



HAL
open science

Analyse par IRMPD de polysaccharides de lichen contenant des unités galactofuranose

Oznur Yeni

► **To cite this version:**

Oznur Yeni. Analyse par IRMPD de polysaccharides de lichen contenant des unités galactofuranose. Chimie-Physique [physics.chem-ph]. Université Claude Bernard - Lyon I, 2023. Français. NNT : 2023LYO10026 . tel-04481941

HAL Id: tel-04481941

<https://theses.hal.science/tel-04481941>

Submitted on 28 Feb 2024

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.



THESE de DOCTORAT DE L'UNIVERSITE CLAUDE BERNARD LYON 1

**Ecole Doctorale N°52
Physique et Astrophysique de Lyon**

Discipline : Physique Chimie

Soutenue publiquement le 06/03/2023, par :
Oznur Yeni

Analyse par IRMPD de polysaccharides de lichen contenant des unités galactofuranose

Devant le jury composé de :

SALVADOR Arnaud, PU, Université Lyon 1

Président

SCUDERI Debora, MCF-HDR, Université Paris Saclay
BARDOR Muriel, PU, Université de Rouen Normandie
FERRIERES Vincent, PU, Université de Rennes/ENSCR

Rapporteure
Rapporteure
Examineur

COMPAGNON Isabelle, MCF-HDR, Université Lyon 1

Directrice de thèse

Remerciements

Je tiens à remercier toutes les personnes qui ont permis de près ou de loin à l'aboutissement de cette thèse.

Je voudrais commencer par exprimer ma gratitude aux membres du jury : Mr Arnaud Salvador, Mr Vincent Ferrières, Mme Muriel Bardor et Mme Debora Scuderi. Je remercie particulièrement ces deux dernières pour la relecture du manuscrit, leurs rapports et les corrections apportées.

Le plus grand merci revient à Isabelle pour sa confiance, sa bienveillance, sa pédagogie, son soutien et ses précieux conseils. Tu m'as tellement apporté durant ces quatre années que je ne te remercierai jamais assez. Mes remerciements vont également au physicien Franck Lépine. Merci à tous les deux de m'avoir fait confiance en toutes circonstances, y compris lorsque j'étais au volant.

Je voudrais ensuite chaleureusement remercier les stagiaires/doctorants/post-doctorants que j'ai côtoyé durant ma thèse. Baptiste 1^{er}, merci de m'avoir formée à l'utilisation de l'instrumentation et pour tes conseils. Gina, merci et plein de bonheur avec tes enfants. Marius, merci pour les gâteaux et ta bonne humeur. Ninja Alexie, arigatô ! Baptiste 2nd, merci pour toutes ces fois où j'ai eu besoin d'attraper des choses en hauteur. Alicia, merci de ta gentillesse. Merci également à Corentin, Yaowei, Sreelakshmi, Amira, Sarah, Clara et les stagiaires que j'ai pu croiser sur ma route.

Je remercie également les permanents de l'équipe pour leurs conseils, encadrement, soutien et disponibilité : Vincent, Richard, Saikat, Eric, Bruno, Christian ainsi que Gulabi. Ce fut un grand plaisir d'effectuer une thèse au sein de votre équipe. Merci aux membres des équipes administratives et techniques de l'ILM.

Je tiens à remercier les membres du consortium ALGAIMS : merci Vincent pour ce beau projet, Laurent, Bénédicte, Camille, Joël et Françoise, merci pour les échantillons sans lesquels cette thèse n'aurait pas été aussi fructueuse. Merci Allouche de m'avoir formée à la théorie et pour votre disponibilité. Merci également aux membres de l'INRAe, Simon, David et Hélène. Merci à tous pour les nombreuses discussions que l'on a eues, j'ai été ravie de travailler sur ce projet avec vous.

Enfin, mes derniers mots vont aux membres de ma famille : mes parents pour m'avoir soutenue durant toutes ces années, mes frères et mon mari pour leur aide. Isa, tu es la plus belle chose qui m'est arrivée durant ces quatre années, merci pour la joie que tu m'apportes chaque jour.

Le lichen

Recouvrant près de 8% des terres émergées, le lichen est un être qui devrait vous être familier. Il pousse là où aucune plante ne peut loger: sur les arbres, les trottoirs, les rochers... Et c'est même le premier signe de vie à occuper les coulées de lave à peine solidifiées.

Il est si petit, si discret qu'il nous semble insignifiant, sans intérêt. Or plus nous l'observons de près, et plus se révèlent ses beautés et ses secrets. En vous approchant vous verrez qu'un lichen, c'est l'association d'une algue et d'un champignon. Ces deux espèces ont décidé de vivre en union dans un monde où règne la division.

Cette bête des trottoirs, du galactofuranose elle peut concevoir.
« Qu'est-ce que c'est ? » me demande mon auditoire.
Si vous voulez le savoir, c'est le nom barbare que l'on donne à un sucre rare.

Sous la direction d'Isabelle, je suis partie à la quête de ce sucre sensationnel.

Ma mission ?

Détecter du galactofuranose dans des extraits naturels
et élucider ses mystères conformationnels.

Mes compagnons ?

La spectroscopie vibrationnelle pour les expérimentations
et la théorie de la densité fonctionnelle pour les simulations.

Ces deux méthodes se définissent par leur réciproque relation:

Autant l'une apporte à l'autre, autant l'autre le lui rend... comme chez le lichen.

Table des matières

Résumé.....	viii
Abstract	ix
Liste des sigles.....	x
Introduction.....	1
1 Généralités sur la spectroscopie moléculaire	1
2 Spectroscopie des biomolécules.....	1
3 Spectroscopie des sucres	2
4 Projet ALGAIMS.....	4
5 Bibliographie	7
CHAPITRE 1 : Méthodes expérimentales et théoriques	9
1 Généralités sur l'IRMPD	10
1.1 Historique	10
1.2 Principe d'une mesure de spectre IRMPD.....	11
1.3 Autres spectroscopies IR d'ions.....	14
2 Instrumentation	17
2.1 Généralités sur le couplage MS-IR.....	17
2.2 Sources d'ionisation : ESI et nanospray.....	18
2.3 Piège à ions.....	19
2.4 Laser infrarouge.....	19
2.5 Couplage de l'IRMPD avec les méthodes séparatives.....	20
3 Outils théoriques.....	22
3.1 Obtention des structures.....	22
3.2 Calculs des spectres vibrationnels.....	23
4 Interprétations.....	24
4.1 Etude conformationnelle.....	24
4.2 Caractérisation de sucres inconnus.....	26
5 Autres méthodes du consortium	29
5.1 Synthèse des composés modèles.....	29
5.2 Préparation des échantillons de lichen	30
5.3 Analyse par mobilité ionique.....	31
6 Bibliographie	32
CHAPITRE 2 : Identification par IRMPD de la taille du cycle du galactose sur de petits modèles synthétiques	33
1 Introduction	34
1.1 Isomérisation.....	34
1.2 Molécule neutre	34
1.3 Cycle flexible.....	35
1.4 Echantillon naturel	35
2 Détails expérimentaux	36
3 Détermination par IRMPD de la taille du cycle sur des modèles monosaccharidiques	37

3.1	Systèmes étudiés	37
3.2	Spectres MS/MS	38
3.3	Spectres IRMPD	39
3.4	Conclusion	42
4	Création d'une base de données avec des disaccharides possédant ou non du Galf	43
4.1	Systèmes étudiés	43
4.2	Spectres MS/MS	44
4.3	Spectres IRMPD	45
4.4	Conclusion	48
5	Conclusion	49
6	Bibliographie	50
CHAPITRE 3 : Impact du cycle et du micro-environnement sur la conformation du galactose en phase gazeuse		51
1	Introduction	52
1.1	Vocabulaire	52
1.2	La configuration	52
1.3	L'environnement	54
1.4	Etude des conformations par synergie théorie/expérience	55
2	Détails expérimentaux et théoriques	56
3	Impact de la taille du cycle sur la conformation du galactose	57
3.1	Une recherche conformationnelle similaire pour les deux cycles du galactose	57
3.2	Calculs des spectres vibrationnels	60
3.3	Interprétation	70
4	Impact du micro-environnement sur la conformation du galactofuranose	72
4.1	Préférences conformationnelles du galactofuranose neutre en phase gazeuse	72
4.2	Interaction du galactofuranose avec une molécule d'eau	77
4.3	Interaction du galactofuranose avec un ion	82
4.4	Interprétation	84
5	Conclusion	87
6	Bibliographie	88
CHAPITRE 4 : Analyse par MS-MS/IR d'oligosaccharides contenant du galactofuranose		89
1	Introduction	90
2	Détails expérimentaux	92
2.1	Séquençage	92
2.2	Criblage	92
3	Séquençage d'oligosaccharides de synthèse contenant du galactofuranose	94
3.1	Développement de la méthode de séquençage sur des trisaccharides synthétiques	94
3.2	Caractérisation d'un trisaccharide par analyse des fragments disaccharidiques	99
3.3	Méthode universelle pour déterminer la taille du cycle par analyse des fragments monosaccharidiques	107
4	Développement d'une méthode de criblage en LC-MS/MS-IR	110
4.1	Méthode de criblage par irradiation IR	110
4.2	Choix des paramètres LC-MS/MS-IR	110
4.3	Criblage par irradiation infrarouge non résonante	111

4.4	Criblage par irradiation infrarouge résonante	115
4.5	Criblage par irradiation infrarouge résonante à longueur d'onde spécifique au Galf....	117
4.6	Interprétation	120
5	Conclusion.....	121
6	Bibliographie	122
CHAPITRE 5 : Détection de galactofuranose dans des extraits de lichen dépolymérisés		123
1	Introduction	124
2	Méthodes	128
2.1	Dépolymérisation avec le protocole FITDOG	128
2.2	Dépolymérisation par hydrolyse acide.....	128
2.3	Analyse par LC-MS.....	128
2.4	Criblage par LC-MS/MS-IR	129
3	Dépolymérisation de polysaccharides modèles par FITDOG	130
3.1	Résultats MS.....	130
3.2	Résultats IRMPD	130
4	Dépolymérisation de polysaccharides modèles par hydrolyse acide	132
4.1	Résultats MS.....	132
4.2	Résultats IRMPD	132
5	<i>Cetraria islandica</i>	134
5.1	Présentation	134
5.2	Résultats de dépolymérisation.....	134
5.3	Séparation des oligosaccharides par LC-MS.....	135
5.4	Criblage des disaccharides avec le laser Firefly seul	136
5.5	Criblage des disaccharides avec le laser CO2 seul.....	137
6	<i>Lasallia pustulata</i>	139
6.1	Présentation	139
6.2	Résultats de dépolymérisation.....	139
6.3	Criblage des disaccharides avec le laser Firefly seul	140
7	<i>Roccella fuciformis</i>	142
7.1	Présentation	142
7.2	Résultats de dépolymérisation.....	142
7.3	Criblage des disaccharides avec laser Firefly seul	143
7.4	Criblage des disaccharides avec laser CO2 seul	143
8	Discussion et perspectives	145
9	Bibliographie	147
Conclusion et perspectives		148
Liste des publications		151

Résumé

Le galactose est l'un des sucres les plus abondants dans la nature. Alors qu'il existe principalement sous forme de cycle à 6 (pyranose), certaines espèces de bactéries, de champignons, de plantes et de lichens sont également capables de synthétiser le galactose avec un cycle à 5 membres (galactofuranose). Cette forme est très rare, elle requiert plus d'énergie pour sa biosynthèse et est très flexible comparée à son analogue pyranose.

Le projet multidisciplinaire ALGAIMS, qui réunit chimistes, lichénologues, massistes et spectroscopistes, a pour objectif de développer des outils pour l'analyse de polysaccharides de lichen contenant du galactofuranose.

Nous verrons dans cette thèse le développement d'une méthode basée sur l'IRMPD, couplage de spectrométrie de masse et de laser infrarouge, pour analyser le galactofuranose.

D'une part, l'IRMPD permet en combinaison avec des outils de chimie quantique de fournir des informations conformationnelles sur le galactofuranose.

D'autre part elle peut être utilisée à visée analytique pour discriminer et analyser des oligosaccharides contenant du galactofuranose. La méthode a d'abord été développée sur des standards synthétiques et ces premiers tests ont démontré que l'IRMPD était capable de détecter et d'analyser spécifiquement le galactofuranose. Nous avons ensuite pu appliquer la méthode développée sur des extraits de lichen. La démarche complète allant de la collecte de lichens en Bretagne à la détection de galactofuranose sera décrite.

Mots-clés : galactofuranose, lichen, spectrométrie de masse, spectroscopie vibrationnelle, IRMPD

Abstract

Galactose is an abundant monosaccharide found in nature. Whereas it essentially exists in 6-membered ring (called pyranose), some bacteria, virus, fungi, lichen species are able to synthesize the 5-membered ring (galactofuranose). This form needs more energy for the synthesis and is highly flexible compared to its pyranose counterpart.

The multidisciplinary project ALGAIMS brings together chemists, lichenologists, massists and spectroscopists and aims at developing analytical tools to analyse galactofuranose containing lichen polysaccharides.

We will see in this manuscript that IRMPD, an innovative method consisting on the coupling of a mass spectrometer and an infrared laser, has been employed with two different approaches to analyse galactofuranose.

On one hand, IRMPD enable with the combination of theoretical tools to study galactofuranose conformational preferences in gas phase.

On the other hand, it can be used as an analytical tool to detect galactofuranose in natural extracts. The method has been developed on synthetic standards. These first tests demonstrate that IRMPD was sensible to this rare sugar. We then applied the method to natural lichen extracts. We will describe in this thesis the complete procedure from lichen collect in Brittany to the detection of galactofuranose.

Key words: galactofuranose, lichen, mass spectrometry, vibrational spectroscopy, IRMPD

Liste des sigles

ADN	Acide DésoxyriboNucléique
ALGAIMS	Analysis of Lichen Galf-containing polysaccharides using IRMPD and ion-mobility MS
ANR	Agence Nationale de la Recherche
CID	Collisional Induced Dissociation (dissociation induite par collision)
DFT	Density Functional Theory (théorie de la fonctionnelle de la densité)
ECD	Electron Capture Dissociation (dissociation par capture d'électrons)
ESI	Electrospray Source Ionization (ionisation par électronébuliseur)
ETD	Electron Transfer Dissociation (dissociation par transfert d'électrons)
FAIMS	Field-Asymmetric Ion Mobility Spectrometry
FEL	Free-Electron Laser (laser à électrons libres)
FITDOG	Fenton's Initiation Toward Defined Oligosaccharide Groups (méthode de dépolymérisation chimique basée sur la réaction de Fenton)
FRET	Fluorescence Resonance Energy Transfer (transfert d'énergie par résonance de fluorescence)
FT-ICR	Fourier Transform Ion Cyclotron Resonance (résonance cyclotronique des ions)
Galf	Galactofuranose
Galp	Galactopyranose
ILM	Institut Lumière Matière
IMS	Ion Mobility Spectrometry (spectrométrie de mobilité ionique)
INRAe	Institut national de recherche pour l'agriculture, l'alimentation et l'environnement
IR	Infrarouge
IRHB	InfraRed Hole Burning
IRMPD	InfraRed Multiple Photon Dissociation
IRPD	InfraRed PhotoDissociation
IVR	Intramolecular Vibrational Redistribution (redistribution vibrationnelle intramoléculaire)
KDP	Phosphate Diacide de Potassium
KTA	Potassium Titanyl Arsenate
KTP	Potassium Titanyl Phosphate
LC	Liquid Chromatography (chromatographie liquide)
LID	Laser Induced Dissociation (dissociation induite par laser)
MALDI	Matrix Assisted Laser Desorption Ionization
Man	Mannose
MS	Mass Spectrometry (spectrométrie de masse) ou Mass spectrum (spectre de masse)
MS/MS	Spectrométrie de masse en tandem ou Spectre de fragmentation
MSn	Spectrométrie de masse en tandem à n étages de fragmentation
Nd:YAG	Neodymium-Doped Yttrium Aluminium Garnet
OPO/OPA	Oscillateur Paramétrique Optique /Amplification Paramétrique Optique
R2PI	Resonant 2 Photons Ionisation (ionisation résonante à 2 photons)
RMN	Résonance Magnétique Nucléaire
SNFG	Symbol Nomenclature for Glycans (nomenclature des symboles pour les glycanes)
UV	UltraViolet
UVHB	UltraViolet Hole Burning
VCD	Vibrational Circular Dichroism (Dichroïsme Circulaire Vibrationnel)

Introduction

1 Généralités sur la spectroscopie moléculaire

Une molécule ne peut exister que dans certains états d'énergie quantifiés avec des degrés de liberté électronique, vibrationnelle et rotationnelle. Ces états peuvent être sondés en excitant la molécule par une onde électromagnétique ou un faisceau de particules, cette interaction onde-molécule donnant alors accès à la différence d'énergie entre deux niveaux énergétiques. C'est le principe de la spectroscopie moléculaire, qui peut avoir différentes finalités : l'identification d'une molécule, sa quantification ou encore l'analyse de ses propriétés physico-chimiques. Ces possibilités sont intéressantes pour de nombreux domaines, allant de l'astrophysique à la biologie.¹

Les méthodes spectroscopiques utilisent les différentes possibilités de transition en employant une large gamme de fréquence devenue accessible avec les développements instrumentaux. Le choix de la méthode spectroscopique dépend de la finalité visée. La RMN, qui mesure l'absorption d'une radiation par un noyau atomique sous l'action d'un champ magnétique, se révèle comme un outil très puissant pour caractériser des structures et pour l'identification de l'espèce. Dans le domaine UV-visible, ce sont les transitions électroniques qui sont observées. Alors que la spectroscopie UV n'est pas très efficace pour l'identification d'une molécule inconnue, elle est très prisée pour des analyses quantitatives d'une molécule connue. Les interactions avec le rayonnement IR induisent des transitions entre niveaux de vibration et de rotation et sont, quant à elles, très performantes pour identifier une espèce inconnue et obtenir des informations sur la structure moléculaire. Ces trois types de spectroscopie et leurs caractéristiques respectives sont récapitulés dans le tableau suivant.

Tableau 1 Les types de spectroscopies et leurs caractéristiques.

Domaine spectral	Type de spectroscopie	Type de transition	Applications
Radiofréquence	RMN		Identification
UV-visible	Spectroscopie UV-visible, Raman, fluorescence	Electronique	Dosage
Infrarouge	Spectroscopie IR	Vibrationnel et rotationnel	Identification, analyse structurelle

2 Spectroscopie des biomolécules

Les méthodes spectroscopiques décrites ci-dessus sont largement employées pour étudier des biomolécules en phase condensée.² Les contraintes liées à l'environnement des biomolécules s'imposent alors, rendant leur analyse difficile : elles sont naturellement présentes sous forme de mélange avec d'autres composants. Ainsi, l'hétérogénéité des échantillons biologiques est un paramètre qui peut affecter l'efficacité des spectroscopies en phase condensée.

Au contraire, la spectrométrie de masse (MS) est une approche non spectroscopique largement utilisée qui, couplée à des méthodes d'ionisation douce (ESI, MALDI...), est particulièrement adaptée à l'analyse d'échantillons complexes.³ La MS permet une séparation selon le rapport m/z des ions formés et peut par ailleurs être couplée en ligne avec des méthodes séparatives et la mobilité ionique, ce qui représente un atout dans le cas de mélange complexe. Un autre avantage que l'on peut lui attribuer

est sa sensibilité, facilitant la détection de composés présents en faible quantité. Cette détection dépend toutefois de la gamme dynamique de l'appareil, c'est-à-dire l'étendue des concentrations relatives détectables (capacité à détecter un composant très minoritaire dans un mélange avec un composant très majoritaire). En ce qui concerne les informations que l'on peut obtenir avec la spectrométrie de masse, c'est avant tout une technique de mesure de masse qui peut nous informer sur la formule brute d'un composé. En outre, une autre technique de spectrométrie de masse, dite en tandem, sert à fragmenter des ions. L'analyse des fragments fournit alors des informations complémentaires.

Ces deux dernières décennies ont été marquées par un effort considérable d'adaptation de schémas spectroscopiques traditionnellement employés en phase condensée afin de les appliquer à des ions sélectionnés dans un analyseur en masse. Ces développements incluent en particulier les mesures d'absorption UV et IR, de fluorescence et de FRET en phase gazeuse.⁴⁻⁶

En revanche, une question cruciale se pose concernant cette approche qui repose sur le transfert des composés vers la phase gazeuse, sous forme ionisée. L'étude en phase gazeuse est directement adaptée pour certains domaines comme l'astrophysique car cela permet de se rapprocher des conditions environnementales mais elle se révèle réductrice. De plus, la molécule de départ est modifiée par l'ionisation forcée nécessaire à la détection en spectrométrie de masse. Ainsi nos questionnements portent sur des biomolécules, neutres ou naturellement chargées, qui se trouvent généralement dans un environnement complexe avec lequel elles interagissent. Pour y répondre, nous analysons ces biomolécules ionisées et isolées, leur empêchant toute interaction avec le milieu environnant. Les propriétés mesurées sont donc celles d'ions en phase gazeuse et l'on peut s'interroger sur leur pertinence et leur transférabilité dans l'environnement naturel de ces molécules. Les études sur les bases de Watson-Crick illustrent parfaitement ce problème : la phase gazeuse et l'ionisation modifient le réseau d'hydrogène entre les bases et la célèbre « forme Watson-Crick » qui est l'arrangement stable des doublets de bases dans l'ADN, ne correspond pas à la forme la plus stable observée en l'absence d'environnement.⁷

3 Spectroscopie des sucres

Parmi les familles de biopolymères, on retrouve les sucres (nommés carbohydrates en anglais) dont les briques élémentaires sont les monosaccharides. Leurs structures moléculaires se distinguent de celles des autres biomolécules par la présence de nombreuses isoméries possibles représentées en Figure 1. Par exemple, le glucose et le galactose, deux monosaccharides très abondants dans la nature, se différencient uniquement par la configuration de leur groupe OH en position 4. On parle spécifiquement d'anomères (α ou β) pour désigner deux isomères se distinguant par la configuration du carbone 1. Un autre type d'isomérie est la taille du cycle (qui sera détaillée dans la prochaine partie). Le nombre d'isoméries augmente dans le cas de disaccharides, association de deux monosaccharides. Différents disaccharides sont obtenus selon la composition monosaccharidique et selon la liaison glycosidique qui unit ces deux monomères. Une liaison glycosidique est définie par sa stéréochimie (configuration α ou β) et sa régiochimie (groupe hydroxyle lié). Dans le cas d'oligosaccharides, l'enchaînement n'est donc pas nécessairement linéaire et peut comporter des ramifications. Or cet enchaînement, que l'on nomme également séquence, est intéressant à caractériser dans le cadre d'études de lien structure/activité car il conduit à des réactions biologiques et des propriétés physico-chimiques différentes.⁸ Les applications qui en découlent sont multiples : santé, agroalimentaire, matériaux et énergie.

3. Spectroscopie des sucres

Il est donc nécessaire d'avoir accès à la structure des sucres jusqu'au détail le plus fin pour mieux cerner leur implication dans les processus biologiques et pour le développement des applications associées.

Comme nous l'avons évoqué plus haut, la spectrométrie de masse apporte des informations sur la masse des ions. On peut par exemple distinguer un monosaccharide d'un disaccharide. Ces informations sont souvent insuffisantes pour l'identification des sucres car la distinction d'isoméris n'est pas possible.

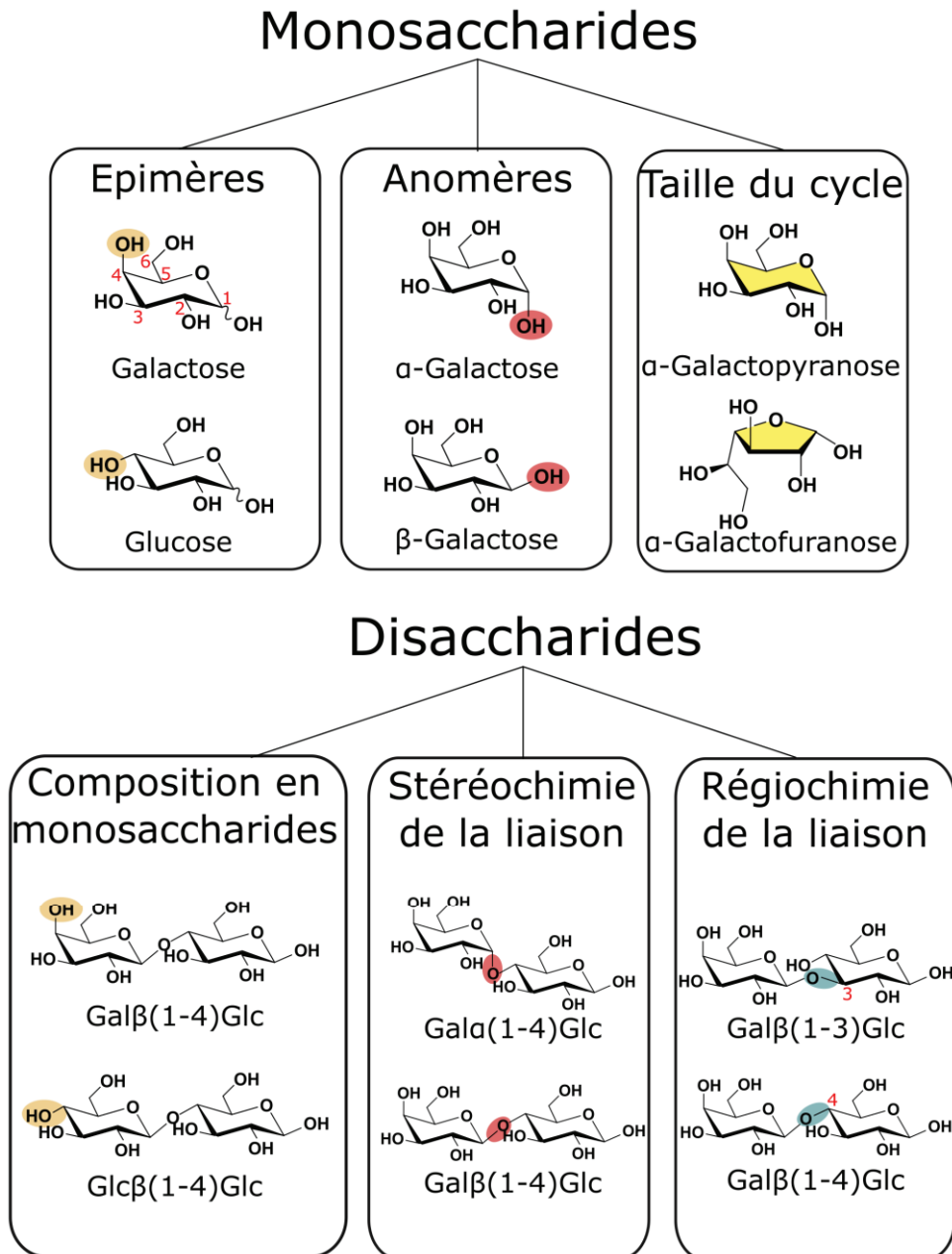


Figure 1 Représentation des types d'isoméris de monosaccharides (en haut) et disaccharides (en bas) de la série D.

Les isomères de sucres se différencient par l'orientation des groupements hydroxyle. Ces derniers présentent des transitions vibrationnelles très caractéristiques dans la gamme des 3 μm . Ainsi, sonder les modes OH dans cette gamme peut être un critère de diagnostic pertinent. Les spectres obtenus en phase condensée dans cette gamme de longueur d'onde sont peu résolus à cause des interactions avec le milieu environnant. Pour s'affranchir de ces effets de solvant et obtenir une meilleure résolution

spectroscopique, une alternative est d'effectuer les spectres IR en phase gazeuse. Il existe différentes stratégies pour passer en phase gazeuse : en chauffant l'échantillon ou par désorption laser mais ces méthodes ne sont pas applicables à tous les sucres et risquent de les détériorer. Comme nous l'avons vu, la spectrométrie de masse entraîne le passage en phase gazeuse et présente de nombreux autres avantages (méthode sensible, possibilité de sélectionner et isoler des ions selon leur rapport m/z , MS en tandem) qui s'avèrent très pertinents pour l'analyse des sucres. La stratégie ici consiste à combiner ces deux méthodes pour tirer les bénéfices de chacune. Alors que la MS nous informe sur la masse moléculaire et permet de former et d'isoler les ions en phase gazeuse, la spectroscopie IR nous renseigne sur la structure du sucre. L'identification du sucre est alors possible avec ces informations. Différentes méthodes reposent sur ce couplage et notre choix s'est porté sur l'IRMPD car il s'avère simple, nécessitant généralement un seul laser et ne requérant pas de refroidissement préalable. Par ailleurs, il offre la possibilité de s'intégrer dans un schéma analytique avec un couplage à une méthode séparative en amont. Ces critères pratiques sont essentiels car les domaines d'application sont multiples et donc des personnes non spectroscopistes pourront être amenées à utiliser cette instrumentation.

Si l'on revient à la question de la pertinence d'étudier en phase gazeuse des biomolécules initialement existantes dans un milieu environnant riche en interactions, il convient de préciser notre objectif. Ce que l'on souhaite mettre en application, c'est le séquençage de sucres. Ainsi même si le réseau d'hydrogène et la conformation (structure tridimensionnelle) sont modifiés pendant le transfert en phase gazeuse, le réseau des liaisons covalentes et donc la nature de la molécule restent les mêmes. La stratégie employée est donc en adéquation avec notre visée. Il est à noter qu'en plus de l'identification, l'IRMPD peut apporter des informations sur les préférences conformationnelles des sucres en phase gazeuse. Il ne s'agit pas de tirer des conclusions sur les conformations des sucres dans leur milieu mais ces travaux nous aident à comprendre leurs propriétés physico-chimiques intrinsèques.

L'IRMPD appliqué aux sucres est une approche originale mise au point par l'équipe DYNAMO. Plusieurs travaux de recherche se basent sur cette méthode. Mon sujet de thèse s'insère dans le projet collaboratif ALGAIMS (financé par l'ANR) qui sera détaillé ci-dessous.

4 Projet ALGAIMS

Une forme d'isomérisation existante est la taille du cycle. Les hexoses sont tous présents sous forme de cycle à 6, c'est-à-dire avec cinq carbones et un oxygène. Le galactose est l'un des monosaccharides qui existe sous deux formes : pyranose (cycle à 6) ou furanose (cycle à 5), représentés en Figure 1. Alors que les mammifères ne synthétisent que le galactopyranose, certaines espèces de bactéries, protozoaires, champignons et lichens peuvent contenir des galactofuranoses.⁹ Cette forme est présente en très peu de quantité et l'intérêt de la détecter est grand, notamment pour développer des thérapies ou pour comprendre les effets biologiques que cette forme implique.¹⁰ Par exemple, les lichens sont des organismes issus de l'association de champignon et d'algue et ont de ce fait des caractéristiques biologiques remarquables (anti-tumoral, résistance à la sécheresse...). Certaines de ces fonctions pourraient être liées aux polysaccharides originaux qu'ils produisent mais les mécanismes ne sont pas toujours bien compris. Par ailleurs, le cycle furanose requiert beaucoup plus d'énergie pour sa biosynthèse que son isomère pyranose.

Toutes ces informations suscitent de nombreux questionnements : pourquoi les espèces en question favorisent le cycle à 5 alors qu'il nécessite plus d'énergie ? Quel est l'impact de la forme furanose sur la conformation de la molécule ? Est-ce que cela a une conséquence sur l'effet biologique ?

4. Projet ALGAIMS

Pour apporter des éléments de réponse, quatre équipes de recherche de domaines différents se sont concertées et de là est né le projet ALGAIMS (Analysis of Lichen Galf-containing polysaccharides using IRMPD and ion-mobility MS). L'objectif du projet est de développer des outils pour l'analyse de polysaccharides naturels contenant des unités galactofuranose. Deux méthodes basées sur la MS et sensibles à l'isomérisation sont employées dans le cadre de ce projet : l'IRMPD par notre équipe DYNAMO à Lyon et la mobilité ionique par l'INRAe de Nantes.

Le projet est divisé en deux parties. Pour la première partie, l'équipe CORINT (Chimie Organique et Interfaces) de l'École Nationale Supérieure de Chimie de Rennes (ENSCR) a synthétisé un grand nombre de standards. La taille des molécules synthétisées va du monosaccharide au trisaccharide. Pour une taille donnée, ces standards se distinguent bien évidemment par la taille du cycle (pyranose/furanose) mais également par la régiochimie des liaisons glycosidiques. Les séquences synthétisées sont les motifs attendus dans les polysaccharides de lichens. L'objectif de la première partie est de développer une stratégie d'analyse (en IRMPD pour l'ILM et en IMS pour l'INRAe) pour caractériser ces modèles synthétiques.

Pour la seconde partie, des lichénologues de l'Institut des Sciences Chimiques de Rennes (ISCR) également membres de l'équipe CORINT se sont chargés de récolter des lichens dans les côtes de Bretagne et d'en extraire les polysaccharides. Plusieurs espèces de lichen non pathogènes ont été récoltées : *Cetraria islandica*, *Lasallia pustulata*, *Roccella fuciformis* et *Lichina pygmaea*. La seconde étape consiste ainsi à appliquer les méthodes développées sur ces échantillons naturels afin d'identifier ceux contenant du galactofuranose.

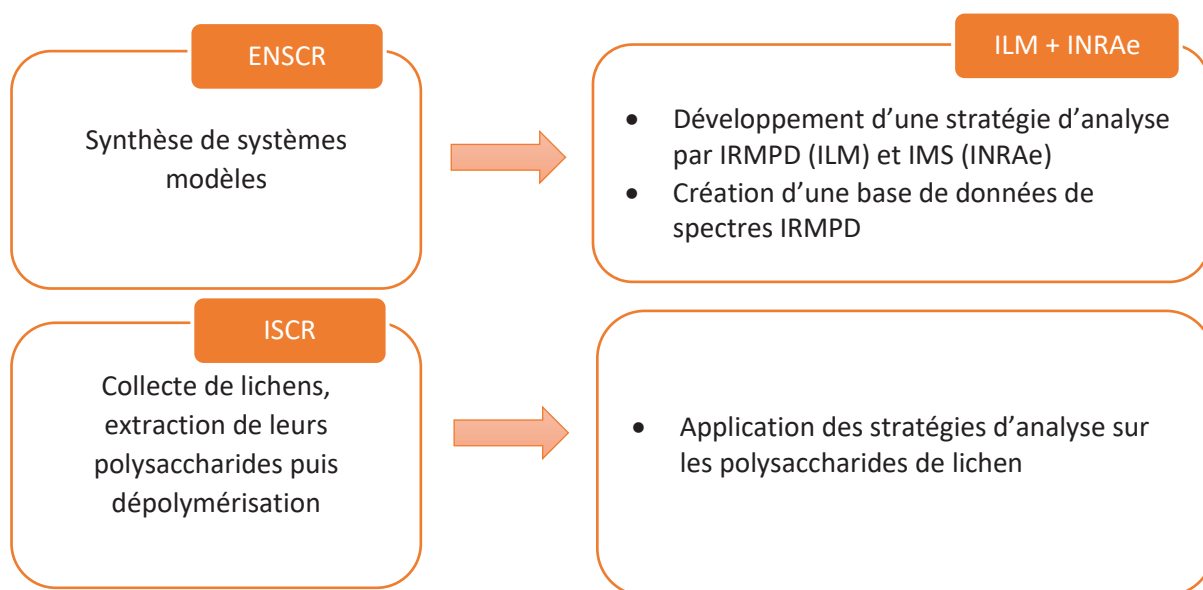


Figure 2 Différentes étapes et collaborateurs du projet ALGAIMS.

En plus de proposer une méthode analytique pour identifier les unités galactofuranose dans des extraits de lichen, l'IRMPD est également utilisé pour répondre à des questions physico-chimiques concernant les furanoses. Les calculs théoriques s'effectuent en collaboration avec l'équipe de physicochimie théorique de l'ILM. Alors que les conformations de monosaccharides sous forme pyranose sont bien connues par les deux équipes,¹¹ celles des furanoses le sont moins. Ces cycles à cinq peuvent être présents sous différentes conformations qui sont séparées par des barrières énergétiques peu élevées, ce qui rend la molécule plus flexible.¹² Nous nous intéresserons donc, avec les simulations théoriques, à cette flexibilité et à l'impact du cycle furanose sur les conformations des molécules.

Introduction

Ce manuscrit est divisé en cinq chapitres. Les méthodes expérimentales et théoriques sont présentées dans le premier chapitre.

Le chapitre 2 est consacré aux premiers résultats obtenus sur de petits composés de synthèse. Nous y verrons que l'IRMPD est une méthode prometteuse pour déterminer la taille du cycle du galactose. Les spectres obtenus peuvent être exploités de deux manières : pour répondre à des questions analytiques (identification de sucres inconnus) ou théoriques (étude conformationnelle de sucres). Ces deux approches seront traitées dans les chapitres suivants.

Le chapitre 3 traite des préférences conformationnelles du galactofuranose tandis que les deux derniers chapitres sont consacrés à l'identification de sucres. En effet, on analysera des petits oligosaccharides de synthèse dans le chapitre 4 et des polysaccharides naturels dans le chapitre 5.

Il s'agit d'une démarche originale et complète de l'utilisation de l'IRMPD sur des éléments issus de ressources biologiques.

5 Bibliographie

- 1 S. Svanberg, *Atomic and molecular spectroscopy: basic aspects and practical applications*, Springer, Berlin ; New York, 4th, rev. ed edn., 2004.
- 2 S. Weiss, Fluorescence Spectroscopy of Single Biomolecules, *Science*, 1999, **283**, 1676.
- 3 J. Fenn, M. Mann, C. Meng, S. Wong and C. Whitehouse, Electrospray ionization for mass spectrometry of large biomolecules, *Science*, 1989, **246**, 64.
- 4 B. Bellina, I. Compagnon, L. Joly, F. Albrieux, A. R. Allouche, F. Bertorelle, J. Lemoine, R. Antoine and Ph. Dugourd, UV spectroscopy of entire proteins in the gas phase, *Int. J. Mass Spectrom.*, 2010, **297**, 36–40.
- 5 A. S. Danell and J. H. Parks, FRET measurements of trapped oligonucleotide duplexes, *Int. J. Mass Spectrom.*, 2003, **229**, 35–45.
- 6 M. S. de Vries and P. Hobza, Gas-Phase Spectroscopy of Biomolecular Building Blocks, *Annu. Rev. Phys. Chem.*, 2007, **58**, 585–612.
- 7 J. Gidden, E. S. Baker, A. Ferzoco and M. T. Bowers, Structural motifs of DNA complexes in the gas phase, *Int. J. Mass Spectrom.*, 2005, **240**, 183–193.
- 8 A. Varki, Biological roles of glycans, *Glycobiology*, 2017, **27**, 3–49.
- 9 Recent knowledge and innovations related to hexofuranosides: structure, synthesis and applications, *Carbohydr. Res.*, 2008, **343**, 1897–1923.
- 10 L. L. Pedersen and S. J. Turco, Galactofuranose metabolism: a potential target for antimicrobial chemotherapy, *Cell. Mol. Life Sci.*, 2003, **60**, 259–266.
- 11 L. Barnes, B. Schindler, S. Chambert, A.-R. Allouche and I. Compagnon, Conformational preferences of protonated N-acetylated hexosamines probed by InfraRed Multiple Photon Dissociation (IRMPD) spectroscopy and ab initio calculations, *Int. J. Mass Spectrom.*, 2017, **421**, 116–123.
- 12 H. A. Taha, M. R. Richards and T. L. Lowary, Conformational Analysis of Furanoside-Containing Mono- and Oligosaccharides, *Chem. Rev.*, 2013, **113**, 1851–1876.

CHAPITRE 1 :

Méthodes expérimentales et théoriques

<u>1</u>	<u>Généralités sur l'IRMPD</u>	10
	<u>1.1 Historique</u>	10
	<u>1.2 Principe d'une mesure de spectre IRMPD</u>	11
	<u>1.3 Autres spectroscopies IR d'ions</u>	14
<u>2</u>	<u>Instrumentation</u>	17
	<u>2.1 Généralités sur le couplage MS-IR</u>	17
	<u>2.2 Sources d'ionisation : ESI et nanospray</u>	18
	<u>2.3 Piège à ions</u>	19
	<u>2.4 Laser infrarouge</u>	19
	<u>2.5 Couplage de l'IRMPD avec les méthodes séparatives</u>	20
<u>3</u>	<u>Outils théoriques</u>	22
	<u>3.1 Obtention des structures</u>	22
	<u>3.2 Calculs des spectres vibrationnels</u>	23
<u>4</u>	<u>Interprétations</u>	24
	<u>4.1 Etude conformationnelle</u>	24
	<u>4.2 Caractérisation de sucres inconnus</u>	26
<u>5</u>	<u>Autres méthodes du consortium</u>	29
	<u>5.1 Synthèse des composés modèles</u>	29
	<u>5.2 Préparation des échantillons de lichen</u>	30
	<u>5.3 Analyse par mobilité ionique</u>	31
<u>6</u>	<u>Bibliographie</u>	32

1 Généralités sur l'IRMPD

1.1 Historique

La spectrométrie de masse permet l'ionisation des composés puis la détermination des rapports masse/charge de ces ions, ce qui peut apporter certaines informations sur les composés d'intérêt. Pour obtenir des informations supplémentaires, il est possible de fragmenter ces ions. L'analyse de fragments est très répandue pour des perspectives d'identification. La méthode de fragmentation employée est souvent la dissociation induite par collision (CID) : l'ion précurseur est soumis à des collisions avec un gaz inerte et se fragmente. En alternative à la fragmentation CID, la fragmentation IRMPD consiste en la dissociation induite par absorption multiple de photons infrarouge à longueur d'onde fixe. Cette méthode a pour la première fois été utilisée pour fragmenter des ions piégés dans des cellules ICR (à résonance cyclotronique ionique)¹ dans lesquelles il n'est pas trivial d'injecter un gaz de collision, puis s'est généralisée sur d'autres types de pièges à ions. Par la suite, le développement de lasers accordables dans une gamme infrarouge étendue a facilité l'évolution de la spectroscopie IRMPD. La spectroscopie IRMPD est une méthode permettant l'acquisition de spectres de fragmentation par balayage de la longueur d'onde (le principe de cette méthode sera détaillé dans la partie 1.2). La fragmentation IRMPD et la spectroscopie IRMPD sont à distinguer et seule la méthode spectroscopique sera traitée. La notation « IRMPD » dans la suite du manuscrit fera donc uniquement référence à la spectroscopie IRMPD.

L'IRMPD a connu différents développements instrumentaux. Les premiers spectres IRMPD ont été réalisés avec un laser CO₂ dans la gamme 900-1100 cm⁻¹.² De nos jours, ces lasers sont peu utilisés en spectroscopie étant donné leur gamme spectrale très limitée et sont remplacés par des lasers accordables dans une gamme plus large. Le développement des lasers à électrons libres (FEL, de l'anglais *Free-Electron Laser*) a permis d'étendre considérablement le champ spectral : Oomens *et al.* ont enregistré des spectres dans la gamme allant de 500 à 1600 cm⁻¹.³ Aujourd'hui, les capacités des FEL permettent d'atteindre la plus large gamme spectrale (66-3600 cm⁻¹ pour FELIX), les infrastructures de CLIO en France et FELIX aux Pays-Bas servent habituellement pour des études d'IRMPD.⁴⁻⁶ Les OPO/OPA accordables (voir description dans la section 2.4) constituent une alternative aux lasers à électrons libres. Leur gamme spectrale et leur puissance sont plus limitées mais leur prix et leur compacité les rendent plus accessibles. Ils sont particulièrement intéressants pour étudier la gamme des 3 μm qui couvre les fréquences d'élongation des liaisons X-H (X=O, N et C).

Les applications de l'IRMPD ont connu un essor important : ce sont d'abord de petites molécules inorganiques qui ont été analysées, puis cette spectroscopie a été largement employée pour étudier la conformation d'ions biomoléculaires⁷ dans une démarche de physico-chimie fondamentale. Plus récemment, le potentiel analytique de la spectroscopie IRMPD s'est révélé, avec en particulier la possibilité de résoudre des isomères de structures moléculaires très voisines. Cette propriété de l'IRMPD ouvre la possibilité d'augmenter les performances de la spectrométrie de masse d'une manière particulièrement pertinente dans les domaines de la métabolomique⁸ et de la glycomique⁹.

Les différents dispositifs d'IRMPD reposent sur un même principe que nous allons détailler.

1.2 Principe d'une mesure de spectre IRMPD

1.2.1 Processus IRMPD

L'IRMPD est une méthode d'analyse structurale en phase gazeuse réalisée sur des ions isolés dans un piège. Notons que ces ions piégés peuvent être des ions précurseurs ou des ions ré-isolés après fragmentation par CID des précurseurs.

Nous pouvons voir en Figure 1 les phénomènes se produisant au niveau du spectromètre de masse (MS), de l'obturateur mécanique (qui permet la synchronisation entre la MS et l'IR), du laser infrarouge (IR) et à plus petite échelle au niveau moléculaire. Nous avons choisi d'illustrer cela avec la glucosamine. Tout d'abord, les ions sont injectés dans la trappe (la formation des ions et leur piégeage seront décrits dans la partie Instrumentation). Le temps d'injection varie selon l'intensité de signal en ions que l'on souhaite. Nous pouvons paramétrer le temps d'injection maximal par séquence MS. Nous avons choisi dans notre exemple un temps d'injection de 50 ms. On isole ensuite uniquement les ions ayant un rapport m/z de $180 \pm 2,5$. La glucosamine protonée, détectée à m/z 180, est alors isolée et les autres ions sont éjectés du piège. Cette étape dure environ 5 ms.

Nous souhaitons ici irradier l'ion précurseur, ce qui ne nécessite pas d'étape préalable de CID après l'isolement de l'ion. L'appareil est synchronisé pour que l'activation et l'entrée du faisceau laser dans la trappe (permise grâce à l'ouverture de l'obturateur mécanique) débutent en même temps.

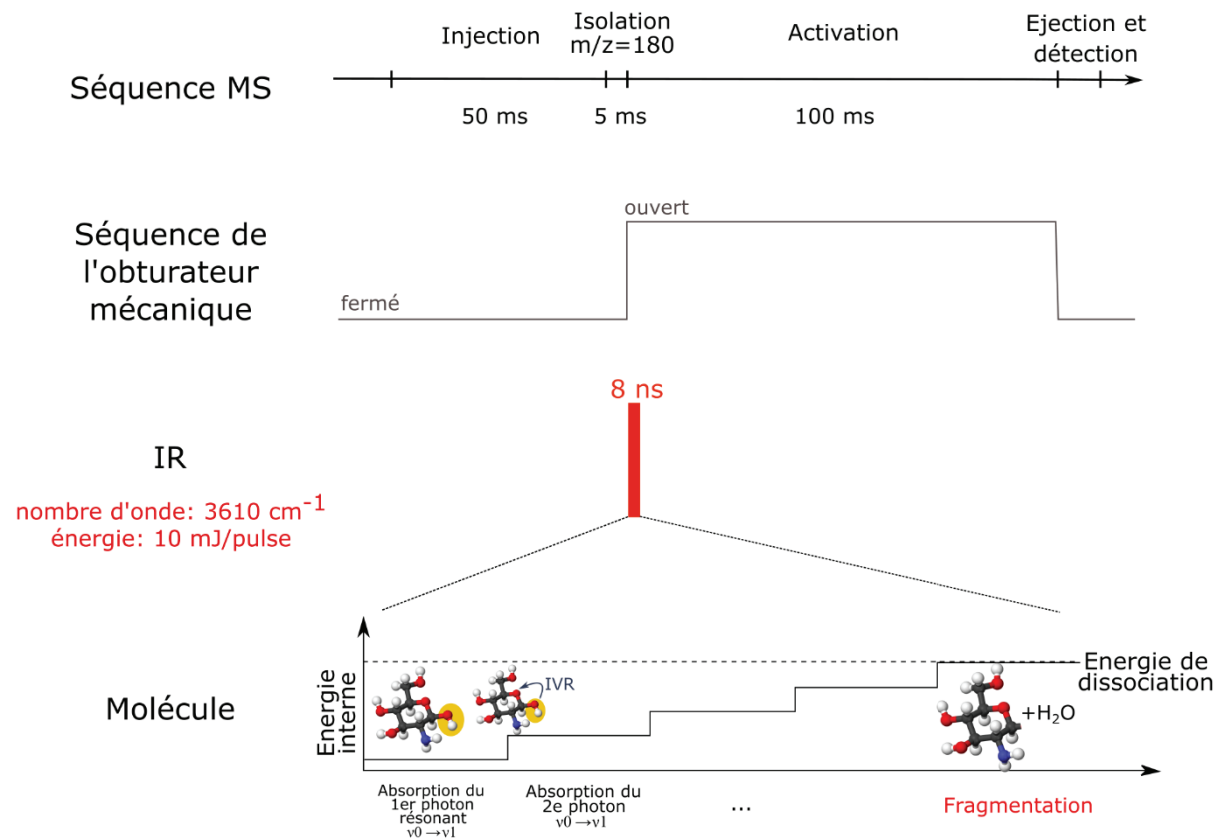


Figure 1 Diagramme temporel du processus IRMPD de la glucosamine avec l'OPO. De haut en bas : séquence MS, séquence de l'obturateur mécanique, séquence IR, énergie de la molécule pendant la durée d'une impulsion laser à une longueur d'onde résonante de la glucosamine.

Une impulsion laser dure 8 ns. L'énergie des photons émis (à 3610 cm^{-1}) étant résonnante avec l'une des transitions vibrationnelles de la glucosamine protonée, celle-ci absorbe un photon. Cela induit la transition vibrationnelle correspondant, soit typiquement la transition $v_0 \rightarrow v_1$ pour un ion thermalisé par le gaz tampon présent dans le piège. Etant donné l'environnement du piège à ions, la molécule ne peut se relaxer et l'énergie est alors redistribuée entre les modes de vibration par le processus de redistribution vibrationnelle intramoléculaire (IVR, de l'anglais *Intramolecular Vibrational Redistribution*) à l'échelle picoseconde. Un autre photon de même énergie pourra alors être réabsorbé sur la transition $v_0 \rightarrow v_1$ et le processus d'IVR sera répété. La durée du pulse laser (8 ns pour l'OPO) est suffisamment longue pour que ces étapes se répètent plusieurs fois. A chaque absorption de photon, l'énergie interne de l'ion augmente jusqu'à atteindre l'énergie de dissociation de la plus faible liaison, puis l'ion se fragmente. La dissociation induite par le laser infrarouge nécessite donc plusieurs photons, d'où l'appellation InfraRed Multiple Photons Dissociation. Ceci implique que l'absorption n'est pas linéaire (à un photon).

Cette fragmentation dépend de différents paramètres de l'appareil. Nous avons déjà évoqué la durée de l'impulsion, qui est une valeur intrinsèque du laser utilisé. Nous avons en revanche une marge de manœuvre sur deux autres paramètres : la focale et la puissance. La focale peut être optimisée afin d'avoir le meilleur recouvrement possible. Quant à la puissance, plus elle sera élevée et plus le nombre de photons sera important, augmentant la probabilité d'atteindre l'énergie de fragmentation. Il faut cependant veiller à ce qu'elle ne soit pas trop excessive au risque de saturer le signal.

Suite à la fragmentation, les fragments produits restent piégés jusqu'à la séquence d'éjection. Leur intensité ainsi que l'intensité de l'ion parent restant sont directement observables grâce à l'analyseur du spectromètre de masse.

1.2.2 Mesure de la photofragmentation

Nous pouvons quantifier la photofragmentation à une longueur d'onde donnée par un taux calculé à partir de l'intensité de l'ion parent et de la somme des intensités des ions fragments selon la formule suivante :

$$\text{photofragmentation} = -\log\left(\frac{I_{\text{parent}}}{I_{\text{parent}} + \sum I_{\text{fragments}}}\right)$$

Dans la partie précédente, nous nous sommes penchés sur un cas simplifié de fragmentation par IRMPD utilisant une seule impulsion laser. La quantité d'ions fragments peut être faible après une seule impulsion. Pour augmenter le signal de photofragmentation, il est possible d'augmenter le nombre d'impulsions laser durant la même période d'activation afin d'irradier les ions parents restants. Pour cela, nous pouvons choisir un laser avec un taux de répétition plus élevé ou, plus simplement, nous pouvons allonger le temps d'activation. La Figure 2 représente le même diagramme temporel que la figure précédente avec un temps d'activation prolongé à 800 ms. Ainsi plus le temps d'activation est augmenté et plus l'intensité des ions fragments est élevée (et celui de l'ion parent bas), augmentant le taux de photofragmentation. Notons que la relaxation est complète en 100 ms et qu'il n'y a donc pas d'effet IRMPD entre deux impulsions laser.

1. Généralités sur l'IRMPD

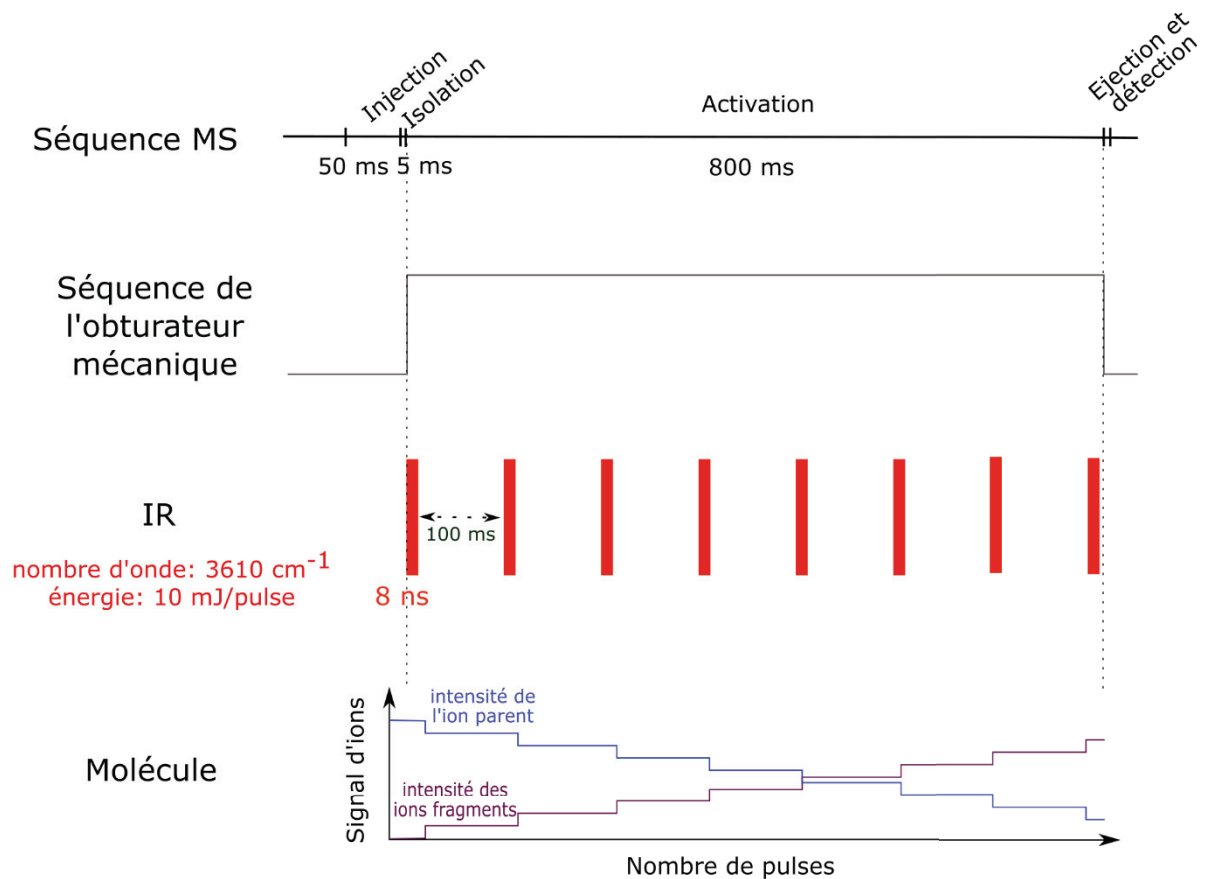


Figure 2 Diagramme temporel du processus IRMPD de la glucosamine avec plusieurs impulsions laser de l'OPO. De haut en bas : séquence MS, séquence de l'obturateur mécanique, séquence IR avec plusieurs impulsions laser à une longueur d'onde résonante de la molécule, évolution de l'intensité des ions parents et fragments selon le nombre d'impulsions laser.

1.2.3 Tracé d'un spectre IRMPD

Pour réaliser un spectre IRMPD, il suffit de choisir un temps d'activation assez long pour avoir un bon signal de photofragmentation. La séquence de spectrométrie de masse est alors répétée dans les mêmes conditions MS mais avec une longueur d'onde du laser différente. En Figure 3, nous pouvons observer dans les encadrés bleu et rouge deux exemples de spectres MS/MS avec fragmentation induite par laser (également nommé LID, de l'anglais *Laser Induced Dissociation*) de la glucosamine protonée respectivement à 3050 et 3610 cm^{-1} . Nous n'observons pas de fragmentation à 3050 cm^{-1} , à l'inverse du spectre MS/MS observé à 3610 cm^{-1} . Le spectre IRMPD est alors tracé en représentant le taux de photofragmentation selon le nombre d'onde.

Un paramètre qui peut être ajusté pour améliorer la qualité du spectre est la vitesse de balayage du laser. Dans l'exemple, il est réglé à 0,4 cm^{-1}/s et il faut 42 minutes pour l'acquisition complète du spectre. Des vitesses plus faibles donnent des spectres avec plus de points mais sont plus lents à acquérir. Un compromis doit donc être trouvé entre durée d'analyse et résolution de spectre.

Enfin, la gamme spectrale étudiée dépend des modes que l'on souhaite analyser. Les sucres étant majoritairement constitués de groupements OH, NH et CH, nous nous intéressons à la gamme des 3 μm qui correspond à leurs modes d'élongation. Les positions et les intensités des bandes observées sur le spectre IRMPD peuvent nous aiguiller sur le type de liaison impliqué. Les OH libres absorbent plutôt vers 3500 – 3700 cm^{-1} . Les bandes OH sont décalées vers de plus basses énergies et sont moins intenses lorsqu'elles sont impliquées dans des liaisons hydrogène intramoléculaires. La gamme des NH se situe vers 3000- 3400 cm^{-1} et celle des CH aux alentours de 2900 cm^{-1} .

Pour être capable d'interpréter plus en détail les bandes observées, des calculs théoriques doivent être réalisés (les outils que nous utilisons sont explicités dans la partie Outils théoriques).

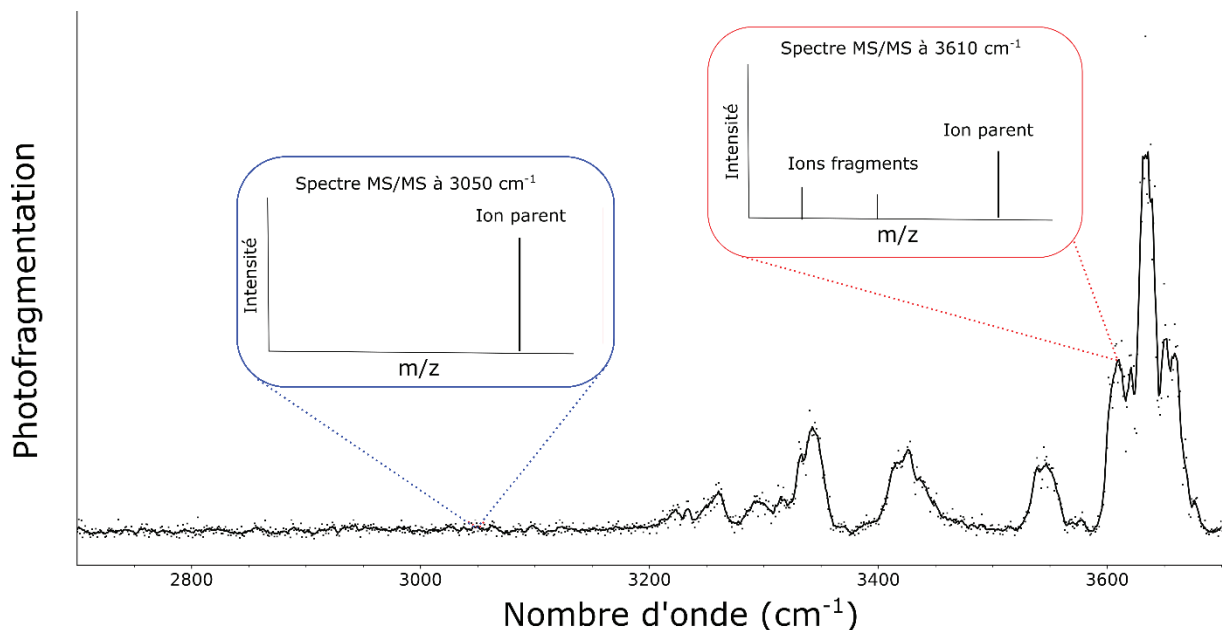


Figure 3 Spectre IRMPD de la glucosamine avec en encadré ses spectres LID à 3050 (en bleu) et 3610 cm^{-1} (en rouge).

Outre l'IRMPD, d'autres méthodes de spectroscopie infrarouge sont également utilisées pour analyser des ions. Nous les détaillerons brièvement comme suit.

1.3 Autres spectroscopies IR d'ions

1.3.1 Spectroscopies R2PI, UVHB et IRHB

Lors d'une collaboration avec Pierre Carçabal (équipe SYSIPHE, ISMO), j'ai eu l'occasion d'utiliser des méthodes d'ionisation résonante à deux photons (R2PI, de l'anglais *Resonant 2 Photons Ionization*), de spectroscopie de dépopulation UVHB (de l'anglais *UltraViolet Hole Burning*) et de spectroscopie de dépopulation IRHB (de l'anglais *InfraRed Hole Burning*).¹⁰ Ces méthodes ont notamment servi à étudier des effets de micro-environnement (les résultats seront présentés dans le chapitre 3). Le principe de chaque spectroscopie est expliqué ci-dessous et illustré en Schéma 1.

Durant le processus R2PI, un premier photon $h\nu_1$ permet la transition électronique $S_0 \rightarrow S_1$ puis un second photon $h\nu_2$ ionise la molécule excitée afin de la détecter par spectrométrie de masse. Il existe une grande différence d'énergie entre les deux photons afin d'éviter que la première irradiation UV crée des ions avec un excès d'énergie (ce qui pourrait entraîner leur fragmentation). Cette méthode permet ainsi d'accéder à la signature électronique du système en faisant balayer la longueur d'onde de la première transition.

Les deux autres types de spectroscopie sont des méthodes dites de dépopulation (ou déplétion), c'est-à-dire qu'une baisse de signal MS est observée lorsque le laser est résonant avec une transition de la molécule. Après la R2PI, la seconde étape consiste à se placer à la longueur d'onde UV la plus intense et ajouter un laser IR en amont pour déterminer les transitions résonantes IR (non représenté sur le schéma). Ensuite, on acquiert le spectre UVHB en positionnant en amont le laser IR fixé à une longueur d'onde résonante et en balayant la gamme UV. On pourra alors comparer avec le spectre R2PI pour voir les bandes n'absorbant pas à la longueur d'onde IR fixe. Si toutes les bandes ne sont pas expliquées, cela signifie la présence d'un autre conformère.

1. Généralités sur l'IRMPD

On répète alors le même processus en fixant le laser IR à une transition qui correspond aux bandes non expliquées. Cette seconde étape nous permet de déterminer le nombre de conformères.

Enfin, on réalise le spectre IR spécifique à chaque conformère en le sélectionnant par l'UV (IRHB). Le laser UV est alors fixé à la transition la plus intense de chaque conformère et le laser IR est balayé en amont.

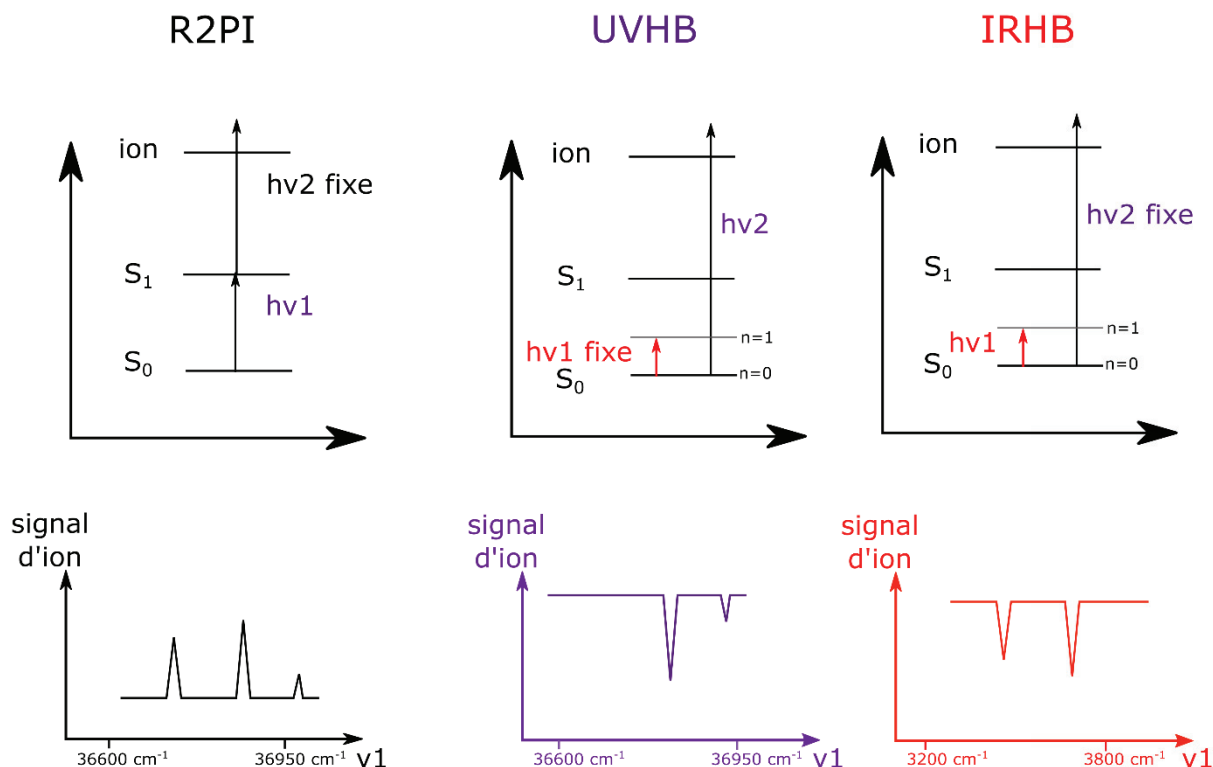


Schéma 1 Principe et exemple de spectre des spectroscopies R2PI, UVHB et IRHB.

