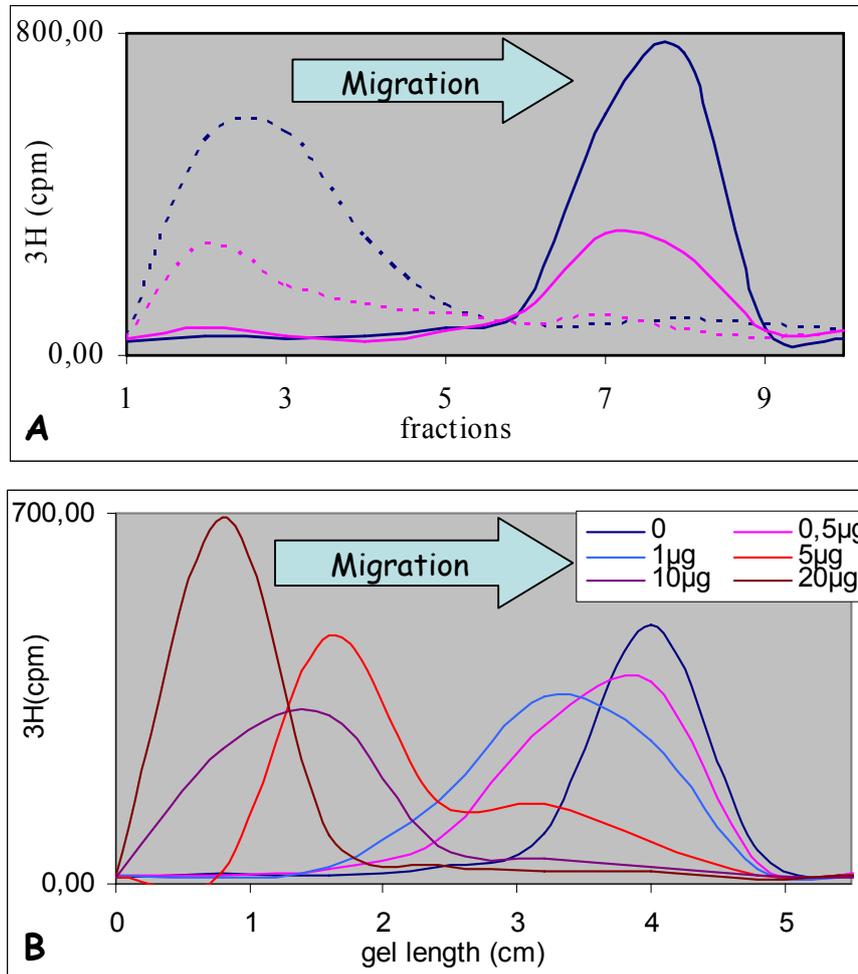
Cette méthode est essentiellement inspirée de celle décrite par Wu et al. (2002) et Holmes et al., (2007). Elle a été adaptée par le Dr Malcolm Lyon pour des applications concernant les interactions entre les GAGs et leurs médiateurs (Lyon et al., 2004). Brièvement, la chaîne de DS radiomarquée ( $^3\text{H}$ ) d'endocan provenant de cellules HUVEC ou de cellules HEK293 est mise en présence (ou non pour le témoin) de facteurs de croissance (FGF-2, HGF/SF) ou de cytokines (SDF-1 $\alpha$ , IFN $\gamma$ ), pendant 15 minutes à température ambiante, à la concentration indiquée dans les légendes des figures, dans un volume total de 7  $\mu\text{L}$  de PBS contenant 25 % de glycérol. Les échantillons sont ensuite déposés dans les puits d'un gel d'agarose 1 % (w/v). La migration des échantillons dans le gel est réalisée dans un système d'électrophorèse horizontal de type Bio-Rad Sub-Cell GT dans un tampon de migration contenant 1 mM d'EDTA, 40 mM Tris / acide acétique, pH 8.0. Dès la migration terminée, pour chaque ligne de migration, le gel est découpé tous les demi centimètres. Chaque fragment de gel est placé dans un eppendorf contenant du PBS pour être fondu à 50°C. La radioactivité ( $^3\text{H}$ ) contenue dans chacun des fragments est alors mesurée par scintillation.

#### Biacore

##### *Préparation des surfaces contenant l'endocan, le core protéique d'endocan, la chaîne DS d'endocan*

Les expériences sont réalisées sur un Biacore 3000, sur une sensorchip CM4 (Biacore). L'endocan, le core protéique d'endocan et les chaînes DS issues d'endocan ont ainsi été immobilisés sur trois canaux différents, et le quatrième et dernier canal est réservé au témoin négatif, où seul le traitement EDC/NHS est effectué.

Le tampon HBS-EP (10 mM HEPES, 150 mM NaCl, 3,4 mM EDTA, 0,05 % de détergent P20, pH 7.4) est utilisé. A un débit de 5  $\mu\text{L}/\text{min}$  et une température fixe de 25°C la surface de la sensorchip est activée avec 50  $\mu\text{L}$  d'un mélange contenant 0,2 M 1-ethyl-3-(diméthylaminopropyl)carbodiimine (EDC) et 0,05 M de N-hydroxy-sulfosuccinimide (NHS). 150  $\mu\text{L}$  d'endocan, core protéique d'endocan ou NH<sub>2</sub>-DS d'endocan (30  $\mu\text{L}$  CTAB 6mM + 120  $\mu\text{L}$  endocan 680  $\mu\text{g}/\text{mL}$ ) sont préparés pour une injection de 100  $\mu\text{L}$ . Les groupements



**Figure 33: Utilisation de la méthode de GMSA pour visualiser l'interaction entre le HGF/SF et la chaîne DS d'endocan**

**A** : L'endocan (225 ng) radiomarqué, produit par les cellules HEK293 ou HUVEC à été mis en présence ou non d'un excès d'HGF/SF (8 µg) pendant 15 min avant d'être déposé sur le gel d'agarose (1 %) utilisé pour le GMSA. Le graphique représente la mesure de la radioactivité tout au long du gel d'agarose (la migration se fait de gauche à droite). Les échantillons suivants ont subi la migration : bleu = endocan produit en HEK293; bleu pointillé = endocan HEK293 + HGF/SF ; rose = endocan produit en cellules HUVEC; rose pointillé = endocan HUVEC + HGF/SF.

**B** : L'endocan d'HUVEC (500 ng) à été mis en présence d'une gamme de d'HGF/SF (0; 0,5; 1; 5; 10; 20 µg) pendant 15 min avant d'être déposé sur le gel d'agarose utilisé pour le GMSA. Le graphique représente la mesure de la radioactivité tout au long du gel d'agarose (la migration se fait de gauche à droite).

NHS qui n'ont pas réagi sont bloqués par une injection à un débit de 5  $\mu\text{L}/\text{min}$  de 50  $\mu\text{L}$  d'éthanolamine 1 M (pH 8.5).

*Criblage de différents facteurs dont l'interaction avec endocan, son core protéique ou sa chaîne de DS a été évaluée*

Le criblage des protéines pouvant potentiellement lier l'endocan (HGF/SF, VEGF165, SDF-1 $\beta$ , RANTES, MCP1, MIP1 $\alpha$ , MCP2, PF-4, IP10, MIP3 $\alpha$ , GRO/MGSA, IGF-1, IFN $\gamma$ , KGF, SDF-1 $\alpha$ , FGF2, TNF $\alpha$  (Preprotech)) a été fait à 200 nM. Pour cela, 150  $\mu\text{L}$  de protéine ont été injectés à un débit de 30  $\mu\text{L}/\text{min}$  sur les surfaces contenant l'endocan, le core protéique uniquement ou la chaîne DS d'endocan uniquement. La dissociation a été observée pendant 240 secondes. La surface a ensuite été régénérée avec une solution de 1,5 M NaCl, 20 mM NaOH.

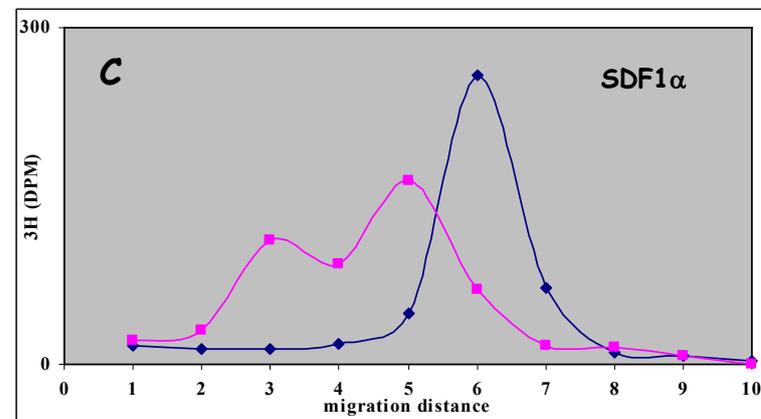
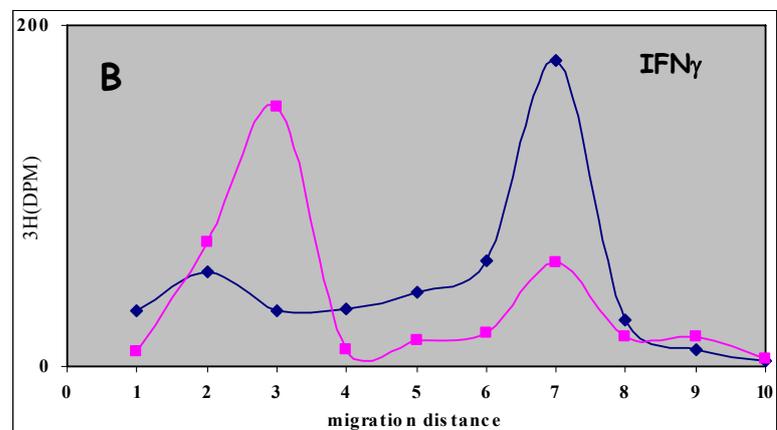
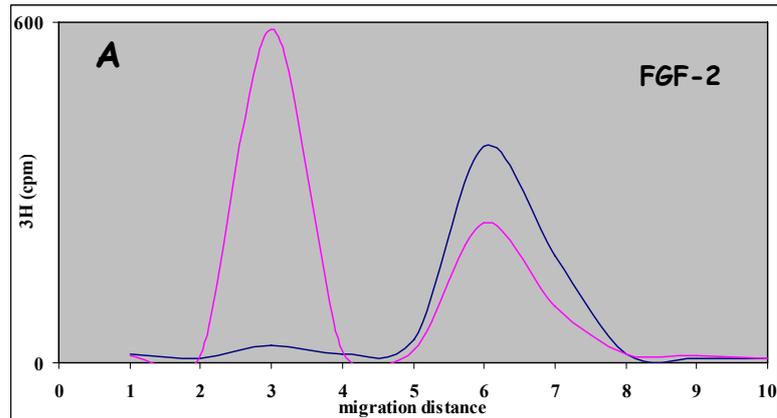
Pour déterminer l'affinité de certains facteurs liant endocan, une gamme de concentration de protéines (de 200 à 0 nM) a été injectée dans les mêmes conditions.

### **3.2.3. Résultats**

*Utilisation de la méthode de GMSA pour étudier quelques exemples d'interactions entre l'endocan et certains médiateurs.*

Endocan est connu depuis 2001 pour interagir avec le facteur HGF/SF directement via sa chaîne de DS (Bécharde et al., 2001) comme cela était connu pour d'autres chaînes de DS (Lyon et al., 1998). Cette interaction avait donc été caractérisée par IAsys dans le laboratoire du Dr David Fernig (Bécharde et al., 2001). Afin de vérifier que la méthode utilisée (GMSA) pouvait fonctionner avec la chaîne DS d'endocan radiomarquée obtenue chez les cellules HUVEC ou des cellules HEK293, les premiers essais ont été réalisés entre le HGF/SF et la chaîne DS d'endocan. Grâce à cette méthode qui a déjà été utilisée pour démontrer l'interaction d'HGF/SF avec les GAGs de type DS (Lyon et al., 2004), il a été possible de montrer que l'endocan, produit par les cellules HEK293 ou HUVEC, est capable de fixer le facteur HGF/SF (Figure 33A). Même si cette méthode ne permet pas de calculer de manière précise des affinités entre les deux molécules, elle est tout de même dépendante des quantités mises en présence (Figure 33B).

De même, pour les médiateurs tels que le FGF-2, l'IFN $\gamma$  ou le SDF-1 $\alpha$ , la méthode de GMSA a été utilisée pour avoir un aperçu des interactions potentielles qu'il pouvait y avoir avec la chaîne DS de l'endocan produit par les HUVEC (Figure 34). Les protéines FGF-2 et



**Figure 34:** Utilisation de la méthode de GMSA pour visualiser l'interaction entre la chaîne DS d'endocan produit par les HUVECs et les facteurs FGF-2, SDF1 $\alpha$ , et IFN $\gamma$

La chaîne DS d'endocan radiomarquée, produit par les cellules HUVEC à été mise en présence (courbes roses) ou non (courbes bleue) de médiateurs tels que le FGF-2 (A), l'IFN $\gamma$  (B) et le SDF1 $\alpha$  (C). Le graphique représente la mesure de la radioactivité tout au long du gel d'agarose (la migration se fait de gauche à droite).

IFN $\gamma$  ont permis d'induire un retard évident dans la migration de la chaîne DS d'endocan. Une interaction entre les deux molécules est donc clairement probable. En revanche, le SDF-1 $\alpha$  a, lui, donné un signal moins facilement interprétable, puisque deux positions de migration, différentes de la migration de la chaîne seule, ont été observées (Figure 34C). Il est difficile d'identifier ces deux spots de migration ainsi créés. Comme le SDF-1 $\alpha$  est connu pour dimériser en présence d'HS (Sadir et al., 2001 ; Veldkamp et al., 2005), cela pourrait rendre compte de nos résultats sans qu'il soit en l'état possible de l'affirmer.

#### *Utilisation du Biacore pour évaluer les interactions entre des médiateurs et endocan*

Après avoir mis en évidence que l'endocan était capable de lier des facteurs de croissance aussi importants que le HGF/SF (Bécharde et al., 2001), le criblage de plusieurs autres protéines a été entrepris par Biacore. Dans cette étude, le but était également de mettre en évidence qui de la chaîne oligosaccharidique ou du corps protéique était impliqué dans ces interactions potentielles. Après de nombreux essais de mise au point et d'optimisation, l'endocan, son corps protéique et même sa chaîne DS ont pu être immobilisés sur des sensorchips Biacore. La méthode d'immobilisation via le détergent "CTAB" utilisée est originale, puisqu'elle n'a été développée qu'à partir de 2004-2005, pour l'immobilisation de fragments d'ADN. L'intérêt de ce détergent cationique est de masquer les charges négatives trop importantes, comme celles que l'on peut trouver sur l'ADN et bien sûr les GAGs, et ainsi d'augmenter le rendement de couplage sur la surface, elle-même anionique.

Le tableau de la Figure 35A liste l'ensemble des facteurs testés jusqu'à présent, à une concentration de 200 nM. Les médiateurs comme MCP1, MIP1 $\alpha$ , IP10, MIP3 $\alpha$ , GRO/MGSA, IGF-I, SDF-1 $\alpha$  et TNF $\alpha$  ne semblent pas interagir avec l'endocan. En revanche, VEGF, SDF-1 $\beta$ , RANTES, MCP2, PF-4, IFN $\gamma$ , KGF (FGF-7), FGF-2 et bien entendu HGF/SF sont des médiateurs qui affichent clairement une réponse correspondant à leur liaison avec l'endocan. Le FGF-2 et le FGF-7 (KGF) sont les deux seuls facteurs qui donnent un signal non spécifique détectable sur la piste contrôle. Le résultat inattendu qui peut être observé (sauf dans le cas de HGF/SF) est la liaison insoupçonnée de ces médiateurs avec le core protéique d'endocan. Pour l'instant, aucune explication ne peut être donnée concernant cette observation et d'autres travaux devront être effectués.

#### *Mesure de l'affinité des facteurs liant endocan*

Parmi les protéines que nous avons criblé, nous avons sélectionné celles qui liaient l'endocan à 200 nM pour étudier leur affinité. Pour ce faire, une gamme de concentration de

Analytes	Ligands		
	endocan	core protéique	chaîne DS
HGF/SF	OUI	NON	OUI
VEGF165	OUI	OUI	OUI
SDF1b	OUI	OUI	OUI
RANTES	OUI	OUI	OUI
MCP1	NON	NON	NON
MIP1a	NON	NON	NON
MCP2	OUI	OUI	OUI
PF4	OUI	OUI	OUI
IP10	NON	NON	NON
MIP3a	NON	NON	NON
GRO/MGSA	NON	NON	NON
IGF-I	NON	NON	NON
IFN $\gamma$	OUI	OUI	OUI
KGF	OUI	OUI	OUI
SDF1a	NON	NON	NON
FGF2	OUI	OUI	OUI
TNF $\alpha$	NON	NON	NON

**Figure 35A:** Tableau récapitulatif du criblage de médiateurs testés dans l'interaction avec endocan, son core protéique ou sa chaîne de DS par Biacore

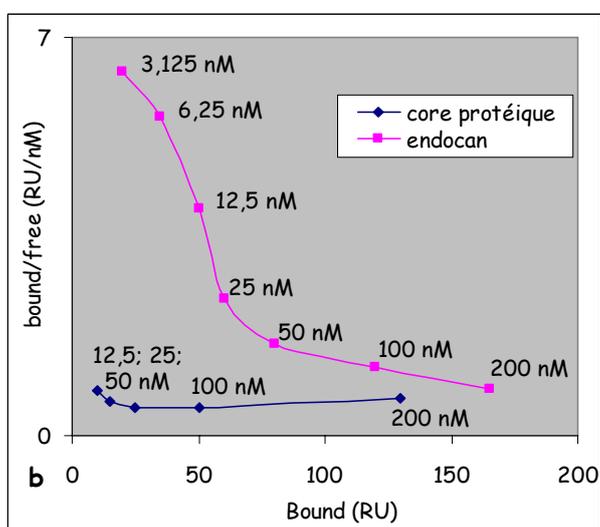
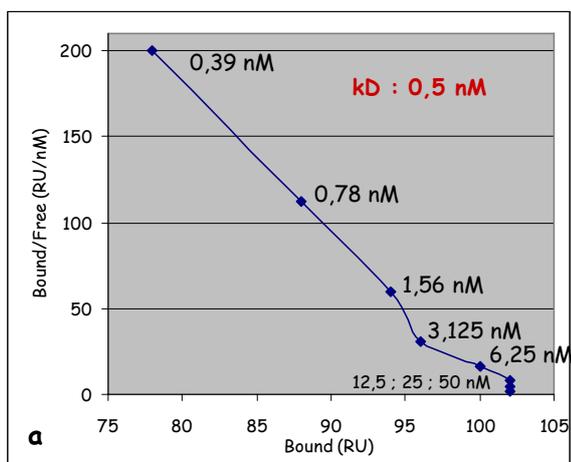
Si une interaction a été observée entre les ligands et les analytes un "OUI" apparaît dans la case correspondante.

protéine est passée en Biacore sur les surfaces où l'endocan, son core protéique et la chaîne de DS d'endocan ont été préalablement immobilisés. A l'équilibre, la valeur de la réponse d'interaction (RU) correspond à la quantité de protéine qui a interagi. Par la méthode de Scatchard, le Kd de l'interaction peut alors être évalué. En utilisant cette méthode, l'affinité de liaison a pu être calculée sur deux des huit protéines criblées. De cette manière, le Kd de l'interaction d'endocan (PG entier) avec HGF/SF a été évalué à 0,5 nM et avec IFN $\gamma$  à 11 nM (Figure 35B). Pour les autres protéines, plus la concentration utilisée est élevée, plus le signal d'interaction est fort. Cependant, les signaux obtenus n'étaient ni interprétables par les logiciels Biacore, ni même par la méthode de Scatchard.

### 3.2.4. Discussion

Il a déjà été montré que, via sa chaîne de DS, l'endocan est capable d'interagir et, de ce fait, influencer l'activité d'un facteur de croissance, le HGF/SF (Bécharde et al., 2001). Nos résultats ont montré l'existence d'interactions possibles entre la chaîne DS d'endocan et le HGF/SF, le FGF-2 et l'IFN $\gamma$  tout d'abord par la méthode GMSA. En revanche, concernant l'interaction avec le SDF-1 $\alpha$ , le profil de migration GMSA pouvait être questionné, comme l'ont confirmé nos études par Biacore. La méthode GMSA est une méthode très consommatrice en protéines recombinantes et n'est finalement que globalement peu informative sur les valeurs biochimiques de l'interaction. Le Biacore est clairement un moyen plus confortable pour l'étude des interactions protéine-GAG bien que l'immobilisation du PG tel qu'endocan à la surface des sensorchips Biacore a nécessité beaucoup d'efforts et de mise au point. Plusieurs méthodes ont ainsi été testées comme la biotinylation sur le core protéique ou sur les chaînes de DS, pour une immobilisation sur une surface streptavidine. Différents types de surface ont été utilisés, avec ajout ou non de contre ions limitant ainsi l'hydrodynamicité de la longue chaîne de DS portée par le PG. Après de nombreux échecs, l'endocan a pu être immobilisé sur les surfaces par un couplage direct (EDC/NHS) en utilisant un détergent cationique particulier, le CTAB, utilisé à la base pour immobiliser des fragments d'ADN. De cette manière, le criblage de différents médiateurs a pu être effectué.

Parmi les médiateurs criblés, ont été sélectionnés des facteurs de croissance, des cytokines ou chimiokines importants dans les phénomènes inflammatoires ou dans le développement tumoral. Les chimiokines jouent un rôle dans la migration cellulaire, lors de l'inflammation ou d'une réponse immunitaire, en recrutant, soit dans le torrent circulatoire, soit dans les tissus environnants, les cellules spécialisées ou participant à ces phénomènes. Les chimiokines qui ont été criblées pour une éventuelle liaison avec l'endocan sont MCP1,



**Figure 35B: Mesure d'interactions entre HGF/SF,  $IFN\gamma$  et endocan et/ou le core protéique d'endocan par Biacore**

**a** : Représentation de Scatchard des données obtenues pour l'interaction entre le HGF/SF et l'endocan. Le  $K_d$  de l'interaction à l'équilibre calculé de cette manière est de 0,5 nM. Aucune interaction sur la piste où seul le core protéique a été immobilisé n'a été observée.

**b** : Représentation de Scatchard des données obtenues pour l'interaction entre l' $IFN\gamma$  et l'endocan (bleu) ou son core protéique (rose). L'interaction entre l'endocan et  $IFN\gamma$  donne une représentation de Scatchard qui possède deux segments de pentes différentes selon la concentration en  $IFN\gamma$  utilisée. Le  $K_d$  calculé à partir du segment le plus pentu ( $[IFN\gamma] = 3,125$  à 25 mM) est de 11 nM.

MCP2, MIP1 $\alpha$ , MIP3 $\alpha$ , IP10, GRO/MGSA, PF-4, RANTES et SDF-1 $\alpha$  et  $\beta$ . Ces chimiokines sont souvent exprimées de manière non constitutive et sont généralement elles-mêmes sous le contrôle d'autres protéines telles que les cytokines. Les cytokines dont nous avons étudié l'interaction avec endocan sont le TNF $\alpha$  et l'IFN $\gamma$ . Ensuite, d'autres facteurs de croissance souvent impliqués dans le développement tumoral ont été criblés, comme le VEGF, le FGF-2, le FGF-7, et bien sûr l'HGF/SF utile comme témoin positif. De plus, comme l'endocan avait été classé comme l'IGFBP-rP6 (insulin-like growth factor binding protein – related protein 6) au regard de son organisation structurale (Hwa et al., 1999) et puisque aucune affinité pour les IGF n'avait été décrite malgré cette classification, l'interaction de l'IGF-1 vis-à-vis d'endocan a été étudiée.

De ce criblage, les médiateurs tels que MCP-1, MIP1 $\alpha$ , IP10, MIP3 $\alpha$ , GRO/MGSA, SDF-1 $\alpha$ , TNF $\alpha$  et IGF-I ne donnent aucun signal d'interaction avec l'endocan. En revanche, les médiateurs suivants : VEGF, SDF-1 $\beta$ , RANTES, MCP-2, PF-4, IFN $\gamma$ , KGF (FGF-7), FGF-2 et HGF/SF ont tous donné un signal d'interaction avec l'endocan ou sa chaîne de DS immobilisée. Ces résultats sont en accord avec la littérature. MCP-1 est ainsi connu pour se lier aux HS (Ali et al., 2001, Lau et al., 2004). Rien n'est réellement décrit pour sa liaison avec des DS, bien qu'une interaction avec les HS (de l'ordre du micromolaire) semble pouvoir être inhibée par des CS ou DS (Kuschert et al., 1999). La non-interaction avec l'endocan à une concentration de 200 nM n'est donc pas réellement surprenante. Le même genre de description peut être fait pour MIP1 alpha ou bêta (Patel et al., 2001). Dans cet article, il est d'ailleurs discuté que l'interaction entre MIP1- $\alpha$  ou MIP1- $\beta$  et l'héparine peut être observée *in vitro* mais pas par immunohistochimie alors que, pour RANTES et TARC, elle le peut. Ils assignent cela à des salinités de mesure *in vitro* plus faibles que la salinité physiologique. L'interaction entre les facteurs MIP1- $\alpha$  et  $\beta$  et les GAGs n'est donc pas si claire et concorde avec nos résultats concernant endocan. De leur côté, IP10 et PF-4 semblent partager le même profil structural d'interaction sur les HS (Luster et al., 1995). Cependant, seule l'interaction du PF-4 vis-à-vis des DS a été décrite à ce jour (Cella et al., 1992 ; Petersen et al., 1998). Les résultats obtenus en Biacore avec endocan concordent donc avec ces différentes données bibliographiques. De même, Witt et Lander (1994) ont montré que PF-4 comme GRO/GMSA étaient capables de se fixer sur des séquences différentes de l'Hp. Cependant pour la chimiokine GRO/MGSA, aucune donnée sur une potentielle interaction avec les DS n'est disponible. Rien n'est non plus disponible sur l'éventuelle interaction des chimiokines MCP-2 ou SDF-1 $\beta$  avec les DS. Concernant la chimiokine SDF-1 $\alpha$ , l'interaction

avec les HS ou l'héparine a été bien caractérisée (Sadir et al., 2001 ; 2004), en revanche l'interaction avec les DS a été observée de manière moins évidente (Natelenbos et al., 2003). Ensuite, concernant les cytokines, l'interaction de l'IFN $\gamma$  avec les HS a été caractérisée et l'interaction avec les DS observée (Lortat-Jacob et al., 1995 ; Brooks et al., 2000 ; Fukui et al., 2002). Pour la cytokine TNF $\alpha$ , son interaction avec l'Hp a été mise en évidence (Lantz et al., 1991) mais ne semble pas être partagée par l'ensemble des spécialistes. Enfin pour les facteurs de croissance tels que VEGF, FGF-2 et FGF-7 l'interaction avec les HS a été largement décrite (Stringer et al., 2006 ; Pye et al., 1998 ; Kreuger et al., 2005). Concernant leur interaction avec les GAGs de type DS, celle des FGF-2 et FGF-7 a été réalisée (Taylor et al., 2005, Li et al., 2007), en revanche l'interaction entre VEGF et des DS n'a pas été clairement mise en évidence pour l'instant. Il semble cependant que les CS/DS liant le VEGF165 doivent être largement sulfatés (Li et al., 2007).

L'affinité des médiateurs liant endocan a été mesurée par la méthode de Scatchard, en faisant passer une gamme de concentration de médiateurs sur les surfaces. Des données interprétables n'ont pu être obtenues que pour les protéines HGF/SF et IFN $\gamma$  pour lesquelles une affinité respectivement de 0,5 et 11 nM a été mesurée vis-à-vis de l'endocan dans son intégralité. L'affinité d'HGF/SF pour endocan avait déjà été mesurée et donnée à 3 nM (Bécharde et al., 2001). Une autre mesure d'interaction entre HGF/SF et les DS a été faite, le Kd obtenu est de l'ordre de 19,7 nM (Lyon et al., 1998). En fonction de la source de DS, cette valeur peut, en effet, varier. Pour des DS d'ascidien, bien plus sulfatés, le Kd de l'interaction est par exemple d'environ 1 nM (Catlow et al., 2003). Concernant l'IFN $\gamma$ , l'interaction avec les DS n'a pas été largement décrite, aucune valeur d'affinité n'est pour l'instant disponible et donc la comparaison n'est pas possible.

De manière bien plus surprenante, pour chacun de ces médiateurs interagissant avec endocan ou sa chaîne de DS (sauf pour HGF/SF), une interaction avec le core protéique d'endocan uniquement a également été observée. Vis-à-vis du core protéique, aucune affinité n'a cependant pu être calculée. Les interactions observées avec le core protéique d'endocan, et le fait que la méthode de Scatchard ne puisse pas être appliquée correctement, posent la question de la spécificité des interactions observées. Il faut bien remarquer que seuls les médiateurs liant la chaîne de DS d'endocan fixent également le core protéique, ce qui n'écarte cependant pas l'hypothèse de l'interaction non-spécifique. Dans ce cas, y aurait-il sur le core protéique d'endocan des motifs que l'on retrouve sur la chaîne de DS qui pourraient être reconnus ? Les O-glycosylations mises récemment en évidence ont-elles un rôle de ce point de vue là ? Toutes ces questions restent encore à étudier et à explorer plus profondément.

### **3.2.5. Perspectives**

Il apparaît clairement qu'il faut continuer ce type d'expériences afin d'améliorer les différentes mesures des interactions, principalement pour évaluer la spécificité avec le core protéique. De plus, afin de définir quelles séquences oligosaccharidiques de type DS sont impliquées dans les interactions, une librairie de tailles et de structures de DS d'endocan pourrait être constituée, et permettrait des expériences d'inhibition de l'interaction.

## **4.ETUDE FONCTIONELLE D'ENDOCAN**

#### 4.1. INTRODUCTION

Endocan ou ESM-1 (pour Endothélial Cell Specific Molécule-1) a été défini originellement comme étant le produit des cellules endothéliales uniquement. Depuis peu, l'expression d'endocan a également été observée *in vitro* dans des cellules telles que les adipocytes (Wellner et al., 2003), les cellules de mélanomes, et certaines cellules rénales (Hendrix et al., 2003 ; Seftor et al., 2002 ; Ross et al., 2000 ; Abid et al., 2006). Et comme nous l'avons découvert au cours de notre travail, d'autres types cellulaires exprimant endocan restent certainement encore à découvrir.

Initialement, Lassalle et coll. (1996) ont montré que certains facteurs pro-inflammatoires tels que le TNF $\alpha$  et l'IL-1 étaient capables d'augmenter l'expression d'endocan sur des cellules HUVEC. L'IFN $\gamma$ , en revanche, permettait d'inhiber l'effet du TNF $\alpha$ . De plus, une étude clinique a montré que, chez les patients atteints de choc septique, le taux d'endocan circulant était accru de manière significative (Scherpereel et al., 2003 ; 2006). Il a également été montré qu'endocan inhibe l'interaction ICAM-1/LFA-1 (Bécharde et al., 2001b). Tous ces éléments permettent de placer endocan comme un acteur potentiel dans les événements inflammatoires dont les fonctions sont encore peu définies.

Récemment, il a été montré que l'expression d'endocan pouvait être augmentée par le VEGF. VEGF augmente également l'expression de PI3K et Hhex qui sont deux facteurs réprimant l'expression d'endocan (Rennel et al., 2007 ; Cong et al., 2006). Cela indique certainement l'existence d'une boucle de régulation d'endocan par le facteur VEGF, impliqué dans l'angiogenèse et donc la progression tumorale. De plus, l'injection de cellules sur-exprimant endocan induit la formation de tumeurs (Scherpereel et al., 2003) alors que l'injection des mêmes cellules, exprimant uniquement le core protéique ou l'endocan muté dans la région riche en phénylalanines, ne permet pas la formation des tumeurs. Chez les patients atteints d'un cancer du poumon, le taux sérique d'endocan est anormalement élevé, ajoutant donc un argument supplémentaire en faveur de l'implication d'endocan dans les processus de tumorigenèse.

Les résultats présentés relèvent tout d'abord d'une étude non exhaustive de lignées cellulaires exprimant ou non l'ARNm d'endocan ou l'endocan lui-même. Puis, dans un second temps, quelques résultats préliminaires concernant l'effet d'endocan sur des médiateurs aussi importants que l'IFN $\gamma$  dans l'inflammation seront exposés.

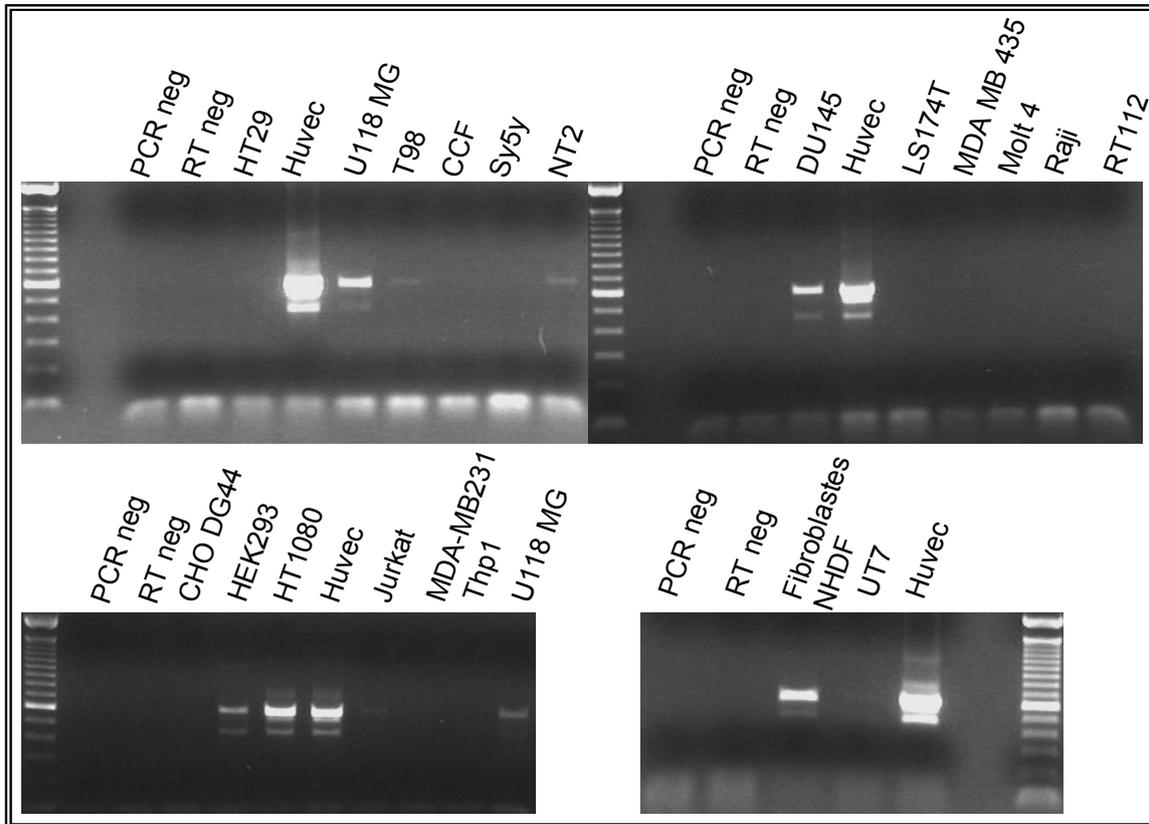
## 4.2. CRIBLAGE DE LIGNEES CELLULAIRES EXPRIMANT OU NON ENDOCAN

### 4.2.1. Matériel et méthodes

#### *Extraction d'ARN et RT-PCR*

**Extraction ARN :** Les ARN totaux des différentes lignées cellulaires ont été isolés selon le protocole d'extraction au RNable® (Eurobio). Les cellules adhérentes sont lavées en PBS, séchées, puis congelées à  $-80^{\circ}\text{C}$  jusqu'à préparation des ARN. Les cellules sont ensuite lysées sur glace à l'aide de 1 ml de solution de RNable®, puis le lysat ainsi obtenu est séparé par centrifugation, après addition de chloroforme, en une phase aqueuse (supérieure) contenant les ARN et une phase organique (inférieure) contenant l'ADN et les protéines. La phase aqueuse est récupérée et les ARN sont précipités par addition d'isopropanol. Après un lavage en éthanol 75 %, les ARN sont séchés par évaporation sous vide (speed-vac) puis resolubilisés dans de l'eau traitée au DEPC (diethylpyrocarbonate). La pureté des ARN ainsi purifiés est vérifiée par mesure de l'absorbance à 260 nm et à 280 nm, et leur qualité est visualisée par électrophorèse sur gel d'agarose 1,5 %, en tampon TBE (Tris Borate EDTA 1X), et en présence de bromure d'éthidium (BET). Les ARN sont ensuite directement utilisés pour la RT-PCR ou stockés à  $-80^{\circ}\text{C}$  pour un usage ultérieur.

**RT-PCR :** La Reverse Transcription (RT) des ARN préparés est réalisée à l'aide du kit SuperScript™ First-Strand Synthesis System for RT-PCR (Invitrogen) selon les recommandations du fournisseur. 2  $\mu\text{g}$  d'ARN sont ainsi mis en présence de l'enzyme reverse transcriptase SuperScript™ II en présence d'un inhibiteur de RNase (RNaseOUT™ Recombinant RNase Inhibitor) et d'amorces oligo(dT)<sub>12-18</sub>. La RT est réalisée à  $42^{\circ}\text{C}$  pendant 50 minutes, puis à  $70^{\circ}\text{C}$  pendant 15 minutes. Les ADN complémentaires (ADNc) ainsi obtenus sont alors traités à la RNase H pendant 20 minutes à  $37^{\circ}\text{C}$ , avant d'être congelés à  $-20^{\circ}\text{C}$  jusqu'à utilisation. 2  $\mu\text{g}$  de chacun des ADNc obtenus sont ensuite amplifiés par PCR (Polymerase Chain Reaction) sur thermocycleur PTC-200 (MJ Research) à l'aide d'amorces spécifiques (100 ng chaque) : amorce sens 5'GAG GCA GCT GGG AAA CAT GAA G3' ; amorce anti-sens 5' GCT TTC TCT CAG AAA TCA CAG3'. La réaction de PCR est réalisée en présence de dNTPs (0.25 mM, Eppendorf), d'enzyme polymérase (HotMaster™ Taq DNA polymerase, 2.5 U, Eppendorf), le tout en tampon de PCR, dans un volume final de 50  $\mu\text{l}$ . La PCR comprend un premier cycle de dénaturation de 2 minutes à



**Figure 36: Gels d'agarose des RT-PCR réalisées pour détecter la présence d'ARNm d'endocan dans les différentes lignées cellulaires**

Après l'extraction des ARN totaux des différentes lignées cellulaires, une RT-PCR est réalisée en utilisant des amorces spécifiques d'endocan (amorce sens 5'-GAG GCA GCT GGG AAA CAT GAA G-3' ; amorce anti-sens 5'-GCT TTC TCT CAG AAA TCA CAG-3'). Sur chacun des gels présentés sont indiqués en haut des lignes de dépôt : la lignée cellulaire étudiée, ainsi que les contrôles (PCR neg et RT neg). Pour chaque gel, un dépôt de la RT-PCR correspondant à la lignée cellulaire HUVEC est fait, cela sert de témoin positif.

94°C, puis 30 cycles d'amplification. Chaque cycle d'amplification comprend 30 secondes de dénaturation à 94°C, puis 30 secondes d'hybridation à 56°C, et enfin 30 secondes d'élongation à 72°C. Enfin, un dernier cycle de 10 minutes à 72°C termine la PCR. Les produits de PCR obtenus sont déposés sur gel d'agarose 1,5 %. La migration se fait en tampon TBE 1X, en présence de BET, pendant 1 heure à 120 V.

#### 4.2.2. Résultats et discussion

Lors de sa découverte, l'endocan, appelé aussi ESM-1 pour "Endothelial Cell Specific Molecule-1", était défini comme un PG exprimé uniquement par les cellules endothéliales (Lassalle et al., 1996). Depuis peu, l'endocan a également été caractérisé comme étant le produit de cellules telles que les adipocytes, les cellules de mélanomes ou certaines cellules rénales (Pour revue, voir Sarrazin et al., 2006). Afin de compléter les connaissances sur les cellules exprimant ou non l'ARNm ou le PG endocan, plusieurs types cellulaires de différents tissus ont été analysés. La Figure 36 présente les différentes RT-PCR qui ont pu être réalisées. Pour chacun des gels, une ligne est réservée aux cellules HUVECs qui servent de témoins positifs. En dessous de la bande principale de ce témoin positif, une bande plus faible peut être observée, qui révèle un épissage alternatif de l'exon 2 d'endocan, phénomène déjà observé (Aitkenhead et al., 2003 ; Depontieu, 2002). Le tableau de la Figure 37 résume, pour chaque lignée cellulaire testée, si l'ARNm et le PG endocan (par ELISA) ont été détectés. Ainsi pour les lignées endothéliales, l'ARNm d'endocan a été retrouvé uniquement chez les cellules HEK293 et les cellules de prostate DU145. La Figure 22 montre que l'expression de l'ARNm dans ces lignées est cependant beaucoup plus faible que chez les HUVECs. Le dosage protéique dans le surnageant de ces lignées n'a malheureusement pas encore été réalisé. Pour la lignée HT1080, qui est une lignée de cellules cancéreuses issues d'un fibrosarcome, la RT-PCR donne un signal aussi important que pour les HUVEC, et de l'endocan a pu être dosé dans les surnageants par dosage ELISA. De manière très intéressante, une étude de l'expression d'endocan dans les surnageants des HT1080 a été faite en fonction de la confluence des cellules. Il apparaît que les cellules dispersées expriment bien plus d'endocan que les cellules à confluence (Vincent Motte. communication personnelle). Aucune des lignées cellulaires d'hémopathies malignes testées lors de cette étude ne semble exprimer l'ARNm ou le core protéique d'endocan. Enfin, différentes lignées cellulaires issues de tissus cérébraux ont été analysées. Il apparaît que les cellules NT2 et T98G donnent un faible signal en RT-PCR. Ce faible signal se traduit pour les T98G par une expression en

Type cellulaire	Lignées	Expression endocan humain (ARNm)	Expression endocan humain (protéine)
<b>Endothéliale</b>			
Cordon ombilical	Huvec	oui	461ng/10 <sup>6</sup> c pour 48h
<b>Epithéliale</b>			
Colon	HT29	non	non
Colon	LS174T	non	non
Prostate	DU145	oui	non testé
Rein	HEK 293	oui	non testé
Sein	MDA-MB-231	non	non
Sein	MDA-MB-435	non	non
Vessie	RT112	non	non
<b>Sarcome</b>			
Fibrosarcome	HT1080	oui	42 ng/ 10 <sup>6</sup> c pour 24h
<b>Hémopathies malignes</b>			
Lymphoblaste	Molt4	non	non
Lymphocyte B	Raji	non	non
Lymphocyte T	Jurkat	non	non
Monocyte	Thp1	non	non
<b>Tissu cérébral divers</b>			
Glioblastome	U118MG	oui	40,45 ng/ 10 <sup>6</sup> c pour 48h
Neuroblastome	Sy5y	non	non
Théatocarcinome	NT2	oui	non testé
Astrogléoblastome	T98G	oui	non détecté
Astrocytome de bas grade	CCF-STTG1	non	non

**Figure 37: Tableau récapitulatif des différentes lignées cellulaires criblées dans le but de détecter l'expression d'ARNm ou de protéines d'endocan**

Le tableau récapitule pour chaque lignée cellulaire testée le type cellulaire et tissulaire auquel elles appartiennent. Pour chacune d'entre elles il est spécifié si l'ARNm d'endocan ou l'endocan lui-même a été détecté (par RT-PCR et ELISA) dans les lysats cellulaires (pour l'ARNm) ou dans le surnageant (pour le PG).

endocan non détectée dans les surnageants. Pour la lignée NT2, le dosage des surnageants n'a pas encore été réalisé. Chez les cellules de type glioblastome (U118 MG), l'endocan a pu être détecté aussi bien au niveau de l'ARNm que de l'expression protéique. Dans des conditions de culture classique, une quantité en protéine dix fois inférieure à celle dosée chez les HUVEC a pu être mesurée. Cependant, cette expression est régulée comme dans le cas des cellules HUVECs. Ces résultats font d'ailleurs l'objet d'un article en préparation qui confirme qu'endocan ne peut plus être cantonné à une expression spécifique des cellules endothéliales (Adam et al., en préparation). Finalement, des ARNm d'endocan ont également été détectés dans les fibroblastes du derme de type NHDF (Figure 36).

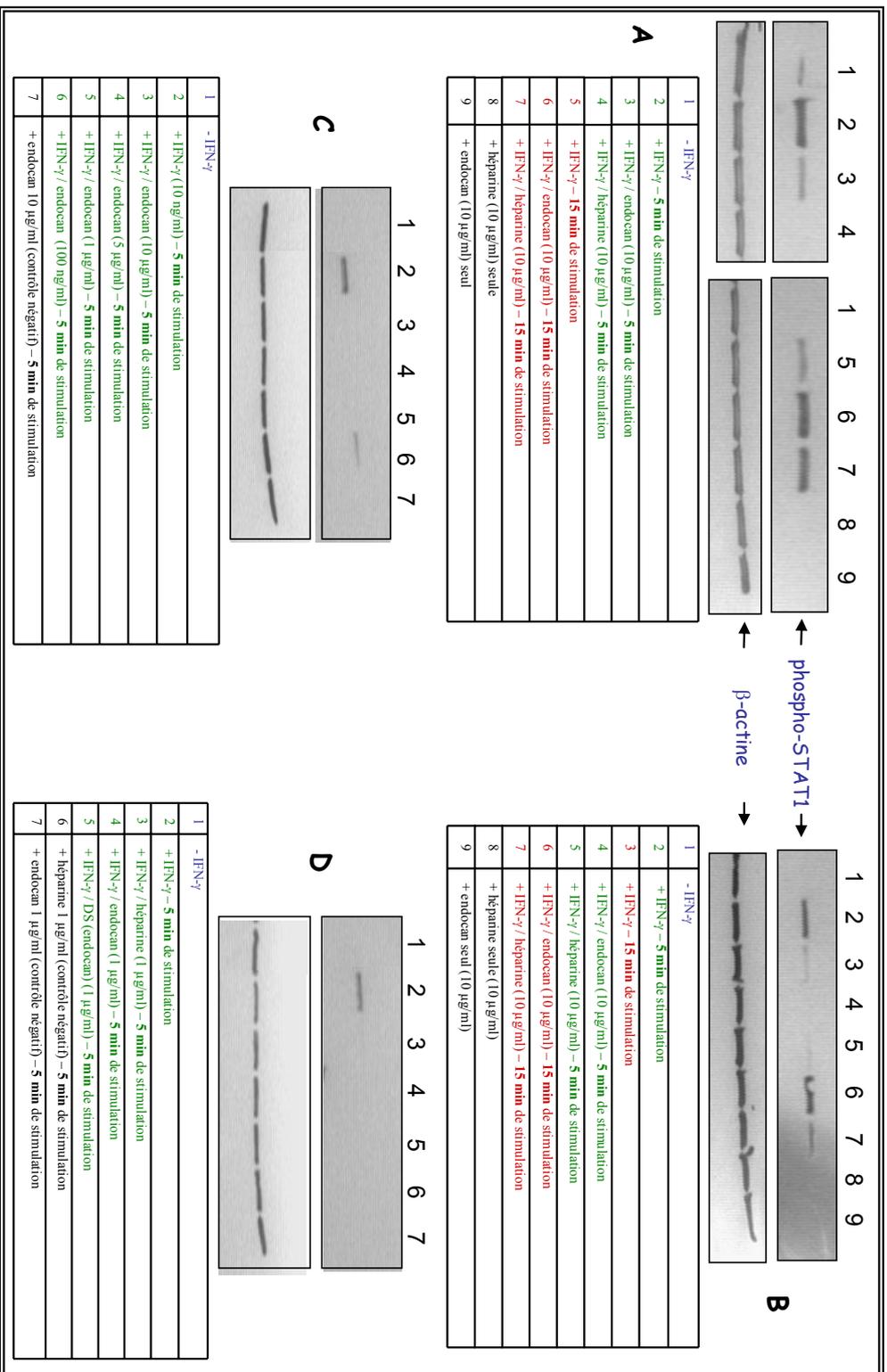
Ces résultats montrent donc clairement que de nombreuses études sont encore nécessaires pour évaluer la localisation et la régulation de l'expression d'endocan dans les différents types cellulaires afin d'appréhender ses fonctions.

### **4.3. EFFET *IN VITRO* D'ENDOCAN SUR LA CYTOKINE IFN $\gamma$**

#### **4.3.1. Matériel et méthodes**

##### *Etude de la signalisation d'IFN $\gamma$ par détection de la phosphorylation du facteur STAT-1*

Sur des cellules Jurkat (ATCC # TIB-152, clone E6-1) ou THP1 (ATCC # TIB-202) en culture ( $5 \times 10^6$  cellules), l'IFN $\gamma$  est ajouté à 10 ng/mL en présence ou non d'endocan, de la chaîne glycanique d'endocan ou d'héparine, en milieu RPMI complet contenant 10 % SVF. Après 5 minutes ou 15 minutes de stimulation à 37°C, les cellules sont lavées en PBS (froid) par centrifugation à 4°C (2000 trs/min, 2 minutes). Le culot cellulaire est repris dans du tampon de lyse (Tris/HCl 250mM pH 6.8, 10 % glycérol, 5 % SDS,  $\beta$ -mercaptoéthanol, 0,025 % bleu de bromophenol, 1 mM phenyl-methylsulfonyl fluoride (PMSF), 1  $\mu$ g/ml leupeptine, 1  $\mu$ g/ml aprotinine, 1 mM orthovanadate de sodium activé, cocktail d'inhibiteurs de protéases (Roche Diagnostics)). Les échantillons sont ensuite bouillis 3 minutes, puis soniqués et stockés à -20°C. 10  $\mu$ L de lysat cellulaire sont déposés sur gel SDS-PAGE 12 %. Après migration, les protéines sont transférées sur membrane de nitrocellulose par Western Blot (45 min ; 100 V). Les membranes de nitrocellulose sont saturées avec du TBS/lait 5 %, pendant 1 heure, puis rincées 3 fois avec du TBS/Tween 0,1 %. Les anticorps primaires, phospho-STAT1 (Tyr 701) (1/1000<sup>ème</sup>) (BD pharmingen) ou  $\beta$ -actine (1/2000<sup>ème</sup>) (Sigma), sont incubés en TBS/lait 5 % toute la nuit, à 4°C, sous agitation. Les membranes sont ensuite rincées 3 fois



**Figure 38 : Étude de la signalisation induite par l'IFN $\gamma$  et effet de l'endocan**  
 Le temps de stimulation (5 et 15 min) ainsi que l'effet de l'endocan et de l'héparine ont été testés sur les lignées cellulaires THP1 (A) et Jurkat (B). Sur les cellules THP1, un effet dose d'endocan (C) ainsi que l'effet de la chaîne glycanique seule (D) ont été testés. Le tableau en dessous du WB indique les conditions de stimulation. La détection de la  $\beta$ -actine est un contrôle du chargement des puits.

avec du TBS/Tween 0,1 %. L'anticorps secondaire,  $\alpha$ -mouse IgG-peroxidase (Amersham), est incubé 1 heure, à température ambiante, au  $1/10000^{\text{ème}}$ , dans du TBS/lait 5 %, sous agitation. La membrane est rincée 3 fois en TBS/Tween 0,1 % puis révélée par chimioluminescence ECL (Amersham).

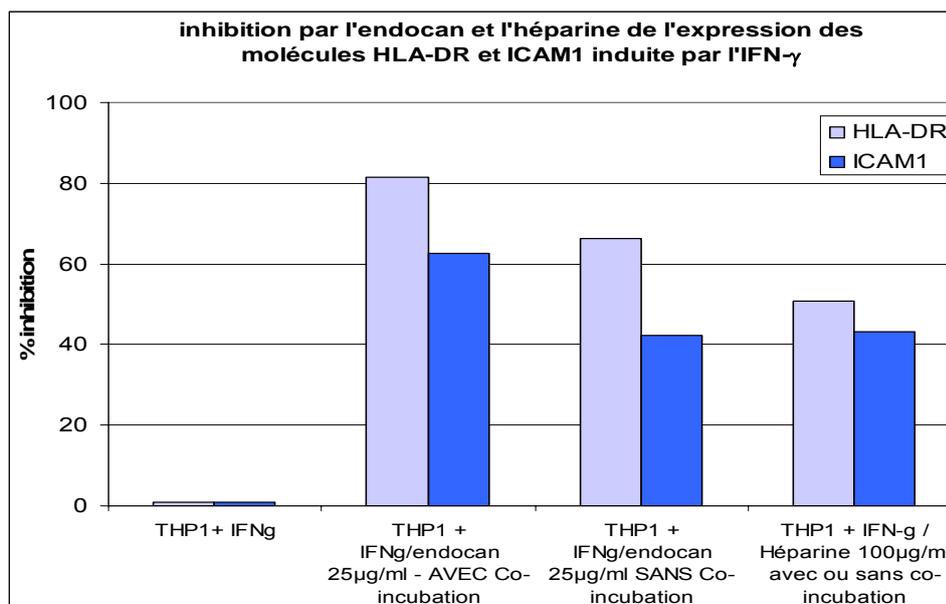
#### *Détection de l'expression des molécules HLA-DR et ICAM-1 induite par l'IFN $\gamma$*

Les cellules THP1 ( $0.1 \times 10^6$  cellules) sont stimulées par de l'IFN $\gamma$  (0-10 ng/mL), en présence ou non d'endocan ou d'héparine avec une pré-incubation ou non d'une heure, en milieu RPMI complet, contenant 10 % SVF. Les cellules sont incubées pendant 72 heures, à 37°C, puis l'expression des molécules HLA-DR et ICAM-1 est mesurée par cytométrie en flux. Pour cela, les cellules sont réparties en plaques 96 puits et lavées deux fois en PBS/SVF 2 % (1600 trs/min, 5 min, 4°C). Les anticorps anti-HLA-DR-FITC ( $1/10^{\text{ème}}$ , BD Pharmingen) et anti-ICAM-1-PE ( $1/10^{\text{ème}}$ , BD Pharmingen), ainsi que les contrôles isotypiques (IgG2a-FITC et IgG2a-PE, BD Pharmingen) sont incubés avec les cellules en PBS/SVF 2 %, pendant 30 minutes, à 4°C et à l'obscurité. Les cellules sont ensuite lavées deux fois en PBS/SVF 2 % (1600 trs/min, 5 min, 4°C), puis elles sont resuspendues dans 200  $\mu$ L de PBS/SVF 2 % et transférées dans des tubes pour FACS (BD Falcon). Les cellules sont analysées par cytométrie en flux à l'aide d'un FACSCalibur et du logiciel CellQuest (Becton Dickinson).

### **4.3.2. Résultats**

#### *Effet d'endocan sur la signalisation cellulaire d'IFN $\gamma$*

Les cellules THP1 ou Jurkat, connues pour répondre à l'IFN $\gamma$ , ont été utilisées pour évaluer si l'endocan avait un effet sur la signalisation d'IFN $\gamma$ . La signalisation d'IFN $\gamma$  passe par la phosphorylation du facteur STAT1. Des anticorps spécifiques de STAT1 phosphorylé ont ainsi pu être utilisés pour mesurer la signalisation liée à la cytokine. Les Figures 38A et B montrent que, pour les deux types cellulaires (THP1 et Jurkat), une stimulation avec l'IFN $\gamma$  pendant 5 minutes permet la phosphorylation de STAT1. En présence d'endocan (10  $\mu$ g/mL) ou d'héparine (10  $\mu$ g/mL), le signal diminue, voire devient nul. Pour une stimulation de 15 minutes, le signal de phosphorylation en réponse à l'IFN $\gamma$  seul est bien moins fort que pour une stimulation de 5 minutes. En revanche, la mise en présence d'IFN $\gamma$  avec l'héparine ou l'endocan permet d'augmenter le signal de phosphorylation au bout de 15 minutes de stimulation et, ainsi, de retrouver un signal au moins aussi fort que celui visualisé pour une stimulation à l'IFN $\gamma$  seul pendant 5 minutes. La Figure 38C présente le signal obtenu lors



***Figure 39: Histogramme récapitulatif des expériences de FACS permettant de détecter les molécules HLA-DR et ICAM-1 à la surface des cellules THP1 soumises à différentes stimulations***

Les cellules THP1 ont été soumises, pendant 72h, à une stimulation par l'IFN $\gamma$  mis en présence ou non d'héparine (100  $\mu$ g/mL) ou d'endocan (25  $\mu$ g/mL), avec ou sans co-incubation préalable. Il a été vérifié que l'héparine ou l'endocan sans IFN $\gamma$  n'induisaient pas l'expression des molécules détectées.

d'une stimulation en IFN $\gamma$  de 5 minutes lorsque celui-ci est mis en présence de différentes concentrations en endocan (10, 5, 1 et 0,1  $\mu\text{g/mL}$ ). Pour une concentration en endocan supérieure ou égale à 1  $\mu\text{g/mL}$ , plus aucun signal de phosphorylation de STAT1 n'est obtenu. A 100 ng/mL d'endocan, une bande plus faible que celle visualisée sans ajout d'endocan est observée. Afin d'évaluer si l'effet d'endocan sur l'IFN $\gamma$  qui a été observée est causée par la chaîne glycanique d'endocan, celle-ci a alors été utilisée seule. Comme le montre la Figure 38D, une incubation de l'IFN $\gamma$  avec soit 1  $\mu\text{g/mL}$  d'endocan, de sa chaîne de DS ou de l'héparine inhibe complètement la signalisation obtenue après 5 minutes de stimulation. Ce résultat met en avant le rôle de la chaîne glycanique d'endocan dans la régulation de l'effet d'IFN $\gamma$ .

#### *Effet d'endocan sur l'expression des molécules HLA-DR ou ICAM-1 induite par une stimulation à l'IFN $\gamma$*

La stimulation des cellules THP1 par l'IFN $\gamma$  conduit à l'expression des molécules du CMH telles que les HLA-DR ou des molécules d'adhésion, telles que ICAM-1. Afin de mesurer l'effet antagoniste d'endocan vis-à-vis des activités d'IFN $\gamma$ , une stimulation a été faite avec de l'IFN $\gamma$  mis en présence ou non d'endocan, avec une préincubation préalable ou non. Les mêmes expériences ont été réalisées avec l'héparine comme témoin positif. Sur l'histogramme de la Figure 39, qui résume les résultats obtenus par FACS, il apparaît que l'endocan exerce un effet inhibiteur vis-à-vis de l'expression des molécules HLA-DR et ICAM-1 induite par l'IFN $\gamma$ . En préincubant endocan et IFN $\gamma$ , l'inhibition s'avère plus importante que sans préincubation (80 contre 65 % d'inhibition de l'expression des HLA-DR et 60 contre 40 % d'inhibition d'ICAM-1). La préincubation n'a, en revanche, pas d'effet lors de l'ajout d'héparine à l'IFN $\gamma$ . Il est intéressant, d'ailleurs, de noter qu'une dose d'endocan de 25  $\mu\text{g/mL}$  permet d'inhiber l'expression des molécules HLA-DR et ICAM-1 d'une manière plus efficace que l'ajout de 100  $\mu\text{g/mL}$  d'héparine (80 contre 50 % pour les HLA-DR et 60 contre 40 % pour ICAM-1).

### **4.3.3. Discussion**

Comme il a déjà été dit, l'endocan semble jouer un rôle dans différents processus cancéreux et inflammatoires. Certaines cytokines pro-inflammatoires telles que le TNF $\alpha$  ou l'IL-1 permettent une expression accrue d'endocan. L'IFN $\gamma$ , en revanche, est capable d'inhiber l'effet du TNF $\alpha$  alors que l'IL-4 n'a aucun effet (Lassalle et al., 1996 ; Bécard et

al., 2000). L'expression d'ARNm d'endocan peut également être augmentée par le TGF- $\beta$ 1 alors que le TGF- $\beta$ 2 n'a pas d'effet (Zaho et al., 2004). Parmi les facteurs pro-tumoraux, le PDGF-BB, l'HGF/SF et l'EGF ne semblent pas avoir d'effet sur l'expression de l'ARNm d'endocan (Rennel et al., 2007). D'après les mêmes travaux, le FGF-2 ne semble pas avoir d'effet sur des cellules endothéliales. Cependant, nos résultats obtenus au sein d'Endotis ainsi que ceux d'Aitkenhead et coll. (2002) montrent un effet inducteur du FGF-2 sur endocan. Plus récemment, sur des cellules issues de glioblastomes, le FGF-2 induit l'expression d'endocan alors que le VEGF est sans effet (Adam et al. En préparation). Ces effets différents pourraient être dus au type cellulaire ou à l'état d'activation des cellules.

Par ailleurs, les effets d'endocan sur la cytokine IFN $\gamma$  ont été évalués. Tout d'abord, le fait qu'IFN $\gamma$  soit un répresseur de l'expression d'endocan dans certains cas laisse penser qu'une boucle de régulation pourrait exister. En effet, il est possible qu'en cas d'inflammation, l'expression d'endocan et d'IFN $\gamma$  soient toutes deux augmentées et que, selon les circonstances, chacune des deux molécules puisse jouer un rôle sur la présence, la disponibilité, la signalisation de l'autre molécule. Par exemple, il est connu qu'IFN $\gamma$  sur des fibroblastes réduit le taux d'IdoA des DS (Praillet et al., 1996). Or, l'IdoA est, dans certains cas, très important pour la liaison avec des facteurs tels que l'HGF/SF. Cela pourrait aller dans le même sens que l'effet répresseur vis-à-vis de l'expression d'endocan observé lors de l'ajout d'IFN $\gamma$  sur des cellules endothéliales ou inflammatoires. C'est pourquoi nous avons essayé d'évaluer l'effet d'endocan sur la signalisation d'IFN $\gamma$  et l'activité de cette cytokine. Ainsi, il apparaît clairement qu'endocan est capable d'inhiber l'expression des molécules HLA-DR et ICAM-1 induite par l'IFN $\gamma$  de manière au moins aussi efficace que peut le faire l'héparine. L'effet d'endocan sur la signalisation directe d'IFN $\gamma$  a également été évalué. Lorsque l'effet est observé sur les 5 premières minutes de stimulation, l'inhibition d'IFN $\gamma$  par endocan via sa chaîne de DS est évidente. En revanche, lorsque l'effet est observé au bout de 15 minutes de stimulation, la signalisation par l'IFN $\gamma$  semble au moins aussi importante en présence d'endocan (ou d'héparine). Il est connu que l'interaction IFN $\gamma$ /HS permet de protéger la cytokine de coupures protéolytiques et de la conserver ainsi active. Cela peut, par transposition avec les DS, expliquer cette dernière observation, qui peut également être liée à un déplacement de l'équilibre, c'est-à-dire que IFN $\gamma$  en présence d'endocan va mettre plus longtemps à atteindre son récepteur et activer pleinement la voie de transduction du signal.

Les résultats obtenus mettent donc en évidence qu'endocan est capable d'inhiber *in vitro* quelques unes des activités de l'IFN $\gamma$  en bloquant la transduction du signal. Il n'est cependant pas évident qu'*in vivo*, dans des phénomènes où l'IFN $\gamma$  est exprimés de manière conjointe

avec l'endocan, les conditions spatio-temporelles et l'implication d'autres facteurs permettent que de telles observations puissent être faites. Néanmoins, nos résultats ouvrent une piste intéressante pour de futures investigations.

## **5. CONCLUSION ET PERSPECTIVES**

Les protéoglycannes sont des macromolécules complexes et variées qui exercent d'indénombrables fonctions, dans une large gamme de processus physiologiques (Iozzo et Cohen, 1993 ; Iozzo et San Antonio, 2001 ; Timar et al., 202 ; Delehedde, 2002 ; Delehedde et al., 2005 ; Delehedde et al., 2006). Pour cela, ils peuvent être localisés dans le compartiment intracellulaire (cas de la serglycine), à la surface des cellules, dans les matrices extracellulaires, voire même être circulants (Delehedde et al., 2006). L'étude des PGs remonte à une cinquantaine d'années environ pour les plus anciennement découverts. Cela a laissé le temps d'appréhender certaines de leurs fonctions de bon nombre d'entre eux, notamment grâce à la création des souris privées du gène codant pour le PG étudié (les souris KO) (Goldberg et al., 2005 ; Gotte et al., 2005 ; Sarraj et al., 2006). Cependant, la caractérisation fonctionnelle de telles macromolécules est loin d'être évidente et leurs fonctions restent souvent assez larges. Les protéoglycannes sont en fait des molécules complexes, biochimiquement composées d'un core protéique sur lequel est lié au moins une chaîne oligosaccharidique de type glycosaminoglycane. De ce fait, leurs fonctions sont tributaires de la partie protéique (régulation de l'expression protéique, localisation, interaction avec des ligands), de leurs chaînes glycaniques (nature, structure, nombre, interactions avec des ligands), et surtout de la combinaison des deux (encore plus difficile à appréhender). Selon les cores protéiques, le nombre et la séquence des sites de liaison aux GAGs varient et donc la nature de ces derniers également. Les kératanes sulfates, les héparanes sulfates, les chondroïtines sulfates et les dermatanes sulfates peuvent tous être liés aux cores protéiques de PGs.

Endocan ou ESM-1 (Endothelial Cell Specific Molecule-1) est un protéoglycane à dermatane sulfate découvert il y a une dizaine d'années dans un laboratoire de l'Institut Pasteur de Lille. Ce protéoglycane est original puisqu'il se compose d'une seule chaîne de DS et a été retrouvé dans le courant circulatoire. Son core protéique ne semble pas avoir d'homologue protéique connu à ce jour et endocan n'a pas pu être classé dans une famille de protéines ou PGs existante (Lassalle et al., 1996), même si on peut remarquer que la bikunine est un CSPG circulant ne portant qu'une seule chaîne de GAG.

Le core protéique d'endocan est constitué de 165 acides aminés (après clivage du signal peptidique de 19 acides aminés) et possède 18 cystéines dans les 110 premiers acides aminés. Pour en apprendre davantage sur les caractéristiques biochimiques de ce core protéique, des expériences de caractérisation ont été entreprises. La mutation de la sérine 137 en alanine a permis d'obtenir, dans un système d'expression eucaryote (HEK293), le core protéique d'endocan dépourvu de sa chaîne glycanique. Les plateformes de production et de

purification mises en place au cours de mon travail de thèse pour l'obtention de cette protéine ont permis, par le biais de deux étapes de chromatographie (Echange d'ions et Immunoaffinité), d'obtenir le core protéique à l'échelle du milligramme, avec un très bon degré de pureté. Concernant la production, l'adaptation originale des cellules HEK293 au bioréacteur de type CL1000 (conçu initialement pour la production d'anticorps) s'est avérée très efficace pour la production d'endocan recombinant, mais peut également constituer un système de production intéressant dans bien d'autres cas. L'analyse par spectrométrie de masse (MALDI-TOF) du core protéique d'endocan purifié révèle un pic unique mais d'une largeur d'environ 4000 Da venant s'ajouter à la masse attendue de 18106 Da. Ce résultat inattendu a logiquement conduit à engager de nouvelles recherches pour évaluer les modifications post-traductionnelles que le core protéique d'endocan aurait pu subir. Des prédictions bioinformatiques n'excluent pas d'éventuelles phosphorylations par la protéine kinase C (S21, T107, S124) ou la caséine kinase II (T27, S51, S67, S129). De telles modifications ne peuvent cependant pas engendrer l'ajout de 4000 Da. C'est pourquoi l'étude de glycosylations de nature non glycosaminoglycannique a été entreprise. Un marquage à la biotine spécifique des glycoprotéines permet de supposer l'existence de glycanes. Les prédictions ne sont pas optimistes quant aux éventuelles N-glycosylations car aucune séquence consensus les permettant n'est réellement présente chez endocan. En revanche, rien ne peut être prédit pour les O-glycosylations. L'analyse par spectrométrie de masse du core protéique traité ou non aux N- ou O-glycosidases indique une réduction de la largeur du pic observé lors du traitement à la O-glycosidase, signe que d'éventuelles modifications de ce genre pourraient être présentes. De plus, lors de la caractérisation de la chaîne glycannique d'endocan par incorporation de glucosamine radioactive, des glycanes de faible poids moléculaire ont été observés en sortie de chromatographie d'exclusion. Ces glycanes ne se fixent pas à la résine DEAE, et ne semblent donc pas être de nature aussi anionique que le DS. Pour compléter cette étude sur les éventuelles glycosylations qui pourraient orner le core protéique d'endocan, une collaboration avec Marie-Christine Slomianny (laboratoire de Jean-Claude Michalski), spécialiste des O- et N-glycanes, a été établie. Cela a permis de révéler la présence d'au moins une O-glycosylation sur endocan et vraisemblablement d'une (ou plusieurs) N-glycosylation(s). D'après ces travaux, il est estimé que le core protéique d'endocan produit en cellule HEK293 possède au moins 1 % de glycanne (rapport massique).

Pour en connaître davantage sur l'organisation structurale de ce core protéique, des expériences de cristallographies ont été réalisées. Après le criblage de 1800 conditions de cristallographie, il s'est avéré qu'aucune d'entre elles n'a donné de résultats exploitables. Ce

fait peut être expliqué par l'hétérogénéité de la protéine qui se traduit en spectrométrie de masse par la large bande observée. Mais encore, de part sa constitution, il est fort probable que cette protéine soit composée d'une région structurée et globulaire dans les 110 premiers acides aminés grâce à la formation de ponts disulfures, et d'une autre région (en C-terminal), bien moins structurée et donc contraignante pour la cristallisation. C'est ce qui ressort de l'analyse bioinformatique du profil structural du polypeptide endocan. Pour avoir une idée de l'organisation structurale des 110 premiers acides aminés et, plus précisément, pour évaluer l'existence et les combinaisons des ponts disulfures possibles, un travail combinant digestion protéique et spectrométrie de masse a été réalisé. Ce travail permet de confirmer l'existence d'un motif EGF-like qui avait pu être mis à jour par bioinformatique. Ce motif EGF-like impliquerait la formation des ponts disulfures suivants : C46-C64 ; C58-C80 et C83-C92. En plus des ponts disulfures du motif EGF-like, les autres liaisons di-thiols les plus probables sont : C9-C32 ; C18-C35 ; C96-C103 ; C98-C110. En intégrant l'ensemble des résultats obtenus lors de ces premiers travaux de caractérisation, un modèle de l'organisation du polypeptide d'endocan a été proposé (Figure 28).

Ensuite, afin de compléter la caractérisation d'endocan, une étude structurale de la chaîne d'endocan a été entreprise. Dans un premier temps, la composition de la chaîne glycanique d'endocan provenant du système d'expression HEK293 ou de l'expression naturelle des cellules HUVECs a pu être étudiée par un marquage radioactif des chaînes. L'objectif était de vérifier que l'endocan obtenu en quantité plus importante (grâce au système de production mis en place dans les cellules HEK293) avait une chaîne glycanique quasiment identique à celle de l'endocan "naturel" de cellules endothéliales (il faut noter que lorsque nous avons commencé en 2003, nous n'avions pas encore connaissance de l'expression d'endocan dans des lignées cellulaires tumorales). Néanmoins, les résultats obtenus sont assez surprenants. A l'exception de la longueur de la chaîne glycanique, qui est en moyenne de 15 kDa pour l'endocan issu des cellules HUVECs et de 40 kDa pour l'endocan issu des HEK293, la composition structurale de deux chaînes glycaniques est quasiment identique. La chaîne glycanique d'endocan est ainsi composée de 30 % d'acides iduroniques avec une majorité de disaccharides mono-4-O-sulfatés (50 %) et mono-6-O-sulfatés (22 %), 10 % des disaccharides ne sont pas sulfatés et environ 18 % sont disulfatés, ce qui diffère des résultats obtenus en 2001 par Bécharde et al. Cependant, au regard de ces résultats, il est normal de se demander si c'est le core protéique d'endocan qui guide de manière très spécifique la composition structurale de sa chaîne glycanique ou si ces résultats sont le fruit du hasard. Afin d'en apprendre davantage sur l'organisation structurale de cette chaîne glycanique, une analyse

par RMN a également été entreprise. Les quantités requises pour ces analyses sont telles, que seule la chaîne glycanique d'endocan purifié à partir du système d'expression HEK293 a pu être analysée. Il a, de cette manière, été possible de confirmer que 30 % de la chaîne d'endocan est composé d'acides iduroniques. Par ailleurs, aucune séquence du type IdoA-GalNAc,4OH n'a été visualisée, mais, en revanche, la présence de séquences IdoA-GalNAc,4SO<sub>3</sub> a pu être observée. En intégrant l'ensemble des résultats et données structurales obtenus, un modèle de l'organisation de la chaîne glycanique d'endocan a été proposé (Figure 14).

L'appréhension de l'organisation structurale d'endocan était un élément important pour mieux comprendre les fonctions de ce PG. En effet, les fonctions d'endocan restent encore mal connues et n'ont pour l'instant pas pu encore être clairement définies. L'endocan semble jouer un rôle dans certains processus inflammatoires ainsi que dans certains cancers (Scherpereel et al., 2006 ; Sarrazin et al., 2006). Tout d'abord, il est possible qu'endocan puisse avoir un rôle dans l'inflammation. En effet, il a été démontré que l'expression d'endocan est sous le contrôle de cytokines pro-inflammatoires que sont le TNF $\alpha$ , l'IL-1 et l'IFN $\gamma$ . De plus, dans les différentes phases de la septicémie, le taux d'endocan circulant varie énormément avec un taux maximal dans les phases les plus graves (Scherpereel et al., 2006). Il a également été démontré *in vivo*, lors de l'induction d'un choc septique, que l'injection d'endocan permettait une survie des souris de l'ordre de 40 % (Depontieu, 2002). S'ajoute à cela le fait qu'endocan est capable d'inhiber l'interaction LFA-1/ICAM-1 (Bécharde et al., 2001). Toutes ces observations sont autant d'éléments qui permettent d'imaginer qu'endocan doit avoir un rôle dans les processus inflammatoires.

Ensuite, plusieurs éléments permettent d'imaginer qu'endocan participe au développement de cancers. Tout d'abord, en explorant la variation du taux d'ARNm d'endocan, de nombreuses études ont montré que l'augmentation de la quantité d'ARNm d'endocan décelable est corrélée avec l'apparition de métastases et un mauvais pronostic dans les cancers du sein, du rein et du poumon. Dans le cancer du sein, l'augmentation du taux d'ARNm d'endocan est synonyme de décès dans les 5 ans. De même, l'augmentation du taux d'ARNm d'endocan dans les cellules de carcinomes rénaux est en relation avec la vascularisation tumorale et l'inflammation du cancer du rein. Cela est conforté par une autre étude sur le même type de cancer qui met en évidence qu'endocan est l'un des gènes qui est le plus surexprimé par les cellules cancéreuses rénales, au même titre que ceux du VEGF et de c-Met (récepteur du HGF/SF). Dans le cancer du poumon, l'endocan a été catégorisé comme l'une

des protéines permettant d'émettre un mauvais pronostic ainsi qu'un risque important de décès dans les 12 mois (Figure 18). Certaines études révèlent un lien étroit entre l'expression d'endocan et les facteurs angiogéniques notamment dans les cancers broncho-pulmonaires où le taux d'endocan circulant atteint des valeurs maximales pour un pronostic très mauvais de la maladie (Grigoriu et al., 2006). L'expression d'endocan semble d'ailleurs être sous le contrôle de VEGF et/ou FGF-2, deux facteurs pro-tumoraux (Rennel et al., 2007). Dans un modèle *in vivo*, l'injection de cellules HEK293 exprimant de l'endocan recombinant, entraîne, chez les souris, la formation de tumeurs (Scherpereel et al., 2003). L'utilisation d'anticorps ayant pour cible la région de l'exon II riche en phénylalanine permettrait de réduire le développement tumoral ainsi induit. Il a également été montré que, via sa chaîne glycanique, l'endocan était capable de réguler l'activité d'HGF/SF, suggérant un rôle potentiel d'endocan dans le contrôle de la croissance tumorale au même titre que dans l'angiogenèse lors de la progression tumorale.

De ces observations, qui sont autant de pistes à approfondir pour cerner les fonctions de ce jeune PG, ont été initiés plusieurs travaux. Tout d'abord, dans le cadre du développement d'un kit de dosage d'endocan, la détermination de l'affinité des anticorps utilisés se devait d'être évaluée rapidement. Le développement d'anticorps anti-endocan potentiellement thérapeutiques a également nécessité une mesure de leurs affinités vis-à-vis d'endocan. Ainsi, les affinités des anticorps EP236, EP208 et EP214 ont pu être mesurées grâce au système SPR Biacore (plutôt que par le système IAsys probablement en raison des problèmes d'organisation structurale) et sont respectivement de 26, 78 et 19 nM.

Dans la continuité des travaux qui avaient été initiés avec l'HGF/SF, l'utilisation du Biacore nous a permis d'évaluer si certains médiateurs pro-inflammatoires ou pro-tumoraux interagissaient avec endocan. Après de nombreux essais d'immobilisation d'endocan en Biacore, l'utilisation d'un détergeant cationique (CTAB), utilisé depuis peu pour immobiliser des fragments d'ADN, a permis d'immobiliser endocan, son core protéique et sa chaîne de DS séparément et, de faire ainsi la première évaluation actuellement disponible de leurs différentes interactions avec les médiateurs criblés. De ce criblage de molécules, les médiateurs tels que MCP1, MIP1 $\alpha$ , IP10, MIP3 $\alpha$ , GRO/MGSA, IGF-I, SDF-1 $\alpha$  et TNF $\alpha$  n'ont donné aucun signal d'interaction avec l'endocan. En revanche, les médiateurs suivants : VEGF, SDF-1 $\beta$ , RANTES, MCP2, PF-4, IFN $\gamma$ , KGF (FGF-7), FGF-2 et bien entendu HGF/SF ont tous donné un signal d'interaction avec l'endocan ou sa chaîne de DS immobilisée. Les résultats obtenus dans cette étude originale sont, d'une manière générale, en accord avec les données de la littérature actuelle concernant les HS ou les DS. De manière

plus surprenante cependant, pour chacun de ces ligands identifiés (sauf pour HGF/SF), une interaction avec le core protéique d'endocan seule a également été observée. Par la méthode de Scatchard, seules les affinités d'HGF/SF et IFN $\gamma$  ont pu être estimées à 0,5 et 11 nM, respectivement. En utilisant la méthode du GMSA, il est apparu qu'une interaction entre endocan et le SDF-1 $\alpha$  pouvait potentiellement exister, mais aucune valeur d'affinité ne peut être donnée par cette méthode uniquement qualitative.

Afin d'approfondir l'effet modulateur que peut avoir endocan sur certains médiateurs, des essais cellulaires *in vitro* ont été mis en place. La cytokine IFN $\gamma$  est une protéine clé dans les phénomènes inflammatoires et les résultats du criblage par Biacore ont permis de définir clairement qu'une interaction avec la chaîne glycanique d'endocan avait lieu. L'effet qu'endocan pouvait avoir *in vitro* sur cette cytokine a donc été testé. Il est apparu que, via sa chaîne de DS, l'endocan était capable d'avoir un effet inhibiteur aussi important que l'Hp vis-à-vis d'IFN $\gamma$  dans des essais mesurant l'effet d'endocan sur la signalisation cellulaire d'IFN $\gamma$  et sur l'expression des molécules HLA-DR et ICAM-1 induite par la cytokine. Ces effets observés sont des éléments importants qui restent néanmoins pour l'instant dans le cadre de tests *in vitro* dépendants d'une seule lignée cellulaire. De telles observations nous indiquent cependant clairement qu'endocan pourrait certainement être un acteur majeur d'événements inflammatoires.

Il avait été initialement établi qu'endocan était un PG sécrété par les cellules endothéliales uniquement. Depuis, plusieurs études ont montré que l'endocan pouvait être exprimé par des adipocytes, des cellules tumorales de mélanomes et certaines cellules tumorales rénales (Wellner et al., 2003 ; Hendrix et al., 2003 ; Seftor et al., 2002 ; Ross et al., 2000 ; Abid et al., 2006). Afin de compléter ces données, des études ont été réalisées par RT-PCR et par ELISA pour mesurer l'expression d'endocan dans différentes lignées cellulaires. L'ARNm d'endocan a, de cette manière, été trouvé dans les cellules épithéliales issues d'une tumeur de la prostate hormonodépendante (DU145) et de cellules de rein (HEK293), dans les cellules issues de fibrosarcomes (HT1080), dans les cellules de cancers issus de tissus cérébraux, comme les astroglioblastomes (T98G), et surtout les glioblastomes (U118MG) (Adam et al., en préparation), dans les cellules issues de tératocarcinome (NT2), ainsi que dans les fibroblastes normaux adultes (NHDF). La détection de l'endocan par ELISA a notamment pu être faite de manière très significative dans les cellules HT1080 et U118MG, mais non chez les T98G. Pour les cellules NT2, DU145 et HEK293, la détection de protéine n'a pas encore été réalisée. Ces données supplémentaires indiquent que les fonctions d'endocan doivent être appréhendées d'une manière plus large que cela avait été fait jusqu'à présent.

Pour résumer, les travaux entrepris ont permis :

- de mettre en place, au sein d'Endotis Pharma, des plateformes de production et de purification permettant d'obtenir de l'endocan ou de son core protéique à l'échelle du milligramme par semaine et avec un bon degré de pureté
- la caractérisation du core protéique d'endocan, avec la mise en évidence
  - de glycosylations autres que la chaîne de DS (O-glycosylations et peut être N-glycosylations) sur de l'endocan recombinant issu de cellules HEK293
  - d'un motif EGF-like
  - de l'organisation des ponts disulfures
  - d'éventuels sites de phosphorylation
- la caractérisation de la chaîne glycanique d'endocan
  - par chromatographie afin de comparer la composition structurale de la chaîne glycanique d'endocan issue d'une production naturelle (HUVEC) ou d'un système de surexpression (HEK293)
  - par RMN pour connaître l'enchaînement structural des disaccharides
  - qui a ainsi donné naissance à un modèle de l'organisation structurale de cette chaîne (Figure 14)
- de mesurer des interactions
  - avec des anticorps par SPR en flux continu
  - avec des ligands potentiels, par GMSA, et surtout par SPR en distinguant les interactions avec le core protéique et la chaîne de DS
- d'évaluer *in vitro* l'effet d'endocan sur la cytokine IFN $\gamma$
- d'élargir les connaissances sur les lignées cellulaires exprimant endocan et plus globalement sur les tissus exprimant endocan.

A partir de ces résultats, plusieurs études complémentaires peuvent être proposées.

Concernant la caractérisation structurale du core protéique d'endocan, il peut être envisagé de poursuivre le criblage des conditions de cristallisation. Pour cela, il semble important que le core protéique soit totalement dépourvu de ses glycosylations. Dans cette optique, nous avons réfléchi, avec le groupe du Dr Philippe Lassalle, à un système d'expression chez E.coli. Cependant, ne possédant pas de test d'activité d'endocan et donc par crainte que les

nombreux ponts disulfures ne se forment pas de la même manière qu'en système eucaryote, cette piste n'a pas été poursuivie. La déglycosylation complète d'endocan pourrait donc passer soit par la création de mutants, soit par une suppression enzymatique de ces glycanes. Il apparaît également important dans la procédure de cristallisation de poursuivre nos travaux concernant la stabilisation de la partie C-terminale d'endocan avec un Fab, comme cela a été initié. Il peut être également proposé de tenter la cristallogénèse du core protéique privé de sa région C-terminale puisque celle-ci semblerait moins globulaire, ou plus labile et désordonnée d'après les prédictions bio-informatiques.

Ensuite, la création de mutants sur les acides aminés que l'on suppose glycosylés d'après nos analyses, pourrait être également envisagée. De plus, l'endocan recombinant ne possédant pas l'exon II, construit produit par le Dr. Florence Depontieu (laboratoire de Philippe Lassalle) pourrait être un outil très intéressant pour cibler l'emplacement de ces glycosylations.

Enfin, l'étude des sites potentiels de phosphorylation sur le core protéique d'endocan pourrait être entreprise grâce à l'utilisation d'anticorps spécifiques. Cela pourrait être un élément précieux pour compléter d'éventuelles pistes de recherche sur les fonctions d'endocan.

En ce qui concerne la chaîne de DS d'endocan, on ne peut pas dire actuellement qu'endocan est capable de guider les enzymes de modification des CS pour posséder une chaîne de composition identique quelque soit la lignée cellulaire qui le produit. Seule une analyse de la composition oligosaccharidique des DS d'endocan radiomarqué, obtenue de cellules exprimant naturellement endocan comme les cellules HUVEC (contrôle), les cellules tumorales U118MG ou encore des cellules HT1080, permettrait d'évaluer si la composition du DS est identique pour chaque lignée. Dans le même but d'étudier la spécificité structurale de la chaîne glycanique d'endocan, une étude comparative de la composition structurale de la chaîne DS d'endocan avec d'autres CS/DS de PG exprimés par les mêmes lignées cellulaires pourrait être très informative.

Ensuite, la seconde question à laquelle il serait intéressant, d'un point de vue fonctionnel de répondre concerne la modulation de l'expression d'endocan par le  $TNF\alpha$ , l'IL-1, l'IFN $\gamma$  ou le VEGF. Est-ce que ces médiateurs ont une influence directe sur la composition de la chaîne DS d'endocan ? Pour cela, une analyse de la composition de la chaîne DS d'endocan radiomarquée, obtenue d'une seule lignée cellulaire exprimant naturellement endocan mais stimulée par l'un ou différents facteurs permettrait d'y répondre. Il a, par exemple, été observé que dans des fibroblastes, un traitement par l'IFN $\gamma$  réduisait le taux d'IdoA dans les chaînes

de DS (Praillet et al., 1996). Réaliser ce genre d'étude avec l'endocan et un DSPG témoin pourrait s'avérer intéressant pour comprendre la régulation d'endocan observé dans certaines pathologies. De même, comme cela a déjà été décrit dans d'autres systèmes (Carter et al., 2003), l'étude de l'influence de ces facteurs sur les enzymes de synthèse ou les modifications des DS pourrait être également informative dans le cas de l'endocan.

Ensuite, un travail complémentaire par RMN de la chaîne DS d'endocan fragmentée, permettrait, par l'assemblage d'informations séquentielles, de vérifier l'exactitude du modèle d'organisation structurale de la chaîne d'endocan que nous avons proposé dans ce travail.

Concernant les fonctions d'endocan, il convient d'un point de vue biochimique de continuer le criblage des médiateurs pouvant être sous le contrôle d'endocan et de caractériser précisément la nature de ces interactions. Cela doit passer par une validation de la méthode au Biacore, c'est-à-dire essentiellement vérifier que les interactions sont bien spécifiques, et tout particulièrement celles impliquant le core protéique d'endocan. Ensuite, la création d'une librairie d'oligosaccharides issus de la fragmentation du DS d'endocan pourrait nous permettre de caractériser plus précisément les séquences glycaniques importantes dans la liaison avec certains de ces médiateurs. De la même manière, l'utilisation d'anticorps monoclonaux anti-endocan pourraient permettre la localisation sur le core protéique des sites impliqués dans l'interaction avec certains médiateurs.

D'un point de vue plus biologique, il faudrait, pour chaque facteur avec lequel endocan interagit, mettre en place un test d'activité fonctionnelle qui permettrait de mesurer l'effet du PG sur ces médiateurs, comme cela a pu être fait vis-à-vis de l'IFN $\gamma$ .

**ETUDE STRUCTURALE ET  
FONCTIONNELLE DE L'INTERACTION  
IFN $\gamma$ /HS : DEVELOPPEMENT D'UN MIME  
OLIGOSACCHARIDIQUE MODULANT  
L'ACTIVITE D'IFN $\gamma$**

# **1. 2O<sub>10</sub> : UN MIMETIQUE DES CHAINES HEPARANES SULFATES POUR MODULER L'ACTIVITE D'IFN $\gamma$**

Après avoir caractérisé le motif "glycosaminoglycannique" des HS reconnu par l'IFN $\gamma$ , une stratégie de synthèse chimique a été développée pour valider la structure du site d'interaction (Lubineau et al., 2004). Pour cela des oligosaccharides de différentes familles (tétra, hexa, ou octasaccharide) ont été liés entre eux par différents bras espaceurs. Le travail ci-après présenté sous forme de publication, met en évidence que c'est la molécule 2O<sub>10</sub> (deux octasaccharides reliés par un bras de dix unités de polyéthylène glycol) qui permet à la fois de mimer au mieux le motif HS impliqué dans l'interaction avec la cytokine, et de moduler *in vitro* l'activité de l'IFN $\gamma$ . Les travaux réalisés ont montré que 2O<sub>10</sub> inhibait de façon très efficace l'interaction d'IFN $\gamma$  avec les HS ainsi qu'avec le récepteur cellulaire IFN $\gamma$ R.

2O<sub>10</sub> représente donc une structure de base intéressante, à partir de laquelle pourrait être développé un inhibiteur de l'IFN $\gamma$ , utilisable dans les désordres auto-immuns ou inflammatoires.

## Heparan Sulfate Mimicry

### A SYNTHETIC GLYCOCONJUGATE THAT RECOGNIZES THE HEPARIN BINDING DOMAIN OF INTERFERON- $\gamma$ INHIBITS THE CYTOKINE ACTIVITY\*<sup>‡</sup>

Received for publication, July 15, 2005, and in revised form, August 29, 2005. Published, JBC Papers in Press, September 9, 2005, DOI 10.1074/jbc.M507729200

Stéphane Sarrazin<sup>‡</sup>, David Bonnaffé<sup>§</sup>, André Lubineau<sup>§</sup>, and Hugues Lortat-Jacob<sup>†1</sup>

From the <sup>†</sup>Institut de Biologie Structurale, Commissariat à l'Energie Atomique-CNRS, Université Joseph Fourier, Unité Mixte de Recherche 5075, 41 Avenue Horowitz, 38027 Grenoble Cedex 01, France and the <sup>‡</sup>Laboratoire de Chimie Organique Multifonctionnelle, CNRS-Université Paris Sud, Unité Mixte de Recherche 8614, 91405 Orsay Cedex, France

Cell-associated heparan sulfate (HS) is endowed with the remarkable ability to bind numerous proteins. As such, it represents a unique system that integrates signaling from circulating ligands with cellular receptors. This polysaccharide is extraordinary complex, and examples that define the structure-function relationship of HS are limited. In particular, it remains difficult to understand the structures by which HS interact with proteins. Among them, interferon- $\gamma$  (IFN $\gamma$ ), a dimeric cytokine, binds to a complex oligosaccharide motif encompassing a *N*-acetylated glucosamine-rich domain and two highly sulfated sequences, each of which binds to one IFN $\gamma$  monomer. Based on this template, we have synthesized a set of glycoconjugate mimetics and evaluated their ability to interact with IFN $\gamma$ . One of these molecules, composed of two authentic *N*-sulfated octasaccharides linked to each other through a 50-Å-long spacer termed 2O<sub>10</sub>, displays high affinity for the cytokine and inhibits IFN $\gamma$ -HS binding with an IC<sub>50</sub> of 35–40 nM. Interestingly, this molecule also inhibits the binding of IFN $\gamma$  to its cellular receptor. Thus, in addition to its ability to delocalize the cytokine from cell surface-associated HS, this compound has direct anti-IFN $\gamma$  activity. Altogether, our results represent the first synthetic HS-like molecule that targets a cytokine, strongly validating the HS structural determinants for IFN $\gamma$  recognition, providing a new strategy to inhibit IFN $\gamma$  in a number of diseases in which the cytokine has been identified as a target, and reinforcing the view that it is possible to create "tailor-made" sequences based on the HS template to isolate therapeutic activities.

Current research increasingly implicates heparan sulfate (HS),<sup>2</sup> a highly sulfated glycosaminoglycan present in the extracellular matrix and at the cell surface, in a plethora of phenomena that include cell proliferation, cell adhesion, matrix assembly, chemotaxis, inflammation, immune response, development, lipid metabolism, angiogenesis, wound healing, and viral attachment (1, 2). Mechanistically, this extensive functional repertoire often relies on the ability of HS to rec-

ognize diverse proteins, the conformation, stability, local concentration, or biological activities of which are modified by the interaction (3–7).

Consistent with its wide protein binding activity, HS is structurally complex. It is composed of strongly anionic domains enriched in *N*-sulfated glucosamines and iduronic acids, typically 3–8 disaccharides long (referred to as NS or heparin-like domains), that bear a variable number of *O*-sulfate moieties and are hypervariable in sequence. These domains are separated by relatively regular regions encompassing a larger area that contain predominantly *N*-acetylated glucosamine and glucuronic acid domains (NA domains) and mixed NA/NS regions that make the transition between NA and NS domains (8). It has been thought that specific information for protein recognition resided within the NS domains in HS and, indeed, a large number of "heparin-binding proteins" interact with such structures. These include, for example, fibroblast growth factors (9), stromal cell-derived factor-1 (10), herpes simplex virus type 1 glycoprotein D (11), or antithrombin III (12). However, of the large number of heparin-protein complexes that can be experimentally demonstrated, the latter (heparin-antithrombin III) is the only one to date for which a specific sequence has been formally defined (13) and reproduced by chemical synthesis (14). Importantly, this achievement led to the development of an approved drug against deep venous thrombosis.

Interferon- $\gamma$  (IFN $\gamma$ ), a multifunctional T cell-secreted cytokine (15, 16), has been identified as a heparin-binding protein some years ago (17). IFN $\gamma$  and growth factors (such as fibroblast growth factors) belong to distinct groups in regard to their regulation by HS. In particular, HS does not promote IFN $\gamma$  association to its cell surface receptor (IFN $\gamma$ R), and IFN $\gamma$ R does not bind to HS.<sup>3</sup> IFN $\gamma$  binding to HS was found to control the blood clearance, subsequent tissue targeting, and local accumulation of the cytokine. It also regulates IFN $\gamma$  activity by a unique mechanism involving a controlled processing of the carboxyl-terminal peptide (18, 19).

In contrast to the above mentioned heparin-binding proteins, IFN $\gamma$  does not interact with isolated NS domains. The binding requires a larger sequence that encompasses an internal NA domain flanked at both sides by two NS domains (20). In such a structure the two external NS regions are believed to interact with the two carboxyl-terminal sequences of an IFN $\gamma$  dimer and bridge the two IFN $\gamma$  monomers by virtue of the internal domain (Fig. 1). Because of the structural heterogeneity of HS, such complex binding motif is obviously impossible to obtain in pure form and in a large quantity from natural sources. We thus used a chemical approach that has been recently developed (21) to obtain compounds that would mimic the IFN $\gamma$  binding site with the goal of validating the model of the IFN $\gamma$ -HS complex initially proposed and obtaining material in homogeneous form and a large amount.

\* This work has been supported by the Commissariat à l'Energie Atomique, the CNRS (Contrat Physique et Chimie du Vivant), the Université Paris-Sud, and la Région Rhône-Alpes (Programme Emergence). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

<sup>‡</sup> The on-line version of this article (available at www.jbc.org) contains supplemental Fig. 1 (<sup>1</sup>H NMR spectra of glycoconjugates 2O<sub>1</sub>, 2O<sub>10</sub>, and 2O<sub>22</sub>) and supplemental Fig. 2 (<sup>1</sup>H-<sup>1</sup>H COSY and <sup>1</sup>H-<sup>13</sup>C HSQC (heteronuclear single quantum coherence) of compound 2O<sub>10</sub>).

<sup>1</sup> To whom correspondence should be addressed. Tel: 33-438-784-4-85; Fax: 33-438-785-4-94; E-mail: Hugues.Lortat-Jacob@ibs.fr.

<sup>2</sup> The abbreviations used are: HS, heparan sulfate; IFN $\gamma$ , interferon- $\gamma$ ; IFN $\gamma$ R, IFN $\gamma$  receptor; NA, *N*-acetylated glucosamine- and glucuronic acid-containing (region); NS, *N*-sulfated glucosamine- and iduronic acid-enriched (region); RU, resonance units.

<sup>3</sup> S. Sarrazin, D. Bonnaffé, A. Lubineau, and H. Lortat-Jacob, unpublished observation.

A set of compounds in which authentic NS domains, ranging from tetrasaccharides to octasaccharides, linked to poly(ethylene glycol)-based spacers of different lengths that would mimic the internal NA domain was produced by chemical synthesis, and the ability of the compounds to interact with IFN $\gamma$  was analyzed. One of these molecules inhibited the IFN $\gamma$ /heparin interaction with an IC $_{50}$  of 35–40 nM. Importantly, this compound also inhibited the binding of IFN $\gamma$  to IFN $\gamma$ R and, consequently, the biological activity of the cytokine. These results provide a new strategy to inhibit IFN $\gamma$  in a number of diseases in which this cytokine has been identified as a target and strongly validate the model proposed for the IFN $\gamma$ -HS complex.

## EXPERIMENTAL PROCEDURES

**IFN $\gamma$  Production and Characterization**—Human IFN $\gamma$  cDNA was cloned into a pET11a expression vector (Novagen) and used to transform *Escherichia coli* strain BL21 Star DE3 (Invitrogen). Cells were grown at 37 °C in Luria broth medium containing 100  $\mu$ g/ml ampicillin and induced with 0.5 mM isopropyl 1-thio- $\beta$ -D-galactopyranoside for 5 h. Purification from inclusion bodies was performed as described (22) with slight modifications. Briefly, inclusion bodies were solubilized in 6 M guanidine HCl, and the protein was refolded by dilution to 0.2 mg/ml into 50 mM phosphate buffer and 0.5 M guanidine HCl, pH 7. IFN $\gamma$  was purified by ion exchange (Mono S HR 5/5 column) and gel filtration (Superdex 75 column) chromatography (Amersham Biosciences) and stored frozen in 10 mM Tris and 10 mg/ml mannitol, pH 6.8. Purified material was characterized by mass spectrometry and amino-terminal sequencing and quantified by amino acid analysis. Biological activities of IFN $\gamma$  (antiviral activity and up-regulation of HLA-DR antigen; see below) were found to be identical to that of commercially available IFN $\gamma$  (Promega), and both samples were used in this study.

**Synthesis of Oligosaccharides and Glycoconjugates**—The general approach used to synthesize HS oligosaccharides and glycoconjugate mimetics is illustrated in Fig. 2 and briefly described here (a fully detailed description of each synthetic step can be found in Ref. 21). A disaccharide building block with orthogonal allyl and *para*-methoxybenzyl protecting groups on the anomeric and 4' position, respectively, was first prepared as described (23, 24). The allyl group was cleaved (25) followed by the activation of the anomeric position in the form of a trichloroacetimidate derivative to obtain a disaccharide donor (88% yield). Treatment of the same building block with 2,3-dichloro-5,6-dicyano-1,4-benzoquinone gave rise to a disaccharide acceptor (83% yield). Donor and acceptor were then condensed in dichloromethane at -40 °C, using *tert*-butyldimethylsilyl triflate as promoter, to yield the desired tetrasaccharide with total  $\alpha$  stereoselectivity in 90% isolated yield. Using the same procedure, the tetrasaccharide was converted into tetrasaccharide acceptors and donors with respective yields of 81 and 87%. Hexasaccharides and octasaccharides were then obtained, using a 2 + 4 or 4 + 4 oligomerization strategy with total  $\alpha$  stereoselectivity. Oligosaccharides were then deacetylated, the azido groups were reduced, the resulting amino groups were sulfated along with the free hydroxyl groups, and the methyl esters were saponified. Hydrogenolysis of the oligosaccharides benzyl groups gave the non-conjugated tetrasaccharides, hexasaccharides, and octasaccharides in 85% to quantitative yields. To obtain mimetics of the IFN $\gamma$  binding site, oligosaccharides were conjugated, before hydrogenolysis, with bis-thio-poly(ethylene glycol) under UV light irradiation (360 nm) as described (26). The benzyl protecting groups were then removed by hydrogenolysis after oxidation of the thioether linkages using pH 7 buffered oxone (27), giving rise to the glycoconjugates 2T $_n$ , 2H $_n$ , and 2O $_n$  (where T, H, and O stand for tetrasaccharide, hexasaccharide, and octasaccharide, respectively,

and 2 indicates that two oligosaccharides are linked through a spacer having  $n$  equal to 5, 10, or 32 ethylene glycol repeats). All samples were sterilized by filtration through 0.22- $\mu$ m filters, quantified as described (28), and analyzed by polyacrylamide gel electrophoresis analysis (29),  $^1$ H NMR, and, for 2T $_5$ , 2H $_{32}$ , and 2O $_{10}$ , heteronuclear single quantum coherence experiments. Size-defined, heparin-derived oligosaccharides were also obtained by enzymatic depolymerization of pig intestinal mucosa heparin as described (10).