Le tableau suivant reprend les principales différences des deux méthodes utilisées dans cette thèse. Notons que la source impacte les conditions expérimentales de l'analyte (charge, température). L'ajout de l'UV requiert un chromophore mais permet de sélectionner les conformères.

Tableau 1 Récapitulatif des principales différences entre l'IRMPD et l'IR-UV.

	IRMPD	IR-UV
source	ESI	jet moléculaire
charge	chargé	neutre
température	ambiante	refroidi
chromophore	pas de modification	greffage chimique
spectre obtenu	somme de toutes les conformations	spécifique à un conformère sélectionné par l'UV

1.3.2 Le « tagging » IRPD

Le « tagging » IRPD (*InfraRed PhotoDissociation*) emploie l'utilisation d'un messenger, généralement un gaz, lié à l'ion à analyser. La liaison est faible et stable uniquement à très basse température. Elle est donc facilement rompue en cas d'absorption résonnante. La méthode a pour la première fois été démontrée sur des ions oxonium liés au H₂¹¹ puis a été utilisée en protéomique et glycomique.^{12,13}

1.3.3 Gouttelettes d'hélium

Cette méthode consiste à piéger les ions dans des gouttelettes d'hélium refroidi. L'hélium interagissant très peu avec les ions, les spectres vibrationnels sont peu affectés par la présence des gouttelettes. De plus, les spectres sont plus résolus avec une basse température car l'état fondamental est majoritairement peuplé.¹⁴

Ces trois méthodes offrent une très bonne résolution spectroscopique mais nécessitent des installations encombrantes et onéreuses. Notons plus particulièrement que l'IRMPD est la seule de ces approches qui ne nécessite pas de refroidissement. Dans le cadre de ma thèse, des dizaines d'analyses ont été effectuées sur des standards puis sur des extraits de lichen. La résolution de l'IRMPD s'est avérée suffisante pour l'analyse des échantillons étudiés, avec des spectres distincts pour les formes furanose et pyranose. L'usage de l'IRMPD dans ce projet est donc adapté car la résolution est satisfaisante et elle est plus facile à mettre en œuvre. Le couplage avec d'autres techniques (LC, mobilité ionique) est également plus accessible avec l'IRMPD. La partie suivante décrit les outils instrumentaux dont on dispose pour appliquer l'IRMPD.

2 Instrumentation

2.1 Généralités sur le couplage MS-IR

L'outil expérimental que nous utilisons, représenté en Figure 4, est constitué de deux instruments commerciaux : un laser infrarouge accordable et un spectromètre de masse composé d'un piège à ions.

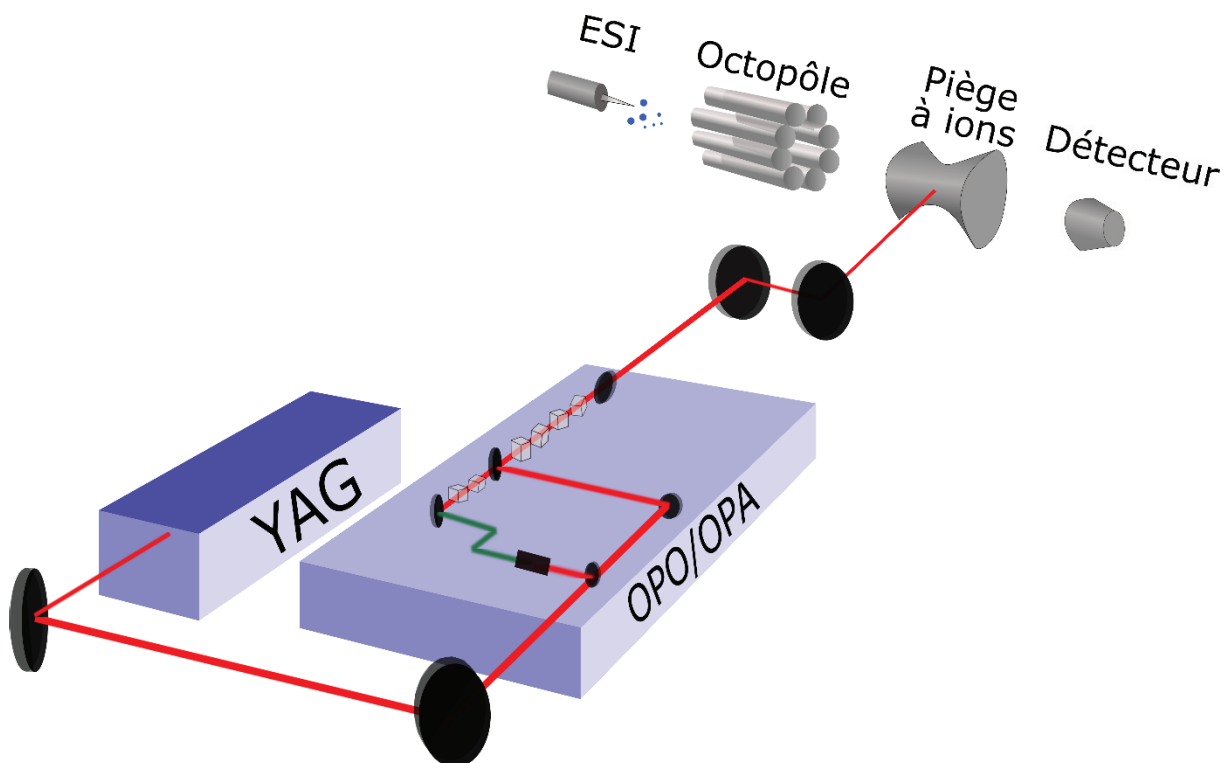


Figure 4 Couplage d'un spectromètre de masse constitué d'un piège à ions avec un laser OPO/OPA pompé par un laser YAG.

Nous disposons de deux types de laser qui seront détaillés en 2.4 : un OPO/OPA (noté « OPO » dans la suite) pompé par YAG commercialisé par LaserVision (qui est représenté ci-dessus) et un système compact Firefly de la société M Squared. À la sortie de la source laser, deux miroirs et un miroir parabolique permettent de focaliser le faisceau au centre du piège à ions. Un obturateur mécanique a également été placé pour synchroniser l'irradiation et l'isolement des ions. Concernant le spectromètre de masse, l'équipe disposait de deux modèles au début de ma thèse : un LCQ et un LTQ XL, tous deux commercialisés par Thermo Finnigan. Le LCQ a ensuite été remplacé par un autre LTQ. Quelques modifications sont effectuées sur les spectromètres de masse afin de procéder au couplage avec l'infrarouge : l'électrode centrale du piège a été percée pour le LCQ, ce qui n'est pas nécessaire pour les LTQ qui possèdent des pièges linéaires dont l'électrode de sortie est déjà percée pour le transfert des ions. Une modification commune aux trois instruments est le placement d'une fenêtre transparente à l'IR afin d'assurer le vide tout en laissant passer le faisceau laser. La synchronisation électronique est obtenue grâce à un obturateur mécanique qui permet l'irradiation pendant l'étape d'isolement. En ce qui concerne l'acquisition des données, une droite de calibration est générée à partir de points de référence enregistrés tout au long de l'acquisition (numéro de scan et nombre d'onde) et permet la conversion du numéro de scan au nombre d'onde correspondant. Ces différents points peuvent également être obtenus automatiquement à l'aide d'un algorithme. Ensuite, ces données sont traitées manuellement (calcul et tracé du taux de photofragmentation en fonction du nombre d'onde) ou automatiquement.

Nous allons à présent détailler les différentes parties de ce couplage : la source d'ions ESI, le piège à ions et le laser infrarouge.

2.2 Sources d'ionisation : ESI et nanospray

La source électrospray, développée par Lydia N. Gall et publiée en 1984 par John Bennett Fenn (prix Nobel de chimie de 2002),¹⁵ est une méthode d'ionisation dite douce qui permet l'analyse de biomolécules. La nanospray est basée sur le même principe mais avec un débit d'injection réduit. Les analytes sont préalablement préparés dans un solvant à une concentration d'environ 50 μM et prélevés dans une seringue. L'injection de la solution dans la source se fait par le biais d'un capillaire en silice à pression atmosphérique avec l'aide d'un pousse-seringue, à un débit d'environ 3 $\mu\text{L}/\text{min}$ (quelques nL/min pour la nanospray). La source électrospray peut être utilisée en mode positif (représenté en Figure 5) ou négatif selon la tension appliquée entre l'extrémité du capillaire et la contre-électrode. Lors de l'exposition à cette haute tension d'environ 4 kV, des charges négatives s'accumulent à la surface du liquide et forment un « cône de Taylor » à l'extrémité du capillaire. Lorsque la tension de surface est atteinte, des microgouttelettes se détachent du cône. Sous l'effet d'une température élevée (d'environ 200 °C) et de la collision avec un gaz inerte injecté (l'azote), l'évaporation du solvant réduit la taille des gouttelettes jusqu'à atteindre un volume critique où, la densité de charge devenant trop importante, les gouttelettes explosent. Des gouttelettes plus petites sont ensuite obtenues et le processus est répété jusqu'à ce que les ions soient complètement dessolvatés. Les ions sont chauds à leur sortie de la source d'ions.

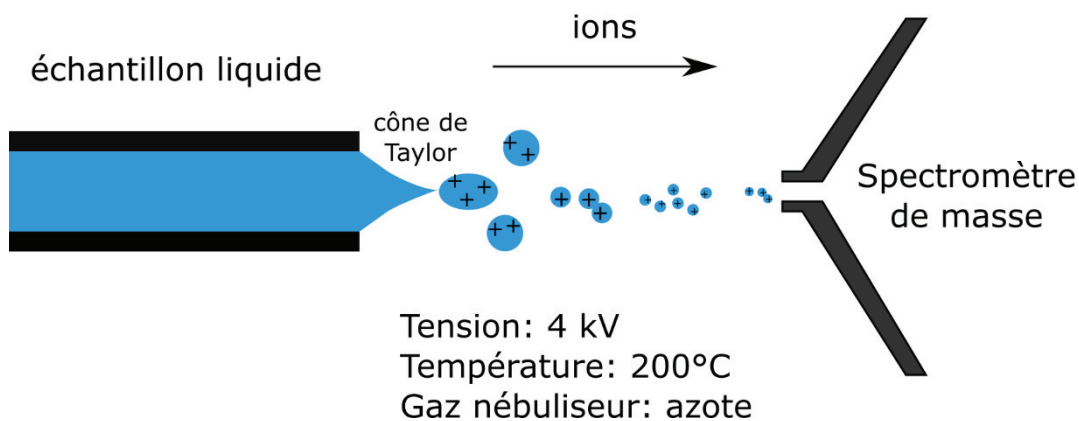


Figure 5 Principe de l'ionisation par électrospray en mode positif.

Ce processus d'ionisation étant continu, il se prête parfaitement à la réalisation de spectres IRMPD. En effet, l'acquisition d'un spectre IRMPD peut être plus ou moins long selon la configuration (entre 2 et 40 minutes dans notre cas). Quelques μg d'échantillon sont suffisants pour acquérir son spectre. Un autre avantage de l'ESI par rapport aux autres sources d'ionisation (MALDI, impact électronique) est sa capacité à générer différents états de charge, notamment en ajoutant des charges différentes (NaCl, LiCl, RbCl...) lors de la préparation d'échantillon. Nous verrons dans les prochains chapitres que le choix de la charge est important en IRMPD.

Ce mode d'ionisation a été continuellement développé et a été associé à plusieurs analyseurs (temps de vol, quadripôle, FT-ICR...). Dans notre installation, la source d'ionisation est couplée à un piège à ions.

2.3 Piège à ions

Une fois les ions produits, ils sont guidés par un ensemble d'optiques ioniques (constitués de deux octapôles pour le LCQ et de deux quadripôles et un octapôle pour le LTQ) dans le piège à ions (tridimensionnel dans le cas du LCQ et linéaire pour le LTQ). Ce piège est constitué d'une électrode circulaire et deux électrodes chapeaux. De l'hélium est introduit dans le piège pour stabiliser la trajectoire des ions et permettre d'en stocker en plus grand nombre. Notons que les multiples collisions faibles entre les ions et le gaz conduisent à la thermalisation des ions. Nous pouvons voir en mode MS les différents ions présents dans le piège (Figure 6, à gauche).

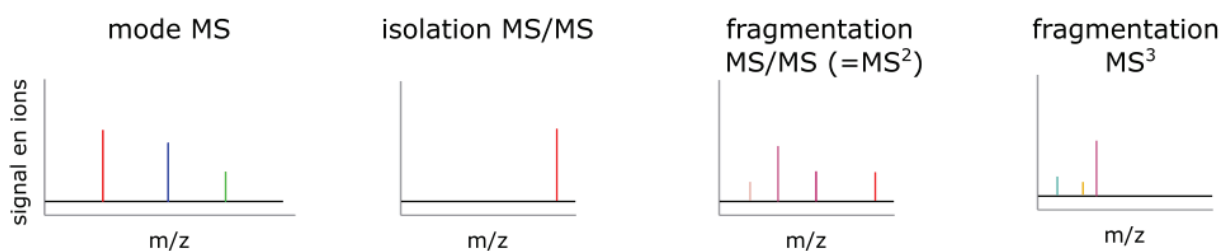


Figure 6 Les différents modes possibles avec un piège un ions : mode MS, isolement MS/MS, fragmentation MS/MS et fragmentation MS³. Des spectres de masse sont donnés en exemple pour chaque mode.

Il est possible d'isoler l'ion d'intérêt selon son rapport m/z et d'éjecter du piège tous les autres ions. En règle générale, nous isolons l'ion d'intérêt selon son rapport m/z avec une fenêtre d'isolement de +/- 2,5 Da. Cette fenêtre d'isolement peut être adaptée en présence d'impuretés avec un rapport m/z proche pour éviter de les isoler.

L'ion isolé peut ensuite être fragmenté par collision avec l'hélium. L'une des grandes particularités des pièges ioniques est la possibilité d'effectuer des analyses en MSⁿ, c'est-à-dire d'isoler et de maintenir piégés successivement les fragments de différentes générations et de les refragmenter. Cet attribut est exploité en IRMPD dans le cas du séquençage, que nous détaillerons en 4.2.2.

Les ions isolés dans le piège, précurseurs ou fragments, peuvent ensuite être irradiés par un laser infrarouge. Les différents lasers utilisés par l'équipe seront présentés ci-dessous.

2.4 Laser infrarouge

L'OPO commercialisé par LaserVision (Figure 4) et le Firefly de M-Squared sont tous deux des lasers de table accordables émettant dans la gamme des 3 μm .

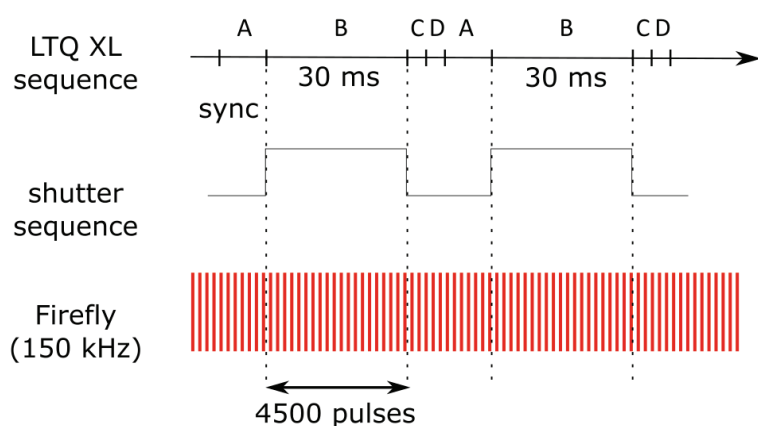
L'OPO est pompé par un laser pulsé Nd :YAG à 1064 nm avec un taux de répétition de 10 Hz et une durée d'impulsion de 8 à 10 ns. La puissance mesurée en sortie du Nd :YAG est d'environ 7 W. Le faisceau émis par le YAG est dirigé vers l'OPO par l'intermédiaire de deux miroirs. Le laser est constitué de deux cavités : l'OPO et l'OPA. Les propriétés d'optique non linéaires de certains cristaux sont exploitées pour permettre l'accordabilité du laser. Le faisceau reçu est scindé en deux faisceaux. La fréquence du premier faisceau est doublée grâce à un cristal KDP et la longueur d'onde atteint 532 nm. Ce faisceau est ensuite aligné avec deux cristaux KTP de la cavité OPO pour obtenir deux signaux : un faisceau accordable de 710 à 880 cm^{-1} et son faisceau complémentaire. Ce dernier ainsi que le second faisceau à 1064 nm sont injectés dans la cavité OPA, constituée de quatre cristaux KTA, produisant des photons de l'ordre de 2700 à 3700 cm^{-1} selon leur angle d'incidence. La puissance en sortie d'OPA est d'environ 100 mW.

Le second laser pulsé utilisé est le Firefly, un laser à diode commercialisé par M-Squared Lasers. Il s'agit d'un oscillateur paramétrique intracavitaire et compact qui permet d'étudier la même gamme de

longueur d'onde que le précédent (2500-3700 cm^{-1}). La durée d'impulsion est inférieure à 10 ns et la puissance en sortie est d'environ 120 mW. Son taux de répétition élevé (130-300 kHz) permet une acquisition plus rapide des spectres IRMPD.¹⁶ En effet, un temps d'irradiation de 30 ms correspond à 4500 impulsions laser avec le Firefly (Figure 7a). En comparaison, les spectres avec l'OPO sont enregistrés avec un temps d'irradiation de 700 ms, ce qui correspond à seulement 7 impulsions laser par séquence (Figure 7b).

Enfin, un laser au CO₂ (Synrad 48-10) est également utilisé dans l'équipe pour palier à la faible puissance du Firefly. Il s'agit d'un laser continu émettant dans les 10 μm . Sa puissance de sortie est de 7 W et peut être modulée avec un contrôleur TTL.

a)



b)

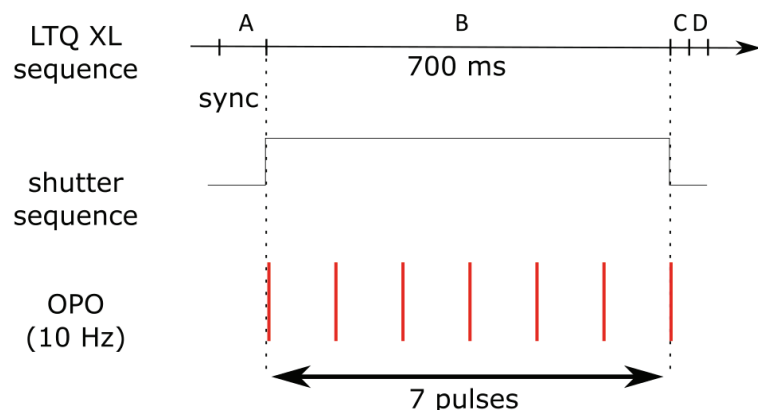


Figure 7 a) Diagramme temporel avec le couplage LTQ XL - Firefly. b) Diagramme temporel avec le couplage LTQ XL - OPO. A= injection des ions dans la trappe, B= irradiation des ions d'intérêt dans la trappe, C= éjection des ions de la trappe et D= détection des ions.(tiré de ¹⁶).

2.5 Couplage de l'IRMPD avec les méthodes séparatives

Nous avons vu que l'IRMPD implique une modification au niveau du piège. Par conséquent, il est possible d'ajouter des méthodes séparatives en amont de la MS-IR sans perturber le processus IRMPD. L'IRMPD peut ainsi s'intégrer dans des schémas analytiques. Cette possibilité est intéressante pour l'analyse d'échantillons biologiques, en l'occurrence des extraits de lichen, afin de séparer les différents constituants. Au sein de l'équipe, des couplages de l'IRMPD ont été réalisés avec la chromatographie liquide et une technique de mobilité ionique (FAIMS).^{17,18}

2. Instrumentation

En ce qui concerne le couplage avec la chromatographie liquide (Figure 8), nous utilisons l'appareil Ultimate 3000 commercialisé par Thermo Fisher Scientific. Les solvants (1) constituent la phase mobile dont la composition est essentielle pour favoriser l'interaction des composés avec la colonne ou entrainer leur élution. Les différents solvants sont d'abord dégazés dans un bain à ultrasons puis le dégazeur installé dans notre système LC entre les solvants et la pompe (2) élimine le gaz restant. Cette étape de dégazage est importante pour éviter la formation de microbulles qui peuvent interférer avec les interactions composé/colonne. Ensuite, une pompe à haute pression (5) contrôle la composition de la phase mobile et permet sa circulation vers la colonne. La vanne d'injection, en fonction de sa position « inject » ou « load », permet d'injecter l'échantillon ou la phase mobile dans la colonne. Une précolonne (8) peut être installée en amont de la colonne pour la protéger. Enfin, la colonne représente la phase stationnaire et est placée dans un four pour garantir la stabilité de la température. La rétention des composés dépend de la phase mobile et de la phase stationnaire.

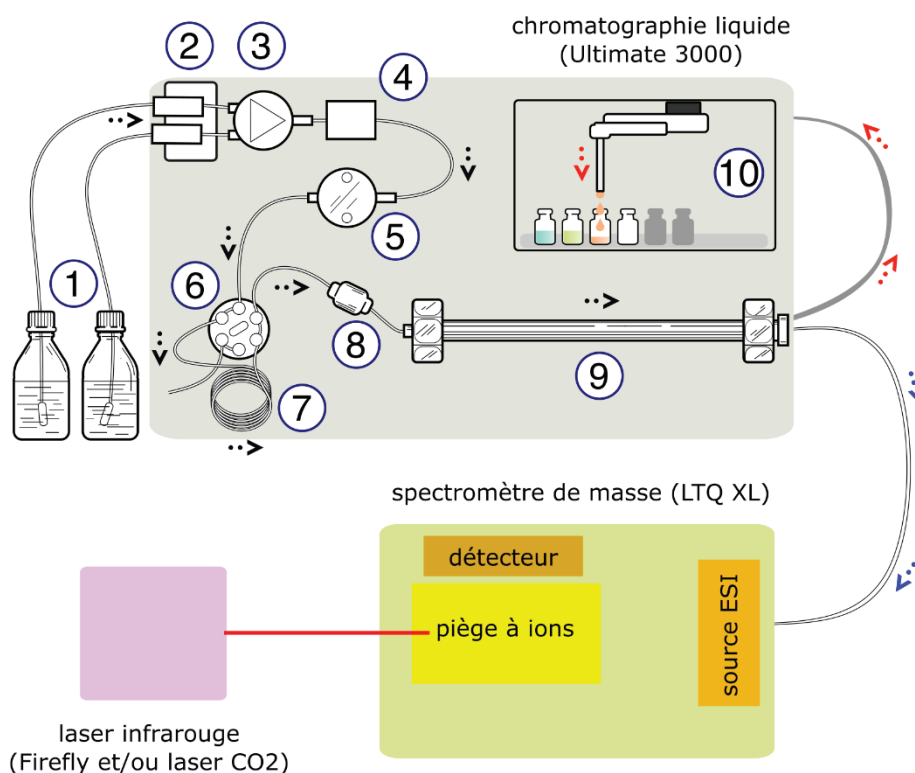


Figure 8 Couplage LC-MS-IR. Légende pour les composants de la chromatographie liquide : (1) Réservoir de solvants, (2) Dégazeur, (3) Valve de gradient d'élution, (4) Doseur de phase mobile ou éluant, (5) Pompe à haute pression, (6) Vanne d'injection, (7) Boucle d'injection de l'échantillon, (8) Précolonne, (9) Colonne, (10) Collecteur de fraction (source pour la partie chromatographie : <https://commons.wikimedia.org>).

En sortie de colonne, le flux peut être dirigé vers la source électrospray (flèches bleues) pour des analyses en spectrométrie de masse avec ou sans couplage avec le laser infrarouge. L'Ultimate 3000 est également équipé d'un collecteur de fraction vers lequel peut être dirigé le flux (flèches rouges) afin de collecter des fractions d'intérêt.

Mes premières analyses ont été réalisées en infusion directe (chapitre 2 et 4.3) puis la chromatographie liquide a progressivement été ajoutée (chapitre 4.4) pour aboutir à l'analyse d'échantillons naturels complexes (chapitre 5.2).

En plus des manipulations expérimentales, des calculs de chimie quantique sont parfois nécessaires pour certains types d'application comme l'analyse structurale des molécules. Les outils théoriques sont détaillés dans la prochaine partie.

3 Outils théoriques

L'équipe DYNAMO travaille en collaboration avec une équipe de théoriciens : Physico-Chimie théorique de l'ILM, dirigée par Abdul-Rahman Allouche. Nous allons détailler les outils mis en place pour simuler les spectres IR puis nous verrons l'usage que nous en faisons.

Les spectres vibrationnels varient pour chaque structure et sont donc différents dans le cas d'isomères, d'anomères et de conformères. Il convient d'effectuer une recherche conformationnelle pour chaque structure moléculaire en sélectionnant et optimisant les conformations les plus pertinentes, puis de calculer leurs spectres vibrationnels. Dans un second temps, ces spectres sont comparés avec les spectres expérimentaux pour en déduire les conformations présentes en phase gazeuse. Le Schéma 2 reprend les différentes étapes d'une recherche conformationnelle.

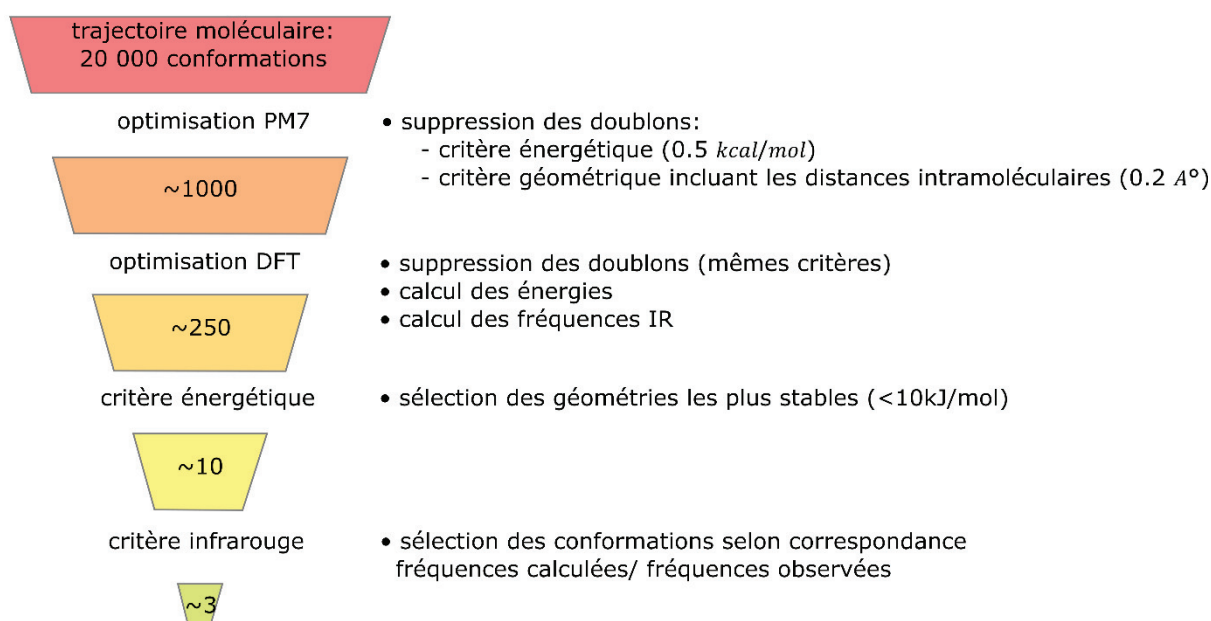


Schéma 2 Les différentes étapes de la recherche conformationnelle avec pour chaque étape les tâches effectuées et le nombre approximatif de conformations qui en résultent.

3.1 Obtention des structures

Dans un premier temps, une recherche conformationnelle est réalisée pour obtenir les différentes conformations d'une molécule. Cette étape pourrait se faire manuellement mais serait très longue. Pour cette raison, nous utilisons des outils de dynamique moléculaire qui simulent l'hétérogénéité conformationnelle des molécules. À partir d'une seule structure, des milliers de géométries sont produites à très haute température (avoisnant les 3000 K) avec des calculs semi-empiriques de PM7 permettant de décrire le mouvement des atomes. Les doublons sont éliminés selon des critères énergétiques et géométriques : lorsque deux structures ont une différence d'énergie inférieure à $0,5 \text{ kcal/mol}$, leurs distances interatomiques sont vérifiées et si ces dernières sont inférieures à $0,2 \text{ \AA}$, les deux structures sont considérées identiques. Nous obtenons alors environ 1000 géométries uniques.

Dans un second temps, nous optimisons les conformations avec la théorie de la fonctionnelle de la densité (DFT, de l'anglais *Density Functional Theory*) qui est une méthode de calcul quantique utilisée pour résoudre la structure électronique de la matière. Les fonctionnelles B3LYP sont couramment utilisées pour l'étude des molécules organiques et sont économiques en temps de calculs avec la base modeste 6-31+G*. De nombreuses autres méthodes ont été testées et la fonctionnelle CAM-B3LYP a

3. Outils théoriques

été sélectionnée car elle a la précision la plus satisfaisante.¹⁹ Ainsi, les 1000 conformations sont optimisées avec les fonctionnelles B3LYP/6-31+G* puis CAM-B3LYP/6-311++G(2df,2pd). Les doublons sont supprimés entre chaque étape. CAM-B3LYP/6-311++G(2df,2pd) permet également le calcul d'énergie et des fréquences IR de chaque conformation, utilisées comme critères de sélection lors de l'étape suivante.

Les géométries les plus stables énergétiquement sont sélectionnées, avec une valeur seuil d'énergie généralement fixée autour de 10 kJ/mol. Les spectres vibrationnels des géométries sélectionnés sont ensuite calculés.

3.2 Calculs des spectres vibrationnels

Le calcul des spectres vibrationnels repose sur la résolution de l'équation de Schrödinger pour la partie nucléaire. Seule la transition $v_0 \rightarrow v_1$ est concernée par l'IRMPD. Or, comme nous pouvons le noter sur la Figure 9, le potentiel d'énergie peut être approximé par un potentiel quadratique pour les plus basses énergies. Ces approximations harmoniques du potentiel calculées par DFT servent à résoudre l'équation de Schrödinger mais conduisent à des écarts entre les valeurs théoriques et expérimentales. L'élargissement du potentiel réel dû à sa limite de dissociation résulte généralement en une valeur de fréquence plus faible que l'approximation quadratique.

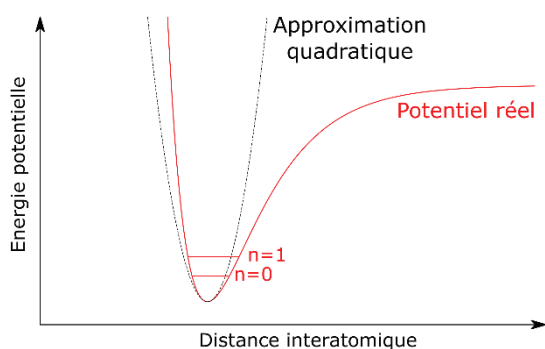


Figure 9 Potentiel réel (courbe rouge) et approximation quadratique (courbe noire) avec représentation des niveaux de vibrations $n=0$ et $n=1$.

La solution la plus simple pour corriger ces écarts est d'appliquer un facteur correctif inférieur à 1 pour que les résultats des calculs théoriques soient plus proches des valeurs réelles. Cette méthode se révèle efficace pour obtenir les positions sur les spectres théoriques proches de celles obtenues expérimentalement et est notamment utilisée dans certains de mes travaux.²⁰

Pour l'interprétation de ces résultats, il est important de constater que les calculs DFT sont obtenus à partir de structures figées alors que l'IRMPD est un processus dynamique d'absorption multiphotonique à haute température. Par conséquent, la théorie ne prédit pas la forme des bandes de photofragmentation mesurée car elle prédit une fréquence unique pour un mode de vibration donné. La comparaison entre le spectre des fréquences de vibrations calculées et le spectre IRMPD s'appuie donc en grand partie sur l'expertise et l'expérience des expérimentateurs et des théoriciens.

Certaines équipes proposent d'autres méthodes pour approcher les conditions expérimentales comme la simulation de la DFT à température ambiante. Une température basse empêche les excitations rotationnelles qui entraînent l'élargissement des bandes.²¹ D'autres approches, plus avancées et que nous ne détaillerons pas ici, prennent en compte l'anharmonicité du potentiel.

Seule ou en combinaison avec la théorie, l'IRMPD permet d'élucider de nombreuses questions de physico-chimie concernant les sucres. Voyons comment interpréter les données collectées.

4 Interprétations

Les spectres expérimentaux IRMPD peuvent être interprétés en utilisant deux approches différentes. D'une part, ils permettent avec les résultats de chimie quantique de l'équipe Physico-chimie théorique de répondre à des questions d'ordre fondamental sur la conformation d'ions (4.1). D'autre part, les expérimentations apportent également des réponses à des questions analytiques concernant l'identification de sucres inconnus (4.2).

4.1 Etude conformationnelle

En combinaison avec les calculs théoriques, l'IRMPD fournit une mine d'informations sur les propriétés structurales des ions en phase gazeuse. Ces analyses sont effectuées en confrontant les données expérimentales et les résultats des calculs théoriques. Nous prenons en compte pour cela différents paramètres : comparaison des positions et des largeurs de bande, attribution de bande spécifique le cas échéant, environnement des groupes fonctionnels dans les différentes conformations, prise en compte des énergies relatives.

Nous pouvons d'abord déterminer les modes représentés sur un spectre IRMPD selon la région d'absorption des bandes expérimentales. En effet, on sait que les bandes OH sont représentées dans la gamme $3450 - 3700 \text{ cm}^{-1}$, les NH vers $3000 - 3400 \text{ cm}^{-1}$ et les CH à des énergies beaucoup plus faibles (vers 2900 cm^{-1}). Ces fréquences ne sont pas figées et varient notamment en cas de liaison hydrogène : on observe alors un déplacement vers les basses énergies et un élargissement des bandes liées. Des bandes OH liées pourront ainsi se chevaucher avec des bandes NH libres et de même les bandes NH liées pourront être confondues dans la gamme des CH. Nous avons recours aux calculs théoriques pour une analyse plus poussée, notamment pour déterminer le groupement de la molécule correspondant à chaque mode.

À titre d'exemple, considérons notre étude conformationnelle menée sur la N-acétylglucosamine O-méthylée acétylée en position 3. Son spectre expérimental est tracé en gris sur la Figure 10.²⁰ La large bande étendue de 2850 à 3200 cm^{-1} correspond aux modes CH et NH et les deux bandes à 3560 et 3660 cm^{-1} à des modes OH. Plus un groupement est lié, plus sa fréquence sera basse. Comme le mode NH est calculé à basse fréquence, on s'attend à un NH fortement lié (les NH libres sont généralement observés vers 3300 cm^{-1}). De même, le groupe OH observé à 3560 cm^{-1} est lié alors que celui à 3660 cm^{-1} est libre. Cependant, nous ne sommes pas capables de déterminer l'hydroxyle de la molécule qui absorbe à 3560 cm^{-1} et celui qui absorbe à 3660 cm^{-1} .

Pour cela, et également pour analyser la conformation, nous avons effectué des calculs théoriques qui indiquent deux conformations répondant au critère énergétique (en dessous de 10 kJ/mol) que l'on nommera glcn1 et glcn2. On notera que glcn1 est beaucoup plus stable énergétiquement que glcn2 (écart de $8,2 \text{ kJ/mol}$).

4. Interprétations

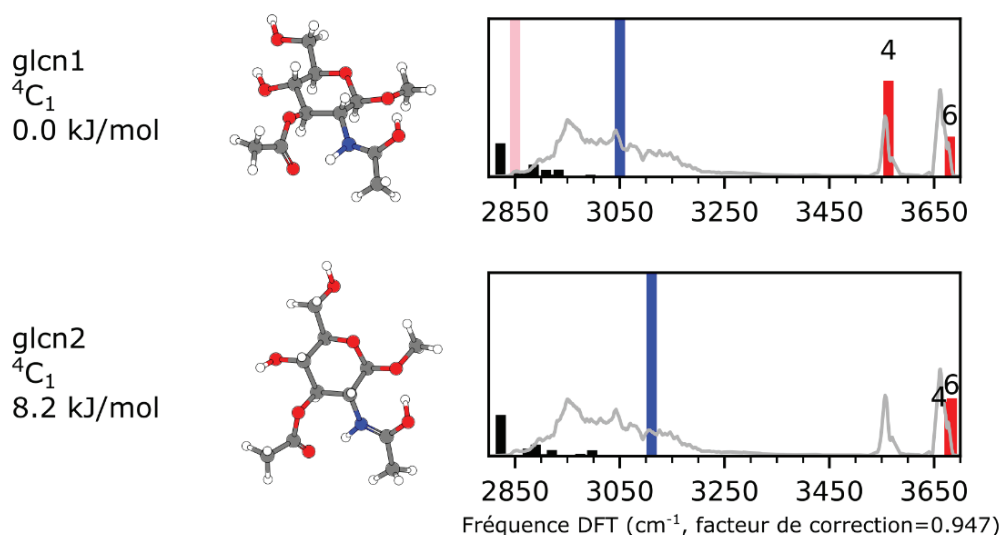


Figure 10 Au milieu : structures des deux conformations les plus stables (glcn1 et glcn2) de la glucosamine O-méthylée acétylée en C3 (3OAcGlcNOMe). À gauche : leur conformation de cycle et leur énergie relative. À droite : leurs spectres théoriques (modes OH en rouge, NH en bleu, CH en noir et OH du N-acétyle en rose) et le spectre expérimental tracé en gris.

Nous analyserons les spectres théoriques des deux conformations dans un premier temps, puis nous les comparerons au spectre expérimental.

Pour la conformation la plus stable glcn1, le groupe O6H est libre et a une fréquence calculée vers 3670 cm^{-1} . Le groupe O4H étant lié à O6H, sa fréquence est décalée vers le rouge (à 3560 cm^{-1}) et est plus intense. Enfin, le mode OH du groupe protoné N-acétyle est fortement lié et calculé avec une forte intensité vers 2850 cm^{-1} . Notons que plus un mode est lié, plus sa fréquence calculée se décale vers les basses énergies et plus son intensité augmente. Expérimentalement, les bandes très liées se déplacent aussi vers les basses énergies mais elles s'élargissent et leur intensité diminue au point de ne plus être observables dans le cas des liaisons fortes. Nous pouvons constater que la bande expérimentale vers 3660 cm^{-1} correspond à O6H et celle à 3560 cm^{-1} à O4H. La bande de droite est plus intense expérimentalement que celle de gauche, alors que c'est l'inverse pour les calculs théoriques. Enfin, le mode OH du groupement N-acétyle (en rose) est très décalé pour être observé expérimentalement.

Pour la seconde conformation glcn2, les deux groupes O4H et O6H sont libres et leurs fréquences calculées vers 3670 cm^{-1} . La bande expérimentale à 3560 cm^{-1} ne peut donc pas être expliquée par cette conformation. Grâce à nos outils théoriques, nous savons maintenant à quel mode correspond chaque bande expérimentale.

L'interprétation des modes très liés est moins directe que celle des modes libres ou faiblement liés. Leurs fréquences calculées se déplacent vers les basses énergies tandis que leur intensité augmente. Expérimentalement, les bandes fortement liées se déplacent également vers les basses énergies mais elles s'élargissent et leur intensité diminue au point de ne plus être observables dans le cas des liaisons fortes. C'est le cas du mode OH du groupement N-acétyle, calculé vers 2850 cm^{-1} avec une forte intensité pour glcn1. Il n'y a pas de bande expérimentale correspondant à ce mode.

La comparaison entre bandes expérimentales et théoriques demande donc une certaine expertise acquise par l'expérience. Outre l'assignation des bandes, les calculs théoriques donnent également des informations sur la structure tridimensionnelle des ions en phase gazeuse. Ici la combinaison de l'IRMPD et des outils théoriques a révélé que la conformation la plus stable était de type 4C_1 , que le proton se positionnait au niveau du groupement N-acétyle et que la conformation était riche en liaisons hydrogène.

Nous verrons dans la prochaine partie que nous pouvons également utiliser l'IRMPD seul, notamment pour répondre à des questions analytiques en glycomique.

4.2 Caractérisation de sucres inconnus

Les travaux antérieurs de l'équipe ont inclus l'analyse de nombreux sucres de différentes tailles, allant du simple monosaccharide à la glycoprotéine issue du sérum humain.

4.2.1 Identification de sucres de petite taille

L'IRMPD s'est avérée être une méthode robuste pour différencier des sucres de petite taille. Des spectres distincts ont été obtenus pour des épimères, anomères, isomères de position, régio- et stéréoisomères.^{13,14,15} Ces spectres servent à construire une librairie de référence.

Lorsque l'on souhaite identifier un sucre inconnu de petite taille, on l'isole dans le piège d'ions selon son rapport m/z (Schéma 3). Les autres ions sont ainsi éjectés du piège et l'on n'observe sur les spectres de masse plus que l'ion avec le m/z d'intérêt. On irradie ensuite l'ion isolé dans la gamme spectrale d'intérêt afin d'acquérir son spectre IRMPD (voir partie 1.2.3).

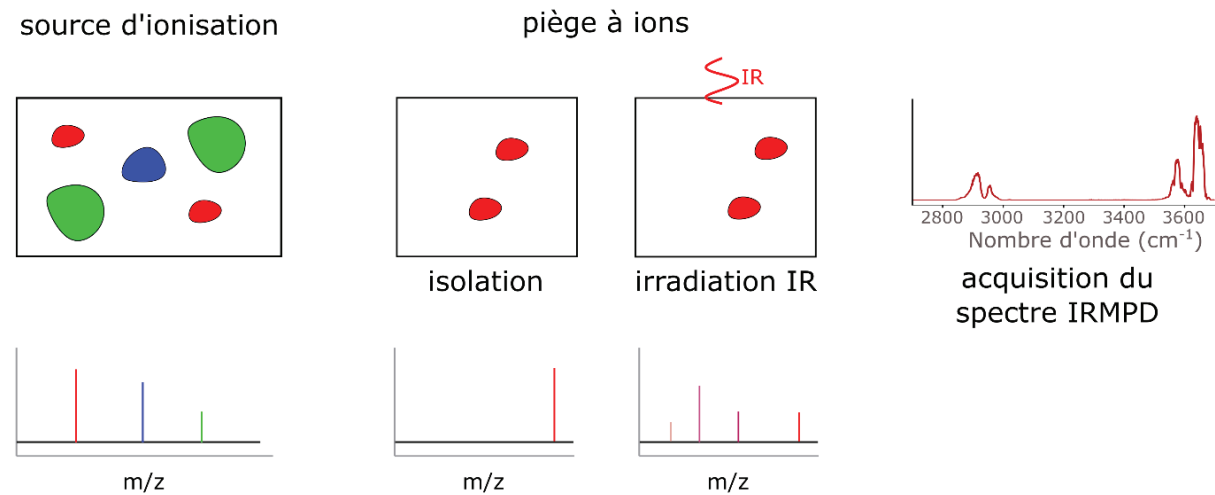


Schéma 3 Différentes étapes de l'irradiation d'un ion parent.

L'interprétation repose sur la comparaison du spectre acquis avec ceux de la base de données. Si on sait ce que l'on cherche (par exemple on sait qu'il s'agit soit du glucose soit du mannose), la comparaison s'effectuera uniquement avec les signatures IR de ces deux composés. Il est possible de ne pas avoir d'a priori sur ce qui est analysé. Dans ce cas, il conviendra de comparer avec tous les spectres de la base de données présentant le même rapport m/z que l'ion à identifier. Les comparaisons s'effectuent visuellement ou à l'aide d'algorithmes de comparaison. La reconnaissance automatisée de spectres IRMPD à l'aide d'outils en intelligence artificielle est en cours d'exploration dans l'équipe.

4. Interprétations

4.2.2 Séquençage

Les problèmes biologiques, en revanche, sont souvent associés à des sucres plus longs que les monosaccharides. Dans notre cas, le galactofuranose se trouve dans de longs polysaccharides. Pour ces sucres de plus grande taille, la sensibilité de la méthode est généralement insuffisante pour détecter des détails fins. Cependant, des analyses de MSⁿ-IR ont révélé que les fragments de CID conservent la mémoire des détails structurels du précurseur,²² ce qui s'avère particulièrement intéressant pour le séquençage.

Ainsi l'IRMPD permet la caractérisation de sucres de plus grande taille en analysant ses fragments pour ensuite remonter à la structure de ce sucre. La procédure d'analyse des fragments est présentée dans le Schéma 4. D'abord, l'ion d'intérêt est isolé selon son rapport m/z puis fragmenté par CID. Parmi les fragments produits, le fragment d'intérêt est ré-isolé et irradié en balayant la gamme de longueur d'onde. Le spectre IRMPD est alors reconstitué en fonction de l'intensité des photo-fragments. Une seule étape de CID est effectuée dans cet exemple, mais plusieurs étapes de fragmentation successives peuvent être effectuées.

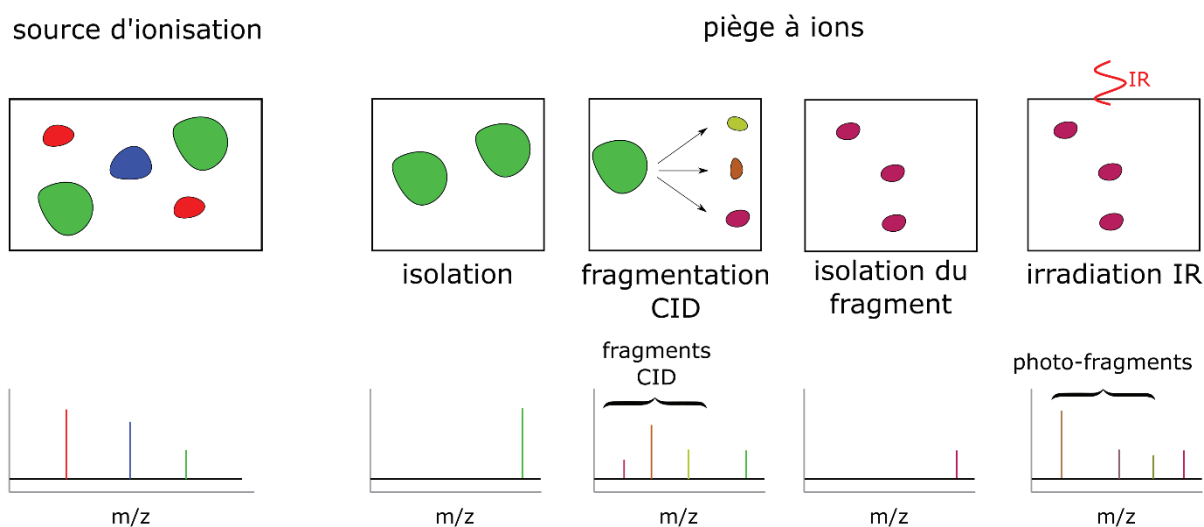


Schéma 4 Différentes étapes de l'irradiation d'un ion fragment.

L'interprétation du spectre se concrétise par comparaison avec les spectres de la base de données. Cette étape s'effectue visuellement ou à l'aide d'algorithme qui attribue des scores de correspondance. L'identification des fragments permet ensuite la reconstitution de la séquence de l'ion parent.

4.2.3 Criblage

Le criblage (*screening* en anglais) est une méthode qui discrimine des molécules ayant des propriétés communes. C'est donc une première étape de tri et les molécules sélectionnées peuvent ensuite faire l'objet d'études complémentaires.

La méthode de criblage que nous avons mise au point est basée sur le couplage de la chromatographie liquide, la spectrométrie de masse et du laser infrarouge (Schéma 5). Les composés sont séparés par la LC, isolés par la MS et irradiés par le laser IR. L'irradiation laser sert à discriminer les molécules partageant des propriétés vibrationnelles communes.

Notons que nous pouvons appliquer la méthode de criblage sur l'ion précurseur (mode LC-MS-IR) mais également sur des ions fragments (LC-MS/MS-IR) si nous souhaitons discriminer les molécules selon les propriétés vibrationnelles de leur fragment.

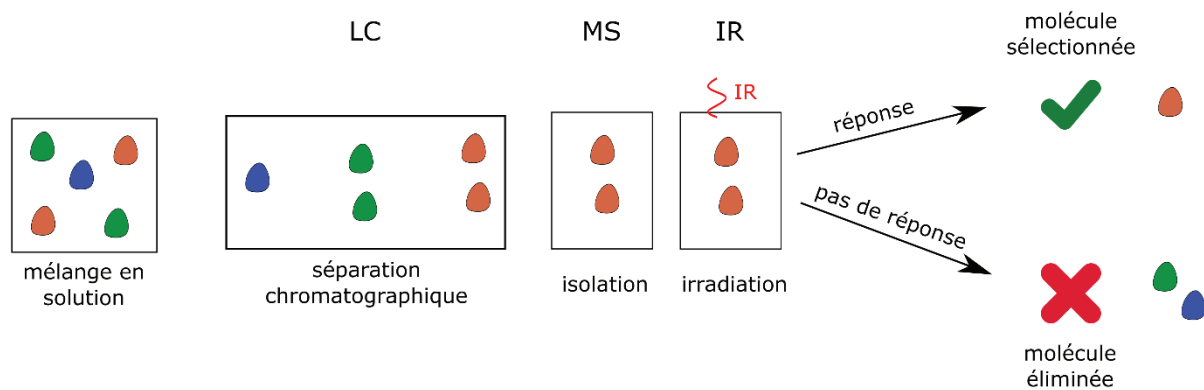


Schéma 5 Différentes étapes de la méthode de criblage.

Nous avons vu dans les parties précédentes les méthodes employées au sein de notre équipe dans le cadre du projet ALGAIMS. Le consortium est constitué de trois autres équipes, chacune apportant son expertise dans son domaine. La partie suivante décrit brièvement ces méthodes.

5 Autres méthodes du consortium

5.1 Synthèse des composés modèles

La première phase du projet nécessite des modèles de sucres de différentes tailles. Des monosaccharides ont été synthétisés lors d'une collaboration avec Laura L. Kiessling de l'Institut de technologie du Massachusetts. Les synthèses des di- et trisaccharides ont été réalisées par l'équipe de l'ENSCR. Les structures ont été sélectionnées en amont du projet selon ce qui était décrit dans la littérature. La stratégie de synthèse consiste à préparer des briques moléculaires qui sont ensuite utilisées dans plusieurs voies de synthèse. Nous pouvons voir dans le Schéma 6 que la synthèse des quatre disaccharides Galf-Man est accomplie avec une seule molécule de départ. Notons par ailleurs que leur synthèse passe par des intermédiaires de synthèse disaccharide Galf-Man avec un groupe thiophényle en extrémité réductrice. Ces composés Gal-ManSPH ont servi lors des manipulations de fragmentation (chapitre 4).

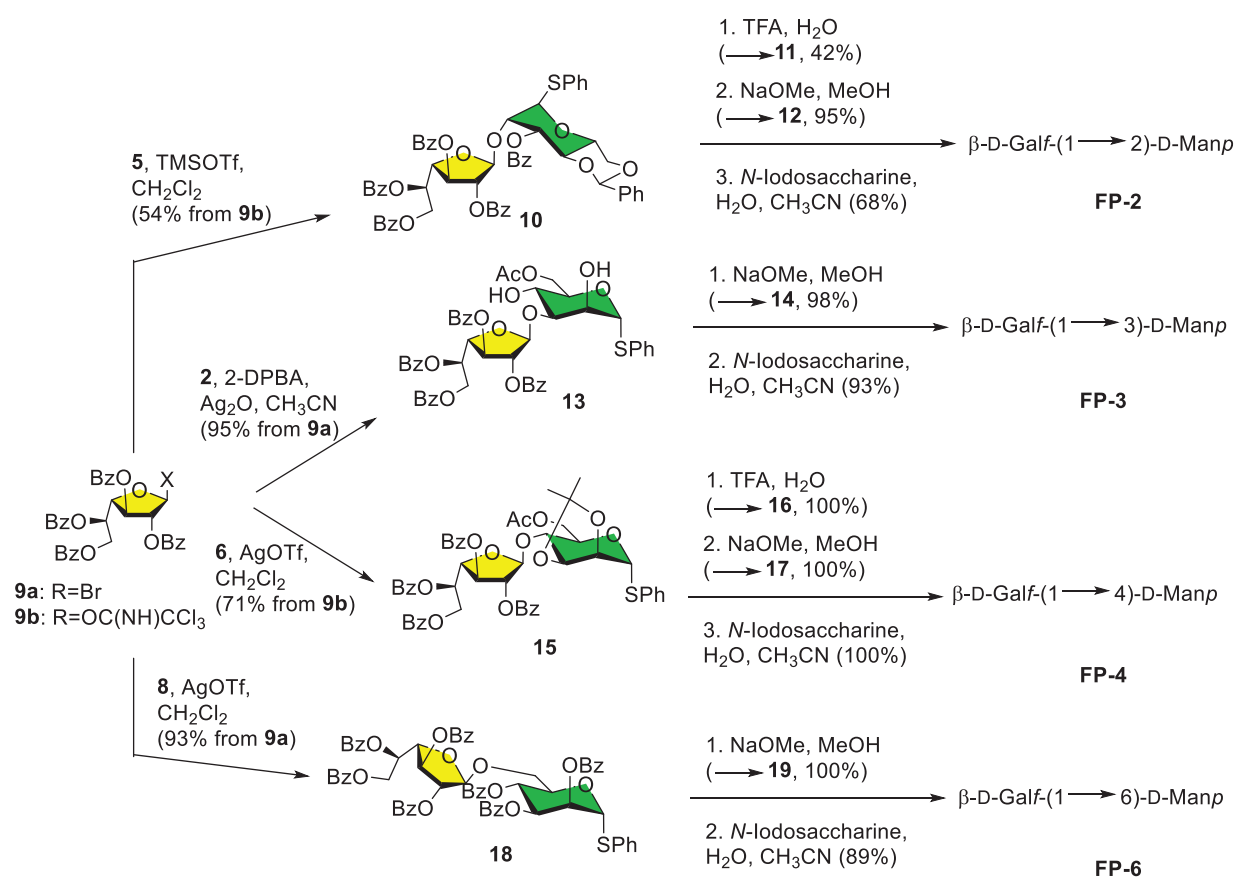


Schéma 6 Synthèse des disaccharides Galf-Man (extrait de [24]).

5.2 Préparation des échantillons de lichen

Les lichens des espèces *Cetraria islandica*, *Lasallia pustulata*, *Lichina pygmaea* et *Roccella fuciformis* ont été collectés en Bretagne par l'équipe CORINT de l'ISCR. Le protocole d'extraction des polysaccharides est composé de plusieurs étapes et est représenté dans le Schéma 7.

Protocole d'extraction

Réalisé à partir de Pereyra, 2008

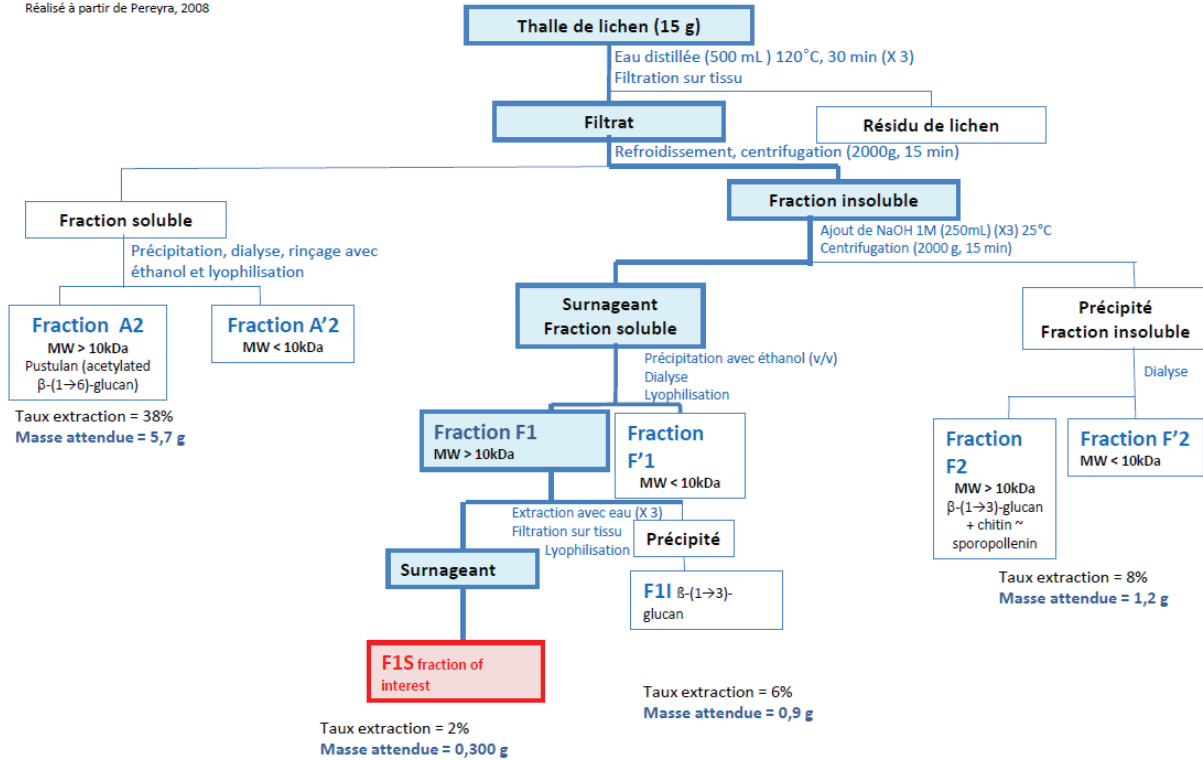


Schéma 7 Protocole d'extraction des polysaccharides de lichen (source : ISCR).

La première étape consiste à filtrer environ 15 grammes de collecte à travers un tissu en faisant passer de l'eau distillée. On voit ensuite plusieurs voies qui conduisent à des fractions différentes. Nous sommes intéressés à la fraction A2 qui présente un fort taux d'extraction. Une digestion enzymatique était ensuite prévue afin d'obtenir des oligosaccharides pouvant être analysés par spectrométrie de masse mais les enzymes n'ont pas fonctionné. Nos efforts se sont alors orientés vers deux méthodes de dépolymérisation chimique : le protocole FITDOG et l'hydrolyse acide. Elles seront décrites dans le dernier chapitre consacré aux polysaccharides où nous verrons que l'IRMPD a permis de sélectionner la méthode appropriée.

5.3 Analyse par mobilité ionique

La spectrométrie de mobilité ionique (IMS, de l'anglais *Ion Mobility Spectrometry*) est une technique d'analyse qui consiste à introduire des ions dans une cellule de dérive dans laquelle ils seront attirés par un champ électrique mais retenus par des collisions avec un gaz. Selon la conformation de l'ion, les collisions seront différentes et donc son temps d'arrivée dans l'analyseur MS sera différent. Le Schéma 8 montre deux ions introduits dans une cellule de dérive, l'un plutôt compact et l'autre étendu. L'ion compact arrive en premier au niveau du détecteur.

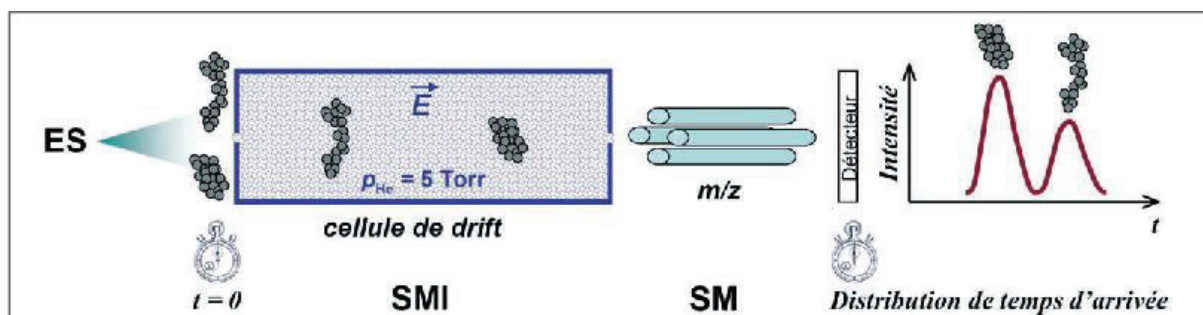


Schéma 8 Principe de la mobilité ionique (extrait de [25]).

Il existe différents appareils commerciaux de mobilité ionique. L'instrument utilisé à l'INRAE de Nantes est le *Select Series Cyclic IMS™* commercialisé par Waters et équipé d'une cellule TWIMS (*Travelling Wave Ion Mobility Spectrometry*).

Ce chapitre contient les principes de chaque méthode utilisée. Les détails expérimentaux et/ou théoriques relatifs à chaque étude seront décrits dans les chapitres correspondants.

6 Bibliographie

- 1 R. L. Woodin, D. S. Bomse and J. L. Beauchamp, *J. Am. Chem. Soc.*, 1978, **100**, 3248–3250.
- 2 C. A. Wight and J. L. Beauchamp, *J. Am. Chem. Soc.*, 1981, **103**, 6499–6501.
- 3 J. Oomens, A. J. A. van Roij, G. Meijer and G. von Helden, *Astrophys. J.*, 2000, **542**, 404–410.
- 4 N. C. Polfer, J. Oomens and R. C. Dunbar, *Phys. Chem. Chem. Phys.*, 2006, **8**, 2744.
- 5 S. H. Yoon, J. Chamot-Rooke, B. R. Perkins, A. E. Hilderbrand, J. C. Poutsma and V. H. Wysocki, *J. Am. Chem. Soc.*, 2008, **130**, 17644–17645.
- 6 L. M. Aleese, A. Simon, T. B. McMahon, J.-M. Ortega, D. Scuderi, J. Lemaire and P. Maître, *Int. J. Mass Spectrom.*, 2006, **249–250**, 14–20.
- 7 J. Oomens, N. Polfer, D. T. Moore, L. van der Meer, A. G. Marshall, J. R. Eyler, G. Meijer and G. von Helden, *Phys. Chem. Chem. Phys.*, 2005, **7**, 1345.
- 8 J. Martens, G. Berden, R. E. van Outersterp, L. A. J. Kluijtmans, U. F. Engelke, C. D. M. van Karnebeek, R. A. Wevers and J. Oomens, *Sci. Rep.*, , DOI:10.1038/s41598-017-03387-4.
- 9 C. J. Gray, L. G. Migas, P. E. Barran, K. Pagel, P. H. Seeberger, C. E. Eyers, G.-J. Boons, N. L. B. Pohl, I. Compagnon, G. Widmalm and S. L. Flitsch, *J. Am. Chem. Soc.*, 2019, **141**, 14463–14479.
- 10 P. Çarçabal, R. A. Jockusch, I. Hünig, L. C. Snoek, R. T. Kroemer, B. G. Davis, D. P. Gamblin, I. Compagnon, J. Oomens and J. P. Simons, *J. Am. Chem. Soc.*, 2005, **127**, 11414–11425.
- 11 M. Okumura, L. I. Yeh, J. D. Myers and Y. T. Lee, *J. Chem. Phys.*, 1986, **85**, 2328–2329.
- 12 M. Z. Kamrath, E. Garand, P. A. Jordan, C. M. Leavitt, A. B. Wolk, M. J. Van Stipdonk, S. J. Miller and M. A. Johnson, *J. Am. Chem. Soc.*, 2011, **133**, 6440–6448.
- 13 N. Khanal, C. Masellis, M. Z. Kamrath, D. E. Clemmer and T. R. Rizzo, *Anal. Chem.*, 2017, **89**, 7601–7606.
- 14 E. Mucha, A. I. González Flórez, M. Marianski, D. A. Thomas, W. Hoffmann, W. B. Struwe, H. S. Hahm, S. Gewinner, W. Schöllkopf, P. H. Seeberger, G. von Helden and K. Pagel, *Angew. Chem. Int. Ed.*, 2017, **56**, 11248–11251.
- 15 M. Yamashita and J. B. Fenn, *J. Phys. Chem.*, 1984, **88**, 4451–4459.
- 16 O. Yeni, B. Schindler, B. Moge and I. Compagnon, *The Analyst*, 2022, **147**, 312–317.
- 17 B. Schindler, G. Laloy-Borgna, L. Barnes, A.-R. Allouche, E. Bouju, V. Dugas, C. Demesmay and I. Compagnon, *Anal. Chem.*, 2018, **90**, 11741–11745.
- 18 B. Schindler, A. D. Depland, G. Renois-Predelus, G. Karras, B. Concina, G. Celep, J. Maurelli, V. Lorient, E. Constant, R. Bredy, C. Bordas, F. Lépine and I. Compagnon, *Int. J. Ion Mobil. Spectrom.*, 2017, **20**, 119–124.
- 19 L. Barnes, B. Schindler, A.-R. Allouche, D. Simon, S. Chambert, J. Oomens and I. Compagnon, *Phys. Chem. Chem. Phys.*, 2015, **17**, 25705–25713.
- 20 O. Yeni, A. Gharbi, S. Chambert, J. Rouillon, A.-R. Allouche, B. Schindler and I. Compagnon, *Phys. Chem. Chem. Phys.*, 2022, **24**, 1016–1022.
- 21 M.-P. Gaigeot, *Phys Chem Chem Phys*, 2010, **12**, 3336–3359.
- 22 B. Schindler, L. Barnes, G. Renois, C. Gray, S. Chambert, S. Fort, S. Flitsch, C. Loison, A.-R. Allouche and I. Compagnon, *Nat. Commun.*, 2017, **8**, 1–7.
- 23 B. Schindler, G. Renois-Predelus, N. Bagdadi, S. Melizi, L. Barnes, S. Chambert, A.-R. Allouche and I. Compagnon, *Glycoconj. J.*, 2017, **34**, 421–425.
- 24 B. Favreau, O. Yeni, S. Ollivier, J. Boustie, F. L. Dévéhat, J.-P. Guégan, M. Fanuel, H. Rogniaux, R. Brédy, I. Compagnon, D. Ropartz, L. Legentil and V. Ferrières, *J. Org. Chem.*, 2021, **86**, 6390–6405.
- 25 V. Gabelica, *Spectra Anal.*, 2006, **251**, 21–27.