**Biotinylation of Heparin and Biacore-based Binding Assay**—Heparin, 1 mM in phosphate-buffered saline, was reacted for 24 h at room temperature with 10 mM biotin-LC-hydrazide. The mixture was extensively dialyzed against water to remove unreacted biotin and freeze-dried (10). Two flow cells of CM3 or CM4 sensor chips were activated with 50  $\mu$ l of 0.2 M *N*-ethyl-*N'*-(diethylaminopropyl)-carbodiimide and 0.05 M *N*-hydroxysuccinimide, after which 50  $\mu$ l of streptavidin (Sigma-Aldrich) at 0.2 mg/ml in 10 mM sodium acetate buffer, pH 4.2, was injected. Remaining activated groups were blocked with 50  $\mu$ l of ethanolamine (1 M), pH 8.5. Typically, this procedure allowed coupling 2000–2500 resonance units (RU) of streptavidin on both flow cells. Biotinylated heparin (5  $\mu$ g/ml) was captured to a level of 50 RU on one surface, and the other one was left untreated and served as negative control. Before use, the chip was submitted to several injections of Hepes-buffered saline containing 1.25 M NaCl and then washed by continuous flow of HBS-P buffer (10 mM Hepes, 0.15 M NaCl, and 0.05% P20 detergent, pH 7.4). For binding assays, IFN $\gamma$ , either alone or in combination with oligosaccharides, was simultaneously injected over both negative control and heparin surfaces for 5 min at 25 °C and 50  $\mu$ l/min. The heparin surface was regenerated with a 250  $\mu$ l pulse of 1.25 M NaCl.

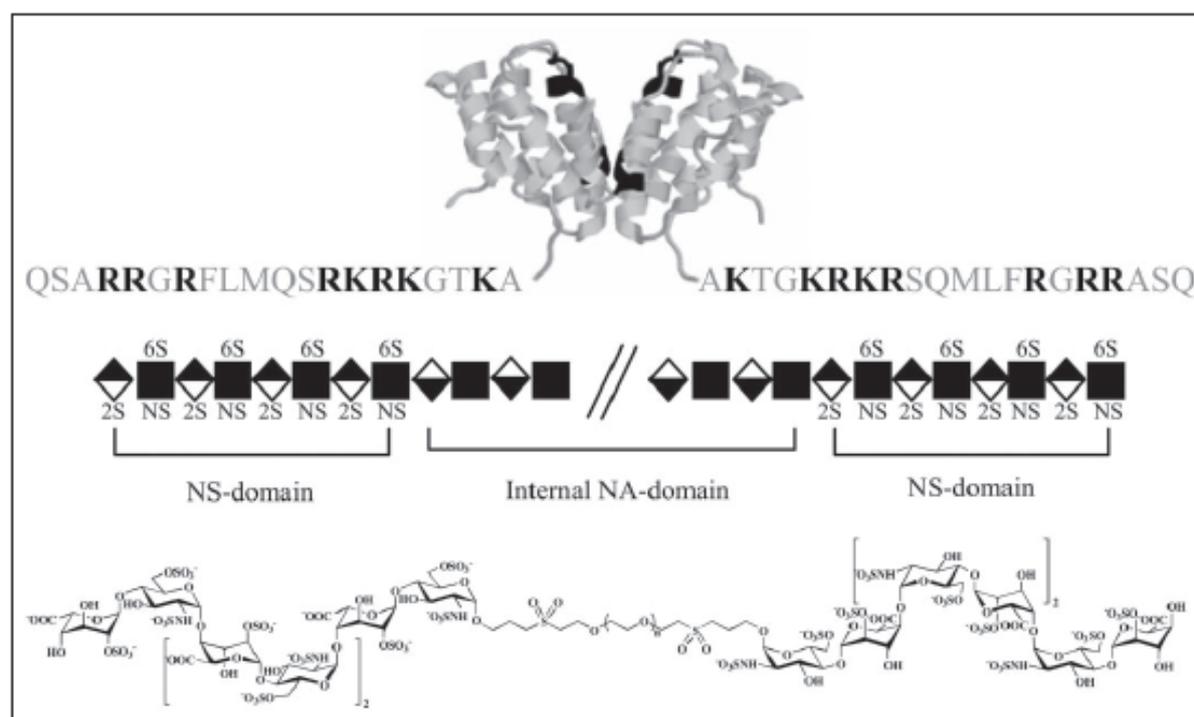
**Biotinylation of Soluble IFN $\gamma$  Receptor and Biacore Base Binding Assay**—Soluble IFN $\gamma$ R (R&D Systems) at 250  $\mu$ g/ml in 20 mM phosphate buffer, pH 6, was reacted for 20 min in the dark at 4 °C, with 10 mM sodium periodate to oxidize glycans of the molecule. The reaction was quenched with 15 mM glycerol, and the sample was dialyzed against the 20 mM phosphate buffer. Biotin-LC-hydrazide was then added to a concentration of 5 mM, and the mixture was incubated for 4 h at 4 °C. Ethanolamine (100 mM final) was then added to the sample, which was then extensively dialyzed against phosphate-buffered saline, pH 7.2. Biotinylated soluble IFN $\gamma$ R (40  $\mu$ g/ml) was captured on top of a streptavidin sensor chip (prepared as described above) to a level of 150–200 RU. Before use, the surface was conditioned with 10 2-min pulses of 10 mM HCl. For binding assay, samples (IFN $\gamma$  either alone or in combination with oligosaccharides) diluted in HBS-P buffer maintained at 25 °C were injected over the IFN $\gamma$ R surface at a flow rate of 50  $\mu$ l/min for 5 min. The IFN $\gamma$ R surface was regenerated with 10 mM HCl.

**IFN $\gamma$  Bioassays**—Antiviral activity was determined in triplicate with a standard microtiter inhibition of cytopathic effect assay against the vesicular stomatitis virus on monolayers of WISH cells essentially as described (30). An IFN $\gamma$ -induced HLA-DR antigen was measured with an enzyme-linked bio-immunoassay (31). For that purpose, Colo 205 cells seeded in a 96-well microtiter plate were treated with IFN $\gamma$  during a 4-h period. After 72 h of additional culture, cells were fixed with ice-cold ethanol, and HLA-DR antigen was detected with an anti-HLA-DR monoclonal antibody (1  $\mu$ g/ml; BD Biosciences) followed by a goat anti-mouse IgG-horseradish peroxidase conjugate (1:2500; Amersham Biosciences), both in Hanks' buffer sodium salt containing 0.02% Tween 20 and 1% bovine serum albumin.

## RESULTS

**Design, Synthesis, and Characterization of the Glycoconjugates**—Our previous studies (20) suggested that in the IFN $\gamma$ -HS complex two NS

## Interferon- $\gamma$ Binding Site Structure In Heparan Sulfate

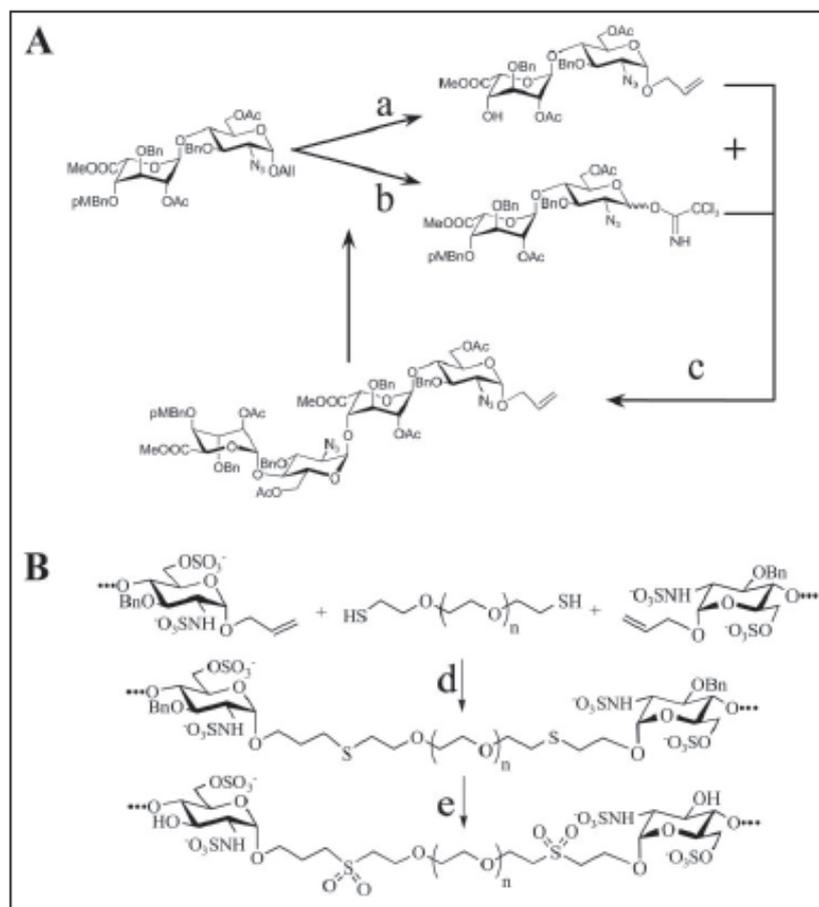


**FIGURE 1.** A model of the IFN $\gamma$ -HS complex. IFN $\gamma$  is a homodimer that contains four clusters of basic residues (boldfaced), two of which (Lys-Thr-Gly-Lys-Arg-Lys-Arg and Arg-Gly-Arg-Arg) are located in the unstructured carboxyl-terminal end of the protein, that function as HS binding sites. The motif that IFN $\gamma$  recognizes along the HS chain consists of two NS domains that directly interact with the two carboxyl termini of the cytokine and are linked to each other by an internal NA domain. Among the nine glycoconjugates that have been synthesized to mimic such structure, the most active compound ( $n = 10$ ) is shown.

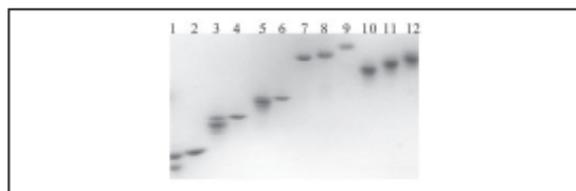
domains directly interact with two IFN $\gamma$  carboxyl termini and bridge the two cytokine monomers through an internal NA domain (Fig. 1). Although the atomic structure of IFN $\gamma$  has been solved by x-ray crystallography, the organization of the carboxyl termini, which are believed to be intrinsically disordered, is unknown downstream of the residue Ala-123 (32). Because the exact distance that spans the two heparin binding domains is thus not defined, we prepared synthetic glycoconjugates in which poly(ethylene glycol)-based spacers of different lengths were used to link chemically obtained NS domains. For that purpose, a disaccharide building block was first prepared and converted into either a disaccharide donor or a disaccharide acceptor, which were condensed together (2 + 2) to give a tetrasaccharide (T). A similar strategy (2 + 4 or 4 + 4) was used to prepare hexasaccharides and octasaccharides (Fig. 2A). The glycoconjugates 2T<sub>5</sub>, 2T<sub>10</sub>, 2T<sub>32</sub>, 2H<sub>5</sub>, 2H<sub>10</sub>, 2H<sub>32</sub>, 2O<sub>5</sub>, 2O<sub>10</sub>, and 2O<sub>32</sub> were obtained through conjugation of the oligosaccharides onto  $\alpha,\omega$ -bis-thio-poly(ethylene glycol) spacers of different lengths (Fig. 2B). <sup>1</sup>H NMR and heteronuclear single quantum coherence spectra were in accordance with the expected structures (data not shown). PAGE analysis (Fig. 3) showed that synthetic tetrasaccharides, hexasaccharides, and octasaccharides (lanes 2, 4, and 6) have the same migration pattern as the corresponding heparin-derived oligosaccharides (lanes 1, 3, and 5) but, in contrast to the natural size-purified mixtures, are homogenous. Conjugation to poly(ethylene glycol) linkers was clearly evidenced by the shift in migration observed for 2O<sub>5</sub> (Fig. 3, lane 7), 2O<sub>10</sub> (lane 8), and 2O<sub>32</sub> (lane 9), each of which migrated at the position of natural octadecasaccharides and larger. Similar migration changes were observed for tetrasaccharide and hexasaccharide conjugation (data not shown). In some case, conjugated oligosaccharides

migrated as apparent doublets on the gel. However, because NMR analyses of these molecules were consistent with single structures, this result should not be caused by sample heterogeneity (see supplemental data, available in the on-line version of this article).

**Inhibition of the IFN $\gamma$ -Heparin Interaction by the Glycoconjugates—**To investigate the ability of the 12 different synthetic compounds to interact with IFN $\gamma$ , an inhibition assay was set up in which the cytokine, either alone or coincubated with each of the 12 molecules (at 37.5, 75, or 150 nM) to be analyzed, was injected over both a heparin-functionalized sensor chip and a streptavidin sensor chip used as a control surface. Injection of IFN $\gamma$  over the heparin surface produced at equilibrium a binding response of 180 RU, whereas a response of 5 RU was observed over the streptavidin surface (data not shown). Results showed (Fig. 4) that whereas the tetrasaccharides and hexasaccharides were completely inactive, the octasaccharide slightly prevented the IFN $\gamma$ -heparin binding (no more than 20% inhibition was observed with the highest dose of octasaccharide), indicating that by themselves these oligosaccharides did not display significant affinity for the cytokine. Conjugation of tetrasaccharides or hexasaccharides to the different poly(ethylene glycol) linkers did not lead to significant inhibition activity (25–30% of inhibition at 150 nM; Fig. 4C). In contrast, octasaccharides conjugated to spacers with 5, 10, or 32 poly(ethylene glycol) repeats (2O<sub>5</sub>, 2O<sub>10</sub>, or 2O<sub>32</sub>) clearly inhibited the binding of IFN $\gamma$  to heparin, the most active molecule (2O<sub>10</sub>) displaying a 50% inhibition at ~35–40 nM (Fig. 4A). These results strongly suggest that the glycoconjugates function by bridging the two IFN $\gamma$  monomers and that proper spacing of the oligosaccharides are important for the binding process. Similar experiments



**FIGURE 2.** Oligosaccharides and glycoconjugates synthesis. **A**, synthesis of the protected tetrasaccharides, hexasaccharides, and octasaccharides. A single disaccharide building block was converted into a disaccharide acceptor (arrow *a*) or a donor (arrow *b*) and then condensed together (2 + 2) to give a tetrasaccharide (arrow *c*). This tetrasaccharide was similarly transformed into either a tetrasaccharide donor (arrow *a*) or acceptor (arrow *b*) that were then converted, using a 2 + 4 or 4 + 4 strategy (arrow *c*), into the expected hexasaccharide and octasaccharide. These compounds were treated as described under "Experimental Procedures" to give the non-conjugated tetrasaccharides, hexasaccharides, and octasaccharides. **B**, general strategy for the preparation of the glycoconjugates. The sulfated and benzylated tetrasaccharides, hexasaccharides, or octasaccharides were conjugated to bis-thio-poly(ethylene glycol) linkers of different lengths ( $n = 5, 10, \text{ or } 32$ ) under UV light irradiation (arrow *d*). The nine benzylated glycoconjugates thus obtained were debenzylated after oxidation of the thioether linkages to give the awaited nine deprotected glycoconjugates (arrow *e*).



**FIGURE 3.** PAGE analysis of oligosaccharides. Heparin-derived (size-purified mixture) or synthetic oligosaccharides (0.5  $\mu\text{g}$  each) were run through a 30% polyacrylamide gel and stained with azure A. Lane 1, heparin-derived tetrasaccharide; lane 2, synthetic tetrasaccharide; lane 3, heparin-derived hexasaccharide; lane 4, synthetic hexasaccharide; lane 5, heparin-derived octasaccharide; lane 6, synthetic octasaccharide; lane 7,  $2\text{O}_5$ ; lane 8,  $2\text{O}_{10}$ ; lane 9,  $2\text{O}_{32}$ ; lane 10, heparin-derived tetradecasaccharide; lane 11, hexadecasaccharide; lane 12, octadecasaccharide.

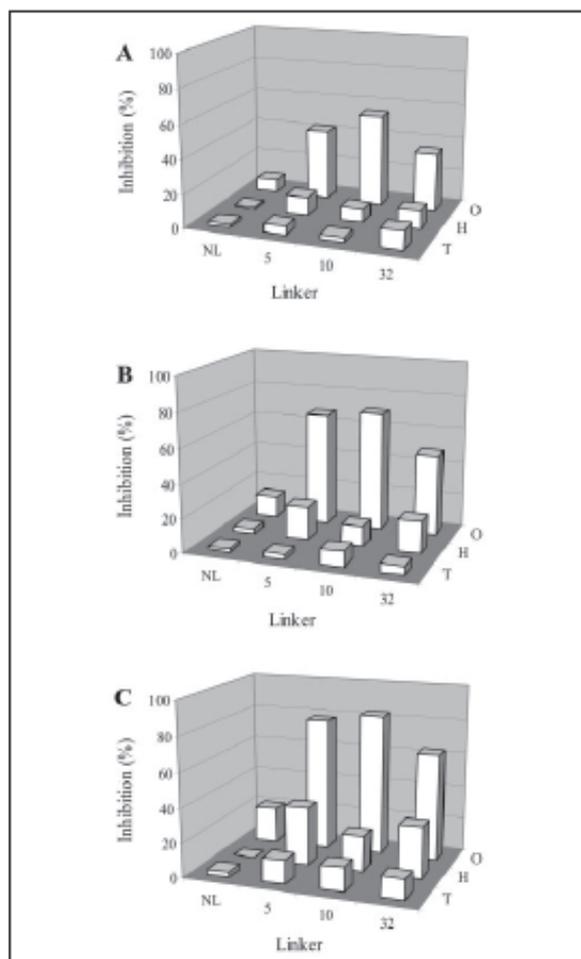
performed with 15- and 5-kDa heparin indicated that 20 and 130 nM were required, respectively, to produce 50% inhibition (data not shown).

**The Glycoconjugates Inhibit the IFN $\gamma$ -IFN $\gamma$ R Interaction**—A number of studies demonstrated that the IFN $\gamma$  carboxyl termini are critical for bioactivity (33). In particular, it has been suggested that residues 128–132 (Lys-Arg-Lys-Arg-Ser) within this domain of the cytokine are required for receptor binding, and a model has been proposed where this basic cluster is located near an acidic patch of the IFN $\gamma$ R (32). We thus investigated whether the synthetic oligosaccharides would inhibit the binding of the cytokine to its receptor. For that purpose, we immo-

bilized the ectodomain of the IFN $\gamma$ R on a Biacore sensor chip and analyzed the ability of the synthetic oligosaccharides to inhibit the IFN $\gamma$ -IFN $\gamma$ R interaction. IFN $\gamma$  was preincubated with increasing concentrations of the different molecules to be analyzed and then injected over the IFN $\gamma$ R surface (Fig. 5). The first two sets of sensorgrams show that unconjugated oligosaccharides (Fig. 5A) and conjugated tetrasaccharides (Fig. 5B) had no or minimal effect on the IFN $\gamma$ -IFN $\gamma$ R interaction. Conjugated hexasaccharides ( $2\text{H}_5$ ,  $2\text{H}_{10}$ , and  $2\text{H}_{32}$ ) yielded a modest inhibition (Fig. 5C) as did the conjugated octasaccharides  $2\text{O}_5$  and  $2\text{O}_{32}$  (Fig. 5D and E), whereas the  $2\text{O}_{10}$  had a marked effect on the binding process (Fig. 5F). The  $2\text{O}_{10}$  compound activity is similar to that of a 6-kDa heparin fragment.

**The  $2\text{O}_{10}$  Glycoconjugate Displays Anti-IFN $\gamma$  Activity**—In view of the above data, we investigated whether  $2\text{O}_{10}$  would inhibit IFN $\gamma$  activity. We thus measured the ability of  $2\text{O}_{10}$  to prevent the induction of the HLA-DR antigen by IFN $\gamma$ , using Colo 205 cells. We found that optimum HLA-DR antigen was observed 72 h after the initial stimulation with 5 ng/ml of IFN $\gamma$ , and these conditions were used in the following experiments. Cells were treated with IFN $\gamma$  that was preincubated with a range of concentration of  $2\text{O}_{10}$  (0–25  $\mu\text{g}/\text{ml}$ ), and HLA-DR expression was quantified 72 h later. Our data (Fig. 6A) showed that  $2\text{O}_{10}$  clearly displayed anti-IFN $\gamma$  activity, with an  $\text{IC}_{50}$  close to 1.5  $\mu\text{g}/\text{ml}$  (250 nM). Analysis of all the other glycoconjugates demonstrated that their anti-

## Interferon- $\gamma$ Binding Site Structure in Heparan Sulfate



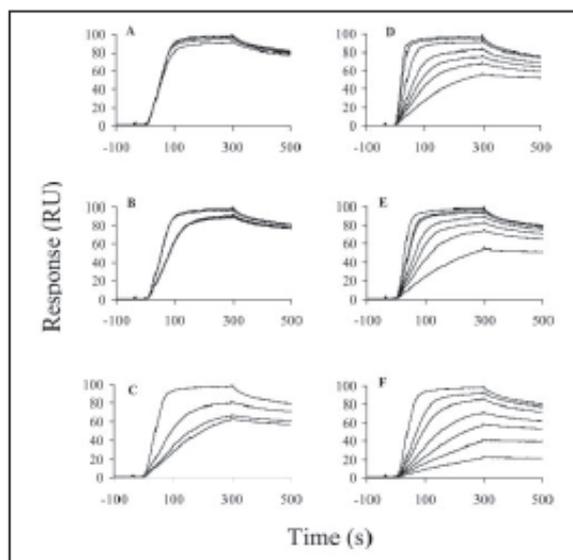
**FIGURE 4.** IFN $\gamma$ -heparin binding inhibition by synthetic oligosaccharides. IFN $\gamma$  (7.5 nM) was preincubated with 37.5 (A), 75 (B), or 150 nM (C) tetrasaccharide (T), hexasaccharide (H), or octasaccharide (O), either unconjugated (NL, no linker) or conjugated with linkers of increasing length ( $n = 5, 10,$  or  $32$  poly(ethylene glycol) repeats). Complexes were injected at  $50 \mu\text{l}/\text{min}$  over a heparin-activated sensor chip, and the binding response was recorded at the end of the injection phase. The response obtained in the absence of oligosaccharides (100%) was equal to 180 RU. Each experiment was repeated 2 or 3 times using two batches of independently synthesized molecules. The SE never exceeded 10% of the mean value.

IFN $\gamma$  activity closely match their ability to inhibit the IFN $\gamma$ -IFN $\gamma$ R interaction (Fig. 6B).

### DISCUSSION

Structural heterogeneity of HS is the basis of the many functions these molecules fulfill, in particular through their unique ability to interact with a large array of proteins (34). Despite increased interest in the field, progress has been hampered by the extraordinary complexity of HS. In particular, the characterization of the protein-HS interface and the isolation of the corresponding binding domains are very difficult, an aspect particularly prominent in the case of IFN $\gamma$ , the binding site of which is not simply contained within a single structural domain of the polysaccharide (20).

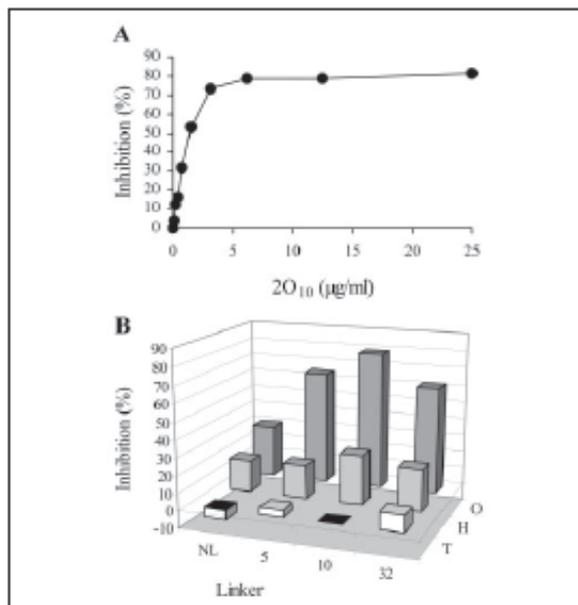
In this context, we chose a chemical strategy to investigate the general motif organization that IFN $\gamma$  recognizes on HS. According to the model



**FIGURE 5.** Inhibition of the IFN $\gamma$ -IFN $\gamma$ R interaction by synthetic oligosaccharides. IFN $\gamma$  (1.5 nM) was preincubated with a range of concentrations of the different oligosaccharides and injected for 5 min over an IFN $\gamma$ -activated sensor chip at  $50 \mu\text{l}/\text{min}$ . The binding responses (in RU) were recorded as a function of time and corresponded to the complexes made with 75 nM tetrasaccharides, hexasaccharides, and octasaccharides (A), 2T<sub>2</sub>, 2T<sub>10</sub>, and 2T<sub>32</sub> (B), 2H<sub>5</sub>, 2H<sub>10</sub>, and 2H<sub>32</sub> (C), or 75, 30, 15, 7.5, 3, 1.5 or 0 nM (from top to bottom curves in each panel) for 2O<sub>5</sub> (D), 2O<sub>10</sub> (E), or 2O<sub>10</sub> (F).

of the IFN $\gamma$ -HS complex (see Fig. 1) that was used as a working hypothesis, different heparin-like oligosaccharides were synthesized and linked to each other using molecular spacers of distinct lengths. Binding studies first pointed out that properly spaced interacting sequences acted in a concerted manner to form a functional unit, because none of the unconjugated oligosaccharides displayed efficient binding activity. Optimal activity was observed when a spacer of 10 poly(ethylene glycol) repeats ( $\sim 50$ -Å-long) was introduced between the two NS domains. Interestingly, the last two defined residues (A1a-123) in the IFN $\gamma$  crystallographic structure are at a distance of 23 Å. The 50-Å spacer should thus function by optimally presenting the two binding NS domains to the IFN $\gamma$  basic clusters (see Fig. 1), in contrast to the 5 and 32 poly(ethylene glycol) repeats that are 33- and 114-Å-long, respectively, and presumably might impose accommodations to meet an appropriate binding conformation. Our results also clearly demonstrated the importance of the oligosaccharide size for the binding, because the glycoconjugates containing tetra-NS or hexa-NS domains, although linked with the optimal spacer, were unable to significantly interact with IFN $\gamma$ . The 2O<sub>10</sub> molecule (two octasaccharides linked with the 10 poly(ethylene glycol) repeat spacer) thus appears to be the best mimic of the IFN $\gamma$  binding domain. Because its central domain is not HS derived, the 2O<sub>10</sub> molecule is much less charged than the natural equivalent molecule. However, its binding activity and its ability to inhibit IFN $\gamma$  are identical to that of heparin, demonstrating that it represents a fully functional binding unit. The C<sub>2</sub> symmetry of this compound may also contribute to its binding activity by fitting the symmetry of the IFN $\gamma$  dimer better than a natural single HS chain.

The achievement of such a molecule has considerable interests for different purposes. First, in contrast to natural derived compounds, this molecule is homogeneous and large amounts of it can be prepared. A possibility that arises from this quality is the development of structural studies of the cytokine in complex with the glycoconjugate, a point that



**FIGURE 6.** Synthetic oligosaccharides and glycoconjugates inhibit the IFN- $\gamma$  activity. **A**, monolayers of Colo 205 cells were treated with IFN- $\gamma$  (5 ng/ml), preincubated with a range of concentrations of 2O<sub>10</sub> (0–25  $\mu$ g/ml), and HLA-DR expression was quantified 72 h later. **B**, monolayers of Colo 205 cells were treated with IFN- $\gamma$  (5 ng/ml) that was preincubated or not with each of the 12 synthetic molecules (25  $\mu$ g/ml), and HLA-DR expression was quantified 72 h later. Results are expressed as a percentage of inhibition. T, tetrasaccharide; H, hexasaccharide; O, octasaccharide.

was not possible before. Like many disordered segments, the carboxyl-terminal peptide of IFN- $\gamma$  may fold on binding to its target (35) and thus might be characterized by x-ray crystallography. Because this part of the protein is a critical regulatory element of the cytokine activity (33), resolving its structure is an important issue. Secondly, the 2O<sub>10</sub> molecule is clearly a prototype with considerable potential for refinement. In particular, we have not yet addressed the importance of the sulfation profile within the NS octasaccharides. Combinatorial synthesis is a possible tool to generate distinct sulfo forms and/or an iduronic/glucuronic pattern (23) and should give rise to molecules with better defined and specific oligosaccharide sequences, a point that can now be investigated in the context of the 2O<sub>10</sub> scaffold.

Another important aspect of the present work is the observation that, once bound to the 2O<sub>10</sub> molecule, IFN- $\gamma$  did not interact anymore with its cell surface receptor and, consequently, is biologically inactive. The importance of the IFN- $\gamma$  carboxyl-terminal domain for receptor binding has been investigated by carboxyl-terminal deletions and site-directed mutagenesis (33). These studies all point out that preservation of the Lys-Arg-Lys-Arg-Ser sequence is critical for activity. Our data are thus consistent with the view that 2O<sub>10</sub> directly interacts with this sequence and blocks binding to both HS and IFN- $\gamma$ R. The fact that a carboxyl-terminal region of the cytokine can bind to two such unrelated molecules as 2O<sub>10</sub> and IFN- $\gamma$ R could be somewhat surprising. However, this part of the molecule is disordered, a feature that has been recently recognized as frequently involved in regulatory functions and that confers functional advantages to the protein, including the ability to bind in different conformations to different targets (35).

IFN- $\gamma$  is a key cytokine involved in the development and maintenance of the Th1 arm of the immune system (16). However, aside from normal functions in host defense, IFN- $\gamma$  production may also contribute to a

number of diseases characterized by pro-inflammatory and autoaggressive Th1 response (36). These include autoimmune pathologies such as rheumatoid arthritis, systemic lupus erythematosus, multiple sclerosis, insulin-dependent diabetes mellitus, psoriasis, and alopecia areata (37) or chronic inflammatory bowel diseases such as Crohn disease and ulcerative colitis (38). IFN- $\gamma$  is also a key contributor to hyperoxia-induced lung injury (39) and, because of its ability to induce a major histocompatibility complex in many tissues, is associated with graft rejection (40). This cytokine is thus considered as an attractive target for the treatment of such diseases, and strategies have been developed to antagonize its biological activity. So far these strategies have focused on the use of soluble IFN- $\gamma$ R (41, 42) or neutralizing antibodies, and in some cases these were found to offer substantial clinical benefits (37, 43). Humanized anti-IFN- $\gamma$  is also currently evaluated in Crohn disease (44). On the other side, administration of a HS binding cationic peptide (derived from the IFN- $\gamma$  carboxyl terminus) has delayed the time of rejection in a mouse model of allogenic skin flap transplantation (45). Most interestingly, the 2O<sub>10</sub> molecule described here inhibits the binding of IFN- $\gamma$  to both HS and IFN- $\gamma$ R, the two known ligands of the cytokine. Such a molecule should thus inhibit both bioactivity and local concentration within tissue and therefore should compare very favorably with neutralizing antibodies that have been selected for their ability to inhibit IFN- $\gamma$  bioactivity but not binding to HS.

There is enormous potential for the development of heparin-like structures as drugs for a range of diseases in addition to the current antithrombotic target (46, 47). Based on the 2O<sub>10</sub> scaffold, our future work will investigate whether such type of molecules could be of some interest in pathologies for which IFN- $\gamma$  has been identified as a target.

*Acknowledgments*—We thank Jean-Pierre Andrieu for amino-terminal sequencing and amino acid analysis, Bernard Dublet for mass spectrometry, and Romain Vives for reading of the manuscript.

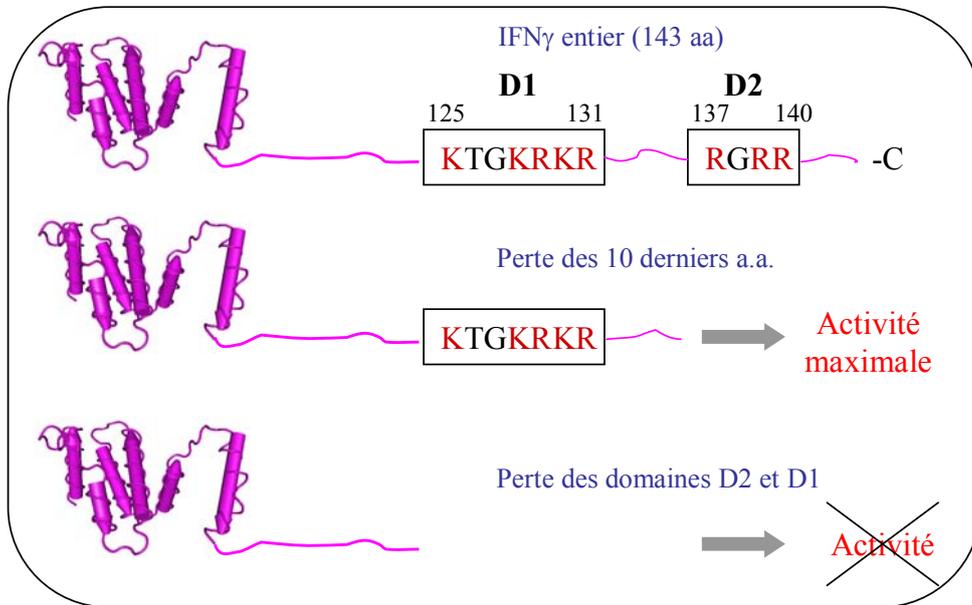
#### REFERENCES

- Lindahl, U., Kusche-Gallberg, M., and Kjellen, L. (1998) *J. Biol. Chem.* 273, 24979–24982
- Bernfield, M., Gotte, M., Park, P. W., Reizes, O., Fitzgerald, M. L., Lincicum, J., and Zako, M. (1999) *Annu. Rev. Biochem.* 68, 729–777
- Tumova, S., Woods, A., and Couchman, J. R. (2000) *Int. J. Biochem. Cell Biol.* 32, 269–288
- Park, P. W., Reizes, O., and Bernfield, M. (2000) *J. Biol. Chem.* 275, 29923–29926
- Gallagher, J. T. (2001) *J. Clin. Invest.* 108, 357–361
- Lortal-Jacob, H., Grosdidier, A., and Imberty, A. (2002) *Proc. Natl. Acad. Sci. U. S. A.* 99, 1229–1234
- Raman, R., Sasisekharan, V., and Sasisekharan, R. (2005) *Chem. Biol.* 12, 267–277
- Murphy, K. J., Merry, C. L., Lyon, M., Thompson, J. E., Roberts, I. S., and Gallagher, J. T. (2004) *J. Biol. Chem.* 279, 27239–27245
- Faham, S., Linhardt, R. J., and Rees, D. C. (1998) *Curr. Opin. Struct. Biol.* 8, 578–586
- Sadir, R., Baloux, F., Grosdidier, A., Imberty, A., and Lortal-Jacob, H. (2001) *J. Biol. Chem.* 276, 8288–8296
- Shukla, D., Liu, J., Blaklock, P., Shwartz, N. W., Bai, X., Esko, J. D., Cohen, G. H., Eisenberg, R. J., Rosenberg, R. D., and Spear, P. G. (1999) *Cell* 99, 13–22
- Lindahl, U., Backstrom, G., Thunberg, L., and Leder, I. G. (1980) *Proc. Natl. Acad. Sci. U. S. A.* 77, 6551–6555
- Patitou, M., Casu, B., and Lindahl, U. (2003) *Biochimie (Paris)* 85, 83–89
- Patitou, M., and van Boeckel, C. A. (2004) *Angew. Chem. Int. Ed. Engl.* 43, 3118–3133
- Boehm, U., Klamp, T., Groot, M., and Howard, J. C. (1997) *Annu. Rev. Immunol.* 15, 749–795
- Schroder, K., Hertzog, P. J., Ravasi, T., and Hume, D. A. (2004) *J. Leukocyte Biol.* 75, 163–189
- Lortal-Jacob, H., Kleinman, H. K., and Grimaud, J. A. (1991) *J. Clin. Invest.* 87, 878–883
- Lortal-Jacob, H., Baltzer, F., and Grimaud, J.-A. (1996) *J. Biol. Chem.* 271, 16139–16143
- Sadir, R., Forest, E., and Lortal-Jacob, H. (1998) *J. Biol. Chem.* 273, 10919–10925
- Lortal-Jacob, H., Turnbull, J. E., and Grimaud, J. A. (1995) *Biochem. J.* 310, 497–505

## Interferon- $\gamma$ Binding Site Structure In Heparan Sulfate

21. Lubineau, A., Lortat-Jacob, H., Gavard, O., Sarrazin, S., and Bonnaffé, D. (2004) *Chemistry* 10, 4265–4282
22. Haelewyn, J., and De Ley, M. (1995) *Biochem. Mol. Biol. Int.* 37, 1163–1171
23. Gavard, O., Hensant, Y., Alais, J., Duverger, V., Dilhas, A., Bascou, A., and Bonnaffé, D. (2003) *Eur. J. Org. Chem.* 18, 3603–3620
24. Dilhas, A., and Bonnaffé, D. (2004) *Tetrahedron Lett.* 45, 3643–3645
25. Oltvoort, J. J., van Boeckel, C. A. A., Koning, J. H., and van Boom, J. H. (1981) *Synthesis* 1981, 305–308
26. Lubineau, A., Escher, S., Alais, J., and Bonnaffé, D. (1997) *Tetrahedron Lett.* 38, 4087–4090
27. Frost, B. M., and Curran, D. P. (1981) *Tetrahedron Lett.* 22, 1287–1290
28. Bitter, T., and Muir, H. (1962) *Anal. Biochem.* 4, 330–334
29. Vives, R. R., Goodger, S., and Pye, D. A. (2001) *Biochem. J.* 354, 141–147
30. Berg, K., Hansen, M. B., and Nilsen, S. E. (1990) *APMIS* 98, 156–162
31. Gibson, U. E., and Kramer, S. M. (1989) *J. Immunol. Methods* 125, 105–113
32. Walter, M. R., Windsor, W. T., Nagabhushan, T. L., Lundell, D. J., Lunn, C. A., Zauodny, P. J., and Narula, S. K. (1995) *Nature* 376, 230–235
33. Lundell, D. L., and Narula, S. K. (1994) *Pharmacol. Ther.* 64, 1–21
34. Esko, J. D., and Salbeck, S. B. (2002) *Annu. Rev. Biochem.* 71, 435–471
35. Dyson, H. J., and Wright, P. E. (2002) *Curr. Opin. Struct. Biol.* 12, 54–60
36. Billiau, A. (1996) *Adv. Immunol.* 62, 61–130
37. Skurkovich, B., and Skurkovich, S. (2003) *Curr. Opin. Mol. Ther.* 5, 52–57
38. Bouma, G., and Strober, W. (2003) *Nat. Rev. Immunol.* 3, 521–533
39. Yamada, M., Kubo, H., Kobayashi, S., Ishizawa, K., and Sasaki, H. (2004) *Am. J. Physiol.* 287, L1042–L1047
40. Hidalgo, L. G., and Halloran, P. F. (2002) *Crit. Rev. Immunol.* 22, 317–349
41. Ozmen, L., Roman, D., Fountoulakis, M., Schmid, G., Ryffel, B., and Garotta, G. (1995) *Eur. J. Immunol.* 25, 6–12
42. Nicoletti, F., Zaccara, P., Di Marco, R., Di Mauro, M., Magro, G., Grasso, S., Mughini, L., Maroni, P., and Garotta, G. (1996) *Endocrinology* 137, 5567–5575
43. Froyen, G., and Billiau, A. (1997) *Biotherapy* 10, 49–57
44. Lim, W. C., and Hanauer, S. B. (2004) *Rev. Gastroenterol. Disord.* 4, 66–85
45. Fernandez-Botran, R., Gorantla, V., Sun, X., Ren, X., Perez-Abadia, G., Crespo, F. A., Oliver, R., Cehun, H. I., Quan, E. E., Maldonado, C., Ray, M., and Barker, J. H. (2002) *Transplantation* 74, 629–629
46. Lever, R., and Page, C. P. (2002) *Nat. Rev. Drug Discov.* 1, 140–148
47. Coombe, D. R., and Kett, W. C. (2005) *Crit. Mol. Life Sci.* 62, 410–424

## **2.ETUDE DE LA REGION C-TERMINALE D'IFN $\gamma$**



**Figure 40: Schéma des domaines D2 et D1 de la région C-terminale d'IFN $\gamma$**

La région C-terminale (aa 123-143) d'IFN $\gamma$  semble être extérieure à la partie globulaire formée par les 7 hélices  $\alpha$ . Elle est constituée de deux domaines (D1 et D2) riches en acides aminés basiques. Ces domaines sont importants pour l'interaction avec les HS (Lortat-Jacob et Grimaud, 1992). Le domaine D1 semble également être impliqué dans l'interaction de la cytokine avec son récepteur spécifique, l'IFN $\gamma$ R (Sadir et al., 1998). Pour l'activité de la cytokine, la perte des 10 derniers acides aminés lui confère une activité maximale, alors que la perte des deux domaines ne lui permet plus de transduire son signal (Döbeli et al., 1988).

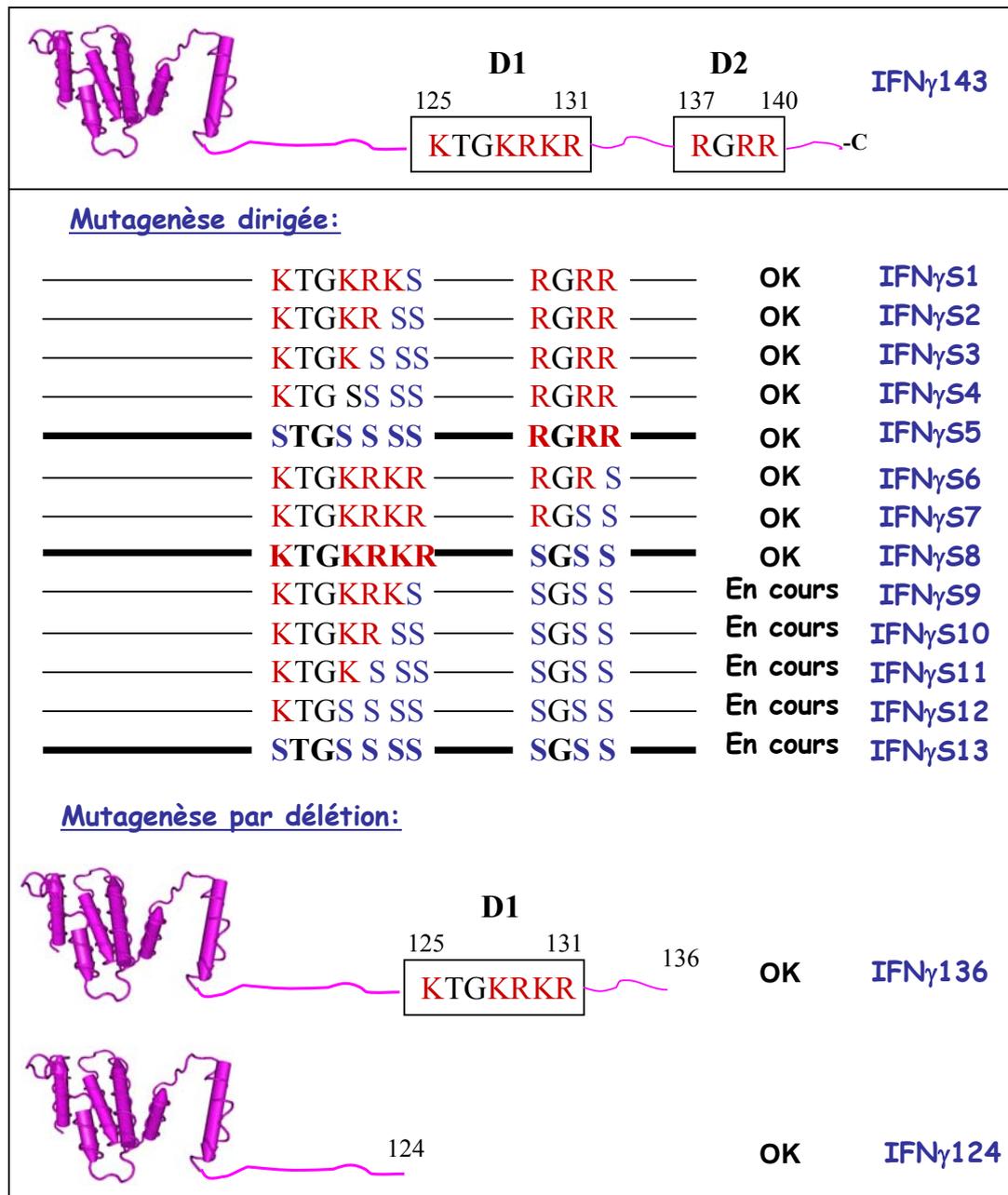
## 2.1. Introduction

L'IFN $\gamma$  possède deux sites d'interactions membranaires de haute affinité, à savoir les protéoglycannes à héparanes sulfates ( $K_d = 1,5$  nM) et le récepteur à IFN $\gamma$ , l'IFN $\gamma$ R ( $K_d = 0,1 - 1$  nM) via lequel il transmet son signal (Lortat-Jacob et al., 1991 ; Anderson et al., 1982).

Comme cela a été exposé dans le rapport bibliographique, le complexe IFN $\gamma$ /IFN $\gamma$ R met en jeu deux récepteurs pour un IFN $\gamma$  (homodimère), en conformité avec la symétrie de la cytokine (Fountoulakis et al., 1992 ; Greenlund et al., 1993). Les régions de l'IFN $\gamma$  impliquées dans l'interaction avec le récepteur sont : la boucle qui fait le lien entre l'hélice A et B (aa 18 à 26), les résidus 108-124 de l'hélice F, et une séquence de la région C-terminale connue pour être flexible (Lundell et Narula, 1994 ; Walter et al., 1995). La structure de l'IFN $\gamma$  a été résolue que jusqu'à l'acide aminé 124, la région C-terminale (acides aminés 125 à 143) apparaissent trop labile pour donner un signal en cristallographie (Ealick et al., 1991 ; Thiel et al., 2000). Cette région C-terminale doit être située à l'extérieur de la partie globulaire de la protéine. Elle est caractérisée par deux domaines (D1 et D2) riches en acides aminés basiques (D1 : résidus 125 à 131, KTGKRKR ; D2 : résidus 137 à 140, RGRR) (Figure 40). Ces domaines sont sensibles à la protéolyse (Trinchieri et al., 1985 ; Döbeli et al., 1988 ; Curling et al., 1990 ; Jarpe et Johnson, 1990) et sont importants pour l'activité de la cytokine. En effet, la perte des 10 derniers acides aminés, et donc du domaine D2, entraîne une hausse de l'activité de la protéine. En revanche, la perte des deux domaines (D1 et D2) rend la protéine inactive (Döbeli et al., 1988). L'interaction entre cette région C-terminale et l'IFN $\gamma$ R n'a pas encore été caractérisée structuralement. En revanche, il a été montré que cette région C-terminale avait également une forte affinité pour les héparanes sulfates. Cette interaction a largement été caractérisée (Lortat-Jacob et Grimaud, 1992 ; Lortat-Jacob et al., 1995). Ainsi, comme pour un grand nombre de protéines liant les HS, il apparaît que ce sont les domaines riches en acides aminés basiques (D1 et D2 ) qui permettent une telle interaction.

Afin de savoir si, comme dans le cas du FGF, les HS jouaient un rôle de co-récepteur pour l'IFN $\gamma$ , Sadir et coll. (1998) ont entrepris différents travaux. Tout d'abord, il a été démontré que l'IFN $\gamma$  ne peut pas se fixer à la fois sur les HS et sur l'IFN $\gamma$ R. En utilisant des anticorps monoclonaux dirigés contre D2 ou D1, il a aussi été mis en évidence que le domaine D1 était impliqué dans l'interaction avec l'IFN $\gamma$ R au même titre qu'avec les HS.

L'IFN $\gamma$  sécrété par les cellules NK ou par les LT va transmettre son signal par le biais de l'IFN $\gamma$ R. Or, à la surface des cellules cibles, il s'avère que les HSPGs occupent un volume



**Figure 41:** Représentation des mutants dans la région C-terminale d'IFN $\gamma$  qui seront utilisés pour étudier l'importance des domaines D1 et D2 dans les interactions avec les deux ligands d'IFN $\gamma$

En haut est représenté le monomère d'IFN $\gamma$  avec sa région C-terminale déployée et ses domaines basiques D1 et D2 qui correspond à l'IFN $\gamma$  sauvage (IFN $\gamma$ 143). Ensuite sont représentés les mutants issus de la mutagenèse dirigée (IFN $\gamma$ S1 à IFN $\gamma$ S13). "OK" signifie qu'ils s'expriment en bactérie. "En cours" signifie que la mutagenèse est encore à réaliser. Et enfin, en-dessous sont présentés les mutants obtenus par délétion, IFN $\gamma$ 124 et IFN $\gamma$ 136, tous les deux s'exprimant également dans le système bactérien BL21(DE3).

hydrodynamique et sont d'une taille plus importante que l'IFN $\gamma$ R. De plus, si une cellule exprime à sa surface environ 500 à 20 000 IFN $\gamma$ R, le nombre de molécules d'HS à cette même surface peut avoisiner les  $10^5$  à  $10^6$ . L'affinité de la cytokine pour ses deux ligands étant du même ordre, avant de rejoindre l'IFN $\gamma$ R, les HS se positionnent comme un passage obligé. Ainsi, les questions posées sont les suivantes : quels sont les mécanismes qui amènent le médiateur à passer du glycocalyx à l'IFN $\gamma$ R? Etant donné l'importance de chacun des domaines D1 et D2, tant au point de vue de l'interaction avec chacun des deux ligands qu'au niveau de l'activité de la cytokine (contrôlée par la protéolyse de l'extrémité C-terminale), quel rôle peut être précisément attribué à chacun des deux domaines ainsi qu'aux acides aminés qui les composent?

Un premier travail a été entrepris pour tenter de répondre à ces interrogations. Le but est d'étudier les interactions de mutants (des domaines D2 et D1 comme cela est présenté en Figure 41) de l'IFN $\gamma$  avec ses deux ligands. Plus précisément, ce sont les valeurs cinétiques des interactions qui vont être mesurées afin d'établir quels sont les acides aminés ou la perte d'acides aminés qui vont favoriser l'interaction avec l'un ou l'autre des deux ligands.

Un second travail a également été initié pour essayer de résoudre la structure de la région C-terminale d'IFN $\gamma$ .

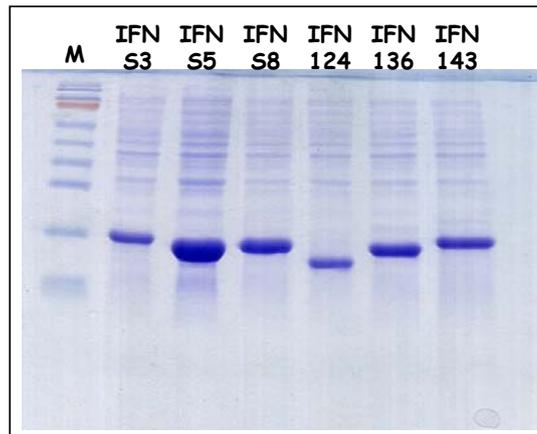
## **2.2. Etude des deux domaines basiques C-terminaux d'IFN $\gamma$**

### **2.2.1. Matériel et méthodes**

#### *Construction des vecteurs d'expression pour les mutants d'IFN $\gamma$*

#### Construction des vecteurs exprimant les IFN $\gamma$ mutants dépourvu de leurs domaine D2 (pET11a/IFN $\gamma$ 136) ou D2 et D1 (pET11a/IFN $\gamma$ 124)

Par PCR, les sites de restriction NdeI et BamHI ainsi qu'un codon stop en position 137 et 125 sont ajoutés aux extrémités 3' et 5' de l'ADNc d'IFN $\gamma$  humain (432 pb). (PCR mix (*Taq* DNA Polymérase, dNTPs, MgCl<sub>2</sub> (Proméga)) + amorces : 4  $\mu$ M + ADN : 10 ng, volume réactionnel : 20  $\mu$ L ; Tm des amorces : 60°C, programme : 1 min 95°C, 1 min 54°C, 1 min 72°C, 30 cycles, 5 min 72°C). Le fragment PCR est inséré dans un vecteur intermédiaire TOPO (Invitrogen). A partir du plasmide TOPO/IFN $\gamma$ 136 ou 124 purifié, le cDNA codant pour l'IFN $\gamma$  mutant peut être excisé par les enzymes NdeI et BamHI (500 ng de plasmide, 5 U d'enzyme de restriction (Roche), volume réactionnel 10  $\mu$ L). Le cDNA purifié par extraction sur gel d'agarose (geneclean kit, BIO101), est alors réintroduit dans le vecteur d'expression



***Figure 42: Gel SDS-PAGE de l'expression de plusieurs mutants d'IFN $\gamma$  en système bactérien***

A côté du marqueur de poids moléculaire (M), les mutants exprimés sous forme de corps d'inclusion ont été chargés avant purification, sur un gel SDS-PAGE 15 %. Dans l'ordre, pour les mutants de substitution il y a l'IFNS3, l'IFNS5, l'IFNS8, pour les mutants de délétion il y a l'IFN124, l'IFN136 et enfin le sauvage (l'IFN143) a également été chargé. La présence des protéines est révélée par une coloration au bleu de Coomassie.

pET11a (Novagen) en aval du promoteur inductible à l'IPTG. (Rapid DNA Ligation Kit (Roche), 50 ng de vecteur, 90 ng d'insert, volume : 10  $\mu$ L). Ces vecteurs pET11a/IFN $\gamma$ 136 et 124 seront alors utilisés pour transformer les bactéries BL21 utilisées pour la production.

### Construction des vecteurs d'expression d'IFN $\gamma$ dont les acides aminés basiques des domaines D1 et D2 sont sélectivement mutés en sérine

Les expériences de mutagenèse dirigée sont réalisées sur le vecteur d'expression pET11a/IFN $\gamma$ 143. Certaines paires de bases des domaines D1 ou D2 sont modifiées afin de changer les acides aminés basiques (K, R) en sérine (S). Pour cela le Kit "Quick Change MultiSite directed Mutagenesis Kit" (Stratagène) est utilisé. Grâce à deux amorces complémentaires comportant les mutations, le vecteur pET11a/IFN $\gamma$ 143 est amplifié dans sa totalité (réaction dans 100  $\mu$ L avec le vecteur matrice : 50 ng, les amorces: 125 ng, le tampon de réaction. Les dNTP 1mM et 1 $\mu$ L de la DNA polymérase pfu turbo sont ajoutés après un cycle de 8 minutes à 95°C; programme : 95°C : 30 sec, 55°C : 1 min, 68°C : 12 min, 18 cycles avec une dernière élongation rallongée de 15 min à 68°C). Le produit de réaction est soumis à la DpnI (enzyme de restriction qui coupe uniquement l'ADN méthylé). Ainsi, seul le vecteur muté est conservé car non digéré. Il est alors utilisé pour transformer des bactéries TOP10, lesquelles vont le produire en double brin.

### *Expression des mutants dans le système bactérien BL21 Star (DE3)*

Le vecteur pET11a, a l'avantage de contenir un promoteur T7, inductible par l'IPTG qui permet la fixation de la T7 polymérase, elle-même produite en présence d'IPTG par les bactéries DE3 telles que les BL21 (DE3). La souche BL21 (DE3) a l'avantage d'être déficiente en mRNAse ce qui permet une plus grande durée de vie des ARNm et ainsi une augmentation de la quantité de protéine recombinante obtenue.

Une culture en fin de croissance est diluée 10 fois dans du milieu LB (Ampicilline) afin d'obtenir (2 heures plus tard) une culture en phase exponentielle de croissance ( $DO_{600nm}=0,6$ ). L'expression des protéines d'intérêt est alors induite par 0,5 mM d'IPTG pendant 5 heures, à 37°C.

## **2.2.2. Premiers résultats et perspectives**

Comme il est indiqué Figure 41, les vecteurs d'expression des mutants de la région C-terminale d'IFN $\gamma$  et plus particulièrement de ses domaines D1 et D2 ont pour la plupart été



***IFN gamma globular domains: 1-128***

QDPYVKEAEN LKKYFNAGHS DVADNGTLFL GILKNWKEES DRKIMQSQIV  
 SFYKLFKNF KDDQSIQKSV ETIKEDMNVK FFNSNKKKR DFEKLTNYSV  
 TDLNVQRKAI HELIQVMAEL SPAAKTGKrk rsqmlfrgrr asq

***Figure 43: Prédiction du désordre dans la structure d'IFN $\gamma$***

Le graphique représente le désordre probable de la protéine (en ordonnée) en fonction des acides aminés (abscisse) (obtenu par le logiciel IUpred). La bande en bleu foncé indique quels sont les résidus d'IFN $\gamma$  qui seraient préférentiellement situés dans un domaine globulaire. Ceci est reporté en dessous du graphique, avec les a.a. appartenant à une région stable en majuscules bleu et ceux qui appartiendraient plutôt à une région désordonnée en minuscules (rouge). Ainsi, la région globulaire serait située entre les acides aminés 1 et 128, ensuite, en C-terminal la protéine serait bien plus désordonnée.

construits (IFN $\gamma$ S1 à S8, IFN $\gamma$ 124 et IFN $\gamma$ 136). Leur expression en système bactérien dans des BL21(DE3) a été vérifiée (Figure 42). Des premiers résultats permettent d'affirmer que la méthode de purification qui va être utilisée, pourra être la même que celle utilisée pour purifier l'IFN $\gamma$  sauvage (Sarrazin et al., 2005). Une fois les IFN $\gamma$  mutants purifiés, leur activité respective sera testée et les interactions de chacun avec l'IFN $\gamma$ R et les HS seront minutieusement mesurées par Biacore. De cette manière, les données cinétiques des interactions pourront être analysées et comparées.

## **2.3. Etude structurale de la région C-terminale d'IFN $\gamma$**

### **2.3.1. Résultats et discussion**

Actuellement, aucune information structurale sur les 19 derniers acides aminés de la cytokine n'est connue (acides aminés 125 à 143 qui correspondent à la région C-terminale englobant les domaines D1 et D2). Cela est certainement dû à la flexibilité de cette région (Ealick et al., 1991 ; Thiel et al., 2000). Le logiciel iupred (<http://iupred.enzim.hu/>) permet d'évaluer le désordre d'une protéine en fonction de sa séquence peptidique. Nous l'avons donc utilisé sur la séquence protéique de l'IFN $\gamma$ . Comme cela a été suggéré, le logiciel iupred fait apparaître l'IFN $\gamma$  comme une protéine globulaire avec une région C-terminale (à partir de l'acide aminé 128) désordonnée (Figure 43).

En cristallographie, l'idée d'utiliser un ligand pour stabiliser une protéine, ou simplement une région flexible d'une protéine dans le but d'offrir plus de stabilité et ainsi d'obtenir des informations structurales est maintenant admise. Pour stabiliser la région C-terminale d'IFN $\gamma$ , nous nous sommes proposés d'utiliser des molécules de types GAGS. L'objectif du travail, en utilisant de l'héparine, des fragments d'héparine ou des mimes oligosaccharidiques synthétiques (donc homogènes) était de stabiliser la région C-terminale d'IFN $\gamma$ , et espérer ainsi établir par cristallographie sa structure.

Dans un premier temps, l'objectif a été de définir les conditions pour lesquelles des cristaux d'IFN $\gamma$  pouvaient être obtenus. Dans les conditions testées (kit Crystal Screen I et II), il apparaît qu'une solution de 2 M de sulfate d'ammonium, pH 8 permet de former des cristaux biréfringents d'IFN $\gamma$  de taille exploitable. Malheureusement, pour la mise en place de la cristallisation d'un complexe IFN $\gamma$ /oligosaccharide, cette solution n'est pas adéquate. En effet, à l'aide du Biacore, nous avons pu vérifier que le sulfate d'ammonium dissociait

instantanément l'interaction IFN $\gamma$ /HS. Une autre condition de cristallisation devra donc être utilisée.

Afin de mettre en place les conditions (concentration à utiliser, purification par filtration sur gel, concentration des échantillons, etc) permettant d'obtenir un complexe IFN $\gamma$ /oligosaccharides pur, des premiers essais de formation de complexes ont été réalisés entre l'IFN $\gamma$  et de l'Hp (6 kDa). Il s'est avéré que la mise en présence d'IFN $\gamma$  concentré (5 mg/mL) et d'Hp dans un rapport molaire 1/1,5 induisait la formation d'agrégats. Le même essai en conditions diluées a également conduit à l'agrégation lors de la concentration. Ce phénomène d'agrégation pourrait être expliqué par le fait qu'un fragment d'Hp peut fixer au moins deux IFN $\gamma$  et ainsi de suite, créant ainsi un réseau de protéine. Afin de défavoriser cette façon d'interagir, plusieurs additifs comme le glycérol, différentes concentrations de sulfate d'ammonium ou de chlorure de sodium ont été testées. Aucune condition n'a été concluante.

Dans l'espoir qu'une molécule oligosaccharidique bien plus spécifique d'IFN $\gamma$  puisse ne pas engendrer le phénomène d'agrégation, la molécule 2O<sub>10</sub> (Sarrazin et al., 2005), a été mise en présence d'IFN $\gamma$ . Cependant, cette molécule peut réellement être considérée comme un mimétique de l'Hp puisque son effet est identique. Pour une concentration en IFN $\gamma$  de 10 mg/mL, l'ajout de seulement 1/8<sup>ème</sup> de mole de la molécule 2O<sub>10</sub> crée l'agrégation.

Toujours dans l'optique de ne plus former ce "réseau protéique", des octasaccharides, hexasaccharides et tétrasaccharides obtenus par synthèse chimique (Lubineau et al., 2004) ont été utilisés pour complexer l'IFN $\gamma$  dans sa région C-terminale. L'ajout d'octasaccharide induit la formation d'agrégats. Cette turbidité est moins présente lors de l'ajout d'un hexasaccharide, et ne l'est plus en présence de tétrasaccharide.

Dans la continuité de ces travaux, un complexe entre l'IFN $\gamma$  et des disaccharides trisulfatés issus de la dépolymérisation de l'Hp réalisée de façon enzymatique a été créé. Ce complexe, formé par l'ajout de 6 disaccharides par dimère d'IFN $\gamma$  (10 mg/mL) ne présente aucune turbidité qui pourrait être reliée à une quelconque agrégation. Sur ce complexe, 96 conditions de cristallisation ont été réalisées. Six d'entre elles ont permis la formation de cristaux biréfringents. L'un d'entre eux n'a pas pu être analysé du fait de sa resolubilisation lors de l'extraction. Les cinq autres, après extraction et immobilisation dans l'azote liquide, ont été exposés à un rayonnement synchrotron, il s'est avéré qu'un seul possédait une carte de diffraction d'un cristal de protéine. Les cristaux obtenus étaient malheureusement trop petits pour en obtenir plus d'informations. Les deux conditions menant au cristal resolubilisé et au cristal positif ont été réitérées de façon manuelle (goutte de 4  $\mu$ L dont 2  $\mu$ L de complexe et 2  $\mu$ L de solution de cristallisation) sans succès pour l'instant.

PDB	Protein	Origin	Nature	Length of oligosaccharide <sup>a,b</sup>	Ref.
2BRS	Eosinophil-granule major basic protein (EMBP)	Human	Lectin	disaccharide	Swaminathan et al. 2005
1G5N	Annexin V	Rat	Extracell. prot.	$\Delta$ tetrasaccharide	Capila et al. 2001
2HYU	Annexin A2	Human	Extracell. prot.	$\Delta$ tetrasaccharide	Shao et al 2006
2HYV	Annexin A2	Human	Extracell. prot.	$\Delta$ hexasaccharide	Shao et al 2006
1XT3	Cardiotoxin A3	Cobra	Toxin	hexasaccharide	Lee et al 2005
1ZA4	Thrombospondin 1 (N-term domain)	Human	Extracell. prot.	pentasaccharide (Arixtra)	Tan et al., 2006
1QQP	Foot-and-mouth disease virus capsid	Viral	Capsid	pentasaccharide	Fry et al 1999
1RID	Vaccinia complement protein	Viral	Complement	hexasaccharide	Ganesh et al 2004
1U4L	RANTES	Human	Chemokine	$\Delta$ disaccharide	Shaw et al 2004
1U4M	RANTES	Human	Chemokine	$\Delta$ disaccharide	Shaw et al 2004
1BFB	Basic fibroblast growth factor (bFGF)	Human	Growth factor	$\Delta$ tetrasaccharide	Faham et al 1996
1BFC	Basic fibroblast growth factor (bFGF)	Human	Growth factor	$\Delta$ hexasaccharide	Faham et al 1996
1AXM	Acidic fibroblast growth factor (aFGF)	Human	Growth factor	$\Delta$ hexasaccharide	DiGabriele et al 1998
2AXM	Acidic fibroblast growth factor (aFGF)	Human	Growth factor	$\Delta$ hexasaccharide	DiGabriele et al 1998
1E00	FGF1/ectodomain of receptor FGFR2	Human	Growth factor/receptor	decamer	Pellegrini et al 2000
1FQ9	FGF2/ectodomain of receptor FGFR1	Human	Growth factor/receptor	decamer	Schlessinger et al 2000
1GMN	N-term domain of hepatocyte growth factor (NK1)	Human	Growth factor	tetrasaccharide	Lietha et al 2001
1GMO	N-term domain of hepatocyte growth factor (NK1)	Human	Growth factor	tetrasaccharide	Lietha et al 2001
2FUT	Heparinase II	<i>Pedobacter heparinus</i>	Enzyme	$\Delta$ disaccharide	Shaya et al 2006
1T8U	3-O-Sulfotransferase 3	Human	Enzyme	$\Delta$ tetrasaccharide	Moon et al 2004
1AZX	Antithrombin	Human	Serpin	synthetic penta	Jin et al 1997
1E03	Antithrombin	Human	Serpin	synthetic penta	Jin et al 1997
1NQ9	Antithrombin	Human	Serpin	pentasaccharide	Johnson et Huntington 2003
1XMN	Thrombin	Human	Protease	hexamer	Carter et al 2003
1TB6	Antithrombin/thrombin	Human	Serpin/protease	heparin mimetic	Li et al 2004
2B5T	Antithrombin/thrombin	Human	Serpin/protease	heparin mimetic (SR123781)	Johnson et al 2006a
1SR5	Antithrombin/anhydrothrombin	Human	Serpin/protease	heparin mimetic	Dementiev et al 2004
2GD4	Antithrombin/factor Xa	Human	Serpin/protease	pentasaccharide (Fondaparinux)	Johnson et al 2006b

<sup>a</sup> The reported length corresponds to the fragment observed in the crystal structure.

<sup>b</sup>  $\Delta$  Indicates the presence of  $\Delta$ UA at the nonreducing end (1,4-dideoxy-5-dehydro glucuronic acid)

**Figure 44: Tableau récapitulatif de toutes les structures cristallines des complexes obtenues entre une protéine et des GAGs (HS, fragment d'Hp ou mimétiques) (d'après Imberty et al., 2007)**

L'intérêt d'obtenir la structure du complexe entre IFN $\gamma$  et des oligosaccharides est multiple. Premièrement, cela permettrait d'avoir une structure complète de la cytokine. Ensuite, la caractérisation de l'interaction entre l'IFN $\gamma$  et les GAGs pourrait ainsi être faite à l'échelle moléculaire, ce qui offrirait des informations utiles pour l'amélioration de mimétiques thérapeutiques telle que la molécule 2O<sub>10</sub> (Lubineau et al., 2004 ; Sarrazin et al., 2005 ; Delehedde et al., 2005).

La résolution de la structure d'un tel complexe n'est pas simple. A ce jour, l'ensemble des structures cristallographiques obtenues pour un complexe entre une molécule de type GAG et une protéine, n'inclut aucune protéine qui possède plus d'un site d'interaction avec les GAGs (Figure 44) (Imberty et al., 2007). Seuls RANTES et SDF-1 $\alpha$  ont été cristallisés en présence de GAGs. Il s'agissait dans les deux cas d'un disaccharide, ne permettant pas l'oligomérisation de la chimiokine (Shaw et al., 2004 ; Murphy et al., 2007). Cela laisse penser que les investigations structurales sur les interactions entre chimiokines et GAGs ont malheureusement abouti aux mêmes résultats que ceux obtenus avec IFN $\gamma$  comme cela avait été décrit par Stura et al. (2002).

En ce sens, il convient bien entendu, d'approfondir les conditions de cristallisation entre l'IFN $\gamma$  et des disaccharides. Par ailleurs, une nouvelle piste de travail a été engagée. Elle consiste à produire un mutant d'IFN $\gamma$  qui ne possède qu'un seul de ses deux domaines C-terminaux. Si l'hypothèse énoncée s'avèrerait juste, la co-cristallisation de fragments oligosaccharidique de taille plus importante (qu'un disaccharide) avec cet IFN $\gamma$  "unijambiste" pourrait être ainsi rendue possible et exploitée.

### **3.CONCLUSION ET PERSPECTIVES**

L'IFN $\gamma$  possède deux ligands biologiques qui sont les protéoglycannes à HS et l'IFN $\gamma$ R, son récepteur spécifique. Par le biais de ce dernier, l'IFN $\gamma$  est capable de réguler au moins 200 gènes, dans les processus de réponses immunitaires et inflammatoires principalement. Sadir et al. (1998) ont montré que l'IFN $\gamma$  ne pouvait pas se fixer à la fois sur ses deux ligands. Cela prouve que les HS ne fonctionnent pas comme des "co-récepteurs" pour la cytokine, comme ils peuvent l'être avec le FGF-2.

Sur l'IFN $\gamma$ , en C-terminal, deux domaines (D1 et D2) riches en acides aminés basiques peuvent être distingués. La perte du domaine D2 confère à la cytokine une activité maximale alors que la perte des domaines D1 et D2 lui enlève toute activité (Döbeli et al., 1988). Ceci étant, il a également été montré que les HS sont capables de protéger le domaine D1 des coupures protéolytiques (Lortat-Jacob et Grimaud, 1994). De tels résultats prouvent l'importance des HS dans la protection de la cytokine et dans la potentialisation de son activité. De plus, des études menées chez l'animal ont montré que la fixation de l'IFN $\gamma$  aux HS régulaient l'élimination plasmatique, la distribution tissulaire et l'accumulation locale de la protéine (Lortat-Jacob et al., 1996b).

La caractérisation de l'interaction entre la cytokine IFN $\gamma$  et les HS se trouve donc être un élément important pour comprendre davantage le mode de fonctionnement de ce médiateur. Dans cette optique, plusieurs travaux permettant de caractériser de façon biochimique cette interaction ont été réalisés (Lortat-Jacob et al., 1991 ; Lortat-Jacob et Grimaud, 1991 ; 1992). De ces travaux, un modèle d'interaction entre les deux molécules a été proposé (Lortat-Jacob et al. ; 1995) (Figure 35 du rapport bibliographique). Afin de vérifier ce modèle et de disposer de quantité importante de molécules pour procéder à différents tests biologiques, une collaboration avec le laboratoire du Professeur David Bonnaffé (Université Paris Sud, Laboratoire de Chimie Organique Multifonctionnelle) a été mise en place et a permis d'obtenir par voie synthétique des oligosaccharides ayant pour but de mimer les HS dans l'interaction avec l'IFN $\gamma$  (Lubineau et al., 2004). Ces mimétiques étaient des tétra, des hexa ou des octasaccharides, reliés ou non entre eux par une répétition de fonctions PEG de différentes tailles (5 ; 10 ou 32 répétitions). La capacité de ces molécules à mimer les HS dans l'interaction avec IFN $\gamma$  a alors pu être testée. C'est l'objet de l'article "Sarrazin et al., (2005)" dans lequel l'effet de ces mimes oligosaccharidiques sur l'interaction IFN $\gamma$ /IFN $\gamma$ R et IFN $\gamma$ /HS a pu être évalué par Biacore. De même, l'effet de ces mimétiques sur l'activité de la cytokine a pu être mesuré. Il en ressort que la molécule 2O<sub>10</sub> semble être un bon mimétique des HS puisqu'elle est capable d'inhiber l'interaction de l'IFN $\gamma$  avec ses deux ligands. Cela se traduit *in vitro* par un effet inhibiteur sur l'activité de la cytokine aussi important que celui qui peut

être observé avec l'Hp. La molécule 2O<sub>10</sub> semble donc valider le modèle d'interaction qui avait été proposé. Le fait que cette molécule soit capable d'inhiber l'activité d'IFN $\gamma$  *in vitro* permet d'imaginer que des applications thérapeutiques peuvent lui être assignées, en particulier dans les pathologies liées à une surexpression de l'IFN $\gamma$ .

La molécule 2O<sub>10</sub> est constituée de deux octasaccharides complètement sulfatés reliés entre eux par dix répétitions PEG. Or, il n'est pas évident que chacun des sulfates soit essentiel à l'interaction. Il est même possible que certains de ces sulfates soient gênants pour l'interaction. Si la molécule venait à être utilisée dans une approche thérapeutique, l'intérêt serait qu'elle ne possède que les sulfates (ou les groupements) qui lui sont réellement nécessaires à inhiber l'IFN $\gamma$ , ceci dans un souci de spécificité et d'affinité maximale.

C'est pourquoi il est apparu important de caractériser encore plus en détail l'interaction entre l'IFN $\gamma$  et 2O<sub>10</sub> ou entre l'IFN $\gamma$  et les HS en général ; cela, afin de savoir de manière précise quels sont les acides aminés de la région C-terminale de l'IFN $\gamma$  qui rentrent en interaction avec quels groupements fonctionnels des oligosaccharides. Pour cela, deux études ont été mises en place.

La première étude, toujours en cours, a pour but d'analyser le rôle des domaines D1 et D2 de l'IFN $\gamma$  et surtout de chacun des acides aminés basiques inclus dans ces domaines. La création de mutants de chacun des acides aminés basiques des domaines D1 et D2 a ainsi été engagée afin d'analyser le rôle de chacun d'entre eux dans la fixation de l'IFN $\gamma$  à l'un ou l'autre de ses ligands. Cette approche doit également permettre d'optimiser 2O<sub>10</sub>, par exemple en supprimant les sulfates positionnés en face d'acides aminés n'ayant qu'un rôle mineur dans l'interaction.

La seconde étude a pour objectif de résoudre la structure cristallographique de la région C-terminale d'IFN $\gamma$  qui n'a pour l'instant pas encore été mise à jour. La région C-terminale d'IFN $\gamma$  semble trop labile pour donner un signal cristallographique (Ealick et al., 1991), mais pourrait être stabilisée en présence de son ligand. L'objectif a donc été d'obtenir un complexe entre IFN $\gamma$  et des oligosaccharides (Hp ou 2O<sub>10</sub>) dans le but d'engager des expériences de cristallographie. Malheureusement, la mise en présence d'Hp ou de 2O<sub>10</sub> avec l'IFN $\gamma$  induit une agrégation quel que soient les conditions que nous avons testées. En revanche, l'utilisation de tétrasaccharides ou de disaccharides n'a pas induit d'agrégats. Le complexe entre IFN $\gamma$  et des disaccharides a été utilisé pour des expériences de cristallographie. Deux conditions de cristallisation ont permis l'obtention de cristaux protéiques, mais de taille insuffisante pour en dégager une information structurale. De ce point de vue, il apparaît important de continuer à cribler des conditions autour de ces résultats encourageants. Par

ailleurs, pour éviter le phénomène d'agrégation, la préparation d'un IFN $\gamma$  ne possédant qu'un seul domaine C-terminal pourrait s'avérer être une approche originale permettant de limiter la formation de réseaux. L'obtention de la structure du domaine C-terminal d'IFN $\gamma$  constituerait une avancée très importante pour l'optimisation d'une molécule comme 2O<sub>10</sub>, en particulier en étudiant l'interaction IFN $\gamma$ /2O<sub>10</sub> par modélisation moléculaire.

# **BIBLIOGRAPHIE**

- Abid, M. R., X. Yi, et al. (2006). "Vascular endocan is preferentially expressed in tumor endothelium." Microvasc Res **72**(3): 136-45.
- Abildgaard, U. (1968). "Highly purified antithrombin 3 with heparin cofactor activity prepared by disc electrophoresis." Scand J Clin Lab Invest **21**(1): 89-91.
- Abrams, D. I., S. Kuno, et al. (1989). "Oral dextran sulfate (UA001) in the treatment of the acquired immunodeficiency syndrome (AIDS) and AIDS-related complex." Ann Intern Med **110**(3): 183-8.
- Adam E., V. Motte, S. Sarrazin and M. Delehedde. (2006). "Stromal Derived Growth Factor-1 : a target for GAG-based therapies?" In "New developments in Therapeutic Glycomics". Eds. Maryse Delehedde and Hugues Lortat-Jacob. Chapter 12: 283-302.
- Adamia, S., C. A. Maxwell, et al. (2005). "Hyaluronan and hyaluronan synthases: potential therapeutic targets in cancer." Curr Drug Targets Cardiovasc Haematol Disord **5**(1): 3-14.
- Adams, Y., C. Freeman, et al. (2006). "Inhibition of Plasmodium falciparum growth *in vitro* and adhesion to chondroitin-4-sulfate by the heparan sulfate mimetic PI-88 and other sulfated oligosaccharides." Antimicrob Agents Chemother **50**(8): 2850-2.
- Ahmed, T., J. Garrigo, et al. (1993). "Preventing bronchoconstriction in exercise-induced asthma with inhaled heparin." N Engl J Med **329**(2): 90-5.
- Ai, X., A. T. Do, et al. (2006). "Substrate specificity and domain functions of extracellular heparan sulfate 6-O-endosulfatases, QSulf1 and QSulf2." J Biol Chem **281**(8): 4969-76.
- Aitkenhead, M., S. J. Wang, et al. (2002). "Identification of endothelial cell genes expressed in an *in vitro* model of angiogenesis: induction of ESM-1, (beta)ig-h3, and NrCAM." Microvasc Res **63**(2): 159-71.
- Albo, D., C. Long, et al. (1996). "Modulation of cranial bone healing with a heparin-like dextran derivative." J Craniofac Surg **7**(1): 19-22.
- Ali, S., A. C. Palmer, et al. (2001). "Multimerization of monocyte chemoattractant protein-1 is not required for glycosaminoglycan-dependent transendothelial chemotaxis." Biochem J **358**(Pt 3): 737-45.
- Allain, F., C. Vanpouille, et al. (2002). "Interaction with glycosaminoglycans is required for cyclophilin B to trigger integrin-mediated adhesion of peripheral blood T lymphocytes to extracellular matrix." Proc Natl Acad Sci U S A **99**(5): 2714-9.
- Allain, F., C. Vanpouille, et al. (2003). "Involvement of GAGs in the activity of pro-inflammatory factors." Adv Exp Med Biol **535**: 95-106.
- Alzona, M., H. M. Jack, et al. (1994). "CD30 defines a subset of activated human T cells that produce IFN-gamma and IL-5 and exhibit enhanced B cell helper activity." J Immunol **153**(7): 2861-7.
- Amatschek, S., U. Koenig, et al. (2004). "Tissue-wide expression profiling using cDNA subtraction and microarrays to identify tumor-specific genes." Cancer Res **64**(3): 844-56.
- Ameje, L. and M. F. Young (2002). "Mice deficient in small leucine-rich proteoglycans: novel *in vivo* models for osteoporosis, osteoarthritis, Ehlers-Danlos syndrome, muscular dystrophy, and corneal diseases." Glycobiology **12**(9): 107R-16R.
- Amiral, J., J. C. Lormeau, et al. (1997). "Absence of cross-reactivity of SR90107A/ORG31540 pentasaccharide with antibodies to heparin-PF-4 complexes developed in heparin-induced thrombocytopenia." Blood Coagul Fibrinolysis **8**(2): 114-7.
- Amiral, J. and D. Meyer (1998). "Heparin-induced thrombocytopenia: diagnostic tests and biological mechanisms." Baillieres Clin Haematol **11**(2): 447-60.

- Anderson, M. E. and T. J. Siahaan (2003). "Targeting ICAM-1/LFA-1 interaction for controlling autoimmune diseases: designing peptide and small molecule inhibitors." Peptides **24**(3): 487-501.
- Anderson, P., Y. K. Yip, et al. (1982). "Specific binding of 125I-human interferon-gamma to high affinity receptors on human fibroblasts." J Biol Chem **257**(19): 11301-4.
- Andersson, L. O., T. W. Barrowcliffe, et al. (1976). "Anticoagulant properties of heparin fractionated by affinity chromatography on matrix-bound antithrombin iii and by gel filtration." Thromb Res **9**(6): 575-83.
- Appella, E., I. T. Weber, et al. (1988). "Structure and function of epidermal growth factor-like regions in proteins." FEBS Lett **231**(1): 1-4.
- Ashworth, J. L., S. Biswas, et al. (2006). "Mucopolysaccharidoses and the eye." Surv Ophthalmol **51**(1): 1-17.
- Bajorath, J. (2000). "Molecular organization, structural features, and ligand binding characteristics of CD44, a highly variable cell surface glycoprotein with multiple functions." Proteins **39**(2): 103-11.
- Bal dit Sollier, C., C. Kang, et al. (2004). "Activity of a synthetic hexadecasaccharide (SanOrg123781A) in a pig model of arterial thrombosis." J Thromb Haemost **2**(6): 925-30.
- Barbouche, R., H. Lortat-Jacob, et al. (2005). "Glycosaminoglycans and protein disulfide isomerase-mediated reduction of HIV Env." Mol Pharmacol **67**(4): 1111-8.
- Barbour, L. A., S. D. Kick, et al. (1994). "A prospective study of heparin-induced osteoporosis in pregnancy using bone densitometry." Am J Obstet Gynecol **170**(3): 862-9.
- Barragan, A., D. Spillmann, et al. (1999). "Role of glycans in Plasmodium falciparum infection." Biochem Soc Trans **27**(4): 487-93.
- Barritault, D. and J. P. Caruelle (2006). "[Regenerating agents (RGTAs): a new therapeutic approach]." Ann Pharm Fr **64**(2): 135-44.
- Bauer, K. A., D. W. Hawkins, et al. (2002). "Fondaparinux, a synthetic pentasaccharide: the first in a new class of antithrombotic agents - the selective factor Xa inhibitors." Cardiovasc Drug Rev **20**(1): 37-52.
- Beauvais, D. M. and A. C. Rapraeger (2004). "Syndecans in tumor cell adhesion and signaling." Reprod Biol Endocrinol **2**: 3.
- Béchar, D., T. Gentina, et al. (2001a). "Endocan is a novel chondroitin sulfate/dermatan sulfate proteoglycan that promotes hepatocyte growth factor/scatter factor mitogenic activity." J Biol Chem **276**(51): 48341-9.
- Béchar, D., V. Meignin, et al. (2000). "Characterization of the secreted form of endothelial-cell-specific molecule 1 by specific monoclonal antibodies." J Vasc Res **37**(5): 417-25.
- Béchar, D., A. Scherpereel, et al. (2001b). "Human endothelial-cell specific molecule-1 binds directly to the integrin CD11a/CD18 (LFA-1) and blocks binding to intercellular adhesion molecule-1." J Immunol **167**(6): 3099-106.
- Bellamy, N., J. Campbell, et al. (2006). "Viscosupplementation for the treatment of osteoarthritis of the knee." Cochrane Database Syst Rev(2): CD005321.
- Bennett, K. L., D. G. Jackson, et al. (1995). "CD44 isoforms containing exon V3 are responsible for the presentation of heparin-binding growth factor." J Cell Biol **128**(4): 687-98.
- Bensadoun, E. S., A. K. Burke, et al. (1997). "Proteoglycans in granulomatous lung diseases." Eur Respir J **10**(12): 2731-7.
- Bernfield, M., M. Gotte, et al. (1999). "Functions of cell surface heparan sulfate proteoglycans." Annu Rev Biochem **68**: 729-77.

- Bezakova, G. and M. A. Ruegg (2003). "New insights into the roles of agrin." Nat Rev Mol Cell Biol **4**(4): 295-308.
- Bhowmick, N. A., E. G. Neilson, et al. (2004). "Stromal fibroblasts in cancer initiation and progression." Nature **432**(7015): 332-7.
- Billiau, A. and K. Vandembroeck. (2001). "IFN $\gamma$ " Cytokine reference
- Bitter, T. and H. M. Muir (1962). "A modified uronic acid carbazole reaction." Anal Biochem **4**: 330-4.
- Bix, G. and R. V. Iozzo (2005). "Matrix revolutions: "tails" of basement-membrane components with angiostatic functions." Trends Cell Biol **15**(1): 52-60.
- Blobe, G. C., W. P. Schiemann, et al. (2001). "Functional roles for the cytoplasmic domain of the type III transforming growth factor beta receptor in regulating transforming growth factor beta signaling." J Biol Chem **276**(27): 24627-37.
- Bobardt, M. D., A. C. Saphire, et al. (2003). "Syndecan captures, protects, and transmits HIV to T lymphocytes." Immunity **18**(1): 27-39.
- Boehm, U., T. Klamp, et al. (1997). "Cellular responses to interferon-gamma." Annu Rev Immunol **15**: 749-95.
- Borcuk, A. C., L. Shah, et al. (2004). "Molecular signatures in biopsy specimens of lung cancer." Am J Respir Crit Care Med **170**(2): 167-74 Epub 2004 Apr 15.
- Border, W. A., N. A. Noble, et al. (1992). "Natural inhibitor of transforming growth factor-beta protects against scarring in experimental kidney disease." Nature **360**(6402): 361-4.
- Borza, D. B. and W. T. Morgan (1998). "Histidine-proline-rich glycoprotein as a plasma pH sensor. Modulation of its interaction with glycosaminoglycans by pH and metals." J Biol Chem **273**(10): 5493-9.
- Bourgeois, P., G. Chales, et al. (1998). "Efficacy and tolerability of chondroitin sulfate 1200 mg/day vs chondroitin sulfate 3 x 400 mg/day vs placebo." Osteoarthritis Cartilage **6 Suppl A**: 25-30.
- Brandacher, G., C. Winkler, et al. (2006). "Antitumoral activity of interferon-gamma involved in impaired immune function in cancer patients." Curr Drug Metab **7**(6): 599-612.
- Brat, D. J., A. A. Castellano-Sanchez, et al. (2004). "Pseudopalisades in glioblastoma are hypoxic, express extracellular matrix proteases, and are formed by an actively migrating cell population." Cancer Res. **64**(3): 920-7.
- Brat, D. J. and E. G. Van Meir (2004). "Vaso-occlusive and prothrombotic mechanisms associated with tumor hypoxia, necrosis, and accelerated growth in glioblastoma." Lab Invest. **84**(4): 397-405.
- Brickman, Y. G., M. D. Ford, et al. (1998). "Structural modification of fibroblast growth factor-binding heparan sulfate at a determinative stage of neural development." J Biol Chem **273**(8): 4350-9.
- Brinkhous, K. M., H.P. Smith, E.D. Warner, W.H. Seegers. (1939). "The inhibition of blood clotting : an unidentified substance which acts in conjunction with heparin to prevent the conversion of prothrombin into thrombin." Am. J. Physiol. **125**: 683-687.
- Brinkman, B. M. (2004). "Splice variants as cancer biomarkers." Clin Biochem **37**(7): 584-94.
- Briscoe, J., N. C. Rogers, et al. (1996). "Kinase-negative mutants of JAK1 can sustain interferon-gamma-inducible gene expression but not an antiviral state." Embo J **15**(4): 799-809.
- Brooks, B., D. M. Briggs, et al. (2000). "Presentation of IFN-gamma to nitric oxide-producing cells: a novel function for mast cells." J Immunol **164**(2): 573-9.
- Brown, L. F., A. J. Guidi, et al. (1999). "Vascular stroma formation in carcinoma in situ, invasive carcinoma, and metastatic carcinoma of the breast." Clin Cancer Res **5**(5): 1041-56.

- Brown, T. A., T. Bouchard, et al. (1991). "Human keratinocytes express a new CD44 core protein (CD44E) as a heparan-sulfate intrinsic membrane proteoglycan with additional exons." *J Cell Biol* **113**(1): 207-21.
- Brunda, M. J. (1994). "Interleukin-12." *J Leukoc Biol* **55**(2): 280-8.
- Burch, G. H., Y. Gong, et al. (1997). "Tenascin-X deficiency is associated with Ehlers-Danlos syndrome." *Nat Genet* **17**(1): 104-8.
- Bureau, J. F., F. Bihl, et al. (1995). "The gene coding for interferon-gamma is linked to the D12S335 and D12S313 microsatellites and to the MDM2 gene." *Genomics* **28**(1): 109-12.
- Burgess, R. W., W. C. Skarnes, et al. (2000). "Agrin isoforms with distinct amino termini: differential expression, localization, and function." *J Cell Biol* **151**(1): 41-52.
- Bush, A. I., W. H. Pettingell, Jr., et al. (1994). "The amyloid beta-protein precursor and its mammalian homologues. Evidence for a zinc-modulated heparin-binding superfamily." *J Biol Chem* **269**(43): 26618-21.
- Butler, J., N. W. Rydell, et al. (1970). "Hyaluronic acid in synovial fluid. VI. Effect of intra-articular injection of hyaluronic acid on the clinical symptoms of arthritis in track horses." *Acta Vet Scand* **11**(2): 139-55.
- Capila, I., M. J. Hernaiz, et al. (2001). "Annexin V--heparin oligosaccharide complex suggests heparan sulfate--mediated assembly on cell surfaces." *Structure* **9**(1): 57-64.
- Carey, D. J. (1997). "Syndecans: multifunctional cell-surface co-receptors." *Biochem J* **327** (Pt 1): 1-16.
- Carpentier, M., F. Allain, et al. (1999). "Two distinct regions of cyclophilin B are involved in the recognition of a functional receptor and of glycosaminoglycans on T lymphocytes." *J Biol Chem* **274**(16): 10990-8.
- Carter, N. M., S. Ali, et al. (2003). "Endothelial inflammation: the role of differential expression of N-deacetylase/N-sulphotransferase enzymes in alteration of the immunological properties of heparan sulphate." *J Cell Sci* **116**(Pt 17): 3591-600.
- Carter, W. J., E. Cama, et al. (2005). "Crystal structure of thrombin bound to heparin." *J Biol Chem* **280**(4): 2745-9.
- Cassatella, M. A., F. Bazzoni, et al. (1990). "Molecular basis of interferon-gamma and lipopolysaccharide enhancement of phagocyte respiratory burst capability. Studies on the gene expression of several NADPH oxidase components." *J Biol Chem* **265**(33): 20241-6.
- Casu, B., P. Oreste, et al. (1981). "The structure of heparin oligosaccharide fragments with high anti-(factor Xa) activity containing the minimal antithrombin III-binding sequence. Chemical and <sup>13</sup>C nuclear-magnetic-resonance studies." *Biochem J* **197**(3): 599-609.
- Casu, B., M. Petitou, et al. (1988). "Conformational flexibility: a new concept for explaining binding and biological properties of iduronic acid-containing glycosaminoglycans." *Trends Biochem Sci* **13**(6): 221-5.
- Caterson, B., C. R. Flannery, et al. (2000). "Mechanisms involved in cartilage proteoglycan catabolism." *Matrix Biol* **19**(4): 333-44.
- Catlow, K., J. A. Deakin, et al. (2003). "Hepatocyte growth factor/scatter factor and its interaction with heparan sulphate and dermatan sulphate." *Biochem Soc Trans* **31**(2): 352-3.
- Cella, G., G. Boeri, et al. (1992). "Interaction between histidine-rich glycoprotein and platelet factor 4 with dermatan sulfate and low-molecular-weight dermatan sulfate." *Angiology* **43**(1): 59-62.
- Chakravarti, S. (2002). "Functions of lumican and fibromodulin: lessons from knockout mice." *Glycoconj J* **19**(4-5): 287-93.

- Chakravarti, S., T. Magnuson, et al. (1998). "Lumican regulates collagen fibril assembly: skin fragility and corneal opacity in the absence of lumican." *J Cell Biol* **141**(5): 1277-86.
- Charalambous, C., T. C. Chen, et al. (2006). "Characteristics of tumor-associated endothelial cells derived from glioblastoma multiforme." *Neurosurg Focus* **20**(4): E22.
- Chen, C. Y. and A. B. Shyu (1995). "AU-rich elements: characterization and importance in mRNA degradation." *Trends Biochem Sci* **20**(11): 465-70.
- Chen, R. L. and A. D. Lander (2001). "Mechanisms underlying preferential assembly of heparan sulfate on glypican-1." *J Biol Chem* **276**(10): 7507-17.
- Cheng, F., D. Heinegard, et al. (1994). "Patterns of uronosyl epimerization and 4-/6-O-sulphation in chondroitin/dermatan sulphate from decorin and biglycan of various bovine tissues." *Glycobiology* **4**(5): 685-96.
- Chitu, V. and E. R. Stanley (2006). "Colony-stimulating factor-1 in immunity and inflammation." *Curr Opin Immunol* **18**(1): 39-48.
- Choay, J., J. C. Lormeau, et al. (1981). "[Low molecular weight oligosaccharides active in plasma against factor Xa (author's transl)]." *Ann Pharm Fr* **39**(1): 37-44.
- Choay, J., J. C. Lormeau, et al. (1980). "Anti-Xa active heparin oligosaccharides." *Thromb Res* **18**(3-4): 573-8.
- Choay, J., J. C. Lormeau, et al. (1981). "Structural studies on a biologically active hexasaccharide obtained from heparin." *Ann N Y Acad Sci* **370**: 644-9.
- Choay, J., M. Petitou, et al. (1983). "Structure-activity relationship in heparin: a synthetic pentasaccharide with high affinity for antithrombin III and eliciting high anti-factor Xa activity." *Biochem Biophys Res Commun* **116**(2): 492-9.
- Clarke, D., O. Katoh, et al. (1995). "Interaction of interleukin 7 (IL-7) with glycosaminoglycans and its biological relevance." *Cytokine* **7**(4): 325-30.
- Clasper, S., S. Vekemans, et al. (1999). "Inducible expression of the cell surface heparan sulfate proteoglycan syndecan-2 (fibroglycan) on human activated macrophages can regulate fibroblast growth factor action." *J Biol Chem* **274**(34): 24113-23.
- Cohen, I., O. Pappo, et al. (2006). "Heparanase promotes growth, angiogenesis and survival of primary breast tumors." *Int J Cancer* **118**(7): 1609-17.
- Cong, R., X. Jiang, et al. (2006). "Hhex is a direct repressor of endothelial cell-specific molecule 1 (ESM-1)." *Biochem Biophys Res Commun* **346**(2): 535-45 Epub 2006 Jun 5.
- Cong, R., X. Jiang, et al. (2006). "Hhex is a direct repressor of endothelial cell-specific molecule 1 (ESM-1)." *Biochem Biophys Res Commun* **346**(2): 535-45.
- Contejean, C. (1895). "Recherches sur les injections intraveineuse du peptone et leur influence sur la coagulabilité du sang chez le chien." *Arch. Physiol. Norm. Pathol.* **7** : 45-53.
- Coombe, D. R. and W. C. Kett (2005). "Heparan sulfate-protein interactions: therapeutic potential through structure-function insights." *Cell Mol Life Sci* **62**(4): 410-24.
- Cornuet, P. K., T. C. Blochberger, et al. (1994). "Molecular polymorphism of lumican during corneal development." *Invest Ophthalmol Vis Sci* **35**(3): 870-7.
- Corsi, A., M. Riminucci, et al. (2001). "Achondrogenesis type IB: agenesis of cartilage interterritorial matrix as the link between gene defect and pathological skeletal phenotype." *Arch Pathol Lab Med* **125**(10): 1375-8.
- Corte, M. D., L. O. Gonzalez, et al. (2006). "Expression and clinical significance of cytosolic hyaluronan levels in invasive breast cancer." *Breast Cancer Res Treat* **97**(3): 329-37.
- Curling, E. M., P. M. Hayter, et al. (1990). "Recombinant human interferon-gamma. Differences in glycosylation and proteolytic processing lead to heterogeneity in batch culture." *Biochem J* **272**(2): 333-7.

- Dahlman, T. C. (1993). "Osteoporotic fractures and the recurrence of thromboembolism during pregnancy and the puerperium in 184 women undergoing thromboprophylaxis with heparin." Am J Obstet Gynecol **168**(4): 1265-70.
- Dahlman, T. C., H. E. Sjoberg, et al. (1994). "Bone mineral density during long-term prophylaxis with heparin in pregnancy." Am J Obstet Gynecol **170**(5 Pt 1): 1315-20.
- Damus, P. S., M. Hicks, et al. (1973). "Anticoagulant action of heparin." Nature **246**(5432): 355-7.
- Danielson, K. G., H. Baribault, et al. (1997). "Targeted disruption of decorin leads to abnormal collagen fibril morphology and skin fragility." J Cell Biol **136**(3): 729-43.
- Danielsson, A., E. Raub, et al. (1986). "Role of ternary complexes, in which heparin binds both antithrombin and proteinase, in the acceleration of the reactions between antithrombin and thrombin or factor Xa." J Biol Chem **261**(33): 15467-73.
- Davis, C. G. (1990). "The many faces of epidermal growth factor repeats." New Biol **2**(5): 410-9.
- De Cat, B. and G. David (2001). "Developmental roles of the glypicans." Semin Cell Dev Biol **12**(2): 117-25.
- De Clercq, E. (1989). "New acquisitions in the development of anti-HIV agents." Antiviral Res **12**(1): 1-19.
- de Kort, M., R. C. Buijsman, et al. (2005). "Synthetic heparin derivatives as new anticoagulant drugs." Drug Discov Today **10**(11): 769-79.
- De Lange Davies C, M. R., Munn LL, Mouta-Carreira C, Jain RK, Boucher Y. (2001). "Decorin inhibits endothelial migration and tube-like structure formation : role of thrombospondin-1." Microvasc Res **62**(1): 26-42.
- De Maeyer, E. and J. De Maeyer-Guignard (1992). "The growth rate of two transplantable murine tumors, 3LL lung carcinoma and B16F10 melanoma, is influenced by Hyal-1, a locus determining hyaluronidase levels and polymorphism." Int J Cancer **51**(4): 657-60.
- De Visser, K. E., A. Eichten, et al. (2006). "Paradoxical roles of the immune system during cancer development." Nat Rev Cancer **6**(1): 24-37.
- Deakin, J. A. and M. Lyon (1999). "Differential regulation of hepatocyte growth factor/scatter factor by cell surface proteoglycans and free glycosaminoglycan chains." J Cell Sci **112**(Pt 12): 1999-2009.
- DeAngelis, P. L. (1999). "Hyaluronan synthases: fascinating glycosyltransferases from vertebrates, bacterial pathogens, and algal viruses." Cell Mol Life Sci **56**(7-8): 670-82.
- DeBaun, M. R., J. Ess, et al. (2001). "Simpson Golabi Behmel syndrome: progress toward understanding the molecular basis for overgrowth, malformation, and cancer predisposition." Mol Genet Metab **72**(4): 279-86.
- Delehedde, M. (2002). "Heparan sulfate as a key player in the regulation of growth factor activities." Recent Res Devel Biol Chem **1**: 133-149.
- Delehedde, M., Allain F., Payne SJ., Borgo R., Vanpouille C., Fernig DG., Deudon E. (2002). "Proteoglycans in inflammation." Curr. Med. Chem. **1**: 89-102.
- Delehedde, M., Lortat-Jacob H., Gallagher JT., Bonaffe D., Adam E., Querolle O., Lequien S., Degove S., Lassalle P., Bécharde D. (2005). "Proteoglycan involvement in inflammatory diseases. new developments in GAG-based therapies." Current Med. Chem.
- Delehedde, M., M. Lyon, et al. (2002). "Fibroblast growth factor-2 binds to small heparin-derived oligosaccharides and stimulates a sustained phosphorylation of p42/44 mitogen-activated protein kinase and proliferation of rat mammary fibroblasts." Biochem J **366**(Pt 1): 235-44.

- Delehedde, M., M. Lyon, et al. (2001). "Proteoglycans: pericellular and cell surface multireceptors that integrate external stimuli in the mammary gland." J Mammary Gland Biol Neoplasia **6**(3): 253-73.
- Delehedde, M., M. Lyon, et al. (2002). "Hepatocyte growth factor/scatter factor binds to small heparin-derived oligosaccharides and stimulates the proliferation of human HaCaT keratinocytes." J Biol Chem **277**(14): 12456-62.
- Delehedde, M., N. Sergeant, et al. (2001). "Hepatocyte growth factor/scatter factor stimulates migration of rat mammary fibroblasts through both mitogen-activated protein kinase and phosphatidylinositol 3-kinase/Akt pathways." Eur J Biochem **268**(16): 4423-9.
- Delehedde, M., M. Seve, et al. (2000). "Fibroblast growth factor-2 stimulation of p42/44MAPK phosphorylation and IkappaB degradation is regulated by heparan sulfate/heparin in rat mammary fibroblasts." J Biol Chem **275**(43): 33905-10.
- Delehedde M., S. Sarrazin, E. Adam, V. Motte, C. Vanpouille. (2006). "Proteoglycans and glycosaminoglycans: Complex Molecules with modulating activity" In "New developments in Therapeutic Glycomics". Eds. Maryse Delehedde and Hugues Lortat-Jacob. Chapter 1.
- Dementiev, A., M. Petitou, et al. (2004). "The ternary complex of antithrombin-anhydrothrombin-heparin reveals the basis of inhibitor specificity." Nat Struct Mol Biol **11**(9): 863-7.
- Dempsey, L. A., T. B. Plummer, et al. (2000). "Heparanase expression in invasive trophoblasts and acute vascular damage." Glycobiology **10**(5): 467-75.
- Denholm, E. M., E. Cauchon, et al. (2000). "Inhibition of human dermal fibroblast proliferation by removal of dermatan sulfate." Eur J Pharmacol **400**(2-3): 145-53.
- Denholm, E. M., Y. Q. Lin, et al. (2001). "Anti-tumor activities of chondroitinase AC and chondroitinase B: inhibition of angiogenesis, proliferation and invasion." Eur J Pharmacol **416**(3): 213-21.
- Depontieu F. (2002). "Study on the variation of the expression of endocan during the course of experimental endotoxin shock in balb/c mouse." DEA in Health and Life Sciences
- Desgranges, P., C. Barbaud, et al. (1999). "A substituted dextran enhances muscle fiber survival and regeneration in ischemic and denervated rat EDL muscle." Faseb J **13**(6): 761-6.
- Desnoyers, L. (2004). "Structural basis and therapeutic implication of the interaction of CCN proteins with glycoconjugates." Curr Pharm Des **10**(31): 3913-28.
- Devos, R., H. Cheroutre, et al. (1982). "Molecular cloning of human immune interferon cDNA and its expression in eukaryotic cells." Nucleic Acids Res **10**(8): 2487-501.
- Diamant, Z., M. C. Timmers, et al. (1996). "Effect of inhaled heparin on allergen-induced early and late asthmatic responses in patients with atopic asthma." Am J Respir Crit Care Med **153**(6 Pt 1): 1790-5.
- Diamond, M. S., R. Alon, et al. (1995). "Heparin is an adhesive ligand for the leukocyte integrin Mac-1 (CD11b/CD1)." J Cell Biol **130**(6): 1473-82.
- DiGabriele, A. D., I. Lax, et al. (1998). "Structure of a heparin-linked biologically active dimer of fibroblast growth factor." Nature **393**(6687): 812-7.
- Dijkmans, R. and A. Billiau (1991). "Interferon-gamma/lipopolysaccharide-treated mouse embryonic fibroblasts are killed by a glycolysis/L-arginine-dependent process accompanied by depression of mitochondrial respiration." Eur J Biochem **202**(1): 151-9.
- Dinarelo, C. A., D. Novick, et al. (1998). "Overview of interleukin-18: more than an interferon-gamma inducing factor." J Leukoc Biol **63**(6): 658-64.

- Distler, J. and A. Anguelouch (2006). "Evidence-based practice: review of clinical evidence on the efficacy of glucosamine and chondroitin in the treatment of osteoarthritis." *J Am Acad Nurse Pract* **18**(10): 487-93.
- Döbeli H., G. R. W. Jucker et al., (1988). "Role of the C-terminal sequence on the biological activity of human interferon- $\gamma$ ." *J. Biotechnol.* 7:199-216.
- Dolowitz, D. A. and T. F. Dougherty (1960). "The use of heparin as an anti-inflammatory agent." *Laryngoscope* **70**: 873-84.
- Dolowitz, D. A. and T. F. Dougherty (1965). "The Use of Heparin in the Control of Allergies." *Ann Allergy* **23**: 309-13.
- Dong, S., G. J. Cole, et al. (2003). "Expression of collagen XVIII and localization of its glycosaminoglycan attachment sites." *J Biol Chem* **278**(3): 1700-7.
- Droguett, R., C. Cabello-Verrugio, et al. (2006). "Extracellular proteoglycans modify TGF-beta bio-availability attenuating its signaling during skeletal muscle differentiation." *Matrix Biol* **25**(6): 332-41.
- Dudhia, J. (2005). "Aggrecan, aging and assembly in articular cartilage." *Cell Mol Life Sci* **62**(19-20): 2241-56.
- Ealick, S. E., W. J. Cook, et al. (1991). "Three-dimensional structure of recombinant human interferon-gamma." *Science* **252**(5006): 698-702.
- Eckert, R. and H. Ragg (2003). "Zinc ions promote the interaction between heparin and heparin cofactor II." *FEBS Lett* **541**(1-3): 121-5.
- Edgren, G., B. Havsmark, et al. (1997). "Glypican (heparan sulfate proteoglycan) is palmitoylated, deglycanated and reglycanated during recycling in skin fibroblasts." *Glycobiology* **7**(1): 103-12.
- Ehara, T., F. A. Carone, et al. (1994). "Basement membrane chondroitin sulfate proteoglycan alterations in a rat model of polycystic kidney disease." *Am J Pathol* **144**(3): 612-21.
- El Habbal, M. H., L. Smith, et al. (1995). "Effect of heparin anticoagulation on neutrophil adhesion molecules and release of IL8: C3 is not essential." *Cardiovasc Res* **30**(5): 676-81.
- Elefteriou, F., J. Y. Exposito, et al. (2001). "Binding of tenascin-X to decorin." *FEBS Lett* **495**(1-2): 44-7.
- Elenius, V., M. Gotte, et al. (2004). "Inhibition by the soluble syndecan-1 ectodomains delays wound repair in mice overexpressing syndecan-1." *J Biol Chem* **279**(40): 41928-35.
- Esko, J. D. and U. Lindahl (2001). "Molecular diversity of heparan sulfate." *J Clin Invest* **108**(2): 169-73.
- Evans, R. C., V. S. Wong, et al. (1997). "Treatment of corticosteroid-resistant ulcerative colitis with heparin--a report of 16 cases." *Aliment Pharmacol Ther* **11**(6): 1037-40.
- Ezura, Y., S. Chakravarti, et al. (2000). "Differential expression of lumican and fibromodulin regulate collagen fibrillogenesis in developing mouse tendons." *J Cell Biol* **151**(4): 779-88.
- Faham, S., R. E. Hileman, et al. (1996). "Heparin structure and interactions with basic fibroblast growth factor." *Science* **271**(5252): 1116-20.
- Falcoff, R. (1972). "Some properties of virus and immune-induced human lymphocyte interferons." *J Gen Virol* **16**(2): 251-3.
- Fennie, E. H., Y. S. Lie, et al. (1988). "Reduced mortality in murine cytomegalovirus infected mice following prophylactic murine interferon-gamma treatment." *Antiviral Res* **10**(1-3): 27-39.
- Fernandez, J. A., J. Petaja, et al. (1999). "Dermatan sulfate and LMW heparin enhance the anticoagulant action of activated protein C." *Thromb Haemost* **82**(5): 1462-8.

- Ferrara, N., L. Damico, et al. (2006). "Development of ranibizumab, an anti-vascular endothelial growth factor antigen binding fragment, as therapy for neovascular age-related macular degeneration." Retina **26**(8): 859-70.
- Ferro, V., K. Fewings, et al. (2001). "Large-scale preparation of the oligosaccharide phosphate fraction of *Pichia holstii* NRRL Y-2448 phosphomannan for use in the manufacture of PI-88." Carbohydr Res **332**(2): 183-9.
- Ferro, V., E. Hammond, et al. (2004). "The development of inhibitors of heparanase, a key enzyme involved in tumour metastasis, angiogenesis and inflammation." Mini Rev Med Chem **4**(6): 693-702.
- Filep, J. G. (2006). "Endocan or endothelial cell-specific molecule-1: a novel prognostic marker of sepsis?" Crit Care Med **34**(2): 574-5.
- Filmus, J. (2001). "Glypicans in growth control and cancer." Glycobiology **11**(3): 19R-23R.
- Fischer, I., J. P. Gagner, et al. (2005). "Angiogenesis in gliomas: biology and molecular pathophysiology." Brain Pathol. **15**(4): 297-310.
- Flannery, C. R. (2006). "MMPs and ADAMTSs: functional studies." Front Biosci **11**: 544-69.
- Fong, L. G., T. A. Fong, et al. (1990). "Inhibition of mouse macrophage degradation of acetyl-low density lipoprotein by interferon-gamma." J Biol Chem **265**(20): 11751-60.
- Forsberg, E., G. Pejler, et al. (1999). "Abnormal mast cells in mice deficient in a heparin-synthesizing enzyme." Nature **400**(6746): 773-6.
- Forster-Horvath, C., L. Meszaros, et al. (2004). "Expression of CD44v3 protein in human endothelial cells *in vitro* and in tumoral microvessels *in vivo*." Microvasc Res **68**(2): 110-8.
- Fountoulakis, M., M. Zulauf, et al. (1992). "Stoichiometry of interaction between interferon gamma and its receptor." Eur J Biochem **208**(3): 781-7.
- Francis, D. J., C. R. Parish, et al. (2003). "Blockade of vascular smooth muscle cell proliferation and intimal thickening after balloon injury by the sulfated oligosaccharide PI-88: phosphomannopentaose sulfate directly binds FGF-2, blocks cellular signaling, and inhibits proliferation." Circ Res **92**(8): e70-7.
- Fraser, J. R., T. C. Laurent, et al. (1997). "Hyaluronan: its nature, distribution, functions and turnover." J Intern Med **242**(1): 27-33.
- Fredj-Reygrobellet, D., D. L. Hristova, et al. (1994). "CMDBS, functional analogue of heparin sulfate as a new class of corneal ulcer healing agents." Ophthalmic Res **26**(6): 325-31.
- Frevel, M. A., T. Bakheet, et al. (2003). "p38 Mitogen-activated protein kinase-dependent and -independent signaling of mRNA stability of AU-rich element-containing transcripts." Mol Cell Biol **23**(2): 425-36.
- Fries, E. and A. M. Blom (2000). "Bikunin--not just a plasma proteinase inhibitor." Int J Biochem Cell Biol **32**(2): 125-37.
- Fry, E. E., S. M. Lea, et al. (1999). "The structure and function of a foot-and-mouth disease virus-oligosaccharide receptor complex." Embo J **18**(3): 543-54.
- Fryer, A., Y. C. Huang, et al. (1997). "Selective O-desulfation produces nonanticoagulant heparin that retains pharmacological activity in the lung." J Pharmacol Exp Ther **282**(1): 208-19.
- Fuki, I. V., R. V. Iozzo, et al. (2000). "Perlécan heparan sulfate proteoglycan: a novel receptor that mediates a distinct pathway for ligand catabolism." J Biol Chem **275**(33): 25742-50.
- Fuki, I. V., M. E. Meyer, et al. (2000). "Transmembrane and cytoplasmic domains of syndecan mediate a multi-step endocytic pathway involving detergent-insoluble membrane rafts." Biochem J **351 Pt 3**: 607-12.

- Fukui, S., T. Feizi, et al. (2002). "Oligosaccharide microarrays for high-throughput detection and specificity assignments of carbohydrate-protein interactions." Nat Biotechnol **20**(10): 1011-7.
- Fultz, M. J., S. A. Barber, et al. (1993). "Induction of IFN-gamma in macrophages by lipopolysaccharide." Int Immunol **5**(11): 1383-92.
- Funderburgh, J. L. (2000). "Keratan sulfate: structure, biosynthesis, and function." Glycobiology **10**(10): 951-8.
- Gaffney, A. and P. Gaffney (1996). "Rheumatoid arthritis and heparin." Br J Rheumatol **35**(8): 808-9.
- Gaffney, P. R., C. T. Doyle, et al. (1995). "Paradoxical response to heparin in 10 patients with ulcerative colitis." Am J Gastroenterol **90**(2): 220-3.
- Gaffney, P. R., J. J. O'Leary, et al. (1991). "Response to heparin in patients with ulcerative colitis." Lancet **337**(8735): 238-9.
- Gallagher, J. T. (2001). "Heparan sulfate: growth control with a restricted sequence menu." J Clin Invest **108**(3): 357-61.
- Ganesh, V. K., S. A. Smith, et al. (2004). "Structure of vaccinia complement protein in complex with heparin and potential implications for complement regulation." Proc Natl Acad Sci U S A **101**(24): 8924-9.
- Gansbacher, B., R. Bannerji, et al. (1990). "Retroviral vector-mediated gamma-interferon gene transfer into tumor cells generates potent and long lasting antitumor immunity." Cancer Res **50**(24): 7820-5.
- Gerritsen, M. E., J. E. Tomlinson, et al. (2003). "Using gene expression profiling to identify the molecular basis of the synergistic actions of hepatocyte growth factor and vascular endothelial growth factor in human endothelial cells." Br J Pharmacol **140**(4): 595-610 Epub 2003 Sep 22.
- Gessani, S. and F. Belardelli (1998). "IFN-gamma expression in macrophages and its possible biological significance." Cytokine Growth Factor Rev **9**(2): 117-23.
- Ghatak, S., S. Misra, et al. (2005). "Hyaluronan constitutively regulates ErbB2 phosphorylation and signaling complex formation in carcinoma cells." J Biol Chem **280**(10): 8875-83.
- Ghiselli, G. and R. V. Iozzo (2000). "Overexpression of bamacan/SMC3 causes transformation." J Biol Chem **275**(27): 20235-8.
- Ghiselli, G., L. D. Siracusa, et al. (1999). "Complete cDNA cloning, genomic organization, chromosomal assignment, functional characterization of the promoter, and expression of the murine Bamacan gene." J Biol Chem **274**(24): 17384-93.
- Ghosh, S., R. Chaudhary, et al. (2006). "Interfering with interferons in inflammatory bowel disease." Gut **55**(8): 1071-3.
- Gilat, D., R. Hershkoviz, et al. (1994). "Regulation of adhesion of CD4+ T lymphocytes to intact or heparinase-treated subendothelial extracellular matrix by diffusible or anchored RANTES and MIP-1 beta." J Immunol **153**(11): 4899-906.
- Godard, S., G. Getz, et al. (2003). "Classification of human astrocytic gliomas on the basis of gene expression: a correlated group of genes with angiogenic activity emerges as a strong predictor of subtypes." Cancer Res. **63**(20): 6613-25.
- Goerges, A. L. and M. A. Nugent (2003). "Regulation of vascular endothelial growth factor binding and activity by extracellular pH." J Biol Chem **278**(21): 19518-25.
- Gold, L. I. (1999). "The role for transforming growth factor-beta (TGF-beta) in human cancer." Crit Rev Oncog **10**(4): 303-60.
- Goldberg, M., D. Septier, et al. (2005). "Targeted disruption of two small leucine-rich proteoglycans, biglycan and decorin, excerpts divergent effects on enamel and dentin formation." Calcif Tissue Int **77**(5): 297-310.

- Gonzalez-Iglesias, R., M. A. Pajares, et al. (2002). "Prion protein interaction with glycosaminoglycan occurs with the formation of oligomeric complexes stabilized by Cu(II) bridges." *J Mol Biol* **319**(2): 527-40.
- Gory, S., J. Dalmon, et al. (1998). "Requirement of a GT box (Sp1 site) and two Ets binding sites for vascular endothelial cadherin gene transcription." *J Biol Chem* **273**(12): 6750-5.
- Gospodarowicz, D. and J. Cheng (1986). "Heparin protects basic and acidic FGF from inactivation." *J Cell Physiol* **128**(3): 475-84.
- Gotte, M., M. Bernfield, et al. (2005). "Increased leukocyte-endothelial interactions in syndecan-1-deficient mice involve heparan sulfate-dependent and -independent steps." *Curr Eye Res* **30**(6): 417-22.
- Gotte, M., A. M. Jousen, et al. (2002). "Role of syndecan-1 in leukocyte-endothelial interactions in the ocular vasculature." *Invest Ophthalmol Vis Sci* **43**(4): 1135-41.
- Grant, D. S., C. Yenisey, et al. (2002). "Decorin suppresses tumor cell-mediated angiogenesis." *Oncogene* **21**(31): 4765-77.
- Gray, P. W., D. W. Leung, et al. (1982). "Expression of human immune interferon cDNA in *E. coli* and monkey cells." *Nature* **295**(5849): 503-8.
- Greenlund, A. C., M. A. Farrar, et al. (1994). "Ligand-induced IFN gamma receptor tyrosine phosphorylation couples the receptor to its signal transduction system (p91)." *Embo J* **13**(7): 1591-600.
- Greenlund, A. C., M. O. Morales, et al. (1995). "Stat recruitment by tyrosine-phosphorylated cytokine receptors: an ordered reversible affinity-driven process." *Immunity* **2**(6): 677-87.
- Greenlund, A. C., R. D. Schreiber, et al. (1993). "Interferon-gamma induces receptor dimerization in solution and on cells." *J Biol Chem* **268**(24): 18103-10.
- Greiling, H. (1994). "Structure and biological functions of keratan sulfate proteoglycans." *Exs* **70**: 101-22.
- Gresser, I. and K. Naficy (1964). "Recovery of an Interferon-Like Substance from Cerebrospinal Fluid." *Proc Soc Exp Biol Med* **117**: 285-9.
- Griffith, G. C., G. Nichols, Jr., et al. (1965). "Heparin Osteoporosis." *Jama* **193**: 91-4.
- Grigoriu, B. D., F. Depontieu, et al. (2006). "Endocan expression and relationship with survival in human non-small cell lung cancer." *Clin Cancer Res* **12**(15): 4575-82.
- Grimaud, J. A. and H. Lortat-Jacob (1994). "Matrix receptors to cytokines: from concept to control of tissue fibrosis dynamics." *Pathol Res Pract* **190**(9-10): 883-90.
- Grootenhuys, P. D., P. Westerduin, et al. (1995). "Rational design of synthetic heparin analogues with tailor-made coagulation factor inhibitory activity." *Nat Struct Biol* **2**(9): 736-9.
- Guerrini, M., A. Bisio, et al. (2001). "Combined quantitative (1)H and (13)C nuclear magnetic resonance spectroscopy for characterization of heparin preparations." *Semin Thromb Hemost* **27**(5): 473-82.
- Guerrini, M., A. Naggi, et al. (2005). "Complex glycosaminoglycans: profiling substitution patterns by two-dimensional nuclear magnetic resonance spectroscopy." *Anal Biochem* **337**(1): 35-47.
- Gupta-Bansal, R., R. C. Frederickson, et al. (1995). "Proteoglycan-mediated inhibition of A beta proteolysis. A potential cause of senile plaque accumulation." *J Biol Chem* **270**(31): 18666-71.
- Guy-Grand, D., J. P. DiSanto, et al. (1998). "Small bowel enteropathy: role of intraepithelial lymphocytes and of cytokines (IL-12, IFN-gamma, TNF) in the induction of epithelial cell death and renewal." *Eur J Immunol* **28**(2): 730-44.

- Haagmans, B. L., P. H. van der Meide, et al. (1994). "Suppression of rat cytomegalovirus replication by antibodies against gamma interferon." J Virol **68**(4): 2305-12.
- Hai, T. and M. G. Hartman (2001). "The molecular biology and nomenclature of the activating transcription factor/cAMP responsive element binding family of transcription factors: activating transcription factor proteins and homeostasis." Gene **273**(1): 1-11.
- Harmer, N. J. (2006). "Insights into the role of heparan sulphate in fibroblast growth factor signalling." Biochem Soc Trans **34**(Pt 3): 442-5.
- Harris, M. (2004). "Monoclonal antibodies as therapeutic agents for cancer." Lancet Oncol **5**(5): 292-302.
- Harvey, E. J. and D. P. Ramji (2005). "Interferon-gamma and atherosclerosis: pro- or anti-atherogenic?" Cardiovasc Res **67**(1): 11-20.
- Heim, M. H., I. M. Kerr, et al. (1995). "Contribution of STAT SH2 groups to specific interferon signaling by the Jak-STAT pathway." Science **267**(5202): 1347-9.
- Heinegard, D., J. Wieslander, et al. (1985). "Separation and characterization of two populations of aggregating proteoglycans from cartilage." Biochem J **225**(1): 95-106.
- Helmlinger, G., F. Yuan, et al. (1997). "Interstitial pH and pO<sub>2</sub> gradients in solid tumors *in vivo*: high-resolution measurements reveal a lack of correlation." Nat Med **3**(2): 177-82.
- Hendrix, M. J., E. A. Seftor, et al. (2003). "Molecular plasticity of human melanoma cells." Oncogene **22**(20): 3070-5.
- Henke, C. A., U. Roongta, et al. (1996). "CD44-related chondroitin sulfate proteoglycan, a cell surface receptor implicated with tumor cell invasion, mediates endothelial cell migration on fibrinogen and invasion into a fibrin matrix." J Clin Invest **97**(11): 2541-52.
- Herbert, J. M., J. P. Hérault, et al. (1998). "Biochemical and pharmacological properties of SANORG 34006, a potent and long-acting synthetic pentasaccharide." Blood **91**(11): 4197-205.
- Hileman, R. E., J. R. Fromm, et al. (1998). "Glycosaminoglycan-protein interactions: definition of consensus sites in glycosaminoglycan binding proteins." Bioessays **20**(2): 156-67.
- Hocking, A. M., R. A. Strugnell, et al. (1996). "Eukaryotic expression of recombinant biglycan. Post-translational processing and the importance of secondary structure for biological activity." J Biol Chem **271**(32): 19571-7.
- Holmes, O., S. Pillozzi, et al. (2007). "Insights into the structure/function of hepatocyte growth factor/scatter factor from studies with individual domains." J Mol Biol **367**(2): 395-408.
- Hommes, D. W., T. L. Mikhajlova, et al. (2006). "Fontolizumab, a humanised anti-interferon gamma antibody, demonstrates safety and clinical activity in patients with moderate to severe Crohn's disease." Gut **55**(8): 1131-7.
- Hoogewerf, A. J., G. S. Kuschert, et al. (1997). "Glycosaminoglycans mediate cell surface oligomerization of chemokines." Biochemistry **36**(44): 13570-8.
- Hoogewerf, A. J., J. W. Leone, et al. (1995). "CXC chemokines connective tissue activating peptide-III and neutrophil activating peptide-2 are heparin/heparan sulfate-degrading enzymes." J Biol Chem **270**(7): 3268-77.
- Hook, M., I. Bjork, et al. (1976). "Anticoagulant activity of heparin: separation of high-activity and low-activity heparin species by affinity chromatography on immobilized antithrombin." FEBS Lett **66**(1): 90-3.
- Hopwood, J., M. Hook, et al. (1976). "Anticoagulant activity of heparin: isolation of antithrombin-binding sites." FEBS Lett **69**(1): 51-4.

- Huang, R., M. J. Merrilees, et al. (2006). "Inhibition of versican synthesis by antisense alters smooth muscle cell phenotype and induces elastic fiber formation *in vitro* and in neointima after vessel injury." Circ Res **98**(3): 370-7.
- Hudson, P. J. and C. Souriau (2003). "Engineered antibodies." Nat Med **9**(1): 129-34.
- Hulett, M. D., C. Freeman, et al. (1999). "Cloning of mammalian heparanase, an important enzyme in tumor invasion and metastasis." Nat Med **5**(7): 803-9.
- Humphries, D. E., G. W. Wong, et al. (1999). "Heparin is essential for the storage of specific granule proteases in mast cells." Nature **400**(6746): 769-72.
- Hwa, V., Y. Oh, et al. (1999). "The insulin-like growth factor-binding protein (IGFBP) superfamily." Endocr Rev **20**(6): 761-87.
- Ijzermans, J. N. and R. L. Marquet (1989). "Interferon-gamma: a review." Immunobiology **179**(4-5): 456-73.
- Innes, J. F., A. R. Barr, et al. (2000). "Efficacy of oral calcium pentosan polysulphate for the treatment of osteoarthritis of the canine stifle joint secondary to cranial cruciate ligament deficiency." Vet Rec **146**(15): 433-7.
- Iozzo, R. V. (1984). "Biosynthesis of heparan sulfate proteoglycan by human colon carcinoma cells and its localization at the cell surface." J Cell Biol **99**(2): 403-17.
- Iozzo, R. V. (1998). "Matrix proteoglycans: from molecular design to cellular function." Annu Rev Biochem **67**: 609-52.
- Iozzo, R. V. and I. Cohen (1993). "Altered proteoglycan gene expression and the tumor stroma." Experientia **49**(5): 447-55.
- Iozzo, R. V. and J. D. San Antonio (2001). "Heparan sulfate proteoglycans: heavy hitters in the angiogenesis arena." J Clin Invest **108**(3): 349-55.
- Isaacs, A. and J. Lindenmann (1987). "Virus interference. I. The interferon. By A. Isaacs and J. Lindenmann, 1957." J Interferon Res **7**(5): 429-38.
- Ishida, K., M. K. Wierzbica, et al. (2004). "Novel heparan sulfate mimetic compounds as antitumor agents." Chem Biol **11**(3): 367-77.
- Itoh, N. and D. M. Ornitz (2004). "Evolution of the Fgf and Fgfr gene families." Trends Genet **20**(11): 563-9.
- Jackson, D. G., J. I. Bell, et al. (1995). "Proteoglycan forms of the lymphocyte homing receptor CD44 are alternatively spliced variants containing the v3 exon." J Cell Biol **128**(4): 673-85.
- Jaffe, M. D. and P. W. Willis, 3rd (1965). "Multiple Fractures Associated with Long-Term Sodium Heparin Therapy." Jama **193**: 158-60.
- Janke, J., S. Engeli, et al. (2006). "Adipose tissue and circulating endothelial cell specific molecule-1 in human obesity." Horm Metab Res **38**(1): 28-33.
- Janke, J., S. Engeli, et al. (2006). "Adipose tissue and circulating endothelial cell specific molecule-1 in human obesity." Horm Metab Res **38**(1): 28-33.
- Jarpe, M. A., M. P. Hayes, et al. (1989). "Causal association of interferon-gamma with tumor regression." J Interferon Res **9**(2): 239-44.
- Jarpe, M. A. and H. M. Johnson (1990). "Structure of an epitope in an immunodominant region of the interferon-gamma molecule that is involved in receptor interaction." J Interferon Res **10**(2): 243-52.
- Jarpe, M. A. and H. M. Johnson (1990). "Topology of receptor binding domains of mouse IFN-gamma." J Immunol **145**(10): 3304-9.
- Jayson, G. C., M. Lyon, et al. (1998). "Heparan sulfate undergoes specific structural changes during the progression from human colon adenoma to carcinoma *in vitro*." J Biol Chem **273**(1): 51-7.
- Jin, L., J. P. Abrahams, et al. (1997). "The anticoagulant activation of antithrombin by heparin." Proc Natl Acad Sci U S A **94**(26): 14683-8.

- Johnson, D. J. and J. A. Huntington (2003). "Crystal structure of antithrombin in a heparin-bound intermediate state." *Biochemistry* **42**(29): 8712-9.
- Johnson, D. J., J. Langdown, et al. (2006). "Crystal structure of monomeric native antithrombin reveals a novel reactive center loop conformation." *J Biol Chem* **281**(46): 35478-86.
- Johnson, D. J., W. Li, et al. (2006). "Antithrombin-S195A factor Xa-heparin structure reveals the allosteric mechanism of antithrombin activation." *Embo J* **25**(9): 2029-37.
- Johnson, P. R. (2001). "Role of human airway smooth muscle in altered extracellular matrix production in asthma." *Clin Exp Pharmacol Physiol* **28**(3): 233-6.
- Joyce, J. A., C. Freeman, et al. (2005). "A functional heparan sulfate mimetic implicates both heparanase and heparan sulfate in tumor angiogenesis and invasion in a mouse model of multistage cancer." *Oncogene* **24**(25): 4037-51.
- Kagaya, K., K. Watanabe, et al. (1989). "Capacity of recombinant gamma interferon to activate macrophages for Salmonella-killing activity." *Infect Immun* **57**(2): 609-15.
- Kainulainen, V., L. Nelimarkka, et al. (1996). "Suppression of syndecan-1 expression in endothelial cells by tumor necrosis factor-alpha." *J Biol Chem* **271**(31): 18759-66.
- Kan, M., F. Wang, et al. (1993). "An essential heparin-binding domain in the fibroblast growth factor receptor kinase." *Science* **259**(5103): 1918-21.
- Kawashima, H., M. Hirose, et al. (2000). "Binding of a large chondroitin sulfate/dermatan sulfate proteoglycan, versican, to L-selectin, P-selectin, and CD44." *J Biol Chem* **275**(45): 35448-56.
- Kelker, H. C., J. Le, et al. (1984). "Three molecular weight forms of natural human interferon-gamma revealed by immunoprecipitation with monoclonal antibody." *J Biol Chem* **259**(7): 4301-4.
- Kim, C. J., J. K. Taubenberger, et al. (1996). "Combination therapy with interferon-gamma and interleukin-2 for the treatment of metastatic melanoma." *J Immunother Emphasis Tumor Immunol* **19**(1): 50-8.
- Kim, H., G. L. Xu, et al. (2003). "The heparan sulfate proteoglycan GPC3 is a potential lung tumor suppressor." *Am J Respir Cell Mol Biol* **29**(6): 694-701.
- Kinsella, M. G., S. L. Bressler, et al. (2004). "The regulated synthesis of versican, decorin, and biglycan: extracellular matrix proteoglycans that influence cellular phenotype." *Crit Rev Eukaryot Gene Expr* **14**(3): 203-34.
- Kirwan, R. P., M. O. Leonard, et al. (2005). "Transforming growth factor-beta-regulated gene transcription and protein expression in human GFAP-negative lamina cribrosa cells." *Glia* **52**(4): 309-24.
- Knox, S. M. and J. M. Whitelock (2006). "Perlécan: how does one molecule do so many things?" *Cell Mol Life Sci* **63**(21): 2435-45.
- Kobayashi, H., T. Yagyu, et al. (2004). "Therapeutic efficacy of once-daily oral administration of a Kunitz-type protease inhibitor, bikunin, in a mouse model and in human cancer." *Cancer* **100**(4): 869-77.
- Koenig, A., K. Norgard-Sumnicht, et al. (1998). "Differential interactions of heparin and heparan sulfate glycosaminoglycans with the selectins. Implications for the use of unfractionated and low molecular weight heparins as therapeutic agents." *J Clin Invest* **101**(4): 877-89.
- Koliakos, G. G., K. Kouzi-Koliakos, et al. (1989). "The binding of heparin to type IV collagen: domain specificity with identification of peptide sequences from the alpha 1(IV) and alpha 2(IV) which preferentially bind heparin." *J Biol Chem* **264**(4): 2313-23.

- Konieczny, B. T., Z. Dai, et al. (1998). "IFN-gamma is critical for long-term allograft survival induced by blocking the CD28 and CD40 ligand T cell costimulation pathways." J Immunol **160**(5): 2059-64.
- Koninger, J., N. A. Giese, et al. (2004). "Overexpressed decorin in pancreatic cancer: potential tumor growth inhibition and attenuation of chemotherapeutic action." Clin Cancer Res **10**(14): 4776-83.
- Kovalszky, I., G. Pogany, et al. (1990). "Altered glycosaminoglycan composition in reactive and neoplastic human liver." Biochem Biophys Res Commun **167**(3): 883-90.
- Kresse, H. and E. Schonherr (2001). "Proteoglycans of the extracellular matrix and growth control." J Cell Physiol **189**(3): 266-74.
- Kreuger, J., P. Jemth, et al. (2005). "Fibroblast growth factors share binding sites in heparan sulphate." Biochem J **389**(Pt 1): 145-50.
- Kreuger, J., D. Spillmann, et al. (2006). "Interactions between heparan sulfate and proteins: the concept of specificity." J Cell Biol **174**(3): 323-7.
- Krusius, T., J. Finne, et al. (1986). "Identification of an O-glycosidic mannose-linked sialylated tetrasaccharide and keratan sulfate oligosaccharides in the chondroitin sulfate proteoglycan of brain." J Biol Chem **261**(18): 8237-42.
- Kulseth, M. A., S. O. Kolset, et al. (1999). "Stimulation of serglycin and CD44 mRNA expression in endothelial cells exposed to TNF-alpha and IL-1alpha." Biochim Biophys Acta **1428**(2-3): 225-32.
- Kuschert, G. S., F. Coulin, et al. (1999). "Glycosaminoglycans interact selectively with chemokines and modulate receptor binding and cellular responses." Biochemistry **38**(39): 12959-68.
- Kussie, P. H., J. D. Hulmes, et al. (1999). "Cloning and functional expression of a human heparanase gene." Biochem Biophys Res Commun **261**(1): 183-7.
- Lainee, P., P. Efron, et al. (2005). "Delayed neutralization of interferon-gamma prevents lethality in primate Gram-negative bacteremic shock." Crit Care Med **33**(4): 797-805.
- Lam, L. H., J. E. Silbert, et al. (1976). "The separation of active and inactive forms of heparin." Biochem Biophys Res Commun **69**(2): 570-7.
- Lamant, M., F. Smih, et al. (2006). "ApoO, a novel apolipoprotein, is an original glycoprotein up-regulated by diabetes in human heart." J Biol Chem **281**(47): 36289-302.
- Lantz, M., H. Thysell, et al. (1991). "On the binding of tumor necrosis factor (TNF) to heparin and the release in vivo of the TNF-binding protein I by heparin." J Clin Invest **88**(6): 2026-31.
- Lardner, A. (2001). "The effects of extracellular pH on immune function." J Leukoc Biol **69**(4): 522-30.
- Lassalle, P., S. Molet, et al. (1996). "ESM-1 is a novel human endothelial cell-specific molecule expressed in lung and regulated by cytokines." J Biol Chem **271**(34): 20458-64.
- Lassalle P., E. Adam, S. Sarrazin et al. (2006). "Proteoglycans and glycosaminoglycans: key players in inflammation" In "New developments in Therapeutic Glycomics". Eds. Maryse Delehedde and Hugues Lortat-Jacob. Chapter 2.
- Melchior A., A. Denys, R. Pakula and F. Allain. (2006). "Applications in therapeutic glycomics: Example of the cyclophilins as novel inflammatory mediators in rheumatoid arthritis" In "New developments in Therapeutic Glycomics". Eds. Maryse Delehedde and Hugues Lortat-Jacob. Chapter 13.
- Lau, E. K., C. D. Paavola, et al. (2004). "Identification of the glycosaminoglycan binding site of the CC chemokine, MCP-1: implications for structure and function *in vivo*." J Biol Chem **279**(21): 22294-305.

- LeBaron, R. G., J. D. Esko, et al. (1988). "Adhesion of glycosaminoglycan-deficient chinese hamster ovary cell mutants to fibronectin substrata." J Cell Biol **106**(3): 945-52.
- Lee, S. C., H. H. Guan, et al. (2005). "Structural basis of citrate-dependent and heparan sulfate-mediated cell surface retention of cobra cardiotoxin A3." J Biol Chem **280**(10): 9567-77.
- Lejeune, F. J., C. Ruegg, et al. (1998). "Clinical applications of TNF-alpha in cancer." Curr Opin Immunol **10**(5): 573-80.
- Lenburg, M. E., L. S. Liou, et al. (2003). "Previously unidentified changes in renal cell carcinoma gene expression identified by parametric analysis of microarray data." BMC Cancer **3**(1): 31.
- Leon, M. L. and S. H. Zuckerman (2005). "Gamma interferon: a central mediator in atherosclerosis." Inflamm Res **54**(10): 395-411.
- Lever, R., J. R. Hoult, et al. (2000). "The effects of heparin and related molecules upon the adhesion of human polymorphonuclear leucocytes to vascular endothelium *in vitro*." Br J Pharmacol **129**(3): 533-40.
- Lever, R. and C. P. Page (2002). "Novel drug development opportunities for heparin." Nat Rev Drug Discov **1**(2): 140-8.
- Levy D.E. (1995). "Interferon induction of gene expression through the Jak-STAT pathway." Semin. Virol. **6**: 181-189.
- Ley, K., M. Cerrito, et al. (1991). "Sulfated polysaccharides inhibit leukocyte rolling in rabbit mesentery venules." Am J Physiol **260**(5 Pt 2): H1667-73.
- Li, D., C. C. Clark, et al. (2000). "Basement membrane zone type XV collagen is a disulfide-bonded chondroitin sulfate proteoglycan in human tissues and cultured cells." J Biol Chem **275**(29): 22339-47.
- Li, F., A. K. Shetty, et al. (2007). "Neuritogenic activity of chondroitin/dermatan sulfate hybrid chains of embryonic pig brain and their mimicry from shark liver. Involvement of the pleiotrophin and hepatocyte growth factor signaling pathways." J Biol Chem **282**(5): 2956-66.
- Li, J. P., M. L. Galvis, et al. (2005). "*In vivo* fragmentation of heparan sulfate by heparanase overexpression renders mice resistant to amyloid protein A amyloidosis." Proc Natl Acad Sci U S A **102**(18): 6473-7.
- Li, W., D. J. Johnson, et al. (2004). "Structure of the antithrombin-thrombin-heparin ternary complex reveals the antithrombotic mechanism of heparin." Nat Struct Mol Biol **11**(9): 857-62.
- Liao, H. X., M. C. Levesque, et al. (1993). "Regulation of human CD44H and CD44E isoform binding to hyaluronan by phorbol myristate acetate and anti-CD44 monoclonal and polyclonal antibodies." J Immunol **151**(11): 6490-9.
- Liaw, P. C., D. L. Becker, et al. (2001). "Molecular basis for the susceptibility of fibrin-bound thrombin to inactivation by heparin cofactor ii in the presence of dermatan sulfate but not heparin." J Biol Chem **276**(24): 20959-65.
- Lietha, D., D. Y. Chirgadze, et al. (2001). "Crystal structures of NK1-heparin complexes reveal the basis for NK1 activity and enable engineering of potent agonists of the MET receptor." Embo J **20**(20): 5543-55.
- Lindahl, U., G. Backstrom, et al. (1980). "Evidence for a 3-O-sulfated D-glucosamine residue in the antithrombin-binding sequence of heparin." Proc Natl Acad Sci U S A **77**(11): 6551-5.
- Lindahl, U., M. Kusche-Gullberg, et al. (1998). "Regulated diversity of heparan sulfate." J Biol Chem **273**(39): 24979-82.
- Linker, A., P. Hoffman, et al. (1958). "Heparitin sulfate." Biochim Biophys Acta **29**(2): 443-4.

- Lipscombe, R. J., A. M. Nakhoul, et al. (1998). "Interleukin-5 binds to heparin/heparan sulfate. A model for an interaction with extracellular matrix." J Leukoc Biol **63**(3): 342-50.
- Liu, Y., H. Yang, et al. (2005). "Effects of vascular endothelial growth factor (VEGF) and chondroitin sulfate A on human monocytic THP-1 cell migration." Colloids Surf B Biointerfaces **43**(3-4): 216-20.
- Lokeshwar, V. B., W. H. Cerwinka, et al. (2005). "HYAL1 hyaluronidase in prostate cancer: a tumor promoter and suppressor." Cancer Res **65**(17): 7782-9.
- Lookene, A., R. Savonen, et al. (1997). "Interaction of lipoproteins with heparan sulfate proteoglycans and with lipoprotein lipase. Studies by surface plasmon resonance technique." Biochemistry **36**(17): 5267-75.
- Lopez-Casillas, F., S. Cheifetz, et al. (1991). "Structure and expression of the membrane proteoglycan betaglycan, a component of the TGF-beta receptor system." Cell **67**(4): 785-95.
- Lortat-Jacob, H., F. Baltzer, et al. (1996). "Heparin decreases the blood clearance of interferon-gamma and increases its activity by limiting the processing of its carboxyl-terminal sequence." J Biol Chem **271**(27): 16139-43.
- Lortat-Jacob, H., C. Brisson, et al. (1996). "Non-receptor-mediated tissue localization of human interferon-gamma: role of heparan sulfate/heparin-like molecules." Cytokine **8**(7): 557-66.
- Lortat-Jacob, H. and J. A. Grimaud (1991). "Interferon-gamma binds to heparan sulfate by a cluster of amino acids located in the C-terminal part of the molecule." FEBS Lett **280**(1): 152-4.
- Lortat-Jacob, H. and J. A. Grimaud (1991). "Interferon-gamma C-terminal function: new working hypothesis. Heparan sulfate and heparin, new targets for IFN-gamma, protect, relax the cytokine and regulate its activity." Cell Mol Biol **37**(3): 253-60.
- Lortat-Jacob, H. and J. A. Grimaud (1992). "Binding of interferon-gamma to heparan sulfate is restricted to the heparin-like domains and involves carboxylic--but not N-sulfated--groups." Biochim Biophys Acta **1117**(2): 126-30.
- Lortat-Jacob, H. and J. A. Grimaud (1994). "[The extracellular matrix: from supporting tissue to regulation of cytokines]." Pathol Biol (Paris) **42**(6): 612-20.
- Lortat-Jacob, H., A. Grosdidier, et al. (2002). "Structural diversity of heparan sulfate binding domains in chemokines." Proc Natl Acad Sci U S A **99**(3): 1229-34.
- Lortat-Jacob, H., H. K. Kleinman, et al. (1991). "High-affinity binding of interferon-gamma to a basement membrane complex (matrigel)." J Clin Invest **87**(3): 878-83.
- Lortat-Jacob, H., J. E. Turnbull, et al. (1995). "Molecular organization of the interferon gamma-binding domain in heparan sulphate." Biochem J **310** ( Pt 2): 497-505.
- Lortat-Jacob, H., P. Fender, R.R. Vivès. (2005). "Virus et héparane sulfate : des mécanismes d'adsorption cellulaire à l'entrée virale." Virologie **9**: 315-25
- Lubineau, A., H. Lortat-Jacob, et al. (2004). "Synthesis of tailor-made glycoconjugate mimetics of heparan sulfate that bind IFN-gamma in the nanomolar range." Chemistry **10**(17): 4265-82.
- Lucchiari, M. A., J. P. Martin, et al. (1991). "Acquired immunity of A/J mice to mouse hepatitis virus 3 infection: dependence on interferon-gamma synthesis and macrophage sensitivity to interferon-gamma." J Gen Virol **72** ( Pt 6): 1317-22.
- Ludwig, R. J., B. Boehme, et al. (2004). "Endothelial P-selectin as a target of heparin action in experimental melanoma lung metastasis." Cancer Res **64**(8): 2743-50.
- Lundell, D. L. and S. K. Narula (1994). "Structural elements required for receptor recognition of human interferon-gamma." Pharmacol Ther **64**(1): 1-21.

- Luster, A. D., S. M. Greenberg, et al. (1995). "The IP-10 chemokine binds to a specific cell surface heparan sulfate site shared with platelet factor 4 and inhibits endothelial cell proliferation." *J Exp Med* **182**(1): 219-31.
- Lyon, M., J. A. Deakin, et al. (2002). "The mode of action of heparan and dermatan sulfates in the regulation of hepatocyte growth factor/scatter factor." *J Biol Chem* **277**(2): 1040-6 Epub 2001 Oct 31.
- Lyon, M., J. A. Deakin, et al. (2004). "The interactions of hepatocyte growth factor/scatter factor and its NK1 and NK2 variants with glycosaminoglycans using a modified gel mobility shift assay. Elucidation of the minimal size of binding and activatory oligosaccharides." *J Biol Chem* **279**(42): 43560-7.
- Lyon, M., J. A. Deakin, et al. (2004). "The interactions of hepatocyte growth factor/scatter factor and its NK1 and NK2 variants with glycosaminoglycans using a modified gel mobility shift assay. Elucidation of the minimal size of binding and activatory oligosaccharides." *J Biol Chem* **279**(42): 43560-7 Epub 2004 Aug 2.
- Lyon, M., J. A. Deakin, et al. (1994). "Interaction of hepatocyte growth factor with heparan sulfate. Elucidation of the major heparan sulfate structural determinants." *J Biol Chem* **269**(15): 11216-23.
- Lyon, M., J. A. Deakin, et al. (1998). "Hepatocyte growth factor/scatter factor binds with high affinity to dermatan sulfate." *J Biol Chem* **273**(1): 271-8.
- Lyon, M. and J. T. Gallagher (1998). "Bio-specific sequences and domains in heparan sulphate and the regulation of cell growth and adhesion." *Matrix Biol* **17**(7): 485-93.
- Ma, P., S. L. Beck, et al. (2006). "Heparanase deglycanation of syndecan-1 is required for binding of the epithelial-restricted prosecretory mitogen lacritin." *J Cell Biol* **174**(7): 1097-106.
- Maheu, E., X. Ayrat, et al. (2002). "A hyaluronan preparation (500-730 kDa) in the treatment of osteoarthritis: a review of clinical trials with Hyalgan." *Int J Clin Pract* **56**(10): 804-13.
- Maheu, E., J. Lamotte, M. Lequesne. (1994). 'Sel sodique de l'acide hyaluronique (hyaluronan) et gonarthrose.' In : De Sèze, S., A. Ryckewaert, MF. Kahn, D. Kuntz, A. Dryll, O. Meyer, Th. Bardin, Cl. Guérin. L'actualité rhumatologique, Paris, Expansion Scientifique Française Ed. 324-339
- Maimone, M. M. and D. M. Tollefsen (1990). "Structure of a dermatan sulfate hexasaccharide that binds to heparin cofactor II with high affinity." *J Biol Chem* **265**(30): 18263-71.
- Marchetti, D. and G. L. Nicolson (2001). "Human heparanase: a molecular determinant of brain metastasis." *Adv Enzyme Regul* **41**: 343-59.
- Marshall, J. L., A. Wellstein, et al. (1997). "Phase I trial of orally administered pentosan polysulfate in patients with advanced cancer." *Clin Cancer Res* **3**(12 Pt 1): 2347-54.
- Martin, T. A., G. Harrison, et al. (2003). "The role of the CD44/ezrin complex in cancer metastasis." *Crit Rev Oncol Hematol* **46**(2): 165-86.
- Mascellani, G., L. Liverani, et al. (1993). "Structure and contribution to the heparin cofactor II-mediated inhibition of thrombin of naturally oversulphated sequences of dermatan sulphate." *Biochem J* **296**(Pt 3): 639-48.
- Mascellani, G., L. Liverani, et al. (1993). "Structure and contribution to the heparin cofactor II-mediated inhibition of thrombin of naturally oversulphated sequences of dermatan sulphate." *Biochem J* **296** ( Pt 3): 639-48.
- Massague, J. (1985). "Subunit structure of a high-affinity receptor for type beta-transforming growth factor. Evidence for a disulfide-linked glycosylated receptor complex." *J Biol Chem* **260**(11): 7059-66.

- Matsuzaki, H., H. Kobayashi, et al. (2004). "Bikunin inhibits lipopolysaccharide-induced tumor necrosis factor alpha induction in macrophages." Clin Diagn Lab Immunol **11**(6): 1140-7.
- Matsuzaki, H., H. Kobayashi, et al. (2005). "Plasma bikunin as a favorable prognostic factor in ovarian cancer." J Clin Oncol **23**(7): 1463-72.
- Maxhimer, J. B., R. M. Quiros, et al. (2002). "Heparanase-1 expression is associated with the metastatic potential of breast cancer." Surgery **132**(2): 326-33.
- Mayer, J., M. L. Woods, et al. (1993). "Gamma interferon-induced nitric oxide production reduces Chlamydia trachomatis infectivity in McCoy cells." Infect Immun **61**(2): 491-7.
- Mazieres, B., G. Loyau, et al. (1992). "[Chondroitin sulfate in the treatment of gonarthrosis and coxarthrosis. 5-months result of a multicenter double-blind controlled prospective study using placebo]." Rev Rhum Mal Osteoartic **59**(7-8): 466-72.
- McAlindon, T. E., M. P. LaValley, et al. (2000). "Glucosamine and chondroitin for treatment of osteoarthritis: a systematic quality assessment and meta-analysis." Jama **283**(11): 1469-75.
- McCaffrey, T. A., D. J. Falcone, et al. (1994). "Protection of transforming growth factor-beta 1 activity by heparin and fucoidan." J Cell Physiol **159**(1): 51-9.
- McEwan, P. A., P. G. Scott, et al. (2006). "Structural correlations in the family of small leucine-rich repeat proteins and proteoglycans." J Struct Biol **155**(2): 294-305.
- McKenzie, E., K. Tyson, et al. (2000). "Cloning and expression profiling of Hpa2, a novel mammalian heparanase family member." Biochem Biophys Res Commun **276**(3): 1170-7.
- McLachlin, A. D., S. E. Carroll, et al. (1970). "Experimental venous thrombectomy." Ann Surg **171**(6): 956-60.
- Mc Lean, J. (1916). "The thromboplastic action of cephalin." Am. J. Physiol. **41**: 250-257.
- Meddahi, A., J. P. Caruelle, et al. (1996). "New concepts in tissue repair: skin as an example." Diabetes Metab **22**(4): 274-8.
- Meddahi, A., H. Lemdjabar, et al. (1995). "Inhibition by dextran derivatives of FGF-2 plasmin-mediated degradation." Biochimie **77**(9): 703-6.
- Meddahi, A., H. Lemdjabar, et al. (1996). "FGF protection and inhibition of human neutrophil elastase by carboxymethyl benzylamide sulfonate dextran derivatives." Int J Biol Macromol **18**(1-2): 141-5.
- Melching, L. I., W. D. Fisher, et al. (2006). "The cleavage of biglycan by aggrecanases." Osteoarthritis Cartilage.
- Melrose, J., S. Smith, et al. (2004). "Perlécan immunolocalizes to perichondrial vessels and canals in human fetal cartilaginous primordia in early vascular and matrix remodeling events associated with diarthrodial joint development." J Histochem Cytochem **52**(11): 1405-13.
- Merle, B., L. Durussel, et al. (1999). "Decorin inhibits cell migration through a process requiring its glycosaminoglycan side chain." J Cell Biochem **75**(3): 538-46.
- Mertens, G., B. Van der Schueren, et al. (1996). "Heparan sulfate expression in polarized epithelial cells: the apical sorting of glypican (GPI-anchored proteoglycan) is inversely related to its heparan sulfate content." J Cell Biol **132**(3): 487-97.
- Messmore, H. L., W. H. Wehrmacher, et al. (2004). "Heparin to pentasaccharide and beyond: the end is not in sight." Semin Thromb Hemost **30 Suppl 1**: 81-8.
- Meyer, K., A. Linker, et al. (1953). "The mucopolysaccharides of bovine cornea." J Biol Chem **205**(2): 611-6.
- Meyer, K. H. and J. Fellig (1950). "[Constitution of hyaluronic acid.]" Experientia **6**(5): 186.
- Meyer, K. and J. W. Palmer (1934). "The polysaccharide in vitreous humor." J Biol. Chem.

107: 629-634

- Middleton, J., S. Neil, et al. (1997). "Transcytosis and surface presentation of IL-8 by venular endothelial cells." *Cell* **91**(3): 385-95.
- Middleton, J., A. M. Patterson, et al. (2002). "Leukocyte extravasation: chemokine transport and presentation by the endothelium." *Blood* **100**(12): 3853-60.
- Miller, M. J., C. E. Costello, et al. (2006). "A tandem mass spectrometric approach to determination of chondroitin/dermatan sulfate oligosaccharide glycoforms." *Glycobiology* **16**(6): 502-13.
- Miller, W. E. and V. G. DeWolfe (1966). "Osteoporosis resulting from heparin therapy. Report of a case." *Cleve Clin Q* **33**(1): 31-4.
- Mitropoulou, T. N., F. Lamari, et al. (2001). "Identification of oligomeric domains within dermatan sulfate chains using differential enzymic treatments, derivatization with 2-aminoacridone and capillary electrophoresis." *Electrophoresis* **22**(12): 2458-63.
- Mohammadi, M., S. K. Olsen, et al. (2005). "A protein canyon in the FGF-FGF receptor dimer selects from an a la carte menu of heparan sulfate motifs." *Curr Opin Struct Biol* **15**(5): 506-16.
- Mohammadi, M., S. K. Olsen, et al. (2005). "Structural basis for fibroblast growth factor receptor activation." *Cytokine Growth Factor Rev* **16**(2): 107-37.
- Mongiati, M., K. Taylor, et al. (2000). "The protein core of the proteoglycan perlecan binds specifically to fibroblast growth factor-7." *J Biol Chem* **275**(10): 7095-100.
- Monkhouse, F. C., E. S. France, et al. (1955). "Studies on the antithrombin and heparin cofactor activities of a fraction adsorbed from plasma by aluminum hydroxide." *Circ Res* **3**(4): 397-402.
- Moon, A. F., S. C. Edavettal, et al. (2004). "Structural analysis of the sulfotransferase (3-O-sulfotransferase isoform 3) involved in the biosynthesis of an entry receptor for herpes simplex virus 1." *J Biol Chem* **279**(43): 45185-93.
- Moore, J. P. and M. Stevenson (2000). "New targets for inhibitors of HIV-1 replication." *Nat Rev Mol Cell Biol* **1**(1): 40-9.
- Moore, K. W., A. O'Garra, et al. (1993). "Interleukin-10." *Annu Rev Immunol* **11**: 165-90.
- Morreale, P., R. Manopulo, et al. (1996). "Comparison of the antiinflammatory efficacy of chondroitin sulfate and diclofenac sodium in patients with knee osteoarthritis." *J Rheumatol* **23**(8): 1385-91.
- Moulard, M., H. Lortat-Jacob, et al. (2000). "Selective interactions of polyanions with basic surfaces on human immunodeficiency virus type 1 gp120." *J Virol* **74**(4): 1948-60.
- Moy, F. J., M. Safran, et al. (1997). "Properly oriented heparin-decasaccharide-induced dimers are the biologically active form of basic fibroblast growth factor." *Biochemistry* **36**(16): 4782-91.
- Mulloy, B. (2005). "The specificity of interactions between proteins and sulfated polysaccharides." *An Acad Bras Cienc* **77**(4): 651-64.
- Munir, M., H. Tjandra, et al. (1980). "Heparin in the treatment of cerebral malaria." *Paediatr Indones* **20**(1-2): 47-50.
- Murakami, M., A. Horowitz, et al. (2002). "Protein kinase C (PKC) delta regulates PKCalpha activity in a Syndecan-4-dependent manner." *J Biol Chem* **277**(23): 20367-71.
- Murphy, J. W., Y. Cho, et al. (2007). "Structural and functional basis of CXCL12 (stromal cell-derived factor-1alpha) binding to heparin." *J Biol Chem* **282**(13): 10018-27.
- Murry, B. P., B. E. Blust, et al. (2006). "Heparanase mechanisms of melanoma metastasis to the brain: Development and use of a brain slice model." *J Cell Biochem* **97**(2): 217-25.
- Nacheva, G., K. Todorova, et al. (2003). "Human interferon gamma: significance of the C-terminal flexible domain for its biological activity." *Arch Biochem Biophys* **413**(1): 91-8.

- Nagano, Y., Y. Kojima, et al. (1954). "[Immunity and interference in vaccinia; inhibition of skin infection by inactivated virus.]" *C R Seances Soc Biol Fil* **148**(7-8): 750-2.
- Naito, Z. (2005). "Role of the small leucine-rich proteoglycan (SLRP) family in pathological lesions and cancer cell growth." *J Nippon Med Sch* **72**(3): 137-45.
- Nakamura, L. T., B. A. Wu-Hsieh, et al. (1994). "Recombinant murine gamma interferon stimulates macrophages of the RAW cell line to inhibit intracellular growth of *Histoplasma capsulatum*." *Infect Immun* **62**(2): 680-4.
- Nandi, A., P. Estess, et al. (2000). "Hyaluronan anchoring and regulation on the surface of vascular endothelial cells is mediated through the functionally active form of CD44." *J Biol Chem* **275**(20): 14939-48.
- Napolitano, L. M., T. E. Warkentin, et al. (2006). "Heparin-induced thrombocytopenia in the critical care setting: diagnosis and management." *Crit Care Med* **34**(12): 2898-911.
- Nara, Y., Y. Kato, et al. (1997). "Immunohistochemical localization of extracellular matrix components in human breast tumours with special reference to PG-M/versican." *Histochem J* **29**(1): 21-30.
- Nasu, R., H. Kimura, et al. (1999). "Blood flow influences vascular growth during tumour angiogenesis." *Br J Cancer* **79**(5-6): 780-6.
- Nathan, C. F., M. L. Karnovsky, et al. (1971). "Alterations of macrophage functions by mediators from lymphocytes." *J Exp Med* **133**(6): 1356-76.
- Nathan, C. F., H. W. Murray, et al. (1983). "Identification of interferon-gamma as the lymphokine that activates human macrophage oxidative metabolism and antimicrobial activity." *J Exp Med* **158**(3): 670-89.
- Nathan, C. F., H. G. Remold, et al. (1973). "Characterization of a lymphocyte factor which alters macrophage functions." *J Exp Med* **137**(2): 275-90.
- Naylor, S. L., A. Y. Sakaguchi, et al. (1983). "Human immune interferon gene is located on chromosome 12." *J Exp Med* **157**(3): 1020-7.
- Nelimarkka, L., V. Kainulainen, et al. (1997). "Expression of small extracellular chondroitin/dermatan sulfate proteoglycans is differentially regulated in human endothelial cells." *J Biol Chem* **272**(19): 12730-7.
- Nelimarkka, L., H. Salminen, et al. (2001). "Decorin is produced by capillary endothelial cells in inflammation-associated angiogenesis." *Am J Pathol* **158**(2): 345-53.
- Nelson, R. M., O. Cecconi, et al. (1993). "Heparin oligosaccharides bind L- and P-selectin and inhibit acute inflammation." *Blood* **82**(11): 3253-8.
- Nenci, G. G. (2002). "Dermatan sulphate as an antithrombotic drug." *Pathophysiol Haemost Thromb* **32**(5-6): 303-7.
- Netelenbos, T., J. van den Born, et al. (2003). "Proteoglycans on bone marrow endothelial cells bind and present SDF-1 towards hematopoietic progenitor cells." *Leukemia* **17**(1): 175-84.
- Netelenbos, T., S. Zuijderduijn, et al. (2002). "Proteoglycans guide SDF-1-induced migration of hematopoietic progenitor cells." *J Leukoc Biol* **72**(2): 353-62.
- Neumann, H., H. Schmidt, et al. (1997). "Interferon gamma gene expression in sensory neurons: evidence for autocrine gene regulation." *J Exp Med* **186**(12): 2023-31.
- Nizet, V., T. Ohtake, et al. (2001). "Innate antimicrobial peptide protects the skin from invasive bacterial infection." *Nature* **414**(6862): 454-7.
- Nizolek, D. J. and K. K. White (1981). "Corticosteroid and hyaluronic acid treatments in equine degenerative joint disease. A review." *Cornell Vet* **71**(4): 355-75.
- Noble, P. W. (2002). "Hyaluronan and its catabolic products in tissue injury and repair." *Matrix Biol* **21**(1): 25-9.
- Noonan, D. M., A. Fulle, et al. (1991). "The complete sequence of perlécan, a basement membrane heparan sulfate proteoglycan, reveals extensive similarity with laminin A

- chain, low density lipoprotein-receptor, and the neural cell adhesion molecule." J Biol Chem **266**(34): 22939-47.
- Nyberg, K., M. Ekblad, et al. (2004). "The low molecular weight heparan sulfate-mimetic, PI-88, inhibits cell-to-cell spread of herpes simplex virus." Antiviral Res **63**(1): 15-24.
- Oikawa, T. and T. Yamada (2003). "Molecular biology of the Ets family of transcription factors." Gene **303**: 11-34.
- Oldberg, A., P. Antonsson, et al. (1996). "Amino-terminal deletions in the decorin core protein leads to the biosynthesis of proteoglycans with shorter glycosaminoglycan chains." FEBS Lett **386**(1): 29-32.
- Oliviero, U., G. P. Sorrentino, et al. (1991). "Effects of the treatment with matrix on elderly people with chronic articular degeneration." Drugs Exp Clin Res **17**(1): 45-51.
- Olsen, H. M., C. R. Parish, et al. (1996). "Histidine-rich glycoprotein binding to T-cell lines and its effect on T-cell substratum adhesion is strongly potentiated by zinc." Immunology **88**(2): 198-206.
- Olson, S. T., I. Bjork, et al. (1992). "Role of the antithrombin-binding pentasaccharide in heparin acceleration of antithrombin-proteinase reactions. Resolution of the antithrombin conformational change contribution to heparin rate enhancement." J Biol Chem **267**(18): 12528-38.
- Ornitz, D. M., A. Yayon, et al. (1992). "Heparin is required for cell-free binding of basic fibroblast growth factor to a soluble receptor and for mitogenesis in whole cells." Mol Cell Biol **12**(1): 240-7.
- Overall, C. M., G. A. McQuibban, et al. (2002). "Discovery of chemokine substrates for matrix metalloproteinases by exosite scanning: a new tool for degradomics." Biol Chem **383**(7-8): 1059-66.
- Parish, C. R. (2006). "The role of heparan sulphate in inflammation." Nat Rev Immunol **6**(9): 633-43.
- Parish, C. R., C. Freeman, et al. (1999). "Identification of sulfated oligosaccharide-based inhibitors of tumor growth and metastasis using novel *in vitro* assays for angiogenesis and heparanase activity." Cancer Res **59**(14): 3433-41.
- Parish, C. R., E. J. Hindmarsh, et al. (1998). "Treatment of central nervous system inflammation with inhibitors of basement membrane degradation." Immunol Cell Biol **76**(1): 104-13.
- Park, P. W., G. B. Pier, et al. (2001). "Exploitation of syndecan-1 shedding by *Pseudomonas aeruginosa* enhances virulence." Nature **411**(6833): 98-102.
- Parrish, R. F. and W. R. Fair (1981). "Selective binding of zinc ions to heparin rather than to other glycosaminoglycans." Biochem J **193**(2): 407-10.
- Partenheimer, A., K. Schwarz, et al. (1995). "Proteoglycan form of colony-stimulating factor-1 (proteoglycan-100). Stimulation of activity by glycosaminoglycan removal and proteolytic processing." J Immunol **155**(12): 5557-65.
- Pavlidis, N., C. Nicolaidis, et al. (1996). "Phase II study of 5-fluorouracil and interferon-gamma in patients with metastatic colorectal cancer. A Hellenic Cooperative Oncology Group Study." Oncology **53**(2): 159-62.
- Peck, R. (1989). "Gamma interferon induces monocyte killing of *Listeria monocytogenes* by an oxygen-dependent pathway; alpha- or beta-interferons by oxygen-independent pathways." J Leukoc Biol **46**(5): 434-40.
- Pellegrini, L., D. F. Burke, et al. (2000). "Crystal structure of fibroblast growth factor receptor ectodomain bound to ligand and heparin." Nature **407**(6807): 1029-34.
- Pellegrini, S. and C. Schindler (1993). "Early events in signalling by interferons." Trends Biochem Sci **18**(9): 338-42.

- Penc, S. F., B. Pomahac, et al. (1999). "Dermatan sulfate activates nuclear factor-kappaB and induces endothelial and circulating intercellular adhesion molecule-1." J Clin Invest **103**(9): 1329-35.
- Penc, S. F., B. Pomahac, et al. (1998). "Dermatan sulfate released after injury is a potent promoter of fibroblast growth factor-2 function." J Biol Chem **273**(43): 28116-21.
- Petersen, F., L. Bock, et al. (1998). "A chondroitin sulfate proteoglycan on human neutrophils specifically binds platelet factor 4 and is involved in cell activation." J Immunol **161**(8): 4347-55.
- Petersen, F., E. Brandt, et al. (1999). "Characterization of a neutrophil cell surface glycosaminoglycan that mediates binding of platelet factor 4." J Biol Chem **274**(18): 12376-82.
- Petitou, M., B. Casu, et al. (2003). "1976-1983, a critical period in the history of heparin: the discovery of the antithrombin binding site." Biochimie **85**(1-2): 83-9.
- Petitou, M., P. Duchaussoy, et al. (1988). "Binding of heparin to antithrombin III: a chemical proof of the critical role played by a 3-sulfated 2-amino-2-deoxy-D-glucose residue." Carbohydr Res **179**: 163-72.
- Petitou, M., P. Duchaussoy, et al. (1986). "Synthesis of heparin fragments. A chemical synthesis of the pentasaccharide O-(2-deoxy-2-sulfamido-6-O-sulfo-alpha-D-glucopyranosyl)-(1-4)-O-(beta-D-glucopyranosyluronic acid)-(1-4)-O-(2-deoxy-2-sulfamido-3,6-di-O-sulfo-alpha-D-glucopyranosyl)-(1-4)-O-(2-O-sulfo-alpha-L-idopyranosyluronic acid)-(1-4)-2-deoxy-2-sulfamido-6-O-sulfo-D-glucopyranose decasodium salt, a heparin fragment having high affinity for antithrombin III." Carbohydr Res **147**(2): 221-36.
- Petitou, M., J. P. Herault, et al. (1999). "Synthesis of thrombin-inhibiting heparin mimetics without side effects." Nature **398**(6726): 417-22.
- Petitou, M. and C. A. van Boeckel (2004). "A synthetic antithrombin III binding pentasaccharide is now a drug! What comes next?" Angew Chem Int Ed Engl **43**(24): 3118-33.
- Peyron, J. G. (1993). "A new approach to the treatment of osteoarthritis: viscosupplementation." Osteoarthritis Cartilage **1**(2): 85-7.
- Peyron, J. G. and E. A. Balazs (1974). "Preliminary clinical assessment of Na-hyaluronate injection into human arthritic joints." Pathol Biol (Paris) **22**(8): 731-6.
- Picker, L. J., M. Nakache, et al. (1989). "Monoclonal antibodies to human lymphocyte homing receptors define a novel class of adhesion molecules on diverse cell types." J Cell Biol **109**(2): 927-37.
- Pluda, J. M., L. E. Shay, et al. (1993). "Administration of pentosan polysulfate to patients with human immunodeficiency virus-associated Kaposi's sarcoma." J Natl Cancer Inst **85**(19): 1585-92.
- Ponta, H., L. Sherman, et al. (2003). "CD44: from adhesion molecules to signalling regulators." Nat Rev Mol Cell Biol **4**(1): 33-45.
- Prat, M., S. Bretti, et al. (1987). "Monoclonal antibodies against murine IFN-gamma abrogate *in vivo* tumor immunity against RSV-induced murine sarcomas." J Immunol **138**(12): 4530-3.
- Proudfoot, A. E., S. Fritchley, et al. (2001). "The BBXB motif of RANTES is the principal site for heparin binding and controls receptor selectivity." J Biol Chem **276**(14): 10620-6.
- Proudfoot, A. E., T. M. Handel, et al. (2003). "Glycosaminoglycan binding and oligomerization are essential for the *in vivo* activity of certain chemokines." Proc Natl Acad Sci U S A **100**(4): 1885-90.

- Proudfoot, A. E., M. C. Peitsch, et al. (1997). "Structure and bioactivity of recombinant human CTAP-III and NAP-2." J Protein Chem **16**(1): 37-49.
- Prydz, K. and K. T. Dalen (2000). "Synthesis and sorting of proteoglycans." J Cell Sci **113**(Pt 2): 193-205.
- Pye, D. A., R. R. Vives, et al. (1998). "Heparan sulfate oligosaccharides require 6-O-sulfation for promotion of basic fibroblast growth factor mitogenic activity." J Biol Chem **273**(36): 22936-42.
- Rabenstein, D. L. (2002). "Heparin and heparan sulfate: structure and function." Nat Prod Rep **19**(3): 312-31.
- Rada, J. A., P. K. Cornuet, et al. (1993). "Regulation of corneal collagen fibrillogenesis *in vitro* by corneal proteoglycan (lumican and decorin) core proteins." Exp Eye Res **56**(6): 635-48.
- Rampengan, T. H. (1991). "Cerebral malaria in children. Comparative study between heparin, dexamethasone and placebo." Paediatr Indones **31**(1-2): 59-66.
- Rapraeger, A. C. (2000). "Syndecan-regulated receptor signaling." J Cell Biol **149**(5): 995-8.
- Rapraeger, A. C. (2001). "Molecular interactions of syndecans during development." Semin Cell Dev Biol **12**(2): 107-16.
- Rapraeger, A. C., A. Krufka, et al. (1991). "Requirement of heparan sulfate for bFGF-mediated fibroblast growth and myoblast differentiation." Science **252**(5013): 1705-8.
- Rauch, U., K. Feng, et al. (2001). "Neurocan: a brain chondroitin sulfate proteoglycan." Cell Mol Life Sci **58**(12-13): 1842-56.
- Rauch, U., L. Karthikeyan, et al. (1992). "Cloning and primary structure of neurocan, a developmentally regulated, aggregating chondroitin sulfate proteoglycan of brain." J Biol Chem **267**(27): 19536-47.
- Reed, C. C., A. Waterhouse, et al. (2005). "Decorin prevents metastatic spreading of breast cancer." Oncogene **24**(6): 1104-10.
- Rees, M. D. and M. J. Davies (2006). "Heparan sulfate degradation via reductive homolysis of its N-chloro derivatives." J Am Chem Soc **128**(9): 3085-97.
- Rees, M. D., T. N. McNiven, et al. (2007). "Degradation of extracellular matrix and its components by hypobromous acid." Biochem J **401**(2): 587-96.
- Reeves, E. P., H. Lu, et al. (2002). "Killing activity of neutrophils is mediated through activation of proteases by K<sup>+</sup> flux." Nature **416**(6878): 291-7.
- Reinisch, W., D. W. Hommes, et al. (2006). "A dose escalating, placebo controlled, double blind, single dose and multidose, safety and tolerability study of fontolizumab, a humanised anti-interferon gamma antibody, in patients with moderate to severe Crohn's disease." Gut **55**(8): 1138-44.
- Renne, T., J. Dedio, et al. (2000). "High molecular weight kininogen utilizes heparan sulfate proteoglycans for accumulation on endothelial cells." J Biol Chem **275**(43): 33688-96.
- Rennick, D. M., M. M. Fort, et al. (1997). "Studies with IL-10<sup>-/-</sup> mice: an overview." J Leukoc Biol **61**(4): 389-96.
- Ricard-Blum, S., O. Feraud, et al. (2004). "Characterization of endostatin binding to heparin and heparan sulfate by surface plasmon resonance and molecular modeling: role of divalent cations." J Biol Chem **279**(4): 2927-36.
- Rider, C. C. (1997). "The potential for heparin and its derivatives in the therapy and prevention of HIV-1 infection." Glycoconj J **14**(5): 639-42.
- Rinderknecht, E., B. H. O'Connor, et al. (1984). "Natural human interferon-gamma. Complete amino acid sequence and determination of sites of glycosylation." J Biol Chem **259**(11): 6790-7.
- Ringel, J., R. Jesnowski, et al. (2001). "CD44 in normal human pancreas and pancreatic carcinoma cell lines." Teratog Carcinog Mutagen **21**(1): 97-106.

- Ronca, F., L. Palmieri, et al. (1998). "Anti-inflammatory activity of chondroitin sulfate." Osteoarthritis Cartilage **6 Suppl A**: 14-21.
- Rong, Y., D. L. Durden, et al. (2006). "'Pseudopalisading' necrosis in glioblastoma: a familiar morphologic feature that links vascular pathology, hypoxia, and angiogenesis." J Neuropathol Exp Neurol **65**(6): 529-39.
- Rosenberg, R. D. and P. S. Damus (1973). "The purification and mechanism of action of human antithrombin-heparin cofactor." J Biol Chem **248**(18): 6490-505.
- Ross, D. T., U. Scherf, et al. (2000). "Systematic variation in gene expression patterns in human cancer cell lines." Nat Genet **24**(3): 227-35.
- Rot, A. (1992). "Endothelial cell binding of NAP-1/IL-8: role in neutrophil emigration." Immunol Today **13**(8): 291-4.
- Rouet, V., Y. Hamma-Kourbali, et al. (2005). "A synthetic glycosaminoglycan mimetic binds vascular endothelial growth factor and modulates angiogenesis." J Biol Chem **280**(38): 32792-800.
- Rouet, V., A. Meddahi-Pelle, et al. (2006). "Heparin-like synthetic polymers, named RGTAs, mimic biological effects of heparin *in vitro*." J Biomed Mater Res A **78**(4): 792-7.
- Roughley, P., D. Martens, et al. (2006). "The involvement of aggrecan polymorphism in degeneration of human intervertebral disc and articular cartilage." Eur Cell Mater **11**: 1-7; discussion 7.
- Roughley, P. J. (2006). "The structure and function of cartilage proteoglycans." Eur Cell Mater **12**: 92-101.
- Roughley, P. J. and E. R. Lee (1994). "Cartilage proteoglycans: structure and potential functions." Microsc Res Tech **28**(5): 385-97.
- Rovetta, G. (1991). "Galactosaminoglycuronoglycan sulfate (matrix) in therapy of tibiofibular osteoarthritis of the knee." Drugs Exp Clin Res **17**(1): 53-7.
- Rovetta, G., P. Monteforte, et al. (2004). "A two-year study of chondroitin sulfate in erosive osteoarthritis of the hands: behavior of erosions, osteophytes, pain and hand dysfunction." Drugs Exp Clin Res **30**(1): 11-6.
- Rudzki, Z. and S. Jothy (1997). "CD44 and the adhesion of neoplastic cells." Mol Pathol **50**(2): 57-71.
- Ruegg, C. (2006). "Leukocytes, inflammation, and angiogenesis in cancer: fatal attractions." J Leukoc Biol **80**(4): 682-4.
- Ruiz-Bustos, E., J. L. Ochoa, et al. (2001). "Isolation and characterisation of putative adhesins from *Helicobacter pylori* with affinity for heparan sulphate proteoglycan." J Med Microbiol **50**(3): 215-22.
- Rupp, W. M., H. B. McCarthy, et al. (1982). "Risk of osteoporosis in patients treated with long-term intravenous heparin therapy." Curr Surg **39**(6): 419-22.
- Rydell, N. and E. A. Balazs (1971). "Effect of intra-articular injection of hyaluronic acid on the clinical symptoms of osteoarthritis and on granulation tissue formation." Clin Orthop Relat Res **80**: 25-32.
- Sadir, R., F. Baleux, et al. (2001). "Characterization of the stromal cell-derived factor-1alpha-heparin complex." J Biol Chem **276**(11): 8288-96.
- Sadir, R., E. Forest, et al. (1998). "The heparan sulfate binding sequence of interferon-gamma increased the on rate of the interferon-gamma-interferon-gamma receptor complex formation." J Biol Chem **273**(18): 10919-25.
- Sadir, R., A. Imberty, et al. (2004). "Heparan sulfate/heparin oligosaccharides protect stromal cell-derived factor-1 (SDF-1)/CXCL12 against proteolysis induced by CD26/dipeptidyl peptidase IV." J Biol Chem **279**(42): 43854-60.
- Safran, M., M. Eisenstein, et al. (2000). "Oligomerization reduces heparin affinity but enhances receptor binding of fibroblast growth factor 2." Biochem J **345 Pt 1**: 107-13.

- Sakatsume, M., K. Igarashi, et al. (1995). "The Jak kinases differentially associate with the alpha and beta (accessory factor) chains of the interferon gamma receptor to form a functional receptor unit capable of activating STAT transcription factors." J Biol Chem **270**(29): 17528-34.
- Sakko, A. J., C. Ricciardelli, et al. (2003). "Modulation of prostate cancer cell attachment to matrix by versican." Cancer Res **63**(16): 4786-91.
- Saksela, O., D. Moscatelli, et al. (1988). "Endothelial cell-derived heparan sulfate binds basic fibroblast growth factor and protects it from proteolytic degradation." J Cell Biol **107**(2): 743-51.
- Salmivirta, M., K. Lidholt, et al. (1996). "Heparan sulfate: a piece of information." Faseb J **10**(11): 1270-9.
- Sanderson, R. D., Y. Yang, et al. (2004). "Heparan sulfate proteoglycans and heparanase--partners in osteolytic tumor growth and metastasis." Matrix Biol **23**(6): 341-52.
- Santra, M., I. Eichstetter, et al. (2000). "An anti-oncogenic role for decorin. Down-regulation of ErbB2 leads to growth suppression and cytodifferentiation of mammary carcinoma cells." J Biol Chem **275**(45): 35153-61.
- Santra, M., C. C. Reed, et al. (2002). "Decorin binds to a narrow region of the epidermal growth factor (EGF) receptor, partially overlapping but distinct from the EGF-binding epitope." J Biol Chem **277**(38): 35671-81.
- Saoncella, S., F. Echtermeyer, et al. (1999). "Syndecan-4 signals cooperatively with integrins in a Rho-dependent manner in the assembly of focal adhesions and actin stress fibers." Proc Natl Acad Sci U S A **96**(6): 2805-10.
- Saphire, A. C., M. D. Bobardt, et al. (2001). "Syndecans serve as attachment receptors for human immunodeficiency virus type 1 on macrophages." J Virol **75**(19): 9187-200.
- Sareneva, T., E. Mortz, et al. (1996). "Biosynthesis and N-glycosylation of human interferon-gamma. Asn25 and Asn97 differ markedly in how efficiently they are glycosylated and in their oligosaccharide composition." Eur J Biochem **242**(2): 191-200.
- Sarraj, B., K. Ludanyi, et al. (2006). "Expression of CD44 and L-selectin in the innate immune system is required for severe joint inflammation in the proteoglycan-induced murine model of rheumatoid arthritis." J Immunol **177**(3): 1932-40.
- Sarrazin, S., E. Adam, et al. (2006). "Endocan or endothelial cell specific molecule-1 (ESM-1): A potential novel endothelial cell marker and a new target for cancer therapy." Biochim Biophys Acta. **1765**(1): 25-37 Epub 2005 Aug 26.
- Sarrazin, S., D. Bonnaffe, et al. (2005). "Heparan sulfate mimicry: a synthetic glycoconjugate that recognizes the heparin binding domain of interferon-gamma inhibits the cytokine activity." J Biol Chem **280**(45): 37558-64.
- Sasisekharan, R., Z. Shriver, et al. (2002). "Roles of heparan-sulphate glycosaminoglycans in cancer." Nat Rev Cancer **2**(7): 521-8.
- Saunders, S., S. Paine-Saunders, et al. (1997). "Expression of the cell surface proteoglycan glypican-5 is developmentally regulated in kidney, limb, and brain." Dev Biol **190**(1): 78-93.
- Savi, P., A. M. Pflieger, et al. (2003). "Comparative effects of two synthetic oligosaccharides on platelet activation induced by plasma from HIT patients." J Thromb Haemost **1**(9): 2008-13.
- Sawhney, R. S., T. M. Hering, et al. (1991). "Biosynthesis of small proteoglycan II (decorin) by chondrocytes and evidence for a procure protein." J Biol Chem **266**(14): 9231-40.
- Scherpereel, A., F. Depontieu, et al. (2006). "Endocan, a new endothelial marker in human sepsis." Crit Care Med **34**(2): 532-7.
- Scherpereel, A., T. Gentina, et al. (2003). "Overexpression of endocan induces tumor formation." Cancer Res **63**(18): 6084-9.

- Schick, B. P., J. F. Gradowski, et al. (2001). "Synthesis, secretion, and subcellular localization of serglycin proteoglycan in human endothelial cells." Blood **97**(2): 449-58.
- Schick, B. P., H. C. Ho, et al. (2003). "Serglycin proteoglycan expression and synthesis in embryonic stem cells." Biochim Biophys Acta **1593**(2-3): 259-67.
- Schindler, C. and J. E. Darnell, Jr. (1995). "Transcriptional responses to polypeptide ligands: the JAK-STAT pathway." Annu Rev Biochem **64**: 621-51.
- Schindler, C., K. Shuai, et al. (1992). "Interferon-dependent tyrosine phosphorylation of a latent cytoplasmic transcription factor." Science **257**(5071): 809-13.
- Schlessinger, J., A. N. Plotnikov, et al. (2000). "Crystal structure of a ternary FGF-FGFR-heparin complex reveals a dual role for heparin in FGFR binding and dimerization." Mol Cell **6**(3): 743-50.
- Schmidtchen, A., I. M. Frick, et al. (2001). "Dermatan sulphate is released by proteinases of common pathogenic bacteria and inactivates antibacterial alpha-defensin." Mol Microbiol **39**(3): 708-13.
- Schonherr, E., C. Sunderkotter, et al. (2005). "Decorin, a novel player in the insulin-like growth factor system." J Biol Chem **280**(16): 15767-72.
- Schor, H., G. G. Vaday, et al. (2000). "Modulation of leukocyte behavior by an inflamed extracellular matrix." Dev Immunol **7**(2-4): 227-38.
- Schroder, K., P. J. Hertzog, et al. (2004). "Interferon-gamma: an overview of signals, mechanisms and functions." J Leukoc Biol **75**(2): 163-89.
- Schwartzmann, G., E. Sprinz, et al. (1996). "Phase II study of pentosan polysulfate (PPS) in patients with AIDS-related Kaposi's sarcoma." Tumori **82**(4): 360-3.
- Schwarz, K., B. Breuer, et al. (1990). "Biosynthesis and properties of a further member of the small chondroitin/dermatan sulfate proteoglycan family." J Biol Chem **265**(35): 22023-8.
- Scott, P. G., C. M. Dodd, et al. (2006). "Crystal structure of the biglycan dimer and evidence that dimerization is essential for folding and stability of class I small leucine-rich repeat proteoglycans." J Biol Chem **281**(19): 13324-32.
- Seftor, E. A., P. S. Meltzer, et al. (2002). "Expression of multiple molecular phenotypes by aggressive melanoma tumor cells: role in vasculogenic mimicry." Crit Rev Oncol Hematol **44**(1): 17-27.
- Seidel, C., M. Borset, et al. (2000). "High levels of soluble syndecan-1 in myeloma-derived bone marrow: modulation of hepatocyte growth factor activity." Blood **96**(9): 3139-46.
- Seidler, D. G., J. Peter-Katalinic, et al. (2007). "Galactosaminoglycan function and oligosaccharide structure determination." ScientificWorldJournal **7**: 233-41.
- Sekiguchi, K., S. Hakomori, et al. (1983). "Binding of fibronectin and its proteolytic fragments to glycosaminoglycans. Exposure of cryptic glycosaminoglycan-binding domains upon limited proteolysis." J Biol Chem **258**(23): 14359-65.
- Seo, N. S., A. M. Hocking, et al. (2005). "Decorin core protein secretion is regulated by N-linked oligosaccharide and glycosaminoglycan additions." J Biol Chem **280**(52): 42774-84.
- Sergeant, N., M. Lyon, et al. (2000). "Stimulation of DNA synthesis and cell proliferation of human mammary myoepithelial-like cells by hepatocyte growth factor/scatter factor depends on heparan sulfate proteoglycans and sustained phosphorylation of mitogen-activated protein kinases p42/44." J Biol Chem **275**(22): 17094-9.
- Shao, C., F. Zhang, et al. (2006). "Crystallographic analysis of calcium-dependent heparin binding to annexin A2." J Biol Chem **281**(42): 31689-95.
- Sharma, B., M. Handler, et al. (1998). "Antisense targeting of perlecan blocks tumor growth and angiogenesis *in vivo*." J Clin Invest **102**(8): 1599-608.

- Shaw, G. and R. Kamen (1986). "A conserved AU sequence from the 3' untranslated region of GM-CSF mRNA mediates selective mRNA degradation." *Cell* **46**(5): 659-67.
- Shaw, J. P., Z. Johnson, et al. (2004). "The X-ray structure of RANTES: heparin-derived disaccharides allows the rational design of chemokine inhibitors." *Structure* **12**(11): 2081-93.
- Shaya, D., A. Tocilj, et al. (2006). "Crystal structure of heparinase II from *Pedobacter heparinus* and its complex with a disaccharide product." *J Biol Chem* **281**(22): 15525-35.
- Siegel, J. P. (1988). "Effects of interferon-gamma on the activation of human T lymphocytes." *Cell Immunol* **111**(2): 461-72.
- Simons, K. and D. Toomre (2000). "Lipid rafts and signal transduction." *Nat Rev Mol Cell Biol* **1**(1): 31-9.
- Skandalis, S. S., D. Kletsas, et al. (2006). "The greatly increased amounts of accumulated versican and decorin with specific post-translational modifications may be closely associated with the malignant phenotype of pancreatic cancer." *Biochim Biophys Acta* **1760**(8): 1217-25.
- Skinner, M. P., C. M. Lucas, et al. (1991). "GMP-140 binding to neutrophils is inhibited by sulfated glycans." *J Biol Chem* **266**(9): 5371-4.
- Skurkovich, B. and S. Skurkovich (2003). "Anti-interferon-gamma antibodies in the treatment of autoimmune diseases." *Curr Opin Mol Ther* **5**(1): 52-7.
- Smorenburg, S. M., B. A. Hutten, et al. (1999). "Should patients with venous thromboembolism and cancer be treated differently?" *Haemostasis* **29** Suppl S1: 91-7.
- Smorenburg, S. M. and C. J. Van Noorden (2001). "The complex effects of heparins on cancer progression and metastasis in experimental studies." *Pharmacol Rev* **53**(1): 93-105.
- Sperinde, G. V. and M. A. Nugent (1998). "Heparan sulfate proteoglycans control intracellular processing of bFGF in vascular smooth muscle cells." *Biochemistry* **37**(38): 13153-64.
- Spillmann, D. (2001). "Heparan sulfate: anchor for viral intruders?" *Biochimie* **83**(8): 811-7.
- St Onge, R., C. Weiss, et al. (1980). "A preliminary assessment of Na-hyaluronate injection into "no man's land" for primary flexor tendon repair." *Clin Orthop Relat Res*(146): 269-75.
- Stanton, G. J., C. Jordan, et al. (1987). "Nondetectable levels of interferon gamma is a critical host defense during the first day of herpes simplex virus infection." *Microb Pathog* **3**(3): 179-83.
- Stark, G. R., I. M. Kerr, et al. (1998). "How cells respond to interferons." *Annu Rev Biochem* **67**: 227-64.
- Steeber, D. A., G. M. Venturi, et al. (2005). "A new twist to the leukocyte adhesion cascade: intimate cooperation is key." *Trends Immunol* **26**(1): 9-12.
- Stringer, S. E. (2006). "The role of heparan sulphate proteoglycans in angiogenesis." *Biochem Soc Trans* **34**(Pt 3): 451-3.
- SundarRaj, N., D. Fite, et al. (1995). "Perlécan is a component of cartilage matrix and promotes chondrocyte attachment." *J Cell Sci* **108** ( Pt 7): 2663-72.
- Svensson, L., A. Aszodi, et al. (1999). "Fibromodulin-null mice have abnormal collagen fibrils, tissue organization, and altered lumican deposition in tendon." *J Biol Chem* **274**(14): 9636-47.
- Svensson, L., I. Narlid, et al. (2000). "Fibromodulin and lumican bind to the same region on collagen type I fibrils." *FEBS Lett* **470**(2): 178-82.
- Swain, S. M., B. Parker, et al. (1995). "Phase I trial of pentosan polysulfate." *Invest New Drugs* **13**(1): 55-62.

- Swaminathan, G. J., D. G. Myszka, et al. (2005). "Eosinophil-granule major basic protein, a C-type lectin, binds heparin." *Biochemistry* **44**(43): 14152-8.
- Sweeney, E. A., H. Lortat-Jacob, et al. (2002). "Sulfated polysaccharides increase plasma levels of SDF-1 in monkeys and mice: involvement in mobilization of stem/progenitor cells." *Blood* **99**(1): 44-51.
- Tan, K., M. Duquette, et al. (2006). "The structures of the thrombospondin-1 N-terminal domain and its complex with a synthetic pentameric heparin." *Structure* **14**(1): 33-42.
- Tanaka, Y., D. H. Adams, et al. (1993). "T-cell adhesion induced by proteoglycan-immobilized cytokine MIP-1 beta." *Nature* **361**(6407): 79-82.
- Tangelder, G. J. and K. E. Arfors (1991). "Inhibition of leukocyte rolling in venules by protamine and sulfated polysaccharides." *Blood* **77**(7): 1565-71.
- Taylor, K. R. and R. L. Gallo (2006). "Glycosaminoglycans and their proteoglycans: host-associated molecular patterns for initiation and modulation of inflammation." *Faseb J* **20**(1): 9-22.
- Taylor, K. R., J. A. Rudisill, et al. (2005). "Structural and sequence motifs in dermatan sulfate for promoting fibroblast growth factor-2 (FGF-2) and FGF-7 activity." *J Biol Chem* **280**(7): 5300-6.
- Taylor, K. R., J. A. Rudisill, et al. (2005). "Structural and sequence motifs in dermatan sulfate for promoting fibroblast growth factor-2 (FGF-2) and FGF-7 activity." *J Biol Chem* **280**(7): 5300-6 Epub 2004 Nov 24.
- Teixeira, L. K., B. P. Fonseca, et al. (2005). "The role of interferon-gamma on immune and allergic responses." *Mem Inst Oswaldo Cruz* **100 Suppl 1**: 137-44.
- Theocharis, A. D. (2002). "Human colon adenocarcinoma is associated with specific post-translational modifications of versican and decorin." *Biochim Biophys Acta* **1588**(2): 165-72.
- Theocharis, A. D., M. E. Tsara, et al. (2000). "Pancreatic carcinoma is characterized by elevated content of hyaluronan and chondroitin sulfate with altered disaccharide composition." *Biochim Biophys Acta* **1502**(2): 201-6.
- Theocharis, A. D., D. H. Vynios, et al. (2003). "Altered content composition and structure of glycosaminoglycans and proteoglycans in gastric carcinoma." *Int J Biochem Cell Biol* **35**(3): 376-90.
- Theroux, P., D. Waters, et al. (1993). "Aspirin versus heparin to prevent myocardial infarction during the acute phase of unstable angina." *Circulation* **88**(5 Pt 1): 2045-8.
- Thorne, R. F., J. W. Legg, et al. (2004). "The role of the CD44 transmembrane and cytoplasmic domains in co-ordinating adhesive and signalling events." *J Cell Sci* **117**(Pt 3): 373-80.
- Thunberg, L., G. Backstrom, et al. (1980). "The molecular size of the antithrombin-binding sequence in heparin." *FEBS Lett* **117**(1): 203-6.
- Thunberg, L., G. Backstrom, et al. (1982). "Further characterization of the antithrombin-binding sequence in heparin." *Carbohydr Res* **100**: 393-410.
- Tiedemann, K., B. Olander, et al. (2005). "Regulation of the chondroitin/dermatan fine structure by transforming growth factor-beta1 through effects on polymer-modifying enzymes." *Glycobiology* **15**(12): 1277-85.
- Timar, J., K. Lapis, et al. (2002). "Proteoglycans and tumor progression: Janus-faced molecules with contradictory functions in cancer." *Semin Cancer Biol* **12**(3): 173-86.
- Tissier, S., S. Lancel, et al. (2004). "Calpain inhibitors improve myocardial dysfunction and inflammation induced by endotoxin in rats." *Shock* **21**(4): 352-7.
- Toole, B. P. (1990). "Hyaluronan and its binding proteins, the hyaladherins." *Curr Opin Cell Biol* **2**(5): 839-44.

- Toole, B. P. (2004). "Hyaluronan: from extracellular glue to pericellular cue." Nat Rev Cancer **4**(7): 528-39.
- Trent, J. M., S. Olson, et al. (1982). "Chromosomal localization of human leukocyte, fibroblast, and immune interferon genes by means of in situ hybridization." Proc Natl Acad Sci U S A **79**(24): 7809-13.
- Trevani, A. S., G. Andonegui, et al. (1999). "Extracellular acidification induces human neutrophil activation." J Immunol **162**(8): 4849-57.
- Trinchieri, G. (1995). "Interleukin-12: a proinflammatory cytokine with immunoregulatory functions that bridge innate resistance and antigen-specific adaptive immunity." Annu Rev Immunol **13**: 251-76.
- Trinchieri, G., and B. Perussia. (1985). "Immune interferon: a pleiotropic lymphokine with multiple effects" Immunol. Today **6**: 131-136
- Trowbridge, J. M. and R. L. Gallo (2002). "Dermatan sulfate: new functions from an old glycosaminoglycan." Glycobiology **12**(9): 117R-25R.
- Trowbridge, J. M., J. A. Rudisill, et al. (2002). "Dermatan sulfate binds and potentiates activity of keratinocyte growth factor (FGF-7)." J Biol Chem **277**(45): 42815-20.
- Tsai, J. C., J. Zhang, et al. (2002). "Cloning and characterization of the human lung endothelial-cell-specific molecule-1 promoter." J Vasc Res **39**(2): 148-59.
- Tsara, M. E., N. Papageorgacopoulou, et al. (1995). "Distribution and changes of glycosaminoglycans in neoplasias of rectum." Anticancer Res **15**(5B): 2107-12.
- Tsara, M. E., A. D. Theocharis, et al. (2002). "Compositional and structural alterations of proteoglycans in human rectum carcinoma with special reference to versican and decorin." Anticancer Res **22**(5): 2893-8.
- Tschernatsch, M. M., B. Mlecnik, et al. (2006). "LPL-mediated lipolysis of VLDL induces an upregulation of AU-rich mRNAs and an activation of HuR in endothelial cells." Atherosclerosis **189**(2): 310-7.
- Tsen, G., W. Halfter, et al. (1995). "Agrin is a heparan sulfate proteoglycan." J Biol Chem **270**(7): 3392-9.
- Tuettenberg, J., C. Friedel, et al. (2006). "Angiogenesis in malignant glioma--a target for antitumor therapy?" Crit Rev Oncol Hematol. **59**(3): 181-93 Epub 2006 Jul 24.
- Tufvesson, E. and G. Westergren-Thorsson (2002). "Tumour necrosis factor-alpha interacts with biglycan and decorin." FEBS Lett **530**(1-3): 124-8.
- Tumova, S., B. A. Hatch, et al. (1999). "Basic fibroblast growth factor does not prevent heparan sulphate proteoglycan catabolism in intact cells, but it alters the distribution of the glycosaminoglycan degradation products." Biochem J **337** ( Pt 3): 471-81.
- Turnbull, J., A. Powell, et al. (2001). "Heparan sulfate: decoding a dynamic multifunctional cell regulator." Trends Cell Biol **11**(2): 75-82.
- Uchimura, K., M. Morimoto-Tomita, et al. (2006). "Measuring the activities of the Sulfs: two novel heparin/heparan sulfate endosulfatases." Methods Enzymol **416**: 243-53.
- Udabage, L., G. R. Brownlee, et al. (2005). "Antisense-mediated suppression of hyaluronan synthase 2 inhibits the tumorigenesis and progression of breast cancer." Cancer Res **65**(14): 6139-50.
- Uebelhart, D., M. Malaise, et al. (2004). "Intermittent treatment of knee osteoarthritis with oral chondroitin sulfate: a one-year, randomized, double-blind, multicenter study versus placebo." Osteoarthritis Cartilage **12**(4): 269-76.
- Uebelhart, D., E. J. Thonar, et al. (1998). "Effects of oral chondroitin sulfate on the progression of knee osteoarthritis: a pilot study." Osteoarthritis Cartilage **6 Suppl A**: 39-46.
- Ueno, M., S. Yamada, et al. (2001). "Structural characterization of heparan sulfate and chondroitin sulfate of syndecan-1 purified from normal murine mammary gland

- epithelial cells. Common phosphorylation of xylose and differential sulfation of galactose in the protein linkage region tétrasaccharide sequence." J Biol Chem **276**(31): 29134-40.
- Vaday, G. G. and O. Lider (2000). "Extracellular matrix moieties, cytokines, and enzymes: dynamic effects on immune cell behavior and inflammation." J Leukoc Biol **67**(2): 149-59.
- van 't Veer, L. J., H. Dai, et al. (2002). "Gene expression profiling predicts clinical outcome of breast cancer." Nature **415**(6871): 530-6.
- van der Voort, R., T. E. Taher, et al. (1999). "Heparan sulfate-modified CD44 promotes hepatocyte growth factor/scatter factor-induced signal transduction through the receptor tyrosine kinase c-Met." J Biol Chem **274**(10): 6499-506.
- van Horsen, J., M. M. Wilhelmus, et al. (2002). "Collagen XVIII: a novel heparan sulfate proteoglycan associated with vascular amyloid depositions and senile plaques in Alzheimer's disease brains." Brain Pathol **12**(4): 456-62.
- Vancheri, C., C. Mastruzzo, et al. (2001). "Intranasal heparin reduces eosinophil recruitment after nasal allergen challenge in patients with allergic rhinitis." J Allergy Clin Immunol **108**(5): 703-8.
- Vanhaverbeke, C., J. P. Simorre, et al. (2004). "NMR characterization of the interaction between the C-terminal domain of interferon-gamma and heparin-derived oligosaccharides." Biochem J **384**(Pt 1): 93-9.
- Vanpouille, C., A. Denys, et al. (2004). "Octasaccharide is the minimal length unit required for efficient binding of cyclophilin B to heparin and cell surface heparan sulphate." Biochem J **382**(Pt 2): 733-40.
- Veldkamp, C. T., F. C. Peterson, et al. (2005). "The monomer-dimer equilibrium of stromal cell-derived factor-1 (CXCL 12) is altered by pH, phosphate, sulfate, and heparin." Protein Sci **14**(4): 1071-81.
- Venkataraman, G., V. Sasisekharan, et al. (1996). "Preferential self-association of basic fibroblast growth factor is stabilized by heparin during receptor dimerization and activation." Proc Natl Acad Sci U S A **93**(2): 845-50.
- Venkatesan, N., T. Ebihara, et al. (2000). "Alterations in large and small proteoglycans in bleomycin-induced pulmonary fibrosis in rats." Am J Respir Crit Care Med **161**(6): 2066-73.
- Verbruggen, G., S. Goemaere, et al. (1998). "Chondroitin sulfate: S/DMOAD (structure/disease modifying anti-osteoarthritis drug) in the treatment of finger joint OA." Osteoarthritis Cartilage **6 Suppl A**: 37-8.
- Veugelers, M., B. De Cat, et al. (1999). "Glypican-6, a new member of the glypican family of cell surface heparan sulfate proteoglycans." J Biol Chem **274**(38): 26968-77.
- Vicenzi, E., A. Gatti, et al. (2003). "Broad spectrum inhibition of HIV-1 infection by sulfated K5 Escherichia coli polysaccharide derivatives." Aids **17**(2): 177-81.
- Viola, M., E. G. Karousou, et al. (2006). "Decorin from different bovine tissues: study of glycosaminoglycan chain by PAGEFS." J Pharm Biomed Anal **41**(1): 36-42.
- Vives, R. R., A. Imberty, et al. (2005). "Heparan sulfate targets the HIV-1 envelope glycoprotein gp120 coreceptor binding site." J Biol Chem **280**(22): 21353-7.
- Vives, R. R., R. Sadir, et al. (2002). "A kinetics and modeling study of RANTES(9-68) binding to heparin reveals a mechanism of cooperative oligomerization." Biochemistry **41**(50): 14779-89.
- Vlodavsky, I., G. Abboud-Jarrous, et al. (2006). "The impact of heparanase and heparin on cancer metastasis and angiogenesis." Pathophysiol Haemost Thromb **35**(1-2): 116-27.
- Vlodavsky, I., Y. Friedmann, et al. (1999). "Mammalian heparanase: gene cloning, expression and function in tumor progression and metastasis." Nat Med **5**(7): 793-802.