CHAPITRE 2 :

Identification par IRMPD de la taille du cycle du galactose sur de petits modèles synthétiques

<u>1</u>	Introduction	34
	1.1 Isomérie	34
	1.2 Molécule neutre	34
	1.3 Cycle flexible	35
	1.4 Echantillon naturel	35
<u>2</u>	Détails expérimentaux	36
<u>3</u>	Détermination par IRMPD de la taille du cycle sur des modèles monosaccharidiques	37
	3.1 Systèmes étudiés	37
	3.2 Spectres MS/MS	38
	3.3 Spectres IRMPD	39
	3.4 Conclusion	42
<u>4</u>	Création d'une base de données avec des disaccharides possédant ou non du galactofuranose	
	43	
	4.1 Systèmes étudiés	43
	4.2 Spectres MS/MS	44
	4.3 Spectres IRMPD	45
	4.4 Conclusion	48
<u>5</u>	Conclusion	49
<u>6</u>	Bibliographie	50

1 Introduction

Le galactose peut exister sous forme de cycle à six membres (galactopyranose) ou à cinq membres (galactofuranose). Ces deux formes, représentées dans le Schéma 1, partagent la même formule brute $C_6H_{12}O_6$ ($M=180$ g/mol) et sont donc des isomères.

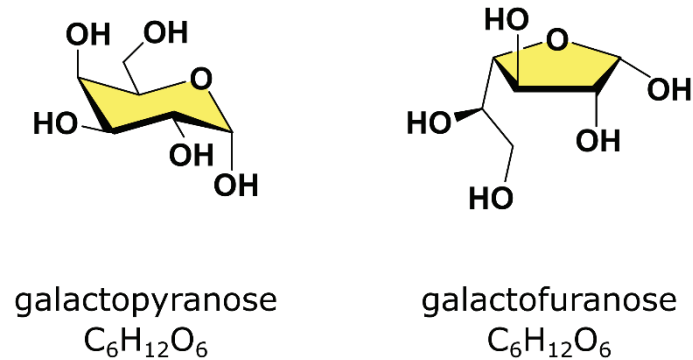


Schéma 1 Structures du galactopyranose et du galactofuranose représentées par la projection Haworth.

La distinction de ces deux formes représente un véritable défi analytique au vu des contraintes existantes. Le Schéma 2 reprend les méthodes qui pourraient être employées, chacune encadrée par ses limitations. Les parties suivantes détailleront ces différentes contraintes.

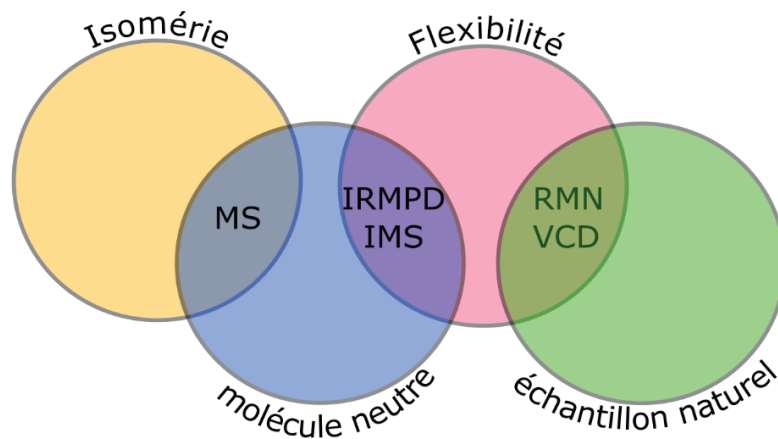


Schéma 2 Contraintes des différentes méthodes analytiques.

1.1 Isomérisation

Les galactofuranose et galactopyranose sont des isomères et sont donc difficiles à distinguer par spectrométrie de masse seule (voir Introduction).

1.2 Molécule neutre

Certains sucres possèdent des groupements basiques, comme le groupe NH de la glucosamine. Cette dernière se protone et peut donc facilement être détectée par spectrométrie de masse (électrospray en mode positif) sous sa forme $[M+H]^+$. D'autres sucres sont composés de groupements acides, c'est le cas des glycosaminoglycane et de leurs groupements sulfates. On peut alors les détecter déprotonés ($[M-H]^-$) en basculant la source électrospray en mode négatif.¹

Le galactose (sous forme furanose ou pyranose) étant une molécule neutre, sa protonation ou déprotonation s'avère difficile. Il existe des alternatives d'ionisation :

1. Introduction

- Ajout d'un groupement basique B qui pourra à la fois se complexer au galactose et se protoner, il sera donc le porteur de proton du complexe $[\text{Gal}+\text{B}+\text{H}]^+$.
- Ajout d'un groupement acide A qui pourra à la fois se complexer au galactose et se déprotoner, il sera donc le porteur de charge du complexe $[\text{Gal}+\text{A}-\text{H}]^-$.
- Ajout d'un métal M alcalin (lithium, sodium, potassium, rubidium...) ou alcalino-terreux (calcium, magnésium...) pour avoir un complexe de type $[\text{Gal}+\text{M}]^{n+}$ ou $[\text{Gal}+\text{M}]^{n-}$.

1.3 Cycle flexible

Le cycle furanose est réputé comme étant très flexible,² c'est-à-dire qu'il existe sous plusieurs conformères et/ou qu'il y a une dynamique entre ces conformères. Ce manque de rigidité rend difficile l'analyse par des techniques spectroscopiques (RMN, VCD et IRMPD). Les spectres IRMPD ont tendance à être congestionnés lorsque plusieurs conformères co-existent. En effet, un élargissement des bandes se produit lorsque l'échelle de la dynamique conformationnelle est plus rapide que celle de la mesure. La flexibilité du galactofuranose complique ainsi son identification structurale.

1.4 Échantillon naturel

Différents travaux ont néanmoins permis d'avancer sur l'identification de la taille du cycle du galactose, notamment par spectroscopie RMN qui est considérée comme la méthode de référence pour l'analyse structurale des glycanes. Une étude menée par Lowary en 2013 vérifie que cette méthode analytique est sensible à la taille du cycle du galactose.⁵ En effet, les valeurs des déplacements chimiques et des constantes de couplage dépendent des atomes d'hydrogène, ces derniers sont disposés différemment pour le furanose et le pyranose. Par ailleurs, nos collaborateurs chimistes ont recours à la RMN pour valider la structure de leurs composés synthétiques. Les spectres RMN sont disponibles dans les publications du consortium ALGAIMS.^{4,5} En 2017, Taniguchi a caractérisé plusieurs monosaccharides furanosides par dichroïsme circulaire vibrationnel (VCD).⁶

Ces méthodes sont très performantes pour identifier la taille du cycle du galactose dans des échantillons simples. Mais, comme indiqué plus tôt, elles sont plus difficiles à appliquer sur des extraits naturels. Or notre visée est la détection de galactofuranose dans des extraits de lichen.

Les méthodes avancées de spectrométrie de masse sont quant à elles pertinentes pour l'étude des mélanges complexes car, comme indiqué dans la partie Méthodes, elles séparent des composés par leur rapport masse-sur-charge, peuvent être couplées avec la chromatographie liquide et nécessitent peu d'échantillons. Deux méthodes de MS avancées (IRMPD et IMS) ont été testées dans le projet ALGAIMS pour identifier la taille du cycle du galactose. Des premiers résultats d'IRMPD sur le Gal β NAc publiés par l'équipe montrent le potentiel de cette méthode pour ce projet.⁷

Dans ce second chapitre seront présentés les résultats IRMPD de petits modèles synthétiques dont les structures pourraient être retrouvées dans des polysaccharides de lichen. Deux études de complexité croissante sont exposées : d'abord sur des monosaccharides isomères (se distinguant par leur anomérie et la taille du cycle) puis sur des disaccharides auxquels s'ajoute la régioisométrie. Ces analyses vérifieront que la structure furanosidique est conservée après le passage en phase gazeuse. Ensuite, les spectres acquis serviront à créer une base de données constituée des rapports m/z et des spectres IRMPD. Chacune de ces deux études a fait l'objet d'un article.^{4,5}

2 Détails expérimentaux

Le dispositif instrumental utilisé pour ce chapitre est le LTQ-OPO (voir page 17).

Les différents composés analysés dans cette partie (mono- et disaccharides) ont été préparés à une concentration de 1 µg/mL dans un mélange eau/méthanol (50/50). Trois échantillons ont été préparés pour chaque composé. On a ajouté au premier 0,1% de sulfate d'ammonium pour obtenir des complexes $[M+NH_4]^+$, au second 0,1% de chlorure de lithium (LiCl) pour former des complexes $[M+Li]^+$ et au troisième 0,1% de chlorure de rubidium (RbCl) pour les complexes $[M+Rb]^+$. Quatre monosaccharides et huit disaccharides ont été analysés dans trois états de charge, on compte donc un total de 36 échantillons.

Ces échantillons sont introduits dans le spectromètre de masse à un débit de 3 µL/min à l'aide d'un pousse-seringue. Les conditions d'électrospray pour le passage en phase gazeuse et l'ionisation des molécules sont les suivantes : tension d'environ 4 kV et température à 200 °C.

Les ions formés sont guidés dans le piège à ions et l'ion d'intérêt est isolé selon son rapport m/z. Les valeurs m/z attendues des différents ions sont indiquées dans le Tableau 1. Pour les monosaccharides (M=194), le rapport m/z sélectionné est à 212 (complexe au NH_4^+), 201 (complexe au Li^+) ou 279 (complexe au Rb^+) avec une fenêtre d'isolement de +/- 2,5 Da. Les mêmes adduits ont été testés sur les disaccharides (M=342), donnant des complexes avec un rapport m/z de 360 (NH_4^+), 349 (Li^+) ou 327 (Rb^+).

Tableau 1 Rapports m/z des différents composés analysés (mono- et disaccharide) selon l'adduit (NH_4^+ , Li^+ ou Rb^+).

Rapport m/z	NH_4^+	Li^+	Rb^+
Monosaccharide GalOMe (M=194)	212	201	279
Disaccharide Gal-Man (M=342)	360	349	327

Pour les spectres MS/MS, le temps d'isolement est de 30 ms et l'hélium est utilisé pour la fragmentation CID.

Pour les spectres IRMPD, les ions sont isolés plus longtemps pour permettre l'irradiation. Ils sont irradiés pendant 700 ms pour chaque séquence MS. La gamme spectrale va de 2700 à 3700 cm^{-1} (sauf indication contraire) avec une vitesse de balayage de 0,4 cm^{-1}/s . Il faut donc 42 minutes pour réaliser un spectre IRMPD. La répétabilité des spectres IRMPD est régulièrement vérifiée sur les standards glucosamine et tryptophane. En plus de cela, certains spectres comme ceux des disaccharides avec adduit lithium ont été réalisés deux fois.

Le taux de photofragmentation est calculé pour chaque séquence MS. Un point sur le tracé des spectres IRMPD correspond à une séquence MS. Une ligne continue correspondant à la moyenne glissante de 5 points est ajoutée pour une meilleure visualisation.

3 Détermination par IRMPD de la taille du cycle sur des modèles monosaccharidiques

3.1 Systèmes étudiés

Une étude par IRMPD a été menée sur des monosaccharides (les plus petites unités saccharidiques) de galactose dont les structures sont présentées en Schéma 3 avec le cycle coloré en jaune. Ce cycle peut avoir deux tailles possibles :

- Furanoses (les deux molécules de gauche) : cycles à 5 membres, avec quatre atomes de carbone et un atome d'oxygène
- Pyranoses (celles de droite) : cycles à 6 membres, avec cinq atomes de carbone et un atome d'oxygène

Pour bloquer la taille du cycle, l'oxygène en position anomère a été méthylé. Cette modification chimique est indispensable car ces formes interconvertissent en solution dans le cas du galactose naturel et ne peuvent pas être isolées. Selon l'orientation du groupement O-méthyle, deux anomères sont possibles :

- Anomères α (molécules du haut) le groupe OMe du carbone anomère est *cis* par rapport à l'oxygène du cycle, dirigé vers le bas dans la projection Haworth
- Anomères β (molécules du bas) : le groupe OMe du carbone anomère est *trans* par rapport à l'oxygène du cycle, dirigé vers le haut dans la projection Haworth

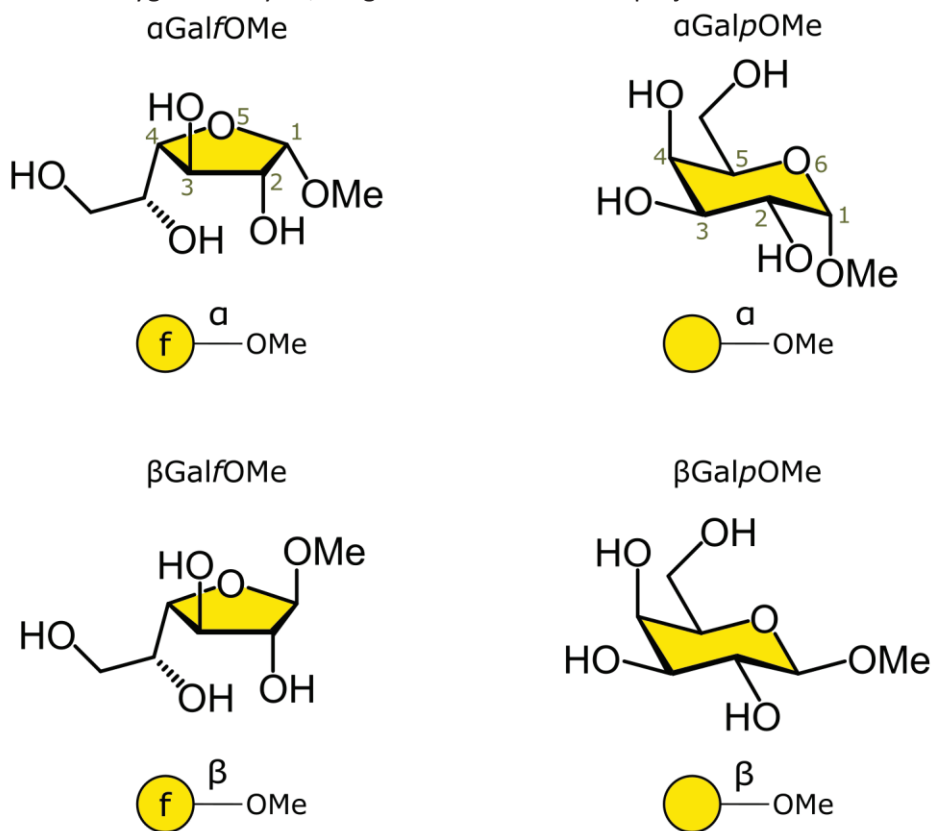


Schéma 3 Structure des monosaccharides α GalfOMe, β GalfOMe, α GalpOMe et β GalpOMe (chacun représenté par la projection Haworth en haut et la nomenclature SNFG en bas).

Ces quatre composés synthétiques sont le α -D-galactofuranoside de méthyle, β -D-galactofuranoside de méthyle, α -D-galactopyranoside de méthyle, β -D-galactopyranoside de méthyle. Ils seront respectivement nommés α GalfOMe, β GalfOMe, α GalpOMe et β GalpOMe pour faciliter la lecture.

Les deux représentations (projection Haworth et SNFG) sont incluses dans ce schéma pour que le lecteur puisse identifier les différences au niveau moléculaire. Par soucis de simplicité, les monosaccharides seront représentés par leurs symboles SNFG.

La détection en spectrométrie de masse nécessite des composés chargés. L'ionisation par électrospray permet d'ajouter différents types de charges et même de former des ions plusieurs fois chargés. Le choix de l'état de charge est critique en IRMPD car il influence fortement les empreintes spectroscopiques. Comme mentionné dans la partie précédente, les hexoses sont difficiles à protoner car ils ne possèdent pas de groupement acide ou basique. En alternative, nous nous sommes intéressés à l'ajout d'adduit NH_4^+ , Li^+ ou Rb^+ . Nous analyserons donc des ions portant une seule charge.

3.2 Spectres MS/MS

La Figure 1 montre les spectres MS/MS des quatre isomères obtenus avec différents adduits. Chaque ligne correspond à un isomère (de haut en bas : α GalfOMe, β GalfOMe, α GalpOMe et β GalpOMe) et chaque colonne à un adduit (de gauche à droite : NH_4^+ , Li^+ et Rb^+).

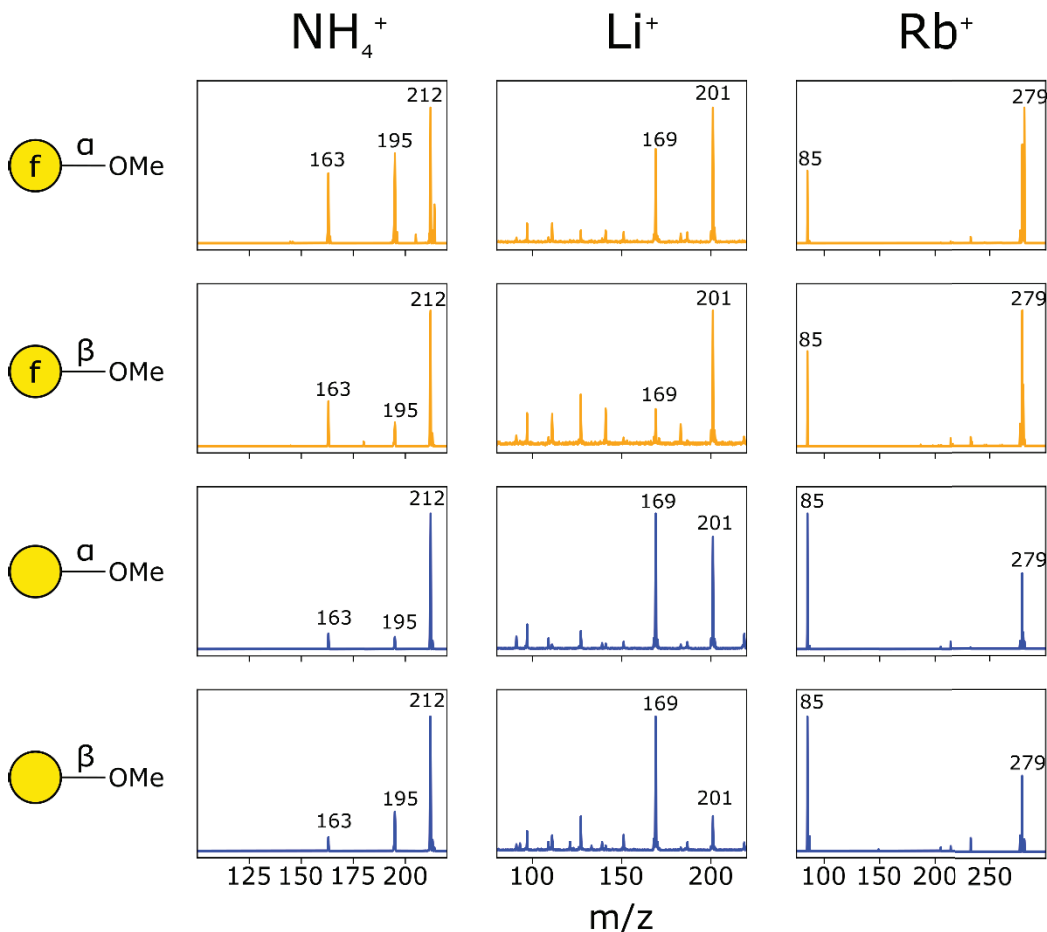


Figure 1 Spectres CID des monosaccharides avec adduit NH_4^+ (gauche), Li^+ (milieu) et Rb^+ (droite).

3. Détermination par IRMPD de la taille du cycle sur des modèles monosaccharidiques

Avec l'adduit NH_4^+ , les quatre isomères de galactose sont détectés à m/z 212. Sous fragmentation par CID, leur fragment principal m/z 195 correspond à la perte neutre du NH_3 . Le second fragment observé à m/z 163 a une perte cumulée du NH_3 et de MeOH .

Le galactose de méthyle complexé au lithium est détecté à m/z 201 et les spectres MS/MS montrent un fragment à m/z 169 correspondant à la perte du groupement MeOH . La liaison la plus faible du monosaccharide galactose de méthyle est donc C1-O1Me . D'autres fragments minoritaires sont également visibles sur les spectres de fragmentation.

Le fragment majoritaire des complexes $[\text{GalOMe}+\text{Rb}]^+$, détectés à m/z 279, correspond au rubidium (m/z 85). On a donc une perte neutre de la molécule.

Ainsi pour les trois adduits testés, les fragments sont comparables pour les quatre isomères. On peut noter des différences d'intensité de ces fragments mais elles sont très sensibles aux conditions expérimentales. Cela suggère que la distinction des cycles furanose/pyranose et celle des anomères α/β n'est pas robuste par spectrométrie de masse seule. Nous verrons dans la partie suivante les spectres IRMPD obtenus par couplage MS/IR pour les mêmes ions.

3.3 Spectres IRMPD

Les spectres IRMPD des quatre monosaccharides sont représentés en Figure 2 : à gauche avec l'adduit NH_4^+ , au milieu avec le Li^+ et à droite avec le Rb^+ . Nous pouvons constater par l'allure des spectres que l'adduit influence fortement l'empreinte IR.

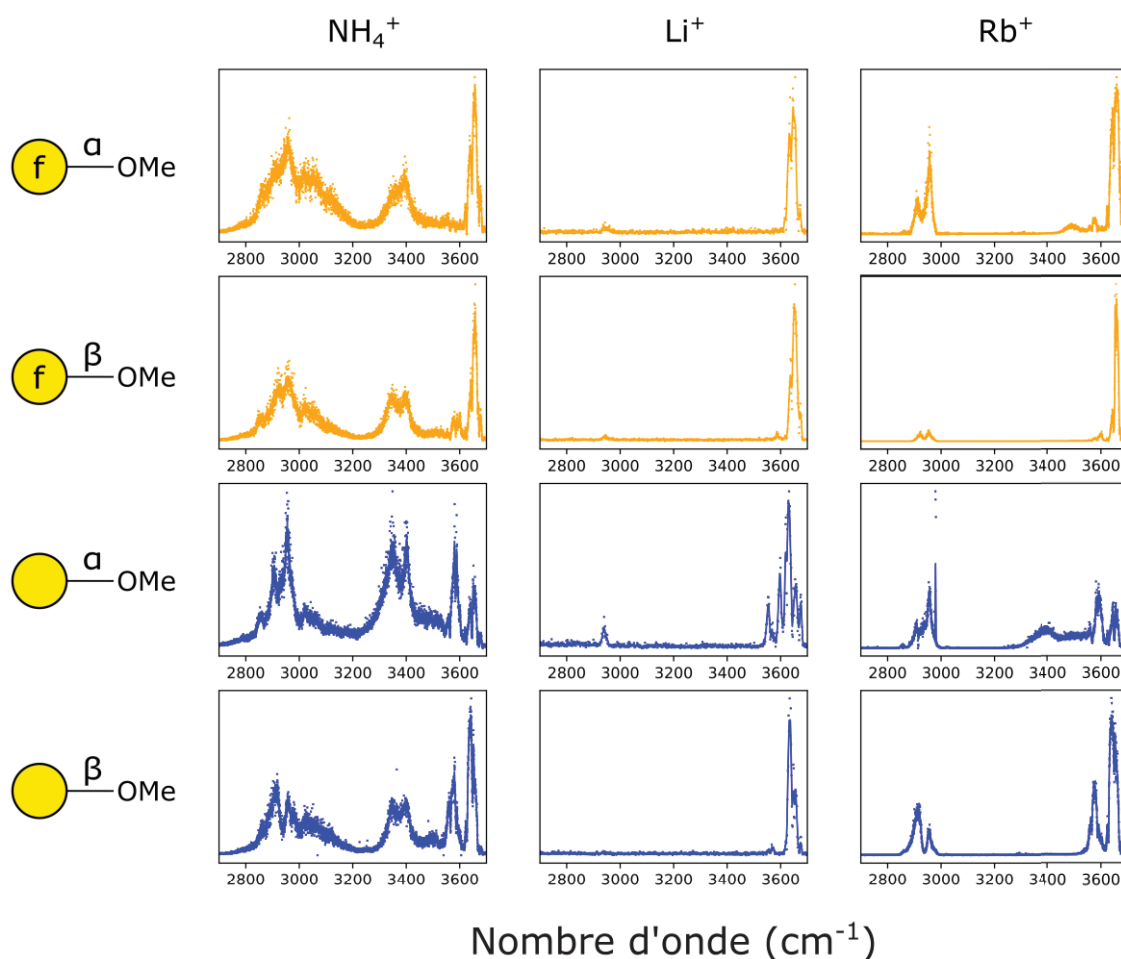


Figure 2 Spectres IRMPD des monosaccharides avec adduit NH_4^+ (gauche), Li^+ (milieu) et Rb^+ (droite).

3.3.1 Spectres obtenus avec l'adduit NH_4^+

Les spectres IRMPD des complexes avec NH_4^+ (Figure 2, à gauche), sont composés de trois régions distinctes : la région des modes CH autour de 2950 cm^{-1} , celle des bandes NH vers 3350 cm^{-1} (le galactose n'ayant pas de groupement NH, celles-ci proviennent de l'adduit NH_4^+) et celle des bandes OH au-dessus de 3400 cm^{-1} . Cette dernière région est composée d'une bande intense à 3650 cm^{-1} pour les deux galactofuranoses (en orange) alors que les galactopyranoses expriment également une seconde bande OH intense à 3580 cm^{-1} . Cette caractéristique est donc très pertinente pour distinguer la taille du cycle du galactose, indépendamment de l'anomérie.

Ces spectres apportent également l'information sur l'anomérie. Alors que les bandes CH sont assez similaires pour les deux galactofuranoses, on peut noter que l'anomère α présente une bande NH vers 3400 cm^{-1} et que l'anomère β possède deux bandes NH mal résolues (à 3380 et 3400 cm^{-1}). La zone des OH est également informative : les deux anomères galactofuranosides ont une bande très intense vers 3650 cm^{-1} mais on peut noter que la forme β présente également une faible bande à 3580 cm^{-1} . Ces deux détails facilitent ainsi la distinction entre les deux anomères pour les formes furanose (en orange).

Les deux anomères pyranose (en bleu) se distinguent par l'intensité relative des deux bandes OH. La forme α a sa bande à 3580 cm^{-1} plus intense que celle à 3650 cm^{-1} , alors que l'inverse est vrai pour l'anomère β . Les bandes OH facilitent donc la distinction des quatre isomères.

3.3.2 Spectres obtenus avec l'adduit Li^+

Sur la colonne du milieu figurent les spectres IRMPD des quatre monosaccharides avec l'adduit Li^+ . Une première observation est que les régions d'absorption sont plus compactes que les spectres avec NH_4^+ . Tout d'abord, il y a très peu d'activité vibrationnelle dans la zone des CH : les trois premiers isomères présentent une bande très faible à 2950 cm^{-1} .

Ensuite, étant donné que le galactose de méthyle ne comporte pas de groupe NH, il n'y a pas d'activité vibrationnelle dans la gamme des NH.

Enfin, la région des OH (au-dessus de 3500 cm^{-1}) est assez similaire mais elle toutefois informative sur le type d'isomérisation. Nous avons représenté dans la Figure 3 les mêmes spectres avec un agrandissement de la gamme des OH. Le spectre de $\alpha\text{Gal}f\text{OMe}$ est caractérisé par une bande unique et intense détectée à 3650 cm^{-1} . Son analogue $\beta\text{Gal}f\text{OMe}$ présente, en plus de cette bande, un pic secondaire à 3590 cm^{-1} . Le motif OH de $\alpha\text{Gal}p\text{OMe}$ est très caractéristique avec cinq bandes entre 3550 et 3680 cm^{-1} . Enfin, le spectre IRMPD de $\beta\text{Gal}p\text{OMe}$ montre une bande secondaire à 3570 cm^{-1} en plus de sa bande OH intense à 3650 cm^{-1} .

3. Détermination par IRMPD de la taille du cycle sur des modèles monosaccharidiques

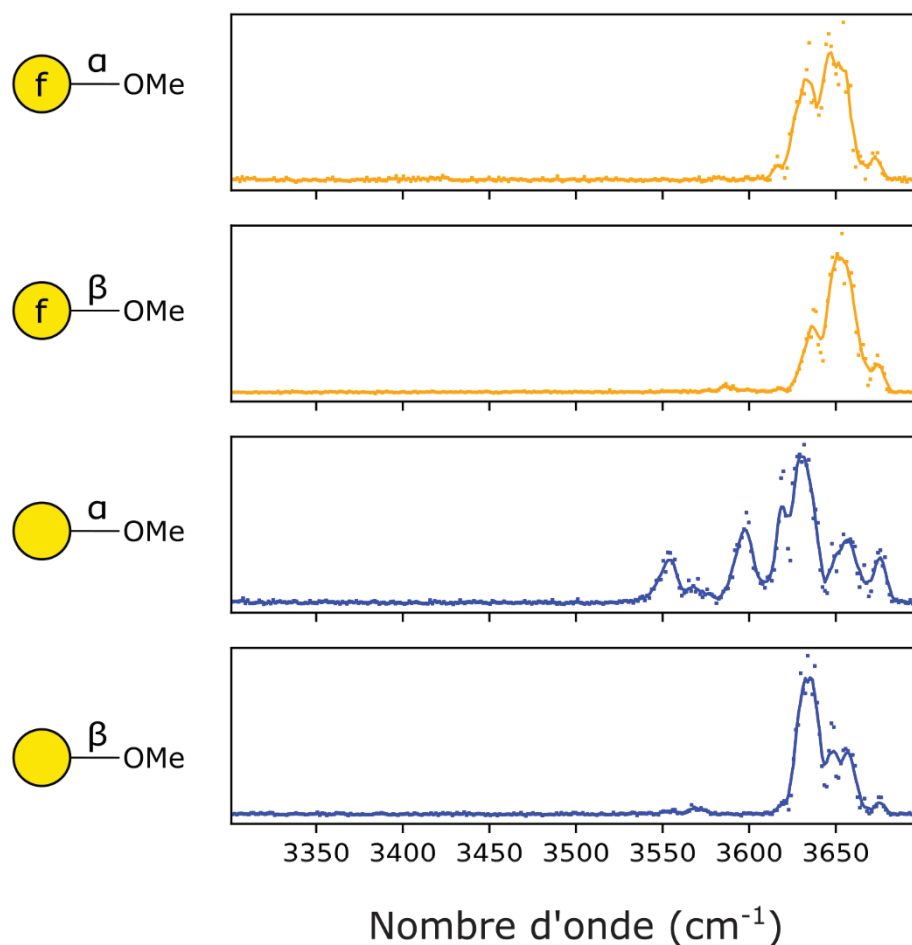


Figure 3 Spectres IRMPD des monosaccharides avec adduit Li^+ dans la gamme $[3300\text{-}3700\text{ cm}^{-1}]$.

Ainsi, bien que ces signatures IR donnent l'impression qu'elles sont très proches les unes des autres, nous constatons que nous pouvons les distinguer avec des bandes spécifiques dans la région des OH.

3.3.3 Spectres obtenus avec l'adduit Rb^+

Les spectres obtenus avec l'ion Rb^+ figurent à droite de la figure. Les spectres des galactopyranosides (en bleu) ont déjà été publiés par une autre équipe.⁸ On distingue la zone des CH entre 2850 et 3000 cm^{-1} et la région des OH entre 3300 et 3650 cm^{-1} . Les quatre isomères montrent deux bandes CH à 2915 et 2950 cm^{-1} . Ces deux bandes sont d'intensité comparable pour $\beta\text{Gal}\rho\text{OMe}$, celle de gauche est plus intense pour $\beta\text{Gal}\rho\text{OMe}$ tandis que celle de droite est plus intense pour les anomères α . Alors que les bandes CH apportent généralement peu d'informations, elles sont intéressantes et bien résolues ici.

La région des OH est également très informative. Les formes furanose présentent une bande intense vers 3650 cm^{-1} et une bande mineure plus décalée vers le rouge (à 3575 cm^{-1} pour la forme α et à 3600 cm^{-1} pour la forme β). $\alpha\text{Gal}f\text{OMe}$ montre également une large bande peu intense de 3450 à 3530 cm^{-1} . Les formes pyranose ont deux bandes OH intenses à 3590 et 3650 cm^{-1} pour l'anomère α et à 3570 et 3545 cm^{-1} pour l'anomère β . La forme α présente également une faible activité vibrationnelle dans la gamme 3320-3570 cm^{-1} .

Ces spectres sont constitués de deux régions distinctes, chacune suffisante pour distinguer les quatre spectres. L'analyse combinée des deux régions confirme le type d'isomérisation.

3.4 Conclusion

Bien que le galactose soit neutre, il est ionisable avec différents états de charge (NH_4^+ , Li^+ , Rb^+).

Les spectres MS/MS des quatre isomères de galactose avec un même état de charge sont très similaires et donc insuffisants pour les distinguer.

À l'inverse, les signatures IRMPD des quatre isomères sont différentes. Cela est valable pour les trois états de charge testés :

- Les spectres avec NH_4^+ sont riches en informations et fournissent un fort diagnostic furanose/pyranose selon le nombre de bandes OH.
- Des différences, bien que mineures, sont également observées avec l'adduit Li^+ dans la région des OH.
- On retrouve également un fort diagnostic par le nombre de bandes OH avec l'adduit Rb^+ . Les autres bandes OH ainsi que la région des CH facilitent ensuite la distinction des anomères α et β .

Grâce à ces différents spectres IRMPD, nous pouvons tirer deux conclusions. D'une part, la RMN a montré que les quatre isomères pouvaient être séparés en solution. L'IRMPD valide que les quatre structures sont également conservées en ESI-MS. D'autre part, l'IRMPD offre une résolution suffisante pour détecter les différences structurelles. On est donc capable de produire des spectres de référence standard.

Nous avons été agréablement surpris d'aboutir à des spectres distincts malgré la flexibilité du galactofuranose qui ne nuit donc pas à sa résolution structurelle. Les spectres sont différents et présentent en plus un diagnostic fort permettant de différencier les galactofuranoses des galactopyranoses.

Le galactose est retrouvé dans la nature lié à d'autres monosaccharides. Il est ainsi essentiel d'étendre l'étude sur des composés plus proches de ce qui est trouvé dans la nature, ce qui augmente par ailleurs la complexité de nos analyses. Nous verrons dans la prochaine partie les résultats de l'étude menée sur les disaccharides.

4 Création d'une base de données avec des disaccharides possédant ou non du galactofuranose

4.1 Systèmes étudiés

Le galactose, sous forme pyranose ou furanose, peut se trouver lié à d'autres sucres tels le glucose, le mannose, etc. Il existe alors plusieurs isomères pour chaque type de sucre. Étant donné l'effort considérable que représente la synthèse de ces modèles, nous avons choisi de restreindre nos disaccharides au motif β -Galactose-Mannose, fréquemment retrouvé dans les polysaccharides de lichen. Le Schéma 4 montre les huit disaccharides synthétisés pour cette seconde étude. Selon la façon dont se lient ces deux monomères, on obtient des isomères différents appelés régioisomères. On observe sur la première ligne la structure des disaccharides Gal β (1-2)Man, c'est-à-dire avec le galactose qui se lie sur l'oxygène porté par le carbone 2 du mannose. Puis sont représentés successivement les Gal β (1-3)Man, Gal β (1-4)Man et Gal β (1-6)Man. À gauche figurent les disaccharides ayant du galactofuranose et à droite ceux avec du galactopyranose.

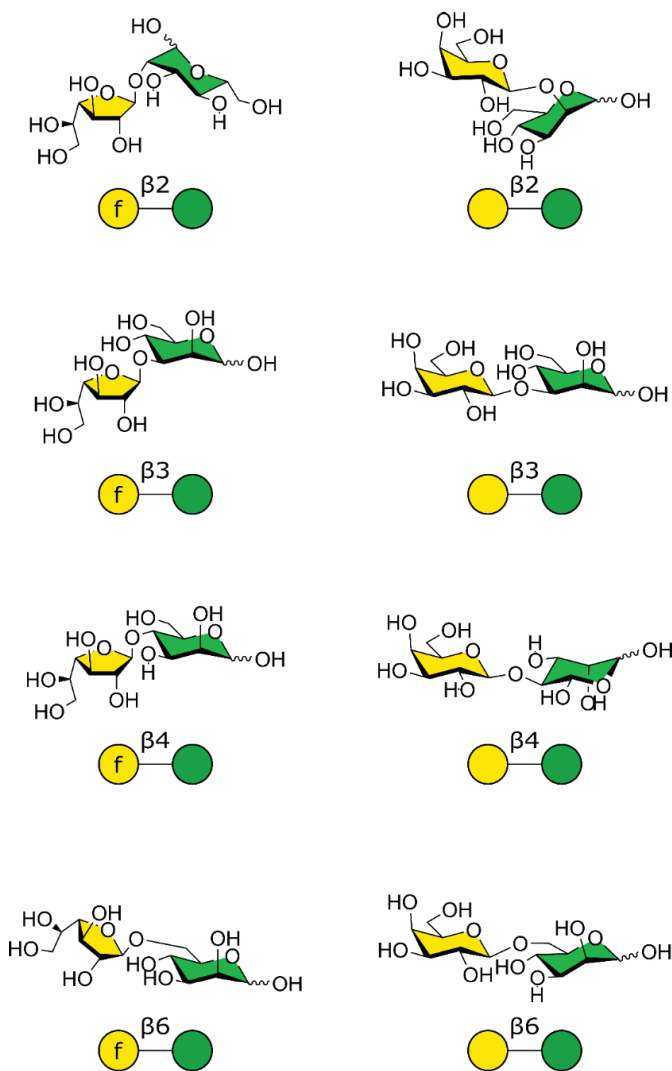


Schéma 4 Structure des disaccharides β -Gal-Man (projection Haworth et nomenclature SNFG).

Tout comme pour les monosaccharides, les disaccharides seront symbolisés selon la nomenclature SNFG sur les figures suivantes.

4.2 Spectres MS/MS

Il est classique en MS/MS d'explorer différents porteurs de charge car les fragments sont différents, ce qui est indicateur pour identifier un ion. Les disaccharides ont été détectés avec les mêmes adduits que les monosaccharides (NH_4^+ , Li^+ et Rb^+). Leurs spectres de fragmentation sont présentés dans la figure ci-dessous.

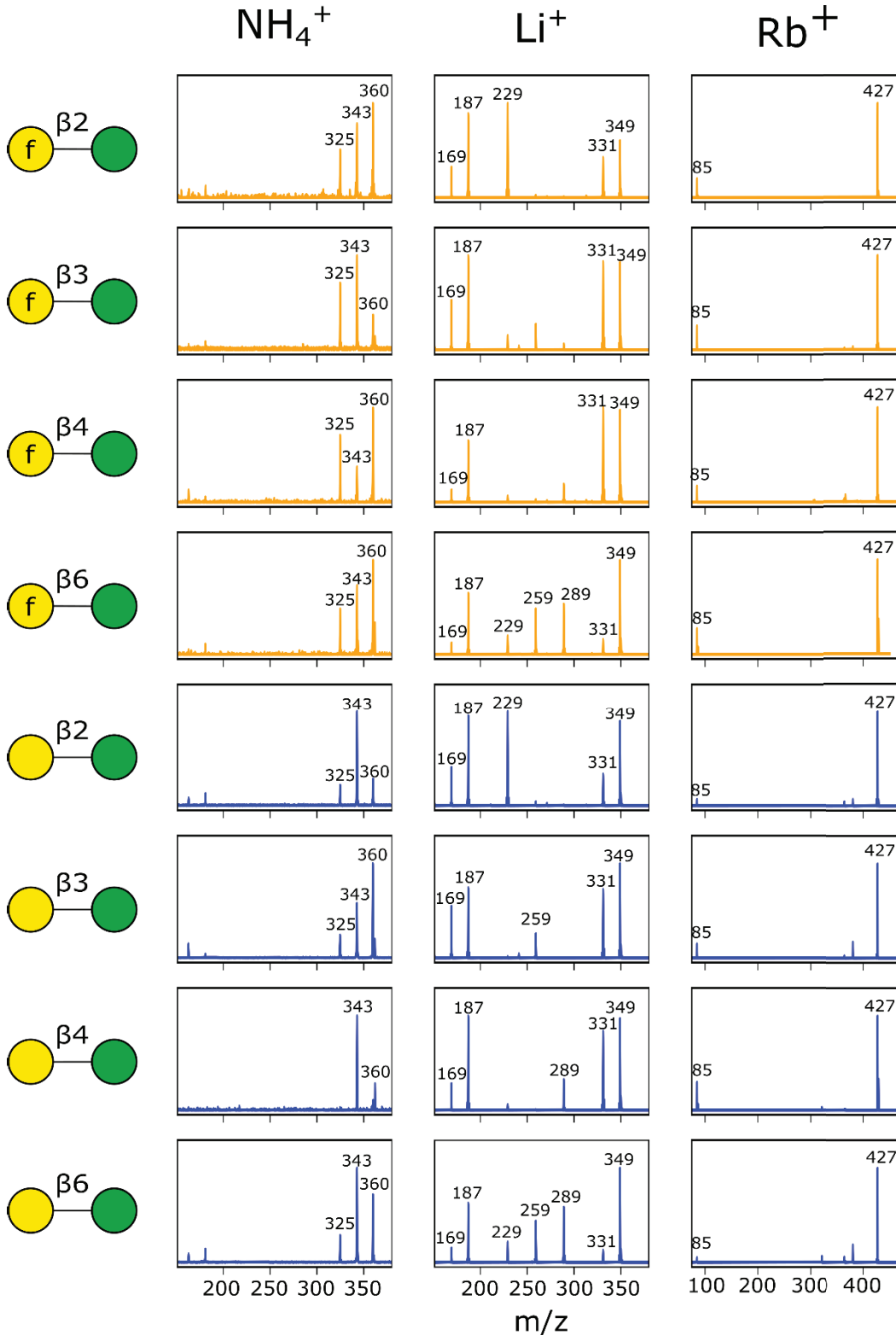


Figure 4 Spectres CID des disaccharides avec adduit NH_4^+ (gauche), Li^+ (milieu) et Rb^+ (droite).

4. Création d'une base de données avec des disaccharides possédant ou non du galactofuranose

Les huit disaccharides sont détectés à m/z 360 lorsqu'ils sont complexés à l'adduit ammonium. Les spectres MS/MS montrent des fragments ayant perdu le groupement NH_3 de l'adduit. Ainsi, les rapports m/z correspondent aux m/z des fragments protonés : le disaccharide protoné intact (m/z 343) ou déshydraté (m/z 325). La fragmentation de certains disaccharides comme Gal β (1-2)Man conduit également à la production de fragments monosaccharidiques (m/z 162 et 180), avec une intensité beaucoup plus faible.

Les spectres MS/MS avec l'adduit lithium sont beaucoup plus riches en fragments. Les rapports m/z des fragments ainsi que leurs correspondances peuvent être consultés dans le Tableau 2. Le fragment m/z 331 (perte de 18 Da) correspond au disaccharide déshydraté. Les fragments m/z 169 et 187 sont obtenus par le clivage de la liaison glycosidique. Les disaccharides ayant une structure symétrique (le galactose et le mannose ont la même masse), le fragment m/z 169 peut aussi bien correspondre au galactose déshydraté qu'au mannose déshydraté. Il en est de même pour le fragment 187 qui porte un groupe OH supplémentaire. Il s'agit donc du galactose intact ou du mannose intact. Les trois autres fragments ont un écart de 30 Da (m/z 229, 259 et 289). Ils correspondent à des fragments obtenus par clivage intra-cyclique avec une perte successive d'un groupe formaldéhyde (CH_2O). Ces derniers fragments sont généralement utilisés en MS/MS pour identifier la régiochimie de la liaison.

Tableau 2 Fragments des disaccharides complexés au lithium et leur correspondance.

Rapport m/z	Type de clivage	Fragment correspondant
331		disaccharide déshydraté
289	intra-cyclique	perte d'un groupement formaldéhyde
259	intra-cyclique	perte de 2 groupements formaldéhyde
229	intra-cyclique	perte de 3 groupements formaldéhyde
187	liaison glycosidique	galactose intact ou mannose intact
169	liaison glycosidique	galactose déshydraté ou mannose déshydraté

Enfin, les disaccharides complexés au rubidium donnent principalement de l'ion Rb^+ (m/z 85), ce qui correspond à une perte neutre du disaccharide.

4.3 Spectres IRMPD

Nous explorerons dans cette partie la sensibilité de la technique IRMPD pour résoudre les deux types d'isomérisation présents dans les modèles disaccharidiques synthétisés : la régiochimie de la liaison Gal-Man et la taille du cycle galactose. Les résultats IRMPD sont tracés en Figure 5.

Les spectres IRMPD ont été acquis avec les mêmes adduits que ceux utilisés précédemment (NH_4^+ , Li^+ et Rb^+). Pour rappel, la région comprise entre 2800 et 3000 cm^{-1} correspond aux modes d'élongation des CH, celle vers 3400 cm^{-1} correspond aux modes NH et enfin les modes OH sont représentés au-dessus de 3400 cm^{-1} .

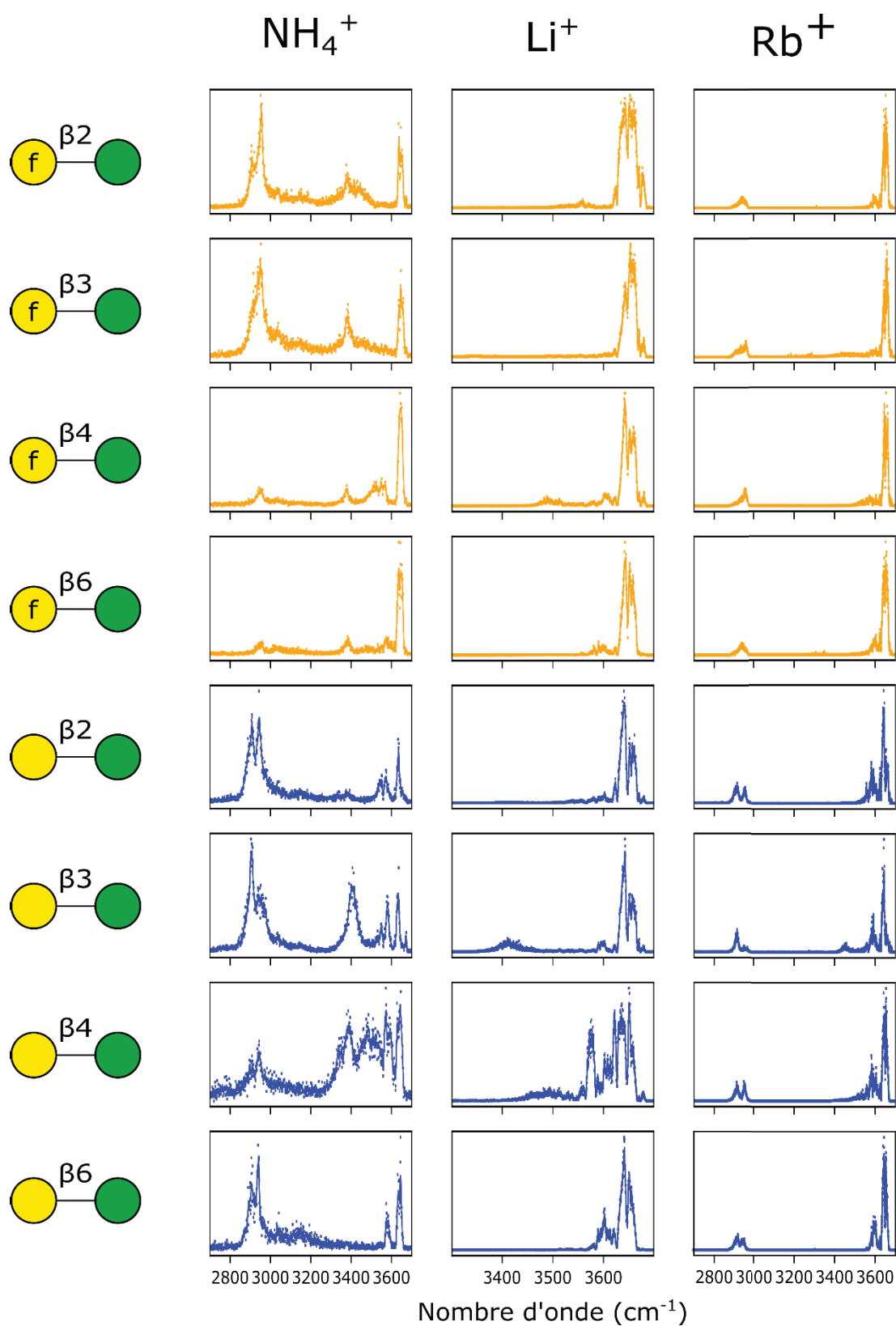


Figure 5 Spectres IRMPD des disaccharides β -Gal-Man avec adduit NH_4^+ (gauche), Li^+ (milieu) et Rb^+ (droite).

4. Création d'une base de données avec des disaccharides possédant ou non du galactofuranose

4.3.1 Spectres obtenus avec l'adduit NH_4^+

La colonne de gauche montre les spectres IRMPD de chaque disaccharide complexé au NH_4^+ . Les huit spectres présentent une bande intense autour de 3650 cm^{-1} . Une seconde bande OH intense est visible à 3575 cm^{-1} uniquement pour les disaccharides contenant du galactopyranose (en bleu). La régioisométrie de la liaison Galp-Man peut ensuite être déterminée selon différents critères : le spectre de Galp β (1-6)Man ne présente pas d'activité vibrationnelle autre que le doublet de OH au-dessus de 3400 cm^{-1} tandis que celui de Galp β (1-2)Man montre une bande supplémentaire à 3547 cm^{-1} ainsi qu'une bande NH peu intense à 3380 cm^{-1} . A cette longueur d'onde, les spectres des deux autres régioisomères Galp β (1-3)Man et Galp β (1-4)Man ont une bande NH large et très intense. Le spectre de Galp β (1-4)Man se distingue des autres par la bande supplémentaire autour de 3500 cm^{-1} , d'intensité modérée.

Les régioisomères avec du galactofuranose (en orange) sont moins distinctifs, mais présentent néanmoins de subtiles différences. En dessous de la bande principale à 3650 cm^{-1} (diagnostique des furanoses), les quatre spectres montrent une bande NH dans la même gamme (entre 3380 et 3390 cm^{-1}). Le spectre IRMPD de Galf β (1-6)Man a une bande de faible intensité à 3580 cm^{-1} . Quant à Galf β (1-4)Man, son spectre IRMPD possède des bandes résolues dans la région 3500 - 3570 cm^{-1} . Les spectres des régioisomères Galf β (1-2)Man et Galf β (1-3)Man ont tous deux un élargissement de leur bande NH allant jusqu'à 3490 cm^{-1} . Ils se différencient par l'intensité relative de cet épaulement qui est plus intense pour Galf β (1-2)Man.

Ainsi, les spectres IRMPD fournissent un diagnostic fort de la taille du cycle : une bande intense à 3650 cm^{-1} est caractéristique des furanoses alors qu'un doublet de bandes à $3575/3650\text{ cm}^{-1}$ est spécifique aux pyranoses. Les signatures IR des disaccharides Galp-Man sont aussi diagnostiques de la régiochimie puisque les quatre spectres sont distincts. La distinction de la régiochimie pour les Galf-Man est moins évidente, en particulier pour Galf β (1-2)Man et Galf β (1-3)Man dont les spectres IRMPD sont assez proches.

Il est à noter que la puissance laser est stable dans la région 3200 - 3700 cm^{-1} mais qu'elle peut varier dans la région 2700 - 3200 cm^{-1} . Cela peut expliquer la variation d'intensité dans la zone des CH (autour de 2900 cm^{-1}). Pour cette raison, les bandes CH n'ont pas été incluses lors des comparaisons de spectres.

4.3.2 Spectres obtenus avec l'adduit Li^+

Comme la gamme des CH avec adduit lithium sur les monosaccharides est peu informative, les spectres disaccharides avec lithium ont été enregistrés dans l'intervalle 3300 - 3700 cm^{-1} . Donc seules les bandes OH, libres ou liées, sont représentées. On remarque que les huit spectres présentent une bande intense entre 3625 et 3665 cm^{-1} . Le spectre IRMPD de Galp β (1-4)Man se distingue par d'autres bandes fortes, notamment à 3620 et 3575 cm^{-1} , et des bandes plus faibles entre 3450 et 3565 cm^{-1} et entre 3580 et 3615 cm^{-1} . Les sept autres isomères peuvent être identifiés par la présence, ou l'absence, de bandes mineures. Ainsi les deux isomères Gal β (1-6)Man se distinguent des autres par l'absence de bande vers 3680 cm^{-1} . Ils ont tous deux une bande entre 3575 et 3625 cm^{-1} , plus intense pour le disaccharide avec le cycle pyranose. Le spectre de Galf β (1-2)Man montre une faible bande à 3555 cm^{-1} , celui de Galf β (1-4)Man de 3480 à 3620 cm^{-1} , celui de Galp β (1-2)Man à 3600 cm^{-1} et celui de Galp β (1-3)Man une bande à 3600 cm^{-1} mais également une bande un peu plus intense à 3400 cm^{-1} .

4.3.3 Spectres obtenus avec l'adduit Rb⁺

Les spectres IRMPD obtenus avec l'adduit rubidium sont tracés à droite en Figure 5. Une première caractéristique est très apparente sur la partie gauche : les spectres des Gal α -Man (en orange) expriment une seule bande CH vers 2940 cm⁻¹ alors que ceux des Gal β -Man (en bleu) en ont deux (vers 2910 et 2950 cm⁻¹). Ainsi, la région des CH présente un critère de diagnostic entre les disaccharides ayant un cycle furanose et ceux ayant un cycle pyranose.

Il n'y a pas d'absorption dans la zone des NH étant donné que la molécule n'a pas de mode NH.

La zone des OH s'étend de 3430 à 3680 cm⁻¹. Les quatre Gal α -Man ont une bande d'absorption très intense de 3630 à 3675 cm⁻¹ ainsi qu'une faible bande de 3550 à 3630 cm⁻¹. L'intensité dans cet intervalle est plus intense et plus large pour les isomères pyranose. Il va de 3550 à 3630 cm⁻¹ pour les régioisomères en β (1-2), β (1-3) et β (1-4). Le spectre de Gal β (1-3)Man possède une bande supplémentaire à 3450 cm⁻¹.

4.4 Conclusion

Les résultats de cette partie indiquent que les disaccharides Gal-Man étudiés sont observables avec les trois adduits testés précédemment sur les monosaccharides.

Le choix de l'adduit est important en MS/MS. En effet, les spectres de fragmentation des huit isomères sont comparables avec certains adduits, mais ils peuvent apporter des informations sur la structure avec d'autres adduits. En particulier, le lithium favorise des fragments intracycliques qui peuvent renseigner sur la régiochimie de la liaison. En revanche, aucun adduit testé ne fournit l'information sur le cycle du galactose.

Bien qu'il y ait plus de composés que dans la partie précédente, les huit spectres IRMPD sont distincts :

- NH₄⁺ : comme pour les monosaccharides, une seule bande OH intense est observée pour les disaccharides contenant du furanose et deux bandes pour leurs analogues avec un cycle pyranose. D'autres éléments des spectres facilitent ensuite la distinction entre les régioisomères.
- Li⁺ : les huit isomères ont une activité intense vers 3650 cm⁻¹, les différentes bandes OH distinguent ensuite les huit spectres.
- Rb⁺ : alors que les spectres obtenus avec l'adduit NH₄⁺ présentent un diagnostic furanose/pyranose dans la région des OH, ceux avec l'adduit Rb⁺ ont un diagnostic fort dans la région des CH. En effet, les disaccharides Gal α -Man expriment une bande CH contre deux pour les Gal β -Man. On remarquera également que les bandes OH liées sont plus intenses pour les disaccharides avec du galactopyranose, ce qui était également observé sur les monosaccharides.

En somme, la MS/MS apporte l'information sur la régiochimie et elle est facile à traiter (présence ou absence de fragments intracycliques). La taille du cycle est en revanche inaccessible par MS/MS. L'IRMPD fournit les deux informations (taille de cycle et régiochimie). Nous pouvons donc utiliser l'IRMPD seule pour avoir la structure complète mais il peut être également intéressant d'inclure la MS/MS en outil complémentaire pour faciliter l'interprétation.

5 Conclusion

La spectrométrie de masse peut faire la distinction entre des composés de masse différente, par exemple les rapports m/z des mono- et disaccharides sont différents. En revanche, elle n'est pas en mesure de différencier les isomères, qu'il s'agisse d'anomères ou d'isomères différents par la taille de cycle.

La spectrométrie de masse en tandem révèle la régiochimie mais est insuffisante pour déterminer la taille du cycle car les fragments obtenus après CID sont comparables entre les furanoses et les pyranoses. Il est à noter que les fragments MS/MS les plus riches sont observés avec l'adduit lithium. Cette information est importante pour l'approche par séquençage qui consiste, rappelons-le, à analyser les fragments obtenus par CID. Nous y reviendrons dans le chapitre 3.

À l'inverse, les signatures IRMPD sont distinctes pour chaque monosaccharide et disaccharide. Ainsi les structures sont conservées après le passage en phase gazeuse et l'IRMPD est sensible à la taille du cycle du galactose, qu'il soit seul ou lié à un autre monosaccharide. Ces différents spectres servent à construire une base de données.

Les spectres avec adduit NH_4^+ sont intéressants car en plus d'être différents, ils présentent un diagnostic fort de la taille du cycle : alors que les composés contenant du galactofuranose présentent une seule bande OH intense, ceux avec du galactopyranose en ont deux.

Les spectres IRMPD présentés dans ce chapitre peuvent être exploités de deux manières différentes. L'approche théorique, avec l'aide d'outils de calcul quantique, contribue à caractériser la conformation adoptée en phase gazeuse et avoir une meilleure compréhension des spectres expérimentaux. Cette première stratégie est employée dans le chapitre 3. L'approche analytique consiste à identifier une molécule inconnue en comparant son spectre IRMPD avec la base de données établie sur des standards connus. On espère ainsi pouvoir identifier ces motifs référencés dans les études de polysaccharides naturels. Cette seconde approche est développée dans le chapitre 4.

6 Bibliographie

- 1 J. Zaia, *Mass Spectrom. Rev.*, 2004, **23**, 161–227.
- 2 H. A. Taha, M. R. Richards and T. L. Lowary, *Chem. Rev.*, 2013, **113**, 1851–1876.
- 3 M. R. Richards, Y. Bai and T. L. Lowary, *Carbohydrate Research*, 2013, **374**, 103–114.
- 4 J. S. Ho, A. Gharbi, B. Schindler, O. Yeni, R. Brédy, L. Legentil, V. Ferrières, L. L. Kiessling and I. Compagnon, *J. Am. Chem. Soc.*, 2021, **143**, 10509–10513.
- 5 B. Favreau, O. Yeni, S. Ollivier, J. Boustie, F. L. Dévéhat, J.-P. Guégan, M. Fanuel, H. Rogniaux, R. Brédy, I. Compagnon, D. Ropartz, L. Legentil and V. Ferrières, *J. Org. Chem.*, 2021, **86**, 6390–6405.
- 6 T. Taniguchi, K. Nakano, R. Baba and K. Monde, *Organic Letters*, 2017, **19**, 404–407.
- 7 B. Schindler, L. Legentil, A.-R. Allouche, V. Ferrières and I. Compagnon, *Physical Chemistry Chemical Physics*, 2019, **21**, 12460–12467.
- 8 W. L. Pearson, C. Contreras, D. Powell, G. Berden, J. Oomens, B. Bendiak and J. R. Eyler, *J. Phys. Chem. B*, 2015, **119**, 12970–12981.

CHAPITRE 3 :

Impact du cycle et du micro-environnement sur la conformation du galactose en phase gazeuse

<u>1</u>	Introduction	52
	1.1 Vocabulaire	52
	1.2 La configuration	52
	1.3 L'environnement	54
	1.4 Etude des conformations par synergie théorie/expérience	55
<u>2</u>	Détails expérimentaux et théoriques	56
<u>3</u>	Impact de la taille du cycle sur la conformation du galactose	57
	3.1 Une recherche conformationnelle similaire pour les deux cycles du galactose	57
	3.2 Calculs des spectres vibrationnels	60
	3.3 Interprétation	70
<u>4</u>	Impact du micro-environnement sur la conformation du galactofuranose	72
	4.1 Préférences conformationnelles du galactofuranose neutre en phase gazeuse	72
	4.2 Interaction du galactofuranose avec une molécule d'eau	77
	4.3 Interaction du galactofuranose avec un ion	82
	4.4 Interprétation	84
<u>5</u>	Conclusion	87
<u>6</u>	Bibliographie	88

1 Introduction

1.1 Vocabulaire

Commençons par définir les mots-clés relatifs à la structure.¹ La structure désigne la constitution, la configuration et la conformation d'une molécule. La constitution d'une molécule fait référence à sa composition en atomes. Deux isomères ont donc la même constitution. Par exemple, le galactopyranose et le galactofuranose ont tous deux la formule brute $C_6H_{12}O_6$. La configuration concerne l'arrangement spatial des atomes, un changement de configuration nécessitant qu'une liaison covalente soit rompue puis reformée. Deux isomères sont donc de configuration différente (par exemple, configuration α et β). Enfin, la conformation fait également référence à l'arrangement spatial des atomes mais aucune liaison covalente n'est rompue pour passer d'un conformère à un autre. En pratique, on s'intéressera à des familles de conformations caractérisées par des conformations de cycle (chaise, bateau...). Les conformères d'une même famille se distinguent par leur réseau de liaisons hydrogène qui dépendent de la rotation des groupements hydroxyle.

La conformation spatiale des sucres est étroitement liée à leurs fonctions biologiques.² Des réactions chimiques et enzymatiques se produisent ou non selon cette conformation tridimensionnelle.^{3,4} Elle dépend de deux facteurs : de la molécule (sa constitution et sa configuration) et de l'environnement qui entoure cette molécule. Comme nous nous intéressons principalement aux effets d'isomérisation dans l'équipe, nous étudions principalement les effets de configuration. Nous détaillerons les deux facteurs qui nous intéressent (la configuration et l'environnement) dans les deux parties suivantes puis nous verrons la stratégie que nous adoptons pour étudier leurs impacts sur la conformation.

1.2 La configuration

Il est connu que la configuration des sucres affecte grandement la conformation. Ainsi selon l'épimérisation, l'anomérisation, la présence de groupements chargés et leur position, la conformation spatiale ne sera pas la même. Le type de configuration considéré ici est la taille du cycle. Nous souhaitons effectivement connaître les préférences conformationnelles du galactose sous ses formes furanose et pyranose. Pour cela, il est indispensable de disposer d'outils pour décrire et visualiser les conformations.

Alors que les projections de Haworth que l'on a utilisées dans le chapitre précédent sont très utiles pour représenter la configuration d'une molécule, elles n'apportent pas d'information sur la conformation tridimensionnelle. En effet, la représentation du furanose est plane alors qu'on a en réalité une torsion du cycle.

1.2.1 Monosaccharides avec un cycle pyranose

Les paramètres de Cremer-Pople décrivent la torsion des cycles pyranosides. Ils sont calculés en fonction des angles de torsion du cycle et servent de coordonnées sphériques. Nous pouvons également visualiser les conformations à l'aide de la projection de Mercator, avec le paramètre θ en axe des abscisses et la valeur θ en axe des ordonnées. C'est cette dernière représentation que nous utiliserons par la suite. Il existe 38 types de conformations décrites dans la nomenclature IUPAC : 2 chaises ($\phi = 0^\circ$ ou 180°), 6 bateaux, 6 bateaux-croisés, 12 enveloppes et 12 demi-chaises.⁵ Les conformations les plus stables sont généralement celles dites en chaise, c'est-à-dire avec un plan formé par quatre atomes (O, C2, C3 et C5) et deux atomes (C1 et C4) en dehors de ce plan, donnant l'impression d'une chaise. On a représenté en Figure 1 les deux familles de conformations en chaise : on parle de conformation de type 4C_1 si l'atome C4 est dirigé vers le haut et C1 vers le bas et *vice versa* pour 1C_4 . Plusieurs conformères peuvent appartenir à une même famille et se distinguent par la

1. Introduction

position des hydroxyles et le réseau des liaisons hydrogène. Dans notre exemple, les conformations 1 et 2 appartiennent à la famille des 4C_1 , leur cycle adopte donc le même type de conformation mais on voit que les hydroxyles sont très différents. Il en est de même pour les conformations 3 et 4 qui appartiennent à la famille des 1C_4 . En pratique, on représentera leurs structures avec un axe parallèle au cycle pour mieux appréhender la position des OH. La torsion du cycle ne sera donc pas visible sous cet angle mais sera précisé à côté de la structure et dans le texte.

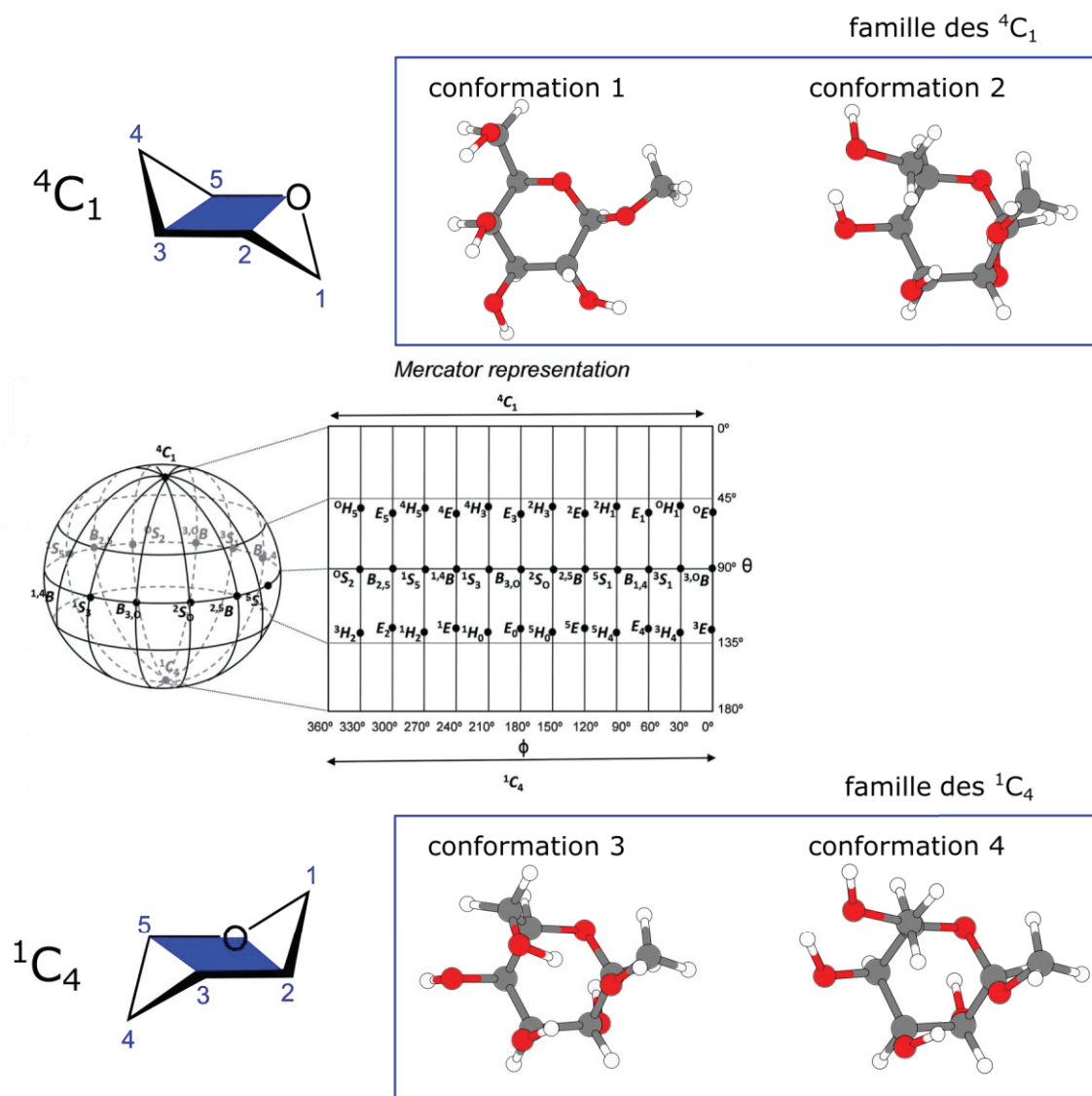


Figure 1 Sphère Cremer-Pople et sa projection Mercator (au milieu, issu de [6]), conformation de cycle 4C_1 et deux conformations de cette famille (en haut), conformation de cycle 1C_4 et deux conformations de cette famille (en bas).

Les nombreux calculs effectués sur les monosaccharides ayant un cycle pyranose ont permis d'accroître l'expertise et la connaissance de l'équipe sur les préférences conformationnelles des pyranoses.^{7,8} On sait qu'ils adoptent généralement des conformations dites chaise, stabilisées par un réseau de liaisons hydrogène et ayant les plus faibles énergies calculées par DFT.

1.2.2 Monosaccharides avec un cycle furanose

Nous utiliserons les paramètres d'Altona-Sundaralingam pour estimer la conformation d'une structure furanose. Ces paramètres servent de coordonnées polaires : la coordonnée radiale représente la phase de pseudorotation ϕ et la coordonnée angulaire correspond à l'amplitude de pseudorotation θ (Figure

2, à gauche). On pourra désigner une conformation selon un point cardinal, par exemple une conformation dite ouest signifie que son paramètre ϕ est proche de 270° .

Le cycle furanose se plisse de deux façons. Il adopte une conformation en enveloppe (E) si un atome est hors du plan. Une conformation est de type 4E si l'atome C4 est au-dessus du plan (Figure 2, à droite) et de type E_4 s'il est en dessous. Si deux atomes sont hors du plan, il s'agit d'une conformation torsadée (T). Une conformation de type 4T_3 (représentée ci-dessous) a l'atome C4 au-dessus du plan et l'atome C3 en dessous du plan formé par O, C1 et C2. On compte au total 10 familles de conformations en enveloppe et 10 familles de conformations torsadées. Comme pour les pyranoses, on peut avoir plusieurs conformations pour chaque famille qui se distinguent par l'orientation des hydroxyles.

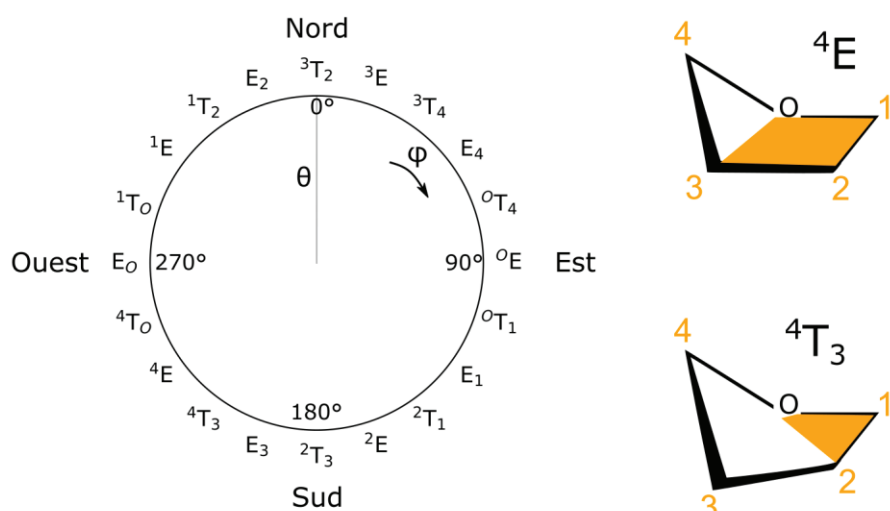


Figure 2 Roue avec les paramètres d'Altona-Sundaralingam (à gauche), structure des conformations 4E et 4T_3 (à droite).

Nos connaissances sur les furanosides sont plus rares d'une part car ils sont moins abondants dans la nature et d'autre part car ils sont plus difficiles à analyser. Une seule étude sur les furanoses avait été menée par l'équipe avant le début de ma thèse. Elle montre que la N-acétylgalactofuranosamine existe en phase gazeuse sous plusieurs conformations situées au sud de la roue (${}^4T_3/E_3/{}^2T_3$).

Alors que le cycle pyranose adopte une seule conformation chaise, on note plusieurs conformations pour le cycle furanose. Ceci est en accord avec la littérature : de faibles barrières énergétiques existent entre les conformères furanosides et facilitent la transition d'une conformation à une autre.⁹ Pour cette raison, le cycle furanose est dit flexible.

1.3 L'environnement

Les préférences conformationnelles d'une molécule dépendent également de l'environnement dans lequel elle se trouve. En solution, la nature du solvant influence fortement la conformation selon les interactions solvant-soluté. En phase gazeuse, la molécule se retrouve isolée de cet environnement. Il n'existe alors plus de liaisons entre les molécules car elles se retrouvent éloignées. On peut en revanche créer des complexes de la molécule (avec des ions métaux, ligands, molécules d'eau...) et étudier les liaisons intermoléculaires. Il est également possible de modifier la température pour évaluer sa contribution sur la conformation. Les études en phase gazeuse présentent l'avantage de pouvoir modifier les conditions locales (micro-environnement) et d'étudier leur influence.¹⁰

1.4 Etude des conformations par synergie théorie/expérience

Les résultats exposés dans le chapitre précédent montrent des spectres IRMPD différents selon l'anomérie, la taille du cycle et la régioisomérie de la molécule. Les bandes infrarouge dépendent ainsi de la configuration de la molécule. Elles sont également liées à sa conformation tridimensionnelle. L'interprétation des spectres vibrationnels nécessite donc une étude conformationnelle de la molécule. Cette étude s'appuie sur de puissants outils théoriques et comprend deux étapes (détaillées page 22). La première étape consiste à obtenir par dynamique moléculaire un grand nombre de candidats avec des conformations différentes. Ensuite, ces conformères sont optimisés et leurs spectres théoriques sont calculés. Les bandes infrarouge de chaque conformation sont individuellement comparées aux bandes expérimentales afin de sélectionner celles expérimentalement présentes.

Ainsi, le spectre expérimental permet de sélectionner les conformations les plus probables par comparaison avec les fréquences calculées. Ces spectres simulés facilitent quant à eux l'interprétation du spectre expérimental. Les deux approches s'enrichissent ainsi mutuellement et la conformation est au cœur de cette synergie (Schéma 1). Cette comparaison apporte en effet un grand nombre d'informations au niveau atomique sur la conformation qu'adopte la molécule en phase gazeuse : la position des atomes, les liaisons intermoléculaires, la position de la charge, la torsion du cycle, etc.

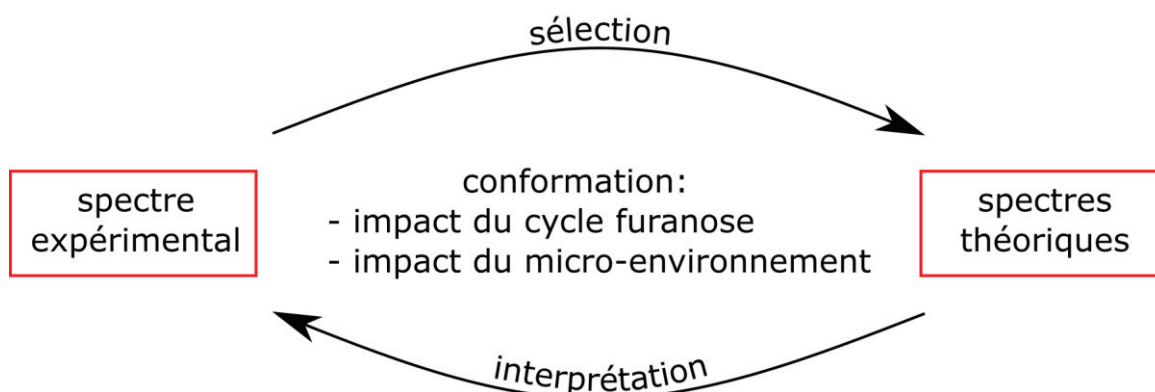


Schéma 1 Enrichissement mutuel des outils expérimentaux et théoriques.

Ce chapitre traite des préférences conformationnelles du galactose. Pour ce faire, nous avons simultanément modifié les paramètres dans les deux approches et examiné l'effet de ce changement sur la conformation. Dans la première partie, nous avons analysé le galactose dans ses deux configurations (pyranose et furanose) pour étudier l'impact de la taille du cycle sur la conformation du galactose. La seconde partie s'intéresse à l'influence du micro-environnement sur la conformation du galactofuranose. Chaque partie a fait l'objet d'une publication.^{11,12}

2 Détails expérimentaux et théoriques

Les spectres IRMPD sont ceux présentés dans le chapitre 2. Les paramètres expérimentaux pour leur acquisition sont décrits page 36.

Les spectres R2PI, UV-UV et UV-IR ont été réalisés avec le dispositif Carlotta de l'ISMO. Le laser UV est un laser à colorant accordable pompé par un Nd :YAG, le laser IR est un OPO/OPA pompé par un Nd :YAG et le spectromètre de masse est un temps de vol vendu par Jordan TOF Products.

Un mélange de β GalpOPh et graphite est étalé sur une barrette (le graphite servant à améliorer l'adhésion). Les molécules sont vaporisées par désorption laser et entraînées dans un jet d'argon pulsé. De l'eau est ajoutée dans l'argon pour les expériences sur le complexe avec la molécule d'eau. Les molécules, isolées ou complexées à une molécule d' H_2O , sont ionisées par ionisation résonante à deux photons (R2PI) et détectées par un spectromètre de masse à temps de vol. La spectroscopie UV-UV sert à différencier les différents conformères par leurs spectres électroniques puis l'IR-UV est utilisée pour l'acquisition de chaque conformère individuellement.

Le Tableau 1 reprend les molécules qui ont fait l'objet d'études théoriques.

Tableau 1 Récapitulatif des molécules étudiées, de leurs adduits respectifs et de la méthode ayant servi à l'étude.

Molécule	Adduit	Méthode
α GalpOMe	NH_4^+	IRMPD
β GalpOMe	NH_4^+	IRMPD
α GalfOMe	NH_4^+	IRMPD
β GalfOMe	NH_4^+	IRMPD
β GalfOMe	Li^+	IRMPD
β GalfOPh	Neutre	IR-UV
β GalfOPh	Neutre complexé à H_2O	IR-UV

La Dynamique Moléculaire est employée pour explorer le potentiel énergétique à la surface de chaque ion avec le potentiel PM7 dans OpenMopac.¹³ Pour chaque structure de départ, une trajectoire de 10 ps à 1000 K a été lancée. Pour les complexes, un piégeage virtuel (de NH_4^+ , Li^+ ou H_2O) a été ajouté en PM7 pour éviter qu'il ne s'évapore durant la trajectoire.¹⁴

Les structures et leurs énergies ont ensuite été affinées avec des calculs DFT à deux niveaux. La première optimisation emploie la méthode B3LYP/6-31G¹⁵⁻¹⁷ et la seconde utilise CAM-B3LYP/6-311++G(2df,2pd).¹⁸ Un facteur de correction établi de manière empirique (à 0,967 pour toutes les études sauf à 0,95 pour le complexe $[M+H_2O]$) a été utilisé pour prendre en compte les effets anharmoniques et pouvoir comparer avec les spectres expérimentaux. Les résultats ont été analysés avec Gabedit.¹⁹

Les simulations de la molécule neutre isolée et celles du complexe $[M+H_2O]$ ont également été réalisées à l'ISMO et ont donné des résultats similaires aux nôtres.

Le calcul des coordonnées (Cremer-Popple pour le cycle pyranose et Altona Sundaralingam pour le cycle furanose) a été effectué avec Gabedit en utilisant les méthodes de calculs décrits dans les références [20,21].

3 Impact de la taille du cycle sur la conformation du galactose

3.1 Une recherche conformationnelle similaire pour les deux cycles du galactose

L'étude théorique a été menée sur les quatre monosaccharides présentés dans le chapitre précédent : methyl α -D-galactofuranoside (α GalfOMe), methyl β -D-galactofuranoside (β GalfOMe), methyl α -D-galactopyranoside (α GalpOMe) et methyl β -D-galactopyranoside (β GalpOMe).

Étant donné que les spectres obtenus avec l'adduit NH_4^+ étaient très caractéristiques, nous avons également opté pour cet état de charge lors de nos simulations. Pour rappel, une des observations importantes que l'on avait eu était la présence d'une bande OH intense à 3650 cm^{-1} pour les furanoses alors que les pyranoses présentaient également une seconde bande OH intense à 3560 cm^{-1} . Ces spectres peuvent être visualisés dans la Figure 3.

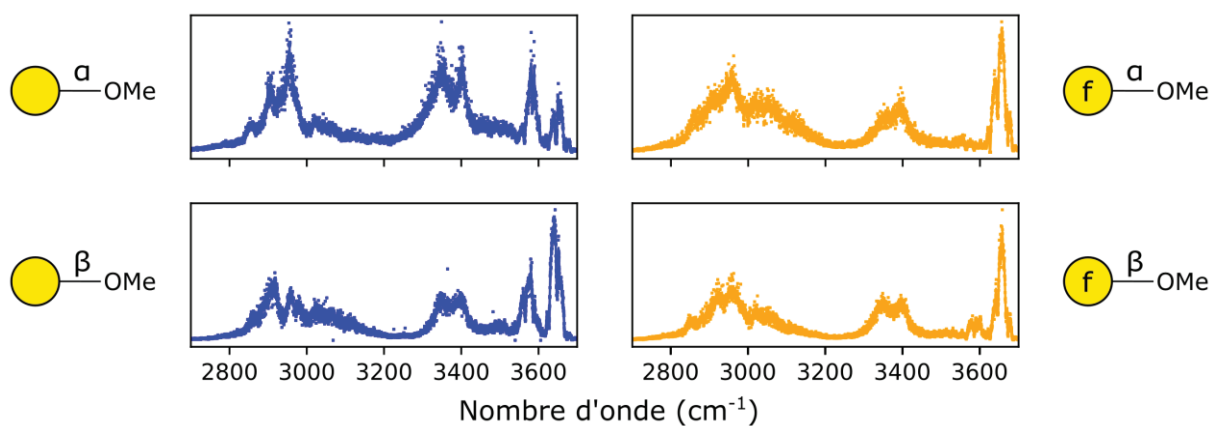


Figure 3 Spectres expérimentaux des monosaccharides complexés au NH_4^+ .

La même méthode a été appliquée pour les deux configurations du galactose : une étape de dynamique moléculaire suivie de calculs DFT.

3.1.1 Recherche conformationnelle des anomères de galactopyranose

Les différentes étapes de la recherche conformationnelle pour les deux anomères galactopyranose se trouvent en Figure 4 (à gauche pour α GalpOMe et à droite pour β GalpOMe). Pour chaque étape, le nombre de conformations est indiquée et les coordonnées de chaque conformation sont représentées. Ainsi, les deux trajectoires moléculaires aboutissent chacune à 20 000 conformations. Une bonne distribution des familles de conformation est observée, validant une exploration conformationnelle complète. La première étape d'optimisation par PM7 a entraîné la suppression d'un grand nombre de doublons : il ne reste plus que 1081 conformations différentes pour α GalpOMe et 1076 pour β GalpOMe. La seconde étape d'optimisation par DFT contribue à raffiner encore plus le nombre de conformations (212 conformations pour α GalpOMe et 256 pour β GalpOMe). À ce stade, nous pouvons constater que les conformations sont réparties sur trois lignes avec $\theta = 0, 90$ ou 180 . Les énergies calculées durant cette étape (avec la base CAMB3LYP) permettent de sélectionner les conformations les plus stables. Le seuil énergétique de $8,3\text{ kJ/mol}$ a été retenu (équivalent à une énergie thermique de 1000 K). Seules trois conformations de chaque anomère ont une énergie inférieure à ce seuil. Elles ont été nommées ap1, ap2 et ap3 (du plus stable au moins stable) pour α GalpOMe et bp1, bp2 et bp3 pour β GalpOMe. Ces six conformères appartiennent à la famille des 4C_1 et diffèrent par la rotation de leurs groupements OH. Leurs structures sont représentées page 61 et 63.

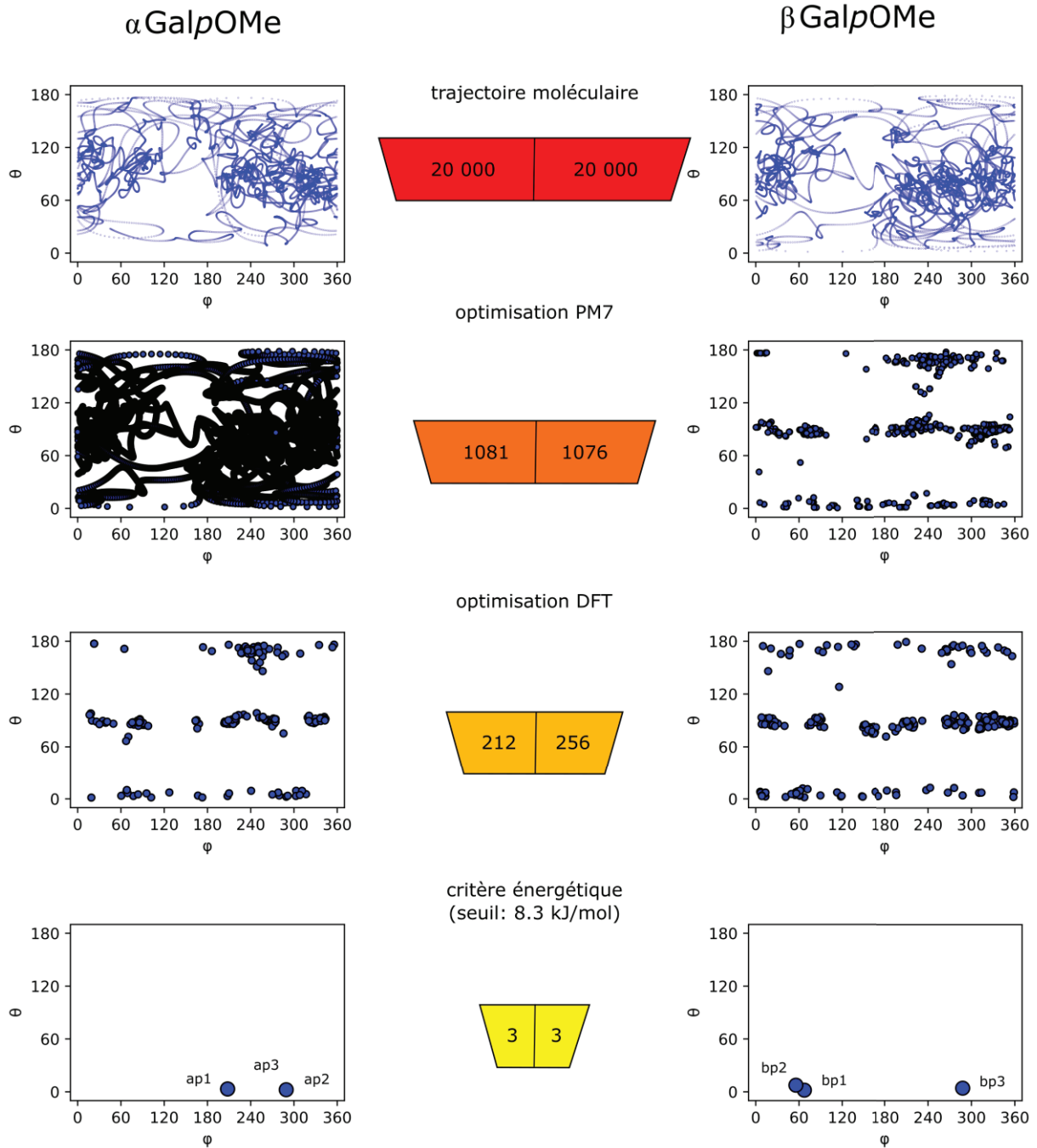


Figure 4 Nombre de conformations à chaque étape de l'étude conformationnelle et leurs coordonnées de Cremer et Popple (α GalpOME à gauche et β GalpOME à droite).

3.1.2 Recherche conformationnelle du galactofuranose

Poursuivons avec les résultats de dynamique moléculaire des deux anomères de galactofuranose. La même représentation a été conservée : on retrouve en Figure 5 l'entonnoir indiquant le nombre de conformations restantes à chaque étape de la dynamique moléculaire. Les conformations adoptées sont ici représentées sous forme d'une roue avec les paramètres d'Altona-Sundaralingam. Les deux trajectoires aboutissent au nombre de conformations préalablement sélectionné (20 000 structures). Un grand nombre de conformères sont obtenus par famille, montrant une exploration complète des furanoses. Une première optimisation grossière par PM7 réduit significativement le nombre de structures différentes (1154 pour α GalpOME et 1291 pour β GalpOME). Ces dernières sont réoptimisées avec les bases DFT B3LYP puis CAMB3LYP, donnant respectivement 250 et 296 conformations. On

3. Impact de la taille du cycle sur la conformation du galactose

retrouve une bonne répartition des structures à l'exception d'une région de la roue. Il ne reste plus à ce niveau d'optimisation de conformations avec O4 devant le cycle (4T_0 et 4E) pour la forme α . De même, des conformations de type E_1 et 0T_1 , c'est-à-dire avec l'oxygène O1 derrière le cycle, ne sont plus présentes pour la forme β . Enfin, la sélection par critère énergétique ($<8,3$ kJ/mol) réduit considérablement le paysage conformationnel. On n'obtient pour l'anomère α plus que des conformations au nord-ouest (E_2 , 1E et E_0) et au sud (2T_3 et 2E). L'anomère β a quant à lui uniquement des conformations au nord-ouest (1T_2 , 1E et 1T_0).

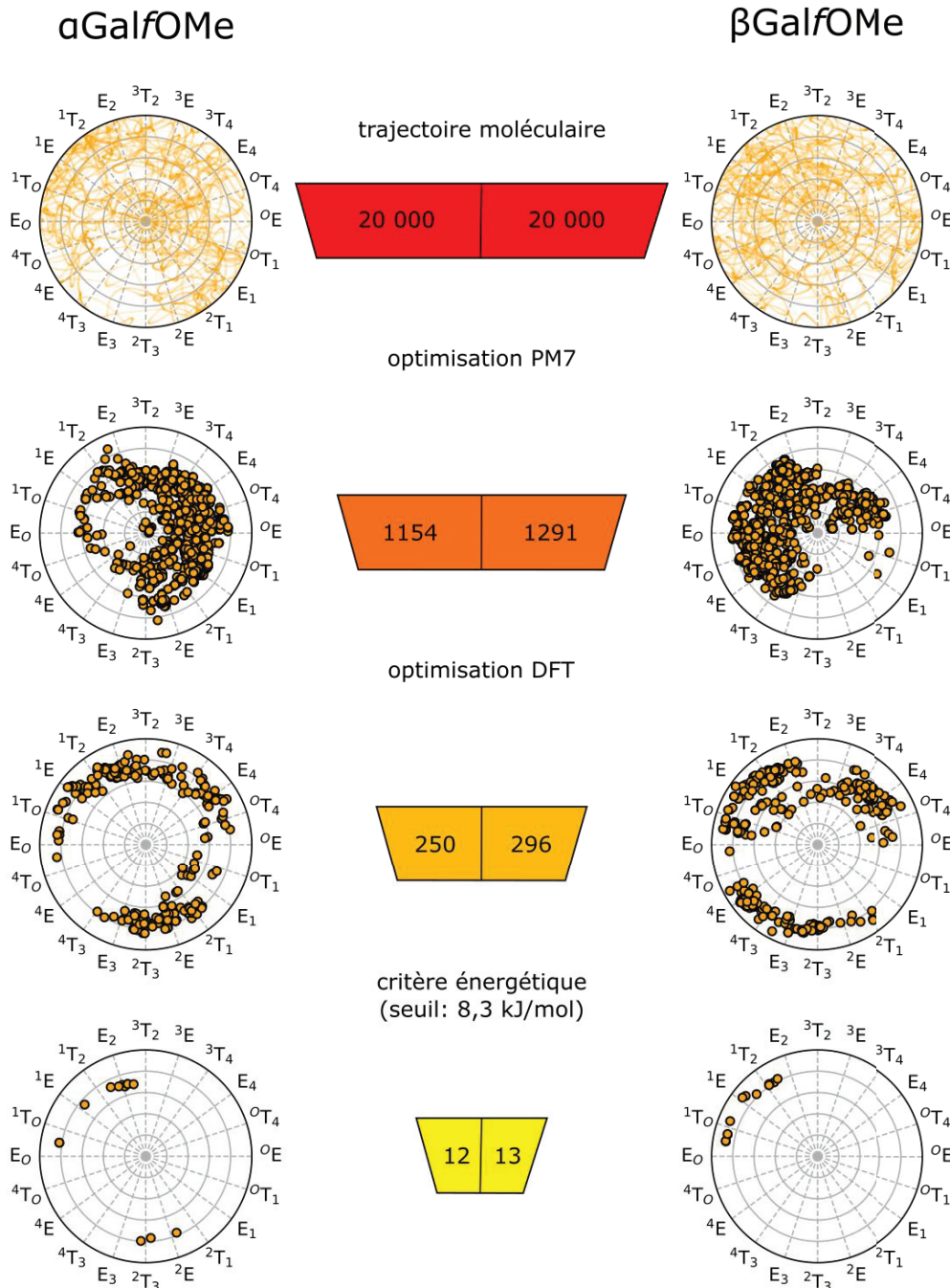


Figure 5 Nombre de conformations et leurs coordonnées Altona Sundaralingam à chaque étape de l'étude conformationnelle (α GalFOMe à gauche et β GalFOMe à droite).

3.1.3 Courbes d'énergie

La Figure 6 reprend les structures sélectionnées selon le critère énergétique (<8,3 kJ/mol). Les structures d'une même molécule sont de la même couleur et leurs énergies relatives sont calculées à partir de l'énergie de la structure la plus stable (qui est une structure de α GalpOMe).

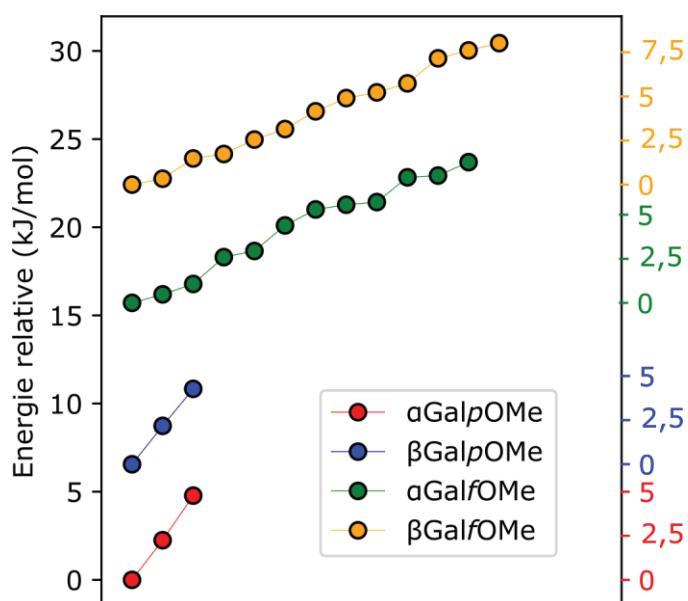


Figure 6 Structures sélectionnées de α GalpOMe (rouge), β GalpOMe (bleu), α GalfOMe (vert) et β GalfOMe (orange) après application du critère énergétique. Energies relatives (niveau CAMB3LYP) par rapport à la structure la plus stable de α GalpOMe (en ordonnée à gauche) ou par rapport à leur conformation de plus basse énergie respective (en ordonnée à droite).

Notons d'abord qu'il existe un nombre de structures répondant au critère énergétique plus important pour le galactofuranose, indépendamment de l'anomérisation. Ensuite, nous pouvons remarquer que les structures furanosidiques sont moins stables. Si l'on compare les énergies de la structure la plus stable pour les quatre molécules étudiées, on constate un écart de 15 kJ/mol entre les furanoses et pyranoses. Ces résultats ne sont pas surprenants puisque le cycle furanose a une contrainte de torsion plus élevée que le cycle pyranose.

La stabilité de la structure dépend de la taille de cycle mais également de l'anomérisation. En effet, avec une différence de 5 kJ/mol, nous avons relevé une meilleure stabilité thermodynamique pour les anomères α .

3.2 Calculs des spectres vibrationnels

Nous avons simulé les spectres infrarouge des conformations sélectionnées lors de l'étape précédente. Les calculs de fréquence sont effectués avec la base CAM-B3LYP/6-311++G(2df,2pd).

Les fréquences calculées sont représentées par des barres verticales dans les figures suivantes. Les modes de vibration des CH sont affichés en noir, ceux des NH en bleu et des OH en rouge. Ces derniers sont également annotés par le numéro de l'oxygène correspondant. Les spectres expérimentaux, représentés par une ligne grise continue, figurent également pour faciliter la comparaison.

3. Impact de la taille du cycle sur la conformation du galactose

3.2.1 Spectres vibrationnels du galactopyranose

La figure suivante représente les spectres vibrationnels des trois conformations les plus probables énergétiquement du complexe $\alpha\text{GalpOMe}+\text{NH}_4^+$.

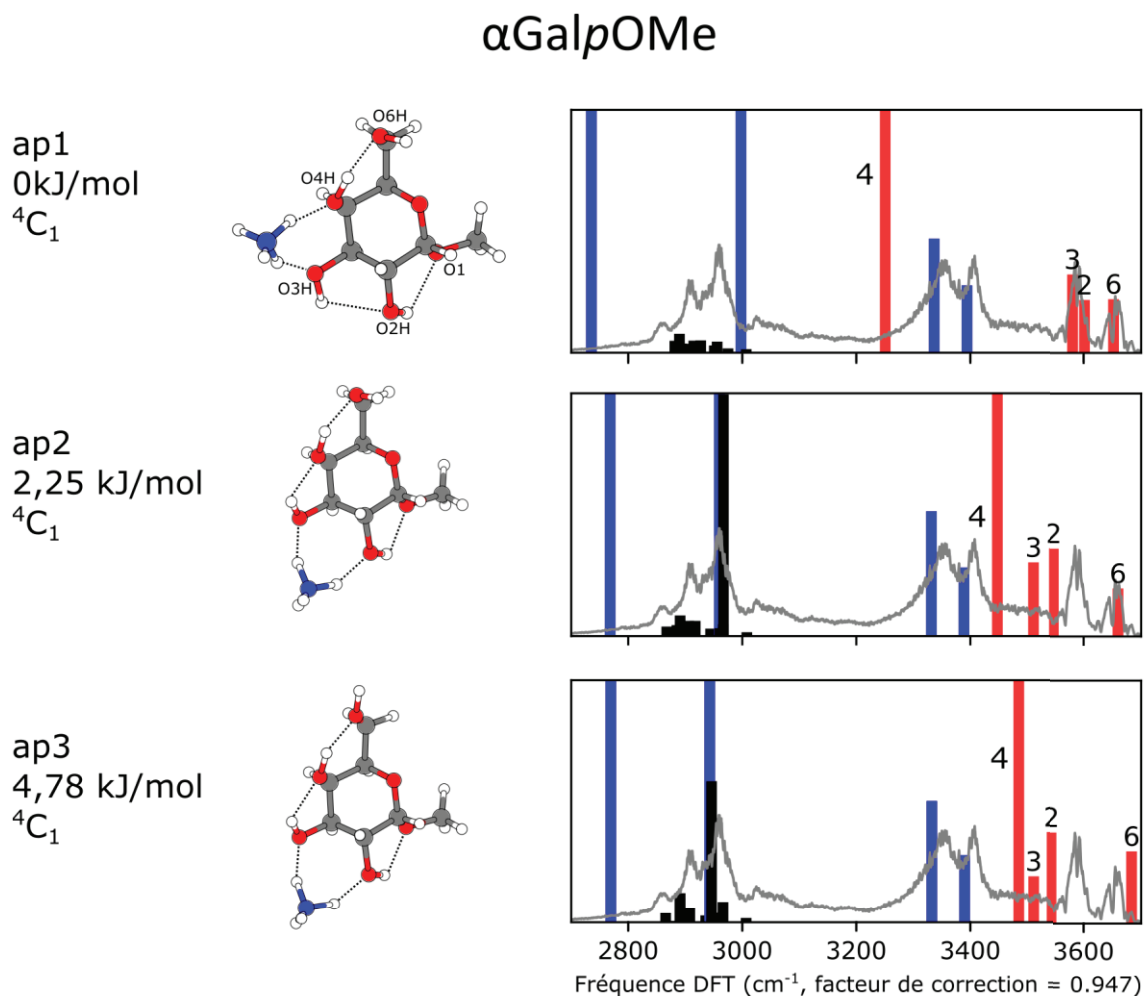


Figure 7 Les trois conformations les plus stables énergétiquement de $\alpha\text{GalpOMe}+\text{NH}_4^+$ (ap1, ap2 et ap3) avec à gauche leurs énergies et conformations de cycle, au milieu leurs structures et à droite leurs spectres IR calculés.

La conformation ap1 est la plus stable énergétiquement. L'adduit NH_4^+ est positionné entre les groupes O3H et O4H et participe au réseau de liaisons hydrogène. En effet, la structure est stabilisée par une série de liaisons $\text{HO6}--\text{HO4}--\text{HNH}--\text{O3H}--\text{O2H}--\text{O1H}$. Seul le groupe O6H est libre de toute interaction, sa fréquence calculée est donc élevée (3650 cm^{-1}). Les groupes O2H et O3H sont impliqués dans des liaisons faibles, leurs fréquences sont ainsi décalées vers la gauche (autour de 3560 cm^{-1}). Enfin, O4H est fortement lié à O6H, sa fréquence est donc davantage décalée (3250 cm^{-1}). Deux modes NH libres sont observés à 3390 et 3400 cm^{-1} et deux autres liés ont été calculés à 2730 et 3020 cm^{-1} . Les modes CH sont détectés dans la gamme $2800\text{--}3000\text{ cm}^{-1}$.

La seconde conformation ap2 a une énergie relative de $2,25\text{ kJ/mol}$. Ici, l'adduit NH_4^+ se situe entre O2H et O3H. Cette différence de position modifie légèrement le réseau de liaisons H : $\text{HO6}--\text{HO4}--\text{HO3}--\text{HNH}--\text{O2H}--\text{O1H}$. Le spectre théorique est affecté : O4H est moins lié et sa fréquence se situe vers 3450 cm^{-1} . À l'inverse, O2H et O3H qui sont à proximité de NH_4^+ se lient plus fortement (fréquences respectives à 3550 et 3510 cm^{-1}). Seul le mode O6H n'est pas affecté puisque, libre, il est calculé à 3650 cm^{-1} .

La troisième conformation la plus stable ap3 a une énergie relative de 4,78 kJ/mol. Nous pouvons remarquer qu'elle est très proche de la conformation précédente, ce qui se traduit par des coordonnées ϕ et θ quasi-identiques. La différence mineure est la rotation du groupement O6H : il est ici orienté vers le haut tandis qu'il se dirige vers la droite pour ap2. La fréquence calculée pour O6H se retrouve légèrement déplacée vers les faibles fréquences (3685 cm^{-1}), et celle de O4H également (3485 cm^{-1}) puisqu'il est lié à O6H.

Les trois conformations font partie de la famille des 4C_1 .

Comparons maintenant les spectres simulés avec celui obtenu expérimentalement (en gris). Nous trouvons une parfaite correspondance avec les fréquences calculées de la conformation ap1. En effet, la bande à 3650 cm^{-1} s'explique par le mode O6H et celle à 3560 cm^{-1} par les modes O3H et O2H. Les deux modes de NH libres correspondent parfaitement à 3390 et 3400 cm^{-1} . Il en est de même pour le motif de CH entre 2800 et 3000 cm^{-1} . Pour ce qui est des modes liés, cela se traduit expérimentalement aussi par un décalage de la bande vers la gauche mais également par une modification de la forme de la bande : celle-ci se retrouve plus étalée et moins intense. Pour cette raison, la bande de O4H est faible et élargie, il en est de même pour les deux NH liés. La conformation la plus stable explique donc à elle seule l'ensemble des bandes expérimentales. Quant aux deux autres conformations, elles peuvent être expérimentalement présentes (ce qui expliquerait par ailleurs la zone d'absorption dans la gamme $3450\text{--}3550\text{ cm}^{-1}$). Cependant, ap2 et ap3 n'expliquent pas à eux seuls le spectre expérimental car il n'y a pas de mode correspondant à la bande à 3560 cm^{-1} . Ainsi nous pouvons conclure que $\alpha\text{GalpOMe}$ est majoritairement présent sous la conformation ap1.

Tout comme pour l'anomère α , $\beta\text{GalpOMe}$ a aussi trois conformations en dessous du seuil énergétique. Leurs structures et leurs spectres vibrationnels sont représentés en Figure 8.

La plus stable conformation bp1 a son groupement NH_4^+ situé entre O3H et O4H. Il en résulte une suite de liaisons intra cycliques HO6--HO4--HNH--O3H--O2H. Le groupe O4H, fortement lié, a une fréquence décalée vers le rouge (3240 cm^{-1}). Plus faiblement lié, O3H a une fréquence calculée à 3595 cm^{-1} . Enfin, les deux autres groupements OH sont libres et leurs fréquences se situent autour de 3650 cm^{-1} . Les modes NH liés sont calculés à 2710 et 3030 cm^{-1} et les deux modes NH libres à 3340 et 3395 cm^{-1} .

La deuxième conformation bp2 a une énergie relative de 2,18 kJ/mol. La position du NH_4^+ est différente (entre les groupements O2H et O3H), ce qui impacte sur la position des groupements et le réseau de liaisons intra-cycliques. Le groupe O4H est lié à O6H mais plus faiblement (fréquence calculée à 3430 cm^{-1}) puisque l'adduit NH_4^+ n'est pas à proximité. A l'inverse, O3H et O2H sont impliqués dans des liaisons plus fortes car l'adduit NH_4^+ se situe entre ces deux groupes. On observe une rotation de O3H vers le haut, il est lié non plus à O2H mais à O4H (fréquence calculée : 3500 cm^{-1}). O2H, quant à lui, est orienté vers O1Me et sa fréquence est décalée (3580 cm^{-1}). Seul le groupe O6H présente une fréquence calculée commune pour les deux conformations (3655 cm^{-1}). Les modes NH libres sont très proches de celles de la conformation précédente, elles sont calculées à 3330 et 3390 cm^{-1} tandis que les deux modes liés apparaissent à 2790 et 2970 cm^{-1} .

En ce qui concerne la conformation bp3, l'adduit NH_4^+ se positionne à l'avant du cycle proche de O5. Les groupements hydroxyle s'orientent dans le sens antihoraire : O6H--O4H--O3H--O2H--OMe. La fréquence de O4H est calculée à 3570 cm^{-1} tandis que celles des autres OH se situent autour de 3650 cm^{-1} . Etant donné la position du NH_4^+ , on note des modes NH peu liés et légèrement déplacés par rapport aux deux conformations précédentes (3300 et 3380 cm^{-1}), tandis que les NH liés ont leurs fréquences à 2940 et 3100 cm^{-1} .

β GalpOMe

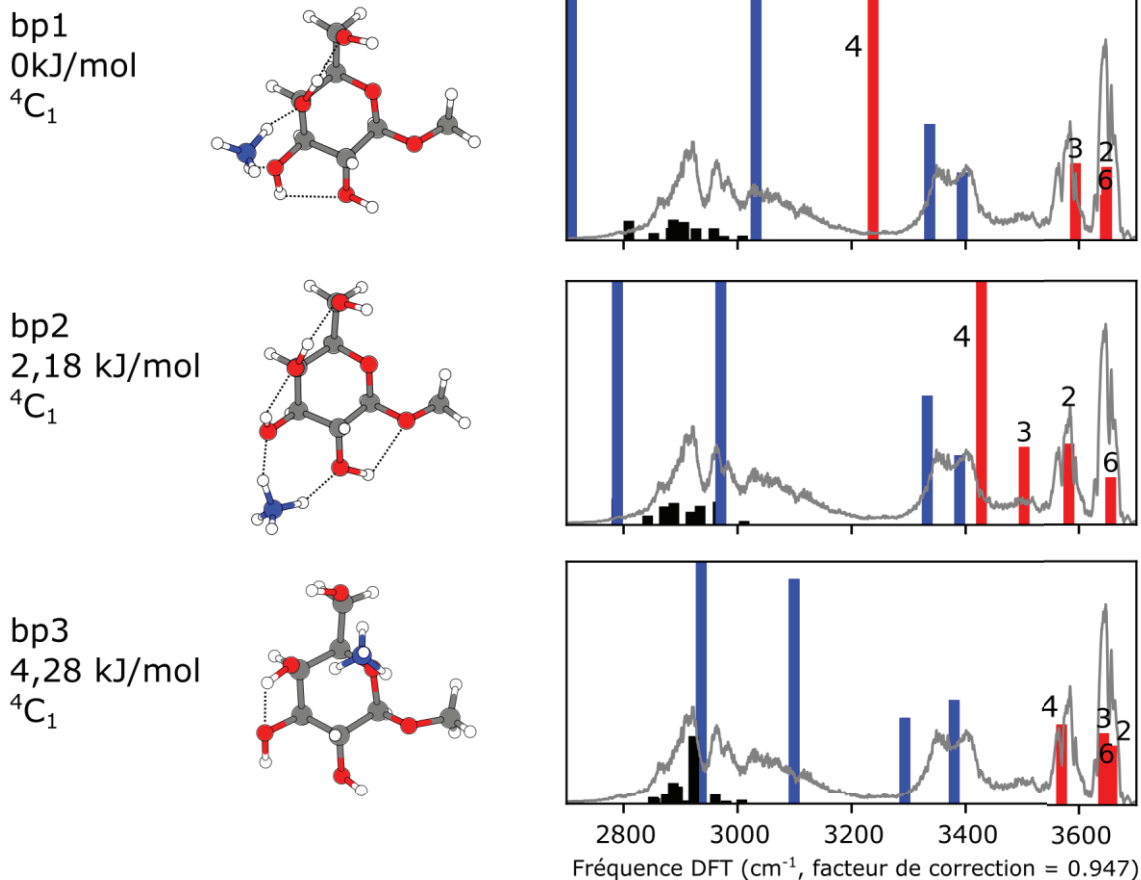


Figure 8 Les trois conformations les plus stables énergétiquement de β GalpOMe+ NH_4^+ (bp1, bp2 et bp3) avec à gauche leurs énergies et conformations de cycle respectives, au milieu leurs structures et à droite leurs spectres IR calculés.

Les conformations bp1 et bp3 ont une bonne correspondance entre leurs fréquences de mode OH calculées et les bandes OH observées expérimentalement. Ainsi la bande à 3580 cm^{-1} correspond à O3H pour bp1 et à O4H pour bp3. Celle à 3650 cm^{-1} correspond à O2H et O6H pour bp1 et O2H, O3H et O6H pour bp3. Il est à noter que le mode O4H de bp1, fortement lié, est expérimentalement peu intense. En revanche, les bandes NH excluent la conformation bp3. En effet, les deux bandes expérimentales à 3350 et 3400 cm^{-1} ne correspondent pas aux deux modes calculés. En ce qui concerne la conformation bp2, on peut envisager qu'elle soit minoritairement présente. Cela expliquerait notamment la faible bande expérimentale à 3500 cm^{-1} qui serait expliquée par O3H.

Les six conformations présentées ci-dessus ont un cycle 4C_1 . La comparaison entre fréquences calculées et spectres expérimentaux montre que les deux conformations les plus stables, ap1 et bp1, sont celles qui prédominent. Les bandes expérimentales peu intenses peuvent s'expliquer par les quatre autres conformations, présentes en minorité.

Ainsi, les deux anomères α -galactopyranose et β -galactopyranose montrent un excellent accord entre leurs spectres expérimentaux et les spectres calculés de la conformation énergétiquement la plus stable. Les deux structures ap1 et bp1 ont des points communs : une position de NH_4^+ identique et un réseau de liaisons hydrogène très similaire. La différence principale est que ce réseau s'étend jusqu'à O1H pour ap1, mais s'arrête à O2H pour bp1. En effet, O1H et O2H sont en *cis* pour ap1 facilitant ainsi leur liaison alors que pour bp1 ces deux groupes sont en *trans*, ce qui complique la liaison hydrogène.

Cela se reflète sur leurs spectres calculés par la fréquence de O2H plus faible pour ap1 que pour bp1. Cette différence peut être observée expérimentalement car la bande à 3600 cm^{-1} est plus intense pour ap1 (due aux groupes O2H et O3H) que pour bp1 (uniquement O3H).

La même analyse sera établie sur les composés galactofuranose afin de mieux appréhender l'impact du changement de cycle sur la conformation.

3.2.2 Spectres vibrationnels du galactofuranose

Nous avons vu lors de la dernière étape de la recherche conformationnelle qu'on obtenait un grand nombre de conformations pour les deux galactofuranoses (12 pour α GalfOMe et 13 pour β GalfOMe) avec des écarts énergétiques très faibles. Elles sont représentées en Figure 9 et Figure 10 pour α GalfOMe et en Figure 11 et Figure 12 pour β GalfOMe. Ces structures sont regroupées par famille plutôt que classées par énergie (les énergies étant trop proches les unes des autres, elles ne sont pas indicatrices).

Pour l'anomère α , on retrouve cinq structures (af1 à af5) adoptant un cycle E_2 . Les quatre premières structures ont une position de NH_4^+ identique et se différencient uniquement par la rotation des groupements hydroxyle. On remarque ainsi la rotation de O6H entre af1 et af2 et celle de O3H entre af2 et af3. Pour af4, c'est O2H qui se dirige vers le bas alors qu'il s'oriente vers la droite pour les autres rotamères. Le seul groupe OH lié est O5H, impliqué dans une forte liaison avec O2H. Les quatre spectres calculés sont ainsi très proches : la fréquence de O5H est à 3435 cm^{-1} alors que celles des autres hydroxyles, libres, sont vers 3650 cm^{-1} . Seul un mode NH est libre, dont la fréquence se situe vers 3380 cm^{-1} . Les autres modes NH sont estimés en dessous de 3200 cm^{-1} . La structure af5 se distingue des quatre précédentes par la localisation de NH_4^+ , situé entre O5H et O6H. Il permet la formation d'une série de liaisons hydrogène O6H--HNH--O5H--O2H--O1Me. Les calculs de cette structure montrent une fréquence de O5H à 3150 cm^{-1} et de O2H à 3555 cm^{-1} . Les autres OH étant libres, leurs fréquences sont vers 3650 cm^{-1} . Deux modes NH libres sont calculés à 3325 et 3390 cm^{-1} .

Les calculs ont abouti à une seule structure avec un cycle 1E (af6). Cette structure est composée uniquement d'hydroxyles libres dont les fréquences sont toutes vers 3650 cm^{-1} . Les modes NH se situent à 2995 , 3115 , 3230 et 3385 cm^{-1} .

On compte deux structures de type 2T_3 (Figure 10) qui se distinguent notamment par la position de l'adduit NH_4^+ , entre O2H et O3H, et le réseau de liaisons hydrogène O5H--O3H--HNH--O2H--OMe. Les groupes O3H et O2H, à proximité de NH_4^+ et liés, ont leurs fréquences affectées (3240 cm^{-1} pour O3H et 3530 cm^{-1} pour O2H). La différence entre les deux conformations est la rotation de O5H, qui affecte peu le spectre théorique obtenu.

On dénombre également deux structures de type E_0 . Tous les OH sont libres sauf O5H qui est lié à O6H alors qu'il était lié à O2H pour les conformations E_2 . Cette liaison est beaucoup plus faible que la liaison O5H--O2H, la fréquence de O2H étant ici calculée à 3560 cm^{-1} .

Enfin, il existe une conformation avec un cycle 2E très caractéristique par la rotation des groupements hydroxyle et la position du NH_4^+ (entre O3H et O6H). Les groupes OH sont tous faiblement liés et leurs fréquences sont décalés vers 3600 cm^{-1} .

3. Impact de la taille du cycle sur la conformation du galactose

α GalfOMe

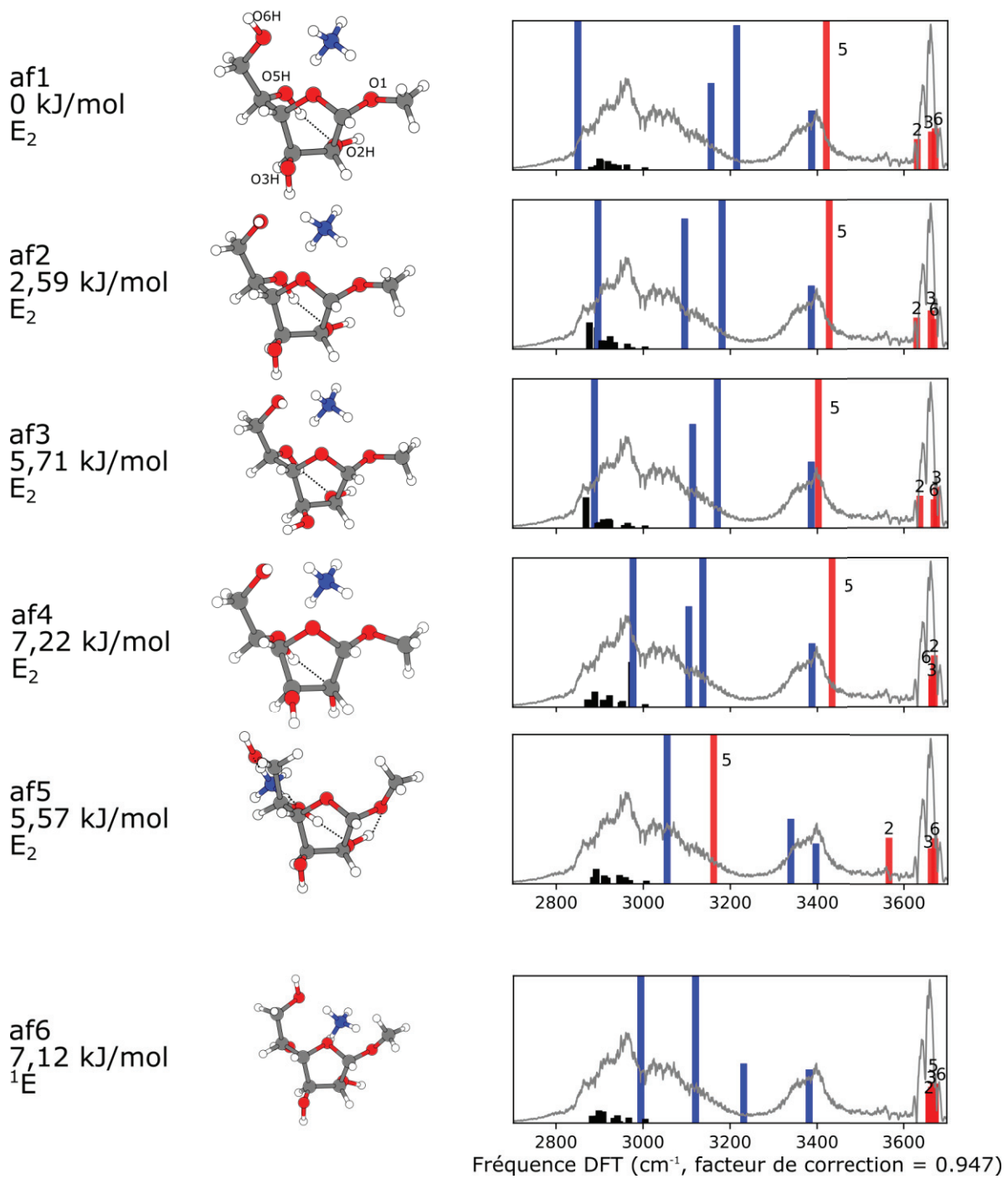


Figure 9 Structures obtenues lors de la recherche conformationnelle du composé α GalfOMe+ NH_4^+ et spectres vibrationnels correspondant (partie 1).

aGalfOMe

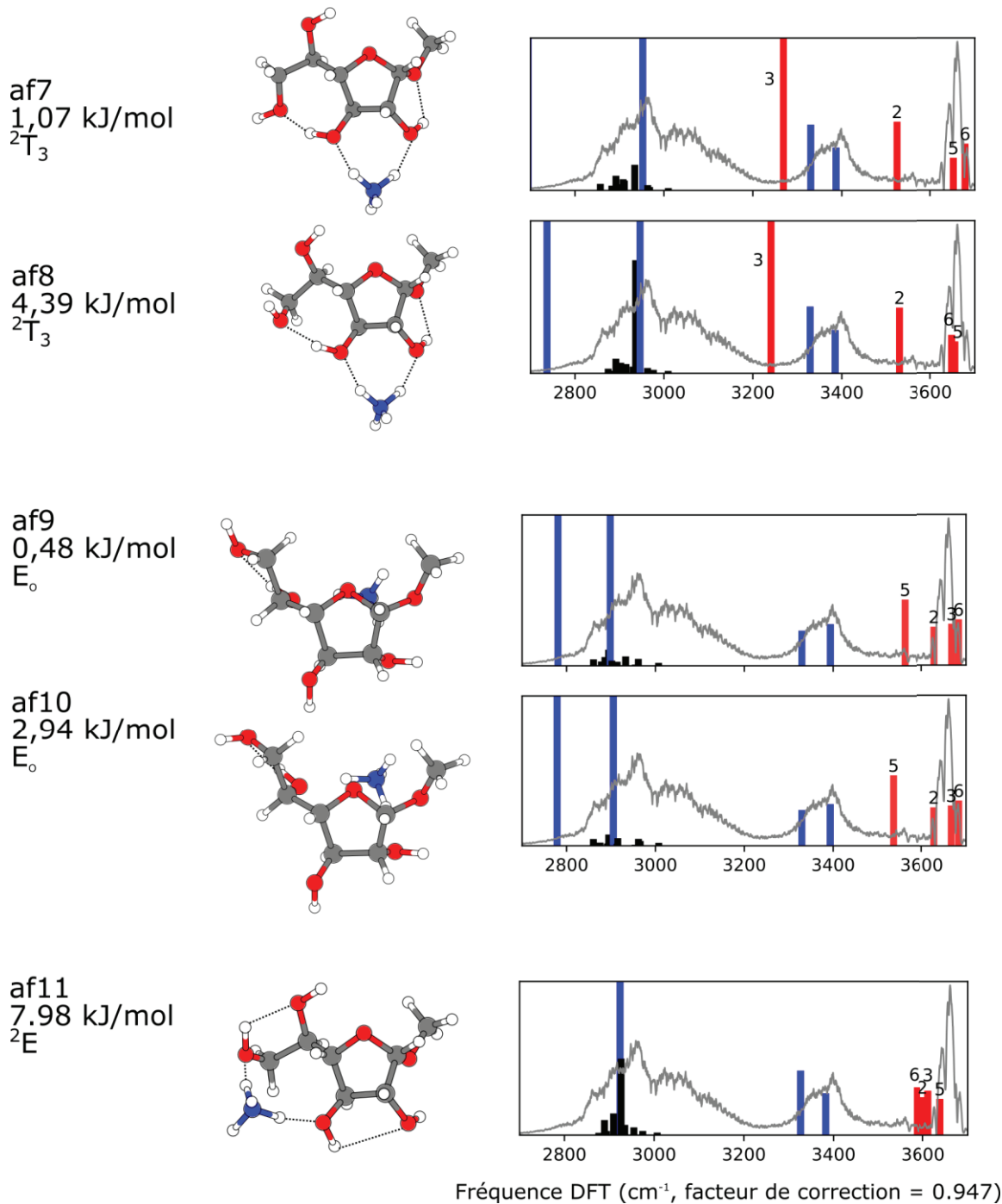


Figure 10 Structures obtenues lors de la recherche conformationnelle du composé aGalfOMe+NH₄⁺ et spectres vibrationnels correspondant (partie 2).

La comparaison entre le spectre expérimental et les spectres simulés valide certaines familles de structure expérimentalement présentes. Les spectres calculés des structures appartenant à la famille des E_2 correspondent : la bande à 3650 cm^{-1} s'explique par les modes OH libres tandis que la large bande vers 3400 cm^{-1} correspond au mode NH libre ainsi qu'au O5H fortement lié. Le motif de 2800 à 3200 cm^{-1} se superpose aux modes CH ainsi qu'aux modes NH liés. La structure de type 1E peut également être expérimentalement présente.

3. Impact de la taille du cycle sur la conformation du galactose

À l'inverse, la comparaison entraîne l'exclusion d'autres familles de structures. Le mode O2H des structures avec un cycle 2T_3 ainsi que le mode O5H des structures avec un cycle E_0 ne coïncident avec aucune bande expérimentale. Les critères spectroscopiques valident la présence de certaines structures situées uniquement au nord-ouest de la roue (E_2 et 1E).

De même sont représentées sur les deux figures suivantes les structures sélectionnées selon le critère énergétique de $\beta\text{Gal}\alpha\text{fOMe}$. On ne retrouve plus que des conformations nord-ouest : cinq structures adoptant un cycle 1T_2 , deux conformations de la famille des 1E (Figure 11) et cinq conformations de type 1T_0 (Figure 12).

La famille des 1T_2 (Figure 11) se caractérise par le bras libre O5H-O6H et l'adduit NH_4^+ assez mobiles. À l'avant du cycle, la liaison hydrogène O3H--OMe se forme. À l'arrière du cycle, les trois premières structures bf1, bf2 et bf3 présentent une forte liaison O5H--O2, impliquant un fort décalage du mode O5H (entre 3260 et 3320 cm^{-1}). Comme O3H est également impliqué dans une liaison intracyclique, sa fréquence est calculée à 3590 cm^{-1} pour ces trois structures. La conformation bf1 a une liaison hydrogène supplémentaire : O6H--O4, ce qui explique que la fréquence de O6H se trouve à 3520 cm^{-1} . L'adduit NH_4^+ de bf1 est situé entre O6H et O4 tandis qu'il est entre O5H et O6H pour les conformères bf2 et bf3. La différence entre ces deux conformères est la rotation de O2H qui n'a pas de répercussion sur la fréquence calculée (3665 cm^{-1}) puisque le groupement est libre de toute interaction dans les deux cas. Les spectres calculés de bf2 et bf3 sont ainsi quasi-identiques. Les deux structures suivantes bf4 et bf5 sont plus étalées, avec un adduit NH_4^+ localisé à gauche du cycle à proximité du groupe O6H. Le bras libre de bf4 est complètement déplié, O5H et O6H sont libres de toute interaction. L'unique OH lié est donc le O3H (lié au OMe comme pour toutes les conformations 1T_2), cette liaison est un peu plus forte étant donné que l'adduit NH_4^+ est à sa proximité. Sa fréquence est estimée à 3490 cm^{-1} . La conformation bf5 se distingue des autres par ses liaisons non covalentes. En effet, une série de liaisons hydrogène O6H--O5H--O2H se forme à l'arrière du cycle. Les fréquences calculées de O5H (3490 cm^{-1}) et O6H (3560 cm^{-1}) se trouvent donc décalées. De même que pour bf4, le mode O3H est plus déplacé que pour les structures bf1 à bf3 en raison de la proximité de NH_4^+ . Ainsi, seul le groupe O2H (3660 cm^{-1}) est libre de toute interaction.

Deux structures de type 1E répondent au critère énergétique (Figure 11). On retrouve une liaison O3H--OMe à l'avant du cycle. À l'arrière du cycle, le groupe O5H est orienté vers la gauche et libre et O2H orienté vers le bas et également libre. La différence entre bf6 et bf7 est la rotation du groupement O6H : alors qu'il se dirige vers le haut et est libre pour bf7 (3670 cm^{-1}), il est impliqué dans une liaison faible pour bf6 (3610 cm^{-1}).

Les cinq structures de la famille des 1T_0 sont très proches (Figure 12). La position du NH_4^+ est semblable, situé derrière le cycle au niveau de O4. On notera également que O5H est lié à O6H, cette liaison est plus faible que quand il est lié à O2H (voir les trois premières structures bf1 à bf3). La fréquence calculée est effectivement faiblement décalée, entre 3530 et 3560 cm^{-1} . Par ailleurs O3H est orienté vers le bas et libre (sauf pour bf8) alors qu'il existait dans la famille des conformations 1T_2 une liaison avec OMe.