- Vlodavsky, I., O. Goldshmidt, et al. (2002). "Mammalian heparanase: involvement in cancer metastasis, angiogenesis and normal development." Semin Cancer Biol **12**(2): 121-9.
- Vogt, A. M., F. Pettersson, et al. (2006). "Release of Sequestered Malaria Parasites upon Injection of a Glycosaminoglycan." PLoS Pathog **2**(9).
- Wadhwa, S., M. Embree, et al. (2005). "Mice deficient in biglycan and fibromodulin as a model for temporomandibular joint osteoarthritis." Cells Tissues Organs **181**(3-4): 136-43.
- Wadhwa, S., M. C. Embree, et al. (2005). "Accelerated osteoarthritis in the temporomandibular joint of biglycan/fibromodulin double-deficient mice." Osteoarthritis Cartilage **13**(9): 817-27.
- Wakiya, K., A. Begue, et al. (1996). "A cAMP response element and an Ets motif are involved in the transcriptional regulation of flt-1 tyrosine kinase (vascular endothelial growth factor receptor 1) gene." J Biol Chem **271**(48): 30823-8.
- Walter, M. R., W. T. Windsor, et al. (1995). "Crystal structure of a complex between interferon-gamma and its soluble high-affinity receptor." Nature **376**(6537): 230-5.
- Wan, M. X., X. W. Zhang, et al. (2001). "Low molecular weight heparin inhibits tumor necrosis factor alpha-induced leukocyte rolling." Inflamm Res **50**(12): 581-4.
- Wang, L., M. Fuster, et al. (2005). "Endothelial heparan sulfate deficiency impairs L-selectin- and chemokine-mediated neutrophil trafficking during inflammatory responses." Nat Immunol **6**(9): 902-10.
- Wang, X., M. J. Brownstein, et al. (1997). "PG25, a pineal-specific cDNA, cloned by differential display PCR (DDPCR) and rapid amplification of cDNA ends (RACE)." J Neurosci Methods **73**(2): 187-91.
- Wang, X. F., H. Y. Lin, et al. (1991). "Expression cloning and characterization of the TGF-beta type III receptor." Cell **67**(4): 797-805.
- Wang, Z., M. Gotte, et al. (2005). "Constitutive and accelerated shedding of murine syndecan-1 is mediated by cleavage of its core protein at a specific juxtamembrane site." Biochemistry **44**(37): 12355-61.
- Warkentin, T. E., M. N. Levine, et al. (1995). "Heparin-induced thrombocytopenia in patients treated with low-molecular-weight heparin or unfractionated heparin." N Engl J Med **332**(20): 1330-5.
- Watanabe, Y., K. Kuribayashi, et al. (1989). "Exogenous expression of mouse interferon gamma cDNA in mouse neuroblastoma C1300 cells results in reduced tumorigenicity by augmented anti-tumor immunity." Proc Natl Acad Sci U S A **86**(23): 9456-60.
- Watt, S. M., J. Williamson, et al. (1993). "The heparin binding PECAM-1 adhesion molecule is expressed by CD34+ hematopoietic precursor cells with early myeloid and B-lymphoid cell phenotypes." Blood **82**(9): 2649-63.
- Waugh, D. F. and M. A. Fitzgerald (1956). "Quantitative aspects of antithrombin and heparin in plasma." Am J Physiol **184**(3): 627-39.
- Wei, Z., M. Lyon, et al. (2005). "Distinct substrate specificities of bacterial heparinases against N-unsubstituted glucosamine residues in heparan sulfate." J Biol Chem **280**(16): 15742-8.
- Weigel, P. H., V. C. Hascall, et al. (1997). "Hyaluronan synthases." J Biol Chem **272**(22): 13997-4000.
- Wellner, M., F. Herse, et al. (2003). "Endothelial cell specific molecule-1--a newly identified protein in adipocytes." Horm Metab Res **35**(4): 217-21.
- Wettreich, A., A. Sebollela, et al. (1999). "Acidic pH modulates the interaction between human granulocyte-macrophage colony-stimulating factor and glycosaminoglycans." J Biol Chem **274**(44): 31468-75.
- Wheelock E.F. (1965). 'Interferon-like virus-inhibitor induced in human leucocytes by

- phytohemagglutinin." Science 149: 310-311
- Whitfield, D. M. and B. Sarkar (1991). "Metal binding to heparin monosaccharides: D-glucosamine-6-sulphate, D-glucuronic acid, and L-iduronic acid." J Inorg Biochem 41(3): 157-70.
- Whitfield, D. M. and B. Sarkar (1992). "Heavy metal binding to heparin disaccharides. II. First evidence for zinc chelation." Biopolymers 32(6): 597-619.
- Wiberg, C., D. Heinegard, et al. (2002). "Biglycan organizes collagen VI into hexagonal-like networks resembling tissue structures." J Biol Chem 277(51): 49120-6.
- Wiberg, C., A. R. Klatt, et al. (2003). "Complexes of matrilin-1 and biglycan or decorin connect collagen VI microfibrils to both collagen II and aggrecan." J Biol Chem 278(39): 37698-704.
- Willett, C. G., Y. Boucher, et al. (2004). "Direct evidence that the VEGF-specific antibody bevacizumab has antivasular effects in human rectal cancer." Nat Med 10(2): 145-7.
- Williams, J. G., G. J. Jurkovich, et al. (1993). "Interferon-gamma: a key immunoregulatory lymphokine." J Surg Res 54(1): 79-93.
- Williams, R. O., D. G. Williams, et al. (1993). "Increased limb involvement in murine collagen-induced arthritis following treatment with anti-interferon-gamma." Clin Exp Immunol 92(2): 323-7.
- Wimmenauer, S., A. Steiert, et al. (1999). "Influence of cytokines on the expression of fas ligand and CD44 splice variants in colon carcinoma cells." Tumour Biol 20(6): 294-303.
- Witt, D. P. and A. D. Lander (1994). "Differential binding of chemokines to glycosaminoglycan subpopulations." Curr Biol 4(5): 394-400.
- Woods, A. (2001). "Syndecans: transmembrane modulators of adhesion and matrix assembly." J Clin Invest 107(8): 935-41.
- Woods, A. and J. R. Couchman (2001). "Syndecan-4 and focal adhesion function." Curr Opin Cell Biol 13(5): 578-83.
- Wu, Y. J., D. P. La Pierre, et al. (2005). "The interaction of versican with its binding partners." Cell Res 15(7): 483-94.
- Wu, Z. L., L. Zhang, et al. (2002). "A new strategy for defining critical functional groups on heparan sulfate." Faseb J 16(6): 539-45.
- Xie, X., H. Thorlacius, et al. (1997). "Inhibitory effect of locally administered heparin on leukocyte rolling and chemoattractant-induced firm adhesion in rat mesenteric venules *in vivo*." Br J Pharmacol 122(5): 906-10.
- Xu, Y., P. D. Carr, et al. (1998). "The crystal structure of bikunin from the inter-alpha-inhibitor complex: a serine protease inhibitor with two Kunitz domains." J Mol Biol 276(5): 955-66.
- Yabe, T., D. Shukla, et al. (2001). "Portable sulphotransferase domain determines sequence specificity of heparan sulphate 3-O-sulphotransferases." Biochem J 359(Pt 1): 235-41.
- Yamaguchi, K., H. Tamaki, et al. (2006). "Detection of oligosaccharide ligands for hepatocyte growth factor/scatter factor (HGF/SF), keratinocyte growth factor (KGF/FGF-7), RANTES and heparin cofactor II by neoglycolipid microarrays of glycosaminoglycan-derived oligosaccharide fragments." Glycoconj J 23(7-8): 513-23.
- Yamaguchi, Y. (2000). "Lecticans: organizers of the brain extracellular matrix." Cell Mol Life Sci 57(2): 276-89.
- Yamashita, T. (2005). "Structural and biochemical properties of fibroblast growth factor 23." Ther Apher Dial 9(4): 313-8.
- Yamauchi, H., P. Desgranges, et al. (2000). "New agents for the treatment of infarcted myocardium." Faseb J 14(14): 2133-4.

- Yang, J., M. A. Price, et al. (2004). "Melanoma chondroitin sulfate proteoglycan enhances FAK and ERK activation by distinct mechanisms." *J Cell Biol* **165**(6): 881-91.
- Yang, Y., V. Macleod, et al. (2005). "Heparanase promotes the spontaneous metastasis of myeloma cells to bone." *Blood* **105**(3): 1303-9.
- Yang, Y., V. Macleod, et al. (2007). "Heparanase Enhances Syndecan-1 Shedding: A NOVEL MECHANISM FOR STIMULATION OF TUMOR GROWTH AND METASTASIS." *J Biol Chem* **282**(18): 13326-33.
- Yang, Y. G., H. Wang, et al. (2005). "Role of Interferon-gamma in GVHD and GVL." *Cell Mol Immunol* **2**(5): 323-9.
- Yayon, A., M. Klagsbrun, et al. (1991). "Cell surface, heparin-like molecules are required for binding of basic fibroblast growth factor to its high affinity receptor." *Cell* **64**(4): 841-8.
- Yeaman, G. R., J. E. Collins, et al. (1998). "IFN-gamma is produced by polymorphonuclear neutrophils in human uterine endometrium and by cultured peripheral blood polymorphonuclear neutrophils." *J Immunol* **160**(10): 5145-53.
- Yip, G. W., M. Smollich, et al. (2006). "Therapeutic value of glycosaminoglycans in cancer." *Mol Cancer Ther* **5**(9): 2139-48.
- Yip, Y. K., B. S. Barrowclough, et al. (1982). "Purification of two subspecies of human gamma (immune) interferon." *Proc Natl Acad Sci U S A* **79**(6): 1820-4.
- Yoneda, A. and J. R. Couchman (2003). "Regulation of cytoskeletal organization by syndecan transmembrane proteoglycans." *Matrix Biol* **22**(1): 25-33.
- Young, H. A. and K. J. Hardy (1995). "Role of interferon-gamma in immune cell regulation." *J Leukoc Biol* **58**(4): 373-81.
- Youngner, J. S. and S. B. Salvin (1973). "Production and properties of migration inhibitory factor and interferon in the circulation of mice with delayed hypersensitivity." *J Immunol* **111**(6): 1914-22.
- Yurt, R. W., R. W. Leid, Jr., et al. (1977). "Native heparin from rat peritoneal mast cells." *J Biol Chem* **252**(2): 518-21.
- Zacharski, L. R. and J. T. Loynes (2002). "The heparins and cancer." *Curr Opin Pulm Med* **8**(5): 379-82.
- Zacharski, L. R. and J. T. Loynes (2003). "Low-molecular-weight heparin in oncology." *Anticancer Res* **23**(3C): 2789-93.
- Zacharski, L. R. and D. L. Ornstein (1998). "Heparin and cancer." *Thromb Haemost* **80**(1): 10-23.
- Zacharski, L. R., D. L. Ornstein, et al. (2000). "Low-molecular-weight heparin and cancer." *Semin Thromb Hemost* **26 Suppl 1**: 69-77.
- Zamfir, A., D. G. Seidler, et al. (2003). "Structural investigation of chondroitin/dermatan sulfate oligosaccharides from human skin fibroblast decorin." *Glycobiology* **13**(11): 733-42.
- Zanetta, J. P., P. Timmerman, et al. (1999). "Determination of constituents of sulphated proteoglycans using a methanolysis procedure and gas chromatography/mass spectrometry of heptafluorobutyrate derivatives." *Glycoconj J* **16**(10): 617-27.
- Zetser, A., Y. Bashenko, et al. (2006). "Heparanase induces vascular endothelial growth factor expression: correlation with p38 phosphorylation levels and Src activation." *Cancer Res* **66**(3): 1455-63.
- Zhao, H., H. Liu, et al. (2006). "Oligomannurinate Sulfate, a Novel Heparanase Inhibitor Simultaneously Targeting Basic Fibroblast Growth Factor, Combats Tumor Angiogenesis and Metastasis." *Cancer Res* **66**(17): 8779-8787.

- Zhao, X., K. E. Ramsey, et al. (2004). "Gene and protein expression changes in human trabecular meshwork cells treated with transforming growth factor-beta." Invest Ophthalmol Vis Sci **45**(11): 4023-34.
- Zhou, Z., J. Wang, et al. (2004). "Impaired angiogenesis, delayed wound healing and retarded tumor growth in perlecan heparan sulfate-deficient mice." Cancer Res **64**(14): 4699-702.
- Zhuo, L., V. C. Hascall, et al. (2004). "Inter-alpha-trypsin inhibitor, a covalent protein-glycosaminoglycan-protein complex." J Biol Chem **279**(37): 38079-82.
- Zhuo, L., A. Salustri, et al. (2002). "A physiological function of serum proteoglycan bikunin: the chondroitin sulfate moiety plays a central role." Glycoconj J **19**(4-5): 241-7.

# **ANNEXES**

## PUBLICATIONS

### Revues :

Delehedde M, **Sarrazin S**, Adam E, Motte V and Vanpouille C (2006) Proteoglycans and glycosaminoglycans : Complex molecules with modulating activity. In *"New developments in Therapeutic Glycomics"*. Eds. Maryse Delehedde and Hugues Lortat-Jacob. pp 1-13.

Lassalle P, Adam E, **Sarrazin S**, Depontieu F, Motte V, Lequien S, Grigoriu B, De Freitas-Carreras N, Vanpouille C and Delehedde M (2006) Proteoglycans and glycosaminoglycans : key players in inflammation. In *"New developments in Therapeutic Glycomics"*. Eds. Maryse Delehedde and Hugues Lortat-Jacob. pp 15-31.

Adam E, Motte V, **Sarrazin S** and Delehedde M (2006) Stromal Derived Growth Factor-1 : a target for GAG-based therapies? In *"New developments in Therapeutic Glycomics"*. Eds. Maryse Delehedde and Hugues Lortat-Jacob. pp 283-302.

**Sarrazin S**, Adam E, Lyon M, Depontieu F, Motte V, Landolfi C, Lortat-Jacob H, Béchard D, Lassalle P and Delehedde M. (2006) Endocan or endothelial cell specific molecule-1 (ESM-1): A potential novel endothelial cell marker and a new target for cancer therapy. *BBA, Review in cancer, vol 1765, pp 25-37 (2006)*

### Articles:

**Sarrazin S**, Bonnaffé D, Lubineau A, Lortat-Jacob H. Heparan sulfate mimicry: A synthetic glycoconjugate that recognizes the heparin-binding domain of IFN $\gamma$  inhibits the cytokine activity. *J Biol Chem. 2005 Nov 11;280(45):37558-64.*

Lubineau A.; Lortat-Jacob H.; Gavard O.; **Sarrazin S.**; Bonnaffé D (2004) Synthesis of Heparan Sulfate neoglycoconjugate mimetics that bind IFN-g in the nanomolar range. *Chemistry.2004;10(17):4265-82*

### Poster publié:

**Sarrazin S.**; Béchard D.; Lassalle P.; Lyon M.; Scherpereel A.; Lortat-Jacob H.; Ouatas T.; Delehedde M (2004) Purification and characterization of endocan (endothelial cell-specific molecule-1), a circulating proteoglycan involved in tumor progression and inflammatory diseases. *International Journal of Experimental Pathology. Vol. 85 no. 4, ppA73.*

## TRAVAUX PRESENTES EN MEETINGS (Posters)

- 13-14 Octobre 2006 TOURS (France). *J2R (Journée de la recherche respiratoire)*:  
"Rôle des sérine-protéases neutrophiles dans le catabolisme d'endocan."  
*De Freitas Caires N., Sarrazin S., Depontieu F., Lortat-Jacob H., Gosset P., Lassalle P.*
- 13-16 Mai 2006 CANNES (France). *6<sup>th</sup> Interdisciplinary Euroconference on Angiogenesis* :  
"Endocan previously named endothelial cell specific molecule-1 is overexpressed in human glioblastoma cells"  
*E. Adam, J.F. Minéo, S. Sarrazin, M. Debunne, C. Villedieu, V. Motte, R.M. Siminski, ML Caillet, P. Lassalle, C.A. Maurage, S. Blond and M. Delehedde.*
- 14 Octobre 2005, REIMS (France). *J2R (Journée de la recherche respiratoire)*:  
"Etude du catabolisme d'endocan : rôle des protéases du polynucléaire neutrophile."  
*De Freitas Caires N. Sarrazin S. Morelle W. Depontieu F. Hauw P. Michalski J-C. Tonnel A-B. Lortat-Jacob H. Lassalle P.*
- 7-11 Septembre 2005 RUNO (Suède). *Proteoglycans in signalling*:  
"“2O<sub>10</sub>” A synthetic glycoconjugate that mimics the IFN $\gamma$ -binding domain in heparin sulphate inhibits the cytokine activity."  
*Lortat-Jacob H. Sarrazin S. Bonnaffé D.*
- 27-30 Mai 2005 SITGES (Espagne). *5<sup>th</sup> Interdisciplinary Euroconference on Angiogenesis*:  
"Structural characterization of a circulating dermatan sulphate proteoglycan called endocan involved in angiogenic growth factor bioactivities."  
*Sarrazin S. Motte V. Adam E. Lassalle P. Lyon M. Béchard D. Delehedde M. Lortat-Jacob H.*  
"Endocan or endothelial cell specific molecule-1 (ESM-1) : a potential new target for cancer therapy."  
*Adam E. Delehedde M. Sarrazin S. Landolfi C. Motte V. Depontieu F. Béchard D. Lassalle P.*
- 22-25 Octobre 2004 CASCAIS (Portugal). *5<sup>th</sup> Interdisciplinary Conference Mechanisms of Cell Death and Disease: Advances in Therapeutic Intervention and Drug Development*:  
"A sandwich ELISA method to determine the expression of human endocan or endothelial cell specific molecule-1 in biological samples."  
*Delehedde M., Adam E., Sarrazin S., Landolfi C., Motte V., Lassalle P., Ouatas T., Béchard D.*
- 26 Mai 2004 LILLE (France). *Journée lilloise de cancérologie*:  
"Involvement of endocan, a circulating proteoglycan carrying a single dermatan sulphate in tumor progression and inflammatory diseases."  
*Sarrazin S.; Béchard D; Lortat-Jacob H.; Depontieu F.; Grigoru B.; Lyon M.; Scherpereel A.; Lassalle P.; Ouatas T.; and Delehedde M.*
- 21-24 Mai 2004 HELSINKI (Finland). *4<sup>th</sup> Interdisciplinary Euroconference on Angiogenesis*:  
"Endocan, a circulating proteoglycan carrying a single dermatan sulphate involved in angiogenesis and tumor progression"  
*Delehedde M.; Sarrazin S.; Lortat-Jacob H.; Adam E.; Lyon M.; Scherpereel A.; Lassalle P.; Ouatas T.; Béchard D.*

5-6 Avril 2004 MANCHESTER (UK). BSMB Meeting : "Grapling the glycome":  
"Purification and characterization of endocan , a circurlating proteglycan involved in  
tumor progression and inflammatory diseases."  
***Sarrazin S.***; *Bécharde D.*; *Lassalle P.*; *Lyon M.*; *Scherpereel A.*; *Lortat-Jacob H.*;  
*Ouatas T.*; *Delehedde M.*

## Synthesis of Tailor-Made Glycoconjugate Mimetics of Heparan Sulfate That Bind IFN- $\gamma$ in the Nanomolar Range<sup>\*,\*\*</sup>

André Lubineau,<sup>[a]</sup> Hugues Lortat-Jacob,<sup>[b]</sup> Ollivier Gavard,<sup>[a, c]</sup> Stéphane Sarrazin,<sup>[b]</sup> and David Bonnaffé<sup>\*,[a]</sup>

**Abstract:** We have recently described the preparation of three building blocks for the combinatorial synthesis of heparan sulfate (HS) fragments. Herein we show that one of these building blocks (disaccharide 4) allows the preparation, in high yields and with total  $\alpha$  stereoselectivity, of tetra-, hexa- and octasaccharides from the heparin (HP) regular region, by using 2+2, 2+4 and 4+4 glycosylation strategies, respectively. These oligosaccharides were processed into sulfated derivatives bearing an allyl moiety in the anomeric position. The UV-promoted conjuga-

tion of these compounds with  $\alpha,\omega$ -bis-(thio)poly(ethylene glycol) spacers of three different lengths allowed us to prepare nine benzylated glycoconjugates. After final deprotection, the glycoconjugates 1a-c, 2a-c and 3a-c were obtained and their ability to inhibit the interaction between IFN- $\gamma$  and HP was tested by using surface plasmon reso-

nance detection. Compound 3b, containing two HP octasaccharides linked by a 50-Å linker was able to inhibit the IFN- $\gamma$ /HP interaction with an IC<sub>50</sub> value of approximately 35 nM. In addition, the nine glycoconjugates were perfect tools in the study to ascertain the topology of the IFN- $\gamma$  binding site on HS. Compounds 1a-c, 2a-c and 3a-c, by mimicking the alternating sulfated and nonsulfated regions found in HS, thus comprise the first example of a library of synthetic HS mimetics giving access to the "second level of molecular diversity" found in HS.

**Keywords:** heparin · combinatorial chemistry · cytokines · glycosylation · oligosaccharides

### Introduction

Heparan sulfate (HS) is a member of the glycosaminoglycan (GAG) family and close in structure to heparin (HP), which is clinically used for its antithrombotic activity. It is a linear sulfated polysaccharide whose basic disaccharide unit is composed of a uronic acid 1,4-linked to a 2-deoxy-2-amino-glucose. HS chains, either at the cell surface or in the extra-

cellular matrix, interact and regulate the activity of numerous proteins, such as growth factors, cytokines, chemokines, viral proteins or coagulation factors.<sup>[1,2]</sup> HS is one of the most heterogeneous biopolymers since the uronic acid may have either the D-gluco or L-ido configuration and various sulfation patterns (sulfoforms) may occur along the chain.<sup>[3-5]</sup> There is growing evidence that the formation of different HS structures is tightly controlled during biosynthesis, with the presumed goal of generating sequences with biological specificity.<sup>[6,7]</sup> Indeed, the anti-factor-Xa property of heparin is linked to the presence of a pentasaccharide in the HP chain with a precise uronic acid content and sulfation patterns that are able to bind with high affinity and specificity to antithrombin III (AT III).<sup>[8,9]</sup> The chemical synthesis of this pentasaccharide has led to the development of Arixtra, a Food and Drug Administration (FDA) and European Medicines Agency (EMA) approved drug against deep-vein thrombosis.<sup>[10]</sup>

Impressive improvements in the synthesis of HS fragments have been made in the last decade and have allowed HS fragments with a large structural variety to be synthesised.<sup>[10-20]</sup> In this regard, we have shown that combinatorial synthesis is an ideal tool to generate all the sulfoforms of the basic disaccharide of chondroitin sulfate (another GAG).<sup>[21]</sup> More recently we have published the synthesis of

[a] Prof. A. Lubineau, Dr. O. Gavard, Prof. D. Bonnaffé  
Laboratoire de Chimie Organique Multifonctionnelle  
UMR CNRS-UPS 8614 "Glycochimie Moléculaire"  
Bat. 420, Université Paris Sud, 91405 Orsay Cedex (France)  
Fax: (+33)169-154-715  
E-mail: david.bonnaffé@icmo.u-psud.fr

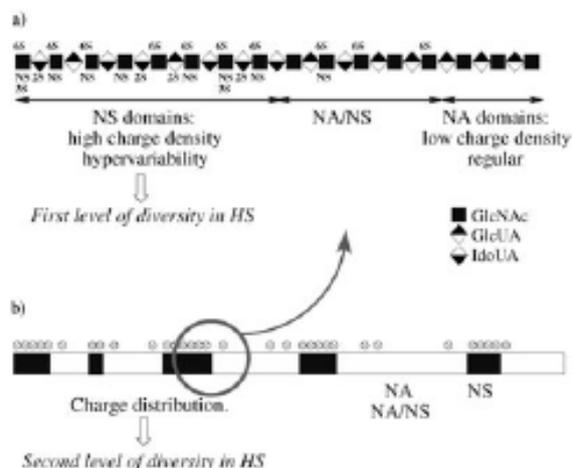
[b] Dr. H. Lortat-Jacob, S. Sarrazin  
Laboratoire d'Enzymologie Moléculaire  
Institut de Biologie Structurale, UMR CNRS-CEA-UJF 5075  
41 rue Horowitz, 38027 Grenoble Cedex 01 (France)

[c] Dr. O. Gavard  
Current address:  
Cerise Chemtech, rue de Stasbourg 5, 1140 Bruxelles (Belgium)

[\*\*] IFN- $\gamma$  =  $\gamma$ -interferon.

Supporting information for this article is available on the WWW under <http://www.chemedj.org/> or from the author.

protected disaccharide building blocks, including compound 4,<sup>[22,23]</sup> which are the basis of a combinatorial approach toward the synthesis of HS fragment libraries.<sup>[24]</sup> However, tight and specific binding between HS and a given protein is not solely regulated by these microheterogeneities, which represent only the first level of molecular diversity in HS chains (Figure 1). In fact, the polymer, typically composed

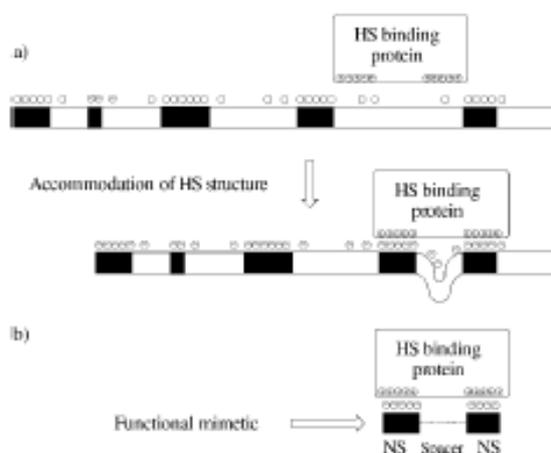


**Figure 1.** Dual molecular diversity in HS chains. a) Microheterogeneities resulting from the various uronic acid and sulfation motifs generate a first level of molecular diversity. b) The charge distribution due to the alternation of NS and NA–NA/NS domains of variable length leads to a second level of molecular diversity.

of 50–200 disaccharide units (25–100 kD), is not fully heterogeneous. Quite regular *N*-acetylated regions (NA domains), mainly composed of *D*-glucuronic acid and *N*-acetylated glucosamine, and thus with low global charge, separate domains rich in *L*-iduronic acid and *N*-sulfated glucosamine (NS domains), which are hypervariable and highly charged (Figure 1a). The latter are typically three–eight disaccharides long, while the NA domains are more regular and encompass a larger area, around 15 disaccharides in length.<sup>[25]</sup> In between, mixed NA/NS regions of variable length make the transition between the NA and NS domains. Thus, in addition to the first level of molecular diversity discussed above, HS presents a second level of diversity, since the various combinations of NS and NA–NA/NS domains generate multiple charge distribution along the polymer chain (Figure 1b).

The primary interaction between HS and a protein is an attraction between the highly negatively charged NS domains and clusters of basic residues at the protein surface, mainly arginines and lysines. In some cases, for example, with AT-III,<sup>[26]</sup> fibroblast growth factors (FGFs)<sup>[1]</sup> or stromal cell-derived factor (SDF-1),<sup>[26]</sup> a single NS domain is sufficient to allow a high-affinity interaction, the specificity of which is then linked to the uronic acid and sulfation pattern of the NS domain. However, with other proteins, such as  $\gamma$ -interferon (IFN- $\gamma$ ),<sup>[27]</sup> platelet factor 4 (PF-4),<sup>[28]</sup> regulated-

on-activation normally T-cell expressed and secreted proteins (RANTES) (9–68)<sup>[29]</sup> or macrophage inflammatory protein (MIP) 1 $\alpha$ ,<sup>[30]</sup> a single NS domain is not sufficient for binding and a longer fragment, including an NA domain, is needed for an efficient interaction. In fact, the proteins cited above are either multimeric in solution or multimerised upon binding to HS and the requirement for two distant NS domains reflects the fact that at least two basic domains of different subunits have to interact with the HS polymer to reach high binding constants. Recently, it has been proposed that the HS chain may adapt its conformation, and thus the distance between NS domains, in order to meet the needs of recognition of a protein (Figure 2a).<sup>[31]</sup> However, this will



**Figure 2.** Model of binding on the HS polymer for a protein needing more than a single NS domain. a) Accommodation of the HS polymer conformation to meet the needs of an HS binding protein. b) Two NS domains linked by a spacer of optimum length should be a functional mimetic of the HS binding site of the protein.

have an enthalpic and entropic cost and the binding of domains with the correct preformed geometry will thus be favoured. In this regard, HS glycoconjugates, in which a linker of an appropriate length separates two NS domains, should be a selective functional mimetic of the binding site of a given protein for HS (Figure 2b). Curiously, it is in the HP field that this approach has been tested, although HP is more rigid than HS and supports much less conformational accommodation.<sup>[31]</sup> HP mimetics containing a thrombin and an AT III binding site, linked by a carbohydrate<sup>[32–34]</sup> or a noncarbohydrate<sup>[35,36]</sup> spacer, have been prepared. These compounds were the first synthetic carbohydrates or mimetics displaying anti-factor-IIa and anti-factor-Xa properties. Moreover, compounds with the full anticoagulant properties of heparin but unable to bind PF-4 and thus devoid of one of the most severe side effects of HP were prepared by using this approach.<sup>[37]</sup> More recently, the “head-to-head” cross-linking of natural HP fragments, with an ethylenediamine linker, has allowed the preparation of mimetics able to bind RANTES with a greater affinity than a natural HP fragment with an equivalent length to the mimetic.<sup>[38]</sup>

Herein we describe a strategy that opens up easy access to libraries of HS functional mimetics by combining HS synthetic fragments and  $\alpha,\omega$ -bis(thio)poly(ethylene glycol) spacers of different lengths (Figure 3b). To our knowledge, this is the first time that a methodology has been developed to address the question of the second level of molecular diversity found in HS chains. This project was started some years ago in order to prepare glycoconjugates mimicking HS and able to bind IFN- $\gamma$  with high affinity and specificity.<sup>[27]</sup> IFN- $\gamma$  is a Th-1 cytokine mainly secreted by NK and T cells. It was originally discovered on the basis of its antiviral activity but was later found to play a key role in the immune response and inflammatory processes. The binding of IFN- $\gamma$  to HS was recognised some years ago<sup>[28]</sup> and was found to control the blood clearance<sup>[29]</sup> and the subsequent tissue targeting, accumulation and localisation of the cytokine.<sup>[40]</sup> Thus, the discovery of compounds able to modulate the activity of

IFN- $\gamma$  could open the way to new immunotherapies,<sup>[40]</sup> especially in the field of Crohn's disease and ankylosing spondylarthropathies.<sup>[42]</sup>

IFN- $\gamma$  is a  $C_2$ -symmetric homodimer in solution and binds to HS by virtue of basic residues located at the C terminus of the two subunits (Figure 3a).<sup>[33]</sup> A fragment of HS polymer, retaining the activity of the full-length polysaccharide, has been isolated by using a protection approach (an enzymatic digestion of HS polymer in the presence of IFN- $\gamma$ ).<sup>[27]</sup> The biochemical analysis of this fragment revealed that it was composed of two small hexa- to octasaccharide NS domains, separated by a more extended NA domain of 15–16 disaccharides (130–140 Å;<sup>[40]</sup> Figure 3a). A model of the interaction between IFN- $\gamma$  and HS was thus proposed in which the KRKR domains<sup>[40]</sup> located at the C terminus of each subunit of the IFN- $\gamma$  dimer, would interact with the highly charged NS domains of the fragment.<sup>[27,39]</sup> However,

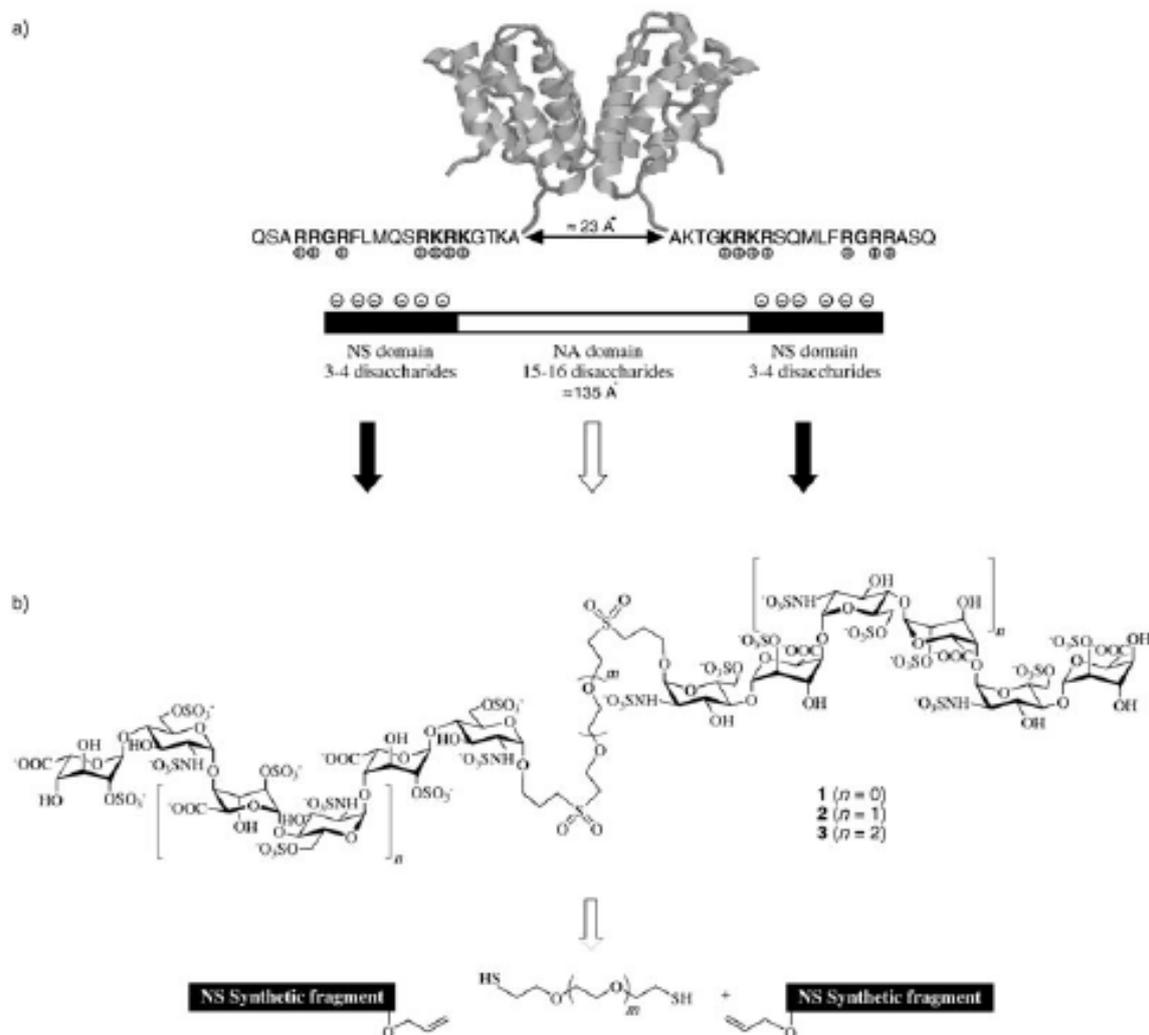


Figure 3. a) Molecular organisation of the HS binding site of IFN- $\gamma$ . b) Glycoconjugates 1–3, prepared by condensation of NS synthetic fragments bearing an allyl group in the anomeric position on bis(thio)PEGs, as functional mimetics of the IFN- $\gamma$  binding site on HS.

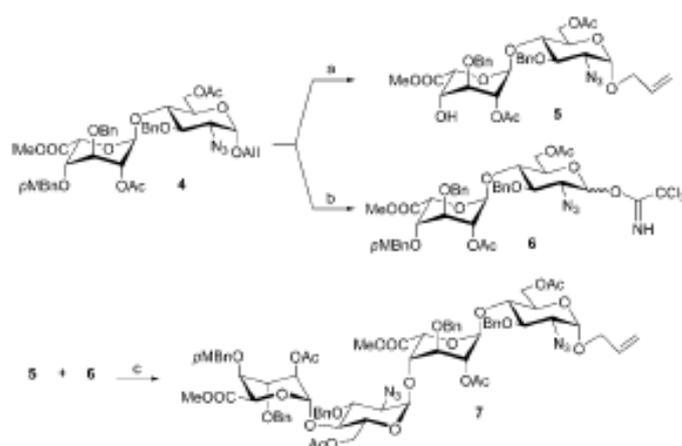
X-ray diffraction crystal structures of C-termini-truncated dimers of IFN- $\gamma$ <sup>[46,47]</sup> reveal a distance of around 23 Å between the two Ala123 residues of the IFN- $\gamma$  subunits; this distance is much shorter than the NA domain length (Figure 3a). This indicates that in this interaction the IFN- $\gamma$  dimer has to adapt the HS chain conformation to its needs. To confirm this hypothesis and to obtain compounds able to bind selectively and with high affinity to IFN- $\gamma$ , we decided to prepare glycoconjugates in which two NS domains, ranging from tetra- to octasaccharides, would be linked by their reducing end to a spacer of the poly(ethylene glycol) (PEG) type (Figure 3b). In a model study toward this goal, we have shown that  $\alpha,\omega$ -bis(thio)PEGs add in high yields to allyl anomeric derivatives to give such  $C_2$ -symmetric glycoconjugates.<sup>[27]</sup> We chose PEG linkers for their water solubility and their availability in different lengths. In this regard, we have shown that, in the synthesis of the  $\alpha,\omega$ -bis(thio)PEG intermediate,  $\alpha,\omega$ -bis(bromo)PEGs may be purified to near homogeneity by silica gel chromatography.<sup>[27]</sup> The use of PEGs as linkers has been criticised since their flexibility leads to lower binding constants than with more rigid spacers, due to the entropic cost needed for their organisation. However, such flexibility may be important to mimic NA domains whose conformational plasticity is thought to be important for their biological role.<sup>[24]</sup> Moreover, PEGs are essentially non-immunogenic and their incorporation into a drug is not a problem from the pharmacological point of view.

## Results and Discussion

The preparation of the glycoconjugates 1–3, relies upon the oligomerisation and functionalisation of the key disaccharide building block 4, for which we have developed efficient, multigram syntheses.<sup>[22,23]</sup> These compounds already contain the L-iduronic moiety protected as methyl esters, in order to avoid oxidation steps that could be troublesome on elaborated material at the end of the synthesis. We opted for a protective-group pattern allowing the same building block to be transformed into a glycosyl donor or acceptor and making clear distinction between the positions to be sulfated and those remaining as free hydroxy groups. Thus, acetates protect the positions to be sulfated while benzyl ethers are used as permanent protection. A *para*-methoxybenzyl group at the 4' position avoids the need for the preparation of a special building block for capping the nonreducing end at the end of the elongation process.<sup>[19,48]</sup> This group will be re-

moved by hydrogenolysis with the benzyl ethers in the last step of the synthesis or, selectively, by oxidation or acidic treatment to furnish a disaccharide acceptor for chain elongation. The anomeric positions are protected as an allyl glycoside that can be removed to give, after activation of the anomeric position, a disaccharide donor.

**Synthesis of tetrasaccharide 7, hexasaccharide 10 and octasaccharide 11:** A 2+2, 2+4 or 4+4 strategy was planned for the preparation of tetrasaccharide 7, hexasaccharide 10 and octasaccharide 11; this strategy required that the *para*-methoxybenzyl group and the allyl group of disaccharide 4 and tetrasaccharide 7 could be removed orthogonally without affecting the other protecting groups. This was indeed done readily by using standard procedures (Scheme 1). The

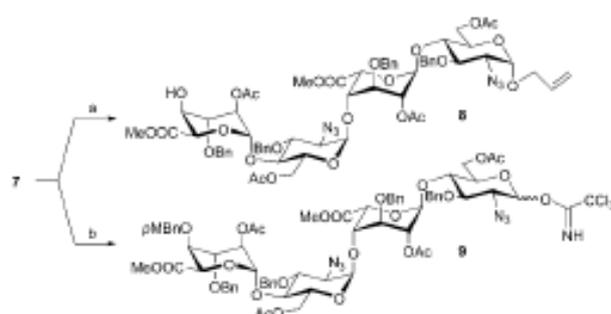


Scheme 1. a) DDQ,  $\text{CH}_2\text{Cl}_2$ , room temperature, 3 h, 83%; b) 1.  $\text{H}_2$ -activated  $[\text{Ir}^{\text{I}}(\text{C}_6\text{H}_5)_2(\text{MePh}_2\text{P})_2]\text{PF}_6$ , THF, room temperature, 2 h; 2.  $\text{HgO}/\text{HgCl}_2$ , acetone/ $\text{H}_2\text{O}$  (9:1), room temperature, 2 h; 3.  $\text{Cl}_3\text{CCN}$ ,  $\text{K}_2\text{CO}_3$ ,  $\text{CH}_2\text{Cl}_2$ , room temperature, 3 h, 88% (two steps); c) TBDMSOTf,  $\text{CH}_2\text{Cl}_2$ ,  $-40$ – $-45^\circ\text{C}$ , 90%. All = allyl, Bn = benzyl, DDQ = 2,3-dichloro-5,6-dicyano-1,4-benzoquinone, pMBn = *para*-methoxybenzyl, TBDMSOTf = *tert*-butyldimethylsilyl trifluoromethanesulfonate, THF = tetrahydrofuran.

*para*-methoxybenzyl group in the 4' position of the known disaccharide with DDQ, in wet  $\text{CH}_2\text{Cl}_2$ , to give the disaccharide acceptor 5<sup>[22]</sup> in 83% yield. For the deallylation of compound 4,  $\text{PdCl}_2/\text{AcONa}$  in aqueous  $\text{AcOH}$ <sup>[49]</sup> was tested first but gave only moderate isolated yields of the expected hemiacetal (60%). An isomerisation of the allyl group by using hydrogen-activated  $[\text{Ir}^{\text{I}}(\text{C}_6\text{H}_5)_2(\text{MePh}_2\text{P})_2]\text{PF}_6$ <sup>[50]</sup> followed by mercuric salt promoted cleavage of the resulting anomeric 2-propenyl, gave better results.<sup>[25]</sup> The resulting hemiacetal was then treated with trichloroacetonitrile and  $\text{K}_2\text{CO}_3$  in  $\text{CH}_2\text{Cl}_2$  to give the imidate 6 as a 60:40  $\alpha/\beta$  mixture, in an overall yield of 88% from 4. The  $\beta$  anomer of the disaccharide imidate 6 has already been prepared, by using a different protecting group strategy. However, it was only used as capping building block at the end of an elongation protocol in order to avoid sulfation of the nonreducing-end 4-OH group of the prepared oligosaccharides.<sup>[46]</sup>

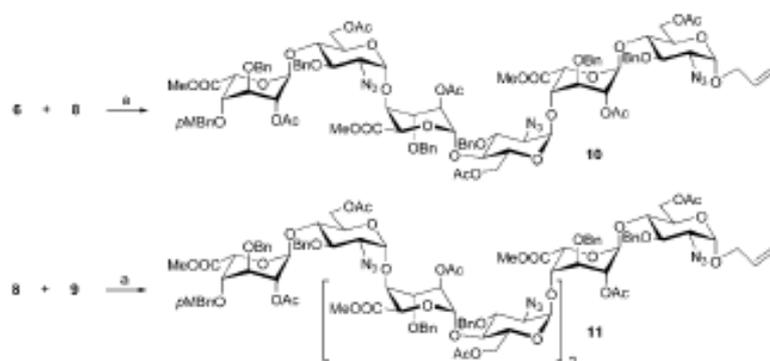
The condensation of the disaccharide acceptor 5 with the disaccharide donor 6, in  $\text{CH}_2\text{Cl}_2$  at  $-40^\circ\text{C}$  and with

TBDMSOTf as a promotor, gave the desired tetrasaccharide **7** in excellent yield (90%) and total  $\alpha$  stereoselectivity,<sup>[25]</sup> as expected for a donor with a nonparticipating group on the C-2 position (Scheme 1). Such a high stereoselectivity is a general tendency when 2-azidoglucose trichloroacetimidate mono- or oligosaccharide donors are condensed onto the 4-OH group of L-iduronyl acceptors.<sup>[10,16,19,48,55]</sup> This has been attributed previously to the conformations adopted by the L-iduronyl ring, which is in equilibrium between the <sup>1</sup>C<sub>4</sub> and <sup>2</sup>S<sub>0</sub> conformations, allowing the C-4 hydroxy group to partly occupy an axial position.<sup>[60]</sup> Indeed, in HP-fragment synthesis, lower diastereoselectivities have been observed mostly when 2-azidoglucose trichloroacetimidate donors were condensed onto D-glucuronyl acceptors that adopt the <sup>4</sup>C<sub>1</sub> conformation.<sup>[16,24,56]</sup> Tetrasaccharide **7** was then converted into tetrasaccharide acceptor **8** and trichloroacetimidate **9** in 81% and 87% yields, respectively, by using the same protocols as described above (Scheme 2). The condensation of



Scheme 2. a) DDQ, CH<sub>2</sub>Cl<sub>2</sub>, room temperature, 3 h, 81%; b) 1. H<sub>2</sub>-activated [Ir<sup>3+</sup>(C<sub>6</sub>H<sub>5</sub>)<sub>3</sub>(MePh<sub>2</sub>P)<sub>2</sub>][PF<sub>6</sub>], THF, room temperature, 2 h; 2. HgO/HgCl<sub>2</sub>, acetone/H<sub>2</sub>O (9:1), room temperature, 2 h; 3. Cl<sub>3</sub>CCN, K<sub>2</sub>CO<sub>3</sub>, CH<sub>2</sub>Cl<sub>2</sub>, room temperature, 3 h, 87% (two steps).

disaccharide and tetrasaccharide donors **6** and **9** with the tetrasaccharide acceptor **8**, under the same conditions as for the preparation of tetrasaccharide **7**, gave hexasaccharide **10** (68%) and octasaccharide **11** (93%) with total  $\alpha$  stereoselectivity<sup>[25]</sup> (Scheme 3). This demonstrates that the single disaccharide **4** may be efficiently used as building block in a



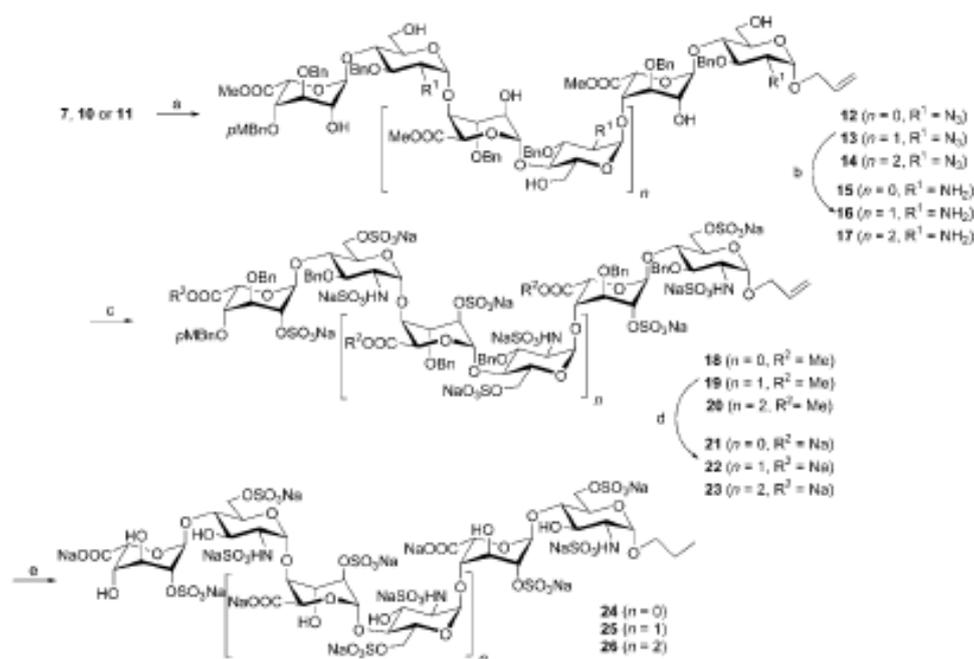
Scheme 3. a) TBDMSOTf, CH<sub>2</sub>Cl<sub>2</sub>, -40–0 °C, 68% (**10**), 93% (**11**).

2*n*+2*m* oligomerisation strategy, thereby alleviating the need for a special building block to cap the nonreducing end.

#### Preparation of the sulfated building blocks **21**, **22** and **23**:

We have previously shown that the condensation of  $\alpha,\omega$ -bis(thio)PEGs onto  $\alpha$ -allyl glucosaminyl derivatives proceeds better in water than in an organic solvent.<sup>[37]</sup> Thus, the coupling reaction has to be performed with water-soluble *O*- and *N*-sulfated oligosaccharides. Since the complete functionalisation and deprotection of the oligosaccharides **7**, **10** and **11** would result in the reduction of the allyl moiety during the last hydrogenolysis step, we chose to condense the  $\alpha,\omega$ -bis(thio)PEGs onto the benzylated, but water-soluble, oligosaccharides **21–23**. To this end, the oligosaccharides **7**, **10** and **11** were first deacetylated in near quantitative yields by using K<sub>2</sub>CO<sub>3</sub> in MeOH (Scheme 4). The next step was to perform the reduction of the azido group without reducing the anomeric allyl moiety. The Staudinger reaction,<sup>[54]</sup> which worked well in model reactions performed on disaccharide **4**, gave a mixture of products when used on larger oligosaccharides. A catalytic reduction, with hydrogen and Lindlar catalyst in the presence of quinoline,<sup>[57]</sup> was tried next but led to concomitant reduction of the allyl group. SmI<sub>2</sub> has been described as an efficient and selective reagent for the reduction of azides<sup>[58]</sup> but, in our case, treatment of tetrasaccharide **7** with a solution of SmI<sub>2</sub> led to a mixture of products. However, 1,3-dithiopropane<sup>[59]</sup> allowed a clean reduction of the azido group in oligosaccharides **12–14** without any reduction or thiol addition on the allyl moiety when performed in the absence of light. Thus, the free-amino-containing tetrasaccharide **15**, hexasaccharide **16** and octasaccharide **17** were obtained in 92, 87 and 84% yields, respectively. The sulfation of both amino and hydroxy functions in compounds **15–17** was performed by using standard conditions. Thus, after treatment with the pyridine-SO<sub>3</sub> complex in pyridine, the sulfated oligosaccharides **18–20** were obtained in 73–96% yield. Saponification of the methyl esters was performed by using 2*M* LiOH in 18% H<sub>2</sub>O<sub>2</sub>. The reaction on tetrasaccharide **18** proceeds smoothly at room temperature and went to completion within 24 h. With hexasaccharide **19** and octasaccharide **20**, the reaction did not go to completion within 24 h at room temperature.

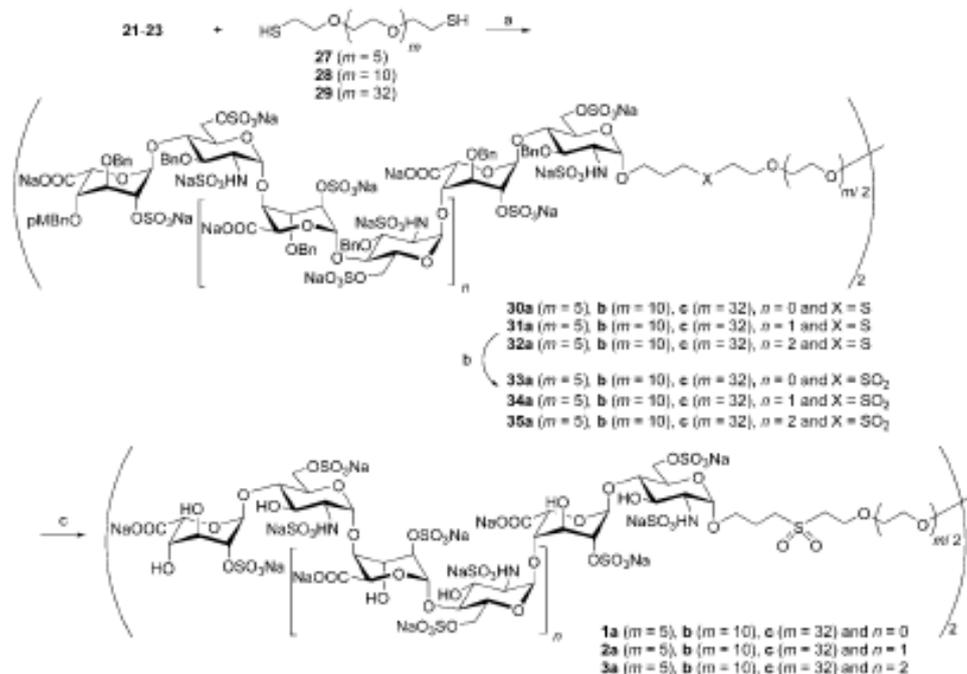
The temperature was thus raised to 37 °C to get reasonable reaction rates. Under these conditions, the saponification of compounds **19** and **20** was complete in 24 h without detectable epimerisation,  $\beta$  elimination or fragmentation. After purification by C-18 reversed-phase chromatography, the desired tetra-, hexa- and octasaccharides **21**, **22** and **23** were obtained in 73, 91 and 80% yields, respectively.



Scheme 4. a)  $\text{K}_2\text{CO}_3$ , MeOH, room temperature, 93% (12), 88% (13), 91% (14); b)  $\text{HS}(\text{CH}_2)_3\text{SH}$ ,  $\text{NEt}_3$ , MeOH, room temperature, 92% (15), 87% (16), 84% (17); c) pyridine $\text{SO}_2$ , pyridine, 24 h, room temperature, then 24 h, 55°C, 73% (18), 95% (19), 96% (20); d)  $\text{LiOH}$ ,  $\text{H}_2\text{O}_2$ , 73% (21), 91% (22), 80% (23); e)  $\text{Pd}(\text{OH})_2\text{C}$ ,  $\text{H}_2$ , phosphate buffer (pH 7.0), 48 h, room temperature, 94% (24), 85% (25), quant. (26).

**Preparation of glycoconjugates 1a–c, 2a–c and 3a–c:** For the syntheses of glycoconjugates 1a–c, 2a–c and 3a–c, almost homogeneous  $\alpha,\omega$ -bis(thio)PEGs 27 ( $m=5$ ) and 28 ( $m=10$ ) were prepared from commercial PEG-300 and PEG-600, and a more polydisperse sample 29 ( $m=32$ ) was prepared from PEG-1500.<sup>[27]</sup> These PEG lengths were chosen because, after conjugation with the allyl group, they should give linkers of different lengths (33, 50 and 114 Å), thereby allowing study of the effect of the distance between the sulfated regions on the biological activity of the glycoconjugates. The first two linkers should restrict the access of the NS domains to the first basic region of the IFN- $\gamma$  C terminus (KRKR domain), while the longer linker should allow their access to the furthest basic region (RGRR domain; Figure 3a). The conditions that we had optimised in our preliminary studies<sup>[27]</sup> were first applied to the preparation of the nine targeted glycoconjugates 30a–c, 31a–c and 32a–c. A mixture of tetrasaccharide 21 and  $\alpha,\omega$ -bis(thio)PEGs 27 was thus irradiated, in a quartz vessel, with the light of a medium-pressure Hg lamp (Scheme 5). However, the results were disappointing since HPLC analyses of the crude reaction mixture indicated that the major product was only formed in 40%. Benzyl protecting groups have been reported recently to be reactive under short-wave UV irradiation.<sup>[69]</sup> The discrepancies between our model experiments and those performed with compounds 21 could, therefore, originate from the benzyl moieties. The coupling reactions were thus performed by using a filter with maximum transmittance at 360 nm. Under these conditions, the proportion of the desired glycoconjugates 30–32 rose to 70%.

Glycoconjugates 30a, 30c, 31a, 31c and 32a–c were then purified by semipreparative RP-18 HPLC because we were not certain that the side products only included compounds containing changes in the number of benzyl groups. This yielded the conjugates in yields of 34–66% and purities ranging from 92–98%<sup>[69]</sup> (Scheme 5). These compounds were unambiguously characterised by ESI-MS mass analyses and  $^1\text{H}$  NMR spectroscopy. In addition, HMQC experiments were performed on compounds 30a, 31a and 32b. In our initial synthetic scheme the last step in the synthesis should have been the hydrogenolysis of the permanent benzyl protecting groups. However, due to the presence of the sulfide linkage this reaction failed. The thioether functions were thus oxidised to sulfones by using oxone,<sup>[62]</sup> to give the expected glycoconjugates 33a, 33c, 34a, 34c and 35a–c (Scheme 5), unfortunately, as a mixture. Indeed products resulting from partial cleavage of the *para*-methoxybenzyl group were obtained, as shown by HPLC and  $^1\text{H}$  NMR analyses. Compounds 30b and 31b were not purified at the thioether stage, but the coupling reaction mixtures were directly oxidised with oxone, and the resulting glycoconjugates were purified by semi-preparative RP-18 HPLC to give 33b and 33c in 21–31% yield. These low yields indicate that, along with the side products arising from the UV irradiation, compounds that had the lost *para*-methoxybenzyl group were removed during the HPLC purification. To our knowledge, there is no report of oxidative removal of a *para*-methoxybenzyl group with oxone. However, the abstraction of a *para*-methoxybenzyl benzylic hydrogen atom by  $\text{KHSO}_5$ , with a mechanism similar to DDQ, cannot be excluded and



Scheme 5. a) hv (360 nm), 34–66%; b) KHSO<sub>5</sub> (oxone), K<sub>2</sub>HPO<sub>4</sub>, (pH 7), 23–31% (combined yields for steps (a) and (b)); c) Pd(OH)<sub>2</sub>/C, H<sub>2</sub>, phosphate buffer (pH 7.0), 96 h, room temperature, 62% to quant. (combined yields for steps (b) and (c)).

would explain the observed side reaction. Excess oxone allows efficient and quick conversion of sulfide into sulfone without detectable sulfoxide intermediates.<sup>[6]</sup> Thus, since partially deprotected side products should lead after hydrogenolysis to the same product as the fully protected ones, we decided to use this reagent in order to avoid uncompleted sulfide oxidation and formation of barely detectable sulfoxides. The final hydrogenolysis was thus performed directly after desalting (PD-10) the reaction mixture of the previous step. The oxidised glycoconjugates were thus obtained in 62% to quantitative yields<sup>[6]</sup> and were shown to be homogeneous by PAGE analysis (Figure 4), <sup>1</sup>H NMR spectroscopy and HSQC experiments. The MS analysis of large poly-

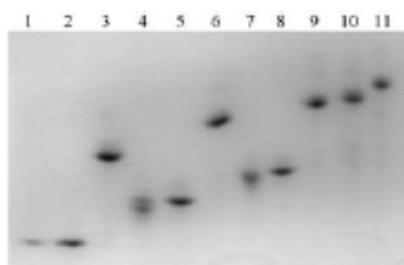


Figure 4. PAGE analysis of the glycoconjugates. Heparin-derived oligosaccharides (prepared as in ref [25]) and glycoconjugates (1 μg each) were run through a 30% polyacrylamide gel and stained with azure A. Lane 1: heparin-derived tetrasaccharide; lane 2: 24; lane 3: 1b; lane 4: heparin-derived hexasaccharide; lane 5: 25; lane 6: 2b; lane 7: heparin-derived octasaccharide; lane 8: 26; lane 9: 3a; lane 10: 3b; lane 11: 3c.

anionic HS- and HP-derived oligosaccharides is still a matter of intensive research.<sup>[64–67]</sup> The product signals are generally low for large oligosaccharides, even in the negative mode, since they poorly ionise. Thus, as expected, the recording of MS data on the free glycoconjugates was difficult even when using ESI techniques. A marked difference was indeed found between the signals of the debenzylated glycoconjugates and the ones obtained with the benzylated glycoconjugates, for which ESI-MS data in accordance with the proposed structures were obtained. However, correct ESI-MS data were obtained at least for compounds 1b, 2b and 3b. PAGE analysis showed that synthetic tetra-, hexa- and octasaccharides 24–26 (lanes 2, 5 and 8, respectively, in Figure 4) have the same migration pattern and are more pure than the corresponding oligosaccharides obtained from natural sources (lanes 1, 4 and 7, respectively). Conjugation to PEG linkers is clearly evidenced by the shift in migration observed for 1b (lane 3), 2b (lane 6) and 3a–c (lanes 9, 10 and 11, respectively). Moreover, the difference in the migration of the glycoconjugate 3a (33 Å linker, lane 9) and the glycoconjugate 3b (50 Å linker, lane 10) demonstrates the high-resolution power of this PAGE technique, since there is only a difference of 220 gmol<sup>-1</sup> (≈ 4%) in molecular weight between the two compounds.

**Biological properties of the glycoconjugates 1a–c, 2a–c, 3a–c and 24–26:** Surface plasmon resonance was used here as a detection system to analyse the ability of the nine glycoconjugates, 1a–c, 2a–c and 3a–c, and the three nonconjugated controls, 24–26, to inhibit the binding of IFN-γ to heparin.

Injection of IFN- $\gamma$  (7.5 nM) over a Biacore sensor chip containing streptavidin and heparin produced a binding response of 175 resonance units (RU) at equilibrium. A response of 5–10 RU was observed with a similar injection over a streptavidin sensor chip, used as a blank surface (not shown). To obtain accurate glycoconjugate concentration data a colorimetric determination of the uronate content was performed.<sup>[60]</sup> The obtained values were in accordance with the weighted values. All the experiments were performed in duplicate or triplicate by using two or three batches of glycoconjugates prepared independently (standard errors were within 8–15% of the mean). Preliminary experiments demonstrated that among the nine glycoconjugates, only 3a–c displayed inhibition activity (not shown). The activities of the three octasaccharide glycoconjugates 3a–c and the monomeric octasaccharide control 26 were first compared. For that purpose, IFN- $\gamma$  (7.5 nM) was preincubated with 150 nM 3a–c and 26 and injected over the heparin surface. Representative sensorgrams are shown in Figure 5A and demonstrate that the glycoconjugates 3a–c strongly inhibit the IFN- $\gamma$ /HP interaction, while the inhibition by the sole octasaccharide 26 is low. Moreover, there is a dependence of the activity on the linker length. Compound 3b (50-Å linker length) is the most potent inhibitor while 3c (114-Å linker length) is the least active and 3a (33-

Å linker length) has medium activity. By using the same binding assay, the concentration dependence of the inhibition by the most active glycoconjugate (3b) was then studied. Figure 5B shows that the activity of 3b is dose dependant with an IC<sub>50</sub> value of approximately 35 nM, that is, within the range of the dissociation constant,  $K_d$ , of IFN- $\gamma$  on HS.<sup>[29]</sup>

These results are consistent with the view that the KRKR domains of the IFN- $\gamma$  C termini are critically involved in the interaction. Moreover, it seems that the 33 Å linker is too small and needs to induce constraints in the IFN- $\gamma$  dimer to allow binding, while the 114 Å linker is too long and has to adapt to the protein. Glycoconjugate 3b, with a 50-Å linker, fits best to the IFN- $\gamma$  dimer geometry and is thus the better functional mimetic of the HS binding site of IFN- $\gamma$ . The optimal linker length is thus two to three times smaller than the NA domain found in the natural IFN- $\gamma$  HS binding site. This confirms the hypothesis that the conformation of the natural HS binding site has to be adapted to IFN- $\gamma$  geometry in order to allow an efficient binding. The C<sub>2</sub> symmetry of compound 3b may also be an important factor for its activity by fitting better with the symmetry of the IFN- $\gamma$  dimer than the natural HS fragment. It is thus possible that, in order to allow efficient binding, the NA domain of the natural fragment has to adopt a conformation allowing a more C<sub>2</sub>-symmetric "head-to-head" presentation of the NS domains. Such a factor has already been found to be important in the binding of RANTES on HP.<sup>[29]</sup> The preparation of glycoconjugates with other symmetry will help to confirm this hypothesis.

Finally, as shown in Figure 5B, our results also point out the importance of the oligosaccharide size for binding. Indeed, the glycoconjugates containing tetrasaccharide (1b) or hexasaccharide (2b) NS domains, although linked with the optimal linker (50 Å), are unable to significantly interact with IFN- $\gamma$ , even at the highest concentration used (150 nM).

Compound 3b is thus a compound on which potential specific inhibitors of IFN- $\gamma$  activity may be constructed.

## Conclusion

We have thus demonstrated that disaccharide 4 is a potent building block for the preparation of oligosaccharides of different lengths deriving from the heparin regular region. From these oligosaccharides, we have optimised a methodology for the preparation of glycoconjugates mimicking the HS binding site of IFN- $\gamma$ . We have thus found that compound 3b was able to inhibit the IFN- $\gamma$ /HP interaction with an IC<sub>50</sub> value of approximately 35 nM. In addition, we have shown that the nine glycoconjugates, prepared with various oligosaccharide and PEG-linker lengths, are perfect tools to define the topology of the IFN- $\gamma$  binding site on HS. However, these results give rise to new intriguing questions, since it is entirely unclear if other linker systems, linking the sugar moiety by the nonreducing end, longer oligosaccharides or modification of the sulfation and uronic acid patterns would show better properties for the exploration of this kind of interaction. We are currently developing the

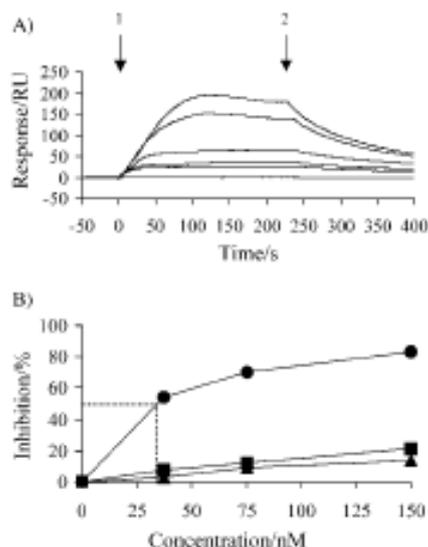


Figure 5. Inhibition of the IFN $\gamma$ /heparin binding by different glycoconjugates. A) IFN $\gamma$  (7.5 nM) was preincubated with 150 nM glycoconjugates and then injected (arrow 1,  $t=0$ ) over a heparin-activated surface, as described in the experimental section. At the end of the injection phase (arrow 2,  $t=240$  s), the formed complexes were washed with running buffer. The binding response (in RU) corresponding to the injection of IFN $\gamma$  (a), IFN $\gamma$  + 26 (b), IFN $\gamma$  + 3c (c), IFN $\gamma$  + 3a (d), IFN $\gamma$  + 3b (e), and plain buffer (f) was recorded as a function of time. B) IFN $\gamma$  (7.5 nM) was preincubated with increasing concentrations (0–150 nM) of 3b (●), 2b (■) or 1b (▲) and injected over a heparin-activated surface as described in the Experimental Section. The level of IFN $\gamma$  bound to the heparin surface at the end of the association phase was recorded and the results were expressed as a percentage of inhibition.

tools needed to answer these questions and the results will be reported in due course.

## Experimental Section

**General procedures:** All moisture-sensitive reactions were performed under an argon atmosphere by using oven-dried glassware. All solvents were dried over standard drying agents<sup>16</sup> and freshly distilled prior to use. Evaporations were performed under reduced pressure. UV irradiation was performed using a 150 W medium-pressure Hg lamp (TQ 150, Heraeus) equipped with a midrange/longwave UV filter (45% transmittance at 360 nm, cut-off at 275 and 475 nm); during irradiation the samples were air cooled. Reactions were monitored by TLC on glass silica gel 60 F<sub>254</sub> plates with detection by UV light at 254 nm and by charring with 5% ethanolic H<sub>2</sub>SO<sub>4</sub> or orcinol reagent<sup>16</sup> for diluted solutions. Flash column chromatography was performed on Silica Gel 60 A.C.C. 6–35  $\mu$  (SDS) or on LiChroprep RP-18 (Merck). HPLC was performed by using a Waters Spherisorb ODS-2 5  $\mu$  C18 250  $\times$  4.6 mm column and UV detection at 220 nm; HPLC purity is given assuming that the major compound and all the impurities have the same molar absorption at 220 nm. The elution was performed at a rate of 1 mL min<sup>-1</sup> with a linear gradient of 10 mM AcOH-NEt<sub>3</sub> buffer (pH 7.0)/CH<sub>2</sub>Cl<sub>2</sub>. Semipreparative HPLC was performed by using a Waters Spherisorb ODS-2 5  $\mu$  C18 250  $\times$  20 mm column, eluting at 20 mL min<sup>-1</sup> with a linear gradient of 5 mM AcOH-NEt<sub>3</sub> buffer (pH 7.0)/CH<sub>2</sub>Cl<sub>2</sub>. Melting points were determined with a Büchi capillary apparatus and are uncorrected. Optical rotations were measured on a Jasco DIP 370 digital polarimeter. NMR spectra were recorded at room temperature with Bruker AC200, AC250, AM250, AM300 or DRX400 spectrometers. Chemical shifts ( $\delta$ ) are given in parts per million (ppm) relative to an internal Me<sub>4</sub>Si reference, solvent signals (CDCl<sub>3</sub>:  $\delta$ (<sup>13</sup>C) = 77.0 ppm) or acetone in D<sub>2</sub>O ( $\delta$ (<sup>1</sup>H) = 2.225 ppm and  $\delta$ (<sup>13</sup>C) = 30.5 ppm). The allyl group carbon atoms are identified in the following way: O-C<sub>1</sub>H<sub>2</sub>-C<sub>2</sub>H=C<sub>3</sub>H<sub>2</sub>; the two protons on C-c are identified as H-c<sub>1</sub>, for the one *cis* to H-b, and H-c<sub>2</sub>, for the one *trans* to H-b. Saccharide units in oligosaccharides are identified alphabetically from the reducing end. For compounds 10–17, spin systems, identified with COSY experiments, were attributed to a monosaccharide unit based on <sup>1</sup>H and <sup>13</sup>C chemical shifts; for compounds 18–20, this attribution was confirmed by using HMBC experiments. Apodizations with Gaussian functions (LB = -1 to -2 Hz and GB = 50%) were used, thereby allowing measurement of coupling constants. COSY, gradient-enhanced COSY, HMBC, HMQC and HSQC experiments were performed by recording 256 FID measurements with 1024 complex data points and using standard Bruker programs. Prior to Fourier transformation, the data were zero filled in the *t*<sub>1</sub> dimension to 1024 points and multiplied with a nonshifted sinebell function in both dimensions for COSY experiments; for HMBC experiments, the data were zero filled in the *t*<sub>1</sub> dimension to 2048 points and an exponential multiplication (LB = 3 Hz) in the *t*<sub>2</sub> dimension and a  $\pi/2$ -shifted squared sinebell function in the *t*<sub>2</sub> dimension were used; for HSQC experiments, the data were zero filled in the *t*<sub>1</sub> dimension to 2048 points and multiplied with a  $\pi/3$ -shifted squared sinebell function in both dimensions. MS spectra were recorded in the positive or negative mode on a Finnigan MAT 95S spectrometer by using electrospray ionisation. IR spectra were recorded on a Fourier transformation Bruker IFS66 apparatus. Elemental analyses were performed at the CNRS (Gif sur Yvette, France).

**Polyacrylamide gel electrophoresis:** PAGE analysis was performed essentially as previously described.<sup>17</sup> Oligosaccharides (1  $\mu$ g) in 20% glycerol were run initially through a stacking gel (8% acrylamide) at a constant current (15 mA), then through a resolving gel (30% acrylamide) at 20 mA, until the Phenol Red marker (applied to a separate lane as the electrophoresis marker) had reached the bottom of the gel. The discontinuous buffer system used comprised 0.125 M tris(hydroxymethyl)aminomethane (Tris)/HCl (pH 6.8) in the stacking gel, 0.375 M Tris/HCl (pH 8.8) in the resolving gel and 25 mM Tris/0.192 M glycine (pH 8.3) in the tank buffer. After electrophoresis, oligosaccharide bands were stained with 0.08% aqueous Azure A for a few minutes under constant agitation. Excess dye was removed by washing the gel in water.

**HN- $\gamma$ /heparin binding analysis by using surface plasmon resonance:** Binding analysis was performed on a Biacore 2000 apparatus, equilibrated with HBS-EP buffer (10 mM 2-[4-(2-hydroxyethyl)-1-piperazinyl]ethanesulfonic acid (HEPES), 0.15 M NaCl, 3 mM ethylenediaminetetraacetate (EDTA), 0.05% surfactant P20, pH 7.4), at 25 °C. Size-defined heparin (9 kDa) was biotinylated at the reducing end and immobilised on a Biacore sensor chip.<sup>18</sup> For this purpose, two flow cells of a Biacore F1 sensor chip were activated with a mixture (50  $\mu$ L) of 0.2 M 3-(3-dimethylaminopropyl)-1-ethylcarbodiimide (EDC)/0.05 M *N*-hydroxysuccinimide (NHS) before injection of streptavidin (50  $\mu$ L; 0.2 mg mL<sup>-1</sup> in 10 mM acetate buffer, pH 4.2). Remaining activated groups were blocked with 1 M ethanolamine (50  $\mu$ L; pH 8.5). Typically, this procedure, performed at a rate of 5  $\mu$ L min<sup>-1</sup>, permitted coupling of approximately 2000–2500 RU of streptavidin. Biotinylated heparin (5  $\mu$ g mL<sup>-1</sup> in HBS-EP buffer) was then injected at a rate of 5  $\mu$ L min<sup>-1</sup> over one of the two surfaces to obtain an immobilisation level of 50–60 RU. The other surface (streptavidin) was left as a blank surface. Flow cells were then conditioned with several injections of 1 M NaCl. For binding inhibition assays, HN- $\gamma$  (7.5 nM in HBS-EP buffer), either alone or preincubated with the oligosaccharides (0–150 nM; see legend of Figure 5), was injected at a rate of 50  $\mu$ L min<sup>-1</sup> over the two surfaces (streptavidin and streptavidin/heparin) for 4 min, after which the complexes formed were washed with buffer. The sensor-chip surface was regenerated at a rate of 50  $\mu$ L min<sup>-1</sup> with a 4-min pulse of 1 M NaCl.

**Allyl (methyl 2-O-acetyl-3-O-benzyl- $\alpha$ -L-idopyranosyluronate)-(1 $\rightarrow$ 4)-O-6-O-acetyl-2-azido-3-O-benzyl-2-deoxy- $\alpha$ -D-glucopyranoside (5):** DDO (415 mg, 1.5 equiv) was added to a solution of disaccharide 4 (1.00 g, 1.2 mmol) in CH<sub>2</sub>Cl<sub>2</sub> saturated with water (15 mL). After 3 h at room temperature, Et<sub>2</sub>O (150 mL) was added and the resulting solution was successively washed with ice-cold satd aq NaHCO<sub>3</sub> solution (50 mL) and water (2  $\times$  50 mL), filtered, dried (MgSO<sub>4</sub>) and concentrated. Flash chromatography of the residue (silica gel, petroleum ether/AcOEt (1:1 $\rightarrow$ 4:6)) gave disaccharide 5 (700 mg, 83%), whose <sup>1</sup>H and <sup>13</sup>C NMR data were identical to those previously reported.<sup>19</sup>

**(Methyl 2-O-acetyl-3-O-benzyl-4-O-(4-methoxybenzyl)- $\alpha$ -L-idopyranosyluronate)-(1 $\rightarrow$ 4)-O-6-O-acetyl-2-azido-3-O-benzyl-2-deoxy- $\alpha$ -D-glucopyranoside trichloroacetimidate (6):** The iridium catalyst (C<sub>12</sub>H<sub>8</sub>(MePh)<sub>2</sub>P)<sub>2</sub>I<sup>+</sup>PF<sub>6</sub><sup>-</sup> (20 mg, 0.024 mmol, 1.3 mol %) was added to a solution of disaccharide 4<sup>20,21</sup> (1.49 g, 1.8 mmol) in THF (30 mL). The mixture was degassed, the Ir catalyst was activated with H<sub>2</sub> as previously described,<sup>19</sup> and, after 2 h at room temperature, the reaction was concentrated. HgO (565 mg, 2.16 mmol, 1.2 equiv) and HgCl<sub>2</sub> (538 mg, 1.98 mmol, 1.1 equiv) were then added to a solution of the residue in acetone/H<sub>2</sub>O (9:1; 40 mL). After 2 h stirring at room temperature, the reaction mixture was filtered over a pad of celite 545 and concentrated. The residue was dissolved in Et<sub>2</sub>O (150 mL) and the resulting solution was successively washed with 10% (w/v) aq KI solution (2  $\times$  100 mL), satd aq Na<sub>2</sub>S<sub>2</sub>O<sub>4</sub> (50 mL) and water (2  $\times$  50 mL), filtered, dried (MgSO<sub>4</sub>) and concentrated. Flash chromatography of the residue (silica gel, toluene/AcOEt (8:2 $\rightarrow$ 6:4)) gave the anticipated hemiacetal (1.26 g, 90%) whose <sup>1</sup>H and <sup>13</sup>C NMR data were identical to those previously reported.<sup>19</sup> This compound was dissolved in CH<sub>2</sub>Cl<sub>2</sub> (2.5 mL) and trichloroacetonitrile (970  $\mu$ L, 9.7 mmol, 6 equiv) was added, followed by potassium carbonate (400 mg, 32 mmol, 2 equiv). After being stirred for 3 h at room temperature, the reaction mixture was directly applied to the top of a flash chromatography column filled with silica gel and eluted (toluene/AcOEt (9:1 $\rightarrow$ 8:2) with 0.1% NEt<sub>3</sub>) to give imidate 6 (1.51 g, 98%) as an  $\alpha/\beta$  (4:6) mixture. Along with the previously described  $\beta$  anomer<sup>19</sup> (60%), we also obtained the  $\alpha$  anomer (40%): <sup>1</sup>H NMR (250 MHz, CDCl<sub>3</sub>):  $\delta$  = 8.73 (s, 1H; NH), 7.40–7.22 (m, 10H; Ph), 7.11 (d, *J* = 8.5 Hz, 2H; PhOMe), 6.82 (d, *J* = 8.5 Hz, 2H; PhOMe), 6.37 (d, *J*<sub>1,2</sub> = 3.5 Hz, 1H; H-1 <sup>$\alpha$</sup> ), 5.26 (d, *J*<sub>1,2</sub> = 4.5 Hz, 1H; H-1 <sup>$\beta$</sup> ), 4.92 (d, *J* = 10.5 Hz, 1H; CH<sub>2</sub>Ph), 4.50 (t, *J*<sub>1,2</sub> = *J*<sub>2,3</sub> = 4.5 Hz, 1H; H-2 <sup>$\beta$</sup> ), 4.75–4.64 (m, 3H), 4.63 (d, *J* = 10.5 Hz, 1H; CH<sub>2</sub>Ph), 4.45 (d, *J* = 11.5 Hz, 1H; CH<sub>2</sub>PhOMe), 4.39 (dd, *J*<sub>4,6a</sub> = 12.5, *J*<sub>4,5</sub> = 2.0 Hz, 1H; H-6 <sup>$\alpha$</sup> ), 4.37 (d, *J* = 11.5 Hz, 1H; CH<sub>2</sub>PhOMe), 4.20 (dd, *J*<sub>6a,6b</sub> = 12.5, *J*<sub>6a,5</sub> = 3.5 Hz, 1H; H-6 <sup>$\beta$</sup> ), 4.11 (dd, *J*<sub>4,5</sub> = 10.0, *J*<sub>4,3</sub> = 9.0 Hz, 1H; H-3 <sup>$\alpha$</sup>  or H-4 <sup>$\alpha$</sup> ), 3.98 (ddd, *J*<sub>3,4</sub> = 10.0, *J*<sub>3,6a</sub> = 3.5, *J*<sub>3,6b</sub> = 2.0 Hz, 1H; H-5 <sup>$\alpha$</sup> ), 4.02 (dd, *J*<sub>1,2</sub> = 10.0, *J*<sub>1,4</sub> = 9.0 Hz, 1H; H-3 <sup>$\beta$</sup>  or H-4 <sup>$\beta$</sup> ), 3.95–3.90 (m with s at  $\delta$  = 3.80, 5H; H-4 <sup>$\alpha$</sup> , H-3 <sup>$\beta$</sup>  and PhOMe), 3.69 (dd, *J*<sub>1,2</sub> = 10.0, *J*<sub>1,3</sub> = 3.5 Hz, 1H; H-2 <sup>$\alpha$</sup> ), 3.54 (s, 3H; COOMe), 2.04 (s, 3H; CH<sub>3</sub>OAc), 2.03 (s, 3H; CH<sub>3</sub>OAc) ppm; <sup>13</sup>C NMR (62.5 MHz, CDCl<sub>3</sub>):  $\delta$  =

170.5, 170.0, 169.8 (C=O), 160.3 (C-NH), 159.4 (C-OMe pMBn), 137.7 (C<sub>quat</sub>), 129.6, 129.3, 128.4, 128.1, 128.0, 127.9, 127.5 (C<sub>arom</sub>), 113.7 (C<sub>arom</sub> pMBn), 98.0 (C-1<sup>o</sup>), 94.4 (C-1<sup>o</sup>), 78.2, 75.2, 75.0, 74.7, 74.4, 73.0, 72.4, 71.8, 70.8, 70.2, 62.8 (C-2<sup>o</sup>), 61.6 (C-6<sup>o</sup>), 55.2 (CH<sub>3</sub> pMBn), 51.7 (CH<sub>3</sub> COOMe), 20.9, 20.6 (CH<sub>3</sub> OAc) ppm.

Allyl [(methyl 2-O-acetyl-3-O-benzyl-4-O-(4-methoxybenzyl)- $\alpha$ -L-idopyranosyluronate)-(1 $\rightarrow$ 4)-O-(6-O-acetyl-2-azido-3-O-benzyl-2-deoxy- $\alpha$ -D-glucopyranoside)-(1 $\rightarrow$ 4)-O-(methyl 2-O-acetyl-3-O-benzyl- $\alpha$ -L-idopyranosyluronate)]-(1 $\rightarrow$ 4)-O-(6-O-acetyl-2-azido-3-O-benzyl-2-deoxy- $\alpha$ -D-glucopyranoside) (7): Acceptor 5 (460 mg, 0.66 mmol) and imidate 6 (785 mg, 0.85 mmol, 1.3 equiv) were azeotropically dried with toluene and dissolved in CH<sub>2</sub>Cl<sub>2</sub> (2 mL). Powdered 4-Å molecular sieves (1.86 g) were added and the mixture was then stirred for 30 min at room temperature. The solution was cooled to -40°C and TBDMSOTf (0.1 M in CH<sub>2</sub>Cl<sub>2</sub>, 1.3 mL, 0.13 mmol, 0.2 equiv) was added. The reaction was stirred at this temperature for 30 min and then the temperature was raised to 0°C over 30 min. The reaction was then quenched with a solution of NEt<sub>3</sub> (0.1 M in CH<sub>2</sub>Cl<sub>2</sub>, 400  $\mu$ L, 0.4 mmol, 0.6 equiv), directly applied to the top of a flash chromatography column filled with silica gel and eluted (CH<sub>2</sub>Cl<sub>2</sub>/AcOEt/petroleum ether (85:10:5 $\rightarrow$ 70:25:5)) to give tetrasaccharide 7 (860 mg, 90%): [ $\alpha$ ]<sub>D</sub><sup>20</sup> = -17 (c = 1.05, CH<sub>2</sub>Cl<sub>2</sub>); <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  = 7.37–7.24 (m, 20H; Ph), 7.13 (d, J = 8.5 Hz, 2H; Ph-OMe), 6.82 (d, J = 8.5 Hz, 2H; Ph-OMe), 5.94 (dddd, J<sub>1,2</sub> = 17.0, J<sub>1,3</sub> = 10.5, J<sub>1,4</sub> = 6.0, J<sub>1,5</sub> = 5.0 Hz, 1H; H-b), 5.36 (dq, J<sub>2,3</sub> = 17.0, J<sub>2,4</sub> = J<sub>2,5</sub> = 1.5 Hz, 1H; H-c), 5.28 (d, J<sub>3,4</sub> = 5.0 Hz, 1H; H-1<sup>o</sup>), 5.26 (dq, J<sub>3,4</sub> = 10.5, J<sub>3,5</sub> = J<sub>3,6</sub> = 1.5 Hz, 1H; H-cc), 5.24 (d, J<sub>3,4</sub> = 3.5 Hz, 1H; H-1<sup>o</sup>), 4.94 (d, J<sub>3,4</sub> = 3.5 Hz, 1H; H-1<sup>o</sup>), 4.93 (dd, J<sub>3,4</sub> = 4.0, J<sub>3,5</sub> = 3.5 Hz, 1H; H-2<sup>o</sup>), 4.91 (d, J<sub>3,4</sub> = 3.5 Hz, 1H; H-1<sup>o</sup>), 4.90 (d, J = 10.5 Hz, 1H; CH<sub>2</sub>Ph), 4.86 (dd, J<sub>3,4</sub> = 5.5, J<sub>3,5</sub> = 5.0 Hz, 1H; H-2<sup>o</sup>), 4.82 (d, J = 10.5 Hz, 1H; CH<sub>2</sub>Ph), 4.73 (s, 2H; CH<sub>2</sub>Ph), 4.69 (d, J = 10.5 Hz, 1H; CH<sub>2</sub>Ph), 4.68 (d, J<sub>3,4</sub> = 4.0 Hz, 1H; H-5<sup>o</sup>), 4.66 (s, 2H; CH<sub>2</sub>Ph), 4.63 (d, J = 10.5 Hz, 1H; CH<sub>2</sub>Ph), 4.59 (d, J<sub>3,4</sub> = 5.0 Hz, 1H; H-5<sup>o</sup>), 4.46 (d, J = 11.5 Hz, 1H; CH<sub>2</sub>PhOMe), 4.42 (d, J = 11.5 Hz, 1H; CH<sub>2</sub>PhOMe), 4.40 (brd, J<sub>4,5</sub> = 12.0 Hz, 1H; H-6<sup>o</sup>), 4.34 (dd, J<sub>4,5</sub> = 12.5, J<sub>4,6</sub> = 2.0 Hz, 1H; H-6<sup>o</sup>), 4.24 (dd, J<sub>4,5</sub> = 12.0, J<sub>4,6</sub> = 3.0 Hz, 1H; H-6<sup>o</sup>), 4.20 (ddt, J<sub>4,5</sub> = 13.0, J<sub>4,6</sub> = 5.0, J<sub>4,7</sub> = 1.5 Hz, 1H; H-a), 4.16 (dd, J<sub>4,5</sub> = 12.5, J<sub>4,6</sub> = 3.0 Hz, 1H; H-6<sup>o</sup>), 4.05 (ddt, J<sub>4,5</sub> = 13.0, J<sub>4,6</sub> = 6.0, J<sub>4,7</sub> = 1.5 Hz, 1H; H-a), 3.98 (dd, J<sub>4,5</sub> = 5.0, J<sub>4,6</sub> = 4.0 Hz, 1H; H-4<sup>o</sup>), 3.95 (dd, J<sub>4,5</sub> = 10.0, J<sub>4,6</sub> = 9.0 Hz, 1H; H-4<sup>o</sup>), 3.92 (dd, J<sub>4,5</sub> = 5.0, J<sub>4,6</sub> = 4.0 Hz, 1H; H-3<sup>o</sup>), 3.90–3.85 (m, 3H; H-3<sup>o</sup>, H-4<sup>o</sup> and H-5<sup>o</sup>), 3.82–3.78 (m with s at  $\delta$  = 3.80, 5H; H-4<sup>o</sup>, H-5<sup>o</sup> and PhOMe), 3.76 (t, J<sub>3,4</sub> = J<sub>3,5</sub> = 5.5 Hz, 1H; H-3<sup>o</sup>), 3.65 (dd, J<sub>3,4</sub> = 10.0, J<sub>3,5</sub> = 9.0 Hz, 1H; H-3<sup>o</sup>), 3.56 (s, 3H; COOMe), 3.49 (s, 3H; COOMe), 3.39 (dd, J<sub>3,4</sub> = 10.0, J<sub>3,5</sub> = 3.5 Hz, 1H; H-2<sup>o</sup>), 3.29 (dd, J<sub>3,4</sub> = 10.0, J<sub>3,5</sub> = 3.5 Hz, 1H; H-2<sup>o</sup>), 2.12 (s, 3H; CH<sub>3</sub> OAc), 2.06 (s, 3H; CH<sub>3</sub> OAc), 2.03 (s, 3H; CH<sub>3</sub> OAc), 2.02 (s, 3H; CH<sub>3</sub> OAc) ppm; <sup>13</sup>C NMR (100.6 MHz, CDCl<sub>3</sub>):  $\delta$  = 170.6, 170.5, 169.9, 169.8, 169.7, 169.3 (C=O), 159.4 (C-OMe pMBn), 137.8, 137.7, 137.6, 137.3 (C<sub>arom</sub>), 133.0 (C-b), 129.6, 129.2, 128.4, 128.3, 128.0, 127.8, 127.6, 127.5, 127.4 (C<sub>arom</sub>), 118.3 (C-c), 113.7 (C<sub>arom</sub> pMBn), 97.9 (C-1<sup>o</sup>), 97.8 (C-1<sup>o</sup>), 97.4 (C-1<sup>o</sup>), 96.5 (C-1<sup>o</sup>), 78.3 (C-3<sup>o</sup>), 77.8 (C-3<sup>o</sup>), 75.7 (C-4<sup>o</sup>), 75.5 (C-4<sup>o</sup>), 75.4 (C-3<sup>o</sup>), 75.1 (C-4<sup>o</sup>), 74.9 (CH<sub>2</sub>Ph), 74.7 (CH<sub>2</sub>Ph), 73.7 (C-3<sup>o</sup>), 73.5 (CH<sub>2</sub>Ph), 73.3 (CH<sub>2</sub>Ph), 72.8 (C-4<sup>o</sup>), 72.5 (CH<sub>2</sub> pMBn), 71.1 (C-5<sup>o</sup>), 70.9 (C-2<sup>o</sup>), 69.6 (C-5<sup>o</sup>), 69.3 (C-5<sup>o</sup>), 69.1 (C-5<sup>o</sup>), 68.8 (C-2<sup>o</sup>), 68.7 (C-a), 68.2 (C-2<sup>o</sup>), 62.8 (C-2<sup>o</sup>), 62.1 (C-6<sup>o</sup>), 61.6 (C-6<sup>o</sup>), 55.2 (CH<sub>3</sub> pMBn), 51.9 (CH<sub>3</sub> COOMe), 51.7 (CH<sub>3</sub> COOMe), 20.8, 20.7, 20.6 (CH<sub>3</sub> OAc) ppm; IR (thin film)  $\nu$  = 3083, 3058, 3024 (C-H<sub>arom</sub>), 2949, 2933, 2866, 2833 (C-H<sub>alk</sub>), 2109 (N<sub>2</sub>), 1740 (C=O), 1613, 1514, 1455, 1438, 1372, 1295, 1245 cm<sup>-1</sup>; elemental analysis: calcd (%) for C<sub>30</sub>H<sub>46</sub>N<sub>4</sub>O<sub>13</sub> (1461.5 g mol<sup>-1</sup>): C 59.99, H 5.79, N 5.75, O 28.46; found: C 60.22, H 5.76, N 5.66, O 28.16.

Allyl [(methyl 2-O-acetyl-3-O-benzyl- $\alpha$ -L-idopyranosyluronate)-(1 $\rightarrow$ 4)-O-(6-O-acetyl-2-azido-3-O-benzyl-2-deoxy- $\alpha$ -D-glucopyranoside)-(1 $\rightarrow$ 4)-O-(methyl 2-O-acetyl-3-O-benzyl- $\alpha$ -L-idopyranosyluronate)]-(1 $\rightarrow$ 4)-O-(6-O-acetyl-2-azido-3-O-benzyl-2-deoxy- $\alpha$ -D-glucopyranoside) (8): Tetrasaccharide 7 (465 mg, 0.32 mmol) was treated with DDQ (144 mg, 0.63 mmol, 2 equiv) in wet CH<sub>2</sub>Cl<sub>2</sub> (10 mL) as described for the preparation of 5. Flash chromatography (silica gel, toluene/AcOEt (8:2 $\rightarrow$ 6:4)) gave compound 8 (370 mg, 81%) along with unreacted starting material (44 mg, 9%): <sup>1</sup>H NMR (250 MHz, CDCl<sub>3</sub>):  $\delta$  = 7.40–7.25 (m, 20H; Ph), 5.95 (dddd, J<sub>1,2</sub> = 17.0, J<sub>1,3</sub> = 10.5, J<sub>1,4</sub> = 6.0, J<sub>1,5</sub> = 5.0 Hz, 1H; H-b), 5.36 (dq, J<sub>2,3</sub> = 17.0, J<sub>2,4</sub> = J<sub>2,5</sub> = 1.5 Hz, 1H; H-c), 5.27 (dq, J<sub>3,4</sub> = 10.5, J<sub>3,5</sub> = J<sub>3,6</sub> = 1.5 Hz, 1H; H-cc), 5.20 (d, J<sub>3,4</sub> = 3.5 Hz, 1H; H-1<sup>o</sup>),

(brd, J<sub>3,4</sub> = 1.0 Hz, 1H; H-1<sup>o</sup>), 4.95 (d, J<sub>3,4</sub> = 3.5 Hz, 1H; H-1<sup>o</sup>), 4.93 (d, J<sub>3,4</sub> = 2.5 Hz, 1H; H-5<sup>o</sup>), 4.92 (d, J<sub>3,4</sub> = 3.5 Hz, 1H; H-1<sup>o</sup>), 4.93 (m, 1H; H-2<sup>o</sup>), 4.83 (d, J<sub>3,4</sub> = 2.5 Hz, 1H; H-5<sup>o</sup>), 4.82 (d, J = 10.5 Hz, 1H; CH<sub>2</sub>Ph), 4.77 (d, J = 11.0 Hz, 2H; CH<sub>2</sub>Ph), 4.75 (t, J<sub>3,4</sub> = J<sub>3,5</sub> = 3.5 Hz, 1H; H-2<sup>o</sup>), 4.71 (d, J = 12.0 Hz, 1H; CH<sub>2</sub>Ph), 4.70 (d, J = 11.0 Hz, 1H; CH<sub>2</sub>Ph), 4.69 (d, J = 11.0 Hz, 1H; CH<sub>2</sub>Ph), 4.62 (d, J = 12.0 Hz, 1H; CH<sub>2</sub>Ph), 4.62 (d, J = 10.5 Hz, 1H; CH<sub>2</sub>Ph), 4.41 (brd, J<sub>4,5</sub> = 12.0, 2H; H-6<sup>o</sup>, H-6<sup>o</sup>), 4.26 (brd, J<sub>4,5</sub> = 12.0 Hz, 1H; H-6<sup>o</sup>), 4.21 (ddt, J<sub>4,5</sub> = 13.0, J<sub>4,6</sub> = 5.0, J<sub>4,7</sub> = 1.5 Hz, 1H; H-a), 4.17 (dd, J<sub>4,5</sub> = 12.0, J<sub>4,6</sub> = 2.5 Hz, 1H; H-6<sup>o</sup>), 4.05 (ddt, J<sub>4,5</sub> = 13.0, J<sub>4,6</sub> = 6.0, J<sub>4,7</sub> = 1.5 Hz, 1H; H-a), 4.03–3.96 (m, 2H), 3.96 (m, 6H), 3.78 (brd, J<sub>3,4</sub> = 9.5 Hz, 1H; H-5<sup>o</sup>), 3.73–3.58 (m, 3H), 3.48 (s, 3H; COOMe), 3.42 (s, 3H; COOMe), 3.39 (dd, J<sub>3,4</sub> = 10.0, J<sub>3,5</sub> = 3.5 Hz, 1H; H-2<sup>o</sup>), 3.28 (dd, J<sub>3,4</sub> = 10.0, J<sub>3,5</sub> = 3.5 Hz, 1H; H-2<sup>o</sup>), 2.59 (d, J = 10.5 Hz, 1H; OH), 2.11 (s, 3H; CH<sub>3</sub> OAc), 2.07 (s, 3H; CH<sub>3</sub> OAc), 2.06 (s, 3H; CH<sub>3</sub> OAc), 2.05 (s, 3H; CH<sub>3</sub> OAc) ppm; IR (thin film)  $\nu$  = 3470 (O-H<sub>arom</sub>), 3088, 3064, 3033 (C-H<sub>arom</sub>), 2978, 2955, 2932, 2908, 2869 (C-H<sub>alk</sub>), 2111 (N<sub>2</sub>), 1749, 1733 (C=O), 1500, 1498, 1455, 1432, 1373, 1304, 1244 cm<sup>-1</sup>; ESI HRMS: m/z calcd for C<sub>30</sub>H<sub>46</sub>N<sub>4</sub>O<sub>13</sub>Na [M + Na]<sup>+</sup>: 1363.4758; found: 1363.4758.

[(methyl 2-O-acetyl-3-O-benzyl-4-O-(4-methoxybenzyl)- $\alpha$ -L-idopyranosyluronate)-(1 $\rightarrow$ 4)-O-(6-O-acetyl-2-azido-3-O-benzyl-2-deoxy- $\alpha$ -D-glucopyranoside)-(1 $\rightarrow$ 4)-O-(methyl 2-O-acetyl-3-O-benzyl- $\alpha$ -L-idopyranosyluronate)]-(1 $\rightarrow$ 4)-O-(6-O-acetyl-2-azido-3-O-benzyl-2-deoxy- $\alpha$ -D-glucopyranoside) trichloroacetimidate (9): Tetrasaccharide 7 (322 mg, 0.22 mmol) was treated first with [I<sup>13</sup>C<sub>6</sub>H<sub>4</sub>(MePh)<sub>2</sub>]<sub>2</sub>PF<sub>6</sub> (3 mg, 3.5  $\mu$ mol, 1.6 mol%) in THF (3.5 mL) and then with HgO (69 mg, 0.26 mmol, 1.2 equiv) and HgCl<sub>2</sub> (67 mg, 0.24 mmol, 1.1 equiv) as described for the preparation of 6. Flash chromatography (silica gel, toluene/AcOEt (70:30 $\rightarrow$ 55:45)) gave the anticipated hemiacetal (298 mg, 81%) whose <sup>1</sup>H NMR spectrum showed the disappearance of the allyl protons. ESI HRMS: m/z calcd for C<sub>30</sub>H<sub>46</sub>N<sub>4</sub>O<sub>13</sub>Na<sub>2</sub> [M + 2Na]<sup>2+</sup>: 733.2459; found: 733.2459. This compound was then dissolved in CH<sub>2</sub>Cl<sub>2</sub> (400  $\mu$ L); trichloroacetonitrile (125  $\mu$ L, 1.25 mmol, 6 equiv) was then added, followed by potassium carbonate (51 mg, 0.42 mmol, 2 equiv). After 1 h stirring at room temperature the reaction mixture was directly applied to the top of a flash chromatography column filled with silica gel and eluted (toluene/AcOEt (9:1 $\rightarrow$ 7:3) with 0.1% NEt<sub>3</sub>) to give imidate 9 (338 mg, 98%) as an oil (4:6) mixture: <sup>1</sup>H NMR (250 MHz, CDCl<sub>3</sub>):  $\delta$  = 8.76 (s, 0.4H; NH), 8.73 (s, 0.6H; NH), 7.40–7.24 (m, 20H; Ph), 7.13 (d, J = 8.5 Hz, 2H; Ph-OMe), 6.82 (d, J = 8.5 Hz, 2H; Ph-OMe), 6.39 (d, J<sub>3,4</sub> = 3.5 Hz, 0.4H; H-1<sup>o</sup>), 5.60 (d, J<sub>3,4</sub> = 8.5 Hz, 0.6H; H-1<sup>o</sup>), 5.28 (d, J<sub>3,4</sub> = 5.0 Hz, 1H; H-1<sup>o</sup>), 5.25 (d, J<sub>3,4</sub> = 3.5 Hz, 0.4H; H-1<sup>o</sup>), 5.22 (d, J<sub>3,4</sub> = 3.5 Hz, 0.6H; H-1<sup>o</sup>), 4.47 (d, J = 11.5 Hz, 1H; CH<sub>2</sub>PhOMe), 4.41 (d, J = 11.5 Hz, 1H; CH<sub>2</sub>PhOMe), 4.45 (brd, J<sub>4,5</sub> = 12.5, 0.4H; H-6<sup>o</sup>), 4.44 (brd, J<sub>4,5</sub> = 12.5, 0.6H; H-6<sup>o</sup>), 4.36 (brd, J<sub>4,5</sub> = 12.5 Hz, 1H; H-6<sup>o</sup>), 4.23 (dd, J<sub>4,5</sub> = 12.5, J<sub>4,6</sub> = 4.0 Hz, 0.6H; H-6<sup>o</sup>), 4.22 (dd, J<sub>4,5</sub> = 12.5, J<sub>4,6</sub> = 4.0 Hz, 0.4H; H-6<sup>o</sup>), 4.17 (dd, J<sub>4,5</sub> = 12.5, J<sub>4,6</sub> = 2.0 Hz, 1H; H-6<sup>o</sup>), 4.04–3.88 (m, 5H), 3.88–3.74 (m with s at  $\delta$  = 3.79, 5H; PhOMe), 3.74–3.58 (m, 3H), 3.35 (s, 3H; COOMe), 3.50 (s, 1.8H; COOMe), 3.49 (s, 1.2H; COOMe), 3.43 (t, J<sub>3,4</sub> = J<sub>3,5</sub> = 8.5 Hz, 0.6H; H-2<sup>o</sup>), 3.30 (dd, J<sub>3,4</sub> = 10.0, J<sub>3,5</sub> = 3.5 Hz, 1H; H-2<sup>o</sup>), 2.11 (s, 1.8H; CH<sub>3</sub> OAc), 2.10 (s, 1.2H; CH<sub>3</sub> OAc), 2.06 (s, 3H; CH<sub>3</sub> OAc), 2.04 (s, 6H; CH<sub>3</sub> OAc) ppm.

Allyl [(methyl 2-O-acetyl-3-O-benzyl-4-O-(4-methoxybenzyl)- $\alpha$ -L-idopyranosyluronate)-(1 $\rightarrow$ 4)-O-(6-O-acetyl-2-azido-3-O-benzyl-2-deoxy- $\alpha$ -D-glucopyranoside)-(1 $\rightarrow$ 4)-O-(methyl 2-O-acetyl-3-O-benzyl- $\alpha$ -L-idopyranosyluronate)]-(1 $\rightarrow$ 4)-O-(6-O-acetyl-2-azido-3-O-benzyl-2-deoxy- $\alpha$ -D-glucopyranoside) (10): Acceptor 8 (150 mg, 0.11 mmol) and imidate 6 (135 mg, 0.15 mmol, 1.3 equiv) were azeotropically dried with toluene and dissolved in CH<sub>2</sub>Cl<sub>2</sub> (360  $\mu$ L). Powdered 4-Å molecular sieves (330 mg) were added and the mixture was then stirred for 30 min at room temperature. The solution was cooled to -40°C and TBDMSOTf (0.1 M in CH<sub>2</sub>Cl<sub>2</sub>, 220  $\mu$ L, 0.022 mmol, 0.2 equiv) was added. The reaction was stirred at this temperature for 30 min and then the temperature was raised to 0°C over 30 min. The reaction was then quenched with a solution of NEt<sub>3</sub> (0.1 M in CH<sub>2</sub>Cl<sub>2</sub>, 63  $\mu$ L, 0.063 mmol, 0.6 equiv), directly applied to the top of a flash chromatography column filled with silica gel and eluted (CH<sub>2</sub>Cl<sub>2</sub>/AcOEt/petroleum ether (85:10:5 $\rightarrow$ 70:25:5)) to give hexasaccharide 10 (210 mg, 68%): [ $\alpha$ ]<sub>D</sub><sup>20</sup> = -17 (c = 1.50, CH<sub>2</sub>Cl<sub>2</sub>); <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  = 7.40–7.23 (m, 30H; Ph), 7.13 (d, J = 8.5 Hz, 2H;

Ph-OMe), 6.82 (d,  $J = 8.5$  Hz, 2H; Ph-OMe), 5.94 (dddd,  $J_{6,4} = 17.0$ ,  $J_{6,5} = 10.5$ ,  $J_{6,7} = 6.0$ ,  $J_{6,8} = 5.0$  Hz, 1H; H-b), 5.36 (dq,  $J_{4,3} = 17.0$ ,  $J_{4,5} = 10.5$ ,  $J_{4,6} = 1.5$  Hz, 1H; H-c), 5.30 (d,  $J_{12} = 5.0$  Hz, 1H; H-1<sup>D</sup>), 5.29 (d,  $J_{12} = 5.0$  Hz, 1H; H-1<sup>B</sup>), 5.27 (dq,  $J_{4,3} = 10.5$ ,  $J_{4,5} = 10.5$ ,  $J_{4,6} = 1.5$  Hz, 1H; H-c), 5.24 (d,  $J_{12} = 3.5$  Hz, 1H; H-1<sup>D</sup>), 4.96 (d,  $J_{12} = 3.5$  Hz, 1H; H-1<sup>B</sup>), 4.94 (d,  $J_{12} = 3.5$  Hz, 1H; H-1<sup>B</sup>), 4.93 (t,  $J_{2,3} = 3.5$  Hz, 1H; H-2<sup>D</sup>), 4.93 (d,  $J_{12} = 3.5$  Hz, 1H; H-1<sup>B</sup>), 4.90 (dd,  $J_{2,3} = 6.0$ ,  $J_{2,1} = 5.0$  Hz, 1H; H-2<sup>D</sup>), 4.90 (d,  $J = 10.5$  Hz, 1H; CH<sub>2</sub>Ph), 4.87 (t,  $J_{2,3} = 5.0$  Hz, 1H; H-2<sup>D</sup>), 4.84 (d,  $J = 10.5$  Hz, 1H; CH<sub>2</sub>Ph), 4.82 (d,  $J = 10.5$  Hz, 1H; CH<sub>2</sub>Ph), 4.77 (d,  $J = 11.5$  Hz, 1H; CH<sub>2</sub>Ph), 4.76 (d,  $J = 11.0$  Hz, 1H; CH<sub>2</sub>Ph), 4.72 (d,  $J = 11.0$  Hz, 1H; CH<sub>2</sub>Ph), 4.70 (d,  $J_{14} = 4.0$  Hz, 1H; H-5<sup>D</sup>), 4.69 (d,  $J = 10.5$  Hz, 1H; CH<sub>2</sub>Ph), 4.69 (d,  $J = 11.5$  Hz, 1H; CH<sub>2</sub>Ph), 4.69 (d,  $J = 12.0$  Hz, 1H; CH<sub>2</sub>Ph), 4.65 (d,  $J = 10.5$  Hz, 1H; CH<sub>2</sub>Ph), 4.65 (d,  $J = 12.0$  Hz, 1H; CH<sub>2</sub>Ph), 4.64 (d,  $J = 10.5$  Hz, 1H; CH<sub>2</sub>Ph), 4.59 (d,  $J_{14} = 5.0$  Hz, 1H; H-5<sup>D</sup>), 4.56 (d,  $J_{14} = 4.5$  Hz, 1H; H-5<sup>D</sup>), 4.47 (d,  $J = 11.5$  Hz, 1H; CH<sub>2</sub>PhOMe), 4.43 (d,  $J = 11.5$  Hz, 1H; CH<sub>2</sub>PhOMe), 4.40 (brd,  $J_{6,4} = 11.5$  Hz, 1H; H-6<sup>A</sup>), 4.37 (dd,  $J_{6,4} = 12.0$ ,  $J_{6,5} = 2.0$  Hz, 1H; H-6<sup>A</sup>), 4.34 (dd,  $J_{6,4} = 12.5$ ,  $J_{6,5} = 2.0$  Hz, 1H; H-6<sup>A</sup>), 4.25 (dd,  $J_{6,4} = 11.5$ ,  $J_{6,5} = 2.5$  Hz, 1H; H-6<sup>A</sup>), 4.20 (ddt,  $J_{6,4} = 13.0$ ,  $J_{6,5} = 5.0$ ,  $J_{6,7} = 1.5$  Hz, 1H; H-a), 4.19 (dd,  $J_{6,4} = 12.0$ ,  $J_{6,5} = 2.5$  Hz, 1H; H-6<sup>A</sup>), 4.16 (dd,  $J_{6,4} = 12.5$ ,  $J_{6,5} = 2.5$  Hz, 1H; H-6<sup>A</sup>), 4.05 (ddt,  $J_{6,4} = 13.0$ ,  $J_{6,5} = 6.0$ ,  $J_{6,7} = 1.5$  Hz, 1H; H-a), 3.99 (t,  $J_{4,3} = 4.0$  Hz, 1H; H-4<sup>D</sup>), 3.98 (dd,  $J_{4,3} = 5.5$ ,  $J_{4,5} = 4.5$  Hz, 1H; H-4<sup>D</sup>), 3.95 (t,  $J_{4,3} = 4.0$  Hz, 1H; H-4<sup>D</sup>), 3.92 (dd,  $J_{4,3} = 5.5$ ,  $J_{4,5} = 6.0$  Hz, 1H; H-3<sup>D</sup>), 3.90 (dd,  $J_{4,3} = 4.0$ ,  $J_{4,5} = 3.5$  Hz, 1H; H-3<sup>D</sup>), 3.89-3.85 (m, 3H; H-3<sup>A</sup>, H-4<sup>A</sup>, H-5<sup>A</sup>), 3.87 (dd,  $J_{4,3} = 10.0$ ,  $J_{4,5} = 9.0$  Hz, 1H; H-4<sup>D</sup>), 3.83 (brt,  $J_{4,3} = 9.0$  Hz, 1H; H-5<sup>A</sup>), 3.82-3.78 (m with s at  $\delta = 3.80$ , 5H; H-4<sup>D</sup>, H-5<sup>D</sup>, PhOMe), 3.77 (t,  $J_{2,3} = 3.5$  Hz, 1H; H-3<sup>D</sup>), 3.66 (dd,  $J_{12} = 10.0$ ,  $J_{14} = 9.0$  Hz, 1H; H-3<sup>D</sup>), 3.62 (dd,  $J_{12} = 10.0$ ,  $J_{14} = 9.0$  Hz, 1H; H-3<sup>D</sup>), 3.57 (s, 3H; COOMe), 3.53 (s, 3H; COOMe), 3.48 (s, 3H; COOMe), 3.39 (dd,  $J_{2,3} = 10.0$ ,  $J_{2,1} = 3.5$  Hz, 1H; H-2<sup>D</sup>), 3.29 (dd,  $J_{2,3} = 10.0$ ,  $J_{2,1} = 3.5$  Hz, 1H; H-2<sup>D</sup>), 3.27 (dd,  $J_{2,3} = 10.0$ ,  $J_{2,1} = 3.5$  Hz, 1H; H-2<sup>D</sup>), 2.12 (s, 3H; CH<sub>3</sub> OAc), 2.09 (s, 3H; CH<sub>3</sub> OAc), 2.06 (s, 3H; CH<sub>3</sub> OAc), 2.03 (s, 3H; CH<sub>3</sub> OAc), 2.02 (s, 3H; CH<sub>3</sub> OAc), 2.01 (s, 3H; CH<sub>3</sub> OAc) ppm; <sup>13</sup>C NMR (62.5 MHz, CDCl<sub>3</sub>):  $\delta = 170.6$ , 170.5 (2C), 169.9 (2C), 169.8, 169.7, 169.5, 169.3 (C=O), 159.5 (C-OMe pMBn), 137.8, 137.7, 137.6, 137.4, 137.3 (C<sub>aromatic</sub>), 133.1 (C-b), 129.6, 129.3, 128.5, 128.4, 128.2, 128.0, 127.9, 127.8, 127.7 (C<sub>aromatic</sub>), 118.3 (C-c), 113.7 (C<sub>p</sub> pMBn), 98.0 (C-1<sup>B</sup>), 97.8 (C-1<sup>A</sup>), 97.4 (C-1<sup>B</sup>), 96.5 (C-1<sup>A</sup>), 78.3 (C-3<sup>A</sup>), 77.8 (C-3<sup>B</sup>), 77.7 (C-3<sup>C</sup>), 75.7 (C-4<sup>A</sup>), 75.6 (C-4<sup>B</sup>), 75.5 (C-4<sup>C</sup>), 75.4 (C-3<sup>A</sup>), 75.3 (C-3<sup>B</sup>), 75.2 (C-3<sup>C</sup>), 74.9 (CH<sub>2</sub>Ph), 74.7 (CH<sub>2</sub>Ph), 74.0 (CH<sub>2</sub>Ph), 73.6 (C-3<sup>D</sup>, CH<sub>2</sub>Ph), 73.4 (CH<sub>2</sub>Ph), 73.3 (CH<sub>2</sub>Ph), 72.6 (C-4<sup>D</sup>, CH<sub>2</sub> pMBn), 71.3 (C-5<sup>B</sup>), 71.1 (C-2<sup>B</sup>), 69.7 (C-5<sup>B</sup>), 69.6 (C-5<sup>C</sup>), 69.3 (C-5<sup>D</sup>), 69.2 (C-5<sup>A</sup>), 68.8 (C-2<sup>B</sup>), 68.7 (C-a), 63.3 (C-2<sup>A</sup>), 62.9 (C-2<sup>B</sup>), 62.7 (C-2<sup>C</sup>), 62.1 (C-6<sup>B</sup>), 61.6 (C-6<sup>C</sup>, C-6<sup>D</sup>), 55.2 (CH<sub>3</sub> pMBn), 51.9 (2 $\times$ CH<sub>3</sub> COOMe), 51.7 (CH<sub>3</sub> COOMe), 20.8, 20.7 (CH<sub>3</sub> OAc) ppm; IR (thin film):  $\nu = 3100$ , 3066, 3033, 2999 (C-H<sub>aromatic</sub>), 2953, 2928, 2870 (C-H<sub>aliphatic</sub>), 2109 (N<sub>2</sub>), 1742 (C=O), 1613, 1515, 1497, 1455, 1438, 1371, 1304, 1235 cm<sup>-1</sup>; elemental analysis: calcd (%) for C<sub>32</sub>H<sub>41</sub>N<sub>3</sub>O<sub>14</sub> (2103.2 g mol<sup>-1</sup>): C 59.39, H 5.70, N 5.99, O 28.91; found: C 59.41, H 5.88, N 5.97.

Allyl [(methyl 2-*O*-acetyl-3-*O*-benzyl-4-*O*-(4-methoxybenzyl)- $\alpha$ -L-idopyranosyluronate)-(1 $\rightarrow$ 4)-*O*-(6-*O*-acetyl-2-azido-3-*O*-benzyl-2-deoxy- $\alpha$ -D-glucopyranoside)-(1 $\rightarrow$ 4)-*O*-(methyl 2-*O*-acetyl-3-*O*-benzyl- $\alpha$ -L-idopyranosyluronate)-(1 $\rightarrow$ 4)-*O*-(6-*O*-acetyl-2-azido-3-*O*-benzyl-2-deoxy- $\alpha$ -D-glucopyranoside)-(1 $\rightarrow$ 4)-*O*-(methyl 2-*O*-acetyl-3-*O*-benzyl- $\alpha$ -L-idopyranosyluronate)-(1 $\rightarrow$ 4)-*O*-(6-*O*-acetyl-2-azido-3-*O*-benzyl-2-deoxy- $\alpha$ -D-glucopyranoside)-(1 $\rightarrow$ 4)-*O*-(6-*O*-acetyl-2-azido-3-*O*-benzyl-2-deoxy- $\alpha$ -D-glucopyranoside)] (II): Acceptor 8 (184 mg, 0.14 mmol) and imidate 9 (298 mg, 0.19 mmol, 1.3 equiv) were azeotropically dried with toluene and dissolved in CH<sub>2</sub>Cl<sub>2</sub> (510  $\mu$ L). Powdered 4-Å molecular sieves (420 mg) were added and the mixture was then stirred for 30 min at room temperature. The solution was cooled to -40 °C and TBDMSOTf (0.1 mL in CH<sub>2</sub>Cl<sub>2</sub>, 288  $\mu$ L, 0.029 mmol, 0.2 equiv) was added. The reaction was stirred at this temperature for 30 min and then the temperature was raised to 0 °C over 30 min. The reaction was then quenched with a solution of NEt<sub>3</sub> (0.1 M in CH<sub>2</sub>Cl<sub>2</sub>, 82  $\mu$ L, 0.082 mmol, 0.6 equiv), directly applied to the top of a flash chromatography column filled with silica gel and eluted (CH<sub>2</sub>Cl<sub>2</sub>/AcOEt/petroleum ether (85:10:5) to give octasaccharide II (343 mg, 93%):  $[\alpha]_D^{25} = -19$  ( $c = 2.53$ , CH<sub>2</sub>Cl<sub>2</sub>); <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>):  $\delta = 7.38$ -7.26 (m, 40H; Ph), 7.13 (d,  $J = 8.5$  Hz, 2H; Ph-OMe),

6.83 (d,  $J = 8.5$  Hz, 2H; Ph-OMe), 5.94 (dddd,  $J_{6,4} = 17.0$ ,  $J_{6,5} = 10.5$ ,  $J_{6,7} = 6.0$ ,  $J_{6,8} = 5.0$  Hz, 1H; H-b), 5.36 (dq,  $J_{4,3} = 17.0$ ,  $J_{4,5} = 10.5$ ,  $J_{4,6} = 1.5$  Hz, 1H; H-c), 5.31 (d,  $J_{12} = 5.0$  Hz, 1H; H-1<sup>D</sup> or H-1<sup>B</sup>), 5.30 (d,  $J_{12} = 6.0$  Hz, 2H; H-1<sup>A</sup>, H-1<sup>D</sup> or H-1<sup>B</sup>), 5.26 (dq,  $J_{4,3} = 10.5$ ,  $J_{4,5} = 10.5$ ,  $J_{4,6} = 1.5$  Hz, 1H; H-c), 5.24 (d,  $J_{12} = 3.5$  Hz, 1H; H-1<sup>B</sup>), 4.98 (d,  $J_{12} = 3.5$  Hz, 1H; H-1<sup>B</sup>), 4.97 (d,  $J_{12} = 3.5$  Hz, 1H; H-1<sup>B</sup>), 4.94 (d,  $J_{12} = 3.5$  Hz, 1H; H-1<sup>A</sup>), 4.93 (t,  $J_{2,3} = 3.5$  Hz, 1H; H-2<sup>D</sup>), 4.93 (d,  $J_{12} = 3.5$  Hz, 1H; H-1<sup>B</sup>), 4.91 (d,  $J = 10.0$  Hz, 1H; CH<sub>2</sub>Ph), 4.90 (m, 2H; H-2<sup>D</sup>, H-2<sup>B</sup>), 4.88 (t,  $J_{2,3} = 6.0$  Hz, 1H; H-2<sup>D</sup>), 4.86 (d,  $J = 10.0$  Hz, 1H; CH<sub>2</sub>Ph), 4.85 (d,  $J = 10.0$  Hz, 1H; CH<sub>2</sub>Ph), 4.83 (d,  $J = 11.0$  Hz, 1H; CH<sub>2</sub>Ph), 4.78 (d,  $J = 11.0$  Hz, 1H; CH<sub>2</sub>Ph), 4.76 (d,  $J = 11.0$  Hz, 1H; CH<sub>2</sub>Ph), 4.74 (s, 2H; CH<sub>2</sub>Ph), 4.73 (d,  $J_{14} = 4.0$  Hz, 1H; H-5<sup>D</sup>), 4.70 (brd,  $J = 11.0$  Hz, 3H; 3 $\times$ CH<sub>2</sub>Ph), 4.69 (d,  $J = 12.0$  Hz, 1H; CH<sub>2</sub>Ph), 4.67 (d,  $J = 10.0$  Hz, 1H; CH<sub>2</sub>Ph), 4.66 (d,  $J = 10.0$  Hz, 1H; CH<sub>2</sub>Ph), 4.65 (d,  $J = 12.0$  Hz, 1H; CH<sub>2</sub>Ph), 4.64 (d,  $J = 10.0$  Hz, 1H; CH<sub>2</sub>Ph), 4.59 (d,  $J_{14} = 5.0$  Hz, 1H; H-5<sup>B</sup>), 4.58 (d,  $J_{14} = 5.0$  Hz, 1H; H-5<sup>D</sup> or H-5<sup>B</sup>), 4.56 (d,  $J_{14} = 5.0$  Hz, 1H; H-5<sup>D</sup> or H-5<sup>B</sup>), 4.47 (d,  $J = 11.5$  Hz, 1H; CH<sub>2</sub>PhOMe), 4.43 (d,  $J = 11.5$  Hz, 1H; CH<sub>2</sub>PhOMe), 4.40 (d,  $J_{6,4} = 12.0$  Hz, 1H; H-6<sup>A</sup>), 4.37 (brd,  $J_{6,4} = 11.0$ , 2H; H-6<sup>B</sup>, H-6<sup>C</sup>), 4.34 (dd,  $J_{6,4} = 12.0$ ,  $J_{6,5} = 1.5$  Hz, 1H; H-6<sup>B</sup>, H-6<sup>C</sup>), 4.25 (dd,  $J_{6,4} = 12.0$ ,  $J_{6,5} = 2.5$  Hz, 1H; H-6<sup>A</sup>), 4.20 (ddt,  $J_{6,4} = 13.0$ ,  $J_{6,5} = 5.0$ ,  $J_{6,7} = 1.5$  Hz, 1H; H-a), 4.22-4.16 (m, 2H; H-6<sup>B</sup>, H-6<sup>C</sup>), 4.17 (dd,  $J_{6,4} = 12.0$ ,  $J_{6,5} = 3.0$  Hz, 1H; H-6<sup>B</sup>, H-6<sup>C</sup>), 4.05 (ddt,  $J_{6,4} = 13.0$ ,  $J_{6,5} = 6.0$ ,  $J_{6,7} = 1.5$  Hz, 1H; H-a), 4.20-3.97 (m, 3H; H-4<sup>D</sup>, H-4<sup>A</sup>, H-4<sup>B</sup>), 3.96 (dd,  $J_{4,3} = 10.0$ ,  $J_{4,5} = 9.0$  Hz, 1H; H-4<sup>D</sup>), 3.95-3.90 (m, 3H; H-3<sup>A</sup>, H-3<sup>B</sup>, H-3<sup>C</sup>), 3.90-3.87 (m, 3H; H-3<sup>A</sup>, H-4<sup>A</sup>, H-5<sup>A</sup>), 3.88 (dd,  $J_{4,3} = 10.0$ ,  $J_{4,5} = 9.0$ , 2H; H-4<sup>D</sup>, H-4<sup>C</sup>), 3.87-3.79 (m with s at  $\delta = 3.80$ , 7H; H-4<sup>D</sup>, H-5<sup>D</sup>, H-5<sup>A</sup>, H-5<sup>B</sup>, H-5<sup>C</sup>, PhOMe), 3.77 (t,  $J_{2,3} = 3.5$  Hz, 1H; H-3<sup>D</sup>), 3.67 (dd,  $J_{12} = 10.0$ ,  $J_{14} = 9.0$  Hz, 1H; H-3<sup>D</sup>), 3.64 (dd,  $J_{12} = 10.0$ ,  $J_{14} = 9.0$  Hz, 1H; H-3<sup>D</sup>), 3.63 (dd,  $J_{12} = 10.0$ ,  $J_{14} = 9.0$  Hz, 1H; H-3<sup>D</sup>), 3.57 (s, 3H; COOMe), 3.54 (s, 3H; COOMe), 3.53 (s, 3H; COOMe), 3.48 (s, 3H; COOMe), 3.39 (dd,  $J_{2,3} = 10.0$ ,  $J_{2,1} = 3.5$  Hz, 1H; H-2<sup>D</sup>), 3.30 (dd,  $J_{2,3} = 10.0$ ,  $J_{2,1} = 3.5$  Hz, 1H; H-2<sup>D</sup>), 3.28 (dd,  $J_{2,3} = 10.0$ ,  $J_{2,1} = 3.5$  Hz, 1H; H-2<sup>D</sup>), 3.27 (dd,  $J_{2,3} = 10.0$ ,  $J_{2,1} = 3.5$  Hz, 1H; H-2<sup>D</sup>), 2.12 (s, 3H; CH<sub>3</sub> OAc), 2.10 (s, 3H; CH<sub>3</sub> OAc), 2.09 (s, 3H; CH<sub>3</sub> OAc), 2.06 (s, 3H; CH<sub>3</sub> OAc), 2.03 (s, 6H; CH<sub>3</sub> OAc), 2.02 (s, 6H; CH<sub>3</sub> OAc) ppm; <sup>13</sup>C NMR (100.6 MHz, CDCl<sub>3</sub>):  $\delta = 170.6$ , 170.5 (3C), 169.9 (2C), 169.7, 169.6 (2C), 169.5, 169.4, 169.2 (C=O), 158.9 (C-OMe pMBn), 137.8, 137.6, 137.5, 137.4, 137.3, 137.2 (C<sub>aromatic</sub>), 132.9 (C-b), 129.6, 129.2, 129.9, 128.4, 128.3, 128.2, 128.1, 128.0, 127.9, 127.8, 127.7, 127.6, 127.5, 127.4 (C<sub>aromatic</sub>), 118.3 (C-c), 113.7 (C<sub>p</sub> pMBn), 97.9 (C-1<sup>B</sup>), 97.8 (C-1<sup>A</sup>), 97.6 (C-1<sup>B</sup>), 97.4 (C-1<sup>A</sup>), 97.1 (C-1<sup>B</sup>), 96.5 (C-1<sup>A</sup>), 78.3 (C-3<sup>A</sup>), 78.2 (C-3<sup>B</sup>), 78.1 (C-3<sup>C</sup>), 75.6, 75.5, 75.4, 75.3, 75.2 (C-4<sup>A</sup>, C-4<sup>B</sup>, C-4<sup>C</sup>, C-4<sup>D</sup> or C-4<sup>E</sup>, C-4<sup>F</sup>, C-4<sup>G</sup>), 74.8 (C-3<sup>D</sup>, CH<sub>2</sub>Ph), 74.7 (2CH<sub>2</sub>Ph), 74.6 (CH<sub>2</sub>Ph), 74.0 (CH<sub>2</sub>Ph), 73.9 (CH<sub>2</sub>Ph), 73.5 (C-3<sup>D</sup>, C-3<sup>E</sup>, CH<sub>2</sub>Ph), 73.3 (C-4<sup>D</sup> or C-4<sup>E</sup>, CH<sub>2</sub>Ph), 72.5 (C-4<sup>F</sup>, CH<sub>2</sub> pMBn), 71.2 (C-5<sup>B</sup>), 71.1 (C-2<sup>B</sup>), 70.5 (C-5<sup>B</sup> or C-5<sup>C</sup>), 70.2 (C-2<sup>C</sup> or C-2<sup>D</sup>, C-5<sup>D</sup> or C-5<sup>E</sup>), 70.0 (C-2<sup>D</sup> or C-2<sup>E</sup>), 69.6-69.5 (3C, C-5<sup>F</sup>, C-5<sup>G</sup>, C-5<sup>H</sup>), 69.2 (C-5<sup>B</sup>), 69.1 (C-5<sup>A</sup>), 68.7 (C-2<sup>B</sup>), 68.6 (C-a), 63.3 (C-2<sup>A</sup>), 62.8, 62.7, 62.6 (C-2<sup>C</sup>, C-2<sup>D</sup>, C-2<sup>E</sup>), 62.0 (C-6<sup>A</sup>), 61.5 (C-6<sup>B</sup>, C-6<sup>C</sup>, C-6<sup>D</sup>), 55.2 (CH<sub>3</sub> pMBn), 51.9 (3 $\times$ CH<sub>3</sub> COOMe), 51.7 (CH<sub>3</sub> COOMe), 20.8, 20.7, 20.6 (CH<sub>3</sub> OAc) ppm; IR (thin film):  $\nu = 3097$ , 3065, 3034 (C-H<sub>aromatic</sub>), 2955, 2932, 2877 (C-H<sub>aliphatic</sub>), 2115 (N<sub>2</sub>), 1749 (C=O), 1613, 1515, 1497, 1455, 1439, 1373, 1304, 1250 cm<sup>-1</sup>; elemental analysis: calcd (%) for C<sub>33</sub>H<sub>43</sub>N<sub>3</sub>O<sub>14</sub> (2744.8 g mol<sup>-1</sup>): C 59.08, H 5.66, N 6.12, O 29.15; found: C 59.02, H 5.43, N 5.96, O 29.13.

Allyl [(methyl 3-*O*-benzyl-4-*O*-(4-methoxybenzyl)-2-*O*-sulfonato- $\alpha$ -L-idopyranosyluronate)-(1 $\rightarrow$ 4)-*O*-(3-*O*-benzyl-2-deoxy-2-sulfamino-6-*O*-sulfonato- $\alpha$ -D-glucopyranoside)-(1 $\rightarrow$ 4)-*O*-(methyl 3-*O*-benzyl-2-*O*-sulfonato- $\alpha$ -L-idopyranosyluronate)-(1 $\rightarrow$ 4)-*O*-(3-*O*-benzyl-2-deoxy-2-sulfamino-6-*O*-sulfonato- $\alpha$ -D-glucopyranoside)-(1 $\rightarrow$ 4)-*O*-(3-*O*-benzyl-2-deoxy-2-sulfamino-6-*O*-sulfonato- $\alpha$ -D-glucopyranoside)] (18): Tetrasaccharide 4 (236 mg, 0.161 mmol) was stirred for 1 h with K<sub>2</sub>CO<sub>3</sub> (11 mg, 0.08 mmol, 0.5 equiv) in MeOH (1 mL). The mixture was then neutralised with BioRad AG50W-X8200 (H<sup>+</sup>) resin, filtered and concentrated. Flash chromatography (silica gel, toluene/AcOEt (1:1-35:65)) gave the deacetylated compound 12 (192 mg, 93%). A portion of this material (52 mg, 40  $\mu$ mol) was dissolved in MeOH (400  $\mu$ L), protected from light and stirred with propane-1,3-dithiol (80  $\mu$ L, 0.8 mmol, 20 equiv) and triethylamine (111  $\mu$ L, 0.8 mmol, 20 equiv) for 2 d. The mixture was then concentrated and purified by silica gel flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH (8:2-9:1) with 0.1% NEt<sub>3</sub>) to give 15 (46 mg, 92%): <sup>13</sup>C NMR (62.5 MHz, CDCl<sub>3</sub>):  $\delta = 170.1$ , 169.8 (C=O), 169.2 (C-OMe pMBn), 138.6, 138.4, 137.5 (C<sub>aromatic</sub>), 133.9 (C-b), 129.8, 128.8, 128.4, 128.2, 128.0, 127.8,

127.4, 127.0 ( $C_{\text{arom}}$ ), 117.2 (C-c), 114.0 ( $C_{\text{arom}}$  pMBn), 101.2, 100.9, 98.7, 96.7 (4×C-1), 81.9, 81.6, 75.6, 75.2, 74.8 ( $CH_2Ph$ ), 74.7, 74.6, 73.7, 72.6 ( $CH_2Ph$ ), 72.5, 72.0, 71.0, 69.7, 69.2, 68.5, 68.3 ( $CH_2Ph$ ), 67.9, 62.1 (2×C-6), 55.9 (OCH<sub>3</sub>), 55.1 (OCH<sub>3</sub>), 51.8 (OCH<sub>3</sub>) ppm; ESI HRMS:  $m/z$  calcd for C<sub>45</sub>H<sub>40</sub>N<sub>2</sub>O<sub>2</sub> [ $M+H$ ]<sup>+</sup>: 1241.5281; found: 1241.5304.

Sulfur trioxide pyridine complex (200 mg, 1.25 mmol, 30 equiv) was added to a solution of 15 (52 mg, 42 μmol) in pyridine (2.5 mL). This mixture was protected from light, stirred for 24 h at room temperature and then heated for 24 h at 55 °C. MeOH (800 μL, 13.8 mmol) and NEt<sub>3</sub> (320 μL, 2.3 mmol) were then added to quench the reaction. The resulting mixture was stirred for 1 h at room temperature and successively purified by Sephadex LH-20 chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH (1:1)) and RP-18 flash chromatography (5 mM AcOH-NEt<sub>3</sub> (pH 7.0)/MeOH (100:0–60:40)), followed by ion exchange on BioRad AG50W-X8 200 (Na<sup>+</sup>, 5 mL) resin after removal of AcOH-NEt<sub>3</sub> salts by lyophilisation (2×1 mL H<sub>2</sub>O). Compound 18 (57 mg, 73%) was thus obtained as a hexa-sodium salt (98% pure by HPLC analysis): <sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O): δ = 7.59–7.41 (m, 9 H; Ph), 7.41–7.33 (m, 6 H; Ph), 7.30–7.21 (m, 5 H; Ph), 7.11 (d,  $J=8.5$  Hz, 2 H; Ph-OMe), 7.00 (d,  $J=8.5$  Hz, 2 H; Ph-OMe), 6.07 (ddd,  $J_{ax}=17.0$ ,  $J_{ax}-J_{ax'}=10.5$ ,  $J_{ax'}=6.0$ ,  $J_{ax}-J_{ax'}=5.0$  Hz, 1 H; H-b), 5.44 (dq,  $J_{ax}=17.0$ ,  $J_{ax}-J_{ax'}=10.5$ ,  $J_{ax'}=6.0$ ,  $J_{ax}-J_{ax'}=5.0$  Hz, 1 H; H-c), 5.37 (brs, 1 H; H-1<sup>a</sup>), 5.31 (dq,  $J_{ax}=10.5$ ,  $J_{ax}-J_{ax'}=10.5$ ,  $J_{ax'}=6.0$ ,  $J_{ax}-J_{ax'}=5.0$  Hz, 1 H; H-c), 5.27 (d,  $J_{12}=3.5$  Hz, 1 H; H-1<sup>b</sup>), 5.20 (d,  $J_{12}=3.5$  Hz, 1 H; H-1<sup>c</sup>), 5.13 (brs, 1 H; H-1<sup>b</sup>), 4.95 (d,  $J_{14}=1.5$  Hz, 1 H; H-5<sup>a</sup>), 4.90 (d,  $J=11.0$  Hz, 1 H; CH<sub>2</sub>Ph), 4.88 (d,  $J_{14}=2.0$  Hz, 1 H; H-5<sup>b</sup>), 4.80 (d,  $J=11.5$  Hz, 2 H; 2×CH<sub>2</sub>Ph), 4.79 (d,  $J=11.0$  Hz, 1 H; CH<sub>2</sub>Ph), 4.78 (d,  $J=10.5$  Hz, 1 H; CH<sub>2</sub>Ph), 4.59 (d,  $J=11.5$  Hz, 1 H; CH<sub>2</sub>Ph), 4.59–4.57 (m, 1 H; H-2<sup>a</sup>), 4.49 (d,  $J=12.0$  Hz, 1 H; CH<sub>2</sub>PhOMe), 4.48–4.45 (m, 1 H; H-2<sup>b</sup>), 4.44 (d,  $J=10.5$  Hz, 1 H; CH<sub>2</sub>Ph), 4.40 (brt,  $J_{13}=J_{13'}=3.0$  Hz, 1 H; H-3<sup>a</sup>), 4.38 (dd,  $J_{ax}=11.5$ ,  $J_{ax'}=2.0$  Hz, 1 H; H-6<sup>a</sup>), 4.33 (d,  $J=11.5$  Hz, 1 H; CH<sub>2</sub>Ph), 4.30 (dd,  $J_{pm}=13.0$ ,  $J_{12}=5.0$ ,  $J_{ax}-J_{ax'}=1.5$  Hz, 1 H; H-a), 4.29 (dd,  $J_{ax}=11.5$ ,  $J_{ax'}=5.0$  Hz, 1 H; H-6<sup>a</sup>), 4.26 (d,  $J=12.0$  Hz, 1 H; CH<sub>2</sub>PhOMe), 4.21 (brs, 1 H; H-6<sup>c</sup>, H-6<sup>b</sup>), 4.19–4.16 (m, 1 H; H-3<sup>b</sup>), 4.16–4.13 (m, 1 H; H-4<sup>a</sup>), 4.13 (ddt,  $J_{pm}=13.0$ ,  $J_{12}=6.0$ ,  $J_{ax}-J_{ax'}=1.5$  Hz, 1 H; H-a), 4.10 (ddd,  $J_{14}=9.5$ ,  $J_{13}=5.0$ ,  $J_{12}=2.0$  Hz, 1 H; H-5<sup>a</sup>), 3.88 (s, 3 H; PhOMe), 3.84 (t,  $J_{13}=J_{13'}=9.5$  Hz, 1 H; H-4<sup>b</sup>), 3.77 (dd,  $J_{13}=10.0$ ,  $J_{12}=9.0$  Hz, 1 H; H-4<sup>c</sup>), 3.71 (brd,  $J_{14}=10.0$  Hz, 1 H; H-5<sup>b</sup>), 3.67 (dd,  $J_{13}=10.0$ ,  $J_{12}=9.5$  Hz, 1 H; H-3<sup>b</sup>), 3.54 (dd,  $J_{12}=10.0$ ,  $J_{11}=9.0$  Hz, 1 H; H-3<sup>c</sup>), 3.53 (brd,  $J_{13}=2.0$  Hz, 1 H; H-4<sup>c</sup>), 3.40 (dd,  $J_{13}=10.0$ ,  $J_{12}=3.5$  Hz, 1 H; H-2<sup>a</sup>), 3.41 (s, 3 H; COOMe), 3.39 (dd,  $J_{13}=10.0$ ,  $J_{12}=3.5$  Hz, 1 H; H-2<sup>b</sup>), 3.13 (s, 3 H; COOMe) ppm; <sup>13</sup>C NMR (100.6 MHz, D<sub>2</sub>O): δ = 171.4, 171.1 (C=O), 159.2 (C-OMe pMBn), 137.7, 137.6, 137.5, 137.4 ( $C_{\text{arom}}$ ), 134.0 (C-b), 131.3, 128.9, 128.8, 128.7, 128.6, 128.5, 128.4, 128.3 ( $C_{\text{arom}}$ ), 118.6 (C-c), 114.3 ( $C_{\text{arom}}$  pMBn), 99.3 (C-1<sup>a</sup>), 98.3 (C-1<sup>b</sup>), 97.8 (C-1<sup>c</sup>), 96.8 (C-1<sup>a</sup>), 78.1 (C-3<sup>a</sup>), 77.7 (C-3<sup>b</sup>), 76.0 (CH<sub>2</sub>Ph), 75.2 (CH<sub>2</sub>Ph), 75.0 (C-4<sup>a</sup>), 74.4 (C-4<sup>b</sup>), 73.9 (C-3<sup>b</sup>), 73.2 (CH<sub>2</sub>Ph), 72.4 (C-4<sup>c</sup>, CH<sub>2</sub>Ph), 72.0 (C-2<sup>a</sup>), 70.6 (C-4<sup>c</sup>), 70.3 (C-2<sup>b</sup>, C-5<sup>a</sup>, CH<sub>2</sub> pMBn), 69.8 (C-3<sup>c</sup>), 69.4 (C-5<sup>b</sup>), 69.3 (Ca), 67.7 (C-5<sup>b</sup>), 67.3 (C-5<sup>c</sup>), 67.2 (C-6<sup>a</sup>), 66.3 (C-6<sup>b</sup>), 58.4 (C-2<sup>a</sup>), 57.9 (C-2<sup>b</sup>), 55.5 (CH<sub>2</sub> pMBn), 53.0 (CH<sub>2</sub> COOMe), 52.6 (CH<sub>2</sub> COOMe) ppm; ESI HRMS:  $m/z$  calcd for C<sub>45</sub>H<sub>40</sub>N<sub>2</sub>O<sub>2</sub>Na<sub>6</sub> [ $M-3Na$ ]<sup>7+</sup>: 594.3945; found: 594.3940.

All-yl [(methyl 3-*O*-benzyl-4-*O*-(4-methoxybenzyl)-2-*O*-sulfonato- $\alpha$ -L-idopyranosyluronate)-(1→4)-*O*-(3-*O*-benzyl-2-deoxy-2-sulfamino-6-*O*-sulfonato- $\alpha$ -D-glucopyranoside)-(1→4)-*O*-(methyl 3-*O*-benzyl-2-*O*-sulfonato- $\alpha$ -L-idopyranosyluronate)-(1→4)-*O*-(3-*O*-benzyl-2-deoxy-2-sulfamino-6-*O*-sulfonato- $\alpha$ -D-glucopyranoside)-(1→4)-*O*-(methyl 3-*O*-benzyl-2-*O*-sulfonato- $\alpha$ -L-idopyranosyluronate)-(1→4)-*O*-(3-*O*-benzyl-2-deoxy-2-sulfamino-6-*O*-sulfonato- $\alpha$ -D-glucopyranoside) nonasodium salt (19): Hexa-saccharide 10 (105 mg, 50 μmol) was treated with K<sub>2</sub>CO<sub>3</sub> (3.5 mg, 25 μmol, 0.5 equiv) in MeOH (1 mL) as described above. Flash chromatography (silica gel, toluene/acetone (8:2–6:4)) gave the deacetylated compound 13 (80 mg, 86%). A portion of this material (59 mg, 32 μmol) was treated with propane-1,3-dithiol (100 μL, 1 mmol, 30 equiv) and triethylamine (140 μL, 1 mmol, 30 equiv) in MeOH (550 μL) as described above. Flash chromatography (silica gel, AcOEt/CH<sub>2</sub>Cl<sub>2</sub>/MeOH (50:40:10–40:40:20)) with 0.1% NEt<sub>3</sub> gave 16 (49 mg, 87%): <sup>13</sup>C NMR (62.5 MHz, CDCl<sub>3</sub>[D<sub>2</sub>O]/MeOH (8:2)): δ = 169.9 (C=O), 159.2 (C-OMe pMBn), 138.1, 137.2 ( $C_{\text{arom}}$ ), 133.5 (C-b), 129.5, 128.1, 128.0, 127.9, 127.7, 127.2, 126.9, 126.4 ( $C_{\text{arom}}$ ), 116.8 (C-c), 113.4 ( $C_{\text{arom}}$  pMBn), 100.6, 100.5, 97.7, 96.4 (6×C-1), 81.4, 80.7, 74.8, 74.7, 74.6, 74.4 (CH<sub>2</sub>Ph), 73.7, 73.6, 72.7 (CH<sub>2</sub>Ph), 72.4, 72.3, 72.1 (CH<sub>2</sub>Ph), 72.0 (CH<sub>2</sub>Ph), 70.9, 68.7,

68.4, 67.9 (CH<sub>2</sub>Ph), 67.8, 67.6, 60.2 (C-6), 60.0 (C-6), 59.8 (C-6), 55.0 (OCH<sub>3</sub>), 54.8 (OCH<sub>3</sub>), 51.5 (OCH<sub>3</sub>) ppm; ESI HRMS:  $m/z$  calcd for C<sub>45</sub>H<sub>40</sub>N<sub>2</sub>O<sub>2</sub> [ $M+2H$ ]<sup>2+</sup>: 886.8719; found: 886.8723.

Compound 16 (35 mg, 18 μmol) dissolved in pyridine (2.5 mL) was treated with sulfur trioxide pyridine complex (130 mg, 0.81 mmol, 45 equiv) as described above to yield 19 (46 mg, 95%) as a nonasodium salt (98% pure by HPLC analysis): <sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O): δ = 7.59–7.20 (m, 30 H; Ph), 7.08 (d,  $J=8.5$  Hz, 2 H; Ph-OMe), 6.99 (d,  $J=8.5$  Hz, 2 H; Ph-OMe), 6.08 (ddd,  $J_{ax}=17.0$ ,  $J_{ax}-J_{ax'}=10.5$ ,  $J_{ax'}=6.0$  Hz, 1 H; H-b), 5.45 (dq,  $J_{ax}=17.0$ ,  $J_{ax}-J_{ax'}=10.5$ ,  $J_{ax'}=6.0$ ,  $J_{ax}-J_{ax'}=5.0$  Hz, 1 H; H-c), 5.42 (brs, 1 H; H-1<sup>a</sup>), 5.32 (dq,  $J_{ax}=10.5$ ,  $J_{ax}-J_{ax'}=10.5$ ,  $J_{ax'}=6.0$ ,  $J_{ax}-J_{ax'}=5.0$  Hz, 1 H; H-c), 5.30 (d,  $J_{12}=3.5$  Hz, 1 H; H-1<sup>b</sup>), 5.27 (d,  $J_{12}=3.5$  Hz, 1 H; H-1<sup>c</sup>), 5.20 (d,  $J_{12}=3.5$  Hz, 1 H; H-1<sup>b</sup>), 5.19 (brs, 1 H; H-1<sup>b</sup>), 4.97 (brd,  $J_{14}=1.5$  Hz, 2 H; H-5<sup>a</sup>, H-5<sup>b</sup>), 4.92 (d,  $J=11.0$  Hz, 1 H; CH<sub>2</sub>Ph), 4.89 (d,  $J_{14}=2.0$  Hz, 1 H; H-5<sup>b</sup>), 4.88 (d,  $J=11.0$  Hz, 1 H; CH<sub>2</sub>Ph), 4.83 (d,  $J=11.0$  Hz, 1 H; CH<sub>2</sub>Ph), 4.82 (d,  $J=11.5$  Hz, 1 H; CH<sub>2</sub>Ph), 4.81 (d,  $J=11.0$  Hz, 1 H; CH<sub>2</sub>Ph), 4.74 (d,  $J=10.0$  Hz, 1 H; CH<sub>2</sub>Ph), 4.71 (d,  $J=10.0$  Hz, 1 H; CH<sub>2</sub>Ph), 4.61–4.59 (m, 1 H; H-2<sup>a</sup>), 4.59 (d,  $J=11.5$  Hz, 1 H; CH<sub>2</sub>Ph), 4.56 (brd,  $J_{13}=3.0$  Hz, 1 H; H-3<sup>a</sup>), 4.47 (d,  $J=12.0$  Hz, 1 H; CH<sub>2</sub>PhOMe), 4.48–4.46 (m, 1 H; H-2<sup>b</sup>), 4.43 (d,  $J=10.0$  Hz, 1 H; CH<sub>2</sub>Ph), 4.42 (brt,  $J_{13}=J_{13'}=3.0$  Hz, 1 H; H-3<sup>a</sup>), 4.40–4.37 (m, 1 H; H-3<sup>b</sup>), 4.37 (d,  $J=11.0$  Hz, 1 H; CH<sub>2</sub>Ph), 4.36 (dd,  $J_{ax}=10.0$ ,  $J_{ax'}=2.5$  Hz, 1 H; H-6<sup>a</sup>), 4.33–4.26 (m, 3 H; H-6<sup>c</sup>, H-6<sup>b</sup>, CH<sub>2</sub>Ph), 4.30 (ddt,  $J_{pm}=13.0$ ,  $J_{12}=6.0$ ,  $J_{ax}-J_{ax'}=1.5$  Hz, 1 H; H-a), 4.26–4.20 (m, 4 H; H-6<sup>c</sup>, H-6<sup>b</sup>, CH<sub>2</sub>PhOMe), 4.19–4.17 (m, 1 H; H-3<sup>b</sup>), 4.15–4.14 (m, 1 H; H-4<sup>a</sup>), 4.14 (ddt,  $J_{pm}=13.0$ ,  $J_{12}=6.0$ ,  $J_{ax}-J_{ax'}=1.5$  Hz, 1 H; H-a), 4.14–4.13 (m, 1 H; H-4<sup>b</sup>), 4.09 (ddd,  $J_{13}=10.0$ ,  $J_{12}=5.0$ ,  $J_{11}=2.5$  Hz, 1 H; H-5<sup>a</sup>), 3.87 (s, 3 H; PhOMe), 3.85 (dd,  $J_{13}=10.0$ ,  $J_{12}=9.0$  Hz, 1 H; H-4<sup>b</sup>), 3.83 (t,  $J_{13}=J_{13'}=10.0$  Hz, 1 H; H-4<sup>c</sup>), 3.78 (dd,  $J_{13}=10.0$ ,  $J_{12}=9.0$  Hz, 1 H; H-4<sup>c</sup>), 3.73 (dt,  $J_{13}=10.0$ ,  $J_{12}=9.0$  Hz, 1 H; H-4<sup>c</sup>), 3.69 (dt,  $J_{13}=10.0$ ,  $J_{12}=9.0$  Hz, 1 H; H-5<sup>a</sup>), 3.68 (dd,  $J_{13}=10.0$ ,  $J_{12}=9.0$  Hz, 1 H; H-3<sup>b</sup>), 3.55 (dd,  $J_{13}=10.0$ ,  $J_{12}=9.0$  Hz, 1 H; H-3<sup>b</sup>), 3.53 (t,  $J_{12}=J_{11}=10.0$  Hz, 1 H; H-3<sup>c</sup>), 3.51–3.49 (m, 1 H; H-4<sup>c</sup>), 3.50 (dd,  $J_{13}=10.0$ ,  $J_{12}=3.5$  Hz, 1 H; H-2<sup>a</sup>), 3.44 (dd,  $J_{13}=10.0$ ,  $J_{12}=3.5$  Hz, 1 H; H-2<sup>b</sup>), 3.40 (dd,  $J_{13}=10.0$ ,  $J_{12}=3.5$  Hz, 1 H; H-2<sup>b</sup>), 3.39 (s, 3 H; COOMe), 3.38 (s, 3 H; COOMe), 3.12 (s, 3 H; COOMe) ppm; <sup>13</sup>C NMR (100.6 MHz, D<sub>2</sub>O): δ = 171.4, 171.2, 171.0 (C=O), 159.2 (C-OMe pMBn), 137.6, 137.4 ( $C_{\text{arom}}$ ), 134.0 (C-b), 131.2, 129.5, 129.4, 129.1, 129.0, 128.8, 128.7, 128.5 ( $C_{\text{arom}}$ ), 118.6 (C-c), 114.2 ( $C_{\text{arom}}$  pMBn), 99.4 (C-1<sup>a</sup>), 98.7 (C-1<sup>b</sup>), 98.2 (C-1<sup>c</sup>, C-1<sup>a</sup>), 97.7 (C-1<sup>b</sup>), 96.7 (C-1<sup>a</sup>), 78.1 (C-3<sup>a</sup>), 77.8 (C-3<sup>b</sup>), 77.7 (C-3<sup>b</sup>), 76.2 (CH<sub>2</sub>Ph), 75.9 (CH<sub>2</sub>Ph), 75.2 (CH<sub>2</sub>Ph), 75.0 (C-4<sup>a</sup>), 74.4 (C-4<sup>b</sup>, C-4<sup>c</sup>), 73.9 (C-3<sup>b</sup>), 73.3 (C-3<sup>b</sup>), 73.2 (CH<sub>2</sub>Ph), 73.1 (CH<sub>2</sub>Ph), 72.9 (C-4<sup>c</sup>), 72.3 (CH<sub>2</sub>Ph), 72.2 (C-4<sup>b</sup>), 72.1 (C-2<sup>a</sup>), 71.4 (C-2<sup>b</sup>), 70.4 (C-4<sup>c</sup>), 70.3 (C-5<sup>a</sup>), 70.2 (C-2<sup>c</sup>, C-5<sup>b</sup>, CH<sub>2</sub> pMBn), 69.5 (C-3<sup>c</sup>), 69.4 (C-5<sup>b</sup>), 69.2 (C-a), 67.7 (C-5<sup>b</sup>, C-5<sup>c</sup>), 67.3 (C-5<sup>c</sup>), 67.1 (C-6<sup>a</sup>), 66.3 (C-6<sup>b</sup>, C-6<sup>c</sup>), 58.4 (C-2<sup>a</sup>), 58.3 (C-2<sup>b</sup>), 57.9 (C-2<sup>c</sup>), 55.7 (CH<sub>2</sub> pMBn), 53.0 (CH<sub>2</sub> COOMe), 52.9 (CH<sub>2</sub> COOMe), 52.5 (CH<sub>2</sub> COOMe) ppm; ESI HRMS:  $m/z$  calcd for C<sub>45</sub>H<sub>40</sub>N<sub>2</sub>O<sub>2</sub>Na<sub>9</sub> [ $M-3Na$ ]<sup>7+</sup>: 873.4034; found: 873.4034.

All-yl [(methyl 3-*O*-benzyl-4-*O*-(4-methoxybenzyl)-2-*O*-sulfonato- $\alpha$ -L-idopyranosyluronate)-(1→4)-*O*-(3-*O*-benzyl-2-deoxy-2-sulfamino-6-*O*-sulfonato- $\alpha$ -D-glucopyranoside)-(1→4)-*O*-(methyl 3-*O*-benzyl-2-*O*-sulfonato- $\alpha$ -L-idopyranosyluronate)-(1→4)-*O*-(3-*O*-benzyl-2-deoxy-2-sulfamino-6-*O*-sulfonato- $\alpha$ -D-glucopyranoside)-(1→4)-*O*-(methyl 3-*O*-benzyl-2-*O*-sulfonato- $\alpha$ -L-idopyranosyluronate)-(1→4)-*O*-(3-*O*-benzyl-2-deoxy-2-sulfamino-6-*O*-sulfonato- $\alpha$ -D-glucopyranoside) dodecasodium salt (20): Octasaccharide 11 (110 mg, 40 μmol) was treated with K<sub>2</sub>CO<sub>3</sub> (6 mg, 40 μmol, 1 equiv) in MeOH (2 mL) as described above. Flash chromatography (silica gel, toluene/acetone (8:2–6:4)) gave the deacetylated compound 14 (88 mg, 91%). This compound was added to another batch of 14, and the combined products (100 mg, 41.5 μmol) were treated with propane-1,3-dithiol (167 μL, 1.66 mmol, 40 equiv) and triethylamine (230 μL, 1.66 mmol, 40 equiv) in MeOH (550 μL) as described above. Flash chromatography (silica gel, AcOEt/CH<sub>2</sub>Cl<sub>2</sub>/MeOH (50:40:10–40:40:20)) with 0.1% NEt<sub>3</sub> gave 17 (82 mg, 84%): <sup>13</sup>C NMR (62.5 MHz, CDCl<sub>3</sub>[D<sub>2</sub>O]/MeOH (8:2)): δ = 169.9 (C=O), 159.2 (C-OMe pMBn), 138.0, 137.4, 137.2, 137.1 ( $C_{\text{arom}}$ ), 133.4 (C-b), 129.5, 128.6, 128.2, 128.1, 128.0, 127.9, 127.6, 127.2, 127.0, 126.9, 126.7 ( $C_{\text{arom}}$ ), 116.9 (C-c), 113.5 ( $C_{\text{arom}}$  pMBn), 100.6, 97.7, 96.1 (8×C-1), 81.4, 80.4, 74.9, 74.8, 74.6, 74.4 (CH<sub>2</sub>Ph), 74.3, 73.7, 73.5, 72.7 (CH<sub>2</sub>Ph), 72.5, 72.2 (CH<sub>2</sub>Ph), 72.1

(CH<sub>2</sub>Ph), 68.8, 68.4, 68.3, 67.9 (CH<sub>2</sub>Ph), 67.8, 67.5, 60.3 (C-6), 60.0 (C-6), 59.8 (C-6), 55.0 (OCH<sub>3</sub>), 54.8 (OCH<sub>3</sub>), 54.7 (OCH<sub>3</sub>), 51.5 (OCH<sub>3</sub>) ppm; ESI HRMS: *m/z* calcd for C<sub>21</sub>H<sub>33</sub>N<sub>4</sub>O<sub>12</sub> [M+2H]<sup>2+</sup>: 1152.4784; found: 1152.4806.

Compound 17 (30 mg, 13 μmol), dissolved in pyridine (2.5 mL), was treated with sulfur trioxide pyridine complex (125 mg, 0.78 mmol, 60 equiv) as described above to yield 20 (44 mg, 95%) as a dodecasodium salt (97% pure by HPLC analysis): <sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O): δ = 7.60–7.54 (m, 6H; Ph), 7.54–7.30 (m, 30H; Ph), 7.30–7.21 (m, 4H; Ph), 7.08 (d, *J* = 8.5 Hz, 2H; Ph-OMe), 6.99 (d, *J* = 8.5 Hz, 2H; Ph-OMe), 6.08 (ddd, *J*<sub>1,2</sub> = 17.0, *J*<sub>1,3</sub> = 10.5, *J*<sub>1,4</sub> = 6.0 Hz, 1H; H-b), 5.45 (dq, *J*<sub>1,2</sub> = 17.0, *J*<sub>1,3</sub> = *J*<sub>1,4</sub> = *J*<sub>2,3</sub> = 1.5 Hz, 1H; H-c), 5.42 (brs, 1H; H-1<sup>a</sup>), 5.40 (brs, 1H; H-1<sup>b</sup>), 5.35 (brs, 1H; H-1<sup>c</sup>), 5.32 (dq, *J*<sub>1,2</sub> = 10.5, *J*<sub>1,3</sub> = *J*<sub>1,4</sub> = *J*<sub>2,3</sub> = 1.5 Hz, 1H; H-cc), 5.31 (d, *J*<sub>1,2</sub> = 3.5 Hz, 1H; H-1<sup>d</sup>), 5.30 (d, *J*<sub>1,2</sub> = 3.5 Hz, 1H; H-1<sup>e</sup>), 5.26 (d, *J*<sub>1,2</sub> = 3.5 Hz, 1H; H-1<sup>f</sup>), 5.21 (d, *J*<sub>1,2</sub> = 3.5 Hz, 1H; H-1<sup>g</sup>), 5.19 (brs, 1H; H-1<sup>h</sup>), 4.98–4.95 (m, 3H; H-5<sup>a</sup>, H-5<sup>b</sup>, H-5<sup>c</sup>), 4.92 (d, *J* = 11.0 Hz, 1H; CH<sub>2</sub>Ph), 4.89 (d, *J*<sub>1,4</sub> = 2.0 Hz, 1H; H-5<sup>d</sup>), 4.89 (d, *J* = 11.0 Hz, 2H; CH<sub>2</sub>Ph), 4.83 (d, *J* = 10.0 Hz, 1H; CH<sub>2</sub>Ph), 4.82 (d, *J* = 12.0 Hz, 1H; CH<sub>2</sub>Ph), 4.81 (d, *J* = 11.0 Hz, 2H; 2 × CH<sub>2</sub>Ph), 4.79 (d, *J* = 11.0 Hz, 1H; CH<sub>2</sub>Ph), 4.76 (d, *J* = 11.0 Hz, 1H; CH<sub>2</sub>Ph), 4.75 (d, *J* = 10.0 Hz, 1H; CH<sub>2</sub>Ph), 4.72 (d, *J* = 10.0 Hz, 1H; CH<sub>2</sub>Ph), 4.70 (d, *J* = 10.5 Hz, 1H; CH<sub>2</sub>Ph), 4.60 (brdd, *J*<sub>1,2</sub> = 3.5, *J*<sub>1,3</sub> = 1.5 Hz, 1H; H-2<sup>a</sup>), 4.59 (d, *J* = 12.0 Hz, 1H; CH<sub>2</sub>Ph), 4.56 (brd, *J*<sub>1,2</sub> = 3.0 Hz, H-2<sup>b</sup>), 4.55 (brd, *J*<sub>1,2</sub> = 3.0 Hz, H-2<sup>c</sup>), 4.47 (d, *J* = 12.5 Hz, 1H; CH<sub>2</sub>PhOMe), 4.48 (brs, 1H; H-2<sup>d</sup>), 4.48 (d, *J* = 10.0 Hz, 1H; CH<sub>2</sub>Ph), 4.43–4.35 (m, 4H; H-3<sup>a</sup>, H-3<sup>b</sup>, H-3<sup>c</sup>, H-6<sup>a</sup>), 4.36 (d, *J* = 10.0 Hz, 1H; CH<sub>2</sub>Ph), 4.37 (brd, *J*<sub>1,2</sub> = 10.0, 1H; H-6<sup>b</sup> or H-6<sup>c</sup>), 4.31 (brd, *J*<sub>1,2</sub> = 10.0, 1H; H-6<sup>d</sup>), 4.30 (ddd, *J*<sub>1,2</sub> = 13.0, *J*<sub>1,3</sub> = 6.0, *J*<sub>1,4</sub> = *J*<sub>2,3</sub> = 1.5 Hz, 1H; H-a), 4.29 (d, *J* = 10.5 Hz, 1H; CH<sub>2</sub>Ph), 4.29 (d, *J* = 10.0 Hz, 1H; CH<sub>2</sub>Ph), 4.26–4.19 (m, 4H; H-6<sup>e</sup>, H-6<sup>f</sup>, H-6<sup>g</sup>, H-6<sup>h</sup>), 4.23 (d, *J* = 12.5 Hz, 1H; CH<sub>2</sub>PhOMe), 4.18 (brt, *J* = 2.5 Hz, 1H; H-3<sup>d</sup>), 4.17–4.11 (m, 4H; H-4<sup>a</sup>, H-4<sup>b</sup>, H-4<sup>c</sup>, H-4<sup>d</sup>), 4.09 (ddd, *J*<sub>1,2</sub> = 10.0, *J*<sub>1,3</sub> = 5.0, *J*<sub>1,4</sub> = 2.5 Hz, 1H; H-5<sup>e</sup>), 3.88 (s, 3H; PhOMe), 3.86 (dd, *J*<sub>1,2</sub> = 10.0, *J*<sub>1,3</sub> = 9.0 Hz, 1H; H-4<sup>e</sup>), 3.85 (t, *J*<sub>1,2</sub> = *J*<sub>1,3</sub> = 10.0 Hz, 1H; H-4<sup>f</sup>), 3.82 (t, *J*<sub>1,2</sub> = *J*<sub>1,3</sub> = 10.0 Hz, 1H; H-4<sup>g</sup>), 3.78 (dd, *J*<sub>1,2</sub> = 10.0, *J*<sub>1,3</sub> = 9.0 Hz, 1H; H-4<sup>h</sup>), 3.73 (brt, *J*<sub>1,2</sub> = 10.0, 1H; H-5<sup>f</sup>), 3.69 (brdd, *J*<sub>1,2</sub> = 10.0, *J* = 2.0 Hz, 2H; H-5<sup>g</sup>, H-5<sup>h</sup>), 3.68 (dd, *J*<sub>1,2</sub> = 10.0, *J*<sub>1,3</sub> = 9.0 Hz, 1H; H-3<sup>e</sup>), 3.54 (dd, *J*<sub>1,2</sub> = 10.0, *J*<sub>1,3</sub> = 9.0 Hz, 1H; H-3<sup>f</sup>), 3.40 (dd, *J*<sub>1,2</sub> = 10.0, *J*<sub>1,3</sub> = 3.5 Hz, 1H; H-3<sup>g</sup>), 3.52–3.30 (m, 1H; H-4<sup>i</sup>), 3.50 (dd, *J*<sub>1,2</sub> = 10.0, *J*<sub>1,3</sub> = 3.5 Hz, 1H; H-2<sup>e</sup>), 3.44 (dd, *J*<sub>1,2</sub> = 10.0, *J*<sub>1,3</sub> = 3.5 Hz, 1H; H-2<sup>f</sup>), 3.43 (dd, *J*<sub>1,2</sub> = 10.0, *J*<sub>1,3</sub> = 3.5 Hz, 1H; H-2<sup>g</sup>), 3.40 (dd, *J*<sub>1,2</sub> = 10.0, *J*<sub>1,3</sub> = 3.5 Hz, 1H; H-2<sup>h</sup>), 3.39 (s, 3H; COOMe), 3.38 (s, 3H; COOMe), 3.35 (s, 3H; COOMe), 3.12 (s, 3H; COOMe) ppm; <sup>13</sup>C NMR (100.6 MHz, D<sub>2</sub>O): δ = 171.4, 171.2, 171.1, 171.0 (C=O), 159.2 (C-OMe pMBn), 137.6, 137.4 (C<sub>quat</sub> pMBn), 134.0 (C-b), 131.2, 129.4, 129.0, 128.8, 128.7, 128.5 (C<sub>arom</sub>), 118.6 (C-c), 114.2 (C<sub>arom</sub> pMBn), 99.3 (C-1<sup>a</sup>), 99.8 (C-1<sup>b</sup>), 98.7 (C-1<sup>c</sup>), 98.2 (C-1<sup>d</sup>, C-1<sup>e</sup>, C-1<sup>f</sup>), 97.7 (C-1<sup>g</sup>), 96.7 (C-1<sup>h</sup>), 78.1 (C-3<sup>a</sup>), 77.9 (C-3<sup>b</sup>, C-3<sup>c</sup>), 77.8 (C-3<sup>d</sup>), 76.2 (CH<sub>2</sub>Ph), 75.9 (CH<sub>2</sub>Ph), 75.2 (CH<sub>2</sub>Ph), 74.9 (C-4<sup>a</sup>), 74.4 (C-4<sup>b</sup>, C-4<sup>c</sup>, C-4<sup>d</sup>), 73.9 (C-3<sup>e</sup> or C-3<sup>f</sup>), 73.4 (C-3<sup>g</sup> or C-3<sup>h</sup>), 73.2 (CH<sub>2</sub>Ph), 73.1 (CH<sub>2</sub>Ph, C-3<sup>i</sup>), 73.0 (CH<sub>2</sub>Ph), 72.8 (C-4<sup>e</sup>), 72.7 (C-4<sup>f</sup>), 72.3 (CH<sub>2</sub>Ph, C-4<sup>g</sup>), 72.1 (C-2<sup>a</sup>), 71.5 (C-2<sup>b</sup>, C-2<sup>c</sup>), 70.4 (C-2<sup>d</sup>, C-4<sup>h</sup>), 70.3 (C-5<sup>a</sup>, C-5<sup>b</sup>, C-5<sup>c</sup>), 70.2 (C-5<sup>d</sup>, CH<sub>2</sub> pMBn), 69.5 (C-5<sup>e</sup>), 69.4 (C-3<sup>j</sup>), 69.2 (C-a), 67.5/67.4 (C-6<sup>a</sup>, C-6<sup>b</sup>, C-6<sup>c</sup>, C-6<sup>d</sup>), 67.3 (C-5<sup>f</sup>), 67.1 (C-6<sup>e</sup>), 66.3 (C-6<sup>f</sup>, C-6<sup>g</sup>, C-6<sup>h</sup>), 58.4/58.2 (C-2<sup>e</sup>, C-6<sup>h</sup>, C-2<sup>f</sup>), 57.9 (C-2<sup>g</sup>), 55.7 (CH<sub>2</sub> pMBn), 53.1 (CH<sub>2</sub> COOMe), 53.0 (CH<sub>2</sub> COOMe), 52.9 (CH<sub>2</sub> COOMe), 52.5 (CH<sub>2</sub> COOMe) ppm; ESI MS: C<sub>21</sub>H<sub>33</sub>N<sub>4</sub>O<sub>12</sub>S<sub>3</sub>Na<sub>12</sub>; *m/z* (%) 5648 (5) [M-6Na]<sup>+</sup> 6, 6825 (10) [M-5Na]<sup>+</sup> 7, 838.8 (43) [M-4Na]<sup>+</sup> 7, 1152.6 (29) [M-3Na]<sup>+</sup> 7.

All<sup>11</sup> [(3-*O*-benzyl-4-*O*-(4-methoxybenzyl)-2-*O*-sulfonato-α-L-idopyranosyluronate)-(1→4)-*O*-(3-*O*-benzyl-2-deoxy-2-sulfamino-6-*O*-sulfonato-α-D-glucopyranoside)-(1→4)-*O*-(3-*O*-benzyl-2-*O*-sulfonato-α-L-idopyranosyluronate)]-(1→4)-*O*-(3-*O*-benzyl-2-deoxy-2-sulfamino-6-*O*-sulfonato-α-D-glucopyranoside) octasodium salt (21): Tetrasaccharide 18 (32 mg, 17.2 μmol) was dissolved in a mixture of LiOH (42 mg, 1 mmol, 52 equiv), water (240 μL) and H<sub>2</sub>O (35% in water, 260 μL, 3 mmol). After 24 h at room temperature, AcOH (45 μL, 0.75 mmol) was added and the mixture was purified by RP-18 flash chromatography (5 min AcOH-NEt<sub>3</sub> (pH 7.0)/MeOH (100:0–60:40)), followed by ion exchange on BioRad AG50W-X8 200 (Na<sup>+</sup>, 5 mL) resin after removal of AcOH-NEt<sub>3</sub> salts by lyophilisation (2 × 1 mL H<sub>2</sub>O). Compound 21 (29 mg, 73%) was thus obtained as an octasodium salt (94% pure by HPLC analysis):

<sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O): δ = 7.30–7.44 (m, 4H; Ph), 7.44–7.32 (m, 16H; Ph), 7.30 (d, *J* = 8.5 Hz, 2H; Ph-OMe), 6.94 (d, *J* = 8.5 Hz, 2H; Ph-OMe), 6.01 (dddd, *J*<sub>1,2</sub> = 17.5, *J*<sub>1,3</sub> = 10.5, *J*<sub>1,4</sub> = 6.0, *J*<sub>2,3</sub> = 5.5 Hz, 1H; H-b), 5.45 (brs, 1H; H-1<sup>a</sup>), 5.38 (dq, *J*<sub>1,2</sub> = 17.5, *J*<sub>1,3</sub> = *J*<sub>1,4</sub> = *J*<sub>2,3</sub> = 1.5 Hz, 1H; H-c), 5.32 (d, *J*<sub>1,2</sub> = 3.5 Hz, 1H; H-1<sup>b</sup>), 5.29 (brs, 1H; H-1<sup>c</sup>), 5.14 (d, *J*<sub>1,2</sub> = 3.5 Hz, 1H; H-1<sup>d</sup>), 5.14 (dq, *J*<sub>1,2</sub> = 10.5, *J*<sub>1,3</sub> = *J*<sub>1,4</sub> = *J*<sub>2,3</sub> = 1.5 Hz, 1H; H-cc), 4.85 (d, *J* = 11.0 Hz, 1H; CH<sub>2</sub>Ph), 4.76 (d, *J* = 11.0 Hz, 3H; CH<sub>2</sub>Ph), 4.76 (d, *J*<sub>1,4</sub> = 2.0 Hz, 1H; H-5<sup>a</sup>), 4.69 (d, *J*<sub>1,4</sub> = 1.5 Hz, 1H; H-5<sup>b</sup>), 4.64 (d, *J* = 11.0 Hz, 1H; CH<sub>2</sub>Ph), 4.64 (d, *J* = 12.0 Hz, 1H; CH<sub>2</sub>Ph), 4.61 (d, *J* = 11.0 Hz, 1H; CH<sub>2</sub>Ph), 4.60 (dd, *J*<sub>1,2</sub> = 3.5, *J*<sub>1,3</sub> = 1.5 Hz, 1H; H-2<sup>a</sup>), 4.51 (d, *J* = 12.0 Hz, 1H; CH<sub>2</sub>Ph), 4.45 (brd, *J*<sub>1,2</sub> = 2.5, 1H; H-2<sup>b</sup>), 4.44 (d, *J* = 11.0 Hz, 1H; CH<sub>2</sub>PhOMe), 4.40 (d, *J* = 11.0 Hz, 1H; CH<sub>2</sub>PhOMe), 4.39 (dd, *J*<sub>1,2</sub> = 11.5, *J*<sub>1,3</sub> = 2.5 Hz, 1H; H-6<sup>a</sup>), 4.34 (dd, *J*<sub>1,2</sub> = 11.5, *J*<sub>1,3</sub> = 5.0 Hz, 1H; H-6<sup>b</sup>), 4.31 (brs, 2H; H-6<sup>c</sup>, H-6<sup>d</sup>), 4.28 (t, *J*<sub>1,2</sub> = *J*<sub>1,3</sub> = 3.5 Hz, 1H; H-3<sup>a</sup>), 4.25 (dd, *J*<sub>1,2</sub> = 3.5, *J*<sub>1,3</sub> = 2.0 Hz, 1H; H-4<sup>a</sup>), 4.24 (ddd, *J*<sub>1,2</sub> = 12.5, *J*<sub>1,3</sub> = 5.5, *J*<sub>1,4</sub> = 1.5 Hz, 1H; H-a), 4.16 (brd, *J*<sub>1,2</sub> = 10.0 Hz, 1H; H-5<sup>a</sup>), 4.06 (ddd, *J*<sub>1,2</sub> = 12.5, *J*<sub>1,3</sub> = 6.0, *J*<sub>1,4</sub> = 1.5 Hz, 1H; H-a), 4.06 (ddd, *J*<sub>1,2</sub> = 9.0, *J*<sub>1,3</sub> = 5.0, *J*<sub>1,4</sub> = 2.5 Hz, 1H; H-5<sup>b</sup>), 3.96 (brt, *J*<sub>1,2</sub> = *J*<sub>1,3</sub> = 2.5 Hz, 1H; H-3<sup>b</sup>), 3.93 (t, *J*<sub>1,2</sub> = *J*<sub>1,3</sub> = 10.0 Hz, 1H; H-4<sup>b</sup>), 3.91 (t, *J*<sub>1,2</sub> = *J*<sub>1,3</sub> = 9.0 Hz, 1H; H-4<sup>c</sup>), 3.91–3.89 (m, 1H; H-4<sup>d</sup>), 3.83 (s, 3H; PhOMe), 3.81 (t, *J*<sub>1,2</sub> = *J*<sub>1,3</sub> = 10.0 Hz, 1H; H-3<sup>c</sup>), 3.73 (dd, *J*<sub>1,2</sub> = 10.0, *J*<sub>1,3</sub> = 9.0 Hz, 1H; H-3<sup>d</sup>), 3.44 (dd, *J*<sub>1,2</sub> = 10.0, *J*<sub>1,3</sub> = 3.5 Hz, 1H; H-2<sup>c</sup>), 3.37 (dd, *J*<sub>1,2</sub> = 10.0, *J*<sub>1,3</sub> = 3.5 Hz, 1H; H-2<sup>d</sup>) ppm; <sup>13</sup>C NMR (100.6 MHz, D<sub>2</sub>O): δ = 180.0, 179.1 (C=O, from HMBC), 163.3 (C-OMe pMBn, from HMBC), 137.9, 137.6, 137.4, 137.2 (C<sub>quat</sub> pMBn), 134.0 (C-b), 130.8, 129.5, 129.2, 128.9, 128.8, 128.6 (C<sub>arom</sub>), 118.4 (C-c), 114.2 (C<sub>arom</sub> pMBn), 98.1 (C-1<sup>a</sup>, C-1<sup>b</sup>), 97.5 (C-1<sup>c</sup>), 96.5 (C-1<sup>d</sup>), 77.3 (C-3<sup>a</sup>), 77.2 (C-3<sup>b</sup>), 75.4 (CH<sub>2</sub>Ph, C-3<sup>a</sup>, C-4<sup>a</sup>), 75.1 (C-4<sup>b</sup>), 75.0 (CH<sub>2</sub>Ph), 73.4 (C-4<sup>c</sup>, C-4<sup>d</sup>), 73.0 (CH<sub>2</sub>Ph), 72.9 (C-2<sup>a</sup>), 72.1 (CH<sub>2</sub> pMBn), 71.7 (CH<sub>2</sub>Ph), 71.6 (C-2<sup>b</sup>), 71.0 (C-3<sup>c</sup>), 69.7 (C-5<sup>a</sup>), 69.2 (C-5<sup>b</sup>), 69.1 (C-5<sup>c</sup>), 69.0 (C-a), 68.5 (C-2<sup>c</sup>), 67.3 (C-6<sup>a</sup>), 66.9 (C-6<sup>b</sup>), 58.4 (C-2<sup>d</sup>), 57.9 (C-2<sup>e</sup>), 55.5 (CH<sub>2</sub> pMBn) ppm; ESI MS: C<sub>40</sub>H<sub>48</sub>N<sub>4</sub>O<sub>26</sub>S<sub>6</sub>Na<sub>8</sub>; *m/z* (%) 438.6 (91) [M-5Na+H]<sup>+</sup> 7, 592.4 (100) [M-4Na+H]<sup>+</sup> 7, 999.7 (76) [M-3Na]<sup>+</sup> 7, 911.0 (32) [M-2Na]<sup>+</sup> 7.

All<sup>12</sup> [(3-*O*-benzyl-4-*O*-(4-methoxybenzyl)-2-*O*-sulfonato-α-L-idopyranosyluronate)-(1→4)-*O*-(3-*O*-benzyl-2-deoxy-2-sulfamino-6-*O*-sulfonato-α-D-glucopyranoside)-(1→4)-*O*-(3-*O*-benzyl-2-*O*-sulfonato-α-L-idopyranosyluronate)]-(1→4)-*O*-(3-*O*-benzyl-2-deoxy-2-sulfamino-6-*O*-sulfonato-α-D-glucopyranoside) nonasodium salt (22): Hexasaccharide 19 (35 mg, 13 μmol) was dissolved in a mixture of LiOH (42 mg, 1 mmol, 78 equiv), water (240 μL) and H<sub>2</sub>O (35% in water, 260 μL, 3 mmol). After 48 h at 37°C, AcOH (45 μL, 0.75 mmol) was added and the mixture was purified as described above to give 22 (32 mg, 91%) as a nonasodium salt (98% pure by HPLC analysis): <sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O): δ = 7.55–7.45 (m, 4H; Ph), 7.45–7.28 (m, 26H; Ph), 7.30 (d, *J* = 8.5 Hz, 2H; Ph-OMe), 6.94 (d, *J* = 8.5 Hz, 2H; Ph-OMe), 6.02 (dddd, *J*<sub>1,2</sub> = 17.0, *J*<sub>1,3</sub> = 10.5, *J*<sub>1,4</sub> = 6.0, *J*<sub>2,3</sub> = 5.5 Hz, 1H; H-b), 5.48 (brs, 1H; H-1<sup>a</sup>), 5.48 (brs, 1H; H-1<sup>b</sup>), 5.38 (dq, *J*<sub>1,2</sub> = 17.0, *J*<sub>1,3</sub> = *J*<sub>1,4</sub> = *J*<sub>2,3</sub> = 1.5 Hz, 1H; H-c), 5.31 (d, *J*<sub>1,2</sub> = 3.5 Hz, 1H; H-1<sup>c</sup>), 5.29 (brs, 1H; H-1<sup>d</sup>), 5.28 (d, *J*<sub>1,2</sub> = 3.5 Hz, 1H; H-1<sup>e</sup>), 5.26 (dq, *J*<sub>1,2</sub> = 10.5, *J*<sub>1,3</sub> = *J*<sub>1,4</sub> = *J*<sub>2,3</sub> = 1.5 Hz, 1H; H-cc), 5.14 (d, *J*<sub>1,2</sub> = 3.5 Hz, 1H; H-1<sup>f</sup>), 4.88 (d, *J* = 10.5 Hz, 1H; CH<sub>2</sub>Ph), 4.84 (d, *J* = 11.0 Hz, 1H; CH<sub>2</sub>Ph), 4.77 (d, *J*<sub>1,4</sub> = 2.0 Hz, 1H; H-5<sup>a</sup>), 4.76 (d, *J* = 11.0 Hz, 1H; CH<sub>2</sub>Ph), 4.75 (d, *J*<sub>1,4</sub> = 2.5 Hz, 1H; H-5<sup>b</sup>), 4.74 (d, *J* = 11.0 Hz, 1H; CH<sub>2</sub>Ph), 4.73 (d, *J* = 11.0 Hz, 1H; CH<sub>2</sub>Ph), 4.72 (d, *J* = 10.5 Hz, 1H; CH<sub>2</sub>Ph), 4.68 (d, *J*<sub>1,4</sub> = 1.5 Hz, 1H; H-5<sup>c</sup>), 4.66 (d, *J* = 11.0 Hz, 1H; CH<sub>2</sub>Ph), 4.64–4.39 (m, 4H; H-2<sup>a</sup>, H-2<sup>b</sup>, 2 × CH<sub>2</sub>Ph), 4.50 (s, 2H; 2 × CH<sub>2</sub>Ph), 4.49 (d, *J* = 12.0 Hz, 1H; CH<sub>2</sub>Ph), 4.46 (brd, *J*<sub>1,2</sub> = 11.0 Hz, 1H; H-6<sup>a</sup>, H-6<sup>b</sup>), 4.43–4.41 (m, 1H; H-2<sup>c</sup>), 4.42 (d, *J* = 12.0 Hz, 1H; CH<sub>2</sub>PhOMe), 4.39 (d, *J* = 12.0 Hz, 1H; CH<sub>2</sub>PhOMe), 4.37 (brs, 2H; H-6<sup>c</sup>, H-6<sup>d</sup>), 4.34 (brd, *J*<sub>1,2</sub> = 11.0 Hz, 1H; H-6<sup>e</sup> or H-6<sup>f</sup>), 4.33 (brd, *J*<sub>1,2</sub> = 10.0 Hz, 1H; H-6<sup>g</sup> or H-6<sup>h</sup>), 4.29 (brt, *J*<sub>1,2</sub> = *J*<sub>1,3</sub> = 3.0 Hz, 1H; H-3<sup>a</sup>), 4.28 (brd, *J*<sub>1,2</sub> = 10.0 Hz, 1H; H-6<sup>i</sup> or H-6<sup>j</sup>), 4.26 (brt, *J*<sub>1,2</sub> = *J*<sub>1,3</sub> = 3.5 Hz, 1H; H-3<sup>b</sup>), 4.26 (dd, *J*<sub>1,2</sub> = 2.0, *J*<sub>1,3</sub> = 3.0 Hz, 1H; H-4<sup>a</sup>), 4.24 (ddd, *J*<sub>1,2</sub> = 13.0, *J*<sub>1,3</sub> = 5.5, *J*<sub>1,4</sub> = 1.5 Hz, 1H; H-a), 4.23–4.21 (m, 1H; H-4<sup>b</sup>), 4.12 (brd, *J*<sub>1,2</sub> = 10.0 Hz, 1H; H-5<sup>a</sup>, H-5<sup>b</sup>), 4.10–4.05 (m, 1H; H-5<sup>c</sup>), 4.06 (ddd, *J*<sub>1,2</sub> = 13.0, *J*<sub>1,3</sub> = 6.0, *J*<sub>1,4</sub> = 1.5 Hz, 1H; H-a), 3.94 (dd, *J*<sub>1,2</sub> = 10.0, *J*<sub>1,3</sub> = 9.0 Hz, 1H; H-4<sup>c</sup>), 3.93 (dd, *J*<sub>1,2</sub> = 10.0, *J*<sub>1,3</sub> = 9.0 Hz, 1H; H-4<sup>d</sup>), 3.94–3.92 (m, 1H; H-3<sup>c</sup>), 3.91 (dd, *J*<sub>1,2</sub> = 10.0, *J*<sub>1,3</sub> = 9.0 Hz, 1H; H-4<sup>e</sup>), 3.90–3.87 (m, 1H; H-4<sup>f</sup>), 3.83 (s, 3H; PhOMe), 3.82 (dd, *J*<sub>1,2</sub> = 10.5,

$J_{1,2}$  = 9.0 Hz, 1H; H-3<sup>b</sup>), 3.74 (dd,  $J_{1,2}$  = 10.5,  $J_{1,4}$  = 9.0 Hz, 1H; H-3<sup>a</sup>), 3.70 (dd,  $J_{1,2}$  = 10.5,  $J_{1,4}$  = 9.0 Hz, 1H; H-3<sup>c</sup>), 3.44 (dd,  $J_{1,2}$  = 10.5,  $J_{1,3}$  = 3.5 Hz, 1H; H-2<sup>b</sup>), 3.39 (dd,  $J_{1,2}$  = 10.5,  $J_{1,3}$  = 3.5 Hz, 1H; H-2<sup>c</sup>), 3.37 (dd,  $J_{1,2}$  = 10.5,  $J_{1,3}$  = 3.5 Hz, 1H; H-2<sup>a</sup>) ppm; <sup>13</sup>C NMR (100.6 MHz, D<sub>2</sub>O):  $\delta$  = 175.4, 174.7 (C=O), 158.9 (C-OMe pMBn), 137.9, 137.7, 137.6, 137.4, 137.2 (C<sub>quat</sub>), 133.9 (C-b), 130.8, 130.0, 129.6, 129.5, 129.4, 129.1, 129.0, 128.9, 128.8, 128.6, 128.2 (C<sub>arom</sub>), 118.3 (C-c), 114.2 (C<sub>arom</sub> pMBn), 98.3 (C-1<sup>c</sup>), 98.1 (C-1<sup>b</sup>), 97.3 (C-1<sup>a</sup>), 97.2 (C-1<sup>b</sup>), 96.5 (C-1<sup>a</sup>), 77.8 (C-3<sup>c</sup>), 77.2 (C-3<sup>b</sup>), 77.1 (C-3<sup>a</sup>), 75.8 (CH<sub>2</sub>Ph), 75.4 (C-3<sup>b</sup>), 75.3 (C-4<sup>a</sup>, CH<sub>2</sub>Ph), 75.1 (C-4<sup>b</sup>, CH<sub>2</sub>Ph), 74.9 (C-4<sup>c</sup>, CH<sub>2</sub>Ph), 74.6 (C-3<sup>b</sup>), 73.3 (C-4<sup>c</sup>, C-4<sup>b</sup>), 73.0 (CH<sub>2</sub>Ph), 72.9 (CH<sub>2</sub>Ph), 72.7 (C-4<sup>c</sup>), 72.3 (C-2<sup>c</sup>), 72.1 (CH<sub>2</sub>Ph), 71.8 (C-2<sup>b</sup>, C-2<sup>c</sup>, CH<sub>2</sub> pMBn), 71.0 (C-3<sup>b</sup>), 69.9 (C-5<sup>c</sup> or C-5<sup>b</sup>), 69.4 (C-5<sup>c</sup> or C-5<sup>b</sup>), 69.3 (C-5<sup>a</sup>), 69.0 (C-a), 68.9 (C-5<sup>b</sup>), 68.5 (C-5<sup>c</sup>), 68.4 (C-5<sup>a</sup>), 67.3 (C-6<sup>a</sup>), 67.0 (C-6<sup>b</sup>), 66.9 (C-6<sup>c</sup>), 58.4 (C-2<sup>c</sup>), 58.3 (C-2<sup>b</sup>), 57.5 (C-2<sup>a</sup>), 55.6 (CH<sub>2</sub> pMBn) ppm; ESI MS: C<sub>20</sub>H<sub>26</sub>N<sub>2</sub>O<sub>10</sub>Na<sub>2</sub>;  $m/z$  (%): 422.0 (53) [M-8Na+2H]<sup>+</sup> 6, 511.0 (100) [M-7Na+2H]<sup>+</sup> 5, 649.8 (58) [M-5Na+H]<sup>+</sup> 4, 881.8 (26) [M-3Na]<sup>+</sup> 3.

**Allyl [(3-O-benzyl-4-O-(4-methyloxybenzyl)-2-O-sulfonato- $\alpha$ -L-idopyranosyluronate)-(1 $\rightarrow$ 4)-O-(3-O-benzyl-2-deoxy-2-sulfoamino-6-O-sulfonato- $\alpha$ -D-glucopyranoside)-(1 $\rightarrow$ 4)-O-(3-O-benzyl-2-deoxy-2-sulfoamino-6-O-sulfonato- $\alpha$ -D-glucopyranoside)-(1 $\rightarrow$ 4)-O-(3-O-benzyl-2-deoxy-2-sulfoamino-6-O-sulfonato- $\alpha$ -D-glucopyranoside)-(1 $\rightarrow$ 4)-O-(3-O-benzyl-2-deoxy-2-sulfoamino-6-O-sulfonato- $\alpha$ -D-glucopyranoside)]-dodecasodium salt (23):** Octasaccharide 19 (37 mg, 105  $\mu$ mol) was dissolved in a mixture of LiOH (42 mg, 1 mmol, 95 equiv), water (240  $\mu$ l) and H<sub>2</sub>O (5% in water, 260  $\mu$ l, 3 mmol). After 48 h at 37°C, AcOH (45  $\mu$ l, 0.75 mmol) was added and the mixture was purified as described above to give 23 (30 mg, 80%) as a nona-sodium salt (96% pure by HPLC analysis): <sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O)  $\delta$  = 7.55–7.45 (m, 8H; Ph), 7.45–7.31 (m, 32H; Ph), 7.29 (d,  $J$  = 8.5 Hz, 2H; Ph-OMe), 6.94 (d,  $J$  = 8.5 Hz, 2H; Ph-OMe), 6.01 (ddt,  $J_{6,5}$  = 17.0,  $J_{6,4}$  = 10.5,  $J_{6,3}$  = 6.0 Hz, 1H; H-b), 5.49 (brs, 1H; H-1<sup>b</sup>), 5.47 (brs, 2H; H-1<sup>b</sup>, H-1<sup>c</sup>), 5.39 (dq,  $J_{4,3}$  = 17.0,  $J_{4,2}$  =  $J_{4,1}$  = 1.5 Hz, 1H; H-c), 5.33 (d,  $J_{1,2}$  = 3.5 Hz, 1H; H-1<sup>a</sup>), 5.29 (brs, 1H; H-1<sup>b</sup>), 5.29 (d,  $J_{1,2}$  = 3.5 Hz, 1H; H-1<sup>c</sup>), 5.28 (d,  $J_{1,2}$  = 3.5 Hz, 1H; H-1<sup>b</sup>), 5.27 (dq,  $J_{6,5}$  = 10.5,  $J_{6,4}$  =  $J_{6,3}$  = 1.5 Hz, 1H; H-c), 5.15 (d,  $J_{1,2}$  = 3.5 Hz, 1H; H-1<sup>a</sup>), 4.89 (d,  $J$  = 10.5 Hz, 1H; CH<sub>2</sub>Ph), 4.88 (d,  $J$  = 10.5 Hz, 1H; CH<sub>2</sub>Ph), 4.84 (d,  $J$  = 11.0 Hz, 1H; CH<sub>2</sub>Ph), 4.79 (d,  $J$  = 11.0 Hz, 1H; CH<sub>2</sub>Ph), 4.77 (d,  $J$  = 11.0 Hz, 1H; CH<sub>2</sub>Ph), 4.78–4.76 (m, 1H; H-5<sup>b</sup>), 4.76 (d,  $J$  = 10.0 Hz, 1H; CH<sub>2</sub>Ph), 4.74–4.72 (m, 2H; H-5<sup>b</sup>, H-5<sup>c</sup>), 4.73 (d,  $J$  = 10.5 Hz, 1H; CH<sub>2</sub>Ph), 4.72 (d,  $J$  = 10.5 Hz, 1H; CH<sub>2</sub>Ph), 4.68 (brs, 1H; H-5<sup>b</sup>), 4.67 (d,  $J$  = 11.0 Hz, 1H; CH<sub>2</sub>Ph), 4.64–4.61 (m, 2H; H-2<sup>b</sup>, H-2<sup>c</sup>), 4.61 (d,  $J$  = 12.0 Hz, 1H; CH<sub>2</sub>Ph), 4.61 (d,  $J$  = 10.0 Hz, 1H; CH<sub>2</sub>Ph), 4.53 (s, 2H; 2 $\times$ CH<sub>2</sub>Ph), 4.50 (d,  $J$  = 12.0 Hz, 1H; CH<sub>2</sub>Ph), 4.49 (s, 2H; 2 $\times$ CH<sub>2</sub>Ph), 4.48 (brd,  $J_{6,5}$  = 11.0 Hz, 1H; H-6), 4.45 (brd,  $J_{6,5}$  = 11.0 Hz, 1H; H-6), 4.44 (d,  $J$  = 12.0 Hz, 1H; CH<sub>2</sub>PhOMe), 4.44–4.42 (m, 1H; H-2<sup>b</sup>), 4.40 (d,  $J$  = 12.0 Hz, 1H; CH<sub>2</sub>PhOMe), 4.38 (brs, 2H; 2 $\times$ H-6), 4.36 (brd,  $J$  = 11.0 Hz, 1H; H-6), 4.33 (brd,  $J_{6,5}$  = 10.0 Hz, 1H; H-6), 4.32–4.20 (m, 8H; H-a, H-3<sup>a</sup>, H-3<sup>b</sup>, H-3<sup>c</sup>, H-4<sup>a</sup>, H-4<sup>b</sup>, H-4<sup>c</sup>, 2 H-6), 4.15–4.04 (m, 5H; H-a', H-5<sup>a</sup>, H-5<sup>b</sup>, H-5<sup>c</sup>, H-5<sup>d</sup>), 3.95 (t,  $J_{4,3}$  =  $J_{4,2}$  = 10.0 Hz, 1H; H-4<sup>b</sup>), 3.94 (dd,  $J_{4,3}$  =  $J_{4,2}$  = 10.0 Hz, 1H; H-4<sup>b</sup>), 3.93–3.92 (m, 1H; H-3<sup>b</sup>), 3.93 (dd,  $J_{4,3}$  =  $J_{4,2}$  = 10.0 Hz, 1H; H-4<sup>b</sup>), 3.93 (dd,  $J_{4,3}$  =  $J_{4,2}$  = 10.0 Hz, 1H; H-4<sup>b</sup>), 3.90–3.88 (m, 1H; H-4<sup>b</sup>), 3.83 (s, 3H; PhOMe), 3.82 (t,  $J_{2,3}$  =  $J_{2,4}$  = 10.0 Hz, 1H; H-3<sup>b</sup>), 3.74 (t,  $J_{2,3}$  =  $J_{2,4}$  = 10.0 Hz, 1H; H-3<sup>a</sup>), 3.71 (t,  $J_{2,3}$  =  $J_{2,4}$  = 10.0 Hz, 1H; H-3<sup>c</sup>), 3.70 (t,  $J_{2,3}$  =  $J_{2,4}$  = 10.0 Hz, 1H; H-3<sup>b</sup>), 3.44 (dd,  $J_{2,3}$  = 10.0,  $J_{2,1}$  = 3.5 Hz, 1H; H-2<sup>b</sup>), 3.41 (dd,  $J_{2,3}$  = 10.0,  $J_{2,1}$  = 3.5 Hz, 1H; H-2<sup>c</sup>), 3.40 (dd,  $J_{2,3}$  = 10.0,  $J_{2,1}$  = 3.5 Hz, 1H; H-2<sup>a</sup>), 3.37 (dd,  $J_{2,3}$  = 10.0,  $J_{2,1}$  = 3.5 Hz, 1H; H-2<sup>b</sup>) ppm; <sup>13</sup>C NMR (100.6 MHz, D<sub>2</sub>O)  $\delta$  = 175.5, 174.6, 174.5 (C=O), 158.9 (C-OMe pMBn), 137.8, 137.7, 137.6, 137.4, 137.3, 137.2 (C<sub>quat</sub>), 133.9 (C-b), 130.8, 130.0, 129.5, 129.4, 129.3, 129.1, 129.0, 128.9, 128.8, 128.7, 128.6, 128.5, 128.2 (C<sub>arom</sub>), 118.3 (C-c), 114.2 (C<sub>arom</sub> pMBn), 98.5 (C-1<sup>b</sup>), 98.3 (C-1<sup>c</sup>), 98.1 (C-1<sup>a</sup>), 98.0 (C-1<sup>b</sup>) or C-1<sup>a</sup>), 97.5 (C-1<sup>b</sup>), 97.4 (C-1<sup>a</sup>), 97.2 (C-1<sup>b</sup>), 96.5 (C-1<sup>a</sup>), 77.8 (C-3<sup>c</sup>), 77.2 (C-3<sup>b</sup>), 77.1 (C-3<sup>a</sup>), 76.9 (CH<sub>2</sub>Ph), 76.8 (CH<sub>2</sub>Ph), 75.5/75.3/75.2/75.1/74.9/74.8 (C-3<sup>a</sup>, C-3<sup>b</sup>, C-3<sup>c</sup>, C-4<sup>a</sup>, C-4<sup>b</sup>, C-4<sup>c</sup>, 2 $\times$ CH<sub>2</sub>Ph), 73.3 (C-4<sup>c</sup>, C-4<sup>b</sup>), 73.0 (2 $\times$ CH<sub>2</sub>Ph), 72.9 (C-2<sup>b</sup>, C-2<sup>c</sup>), 72.8 (CH<sub>2</sub>Ph), 72.7 (C-4<sup>b</sup>), 72.1 (CH<sub>2</sub> pMBn), 71.8 (C-2<sup>a</sup>, C-2<sup>b</sup>, C-2<sup>c</sup>, CH<sub>2</sub>Ph), 71.0 (C-3<sup>b</sup>), 70.0/69.9/69.4/69.3 (C-5<sup>a</sup>, C-5<sup>b</sup>, C-5<sup>c</sup>, C-5<sup>d</sup>), 69.1 (C-5<sup>b</sup>), 69.0 (C-a), 68.6 (C-5<sup>b</sup> or C-5<sup>c</sup>), 68.5

(C-5<sup>b</sup>, C-5<sup>c</sup> or C-5<sup>d</sup>), 67.3/67.1/67.0/67.9 (4 $\times$ C-6), 58.4/58.3 (C-2<sup>c</sup>, C-2<sup>b</sup>, C-2<sup>a</sup>), 57.5 (C-2<sup>a</sup>), 55.6 (CH<sub>2</sub> pMBn); ESI MS: C<sub>20</sub>H<sub>26</sub>N<sub>2</sub>O<sub>10</sub>Na<sub>2</sub>;  $m/z$  (%): 479.2 (64) [M-9Na+2H]<sup>+</sup> 7, 565.5 (100) [M-7Na+H]<sup>+</sup> 6, 683.5 (43) [M-6Na+H]<sup>+</sup> 5, 867.1 (28) [M-4Na]<sup>+</sup> 4.

**1-Propyl [(2-O-sulfonato- $\alpha$ -L-idopyranosyluronate)-(1 $\rightarrow$ 4)-O-(2-deoxy-2-sulfoamino-6-O-sulfonato- $\alpha$ -D-glucopyranoside)-(1 $\rightarrow$ 4)-O-(2-O-sulfonato- $\alpha$ -L-idopyranosyluronate)-(1 $\rightarrow$ 4)-O-(2-deoxy-2-sulfoamino-6-O-sulfonato- $\alpha$ -D-glucopyranoside octasodium salt (24):** Pd(OH)<sub>2</sub> (20% on charcoal, 10 mg) was added to a solution of tetrasaccharide 21 (25 mg, 13  $\mu$ mol) in phosphate buffer (20 mM, pH 7.0, 300  $\mu$ l). The mixture was degassed and stirred for 48 h under hydrogen (1 atm), filtered and desalted on a PD-10 column (Pharmacia) to give 24 (1.7 mg, 94%): <sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O)  $\delta$  = 5.43 (d,  $J_{1,2}$  = 3.5 Hz, 1H; H-1<sup>b</sup>), 5.20 (d,  $J_{1,2}$  = 3.5 Hz, 1H; H-1<sup>b</sup>), 5.17 (brs, 1H; H-1<sup>b</sup>), 5.13 (d,  $J_{1,2}$  = 3.5 Hz, 1H; H-1<sup>a</sup>), 4.86 (d,  $J_{4,3}$  = 1.5 Hz, 1H; H-5<sup>b</sup>), 4.74 (d,  $J_{4,3}$  = 3.0 Hz, 1H; H-5<sup>a</sup>), 4.35 (brs, 2H; H-6<sup>a</sup>, H-6<sup>b</sup>), 4.33–4.28 (m, 3H; H-2<sup>b</sup>, H-2<sup>c</sup>, H-6<sup>c</sup>), 4.26 (brd,  $J_{6,5}$  = 11.5 Hz, 1H; H-6<sup>c</sup>), 4.17 (dd,  $J_{4,3}$  = 6.5,  $J_{4,2}$  = 3.5 Hz, 1H; H-3<sup>b</sup>), 4.08 (brt,  $J$  = 3.5 Hz, 2H; H-3<sup>b</sup>, H-4<sup>a</sup>), 4.04–3.99 (m, 2H; H-5<sup>a</sup>, H-5<sup>b</sup>), 3.99–3.96 (m, 1H; H-4<sup>b</sup>), 3.80–3.72 (m, 2H; H-4<sup>a</sup>, H-4<sup>c</sup>), 3.72–3.64 (m, 3H; H-a, H-3<sup>a</sup>, H-3<sup>c</sup>), 3.49 (dt,  $J_{6,5}$  = 10.0,  $J_{6,4}$  = 6.5 Hz, 1H; H-a'), 3.27 (dd,  $J_{2,3}$  = 10.0,  $J_{2,1}$  = 3.5 Hz, 1H; H-2<sup>a</sup>), 3.26 (dd,  $J_{2,3}$  = 10.0,  $J_{2,1}$  = 3.5 Hz, 1H; H-2<sup>b</sup>), 1.62 (brs,  $J$  = 7.0 Hz, 2H; H-b), 0.92 (t,  $J$  = 7.0 Hz, 3H; H-c) ppm; <sup>13</sup>C NMR from HMQC (100.6 MHz, D<sub>2</sub>O):  $\delta$  = 99.8 (C-1<sup>a</sup>, C-1<sup>b</sup>), 97.2 (C-1<sup>a</sup>), 96.7 (C-1<sup>b</sup>), 77.1 (C-3<sup>a</sup> or C-3<sup>b</sup>), 77.0 (C-2<sup>b</sup>), 76.2 (C-4<sup>b</sup>), 74.4 (C-2<sup>b</sup>), 70.9 (C-4<sup>a</sup>, C-4<sup>c</sup>), 70.4 (C-a), 70.1 (C-3<sup>a</sup>), 68.8 (C-5<sup>a</sup>, C-5<sup>b</sup>), 67.2 (C-6<sup>a</sup>), 66.4 (C-6<sup>b</sup>), 58.0 (C-2<sup>c</sup>, C-2<sup>a</sup>), 22.3 (C-b), 10.2 (C-c) ppm; ESI MS: C<sub>23</sub>H<sub>30</sub>N<sub>2</sub>O<sub>10</sub>Na<sub>2</sub>;  $m/z$  (%): 324.5 (62) [M-4Na]<sup>+</sup> 7, 440.3 (100) [M-3Na]<sup>+</sup> 7, 672.0 (75) [M-2Na]<sup>+</sup> 7.