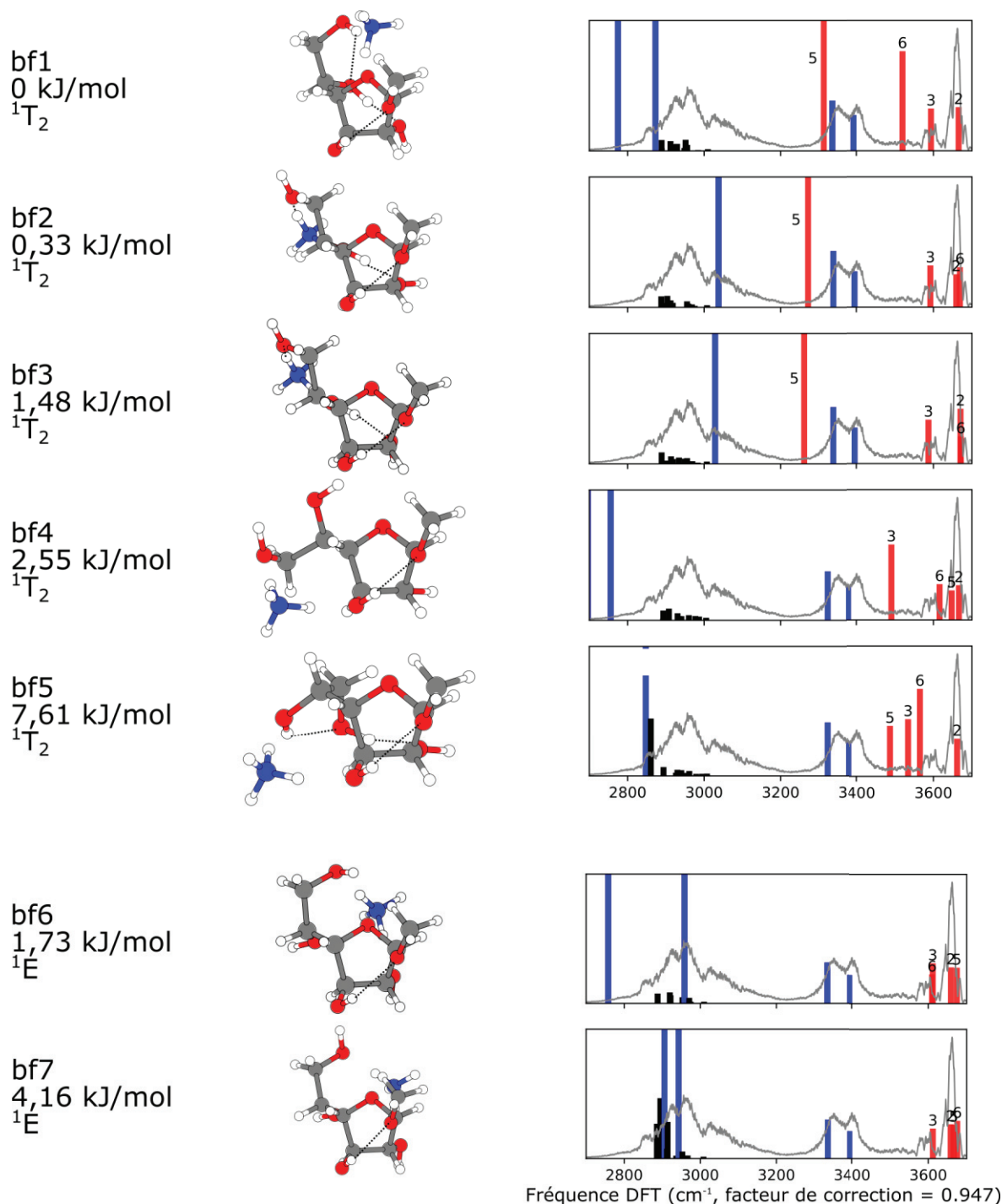
β GalfOMe

Figure 11 Structures obtenues lors de la recherche conformationnelle du composé β GalfOMe+NH $_4^+$ et spectres vibrationnels correspondant (partie 1).

Les modes NH libres sont calculés à 3330 et 3390 cm^{-1} pour toutes les conformations sauf bf6 et bf7 pour lesquels il y a un léger décalage de 10 cm^{-1} (3320 et 3380 cm^{-1}). Elles correspondent à la bande expérimentale de 3300 à 3430 pour laquelle on distingue deux maxima à 3350 et 3400 cm^{-1} . Les modes NH liés (allant de 2700 à 3050 cm^{-1}) et les modes CH (de 2870 à 3020 cm^{-1}) se chevauchent et expliquent la large bande de 2800 à 3150 cm^{-1} . Ce sont surtout les modes OH qui nous permettent de sélectionner les conformations expérimentalement présentes. On peut en effet exclure toutes les structures adoptant une conformation 1T_0 car le mode O5H calculé ne correspond à aucune bande expérimentale.

3. Impact de la taille du cycle sur la conformation du galactose

De même, des structures 1T_2 sont également exclues : bf1 pour son mode O6H, bf4 pour le mode O3H et bf5 pour les modes O5H, O3H et O6H. Les structures restantes (bf2, bf3, bf6 et bf7) expliquent les bandes expérimentales observées : la bande intense à 3660 cm^{-1} s'explique par tous les modes OH libres et la faible bande à 3590 cm^{-1} par le mode O3H qui est lié à OMe pour toutes les conformations sélectionnées. Nous pouvons ajouter que le mode O5H peut être fortement lié comme dans le cas de bf2 et bf3 et que la bande expérimentale est très peu intense du fait du décalage et chevauche avec les bandes NH.

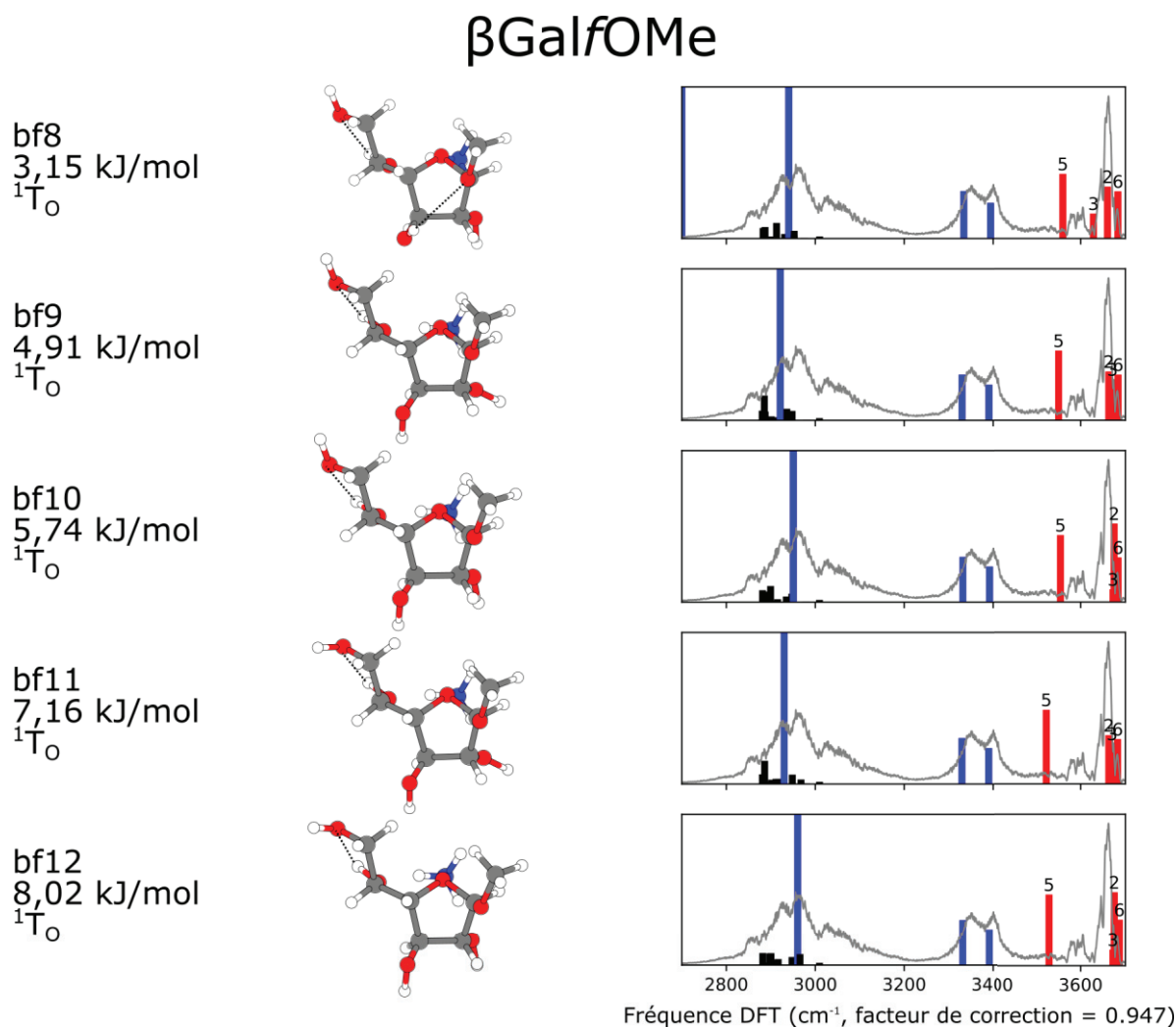


Figure 12 Structures obtenues lors de la recherche conformationnelle du composé β GalfOMe+NH $_4^+$ et spectres vibrationnels correspondant (partie 2).

En Figure 13 sont représentées les structures qui satisfont au critère énergétique ($<8,3\text{ kJ/mol}$). Ces résultats suggèrent que α GalfOMe est présent dans les régions nord-ouest (E_2 , E_0 et 1E) et sud (2T_3 et 2E), un résultat en accord avec une étude en RMN indiquant la présence de la conformation 2E .²² Étonnement, cette même étude n'a pas été conclusive sur la forme β GalfOMe qui, selon nos résultats, est limité à la région nord-ouest (1T_2 , 1T_0 et 1E). Dans un second temps, nous avons ajouté un critère spectroscopique (concordance entre le spectre expérimental et le spectre théorique). Les structures ne répondant pas à ce critère ont été décolorées sur la figure.

Pour α GalfOMe, les conformations au sud (2T_3 et 2E) ainsi que la conformation E_0 sont exclues. Il ne reste plus que des conformations de type E_2 et 1E . En ce qui concerne β GalfOMe, nous avons exclu les

conformations de cycle 1T_0 et validé les familles 1E et 1T_2 . En somme, seule la région nord-ouest est peuplée pour les deux galactofuranoses.

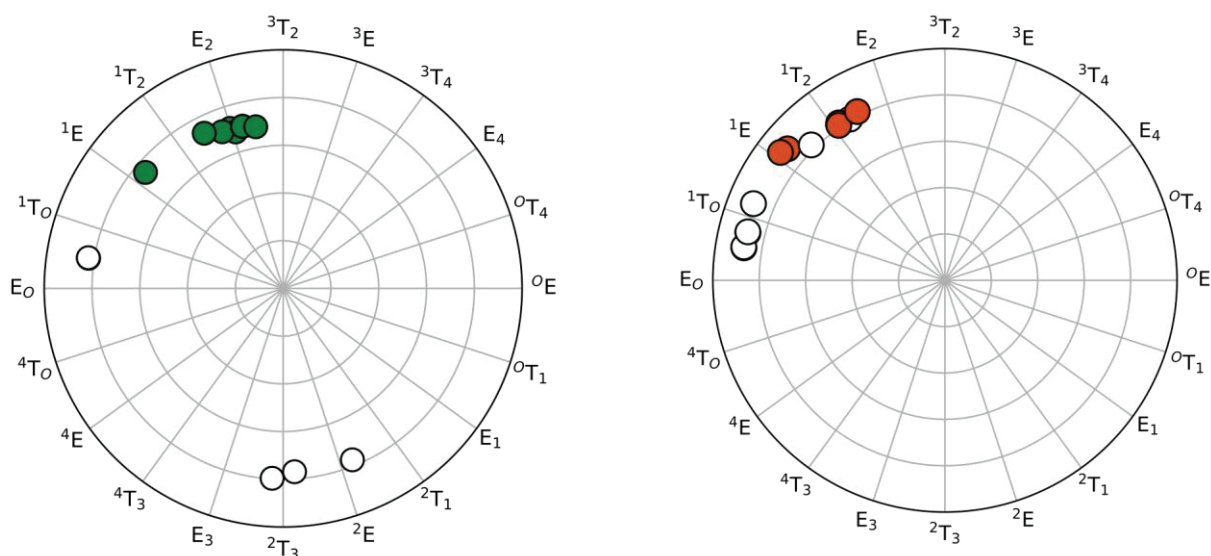


Figure 13 Coordonnées polaires des structures α GalpOMe (à gauche) et β GalpOMe (à droite) répondant au critère énergétique ($<8,3$ kJ/mol) représentés par des ronds. Les structures répondant au critère spectroscopique (bonne correspondance avec le spectre expérimental) sont en couleur.

3.3 Interprétation

Le galactopyranose nous a servi de molécule modèle pour notre approche complémentaire théorie/expérimentale, puis nous avons employé la même méthode pour étudier les propriétés intrinsèques du galactofuranose. Les conditions expérimentales sont donc les mêmes : des ions isolés dans une trappe à température ambiante. De même, nous avons employé la même procédure pour les simulations théoriques, avec les mêmes paramètres de dynamique moléculaire et de base DFT CAM-B3LYP/6-311++G(2df,2pd).

Les molécules flexibles posent un défi majeur aux théoriciens et spectroscopistes car plusieurs conformations co-existent et interconvertissent. On s'attend à un élargissement des bandes spectrales et donc à une perte de l'information diagnostique. Ceci est particulièrement vrai en solution car le solvant a tendance à déstabiliser le réseau de liaisons intramoléculaires et à faciliter l'interconversion. Les liaisons hydrogène n'ont pas été considérées pour cette raison lors de l'interprétation des données RMN dans l'article sur les galactofuranoses cité précédemment.²² Un des avantages de la spectroscopie en phase gazeuse est la révélation des propriétés intrinsèques de la molécule en absence d'environnement et la possibilité d'analyser les changements en présence d'un micro-environnement (par exemple la présence de NH_4^+ dans cette étude).

Contre toute attente, nous avons obtenu des spectres hautement résolus qui de plus sont diagnostiques de la taille du cycle : une bande OH intense à 3650 cm^{-1} est détectée pour les furanoses contre deux (à 3650 et 3600 cm^{-1}) pour les pyranoses.

Cette étude a également démontré que les deux anomères de galactopyranose sont impliqués dans des réseaux de liaisons hydrogène, ce qui explique les deux bandes OH expérimentales. L'une correspond aux OH libres et l'autre, décalée vers le rouge, aux groupes OH peu liés. L'intensité de ces deux bandes dépend de l'anométrie : l'anomère α a un rapport d'intensité OH liés/OH libres plus important que β GalpOMe car la liaison O2H--OMe n'existe que pour l'anomère α . La présence de peu/pas de liaisons intramoléculaires dans les structures furanose se traduit par une bande intense

3. Impact de la taille du cycle sur la conformation du galactose

correspondant aux OH libres. Une faible bande est présente à 3600 cm^{-1} pour l'anomère β uniquement car, tout comme pour les pyranoses, l'anomérisation affecte les liaisons hydrogènes (seule la forme β a une liaison O3H--OMe).

Dans cette étude, une combinaison de critères énergétique et spectroscopique nous a permis d'élucider les préférences conformationnelles intrinsèques et de mieux comprendre le diagnostic direct de la taille de cycle. Nous avons drastiquement affiné le paysage conformationnel des deux anomères de galactofuranose à la région nord-ouest. Ainsi il existe une coexistence et/ou interconversion des conformations validées uniquement sur cette partie de la roue.

Passons à l'étude théorique suivante qui consiste à analyser le galactofuranose dans différents micro-environnements.

4 Impact du micro-environnement sur la conformation du galactofuranose

Nous avons vu dans la partie précédente que le cycle pyranose présentait un squelette relativement rigide en forme de chaise. Les analyses en solution ou dans des cristaux vont dans ce sens. À l'inverse, le comportement conformationnel du furanose est très différent. On peut alors se demander si la flexibilité reflète les propriétés intrinsèques du furanose si elle est imposée par l'environnement. Développer des outils théoriques et expérimentaux pour rationaliser le rôle des facteurs intrinsèques et extrinsèques demeure encore extrêmement complexe.

La spectroscopie en phase gazeuse dans la gamme des 3 μm est dans ce contexte un véritable atout. En sondant les modes vibrationnels des OH, elle révèle la présence et la force des liaisons intra- mais également intermoléculaires (par exemple, avec l'adduit NH_4^+ dans la partie précédente). Ainsi, les préférences conformationnelles intrinsèques à la molécule et les effets du micro-environnement peuvent être spécifiquement étudiés par spectroscopie en phase gazeuse.

Plus particulièrement, la spectroscopie à double résonance est avantageuse dans le cas des molécules flexibles puisque le spectre vibrationnel de chaque conformation est acquis individuellement.²³ C'est cette technique que nous avons utilisé en collaboration avec l'ISMO. La molécule étudiée doit posséder un chromophore pour être détectable en UV, ce qui n'est pas le cas du modèle analysé en IRMPD ($\beta\text{GalfOMe}$). Le O-phényl- β -galactofuranose ($\beta\text{GalfOPh}$) a été synthétisé par nos collaborateurs chimistes de Rennes et a servi de modèle pour les analyses en IR-UV. Des études sur des pyranoses avaient préalablement démontré que la présence du groupement phényle affectait peu la conformation.²⁴

Notre étude sur l'effet du micro-environnement concerne uniquement l'anomère β du galactofuranoside (blocage de l'anomérisation par le groupement méthyle ou phényle selon les expériences). La molécule a été analysée dans différentes conditions locales que nous détaillerons ci-dessous.

4.1 Préférences conformationnelles du galactofuranose neutre en phase gazeuse

4.1.1 Recherche conformationnelle du galactofuranose neutre

Une dynamique moléculaire a été simulée sur le $\beta\text{GalfOPh}$ neutre. Le nombre de conformations obtenues après chaque étape d'optimisation ainsi que leurs coordonnées sont représentées en Figure 14. La trajectoire aboutit à 20 000 structures. L'optimisation PM7 et la suppression des doublons ne laissent plus que 1676 structures différentes. L'optimisation par DFT comprend deux étapes : une première optimisation avec la base B3LYP (après laquelle on obtient 202 structures) puis une seconde avec la base CAMB3LYP. Cette dernière permet de raffiner les structures et les énergies, mais il faut savoir que les énergies obtenues par B3LYP sont également indicatrices (ce qui n'est pas le cas des énergies PM7 qui sont trop grossières). Ainsi pour réduire le temps de calcul, nous avons fait le choix de sélectionner les 50 structures les plus stables B3LYP pour les optimiser par CAMB3LYP. Le critère énergétique ($<8,3$ kJ/mol) a ensuite été appliqué aux énergies CAMB3LYP et seules trois structures ont été validées.

4. Impact du micro-environnement sur la conformation du galactofuranose

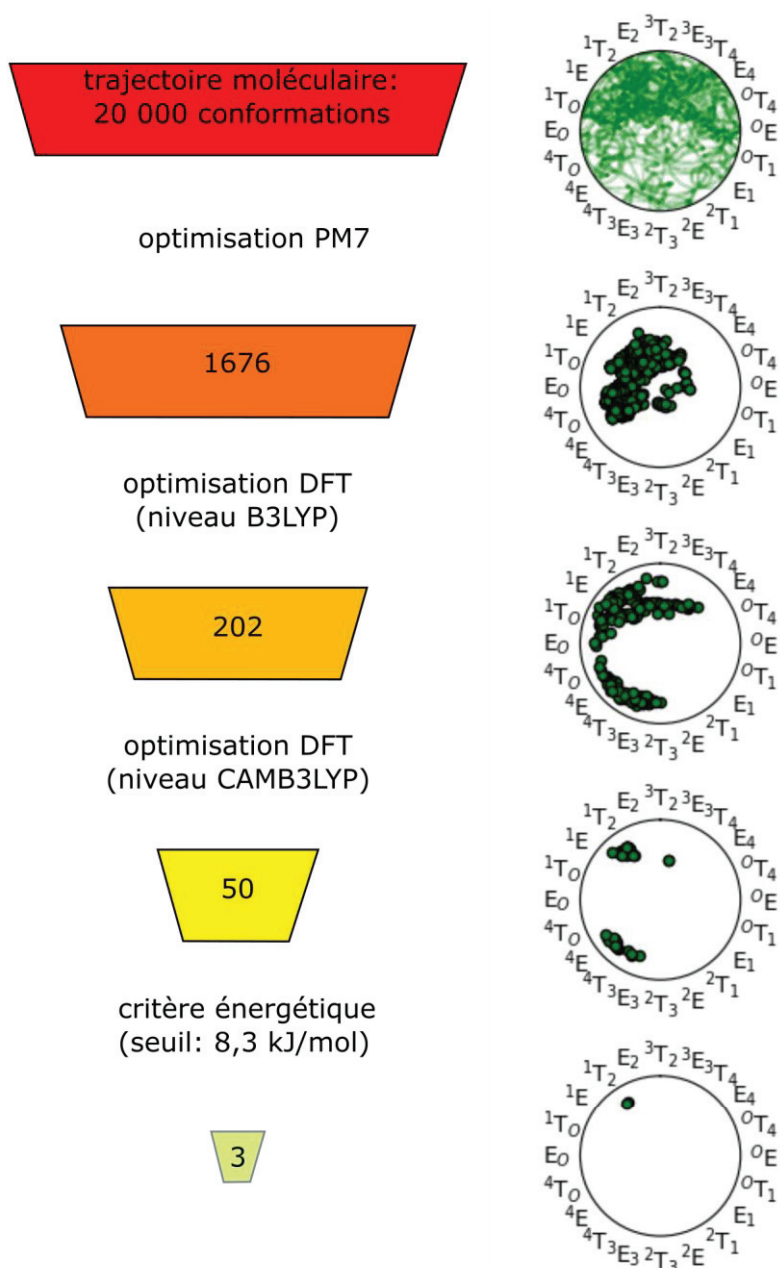


Figure 14 Nombre de conformations de β Gal α Ph neutre à chaque étape de l'étude conformationnelle et leurs coordonnées Altona Sundaralingam.

Nous pouvons observer que le paysage conformationnel se raffine au fil des optimisations. Alors que toutes les familles de conformations (et avec une amplitude très variable) sont représentées lors de la trajectoire moléculaire, on voit un premier regroupement lors de l'optimisation PM7. En effet, les conformations se situent plutôt au nord-ouest de la roue. Ensuite, l'optimisation DFT au niveau B3LYP supprime les conformations sud-est (0T_4 à 2E). L'optimisation CAMB3LYP sur les structures les plus stables entraîne la formation de deux groupes distincts (au nord-ouest et au sud-ouest). Les trois structures répondant au critère énergétique adoptent quant à elles uniquement une conformation de cycle 1T_2 .

4.1.2 R2PI et UVHB de β GalfOPh

L'ion moléculaire de β GalfOPh est détecté à m/z 256. Son signal est suivi pour le tracé des spectres R2PI et UVHB (Figure 15).

Dans un premier temps, la spectroscopie UV indique le nombre de conformères. La gamme 36600 - 36900 cm^{-1} est balayée par un premier laser et on observe du signal aux longueurs d'onde correspondant à une transition $S_0 \rightarrow S_1$. Le second laser est fixé à 32878 cm^{-1} (longueur d'onde d'absorption du chromophore phényle). Il ionise la molécule seulement si la transition a été réalisée. On relève deux signaux intenses à 36760 et 36700 cm^{-1} . Une étape intermédiaire nous permet de déterminer que ces deux signaux proviennent de deux conformères différents car les bandes d'absorption IR ne sont pas les mêmes (l'un absorbe à 3690 et l'autre à 3470 cm^{-1}).

Ensuite, une déplétion UVHB est réalisée. La longueur d'onde du laser IR est fixée (à 3690 cm^{-1} pour le spectre a et à 3470 cm^{-1} pour le spectre b) et on ajoute en amont un second laser UV qui balaie la gamme. On notera dans ce cas une déplétion de signal lorsque la longueur d'onde du premier laser correspond à une transition électronique du conformère recherché.

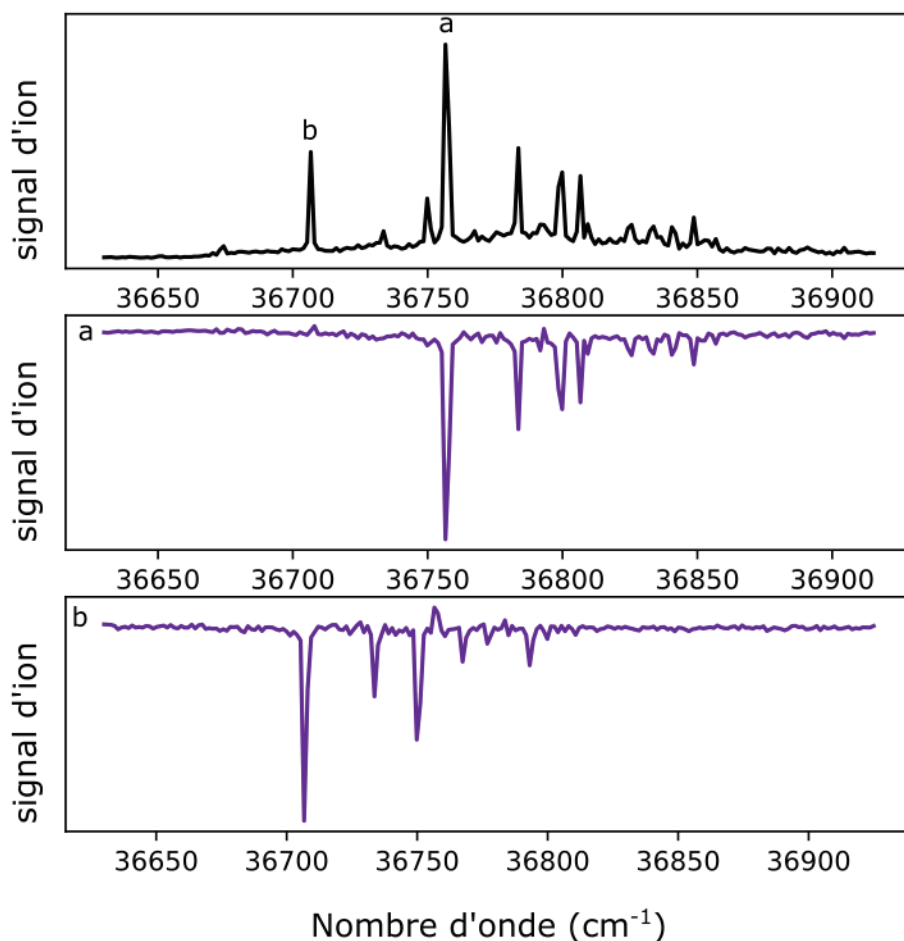


Figure 15 Spectre R2PI de β GalfOPh (en haut), spectre UVHB du conformère a (milieu) et du conformère b (bas).

La superposition des spectres UVHB des conformères a et b permet de recomposer entièrement le spectre R2PI. Nous avons ainsi expérimentalement deux conformations possédant des spectres UV distincts. Il est possible que plusieurs conformères possèdent le même spectre UV et ne peuvent donc pas être différenciés à cette étape. Nous pouvons donc affirmer que l'on a au moins deux conformations, mais nous ne pouvons pas en déduire leur nombre précisément. Nous avons pour cela besoin de l'ajout de l'infrarouge.

4. Impact du micro-environnement sur la conformation du galactofuranose

4.1.3 Calcul des spectres vibrationnels

Les spectres IR des deux conformères ont été acquis et sont tracés en Figure 16. Le conformère a possède quatre bandes expérimentales : une bande plutôt large à 3470 cm^{-1} , une plus fine à 3540 cm^{-1} et un doublet à 3590 et 3610 cm^{-1} . Le conformère b présente également quatre bandes : une bande large décalée à 3500 cm^{-1} , un doublet de bandes à 3600 et 3620 cm^{-1} et une bande à 3695 cm^{-1} . Nous pouvons remarquer que les intensités des différentes bandes sont comparables. Les trois structures les plus stables (nommées g1, g2 et g3) et leurs spectres simulés sont aussi représentés en Figure 16.

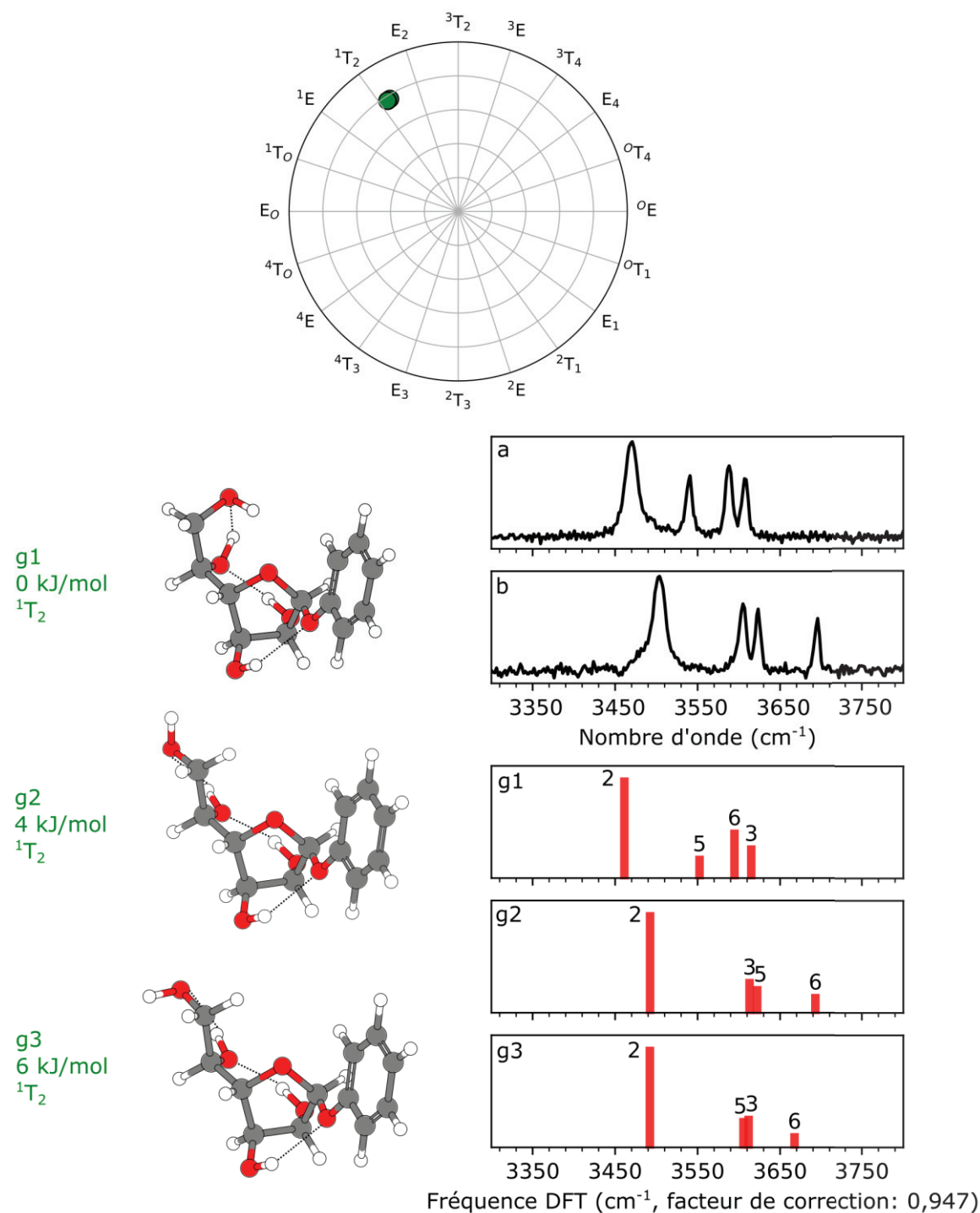


Figure 16 En haut : coordonnées polaires des trois conformations les plus stables de β GalFOPh en phase gazeuse. En bas : noms, énergies et conformations de cycle de ces trois conformations (à gauche), leurs structures (au milieu) et leurs spectres simulés (à droite). Les spectres expérimentaux des conformères a et b sont également représentés.

La conformation g1 est la plus stable énergétiquement. Elle présente une liaison non covalente O3H--OPh à l'avant du cycle et une chaîne de liaisons hydrogène O2H--O5H--O6H--O4 à l'arrière. De ce fait, la fréquence de O3H est calculée à 3615 cm⁻¹, celle de O2H à 3460 cm⁻¹, de O5H à 3550 cm⁻¹ et de O6H à 3595 cm⁻¹.

La seconde conformation g2 a une énergie relative de 4 kJ/mol. Elle présente le réseau de liaisons O2H--O5H--O6H à l'arrière et O3H--OPh à l'avant du cycle. On observe une rotation des groupes O5H et O6H, ce qui modifie les fréquences des OH. On a respectivement 3490, 3620 et 3690 cm⁻¹ pour les modes O2H, O5H et O6H. La fréquence de O3H n'est pas affectée (3615 cm⁻¹).

La dernière conformation en dessous du seuil énergétique est g3, dont l'énergie relative est de 6 kJ/mol. Elle présente une rotation de O6H par rapport à g2, et donc un petit décalage des modes O5H (3600 cm⁻¹) et O6H (3670 cm⁻¹).

Le spectre expérimental du conformère a ne possédant pas de bande expérimentale au-dessus de 3650 cm⁻¹, il ne correspond pas aux structures g2 et g3 qui ont leur mode O6H au-dessus de cette valeur. Seule la conformation la plus stable g1 répond à ce critère. Il est par ailleurs le seul à avoir un mode calculé autour de 3540 cm⁻¹. La structure g1 correspond donc au conformère a. Ainsi la bande à 3470 cm⁻¹ s'explique par le mode O2H, celle à 3540 cm⁻¹ par le mode O5H et le doublet par O6H et O3H.

Le conformère b a une bande expérimentale à 3695 cm⁻¹ qui est à la même position que le mode O6H de g2. Le mode O6H de g3 présente un écart trop important (25 cm⁻¹) pour expliquer cette bande. Les trois autres bandes de ce conformère correspondent aussi à des modes calculés de g2 : la bande à 3620 cm⁻¹ s'explique par O5H, celle à 3600 cm⁻¹ par O3H et à 3500 cm⁻¹ par O2H.

La comparaison des spectres montre une bonne correspondance du spectre expérimental a avec le spectre théorique de g1 et du spectre expérimental b avec le spectre théorique de g2.

4.1.4 Impact du passage de la solution à la phase gazeuse

Passons en revue les résultats de l'étude menée sur le βGalfOMe en phase liquide aboutissant à des résultats différents.²² Les deux anomères α et β de GalfOMe ont été analysés avec d'une part des simulations DFT avec un modèle implicite de solvatation et d'autre part en utilisant le programme PSEUROT (calculs lancés quatre fois). Les simulations DFT et les quatre simulations PSEUROT donnent des résultats cohérents sur la forme α : une conformation de cycle ²E est adoptée par la molécule.

En revanche, il existe quelques différences pour les résultats sur βGalfOMe : alors que la DFT prédit une conformation de cycle ⁴E, les quatre simulations sur PSEUROT donnent des résultats différents (⁴E, ⁴E, E₁ et ⁰E/⁰T₁). Ces simulations semblent indiquer une conformation de cycle ⁴E mais toutes les simulations n'aboutissent pas à ce même résultat.

En phase gazeuse, les différentes simulations du β-galactofuranose que nous avons effectuées donnent des résultats similaires pour les deux équipes. En effet, l'équipe de l'ISMO a également réalisé des calculs sur la molécule neutre et a obtenu des données similaires aux nôtres (deux conformations de type ¹T₂ avec les mêmes liaisons hydrogène). L'obtention de résultats identiques par deux équipes de recherche ayant chacune sa propre méthode de calculs quantiques est très convaincante. Les études en phase gazeuse sont donc particulièrement intéressantes pour les molécules flexibles car elles simplifient leur analyse. En effet, la molécule se retrouve isolée de tout environnement et n'a donc pas d'interaction avec le solvant ou avec une autre molécule.

Nous pouvons alors étudier les facteurs intrinsèques et extrinsèques à la molécule. Alors que le galactofuranose est très flexible en solution, seules deux conformations sont observées en phase gazeuse et elles appartiennent toutes deux à la famille des ¹T₂. Une série de liaisons hydrogène n'est

4. Impact du micro-environnement sur la conformation du galactofuranose

pas envisageable comme dans le cas des pyranoses en raison de la configuration du furanose. On observe néanmoins un nombre important de liaisons hydrogène (3 pour chaque conformère) qui stabilisent la structure. Il est à noter que les deux conformères se distinguent par une faible rotation de deux groupements OH. On voit que l'IR est sensible à chaque rotation et est donc très avantageux pour étudier le comportement du galactofuranose en phase gazeuse.

4.2 Interaction du galactofuranose avec une molécule d'eau

4.2.1 Recherche conformationnelle du galactofuranose complexée à une molécule d'eau

Nous avons ensuite étudié le complexe formé par β Gal f OPh et une molécule d'eau. L'analyse par IR de molécules complexées à une ou plusieurs molécules d'eau est un domaine d'expertise de l'équipe SYSIPHE (ISMO). La complexation est permise par ajout de H₂O dans l'argon. Les études de ces complexes, y compris sur le galactopyranose, démontrent que la molécule d'eau se positionne préférentiellement là où se trouve la plus faible liaison et stabilise le complexe.²⁴ Les différentes étapes du complexe [β Gal f OPh+H₂O] sont affichées en Figure 17.

Notons que les optimisations par PM7 et B3LYP permettent de faire un grand tri des conformations. Nous avons fait le choix de n'optimiser par CAMB3LYP que les 100 conformations les plus stables obtenues par B3LYP. Cette optimisation donne 89 conformères différents et certaines familles de conformation (à l'est, de ⁰T₄ à ²T₁) ne sont plus représentées. Parmi ces conformations, cinq ont une énergie calculée sous la valeur seuil (8,7 kJ/mol). Il s'agit de quatre conformères nord-ouest (¹T₂ ou E₂) et un conformère sud-ouest (⁴E).

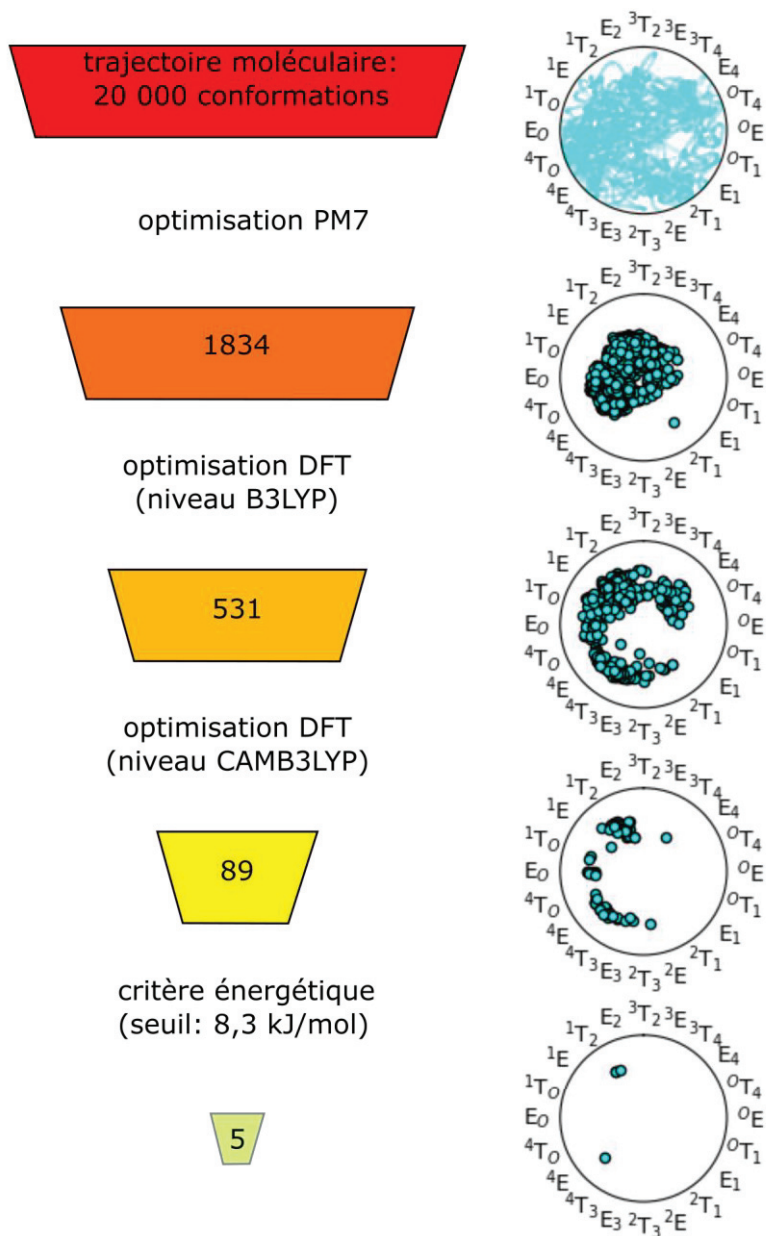


Figure 17 Nombre de conformations de β GalfOPh complexé à une molécule d'eau après chaque étape de l'étude conformationnelle et leurs coordonnées Altona Sundaralingam.

4. Impact du micro-environnement sur la conformation du galactofuranose

4.2.2 R2PI du complexe

L'ion moléculaire $[\beta\text{Gal}f\text{OPh}+\text{H}_2\text{O}]^+$ est détecté à m/z 274. Son signal a été suivi pendant le balayage du laser UV pour l'acquisition du spectre R2PI (Figure 18).

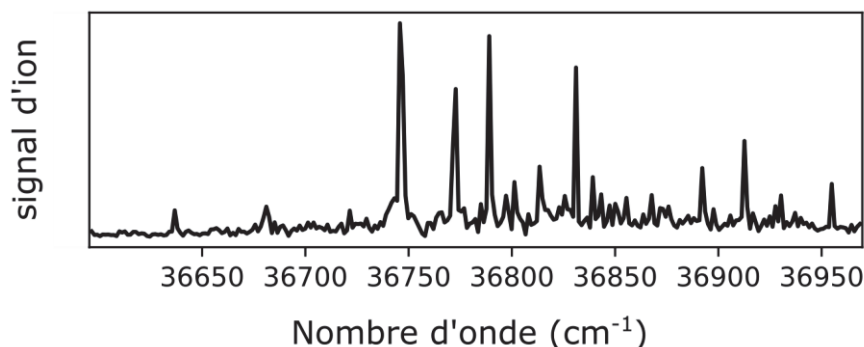


Figure 18 Spectre R2PI du complexe $[\beta\text{Gal}f\text{OPh}+\text{H}_2\text{O}]$.

Plusieurs bandes UV sont observées, notamment à 36745 ; 36770 ; 36790 et 36830 cm^{-1} . Nous pouvons noter que le profil est proche de celui du conformère a dans la Figure 15 qui présente également 4 pics intenses à des valeurs proches. Une seule conformation a été observée pour ce complexe.

4.2.3 Calcul des spectres vibrationnels

Le spectre IR expérimental a été acquis en fixant le laser UV. Il possède six bandes : un doublet à 3400 et 3440 cm^{-1} , une large bande à 3485 cm^{-1} puis trois bandes résolues à 3545 ; 3605 et 3730 cm^{-1} (Figure 19). L'expérimentation a également été menée sur le complexe $[\beta\text{Gal}f\text{OMe}+\text{D}_2\text{O}]^+$ en ajoutant de l'eau deutérée dans l'argon. Les bandes à 3485 et 3730 cm^{-1} disparaissent alors, suggérant que ces deux bandes correspondent à des modes OH de la molécule d' H_2O .

Le spectre expérimental a ensuite été comparé aux cinq conformères répondant au critère énergétique.

Les deux structures les plus stables w1 et w2 appartiennent à la famille des ${}^1\text{T}_2$. On retrouve la liaison O3H--OPh à l'avant du cycle et un large réseau O2H--O5H-- H_2O --O6H à l'arrière. Les deux structures se différencient par une rotation de la molécule d'eau, ce qui a un impact mineur sur l'énergie (écart de 1,87 kJ/mol) et les fréquences calculées. On trouve pour w1 les deux modes OH de l' H_2O à 3725 et 3445 cm^{-1} , le mode O3H à 3630 cm^{-1} , le mode O6H à 3555 cm^{-1} , celui de O2H à 3425 et de O5H (fortement lié à H_2O) à 3360 cm^{-1} . Les valeurs pour w2 sont respectivement 3730 et 3450 cm^{-1} (H_2O), 3625 cm^{-1} (O3H), 3545 cm^{-1} (O6H), 3440 cm^{-1} (O2H) et 3375 cm^{-1} (O5H).

Les trois conformations suivantes ont un écart énergétique important par rapport à la molécule la plus stable. La structure w3 a un écart énergétique de 7,27 kJ/mol et adopte une conformation de type E2. La même liaison O3H--OPh est présente, et donc le mode calculé de O3H est similaire à celui des conformères précédents (3625 cm^{-1}). On note un réseau hydrogène qui se forme à l'arrière du cycle : O2H--O5H-- H_2O --O6H--O4. À l'inverse de w1 et w2, O6H est le donneur de protons et H_2O le receveur. Les modes O2H, O5H et O6H sont déplacés à respectivement 3410 ; 3500 et 3430 cm^{-1} . Les deux modes OH de H_2O sont à 3480 et 3730 cm^{-1} .

La structure w4 (7,88 kJ/mol) est très différente des autres conformères. Notons d'abord que son cycle se plisse en ${}^4\text{E}$ et est le seul situé au sud-ouest de la roue. Ensuite la position de H_2O est également très différente, situé entre O2H et O3H. De ce fait on obtient une suite de liaisons intramoléculaires O2H-- H_2O --O3H--O6H. Les modes de O6H et O5H, libres de liaison, sont calculés à 3690 et 3650 cm^{-1} . Au contraire, les groupes O2H et O3H, fortement liés et à proximité de H_2O , ont leurs fréquences

respectives à 3480 et 3490 cm^{-1} . Enfin, la molécule d'eau a deux modes OH à 3725 et 3420 cm^{-1} . Cette structure fait penser à af7 et af8 (page 66) pour lesquelles le groupement NH_4^+ était positionné comme la molécule H_2O , aboutissant à un réseau de liaisons hydrogènes proche. Ces deux structures adoptaient également des conformations sud mais avaient été éliminées par le critère spectroscopique.

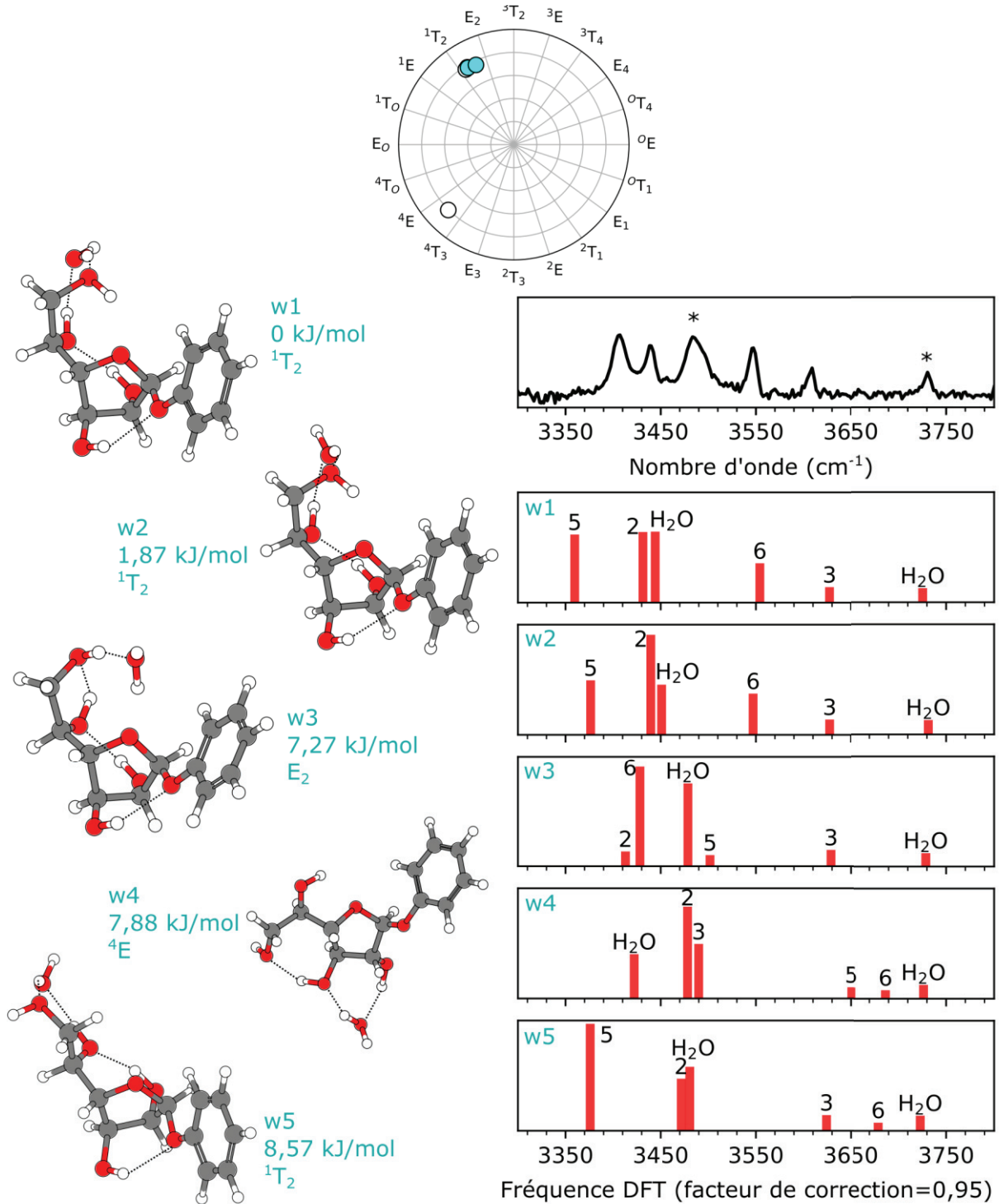


Figure 19 En haut : coordonnées polaires de la conformation la plus stable de $[\beta\text{GalFOPh}+\text{H}_2\text{O}]$ en phase gazeuse. En bas : nom, énergie et conformation de cycle de cette conformation (à gauche), sa structure (au milieu) et son spectre simulé (à droite). Le spectre expérimental du conformère est également représenté par une ligne noire continue (les étoiles correspondent aux bandes qui disparaissent en remplaçant le H_2O par du D_2O).

4. Impact du micro-environnement sur la conformation du galactofuranose

Enfin, w5 appartient à la même famille que w1 et w2 et forme les mêmes liaisons hydrogène. Cette structure est marquée par une rotation de O6H et un léger déplacement de la molécule d'eau.

Passons à la sélection des conformères selon le critère spectroscopique en comparant avec le spectre expérimental. La bande à 3730 cm^{-1} correspond à une bande de H_2O pour les cinq structures, ce qui est d'ailleurs validé par le spectre obtenu avec le complexe deutéré. Ensuite, une seconde bande expérimentale est mesurée vers 3610 cm^{-1} . Cette dernière correspond au mode O3H des conformères w1, w2, w3 et w5. L'écart avec le mode O5H de w4 est trop important (3650 cm^{-1}). Nous pouvons ainsi éliminer ce conformère. Par ailleurs, le mode O6H de w5 ne correspondant à aucune bande expérimentale, nous pouvons également exclure ce conformère.

Parmi les trois conformations restantes, seules w1 et w2 expliquent la bande à 3545 cm^{-1} qui correspondrait à O6H. Le conformère w3, qui par ailleurs présentait une énergie relative trop élevée, peut donc être éliminé car aucun de ses modes calculés n'explique la bande expérimentale à 3545 cm^{-1} . On sait que la bande à 3485 cm^{-1} est une bande de l'eau grâce au spectre obtenu avec le deutérium. Les modes OH de H_2O calculés (à 3445 cm^{-1} pour w1 et 3450 cm^{-1} pour w2) ne correspondent donc pas parfaitement.

Enfin, la bande expérimentale à 3440 cm^{-1} est due au mode O2H et celle à 3400 cm^{-1} à O5H. On notera également un décalage entre la bande calculée et expérimentale de O5H, qui est plus faible pour w2. Il est difficile de choisir entre les deux conformères dont les énergies sont très proches. Les deux structures répondent donc aux critères énergétique et spectroscopique. Comme les structures sont proches (rotation de H_2O), il est normal que l'UV n'ait pas pu les distinguer.

4.2.4 Impact de l'ajout d'une molécule d'eau

La molécule neutre $\beta\text{GalFOPh}$ adopte un cycle ${}^1\text{T}_2$. Ceci est le cas qu'elle soit isolée ou complexée à une molécule d'eau. Lorsqu'elle est isolée, deux conformères coexistent. Seule la plus stable est présente lors de la complexation avec une molécule d'eau. Elle existe alors sous plusieurs formes mais avec la molécule figée et seulement une rotation de H_2O . La molécule d'eau stabilise donc la molécule. La molécule H_2O s'intercale entre les groupements O5H et O6H. Pour cette raison, les modes O3H de la molécule isolée et de la molécule complexée à H_2O sont identiques. Au contraire, les modes O5H et O6H sont fortement déplacés pour le complexe car elles interagissent avec la molécule d'eau.

Tout comme dans le cas du galactopyranose,²⁴ la molécule d'eau s'intercale au niveau de la liaison la plus faible pour former deux liaisons hydrogène fortes.

4.3 Interaction du galactofuranose avec un ion

4.3.1 Recherche conformationnelle du galactofuranose complexé au lithium

La trajectoire moléculaire aboutit à 20 000 conformations qui sont optimisées par PM7. Il ne reste alors plus que 704 conformères après suppression des doublons puis 184 conformères après l'étape DFT (B3LYP puis CAMB3LYP). Quatre structures ont une énergie (niveau CAMB3LYP) inférieure à 8,3 kJ/mol. La plus stable a un cycle en 1T_2 et les trois autres appartiennent à la famille des 1E (Figure 20).

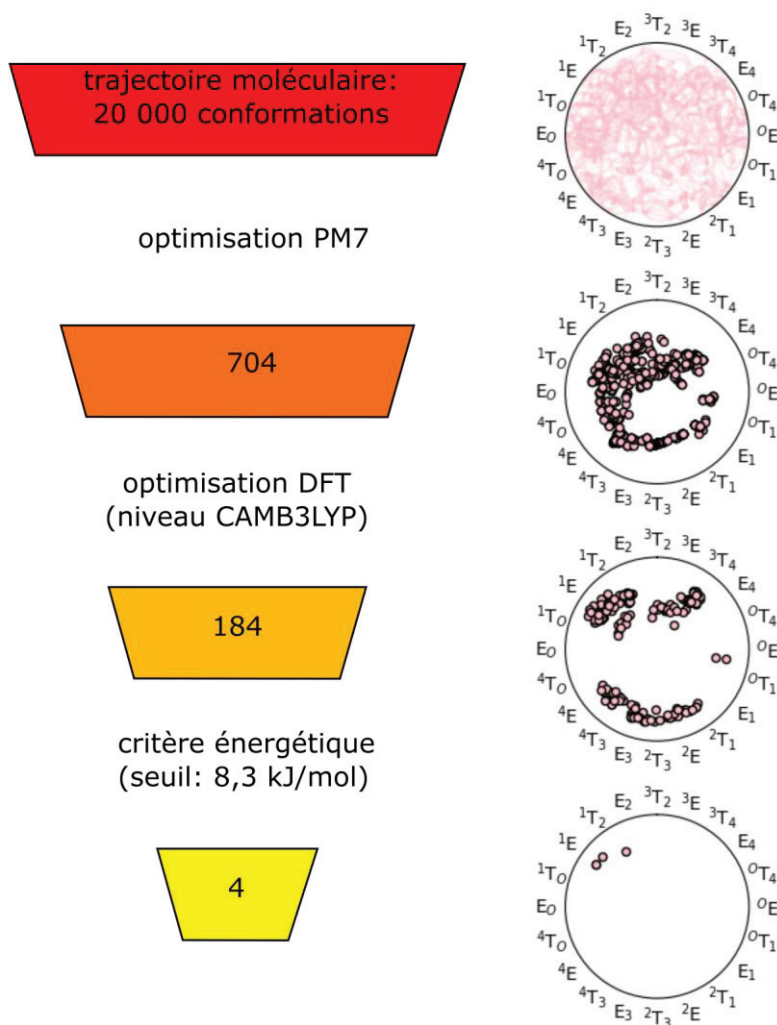


Figure 20 Nombre de conformations de $[\beta\text{GalfOMe}+\text{Li}]^+$ à chaque étape de l'étude conformationnelle et leurs coordonnées Altona Sundaralingam.

4.3.2 Calculs des spectres vibrationnels

Le spectre expérimental du complexe $[\beta\text{GalfOMe}+\text{Li}]^+$ a déjà été décrit dans le chapitre 2 (page 40). Il est repris par une ligne continue grise en Figure 21. Nous avons mené l'étude théorique sur ce complexe qui a abouti à quatre conformations en dessous de 8,3 kJ/mol. Alors que la plus stable l1 adopte une conformation de cycle 1T_2 , les trois autres appartiennent à la famille des 1E . Les structures sont représentées en Figure 21.

Nous pouvons noter que les quatre structures présentent une liaison O3H--OMe favorisée par l'anomérie en β , la fréquence de O3H est ainsi décalée entre 3600 et 3625 cm^{-1} . Le mode O3H explique donc la bande expérimentale à 3595 cm^{-1} .

4. Impact du micro-environnement sur la conformation du galactofuranose

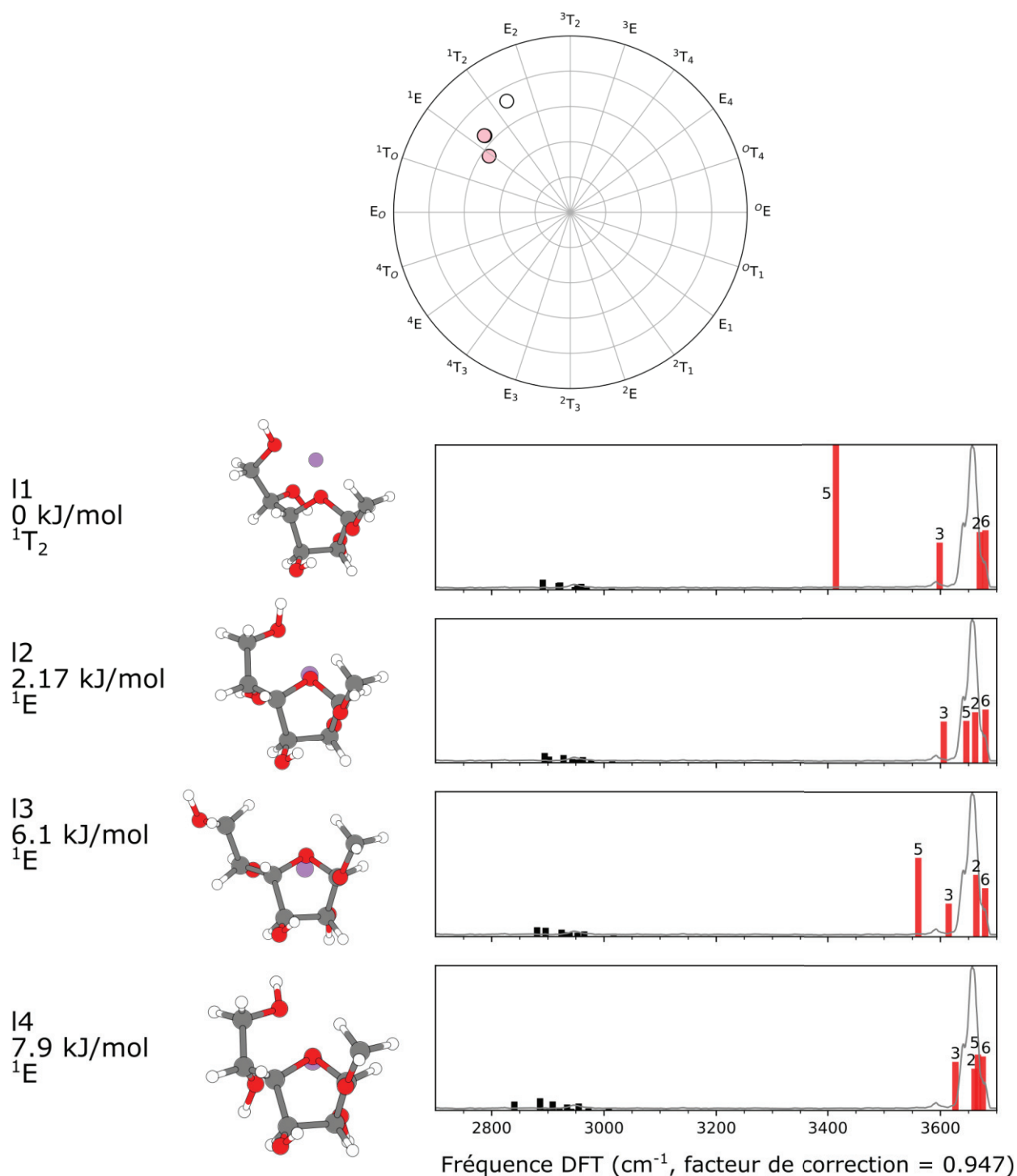


Figure 21 En haut : coordonnées polaires des conformations les plus stables de β GalFOPh complexée au lithium. En bas : noms, énergies et conformations de cycle de ces conformations (à gauche), leurs structures (au milieu) et leurs spectres simulés (à droite).

La structure la plus stable I1 est la seule à adopter une conformation de type 1T_2 . Elle se caractérise par la présence d'une forte liaison O5H--O2 à l'arrière du cycle. La fréquence calculée du mode O5H est à 3410 cm^{-1} . Comme il n'y a pas d'absorption à cette longueur d'onde, nous pouvons exclure la conformation de cycle 1T_2 . Les trois autres conformères adoptent un cycle 1E . La structure I3 (6,1kJ/mol) forme une faible liaison O5H--O6H, son mode O5H est décalé à 3560 cm^{-1} . Or il n'y a pas de bande expérimentale correspondante, nous pouvons donc également éliminer ce conformère. Les deux autres conformères se différencient par une rotation de O5H qui est libre dans les deux cas. Ils

n'ont pas de liaison intracyclique à l'arrière du cycle et ont donc les modes OH hors O3H entre 3640 et 3680 cm^{-1} . Cela correspond parfaitement à la bande expérimentale vers 3660 cm^{-1} .

D'une part, les outils théoriques nous donnent l'explication des bandes expérimentales : la faible bande à 2950 cm^{-1} correspond aux modes CH, celle à 3595 cm^{-1} s'explique par le mode O3H qui est faiblement lié et la bande intense vers 3660 cm^{-1} s'apparente aux autres OH qui sont tous libres.

D'autre part, cette comparaison valide les structures expérimentalement présentes : I2 et I4. Alors que le critère énergétique sélectionne des cycles 1T_2 et 1E , on se restreint uniquement à la famille des 1E en ajoutant le critère spectroscopique.

4.3.3 Impact de l'ajout d'une charge

Comme nous souhaitons étudier une espèce chargée, l'utilisation de l'IR-UV n'est plus adaptée. Le changement de l'instrumentation implique en conséquence des modifications forcées de paramètres outre celui que l'on voulait volontairement changer pour analyser son impact (l'ajout d'une charge). La température n'est plus la même et on sait qu'elle a un effet important sur la flexibilité.

L'ajout du lithium semble légèrement impacter la conformation du cycle par rapport à la molécule neutre : elle adopte la conformation de cycle voisine 1E . Le lithium se place à l'arrière du cycle, là où le nombre de liaisons hydrogène potentielles est plus important. Il participe ainsi à l'échange de protons et stabilise le conformère.

4.4 Interprétation

Le $\beta\text{GalfOMe}$ nous a servi de modèle pour notre étude complémentaire théorie/expérience pour rationaliser les propriétés intrinsèques et extrinsèques du galactofuranose. Les simulations théoriques ayant servi à l'interprétation des résultats IRMPD et IR-UV sont identiques : les mêmes paramètres de dynamique moléculaire et de base de DFT (CAMB3LYP) sont employés. Les conditions expérimentales sont en revanche différentes selon l'instrumentation. Tandis que les analytes irradiés par IRMPD sont chargés et à température ambiante, ils sont neutres et refroidis lors des analyses IR-UV. Il existe également une différence chimique entre les deux molécules (groupe méthyle ou phényle).

Enfin, une particularité importante est que l'on irradie l'ensemble de la population en IRMPD alors qu'il y a une sélection de conformère en IR-UV. Ceci est particulièrement intéressant dans le cas de molécule flexible où plusieurs conformations co-existent.

Nous avons étudié le comportement du galactofuranose dans différents environnements. L'utilisation de l'infrarouge pour répondre à cette problématique est adaptée puisqu'il est sensible aux variations conformationnels d'une molécule au détail atomique et également parce qu'on peut contrôler les paramètres du micro-environnement.

En solution, la conformation de cycle 4E était prédite même si tous les calculs n'aboutissaient pas forcément à ce résultat. Les conformations adoptées en phase gazeuse dans les différents micro-environnements sont reprises en Figure 22. Lorsque l'on isole le β -galactofuranose en phase gazeuse avec ajout d'une molécule d'eau, on voit que la dynamique moléculaire aboutit majoritairement à des conformations de cycle 1T_2 mais que l'on a également une conformation de type 4E . Il est donc énergétiquement possible que le complexe avec la molécule d'eau adopte un cycle 4E , comme en solution. Cette possibilité est cependant éliminée lorsque l'on ajoute le critère spectroscopique. En revanche, la dynamique moléculaire réalisée sur la molécule sans aucune molécule d'eau n'aboutit qu'à des conformères de la famille des 1T_2 . Le cycle 4E n'est donc énergétiquement plus possible.

4. Impact du micro-environnement sur la conformation du galactofuranose

Ainsi qu'il soit isolé ou complexé à une molécule d'eau, le β -galactofuranose adopte un cycle de type 1T_2 . La molécule d'eau se place au niveau du bras C5-C6 et n'impacte pas la conformation adoptée par le cycle. Lorsque l'on complexe la molécule à un ion NH_4^+ , ce dernier peut se placer à deux positions. S'il est comme la molécule d'eau proche du bras, on retrouve un cycle 1T_2 . S'il se positionne à proximité du cycle, le cycle adopte la conformation voisine 1E . Enfin, si l'on charge la molécule avec du Li^+ , ce dernier se place à proximité du cycle derrière O4 et on n'obtient plus que des conformations de type 1E .