**1-Propyl [(2-O-sulfonato- $\alpha$ -L-idopyranosyluronate)-(1 $\rightarrow$ 4)-O-(2-deoxy-2-sulfoamino-6-O-sulfonato- $\alpha$ -D-glucopyranoside)-(1 $\rightarrow$ 4)-O-(2-O-sulfonato- $\alpha$ -L-idopyranosyluronate)-(1 $\rightarrow$ 4)-O-(2-deoxy-2-sulfoamino-6-O-sulfonato- $\alpha$ -D-glucopyranoside)-(1 $\rightarrow$ 4)-O-(2-O-sulfonato- $\alpha$ -L-idopyranosyluronate)]-dodecasodium salt (25):** Hexasaccharide 22 (2.5 mg, 0.9  $\mu$ mol) was debenzylated as described above to give 25 (1.6 mg, 85%): <sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O)  $\delta$  = 5.44 (d,  $J_{1,2}$  = 3.5 Hz, 1H; H-1<sup>b</sup>), 5.41 (d,  $J_{1,2}$  = 3.5 Hz, 1H; H-1<sup>b</sup>), 5.21 (brd,  $J_{2,3}$  = 3.5 Hz, 1H; H-1<sup>b</sup>), 5.20 (brd,  $J_{2,3}$  = 3.5 Hz, 1H; H-1<sup>b</sup>), 5.17 (brs, 1H; H-1<sup>b</sup>), 5.13 (d,  $J_{1,2}$  = 3.5 Hz, 1H; H-1<sup>a</sup>), 4.86 (d,  $J_{4,3}$  = 1.5 Hz, 1H; H-5<sup>b</sup>), 4.82–4.84 (under HOD peak, H-5<sup>b</sup>), 4.75 (d,  $J_{4,3}$  = 3.0 Hz, 1H; H-5<sup>a</sup>), 4.41 (brd,  $J_{6,5}$  = 11.5 Hz, 1H; H-6<sup>a</sup> or H-6<sup>c</sup> or H-6<sup>b</sup>), 4.37–4.28 (m, 6H; H-2<sup>b</sup>, H-2<sup>c</sup>, H-2<sup>a</sup>, 2 $\times$ H-6<sup>a</sup> or H-6<sup>c</sup> or H-6<sup>b</sup>), 4.37–4.28 (m, 6H; H-2<sup>b</sup>, H-2<sup>c</sup>, H-2<sup>a</sup>, 2 $\times$ H-6<sup>a</sup> or H-6<sup>c</sup> or H-6<sup>b</sup>), 4.37–4.28 (m, 6H; H-2<sup>b</sup>, H-2<sup>c</sup>, H-2<sup>a</sup>, 2 $\times$ H-6<sup>a</sup> or H-6<sup>c</sup> or H-6<sup>b</sup>), 4.19 (dd,  $J_{4,3}$  = 6.5,  $J_{4,2}$  = 3.5 Hz, 1H; H-3<sup>b</sup>), 4.17 (dd,  $J_{4,3}$  = 6.5,  $J_{4,2}$  = 3.5 Hz, 1H; H-3<sup>b</sup>), 4.13–4.08 (m, 3H; H-3<sup>a</sup>, H-4<sup>a</sup>, H-4<sup>b</sup>), 4.08–3.99 (m, 3H; H-5<sup>a</sup>, H-5<sup>b</sup>, H-5<sup>c</sup>), 3.99–3.96 (m, 1H; H-4<sup>a</sup>), 3.81–3.71 (m, 3H; H-4<sup>a</sup>, H-4<sup>b</sup>, H-4<sup>c</sup>), 3.71–3.61 (m, 4H; H-a, H-3<sup>a</sup>, H-3<sup>c</sup>, H-3<sup>b</sup>), 3.49 (dt,  $J_{6,5}$  = 9.5,  $J_{6,4}$  = 6.5 Hz, 1H; H-a'), 3.27 (dd,  $J_{2,3}$  = 10.0,  $J_{2,1}$  = 3.5 Hz, 1H; H-2<sup>b</sup>), 3.26 (dd,  $J_{2,3}$  = 10.0,  $J_{2,1}$  = 3.5 Hz, 2H; H-2<sup>a</sup>, H-2<sup>c</sup>), 1.63 (brs,  $J$  = 7.0 Hz, 2H; H-b), 0.92 (t,  $J$  = 7.0 Hz, 3H; H-c) ppm; <sup>13</sup>C NMR from HMQC (100.6 MHz, D<sub>2</sub>O):  $\delta$  = 99.6 (C-1<sup>a</sup>, C-1<sup>b</sup>), 98.8 (C-1<sup>a</sup>), 97.2 (C-1<sup>a</sup>, C-1<sup>b</sup>), 96.6 (C-1<sup>a</sup>), 77.2 (C-2<sup>b</sup> or C-2<sup>c</sup> or C-2<sup>a</sup>), 77.0 (C-3<sup>a</sup> or C-3<sup>b</sup> or C-3<sup>c</sup>), 76.7 (C-3<sup>a</sup> or C-3<sup>b</sup> or C-3<sup>c</sup>), 76.1 (C-2<sup>b</sup> or C-2<sup>c</sup> or C-2<sup>a</sup>, C-4<sup>a</sup> or C-4<sup>b</sup>), 74.3 (C-2<sup>b</sup> or C-2<sup>c</sup> or C-2<sup>a</sup>), 70.4 (C-a, C-5<sup>a</sup> or C-5<sup>b</sup> or C-5<sup>c</sup>), 70.6 (C-3<sup>b</sup>), 69.5 (C-3<sup>b</sup>), C-5<sup>a</sup> or C-5<sup>c</sup> or C-5<sup>b</sup>), 69.1 (C-3<sup>a</sup>), 69.0 (C-5<sup>b</sup>), 68.8 (C-5<sup>a</sup> or C-5<sup>c</sup> or C-5<sup>b</sup>), 67.2/66.5/66.3 (C-6<sup>a</sup>, C-6<sup>b</sup>, C-6<sup>c</sup>), 58.3 (C-2<sup>a</sup>, C-2<sup>c</sup>, C-2<sup>b</sup>), 22.2 (C-b), 10.4 (C-c) ppm; ESI MS: C<sub>26</sub>H<sub>34</sub>N<sub>2</sub>O<sub>10</sub>Na<sub>2</sub>;  $m/z$  (%): 383.5 (87) [M-6Na+H]<sup>+</sup> 5, 388.0 (73) [M-5Na]<sup>+</sup> 5, 490.7 (100) [M-4Na]<sup>+</sup> 4, 662.0 (78) [M-3Na]<sup>+</sup> 3.

**1-Propyl [(2-O-sulfonato- $\alpha$ -L-idopyranosyluronate)-(1 $\rightarrow$ 4)-O-(2-deoxy-2-sulfoamino-6-O-sulfonato- $\alpha$ -D-glucopyranoside)-(1 $\rightarrow$ 4)-O-(2-O-sulfonato- $\alpha$ -L-idopyranosyluronate)-(1 $\rightarrow$ 4)-O-(2-deoxy-2-sulfoamino-6-O-sulfonato- $\alpha$ -D-glucopyranoside)-(1 $\rightarrow$ 4)-O-(2-O-sulfonato- $\alpha$ -L-idopyranosyluronate)]-hexadecasodium salt (26):** Octasaccharide 23 (2.5 mg, 0.7  $\mu$ mol) was debenzylated as described above to give 26 (1.9 mg, quant): <sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O):  $\delta$  = 5.42 (d,  $J_{1,2}$  = 3.5 Hz, 1H; H-1<sup>b</sup>), 5.41 (d,  $J_{1,2}$  = 3.5 Hz, 1H; H-1<sup>b</sup>), 5.40 (d,  $J_{1,2}$  = 3.5 Hz, 1H; H-1<sup>b</sup>), 5.26 (brd,  $J_{2,3}$  = 2.5 Hz, 1H; H-1<sup>b</sup>), 5.23 (brd,  $J_{2,3}$  = 3.5 Hz, 1H; H-1<sup>b</sup>), 5.22 (brd,  $J_{2,3}$  = 3.5 Hz, 1H; H-1<sup>b</sup>), 5.18 (brs, 1H; H-1<sup>b</sup>), 5.13 (d,  $J_{1,2}$  = 3.5 Hz, 1H; H-1<sup>a</sup>), 4.87–4.84 (partly

under HOD peak, H-5<sup>D</sup>, H-5<sup>F</sup>, H-5<sup>G</sup>), 4.76 (d,  $J_{1,2}$  = 2.5 Hz, 1H; H-5<sup>RO</sup>), 4.40 (brd,  $J_{6,5}$  = 11.5 Hz, 2H; 2 × H-6), 4.36-4.32 (m, 5H; H-2<sup>A</sup>, H-2<sup>RO</sup>), 2 × H-6<sub>a</sub>, H-6<sub>b</sub>), 4.32-4.29 (m, 2H; H-2<sup>RO</sup>, H-2<sup>B</sup>), 4.26 (brd,  $J_{6,5}$  = 11.5, 3H; 3 × H-6), 4.22-4.18 (m, 2H; H-3<sup>A</sup>, H-3<sup>RO</sup>), 4.18 (dd,  $J_{1,2}$  = 6.5,  $J_{3,4}$  = 3.5 Hz, 1H; H-3<sup>RO</sup>), 4.12-4.08 (m, 4H; H-3H, H-4<sup>A</sup>, H-4<sup>F</sup>, H-4<sup>G</sup>), 4.08-3.99 (m, 4H; 4 × H-5), 3.98 (brt,  $J$  = 2.5 Hz, 1H; H-4<sup>D</sup>), 3.82-3.72 (m, 4H; 4 × H-4), 3.72-3.64 (m, 5H; H-a, 4 × H-3), 3.49 (dt,  $J_{6,5}$  = 9.5,  $J_{2,3}$  = 6.5 Hz, 1H; H-a'), 3.27 (dd,  $J_{1,2}$  = 10.0,  $J_{2,3}$  = 3.5 Hz, 2H; 2 × H-2), 3.26 (dd,  $J_{2,3}$  = 10.0,  $J_{2,1}$  = 3.5 Hz, 2H; 2 × H-2), 1.63 (bsexet,  $J$  = 7.0 Hz, 2H; H-b), 0.92 (t,  $J$  = 7.0 Hz, 3H; H-c) ppm; ESI MS:  $C_{22}H_{40}N_4O_{12}S_2Na_4$ ;  $m/z$  (%): 408.6 (100) [ $M$ -12Na+6H]<sup>7+</sup> 76, 408.6 (33) [ $M$ -11Na+5H]<sup>7+</sup> 76, 426.8 (69) [ $M$ -7Na+4H]<sup>7+</sup> 76.

**Glycoconjugate 30a:** In a glass screw-capped tube bis-thioPEG 27 (8.5 μL of a 0.2 M solution in degassed water, 1.7 μmol) was added to lyophilised tetrasaccharide 21 (8 mg, 4.28 μmol, 2.5 equiv). The resulting viscous solution was irradiated with UV for 4 h with periodical centrifugation to allow stirring of the solution, then diluted with water (500 μL). An Ellman's test<sup>17</sup> performed on an aliquot of this solution confirmed completion of the reaction. The glycoconjugate was then directly purified by using reversed-phase semipreparative HPLC, followed by ion exchange on BioRad AG50W-X8 300 (Na<sup>+</sup>, 1.5 mL) resin after removal of AcOH-NEt<sub>3</sub> salts by lyophilisation (2 × 1 mL H<sub>2</sub>O) and desalting with a Pharmacia PD-10 prepac column. Compound 30a (4.6 mg, 66%) was obtained as a hexadecasodium salt (95% pure by HPLC analysis): <sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O) δ = 7.53-7.43 (m, 8H; Ph), 7.43-7.31 (m, 32H; Ph), 7.29 (d,  $J$  = 8.5 Hz, 4H; Ph-OMe), 6.94 (d,  $J$  = 8.5 Hz, 4H; Ph-OMe), 5.42 (brs, 2H; 2 × H-1<sup>A</sup>), 5.32 (d,  $J_{1,2}$  = 3.5 Hz, 2H; 2 × H-1<sup>F</sup>), 5.31 (brs, 2H; 2 × H-1<sup>D</sup>), 5.06 (d,  $J_{1,2}$  = 3.5 Hz, 2H; 2 × H-1<sup>B</sup>), 4.85 (d,  $J$  = 11.0 Hz, 2H; 2 × CH<sub>2</sub>Ph), 4.82-4.79 (under HOD signal, 2 × H-5<sup>F</sup>), 4.77 (d,  $J$  = 11.0 Hz, 4H; 4 × CH<sub>2</sub>Ph), 4.73 (d,  $J$  = 11.0 Hz, 2H; 2 × CH<sub>2</sub>Ph), 4.69 (d,  $J_{1,2}$  = 2.0 Hz, 1H; 2 × H-5<sup>D</sup>), 4.64 (d,  $J$  = 12.5 Hz, 2H; 2 × CH<sub>2</sub>Ph), 4.62 (d,  $J$  = 11.0 Hz, 2H; 2 × CH<sub>2</sub>Ph), 4.62-4.60 (m, 2H; 2 × H-2<sup>A</sup>), 4.58 (d,  $J$  = 11.0 Hz, 2H; 2 × CH<sub>2</sub>Ph), 4.50 (d,  $J$  = 12.5 Hz, 2H; 2 × CH<sub>2</sub>Ph), 4.45 (brs, 2H; 2 × H-2<sup>D</sup>), 4.42 (s, 4H; 4 × CH<sub>2</sub>PhOMe), 4.40 (brd,  $J_{6,5}$  = 12.5, 2H; 2 × H-6<sup>A</sup>), 4.33 (dd,  $J_{6,5}$  = 12.5,  $J_{6,3}$  = 5.0 Hz, 2H; 2 × H-6<sup>B</sup>), 4.32-4.26 (m, 6H; 2 × H-3<sup>A</sup>, 2 × H-6<sup>C</sup>, 2 × H-6<sup>D</sup>), 4.25 (brs, 2H; 2 × H-4<sup>F</sup>), 4.16 (brd,  $J_{1,4}$  = 10.0 Hz, 2H; 2 × H-5<sup>F</sup>), 4.03 (ddd,  $J_{1,4}$  = 10.0,  $J_{1,6}$  = 5.0,  $J_{1,2}$  = 3.0 Hz, 2H; 2 × H-5<sup>D</sup>), 3.98-3.92 (m, 4H; 2 × H-3<sup>D</sup>, 2 × H-4<sup>F</sup>), 3.92-3.86 (m, 4H; 2 × H-4<sup>A</sup>, 2 × H-4<sup>F</sup>), 3.86-3.78 (m with s at δ = 3.83, 10H; 2 × H-3<sup>F</sup>, 2 × H-a, 2 × PhOMe), 3.74-3.58 (m, 26H; 2 × OCH<sub>2</sub>CH<sub>2</sub>S, CH<sub>2</sub> PEG, 2 × H-3<sup>B</sup>), 3.52 (dt,  $J_{6,5}$  = 10.0,  $J_{2,3}$  = 7.0, 2H; 2 × H-a'), 3.44 (dd,  $J_{2,3}$  = 10.5,  $J_{2,1}$  = 3.5 Hz, 2H; 2 × H-2<sup>F</sup>), 3.34 (dd,  $J_{2,3}$  = 10.5,  $J_{2,1}$  = 3.5 Hz, 2H; 2 × H-2<sup>D</sup>), 2.76 (t,  $J$  = 6.5 Hz, 4H; 2 × OCH<sub>2</sub>CH<sub>2</sub>S), 2.70 (brt,  $J$  = 7.0 Hz, 4H; 4 × H-c), 1.95 (brdq,  $J$  = 14.0,  $J$  = 7.0 Hz, 2H; 2 × H-b), 1.90 (brdq,  $J$  = 14.0,  $J$  = 7.0 Hz, 2H; 2 × H-b') ppm; <sup>13</sup>C NMR from HMQC (100.6 MHz, D<sub>2</sub>O); aromatic protons were not included in the spectral window, due to a limited signal-to-noise ratio, some cross-peaks were not detected: δ = 98.3 (C-1<sup>A</sup>), 98.4 (C-1<sup>F</sup>, C-1<sup>D</sup>), 97.2 (C-1<sup>B</sup>), 75.3 (C-4<sup>A</sup>), 75.2 (2 × CH<sub>2</sub>Ph), 75.1 (C-3<sup>A</sup>), 75.6 (C-4<sup>F</sup>), 75.3 (C-4<sup>D</sup>), 72.8 (CH<sub>2</sub>Ph), 72.3 (C-2<sup>A</sup>), 72.0 (CH<sub>2</sub> pMBn), 71.9 (CH<sub>2</sub>Ph), 71.6 (C-2<sup>F</sup>), 71.0 (C-3<sup>B</sup>), 70.7-69.0 (C PEG), 68.8 (C-5<sup>A</sup>), 68.6 (C-5<sup>D</sup>), 67.3 (C-6<sup>A</sup>), 67.0 (C-6<sup>F</sup>), 58.5 (C-2<sup>A</sup>), 57.7 (C-2<sup>F</sup>), 55.6 (CH<sub>2</sub> pMBn), 31.0 (OCH<sub>2</sub>CH<sub>2</sub>S), 29.0 (C-c), 9.5 (C-b) ppm; ESI MS:  $C_{26}H_{48}N_4O_{16}S_2Na_4$ ;  $m/z$  (%): 483.5 (45) [ $M$ -10Na+2H]<sup>7+</sup> 78, 559.1 (80) [ $M$ -8Na+4H]<sup>7+</sup> 77, 656.0 (100) [ $M$ -7Na+3H]<sup>7+</sup> 76, 796.2 (32) [ $M$ -5Na+H]<sup>7+</sup> 75.

**Glycoconjugate 30c:** Bis(thio)PEG 29 (6.5 μL of a 0.2 M solution in degassed water, 1.3 μmol) and lyophilised tetrasaccharide 21 (7 mg, 3.74 μmol, 2.5 equiv) were treated as described above to give compound 30c (2.7 mg, 39%) as a hexadecasodium salt (97% pure by HPLC analysis): The <sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O) spectrum of compound 30c was superimposable on that of 30a except at δ = 3.75-3.60 (m, 134H; 2 × OCH<sub>2</sub>CH<sub>2</sub>S, CH<sub>2</sub> PEG, 2 × H-3<sup>A</sup>) ppm; ESI MS:  $C_{24}H_{44}N_4O_{16}S_2Na_4$ ;  $m/z$  (%): 551.8 (55) [ $M$ -14Na+5H]<sup>7+</sup> 79, 623.9 (79) [ $M$ -13Na+5H]<sup>7+</sup> 78, 632.1 (100) [ $M$ -10Na+2H]<sup>7+</sup> 78, 713.0 (95) [ $M$ -13Na+6H]<sup>7+</sup> 77, 716.1 (91) [ $M$ -12Na+5H]<sup>7+</sup> 77, 846.5 (67) [ $M$ -9Na+3H]<sup>7+</sup> 76, 853.7 (61) [ $M$ -7Na+4H]<sup>7+</sup> 76, 1024.7 (29) [ $M$ -7Na+2H]<sup>7+</sup> 75.

**Glycoconjugate 31a:** Bis(thio)PEG 27 (6.0 μL of a 0.2 M solution in degassed water, 1.2 μmol) and lyophilised hexasaccharide 22 (9 mg, 3.30 μmol, 2.8 equiv) were treated as described above to give compound 31a (3.7 mg, 54%) as a tetracosasodium salt (96% pure by HPLC analysis): <sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O) δ = 7.56-7.49 (m, 8H; Ph), 7.49-7.31 (m, 52H; Ph), 7.28 (d,  $J$  = 8.5 Hz, 4H; Ph-OMe), 6.94 (d,  $J$  = 8.5 Hz, 4H; Ph-

OMe), 5.46 (brs, 4H; 2 × H-1<sup>A</sup>, 2 × H-1<sup>D</sup>), 5.31 (d,  $J_{1,2}$  = 3.5 Hz, 2H; 2 × H-1<sup>RO</sup>), 5.30-5.28 (m, 2H; 2 × H-1<sup>F</sup>), 5.27 (d,  $J_{1,2}$  = 3.5 Hz, 2H; 2 × H-1<sup>RO</sup>), 5.08 (d,  $J_{1,2}$  = 3.5 Hz, 2H; 2 × H-1<sup>B</sup>), 4.89 (d,  $J$  = 11.0 Hz, 2H; 2 × CH<sub>2</sub>Ph), 4.86 (d,  $J$  = 11.0 Hz, 2H; 2 × CH<sub>2</sub>Ph), 4.84 (brs, 2H; 2 × H-5<sup>RO</sup>), 4.77 (d,  $J$  = 11.0 Hz, 2H; 2 × CH<sub>2</sub>Ph), 4.79 (brs, 2H; 2 × H-5<sup>RO</sup>), 4.77-4.72 (under HOD signal, 4 × CH<sub>2</sub>Ph), 4.72 (d,  $J$  = 11.0 Hz, 2H; 2 × CH<sub>2</sub>Ph), 4.69 (brs, 2H; 2 × H-5<sup>F</sup>), 4.68 (d,  $J$  = 11.0 Hz, 2H; 2 × CH<sub>2</sub>Ph), 4.65-4.59 (m, 10H; 2 × H-2<sup>A</sup>, 2 × H-2<sup>D</sup>, 6 × CH<sub>2</sub>Ph), 4.56 (d,  $J$  = 11.0 Hz, 2H; 2 × CH<sub>2</sub>Ph), 4.48 (d,  $J$  = 12.5 Hz, 2H; 2 × CH<sub>2</sub>Ph), 4.46 (m, 10H; 2 × H-2<sup>F</sup>, 4 × H-6<sub>a</sub>, 4 × CH<sub>2</sub>PhOMe), 4.38 (d,  $J$  = 11.0 Hz, 2H; 2 × CH<sub>2</sub>Ph), 4.33 (dd,  $J_{6,5}$  = 11.0,  $J_{6,3}$  = 5.5 Hz, 2H; 2 × H-6<sub>b</sub>), 4.33-4.27 (m, 6H; 2 × H-3<sup>A</sup>, 2 × H-3<sup>D</sup>, 2 × H-6<sub>c</sub>), 4.27-4.21 (m, 6H; 2 × H-4<sup>RO</sup>, 4 × H-6), 4.16 (brs, 2H; 2 × H-4<sup>RO</sup>), 4.08-4.00 (m, 6H; 2 × H-5<sup>A</sup>, 2 × H-5<sup>C</sup>, 2 × H-5<sup>F</sup>), 3.96 (brt,  $J$  = 2.0 Hz, 2 × H-3<sup>D</sup>), 3.95-3.92 (m, 6H; 2 × H-4<sup>A</sup>, 2 × H-4<sup>C</sup>, 2 × H-4<sup>F</sup>), 3.86 (brs, 2H; 2 × H-4<sup>F</sup>), 3.86-3.79 (m with s at δ = 3.83, 8H; 2 × H-a, 2 × PhOMe), 3.77 (brt,  $J$  = 10.5, 2H; 2 × H-3<sup>RO</sup>), 3.71 (t,  $J$  = 6.5 Hz, 4H; 2 × OCH<sub>2</sub>CH<sub>2</sub>S), 3.74-3.60 (m, 24H; CH<sub>2</sub> PEG, 2 × H-3<sup>A</sup>, 2 × H-3<sup>RO</sup>), 3.56 (dt,  $J_{6,5}$  = 10.0,  $J_{2,3}$  = 7.0, 2H; 2 × H-a'), 3.43 (dd,  $J_{2,3}$  = 10.5,  $J_{2,1}$  = 3.5 Hz, 2H; 2 × H-2<sup>RO</sup>), 3.41 (dd,  $J_{2,3}$  = 10.5,  $J_{2,1}$  = 3.5 Hz, 2H; 2 × H-2<sup>RO</sup>), 3.37 (dd,  $J_{2,3}$  = 10.5,  $J_{2,1}$  = 3.5 Hz, 2H; 2 × H-2<sup>A</sup>), 2.77 (t,  $J$  = 6.5 Hz, 4H; 2 × OCH<sub>2</sub>CH<sub>2</sub>S), 2.72 (brt,  $J$  = 7.0 Hz, 4H; 4 × H-c), 1.96 (brdq,  $J$  = 14.0,  $J$  = 7.0 Hz, 2H; 2 × H-b), 1.91 (brdq,  $J$  = 14.0,  $J$  = 7.0 Hz, 2H; 2 × H-b') ppm; <sup>13</sup>C NMR from HMQC (100.6 MHz, D<sub>2</sub>O); the spectral width did not include aromatic protons; due to a limited signal-to-noise ratio, some cross-peaks were not detected: δ = 98.7 (C-1<sup>C</sup>, C-1<sup>F</sup>), 98.5 (C-1<sup>A</sup>, C-1<sup>D</sup>), 97.3 (C-1<sup>B</sup>), 75.3 (C-4<sup>RO</sup>), 75.1 (C-4<sup>RO</sup>), 75.2 (2 × CH<sub>2</sub>Ph), 75.3 (C-4<sup>RO</sup>), 75.1 (C-3<sup>RO</sup>, C-4<sup>RO</sup>), 74.8 (C-3<sup>RO</sup>), 73.1 (C-4<sup>F</sup>), 72.9 (2 × CH<sub>2</sub>Ph), 72.0 (C-2<sup>F</sup>, CH<sub>2</sub> pMBn), 72.1 (CH<sub>2</sub>Ph, C-2<sup>F</sup>, C-2<sup>F</sup>), 71.0 (C-3<sup>F</sup>), 70.7-69.0 (C PEG), 68.7 (C-a), 68.6 (C-5<sup>RO</sup>, C-5<sup>F</sup>), 68.6 (C-5<sup>F</sup>), 66.9 (C-6), 55.6 (CH<sub>2</sub> pMBn), 31.1 (OCH<sub>2</sub>CH<sub>2</sub>S), 29.0 (C-c) ppm; ESI MS:  $C_{28}H_{50}N_4O_{18}S_2Na_4$ ;  $m/z$  (%): 553.0 (43) [ $M$ -11Na+H]<sup>7+</sup> 10, 617.0 (84) [ $M$ -10Na+H]<sup>7+</sup> 9, 697.2 (100) [ $M$ -9Na+2H]<sup>7+</sup> 8, 803.4 (68) [ $M$ -7Na]<sup>7+</sup> 7, 940.1 (32) [ $M$ -6Na]<sup>7+</sup> 6.

**Glycoconjugate 31c:** Bis(thio)PEG 29 (6.0 μL of a 0.2 M solution in degassed water, 1.2 μmol) and lyophilised hexasaccharide 22 (9 mg, 3.30 μmol, 2.8 equiv) were treated as described above to give compound 31c (4.7 mg, 56%) as a tetracosasodium salt (96% pure by HPLC analysis): The <sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O) spectrum of compound 31c was superimposable on that of 31a except at δ = 3.76-3.60 (m, 136H; CH<sub>2</sub> PEG, 2 × OCH<sub>2</sub>CH<sub>2</sub>S, 2 × H-3<sup>A</sup>, 2 × H-3<sup>RO</sup>) ppm; ESI MS:  $C_{28}H_{50}N_4O_{18}S_2Na_4$ ;  $m/z$  (%): 547.3 (35) [ $M$ -18Na+6H]<sup>7+</sup> 12, 596.8 (59) [ $M$ -18Na+7H]<sup>7+</sup> 11, 660.7 (81) [ $M$ -16Na+6H]<sup>7+</sup> 10, 730.3 (100) [ $M$ -18Na+9H]<sup>7+</sup> 9, 737.1 (92) [ $M$ -15Na+6H]<sup>7+</sup> 9, 746.9 (80) [ $M$ -11Na+2H]<sup>7+</sup> 9, 840.5 (65) [ $M$ -11Na+3H]<sup>7+</sup> 8, 851.1 (35) [ $M$ -14Na+7H]<sup>7+</sup> 7.

**Glycoconjugate 32a:** Bis(thio)PEG 27 (4.5 μL of a 0.2 M solution in degassed water, 0.9 μmol) and lyophilised octasaccharide 23 (9 mg, 2.52 μmol, 2.8 equiv) were treated as described above to give compound 32a (2.3 mg, 34%) as a dotriacontasodium salt (92% pure by HPLC analysis): The <sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O) spectrum of compound 32a was superimposable on that of 32b except at δ = 3.75-3.60 (m, 26H; CH<sub>2</sub> PEG, 2 × H-3<sup>A</sup>, 4 × H-3<sup>C</sup> or H-3<sup>B</sup> or H-3<sup>F</sup>) ppm; ESI MS:  $C_{34}H_{62}N_4O_{20}S_2Na_4$ ;  $m/z$  (%): 594.1 (50) [ $M$ -15Na+3H]<sup>7+</sup> 12, 652.4 (88) [ $M$ -13Na+2H]<sup>7+</sup> 11, 719.7 (81) [ $M$ -12Na+2H]<sup>7+</sup> 10, 724.4 (77) [ $M$ -10Na]<sup>7+</sup> 10, 807.4 (100) [ $M$ -9Na]<sup>7+</sup> 9, 911.0 (76) [ $M$ -8Na]<sup>7+</sup> 8.

**Glycoconjugate 32b:** Bis(thio)PEG 28 (12 μL of a 0.2 M solution in degassed water, 2.4 μmol) and lyophilised octasaccharide 23 (34 mg, 6.70 μmol, 2.8 equiv) were treated as described above to give compound 32b (94 mg, 51%) as a dotriacontasodium salt (97% pure by HPLC analysis): <sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O) δ = 7.58-7.49 (m, 8H; Ph), 7.43-7.26 (m, 76H; Ph, Ph-OMe), 6.94 (d,  $J$  = 8.5 Hz, 4H; Ph-OMe), 5.49 (brs, 2H; 2 × H-1<sup>RO</sup>), 5.46 (brs, 2H; 2 × H-1<sup>RO</sup>), 5.45 (brs, 2H; 2 × H-1<sup>RO</sup>), 5.32 (d,  $J_{1,2}$  = 3.5 Hz, 2H; 2 × H-1<sup>RO</sup>), 5.28 (brs, 6H; 2 × H-1<sup>A</sup>, 2 × H-1<sup>RO</sup>), 5.08 (d,  $J_{1,2}$  = 3.5 Hz, 2H; 2 × H-1<sup>B</sup>), 4.91 (d,  $J$  = 10.0 Hz, 2H; 2 × CH<sub>2</sub>Ph), 4.88 (d,  $J$  = 10.0 Hz, 2H; 2 × CH<sub>2</sub>Ph), 4.86 (d,  $J$  = 11.5 Hz, 2H; 2 × CH<sub>2</sub>Ph), 4.85-4.79 (under HOD signal, 2 × H-5<sup>F</sup>, 2 × H-5<sup>D</sup>, 8 × CH<sub>2</sub>Ph), 4.78 (d,  $J$  = 11.5 Hz, 2H; 2 × CH<sub>2</sub>Ph), 4.74 (d,  $J$  = 10.0 Hz, 2H; 2 × CH<sub>2</sub>Ph), 4.73 (d,  $J$  = 10.0 Hz, 2H; 2 × CH<sub>2</sub>Ph), 4.72 (d,  $J$  = 10.0 Hz, 2H; 2 × CH<sub>2</sub>Ph), 4.69 (brs, 2H; 2 × H-5<sup>RO</sup>), 4.67 (d,  $J$  = 10.0 Hz, 2H; 2 × CH<sub>2</sub>Ph), 4.65-4.59 (m, 8H; H-2<sup>A</sup>, H-2<sup>D</sup>, H-2<sup>F</sup>, 2 × CH<sub>2</sub>Ph), 4.59 (d,  $J$  = 10.0 Hz, 4H; 4 × CH<sub>2</sub>Ph), 4.55-4.49 (m, 2H; 2 × H-6<sub>a</sub>), 4.51 (d,  $J$  = 12.0 Hz, 2H; 2 × CH<sub>2</sub>Ph), 4.46-4.39 (m, 10H; 2 × H-2<sup>F</sup>, 4 × H-6<sub>b</sub>, 2 × CH<sub>2</sub>PhOMe)

4.30–4.21 (m, 18H; 2×H-3<sup>a</sup>, 2×H-3<sup>b</sup>, 2×H-3<sup>c</sup>, 2×H-4<sup>OPh</sup>), 10×H-6), 4.20–4.14 (brs, 4H; 2×H-4<sup>OPh</sup>), 2×H-4<sup>OPh</sup>), 4.08–4.00 (m, 8H; 2×H-5<sup>a</sup>, 2×H-5<sup>b</sup>, 2×H-5<sup>c</sup>, 2×H-5<sup>d</sup>), 3.96 (brs, 2H; 2×H-3<sup>d</sup>), 3.96–3.86 (m, 8H; 2×H-4<sup>a</sup>, 2×H-4<sup>b</sup>, 2×H-4<sup>c</sup>, 2×H-4<sup>d</sup>), 3.86 (brs, 2H; 2×H-4<sup>e</sup>), 3.86–3.79 (m with s at δ=3.83, 8H; 2×H-a, 2×PhOMe), 3.76 (brt, J=10.0, 2H; 2×H-3<sup>c</sup> or H-3<sup>b</sup> or H-3<sup>d</sup>), 3.72 (t, J=6.5 Hz, 4H; 2×OCH<sub>2</sub>CH<sub>2</sub>S), 3.75–3.60 (m, 46H; CH<sub>2</sub> PEG; 2×H-3<sup>a</sup>, 4×H-3<sup>c</sup> or H-3<sup>b</sup> or H-3<sup>d</sup>), 3.56 (brd, J<sub>gem</sub>=10.0, J<sub>ax</sub>=7.0 Hz, 2H; 2×H-a), 3.45–3.38 (m, 6H; 2×H-2<sup>a</sup>, 2×H-2<sup>b</sup>, 2×H-2<sup>c</sup>), 3.37 (dd, J<sub>1,2</sub>=10.0, J<sub>2,3</sub>=3.5 Hz, 2H; 2×H-2<sup>a</sup>), 2.78 (t, J=6.5 Hz, 4H; 2×OCH<sub>2</sub>CH<sub>2</sub>S), 2.73 (brt, J=7.0 Hz, 4H; 4×H-c), 1.97 (brdq, J=14.0, J=7.0 Hz, 2H; 2×H-b), 1.92 (brdq, J=14.0, J=7.0 Hz, 2H; 2×H-b) ppm; <sup>13</sup>C NMR from HMQC (100.6 MHz, D<sub>2</sub>O); due to a limited signal-to-noise ratio, some cross-peaks were not detected: δ=130.8 (C<sub>q</sub> pMBn), 130.5–128.0 (C<sub>aromatic</sub>), 114.3 (C<sub>q</sub> pMBn), 98.9 (C-1<sup>OPh</sup>), 98.6 (C-1<sup>OPh</sup>), C-1<sup>OPh</sup>), 97.3 (C-1<sup>a</sup>), 75.5 (CH<sub>2</sub>Ph), 75.2 (C-4<sup>OPh</sup>), 75.1 (C-3<sup>OPh</sup>), C-4<sup>OPh</sup>), C-4<sup>OPh</sup>), CH<sub>2</sub>Ph), 75.0 (C-3<sup>OPh</sup>), 74.9 (C-3<sup>OPh</sup>), 73.6–72.4 (CH<sub>2</sub>Ph), 73.3 (C-4<sup>b</sup>), 72.1 (C-2<sup>a</sup>, C-2<sup>b</sup>, C-2<sup>c</sup>), 72.0 (C-2<sup>d</sup>, CH<sub>2</sub> pMBn), 71.8 (CH<sub>2</sub>Ph), 71.0 (C-3<sup>b</sup>), 70.7–69.0 (C PEG), 68.6 (C-5<sup>OPh</sup>), C-5<sup>b</sup>), 68.5 (C-5<sup>OPh</sup>), C-5<sup>OPh</sup>), 67.0 (C-6), 58.758.5 (C-2<sup>a</sup>, C-2<sup>c</sup>, C-2<sup>e</sup>), 55.5 (CH<sub>2</sub> pMBn), 31.0 (OCH<sub>2</sub>CH<sub>2</sub>S), 29.0 (C-q) ppm; ESI MS: C<sub>22</sub>H<sub>30</sub>N<sub>4</sub>O<sub>26</sub>S<sub>4</sub>Na<sub>2</sub>; m/z (%): 6149 (74) [M-14Na+2H]<sup>2+</sup>/12, 6721 (95) [M-13Na+2H]<sup>2+</sup>/11, 7444 (100) [M-11Na+H]<sup>+</sup>/10, 8321 (96) [M-9Na]<sup>+</sup>/8, 9362 (45) [M-9Na+H]<sup>+</sup>/8, 9392 (47) [M-8Na]<sup>+</sup>/8.

Glycoconjugate 32c: Bis(thio)PEG 29 (4.5 μL of a 0.2 M solution in de-gassed water, 0.9 μmol) and lyophilised octasaccharide 23 (9 mg, 2.52 μmol, 2.8 equiv) were treated as described above to give compound 32c (50 mg, 64%) as a dodecaoctasodium salt (97% pure by HPLC analysis). The <sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O) spectrum of compound 32c was superimposable on that of 30a except at δ=3.77–3.60 (m, 138H; 2×OCH<sub>2</sub>CH<sub>2</sub>S, CH<sub>2</sub> PEG, 2×H-3<sup>a</sup>, 4×H-3<sup>c</sup> or H-3<sup>b</sup> or H-3<sup>d</sup>) ppm; ESI MS: C<sub>22</sub>H<sub>30</sub>N<sub>4</sub>O<sub>26</sub>S<sub>4</sub>Na<sub>2</sub>; m/z (%): 6881 (91) [M-18Na+6H]<sup>2+</sup>/12, 7565 (100) [M-15Na+4H]<sup>2+</sup>/11, 8251 (100) [M-18Na+8H]<sup>2+</sup>/10, 8367 (75) [M-13Na+3H]<sup>2+</sup>/10, 9376 (59) [M-10Na+H]<sup>+</sup>/7.

Glycoconjugate 33a: Bis(thio)PEG 28 (7 μL of a 0.2 M solution in de-gassed water, 1.4 μmol) and lyophilised tetrasaccharide 21 (6.4 mg, 3.50 μmol, 2.5 equiv) were irradiated with UV as described above. The intermediate glycoconjugate 30b was diluted with water (60 μL) and K<sub>2</sub>HPO<sub>4</sub> (30 μL, 0.8 mM in water) and oxidised with oxone (potassium monopersulfate triple salt, 56 μL, 0.2 M in water, 11.2 μmol, 8 equiv). After 4 h at room temperature, excess oxidant was reduced with Na<sub>2</sub>S<sub>2</sub>O<sub>4</sub> (84 μL, 0.4 M in water). The mixture was then directly purified by using reversed-phase semipreparative HPLC, followed by ion exchange on BioRad AG50W-X8200 (Na<sup>+</sup>, 1.5 mL) resin after removal of AcOH-NH<sub>4</sub> salts by lyophilisation (2×1 mL H<sub>2</sub>O) and desalting by using a Pharmacia PD-10 prepac column. Compound 33a (1.4 mg, 23%) was obtained as a hexadecasodium salt (96% pure by HPLC analysis): <sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O): δ=7.53–7.45 (m, 8H; Ph), 7.45–7.31 (m, 32H; Ph), 7.30 (d, J=8.5 Hz, 4H; Ph-OMe), 6.94 (d, J=8.5 Hz, 4H; Ph-OMe), 5.44 (brs, 2H; 2×H-1<sup>OPh</sup>), 5.37 (brs, 2H; 2×H-1<sup>OPh</sup>), 5.31 (brs, 2H; 2×H-1<sup>c</sup>), 5.09 (brs, 2H; 2×H-1<sup>a</sup>), 4.80 (under HOD signal, 2×H-5<sup>a</sup>, 2×H-5<sup>b</sup>, 2×CH<sub>2</sub>Ph), 4.76 (d, J=11.0 Hz, 4H; 4×CH<sub>2</sub>Ph), 4.73 (d, J=12.0 Hz, 2H; 2×CH<sub>2</sub>Ph), 4.65 (d, J=11.0 Hz, 2H; 2×CH<sub>2</sub>Ph), 4.63 (d, J=12.0 Hz, 2H; 2×CH<sub>2</sub>Ph), 4.60 (brs, 2H; 2×H-2<sup>OPh</sup>), 4.60 (d, J=12.0 Hz, 2H; 2×CH<sub>2</sub>Ph), 4.51 (d, J=12.0 Hz, 2H; 2×CH<sub>2</sub>Ph), 4.45 (brs, 2H; 2×H-2<sup>OPh</sup>), 4.42 (m, 4H; 4×CH<sub>2</sub>PhOMe), 4.40 (brd, J<sub>ax</sub>=11.0, 2H; 2×H-6<sup>a</sup>), 4.33 (dd, J<sub>ax</sub>=11.0, J<sub>ax</sub>=5.0 Hz, 2H; 2×H-6<sup>a</sup>), 4.32–4.26 (m, 6H; 2×H-3<sup>OPh</sup>), 4×H-6), 4.24 (brs, 2H; 2×H-4<sup>OPh</sup>), 4.19 (brs, 2H; 2×H-5), 4.05–3.94 (m, 10H; 2×H-3<sup>OPh</sup>, 2×H-4<sup>c</sup>, 2×H-5<sup>a</sup>, 2×OCH<sub>2</sub>CH<sub>2</sub>SO<sub>2</sub>), 3.94–3.84 (m, 8H; 2×H-3<sup>c</sup>, 2×H-4<sup>b</sup>, 2×H-4<sup>d</sup>), 2×H-a), 3.83 (s, 6H; 2×PhOMe), 3.73–3.63 (m, 42H; CH<sub>2</sub> PEG, 2×H-3<sup>a</sup>), 3.59 (dt, J<sub>gem</sub>=10.0, J<sub>ax</sub>=6.0, 2H; 2×H-a), 3.51 (t, J=5.0 Hz, 4H; 2×OCH<sub>2</sub>CH<sub>2</sub>SO<sub>2</sub>), 3.45 (brd, J<sub>1,2</sub>=10.0 Hz, 2H; 2×H-2<sup>c</sup>), 3.39 (brt, J=7.0 Hz, 4H; 4×H-c), 3.36 (brd, J<sub>1,2</sub>=10.0 Hz, 2H; 2×H-2<sup>a</sup>), 2.22–2.10 (m, 4H; 4×H-b) ppm; ESI MS: C<sub>22</sub>H<sub>30</sub>N<sub>4</sub>O<sub>26</sub>S<sub>4</sub>Na<sub>2</sub>; m/z (%): 5182 (65) [M-10Na+2H]<sup>2+</sup>/8, 5994 (93) [M-8Na+H]<sup>+</sup>/7, 6021 (100) [M-7Na]<sup>+</sup>/7, 7034 (75) [M-7Na+H]<sup>+</sup>/6, 7064 (87) [M-6Na]<sup>+</sup>/6, 8527 (58) [M-5Na]<sup>+</sup>/5.

Glycoconjugate 34b: Bis(thio)PEG 28 (6.5 μL of a 0.2 M solution in de-gassed water, 1.3 μmol) and lyophilised hexasaccharide 22 (8.0 mg, 2.9 μmol, 2.3 equiv) were treated as described for compound 33a. Glycoconjugate 34b (24 mg, 31%) was thus obtained as a tetracosasodium salt (98% pure by HPLC analysis): <sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O): δ=7.55–

7.45 (m, 12H; Ph), 7.45–7.32 (m, 48H; Ph), 7.29 (d, J=8.5 Hz, 4H; Ph-OMe), 6.93 (d, J=8.5 Hz, 4H; Ph-OMe), 5.48 (brs, 2H; 2×H-1<sup>OPh</sup>), 5.42 (brs, 2H; 2×H-1<sup>OPh</sup>), 5.32 (brs, 4H; 2×H-1<sup>OPh</sup>, 2×H-1<sup>b</sup>), 5.28 (brs, 2H; 2×H-1<sup>OPh</sup>), 5.09 (brs, 2H; 2×H-1<sup>a</sup>), 4.88 (d, J=12.0 Hz, 2H; 2×CH<sub>2</sub>Ph), 4.86–4.79 (under HOD signal, 2×H-5<sup>a</sup>, 2×H-5<sup>b</sup>, 2×H-5<sup>c</sup>, 4×CH<sub>2</sub>Ph), 4.78 (d, J=11.0 Hz, 2H; 2×CH<sub>2</sub>Ph), 4.75 (d, J=12.0 Hz, 2H; 2×CH<sub>2</sub>Ph), 4.72 (d, J=12.0 Hz, 2H; 2×CH<sub>2</sub>Ph), 4.66 (d, J=11.0 Hz, 2H; 2×CH<sub>2</sub>Ph), 4.64–4.58 (m, 8H; 2×H-2<sup>a</sup>, 2×H-2<sup>b</sup>, 4×CH<sub>2</sub>Ph), 4.57 (d, J=12.0 Hz, 2H; 2×CH<sub>2</sub>Ph), 4.49 (d, J=11.0 Hz, 4H; 4×CH<sub>2</sub>Ph), 4.46–4.38 (m, 10H; 2×H-2<sup>c</sup>, 2×H-6<sup>a</sup>, 2×H-6<sup>c</sup> or H-6<sup>b</sup>, 4×CH<sub>2</sub>PhOMe), 4.33 (dd, J<sub>ax</sub>=12.0, J<sub>ax</sub>=5.0 Hz, 2H; 2×H-6<sup>a</sup>), 4.32–4.23 (m, 12H; 2×H-3<sup>c</sup>, 2×H-3<sup>d</sup>, 6×H-6), 4.22 (brs, 2H; 2×H-4<sup>OPh</sup>), 4.16–4.08 (m, 4H; 2×H-5<sup>c</sup>, 2×H-5<sup>b</sup>), 4.04–4.00 (m, 2H; 2×H-5<sup>a</sup>), 3.99 (brt, J=5.0 Hz, 4H; 2×OCH<sub>2</sub>CH<sub>2</sub>SO<sub>2</sub>), 3.98–3.84 (m, 14H; 2×H-3<sup>d</sup>, 2×H-4<sup>a</sup>, 2×H-4<sup>c</sup>, 2×H-4<sup>e</sup>, 2×H-a, 2×H-4<sup>b</sup>, 2×H-3<sup>OPh</sup>), 3.82 (s, 6H; 2×PhOMe), 3.73–3.63 (m, 44H; CH<sub>2</sub> PEG, 2×H-3<sup>a</sup>, 2×H-3<sup>c</sup>), 3.58 (dt, J<sub>gem</sub>=10.0, J<sub>ax</sub>=6.0 Hz, 2H; 2×H-a), 3.51 (t, J=5.0 Hz, 4H; 2×OCH<sub>2</sub>CH<sub>2</sub>SO<sub>2</sub>), 3.44 (brd, J<sub>1,2</sub>=10.0 Hz, 2H; 2×H-2<sup>OPh</sup>), 3.38 (brt, J=7.0 Hz, 4H; 4×H-c), 3.38 (brd, J<sub>1,2</sub>=10.0 Hz, 2H; 2×H-2<sup>OPh</sup>), 3.36 (brd, J<sub>1,2</sub>=10.0 Hz, 2H; 2×H-2<sup>a</sup>), 2.21–2.10 (m, 4H; 4×H-b) ppm; ESI MS: C<sub>22</sub>H<sub>30</sub>N<sub>4</sub>O<sub>26</sub>S<sub>4</sub>Na<sub>2</sub>; m/z (%): 581.0 (64) [M-11Na+H]<sup>+</sup>/10, 6488 (100) [M-10Na+H]<sup>+</sup>/9, 735.4 (85) [M-8Na]<sup>+</sup>/8, 841.3 (68) [M-8Na+H]<sup>+</sup>/7.

Glycoconjugate 1a: Compound 30a (4.65 mg, 0.88 μmol) was oxidised as described for compound 30b to give compound 33a. After desalting, Pd(OH)<sub>2</sub> (20% on charcoal, 25 mg) was added to a solution of 33a in phosphate buffer (100 mM, pH 7.0, 300 μL) and MeOH (200 μL). The mixture was degassed and stirred for 6 d under hydrogen (1 atm). The mixture was filtered over ultrafree-MC filters (Amicon) and the charcoal was rinsed with CH<sub>3</sub>CN/H<sub>2</sub>O (1:1, 3×200 μL) in order to desorb compound 33a from the charcoal. The resulting solution was desalted on a PD-10 column (Pharmacia) and lyophilised to give 33a (2.4 mg, 85%): <sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O): δ=5.42 (d, J<sub>1,2</sub>=3.5 Hz, 2H; 2×H-1<sup>c</sup>), 5.19 (d, J<sub>1,2</sub>=2.5 Hz, 2H; 2×H-1<sup>b</sup>), 5.17 (brs, 2H; 2×H-1<sup>d</sup>), 5.15 (d, J<sub>1,2</sub>=3.5 Hz, 2H; 2×H-1<sup>a</sup>), 4.81 (under HOD peak, 2×H-5<sup>a</sup>), 4.75 (brs, 2H; 2×H-5<sup>b</sup>), 4.38–4.33 (m, 4H; 2×H-6<sup>a</sup>, 2×H-6<sup>c</sup>), 4.33–4.28 (m, 6H; 2×H-2<sup>a</sup>, 2×H-2<sup>b</sup>, 2×H-6<sup>c</sup>), 4.26 (brd, J<sub>ax</sub>=11.5 Hz, 2H; 2×H-6<sup>a</sup>), 4.18 (dd, J<sub>1,2</sub>=6.0, J<sub>1,4</sub>=4.0 Hz, 2H; 2×H-3<sup>c</sup>), 4.13–4.08 (m, 4H; 2×H-3<sup>d</sup>, 2×H-4<sup>b</sup>), 4.07–3.96 (m, 10H; 2×H-5<sup>c</sup>, 2×OCH<sub>2</sub>CH<sub>2</sub>SO<sub>2</sub> (δ=4.02, t, J=5.5 Hz), 2×H-5<sup>a</sup>, 2×H-4<sup>d</sup>), 3.87 (dt, J<sub>gem</sub>=10.0, J<sub>ax</sub>=5.5 Hz, 2H; 2×H-a), 3.80–3.60 (m, 30H; 2×H-4<sup>a</sup>, 2×H-4<sup>c</sup>, 2×CH<sub>2</sub>PEG, 2×H-3<sup>a</sup>, 2×H-3<sup>c</sup>, 2×H-a), 3.55 (t, J=5.5 Hz, 4H; 2×OCH<sub>2</sub>CH<sub>2</sub>SO<sub>2</sub>), 3.42 (dd, J=9.0, 7.0 Hz, 4H; 4×H-c), 3.28 (dd, J<sub>1,2</sub>=10.0, J<sub>1,3</sub>=3.5 Hz, 2H; 2×H-2<sup>a</sup>), 3.26 (dd, J<sub>1,2</sub>=10.0, J<sub>1,3</sub>=3.5 Hz, 2H; 2×H-2<sup>b</sup>), 2.21–2.11 (m, 4H; 4×H-b) ppm; <sup>13</sup>C NMR from HMQC (100.6 MHz, D<sub>2</sub>O): δ=99.7 (C-1<sup>b</sup>), 99.4 (C-1<sup>c</sup>), 97.3 (C-1<sup>a</sup>), 96.6 (C-1<sup>d</sup>), 76.9 (C-4<sup>c</sup>, C-4<sup>d</sup>), 76.5 (C-2<sup>OPh</sup>), 76.1 (C-4<sup>a</sup>), 74.1 (C-2<sup>b</sup>), 69.9 (CH<sub>2</sub> PEG, C-5<sup>a</sup>, C-4<sup>a</sup>, C-4<sup>b</sup>), 69.6 (C-3<sup>b</sup>), 69.4 (C-5<sup>b</sup>), 69.3 (C-4<sup>b</sup>, C-5<sup>b</sup>), 69.1 (C-3<sup>c</sup>), 69.0 (C-5<sup>c</sup>), 67.1 (C-6<sup>OPh</sup>), 66.6 (C-6<sup>OPh</sup>), 66.3 (C-a), 63.7 (OCH<sub>2</sub>CH<sub>2</sub>SO<sub>2</sub>), 58.2 (C-2<sup>c</sup>, C-2<sup>e</sup>), 52.3 (OCH<sub>2</sub>CH<sub>2</sub>SO<sub>2</sub>), 51.2 (C-c), 21.5 (C-b) ppm.

Glycoconjugate 1b: Compound 33b (0.9 mg, 0.2 μmol) was de-benzylated and purified as described for 33a to give 1b (0.7 mg, quant.). The <sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O) spectrum of compound 1b was superimposable on that of 1a except at δ=3.82–3.60 (m, 50H; 2×H-4<sup>a</sup>, 2×H-4<sup>c</sup>, CH<sub>2</sub> PEG, 2×H-3<sup>a</sup>, 2×H-3<sup>c</sup>, 2×H-a) ppm; ESI MS: C<sub>12</sub>H<sub>12</sub>N<sub>4</sub>O<sub>17</sub>S<sub>4</sub>Na<sub>2</sub>; m/z (%): 401.8 (80) [M-9Na+H]<sup>+</sup>/7, 4593 (100) [M-9Na+2H]<sup>2+</sup>/7, 546.9 (59) [M-6Na]<sup>+</sup>/6.

Neoglycoconjugate 1c: Compound 30c (2.6 mg, 0.48 μmol) was oxidised and de-benzylated as described for compound 30a to give 1c (2.2 mg, quant.). The <sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O) spectrum of compound 1c was superimposable on that of 1a except at δ=3.82–3.60 (m, 138H; 2×H-4<sup>a</sup>, 2×H-4<sup>c</sup>, CH<sub>2</sub> PEG, 2×H-3<sup>a</sup>, 2×H-3<sup>c</sup>, 2×H-a) ppm.

Neoglycoconjugate 2a: Compound 31a (3.0 mg, 0.51 μmol) was oxidised and de-benzylated as described for compound 30a to give 2a (1.7 mg, 74%). The <sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O) spectrum of compound 2a was superimposable on that of 2c except at δ=3.82–3.60 (m, 34H; 2×H-4<sup>a</sup>, 2×H-4<sup>c</sup>, 2×H-4<sup>b</sup>, CH<sub>2</sub> PEG, 2×H-3<sup>a</sup>, 2×H-3<sup>b</sup>, 2×H-a) ppm.

Glycoconjugate 2b: Compound 34b (1.6 mg, 0.27 μmol) was de-benzylated and purified as described for 33a to give 2b (0.9 mg, 71%). The <sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O) spectrum of compound 2b was superimposable on that of 2c except at δ=3.82–3.60 (m, 54H; 2×H-4<sup>a</sup>, 2×H-4<sup>c</sup>, 2×

H-4<sup>b</sup>, CH<sub>2</sub> PEG, 2×H-3<sup>a</sup>, 2×H-3<sup>c</sup>, 2×H-3<sup>d</sup>, 2×H-a) ppm; ESI MS: C<sub>28</sub>H<sub>32</sub>N<sub>2</sub>O<sub>20</sub>S<sub>2</sub>Na<sub>2</sub>; m/z (%): 450.3 (66) [M-11Na+4H]<sup>7+</sup>/10, 501.3 (100) [M-10Na+4H]<sup>8+</sup>/9, 569.9 (83) [M-9Na+4H]<sup>9+</sup>/8.

Neoglycoconjugate 2c: Compound 31c (4.2 mg, 0.60 μmol) was oxidised and debenzylated as described for compound 30a to give 2c (2.6 mg, 76%); <sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O): δ=5.43 (d, J<sub>1,2</sub>=3.5 Hz, 2H; 2×H-1<sup>0b</sup>), 5.41 (d, J<sub>1,2</sub>=3.5 Hz, 2H; 2×H-1<sup>0c</sup>), 5.21 (brd, J<sub>1,2</sub>=3.0 Hz, 2H; 2×H-1<sup>0a</sup>), 5.19 (brd, J<sub>1,2</sub>=3.0 Hz, 2H; 2×H-1<sup>0d</sup>), 5.17 (brs, 2H; 2×H-1<sup>0</sup>), 5.15 (d, J<sub>1,2</sub>=3.5 Hz, 2H; 2×H-1<sup>a</sup>), 4.81 (under HOD peak, 2×H-5<sup>b</sup>), 4.80 (under HOD peak, 2×H-5<sup>c</sup>), 4.75 (brs, 2H; 2×H-5<sup>a</sup>), 4.41 (brd, J<sub>4,5</sub>=11.5 Hz, 2H; 2×H-6), 4.38–4.28 (m, 12H; 2×H-2<sup>a</sup>, 2×H-2<sup>b</sup>, 2×H-2<sup>c</sup>, 6×H-6), 4.26 (brd, J<sub>4,5</sub>=11.5 Hz, 2H; 4×H-6<sup>a</sup> or H-6<sup>c</sup> or H-6<sup>b</sup>), 4.19 (dd, J<sub>1,2</sub>=6.0, J<sub>3,4</sub>=4.0 Hz, 2H; 2×H-3<sup>b</sup>), 4.18 (dd, J<sub>1,2</sub>=6.0, J<sub>3,4</sub>=4.0 Hz, 2H; 2×H-3<sup>a</sup>), 4.13–4.08 (m, 6H; 2×H-3<sup>c</sup>, 2×H-4<sup>a</sup>, 2×H-4<sup>b</sup>), 4.08–3.95 (m, 12H; 2×OCH<sub>2</sub>CH<sub>2</sub>SO<sub>2</sub> (β=4.0), t, J=5.5 Hz), 2×H-5<sup>a</sup>, 2×H-5<sup>c</sup>, 2×H-5<sup>b</sup>, 2×H-4<sup>c</sup>), 3.87 (dt, J<sub>pm</sub>=10.0, J<sub>4,5</sub>=5.5 Hz, 2H; 2×H-a), 3.83–3.60 (m, 14H; 2×H-4<sup>a</sup>, 2×H-4<sup>c</sup>, 2×H-4<sup>b</sup>, CH<sub>2</sub> PEG, 2×H-3<sup>a</sup>, 2×H-3<sup>c</sup>, 2×H-3<sup>b</sup>, 2×H-a), 3.54 (t, J=5.5 Hz, 4H; 2×OCH<sub>2</sub>CH<sub>2</sub>SO<sub>2</sub>), 3.43 (brt, J=7.5 Hz, 4H; 4×H-c), 3.28 (dd, J<sub>2,3</sub>=10.0, J<sub>2,1</sub>=3.5 Hz, 2H; 2×H-2<sup>a</sup>), 3.26 (dd, J<sub>2,3</sub>=10.0, J<sub>2,1</sub>=3.5 Hz, 2H; 2×H-2<sup>b</sup>), 3.25 (dd, J<sub>2,3</sub>=10.0, J<sub>2,1</sub>=3.5 Hz, 2H; 2×H-2<sup>c</sup>), 2.21–2.11 (m, 4H; 4×H-b) ppm; <sup>13</sup>C NMR from HMQC (100.6 MHz, D<sub>2</sub>O): δ=99.7 (C-1<sup>a</sup>, C-1<sup>b</sup>), 99.4 (C-1<sup>c</sup>), 97.3 (C-1<sup>d</sup>), 96.9 (C-1<sup>e</sup>), 96.6 (C-1<sup>0a</sup>), 76.3 (C-2<sup>a</sup>), 76.2 (C-2<sup>b</sup>), C-4<sup>a</sup>, C-4<sup>c</sup>, C-4<sup>b</sup>, 74.3 (C-2<sup>0a</sup>), 72.0 (C-3<sup>a</sup> or C-3<sup>c</sup> or C-3<sup>b</sup>), 69.9 (CH<sub>2</sub> PEG, C-3<sup>a</sup> or C-3<sup>c</sup> or C-3<sup>b</sup>), 69.9 (C-5<sup>0a</sup>), 69.7 (C-5<sup>0b</sup>), 69.6 (C-3<sup>a</sup>, C-3<sup>b</sup>), 69.569.3 (C-5<sup>a</sup>, C-5<sup>c</sup>, C-5<sup>b</sup>, C-4<sup>a</sup>), 69.1 (C-3<sup>c</sup>), 67.0 (C-5<sup>c</sup>), 67.166.6 (C-6<sup>a</sup>, C-6<sup>c</sup>, C-6<sup>b</sup>), 66.4 (C-a), 63.8 (OCH<sub>2</sub>CH<sub>2</sub>SO<sub>2</sub>), 58.2 (C-2<sup>a</sup>, C-2<sup>c</sup>, C-2<sup>b</sup>), 52.4 (OCH<sub>2</sub>CH<sub>2</sub>SO<sub>2</sub>), 51.3 (C-c), 21.5 (C-b) ppm.

Neoglycoconjugate 3a: Compound 32a (1.5 mg, 0.19 μmol) was oxidised and debenzylated as described for compound 30a to give 3a (0.6 mg, 90%); The <sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O) spectrum of compound 3a was superimposable on that of 3b except at δ=3.82–3.60 (m, 38H; 2×H-4<sup>a</sup>, 2×H-4<sup>c</sup>, 2×H-4<sup>b</sup>, 2×H-4<sup>d</sup>, CH<sub>2</sub> PEG, 2×H-3<sup>a</sup>, 2×H-3<sup>c</sup>, 2×H-3<sup>b</sup>, 2×H-a).

Neoglycoconjugate 3b: Compound 32b (9.3 mg, 1.20 μmol) was oxidised and debenzylated as described for compound 30a to give 3b (4.5 mg, 62%); <sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O): δ=5.46–5.38 (m, 6H; 2×H-1<sup>0</sup>, 2×H-1<sup>c</sup>, 2×H-1<sup>b</sup>), 5.25–5.18 (m, 6H; 2×H-1<sup>a</sup>, 2×H-1<sup>d</sup>, 2×H-1<sup>e</sup>), 5.16 (brs, 2H; 2×H-1<sup>0</sup>), 5.14 (d, J<sub>1,2</sub>=3.5 Hz, 2H; 2×H-1<sup>a</sup>), 4.81 (under HOD peak, 2×H-5<sup>b</sup>), 4.80 (under HOD peak, 2×H-5<sup>c</sup>), 4.76 (d, J<sub>4,5</sub>=2.5 Hz, 2H; 2×H-5<sup>0a</sup>), 4.41 (brd, J<sub>4,5</sub>=11.0 Hz, 4H; 4×H-6), 4.37–4.28 (m, 14H; 2×H-2<sup>a</sup>, 2×H-2<sup>b</sup>, 2×H-2<sup>c</sup>, 2×H-2<sup>d</sup>, 6×H-6), 4.26 (brd, J<sub>4,5</sub>=11.0 Hz, 6H; 6×H-6), 4.22–4.15 (m, 6H; H-3<sup>a</sup>, H-3<sup>b</sup>, H-3<sup>c</sup>), 4.13–4.08 (m, 8H; 2×H-3H; 2×H-4<sup>a</sup>, 2×H-4<sup>c</sup>, 2×H-4<sup>b</sup>), 4.08–3.96 (m, 14H; 4×OCH<sub>2</sub>CH<sub>2</sub>SO<sub>2</sub> (β=4.0), t, J=5.5 Hz), 2×H-5<sup>a</sup>, 2×H-5<sup>c</sup>, 2×H-5<sup>b</sup>, 2×H-4<sup>c</sup>), 3.87 (dt, J<sub>pm</sub>=10.0, J<sub>4,5</sub>=5.0 Hz, 2H; 2×H-a), 3.83–3.60 (m, 58H; 2×H-4<sup>a</sup>, 2×H-4<sup>c</sup>, 2×H-4<sup>b</sup>, CH<sub>2</sub> PEG, 2×H-3<sup>a</sup>, 2×H-3<sup>c</sup>, 2×H-3<sup>b</sup>, 2×H-3<sup>d</sup>, 2×H-a), 3.54 (t, J=5.5 Hz, 4H; 4×OCH<sub>2</sub>CH<sub>2</sub>SO<sub>2</sub>), 3.42 (dd, J=9.0, 7.5 Hz, 4H; 4×H-c), 3.32–3.23 (m, 8H; 2×H-2<sup>a</sup>, 2×H-2<sup>c</sup>, 2×H-2<sup>b</sup>), 2.21–2.11 (m, 4H; 4×H-b) ppm; <sup>13</sup>C NMR from HMQC (100.6 MHz, D<sub>2</sub>O): δ=99.6 (C-1<sup>a</sup>, C-1<sup>b</sup>, C-1<sup>c</sup>), 99.4 (C-1<sup>d</sup>), 97.2 (C-1<sup>e</sup>), 96.7 (C-1<sup>c</sup>, C-1<sup>a</sup>, C-1<sup>b</sup>), 76.5 (C-2<sup>a</sup>), 76.8/76.2 (C-4<sup>a</sup>, C-4<sup>c</sup>, C-4<sup>b</sup>, C-4<sup>d</sup>), 76.2 (C-2<sup>b</sup>), C-4<sup>a</sup>, C-4<sup>c</sup>, C-4<sup>b</sup>), 74.1 (C-2<sup>0a</sup>), C-2<sup>0b</sup>), 71.9 (C-3<sup>a</sup> or C-3<sup>c</sup> or C-3<sup>b</sup>), 69.8 (CH<sub>2</sub> PEG, C-3<sup>a</sup> or C-3<sup>c</sup> or C-3<sup>b</sup>), 69.8 (C-5<sup>0a</sup>), 69.6 (C-5<sup>0b</sup>), C-5<sup>0c</sup>), 69.5 (C-3<sup>a</sup>, C-3<sup>c</sup>, C-3<sup>b</sup>), 69.569.3 (C-5<sup>a</sup>, C-5<sup>c</sup>, C-5<sup>b</sup>, C-4<sup>a</sup>), 69.1 (C-3<sup>b</sup>), 69.1 (C-5<sup>b</sup>), 67.166.6 (C-6<sup>a</sup>, C-6<sup>c</sup>, C-6<sup>b</sup>, C-6<sup>d</sup>), 66.4 (C-a), 63.4 (OCH<sub>2</sub>CH<sub>2</sub>SO<sub>2</sub>), 58.1 (C-2<sup>a</sup>, C-2<sup>c</sup>, C-2<sup>b</sup>, C-2<sup>d</sup>), 52.4 (OCH<sub>2</sub>CH<sub>2</sub>SO<sub>2</sub>), 51.3 (C-c), 21.6 (C-b) ppm; ESI MS: C<sub>28</sub>H<sub>32</sub>N<sub>2</sub>O<sub>20</sub>S<sub>2</sub>Na<sub>2</sub>; m/z (%): 480.2 (79) [M-14Na+2H]<sup>7+</sup>/12, 487.7 (100) [M-12Na]<sup>7+</sup>/12, 524.1 (84) [M-14Na+3H]<sup>7+</sup>/11, 576.3 (86) [M-14Na+4H]<sup>7+</sup>/10, 580.9 (63) [M-12Na+2H]<sup>7+</sup>/10.

Neoglycoconjugate 3c: Compound 32c (3.8 mg, 0.44 μmol) was oxidised and debenzylated as described for compound 30a to give 3c (2.4 mg, 77%); The <sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O) spectrum of compound 3c was superimposable on that of 3b except at δ=3.82–3.60 (m, 146H; 2×H-4<sup>a</sup>, 2×H-4<sup>c</sup>, 2×H-4<sup>b</sup>, 2×H-4<sup>d</sup>, CH<sub>2</sub> PEG, 2×H-3<sup>a</sup>, 2×H-3<sup>c</sup>, 2×H-3<sup>b</sup>, 2×H-3<sup>d</sup>, 2×H-a) ppm.

## Acknowledgments

We thank the CNRS (Physique Chimie du Vivant programme), the Ministère de l'Éducation Nationale et de la Recherche (OG), the Paris Sud University and the Région Rhone Alpes.

- [1] I. Capila, R. J. Linhardt, *Angew. Chem.* 2002, 114, 426–450; *Angew. Chem. Int. Ed.* 2002, 41, 390–412.
- [2] H. Lortat-Jacob, A. Grosdidier, A. Imberty, *Proc. Natl. Acad. Sci. USA* 2002, 99, 1229–1234.
- [3] J. D. Esko, U. Lindhal, *J. Clin. Invest.* 2001, 108, 169–173.
- [4] U. Lindhal, M. Kusche-Gullberg, L. I. Kjellen, *J. Biol. Chem.* 1998, 273, 24979–24982.
- [5] J. T. Gallagher, *Biochem. Soc. Trans.* 1997, 25, 1205–1209.
- [6] M. A. B. A. Dennissen, G. I. Jenniskens, M. Pieffers, E. M. M. Versteeg, M. Petitou, J. H. Veerkamp, T. van Kuppevelt, *J. Biol. Chem.* 2002, 277, 10982–10986.
- [7] K. Sugahara, H. Kitagawa, *Curr. Opin. Struct. Biol.* 2000, 10, 518–527.
- [8] U. Lindahl, G. Bäckström, L. Thunberg, I. G. Leder, *Proc. Natl. Acad. Sci. USA* 1980, 77, 6551–6555.
- [9] B. Casu, P. Oreste, G. Torri, G. Zoppetti, J. Choay, J. C. Lormeau, M. Petitou, P. Sinaj, *Biochem. J.* 1981, 197, 599–609.
- [10] C. A. A. van Boeckel, M. Petitou, *Angew. Chem.* 1993, 105, 1741–1761; *Angew. Chem. Int. Ed. Engl.* 1993, 32, 1671–1818.
- [11] J. Kowenzky, J. M. Mallet, J. Esnault, P. A. Driguez, P. Siran, I. P. Héroult, J. M. Herbert, M. Petitou, P. Sinaj, *Eur. J. Org. Chem.* 2002, 3395–3603.
- [12] J. L. de Paz, J. Angulo, J. M. Lassaletta, P. M. Nieto, M. Redondo-Horcoja, R. M. Lozano, G. Giménez-Gallego, M. Martín-Lomas, *ChemBioChem* 2001, 2, 673–685.
- [13] R. Ojeda, J. Angulo, P. M. Nieto, M. Martín-Lomas, *Can. J. Chem.* 2002, 80, 917–936.
- [14] J. D. C. Codée, G. A. van der Mare, C. A. A. van Boeckel, J. H. van Boom, *Eur. J. Org. Chem.* 2002, 3954–3965.
- [15] L. Poletti, M. Fleischer, C. Vogel, M. Guerrini, G. Torri, L. Lay, *Eur. J. Org. Chem.* 2001, 2727–2734.
- [16] H. A. Orgueira, A. Bartolozzi, P. Schell, R. E. J. N. Litjens, E. R. Palmacci, P. H. Seeberger, *Chem. Eur. J.* 2003, 9, 143–169.
- [17] L. Poletti, L. Lay, *Eur. J. Org. Chem.* 2003, 2999–3024.
- [18] N. A. Karst, R. J. Linhardt, *Curr. Med. Chem.* 2003, 10, 1993–2031.
- [19] J. L. de Paz, R. Ojeda, N. Reichardt, M. Martín-Lomas, *Eur. J. Org. Chem.* 2003, 3308–3324.
- [20] A. Prabhu, A. Venot, G. J. Booms, *Org. Lett.* 2003, 5, 4975–4978.
- [21] A. Lubineau, D. Bonnaffé, *Eur. J. Chem.* 1999, 2523–2532.
- [22] O. Gavard, Y. Hessant, I. Alais, V. Dauverger, A. Dilhas, A. Bascou, D. Bonnaffé, *Eur. J. Org. Chem.* 2003, 3603–3620.
- [23] A. Dilhas, D. Bonnaffé, *Tetrahedron Lett.* 2004, 45, 3643–3645.
- [24] R. Lucas, D. Hamza, A. Lubineau, D. Bonnaffé, *Eur. J. Org. Chem.* 2004, in press.
- [25] J. E. Turnbull, J. T. Gallagher, *Biochem. J.* 1991, 271, 553–559.
- [26] R. Sadir, F. Baleux, A. Grosdidier, A. Imberty, H. Lortat-Jacob, *J. Biol. Chem.* 2001, 276, 8288–8296.
- [27] H. Lortat-Jacob, J. E. Turnbull, J. A. Grimaud, *Biochem. J.* 1995, 310, 497–505.
- [28] S. E. Stringer, J. T. Gallagher, *J. Biol. Chem.* 1997, 272, 20508–20514.
- [29] R. R. Vivès, R. Sadir, A. Imberty, A. Rencurosi, H. Lortat-Jacob, *Biochemistry* 2002, 41, 14779–14789.
- [30] S. E. Stringer, M. J. Forster, B. Mulloy, C. R. Bishop, G. J. Graham, J. T. Gallagher, *Blood* 2002, 100, 1546–1550.
- [31] J. T. Gallagher, *J. Clin. Invest.* 2001, 108, 357–361.
- [32] M. Petitou, P. Duchaussoy, P. A. Driguez, G. Jurand, J. P. Héroult, J. C. Lormeau, C. A. A. van Boeckel, J. M. Herbert, *Angew. Chem.* 1998, 110, 3186–3191; *Angew. Chem. Int. Ed. Engl.* 1998, 37, 3009–3014.
- [33] M. Petitou, J. P. Héroult, A. Bernat, P. A. Driguez, P. Duchaussoy, J. C. Lormeau, J. M. Herbert, *Nature* 1999, 398, 417–422.

- [34] M. Petitou, A. Imberty, P. Duchaussoy, P.-A. Driguez, M.-L. Coccato, F. Gouvenec, P. Sizun, J.-P. Héroult, S. Perez, J.-M. Herbert, *Chem. Eur. J.* 2001, 7, 858–873.
- [35] P. D. J. Grootenhuis, P. Westerduin, D. Meuleman, M. Petitou, C. A. A. van Boeckel, *Nat. Struct. Biol.* 1995, 2, 736–739.
- [36] P. Westerduin, J. E. M. Basten, M. A. Broekhoven, V. de Kimppe, W. H. A. Kuipers, C. A. A. van Boeckel, *Angew. Chem.* 1996, 108, 339–342; *Angew. Chem. Int. Ed. Engl.* 1996, 35, 331–333.
- [37] A. Lubineau, S. Escher, J. Alais, D. Bonnaffé, *Tetrahedron Lett.* 1997, 38, 4087–4090.
- [38] H. Lortat-Jacob, H. K. Kleinman, J. A. Grimaud, *J. Clin. Invest.* 1991, 87, 878–883.
- [39] H. Lortat-Jacob, F. Baltzer, J. A. Grimaud, *J. Biol. Chem.* 1996, 271, 16139–16143.
- [40] H. Lortat-Jacob, C. Brisson, S. Guerret, G. Morel, *Cytokine* 1996, 8, 557–566.
- [41] B. Skurkovich, S. Surkovich, *Curr. Opin. Mol. Therapeutics* 2003, 5, 52–57.
- [42] P. Parronchi, P. Romagnani, F. Annunziato, S. Sampognaro, A. Becchio, L. Giannarini, E. Maggi, C. Pupilli, F. Tonelli, S. Romagnani, *Am. J. Pathol.* 1997, 150, 823–832.
- [43] H. Lortat-Jacob, J. A. Grimaud, *FEBS Lett.* 1991, 280, 152–154.
- [44] This length was calculated for a conformation of the NA domain close to the helical one of heparin; B. Mulloy, M. J. Forster, C. Jones, D. B. Davies, *Biochem. J.* 1993, 293, 849–858.
- [45] Competition analysis between synthetic peptides encompassing the KRKR domain or the RGRR domain and full-length IPN- $\gamma$  suggests that the KRKR sequence is the main contributor in the binding; H. Lortat-Jacob, J. A. Grimaud, *FEBS Lett.* 1991, 280, 152–154.
- [46] PDB entry 1HIG: S. E. Ealick, W. J. Cook, S. Vijay-Kumar, M. Carson, T. L. Nagabhushan, P. P. Trotta, C. E. Bugg, *Science* 1991, 252, 698–702.
- [47] PDB entry 1F9G: D. J. Thiel, M. L. le Du, R. L. Walter, A. D'Arcy, C. Chene, M. Fountoulakis, G. Garotta, F. K. Winkler, S. E. Ealick, *Struct. Fold. Des.* 2000, 8, 927–936.
- [48] C. Tabour, J. M. Mallet, F. Bono, J. M. Herbert, M. Petitou, P. Sinaÿ, *Bioorg. Med. Chem.* 1999, 7, 2003–2012.
- [49] T. Ogawa, S. Nakabayashi, T. Kitajima, *Carbohydr. Res.* 1983, 114, 225–236.
- [50] C. Lamberth, M. D. Bednarski, *Tetrahedron Lett.* 1991, 32, 7369–7372.
- [51] J. I. Oltvoort, C. A. A. van Boeckel, J. H. Koning, J. H. van Boom, *Synthesis* 1981, 305–308.
- [52] The  $^1\text{H}$  NMR  $J_{1,2}$  coupling constant (3.5 Hz) as well as the  $^{13}\text{C}$  chemical shifts ( $\delta=97.4$  ppm) for 7 at the anomeric centre of the newly formed glycosidic linkage are consistent with the proposed  $\alpha$  stereochemistry.
- [53] J. Kovensky, P. Duchaussoy, F. Bono, M. Salmivirta, P. Sizun, J.-M. Herbert, M. Petitou, P. Sinaÿ, *Bioorg. Med. Chem.* 1999, 7, 1567–1580.
- [54] P. Duchaussoy, G. Jaumand, P.-A. Driguez, I. Lederman, F. Gourvenec, J.-M. Strasel, P. Sizun, M. Petitou, J.-M. Herbert, *Carbohydr. Res.* 1999, 317, 63–84.
- [55] As for the tetrasaccharide 7, the  $^1\text{H}$  NMR  $J_{1,2}$  coupling constant (3.5 Hz for 10 and 11) as well as the  $^{13}\text{C}$  chemical shifts ( $\delta=97.4$  ppm for 10 and  $\delta=97.4$  or  $97.8$  ppm for 11) at the anomeric centres of the newly formed glycosidic linkages are consistent with the proposed  $\alpha$  stereochemistry.
- [56] W. S. Mungall, G. L. Greene, G. A. Heavner, R. L. Letsinger, *J. Org. Chem.* 1975, 40, 1659–1662.
- [57] E. J. Corey, K. C. Nicolaou, R. D. Balanson, Y. Machida, *Synthesis* 1975, 590–591.
- [58] L. Benati, P. C. Montecocchi, D. Nanni, P. Spagnolo, M. Volta, *Tetrahedron Lett.* 1995, 36, 7313–7314; C. Goulaouic-Dubois, M. Hesse, *Tetrahedron Lett.* 1995, 36, 7427–7430.
- [59] H. Bayley, D. N. Standing, J. R. Knowles, *Tetrahedron Lett.* 1978, 19, 3633–3634.
- [60] M. Nitz, B. W. Purse, D. R. Bundle, *Org. Lett.* 2000, 2, 1939–1942.
- [61] The glycoconjugate purity was measured by HPLC with UV detection at 220 nm and by assuming that compounds and impurities had the same absorbance at this wavelength.
- [62] B. M. Trost, D. P. Curran, *Tetrahedron Lett.* 1981, 22, 1287–1290.
- [63] It is important to note that the unprotected glycoconjugates 1–3 bind to the  $\text{Pd}(\text{OH})_2/\text{C}$  catalyst in the hydrogenolysis step; it is thus of the utmost importance to wash the  $\text{Pd}(\text{OH})_2/\text{C}$  with an acetone/trile/water (50:50) mixture to recover the debenzylated compounds in good yields.
- [64] W. Chai, J. Luo, A. M. Lawson, *Anal. Chem.* 1998, 70, 2060–2066.
- [65] A. J. Rhombert, S. Ernst, R. Sasisekharan, K. Biemann, *Proc. Natl. Acad. Sci. USA*, 1998, 95, 4176–4181.
- [66] B. Kuberan, M. Lech, L. Zhang, Z. L. Wu, D. L. Beiler, R. D. Rosenberg, *J. Am. Chem. Soc.* 2002, 124, 8707–8718.
- [67] C. Thanawiroon, K. G. Rice, T. Toida, R. J. Linhardt, *J. Biol. Chem.* 2004, 279, 2608–2615.
- [68] T. Bitter, H. M. Muir, *Anal. Biochem.* 1962, 4, 330–334.
- [69] D. D. Perrin, W. L. F. Armarego, *Purification of Laboratory Chemicals*, 3rd ed., Pergamon Press, Oxford, 1988.
- [70] T. Feizi, M. S. Stoll, C. -T. Yuen, W. Chai, A. M. Lawson, *Methods Enzymol.* 1994, 230, 484–519.
- [71] R. R. Vives, S. Goodger, D. A. Pye, *Biochem. J.* 2001, 354, 141–147.
- [72] P. W. Riddles, R. L. Blackley, B. Zerner, *Anal. Biochem.* 1979, 94, 75–81.

Received: December 23, 2003

Revised: April 23, 2004

Published online: ■■■■, 2004



Review

# Endocan or endothelial cell specific molecule-1 (ESM-1): A potential novel endothelial cell marker and a new target for cancer therapy

S. Sarrazin <sup>a,d</sup>, E. Adam <sup>a,b</sup>, M. Lyon <sup>c</sup>, F. Depontieu <sup>b</sup>, V. Motte <sup>a</sup>, C. Landolfi <sup>a</sup>, H. Lortat-Jacob <sup>d</sup>, D. Bechard <sup>a</sup>, P. Lassalle <sup>b</sup>, M. Delehedde <sup>a,\*</sup>

<sup>a</sup> ENDOTIS PHARMA, Parc Eurasanté, 70 rue du Dr. Yersin, 59120 Loos, France

<sup>b</sup> INSERM U416, Institut Pasteur de Lille, 1 rue du Dr. A. Calmette, 59019 Lille, France

<sup>c</sup> Cancer Research UK Department of Medical Oncology, University of Manchester and Christie Hospital NHS Trust, Wilmslow Road, Manchester M20 4BX, UK

<sup>d</sup> Institut de Biologie Structurale, 41 rue Jules Horowitz, 38027 Grenoble, France

Received 13 May 2005; received in revised form 8 August 2005; accepted 10 August 2005

Available online 26 August 2005

## Abstract

Endocan, previously called endothelial cell specific molecule-1, is a soluble proteoglycan of 50 kDa, constituted of a mature polypeptide of 165 amino acids and a single dermatan sulphate chain covalently linked to the serine residue at position 137. This dermatan sulphate proteoglycan, which is expressed by the vascular endothelium, has been found freely circulating in the bloodstream of healthy subjects. Experimental evidence is accumulating that implicates endocan as a key player in the regulation of major processes such as cell adhesion, in inflammatory disorders and tumor progression. Inflammatory cytokines such as TNF- $\alpha$ , and pro-angiogenic growth factors such as VEGF, FGF-2 and HGF/SF, strongly increased the expression, synthesis or the secretion of endocan by human endothelial cells. Endocan is clearly overexpressed in human tumors, with elevated serum levels being observed in late-stage lung cancer patients, as measured by enzyme-linked immunoassay, and with its overexpression in experimental tumors being evident by immunohistochemistry. Recently, the mRNA levels of endocan have also been recognized as being one of the most significant molecular signatures of a bad prognosis in several types of cancer including lung cancer. Overexpression of this dermatan sulphate proteoglycan has also been shown to be directly involved in tumor progression as observed in mouse models of human tumor xenografts. Collectively, these results suggest that endocan could be a biomarker for both inflammatory disorders and tumor progression as well as a validated therapeutic target in cancer. On the basis of the recent successes of immunotherapeutic approaches in cancer, the preclinical data on endocan suggests that an antibody raised against the protein core of endocan could be a promising cancer therapy.

© 2005 Elsevier B.V. All rights reserved.

**Keywords:** Endocan; ESM-1; Proteoglycan; Dermatan; Cancer; Inflammation; Immunotherapy

## Contents

1. Introduction . . . . .	26
2. Genomic organization and transcriptional control of endocan . . . . .	26
3. Structure of endocan core protein. . . . .	28
4. Endocan is a dermatan sulfate proteoglycan (DSPG) . . . . .	29
5. Endocan is a circulating DSPG originally described as a product of endothelial cells . . . . .	32

**Abbreviations:** ESM-1, endothelial cell-specific molecule-1; PGs, proteoglycans; GAG, glycosaminoglycan; CS, chondroitin sulphate; DS, dermatan sulphate; HS, heparan sulphate; KS, keratan sulphate; HA, hyaluronic acid; TNF- $\alpha$ , tumor necrosis factor alpha; HGF/SF, hepatocyte growth factor/scatter factor; VEGF, vascular endothelial growth factor; FGF-2, fibroblast growth factor 2; PF-4, platelet factor 4; IL-1, interleukin 1; IFN- $\gamma$ , interferon gamma; IGF, insulin-like growth factor; IGFBP, insulin-like growth factor binding-protein; SLRP, small leucine-rich repeat proteoglycan; HUVEC, human umbilical vein endothelial cells; CKII, casein kinase II; EGF, epidermal growth factor; EGFR, epidermal growth factor receptor; GalNAc, N-acetyl galactosamine; GlcA, glucuronic acid; IdoA, iduronic acid

\* Corresponding author. Tel.: +33 328555054; fax: +33 328555128.

E-mail address: marysodelehedde@endotis.com (M. Delehedde).

0304-419X/\$ - see front matter © 2005 Elsevier B.V. All rights reserved.  
doi:10.1016/j.bbcan.2005.08.004

6. Endocan in tumor progression . . . . .	32
Acknowledgements . . . . .	34
References . . . . .	34

## 1. Introduction

Proteoglycans (PGs) are complex macromolecules present primarily at the cell surface, in the extracellular matrix which surrounds most mammalian cell types and also in body fluids. By virtue of the multiplicity of their protein binding partners (e.g., growth factors, chemokines, enzymes and other extracellular matrix proteins), they have been shown to play major roles in regulating normal cellular processes, such as proliferation, remodelling, migration or angiogenesis, that can become dysregulated in disease states and during cancer progression [1–6]. Structurally, PGs consist of a core protein, of variable size and structure, and one or more associated glycosaminoglycan (GAG) chains. These GAGs can be of 5 types: chondroitin sulphate (CS), dermatan sulphate (DS), heparan sulphate (HS), heparin or keratan sulphate (KS). Growing evidence suggests that DS, like the better studied heparin and HS, is an important co-factor in a variety of cell behaviours. In this article, we will review the rapidly expanding data on endocan, also called endothelial cell-specific molecule 1 (ESM-1), a dermatan sulphate PG, which is relatively unusual in that it freely circulates in the blood and carries only one GAG chain [7–9]. Endocan, being a circulating PG, would be present at sites of inflammation, as well as at tumor sites. Soluble PGs or free GAGs themselves have been shown to be critical regulators of chemokines and growth factors that influence inflammatory events and tumor behaviour. In myeloma patients, soluble shed syndecan-1 HSPG has been shown to be associated via its HS chains with hepatocyte growth factor/scatter factor (HGF/SF) [10]. Furthermore, DS and DSPGs were found to be abundant in human wound fluids and to regulate the activities of local growth factors [11]. Interestingly, a significantly increased level of circulating endocan has been observed not only in the serum from septic patients but also in cancer patients [12,13]. Structurally and functionally, endocan has recently appeared as a potentially key player in cancer and inflammation diseases. For instance, adhesion of leucocytes, or of circulating cancerous cells, to the endothelial vasculature is essential during the process of invasion. Direct binding of circulating endocan to integrins has been shown to result in reduction of leucocyte interactions [8]. Moreover, there has been evidence for angiogenic growth factor effects on endocan gene regulation, and also for direct interactions between endocan and those growth factors leading to regulation of their activities. Furthermore, the expression of angiogenic markers has been shown to be concomitant with that of endocan at the mRNA levels as revealed by microarray analysis [14]. Tumor angiogenesis, or the formation of new blood vessels, is a prerequisite for supplying oxygen and nutrients to a growing tumor. We will thus need soon to consider the interplay of

angiogenic factors such as HGF/SF and vascular endothelial growth factor (VEGF) with endocan.

In this review, we will first describe the general structure of the endocan gene as well as of the translated core protein. The nature of its DSPG form will be then detailed and analyzed in order to better understand the presently-known binding interactions and biological activities of endocan. We will then focus on the involvement of endocan in tumor progression, with its major role described so far of being a co-receptor for HGF/SF, a well-known angiogenic growth factor [9]. Moreover, endocan has been recently described as a marker of tumor aggressiveness in several clinical settings. Illustrations will be presented of the recent successes of immunotherapy approaches to cancer treatment. As highlighted by results obtained from preclinical data, endocan should be a validated therapeutic target in cancer and an appropriate antibody raised against endocan could be promising in cancer therapy.

## 2. Genomic organization and transcriptional control of endocan

Endocan, previously called endothelial cell specific molecule-1 (ESM-1), was originally cloned from a human endothelial cell cDNA library in 1996 by Lassalle and collaborators [7]. This newly described molecule is the product of a single gene, localized on human chromosome 5 at the position 5q11.2, that is organized into 3 exons separated by 2 introns [7]. Exon 1 is encoded by 362 bp, exon 2 by 150 bp and exon 3 by 1560 bp. Exon 1 and part of exon 2 encode for an N-terminal cysteine-rich region of 110 amino acids (Fig. 1). Exon 2 also encodes for a particular domain of interest, the functionally-defined F-rich region (<sup>113</sup>FPFFQY<sup>118</sup>), which has been shown to be involved in the function of endocan as we will discuss later [13]. Finally, the large exon 3 encodes for a short 33 amino-acid long C-terminal region (due to an early stop codon within the exon) which includes the unique O-glycosylation site at serine 137 [9]. Analysis of the endocan cDNA shows a large 3'-UTR containing a series of AU-rich elements (AREs), mainly identified as *cis*-elements that affect mRNA stability and which can be regulated by p38 mitogen activated protein kinase [15,16]. This 3'-UTR also contains five putative polyadenylation sites, but only the last one appears functional [7]. Furthermore, endocan has an open reading frame of 552 nucleotides [7]. The promoter of human endocan has been cloned and sequenced by Tsai et al. [17]. Structurally, this promoter is composed of 3888 bp at the 5' flanking region of the human endocan gene and contains a classical TATA box [17]. Interestingly, certain HSPGs such as perlecan, and CS/DSPGs such as biglycan, have been shown to lack canonical TATA boxes [18]. The promoter of human endocan also contains a number of potential transcription factor binding sites

```

-19 MKSVLLLTLL LVPALVAAW  SNNYAVDCPQ  HCDSECKSS  FRCKRVLDD  CGCCRVCAAG

42 RGETYRTVS  GMDGMKGGP  LRQPSNGED  PFGEEFGI-K  DFPYDTFQMD  QREICNCGSQ

102 ICDRGTGKCL  KFFFFQYSVT  KSSNRFVSLT  EHDMAQGDGN  IVREEVYKEN  AAGSPVMRWK

162 LNPR

```

Fig. 1. Amino acid sequence of full-length endocan. The endocan amino acid sequence is numbered from residues –19 to 165, where the sequence from –19 to 0 corresponds to the hydrophobic N-terminal signal peptide. The cysteines involved in the EGF-like motif are shown in grey, the serine 137, where the dermatan sulphate chain is attached, is underlined, and the F-rich region is denoted by italics. The EGF-like consensus sequence is based on the following consensus profile: xxxxCx(0,48)Cx(3,12)Cx(1,70)Cx(1,6)CxxGxx(0,21)GxxxCx with C : conserved cysteine involved in a disulfide bond, G : often conserved glycine, a : often conserved aromatic amino acid, x : any residue.

including three Ets motifs (–77 to –74, –63 to –60 and +16 to +19), a CRE-like element (–68 to –62) and a 6-bp palindromic sequence (GCATGC) (–58 to –53) [17]. Six of the nucleotides in the CRE-like site (TGACGGGA) match the 8-bp cAMP-response-element (CRE-like) consensus site (TGACGTCA) [19]. The Ets motifs may act as functional *cis*-regulatory elements [17]. As revealed by mutational analysis, the Ets-binding motif located between –77 and –74, and the CRE-like motif located between –68 and –62, may mediate high level expression of endocan in endothelial cells. As shown by DNA footprinting of the complete promoter, the second Ets site (–63 to –60) and the palindromic 6-bp sequence (–58 to –53) can inhibit the expression of endocan and therefore function as transcriptional repressors [17]. Interestingly, Ets-1 factors have been shown to regulate the expression of several angiogenic and extracellular matrix remodelling factors that promote invasive phenotype [20]. For example, binding sites for the Ets domain family of transcription factors are found in the promoters of the matrix metalloproteinase family of matrix degrading enzymes. Moreover, functional Ets binding sites are required for the maximal activity of promoters of molecules involved in angiogenesis such as Flt-1, Tie-1, Tie-2 and vascular endothelial cadherin [21,22]. Among the other sites identified within the endocan promoter were 3 GATA motifs, 3 AP4 motifs as well as 2 AP1 motifs [17]. Originally, only one endocan mRNA product was observed by northern blot analysis of mRNA from human umbilical vein endothelial cells (HUVEC) [7]. However, RT-PCR analysis revealed an alternatively spliced form of endocan mRNA, with an internal deletion of 150 bp, from human capillary endothelial cells [23]. In addition, two distinct translated products of endocan were clearly detected in lysates from HUVEC, which clearly derive from alternative splicing of the endocan gene [12] (Deponitieu and Lassalle, unpublished data). Alternative splicing, which is a major mode of genetic regulation in higher eukaryotes, occurs quite frequently among PGs including the hyalactans (versican, aggrecan and brevican) and the Small Leucine-rich Repeat Proteoglycans or SLRPs (such as decorin and biglycan) [3,18,24]. Alternative splicing can be a major mode of regulation of biological activities of the protein [25], notably by the modulation of their glycanation. For example, the various CD44 isoforms, which derive from the differential utilization of 10 variable exons (CD44v1–v10), exhibit different post-transcriptional modifications such as N-

glycosylation, O-glycosylation and glycanation [3,26]. Many isoforms of CD44 are glycanated by CS, however, the v3 exon directs glycanation towards HS [3,26].

Several lines of evidence have demonstrated that *in vitro*, the levels of endocan mRNA are highly regulated by the environment. Inflammatory cytokines, such as tumor necrosis factor alpha (TNF- $\alpha$ ) and interleukin 1 (IL-1), have been shown to strongly upregulate endocan mRNA in HUVEC, and in a time-dependent manner [7]. An accumulation of endocan mRNA was observed after only 1 h of exposure to TNF- $\alpha$ , which peaked at 18 h [7]. Upregulation of endocan mRNA levels has also been observed after stimulation of adipocytes with either PMA or retinoic acid [27]. In contrast, IFN- $\gamma$ , another inflammatory mediator, has no marked effect on endocan expression at the mRNA level after 24 h of treatment [7]. However, IFN- $\gamma$  was shown to totally prevent the increased endocan mRNA levels induced by TNF- $\alpha$  treatment of endothelial cells [7]. These results favour a high degree of control of endocan expression at the mRNA level during the inflammatory process. Other molecules involved in growth and angiogenesis have been shown to also modify endocan gene expression. A recent study using high density oligonucleotide arrays and the screening of changes in the gene expression profile showed that a combination of HGF/SF and VEGF induced an upregulation of endocan in HUVEC after 24 h [14]. This upregulation by two potent endothelial mitogens and all-round angiogenic factors was one of the most significant seen (2.1-fold increase) among the upregulated genes [14]. Interestingly, VEGF itself, as well as Met, the specific signalling receptor for HGF/SF, were concomitantly upregulated [14]. These results clearly linked endocan expression with the angiogenic process during tumor progression. In the presence of VEGF and FGF-2, endocan mRNA levels were also shown to be upregulated 4-fold in an *in vitro* model of angiogenesis, whereby human capillary endothelial cells form tubes in three-dimensional collagen gels [23]. These data are thus in favour of a regulation of endocan mRNA expression occurring during cancer progression under stimulation by angiogenic factors as originally observed with inflammatory mediators [7]. A recent cDNA microarray study carried out to compare alterations in gene expression in highly aggressive versus poorly aggressive melanoma cells, also showed a 44-fold increase in endocan mRNA levels in the former [28]. Endocan appeared to be one of the genes that was most significantly increased along with other key players in

angiogenesis such as tie-1 (100-fold increase) and in lymphogenesis such as VEGF-C (30-fold increase) [28,29]. The widespread expression of the endogenous endocan gene in multiple subtypes of cultured endothelial cells in vitro [7–9,12,17,23] contrasts with the more limited distribution so far observed in vivo [7,17]. Up to now, expression of endocan was only detected by northern blot in normal tissues in kidney and in lung, but, unexpectedly, not in the heart, placenta, pancreas or brain, which are heavily vascularized organs [7,23]. Taken together, these data strongly suggest that endocan itself may be highly regulated by the normal or pathological environment, similarly to what has been observed with inflammatory mediators and angiogenic growth factors [7,12,14,23].

### 3. Structure of endocan core protein

Structurally, endocan is constituted of a mature polypeptide of 165 amino acids in length [7]. At the NH<sub>2</sub>-terminus, endocan possesses a typical peptide signal sequence of 19 amino acids in length and rich in hydrophobic residues [7]. No clear structural evidence leads to placement of endocan in any existing family of PGs, especially those of the secreted CS/DSPG families: no lectin site is present as observed in hyalactans, and there is no central domain containing leucine-rich repeats, a characteristic of the SLRPs [18]. However, endocan is cysteine-rich with 10.9% of the protein comprising cysteine residues [7]. Interestingly, all the 18 cysteines are concentrated in the N-terminal region of the protein (Fig. 1). Homologies can however be found with other families of proteins, more specifically within the N-terminal domain. Endocan shows a 15 to 28% identity with members of the Insulin-like Growth Factor Binding-Protein (IGFBP) superfamily, which are characterized by their cysteine-rich content and their affinity for IGF [7,30]. Endocan has thus been classified as IGFBP-rP6 within the more than 10 IGFBP-related proteins described until now [30]. However, no evidence for the binding of IGF by endocan has been demonstrated yet. In addition, members of the CCN family (including CTGF or IGFBP-rP2, Cyr-61 or IGFBP-rP4, and NOV protein or IGFBP-rP3), which contain an N-terminal domain structurally related to IGFBP and a cysteine-rich domain, also possess 13 to 20% identity with endocan [30,31]. CCN proteins are also mainly secreted proteins involved in a wide variety of biological activities such as proliferation and adhesion, as well as invasion and angiogenesis [31]. Finally, there is a 74% amino acid sequence homology with PG25, a protein originally identified in rat pineal gland [32], which subsequently has been classified as rat ESM-1 or rat endocan. Even though the N-terminal domain of endocan shares similarities with IGFBP and IGFBP-related proteins, no significant similarities with other proteins have yet been described for the C-terminus of endocan.