En poussant la description à une analyse plus fine, nous remarquerons que deux rotamères existent pour la molécule isolée. Seul le plus stable énergétiquement est conservé par ajout d'une molécule d'eau. La molécule d'eau s'intercale au niveau de la plus faible liaison et stabilise le conformère. La présence d'une charge tend à l'inverse à un nombre un peu plus élevés de conformères. Nous avons pu décrire chaque conformère au niveau atomique (Figure 22).

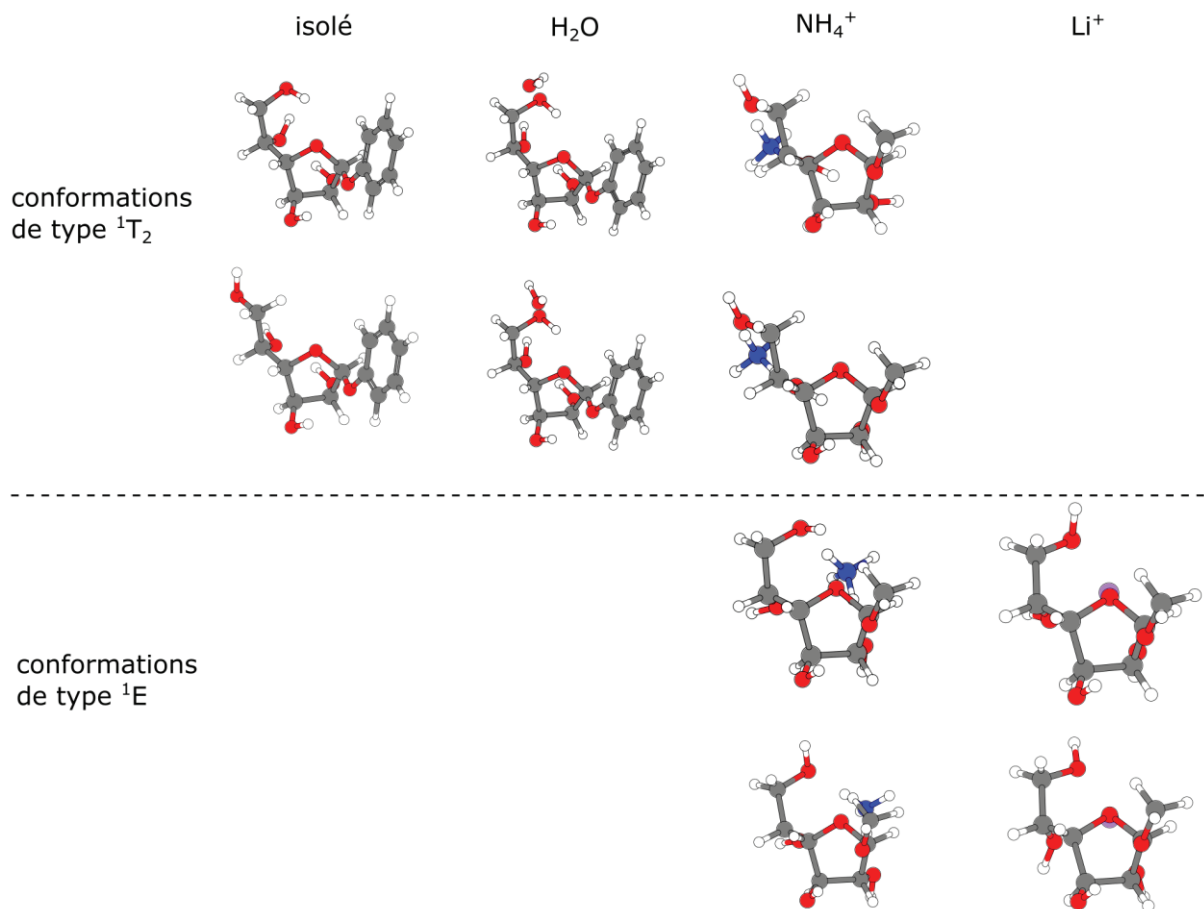


Figure 22 Conformations retenues pour les quatre études du β -galactofuranose (isolé, hydraté, avec adduit NH_4^+ et avec adduit Li^+)

Chapitre 3

Nous pouvons en déduire que la flexibilité du galactofuranose est surtout liée à des conditions extrinsèques à la molécule, lorsque celle-ci est placée en solution. Nous avons vu en étudiant la molécule au détail atomique qu'elle stabilisait en phase gazeuse avec la formation de deux chaînes de liaisons hydrogène, une à l'avant du cycle (avec une seule liaison) et l'autre à l'arrière (avec une suite de deux liaisons). Nous pouvons penser que cela n'est pas suffisant pour stabiliser la molécule et qu'en présence de molécules de H₂O, celles-ci viennent s'intercaler et modifier la conformation du cycle. À l'inverse, le galactopyranose en phase gazeuse est très stable avec quatre liaisons successives d'hydrogène. En présence d'eau, les molécules d'H₂O tendent à stabiliser les systèmes pyranose en coopérant au réseau hydrogène et sans affecter la conformation du cycle.

5 Conclusion

Il existe des milliers de conformères théoriques pour un monosaccharide étant donné que les groupes hydroxyle ont une liberté rotationnelle. En prenant en considération les liaisons intramoléculaires, on réduit drastiquement le nombre de conformères probables. Ces liaisons contribuent à réduire l'énergie interne d'une structure et modifient les fréquences IR. Notre combinaison théorie/expérience a pour but de sélectionner les structures les plus probables selon ces deux critères (énergétique et spectroscopique).

À l'aide de ces outils, nous avons dans un premier temps étudié l'impact de la taille de cycle sur le galactose. La description des liaisons intramoléculaires a été rendue possible en utilisant le niveau de DFT CAM-B3LYP/6-311++G(2df,2pd). Avec un cycle à 6, le galactose adopte une conformation chaise stabilisée par des liaisons hydrogène successives. Un cycle à 5 ne permet pas de réaliser une chaîne continue de liaisons intramoléculaires, seules certaines liaisons sont formées. Les données ont par ailleurs permis de rationaliser le diagnostic des spectres IRMPD sur la taille de cycle.

Dans un second temps, nous nous sommes intéressés à l'impact des conditions locales sur les préférences conformationnelles du galactofuranose, très flexible en solution. Son paysage conformationnel en phase gazeuse est drastiquement réduit. Nous sommes parvenus, à l'aide d'outils complémentaires basés sur l'UV, à estimer le nombre de conformères présents dans différents micro-environnements.

La combinaison de la spectroscopie IR en phase gazeuse et de chimie quantique à haut niveau offre une description très détaillée de la conformation du galactofuranose. Étudier la conformation de galactofuranose dans des polysaccharides en solution est un défi colossal mais d'importance biologique importante car cela faciliterait la compréhension de la relation structure/activité. Notre travail sur le monosaccharide galactofuranoside isolé en phase gazeuse est une première étape. Les calculs théoriques sur de plus grands systèmes sont une perspective mais le haut niveau de calculs rend cela extrêmement complexe à l'heure actuelle.

6 Bibliographie

- 1 V. S. R. Rao, *Conformation of Carbohydrates*, CRC Press, London, 1998.
- 2 A. Varki, *Glycobiology*, 2017, **27**, 3–49.
- 3 A. R. Ionescu, A. Bérces, M. Z. Zgierski, D. M. Whitfield and T. Nukada, *J. Phys. Chem. A*, 2005, **109**, 8096–8105.
- 4 G. J. Davies, A. Planas and C. Rovira, *Acc. Chem. Res.*, 2012, **45**, 308–316.
- 5 *Eur. J. Biochem.*, 1980, **111**, 295–298.
- 6 J. Iglesias-Fernández, L. Raich, A. Ardèvol and C. Rovira, *Chem. Sci.*, 2015, **6**, 1167–1177.
- 7 L. Barnes, B. Schindler, S. Chambert, A.-R. Allouche and I. Compagnon, *Int. J. Mass Spectrom.*, 2017, **421**, 116–123.
- 8 L. Barnes, A.-R. Allouche, S. Chambert, B. Schindler and I. Compagnon, *Int. J. Mass Spectrom.*, 2020, **447**, 116235.
- 9 H. A. Taha, M. R. Richards and T. L. Lowary, *Chem. Rev.*, 2013, **113**, 1851–1876.
- 10 A. M. Rijs and J. Oomens, in *Gas-Phase IR Spectroscopy and Structure of Biological Molecules*, eds. A. M. Rijs and J. Oomens, Springer International Publishing, Cham, 2014, vol. 364, pp. 1–42.
- 11 O. Yeni *et al.*, Conformational preferences of the flexible galactofuranose sugar in gas-phase. (soumis)
- 12 O. Yeni *et al.*, Galactofuranose flexibility in different micro-environment. (en rédaction)
- 13 J. J. P. Stewart, *MOPAC2012*, 2012.
- 14 C. J. Gray, B. Schindler, L. G. Migas, M. Pičmanová, A. R. Allouche, A. P. Green, S. Mandal, M. S. Motawia, R. Sánchez-Pérez, N. Bjarnholt, B. L. Møller, A. M. Rijs, P. E. Barran, I. Compagnon, C. E. Eyers and S. L. Flitsch, *Anal. Chem.*, 2017, **89**, 4540–4549.
- 15 A. D. Becke, *Phys. Rev. A*, 1988, **38**, 3098–3100.
- 16 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 1372–1377.
- 17 C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B*, 1988, **37**, 785–789.
- 18 T. Yanai, D. P. Tew and N. C. Handy, *Chem. Phys. Lett.*, 2004, **393**, 51–57.
- 19 A.-R. Allouche, *J. Comput. Chem.*, 2011, **32**, 174–182.
- 20 D. Cremer and J. A. Pople, *J. Am. Chem. Soc.*, 1975, **97**, 1354–1358.
- 21 C. Altona and M. Sundaralingam, *J. Am. Chem. Soc.*, 1972, **94**, 8205–8212.
- 22 M. R. Richards, Y. Bai and T. L. Lowary, *Carbohydr. Res.*, 2013, **374**, 103–114.
- 23 E. Gloaguen, V. Brenner, M. Alauddin, B. Tardivel, M. Mons, A. Zehnacker-Rentien, V. Declerck and D. J. Aitken, *Angew. Chem. Int. Ed.*, 2014, **53**, 13756–13759.
- 24 P. Çarçabal, R. A. Jockusch, I. Hünig, L. C. Snoek, R. T. Kroemer, B. G. Davis, D. P. Gamblin, I. Compagnon, J. Oomens and J. P. Simons, *J. Am. Chem. Soc.*, 2005, **127**, 11414–11425.

CHAPITRE 4 :

Analyse par MS-MS/IR d'oligosaccharides contenant du galactofuranose

<u>1</u>	Introduction	90
<u>2</u>	Détails expérimentaux	92
	<u>2.1 Séquençage</u>	92
	<u>2.2 Criblage</u>	92
<u>3</u>	Séquençage d'oligosaccharides de synthèse contenant du galactofuranose	94
	<u>3.1 Développement de la méthode de séquençage sur des trisaccharides synthétiques</u>	94
	<u>3.2 Caractérisation complète d'un dp3 par analyse de fragments disaccharidiques</u>	99
	<u>3.3 Méthode universelle pour déterminer la taille du cycle par analyse des fragments monosaccharidiques</u>	107
<u>4</u>	Développement d'une méthode de criblage en LC-MS/MS-IR	110
	<u>4.1 Méthode de criblage par irradiation IR</u>	110
	<u>4.2 Choix des paramètres LC-MS/MS-IR</u>	110
	<u>4.3 Criblage par irradiation infrarouge non résonante</u>	111
	<u>4.4 Criblage par irradiation infrarouge résonante</u>	115
	<u>4.5 Criblage par irradiation infrarouge résonante à longueur d'onde spécifique au galactofuranose</u>	117
	<u>4.6 Interprétation</u>	120
<u>5</u>	Conclusion	121
<u>6</u>	Bibliographie	122

1 Introduction

D'après notre étude sur les mono- et disaccharides (chapitre 2), nous savons que les structures du galactofuranose et du galactopyranose sont conservées en ESI-MS. Par conséquent, la MS-IR convient pour révéler la taille de cycle d'un petit composé. Il est en revanche difficile de l'appliquer à des sucres de plus grande taille. D'une part, il faudrait synthétiser toutes les séquences possibles, sachant que la complexité de synthèse et le nombre de standards augmentent avec la taille du composé. D'autre part, la résolution structurale devra être suffisante pour distinguer les différents sucres synthétisés. Notons que la résolution structurale diminue avec la taille en IRMPD car un sucre plus grand possède plus de groupements hydroxyle et donc plus de bandes infrarouge dans la même gamme, ce qui congestionne le spectre.

L'introduction de la MS en tandem est avantageuse car on n'analyse pas l'ion dans sa globalité mais ses fragments. L'identification des fragments permet ensuite de remonter à la structure de l'ion parent. Cette approche présente deux atouts : d'une part, il y a moins de références à comparer et d'autre part la résolution structurale est meilleure car le fragment analysé est de plus petite taille que l'ion précurseur.

La nomenclature proposée par Domon et Costello indique les différents types de fragments obtenus par spectrométrie de masse en tandem sur les carbohydrates.¹ Les fragments les plus abondants en CID sont obtenus lorsque le clivage a lieu au niveau de la liaison glycosidique :

- Les fragments b_x (en extrémité non réductrice) et leurs complémentaires y_x (en extrémité réductrice) si la cassure a lieu avant la liaison glycosidique,
- Les fragments c_x (en extrémité non réductrice) et leurs complémentaires z_x (en extrémité réductrice) si la cassure se situe après la liaison glycosidique.

Le nombre en indice x correspond au nombre d'unités monosaccharidiques restant dans le fragment. Les fragments que l'on peut obtenir en fragmentant un trisaccharide (le Gal-Man-Glc) figurent en Schéma 1.

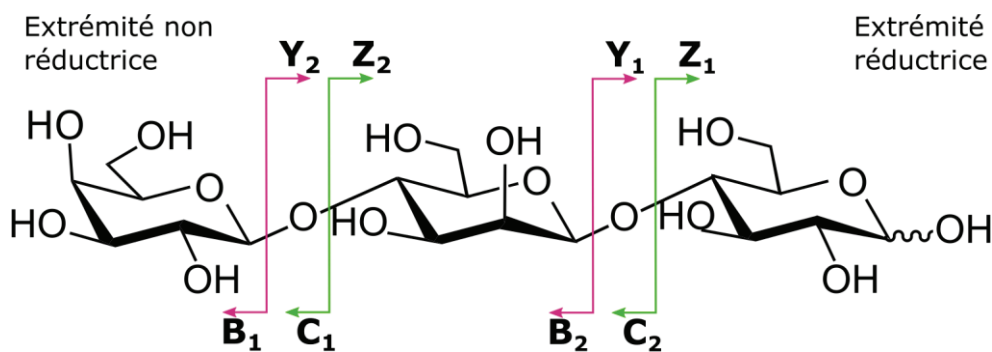


Schéma 1 Dénomination des fragments obtenus par MS/MS de Gal-Man-Glc selon la nomenclature de Domon et Costello.

Ainsi le fragment c_1 correspond au Gal et le fragment c_2 au Gal-Man. Le fragment y_1 est composé de Glc et le fragment y_2 de Man-Glc. Tandis que les fragments c_x et y_x sont constitués de sucres intacts, les fragments b_x et z_x sont composés d'un sucre déshydraté (perte de 18 Da). Par exemple, le fragment b_1 correspond au Gal déshydraté.

Des clivages intra-cycliques sont produits en abondance par des méthodes de fragmentation avec activation électronique (ECD, ETD...), mais ils sont peu observés en CID et ne seront donc pas exploités dans ce chapitre.

1. Introduction

Deux approches basées sur la MS-MS/IR ont été développées pour l'analyse des oligosaccharides contenant du galactofuranose :

- La méthode de séquençage (voir page 27), mise au point par l'équipe précédemment² et brevetée. Elle sert à élucider la séquence complète d'un oligosaccharide mais son application sur les oligosaccharides contenant des cycles furanose n'a pas été vérifiée.
- La méthode de criblage (voir page 27) développée lors de cette thèse.

Lors de la fragmentation par CID, l'ion isolé dans la trappe est accéléré par un potentiel électrique et acquiert une énergie cinétique. Sa collision avec un gaz neutre (l'hélium dans nos expérimentations) provoque une augmentation de son énergie interne puis sa fragmentation. Ce mécanisme peut être à l'origine de réarrangements structuraux, comme la migration de fucose, qui induisent des résultats ambigus et difficiles à interpréter.^{3,4}

En raison de la faible stabilité du cycle furanose, la fragmentation par CID risque également d'induire des réarrangements structurels. Il convient donc de vérifier que l'information sur la taille du cycle est conservée après fragmentation par CID afin d'appliquer les méthodes de séquençage et de criblage.

Dans ce chapitre, nous présenterons dans un premier temps les spectres IRMPD des fragments obtenus par CID. On vérifiera ainsi que les structures galactofuranose et galactopyranose peuvent être différenciées après fragmentation. Une fois cette étape validée, les signatures IR des fragments seront analysées pour réaliser le séquençage d'un oligosaccharide contenant du galactofuranose. Dans une seconde partie, une méthode de criblage sera mise au point pour discriminer les oligosaccharides contenant du galactofuranose. Deux articles ont été rédigés à partir des résultats présentés dans ce chapitre.^{5,6}

2 Détails expérimentaux

2.1 Séquençage

Le dispositif instrumental utilisé est le LTQ-OPO. (figure page 17)

Préparation d'échantillon : nous utiliserons dans cette partie les disaccharides Gal-Man, leurs intermédiaires de synthèse Gal-ManSPh, des disaccharides Man-Man et des trisaccharides Gal-Man-Man (les protocoles de synthèse sont détaillés dans les références [7-9]). De l' H_2^{18}O marquée à 97% a été utilisée pour le marquage ^{18}O des trisaccharides. Pour cela, chaque trisaccharide a été incubé à température ambiante pendant au moins une semaine à une concentration de 2 μg par μL d' H_2^{18}O . Les disaccharides, disaccharides avec un groupement thiophényle et trisaccharides ^{18}O ont été préparés à une concentration de 1 $\mu\text{g}/\text{mL}$ dans de l'eau/méthanol (50/50), 0,1% de LiCl ont été ajoutés à la solution pour former les adduits au lithium.

Spectres MS/MS : Les disaccharides-SPh sont détectés à m/z 441. Ils sont isolés avec une fenêtre d'isolement de +/- 2,5 Da et fragmentés par CID pendant 30 ms. Les trisaccharides ^{18}O sont détectés à m/z 513. La fenêtre d'isolement a été réduite à 3 Da pour ne pas isoler les trisaccharides non marqués (car l'eau ^{18}O est pure à 97%). Les spectres de fragmentation ont été obtenus en fragmentant l'ion précurseur par CID pendant 30 ms.

Spectres IRMPD : la puissance de sortie de l'OPO est de 350 mW. Les spectres des disaccharides Gal-Man (présentés dans le chapitre 2) et Man-Man avec adduit lithium ont été obtenus en isolant m/z 349 et en l'irradiant pendant 400 ms. Le fragment m/z 331 des disaccharides-SPh est ensuite réisolé et irradié pendant 700 ms pour l'obtention des spectres IRMPD. Le fragment b_2 des trisaccharides ^{18}O (détecté à m/z 331) est réisolé et le temps d'irradiation pour l'acquisition de son spectre IRMPD est de 700 ms. Le fragment c_2 (détecté à m/z 349) est réisolé et le temps d'irradiation est de 400 ms.

2.2 Criblage

Le dispositif instrumental utilisé pour les courbes de puissance est le MS (LTQ XL) - IR (Firefly et/ou laser CO₂, les informations relatives à ces deux lasers sont disponibles page 19).

Préparation d'échantillon : le disaccharide testé est préparé à 1 $\mu\text{g}/\text{mL}$ dans un mélange eau/méthanol (50/50), on y ajoute 0,1% de LiCl pour former le complexe au lithium.

Infusion directe : les solutions diluées de disaccharides purs ont été injectées en infusion directe à l'aide d'un pousse-seringue dont le débit est à 3 $\mu\text{L}/\text{min}$.

Conditions MS : l'ion m/z 349 est isolé et fragmenté par CID pendant 30 ms, le fragment m/z 169 est réisolé pendant 150 ms. 50 scans sont enregistrés pour chaque puissance laser.

Conditions laser : la synchronisation entre la MS et l'IR permet l'irradiation du fragment m/z 169. La puissance laser du CO₂ est augmentée tous les 50 scans. Certaines expérimentations ont été réalisées avec le laser CO₂ seul et d'autres en combinaison avec le Firefly (3410 cm^{-1} , 110 mW).

Traitement de données : le taux de fragmentation $I\left(\frac{m}{z} = [50; 165]\right) / I\left(\frac{m}{z} = [167; 171]\right)$ est calculé pour chaque scan, puis la moyenne est établie sur les 50 scans réalisés dans les mêmes conditions de puissance laser. Les courbes de puissance sont obtenues quotidiennement. La comparaison des courbes de puissance des disaccharides Gal ρ -Man et Gal f -Man aide à choisir la puissance optimale et également à normaliser les taux de photofragmentation obtenus lors des manipulations de criblage.

2. Détails expérimentaux

Le dispositif instrumental utilisé pour le criblage est le LC (U3000) - MS (LTQ XL) - laser IR (Firefly et/ou laser CO₂). Le schéma du couplage figure page 21.

Préparation d'échantillon : 3 µL de chaque disaccharide (à environ 1 mg/ml) ont été dilués dans 500 µL de H₂O. Quatre mélanges ont été réalisés :

- Galpβ(1-2)Man + Galpβ(1-3)Man,
- Galfβ(1-2)Man + Galfβ(1-3)Man,
- Galpβ(1-2)Man + Galpβ(1-3)Man + Galpβ(1-4)Man + Galpβ(1-6)Man,
- Galfβ(1-2)Man + Galfβ(1-3)Man + Galfβ(1-4)Man + Galfβ(1-6)Man.

Conditions chromatographie : l'instrument utilisé est l'Ultimate 3000 (Thermo Fischer Scientific). Deux colonnes Hypercarb ont permis les séparations. Les spécifications de la première colonne (référence : 35003-101030) sont les suivantes : longueur 100 mm, diamètre 1 mm, granulométrie 3 µm. Le volume injecté est de 10 µL, le débit à 0,1 mL/min et la température à 50 °C. Le gradient est le suivant (avec A= 100% H₂O, B=100% ACN, C= 10 µM LiCl dans H₂O) :

0-8 minutes : 90% A, 0% B, 10% C. 8-31 minutes : gradient de 0 à 75% B, 10% C. 31-42 minutes : 10% A, 80% B, 10% C. 42-47 : 100% A.

La seconde colonne a un volume beaucoup plus important (150 mmx4.6 mm, 5 µm, référence : 35005-154630). Le volume injecté est de 10 µL. Le débit a été augmenté à 0,4 mL/min et pour qu'il soit compatible avec la source électrospray, un T a été ajouté en sortie de chromatographie (donc débit entrant dans la MS : environ 0,2 mL/min). La température est réglée à 40°C et le gradient est le suivant (avec A= 100% H₂O, B=100% ACN, C= 10 µM LiCl dans H₂O):

0-16 minutes : 92% A, 5% B, 3% C. 0-26 minutes : gradient de 5% à 75% B, 3% C. 26-29 minutes : 75% B et 3% C. 29-35 minutes : rinçage avec 92% A, 5% B, 3% C.

Conditions MS : l'ion m/z 349 est isolé et fragmenté par CID pendant 30 ms, le fragment m/z 169 est réisolé pendant 150 ms.

Conditions laser : trois conditions laser ont été testées avec un temps d'irradiation identique de 130 ms. La première consiste à utiliser le laser CO₂ seul à la puissance optimale (déterminée par les courbes de puissance), la seconde emploie le Firefly seul à 3650 cm⁻¹ et la troisième repose sur une combinaison des deux lasers (Firefly émettant à 3410 cm⁻¹ et le laser CO₂ à sa puissance optimale).

Traitement de données : le chromatogramme de l'ion m/z 169 est normalisé selon la valeur maximale de la moyenne glissante sur 10 points. Le tracé noir représente la moyenne glissante normalisée sur 10 points. La somme des fragments dans l'intervalle m/z [50 ;165] est calculée et divisée par le ratio de fragmentation galactofuranose du jour (obtenue en infusion directe). L'intensité des fragments est représentée par une aire colorée uniquement pour les scans où l'intensité de l'ion m/z 169 normalisée est supérieure à 5. Le ratio de l'intensité des fragments sur l'intensité totale est calculé pour chaque scan où l'intensité normalisée de l'ion parent est supérieure à 5. Une moyenne est réalisée pour chaque pic chromatographique.

3 Séquençage d'oligosaccharides de synthèse contenant du galactofuranose

3.1 Développement de la méthode de séquençage sur des trisaccharides synthétiques

Quatre trisaccharides Gal-Man-Man ont été synthétisés par nos collaborateurs de l'ENSCR. Leurs structures sont représentées dans le Schéma 2. Ils diffèrent par la taille du cycle du galactose et la régiochimie des deux liaisons glycosidiques. Nous testerons la méthode de séquençage sur ces trisaccharides.

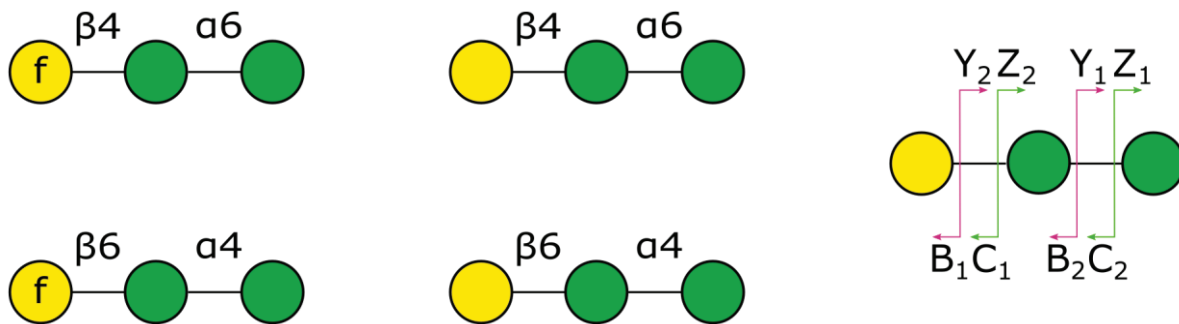


Schéma 2 Structure des quatre trisaccharides $\text{Gal}\beta(1-4)\text{Man}\alpha(1-6)\text{Man}$, $\text{Gal}\beta(1-6)\text{Man}\alpha(1-4)\text{Man}$, $\text{Gal}\beta(1-4)\text{Man}\alpha(1-6)\text{Man}$ et $\text{Gal}\beta(1-6)\text{Man}\alpha(1-4)\text{Man}$.

Nous remarquerons que les disaccharides Gal-Man sont symétriques, c'est-à-dire que le fragment b_1 (qui correspond à l'unité d'intérêt galactose déshydraté) et le fragment z_1 (correspondant au mannose déshydraté) ont une masse identique de 162 Da. Il en est de même pour les fragments c_1 (galactose) et y_1 (mannose) qui possèdent une masse commune de 180 Da. Les différents fragments sont résumés dans la première partie du Tableau 1.

Les trisaccharides présentés ci-dessus sont également symétriques (Tableau 1, partie 3). Les fragments b_1 (galactose déshydraté) et z_1 (mannose déshydraté) ont la même masse (162 Da). Les fragments c_1 (galactose) et y_1 (mannose) ont quant à eux une masse identique de 180 Da. Dans le cas des trisaccharides, on peut également obtenir des fragments disaccharidiques pour lesquels les masses se confondent également. Ainsi, la masse 324 Da peut aussi bien correspondre au fragment b_2 (Gal-Man déshydraté, comportant donc le monosaccharide d'intérêt galactose) qu'au fragment z_2 (Man-Man déshydraté). Quant à la masse 342 Da, il peut s'agir du fragment c_2 (Gal-Man) ou y_2 (Man-Man).

Cette symétrie est problématique pour l'application du séquençage. En effet, en sélectionnant une des masses données ci-dessus, les deux fragments seront isolés conjointement dans le piège à ions. Le spectre IRMPD correspondra alors à un mélange des deux signatures IR et l'identification sera difficile.

Pour pallier ce phénomène, nous avons adopté deux stratégies facilitant la distinction des fragments en extrémité réductrice ou non réductrice. Dans le cas des disaccharides, leur synthèse passe par un intermédiaire possédant un groupement thiophényle ($M=92$ g/mol) en extrémité terminale. Nous avons récupéré l'intermédiaire de synthèse de chaque disaccharide pour notre étude sur le séquençage. Leurs structures sont représentées dans le Schéma 3 et leurs fragments sont indiquées dans la seconde partie du Tableau 1. Ainsi les fragments en extrémité non réductrice contenant le galactose conservent les mêmes masses que ceux des disaccharides Gal-Man (162 Da pour b_1 et 180 Da pour c_1).

3. Séquençage d'oligosaccharides de synthèse contenant du galactofuranose

En revanche, comme le mannose est greffé d'un groupement SPh, 92 Da s'ajoutent aux fragments en extrémité réductrice (254 Da pour z_1 et 272 Da pour y_1). Nous pouvons donc clairement différencier les fragments b_1 et z_1 , ainsi que les fragments c_1 et y_1 .

La stratégie appliquée sur les trisaccharides est le marquage à ^{18}O . Il est reconnu que le marquage se fait sur l'OH en position anomère de l'extrémité réductrice.¹⁰ Les fragments z_1 et y_1 (respectivement 164 et 182 Da) correspondent ainsi au mannose marqué à ^{18}O et ont donc 2 Da de plus que les fragments non terminaux b_1 et c_1 qui contiennent le galactose (voir Tableau 1 partie 4). En ce qui concerne les fragments disaccharidiques, les fragments z_2 et y_2 (326 et 344 Da) possèdent deux unités de mannose dont celui en extrémité terminale est marqué (Man-Man ^{18}O) et ont donc également 2 Da de plus que les fragments b_2 et c_2 qui correspondent au disaccharide galactose-mannose non marqué (Gal-Man). En ajoutant 2 Da au mannose terminal, l'oxygène 18 rend donc possible la différenciation des fragments en extrémité réductrice et non réductrice.

Tableau 1 Fragments des Gal-Man, Gal-ManSPh, Gal-Man-Man et Gal-Man-Man ^{18}O issus de la rupture de la liaison glycosidique selon la nomenclature de Domon et Costello. Pour chaque fragment, la masse et la composition sont indiquées.

1. Fragments des disaccharides Gal-Man		
Type de fragment	Masse (en Da)	Composition
b_1	162	Gal - H ₂ O
c_1	180	Gal
z_1	162	Man - H ₂ O
y_1	180	Man
2. Fragments des disaccharides avec groupement thiophényle Gal-ManSPh		
Type de fragment	Masse (en Da)	Composition
b_1	162	Gal - H ₂ O
c_1	180	Gal
z_1	254	ManSPh - H ₂ O
y_1	272	ManSPh
3. Fragments des trisaccharides Gal-Man-Man		
Type de fragment	Masse (en Da)	Composition
b_1	162	Gal - H ₂ O
c_1	180	Gal
z_1	162	Man - H ₂ O
y_1	180	Man
b_2	324	Gal-Man - H ₂ O
c_2	342	Gal-Man
z_2	324	Man-Man - H ₂ O
y_2	342	Man-Man
4. Fragments des trisaccharides Gal-Man-Man marqués à ^{18}O		
Type de fragment	Masse (en Da)	Composition
b_1	162	Gal - H ₂ O
c_1	180	Gal
z_1	164	Man ^{18}O - H ₂ O
y_1	182	Man ^{18}O
b_2	324	Gal-Man - H ₂ O
c_2	342	Gal-Man
z_2	326	Man-Man ^{18}O - H ₂ O
y_2	344	Man-Man ^{18}O

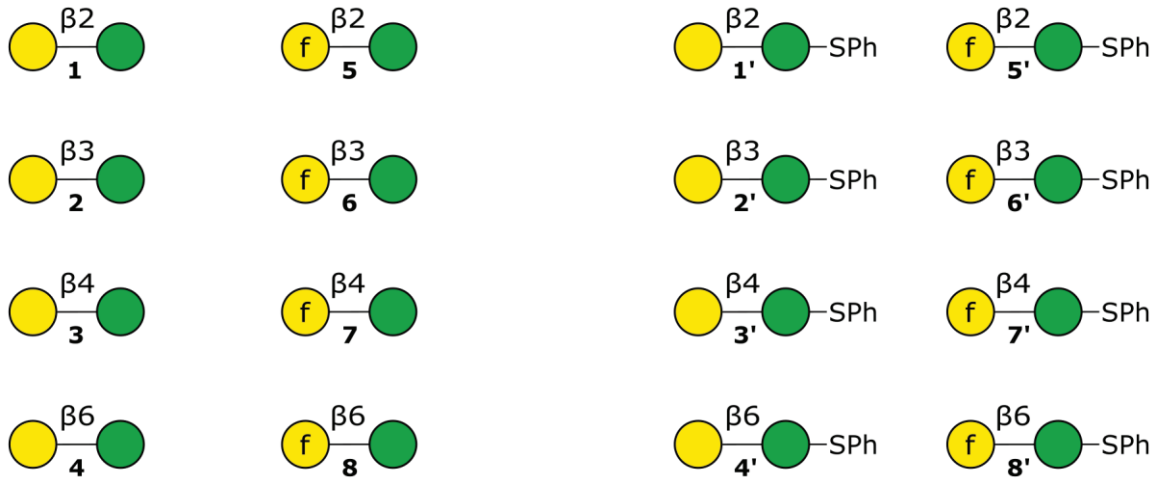


Schéma 3 Structure des huit disaccharides Gal-Man (numérotés de 1 à 8) et des huit disaccharides Gal-Man avec groupement thiophényle en extrémité réductrice (numérotés de 1' à 8').

Nous avons sélectionné les standards à analyser, il nous reste maintenant à choisir un adduit pour les observer en spectrométrie de masse. Cet adduit doit permettre d'obtenir des fragments MS/MS d'intérêt pour le séquençage. Par ailleurs, les spectres IRMPD de ces fragments complexés à l'adduit choisi doivent être informatifs. Nous avons vu dans le chapitre 2 que les spectres de fragmentation avec le Li^+ étaient riches en fragments, notre choix s'est donc naturellement porté sur cet adduit. Comme $z=1$, le rapport masse sur charge d'un fragment complexé au lithium est égal à la masse du fragment (donné dans le tableau précédent) à laquelle s'ajoutent les 7 Da du Li^+ .

Les spectres MS/MS des Gal-ManSPh complexés au lithium, détectés à m/z 441, sont représentés en Figure 1.

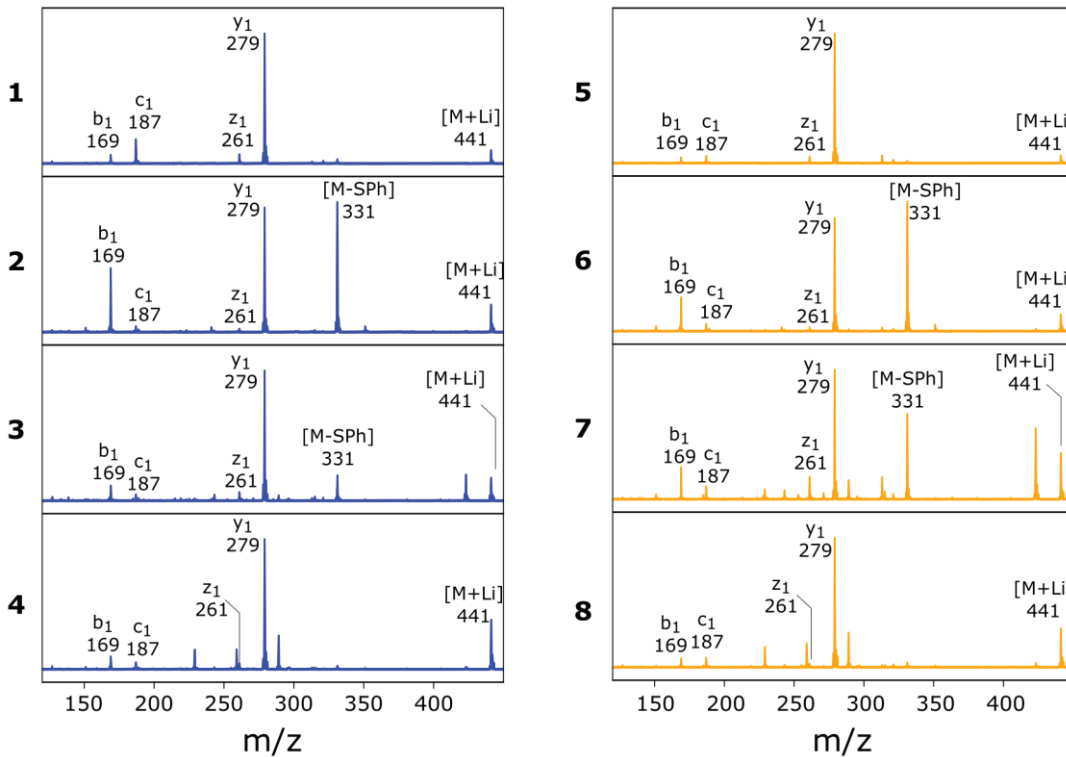


Figure 1 Spectres MS/MS des huit Gal-ManSPh avec adduit Li^+ .

3. Séquençage d'oligosaccharides de synthèse contenant du galactofuranose

Les huit spectres montrent les fragments d'intérêt b_1 et c_1 , respectivement à m/z 169 et m/z 187, qui se distinguent nettement des fragments en extrémité réductrice avec une différence de 92 Da (z_1 à m/z 261 et y_1 à m/z 279). Enfin, nous pouvons noter la présence d'un fragment à m/z 331 qui est l'ion avec perte neutre du thiophényle.

Les trisaccharides Gal-Man-Man complexés au lithium sont détectés à m/z 511. Lorsqu'ils sont marqués au ^{18}O , leur rapport m/z est alors de 513 Da. Comme l' H_2^{18}O est marqué à 97%, il reste une faible proportion de trisaccharide non marqué (à m/z 511). La fenêtre d'isolement a été sélectionnée de manière à isoler uniquement l'ion m/z 513. Les spectres de fragmentation des trisaccharides marqués à l' ^{18}O sont affichés en Figure 2.

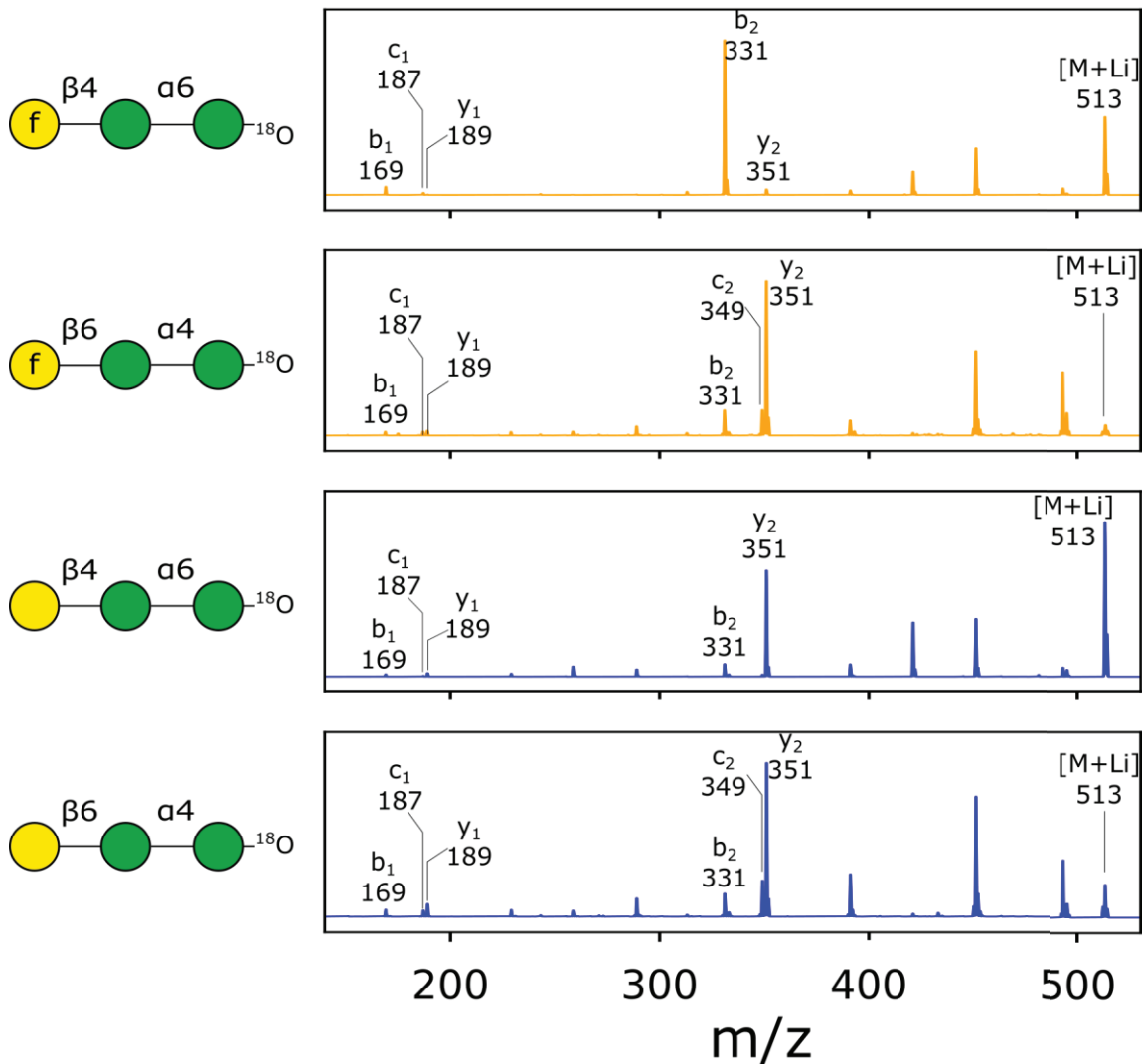


Figure 2 Spectre MS/MS des quatre trisaccharides marqués à l' ^{18}O complexés au Li^+ .

On y distingue le fragment c_1 (galactose, détecté à m/z 187) du fragment y_1 (mannose marqué, détecté à m/z 189). Les fragments z_1 ne sont pas observés mais ont un rapport m/z théorique de 171 Da, ce qui les distingue des fragments b_1 observés à m/z 169. Il en est de même pour les fragments disaccharidiques : les fragments c_2 (Gal-Man visible à m/z 349) et y_2 (Man-Man ^{18}O détecté à m/z 351) ont un écart de 2 Da, ainsi que les fragments b_2 (Gal-Man déshydraté observé à m/z 331) et les fragments z_2 (Man-Man ^{18}O déshydraté avec un rapport m/z théorique de 333 Da).

Pour les expériences de séquençage, on sait que l'analyse des fragments monosaccharidiques est efficace pour identifier le contenu en monosaccharides. L'étude sur les fragments disaccharidiques fournit également l'information sur la stéréochimie (configuration α ou β) et la régiochimie (position) de la liaison glycosidique.² La stéréochimie de la liaison glycosidique entre le galactose et le mannose est de configuration β pour tous nos composés, nous nous intéressons uniquement à la régiochimie.

Les expériences de séquençage menées à ce jour par l'équipe ne concernent que des hexoses incluant le cycle pyranose. Nous ne savons donc pas si les fragments mono- et disaccharidiques portent l'information sur la taille du cycle. Cette question se pose surtout dans le cas des fragments b pour lesquels il y a une restructuration complète de la structure. Nous mènerons donc la première étude de séquençage par IRMPD sur des systèmes modèles contenant du galactofuranose. Nous disposons de plusieurs molécules que nous fragmenterons par CID pour analyser les fragments d'intérêt (c'est-à-dire ceux possédant l'unité galactose). Ces molécules et leurs fragments sont résumés dans le Tableau 2.

Tableau 2 Liste des fragments d'intérêt des Gal-ManSPh et Gal-Man-Man¹⁸O, leur rapport m/z et leur composition.

Parent	Fragment d'intérêt	m/z (avec adduit Li⁺)	Composition
Gal-ManSPh	b ₁	169	Gal-18
Gal-ManSPh	c ₁	187	Gal
Gal-ManSPh	M-SPh	331	Gal-Man-18
Gal-Man-Man ¹⁸ O	b ₁	169	Gal-18
Gal-Man-Man ¹⁸ O	c ₁	187	Gal
Gal-Man-Man ¹⁸ O	b ₂	331	Gal-Man-18
Gal-Man-Man ¹⁸ O	c ₂	349	Gal-Man

Nous possédons les disaccharides Gal β (1-X)Man avec toutes les combinaisons de régiochimie et de taille de cycle possibles (schéma page 96). Leurs spectres IRMPD ont été collectés et présentés dans le chapitre 2 (page 46). Ils serviront de références pour l'analyse des fragments c₂. Les références pour l'étude des autres fragments seront acquises et présentées dans les parties suivantes.

La première partie porte sur les fragments disaccharides (b₂ et c₂) des trisaccharides. Les signatures IRMPD des fragments b₂ seront comparées aux fragments m/z 331 des Gal-ManSPh. Nous avons vu que certains trisaccharides donnent également des fragments c₂ : les signatures IRMPD de ces derniers seront obtenues et comparées aux spectres des disaccharides Gal-Man. Dans une seconde partie, nous allons comparer les fragments b₁ et c₁ des Gal-Man-Man¹⁸O avec ceux des Gal-ManSPh. L'objectif est de vérifier que l'information sur la taille de cycle présent dans le précurseur est conservée après le réarrangement moléculaire subi lors de la fragmentation de la liaison glycosidique entre le galactose et le mannose.

3.2 Caractérisation complète d'un trisaccharide par analyse des fragments disaccharidiques

Pour séquencer un trisaccharide, nous aurons d'abord besoin de compléter notre base de données. Ensuite, nous allons acquérir les spectres des fragments issus du trisaccharide à séquencer. Enfin, nous comparerons ces spectres avec ceux des références. Un système de comparaison à scores sera instauré par la suite pour faciliter la comparaison.

3.2.1 Acquisition des références pour la construction de la base de données

Nous avons déjà implémenté dans notre base de données les spectres IRMPD des références Gal-Man. Ces spectres, décrits dans le chapitre 2 (page 47), sont repris à gauche dans la Figure 3. Ils serviront à identifier les fragments c_2 des trisaccharides.

Pour les références b_2 , la synthèse de Gal-Man déshydratés n'est pas possible, nous les avons formés par fragmentation CID. Pour cela, les disaccharides Gal-ManSPH sont fragmentés et le fragment qui perd le groupement thiophényle (m/z 331) est isolé. Les spectres des fragments m/z 331 des huit Gal-ManSPH ont été acquis et sont tracés à droite de la Figure 3.

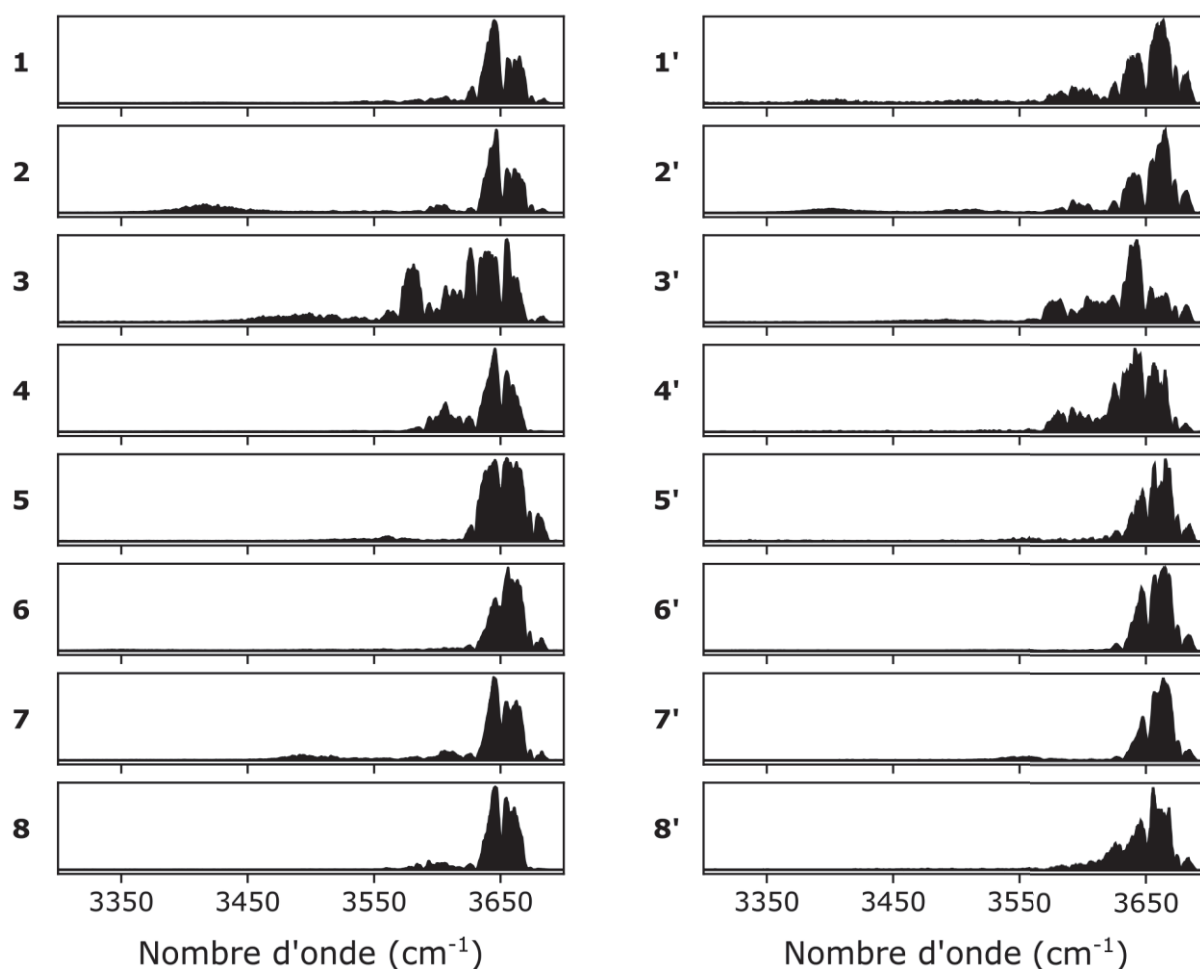


Figure 3 Spectres IRMPD des références Gal-Man numérotées de 1 à 8 à gauche, spectres IRMPD du fragment m/z 331 des références 1' à 8' à droite (adduit Li^+).

Les quatre premiers spectres correspondent à ceux possédant le cycle pyranose. Une activité vibrationnelle est observée entre 3570 et 3700 cm^{-1} . Les spectres IRMPD des composés 1' et 2' ont leur

maximum d'intensité à 3660 cm^{-1} et montrent deux faibles bandes à 3400 et 3500 cm^{-1} . Ils sont très proches entre eux et ne peuvent être distingués que par l'intensité relative des différentes bandes. Les deux spectres suivants (**3'** et **4'**) sont assez proches avec un maximum d'absorption à 3640 cm^{-1} mais l'intensité relative de cette bande à celle de 3660 cm^{-1} permet de distinguer ces deux composés (la bande à 3660 cm^{-1} étant plus intense pour **4'**). Ainsi, les spectres des régioisomères Gal β -ManSPH sont distincts les uns des autres.

Les quatre spectres suivants (**5'** à **8'**) sont ceux des isomères Gal f -ManSPH. On observe qu'ils sont assez proches avec une activité vibrationnelle intense entre 3620 et 3690 cm^{-1} . Nous pouvons néanmoins les distinguer par leurs caractéristiques particulières. Le spectre de **8'** est le seul à présenter de l'activité vibrationnelle entre 3580 et 3620 cm^{-1} . Pour **7'**, une faible bande est visible à 3550 cm^{-1} . Les spectres de **5'** et **6'** sont les plus similaires. Seule une différence d'intensité de la bande à 3640 cm^{-1} , plus élevée pour **6'**, peut les distinguer.

Ainsi, les signatures des fragments b_2 des Gal f -Man et des Gal β -Man sont très différents. La détermination de la taille du cycle est assez facile. La résolution de la régiochimie est également possible et nécessite une analyse plus approfondie des spectres.

Les spectres IRMPD des fragments m/z 331 ont été ajoutés dans notre base de données qui contenait déjà les spectres IRMPD des disaccharides (m/z 349). Les premiers vont servir à identifier les fragments b_2 (m/z 331) et les seconds à analyser les fragments c_2 (m/z 349) des trisaccharides.

3.2.2 Acquisition des spectres IRMPD d'intérêt

Nous avons notre base de données complète, nous allons passer à l'acquisition et l'analyse des spectres sondés. Deux fragments disaccharidiques contiennent l'unité galactose : les fragments c_2 (détectés à m/z 349) et les fragments b_2 (m/z 331). Les spectres de fragmentation des trisaccharides Gal-Man-Man¹⁸O montrent la présence de ces deux fragments (page 97). Nous avons acquis leurs spectres IRMPD qui sont représentés en Figure 4.

Pour le trisaccharide Gal $f\beta(1-6)Man\alpha(1-4)Man$, le spectre IRMPD du fragment c_2 montre une forte bande d'absorption de 3630 à 3670 cm^{-1} ainsi qu'une faible bande de 3550 à 3630 cm^{-1} . Le spectre de son fragment b_2 a également une bande intense de 3630 à 3670 cm^{-1} , une petite bande autour de 3680 cm^{-1} et une bande moyenne de 3590 à 3630 cm^{-1} .

Passons au trisaccharide Gal $\beta(1-6)Man\alpha(1-4)Man$, qui se distingue du premier par la taille du cycle du galactose. Le spectre IRMPD de son fragment c_2 montre une absorption intense allant de 3630 à 3670 cm^{-1} et une bande plus faible de 3590 à 3630 cm^{-1} avec un maximum d'absorption à 3610 cm^{-1} . Le spectre IRMPD de son fragment b_2 exprime quant à lui deux bandes intenses à 3660 et 3640 cm^{-1} , une bande d'intensité modérée à 3625 cm^{-1} et une bande faible et large de 3570 à 3615 cm^{-1} .

La fragmentation CID des deux trisaccharides suivants donne des fragments b_2 mais pas de fragments c_2 (page 97). Seuls les spectres IRMPD des fragments b_2 ont donc été acquis. Pour le Gal $f\beta(1-4)Man\alpha(1-6)Man$, le spectre IRMPD du fragment b_2 est très simple avec une bande intense de 3630 à 3675 cm^{-1} et une bande mineure à 3680 cm^{-1} . Il est également composé d'une très faible bande vers 3540 cm^{-1} . Pour le trisaccharide Gal $\beta(1-4)Man\alpha(1-6)Man$, le fragment b_2 montre un spectre IRMPD beaucoup plus complexe avec une absorption faible qui va de 3440 à 3550 cm^{-1} puis une large bande de 3550 à 3690 cm^{-1} . Cette seconde région est composée de plusieurs bandes dont deux très intenses à 3635 et 3645 cm^{-1} .

3. Séquençage d'oligosaccharides de synthèse contenant du galactofuranose

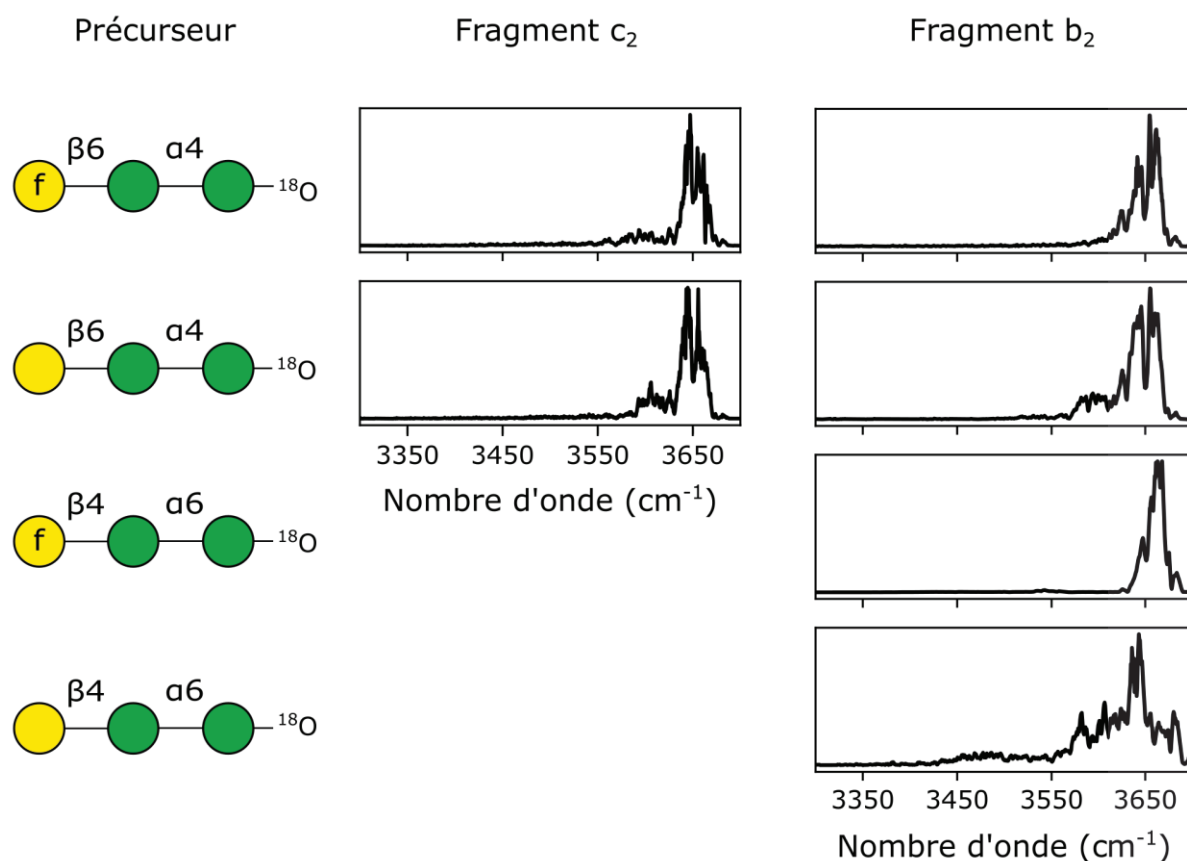


Figure 4 Spectres IRMPD des fragments c_2 et b_2 des trisaccharides marqués à ^{18}O .

Nous allons à présent comparer ces empreintes IR avec les références de notre base de données.

3.2.3 Comparaison visuelle avec la base de données

Commençons par l'analyse des fragments du trisaccharide $Gal\beta(1-6)Man\alpha(1-4)Man$. Les spectres IRMPD de ses fragments sont représentés par une ligne continue noire en Figure 5. Le spectre du fragment c_2 apparaît dans la colonne de gauche et celui du fragment b_2 dans la colonne de droite (un même spectre est répété huit fois pour la comparaison avec la base de données). Les spectres de référence ont été retracés avec une aire colorée. La couleur est verte s'il s'agit de la séquence attendue, ici $Gal\beta(1-6)Man$, et en rouge sinon.

En comparant le spectre du fragment c_2 avec la base de données, peuvent être éliminés dans un premier temps les spectres de référence qui ont des bandes à des positions où n'absorbe pas le fragment analysé. Par exemple, le spectre de **2** a une bande à 3410 cm^{-1} que ne possède pas le spectre d'intérêt, de même que le spectre de **3** avec sa bande de 3460 à 3530 cm^{-1} et que le spectre de **7** avec sa bande à 3500 cm^{-1} . Ensuite, on a le cas inverse où le fragment d'intérêt absorbe mais pas le fragment de référence. C'est le cas des spectres de **5** et **6** qui n'ont pas de bande vers 3600 cm^{-1} . La comparaison des positions permet donc d'éliminer cinq références. Dans un second temps, on compare l'intensité des bandes pour les références restantes. Le spectre de **1** a une plus faible intensité à 3650 cm^{-1} que le spectre d'intérêt, le spectre de **4** présente au contraire une bande plus forte vers 3600 cm^{-1} . Par élimination il ne reste plus que le spectre du composé **8** pour lequel une parfaite concordance est constatée. Cette comparaison est suffisante pour identifier la séquence d'intérêt mais nous avons également mené l'analyse sur le fragment b_2 .

Le spectre IRMPD du fragment b_2 est comparé à ceux des références **1'** à **8'**. Les spectres de **1'** et **2'** ont une faible absorption à 3400 cm^{-1} que ne possède pas le spectre d'intérêt. A l'inverse, les spectres

des composés **5'** à **7'** ne montrent pas d'absorption de 3580 à 3610 cm^{-1} . Ces cinq spectres ne correspondent donc pas au spectre d'intérêt. Les spectres de **3'** et **4'** révèlent bien de l'absorption entre 3580 et 3610 cm^{-1} mais elle est beaucoup plus intense que celle du spectre à analyser. Il ne reste donc plus que le spectre de **8'** pour lequel on observe encore une fois une parfaite correspondance.

La même séquence (**8** et **8'**, qui correspondent au $\text{Gal}\beta(1-6)\text{Man}$) est obtenue avec l'analyse des deux fragments, ce qui permet une double vérification. Notons qu'il s'agit d'un trisaccharide avec du galactofuranose, les fragments disaccharidiques conservent donc bien l'information sur la taille du cycle du galactose après fragmentation CID. Menons l'analyse sur le trisaccharide composé du galactopyranose.

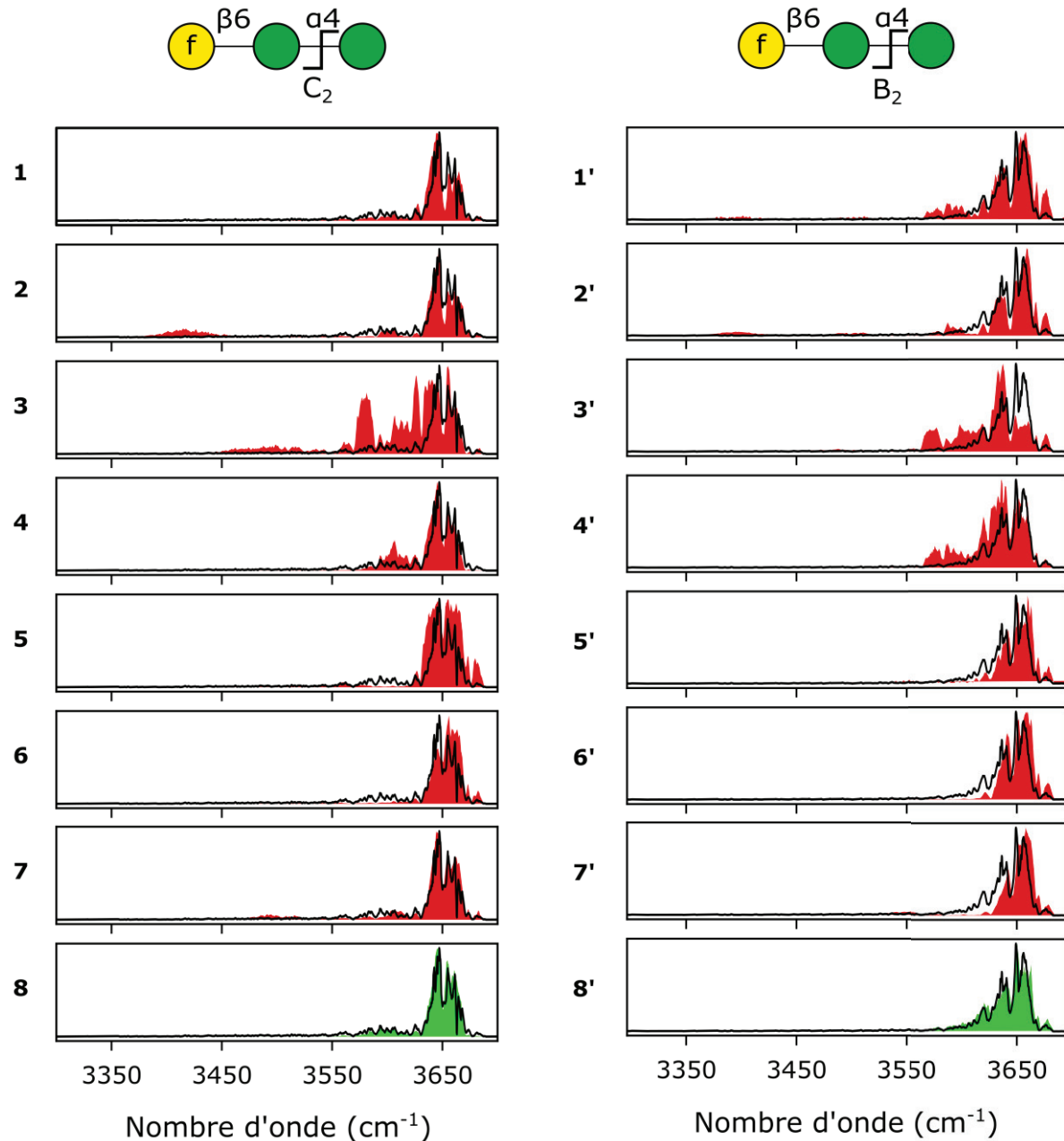


Figure 5 Comparaison visuelle des spectres IRMPD des fragments issus du trisaccharide $\text{Gal}\beta(1-6)\text{Man}\alpha(1-4)\text{Man}$ avec la base de données. Les spectres IRMPD des fragments du trisaccharide analysé sont tracés en ligne noire continue (fragment c_2 à gauche et b_2 à droite). Les spectres de référence (numérotés de 1 à 8 pour les fragments c_2 et de 1' à 8' pour les fragments b_2) sont tracés avec une aire colorée en vert s'il s'agit de la séquence attendue et en rouge le cas échéant.

3. Séquençage d'oligosaccharides de synthèse contenant du galactofuranose

Les fragments c_2 et b_2 du trisaccharide $\text{Gal}\beta(1-6)\text{Man}\alpha(1-4)\text{Man}$ sont analysés en Figure 6. En comparant les données sur la colonne de gauche (fragment c_2), nous pouvons exclure le composé **2** qui absorbe à 3415 cm^{-1} et les composés **3** et **7** qui montrent une bande infrarouge vers 3490 cm^{-1} . Les composés **5** et **6** n'absorbent pas à 3600 cm^{-1} et les spectres des standards **1** et **8** présentent une bande à cette longueur d'onde mais beaucoup plus faible que le spectre IRMPD du fragment analysé. On note à l'inverse une parfaite juxtaposition du spectre sondé avec le spectre du standard **4**.

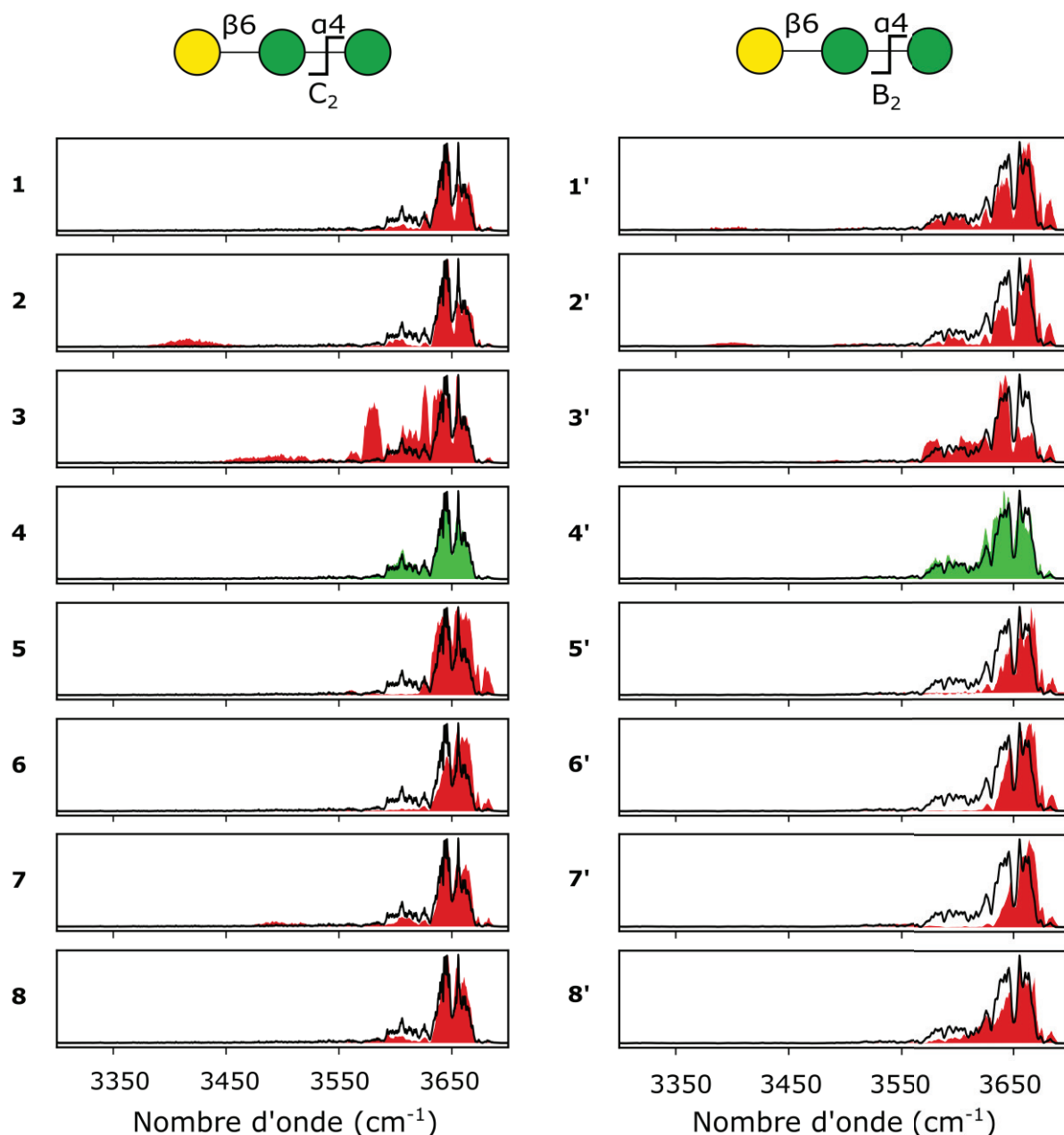


Figure 6 Comparaison visuelle des spectres IRMPD des fragments issus du trisaccharide $\text{Gal}\beta(1-6)\text{Man}\alpha(1-4)\text{Man}$ avec la base de données. Les spectres IRMPD des fragments du trisaccharide analysé sont tracés en ligne noire continue (fragment c_2 à gauche et b_2 à droite). Les spectres de référence (numérotés de 1 à 8 pour les fragments c_2 et de 1' à 8' pour les fragments b_2) sont tracés avec une aire colorée en vert s'il s'agit de la séquence attendue et en rouge le cas échéant.

Les données sur le fragment b_2 figurent à droite de la Figure 6. Les références **1'** et **2'** peuvent être exclues car leurs spectres indiquent de l'absorption à 3400 cm^{-1} . La présence d'une bande assez intense de 3575 à 3600 cm^{-1} permet d'éliminer les références **5'** à **8'**. Le spectre de la référence **3'** a bien une bande d'absorption intense dans cette gamme, en revanche sa bande à 3660 cm^{-1} est beaucoup plus

faible que celle du spectre sondé. Il ne reste donc plus que le composé **4'** dont le spectre IRMPD se superpose parfaitement avec le spectre du fragment b_2 . Cela correspond bien à la séquence non réductrice du trisaccharide analysé.

Rappelons qu'on a acquis uniquement les spectres des fragments b_2 pour les deux trisaccharides Gal β (1-4)Man α (1-6)Man. On va donc effectuer une seule comparaison afin de déterminer la séquence de la partie non terminale. La Figure 7 permet la comparaison des fragments b_2 , du Gal β (1-4)Man α (1-6)Man à gauche et du Gal β (1-4)Man α (1-6)Man à droite, avec la base de données comportant les références **1'** à **8'**.

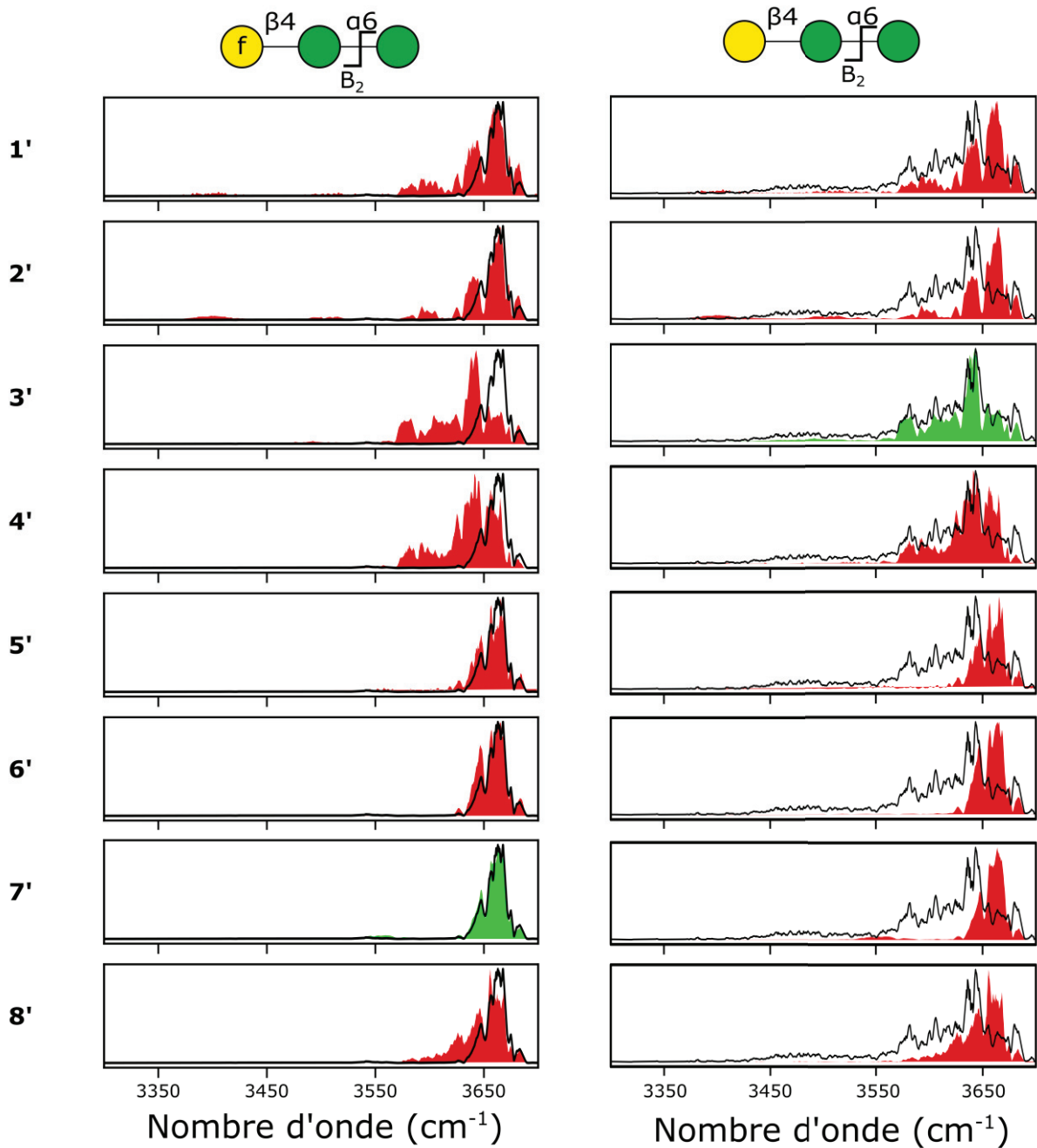


Figure 7 Comparaison du spectre IRMPD du fragment b_2 (Gal β (1-4)Man α (1-6)Man à gauche et Gal β (1-4)Man α (1-6)Man à droite) représenté par une ligne continue avec les spectres IRMPD des références numérotées de **1'** à **8'** et avec l'aire colorée (en vert s'il s'agit de la séquence correspondante, en rouge sinon).

3. Séquençage d'oligosaccharides de synthèse contenant du galactofuranose

Pour le trisaccharide avec le galactofuranose (à gauche), le spectre IRMPD de son fragment b_2 présente peu de bande IR : un motif de OH entre 3630 et 3680 cm^{-1} et une très faible bande à 3550 cm^{-1} . Tous les spectres de référence présentant de l'absorption entre 3560 et 3630 cm^{-1} peuvent être éliminés : les quatre premières références contenant du galactopyranose **1'** à **4'** ainsi que **8'**. Les références **5'** et **6'** ont des bandes dont l'intensité ne correspond pas au spectre sondé (celle à 3660 cm^{-1} pour **5'** et celle à 3645 cm^{-1} pour **6'**). Il ne reste plus que le composé **7'** dont le spectre montre un motif de 3630 à 3680 cm^{-1} et une faible bande à 3550 cm^{-1} , tout comme le spectre sondé.

Le spectre IRMPD du fragment b_2 pour le trisaccharide homologue contenant du galactopyranose (à droite) est à l'inverse très riche en bandes avec un domaine d'absorption élargi jusqu'à 3430 cm^{-1} . On peut d'office éliminer les quatre dernières références possédant du galactofuranose **5'** à **8'** qui n'ont pas/peu d'absorption entre 3570 et 3610 cm^{-1} . Il ne reste donc plus que les régioisomères contenant du galactopyranose (**1'** à **4'**). Les motifs spectraux de **1'**, **2'** et **4'** ne correspondent pas, par exemple la région entre 3570 et 3610 cm^{-1} n'est pas suffisamment intense (**1'** et **2'**) ou au contraire la bande à 3650 est trop intense (**4'**). À l'inverse le spectre de **3'** se superpose bien au spectre sondé.

Une seule comparaison a été effectuée pour ces deux trisaccharides qui n'ont pas de fragment c_2 . En premier abord, les références qui ne possèdent pas le cycle du galactose attendu sont éliminées. L'analyse des fragments disaccharidiques permet donc de déterminer rapidement et efficacement la taille du cycle. La comparaison est plus approfondie pour les régioisomères et conduit à la régiochimie correcte.

En somme, la séquence attendue du disaccharide non terminal est retrouvée par analyse des fragments disaccharidiques. En revanche, cette comparaison visuelle est qualitative et nous avons besoin d'outils pour une analyse quantitative des données. Nous avons instauré un système de calculs de scores pour cela.

3.2.4 Calcul de scores pour l'identification de fragments

Afin de quantifier la ressemblance entre deux spectres, un calcul de score est proposé. Les calculs sont effectués sur les moyennes glissantes des valeurs de photofragmentation obtenues dans la gamme 3300-3700 cm^{-1} . Les deux jeux de données sont d'abord normalisés par rapport à la valeur maximale. Les spectres de référence sont composés d'environ 1270 points et ceux des trisaccharides en ont environ 1030. Cet échantillonnage varie en fonction des paramètres MS et justifie une interpolation. Les valeurs normalisées ont donc été interpolées (I_1 et I_2). La somme des carrés des écarts entre les deux spectres interpolés est ensuite calculée selon la formule :

$$MSE = \frac{100}{3} \sum (I_1 - I_2)^2$$

Plus cette valeur est faible, plus les deux spectres sont similaires. Les scores de ressemblance ont été calculés pour les fragments b_2 (et c_2 le cas échéant) pour les quatre trisaccharides. Ils sont représentés par des diagrammes à barres en Figure 8. Comme pour les figures précédentes, la séquence correcte est indiquée en vert. La bonne séquence est celle qui a le plus faible score pour chaque fragment, à une exception près. Le fragment c_2 de Gal β (1-6)Man α (1-4)Man (5^e ligne) a un score légèrement plus faible avec **2** (140) qu'avec **8** (159). Notons que pour ce même trisaccharide, le fragment b_2 a bien le meilleur score avec la séquence correcte (219). Il est étonnant que le score sur le fragment c_2 ait failli car la comparaison visuelle était convaincante. En effet, le spectre de **2** absorbe vers 3420 cm^{-1} alors que le fragment sondé ne présente pas de bande à cette longueur d'onde (page 102). Cette erreur met en évidence l'intérêt de développer différentes approches, comme normaliser sur l'aire au lieu de l'intensité maximale ou encore déployer des outils d'intelligence artificielle.

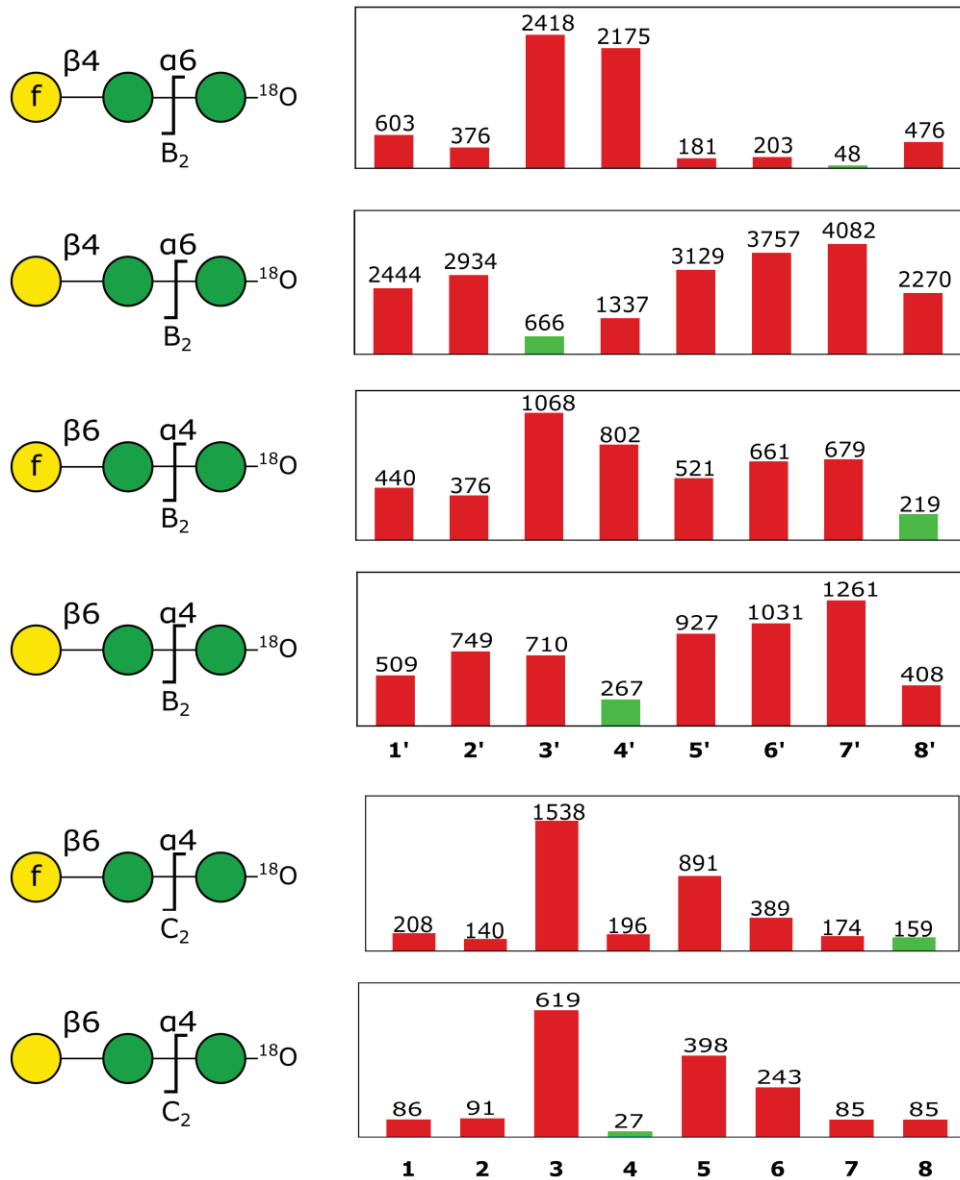


Figure 8 Scores de ressemblance entre les fragments des trisaccharides marqués à l'¹⁸O et les références disaccharidiques. Le score de la séquence correcte est coloré en vert.

À une exception près, le score est donc un bon indice de similarité et rend le choix quantitatif. Nous emploierons cette méthode pour déterminer la séquence complète d'un trisaccharide.

3.2.5 Détermination de la séquence d'un trisaccharide par comparaison des scores

Nous détaillerons ici la procédure par étapes pour le séquençage d'un trisaccharide contenant du galactofuranose. Il faut dans un premier temps marquer le trisaccharide à l'¹⁸O pour lever le verrou de la symétrie. Ce trisaccharide marqué est détecté à m/z 513 lorsqu'il est complexé au lithium (Figure 9). Ensuite, une étape de dissociation induite par collision est réalisée. Plusieurs fragments sont observés dont un à m/z 331 correspondant au fragment b₂ et un autre à m/z 351 qui est le fragment y₂. Ces deux fragments contiennent l'information sur la séquence complète et sont donc suffisants pour retracer la séquence de l'ion parent. L'analyse des empreintes IR de ces deux fragments indique le meilleur score pour le composé **7'** qui est Galβ(1-4)Man en extrémité non réductrice et le composé **18**, c'est-à-dire Man α(1-6)Man, pour l'extrémité réductrice.

3. Séquençage d'oligosaccharides de synthèse contenant du galactofuranose

Ces informations nous permettent de remonter à la séquence de l'ion parent qui est donc Gal β (1-4)Man α (1-6)Man. Cela correspond bien au composé injecté. Notons que la base de données pour les fragments y_2 est incomplète, avec seulement deux références disponibles. Il s'agit là de montrer la procédure et elle reste la même avec une base de données plus complète.

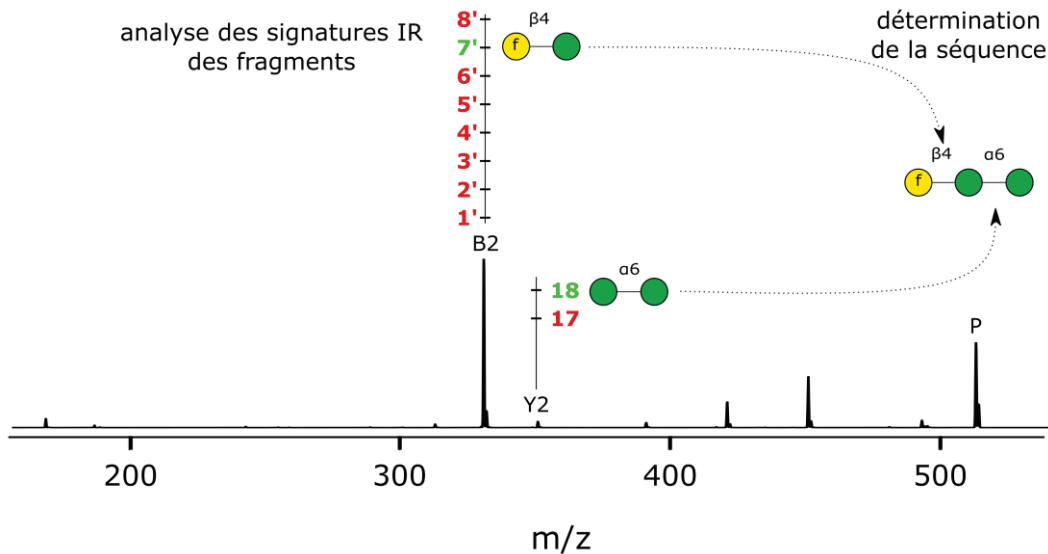


Figure 9 Séquençage du trisaccharide Gal β (1-4)Man α (1-6)Man¹⁸O. En bas, le spectre CID du complexe [M+Li]⁺. Au-dessus des fragments b_2 (m/z 331) et y_2 (m/z 351) : correspondance du spectre IRMPD du fragment sélectionné et de la base de données, le meilleur score est indiqué en vert. Au-dessus de l'ion parent (m/z 513) : structure complète déterminée par l'analyse des ions fragments.

Ainsi, la méthode de séquençage développée par l'équipe² peut être étendue aux oligosaccharides contenant des cycles furanose. L'analyse des fragments disaccharidiques permet en effet de retrouver la séquence de l'ion parent. L'hypothèse de départ est donc validée : les fragments disaccharides obtenus par CID conservent l'information sur la taille de cycle. Nous sommes partis de l'hypothèse d'avoir une séquence Gal-Man-Man. Or, il est possible que le galactose soit présent dans un ordre différent (par exemple Man-Gal-Man, nécessitant ainsi des références Man-Gal) ou encore lié à d'autres sucres (Gal-Glc par exemple). Il n'est pas évident d'obtenir une chimiothèque complète (nous n'avons pas en possession tous les standards Man-Man par exemple). Une alternative est de réduire la taille du fragment analysé, en procédant à l'étude des fragments monosaccharides. Étant donné que la fragmentation par CID se produit immédiatement après le galactose, il existe un risque de perte d'information. Nous verrons dans la partie suivante les résultats obtenus sur les fragments monosaccharides en extrémité non terminale (b_1 et c_1).

3.3 Méthode universelle pour déterminer la taille du cycle par analyse des fragments monosaccharidiques

Comme pour les fragments disaccharidiques, il existe deux types de fragments monosaccharidiques à l'extrémité non réductrice : b_1 et c_1 , obtenus respectivement avant ou après la liaison glycosidique (schéma page 90). Le fragment b_1 (m/z 169) correspond donc au galactose déshydraté et le fragment c_1 (m/z 187) au galactose intact. La fragmentation se produit donc immédiatement après le monosaccharide contenant l'information sur la taille du cycle. Il existe donc un risque de réarrangement moléculaire qui aboutirait à un produit commun pour le galactofuranose et le galactopyranose. Ainsi, les spectres IRMPD étant identiques pour les deux tailles de cycle, il ne serait pas possible de les distinguer par analyse des fragments monosaccharidiques. Vérifions donc si l'information sur la taille du cycle est conservée après fragmentation CID.

Nous avons sélectionné un couple de disaccharides se distinguant uniquement par la taille du cycle du galactose : Gal β (1-2)ManSPh et Gal α (1-2)ManSPh. Les spectres IRMPD de leurs fragments b_1 et c_1 sont représentés dans la Figure 10. Le fragment c_1 du Gal β (1-2)ManSPh (en haut, à gauche) montre une bande plus intense que son isomère Gal α (1-2)ManSPh (en bas, à gauche) à 3600 cm^{-1} et au contraire moins intense vers 3670 cm^{-1} . Les deux spectres sont distincts, ce qui vérifie que la fragmentation après la liaison glycosidique entraîne des fragments différents pour les deux disaccharides.

Les fragments b_1 sont encore plus pertinents : non seulement ils sont distincts mais en plus cette différence consiste en des gammes d'absorption différentes. Alors que le spectre du fragment furanose (en haut, à droite) se distingue par une présence de bande à 3410 cm^{-1} et une absence de bande à 3500 cm^{-1} , c'est l'inverse pour son isomère (en bas, à droite). Ces deux zones d'absorption présentent donc un diagnostic fort de la taille du cycle. Nous nous sommes donc principalement intéressés au fragment b_1 pour la suite de l'étude sur les fragments de taille monosaccharide.

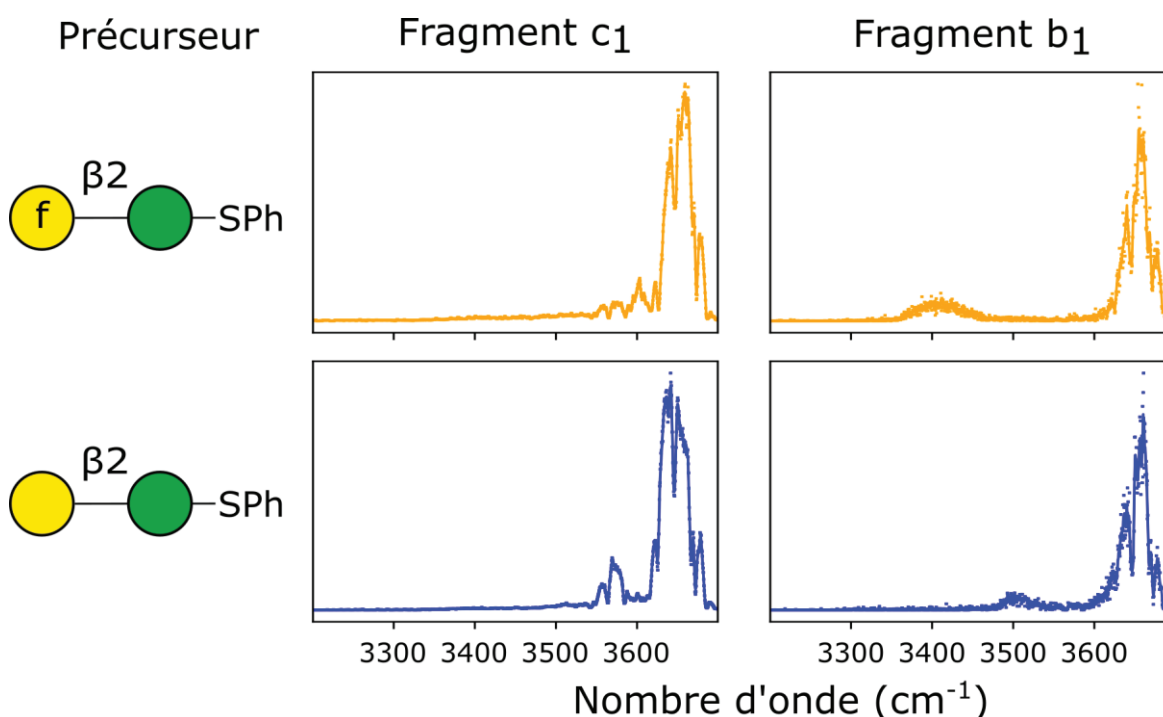


Figure 10 Spectres IRMPD des fragments c_1 (gauche) et b_1 (droite) du disaccharide Gal β (1-2)ManSPh (orange) et de son isomère Gal α (1-2)ManSPh (bleu) complexés au lithium.

Les spectres IRMPD des fragments b_1 ont été acquis pour les huit disaccharides Gal-ManSPh, uniquement dans la gamme d'intérêt [3350 - 3550] cm^{-1} . Ils sont tracés en Figure 11, chaque ligne correspondant à une régioisomérisation différente. Les disaccharides Gal f -ManSPh, représentés à gauche et dont les signatures IRMPD sont colorées en orange, présentent les mêmes caractéristiques que ceux évoqués précédemment : une forte absorption de 3350 à 3475 cm^{-1} (avec un maximum à 3410 cm^{-1}) et pas/très peu d'absorption autour de 3500 cm^{-1} . Les Gal p -ManSPh (à droite, en bleu) valident l'effet inverse : ils absorbent de 3475 à 3550 cm^{-1} et pas autour de 3410 cm^{-1} .

Deux trisaccharides ont également été testés. La molécule étant plus grande, la fragmentation peut se produire au niveau des différentes liaisons et il est donc plus difficile d'obtenir des fragments b_1 . Malgré le faible signal de m/z 169 (correspondant au fragment b_1), il a été isolé et son empreinte IR a été enregistrée dans la gamme d'intérêt (Figure 11, dernière ligne). Comme pour les Gal-ManSPh, le fragment b_1 galactofuranose absorbe dans la région de gauche autour de 3410 cm^{-1} et le fragment b_1 galactopyranose absorbe majoritairement vers 3500 cm^{-1} .

3. Séquençage d'oligosaccharides de synthèse contenant du galactofuranose

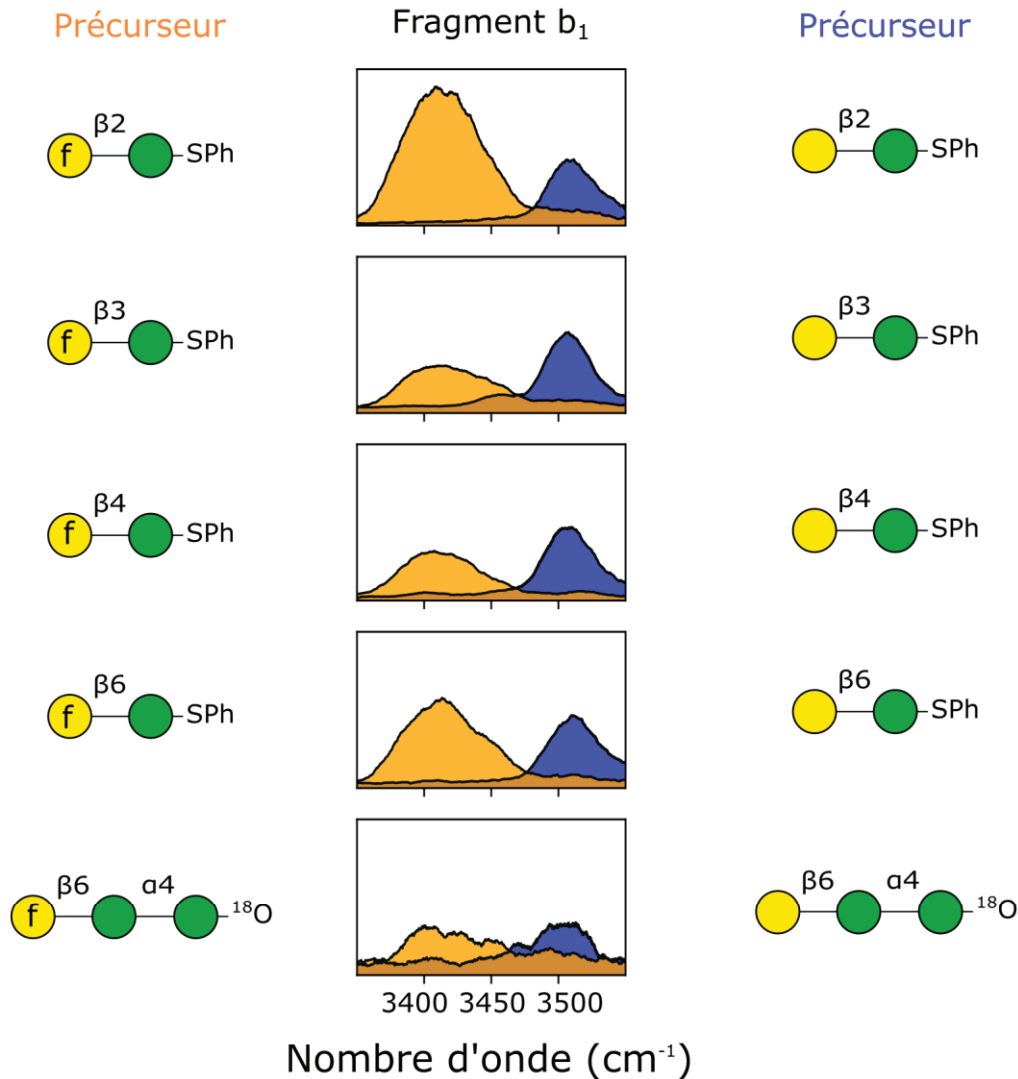


Figure 11 Spectres IRMPD des fragments b_1 des disaccharides Gal-ManSPh (les quatre premières lignes, chacune correspondant à un régioisomère) et des trisaccharides Gal β (1-6)Man α (1-4)Man ^{18}O (dernière ligne). Les spectres des isomères contenant du galactofuranose sont représentés en orange et ceux contenant du galactopyranose en bleu.

La littérature montre que la fragmentation de type b conduit à d'importants réarrangements moléculaires. Nous nous attendions à perdre l'information sur la taille du cycle après ce type de fragmentation. Or notre étude sur les fragments b_1 montre de manière inattendue que l'on maintient, malgré les réarrangements moléculaires, des structures distinctes selon la taille du cycle du précurseur. Ces deux structures peuvent être distinguées par IRMPD. Cela a été vérifié sur des disaccharides avec différentes régiochimies de liaisons et également sur des trisaccharides. Ces résultats sont très avantageux pour la suite du projet puisque cette approche a l'avantage de ne nécessiter que deux références : la signature IR des b_1 galactofuranose et b_1 galactopyranose. Ainsi, cette approche peut être appliquée indépendamment de la régioisométrie et même du monosaccharide suivant.

En outre, les signatures des fragments b_1 s'avèrent intéressantes pour des expérimentations de criblage permettant la détection de composés contenant du galactofuranose. En effet, les composés avec du galactofuranose en extrémité non réductrice peuvent être distingués car la signature IR présente une zone d'absorption spécifique. Nous développerons dans la partie suivante une méthode de criblage permettant une sélection rapide des composés contenant du galactofuranose.

4 Développement d'une méthode de criblage en LC-MS/MS-IR

4.1 Méthode de criblage par irradiation IR

Le criblage (*screening* en anglais) fait référence aux méthodes d'investigation qui permettent de trier des molécules selon un critère donné. Les molécules répondant à ce critère sont discriminées et font l'objet d'études complémentaires.¹¹

Du fait du grand nombre de molécules à analyser, nous nous dirigeons vers des stratégies de criblage qui sont rapides. Le test déployé doit être sensible car il doit exclure les molécules ne répondant pas au critère. La spécificité, c'est-à-dire la capacité de confirmer la caractéristique, est moins critique à ce niveau car d'autres étapes analytiques suivent le criblage. Enfin, le critère doit idéalement être quantitatif pour un gain de temps lors de la prise de décision. Ce critère sert à indiquer rapidement si la molécule passe à l'étape suivante ou est éliminée.

Le criblage permet ainsi de distinguer des molécules ayant des propriétés communes. La photo-SRM est une méthode mise au point par l'équipe SpectroBio de l'ILM.¹² Elle a notamment été utilisée pour discriminer des peptides contenant de la cystéine par couplage LC-MS-UV.¹³ Dans notre cas, nous souhaitons discriminer des oligosaccharides contenant du galactofuranose. Notre étude MS/MS-IR montre que les fragments b₁ des oligosaccharides contenant du galactofuranose partagent des propriétés vibrationnelles communes (photofragmentation à 3410 cm⁻¹) que ne possèdent pas leurs analogues avec du galactopyranose. Par conséquent, nous allons développer un couplage basé sur la LC-MS/MS-IR pour mettre au point notre stratégie de criblage. La LC permet de séparer les composés, une double sélectivité est obtenue par MS/MS puis la spécificité est apportée par l'ajout de l'IR.

4.2 Choix des paramètres LC-MS/MS-IR

4.2.1 Méthode LC

Nous avons utilisé deux colonnes Hypercarb constituées de carbone graphité poreux.¹⁴ La rétention des composés par la colonne repose sur la combinaison de deux mécanismes :

- 1) Des interactions dispersives analyte/phase mobile et analyte/surface graphite qui augmentent la rétention des molécules hydrophobes
- 2) Des interactions charge-dipôle induites par la surface polarisable du graphite

La force des interactions dépend de la surface moléculaire en contact avec le graphite ainsi que des positions des groupes fonctionnels interagissant avec le carbone graphité. Il s'agit ainsi d'une colonne sensible aux isomères.

4.2.2 Méthode MS/MS

Nous avons vu dans la partie séquençage que le fragment b₁ galactofuranose absorbe à une longueur d'onde spécifique à laquelle le fragment b₁ galactopyranose ne réagit pas. Cette caractéristique est très intéressante pour le criblage, notre méthode a donc été développée sur le fragment m/z 169.

Pour le séquençage, deux stratégies ont été déployées dans la partie précédente pour isoler spécifiquement le fragment b₁. La présence de groupement thiophényle est un artefact de synthèse et n'est pas transférable sur des molécules naturelles. Quant au marquage ¹⁸O, il faut attendre plusieurs jours pour qu'il se fasse. Alors que l'isolement spécifique du fragment b₁ est indispensable pour le séquençage, ce n'est pas le cas pour le criblage. En effet, nous souhaitons détecter du galactofuranose, la présence conjointe de mannose n'interfère pas à condition qu'il ne soit pas majoritaire. Or nous

4. Développement d'une méthode de criblage en LC-MS/MS-IR

avons vu dans les spectres MS/MS que le fragment z_1 est minoritaire comparé au fragment b_1 . Ainsi nous effectuerons une fragmentation CID sur les disaccharides Gal-Man, puis nous isolerons le fragment m/z 169 (correspondant aux fragments b_1 et z_1).

4.2.3 Méthode IR

Avec les dispositifs instrumentaux installés, le couplage LC-MS-IR n'est possible qu'avec le Firefly et le laser CO₂ (le laser OPO n'est pas couplé à la LC-MS). Le cas idéal est donc de reproduire les conditions OPO avec le Firefly, c'est-à-dire longueur d'onde et puissance comparable (3410 cm⁻¹, 350 mW). Or le Firefly est un laser compact à faible puissance. Nous avons effectué des tests qui ont montré que la puissance du Firefly n'est pas suffisante pour obtenir une photofragmentation du fragment b_1 galactofuranose. Pour y remédier, un laser non résonant CO₂ a été ajouté au dispositif, dont le rôle est d'augmenter l'énergie interne des ions irradiés afin d'approcher la barrière de dissociation.¹⁵ Ce couplage a donc l'avantage d'être spécifique (longueur d'onde du fragment b_1 galactofuranose) mais en contrepartie le besoin d'aligner deux lasers dans le piège le rend plus difficile à utiliser. Une autre propriété du galactofuranose exploitée dans cette partie est non pas la résonance spécifique à 3410cm⁻¹ mais le fait qu'il soit plus fragile. En effet, il s'est avéré que la photofragmentation du b_1 galactofuranose était plus intense que le b_1 galactopyranose dans certaines conditions (laser CO₂ seul). Les avantages et inconvénients des différentes conditions laser sont récapitulés dans le Tableau 3.

Tableau 3 Récapitulatif des choix de conditions laser pour la méthode de criblage avec leurs avantages et inconvénients.

Système	Avantages	Inconvénients
Firefly seul à 3410 cm ⁻¹	Un seul laser Spécifique au b_1 galactofuranose	Photofragmentation insuffisante
Firefly 3410 cm ⁻¹ + laser CO ₂	Spécifique au b_1 galactofuranose	Montage difficile impliquant deux lasers
Firefly seul à 3650 cm ⁻¹	Un seul laser Fragmentation suffisante	Longueur d'onde non spécifique au galactofuranose
Laser CO ₂ seul	Un seul laser Choix de la puissance	Non résonant

Concrètement, on appellera précurseur le disaccharide standard, fragment le fragment m/z 169 détecté en LC-MS/MS et photofragment l'observable par IRMPD, c'est-à-dire l'ensemble des fragments IRMPD obtenus en LC-MS/MS-IR. Le ratio de photofragmentation est calculé en divisant l'intensité des photofragments par l'intensité totale (photofragments+fragment).

4.3 Criblage par irradiation infrarouge non résonnante

4.3.1 Optimisation des conditions laser pour la séparation des composés

Nous avons choisi dans un premier temps de vérifier s'il était possible de faire la distinction entre des composés avec et sans galactofuranose en utilisant un laser non résonnant. Pour ce faire, nous avons injecté un disaccharide en infusion direct que nous avons fragmenté par CID. Le fragment m/z 169 est ensuite isolé et irradié avec le laser non résonnant à des puissances différentes. La courbe de puissance est obtenue en calculant le ratio intensité photofragments/intensité totale pour chaque puissance. La même expérience MS/MS-IR a été reproduite pour les huit disaccharides et les courbes de puissance sont tracées en Figure 12. À partir de 3 W, les fragments des disaccharides Galf-Man commencent à photofragmenter proportionnellement à la puissance laser.

Les ratios de photofragmentation pour les quatre composés sont très proches à certaines puissances comme aux points acquis à 4,7 et 5 W alors qu'ils sont variables à d'autres points comme à 5,3 W. La photofragmentation des disaccharides contenant du galactopyranose est en revanche nulle jusqu'à

5 W où peu de photofragments sont observés. Ces résultats démontrent que la distinction entre Gal β -Man et Gal ρ -Man est possible avec un laser non résonnant. Cela s'explique par la plus grande fragilité du galactofuranose. Ainsi la présence de fragment z_1 n'interfère pas, et nous sommes encore capable de différencier le fragment b_1 galactofuranose du fragment b_1 galactopyranose. Ces résultats sont encourageants pour le développement de la stratégie de criblage.

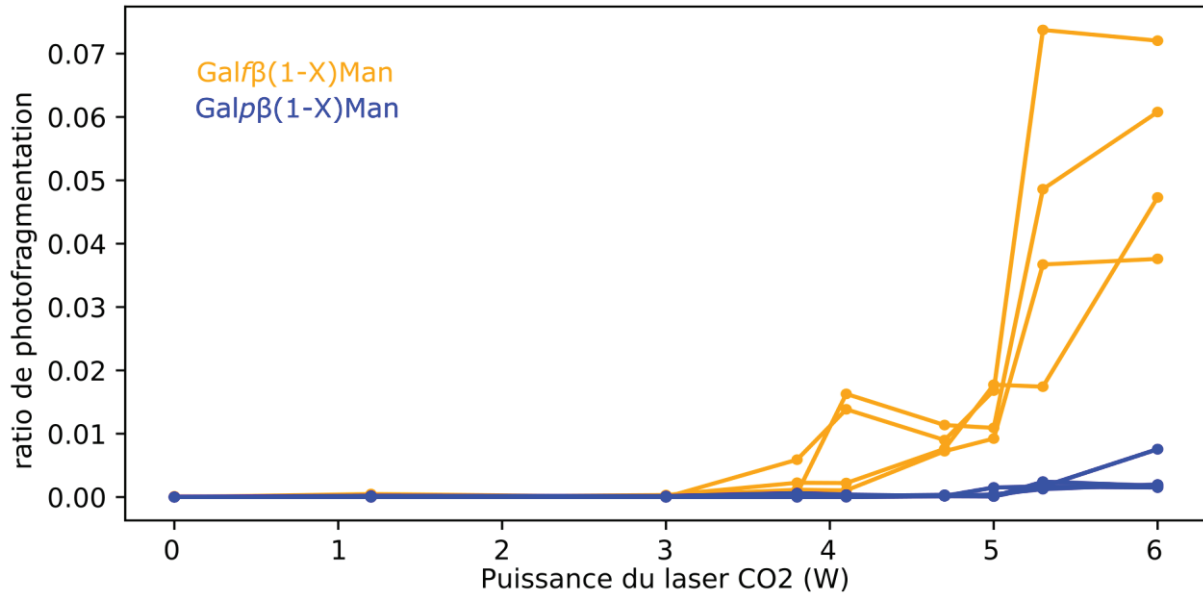


Figure 12 Courbes de puissance du fragment m/z 169 des quatre Gal β -Man complexés au lithium (en orange) et des quatre Gal ρ -Man complexés au lithium (en bleu) avec le laser CO2 seul (jour J).

Ces courbes sont acquises quotidiennement pour deux raisons. D'une part, elles aident à choisir la puissance optimale pour les expériences de criblage par LC-MS/MS-IR. En effet, les expérimentations à puissance fixe sont suffisantes pour distinguer les deux groupes mais il faut choisir la puissance optimale. Dans l'exemple ci-dessus, les écarts entre les deux groupes sont trop proches entre 3 et 4 W. À 6 W, l'écart est plus important mais notons que le fragment des Gal ρ -Man commence à photofragmenter à haute puissance. Il est préférable d'opter pour une puissance où leur signal est quasi-nul pour mieux distinguer les deux groupes. D'autre part, une fois la puissance sélectionnée, le ratio calculé à cette puissance nous permet de normaliser les valeurs acquises ce jour. En effet, les ratios de photofragmentation sont très sensibles à l'alignement du laser dans le piège. Bien que cet alignement soit vérifié quotidiennement, il est difficile de reproduire les mêmes conditions.

La Figure 13 indique les courbes de puissance du couple Gal β (1-3)Man (en bleu pour la forme pyranose et en orange pour la forme furanose) acquises plusieurs semaines après les spectres décrits ci-dessus. L'alignement du faisceau laser a été optimisé entre l'acquisition des résultats présentés entre les deux figures. Un écart important peut être constaté, une plus faible puissance est nécessaire pour la photofragmentation. Cela justifie une normalisation des données par des valeurs acquises quotidiennement. Une puissance de 1,7 W a été sélectionnée ce jour, valeur pour laquelle le fragment des Gal ρ -Man ne photofragmentent pas. Le ratio de photofragmentation correspondant à 1,7 s'élève à 0,016. Les données LC-MS/MS-IR acquises ce jour l'ont été avec une puissance de 1,7 et ont été normalisées en les divisant par 0,016.

4. Développement d'une méthode de criblage en LC-MS/MS-IR

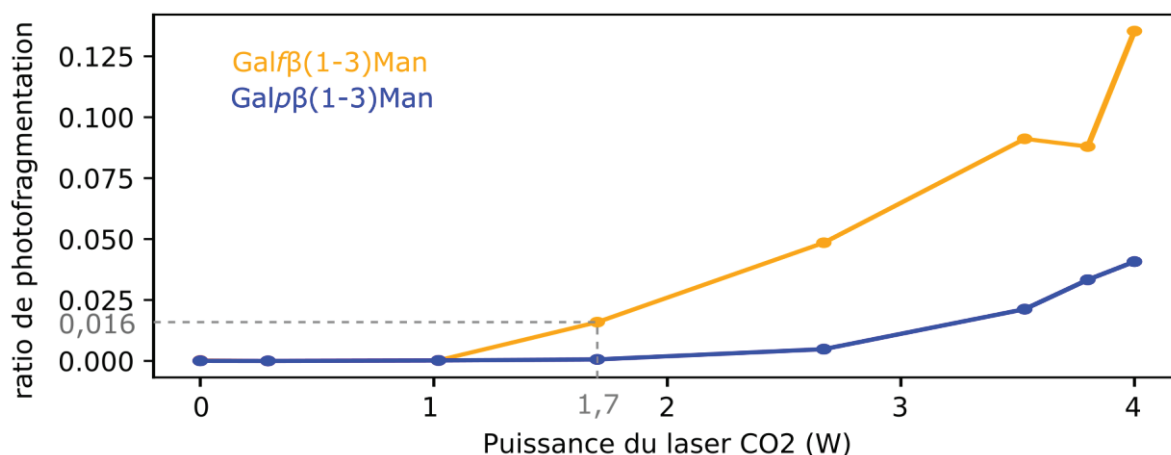


Figure 13 Courbes de puissance de Galfβ(1-3)Man en orange et Galpβ(1-3)Man en bleu avec le laser CO2 seul (jour J+15).

4.3.2 Résultats

Les chromatogrammes de chaque disaccharide ont été enregistrés selon les conditions LC et MS/MS décrites précédemment. Le signal de l'ion fragment m/z 169 est représenté par une ligne noire continue en Figure 14. Notons qu'il correspond également au signal de l'ion parent qui n'a pas été enregistré. Les quatre isomères contenant du galactofuranose ont des temps de rétention assez différents : 8-9,5 min pour Galfβ(1-2)Man, 14,1-16 min pour Galfβ(1-3)Man, 24,5-25 min pour Galfβ(1-4)Man et 9,5-10 min pour Galfβ(1-6)Man. Les temps de rétention des isomères Galp-Man sont plus rapprochés : 7,7-8,9 min pour Galpβ(1-2)Man, 13,7-14,2 min pour Galpβ(1-3)Man, 10-10,9 min pour Galpβ(1-4)Man et 9,2-10,4 min pour Galpβ(1-6)Man. Certains composés, en particulier les couples de régioisomères, ont des temps de rétention très proches.

Les pics chromatographiques ont une durée d'au moins 30 secondes et la durée d'irradiation par séquence MS est de 130 ms, ce qui donne plus d'une centaine de séquences LC-MS/MS-IR par composé. L'intensité de la photofragmentation est représentée en couleur (orange pour les Galf-Man et bleu pour les Galp-Man). Comme elle a été normalisée, notons que la photofragmentation peut paraître plus intense que l'ion précurseur qui n'a pas été normalisé. Les photofragments sont plus intenses pour les Galf-Man que pour les Galp-Man.

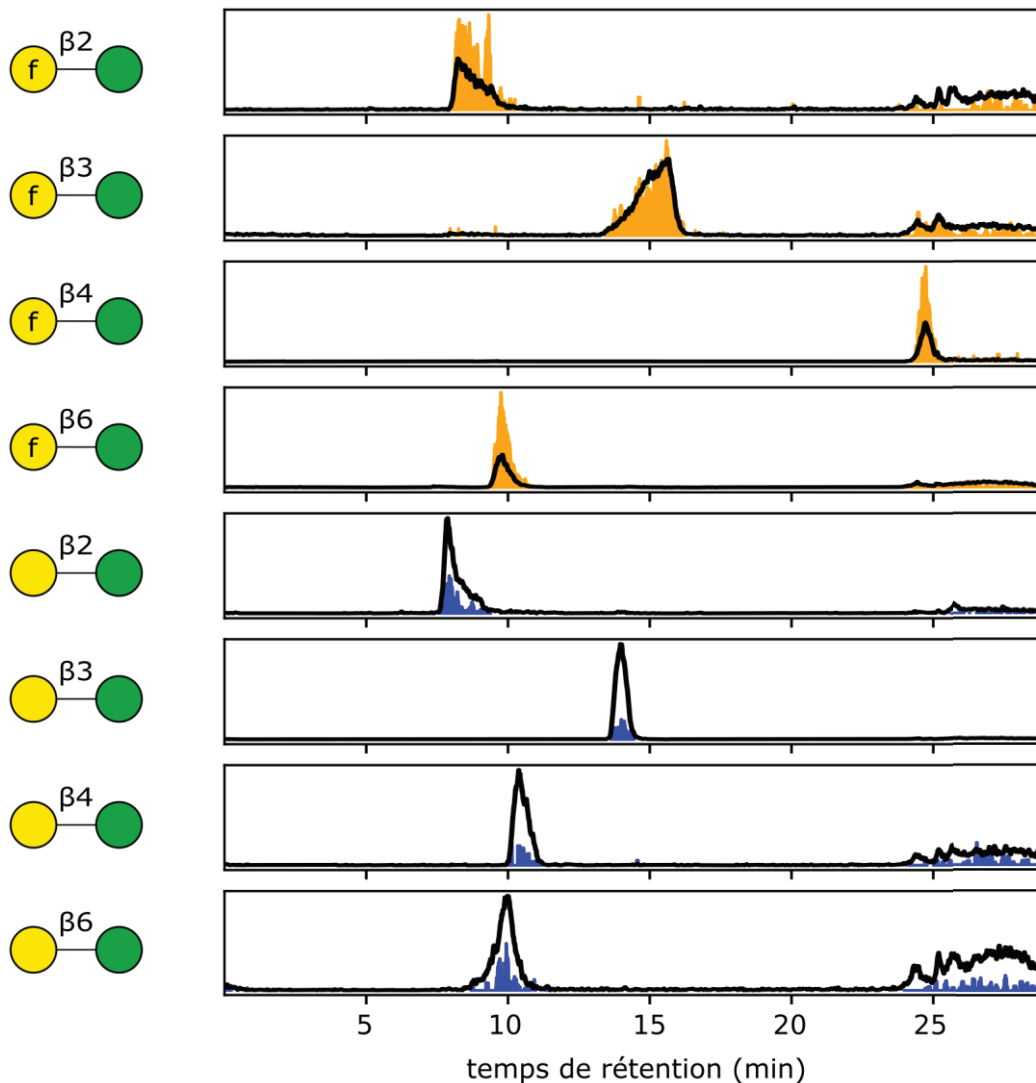


Figure 14 Criblage de chaque disaccharide au laser CO₂ (la structure du disaccharide est à gauche). La ligne continue correspond à l'intensité de l'ion m/z 169 et l'aire colorée (orange pour les Galf-Man et bleu pour les Galp-Man) correspond à l'intensité normalisée des photofragments.

Les données de criblage sont collectées, nous allons à présent passer à leur traitement. Deux éléments sont nécessaires pour distinguer les composés Galf-Man et Galp-Man : des valeurs quantitatives et un critère pour classer le composé.

4.3.3 Calcul de ratio

Pour obtenir des valeurs quantitatives, les rapports de photofragmentation en LC-MS/MS-IR ont été calculés. Ces ratios sont moyennés sur la période de rétention. Les moyennes sont représentées dans la Figure 15. Nous pouvons noter que les moyennes de photofragmentation des Galp-Man (en bleu) sont inférieures à 0,12 et celles des Galf-Man (en orange) sont comprises entre 0,33 et 0,66. Le fragment m/z 169 des Galf-Man fragmente en moyenne 5,5 fois plus que celui des Galp-Man.

4. Développement d'une méthode de criblage en LC-MS/MS-IR

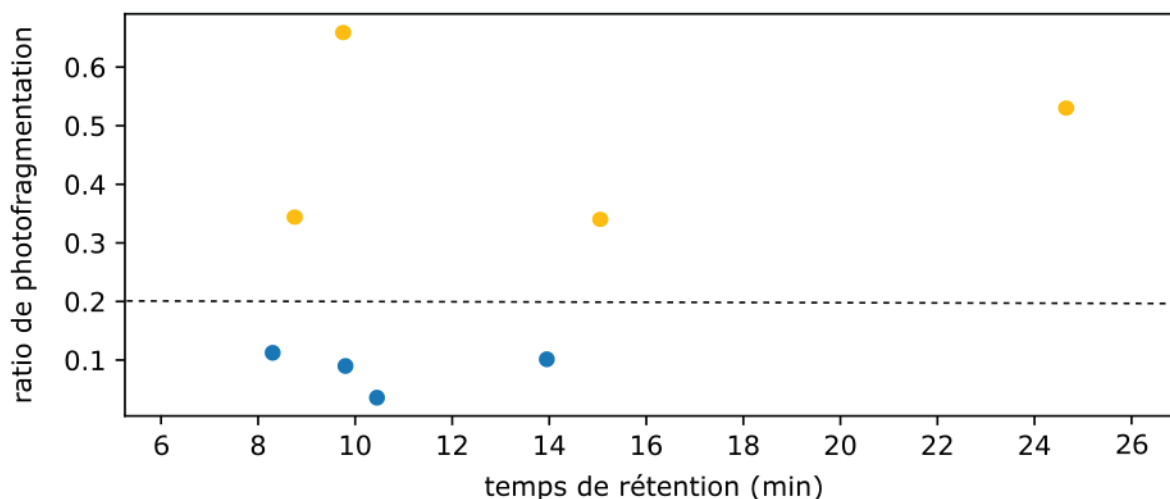


Figure 15 Ratio de l'intensité des photofragments sur l'intensité de l'ion fragment m/z 169 des huit composés avec le laser CO_2 seul. Les ratios des composés Galp-Man sont en orange et ceux des Galf-Man en bleu.

4.3.4 Sélection de la valeur seuil

Enfin, la dernière étape du processus de criblage consiste à définir un critère de classification. Notre critère de sélection des disaccharides contenant du galactofuranose sera ainsi : « si le ratio de photofragmentation est supérieur à la valeur seuil, le disaccharide peut contenir du galactofuranose ». Il faut donc définir un seuil. Cette valeur affecte la sensibilité et la spécificité du test. Si elle est trop faible, le test sera sensible mais peu spécifique, c'est-à-dire qu'il sera en mesure de valider le galactofuranose mais validera aussi à tort des composés avec du galactopyranose. La valeur du seuil dépend donc de l'utilisation souhaitée. Notons qu'une meilleure sensibilité est à privilégier dans le cas du criblage afin de ne pas éliminer à tort des composés contenant du galactofuranose. Le cas inverse (sélectionner des composés contenant du galactopyranose) est moins grave puisque le criblage est suivi d'autres étapes d'analyse qui permettront d'éliminer ces faux positifs.

Nous pouvons voir que nous sommes dans un cas de test parfait puisque certaines valeurs permettent d'avoir une sensibilité et une spécificité toutes deux égales à 1. Ce seuil peut être choisi entre 0,12 et 0,33. Une valeur faible est préférable pour éviter d'exclure faussement les composés avec du galactofuranose (des faux négatifs). Nous avons fixé le seuil à 0,2 pour l'analyse des composés de lichen dans le chapitre suivant. Ainsi les composés qui auront un ratio de photofragmentation supérieur à 0,2 avec ces conditions laser seront classés comme contenant potentiellement du galactofuranose.

4.4 Criblage par irradiation infrarouge résonante

4.4.1 Résultats

Voyons ce que nous observons avec le laser résonnant seul. Comme mentionné précédemment, le Firefly n'est pas assez puissant pour photofragmenter à 3410 cm^{-1} , longueur d'onde spécifique au fragment b_1 des Galp-Man. Une alternative est de choisir une longueur d'onde où l'absorption est plus forte, comme à 3650 cm^{-1} . Le fragment b_1 des Galp-Man absorbe également à cette longueur d'onde mais plus faiblement que celui des Galf-Man. L'hypothèse testée est donc la possibilité d'effectuer du criblage avec le laser Firefly seul en comparant les intensités de photofragmentation.

Nous avons injecté un mélange des quatre disaccharides Galp-Man (Figure 16, en haut) puis un mélange des quatre disaccharides Galf-Man (Figure 16, en bas).

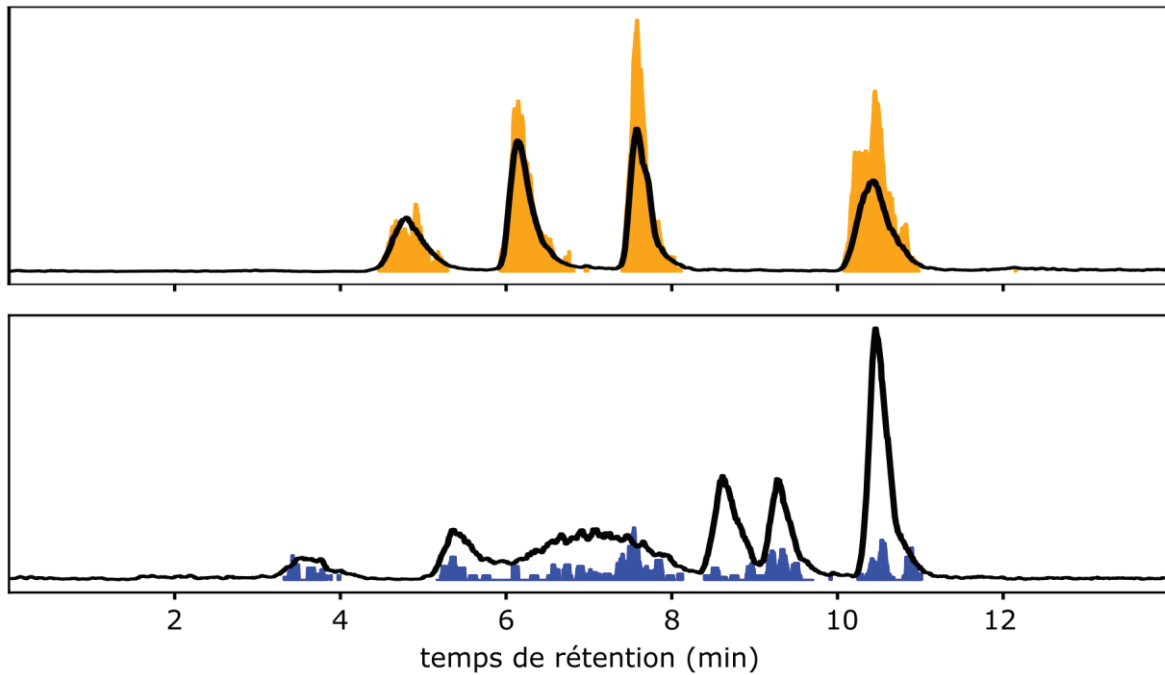


Figure 16 Criblage du mélange des quatre disaccharides Galf-Man (en haut) et des quatre disaccharides Galp-Man (en bas). La ligne continue correspond au signal de l'ion fragment m/z 169 et l'aire colorée à l'intensité normalisée des photofragments.

Le chromatogramme en ligne noire continue nous indique la présence de quatre pics résolus pour le mélange des Galf-Man : 4,5-5 minutes, 6-6,5 minutes, 7,4-7,9 minutes et 10,1-10,8 minutes. Le chromatogramme des Galp-Man est plus complexe puisque six pics sont observés : un plutôt faible entre 3,3 et 3,9 minutes, deux pics non résolus à 5,2-5,8 et 6,1-8 minutes, puis trois autres pics plus intenses et fins à 8,4-8,9, 9,1-9,5 et 10,3-10,9 minutes. Ce second mélange montre plus de pics que le nombre de composés injectés. Cela peut s'expliquer par la présence d'impuretés ou par l'interconversion α/β des disaccharides dont l'extrémité réductrice est libre.

La photofragmentation du fragment m/z 169 est représentée en orange pour les disaccharides Galf-Man et en bleu pour les Galp-Man. Nous pouvons observer qu'ils absorbent tous à 3650 cm^{-1} , ce qui est en accord avec les spectres des fragments b_1 (figure page 108). L'intensité d'absorption des différents composés d'un même mélange est comparable. En revanche, le fragment des Galf-Man absorbe beaucoup plus que le fragment des Galp-Man. Ces résultats suggèrent que l'irradiation du fragment m/z 169 à 3650 permet de distinguer les disaccharides contenant du galactofuranose ou du galactopyranose.

4.4.2 Calcul de ratio

Les ratios de fragmentation ont été calculés et moyennés sur la période de rétention (Figure 17). Les moyennes sont assez proches pour une taille de cycle donnée : entre 0,4 et 0,56 pour le premier mélange contenant des Galf-Man (en orange) et entre 0,03 et 0,12 pour le second mélange contenant des Galp-Man. Les intensités de photofragmentation entre les deux groupes sont en revanche très différentes : 5,8 fois supérieur pour les Galf-Man. Nous pouvons ainsi clairement discriminer les deux groupes selon leur taille de cycle. Le contraste est comparable à celui obtenu précédemment avec le laser CO_2 seul.

4. Développement d'une méthode de criblage en LC-MS/MS-IR

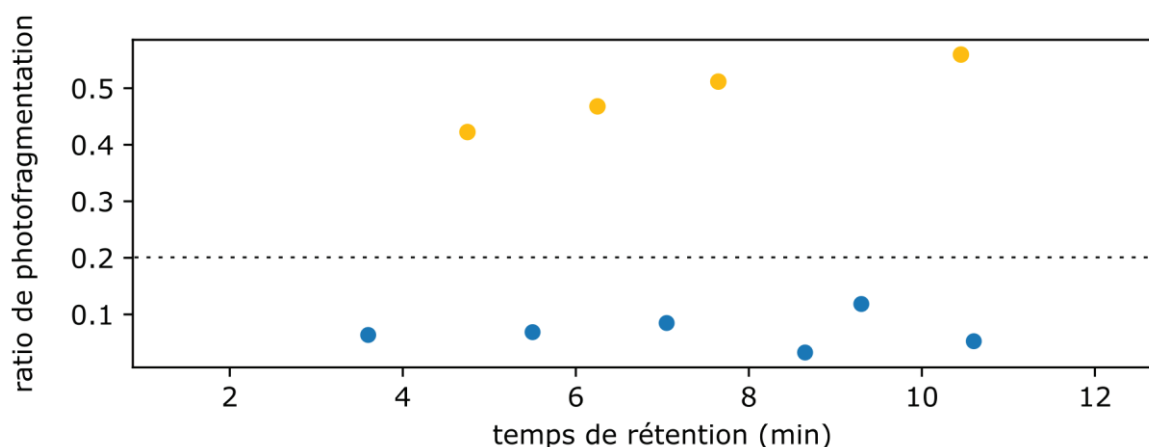


Figure 17 Ratio de l'intensité des photofragments sur l'intensité de l'ion fragment m/z 169 des composés détectés précédemment avec le laser Firefly seul à 3650 cm^{-1} . Les ratios des composés Galp-Man sont en orange et ceux des Galf-Man en bleu.