As a result of post-translational modifications, mature full-length endocan appears at 50 kDa on SDS-PAGE [8,9,12]. A single GAG chain of the DS type is covalently linked to the serine at position 137, leading to the generation of a DSPG, as will be described in more detail later (Fig. 1). By a computational search, 3 putative sites of O-glycosylation were

originally identified, corresponding to serine 16, threonine 120 and serine 137 (Fig. 1). No site for N-glycosylation was detected for endocan [7,9]. N-glycosylations have been described for other DSPGs, e.g., there are 3 N-linked oligosaccharides described on decorin, which may have a role in the conformation of the protein [33]. Of the three putative O-glycosylation sites on endocan, it is known that a GAG chain is covalently attached to serine 137 [9]. Furthermore, a mutation of threonine 120 to alanine gives a form which also migrates at 50 kDa, as wild type endocan, ruling out the possibility of any glycanation at this potential site [9]. No role for this other putative O-glycosylated site has yet been proposed. By working through databases, other sites of possible post-translational modifications have been detected (Fig. 2). Among them, one N-myristoylation site can be detected (125 to 130: GTGKCL) inside the exon 2-derived sequence (Fig. 2B). Putative phosphorylation sites can also be identified. Among the latter, 3 putative sites for protein kinase C recognition have been observed (Fig. 2B). Moreover, 5 sites for casein kinase II (CKII) can be detected in the core protein (Fig. 2B). No role or functionality of these latter motifs has been identified yet. The presence of all these putative sites should be further revealed by structural data obtained after the crystallisation of the core protein of endocan. By sequence analysis, it was also possible to detect a sequence corresponding to an epidermal growth factor (EGF)-like motif between amino acids 46 to 90 (Fig. 2B). An EGF-like domain is mainly identified by the characteristic disposition of six cysteine residues which are generally involved in disulphide bonding. Subdomains between the conserved cysteines will then vary in length between different proteins. The general sequence observed is: X(4)-C-X(0,48)-C-X(3,12)-C-X(1,70)-C-X(1,6)-C-X(2)-G-a-X(0,21)-G-X(2)-C-, with "C" corresponding to the conserved cysteine involved in a disulfide bond, with "G" corresponding to the often conserved glycine, with "a" corresponding to a conserved aromatic amino acid, with "X" calling for any residue (Fig. 1). These EGF-like domains have been described within various extracellular matrix PGs including perlecan and agrin, two of the major HSPGs of basement membranes, and among all the members of the hyalactan family [18]. Interestingly, EGF-like motifs are not present in SLRP members such as decorin [18]. In the case of perlecan, which contains 4 EGF-like repeats, they have been proposed to have a function in oligomerization, which occurs via domain V of perlecan, where these EGF-like motifs occur [18]. Hyalactans (including versican, aggrecan, neurocan and brevican) all possess EGF-like motifs adjacent to their characteristic lectin domains. Brevican (the shortest member of the hyalactan family with a core protein of 100 kDa and 1 to 3 CS chains) possesses only one EGF-like motif like endocan. The functional significance of EGF-like domains in what appear to be unrelated proteins is not yet fully understood. Finally, a 3D model of endocan based on protein homologues of known structure (CPH models-2.0 Server-Technical, University of Denmark) predicts that charged residues of the endocan polypeptide are segregated on its surface such that there would be an electronegative face and an adjacent electropositive face, more or less like a dipole (Sarrazin, personal communication).

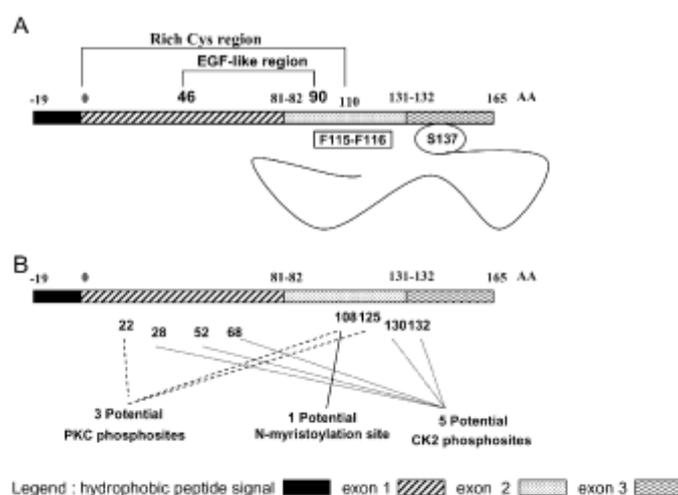


Fig. 2. Structural determinants and composition of endocan. Endocan is represented with its hydrophobic signal peptide (black box), followed by its exon 1- (hatched box), exon 2- (punctuated box), and exon 3- (waved box)- derived sequences. The amino-acid positions of the exon boundaries are marked. (A) Domain structure of endocan which contains an N-terminal cysteine-rich region (18 Cys / 110 aa), a putative EGF-like domain (aa 46 to 90), a phenylalanine-rich region (aa 115–116), and the S137 residue which is the site of glycosylation with a DS chain. (B) Positions of additional potential modifications sites; 3 potential sites of phosphorylation for PKC, 1 potential N-myristoylation site and 5 potential phosphorylation sites for CK2.

Inflammatory cytokines such as TNF- $\alpha$  increase the synthesis and the secretion of endocan from HUVEC [12]. In contrast, IFN- $\gamma$  clearly down-regulates the secretion of endocan whereas IL-4 has no effect [12]. The secreted, blood-borne endocan has been shown to be in the PG form in sera from healthy subjects [8,9,12,13]. Interestingly, the amount of the circulating PGs is markedly increased in patients with acute and severe sepsis. Endocan appears then to be also regulated *in vivo* by pro-inflammatory cytokines. Various PGs have been found circulating in the blood [6]. Their presence in the blood flow could enhance their physiological role as a co-receptor of multiple partners such as growth factors, cytokines or enzymes. Bikunin, for example, is a small CSPG that occurs in blood as the light chain of inter-alpha-trypsin inhibitor family members [34]. PG-100, also called CSF-1 (colony-stimulating factor-1) or M-CSF (macrophage colony-stimulating factor-1), is a growth factor which can be present in blood as a circulating PG form, carrying a single CS chain. Interestingly, PG-100 expression is also markedly upregulated in cultured human endothelial cells in the presence of TNF- $\alpha$  [35], comparably to what has been observed for endocan [7,12]. Among circulating PGs, the proteolytically-shed ectodomains of syndecan-1 and syndecan-4 (primarily viewed as purely transmembrane HSPGs) have been shown to have a role in complexing active molecules such as growth factors, enzyme or cytokines [10,36]. For instance, soluble syndecan-1 harvested from myeloma patients contains HGF/SF bound to its HS chains and the level of syndecan-1 ectodomain determines whether it potentiates or inhibits HGF/SF activity [10]. Similarly, HGF/SF function can also be modulated by altering the level of

soluble endocan, suggesting that, in general, the concentration of soluble PGs or GAGs may be a critical regulator of HGF/SF activities, and possibly other growth factors activities in regulation of tumor behaviour. DS and DSPGs have been shown to be abundant in human wound fluids. This soluble DS has been shown to support FGF-2-mediated cell proliferation and is likely to be an important molecule involved in the regulation of wound repair [11]. Interestingly, a significantly increased level of serum endocan has been observed not only in the serum from sepsis patients, but also in cancer patients as we will develop later.

#### 4. Endocan is a dermatan sulfate proteoglycan (DSPG)

In 2001, ESM-1 from cultured HUVEC and over-expressing HEK cells, as well as from normal human blood, was shown to be a dermatan sulphate proteoglycan (DSPG). At this point it thereby joined the “big family of proteoglycans”, and hence the name of “endocan” was given [9]. Degradation by the specific GAG depolymerising enzyme chondroitinase ABC *in vitro* reduced the size of secreted endocan on SDS-PAGE from approximately 50 kDa to 20 kDa, the latter size corresponding to the deglycosylated core protein [9]. This confirmed that the GAG chain belonged to the CS/DS sub-family of GAGs (Fig. 3). Further and more selective enzyme treatments revealed that the GAG was partially resistant to chondroitinase AC1, and partially susceptible to chondroitinase B, indicating that the GAG chain was a dermatan sulphate. Endocan from all sources investigated has never showed sensitivity to heparinase enzymes, indicating that it is always produced as a CS/DSPG and not a HSPG.

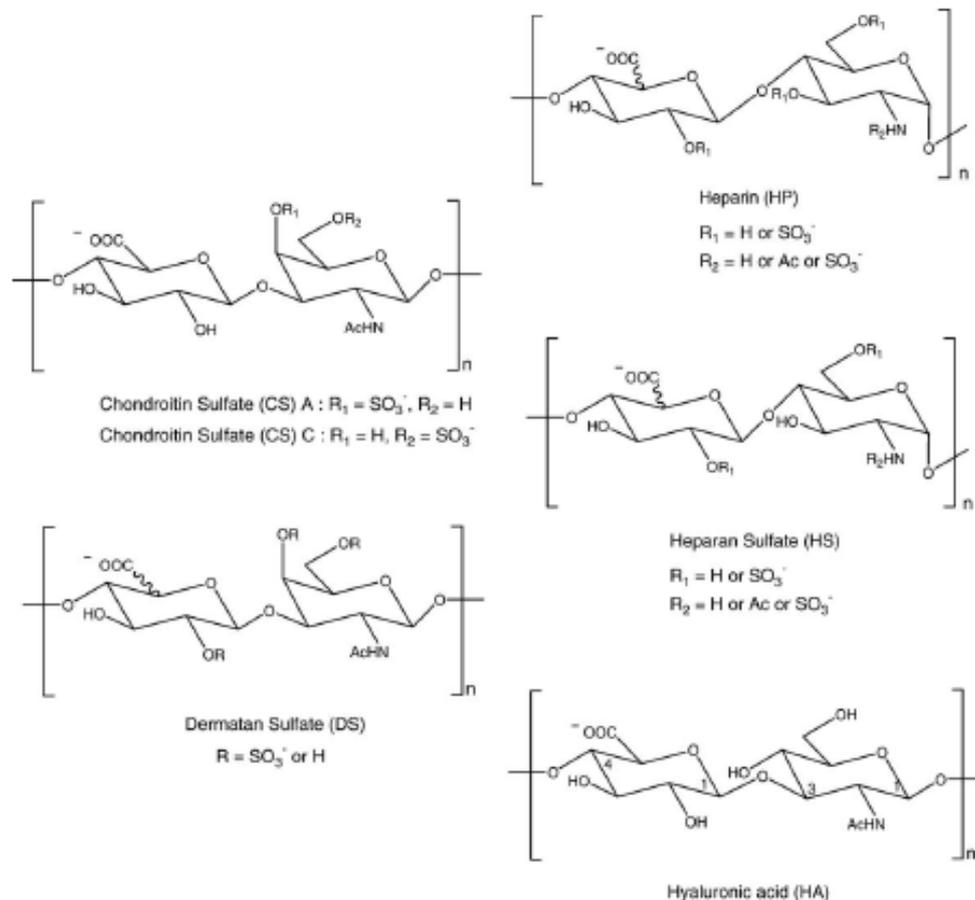


Fig. 3. Structural determinants in glycosaminoglycans. Structurally, glycosaminoglycans are linear polysaccharides consisting of repeating disaccharide units, in most cases composed of a uronic acid alternating with an N-substituted hexosamine, either glucosamine or galactosamine. HA is the simplest as it does not undergo further modifications. KS has the distinction of not containing a uronic acid residue, but having galactose instead. Heparin and HS are composed of alternating  $\alpha$ -D-glucosamine and hexuronate ( $\alpha$ -L-iduronate or  $\beta$ -D-glucuronate) residues. Dermatan sulphate (DS) like chondroitin sulphate (CS) is assembled from  $\beta$ 1,4-linked N-acetylgalactosamine (GalNAc) alternating with  $\beta$ 1,3-linked glucuronic acid (GlcA). Structurally, the common feature of CS/DS is thus the presence of GalNAc.

Structurally, the GAG family is comprised of four subfamilies of linear polysaccharides consisting of repeating disaccharide units, in most cases composed of a uronic acid alternating with an N-substituted hexosamine, either glucosamine or galactosamine (Fig. 3). These four subfamilies are HS/heparin, CS/DS, HA and KS. HA is the simplest as it does not undergo further modifications. Moreover, HA is synthesized as a free polysaccharide chain and never on a protein core to form a PG [24]. Heparin and HS are composed of alternating  $\alpha$ -D-glucosamine and hexuronate ( $\alpha$ -L-iduronate or  $\beta$ -D-glucuronate) residues (Fig. 3). KS has the distinction of not containing a uronic acid residue, but having galactose instead. With the exception of HA, all other GAGs experience a complex series of post-polymeric modifications which can include epimerisation of glucuronic acid to iduronic acid (IdoA) and possible sulphations at various O- and N-positions on either monosaccharide unit.

Dermatan sulphate (DS) like chondroitin sulphate (CS) is assembled from  $\beta$ 1,4-linked N-acetylgalactosamine (GalNAc) alternating with  $\beta$ 1,3-linked glucuronic acid (GlcA). Structurally, the common feature of CS/DS is thus the presence of GalNAc (Fig. 3). However, conversion of a proportion of glucuronate residues to iduronate defines a DS, and distinguishes it from chondroitin 4-O-sulphate and chondroitin 6-O-sulphate which both preserve a 100% GlcA content. This feature of DS therefore likens it to heparin and HS which also contain IdoA residues. IdoA appears to play a key role in binding site specificity for HS/heparin-binding proteins (reviewed in [4]), which probably reflects the role of the flexibility of this sugar ring in mediating interactions [37]. It is likely that this is the same for DS as we are now seeing a greater recognition of the protein-binding properties of DS, relative to the apparently much less interactive CS. The sequence information within a DS chain is thus composed of

three potential variables at the uronic acid position (i.e., GlcA, IdoA or 2-*O*-sulphated IdoA) and four variables at the hexosamine position (i.e., GalNAc, 4-*O*-sulphated GalNAc, 6-*O*-sulphated GalNAc, or 4-*O*-, 6-*O*-disulphated GalNAc) (Fig. 3). The uronate epimerization reaction, responsible for conversion of GlcA to IdoA, and the subsequent sulfation reactions, are not random but reflect a controlled enzymatic system for encrypting functional information into the GAG moiety [24].

Therefore, to understand the functional significance of endocan, one should consider both the core protein and the GAG structure that together contribute to making the mature DSPG. The DS chain isolated from endocan produced by over-expressing HEK293 cells has been characterized in greatest detail [9]. Disaccharide analysis by strong anion-exchange HPLC revealed a clear preponderance of 4-*O*-sulphated GalNAc (66.5%) over 6-*O*-sulphated GalNAc residues (19.4%), with the remainder being non- and disulphated species [9]. Surprisingly, the IdoA content was relatively low (6%). Moreover, the results of specific enzyme digestions and analyses of resulting fragmentation patterns indicate that iduronates occur singly, or in small and relatively widely spaced clusters in this endocan [9]. These IdoA however appeared to be critical residues in modulating the binding abilities and therefore the biological activities of PGs [4,38,39]. One of the first biological activities described for endocan was attributed to the DS chain, which acts as a co-factor for HGF/SF [9]. HGF/SF is mainly produced by fibroblasts and vascular smooth muscle cells and acts mostly on epithelial and endothelial cells, but also on some stromal cells themselves [40,41]. Endocan strongly increased the HGF/SF-induced proliferation of both HEK 293 cells [9] and HaCat keratinocytes (Delehedde, Femig and Lassalle, personal communication), as detected by [<sup>3</sup>H]thymidine incorporation. An endocan mutant which does not possess the DS chain had no effect on HGF/SF-induced proliferation [9]. Similarly to intact endocan, the purified DS chain from endocan was able to stimulate proliferation, clearly highlighting the role of the DS moiety in driving some of the biological activities of endocan [9]. The DS chain does indeed interact directly with HGF/SF. The kinetics of the interaction between endocan and HGF/SF were analyzed on an optical biosensor and a  $K_d$  of 3–6 nM was determined [9]. HGF/SF has previously been shown to bind both HS/heparin and DS [38,42–46]. The minimal binding site in HS/heparin was identified as a tetrasaccharide [44,46], with a  $K_d$  for the heparin of 9–16 nM, as measured on a biosensor [44]. For DS, this minimal size is a hexasaccharide as characterized by gel mobility shift assay [46]. At present, the basis for this size difference remains unexplained. However, the fact that the minimal HS and DS oligosaccharides do effectively compete with each other for binding to HGF/SF does suggest that they both occupy approximately the same binding site on the protein. A combination of iduronate and hexosamine *O*-sulphation would appear to be critical for HGF/SF binding, with the nature of the *N*-substituent and the presence of 2-*O*-sulphation not being important [42,43,46]. The importance of

the iduronates is reflected by the fact that a CS of similar sulphation level to DS, but of course lacking IdoA, failed to bind HGF/SF [43]. Furthermore, studies with an invertebrate DS, in which the GalNAc residues are 6-*O*-sulphated rather than possessing the 4-*O*-sulphation encountered in mammalian DS species, is equally effective at binding to HGF/SF (Catlow, Deakin, Delehedde, Femig, Pavão and Lyon, unpublished data). This, surprisingly, demonstrates a tolerance for *O*-sulphation at either 4-*O*- or 6-*O*-positions in DS. Interestingly in the endocan produced by HEK cells, only 6% of the uronates are IdoA and yet they support high affinity binding to HGF/SF [9], consistent with the idea that the few IdoA residues are probably clustered, and that only a relatively short consecutive sequence of them is required to constitute a binding site for HGF/SF. A high-affinity interaction of GAGs with HGF/SF is critical for its activation of the Met receptor and subsequent generation of a sustained downstream signal leading to a cellular response [38,44, 47,48]. Interaction of endocan with HGF/SF, and promotion of its activity, are thus of potential importance, as aberrant expression of either HGF/SF or Met, or both, has been implicated in tumorigenesis, angiogenesis and metastasis [41]. Furthermore, in addition to the recently documented effects on HGF/SF, DS has now been shown to bind to a diverse range of other proteins including growth factors, cytokines, chemokines, matrix macromolecules and protease inhibitors, reminiscent of the situation with HS/heparin [4,39]. Among the chemokines, RANTES [49] and platelet factor-4 [50,51] have been shown to bind to DS, as does the cytokine IFN- $\gamma$  [52]. Among growth factors, DS has been shown to bind both FGF-2 and FGF-7 [11,53,54]. The interaction with FGF-2 is of low affinity with a  $K_d$  of 2.48  $\mu$ M, as assessed by affinity co-electrophoresis [11]. Moreover, similar to what has been previously shown for heparin and HS (reviewed in [4]), DS not only binds but also promotes FGF-2 and FGF-7 dependant cellular proliferation [11,53,54]. Among proteins of the extracellular matrix, DS was shown to bind to collagen and fibronectin [18,55]. Furthermore, DS was shown to bind heparin cofactor II [56–58]. Interestingly, DS selectively catalyzes the inactivation of thrombin by heparin cofactor II by an allosteric mechanism, but without interacting with antithrombin III. Therefore, DS differs from heparin in that it acts on the coagulation system as a selective thrombin inhibitor and has been therefore tested as an effective antithrombotic drug [58]. However, the precise binding sites for many of these protein interactions have not been fully identified yet, nor have the binding affinities for different structural variants of DS. It is thus clear that endocan has the potential to interact via its DS chain with many of the proteins already mentioned as well as potentially many, as yet, undescribed ones, as the binding properties of DS have been relatively underinvestigated in comparison with those of heparin/HS. Clearly, much more investigation needs to be undertaken, especially with those proteins involved in inflammatory processes.

In parallel to the binding properties of the DS chain, the core protein of endocan can bind directly in vitro to the

surface of human blood leukocytes, including both monocytes and lymphocytes [8]. Among other DSPGs, there are also a few examples of where interactions separately mediated by the DS chain and the core protein co-operate in a biological function of the intact PG, which may also be of relevance to the function of endocan. For example, in matrix organization, decorin binds collagen via its protein core, and tenascin via its DS chain [18,59]. The core protein of endocan has been shown to bind to the integrin CD11a/CD18, also called LFA-1 (Lymphocyte Function-associated Antigen 1). The core protein of endocan and LFA-1 dynamically interact with high affinity and a  $K_d$  of 18.7 nM as determined on a biosensor [8]. Moreover, endocan inhibits the specific binding of soluble intercellular adhesion molecule 1 (ICAM-1) to lymphocytes in a dose-dependent manner [8] suggesting that endocan can inhibit the adhesion and/or the activation of leukocytes that is depending on the LFA-1 pathway. Increased levels of circulating endocan in blood were also shown to be able to prevent the adhesion of leukocytes to the endothelium in a rat model of inflammation [60]. Control of the circulating endocan involved in the regulation of these processes should then be of prime importance in the control of inflammatory responses, and also in tumor development-associated events.

#### 5. Endocan is a circulating DSPG originally described as a product of endothelial cells

Endocan was originally described as being secreted by human umbilical vein endothelial cells (HUVEC) [7–9, 12,14,17]. Endocan transcripts were also detected from cultured endothelial cells of different origins, i.e., in HCAEC (human coronary artery endothelial cells), HPAEC (human pulmonary artery endothelial cells), HDMVEC (human dermal microvascular endothelial cells) and HUCE (human capillary endothelial cells purified from adipose tissues) [17,23]. Other secreted PGs have been described from human endothelial cells. For instance, the large modular HSPG called perlecan is abundant in the extracellular matrix and basement membranes of normal tissues and blood vessels [18]. In addition, CS/DSPGs, including biglycan and PG-100 have also been shown to be produced by human endothelial cells [35]. In contrast to what has been observed with endocan [7,12], TNF- $\alpha$  has been shown to negatively regulate the expression of biglycan in cultured human endothelial monolayers [35]. Among transmembrane PG, the different members of the syndecan family (syndecan 1 to syndecan 4) as well as CD44 isoforms have been also shown to be largely expressed by endothelial cells [61–65]. TNF- $\alpha$  specifically, and dose-dependently, also suppressed syndecan-1 HSPG expression at both the mRNA and proteins levels in these same cells [63]. Moreover, PGs of the vascular endothelium have been described for years now as receptors of molecules involved in leukocyte recruitment, which is the earliest event observed in an inflammatory response [6,36,66]. This recruitment into tissue is regulated by a sequence of interactions between the circulating leukocytes and the endothelium, and is mediated by adhesion receptors which include integrins,

selectins and cell surface PGs [67,68]. Leukocyte adhesion is partially mediated by binding of Mac-1 to HSPGs on endothelial cells [69]. During the recruitment of leukocytes, HSPGs on endothelial cells immobilize and present chemokines such as macrophage inflammatory protein 1 $\beta$  (MIP-1 $\beta$ ), interleukin 8 (IL-8) and platelet factor 4 (PF-4), which consequently induce the activation of leukocyte integrins, and then mediate stable leukocyte adhesion and transmigration of the cells across the vessel wall [66,70]. In this context, there are various examples of how both soluble GAGs and soluble PGs are able to inhibit the recruitment of leukocytes [68,71,72]. For instance, heparin has been shown to interfere with the initial selectin–ligand interaction [71,72]. Interestingly, a marked attenuation of leukocyte–endothelium interactions was associated with an increase in circulating endocan in a rat model of inflammation [60]. Furthermore, PGs have been implicated in angiogenesis and vasculogenesis [2,3,5,73], with most of their effects being due to their interactions with angiogenic GAG-binding growth factors such as FGF-2, VEGF and HGF/SF [4,6,44]. Moreover, such angiogenic growth factors also seem to be involved in the regulation of endocan mRNA expression *in vitro* as well as endocan secretion [14,23] (Adam, Sarrazin, Depontieu, Lassalle and Delehedde, personal communication).

In endothelial cells such as HUVEC, endocan is also present in cell lysates with an apparent molecular weight of 20 kDa, as shown by SDS-PAGE. This mass corresponds to the mature polypeptide of 165 amino acids without any GAG chain [8,9,12]. However, no role of this isoform present in the endothelial cell lysates has been proposed yet. Although endocan expression was originally shown to be restricted to endothelial cells, leading to the original naming as endothelial cell specific molecule-1 [7,17,23], and then subsequently endocan, there have been lately few exceptions seen in *in vitro* cultures. These are adipocytes [27], melanoma cells involved in the process of vasculomimicry, which is the process describing the ability of aggressive melanoma cells to express endothelial-associated genes, and also glioblastoma cells and renal cells [28,29,74], (Adam et al., unpublished). By RT-PCR, human primary adipocytes, as well as adipocytes isolated from fat depots of Sprague–Dawley rats, were shown to express endocan. This expression was stimulated in a dose-dependent manner by phorbol ester, an activator of protein kinase C [27]. Up to now, no other experimental or clinical data has yet been forthcoming to further characterize these differential expressions of endocan in normal, vascularized or tumoral tissues.

#### 6. Endocan in tumor progression

Recently, the consequences of the overexpression of endocan on tumor progression were described [13]. The *in vitro* activities of endocan, such as regulation of HGF/SF, suggested a potential role for endocan in the control of tumor growth and development, as well as angiogenesis during tumor progression. PGs have been proposed as critical regulators of tumor invasion and metastasis [2,3,5,73,75]. The tumor growth promoting properties of endocan were first observed in a mouse model of human tumor xenografts [13]. When HEK293

cells engineered to over-express endocan were injected subcutaneously into SCID mouse, tumors were formed at the site of inoculation in a marked contrast to their complete lack of effect with the normal counterpart HEK293 cells. In another experiment using the naturally tumorigenic colon carcinoma HT29 cells, those cells over-expressing endocan presented tumors of twice the volume of those produced by wild type HT29 cells, at 8 weeks after subcutaneous injection into the flank of SCID mice. No satellite lymph node invasion or metastatic dissemination were observed in both cases [13]. Moreover, endocan with the mutation at serine 137, which disrupts the DS glycanation site, did not give rise to tumors in the HEK293 cell tumor model, emphasizing the role of the GAG moiety [13]. By virtue of the multiplicity of their protein binding partners (e.g., growth factors, cytokines, chemokines, enzymes, etc.), it has been known for many years that PGs play a major role in regulating disease processes such as proliferation, remodelling, metastasis or angiogenesis during cancer progression [1,2,4,5]. Together, with the quantitative changes in PG expression observed during cancer progression, associated and significant structural alterations within GAGs have been demonstrated regarding their length, their sulphation pattern and the extent of uronate epimerization. Marked structural alterations of GAGs, relative to normal counterparts, have been recorded in tissues biopsies from rectal, colon, gastric, liver and pancreatic carcinomas [76–82]. An increase in HA is also a common phenomenon in the matrix of malignant tissue with a proposed role in facilitating cancer cell migration and the tumor HA content does correlate with metastatic ability [83]. Within HS, it has been shown that specific structural changes take place *in vitro* during the tumorigenic progression from human colon adenoma to carcinoma [78]. Similarly, in colon cancer, the sulfation pattern of tumor-associated CSPGs also appears to be altered, with increased levels of non- and 6-*O*-sulfated disaccharide units [80,82]. A switch from the synthesis of DS to CS chains on versican and decorin has been observed in tissue biopsies from patients with colon cancer [80]. Such structural modifications during cancer progression, if also occurring within the GAG chain of endocan, could be of great importance in how endocan regulates the biological activities of various mediators such as the angiogenic growth factors (including HGF/SF). Such studies need to be addressed to fully understand the role of this dually functional molecule.

When cells overexpressing the variant of endocan bearing a mutation in serine 137, which disrupts the glycanation site, are injected subcutaneously into SCID mice, they did not give rise to tumors [13]. Other PGs have been shown to have a tumor suppressor role, such as the HSPGs, glypican 3 and syndecan-1 [84] and such as the CS/DSPG decorin [73,75,85]. Decorin inhibits cell adhesion, migration and tube-like formation of endothelial cells *in vitro* [55,86]. Tumor cells engineered to over-express decorin show decreased vascularization of tumors *in vivo*, highlighting that decorin can prevent angiogenesis [75]. Moreover, Iozzo and collaborators have demonstrated that the expression of a decorin without its GAG chains can prevent metastatic spreading of breast cancer [73]. Treatment with the

decorin core protein reduced primary tumor growth by 70% and eliminated the occurrence of metastases [73]. Interestingly, within the protein core of endocan, there exists another peptide region, separate from the glycanation site that seems also to be relevant for its anti-tumoral effect. Cells over-expressing a mutant endocan in which there has been specific mutagenesis of both F115 and F116 to alanine, failed also to give rise to tumors in the xenograft mouse models [13]. Further investigations need to be done to better understand the contribution of this particular region of endocan to its tumorigenic activity in the xenograft models.

In human tissues, high endocan mRNA levels have been shown to correlate with poor prognosis and metastasis in several types of cancers, including breast, renal and lung cancer (Table 1). In breast cancer, a study of 78 patients, with the aim of defining the optimal prognosis classifier, was performed on 70 genes according to standard prognostic criteria [87]. They described endocan overexpression in breast cancer as being associated with a higher risk of metastasis and death within 5 years [87]. Furthermore, among 1234 genes identified as being differentially expressed in renal cell carcinoma, endocan mRNA levels showed a marked 3-fold increase in a parametric analysis of microarray data [88]. This upregulation of endocan expression also correlated with increases in both tumor vasculature and inflammation in renal cancer, which is actually the ninth most common malignancy in Western countries, with no present effective treatment [88]. Similarly, in a recent extensive hybridization study, the endocan gene was one of those most highlighted, with at least a 2-fold upregulation in all of the 8 renal cell carcinoma samples analyzed, compared to normal tissues [89]. Interestingly, a parallel upregulation was also revealed for VEGF and the c-Met proto-oncogene receptor for HGF/SF, both of which are heavily implicated in angiogenesis [89]. In a comparable study, by dot blotting and hybridization, endocan was again dramatically upregulated in several (5/14) renal cell carcinoma biopsies, and was correlated with both VEGF and VEGF receptor gene expressions [23]. In lung cancer, a gene profiling study of tissues from 23 patients identified endocan as being one of the significant poor prognosis classifiers among the 42 genes associated with high risk for cancer death [90]. High endocan mRNA levels

Table 1  
Endocan expression in human tumors

Tumors	Number of patients included in the study	Associated markers	References
Breast	78	Higher risk of metastasis	[87]
Kidney	8	VEGF, c-met	[89]
	9	Tumor vasculature increase and inflammation	[88]
Uterus, stomach, rectum, breast, kidney	14	VEGF, VEGF-R	[23]
Lung	23	Higher risk of metastasis	[90]

Endocan mRNA levels have been analyzed in several types of cancers including breast, renal and lung cancer, and correlated with other associated markers.

correlated also with poor prognosis and death within 12 months in lung carcinoma [90]. Finally, a more limited study has again shown that endocan mRNA overexpression can also be associated with angiogenesis (Table 1), in view of its correlation with different markers of angiogenesis such as VEGF and VEGF receptors in breast, uterine, and rectal human tumor samples [23].

As described above, endocan mRNA is clearly overexpressed in tumors (human tissue samples and xenogenic models). Moreover, blood levels of circulating endocan, as measured by ELISA, were found to be markedly increased in the sera of late-stage lung cancer patients [13]. Taken together, these various data on endocan highlight the potential of this molecule as a marker of tumor aggressiveness. Nowadays, biomarkers are essential to assess pharmacodynamics (setting the dosage) and for selecting those patients most likely to respond to a particular treatment. Such markers have yet to be identified and validated for use in clinical trials of drugs. Endocan could be an effective biomarker to be developed for clinical settings in cancer. If endocan represents a potential target for novel drugs to treat cancer, then one strategic therapeutic approach could be to target soluble endocan in the pathological environment using a specific monoclonal antibody. Diverse immunotherapy approaches have achieved success in controlling individual aspects of immune responses in animal models. Correspondingly, there have been few successes in the use of monoclonal antibodies in cancer therapy from the transfer of such immunotherapies to clinical trials [91]. It has taken a while since the development of monoclonal antibodies by Kohler and Milstein for these reagents to fulfil their potential promise in immunotherapy. Application of innovative recombinant DNA technology, such as chimerization and humanization, have enhanced the clinical efficiency of mouse antibodies and has led to approvals by the Food and Drug Administration (FDA) for the clinical use of therapeutic immunoglobulins and Fab molecules (monovalent antibody fragments produced by proteolysis) [92]. A monoclonal antibody against CD-20, called rituximab, was the first to be approved to treat malignancy, and it showed an overall response rate of 57% for patients with relapsed low grade non Hodgkin lymphoma [91,92]. Another example has been the humanized antibody against the HER/neu tyrosine kinase receptor, called trastuzumab (or Herceptin), which exhibited an overall response rate of 15% in breast cancer patients. Both rituximab and Herceptin have also been shown to improve the overall survival of patients when used to supplement standard chemotherapy in randomized trials. The VEGF-specific antibody bevacizumab (Avastin, Genentech) was shown to have antivasculature effects in human rectal cancer in phase I clinical trials [93]. Recombinant antibodies have also been fused with radionuclides, toxins, enzymes, liposomes or viruses to try and increase their therapeutic impact [91]. There will probably be further developments in the near future to increase the sensitivity and efficacy of these kinds of therapeutic agents. Interestingly, Scherpereel and collaborators have clearly demonstrated the effectiveness of a monoclonal antibody against endocan in an *in vivo* model system [13]. This antibody

decreased tumor incidence and tumor growth, compared to a control antibody, when injected *i.p.* at 200  $\mu\text{g}/\text{week}$  into SCID mice that had previously received a subcutaneous injection of endocan-overexpressing HEK293 cells, which normally leads to tumor growth. Such blocking antibodies clearly target an epitope in endocan that is (or is close to) a critical determinant for tumor growth, which still needs to be located. In contrast, monoclonal antibodies that target the C-terminal of endocan have no blocking effect on tumor growth [13].

In conclusion, since its cloning in 1996, endocan (ESM-1) has already been demonstrated to have important roles in several pathophysiological processes including inflammatory disorders and tumor progression. More experimental evidence is continually appearing on the involvement of endocan in the control of fundamental cellular processes such as adhesion, migration and angiogenesis, and no doubt that we will soon be learning more on the mechanism of this interesting proteoglycan. Already, the existing set of data suggests that endocan should be considered as a valid therapeutic target in cancer.

#### Acknowledgements

Stéphane Sarrazin is partly supported by a fellowship from the ANRT (Association Nationale de la Recherche Technique, France). Florence Depontieu is a recipient of a fellowship from the MRT (Ministère de la Recherche et de la Technologie, France). We are grateful to Stéphane Lequien and Olivier Querolle for the critical reading of the manuscript.

#### References

- [1] R.V. Iozzo, I. Cohen, Altered proteoglycan gene expression and the tumor stroma, *Experientia* 49 (1993) 447–455.
- [2] R.V. Iozzo, J.D. San Antonio, Heparan sulfate proteoglycans: heavy hitters in the angiogenesis arena, *J. Clin. Invest.* 108 (2001) 349–355.
- [3] J. Timar, K. Lapis, J. Dudas, A. Sebestyén, L. Kopper, I. Kovalszky, Proteoglycans and tumor progression: Janus-faced molecules with contradictory functions in cancer, *Semin. Cancer Biol.* 12 (2002) 173–186.
- [4] M. Delehedde, Heparan sulfate as a key player in the regulation of growth factor activities, *Recent Res. Dev. Biol. Chem.* 1 (2002) 133–149.
- [5] R.D. Sanderson, Y. Yang, L.J. Suva, T. Kelly, Heparan sulfate proteoglycans and heparanase—Partners in osteolytic tumor growth and metastasis, *Matrix Biol.* 23 (2004) 341–352.
- [6] M. Delehedde, H. Lortat-Jacob, J.T. Gallagher, D. Bonaffé, E. Adam, O. Querolle, S. Lequien, S. Degove, P. Lassalle, D. Becharard, Proteoglycan involvement in inflammatory diseases. New developments in GAG-based therapies, *Curr. Med. Chem.* 2 (2005) 345–357.
- [7] P. Lassalle, S. Molet, A. Janin, J.V. Heyden, J. Tavemier, W. Fiers, R. Devos, A.B. Tonnel, ESM-1 is a novel human endothelial cell-specific molecule expressed in lung and regulated by cytokines, *J. Biol. Chem.* 271 (1996) 20458–20464.
- [8] D. Becharard, A. Scherpereel, H. Hammad, T. Gentina, A. Tsicopoulos, M. Aumercier, J. Pestel, J.P. Dessaint, A.B. Tonnel, P. Lassalle, Human endothelial-cell specific molecule-1 binds directly to the integrin CD11a/CD18 (LFA-1) and blocks binding to intercellular adhesion molecule-1, *J. Immunol.* 167 (2001) 3099–3106.
- [9] D. Becharard, T. Gentina, M. Delehedde, A. Scherpereel, M. Lyon, M. Aumercier, R. Vazeux, C. Richet, P. Degand, B. Jude, A. Janin, D.G. Ferrig, A.B. Tonnel, P. Lassalle, Endocan is a novel chondroitin sulfate/dermatan sulfate proteoglycan that promotes hepatocyte growth factor/scatter factor mitogenic activity, *J. Biol. Chem.* 276 (2001) 48341–48349.

- [10] C. Seidel, M. Borset, O. Hjertner, D. Cao, N. Abildgaard, H. Hjorth-Hansen, R.D. Sanderson, A. Waage, A. Sundan, High levels of soluble syndecan-1 in myeloma-derived bone marrow: modulation of hepatocyte growth factor activity, *Blood* 96 (2000) 3139–3146.
- [11] S.F. Penc, B. Pomahac, T. Winkler, R.A. Dorschner, E. Eriksson, M. Herndon, R.L. Gallo, Dermatan sulfate released after injury is a potent promoter of fibroblast growth factor-2 function, *J. Biol. Chem.* 273 (1998) 28116–28121.
- [12] D. Bechard, V. Meignin, A. Scherpercel, S. Oudin, G. Kervoaze, P. Bertheau, A. Janin, A. Tonnel, P. Lassalle, Characterization of the secreted form of endothelial-cell-specific molecule 1 by specific monoclonal antibodies, *J. Vasc. Res.* 37 (2000) 417–425.
- [13] A. Scherpercel, T. Gentina, B. Grigoriu, S. Senechal, A. Janin, A. Tsicopoulos, F. Plenat, D. Bechard, A.B. Tonnel, P. Lassalle, Overexpression of endocan induces tumor formation, *Cancer Res.* 63 (2003) 6084–6089.
- [14] M.E. Gerritsen, J.E. Tomlinson, C. Zlot, M. Ziman, S. Hwang, Using gene expression profiling to identify the molecular basis of the synergistic actions of hepatocyte growth factor and vascular endothelial growth factor in human endothelial cells, *Br. J. Pharmacol.* 140 (2003) 595–610.
- [15] C.Y. Chen, A.B. Shyu, AU-rich elements: characterization and importance in mRNA degradation, *Trends Biochem. Sci.* 20 (1995) 465–470.
- [16] M.A. Fzevel, T. Bakheet, A.M. Silva, J.G. Hissong, K.S. Khabar, B.R. Williams, p38 Mitogen-activated protein kinase-dependent and -independent signaling of mRNA stability of AU-rich element-containing transcripts, *Mol. Cell. Biol.* 23 (2003) 425–436.
- [17] J.C. Tsai, J. Zhang, T. Minami, C. Voland, S. Zhao, X. Yi, P. Lassalle, P. Oettgen, W.C. Aird, Cloning and characterization of the human lung endothelial-cell-specific molecule-1 promoter, *J. Vasc. Res.* 39 (2002) 148–159.
- [18] R.V. Iozzo, Matrix proteoglycans: from molecular design to cellular function, *Annu. Rev. Biochem.* 67 (1998) 609–652.
- [19] T. Hai, M.G. Hartman, The molecular biology and nomenclature of the activating transcription factor/cAMP responsive element binding family of transcription factors: activating transcription factor proteins and homeostasis, *Gene* 273 (2001) 1–11.
- [20] T. Oikawa, T. Yamada, Molecular biology of the Ets family of transcription factors, *Gene* 303 (2003) 11–34.
- [21] K. Wakiya, A. Begue, D. Sebodin, M. Shibuya, A cAMP response element and an Ets motif are involved in the transcriptional regulation of fib-1 tyrosine kinase (vascular endothelial growth factor receptor 1) gene, *J. Biol. Chem.* 271 (1996) 30823–30828.
- [22] S. Goy, J. Dalmon, M.H. Prandini, T. Kostulewski, Y. de Launoit, P. Huber, Requirement of a GT box (Sp1 site) and two Ets binding sites for vascular endothelial cadherin gene transcription, *J. Biol. Chem.* 273 (1998) 6750–6755.
- [23] M. Aitkenhead, S.J. Wang, M.N. Nakatsu, J. Mestas, C. Heard, C.C. Hughes, Identification of endothelial cell genes expressed in an in vitro model of angiogenesis: induction of ESM-1, (beta)ig-h3, and NCAM, *Microvasc. Res.* 63 (2002) 159–171.
- [24] K. Prydz, K.T. Dalen, Synthesis and sorting of proteoglycans, *J. Cell. Sci.* 113 (2000) 193–205.
- [25] B.M. Brinkman, Splice variants as cancer biomarkers, *Clin. Biochem.* 37 (2004) 584–594.
- [26] Z. Rudzki, S. Jothy, CD44 and the adhesion of neoplastic cells, *Mol. Pathol.* 50 (1997) 57–71.
- [27] M. Welner, F. Herse, J. Janke, K. Gorzelnik, S. Engeli, D. Bechard, P. Lassalle, F.C. Luft, A.M. Sharma, Endothelial cell specific molecule-1—A newly identified protein in adipocytes, *Horm. Metab. Res.* 35 (2003) 217–221.
- [28] M.J. Hendrix, E.A. Sefror, A.R. Hess, R.E. Sefror, Molecular plasticity of human melanoma cells, *Oncogene* 22 (2003) 3070–3075.
- [29] E.A. Sefror, P.S. Meltzer, G.C. Schatteman, L.M. Guman, A.R. Hess, D.A. Kirschmann, R.E. Sefror, M.J. Hendrix, Expression of multiple molecular phenotypes by aggressive melanoma tumor cells: role in vasculogenic mimicry, *Crit. Rev. Oncol. Hematol.* 44 (2002) 17–27.
- [30] V. Hwa, Y. Oh, R.G. Rosenfeld, The insulin-like growth factor-binding protein (IGFBP) superfamily, *Endocr. Rev.* 20 (1999) 761–787.
- [31] L. Desmoyers, Structural basis and therapeutic implication of the interaction of CCN proteins with glycoconjugates, *Curr. Pharm. Des.* 10 (2004) 3913–3928.
- [32] X. Wang, M.J. Brownstein, W.S. Young, PG25, a pineal-specific cDNA, cloned by differential display PCR (DDPCR) and rapid amplification of cDNA ends (RACE), *J. Neurosci. Methods* 73 (1997) 187–191.
- [33] R.S. Sawhney, T.M. Hering, L.J. Sandell, Biosynthesis of small proteoglycan II (decorin) by chondrocytes and evidence for a procure protein, *J. Biol. Chem.* 266 (1991) 9231–9240.
- [34] L. Zhao, A. Salusti, K. Kimata, A physiological function of serum proteoglycan bikunin: the chondroitin sulfate moiety plays a central role, *Glycoconjugate J.* 19 (2002) 241–247.
- [35] L. Nelmarikka, V. Kainulainen, E. Schonherr, S. Moisaner, M. Jortikka, M. Lammi, K. Elenius, M. Jalkanen, H. Jarvelainen, Expression of small extracellular chondroitin/dermatan sulfate proteoglycans is differentially regulated in human endothelial cells, *J. Biol. Chem.* 272 (1997) 12730–12737.
- [36] M. Delehedde, F. Allain, S.J. Payne, R. Borgo, C. Vanpouille, D.G. Fernig, E. Deudon, Proteoglycans in inflammation, *Curr. Med. Chem.* 1 (2002) 89–102.
- [37] B. Casu, M. Petitou, M. Provasoli, P. Sinay, Conformational flexibility: a new concept for explaining binding and biological properties of iduronic acid-containing glycosaminoglycans, *Trends Biochem. Sci.* 13 (1988) 221–225.
- [38] M. Lyon, J.A. Deakin, J.T. Gallagher, The mode of action of heparan and dermatan sulfates in the regulation of hepatocyte growth factor/scatter factor, *J. Biol. Chem.* 277 (2002) 1040–1046.
- [39] J.M. Townbridge, R.L. Gallo, Dermatan sulfate: new functions from an old glycosaminoglycan, *Glycobiology* 12 (2002) 117–125.
- [40] M. Delehedde, N. Sergeant, M. Lyon, P.S. Rudland, D.G. Fernig, Hepatocyte growth factor/scatter factor stimulates migration of mammary fibroblasts through both mitogen-activated protein kinase and phosphatidylinositol 3-kinase/Akt pathways, *Eur. J. Biochem.* 268 (2001) 4423–4429.
- [41] N.A. Bhowmick, E.G. Neilson, H.L. Moses, Stromal fibroblasts in cancer initiation and progression, *Nature* 432 (2004) 332–337.
- [42] M. Lyon, J.A. Deakin, K. Mizuno, T. Nakamura, J.T. Gallagher, Interaction of hepatocyte growth factor with heparan sulfate. Elucidation of the major heparan sulfate structural determinants, *J. Biol. Chem.* 269 (1994) 11216–11223.
- [43] M. Lyon, J.A. Deakin, H. Rahmouni, D.G. Fernig, T. Nakamura, J.T. Gallagher, Hepatocyte growth factor/scatter factor binds with high affinity to dermatan sulfate, *J. Biol. Chem.* 273 (1998) 271–278.
- [44] M. Delehedde, M. Lyon, R. Vidyasagar, T.J. McDonnell, D.G. Fernig, Hepatocyte growth factor/scatter factor binds to small heparin-derived oligosaccharides and stimulates the proliferation of human HaCaT keratinocytes, *J. Biol. Chem.* 277 (2002) 12456–12462.
- [45] K. Catlow, J.A. Deakin, M. Delehedde, D.G. Fernig, J.T. Gallagher, M.S. Pavao, M. Lyon, Hepatocyte growth factor/scatter factor and its interaction with heparan sulphate and dermatan sulphate, *Biochem. Soc. Trans.* 31 (2003) 352–353.
- [46] M. Lyon, J.A. Deakin, D. Lietha, E. Gherardi, J.T. Gallagher, The interactions of hepatocyte growth factor/scatter factor and its NK1 and NK2 variants with glycosaminoglycans using a modified gel mobility shift assay. Elucidation of the minimal size of binding and activatory oligosaccharides, *J. Biol. Chem.* 279 (2004) 43560–43567.
- [47] J.A. Deakin, M. Lyon, Differential regulation of hepatocyte growth factor/scatter factor by cell surface proteoglycans and free glycosaminoglycan chains, *J. Cell. Sci.* 112 (1999) 1999–2009.
- [48] N. Sergeant, M. Lyon, P.S. Rudland, D.G. Fernig, M. Delehedde, Stimulation of DNA synthesis and cell proliferation of human mammary myoepithelial-like cells by hepatocyte growth factor/scatter factor depends on heparan sulfate proteoglycans and sustained phosphorylation of mitogen-activated protein kinases p42/44, *J. Biol. Chem.* 275 (2000) 17094–17099.
- [49] G.S. Kuschert, F. Coulin, C.A. Power, A.E. Proudfoot, R.E. Hubbard, A.J. Hoogwerf, T.N. Wells, Glycosaminoglycans interact selectively with

- chemokines and modulate receptor binding and cellular responses, *Biochemistry* 38 (1999) 12959–12968.
- [50] G. Cella, G. Boeri, G. Saggiaro, R. Paolini, G. Luzzatto, V.I. Terribile, Interaction between histidine-rich glycoprotein and platelet factor 4 with dermatan sulfate and low-molecular-weight dermatan sulfate, *Angiology* 43 (1992) 59–62.
- [51] F. Petersen, E. Brandt, U. Lindahl, D. Spillmann, Characterization of a neutrophil cell surface glycosaminoglycan that mediates binding of platelet factor 4, *J. Biol. Chem.* 274 (1999) 12376–12382.
- [52] B. Brooks, D.M. Briggs, N.C. Eastmond, D.G. Fernig, J.W. Coleman, Presentation of IFN-gamma to nitric oxide-producing cells: a novel function for mast cells, *J. Immunol.* 164 (2000) 573–579.
- [53] J.M. Trowbridge, J.A. Rudisill, D. Ron, R.L. Gallo, Dermatan sulfate binds and potentiates activity of keratinocyte growth factor (FGF-7), *J. Biol. Chem.* 277 (2002) 42815–42820.
- [54] K.R. Taylor, J.A. Rudisill, R.L. Gallo, Structural and sequence motifs in dermatan sulfate for promoting fibroblast growth factor-2 (FGF-2) and FGF-7 activity, *J. Biol. Chem.* 280 (2005) 5300–5306.
- [55] B. Merle, L. Duressat, P.D. Delmas, P. Clezardin, Decorin inhibits cell migration through a process requiring its glycosaminoglycan side chain, *J. Cell. Biochem.* 75 (1999) 538–546.
- [56] M.M. Maimone, D.M. Tollefsen, Structure of a dermatan sulfate hexasaccharide that binds to heparin cofactor II with high affinity, *J. Biol. Chem.* 265 (1990) 18263–18271.
- [57] G. Mascollani, L. Livemani, P. Bianchini, B. Parma, G. Torri, A. Bisio, M. Guernini, B. Casu, Structure and contribution to the heparin cofactor II-mediated inhibition of thrombin of naturally oversulphated sequences of dermatan sulphate, *Biochem. J.* 296 (1993) 639–648.
- [58] G.G. Nenci, Dermatan sulphate as an antithrombotic drug, *Pathophysiol. Haemost. Thromb.* 32 (2002) 303–307.
- [59] F. Eleftheriou, J.Y. Exposito, R. Garrone, C. Lethias, Binding of tenascin-X to decorin, *FEBS Lett.* 495 (2001) 44–47.
- [60] S. Tissier, S. Lancel, X. Marechal, S. Mordon, F. Depontieu, A. Scherpereel, C. Chopin, R. Neviere, Calpain inhibitors improve myocardial dysfunction and inflammation induced by endotoxin in rats, *Shock* 21 (2004) 352–357.
- [61] M. Murakami, A. Horowitz, S. Tang, J.A. Ware, M. Simons, Protein kinase C (PKC) delta regulates PKCalpha activity in a Syndecan-4-dependent manner, *J. Biol. Chem.* 277 (2002) 20367–20371.
- [62] C.A. Henke, U. Roongta, D.J. Mickelson, J.R. Knutson, J.B. McCarthy, CD44-related chondroitin sulfate proteoglycan, a cell surface receptor implicated with tumor cell invasion, mediates endothelial cell migration on fibrinogen and invasion into a fibrin matrix, *J. Clin. Invest.* 97 (1996) 2541–2552.
- [63] V. Kainulainen, L. Nelimäki, H. Jarvelainen, M. Laato, M. Jalkanen, K. Elenius, Suppression of syndecan-1 expression in endothelial cells by tumor necrosis factor- $\alpha$ , *J. Biol. Chem.* 271 (1996) 18759–18766.
- [64] M.A. Kulbeth, S.O. Kolset, T. Ranheim, Stimulation of senglycin and CD44 mRNA expression in endothelial cells exposed to TNF- $\alpha$  and IL-1 $\alpha$ , *Biochim. Biophys. Acta* 1428 (1999) 225–232.
- [65] C. Forster-Horvath, L. Meszaros, E. Raso, B. Dome, A. Ladanyi, M. Morini, A. Albini, J. Timar, Expression of CD44v3 protein in human endothelial cells in vitro and in tumoral microvessels in vivo, *Microvasc. Res.* 68 (2004) 110–118.
- [66] D.A. Steeber, G.M. Venturi, T.F. Tedder, A new twist to the leukocyte adhesion cascade: intimate cooperation is key, *Trends Immunol.* 26 (2005) 9–12.
- [67] G.G. Vaday, O. Lider, Extracellular matrix moieties, cytokines, and enzymes: dynamic effects on immune cell behavior and inflammation, *J. Leukocyte Biol.* 67 (2000) 149–159.
- [68] H. Schor, G.G. Vaday, O. Lider, Modulation of leukocyte behavior by an inflamed extracellular matrix, *Dev. Immunol.* 7 (2000) 227–238.
- [69] M.S. Diamond, R. Alon, C.A. Parkos, M.T. Quinn, T.A. Springer, Heparin is an adhesive ligand for the leukocyte integrin Mac-1 (CD11b/CD18), *J. Cell Biol.* 130 (1995) 1473–1482.
- [70] Y. Tanaka, D.H. Adams, S. Hubocher, H. Hirano, U. Siebenlist, S. Shaw, T-cell adhesion induced by proteoglycan-immobilized cytokine MIP-1 beta, *Nature* 361 (1993) 79–82.
- [71] R.M. Nelson, O. Cecconi, W.G. Roberts, A. Aruffo, R.J. Linhardt, M.P. Bevilacqua, Heparin oligosaccharides bind L- and P-selectin and inhibit acute inflammation, *Blood* 82 (1993) 3253–3258.
- [72] R. Lever, J.R. Hoult, C.P. Page, The effects of heparin and related molecules upon the adhesion of human polymorphonuclear leucocytes to vascular endothelium in vitro, *Br. J. Pharmacol.* 129 (2000) 533–540.
- [73] C.C. Reed, A. Watenhouse, S. Kirby, P. Kay, R.T. Owens, D.J. McQuillan, R.V. Iozzo, Decorin prevents metastatic spreading of breast cancer, *Oncogene* 24 (2005) 1104–1110.
- [74] D.T. Ross, U. Scherf, M.B. Eisen, C.M. Perou, C. Rees, P. Spellman, V. Iyer, S.S. Jeffrey, M. Van de Rijn, M. Waltham, A. Pergamenschikov, J.C. Lee, D. Lashkari, D. Shalon, T.G. Myers, J.N. Weinstein, D. Botstein, P.O. Brown, Systematic variation in gene expression patterns in human cancer cell lines, *Nat. Genet.* 24 (2000) 227–235.
- [75] D.S. Grant, C. Yenisey, R.W. Rose, M. Tootell, M. Santm, R.V. Iozzo, Decorin suppresses tumor cell-mediated angiogenesis, *Oncogene* 21 (2002) 4765–4777.
- [76] I. Kovalszky, G. Pogany, G. Molnar, A. Jeney, K. Lapis, S. Kamcsanyi, A. Szecseny, R.V. Iozzo, Altered glycosaminoglycan composition in reactive and neoplastic human liver, *Biochem. Biophys. Res. Commun.* 167 (1990) 883–890.
- [77] M.E. Tsam, N. Papageorgacopoulou, D.D. Karavias, D.A. Theocharis, Distribution and changes of glycosaminoglycans in neoplasias of rectum, *Anticancer Res.* 15 (1995) 2107–2112.
- [78] G.C. Jayson, M. Lyon, C. Paraskeva, J.E. Turnbull, J.A. Deakin, J.T. Gallagher, Heparan sulfate undergoes specific structural changes during the progression from human colon adenoma to carcinoma in vitro, *J. Biol. Chem.* 273 (1998) 51–57.
- [79] A.D. Theocharis, M.E. Tsara, N. Papageorgacopoulou, D.D. Karavias, D.A. Theocharis, Pancreatic carcinoma is characterized by elevated content of hyaluronan and chondroitin sulfate with altered disaccharide composition, *Biochim. Biophys. Acta* 1502 (2000) 201–206.
- [80] A.D. Theocharis, Human colon adenocarcinoma is associated with specific post-translational modifications of versican and decorin, *Biochim. Biophys. Acta* 1588 (2002) 165–172.
- [81] M.E. Tsara, A.D. Theocharis, D.A. Theocharis, Compositional and structural alterations of proteoglycans in human rectum carcinoma with special reference to versican and decorin, *Anticancer Res.* 22 (2002) 2893–2898.
- [82] A.D. Theocharis, D.H. Vynios, N. Papageorgacopoulou, S.S. Skandalis, D.A. Theocharis, Altered content composition and structure of glycosaminoglycans and proteoglycans in gastric carcinoma, *Int. J. Biochem. Cell Biol.* 35 (2003) 376–390.
- [83] S. Adamia, C.A. Maxwell, L.M. Pilanski, Hyaluronan and hyaluronan synthase: potential therapeutic targets in cancer, *Curr. Drug Targets Cardiovasc. Haematol. Disord.* 5 (2005) 3–14.
- [84] H. Kim, G.L. Xu, A.C. Borczak, S. Busch, J. Filmas, M. Capurro, J.S. Brody, J. Lange, J.M. D'Armiento, P.B. Rothman, C.A. Powell, The heparan sulfate proteoglycan GPC3 is a potential lung tumor suppressor, *Am. J. Respir. Cell Mol. Biol.* 29 (2003) 694–701.
- [85] J. Koniger, N.A. Giese, F.F. di Mola, P. Berberat, T. Giese, I. Esposito, M.G. Bachem, M.W. Buchler, H. Friess, Overexpressed decorin in pancreatic cancer: potential tumor growth inhibition and attenuation of chemotherapeutic action, *Clin. Cancer Res.* 10 (2004) 4776–4783.
- [86] M.R. De Lange, C. Davies, L.L. Munn, C. Mouta-Carreira, R.K. Jain, Y. Boucher, Decorin inhibits endothelial migration and tube-like structure formation: role of thrombospondin-1, *Microvasc. Res.* 62 (2001) 26–42.
- [87] L.J. van 't Veer, H. Dai, M.J. van de Vijver, Y.D. He, A.A. Hart, M. Mao, H.L. Peterse, K. van der Kooy, M.J. Manton, A.T. Witteveen, G.J. Schreiber, R.M. Kerhovan, C. Roberts, P.S. Linsley, R. Bernards, S.H. Friend, Gene expression profiling predicts clinical outcome of breast cancer, *Nature* 415 (2002) 530–536.
- [88] M.E. Lenburg, L.S. Liou, N.P. Geary, G.M. Frampton, H.T. Cohen, M.F. Christman, Previously unidentified changes in renal cell carcinoma gene expression identified by parametric analysis of microarray data, *BMC Cancer* 3 (2003) 31.
- [89] S. Amatschek, U. Koenig, H. Auer, P. Steinlein, M. Pachor, A.

- Gruenfelder, G. Dekan, S. Vogl, E. Kubista, K.H. Heider, C. Stratowa, M. Schreiber, W. Sommergruber, Tissue-wide expression profiling using cDNA subtraction and microarrays to identify tumor-specific genes, *Cancer Res.* 64 (2004) 844–856.
- [90] A.C. Boczuk, L. Shah, G.D. Pearson, K.L. Walter, L. Wang, J.H. Austin, R.A. Friedman, C.A. Powell, Molecular signatures in biopsy specimens of lung cancer, *Am. J. Respir. Crit. Care Med.* 170 (2004) 167–174.
- [91] M. Harris, Monoclonal antibodies as therapeutic agents for cancer, *Lancet Oncol.* 5 (2004) 292–302.
- [92] P.J. Hudson, C. Souriau, Engineered antibodies, *Nat. Med.* 9 (2003) 129–134.
- [93] C.G. Willett, Y. Boucher, E. di Tomaso, D.G. Duda, L.L. Munn, R.T. Tong, D.C. Chung, D.V. Sahani, S.P. Kalva, S.V. Kozin, M. Mino, K.S. Cohen, D.T. Scadden, A.C. Hartford, A.J. Fischman, J.W. Clark, D.P. Ryan, A.X. Zhu, L.S. Blaszkowsky, H.X. Chen, P.C. Shellito, G.Y. Lauwers, R.K. Jain, Direct evidence that the VEGF-specific antibody bevacizumab has antivasculature effects in human rectal cancer, *Nat. Med.* 10 (2004) 145–147.



## Proteoglycans and glycosaminoglycans: Complex molecules with modulating activity

Maryse Delehedde<sup>1</sup>, Stéphane Sarrazin<sup>1,2</sup>, Estelle Adzam<sup>1</sup>  
Vincent Motte<sup>1</sup> and Christophe Vampouille<sup>1</sup>  
ENDOTIS PHARMA, Parc Eurasat, 70 rue du Dr YERSIN, 59120 Loos  
France; <sup>1</sup>LEM GAG, Institut de Biologie Structurale, 38000 Grenoble, France  
Bethesda, USA; <sup>2</sup>National Institute of Health, Bethesda, USA

### Abstract

*Proteoglycans (PG) and glycosaminoglycans (GAG) have been shown to be key molecules in physiology and consequently their modulations have been observed in various pathologies. This review and the following chapters will describe the multiple facets of their structures and their functions, which have raise new opportunities in drug developments. PG and GAG are mainly classified according their*

**Correspondence/Reprint request:** Dr. Maryse Delehedde, ENDOTIS PHARMA, Parc Eurasat, 70 rue du Dr YERSIN, 59120 Loos, France. E-mail: [md@bioo.fr](mailto:md@bioo.fr) or [maryse.delehedde@endotis.com](mailto:maryse.delehedde@endotis.com)

localisation at the cell membrane, in the extracellular matrix, being intracellular or circulating in the blood. Structurally, PG are macromolecules which consist of a core protein and associated glycosaminoglycan (GAG) chains. The different GAG chains of PG, heparan sulfate/heparin, dermatan/chondroitin sulfate, keratan sulfate are synthesized as long polymers of repeating disaccharide units. They are strongly implicated in the different stages of development, cancer or inflammation. For instance, the repertoire of PG and GAG (natures and structures) has been shown to change continuously and dynamically during the progression of processes such as the inflammatory response or the development of tumours. These changes include the modulation of the activity of GAG-binding cytokines, growth factors, proteases and protease inhibitors. All the interactions of regulatory proteins (eg. growth factors, chemokines, enzymes) with GAGs provide much of the focus for GAG-based therapeutic targets. GAG mimetics should have therefore clinical applications as modulators of cytokine, growth factors or enzymes function in various diseases and pathologies.

### List of abbreviations

AT-III : antithrombin III ; CS : chondroitin sulfate ; CSPG : chondroitin sulfate proteoglycan ; DS : dermatan sulfate ; DSPG : dermatan sulfate proteoglycan ; ECM : extracellular matrix ; FGF-2 : fibroblast growth factor 2 ; GAG : glycosaminoglycan ; HA : hyaluronic acid ; HGF/SF : hepatocyte growth factor / scatter factor ; HS : heparan sulfate ; HSPG : heparan sulfate proteoglycan ; ITI : inter-alpha-trypsin inhibitor ; KS : keratan sulfate ; KSPG : keratan sulfate proteoglycan ; LMWH : low molecular weight heparin ; M-CSF : macrophage colony stimulating factor ; NDST : N-deacetylase/N-sulfotransferase ; PG : proteoglycan ; SHAP : Serum-derived Hyaluronan-Associated Protein ; SLRP : small leucin-rich proteoglycan ; TGFβ : transforming growth factor β ; TNFα : tumor necrosis factor α ; UFH : unfractionated heparin.

### Introduction

Proteoglycans (PG) and glycosaminoglycans (GAG) have been shown to be key molecules in physiology and consequently their modulations have been observed in various pathologies (1-7). PG and GAG are mainly classified according their localisation at the cell membrane, in the extracellular matrix, being intracellular or circulating in the blood. Structurally, PG are macromolecules which consist of a core protein and associated glycosaminoglycan (GAG) chains. The different GAG chains of PG, heparan sulphate (HS)/heparin, dermatan sulphate (DS), chondroitin sulphate (CS), keratan sulfate (KS) are synthesized as long polymers of repeating

disaccharide units (1-4, 7). They are strongly implicated in the different stages of development, wound healing, inflammation, angiogenesis, cancer and metastasis (7). The repertoire of PG and GAG (natures and structures) has therefore been shown to change continuously and dynamically during the progression of processes such as the inflammatory response or the development of tumors. These changes include the modulation of the activity of GAG-binding cytokines, growth factors, proteases and protease inhibitors. All the interactions of regulatory proteins (eg growth factors, chemokines, enzymes) with GAGs provide much of the focus for GAG-based therapeutic targets as we will described later.

### Structure of glycosaminoglycan chains

GAG are linear polysaccharides that consist of repeating disaccharide units composed of an uronic acid and an N-substituted hexosamine, glucosamine or galactosamine. The distinct types of GAG are then described as follows: chondroitin sulphate (CS), dermatan sulphate (DS), heparan sulphate (HS), heparin, keratan sulphate (KS) and hyaluronic acid (HA). HA, also called hyaluronan, is composed of repeating units of D-glucuronic acid and N-acetyl-D-glucosamine. HA is unique among GAG in that it is not sulfated, not attached covalently to a core protein and also that it is synthesized at the plasma membrane rather than in the Golgi (1-3). HA is synthesized in a wide range of sizes. Functional properties of HA include the maintenance and hydration of the extracellular matrix (ECM), and the binding to various molecules such as PG and CD44 receptors (3). KS possesses galactose in the disaccharide repeat characterizing their polymeric structure. KS has also the particularity among GAG that is not necessarily linear (4). KS may thus contain branching fucose residues, and may also be end-capped by various sugars such as sialic acids (4). Moreover KS are generally shorter than HA and the proteoglycan-forming GAG (4). CS consist typically of repeating disaccharide of glucuronic acid and N-acetylglucosamine. They are classified by the position of the sulfation of their hexosamine as C-4S or C-6S. As CS and DS are sharing the same biosynthetic polymer precursor, the presence of iduronic acid residues defines a GAG as a DS rather than a CS (1, 5, 6). However, PG such as aggrecan only contains pure CS whereas others, such as decorin may also contain CS or DS. According GAG nomenclature, heparin and HS are defined as the polymers composed of alternating D-glucosamine and hexuronate ( $\alpha$ -L-iduronic acid or  $\beta$ -D-glucuronate) residues (1, 7). The synthesis of heparin is largely restricted to mast cells which produce and release it with histamine. In contrast, HS is expressed as a part of a proteoglycan ubiquitously, on cell surfaces and throughout extracellular matrix and basement membranes or even circulating in the blood (7).

Table 1. The main characteristics of the glycosaminoglycans found in human

	Amino sugar	Uronic Acid or Galactose	Sulfation	Linked to a core protein
Heparan sulfate	N-acetyl-D-glucosamine (D-GlcNAc)	L-iduronic acid or D-glucuronate	Yes	Yes
Heparin	N-acetyl-D-glucosamine (D-GlcNAc)	L-iduronic acid	Yes	No
Chondroitin sulfate	N-acetyl-D-glucosamine (D-GlcNAc)	D-glucuronate	Yes	Yes
Dermatan sulfate	N-acetyl-D-glucosamine (D-GlcNAc)	L-iduronic acid or D-glucuronate	Yes	Yes
Keratan sulfate	N-acetyl-D-glucosamine (D-GlcNAc)	Galactose	Yes	Yes
Hyaluronan/hyaluronic acid	N-acetyl-D-glucosamine (D-GlcNAc)	D-glucuronate	No	No

Synthesis of HS/heparin as well as of CS/DS is initiated by the formation of a particular tetrasaccharide linker, composed of a xylose, galactose, galactose, and a glucuronic acid, with the xylose being the sugar covalently linked to a serine residue (1-7). Then, the sequential polymerisation of repeating units of glucuronic acid and N-acetylglucosamine (CS/DS) or N-acetyl glucosamine (HS/heparin) results a linear polysaccharide of the disaccharide glucuronic acid- $\beta$ 1-3 N-acetylglucosamine (CS/DS), or of the disaccharide glucuronic acid- $\beta$ 1-4 N-acetyl glucosamine (HS/heparin) in the initial product of GAG biosynthesis. Consecutive series of modification reactions act then on the monotonous initial polymer of HS/heparin or CS/DS and allow the cell to generate a large number of different structures within the chains of these GAG (5-7). In HS/heparin, the first modification is the N-deacetylation/N-sulfation reaction, which is catalyzed by the bifunctional N-deacetylase/N-sulfotransferases (NDST) (5, 8, 9). The N-sulfation has been shown to be critical for the domain structure and consecutive O-sulfation pattern in sulfated GAG. Clearly, 6-O-sulfation occurs preferentially on N-sulfated glucosamines or on N-acetylated glucosamines that are immediate neighbours of an N-sulfated glucosamine (5). Epimerisation of glucuronic acid to iduronic acid only occurs in the uronic acid linked to C-1 of an adjacent N-sulfated glucosamine (5,10,11). When iduronic acid is flanked on both sides by N-sulfates, it may then itself be 2-O sulfated. In contrast, the glucuronate residue is rarely 2-O sulfated. Moreover, 3-O-sulfation of hexosamines only occurs as a final modification of highly sulfated disaccharides and appears to be a rare event (5, 7, 12). However, the 3-O sulfation has been shown to be critical for the anticoagulant activity of HS/heparin (12). Over 85% of glucosamines are N-sulfated in heparin, whereas only a fraction (usually 40-50%) of the total

potential N-sulfation reactions occur in HS (10). Consequently, the major structural difference between HS and heparin result then from these sulfation occurring during the post-polymerization modification process. Heparin is a polymer more fully and evenly sulphated than HS (10). In contrast, the polysaccharide chains of HS have a distinct domain structure, in which domains of low sulfation separate highly modified sulfated domains, called S-domains, that are enriched in both sulfate groups and iduronic acid (7,11). The sugar ring of the iduronic acid has been shown to be particularly flexible, and then to allow a substantial conformational change required for physical interactions with proteins (13,14). CS and DS are both modified by O-sulfation and also by the epimerisation of glucuronic acid to iduronic acid in the case of DS (5,6). As observed from years with heparin/HS, an increasing number of proteins have been recently shown to interact with DS which underlines the relative importance of the flexible iduronic acid sugar ring in mediating interactions between GAG and proteins (7,14,15).

## A large diversity of the core protein

The complexity of PG (structure, variety of shape and size) has led to a broad classification (Table 2). Up to now, these macromolecules have been gathered together in subfamilies according to their localisation into intracellular PG, cell surface PG, circulating PG and finally extracellular matrix PG (PG secreted into the pericellular matrix). Syndecan-1 is for instance the prototype of the cell surface PG family. All the four syndecans (called syndecan-1 to syndecan-4) possess a cleavable signal peptide, an extracellular N-terminal domain containing GAG attachment sites, a single transmembrane domain and a short cytoplasmic domain (16). They may carry two-to-three HS chains at the main glycanation site near the N-terminus although they can also contain CS chains at a minor glycanation site near to the plasma membrane as observed in syndecan-1 (16). The extracellular domain of the syndecans is called the ectodomain and can be shed intact by proteolytic cleavage (16, 17). The HS chains of the freely circulating ectodomain retain the protein-binding properties of their cell surface precursors (16-20). The glypicans are another type of cell surface HSPG (16, 21, 22). Glypicans are another well-studied subfamily of cell surface proteoglycans: they have the particularity to be attached to the cell surface by a glycosylphosphatidylinositol lipid anchor (21, 22). They then do not possess a transmembrane and cytoplasmic domains and the glycanation sites in glypicans are all in close proximity to the plasma membrane (21, 22). Other cell surface proteins have been reported to be glycanated occasionally with GAG chains. For example, the betaglycan also called the type III transforming growth factor  $\beta$  (TGF- $\beta$ ) receptor, is a proteoglycan which can bind TGF $\beta$  by its core protein as well as

the HS chains, and presents TGF $\beta$  to the type II, serine/threonine kinase receptors (which are signalling receptors) (23, 24). The CD44 receptor, a ubiquitous cell surface adhesion molecule involved in cell-cell and cell-matrix interactions, may also carry GAG chains occasionally (25-29). The standard CD44 (also called CD44s) is the smallest isoform (85-95 kDa) found in the body (27, 29). However, the different utilization of 10 variable exons (CD44v1 to CD44v10), as well as the various possibilities of N-glycosylation, O-glycosylation and glycanation by HS and CS generate multiple isoforms of varying molecular weights (85-230 kDa). The splice variant called CD44E (also called CD44v3 as it contains the v3 exon) carries HS chains (27-29). CD44s and the high molecular mass proteoglycan-like variants have been identified in lymphocytes and parenchymal cells (30, 31) and were shown to be involved together with the non sulfated HA (their ligand), in several inflammatory diseases and in cancer (25-27).

Matrix proteoglycans represent a second major family of PG and are defined as PG secreted into the pericellular matrix (Table 2). The perlecan, present in most basement membranes is a large PG (400 kDa) carrying three GAG chains, which can be HS or CS nature, though HS is largely favored (16, 32). Versican is then the largest member (as a core protein) of the so-called hyalactan or lectican family, which define the matrix PG interacting with HA and lectins (32). Structurally, versican possesses 3 characteristic domains, with domain I being involved in the binding to HA (32, 33), domain II containing up to 30 attachment sites for CS and at the C terminus, the domain III that binds specifically lectin (32). The high concentration of CS side chains in versican and the formation of large aggregates with HA increase the water content

Table 2. Classification of the main PG according to their localization

Cell surface PG	Name	Type of GAG
Matrix PG	Syndecans	Mainly HS and CS
	Glypicans	Mainly HS
	Betaglycan	Mainly HS
	CD44 isoforms	HS or CS
	Versican	Mainly CS
	Aggrecan	CS and KS
	Perlecan	Mainly HS
	Decorin	DS
	Biglycan	DS or CS
	Lumican	KS
Fibronectin	KS	
Intracellular PG	Serpinin	HS or CS
Circulating PG	Bilamin	CS
	Endocan	Single chain of DS
	PGI00 or MCSF1	Single chain of CS

of the extracellular matrix, notably during the cell remodelling process occurring after injury and inflammation. Enhanced deposits of versican have been observed to proliferating tissues, in particular in the HA-rich areas that are associated with inflammation (34-38). Aggrecan, another particular matrix PG containing both CS and KS chains (32), forms various and large multimolecular complexes with hyaluronan (32) and is therefore also called "aggregating PG" (32). Aggrecan is the main proteoglycan found in the cartilaginous tissues and seems somewhat responsible for the mechanical properties of the cartilage (32, 39). Decorin is then the prototype of the third family of PG found in the matrix, called the small leucine-rich proteoglycans (SLRP) (16, 32, 40, 41). These SLRP are also called "small DSPG" or "non-aggregating PG" (compared to aggrecan or versican), because they contained CS or DS and did not interact with HA (32, 40). They are also characterized by their central domain of containing several leucine-rich repeats and presented small cysteine clusters at each side (40). Decorin contains a single GAG chain of DS type (16). This extracellular small DSPG is involved in formation of collagen fibrils mainly due to the interaction with the core protein since the DS-depleted core protein still possesses binding capacity to collagen (40). Decorin has also been shown to be involved in angiogenesis *in vivo*, especially in that associated with inflammation (41). Biglycan is another member of this family of SLRP, carrying two DS or CS chains (16, 40). The core protein of biglycan interacts with several different matrix proteins, including fibronectin and type I and VI collagens (40,42). Interestingly, biglycan as well as perlecan have been reported to be the possible substrates for lipoprotein entrapment in atherosclerotic lesions (43). Lumican is also another SLRP mainly present in the cornea (4, 16, 40-45) which carries characteristically KS chains (4,40). We know now that lumican is closely involved in the regulation of the collagen fibrillogenesis (44), a process critical to corneal transparency (45). Interestingly, mice lacking lumican exhibit an age-dependent corneal opacity and present progressively a high proportion of abnormally thick collagen fibers (45). Besides, lumican is also widely present as a non- or low-sulfated glycoprotein in connective tissues of many other organ systems, e.g. skeleton, heart, kidney and lung (40,45). Lumican that exhibit a low sulfate content (undersulfated KS or low glycanation content) is present in pathologic corneas and may act to localize inflammatory cells such as the recruited macrophages into regions of inflammation (46). As a last example, fibromodulin is another SLRP whose core protein has attachment sites for KS chains (40). An altered fibromodulin expression has been shown to affect the organization of collagen in pulmonary fibrosis (37, 42, 47). As shown for decorin, fibromodulin can bind to type I and type II collagens and consequently modulates the structure of fibrils during the process of the fibrillogenesis (44, 47).

Intracellular PG represent another family of PG (7, 16, 48). Serglycin is for instance a small PG presenting Ser-Gly dipeptide repeats, which is abundant in the secretory granules of mast cells, basophils and natural killer cells (48). This particular PG carries mainly two types of carbohydrate chains, HS/heparin or CS (48). In the secretory granules, serglycin helps stabilizing some of the main mediators of inflammation including cytokines and proteases and transporting them to target sites (48). Finally, PG have been found freely circulating in the blood and they are then called "circulating proteoglycans" (Table 2). Their presence in the blood flow can enhance their physiological role as co-receptor of different protein partners largely described in the literature (chemokines, growth factors, enzymes, coagulation protein or cytokines). Bikunin is thus a small chondroitin sulfate proteoglycan that occurs in blood as the light chain of inter-alpha-trypsin inhibitor (ITI) family members, that are synthesized in hepatocytes and release into circulation (49, 50). The relatively short CS chain of bikunin shows a characteristic pattern of sulfation in both the linkage region and the chondroitin sulfate backbone. The GAG chain may be switched from CS to hyaluronan by a transglycosylation process to form the Serum-derived Hyaluronan-Associated Protein (SHAP)-hyaluronan complex (49, 50). Endocan is another soluble proteoglycan of 50 kDa, constituted of a mature polypeptide of 165 aminoacids and a single DS chain (15, 51-53). Endocan was detected in the human serum and plasma of healthy subjects (52). Inflammatory cytokines such as TNF- $\alpha$  increased the synthesis and the secretion of endocan from endothelial cells acting both at the mRNA and the protein levels (15, 51-53). The DS chain of endocan has been reported to interact to growth factors such as the angiogenic factor called HGF/SF (15). Among circulating proteoglycans, the presence of the ectodomains of syndecan-1 and syndecan-4 have been reported in blood, which seems to have a role in complexing active molecules such as growth factors, enzyme or cytokines (55). Furthermore, PG-100, also called CSF-1 or MCSF (macrophage colony-stimulating factor-1) is a growth factor which can be present as a circulating proteoglycan form in blood, carrying one single CS chain (56, 57).

## Heparins : An example of a therapeutic use of GAG

Being implicated in the regulation of cytokine, growth factor and enzymes activities (1, 7, 32, 58-69), PG and GAG are clearly involved in cell fate. Changes in GAG fine structure have been observed during development, cell differentiation or pathological processes influence their ligand binding abilities and subsequently the physiological activities of the ligand (32, 58, 61, 62, 66). Moreover, cells undergoing pathogenic processes such as for instance tumorigenic processes, often modify GAG sulfation, leading to altered

adhesion properties and different abilities to interact with growth factors. Furthermore, all these interactions of regulatory proteins (eg growth factors, chemokines, enzymes) with GAG described above, provide much of the focus for GAG-based therapeutic targets. For many years now, heparins have been used in clinics as the standard treatment for venous thromboembolism (1, 7, 70-77). Historically, the anticoagulant activity of heparin was first described in 1916 by Jay MacLean at the John Hopkins University (USA), which found that an extract of dog liver prolonged in a tube, the time required for plasma to clot (71). Structurally, a characteristic heparin pentasaccharide termed DEFGH (with a crucial 3-O-sulfate group in unit F) was progressively defined and was shown to be necessary for binding to antithrombin, an interaction required for the anticoagulant activity of heparin. For clinical use, heparins are natural products, extracted mostly from pig intestine or from beef lung, mainly by biochemical procedures. Two class of heparins are found: the class of so called "unfractionated heparin" or UFH corresponding to the mixture of GAG chains obtained (mainly of 200 to 300 saccharide units), and the class of the "low molecular weight heparin" or LMWH, corresponding to low molecular weight fragments of UFH produced by controlled enzymatic or chemical depolymerization, which yields chains that are less than 18 saccharide units long (70, 72, 77). UFH and LMWH exerts their anticoagulant effects by activating the coagulation inhibitor called antithrombin, which in turn neutralize many of the serine proteases involved in the coagulation system, particularly thrombin and activated X (Xa). Around 50 different LMWH has been prepared and tested as anticoagulant drugs (ardeparin®, enoxaparin®, reviparin®, certoparin®, vasaflux®...). Heparins exert also non anticoagulant actions and therefore have several drawbacks for more specific applications. Other natural GAG have also been used as anticoagulants. For instance, Danaparoid also called ORGARAN® is composed of dermatan and chondroitin sulphate chains extracted from pig intestine tract (75). Like LMWH, danaparoid inhibits preferentially the factor Xa and not the factor IIa, and as its half life reach 25 hours, only a single daily intravenous injection is required (75). However, the need for more specific and homogeneous sequences rose of the complexity of the heparin and GAG structures from natural sources. Recently, Arixtra® (fondaparinux sodium), a compound completely issued from sugar chemistry, have been successfully launched in the treatment of venous thrombosis and pulmonary embolism (70-74). This compound is in fact the first of new class of anticoagulants that inhibit Factor Xa selectively (74). Moreover, this totally synthetic pentasaccharide presents a high degree of purity and homogeneity in contrast of all the LMWH on the market which are animal sourced. Theoretically eliminating the risk of pathogen contamination, and allowing the batch to batch consistency compared to the heparins, fondaparinux has therefore been a major advance for GAG-based therapy

strategies (74). Furthermore, manipulation of the sulfate groups in GAG has been a potential strategy to generate low anticoagulant molecules with specifically retained pharmacological properties and a great degree of biological specificity (70, 72, 76, 77). Other synthetic polysaccharides with different structure and sulphate positions are currently under clinical investigations and new drugs will be probably developed soon to pursue this major technical and therapeutic advance (72, 74-77). In a near future, GAG mimetics should have therefore more and more clinical applications as modulators of cytokine, growth factors or enzymes function in various diseases and pathologies.

## Acknowledgements

Stéphane Sarrazin is partly supported by a fellowship from the ANRT (Association Nationale de la Recherche Technique, France). We are grateful to Stéphane Lequien for the critical reading of the manuscript.

## References

1. Pihlm, P. 1984, *Biochem. J.*, 220, 597.
2. Weigel, P.H., Hascall, V.C., and Tammi, M. 1997, *J. Biol. Chem.*, 272, 13997.
3. Fraser, J.R., Laurent, T.C., and Laurent, U.B. 1997, *J. Intern. Med.*, 242, 27.
4. Funderburgh, J.L. 2000, *Glycobiology*, 10, 951.
5. Habuchi, O. 2000, *Biochim. Biophys. Acta*, 1474, 115.
6. Sugihara, K., and Kitagawa, H. 2000, *Curr. Opin. Struct. Biol.*, 10, 518.
7. Gallagher, J.T., and Lyon, M. 2000, *Proteoglycans: Structure, Biology and Molecular Interactions* Iozzo, and R.V. (Ed), Marcel Dekker, New York, 27.
8. Pkasz, D.S., Eriksson, I., and Kjellen, L. 2000, *Biochemistry*, 39, 4552.
9. Ringvall, M., Ledin, J., Holmborn, K., van Kuppevelt, T., Ellin, F., Eriksson, I., Olofsson, A.M., Kjellen, L., and Forsberg, E. 2000, *J. Biol. Chem.*, 275, 25926.
10. Lyon, M., and Gallagher, J.T. 1998, *Matrix. Biol.*, 17, 485.
11. Gallagher, J.T. 2001, *J. Clin. Invest.*, 108, 357.
12. Thunberg, L., Backstrom, G., and Lindahl, U. 1982, *Carbohydr. Res.*, 100, 393.
13. Casu, B., Petitou, M., Provasoli, M., and Sinay, P. 1988, *Trends Biochem. Sci.*, 13, 221.
14. Mulloy, B., and Forster, M.J. 2000, *Glycobiology*, 10, 1147.
15. Bechard, D., Gentina, T., Delebede, M., Scherpereel, A., Lyon, M., Aumerier, M., Vazeux, R., Rechet, C., Degand, P., Jude, B., Janin, A., Fering, D.G., Tomel, A.B., and Lassalle, P. 2001, *J. Biol. Chem.*, 276, 48341.
16. Bernfield, M., Gotte, M., Park, P.W., Reizes, O., Fitzgerald, M.L., Lincecum, J., and Zako, M. 1999, *Annu. Rev. Biochem.*, 68, 729.
17. Kato, M., Wang, H., Kainulainen, V., Fitzgerald, M.L., Ledbetter, S., Ornitz, D.M., and Bernfield, M. 1998, *Nat. Med.*, 4, 691.
18. Kainulainen, V., Wang, H., Schick, C., and Bernfield, M. 1998, *J. Biol. Chem.*, 273, 11563.

19. Cizmeci-Smith, G., Langan, E., Youkey, J., Showalter, L.J., and Carey, D.J. 1997. Arteriolect. *Thromb. Vasc. Biol.*, 17, 172.
20. Gallo, R.L., Ono, M., Powsic, T., Page, C., Eriksson, E., Klagesbrun, M., and Bernfield, M. 1994. *Proc. Natl. Acad. Sci. U.S.A.*, 91, 11035.
21. Filmus, J. 2001. *Glycobiology*, 11, 19.
22. Filmus, J., and Selbeck, S.B. 2001. *J. Clin. Invest.*, 108, 497.
23. Andres, J.L., Stanley, K., Cheifetz, S., and Massague, J. 1989. *J. Cell. Biol.*, 109, 3137.
24. Lopez-Casillas, F., Payne, H.M., Andres, J.L., and Massague, J. 1994. *J. Cell Biol.*, 124, 557.
25. Pure, E., and Cuff, C.A. 2001. *Trends Mol. Med.*, 7, 213.
26. Goodison, S., Urquidí, V., and Tarin, D. 1999. *Mol. Pathol.*, 52, 189.
27. Herrlich, P., Morrison, H., Sleeman, J., Orsian-Rousseau, V., König, H., Wegremers, S., and Ponta, H. 2000. *Ann. N.Y. Acad. Sci.*, 910, 106.
28. Brown, T.A., Bouchard, T., St John, T., Weyerer, E., and Carter, W.G. 1991. *J. Cell Biol.*, 113, 207.
29. Kalish, E.D., Iida, N., Moffat, F.L., and Bourguignon, L.Y. 1999. *Front. Biosci.*, 4, 1.
30. Jackson, D.G., Bell, J.I., Dickinson, R., Timans, J., Shields, J., and Whittle, N. 1995. *J. Cell Biol.*, 128, 673.
31. Jones, M., Tussey, L., Athanasou, N., and Jackson, D.G. 2000. *J. Biol. Chem.*, 275, 7964.
32. Izzo, R.V. 1998. *Annu. Rev. Biochem.*, 67, 609.
33. LeBaron, R.G., Zimmermann, D.R., and Ruoslahti, E. 1992. *J. Biol. Chem.*, 267, 10003.
34. Nara, Y., Kato, Y., Torii, Y., Nakagaki, S., Goto, S., Isobe, H., Nakashima, N., and Takeuchi, J. 1997. *Histochem. J.*, 29, 21.
35. Brown, L.F., Guidi, A.J., Schnitt, S.J., Van De Water, L., Iruela-Arispe, M.L., Yeo, T.K., Tognazzi, K., and Dvorak, H.F. 1999. *Clin. Cancer Res.*, 5, 1041.
36. Johnson, P.R. 2001. *Clin. Exp. Pharmacol. Physiol.*, 28, 233.
37. Venkatesan, N., Ebihara, T., Roughley, P.J., and Ludwig, M.S. 2000. *Am. J. Respir. Crit. Care Med.*, 161, 2066.
38. Bensoudon, E.S., Burke, A.K., Hogg, J.C., and Roberts, C.R. 1997. *Eur. Respir. J.*, 10, 2731.
39. Cameron, B., Flannery, C.R., Hughes, C.E., and Little, C.B. 2000. *Matrix Biol.*, 19, 333.
40. Hoeking, A.M., Shinomura, T., and McQuillan, D.J. 1998. *Matrix Biol.*, 17, 1.
41. Nelimarkka, L., Salminen, H., Kuopio, T., Nikkari, S., Ekfors, T., Laine, J., Pelliniemi, L., and Jarvelainen, H. 2001. *Am. J. Pathol.*, 158, 345.
42. Hildebrand, A., Romaris, M., Rasmussen, L.M., Heinegard, D., Twardzik, D.R., Border, W.A., and Ruoslahti, E. 1994. *Biochem. J.*, 302, 527.
43. Kunjathoor, V.V., Chiu, D.S., O'Brien, K.D., and LeBoeuf, R.C. 2002. *Arterioscler. Thromb. Vasc. Biol.*, 22, 462.
44. Ezura, Y., Chakravarti, S., Oldberg, A., Chervoneva, I., and Birk, D.E. 2000. *J. Cell Biol.*, 151, 779.
45. Chakravarti, S., Magnuson, T., Lass, J.H., Jepsen, K.J., LaMantia, C., and Carroll, H. 1998. *J. Cell Biol.*, 141, 1277.
46. Funderburgh, J.L., Mitschler, R.R., Funderburgh, M.L., Roth, M.R., Chapes, S.K., and Conrad, G.W. 1997. *Invest. Ophthalmol. Vis. Sci.*, 38, 1159.
47. Svensson, L., Azodi, A., Reinhold, F.P., Fassler, R., Heinegard, D., and Oldberg, A. 1999. *J. Biol. Chem.*, 274, 9636.
48. Schick, B.P., Gradowski, J.F., and San Antonio, J.D. 2001. *Blood*, 97, 449.
49. Zhuo, L., Sahasri, A., and Kimata, K. 2002. *Glycoconj. J.*, 19, 241.
50. Fries, E., and Kaczmarczyk, A. 2003. *Acta Biochim. Pol.*, 50, 735.
51. Bechar, D., Scherperel, A., Hammad, H., Gentina, T., Tsiopoulou, A., Aumerker, M., Pestel, J., Dessaint, J.P., Tonnell, A.B., and Lassalle, P. 2001. *J. Immunol.*, 167, 3099.
52. Bechar, D., Meignin, V., Scherperel, A., Oudin, S., Kervouza, G., Bertheau, P., Janin, A., Tonnell, A., and Lassalle, P. 2000. *J. Vasc. Res.*, 37, 417.
53. Lassalle, P., Molet, S., Janin, A., Heyden, J.V., Tavemier, J., Fleis, W., Devos, R., and Tonnell, A.B. 1996. *J. Biol. Chem.*, 271, 20458.
54. Scherperel, A., Gentina, T., Grigoriu, B., Senechal, S., Janin, A., Tsiopoulou, A., Penat, F., Bechar, D., Tonnell, A.B., and Lassalle, P. 2003. *Cancer Res.*, 63, 6084.
55. Marshall, L.J., Ramdin, L.S., Brooks, T., PC, D.P., and Shure, J.K. 2003. *J. Immunol.*, 171, 2057.
56. Ohtsuki, T., Suzu, S., Hatake, K., Nagata, N., Miura, Y., and Motoyoshi, K. 1993. *Biochem. Biophys. Res. Commun.*, 190, 215.
57. Parsenheimer, A., Schwanz, K., Wrocklage, C., Kolsch, E., and Kresse, H. 1995. *J. Immunol.*, 155, 5557.
58. Delebedde, M., Lyon, M., Sergeant, N., Rahmoune, H., and Fernig, D. G. 2001. *J. Mammary Gland Biol. Neoplasia*, 6, 273.
59. Sergeant, N., Lyon, M., Rudland, P. S., Fernig, D. G., and Delebedde, M. 2000. *J. Biol. Chem.*, 275, 17094.
60. Delebedde, M., Sergeant, N., Lyon, M., Rudland, P. S., and Fernig, D. G. 2001. *Eur. J. Biochem.*, 269, 4423.
61. Vlodiavsky, I., and Friedmann, Y. 2001. *J. Clin. Invest.*, 108, 341.
62. Delebedde, M., Deudon, E., Botly, B., and Hondermarck, H. 1996. *Exp. Cell Res.*, 229, 398.
63. Rapraeger, A. C., Krufka, A., and Ohlwin, B. B. 1991. *Science*, 252, 1705.
64. Delebedde, M., Seve, M., Sergeant, N., Lyon, M., Rudland, P. S., and Fernig, D. G. 2000. *J. Biol. Chem.*, 275, 35905.
65. Guimond, S., Maccarana, M., Ohlwin, B. B., Lindahl, U., and Rapraeger, A. C. 1993. *J. Biol. Chem.*, 268, 23906.
66. Lorat-Jacob, H., Grosdidier, A., and Imberty, A. 2002. *Proc. Natl. Acad. Sci. U.S.A.*, 99, 1229.
67. Delebedde, M. 2002. *Recent Res. Devel. Biol. Chem.*, 1, 133.
68. Delebedde, M., Lorat-Jacob, H., Gallagher, J., Bonaffé, D., Adam, E., Querolle, O., Lequien, S., Degove, S., Lassalle, P., and Bechar, D. 2005. *Current Med. Chem.*, 2, 345.
69. Delebedde, M., Lyon, M., Gallagher, J.T., Rudland, P.S., and Fernig, D.G. 2002. *Biochem. J.*, 366, 235.
70. Bauer, K. A. 2003. *Chest* 124, 364.
71. McLean, J., 1959. *Circulation*, 19, 75.

72. Petitou, M., Driquez, P. A., Duchaussoy, P., Hérault, J. P., Lormeau, J. C., and Herbert, J. M., 1999, *Bioorg. Med. Chem. Lett.*, 9, 1161.
73. Petitou, M., Duchaussoy, P., Driqez, P. A., Hérault, J. P., Lormeau, J. C., and Herbert, J. M., 1999, *Bioorg. Med. Chem. Lett.*, 9, 1155.
74. Petitou, M., Hérault, J. P., Bernat, A., Driqez, P. A., Duchaussoy, P., Lormeau, J. C., and Herbert, J. M., 1999, *Nature*, 398, 417.
75. Acostamadeo, J.M., Lyeer, U.G., and Owen, J., 2000, *Expert Opin. Pharmacoth* 1, 803.
76. Dilhas, A., and Bonnafie, D., 2003, *Carbohydr. Res.*, 338, 68
77. Noti, C., and Seeberger, P.H., 2005, *Chem. Biol.*, 12, 731.

Research Signpost  
371 661 (2), Fort P.O., Trivandrum-695 023, Kerala, India



New Developments in Therapeutic Glycomics, 2006: ISBN: 81-7736-266-6  
Editors: Maryse Delehedde and Hugues Lortat-Jacob

## 2 Proteoglycans and glycosaminoglycans : Key players in inflammation

Philippe Lassoik<sup>1</sup>, Estelle Adam<sup>1</sup>, Stéphanie Sarrazin<sup>1</sup>  
Florence Depontieu<sup>1</sup>, Vincent Motte<sup>2</sup>, Christelle Villedieu<sup>1</sup>  
Stéphanie Lequien<sup>1</sup>, Bogdan Grigoriu<sup>1</sup>, Nathalie de Freitas-Carreras<sup>1</sup>  
Christophe Vangouille<sup>1</sup> and Maryse Delehedde<sup>1</sup>  
Inserm, U774, Lille, F-59019, France; <sup>1</sup>Endotis Pharma, Parc Eurasané, Loos  
59120 France; <sup>2</sup>Université de Valenciennes, Cambrai, F-59600, France  
<sup>1</sup>National Institute of Health, Bethesda, USA

### Abstract

*Proteoglycans (PG) are complex macromolecules which consist structurally of a core protein and associated glycosaminoglycan (GAG) chains. There are several GAG families: heparan sulfate, chondroitin sulfate, dermatan sulfate, keratan sulfate and hyaluronic acid. Among GAG, hyaluronic acid is unique in that it is not sulfated and not attached covalently to a core protein. PG and GAG in general*

*are found in various tissues and cells, in the extracellular matrix, at the cell membrane, being intracellular or also freely circulating in blood. By their variety of localisations and structures, and consequently the multiplicity of their protein binding partners (growth factors, chemokines, cytokines, enzymes), PG and GAG are strongly implicated in the different stages of inflammation from the recruitment of inflammatory cells to the release of mediators of inflammation by infiltrating leukocytes and the turnover of extracellular matrix. Therefore, PG and GAG are involved in the regulation of the inflammatory microenvironment, which has to change continuously and dynamically during the progression of the inflammatory response. The main changes include the modulation of the activity of their GAG-binding partners: cytokines, growth factors, proteases and protease inhibitors. GAG mimetics will have in a very near future clinical applications as the specific interactions of these regulatory proteins with GAG provide much of the focus for GAG-based therapy.*

### List of abbreviations

ADAMTS: a disintegrin and metalloproteinase with thrombospondin motifs ; AT-III : antithrombin III ; CS : chondroitin sulfate ; CSPG : chondroitin sulfate proteoglycan ; CyPA, CyPB : cyclophilins A, B ; DS : dermatan sulfate ; DSPG : dermatan sulfate proteoglycan ; ECM : extracellular matrix ; EGF : epithelial growth factor ; FGF-2 : fibroblast growth factor 2 ; FGFR : FGF receptor ; GAG : glycosaminoglycan ; HA : hyaluronic acid ; HGF/SF : hepatocyte growth factor / scatter factor ; HS : heparan sulfate ; HSPG : heparan sulfate proteoglycan ; IFN- $\gamma$  : interferon  $\gamma$  ; IL-1 $\beta$ , IL-6, IL-8, IL-10 : interleukin-1  $\beta$ , -6, -8, -10 ; KS : keratan sulfate ; KSPG : keratan sulfate proteoglycan ; M-CSF : macrophage colony stimulating factor ; MIP-1 $\alpha$ , MIP-1 $\beta$  : macrophage inflammatory protein-1 $\alpha$ , -1 $\beta$  ; MMP : metalloproteinases ; NDST : N-deacetylase/N-sulfotransferase ; PF-4 : platelet factor ; PG : proteoglycan ; RANTES : Regulated upon Activation, Normal T Expressed and presumably Secreted ; SLRP : small leucine-rich proteoglycan ; TGF- $\beta$  : transforming growth factor  $\beta$  ; TNF- $\alpha$  : tumor necrosis factor  $\alpha$ .

### I. Introduction

Proteoglycans (PG) are ubiquitously present in the organism: from the cell surface to the extracellular matrix, as well as circulating in blood. Structurally, PG consists of a specific core protein and one or more associated glycosaminoglycan (GAG) chains. Structurally, GAGs are linear polysaccharides that consist of repeating disaccharide units composed of an uronic acid and an N-substituted hexosamine, glucosamine or galactosamine. The distinct types of GAG are described as follows: chondroitin sulfate (CS),

Correspondence/Reprint request: Dr. Maryse Delehedde, Endotis Pharma, Parc Eurasané, 70 rue du Dr Yersin 59120 Loos, France. E-mail: maryse.delehedde@endotis.com

dermatan sulfate (DS), heparan sulfate (HS), heparin, keratan sulfate (KS) and hyaluronic acid (HA). The GAG hyaluronic acid, also called hyaluronan, is not normally covalently attached to a core protein and is also unique in that it is not sulfated (1-3).

This large variety of structures confers to them a large variety of functions as homeostasis maintaining or regulation of a large number of pathophysiological processes. A large number of studies have shown that PG are involved in the different stages of inflammation such as the recruitment of inflammatory cells, the release and diffusion of mediators of inflammation by infiltrating leukocytes and inflammatory cells and the turnover of extracellular matrix. The ability of these various mediators of inflammation to bind to and be immobilized by PG on either the surface of cells or within the extracellular matrix may accentuate their activity by promoting the accumulation of active molecules at high concentrations in the appropriate location to encounter their target cells. Moreover, PG are involved in promoting the conformation-dependent association or the polymerization of chemokines/cytokines or growth factors and their cell-surface signalling receptors. PG can then facilitate the assembly of the appropriate molecular complex to initiate signal transduction. The structural basis of the binding of fibroblast growth factor (FGF-2) to HS is up to now the best-defined FGF-glycosaminoglycan interaction (3-5). To stimulate cell proliferation, FGF-2 must interact with a dual receptor system, consisting of a low affinity HS co-receptor and a high affinity receptor tyrosine kinase FGFR (3,5-7). Similar dependency on HSPG has also been reported for a growing number of growth factors such as hepatocyte growth factor/scatter factor (HGF/SF) and vascular endothelial growth factor (VEGF). HSPG have also an effect upon chemokines/cytokines such as stromal derived factor (SDF-1), RANTES or interferon gamma (IFN- $\gamma$ ) (8-13). PG have also been shown to have an important role of protecting mediators of inflammation such as cytokines from degradation by extracellular proteases which are particularly abundant at the sites of cell invasion and inflammation (14). The function of the mediators of inflammation can also be amplified or efficiently modulated by their interaction with the PG of the matrix and/or the cell surface. PG thus represent a powerful central regulatory element in the pericellular environment of tissues that influence the variety of events involved in the inflammatory response. Furthermore a proteoglycan named endocan, since it has been originally described as a product of endothelial cells, appears to be strongly implicated in inflammation, but in this case, the PG is released into the blood circulation rather than being fixed in the cell periphery (15-18). Many of the PG-interacting proteins require distinct sequences of oligosaccharides within the GAG chain for binding. Changes in GAG fine structure during inflammatory and pathological processes will influence their ligand binding abilities and subsequently the physiological

activities of the ligand during the inflammatory response. The interactions of the regulatory proteins with specific oligosaccharide sequences in GAG provides much of the focus for GAG-based therapeutic targets (19-23). Since PG through their GAG chains bind and modulate the activity of cytokines that regulate the inflammatory response, several anti-inflammatory strategies targeting GAG have been attempted. They included the prevention of inflammatory cell recruitment and cell adhesion, the prevention of growth factor/cytokine/chemokine activation or the inhibition of the activity of enzymes such as heparanases. Therefore, development of GAG mimetics as antagonists or agonists of particular activities may have clinical applications (8,10,20).

## II. PG and GAG in the recruitment of inflammatory cells

PG and GAG have been shown to be involved in the different stages of inflammation. The recruitment of inflammatory cells is the earliest event observed in an inflammatory response. The recruitment of leukocytes into tissue is regulated by a sequence of interactions between the circulating leukocytes and the endothelium, and is mediated by adhesion receptors, which include integrins, selectins and cell surface PG (24-26). Some findings suggest a direct role of HSPG and heparin in the adhesion of leukocytes to vascular endothelium (27-31). Leukocyte adhesion is partially mediated by binding of Mac-1 (32) or L-selectin (24) to HSPG from endothelial cells. Moreover, soluble heparin can inhibit Mac-1-dependent leukocyte adhesion on endothelial cells and can attenuate inflammatory and proliferative stimuli associated with the emigration of leukocytes (24). There are various examples of how HS/heparin are able to inhibit the recruitment of leukocytes to various tissues such as lung, skin and gut (27-30,33-35). Non-anticoagulant derivatives of heparin have been shown to mimic heparin and prevent leukocytes recruitment (27,28). One of the crucial role of PG in the recruitment of leukocytes is played by HSPG on endothelial cells, which immobilize and present chemokines such as macrophage inflammatory protein-1 $\beta$  (MIP-1 $\beta$ ), interleukin-8 (IL-8) and platelet factor 4 (PF-4). These chemokines induce the activation of leukocyte integrins and mediate stable adhesion and transmigration of the cells across the vessel wall (36,37).

The different cellular partners involved in the early steps of inflammation (endothelial cells, lymphocytes, macrophages) have been shown to differentially express and synthesize cell surface PG as well as intracellular PG. Endothelial cells are the major source of HSPG (2,38-40). Their HSPG has been implicated in T-cell adhesion *in vitro*, where it was shown to capture the chemokine MIP-1 $\beta$  and activate T-cell integrins (37). Peripher-

blood monocytes seem to not express any PG on their cell surfaces. However, during the recruitment process, blood-derived monocytes undergo progressive differentiation to tissue macrophages characterized by changes in cell surface antigen expression, increased phagocytic activity, and most notably the acquisition of the capacity to synthesize an array of different cytokines and growth factors (26,41). The expression of PG, and more particularly of HSPG on these activated macrophages is a dynamic and highly regulated process that depends on the state of cellular differentiation and activation (42,43).

### III. PG and mediators of inflammation

Multiple factors orchestrating the inflammatory response include cytokines, growth factors, cyclophilins, proteases, antiproteases and endoglycosidases, which are derived from cells coming from or invading the inflammation site as well as from plasma (25). Specific interactions with PG confer to these molecules properties which are crucial in the inflammation response and inflammatory diseases. The binding to PG protects cytokines and growth factors from degradation by extracellular proteases, which are particularly abundant at the sites of inflammation. For example, endothelial HSPG as well as soluble heparin will protect IL-8 from a rapid degradation by elastase which is released by activated neutrophils (44). Moreover, direct feedback loops, such as the regulation of the synthesis of PG and the induction of enzymes involved in the degradation of matrix and cell surface PG by a number of pro-inflammatory molecules provide the means to fine-tune the system.

Cytokines, growth factors and chemokines are soluble proteins, which are involved in the regulation of the main processes regulated during inflammation: survival, proliferation, differentiation and cell migration. Interestingly, these molecules interact either with the GAG chains (HS, CS, DS or KS) or with the core protein of the PG (2-4,10-13,45). Basement membrane PG play a critical role in modulating the activity of growth factors and cytokines by being a reservoir for such molecules, an effective means to limit their diffusion and their activities (38,46,47). For example, HS chains have been shown to sequester IL-8 and prevent the diffusion of the secreted chemokine from an inflammatory site, as well as its clearance from the blood stream through nonspecific chemokine receptors on erythrocytes (44). Increased shedding of IL-8/heparan sulfate/syndecan-1 complexes is accompanied by inhibition of neutrophil transendothelial migration (48). The macrophage colony-stimulating factor-1 (M-CSF) which can be present as a proteoglycan form (49), can be immobilized in tissues via its GAG chain as a means of storage in the extracellular matrix (49,50). This "immobilising"

function of the GAG chain may lead to the attenuation of ongoing tissue damage induced by M-CSF. The action of other pro-inflammatory mediators released from inflammatory cells may also be dampened by binding to HS in the ECM. The release of these mediators can be then triggered, when required, by the degradation of PG itself by either glycosidases or proteases. Endoglycosidases such as heparanases, which are expressed by platelets and inflammatory cells (51,52), for example release an active complex of FGF-2 and HS fragment from the basement membrane and the extracellular matrix HSPG (53,54). Exogenous leukocyte elastase, as well as endogenous vascular elastase, also release FGF-2 stored on HSPG (55,56). The released FGF-2 is a potent mitogen for the surrounding smooth muscle cells (55) and downregulates the transcription activity of the elastin gene itself, which can by doing so influence the repair mechanisms during lung injury (56).

PG are also involved in the presentation of released mediators of inflammation such as cytokines and growth factors to their cognate receptors. For example, HS can promote IL-8 mediated attraction of neutrophils by trapping the secreted chemokine on the luminal surface of the endothelium at the sites of inflammation (44). This interaction with GAG results in the effective presentation of chemokines to circulating leukocytes, thereby promoting their adhesion and their migration through the vessel and the extracellular matrix (37). The mechanism of presentation of regulatory proteins could involve PG from the cell surface, extracellular matrix PG, circulating PG or also intracellular PG. Following activation, cytotoxic T lymphocytes will secrete significant amounts of chemokines (such as RANTES, MIP-1 $\alpha$  and MIP-1 $\beta$ ) together with granular PG (57). Intracellular packaging of chemokines bound to GAG has also been described for PF-4 and platelet basic protein, which are stored bound to PG and are released following platelet activation. The storage of chemokines in specialized granules bound to PG could be a more general phenomenon for cellular effectors of inflammation. Furthermore, the dependency on HSPG for the stimulation of cell growth by FGF-2 has been largely described (3,4,6,7,58). Macrophages expressing syndecan-2 bind FGF-2 and consequently promote the proliferation of neighbouring cells that possess the FGF receptor (FGFR) tyrosine kinase but lack HSPG (59). Dependency on HSPG has been then described for the regulation of various growth factors and cytokines to their respective tyrosine kinase-linked receptors, consistent with a general role for HSPG as components of a dual receptor system for heparin-binding growth factors and cytokines (3-7,58,60). Hepatocyte growth factor/scatter factor (HGF/SF) regulates integrin-mediated lymphocyte adhesion and migration, by binding to its signaling receptor, MET, and the HS chains of CD44v3 (61). Among cytokines, IL-10 requires on the cell surface of the monocytes, the presence of HSPG, which may in turn facilitate interaction with high-affinity receptor on

these cells and/or participate in the signaling event (62). Other GAG such as CSPG and DSPG bind also chemokines and play important roles in the regulation of chemokine functions during inflammatory response. CSPG have been reported to promote the dimerization of chemokines and facilitate their presentation to specific receptors (63-65). For instance, CSPG on human neutrophils specifically bind platelet factor 4 and are involved in the activation of neutrophils (63,64).

Enzymes play also a key role in rearranging the microenvironment surrounding the inflammation site. For example, incoming lymphocytes will secrete extracellular matrix degrading enzymes, such as metalloproteinases and endoglycosidases, which facilitate the crossing of basement membrane by the cells (14). Emigrating leukocytes and activated platelets release proteolytic enzymes into the injury site, which include neutrophil derived elastase, cathepsin G and thrombin (66-68). These enzymes are involved in the inflammatory response by degrading extracellular proteins, releasing cytokines, chemokines and growth factors from extracellular matrix, and, in conjunction with *de novo* synthesis of matrix components, remodeling the matrix (14). However, they can also destroy tissues when proteolysis is prolonged, inappropriate, or excessive. The balance of proteases and antiproteases at the site of injury can regulate the extent of the inflammatory response. Antiproteases, such as serine protease inhibitors (also called serpins), mostly derived from plasma, act by forming covalent complex with protease, which results in the inactivation of the protease. Elastases and cathepsin G has been shown to bind to heparin and HS (66,67,69). The soluble syndecan-1 ectodomain, which is found in wound fluids (70,71), can reduce the affinity of these proteases for their physiological inhibitors, the plasma-derived serpins  $\alpha$ 1-antitrypsin and  $\alpha$ 1-antichymotrypsin (72). Therefore, ectodomains of syndecan-1 represent a local control of the growth factor activities and proteolytic balance at the wound site, when shed from the cells in the site of inflammation (73). Moreover, soluble syndecan-4 is also (or as well as syndecan-1) found in the cells of the wounded tissue in case of experimental injuries (74,75).

Other proteases have been reported to degrade specific core protein of PG. Several metalloproteinases (MMP) have been shown to be secreted by various inflammatory cells including neutrophils and macrophages. Matrilysin, which is a matrix metalloproteinase expressed by macrophages, specifically degrades versican at the sites of potential rupture in atherosclerotic lesions (76). In rheumatoid arthritis, the degradation of aggrecan is driven by several proteases including specific members of the MMP-family and aggrecanases (77,78). Aggrecanase-1 (ADAMTS4) and aggrecanase-2 (ADAMTS11) are members of the ADAMTS (a disintegrin and metalloproteinase with thrombospondin motifs) family that cleave aggrecan at the Glu373-Ala374

bond, while MMP cleave aggrecan at the Asn341-Phe342 site. Moreover an imbalance between matrix metalloproteinases and their specific inhibitors accelerates the turnover of the cartilage proteoglycan and collagen turnover resulting in the appearance of breakdown products in the synovial fluids from patients with rheumatoid arthritis (78-81).

A balance between proteases and antiproteases such as thrombin and antithrombin III (AT-III) also play major roles in perpetuating inflammation notably through the activation of the coagulation system (68). The thrombin formed at a site of vascular damage activates numerous cells involved in the inflammatory responses including monocytes, blood T lymphocytes, platelets and mast cells located along the blood vessels (68). Antithrombin III (AT-III) exerts its anti-inflammatory effects by inhibiting thrombin (82,83). Heparin (synthesized by mast cells) and HSPG potentiate this inhibition (68). Heparin interacts with AT-III to form a complex that inhibits thrombin and factor Xa much more effectively than AT-III alone (82-85). Heparin binds to AT-III via a specific high affinity pentasaccharide sequence that represents only a minor portion of any heparin chain (82,85). AT-III also exhibits thrombin-independent effects on leukocytes and endothelial cells. In ischemia, AT-III prevents leukocyte recruitment and inhibits lymphocyte activation (86,87). AT-III was also demonstrated to deactivate migration of neutrophils towards the chemokine IL-8 (88). Thus AT-III mediated deactivation of IL-8 was dependent on the interaction between the heparin-binding site of the AT-III and syndecan-4 on neutrophils (89).

Heparanases are endoglycosidases, which degrade HS to small oligosaccharides (51). Heparanase activities are expressed by a large number of cells including platelets, tumor, and inflammatory cells and have been shown to play a major role in cell diapedesis and trafficking into tissues (90,91). This enzyme is also involved in remodeling basement membranes after injury or at inflammation sites (91-94). Molecular cloning and sequencing indicate that mammalian cells produce a single heparanase that undergoes endoproteolytic cleavage to produce an active heterodimer (52,95). Expression of these enzymes was found to be associated with the ability of activated T lymphocytes to penetrate blood vessel walls and then accumulate in target organs (90,96). Moreover, the biological and enzymatic activity of heparanase was found to be dependent on the local acidity associated with inflammation, thus restricting the degradative activity to the inflamed tissues (97). Neutrophil heparanase is then secreted into wound fluids and partially degrades the HS chain to release the S domains that can markedly stimulate FGF-2 mitogenicity. Significant concentrations of HS fragments, that resemble S-domains are usually found in acute wound fluids (73). The short HS oligosaccharides generated by either extracellular or intracellular heparanases may regulate the binding of ligands to cell surface or extracellular

HSPG, stabilize the ligand (98-100), facilitate the transport of the ligand to other sites of action (101,102) or dilute the cytokines by promoting their diffusion and so reduce their activity. Degradation of HSPG by heparanases may also be an important mechanism to prevent PG or native HS chains from associating with molecules that would be deleterious to the cell, such as the  $\beta$ -amyloid-HSPG complex deposited in senile plaques of Alzheimer's disease (103,104).

#### IV. Regulation of PG metabolism by the mediators of inflammation

Cytokines and growth factors involved in the inflammation process can modify the synthesis of PG and consequently regulate the cellular response to inflammation. They have been shown to regulate PG metabolism in different ways. Cytokines and chemokines can induce different PG according to the target cell type. For example, TNF- $\alpha$  induces the expression of syndecan-2 on activated macrophages, which allows the activation of FGF-2 (59). Peripheral blood monocytes show a striking induction of cell surface HSPG including CD44-HSPG forms in response to TNF- $\alpha$  released by cells on the site of inflammation (40,105). These newly synthesized HSPG can then bind cytokines and thus enable recruitment and activation of new cells. In contrast, TNF- $\alpha$  caused a dose-dependent decrease in mRNA levels for the core proteins of perlecan and biglycan in endothelial cells in culture. TNF- $\alpha$  also transcriptionally inhibit decorin gene expression in fibroblasts (106). One cytokine can differentially modulate the biosynthesis of different PG in response to inflammatory stimuli according to the tissue. For example, IL-1 $\beta$  caused a significant decrease in synthesis and deposition of versican in connective tissue during inflammation and wound healing (107). A pronounced decrease in the steady-state mRNA levels of versican has been observed after IL-1 $\beta$  treatment of fibroblasts (107). In contrast, IL-1 $\beta$  induced an up-regulation of perlecan mRNA in hippocampus but not in the striatum after a brain injury (108). The pro-inflammatory cytokine IL-1 $\beta$  was also found to induce macrophage syndecan-2 expression *in vitro* (59).

PG synthesis can be diversely regulated at the site of injury by a finely balanced release of active cytokines from the inflammatory cells. For example, TGF- $\beta$  mainly down-regulates decorin gene expression, whereas it induces an upregulation of biglycan (109,110). Released cytokines can also act synergistically to modulate the synthesis of PG during the inflammation response. TNF- $\alpha$  and TGF- $\beta$  together inhibit decorin gene expression. In contrast, TNF- $\alpha$  counteracts the effects of TGF- $\beta$  and IL-6 in controlling biglycan expression (110,111). In endothelial cells, TNF- $\alpha$  and IL-1 upregulates

the synthesis and the secretion of endocan whereas IFN- $\gamma$  counteracts the effects of TNF- $\alpha$  in the synthesis of this freely circulating DSPG (16,18).

The synthesis of one particular GAG within PG can modify the inflammatory response by preventing interactions or removing cytokine from the activating receptors. In lymphoid tissue, the CS and DS chains of versican are able to bind chemokines and thereby negatively regulate the activation of integrin and Ca<sup>2+</sup> mobilization (112). CS chains within versican can then counteract the effect of HS chains in recruiting leukocytes (112). Another role of the pro-inflammatory cytokines and growth factors is to regulate the levels of proteases and endoglycosidases involved in the metabolism of PG. IL-1 or IL-6 induced *in vitro* the breakdown of aggrecan, which has been associated with an increased of aggrecanase activity (78,113,114). Elevated concentrations of IL-6 in the synovial fluids and serum of patients with arthritis have been implicated in the joint tissue destruction (114).

#### V. PG and GAG in anti-inflammatory strategies

Prevention of inflammatory cell recruitment and cell adhesion, prevention of growth factor/cytokine/chemokine activation and inhibition of the activity of enzymes represent actually three of the main anti-inflammatory strategies that can target GAG. Using GAG to block the cell adhesion and the recruitment of cells was probably the simple approach to be attempted. As being well documented, heparin and other GAG has been shown to reduce the accumulation of leukocytes in the lung diseases such as asthma (115-117). Moreover, when administered directly to the airways, heparin was also shown to inhibit allergen-induced eosinophil infiltration (116). Other GAGs, such as CS and HS were also able to influence the extent of allergen-induced eosinophil infiltration when DS had no effect (116). The sulfation level of heparin seems to correlate with the inhibition of leukocyte recruitment since the use of lower-sulfated derivatives (such as N-desulfated heparin or 2-O-desulfated heparins) were less efficient than heparin to block leukocyte recruitment and adhesion (28). Small heparin oligosaccharides (tetrasaccharides and longer) bind L- and P-selectin and inhibit their function *in vitro* (118,119). Heparin can interfere directly with the selectin-ligand interaction which is the first step of leukocyte recruitment from the vasculature to a site of inflammation (27-31,118). In animal models, heparin oligosaccharides, including non-anticoagulant tetrasaccharides, are effective L- and P-selectin inhibitors *in vitro* and have been described to display anti-inflammatory activity (118,119).

The blockade of enzyme activities represents the second main anti-inflammatory strategies. As we described, heparin can prevent T cell migration by inhibiting T-lymphocyte heparanase (96). HS can also inhibit matrix

metalloproteinase activity (127). Heparin and a number of heparin derivatives such as 2-O-desulfated heparin and N-desulfated heparin can block protease activities such as the platelet aggregation induced by cathepsin G (128,129). Heparin and heparan oligosaccharides, as well as over-O-sulfated oligosaccharides and related polysulfated compounds inhibit human leukocyte elastase (67,130). Moreover, a trisulfated disaccharide can inhibit delayed-type hypersensitivity reactivity at much lower concentrations than heparin itself (131). Circulating heparin may also replace endothelial HS, which has been lost through digestion by leukocyte-derived enzymes. Such replacement has been found to correlate directly with sites of active inflammation in Crohn's disease and ulcerative colitis (132). Heparin is a potent anticoagulant and this is also true of HS chains that contain the antithrombin III binding site. The blockade of all these enzymes which are responsible for the spatiotemporal remodelling of ECM (trafficking of leukocytes and extravasation) and the liberation of HSPG-bound cytokines/growth factors would have major effects on the response to inflammatory stimuli.

Finally, the blockade of cytokine, chemokine or growth factor activities is probably the most challenging of these GAG therapy approaches. According to the literature, the biological activities and functions of GAG depend clearly on the saccharide composition, molecular weight, functional group distribution, overall charge, sulfate content and also their distribution along the chain. Relatively short domains within GAG chains such as HS and DS have been shown to act as specific co-receptors for growth factors and cytokines, that directly modulate the activation of the cognate signaling receptors (3,4). Manipulation of the sulfate groups in GAG has been a potential strategy to generate low anticoagulant molecules with specifically retained pharmacological properties and a great degree of biological specificity. In the treatment of deep-vein thrombosis, low molecular weight heparins have been used now for many years. A compound issued from sugar chemistry called Anixtra® (fondaparinux sodium), have been successfully launched in the treatment of venous thrombosis and pulmonary embolism and has been a major advance for GAG-based therapy strategies (19,120). This fully synthetic pentasaccharide present a high degree of purity in contrast of LMWH which are animal sourced (120). More recently, Idraparinux sodium which also belongs to the synthetic oligosaccharide family has been launched in several clinical trials. New molecules will be probably developed soon to pursue this major technical and therapeutic advance (120).

GAG strategies could offer great opportunities and challenges for developing small molecules that selectively inhibit several events, e.g., cell adhesion and cytokine activity, which, due to their several activities and may then be particularly useful in treating inflammatory responses in various

pathologies. Non-sulfated and N-acetylated domains present in the HS chain, called bridging domains, may also offer a good target for drug design strategy. The recognition of specific sequence in the HS chains by IL-8 has been proposed to regulate recruitment of neutrophils (36). Dimeric IL-8 (121) and tetrameric PF-4 (122) can bind two short S-domains bridged by an N-acetylated sequence. The bridging domain of PF-4 seems to be similar in size to the one identified for IL-8, although the domain is suggested to stretch over the intersection between two dimers and not over two monomers as in the suggested model for IL-8 (122). In the case of interferon  $\gamma$ , the bridging domain is considerably longer (15-16 disaccharide units) than that associated with IL-8 binding (8,9,123,124). Regulating the interaction of IL-8 with a cell surface PG with GAG antagonists, could for instance limit neutrophil emigration from the vessel by preventing integrin upregulation or by inducing premature activation of circulating granulocytes prior to their tethering to the endothelium (125,126). Moreover, heparin can compete with endogenous HS on aortic endothelium as a receptor for monocyte L-selectin and then lowering the probability of monocyte rolling (41). Depending on the activity and nature of the GAG antagonist, the adhesion of lymphocyte to ECM will be reduced due to a decrease in receptor occupancy, consequently disrupting chemotaxis of inflammatory cells. Finally, there is no doubt that the advent of synthetic routes to such oligosaccharides means that GAG-based therapeutic approaches could become more common.

## References

1. Iozzo, R.V. 1998, *Annu. Rev. Biochem.*, 67, 609.
2. Bernfield, M., Gotte, M., Park, P.W., Reizes, O., Fitzgerald, M.L., Lincecum, J., and Zako, M. 1999, *Annu. Rev. Biochem.*, 68, 729.
3. Delehsdè, M., Lyon, M., Sergeant, N., Rahmouné, H., and Fernig, D.G. 2001, *J. Mammary Gland Biol Neoplasia*, 6, 253.
4. Gallagher, J.T., and Lyon, M. 2000, *Proteoglycans: Structure, Biology and Molecular Interactions* Iozzo, and R.V. (Ed), Marcel Dekker, New York, 27.
5. Delehsdè, M., Lyon, M., Vidyasagar, R., McDonnell, T.J., and Fernig, D.G. 2002, *J. Biol. Chem.*, 277, 12456.
6. Delehsdè, M., Seve, M., Sergeant, N., Wanalle, I., Lyon, M., Rudland, P.S., and Fernig, D.G. 2000, *J. Biol. Chem.*, 275, 33905.
7. Delehsdè, M., Deudon, E., Boilly, B., and Hondermarck, H. 1996, *Exp. Cell Res.*, 229, 398.
8. Lortat-Jacob, H., and Grimaud, J.A. 1991, *FEBS Lett.*, 280, 152.
9. Lortat-Jacob, H., Balzer, F., and Grimaud, J.A. 1996, *J. Biol. Chem.*, 271, 16139.
10. Lortat-Jacob, H., Grosdidier, A., and Imberty, A. 2002, *Proc Natl Acad Sci U S A*, 99, 1229.
11. Vives, R.R., Sadir, R., Imberty, A., Rencurosi, A., and Lortat-Jacob, H. 2002, *Biochemistry*, 41, 14779.

12. Sadr, R., Bakux, F., Grosdidier, A., Imberby, A., and Lortat-Jacob, H. 2001, *J Biol Chem.*, 276, 8288.
13. Sadr, R., Forest, E., and Lortat-Jacob, H. 1998, *J Biol Chem.*, 273, 10919.
14. Vaday, G.G., and Lider, O. 2000, *J. Leukoc. Biol.*, 67, 149.
15. Bechard, D., Schepereel, A., Hammad, H., Gentina, T., Tsiocopoulos, A., Aumercier, M., Pestel, J., Dessaint, J.P., Tomasi, A.B., and Lassalle, P. 2001, *J. Immunol.*, 167, 3099.
16. Bechard, D., Meignin, V., Schepereel, A., Oudin, S., Kervozze, G., Bertheau, P., Janin, A., Tommel, A., and Lassalle, P. 2000, *J. Vasc. Res.*, 37, 417.
17. Bechard, D., Gentina, T., Delehedde, M., Schepereel, A., Lyon, M., Aumercier, M., Vazeux, R., Richet, C., Degand, P., Jude, B., Janin, A., Fernig, D.G., Tommel, A.B., and Lassalle, P. 2001, *J. Biol. Chem.*, 276, 48341.
18. Lassalle, P., Moket, S., Janin, A., Heyden, J.V., Tavemier, J., Fiers, W., Devos, R., and Tommel, A.B. 1996, *J Biol Chem.*, 271, 20458.
19. Petitou, M., Driguez, P.A., Duchaussoy, P., Herault, J.P., Lormeau, J.C., and Herbert, J.M. 1999, *Bioorg. Med. Chem Lett.*, 9, 1161.
20. Petitou, M., Duchaussoy, P., Driguez, P.A., Herault, J.P., Lormeau, J.C., and Herbert, J.M. 1999, *Bioorg. Med. Chem Lett.*, 9, 1155.
21. Petitou, M., Herault, J.P., Bernat, A., Driguez, P.A., Duchaussoy, P., Lormeau, J.C., and Herbert, J.M. 1999, *Nature*, 398, 417.
22. Dilhas, A., and Bonnaffé, D. 2003, *Carbohydr Res.*, 338, 681.
23. Karst, N.A., Islam, T.F., and Linhardt, R.J. 2003, *Org Lett.*, 5, 4839.
24. El Habbal, M.H., Smith, L., Elliott, M.J., and Strobel, S. 1995, *Cardiovasc. Res.*, 30, 676.
25. Vaday, G.G., and Lider, O. 2000, *J. Leukoc. Biol.*, 67, 149.
26. Schor, H., Vaday, G.G., and Lider, O. 2000, *Dev. Immunol.*, 7, 227.
27. Lever, R., Hout, J.R., and Page, C.P. 2000, *Br J. Pharmacol.*, 129, 533.
28. Silvestro, L., Viano, I., Macario, M., Colangelo, D., Montrocchio, G., Panico, S., and Fantozzi, R. 1994, *Semin. Thromb. Hemost.*, 20, 254.
29. Bazzoni, G., Beltran Nunez, A., Mascellani, G., Bianchini, P., Dejana, E., and Del Maschio, A. 1993, *J. Lab. Clin. Med.*, 121, 268.
30. Smallegovic, A., Lever, R., and Page, C.P. 2001, *Br J. Pharmacol.*, 134, 827.
31. Tyrrell, D.J., Killeather, S., and Page, C.P. 1995, *Trends Pharmacol. Sci.*, 16, 198.
32. Diamond, M.S., Alon, R., Parkos, C.A., Quinn, M.T., and Springer, T.A. 1995, *J. Cell. Biol.*, 130, 1473.
33. Teixeira, M.M., and Hellewell, P.G. 1993, *Br J. Pharmacol.*, 110, 1496.
34. Sy, M.S., Schneberger, E., McChuskey, R., Greene, M.L., Rosenberg, R.D., and Benacerraf, B. 1983, *Cell. Immunol.*, 82, 23.
35. Tangelier, G.J., and Arfors, K.E. 1991, *Blood*, 77, 1565.
36. Rot, A. 1992, *Immunol Today*, 13, 291.
37. Tanaka, Y., Adams, D.H., Hubscher, S., Hirano, H., Siebenlist, U., and Shaw, S. 1993, *Nature*, 79.
38. Whitelock, J.M., Murdoch, A.D., Iozzo, R.V., and Underwood, P.A. 1996, *J Biol Chem.*, 271, 10079.
39. Goodison, S., Urquidí, V., and Tarrin, D. 1999, *Mol. Pathol.*, 52, 189.
40. Jones, M., Tussey, L., Athanasou, N., and Jackson, D.G. 2000, *J Biol. Chem.*, 275, 7964.
41. Ley, K. 1993, *Eur Heart J.*, 14, 68.
42. Yeaman, C., and Rapraeger, A.C. 1993, *J. Cell. Biol.*, 122, 941.
43. Yeaman, C., and Rapraeger, A.C. 1993, *J. Cell. Physiol.*, 157, 413.
44. Webb, L.M., Ehrengruber, M.U., Clark-Lewis, I., Baggiolini, M., and Rot, A. 1993, *Proc. Natl. Acad. Sci. USA*, 90, 7158.
45. Sweeney, E.A., Lortat-Jacob, H., Priesky, G.V., Nakamoto, B., and Papayannopoulou, T. 2002, *Blood*, 99, 44.
46. Flaumenhaft, R., Moscatelli, D., and Rifkin, D.B. 1990, *J. Cell Biol.*, 111, 1651.
47. Vlodavsky, I., Foks, Z., Ishai-Michaeli, R., Bashkin, P., Levi, E., Korner, G., Bar-Shavit, R., and Klagsbrun, M. 1991, *J. Cell. Biochem.*, 45, 167.
48. Marshall, L.J., Ramdin, L.S., Brooks, T., PC, D.P., and Shune, J.K. 2003, *J Immunol.*, 171, 2057.
49. Ohitsuki, T., Suzu, S., Harake, K., Nagata, N., Miura, Y., and Motoyoshi, K. 1993, *Biochem. Biophys. Res. Commun.*, 190, 215.
50. Parenbsimer, A., Schwarz, K., Wrocklage, C., Kolsch, E., and Kresse, H. 1995, *J. Immunol.*, 155, 5557.
51. Bame, K.J. 2001, *Glycobiology*, 11, 91.
52. Vlodavsky, I., and Friedmann, Y. 2001, *J. Clin. Invest.*, 108, 341.
53. Ishai-Michaeli, R., Eldor, A., and Vlodavsky, I. 1990, *Cell Regul.*, 1, 833.
54. Vlodavsky, I., Miso, H.O., Medalion, B., Danagher, P., and Ron, D. 1996, *Cancer Metastasis Rev.*, 15, 177.
55. Thompson, K., and Rabinovitch, M. 1996, *J. Cell. Physiol.*, 166, 495.
56. Rich, C.B., Nugent, M.A., Stone, P., and Foster, J.A. 1996, *J. Biol. Chem.*, 271, 23043.
57. Wagner, L., Yang, O.O., Garcia-Zepeda, E.A., Ge, Y., Kalam, S.A., Walker, B.D., Pasternack, M.S., and Luster, A.D. 1998, *Nature*, 391, 908.
58. Delehedde, M., Lyon, M., Gallagher, J.T., Rudland, P.S., and Fernig, D.G. 2002, *Biochem J.*, 366, 235.
59. Clasper, S., Vellemans, S., Fore, M., Plebanski, M., Wordsworth, P., David, G., and Jackson, D.G. 1999, *J. Biol Chem.*, 274, 24113.
60. Sergeant, N., Lyon, M., Rudland, P.S., Fernig, D.G., and Delehedde, M. 2000, *J. Biol. Chem.*, 275, 17094.
61. van der Voort, R., Keelren, R.M., Beuling, E.A., Spaargaren, M., and Pals, S.T. 2000, *J. Exp. Med.*, 192, 1115.
62. Salek-Ardakani, S., Arrand, J.R., Shaw, D., and Maclell, M. 2000, *Blood*, 96, 1879.
63. Petersen, F., Bock, L., Flad, H.D., and Brandt, E. 1998, *J. Immunol.*, 161, 4347.
64. Petersen, F., Brandt, E., Lindahl, U., and Spillmann, D. 1999, *J. Biol. Chem.*, 274, 12376.
65. Wolff, E.A., Greenfield, B., Taub, D.D., Murphy, W.J., Bennett, K.L., and Aruffo, A. 1999, *J Biol Chem.*, 274, 2518.
66. Ermoleff, J., Boudier, C., Laine, A., Meyer, B., and Bieth, J.G. 1994, *J. Biol Chem.*, 269, 29502.
67. Redini, F., Tixier, J.M., Petitou, M., Choisy, J., Robert, L., and Homebeck, W. 1988, *Biochem J.*, 252, 515.

68. Strukova, S.M. 2001, *Biochemistry (Mosco)*, 66, 8.
69. Ernolteff, J., Duranton, J., Petitou, M., and Blith, J.G. 1998, *Biochem. J.*, 330, 1369.
70. Penc, S.F., Pomahac, B., Winkler, T., Dorschner, R.A., Eriksson, E., Herndon, M., and Gallo, R.L. 1998, *J. Biol. Chem.*, 273, 28116.
71. Subramanian, S.V., Fitzgerald, M.L., and Bernfield, M. 1997, *J. Biol. Chem.*, 272, 14713.
72. Karunainen, V., Wang, H., Schick, C., and Bernfield, M. 1998, *J. Biol. Chem.*, 273, 11563.
73. Kato, M., Wang, H., Kaimulainen, V., Fitzgerald, M.L., Ledbetter, S., Ornitz, D.M., and Bernfield, M. 1998, *Nat. Med.*, 4, 691.
74. Cizmeci-Smith, G., Langan, E., Youkey, J., Showalter, L.J., and Carey, D.J. 1997, *Arterioscler. Thromb. Vasc. Biol.*, 17, 172.
75. Gallo, R.L., Chao, M., Powsic, T., Page, C., Eriksson, E., Klagsbrun, M., and Bernfield, M. 1994, *Proc. Natl. Acad. Sci. U.S.A.*, 91, 11035.
76. Halpert, I., Sires, U.L., Roby, J.D., Potter-Perigo, S., Wight, T.N., Shapiro, S.D., Weigus, H.G., Wickline, S.A., and Parks, W.C. 1996, *Proc. Natl. Acad. Sci. U.S.A.*, 93, 9748.
77. Billinghamurst, R.C., Dahlberg, L., Ionescu, M., Reiner, A., Boume, R., Ronabesk, C., Mitchell, P., Hambor, J., Diekmann, O., Tschesche, H., Chen, J., Van Wart, H., and Poole, A.R. 1997, *J. Clin. Invest.*, 99, 1534.
78. Caterson, B., Flannery, C.R., Hughes, C.E., and Little, C.B. 2000, *Matrix Biol.*, 19, 333.
79. Ishiguro, N., Ito, T., Oguchi, T., Kojima, T., Iwata, H., Ionescu, M., and Poole, A.R. 2001, *Arthritis Rheum.*, 44, 2503.
80. Kojima, T., Mwale, F., Yasuda, T., Girard, C., Poole, A.R., and Lavery, S. 2001, *Arthritis Rheum.*, 44, 120.
81. Chevner, A., Mott, J.S., Crine, P., Hoemann, C.D., and Buschmann, M.D. 2001, *Arch. Biochem. Biophys.*, 396, 178.
82. Thunberg, L., Backstrom, G., and Lindahl, U. 1982, *Carbohydr. Res.*, 100, 393.
83. De sai, U.R., Petitou, M., Bjork, I., and Olson, S.T. 1998, *Biochemistry*, 37, 13033.
84. Hricovini, M., Guerrini, M., Bisio, A., Torri, G., Petitou, M., and Casu, B. 2001, *Biochem. J.*, 359, 265.
85. Lindahl, U., Backstrom, G., and Thunberg, L. 1983, *J. Biol. Chem.*, 258, 9826.
86. Ostrowsky, L., Woodman, R.C., Payne, D., Teoh, D., and Kubas, P. 1997, *Circulation*, 96, 2302.
87. Zuo, X.J., Okada, Y., Nicolaidou, E., Toyoda, T., Marchevsky, A., Matloff, J.M., and Jordan, S.C. 1999, *Transplant. Proc.*, 31, 847.
88. Marie, C., Muret, J., Fitting, C., Lossier, M.R., Payen, D., and Cavillon, J.M. 1998, *Blood*, 91, 3439.
89. Dunzendorfer, S., Koneider, N., Rabensteiner, A., Meirhofer, C., Reinisch, C., Romisch, J., and Wiedermann, C.J. 2001, *Blood*, 97, 1079.
90. Vlodavsky, I., Eldor, A., Hammovitz-Friedman, A., Matzner, Y., Ishai-Michaeli, R., Lider, O., Napatnik, Y., Cohen, I.R., and Fuks, Z. 1992, *Invasion Metastasis*, 12, 112.
91. Brucke, N.S., Parker, W., Reissner, K.J., and Platt, J.L. 1998, *J. Cell. Physiol.*, 175, 255.
92. Hoogewerf, A.J., Leone, J.W., Reardon, I.M., Howe, W.J., Assa, D., Heinrichson, R.L., and Ledbetter, S.R. 1995, *J. Biol. Chem.*, 270, 3268.
93. Parish, C.R., Hindmarsh, E.J., Bartlett, M.R., Staykova, M.A., Cowden, W.B., and Willenborg, D.O. 1998, *Immunol. Cell. Biol.*, 76, 104.
94. Dempsey, L.A., Plummer, T.B., Coombes, S.L., and Platt, J.L. 2000, *Glycobiology*, 10, 467-475.
95. Parish, C.R., Freeman, C., and Hulett, M.D. 2001, *Biochim. Biophys. Acta*, 1471, 99.
96. Lider, O., Melkoni, Y.A., Miller, T., Bar-Tana, R., Vlodavsky, I., Baharav, E., Cohen, I.R., and Napatnik, Y. 1990, *Eur. J. Immunol.*, 20, 493.
97. Gilat, D., Hershkovitz, R., Goldkorn, I., Cahalon, L., Korner, G., Vlodavsky, I., and Lider, O. 1995, *J. Exp. Med.*, 181, 1929.
98. Pillarsetti, S., Paka, L., Obunike, J.C., Berglund, L., and Goldberg, I.J. 1997, *J. Clin. Invest.*, 100, 867.
99. Pillarsetti, S., Paka, L., Sasaki, A., Vanni-Reyes, T., Yin, B., Parthasarathy, N., Wagner, W.D., and Goldberg, I.J. 1997, *J. Biol. Chem.*, 272, 15753.
100. Moscaelli, D. 1992, *J. Biol. Chem.*, 267, 25803.
101. Spermb, G.V., and Nugent, M.A. 1998, *Biochemistry*, 37, 13153.
102. Tunova, S., Hatch, B.A., Law, D.J., and Bame, K.J. 1999, *Biochem. J.*, 337, 471.
103. Bame, K.J., Danda, J., Hassall, A., and Tunova, S. 1997, *J. Biol. Chem.*, 272, 17005.
104. Snow, A.D., Sekiguchi, R.T., Nochin, D., Kalaria, R.N., and Kimata, K. 1994, *Am. J. Pathol.*, 144, 337.
105. Jackson, D.G. 1997, *Biochem. Soc. Trans.*, 25, 220.
106. Mauviel, A., Santra, M., Chen, Y.Q., Uitto, J., and Iozzo, R.V. 1995, *J. Biol. Chem.*, 270, 11692.
107. Qvarnstrom, E.E., Jarvelainen, H.T., Kinsella, M.G., Ostberg, C.O., Sandell, L.J., Page, R.C., and Wight, T.N. 1993, *Biochem. J.*, 294, 613.
108. Garcia de Yebenes, E., Ho, A., Damani, T., Fillit, H., and Blum, M. 1999, *J. Neurochem.*, 73, 812.
109. Romaris, M., Heredia, A., Molist, A., and Bassols, A. 1991, *Biochim. Biophys. Acta*, 1093, 229.
110. Kahari, V.M., Larjava, H., and Uitto, J. 1991, *J. Biol. Chem.*, 266, 10608.
111. Kahari, V.M., Hakkinen, L., Westermarck, J., and Larjava, H. 1995, *J. Invest. Dermatol.*, 104, 503.
112. Hirose, J., Kawashima, H., Yoshie, O., Tashiro, K., and Miyasaka, M. 2001, *J. Biol. Chem.*, 276, 5228.
113. Cawston, T., Billington, C., Cleaver, C., Elliott, S., Hui, W., Koshy, P., Shingleton, B., and Rowan, A. 1999, *Ann. N.Y. Acad. Sci.*, 878, 120.
114. Flannery, C.R., Little, C.B., Hughes, C.E., Curtis, C.L., Caterson, B., and Jones, S.A. 2000, *Matrix Biol.*, 19, 549.
115. Page, C.P. 1997, *Thorax*, 52, 924.
116. Seeds, E.A., Horne, A.P., Tyrrell, D.J., and Page, C.P. 1995, *Pulm. Pharmacol. Ther.*, 8, 97.
117. Seeds, E.A., and Page, C.P. 2001, *Pulm. Pharmacol. Ther.*, 14, 111.
118. Nelson, R.M., Cecconi, O., Roberts, W.G., Anuffo, A., Linhardt, R.J., and Bevilacqua, M.P. 1993, *Blood*, 82, 3253.

119. Nelson, R.M., Dolich, S., Aruffo, A., Cecconi, O., and Bevilacqua, M.P. 1993, *J. Clin. Invest.*, 91, 1157.
120. Bauer, K.A. 2003, *Chest*, 124, 364S.
121. Witt, D.P., and Lander, A.D. 1994, *Curr. Biol.*, 4, 394.
122. Stringer, S.E., and Gallagher, J.T. 1997, *J. Biol. Chem.*, 272, 20508.
123. Lubineau, A., Lortat-Jacob, H., Gavaud, O., Sarrazin, S., and Bonnafant, D. 2004, *Chemistry*, 10, 4265.
124. Sarrazin, S., Bonnafant, D., Lubineau, A., and Lortat-Jacob, H. 2005, *J Biol Chem*, 280, 37558.
125. Ley, K., Baker, J.B., Cybulsky, M.L., Gimbrone, M.A., Jr., and Luscinskas, F.W. 1993, *J. Immunol.*, 151, 6347.
126. Simonet, W.S., Hughes, T.M., Nguyen, H.Q., Trebaszky, L.D., Danilenko, D.M., and Medlock, E.S. 1994, *J. Clin. Invest.*, 94, 1310.
127. Kennagy, R.D., Nikkari, S.T., Welgus, H.G., and Clowes, A.W. 1994, *J. Clin. Invest.*, 93, 1987.
128. Evangelista, V., Piccardoni, P., Mauge ri, N., De Gaetano, G., and Gerletti, C. 1992, *Eur. J. Pharmacol.*, 216, 401.
129. Cerletti, C., Rajtar, G., Marchi, E., and de Gaetano, G. 1994, *Semin. Thromb Hemost.*, 20, 245.
130. Lemini, A., Ternai, B., and Ghosh, P. 1985, *Biochem. Int.*, 10, 221.
131. Lóder, O., Cahalon, L., Gilat, D., Hershkoviz, R., Siegel, D., Margalit, R., Shoseyov, O., and Cohen, I.R. 1995, *Proc. Natl. Acad. Sci. U S A*, 92, 5037.
132. Murch, S.H., MacDonald, T.T., Walker-Smith, J.A., Levin, M., Lionetti, P., and Klein, N.J. 1993, *Lancet*, 341, 711.



## Stromal derived growth factor-1: A target for GAG-based therapies?

# I2

Estelle Adam, Vincent Motte, Stéphane Sarrazin and Maryse Delehedde  
ENDOTIS PHARMA, Parc Eurasanté, 70 rue du Dr YERSIN, 59120 Loos, France

### Abstract

*Stromal cell-derived factor-1 (SDF-1 $\alpha$ )/CXCL12, a member of the CX12 chemokine subfamily, is a secreted protein of 68 amino acids, which binds to CXCR4, a G-protein-coupled receptor. Together SDF-1 and CXCR4 coordinate a diverse and large variety of cellular functions, including the tight control of basal trafficking and homing of the various subsets of hematopoietic cells from the blood stream to specific anatomical sites of the body. Interestingly SDF-1 strongly interacts with glycosaminoglycans (GAGs) present on the cell surface, and more precisely with heparan sulphates (HS), which have been*

*known for years now, to be implicated in the control of the activation and localization of various growth factors and cytokines. SDF-1 and CXCR4 have been shown to be clearly the main players involved in the homing of hematopoietic stem and progenitor cells within the bone marrow, attracting them and regulating their engraftment into specific niches for correct hematopoiesis to take place. A large number of studies have demonstrated that SDF-1 $\alpha$  is implicated outside the regulation of leukocyte migration, in major biological processes such as normal embryonic development (especially in vascular and neuronal systems) and in the control of normal adult homeostasis. The SDF-1/CXCR4 axis plays also a pivotal role in pathological conditions, more precisely in cancer where they both determine the destination of metastatic tumor cells, or in inflammatory diseases where they can regulate leukocyte trafficking. Such implications in pathological processes have suggested since few years now, that both SDF-1 and CXCR4 could be potential targets for development of new therapy approaches. Whereas the main drugs currently under investigations directly target the SDF-1/CXCR4 interactions, new challenging opportunities in drug development could come from strategy directly targeting the SDF-1/HS interactions both in the normal and pathological conditions.*

### Abbreviations

SDF-1, Stromal cell derived factor-1; GPCR, G-Protein Coupled Receptor; GAGs, Glycosaminoglycans; HS, Heparan Sulfate; PG, Proteoglycans; PBSF, Pre-B-cell growth Stimulatory Factor; MMP, Matrix Metalloproteases; HIV, Human Immunodeficiency Virus; KO, Knock-Out; FGF, Fibroblast Growth Factor; HSC, Hematopoietic Stem Cells; BM, Bone Marrow; BMEC, Bone Marrow Endothelial Cells; G-CSF, Granulocyte-Colony Stimulating Factor; TGF, Tumor Growth Factor; VEGF, Vascular Endothelial Growth Factor; RA, Rheumatoid Arthritis.

### Introduction

As part of the vast family of chemoattractant cytokines, stromal cell-derived factor-1 $\alpha$  (SDF-1 $\alpha$ )/CXCL12, is a secreted protein constitutively expressed by several cell types. SDF-1 $\alpha$  binds to CXCR4, a G-protein-coupled receptor (GPCR) and coordinates a diverse array of cellular functions (1). Moreover, SDF-1 is able to bind to heparan sulphate proteoglycans (HSPG) on the cell surface, and their interactions have now been clearly implicated in the regulation of the SDF-1 activity (2,3). The SDF-1 chemokine and its receptor CXCR4 have been shown to be involved in numerous normal as well as pathologic processes such as homing of hematopoietic stem cells into the bone marrow, homeostasis, angiogenesis, cancer and inflammatory diseases.

Recently, SDF-1 has been validated as a good target in various pathologies and several SDF-1 or CXCR4 inhibitors are currently in clinical trials (4,5). New opportunities in drug development now could come from our better knowledge of the biological significance of SDF-1/HS interactions, thus offering the possibility to develop new compounds interfering with HS binding.

### I. SDF-1 or CXCL12

Stromal cell-derived factor-1 (SDF-1) or CXCL12 is a member of the CXCR chemokine family. Whereas most chemokine expression is induced by inflammatory cytokines, SDF-1 is constitutively secreted from several cell types. This chemokine was initially described in 1993 by *Tachiro et al* as part of a cDNA library constructed from a mouse bone marrow stromal cell line (6). The chemokine was concomitantly reported by another group as a pre-B-cell growth stimulatory factor (PBSEF), by virtue of its ability to stimulate human early B cell progenitors and was then cloned from human stromal cell line (7). In a third study, *Jiang et al*, working on cDNA library from mouse embryo fibroblasts treated with TPA, isolated a partial cDNA sequence, named TPARI (TPA repressed gene 1), which was found to be SDF-1 (8).

Chemokines constitute a growing family of small inducible secreted proteins (8-14 kDa) involved in a variety of inflammatory and immune responses, best known for their ability to mediate leukocyte activation and trafficking (9,10). Chemokines classification is based upon conserved cysteine residues involved in intramolecular disulfide bond formation and four subfamilies have thus been described: C chemokines, CC chemokines, C-X-C chemokines and C-X<sub>2</sub>-C chemokines. SDF-1 $\alpha$  belongs to the C-X-C chemokines (also known as  $\alpha$ -chemokines) in which the first two cysteine residues are separated by a single amino acid, and is then called CXCL12.

Two SDF-1 isoforms, SDF-1 $\alpha$  and SDF-1 $\beta$ , were originally identified in human and mouse (11). Both isoforms are produced by alternative splicing of a single gene. Unlike others C-X-C chemokine genes which have been located on human chromosome 4q, the human SDF-1 gene maps to chromosome 10q (11) and spans over 88 kilobase-pairs. SDF-1 $\alpha$  and SDF-1 $\beta$  cDNAs encode protein of 68 amino acids and 72 amino acids respectively. Both proteins are identical except for the four amino acids present in the carboxy terminal end of SDF-1 $\beta$  and not in SDF-1 $\alpha$ . SDF-1 show highly conserved amino acid sequence between species, with only one amino acid substitution (isoleucine converted to Valine at position 18) between human and mouse proteins. Unlike other chemokines which are mainly induced by pro-inflammatory cytokines, SDF-1 mRNA is ubiquitously expressed. SDF-1 $\alpha$  is produced by virtually all tissues, in particular by bone marrow stromal cells, and SDF-1 $\alpha$  appears to be

the predominant isoform (12). SDF-1 $\alpha$  and SDF-1 $\beta$  isoforms shared the same similar expression pattern with one exception. SDF-1 $\alpha$  can be detected in adult human brain when SDF-1 $\beta$  mRNA cannot. In human, four additional isoforms have been recently identified (SDF-1 $\gamma$ , SDF-1 $\delta$ , SDF-1 $\epsilon$  and SDF-1 $\phi$ ). These isoforms result also from alternative splicing events of the same gene on chromosome 10 (12). SDF-1 $\gamma$  isoform was first identified in rat with high expression in the brain (13), and this isoform is interestingly induced in rat heart during ischemia injury (14). In human, SDF-1 $\gamma$  was mainly detected in the adult heart but not in brain (12). The physiological significance of these different splice variants have still yet to be determined.

The three-dimensional structure of the SDF-1 chemokine was first determined by Crump *et al* by NMR spectroscopy. From this study, it results that SDF-1 $\alpha$  shows a chemokine-like structure with an N-terminal region, a central region of 3 anti-parallel  $\beta$ -strands and a C terminal  $\alpha$ -helix. (15). Concerning the quaternary structure of SDF-1, *Veidkamp et al* demonstrated that SDF-1 $\alpha$  exists in a monomer-dimer equilibrium depending on the solution conditions, with particular sensitivity to pH and a major role of glycosaminoglycans in the oligomerization of the molecule (16). Structural analysis revealed a positive charged surface located in the first beta strand, which may be important for binding to the negatively charged extracellular loop of CXCR4 (17). Moreover, the N-terminus of SDF-1 was described to be particularly important for CXCR4 activation and thus for the chemokine activity. Structurally, SDF-1 $\alpha$  protein is also more particularly degraded at its N terminus region by proteolytic enzymes such as elastases (18-20), cathepsin G (21), cathepsin K (22), CD26 (also known as dipeptidylpeptidase IV) (23) and matrix metalloproteinases (MMP) (24).

### II. Receptors of SDF-1

The main receptor for SDF-1 is CXCR4 (also known as Fusin or LESTR) for which SDF-1 is the only known ligand (17,25). CXCR4 was originally discovered in the 1990's as an orphan receptor (LESTR for Leucocyte-derived seven transmembrane domain receptor) which shares structural similarity to chemokine receptors, in particular with interleukin-8 (IL-8) receptors (25,26). Before its ligand identification, CXCR4 was also identified in human immunodeficiency virus (HIV) infection as a key element for HIV entry into CD4+ T cells, and therefore called Fusin (27). Like other chemokine receptors, CXCR4 is a seven transmembrane domain, G protein-coupled receptor, with activity involving intracellular calcium mobilization, phosphorylation events and internalization (26,28). Interestingly, CXCR4 knock-out (KO) mice showed the same defects as SDF-1 KO mice, leading to the hypothesis of a monogamous ligand-receptor pair (29), which last until 2005 (30).

The activities of SDF-1 $\alpha$  are numerous and varied. This chemokine, more than any other chemokine identified until now, has a wide range of biological functions, especially outside of the immune system. Consistent with the SDF-1 pleiotropic functions, CXCR4 is widely expressed on both non-hematopoietic and hematopoietic cells, whereas others chemokine receptors expression is more restricted (1). CXCR4 is then detected on leukocytes, i.e. lymphocytes, monocytes, B cells, NK cells, mature dendritic cells, megakaryocytes, as well as on endothelial cells, adult bone marrow stem cells and progenitor cells (31), and brain cells (32). SDF-1 $\alpha$  binding and interaction with CXCR4 induce conformational changes on CXCR4, resulting in signal transduction. As other G-protein coupled receptor, CXCR4 activation can induce phospholipase C- $\gamma$  (PLC- $\gamma$ ) activation and subsequent intracellular calcium mobilization, as well as activation of the mitogen-activated protein kinase (MAPK) pathway and the phosphoinositide-3 kinase (PI3K) pathway (26,28). Activation and inhibition signals from these different pathways are integrated, conducting to cellular responses such as proliferation, survival and migration (1,33).

Recently, it was reported that SDF-1 can also bind to and signal through RDC-1, which was then renamed as CXCR7 (30). RDC-1 is an orphan G-protein coupled receptor with seven transmembrane domains, sharing high homologies with known chemokine receptors. This RDC-1 receptor was cloned in 1995 (34). However the functional significance of the SDF-1 $\alpha$  binding on this new receptor remains to be elucidated.

### III. Interaction of SDF-1 with glycosaminoglycans (GAGs)

Among the molecules that regulate chemokine activities, cellular proteoglycans (PG) and particularly PG carrying heparan sulfate chain have been shown to play strategic roles, being present at the cell surface, freely circulating in blood or in the extracellular matrix of most mammalian cell types (35-37). These heparan sulphate proteoglycans (HSPG) structurally consist of repeating disaccharides units, synthesized by alternating sulphated glucosamine and glucuronic residues (35,37). The disaccharide components are modified with sulfation at various positions, i.e. 2-O and 6-O sulphate, and HS are thus highly negatively charged structure (37,38). Sulfation pattern defines regions of variable lengths and variable sulfation separated by regions of low sulfation and relatively uniform structure. HS chain thus presents high structural diversity, inside the HS chain, but also in HS from different cells and tissues. Moreover, the sulfation pattern along the HS chain is thought to be important in defining the specificity of the protein/HS interactions (3,55). For instance, distinct structural requirements for binding of HS to fibroblast growth factor-2 (FGF-2) (2-O-sulfate but no 6-O-sulfate) vs FGF-1 (2-O-sulfate and 6-

O-sulfate) was demonstrated (39), and was furthermore associated with distinct cellular responses to FGF (40).

HS are found ubiquitously at the cell surface and in the extracellular matrix. An increasing number of growth factors, cytokines or chemokines has been shown to bind to HS structure, with such interactions increasing their stability, regulating their activity, as well as serving of storage reservoirs (37,38). Different studies report that HS immobilize and enhance local concentrations of chemokines (38,41), protect chemokines from degradation (41,42), promote their oligomerization (43,44) and facilitate their presentation to the receptors (43,45). Moreover, these chemokines/GAG interactions were found to be important for the *in vivo* activities of some inflammatory molecules (46).

SDF-1 $\alpha$  have been shown to bind to heparan sulfate through his first beta-strand (47,48) with high affinity ( $K_d$  of 30 nM) to the sulphated regions of HS (called the S-domains) (2,49). More particularly in SDF-1 $\alpha$ , amino acids Lys<sup>44</sup> and Lys<sup>37</sup> have a dominant role in the binding to HS. The residues Lys<sup>1</sup>, Arg<sup>41</sup>, and Lys<sup>48</sup> participate in the recognition process but are not strictly required for the interaction to take place (49). The HS interaction site for SDF-1 binding has not yet been more precisely defined. However a minimal length of four disaccharides units localized in a highly sulphated domain was shown to be sufficient to bind SDF-1 (49). A major role of 6-O sulfation was also described for this interaction to take place (49,50). HSPG was shown, within tissues, to bind and concentrate SDF-1 at the surface of a number of cells, including bone marrow endothelial cells (51) and rheumatoid endothelial cells (52). SDF-1 binding to cell surface HS has been demonstrated to be involved in triggering specific arrest of Kaposi's sarcoma herpes virus-infected cells (53), enhancing the chemokine-induced migration of cells by the formation of haptotactic gradient (54,55) or facilitating the persistent occupancy and sustained down-regulation of CXCR4 used by HIV as a co-receptor, thus contributing to the HIV inhibitory effect of SDF-1 (56). Moreover, SDF-1 binding to HS was found to facilitate SDF-1 binding to CXCR4 at the cell surface of primary macrophages (57). As part of the physiological roles of SDF-1 $\alpha$ /HS interaction, the GAG binding C-terminus region of SDF-1 $\alpha$  can thus increase the biological effect mediated by its N-terminus part, and this effect can be inhibited by addition of soluble heparin (58). Another role of HS in the regulation of SDF-1 $\alpha$  was demonstrated by *Sadrir et al.*, showing that HS are able to protect SDF-1 $\alpha$  from proteolysis mediated by CD26 (2), directly affecting the chemokine activity. As largely described for FGF-2 (37,59), there have been several studies directly demonstrating the role of HS in mediating SDF-1 activities. For instance in hematopoietic stem cells (HSC) mobilization, the degradation of HS by specific enzymes such as heparinase totally prevent the migration of progenitors (54). Moreover using a competitive inhibitor of

sulfation (called sodium chorate directly added to the cell culture medium), also prevent the adhesion of progenitor cells on endothelium (51). Furthermore, the role of HS in SDF-1 activities was demonstrated on other cellular models of different diseases. A pretreatment of endothelial cells with heparinase can reduce cell surface SDF-1 expression on endothelial cells and consequently virtually abolished cell arrest onto endothelial cells under conditions of shear flow as what can be observed during metastasis. By contrast, a treatment of endothelial cells with enzymes specific of other glycosaminoglycans (such as the chondroitinase ABC) had no or little effect (53). Desulfation of HSPGs by pretreatment with sodium chlorate decreased or abrogated the ability of the SDF-1 recruited cells to bind to endothelial cells as it is observed in rheumatoid arthritis (52). Binding to cell surface heparan sulfate proteoglycans determines the capacity of SDF-1 $\alpha$  to prevent the fusogenic activity of HIV-1 isolates in leukocytes (56). SDF-1 $\alpha$  mutants lacking the capacity to interact with HSPGs showed a substantially reduced capacity to prevent cell-to-cell fusion mediated by X4 HIV envelope glycoproteins. Moreover, enzymatic removal of cell surface HS diminishes also the HIV-inhibitory capacity of SDF-1 to the levels shown by the HS-binding-disabled mutant counterparts (56). However, the exact role of the SDF-1/GAG interactions in mediating *in vitro* and *in vivo* biological functions of the chemokine remains yet to be elucidated.

#### IV. Main biological activities of SDF-1

Signaling through CXCR4 induces chemotactic migration and adhesion of receptor-bearing cells to integrin ligands, as is typical for chemokine receptors. However, unlike other chemokine receptors which are more particularly expressed or up-regulated under inflammatory conditions, CXCR4 is expressed on all types of leukocytes and a role of this receptor in the turnover and homing of blood leukocytes is suggested (1). CXCR4 does not only regulate leukocyte trafficking to sites of inflammation, but is also implicated in normal physiological events (Fig.1).

##### 1. Embryogenesis, vasculogenesis and neuronal development

An essential role for SDF-1 in embryonic development was demonstrated in studies of mouse embryos that were genetically engineered to be deficient either in SDF-1 or in CXCR4. Deletion of either gene was lethal before or shortly after birth, and mutant embryos showed severe defects in B lymphopoiesis both in fetal liver and bone marrow, as well as high defects in bone marrow myelopoiesis (29,60). In different studies, SDF-1 or CXCR4 deficient mice were also shown to have defective ventricular septum formation of the heart (29,60), and abnormal large vessels formation in the

gastrointestinal tract (61). Moreover CXCR4 knockout mice exhibit gross defects in vascular development, suggesting an important role for SDF-1 and CXCR4 in normal organ vascularization (61). Interestingly, both knock-out mice showed impaired neuronal cells migration leading to impaired cerebellum development (32,60), which directly involves CXCR4 and SDF-1 in the central nervous system development, which was then confirmed in numerous studies. Taken together, these results implicated the SDF-1 and CXCR4 couple in several essential functions in embryonic development, with particular actions on vascular and neuronal developments.

##### 2. Adult angiogenesis and vasculogenesis

More recently, SDF-1 $\alpha$  has been clearly involved in normal adult vascular development, as well as in pathological conditions. SDF-1 $\alpha$  seems to play a role both in angiogenesis, i.e. sprouting of resident endothelial cells from preexisting blood vessels, and in vasculogenesis, i.e. new blood vessel formation by endothelial progenitor cells recruitment (62). It is now clear that bone marrow (BM) derived cells are implicated in the vasculogenesis process, even if the exact nature of these bone marrow derived cells remains unclear. More particularly whether one or several type of cells are implicated, is still under investigation. In post-natal vasculogenesis, these BM derived cells are recruited to site of neovascularization. SDF-1, by its chemotactic properties, was shown to play an important role in such cell recruitment. *Cevadini et al* thus demonstrated that SDF-1 enhances endothelial progenitor adhesion to ischemic endothelium (63), suggesting that SDF-1 could act as a targeting and homing signal for these cells. In the same study, the blockade of either SDF-1 or CXCR4 dramatically reduces the recruitment of progenitor cells, with a consecutive reduction of their presence in blood flow and in the neovascularization process (63). Other studies refer to the fact that the BM derived cells could act as accessory cells releasing pro-angiogenic factors for new blood vessels formation. *Grunewald et al* have recently highlighted the essential role of SDF-1 $\alpha$  in the perivascular homing and retention of these CXCR4+ BM derived cells, directly implicating SDF-1 in the angiogenic process (64). Furthermore, there have been reports of the implication of SDF-1 and CXCR4 in the mobilization, migration, homing, as well as in the proliferation of hematopoietic cells, mediating new blood vessel formation under ischemic conditions (1). Clear evidence for SDF-1 role in angiogenesis also comes from tumorigenesis studies, underlining SDF-1 and CXCR4 involvement in tumor angiogenesis.

##### 3. Mononuclear cells trafficking and homeostasis

SDF-1 was firstly described as PBSF, for "pre-B-cell growth stimulatory factor" as it was able to support the proliferation of a stromal cell-dependent

pre-B cell clone (7). This chemokine mainly acts as chemoattractant for B homing, CD44 and hyaluronic acid (a non sulphated glycosaminoglycan) were progenitor cells (65), naive and memory B cells, and is implicated in B found to be implicated in the transendothelial migration and anchorage of HSC lymphocytes trafficking to secondary lymphoid organs, as well as their (73). Finally, SDF-1 $\alpha$  was also shown to mediate antiapoptotic and specific localization within these secondary lymphoid organs (66). SDF-1 $\alpha$  is proliferative effects in the BM, acting as an important survival and mitogenic also a highly potent T cell chemoattractant (67) and is involved in basal factor (74,75).

extravasation of T cells by regulating rolling, adhesion and endothelial transmigration. As for B cells, SDF-1 regulates T cells trafficking and their marrow, the SDF-1/CXCR4 axis seems to be a good target for hematopoietic localization inside the secondary lymphoid organs, such as the lymph nodes. stem cells mobilization to the peripheral blood. Mechanisms implicated in However SDF-1 does not contain the conserved ELR motif (a glutamic acid-HSC mobilization, i.e HSC egress from bone marrow to peripheral blood, are leucine-arginine sequence motif) found in most C-X-C chemokines, acting less well understood. We know that HSC mobilization occur concomitantly then predominantly on lymphocytes and monocytes and not on neutrophils when plasmatic levels of SDF-1 $\alpha$  increased, for instance after adenovirus mediated overexpression (76), or following an *in vivo* administration of SDF-1 $\alpha$  analog (77), or an administration of sulphated polysaccharides (78).

#### 4. Hematopoietic stem cells homing to the bone marrow and mobilization to the peripheral blood

SDF-1 $\alpha$  and CXCR4 implication in the homing of hematopoietic stem and progenitor (HPC) cells in the bone marrow, as well as in their mobilization to peripheral blood is generally induced to facilitate stem cells collection. The peripheral blood, is now well established (1,22,68). Homing of circulating collected stem cells are then used for autologous transplant, a process aiming HSC within the BM is a multi-step process in which hematopoietic cells or to restore the immune system destroyed after chemotherapy. Among the transplanted cells circulating in the peripheral blood are attracted and retained molecules that promote HSC mobilization, granulocyte colony-stimulating to the bone marrow by SDF-1 $\alpha$  gradient. SDF-1 $\alpha$  is secreted by bone marrow factor (G-CSF) is the most commonly used agent, both in patients and normal endothelial cells (BMEC) as well as by osteoblasts and stromal cells (25,69). donors. The molecular and cellular mechanisms leading to mobilization in SDF-1 $\alpha$  secreted by BMEC is presented on their cell surface via SDF-1/HIS response to G-CSF are again poorly understood, but G-CSF administration was interactions (51), which allow for optimal interactions between endothelial associated with a marked reduction in SDF-1 concentration in the bone cells and circulating CD34+ CXCR4+ hematopoietic stem cells. CXCR4 marrow, accompanied by N-terminus CXCR4 truncation (19,79). Moreover, engagement on CD34+ stem cells up-regulate adhesion molecules expression this decrease in SDF-1 levels was found to be caused by proteolytic such as VLA-4 (Very Late Antigen) and LFA-1 (Leucocyte Function- degradation of the molecule by neutrophil elastase in response to G-CSF (19). associated Antigen) (70), and mediate interactions with VCAM-1 (Vascular Cell Adhesion Molecule) AMD3100, a bicyclam molecule first described as a HIV inhibitor (5). Cell Adhesion Molecule) and ICAM-1 (InterCellular Adhesion Molecule) AMD3100 is a specific CXCR4 inhibitor, which binds to the second constitutively expressed on the endothelium, resulting in HSC rolling and firm AMD3100 is a specific CXCR4 inhibitor, which binds to the second adhesion (70). The arrested CD34+ cells can localize then to the bone marrow extracellular loop of CXCR4, thus preventing SDF-1 binding and CXCR4 niches through endothelial transmigration mechanism. SDF-1 $\alpha$  secreted by activity. Blocking the SDF-1/CXCR4 interactions in the bone marrow by osteoblasts and stromal cells is retained in the bone marrow environment AMD3100 administration was shown to be effective for hematopoietic stem thanks to its binding to local HSPG. This particular localization of SDF-1 $\alpha$  on cells mobilization, both when injected alone (80-82) or in combination with G-HS within the bone marrow allow for establishment of chemotactic gradients, CSF (83,84).

which polarizes the migration of stem cells up to the hematopoietic niches (71). SDF-1 $\alpha$  can also induce an increase in actin polymerization for an 5. Tissue regeneration and repair efficient migration of CD34+ cells to their specific location (72). Moreover, SDF-1 and CXCR4 have been shown to be implicated in tissue SDF-1 $\alpha$  is implicated in the anchoring and retention of the stem cells within regeneration, mediating the stem cell driven repair process, by regulating tissue the niches by up-regulating adhesion molecules for interaction of the CD34+ specific homing, proliferation and differentiation of progenitor cells following cells with osteoblasts (22). Among the adhesion molecules implicated in the tissue damage.

cancer models (95). High expression of SDF-1 was also found in particular tissues such as bone which are the preferential sites for the metastasis spreading of breast and prostate tumors. Moreover, *in vitro* studies using various tumor cell lines demonstrated the critical role of SDF-1 and CXCR4 in the proliferation, migration and invasion of malignant cells (1).

It is now well established that tumor progression and metastasis development are under influence of the tumor environment. Stromal tumor microenvironment is a rich and complex structure, including extracellular matrix components, soluble growth factors, cytokines and chemokines, as well as non epithelial cell types, i.e. endothelial cells, pericytes, leucocytes and fibroblasts. A recent study by *Oryno et al* reports that carcinoma associated fibroblasts (96) actively synthesized and secreted SDF-1, which was shown to promote the growth of tumor cells by activating CXCR4, as well as the tumor angiogenesis by inducing the recruitment of endothelial progenitor cells (96). To survive, tumor cells require indeed sufficient supply of oxygen and nutrients to maintain their own growth and to spread by metastasis. Therefore the role of SDF-1 in the formation of new blood vessels could appear critical for the tumor development.

## 2. Rheumatoid arthritis

Rheumatoid arthritis (RA) is a chronic inflammatory disease, with persistent activation of the immune system in the synovial joints, resulting in tissue destruction of bone and cartilage. Several reports have suggested a role for SDF-1 in RA development, and SDF-1 mRNA expression as well as SDF-1 protein secretion were detected in RA synovial fibroblasts (97). SDF-1 was shown to control the development of RA at different stage of the inflammatory process. This chemokine can be involved in the accumulation of CD4+ T cells in the joint (97), by mediating T cell recruitment to the inflammatory sites, i.e. by mediating chemoattraction and upregulation of adhesion molecules as described for normal trafficking of leucocytes. Moreover, the matrix metalloproteinases are proteolytic enzymes shown to be involved in the degradation of the extracellular matrix and cartilage in the joint in RA (98). SDF-1 also induces the production of these MMPs and their activation by synovial fibroblasts from RA patients, thus participating actively in bone destruction (98).

SDF-1 $\alpha$  displayed on the endothelium was also shown to induce angiogenesis (99), then favouring the arrival of the lymphocytes to the synovial tissue. A critical role of the SDF-1 $\alpha$ /HS interactions in this process was demonstrated by *Pablos* and co-workers in different studies, with HS immobilizing the synovioocyte-released SDF-1 $\alpha$  on the endothelial cells surface (52,99). The authors suggested that this HS-dependent immobilization of SDF-1 on endothelial surfaces offer new approach opportunities for therapeutic intervention.

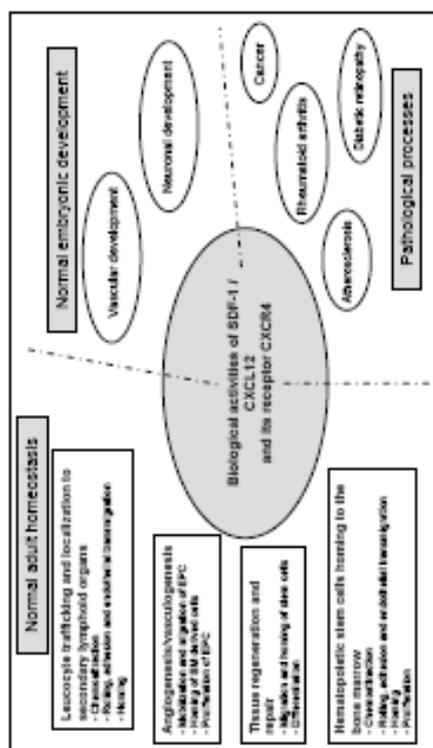


Figure 1. Main biological activities of the SDF-1 and CXCR4 ligand-receptor pair.

## V. Pathophysiological processes implicating SDF-1

In addition to their implication in normal development and homeostasis, SDF-1 and CXCR4 play an important role in several pathological processes, with main involvement in cancer and inflammatory diseases such as rheumatoid arthritis (Fig.1).

### 1. Cancer

SDF-1 $\alpha$  has been now clearly implicated in cancer and metastasis, promoting tumor progression and directing the migration of metastatic cancer cells to specific tissues (1,33). Numerous studies have demonstrated a significant role of CXCR4 expression in metastasis. CXCR4 expression was studied in numerous cell lines as well as in tumor tissues, and CXCR4 was there mostly found to be overexpressed (1,33). CXCR4 has been shown to be up-regulated in many cancers in both hematopoietic malignancies (such as acute myeloid leukaemia (AML) (85), multiple myeloma (86), chronic myelogenous leukaemia (87), non-Hodgkin lymphoma (88)), and non-hematopoietic malignancies (such as gliomas (89), breast tumors (90), prostate cancer (91) and small cell lung cancer (92)). Growth factors such as TGF- $\beta$  (Tumor Growth Factor) in ovary cells (93) or VEGF (Vascular Endothelial Growth Factor) in breast cancer cells (94) have been shown to be involved in the upregulation of CXCR4 in these cancers. Moreover, transformation of tumor cells to overexpress CXCR4 can accelerate tumor growth and invasion, promote tumor vascularization and accelerate metastasis in prostate and breast

### 3. Diabetic retinopathy

Diabetic retinopathy is the leading cause of blindness in adults, and is caused by oxygen starvation in the retina, leading to aberrant formation of blood vessels that destroy retinal architecture. An implication of SDF-1 in diabetic retinopathy was also evocated, as SDF-1 concentration in the vitreous increases as proliferative diabetic retinopathy progresses (100), helping in the recruitment of hematopoietic and endothelial progenitor cells. Moreover, anti-SDF-1 antibodies blockade in a murine model of proliferative adult retinopathy prevents retinal neovascularization (100).

### 4. Atherosclerosis

Another pathological process involving SDF-1 is atherosclerosis, a progressive inflammatory disease where different cells, lipids and extracellular matrix accumulate in the arterial wall. Leucocytes, i.e. lymphocytes and macrophages, and platelets are known to be important sources of cytokines and growth factors involved in the pathological process. As part of these mediators, SDF-1 was shown to be highly expressed in endothelial cells and macrophages of human atherosclerotic plaques (101), and to induce platelet aggregation (101).

## VI. GAG-based strategies to combat diseases by interference with SDF-1 activities?

Availability of specific inhibitors of the SDF-1/CXCR4 interactions has allowed for the demonstration that targeting this monogamous receptor-ligand pair is relevant for new therapy development (Table 1).

Table 1. Status of SDF-1 and CXCR4 inhibitors in development.

Molecule	Nature	Clinical indication	Clinical trials
<b>SDF-1 analogs</b>			
CTCE-0214	Small peptide SDF-1 agonist	- Hematopoietic stem cells mobilization - Cancer	Phase I. Chemokine Therapeutics Phase II/III. Chemokine Therapeutics
CTCE-9908	Small peptide SDF-1 antagonist		
<b>CXCR4 inhibitors</b>			
AMD3100	Bicyclicam	- Hematopoietic stem cells mobilization - HIV infection	Phase III. AzoMed
T140	Small peptide, polyphenolamine analog	- HIV infection - Cancer - Rheumatoid arthritis	Pre-clinical experiments
Monoclonal antibody	Anti-CXCR4 antibody		Pre-clinical experiments

As part of the inhibitors now available, AMD3100 is the more used and efficient molecule, firstly developed in HIV infection treatment, and now tested for hematopoietic stem cells mobilization (5). Whereas CXCR4 inhibitors or SDF-1 agonists directly target the SDF-1/CXCR4 interaction by inhibiting SDF-1 binding on CXCR4 and its subsequent activation, new molecules targeting the SDF-1/HS interactions are now thought to have potential efficacy.

First track for therapeutic intervention targeting SDF-1/HS come from studies using heparin administration in cancer models. Since the first experiments in the 1960's, numerous studies report inhibitory effects of heparin on growth and metastasis of tumor cells in various experimental models (102-104). Heparin used for its anti-coagulant activity was also shown to ameliorate survival of cancer patients. The mechanisms of action by which heparin mediate its effects, outside of its anti-coagulant part, are still unclear but data suggest that heparin may inhibit tumor cells attachment to extracellular matrix components by competing with HSPG. Moreover, unfractionated heparin has been reported to interfere also with the immune system in inflammatory diseases (105), as well as with angiogenesis (102). As it was shown in various *in vitro* and *in vivo* experiments, heparin exhibits a broad spectrum of immunomodulating and anti-inflammatory properties, by inhibiting the recruitment of leucocytes and reducing pro-inflammatory cytokines.

Sweeney *et al* have more recently demonstrated that use of sulphated glycans may represent a new method for inducing hematopoietic stem cells mobilization, as *in vivo* administration of fucoidan induces rapid stem and progenitor cells mobilization (106,107). Another work by Sweeney *et al* showed that this sulphated polysaccharides-induced mobilization is associated with increased plasma levels of SDF-1 both in mice and monkeys, with concomitant SDF-1 decrease in the bone marrow (78). The authors also clearly showed that the sulphated glycans specifically displace the SDF-1 pool normally retained in the BM by interaction with HS, conducting to its release into the peripheral blood. The consequent increase in both leucocytes and stem and progenitor cells observed in the blood of mice and non human primates suggests that the released SDF-1 chemokine remains functional.

## Conclusion

As part of the vast family of chemoattractant cytokines, also named chemokines, SDF-1 $\alpha$  regulates leucocyte trafficking, as well as homing and attraction of numerous cell types, both in normal development and diseases. SDF-1  $\alpha$  is now also known to be involved in new blood vessel formation by directing the migration of endothelial progenitor cells, and this effect on

vasculogenesis implicate SDF-1 on various pathological processes. Inhibition of the SDF-1/CXCR4 axis has recently been proved to be a relevant strategy for stem cell mobilization to the peripheral blood, and SDF-1 analogs or CXCR4 inhibitors are currently under investigation for different clinical applications. The increasing evidence for a critical role of the SDF-1 binding to heparan sulphate proteoglycans for efficient biological activity of the chemokine now offers new opportunities for therapeutic intervention. *In vivo* studies by Sweeney *et al* using administration of sulphated polysaccharides (Fucoidan) clearly demonstrated the relevance of such strategy. Determination of the sulfation pattern implicated in the binding of SDF-1 to HS will probably help to design more specific inhibitory molecules than those observed in natural products such as heparin or fucoidan. Finally, the development of safe and efficient sulphated compounds will also help to understand the specific interaction of the targeted proteins, i.e growth factors, cytokines or chemokines, and heparan sulphate chains.

## References

- Burger, J.A., and Kipps, T.J. 2006. *Blood*, 107, 1761.
- Sadir, R., Imbery, A., Baleux, F., and Lorrain-Jacob, H. 2004. *J Biol Chem*, 279, 43854.
- Lorrain-Jacob, H., Grosdidier, A., and Imbery, A. 2002. *Proc Natl Acad Sci U S A*, 99, 1229.
- Tanamura, H., and Fujii, N. 2005. *Expert Opin Ther Targets*, 9, 1267.
- De Clercq, E. 2005. *Mini Rev Med Chem*, 5, 805.
- Teshiro, K., Tada, H., Heilker, R., Shirozu, M., Nakano, T., and Honjo, T. 1993. *Science*, 261, 600.
- Nagasawa, T., Kikutani, H., and Kishimoto, T. 1994. *Proc Natl Acad Sci U S A*, 91, 2305.
- Jiang, W., Zhou, P., Kahn, S.M., Tomita, N., Johnson, M.D., and Weinstein, I.B. 1994. *Exp Cell Res*, 215, 284.
- Moser, B., and Willmann, K. 2004. *Ann Rheum Dis*, 63, i84.
- Chen, J.F., and Rausohoff, R.M. 2006. *N Engl J Med*, 354, 610.
- Shirozu, M., Nakano, T., Inazawa, J., Tashiro, K., Tada, H., Shimohara, T., and Honjo, T. 1995. *Genomics*, 28, 495.
- Yu, L., Cecil, J., Peng, S.B., Schremm, J., Kovacevic, S., Paul, D., Su, E.W., and Wang, J. 2006. *Gene*, 18, 18.
- Geichmann, M., Gillen, C., Czardybon, M., Bossa, F., Greiner-Petter, R., Auer, J., and Muller, H.W. 2000. *Eur J Neurosci*, 12, 1857.
- Pillarisetti, K., and Gupta, S.K. 2001. *Inflammation*, 25, 293.
- Crump, M.P., Gong, J.H., Loetscher, P., Rajaraman, K., Amara, A., Arenzana-Seisdedos, F., Virelizier, J.L., Baggiolini, M., Sykes, B.D., and Clark-Lewis, I. 1997. *Embo J*, 16, 6996.
- Velickamp, C.T., Peterson, F.C., Peizek, A.J., and Volkman, B.F. 2005. *Protein Sci*, 14, 1071.
- Denhys, C., Fernandez, E.J., Thompson, D.A., Simon, R.J., Siani, M.A., and Lolis, E. 1998. *Proc Natl Acad Sci U S A*, 95, 6641.
- Valenzuela-Fernandez, A., Planchena, T., Baleux, F., Staropoli, I., Le-Barillic, K., Leduc, D., Delannay, T., Lazarini, F., Virelizier, J.L., Chignard, M., Pidard, D., and Arenzana-Seisdedos, F. 2002. *J Biol Chem*, 277, 15677.
- Petit, I., Szyper-Kravitz, M., Nagler, A., Lahav, M., Peled, A., Habler, L., Ponomarev, T., Taichman, R.S., Arenzana-Seisdedos, F., Fujii, N., Sandbank, J., Zipori, D., and Lapidot, T. 2002. *Nat Immunol*, 3, 687.
- Rao, R.M., Beitz, T.V., Lamont, D.J., Kim, M.B., Shaw, S.K., Frolo, R.M., Baleux, F., Arenzana-Seisdedos, F., Alon, R., and Luskinas, F.W. 2004. *J Exp Med*, 200, 713.
- Delgado, M.B., Clark-Lewis, I., Loetscher, P., Langen, H., Thelen, M., Baggiolini, M., and Wolf, M. 2001. *Eur J Immunol*, 31, 699.
- Kollet, O., Dar, A., Shirval, S., Kalkbrenner, A., Lapid, K., Szranberg, Y., Tesio, M., Samstein, R.M., Goichberg, P., Spiegel, A., Elson, A., and Lapidot, T. 2006. *Nat Med*, 21, 21.
- Lambre, A.M., Proost, P., Durinx, C., Bal, G., Senten, K., Augustyns, K., Scharpe, S., Van Damme, J., and De Meester, I. 2001. *J Biol Chem*, 276, 29839.
- McQuibban, G.A., Butler, G.S., Gong, J.H., Bendall, L., Power, C., Clark-Lewis, I., and Overall, C.M. 2001. *J Biol Chem*, 276, 43503.
- Bleul, C.C., Farzan, M., Choe, H., Parolin, C., Clark-Lewis, I., Sodroski, J., and Springer, T.A. 1996. *Nature*, 382, 829.
- Oberlin, E., Amara, A., Bachelier, F., Bessia, C., Virelizier, J.L., Arenzana-Seisdedos, F., Schwartz, O., Heard, J.M., Clark-Lewis, I., Legler, D.F., Loetscher, M., Baggiolini, M., and Moser, B. 1996. *Nature*, 382, 833.
- Feng, Y., Broder, C.C., Kennedy, P.E., and Berger, E.A. 1996. *Science*, 272, 872.
- Kristiansen, K. 2004. *Pharmacol Ther*, 103, 21.
- Nagasawa, T., Hirota, S., Tachibana, K., Takakura, N., Nishikawa, S., Kinoshita, Y., Yoshida, N., Kitamura, H., and Kishimoto, T. 1996. *Nature*, 382, 635.
- Balabanian, K., Leguina, B., Infarinso, S., Chow, K.Y., Harraguet, J., Moepps, B., Arenzana-Seisdedos, F., Thelen, M., and Bachelier, F. 2005. *J Biol Chem*, 280, 35760.
- Ainri, A., Webb, J.J., Bleul, C., Springer, T., and Gutierrez-Ramos, J.C. 1997. *J Exp Med*, 185, 111.
- Zou, Y.F., Kortmann, A.H., Kuroda, M., Tanuchi, I., and Littman, D.R. 1998. *Nature*, 393, 595.
- Balkwill, F. 2004. *Semin Cancer Biol*, 14, 171.
- Kapas, S., and Clark, A.J. 1995. *Biochem Biophys Res Commun*, 217, 832.
- Delehedde, M. 2002. *Recent Res Devel Biol Chem*, 1, 133.
- Delehedde, M., Lorrain-Jacob, H., Gallagher, J., Bonaffe, D., Adam, E., Querolle, O., Lequien, S., Degove, S., Lassalle, P., and Becharard, D. 2005. *Current Med Chem*, 2, 345.
- Whitelock, J.M., and Iozzo, R.V. 2005. *Chem Rev*, 105, 2745.
- Bernfield, M., Gotte, M., Park, P.W., Reizes, O., Fitzgerald, M.L., Lincoff, J., and Zako, M. 1999. *Annu Rev Biochem*, 68, 729.

39. Faham, S., Hileman, R.E., Fromm, J.R., Linhardt, R.J., and Rees, D.C. 1996. *Science*, 271, 1116.
40. Nurcombe, V., Smart, C.E., Chipperfield, H., Cool, S.M., Bouilly, B., and Hondermarck, H. 2000. *J Biol Chem*, 275, 30009.
41. Tanaka, Y., Adams, D.H., and Shaw, S. 1993. *Immunol Today*, 14, 111.
42. Lortat-Jacob, H., Balzer, F., and Grimaud, J.A. 1996. *J Biol Chem*, 271, 16139.
43. Hoogewarf, A.J., Kuscher, G.S., Proudfoot, A.E., Borlat, F., Clark-Lewis, I., Power, C.A., and Wells, T.N. 1997. *Biochemistry*, 36, 13570.
44. McCormack, M.A., Boren, D.M., and LiWang, P.J. 2004. *Biochemistry*, 43, 10090.
45. Kuscher, G.S., Coulin, F., Power, C.A., Proudfoot, A.E., Hubbard, R.E., Hoogewarf, A.J., and Wells, T.N. 1999. *Biochemistry*, 38, 12659.
46. Proudfoot, A.E., Handel, T.M., Johnson, Z., Lau, E.K., LiWang, P., Clark-Lewis, I., Borlat, F., Wells, T.N., and Kosco-Vilbois, M.H. 2003. *Proc Natl Acad Sci U S A*, 100, 1885.
47. Amara, A., Lorthioir, O., Valenzuela, A., Magerus, A., Thelen, M., Montes, M., Vretilizer, J.L., Deleplanne, M., Balleux, F., Lortat-Jacob, H., and Arenzana-Seisdedos, F. 1999. *J Biol Chem*, 274, 23916.
48. Mbenba, E., Gluckman, J.C., and Gattegno, I. 2000. *Glycobiology*, 10, 21.
49. Sadrir, R., Balleux, F., Grosdidier, A., Imberty, A., and Lortat-Jacob, H. 2001. *J Biol Chem*, 276, 8288.
50. Uchimura, K., Morimoto-Tomin, M., Bistrup, A., Li, J., Lyon, M., Gallagher, J., Werb, Z., and Rosen, S.D. 2006. *BMC Biochem*, 7, 2.
51. Netelebos, T., van den Born, J., Kessler, F.L., Zweegman, S., Merie, P.A., van Oostrveen, J.W., Zwarginga, J.J., Huijgens, P.C., and Drager, A.M. 2003. *Leukemia*, 17, 175.
52. Santiago, B., Balleux, F., Palao, G., Gutierrez-Canaas, I., Ramirez, J.C., Arenzana-Seisdedos, F., and Pablos, J.L. 2006. *Arthritis Res Ther*, 8, R43.
53. Yoo, I., Salvucci, O., Cardenas, A.R., Hwang, S.T., Aoki, Y., De La Luz Sierra, M., Sajewicz, A., Pittaluga, S., Yarchona, R., and Tosato, G. 2003. *Blood*, 102, 3900.
54. Netelebos, T., Zuijderduijn, S., Van Den Born, J., Kessler, F.L., Zweegman, S., Huijgens, P.C., and Drager, A.M. 2002. *J Leukoc Biol*, 72, 353.
55. Raisz, K., Mendem, R., Sievers, J., and Harman, D. 2002. *Neuroscience*, 115, 295.
56. Valenzuela-Fernandez, A., Palanca, T., Amara, A., Magerus, A., Altmeyer, R., Delaunay, T., Vretilizer, J.L., Balleux, F., Galzi, J.L., and Arenzana-Seisdedos, F. 2001. *J Biol Chem*, 276, 26550.
57. Hamon, M., Mbenba, E., Charneau, N., Slimani, H., Brule, S., Saifir, L., Vassy, R., Prost, C., Ljevre, N., Starzec, A., and Gattegno, I. 2004. *Glycobiology*, 14, 311.
58. Luo, J., Luo, Z., Zhou, N., Hall, J.W., and Huang, Z. 1999. *Biochem Biophys Res Commun*, 264, 42.
59. Delebede, M., Lyon, M., Gallagher, J.T., Rudland, P.S., and Fernig, D.G. 2002. *Biochem J*, 366, 235.
60. Ma, Q., Jones, D., Borghesani, P.R., Segal, R.A., Nagasawa, T., Kishimoto, T., Bronson, R.T., and Springer, T.A. 1998. *Proc Natl Acad Sci U S A*, 95, 9448.
61. Tachibana, K., Hirota, S., Izasa, H., Yoshida, H., Kawabata, K., Kataoka, Y., Kitamura, Y., Matsushima, K., Yoshida, N., Nishikawa, S., Kishimoto, T., and Nagasawa, T. 1998. *Nature*, 393, 391.
62. Asahara, T., Murohara, T., Sullivan, A., Silver, M., van der Zee, R., Li, T., Witzenbichler, B., Schemm, G., and Isner, J.M. 1997. *Science*, 275, 964.
63. Ceradini, D.J., Kulkarni, A.R., Callaghan, M.J., Tepper, O.M., Bastidas, N., Kleinman, M.E., Capla, J.M., Galiano, R.D., Levine, J.P., and Gurtner, G.C. 2004. *Nat Med*, 10, 858.
64. Grunewald, M., Avraham, I., Dor, Y., Bachar-Lustig, E., Inn, A., Yung, S., Chimenti, S., Landsman, L., Abramovitch, R., and Keshet, E. 2006. *Cell*, 124, 175.
65. D'Apuzzo, M., Rolink, A., Loetscher, M., Hoxie, J.A., Clark-Lewis, I., Melchers, F., Baggiolini, M., and Moser, B. 1997. *Eur J Immunol*, 27, 1788.
66. Bleul, C.C., Schmitz, J.L., and Springer, T.A. 1998. *J Exp Med*, 187, 753.
67. Bleul, C.C., Fuhrbringer, R.C., Casasnovas, J.M., Alant, A., and Springer, T.A. 1996. *J Exp Med*, 184, 1101.
68. Jung, Y., Wang, J., Schneider, A., Sun, Y.X., Koh-Paige, A.J., Osman, N.I., McCauley, L.K., and Teichman, R.S. 2005. *Bone*, 4, 4.
69. Ponomarev, T., Peled, A., Petit, I., Taichman, R.S., Habler, L., Sandbank, J., Arenzana-Seisdedos, F., Magerus, A., Caruz, A., Fujii, N., Nagler, A., Lahav, M., Szyper-Kravitz, M., Zipori, D., and Lapidot, T. 2000. *J Clin Invest*, 106, 1331.
70. Peled, A., Grabovsky, V., Habler, L., Sandbank, J., Arenzana-Seisdedos, F., Petit, I., Ben-Hur, H., Lapidot, T., and Aloni, R. 1999. *J Clin Invest*, 104, 1199.
71. Netelebos, T., van den Born, J., Kessler, F.L., Zweegman, S., Huijgens, P.C., and Drager, A.M. 2003. *J Leukoc Biol*, 74, 1035.
72. Voerms, C., Anthony, E.C., Mul, E., van der Schoot, E., and Hordijk, P. 2001. *Exp Hematol*, 29, 1456.
73. Avigdor, A., Goitberg, P., Shivrit, S., Dor, A., Peled, A., Samra, S., Kollet, O., Harshkovitz, R., Aloni, R., Hardan, I., Ben-Hur, H., Naor, D., Nagler, A., and Lapidot, T. 2004. *Blood*, 103, 2981.
74. Broxmeyer, H.E., Cooper, S., Kohli, L., Hangoc, G., Lee, Y., Mantel, C., Clapp, D.W., and Kim, C.H. 2003. *J Immunol*, 170, 421.
75. Broxmeyer, H.E., Kohli, L., Kim, C.H., Lee, Y., Mantel, C., Cooper, S., Hangoc, G., Shaheen, M., Li, X., and Clapp, D.W. 2003. *J Leukoc Biol*, 73, 630.
76. Hattori, K., Heissig, B., Tashiro, K., Houjo, T., Taniuchi, M., Shieh, J.H., Hackett, N.R., Quitoriano, M.S., Crystal, R.G., Rafii, S., and Moore, M.A. 2001. *Blood*, 97, 3354.
77. Pelus, L.M., Bian, H., Fukuda, S., Wong, D., Merzouk, A., and Salari, H. 2005. *Exp Hematol*, 33, 295.
78. Sweeney, E.A., Lortat-Jacob, H., Priestley, G.V., Nakamoto, B., and Papayanopoulos, T. 2002. *Blood*, 99, 44.
79. Levesque, J.P., Hendy, J., Takamatsu, Y., Simmons, P.J., and Bendall, L.J. 2003. *J Clin Invest*, 111, 187.
80. Broxmeyer, H.E., Orsichel, C.M., Clapp, D.W., Hangoc, G., Cooper, S., Platt, P.A., Liles, W.C., Li, X., Graham-Evans, B., Campbell, T.B., Calandra, G., Bridger, G., Dale, D.C., and Strou, E.F. 2005. *J Exp Med*, 201, 1307.
81. Liles, W.C., Broxmeyer, H.E., Rodger, E., Wood, B., Hubel, K., Cooper, S., Hangoc, G., Bridger, G.J., Henson, G.W., Calandra, G., and Dale, D.C. 2003. *Blood*, 102, 2738.

82. Devine, S.M., Flomenberg, N., Vesole, D.H., Liesveld, J., Weisdorf, D., Bode, K., Calandra, G., and DiPersio, J.F. 2004, *J Clin Oncol*, 22, 1095.
83. Flomenberg, N., Devine, S.M., Dipersio, J.F., Liesveld, J.L., McCarty, J.M., Rowley, S.D., Vesole, D.H., Bode, K., and Calandra, G. 2005, *Blood*, 106, 1867.
84. Liles, W.C., Rodeger, E., Broxmeyer, H.E., Dehner, C., Bode, K., Calandra, G., Christensen, J., Wood, B., Price, T.H., and Dale, D.C. 2005, *Transfusion*, 45, 295.
85. Mothle, R., Schirrenhelm, M., Fallenschmid, C., Bantz, F., Kratz-Albers, K., Serve, H., Brugger, W., and Kanz, L. 2000, *Br J Haematol*, 110, 563.
86. Moller, C., Stromberg, T., Juremalin, M., Nilsson, K., and Nilsson, G. 2003, *Leukemia*, 17, 203.
87. Peled, A., Hardan, I., Trakhenbrot, L., Gur, E., Magid, M., Darash-Yahana, M., Cohen, N., Grabovsky, V., Framizra, S., Kollit, O., Lider, O., Aloni, R., Rechavi, G., and Lapidot, T. 2002, *Stem Cells*, 20, 259.
88. Weng, A.P., Shalshafat, A., and Dorfman, D.M. 2003, *Am J Clin Pathol*, 119, 424.
89. Rempel, S.A., Dudas, S., Ge, S., and Gutierrez, J.A. 2000, *Clin Cancer Res*, 6, 102.
90. Muller, A., Homey, B., Soto, H., Ge, N., Cao, D., Cameron, D., Buchanan, M.E., McClanahan, T., Murphy, E., Yuan, W., Wagner, S.N., Barrera, J.L., Mohr, A., Verastegui, E., and Zlotnik, A. 2001, *Nature*, 410, 50.
91. Sun, Y.X., Wang, J., Shelburne, C.E., Lopatin, D.E., Chinnaiyan, A.M., Rubin, M.A., Pientz, K.J., and Taichman, R.S. 2003, *J Cell Biochem*, 89, 462.
92. Phillips, R.J., Burdick, M.D., Lutz, M., Belperio, J.A., Keane, M.P., and Strieter, R.M. 2003, *Am J Respir Crit Care Med*, 167, 1676.
93. Scotton, C.J., Wilson, J.L., Scott, K., Stamp, G., Wilbanks, G.D., Fricker, S., Bridger, G., and Balkwill, F.R. 2002, *Cancer Res*, 62, 5930.
94. Bachelier, R.E., Wendt, M.A., and Mercurio, A.M. 2002, *Cancer Res*, 62, 7203.
95. Darash-Yahana, M., Pkarsky, E., Abramovitch, R., Zeira, E., Pal, B., Karpus, R., Beider, K., Avniel, S., Kosem, S., Galun, E., and Peled, A. 2004, *Faseb J*, 18, 1240.
96. Orimo, A., Gupta, P.B., Sgroi, D.C., Arezana-Seisdedos, F., Delaunay, T., Nosen, R., Carey, V.J., Richardson, A.L., and Weinberg, R.A. 2005, *Cell*, 121, 335.
97. Nanki, T., Hayashida, K., El-Gabalawy, H.S., Suson, S., Shi, K., Girschick, H.J., Yavuz, S., and Lipsky, P.E. 2000, *J Immunol*, 165, 6590.
98. Garcia-Vicuna, R., Gomez-Gavino, M.V., Dominguez-Luis, M.J., Pec, M.K., Gonzalez-Alvaro, I., Alvaro-Gracia, J.M., and Diaz-Gonzalez, F. 2004, *Arthritis Rheum*, 50, 3866.
99. Peblos, J.L., Sainago, B., Galindo, M., Torres, C., Brelmer, M.T., Blanco, F.J., and Garcia-Lazaro, F.J. 2003, *J Immunol*, 170, 2147.
100. Butler, J.M., Guthrie, S.M., Koc, M., Afzal, A., Caballero, S., Brooks, H.L., Mames, R.N., Segal, M.S., Grant, M.B., and Scott, E.W. 2005, *J Clin Invest*, 115, 86.
101. Abi-Younes, S., Saury, A., Mach, F., Sukhova, G.K., Libby, P., and Luster, A.D. 2000, *Circ Res*, 86, 131.
102. Smorenburg, S.M., and Van Noorden, C.J. 2001, *Pharmacol Rev*, 53, 93.
103. Ornstein, D.L., and Zacharski, L.R. 2002, *Clin Hemorheol Microcirc*, 26, 33.
104. Kragh, M., and Loechel, F. 2005, *Int J Oncol*, 27, 1159.
105. Tyrrell, D.J., Horne, A.P., Holme, K.R., Preuss, J.M., and Page, C.P. 1999, *Adv Pharmacol*, 46, 151.
106. Sweeney, E.A., Priestley, G.V., Nakamoto, B., Collins, R.G., Beaudet, A.L., and Papayannopoulou, T. 2000, *Proc Natl Acad Sci U S A*, 97, 6544.
107. Fresno, P.S., and Weiss, L. 2000, *Blood*, 96, 2460.

## RESUME

Les protéoglycannes exercent de nombreuses fonctions par le biais de leur partie protéique ou de leurs glycosaminoglycannes. La caractérisation des interactions entre les glycosaminoglycannes et des protéines a ouvert de larges champs d'applications. Dans ce cadre, deux thèmes de recherche ont été développés.

Premièrement, nous nous sommes intéressés à un nouveau protéoglycanne appelé endocan. Le développement des capacités de production et de purification de cette macromolécule, nous a permis par différentes approches de déterminer le profil structural de sa chaîne glycanique et de sa partie protéique, mais aussi d'étudier les interactions avec plusieurs de ses partenaires protéiques dont l'interféron  $\gamma$  et l'hépatocyte growth factor, impliqués respectivement dans l'inflammation et le développement tumoral. Parallèlement, une étude structurale et fonctionnelle de l'interaction entre l'interféron  $\gamma$  et des glycosaminoglycannes de type héparanes sulfates a conduit au développement de mimes oligosaccharidiques obtenus par synthèse chimique. Parmi ces molécules, certaines permettent de moduler *in vitro* l'activité de la cytokine, et constituent une base possible pour le développement de nouveaux médicaments.

## ABSTRACT

Proteoglycans are complex macromolecules which consist structurally of a core protein and one or more covalently linked glycosaminoglycan chains. Both components can interact with a variety of binding proteins known to be involved in various pathophysiological processes such as inflammation and cancer. The characterization of the specific interactions between glycosaminoglycans and their binding partners should raise new therapeutic opportunities in drug development.

In this context, we first studied a recently discovered proteoglycan called endocan. By developing new processes for high yield production and purification, we have been able to further characterise the structure of endocan using different biochemical approaches and also to study the interactions with various binding partners including interferon  $\gamma$  and hepatocyte growth factor. In the same time, the characterisation of the specific interaction of interferon  $\gamma$  and the particular glycosaminoglycan of heparan sulphate type allowed the synthesis by sugar chemistry of glycosaminoglycan mimetics. Among these molecules, a glycoconjugate modulates the activity of the cytokine and could lead to the development of a new type of drugs.