4.4.3 Sélection de la valeur seuil

Comme pour le cas précédent, nous avons ici un cas parfait qui permet de distinguer sans ambiguïté les Galp-Man et les Galf-Man. La valeur seuil séparant les deux groupes peut être fixée entre 0,12 et 0,4 pour avoir une sensibilité et une spécificité maximale. Nous avons choisi un seuil de 0,2.

4.5 Criblage par irradiation infrarouge résonante à longueur d'onde spécifique au galactofuranose

4.5.1 Optimisation des conditions laser pour la séparation des composés

La longueur d'onde de 3410 cm^{-1} est spécifique du b_1 galactofuranose, contrairement à la longueur d'onde de 3650 cm^{-1} à laquelle absorbent les fragments b_1 galactofuranose et galactopyranose. Bien que la discrimination soit très satisfaisante à 3650 cm^{-1} , il est intéressant d'effectuer les analyses à 3410 cm^{-1} pour évaluer l'efficacité dans ces conditions. En revanche, comme la puissance du laser résonant Firefly est insuffisante, nous aurons recours à l'ajout du laser CO₂.

Les courbes de puissance des huit disaccharides ont été enregistrées avec les deux lasers (Firefly à 3410 cm^{-1} et puissance du laser CO₂ variable). Dans la Figure 18, l'axe des abscisses représente la puissance du laser CO₂, celle du laser Firefly étant fixe. Le ratio de fragmentation en ordonnée correspond à la fragmentation obtenue conjointement par les deux lasers. Nous pouvons remarquer une photofragmentation du fragment m/z 169 des Galp-Man (en orange) dès 1 W de laser CO₂ qui augmente ensuite linéairement avec la puissance du laser. Le fragment d'intérêt pour les composés Galp-Man (en bleu) ne fragmente qu'à partir de 4,9 W et cette fragmentation n'est probablement pas résonante et due uniquement à la puissance élevée du laser CO₂.

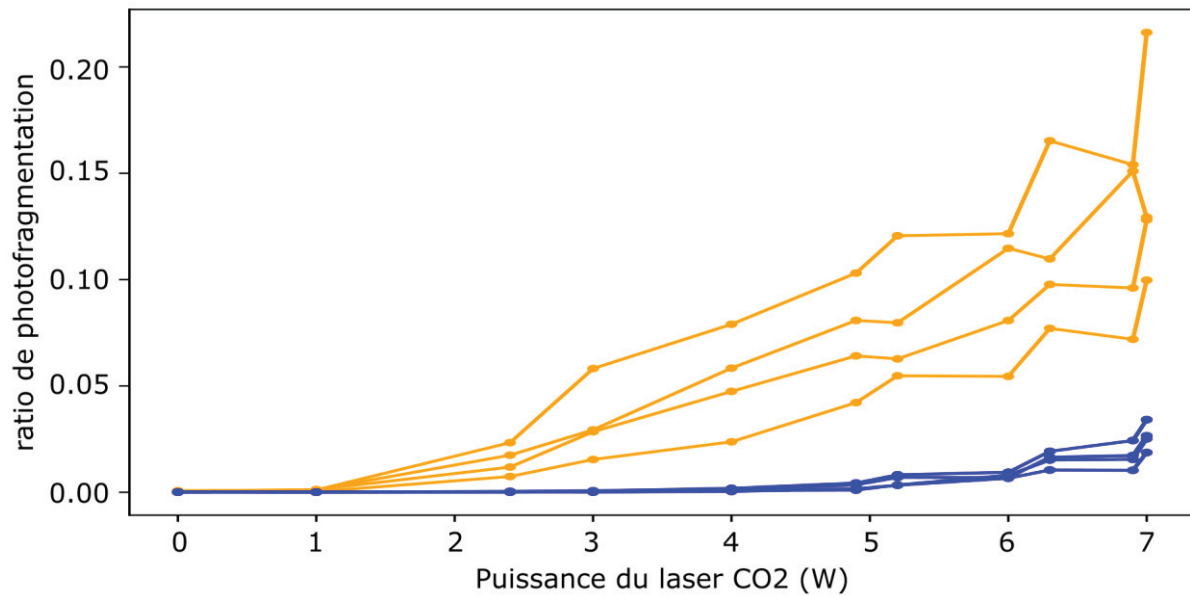


Figure 18 Courbes de puissance du fragment m/z 169 des quatre Gal α -Man complexés au lithium (en orange) et des quatre Gal ρ -Man complexés au lithium (en bleu) avec le laser Firefly à 3410 cm^{-1} et le laser CO2.

4.5.2 Résultats

Nous avons injecté le mélange de deux disaccharides Gal β (1-2)Man et Gal β (1-3)Man (Figure 19, en haut) puis un mélange de leurs analogues Gal ρ (1-2)Man et Gal ρ (1-3)Man (Figure 19, en bas).

Pour le mélange Gal α -Man, nous pouvons noter la présence d'un premier pic chromatographique à 7,3-8,5 minutes puis un second, beaucoup plus intense et plus fin de 11,1 à 1,6 minutes. Une faible impureté apparaît à la 20^e minute. Le profil du mélange Gal ρ -Man est très similaire, avec un premier pic peu intense à 8,5-9,2 minutes puis un second plus intense à 11,2-11,8 minutes. Une impureté est détectée à 5,4 minute. Dans la partie 4.3.2, nous avons vu que la taille du cycle modifiait peu la rétention pour les composés Gal β (1-2)Man et Gal β (1-3)Man. Nous pouvons donc supposer que les deux composés avec un temps de rétention de 8 min sont des régioisomères et de même pour les deux composés élués vers la 11^e minute.

Alors que les profils chromatographiques sont similaires pour ces régioisomères, leurs profils d'absorption sont distincts. Une très forte absorption est observée pour le fragment des Gal α -Man alors qu'il n'y a pas d'absorption pour les Gal ρ -Man. Ce résultat est attendu puisqu'on a constaté que seul le fragment b_1 galactofuranose présente une activité vibrationnelle à 3410 cm^{-1} .

4. Développement d'une méthode de criblage en LC-MS/MS-IR

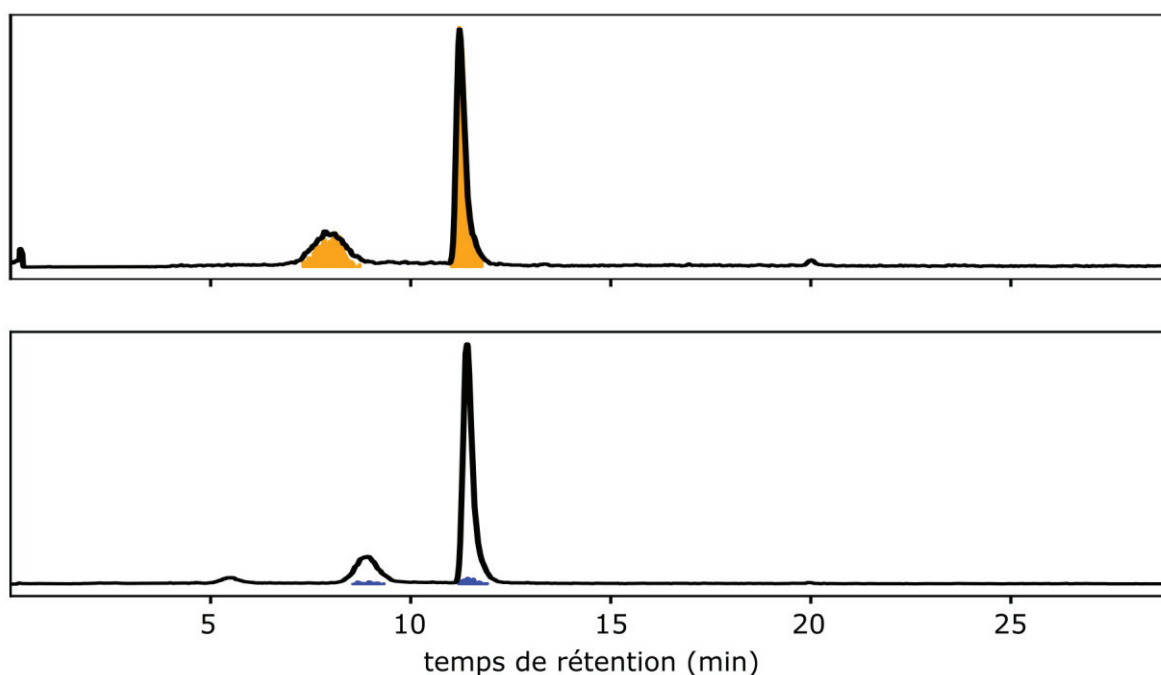


Figure 19 Criblage du mélange Galf β (1-2)Man et Galf β (1-3)Man (en haut) et du mélange Galp β (1-2)Man et Galp β (1-3)Man (en bas). La ligne continue correspond au signal de l'ion fragment m/z 169 et l'aire colorée à l'intensité normalisée des photofragments.

4.5.3 Calcul de ratio

Les ratios de fragmentation ont été calculés et leurs moyennes sont représentées par des points orange pour les Galf-Man et des points bleus pour les Galp-Man (Figure 20). Les moyennes sont comprises entre 0,3 et 0,5 pour les premiers et en dessous de 0,05 pour les seconds. Nous pouvons remarquer qu'il s'agit des conditions laser où le taux de photofragmentation pour les Galp-Man est le plus faible. Le contraste est de 25, il s'agit du meilleur rapport pour différencier les deux tailles de cycle. Mais rappelons que c'est également l'installation la plus difficile à mettre en place avec l'alignement de deux lasers dans le piège.

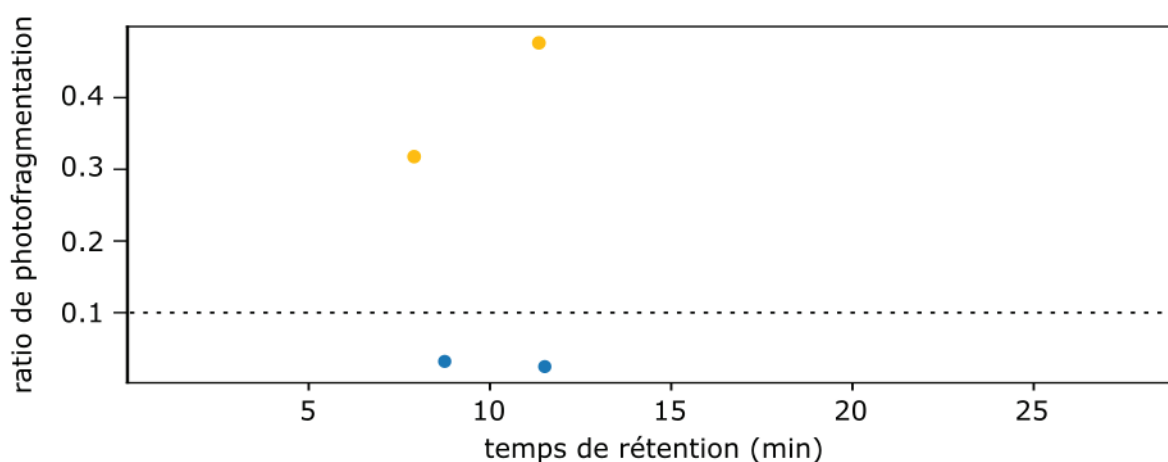


Figure 20 Ratio de l'intensité des photofragments sur l'intensité de l'ion fragment m/z 169 des composés détectés précédemment avec le couplage du laser Firefly à 3410 cm^{-1} et du laser CO₂. Les ratios des composés Galp-Man sont en orange et ceux des Galf-Man en bleu.

La valeur seuil, à choisir dans l'intervalle de 0,05 et 0,3, a été fixée à 0,1 avec ces conditions laser. Tout comme dans les cas précédents, la sensibilité et la spécificité du test sont maximales.

4.6 Interprétation

Le Tableau 4 reprend les trois conditions laser développées pour le criblage. Le contraste, c'est-à-dire le rapport entre les ratios de photofragmentation furanose/pyranose, est comparable pour les conditions laser CO₂ seul (5,5) et laser Firefly à 3650 cm⁻¹ seul (5,8). La combinaison des deux lasers (Firefly à 3410 cm⁻¹ assistée par le laser CO₂) porte ce contraste à 25. Il est donc tentant de choisir ces conditions laser. Cependant nous pouvons noter que pour chaque condition laser, en fixant la valeur seuil définie dans le tableau, la sensibilité est à 1 (tous les Galp-Man ont un ratio de photofragmentation au-dessus de cette valeur) et la spécificité est aussi à 1 (tous les Galp-Man ont un ratio de photofragmentation en dessous de la valeur seuil). Ainsi les trois conditions répondent à notre besoin analytique sans erreur. Par ailleurs, le couplage des deux lasers est lourd à mettre en place et ne représente donc pas forcément la meilleure option.

Tableau 4 Récapitulatif des choix de conditions laser pour la méthode de criblage et leurs caractéristiques respectives

Système	Contraste	Valeur seuil	Sensibilité	Spécificité
Laser CO ₂ seul	5,5	0,2	100%	100%
Firefly seul à 3650 cm ⁻¹	5,8	0,2	100%	100%
Firefly 3410 cm ⁻¹ + laser CO ₂	25	0,1	100%	100%

En somme, la méthode de criblage LC-MS/MS-IR développée dans le cadre de cette thèse est très robuste. La préparation d'échantillon est rapide (pas de marquage à effectuer) et une faible quantité d'échantillon est nécessaire. Les composés sont d'abord séparés par chromatographie liquide. Ensuite, les ions précurseurs m/z 349 sont isolés et fragmentés par CID pour isoler le fragment m/z 169. Cette étape de MS/MS apporte donc une double spécificité et les substances autres que des disaccharides sont éliminées. Enfin, l'infrarouge permet de discriminer les galactofuranoses. Les trois conditions laser proposées montrent d'excellentes performances.

5 Conclusion

Nous savons que les fragments intracycliques fournissent l'information sur la régiochimie mais que l'information sur la taille du cycle est perdue puisque la fragmentation a lieu au niveau du cycle.

Les fragments issus de rupture de la liaison glycosidique portent quant à eux l'information sur la régiochimie (fragments disaccharidiques) mais également sur la taille du cycle (fragments mono- et disaccharidiques). La résolution de l'IRMPD est suffisante pour distinguer les différents fragments.

Par conséquent, la MS/MS ou l'IRMPD sur les fragments disaccharidiques peuvent être appliqués pour déterminer la régiochimie. Ces deux méthodes sont compétitives car les spectres MS/MS pas toujours faciles à interpréter. Nous avons vu qu'un système automatisé pouvait être développé pour assurer une interprétation rapide des spectres IRMPD (calcul de score ou intelligence artificielle).

Pour obtenir l'information sur la taille du cycle, l'IRMPD sur les fragments disaccharidiques est envisageable. Contre toute attente, l'IRMPD sur les fragments monosaccharidiques permet également de déterminer la taille du cycle. L'avantage de cette dernière méthode est qu'elle ne nécessite que deux références qui sont très caractéristiques, l'interprétation est donc rapide.

Enfin, nous avons déployé une méthode de criblage par LC-MS/MS-IR sur les fragments monosaccharidiques pour discriminer les disaccharides contenant du galactofuranose.

En somme, trois stratégies basées sur la MS/MS-IR ont été développées sur des oligosaccharides de synthèse dans ce chapitre. Selon l'information souhaitée, le séquençage sur les monosaccharides, le séquençage sur les disaccharides ou le criblage peuvent être explorés sur des oligosaccharides naturels. L'application de ces stratégies sur des polysaccharides est envisageable et nécessite au préalable une dépolymérisation (procédé permettant d'obtenir des oligosaccharides à partir de polysaccharides). Il n'y a donc qu'un pas entre notre développement sur les oligosaccharides synthétiques et l'application sur les polysaccharides naturels. Cela ouvre entièrement des champs de possibilités sur les polysaccharides de lichen.

6 Bibliographie

- 1 B. Domon and C. E. Costello, *Glycoconjugate J*, 1988, **5**, 397–409.
- 2 B. Schindler, L. Barnes, G. Renois, C. Gray, S. Chambert, S. Fort, S. Flitsch, C. Loison, A.-R. Allouche and I. Compagnon, *Nat Commun*, 2017, **8**, 1–7.
- 3 D. P. Demarque, A. E. M. Crotti, R. Vessecchi, J. L. C. Lopes and N. P. Lopes, *Nat. Prod. Rep.*, 2016, **33**, 432–455.
- 4 D. J. Harvey, T. S. Mattu, M. R. Wormald, L. Royle, R. A. Dwek and P. M. Rudd, *Analytical Chemistry*, 2002, **74**, 734–740.
- 5 O. Yeni, S. Ollivier, D. Ropartz, H. Rogniaux, F. Le Dévéhat, J. Boustie, L. Legentil, V. Ferrières, B. Moge and I. Compagnon. Ring-size memory of galactose-containing MS/MS fragment and application to the detection of galactofuranose in oligosaccharides, and their sequencing (soumis)
- 6 O. Yeni *et al.*, Screening using LC-MS/MS-IR to detect galactofuranose (en rédaction).
- 7 B. Favreau, O. Yeni, S. Ollivier, J. Boustie, F. L. Dévéhat, J.-P. Guégan, M. Fanuel, H. Rogniaux, R. Brédy, I. Compagnon, D. Ropartz, L. Legentil and V. Ferrières, *J. Org. Chem.*, 2021, **86**, 6390–6405.
- 8 S. Ollivier, L. Legentil, O. Yeni, L.-P. David, V. Ferrières, I. Compagnon, H. Rogniaux and D. Ropartz, *J. Am. Soc. Mass Spectrom.*, 2023, **34**, 627–639.
- 9 L-P David, S. Ferron, B. Favreau, D. Ropartz, O. Yeni, I. Compagnon, Vincent Ferrières, F. Le Dévéhat and L. Legentil. Lichenic oligosaccharides: synthesis of galactomannans fragments to help NMR assignment for extracted polysaccharides (soumis).
- 10 T. T. Fang and B. Bendiak, *J. Am. Chem. Soc.*, 2007, **129**, 9721–9736.
- 11 M. Valcárcel, S. Cárdenas and M. Gallego, *TrAC Trends in Analytical Chemistry*, 1999, **18**, 685–694.
- 12 Q. Enjalbert, R. Simon, A. Salvador, R. Antoine, S. Redon, M. M. Ayhan, F. Darbour, S. Chambert, Y. Bretonnière, P. Dugourd and J. Lemoine, *Rapid Communications in Mass Spectrometry*, 2011, **25**, 3375–3381.
- 13 Q. Enjalbert, M. Girod, R. Simon, J. Jeudy, F. Chiro, A. Salvador, R. Antoine, P. Dugourd and J. Lemoine, *Anal Bioanal Chem*, 2013, **405**, 2321–2331.
- 14 Thermo Scientific Hypercarb Columns: Applications Notebook, Issue 1, June 2009, https://assets.thermofisher.com/TFS-Assets/CMD/Application-Notes/ANGSCHYPERCARB0609-hypercarb_appnotebook.pdf.
- 15 B. Moge, O. Yeni, A. Infantino and I. Compagnon, CO₂ laser enhanced rapid IRMPD spectroscopy for glycan analysis, *International Journal of Mass Spectrometry*, 2023, **490**, 117071.

CHAPITRE 5 :

Détection de galactofuranose dans des extraits de lichen dépolymérisés

<u>1</u>	Introduction	124
<u>2</u>	Méthodes	128
	2.1 Dépolymérisation avec le protocole FITDOG	128
	2.2 Dépolymérisation par hydrolyse acide	128
	2.3 Analyse par LC-MS	128
	2.4 Criblage par LC-MS/MS-IR	129
<u>3</u>	Dépolymérisation de polysaccharides modèles par FITDOG	130
	3.1 Résultats MS	130
	3.2 Résultats IRMPD	130
<u>4</u>	Dépolymérisation de polysaccharides modèles par hydrolyse acide	132
	4.1 Résultats MS	132
	4.2 Résultats IRMPD	132
<u>5</u>	Cetraria islandica	134
	5.1 Présentation	134
	5.2 Résultats de dépolymérisation	134
	5.3 Séparation des oligosaccharides par LC-MS	135
	5.4 Criblage des disaccharides avec le laser Firefly seul	136
	5.5 Criblage des disaccharides avec le laser CO2 seul	137
<u>6</u>	Lasallia pustulata	139
	6.1 Présentation	139
	6.2 Résultats de dépolymérisation	139
	6.3 Criblage des disaccharides avec le laser Firefly seul	140
<u>7</u>	Roccella fuciformis	142
	7.1 Présentation	142
	7.2 Résultats de dépolymérisation	142
	7.3 Criblage des disaccharides avec laser Firefly seul	143
	7.4 Criblage des disaccharides avec laser CO2 seul	143
<u>8</u>	Discussion et perspectives	145
<u>9</u>	Bibliographie	147

1 Introduction

Il convient de rappeler l'objectif de cette thèse dans ce dernier chapitre. Nous souhaitons développer et proposer une méthode analytique permettant de détecter et identifier des motifs saccharidiques contenant des unités galactofuranose à partir d'échantillons de lichen. Pour ce faire, nous proposons une démarche complète et interdisciplinaire en sept étapes (illustrées dans le Schéma).

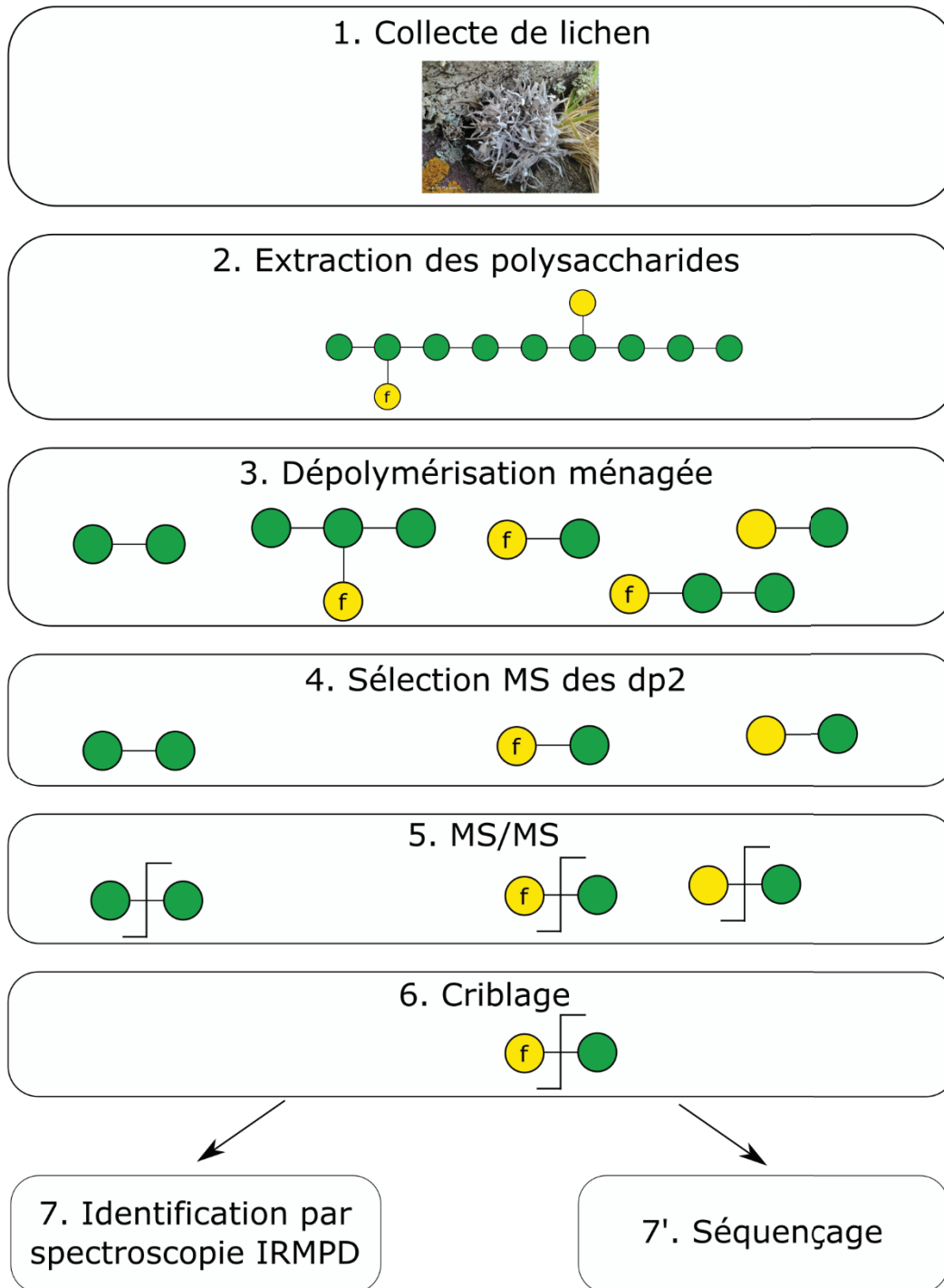


Schéma 1 Démarche complète en 7 étapes pour détecter et identifier des oligosaccharides contenant du galactofuranose à partir de lichen.

1. Introduction

Étapes 4 à 7 : développées sur les standards synthétiques

Nous avons démontré la faisabilité des étapes 4 à 7 sur les échantillons standards fournis par l'équipe de chimie de Rennes. Les résultats présentés dans les chapitres précédents montrent que l'on peut injecter en électrospray des oligosaccharides contenant des unités galactofuranose sans induire de modification structurale. Par conséquent, des oligosaccharides contenant du galactofuranose peuvent être examinés en MS. Cet instrument peut être utilisé pour isoler des oligosaccharides d'une taille donnée. Bien que des travaux aient été menés sur des mono- et trisaccharides, notre méthode a été principalement développée sur les disaccharides. En effet, nous avons créé une base de données avec les motifs disaccharidiques et la méthode de criblage a également été développée sur les disaccharides. Nous isolons donc spécifiquement les dp2 en MS lors de l'étape 4.

Ces dp2 isolés peuvent correspondre à plusieurs disaccharides. Nous comptons huit combinaisons de dp2 β -Gal-Man (variant par la taille du cycle et la régiochimie) auxquelles s'ajoutent d'autres dp2 (Man-Man, Man-Gal...). L'analyse de fragment monosaccharidique plutôt que du dp2 s'avère donc intéressante puisque l'on réduit encore la complexité en s'affranchissant du type de liaison. Il ne reste alors plus que trois références : Gal f , Gal p et Man (la conservation de l'information sur le galactofuranose après une fragmentation CID a été vérifiée sur les standards). L'étape 5 consiste donc à fragmenter par CID les disaccharides et à isoler le fragment monosaccharidique.

La méthode de criblage que nous avons mise au point s'est avérée très efficace sur les produits fournis pour discriminer ceux contenant une unité galactofuranose. La stratégie de criblage peut donc être appliquée sur le fragment monosaccharidique des disaccharides issus d'échantillons naturels afin de discriminer ceux possédant du galactofuranose (étape 6).

Les fractions d'intérêt seront alors identifiées à partir de la base de données créée (étape 7) ou par séquençage (étape 7'). Ceci est la démarche proposée pour détecter puis caractériser les disaccharides contenant du galactofuranose.

Mais alors, comment obtenir des disaccharides à partir de lichens présents sur des rochers en bord de mer, des troncs d'arbre ou autre lieu insolite ? Remontons aux étapes 1 à 3 pour y répondre.

Étapes 1 et 2 : effectuées par nos collaborateurs lichénologues

Les premières étapes nécessitent une bonne connaissance des lichens et de leur manipulation. Elles ont été réalisées par l'équipe de recherche CORINT (de l'ISCR). Diverses espèces de lichen avaient été sélectionnées en amont du projet : *Lasallia pustulata*, *Roccella fuciformis*, *Cetraria islandica* et *Lichina pygmaea*. La première étape consiste à reconnaître et collecter des échantillons de ces espèces. La collecte a été réalisée en Bretagne. Pour chaque espèce, il existe plusieurs types de polysaccharides.

Lors de la seconde étape, ces polysaccharides sont extraits séparément selon le protocole d'extraction qui figure page 30. Dans la version initiale du projet, cette même équipe devait réaliser une digestion enzymatique des polysaccharides avec des enzymes commerciaux reconnaissant spécifiquement le mannose (α -mannosidase et α -mannanase)¹. Ainsi cette étape allait conduire à des oligosaccharides de plus petite taille. L'obtention d'oligosaccharides est primordiale car rechercher un détail structural tel la taille de cycle du galactose dans de grands polysaccharides (plus de 10 kDa pour certaines espèces) s'avère extrêmement complexe. Or la digestion enzymatique n'a pas fonctionné. Il a donc fallu proposer et tester des alternatives pour que les échantillons de lichen puissent être analysés. Cette étape 3 a finalement été réalisée en collaboration des différentes équipes.

Étape 3 : effort conjoint des différentes équipes

Nous nous sommes orientés vers deux méthodes de dépolymérisation chimique :

- L'hydrolyse acide, une méthode ancienne et bien maîtrisée²
- FITDOG (Fenton's Initiation Toward Defined Oligosaccharide Groups), une méthode publiée au cours du projet ALGAIMS (en 2020)³ sur laquelle nous n'avons donc pas beaucoup de recul.

Il s'agit là de méthodes applicables sur différents types de polysaccharides. Le choix de la bonne méthode est primordial car il faut veiller à ce que les conditions chimiques ne dégradent pas la structure et plus précisément la structure furanoside qui est celle que nous recherchons.

Il est nécessaire d'explorer ces méthodes chimiques de dépolymérisation sur un système modèle pour vérifier que les structures sont conservées. Nous avons sélectionné le galactomannane issu de la gomme de caroube dont la séquence est représentée dans le Schéma 2. Il s'agit d'une chaîne de mannoses liés en $\beta(1-4)$ avec des ramifications de galactose en $\alpha(1-6)$. On compte en moyenne un galactose pour quatre unités de mannose. La composition est proche des polysaccharides de lichen qui contiennent en abondance du mannose et du galactose. Nous souhaitons ainsi vérifier si les conditions chimiques des protocoles de dépolymérisation impactent la structure. Nous avons au sein de notre laboratoire effectué une dépolymérisation complète du galactomannane par FITDOG ou par hydrolyse acide. Ensuite, nous avons analysé la structure des monosaccharides obtenus par MS-IR pour valider la structure du produit de dépolymérisation. Nous verrons dans les résultats que FITDOG n'a pas été retenu. Nos collaborateurs ont alors réalisé l'hydrolyse acide des polysaccharides de lichen.

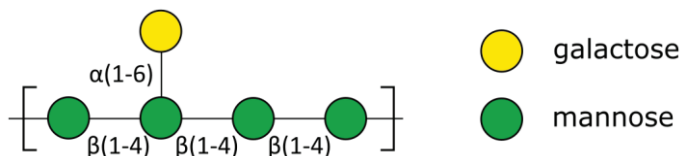


Schéma 2 Motif du galactomannane issu de gomme de caroube selon la nomenclature SNFG (légende à droite).

La question de structure se pose sur les furanoses : est ce que les conditions de l'hydrolyse acide (notamment un pH très acide) dégradent la structure furanoside ? Nous n'avons ici pas de système modèle de polysaccharide contenant du galactofuranose donc nous avons mené notre étude directement sur les polysaccharides de lichen. En cas de dépolymérisation complète, il est fort probable que la structure galactofuranoside soit perdue. En effet, les monosaccharides sont libérés dans le cas de dépolymérisation complète et les unités galactofuranosidiques isomérisent alors naturellement. Cela justifie que les analyses soient effectuées sur des oligosaccharides qui peuvent être obtenus par dépolymérisation ménagée. En plus d'avoir moins de risque de perdre l'information sur le galactofuranose, nous pouvons retrouver l'information sur les liaisons glycosidiques puisqu'elles ne sont pas rompues (toutes les liaisons sont rompues lors de la dépolymérisation totale et donc la régiochimie et la stéréochimie de la liaison sont perdues).

Notre choix de taille s'est plus spécifiquement porté sur les dp2 pour l'étude en IRMPD étant donné que nos résultats précédents (création de la base de données, criblage) ont pour la plupart été obtenus sur les modèles disaccharidiques.

1. Introduction

Nos collaborateurs de l'ISCR ont réalisé une hydrolyse acide des polysaccharides extraits avec différents temps de réaction. Le récapitulatif des échantillons fournis figure dans le Tableau 5.

Tableau 5 Échantillons fournis par l'ISCR (seuls ceux marqués par un astérisque ont été analysés à l'ILM).

Espèce de lichen	Temps de l'hydrolyse acide
<i>Cetraria islandica</i>	2 heures
	6 heures*
	8 heures
	24 heures*
<i>Lasallia pustulata</i>	2 heures
	6 heures*
	8 heures
	24 heures*
<i>Roccella fuciformis</i>	15 minutes
	45 minutes
	2 heures*
	3 heures*

Ce dernier chapitre s'intéresse à l'analyse de polysaccharides. Dans une première partie, les différentes méthodes de dépolymérisation sont testées sur des polysaccharides modèles. Nous verrons que l'hydrolyse acide a été retenue. Dans la seconde partie, les polysaccharides de lichen hydrolysés sont analysés. Les résultats présentés dans ce chapitre font l'objet d'un article et certains des résultats ont été publiés dans l'article de criblage.^{4,5}

2 Méthodes

2.1 Dépolymérisation avec le protocole FITDOG

La réaction FITDOG est une méthode de dépolymérisation chimique récemment publiée, dont le protocole est détaillé dans la référence [3]. La première étape consiste à incuber le polysaccharide à 1mg/mL dans un mélange d'acétate de sodium, d'acide acétique, de peroxyde d'hydrogène et de sulfate de fer (III) à 100 °C pendant une heure. La réaction est arrêtée par ajout de NaOH et la solution est ensuite neutralisée. Lors de la seconde étape, les oligosaccharides obtenus sont réduits par incubation avec du NaBH₄. Enfin, les oligosaccharides sont purifiés avec des colonnes SPE. Cette méthode chimique a été testée sur des polysaccharides modèles (cellulose de Thermo, référence 10065563 et galactomannane issu de gomme de caroube commercialisé par Sigma-Aldrich, référence G0753).

2.2 Dépolymérisation par hydrolyse acide

Il existe différents protocoles d'hydrolyse acide dont les paramètres variables sont le choix de l'acide (TFA, HCl, H₂SO₄...), sa concentration, le temps et la température d'incubation. Les tests réalisés à Lyon se basent sur le protocole suivant : 500 µL de polysaccharides à 1 mg/mL et 500 µL d'HCl à 6 mol/L sont incubés à 100 °C pendant 30 minutes. Les échantillons envoyés par Rennes suivent ce protocole : sont ajoutés dans 10mL de TFA à 10 mM 20 mg de polysaccharide dilués dans 10 mL d'eau. Le mélange est incubé à 100 °C pendant une durée variable entre 0 et 24 h (le temps de réaction est indiqué pour chaque échantillon) puis 10mL de NaOH sont ajoutés pour arrêter l'hydrolyse. Une purification par colonne SPE est ensuite mise en place. La solution éluée est séchée au SpeedVac.

Pour les polysaccharides de lichen, les profils de distribution des oligosaccharides ont été analysés en MS avec deux instruments qui ont des fenêtres de rapports m/z différentes : le MALDI-MS de Nantes permet de détecter les rapports m/z entre 700 et 3000 (donc entre dp4 et dp19) et l'ESI-MS que nous avons à disposition permet de détecter les plus petits rapports m/z entre 200 et 2000 (de dp2 à dp12).

2.3 Analyse par LC-MS

Préparation d'échantillon : les extraits hydrolysés ont été mis en solution à une concentration de 1 mg/mL.

Conditions LC : Les séparations ont été réalisées avec le dispositif LC (U3000)-MS (LTQ XL). Le volume injecté est de 10 µL.

Deux colonnes Hypercarb ont permis les séparations. Les conditions suivantes ont été utilisées avec la colonne 1 (référence : 35003-101030) : 10 µL de volume injecté (à 1 mg/mL), débit à 0,1 mL/min, température à 50 °C et la phase mobile suit le gradient (avec A= 100% H₂O et B=100% ACN):

0-8 minutes : 90% A, 0% B, 10% C. 8-11 min : gradient de 0 à 75% B, 10% C. 11-13 min : 25% A, 75% B. 13-18 : 100% A, 0% B. Pour la formation d'adduit Li⁺, nous avons remplacé 10% de A par du LiCl 10 µM.

Pour la colonne 2 (référence : 35005-154630), on injecte 10 µL, le débit est de 0,4 mL/min (T en sortie de chromatographie, donc débit entrant dans la MS d'environ 0,2 mL/min), la température est à 40 °C et le gradient est le suivant (avec A= 100% H₂O, B=100% ACN, C= 10 µM LiCl dans H₂O) :

2. Méthodes

0-16 min : 92% A, 5% B, 3% C. 0-26 min : gradient de 5% à 75% B, 3% C. 26-29 min : 75% B et 3% C.
29-35 minutes : rinçage avec 92% A, 5% B, 3% C.

2.4 Criblage par LC-MS/MS-IR

Le dispositif instrumental utilisé pour le criblage est le LC (U3000) - MS (LTQ XL) - laser IR (Firefly et/ou laser CO₂). Le schéma du couplage figure page 21.

Préparation d'échantillon et conditions LC : voir partie précédente.

Conditions MS : l'ion m/z 349 correspondant au complexe $[dp_2+Li]^+$ est isolé et fragmenté par CID. Le fragment m/z 169 est alors réisolé pendant 150 ms et photofragmenté selon les conditions IR décrites dans le paragraphe suivant. Cet ion représente les fragments b_1 et z_1 . Pour rappel, la symétrie des disaccharides n'est pas un obstacle dans une perspective de criblage car notre objectif est de détecter le fragment galactofuranose. Le mélange de celui-ci avec un autre fragment n'empêche pas sa détection à condition que l'autre fragment ne soit pas majoritaire.

Conditions laser : Les mesures ont été obtenues avec deux conditions laser. La première condition laser est le Firefly seul qui émet à 3650 cm^{-1} . Dans le second, le laser CO₂ irradie seul à la puissance définie par les courbes quotidiennes (environ 1,7 W). La troisième condition développée dans le chapitre précédent (couplage du laser Firefly à 3410 cm^{-1} et du laser CO₂) n'a pas été appliquée sur les échantillons de lichen.

3 Dépolymérisation de polysaccharides modèles par FITDOG

Le galactomannane a été dépolymérisé selon le protocole FITDOG. Nous verrons les résultats MS puis MS-IR.

3.1 Résultats MS

Le spectre MS du galactomannane après réaction FITDOG est représenté en Figure 21. Des oligosaccharides chargés au Na^+ sont observés. Ils ont une taille allant du monosaccharide au dp8, avec une distribution centrée sur le trisaccharide (dp3). Le protocole FITDOG a donc bien permis de dépolymériser le galactomannane.

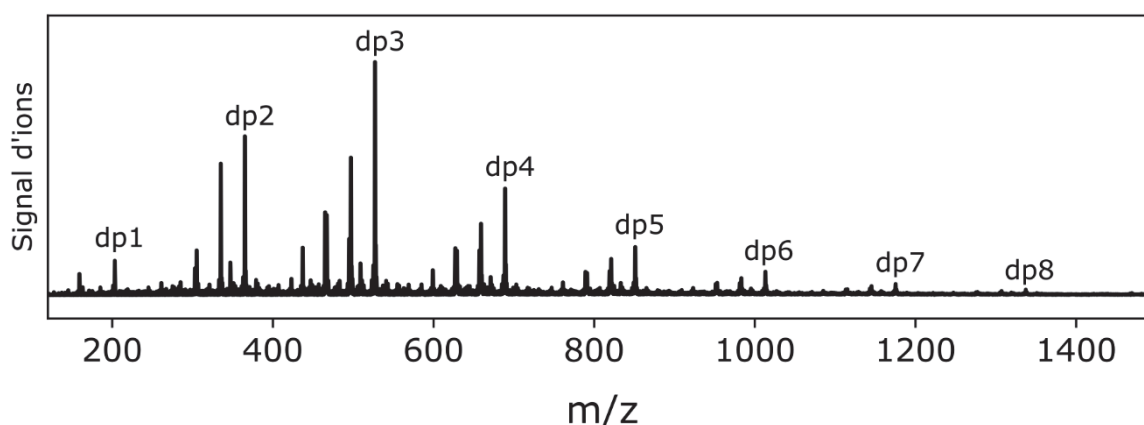


Figure 21 Spectres de masse après la dépolymérisation par le protocole FITDOG sur la gomme de caroube. Les oligosaccharides sont détectés complexés à un ion Na^+ .

En revanche, la MS ne fournit pas d'information sur la conservation des structures pendant les réactions chimiques. Pour vérifier cela, nous avons mené une étude par IRMPD.

3.2 Résultats IRMPD

Nous avons besoin de références pour vérifier une structure par IRMPD. Lors de la dépolymérisation des galactomannanes, les disaccharides obtenus correspondent à un mélange de $\text{Gal}\alpha(1-6)\text{Man}$ et $\text{Man}\beta(1-4)\text{Man}$. Or nous ne disposons pas de ces standards disaccharidiques, ni de ceux des oligosaccharides de taille supérieure. L'étude IRMPD a donc été menée sur les monosaccharides pour lesquels nous avons les références (Gal et Man).

Nous avons choisi la taille du composé, sélectionnons maintenant l'adduit. Les spectres IRMPD de complexe avec le sodium sont généralement congestionnés, il est préférable d'induire la formation d'autres complexes pour avoir des spectres IRMPD de meilleure qualité. Nous avons vu que les spectres IRMPD avec l'adduit NH_4^+ étaient résolus mais les complexes $\text{dp}+\text{NH}_4^+$ après ajout de sulfate d'ammonium n'ont pas pu être formés. La raison principale est l'utilisation de NaOH en grande quantité lors du protocole FITDOG. L'ammonium ajouté n'a donc pas pu remplacer le sodium déjà présent dans la solution. Nous avons testé deux alternatives (le lithium et le rubidium) pour lesquelles les complexes ont pu être formés par ajout de LiCl ou de RbCl.

Dans la première ligne de la Figure 22 sont tracés les références du galactose complexé au Li^+ (à gauche) ou au Rb^+ (à droite). Figurent sur la deuxième ligne les empreintes infrarouge du mannose avec ces mêmes adduits. Enfin, le galactomannane provenant de la gomme de caroube est constitué de 20% de galactose et de 80% de mannose. La mise au point d'une méthode de dosage par IRMPD

3. Dépolymérisation de polysaccharides modèles par FITDOG

avait déjà été effectuée, on sait que le spectre d'un mélange correspond aux spectres individuels corrélés par leur proportion.⁶ Nous avons donc simulé les spectres théoriques du mélange 20% galactose et 80% mannose à partir des spectres précédents. Ces spectres sont observables sur la dernière ligne en noir. Les spectres expérimentaux des monosaccharides issus de la dépolymérisation sont quant à eux représentés avec des aires colorées.

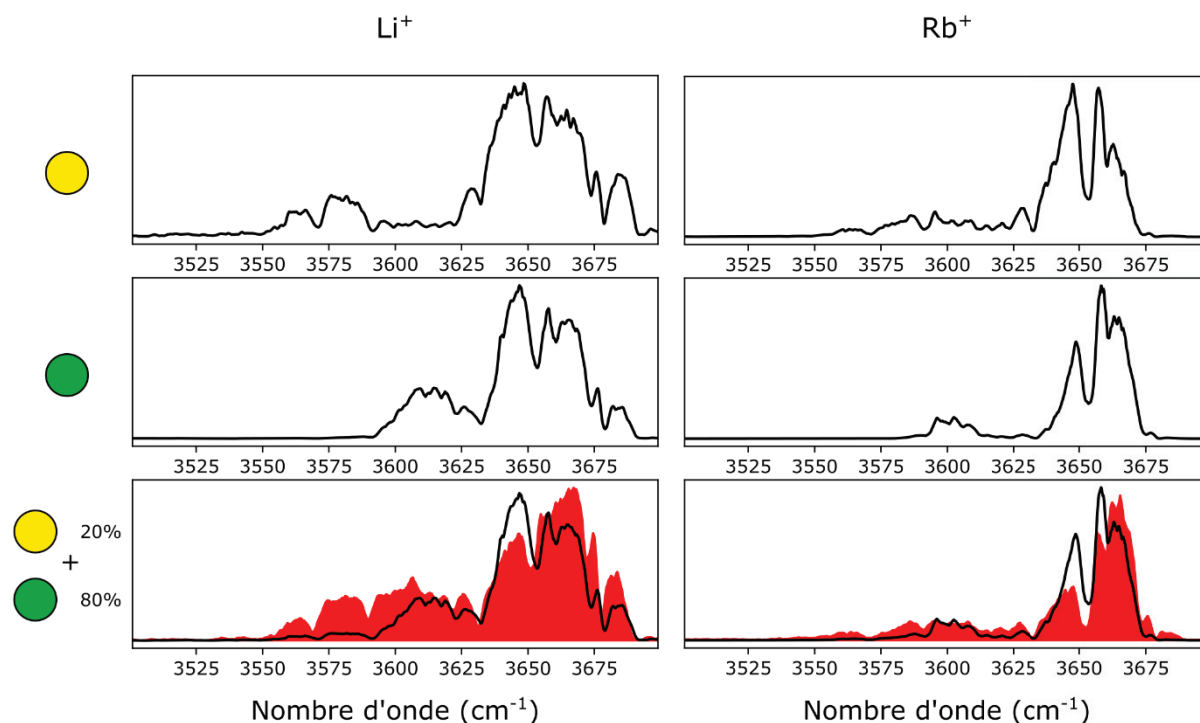


Figure 22 En ligne noire continue : spectres IRMPD du galactose (en haut), du mannose (au milieu) et du spectre simulé avec 20% de galactose et 80% de mannose (en bas). En couleur : spectre IRMPD du $\text{dp1}+\text{NH}_4^+$ de la gomme de caroube après réaction FITDOG. Les spectres à gauche ont été obtenus avec l'adduit Li^+ et à droite avec l'adduit Rb^+ .

Les positions des bandes correspondent entre les spectres sondés et simulés. Nous observons effectivement de l'absorption entre 3550 et 3690 cm^{-1} pour les quatre spectres et une absence d'absorption en dehors de cette zone. En revanche les intensités relatives sont très différentes. Pour l'adduit Li^+ , le spectre coloré en rouge présente de l'absorption plus intense que le spectre de référence de 3550 à 3530 cm^{-1} et de 3660 à 3690 cm^{-1} . La bande à 3640 cm^{-1} est au contraire plus faible. Il en est de même avec l'adduit Rb^+ pour lequel certaines bandes sont plus intenses (notamment à 3580 et 3665 cm^{-1}) et d'autres plus faibles (3650 et 3660 cm^{-1}) que le spectre simulé du mélange. Nous remarquerons de manière générale que les motifs des spectres sondés et de référence ne sont pas les mêmes. Ainsi la structure du monosaccharide obtenu par la dépolymérisation FITDOG de la gomme de caroube ne correspond pas à la composition en monosaccharide attendue.

L'explication la plus probable est un phénomène de réarrangement qui a pu se produire durant le procédé chimique. En effet, le protocole FITDOG est inspiré de la réaction de Fenton pour laquelle des réarrangements sont observés. Cette réaction chimique n'a pas été retenue pour l'application sur les polysaccharides de lichen.

4 Dépolymérisation de polysaccharides modèles par hydrolyse acide

La même étude a été menée sur les oligosaccharides obtenus par hydrolyse acide de la gomme de caroube. Les résultats sont présentés ci-dessous.

4.1 Résultats MS

La dépolymérisation par hydrolyse acide donne des oligosaccharides de différentes tailles, allant du monosaccharide au dp7 (Figure 23). Ils sont naturellement observés sous forme de complexe avec du Na^+ . La dépolymérisation par hydrolyse acide a donc fonctionné.

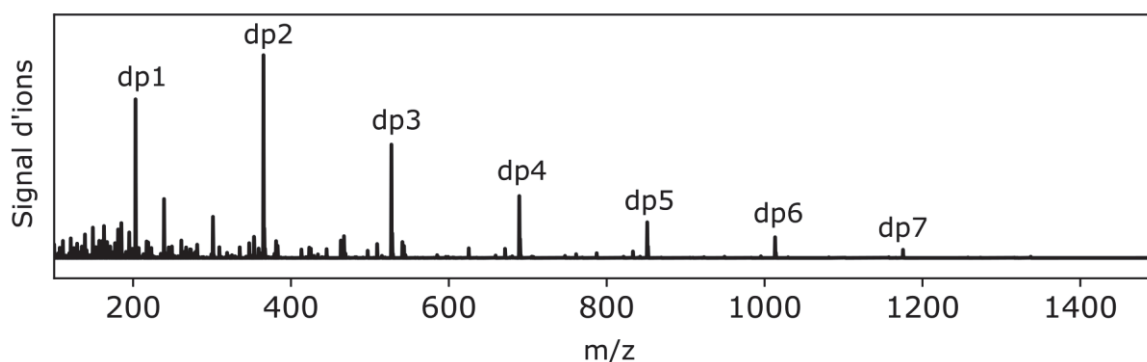


Figure 23 Spectres de masse après hydrolyse acide de la gomme de caroube. Les oligosaccharides sont détectés complexés à un ion Na^+ .

4.2 Résultats IRMPD

Pour les mêmes raisons que précédemment (page 130), nous analyserons les monosaccharides. À l'inverse des solutions FITDOG où du sodium se trouve en excès, la solution après hydrolyse acide ne contient pas ou peu de sodium. Il a ainsi été possible de former des complexes avec du NH_4^+ par ajout de sulfate d'ammonium. Les spectres IRMPD sont représentés en Figure 24 : le spectre du complexe $[\text{Gal}+\text{NH}_4]^+$ figure en haut et celui de $[\text{Man}+\text{NH}_4]^+$ au milieu. Le spectre simulé correspondant à 80% de $[\text{Gal}+\text{NH}_4]^+$ et à 20% de $[\text{Man}+\text{NH}_4]^+$ figure en bas, en ligne noire continue. Ce spectre est riche en bandes : la région des OH est composée d'une bande intense à 3640 cm^{-1} et d'une très faible bande à 3570 cm^{-1} , un motif constitué de deux maxima à 3350 et 3390 cm^{-1} se trouve dans la région des NH et une large et intense bande CH est observée de 2890 à 2980 cm^{-1} . Nous avons acquis le spectre IRMPD du monosaccharide issu de l'hydrolyse acide du galactomannane complexé au NH_4^+ . Sa signature IRMPD est également représentée en bas de la Figure 24 avec une aire colorée pour le différencier du spectre simulé. Nous pouvons noter qu'il possède des bandes aux mêmes positions que le spectre simulé et que les intensités relatives sont également comparables.

Nous avons vérifié grâce à l'IRMPD que les structures étaient conservées lors de l'hydrolyse acide.

4. Dépolymérisation de polysaccharides modèles par hydrolyse acide

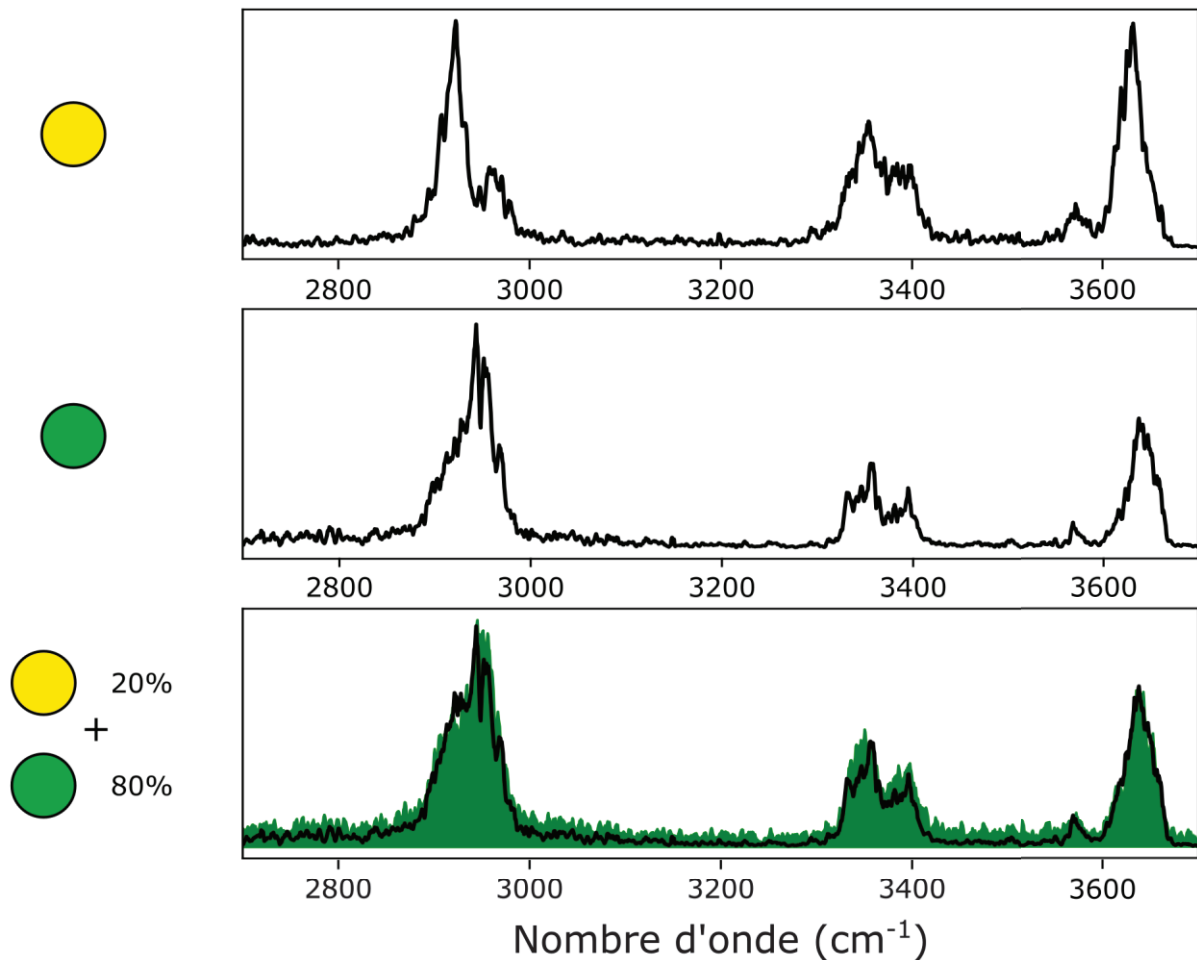


Figure 24 En ligne noire continue : spectres IRMPD du galactose+ NH_4^+ (en haut), du mannose+ NH_4^+ (au milieu) et du spectre simulé avec 20% de galactose et 80% de mannose avec NH_4^+ (en bas). En couleur : spectre IRMPD du $\text{dp1}+\text{NH}_4^+$ de la gomme de caroube après hydrolyse acide.

L'hydrolyse acide est ainsi plus prometteuse car elle préserve la structure des produits obtenus, ce qui n'est pas le cas de la méthode FITDOG. Outre la conservation de la structure qui est notre critère le plus important, l'hydrolyse acide présente d'autres avantages par rapport à FITDOG :

- Un meilleur rendement : pour une même quantité de polysaccharides, plus de produits sont obtenus par l'hydrolyse acide. Cette perte n'est pas critique pour nos polysaccharides modèles présents en abondance mais elle est à considérer pour les produits naturels disponibles en quantité limitée.
- Plus rapide et plus facile (moins d'étapes)
- Polyvalent : un autre polysaccharide (la cellulose) a été dépolymérisé par hydrolyse acide alors que la méthode FITDOG n'a pas fonctionné
- Moins toxique : un acide et d'une base sont nécessaires pour l'hydrolyse acide alors que FITDOG requiert également d'autres produits chimiques dont certains sont toxiques (NaBH_4).

Les polysaccharides de lichen ont donc été dépolymérisés par hydrolyse acide. Il est à noter que le protocole choisi par l'équipe de Rennes utilise du TFA et est décrit dans la référence [7]. Les oligosaccharides obtenus sont analysés dans les parties suivantes (une partie par espèce de lichen).

5 Cetraria islandica

5.1 Présentation

Également appelé mousse d'Islande, *Cetraria Islandica* (Figure 25) est un lichen très abondant poussant dans les régions froides. Une étude de ses polysaccharides a montré la présence de galactomannanes avec une chaîne principale de α -(1 \rightarrow 6)mannose.⁸ Les polysaccharides attendus sont de grande taille (>10 kDa).



Figure 25 *Cetraria islandica* (source photo : ISCR).

5.2 Résultats de dépolymérisation

Différents temps de réaction ont été appliqués pour l'hydrolyse acide : 2h, 6h, 8h ou 24h. Nous avons effectué le criblage uniquement sur les échantillons après 6h et 24h. Les spectres MS pour ces deux conditions sont en Figure 26 (après 6h d'hydrolyse en haut et après 24h en bas).

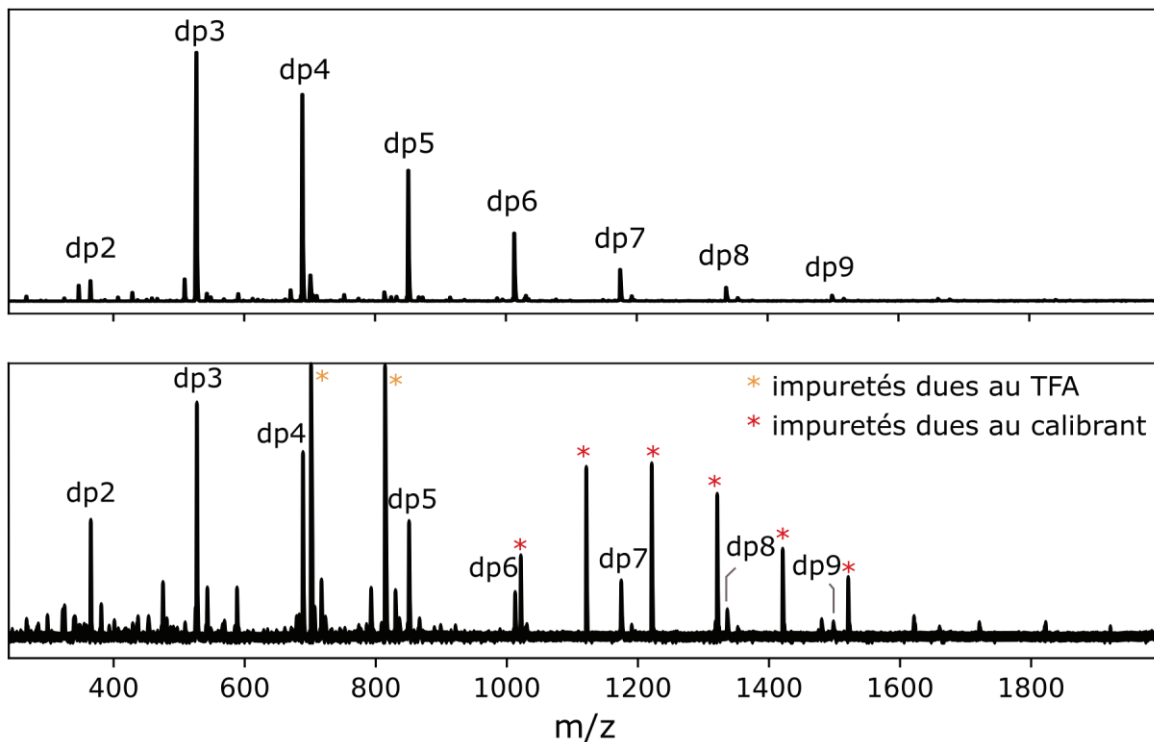


Figure 26 Profil de distribution des oligosaccharides selon leur rapport m/z après 6 h (en haut) ou 24 h (en bas) d'hydrolyse acide sur les polysaccharides de *Cetraria islandica*.

Des oligosaccharides de dp2 à dp9 sont visibles (complexés au Na^+). Leur distribution est centrée sur dp3 dans les deux cas. Signalons que cela ne signifie pas forcément que les trisaccharides sont en majorité dans la solution injectée en spectrométrie de masse. En effet, le signal d'un ion peut être optimisé en MS selon le rapport m/z attendu.

En revanche, les intensités relatives des spectres obtenus dans les mêmes conditions MS peuvent être comparées. Sachant que les paramètres d'optimisation n'ont pas été modifiés entre l'acquisition des deux spectres, le signal des dp2 plus important après 24h signifie que plus de disaccharides sont présents à des temps de réaction plus longs. Ce résultat est attendu car le nombre de liaisons glycosidiques rompues augmente avec le temps d'hydrolyse, donnant des oligosaccharides plus petits.

5.3 Séparation des oligosaccharides par LC-MS

Pour une taille donnée d'oligosaccharides, plusieurs composés sont attendus (des isomères se distinguant par leur séquence). Une séparation en chromatographie liquide a été réalisée pour les distinguer (Figure 27).

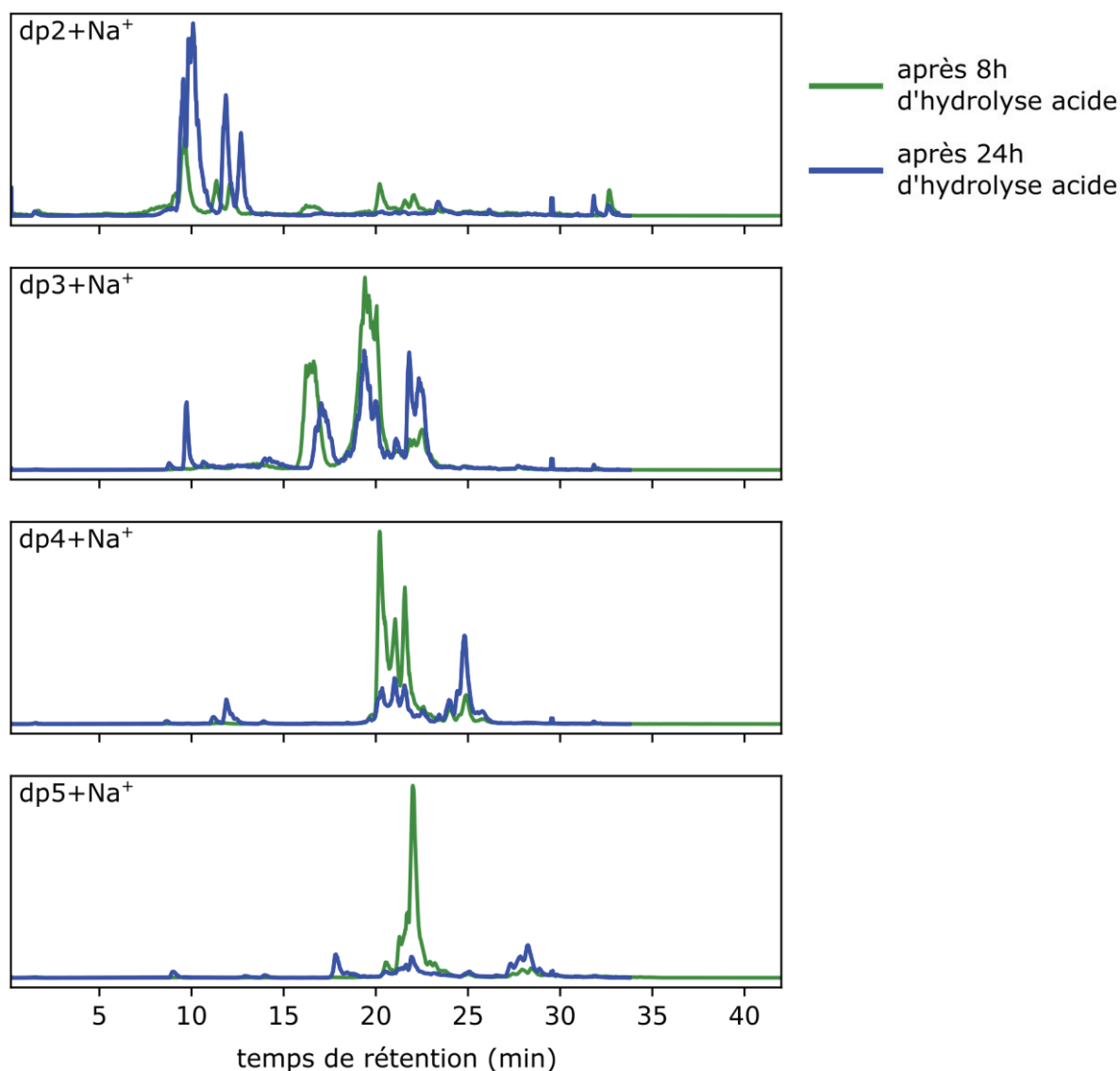


Figure 27 Chromatogrammes des dp2 (m/z 365), dp3 (m/z 527), dp4 (m/z 689) et dp5 (m/z 851) avec adduit Na^+ après hydrolyse acide des polysaccharides de *Cetraria islandica* pendant 8 h (en vert) ou 24 h (en bleu).

Trois dp2 sont observés après 8 h d'hydrolyse acide ($t_r=9,6$; 11,3 et 12,11 min). Ils sont également visibles, et plus intenses, après 24 h de réaction. Un léger décalage des temps de rétention est observé entre les deux chromatogrammes. D'autres dp2 sont détectés à 16,4 ; 20,2 ; 21,5 et 22 minutes seulement après 8 h de réaction mais ils proviennent probablement de la fragmentation en source de dp3 car leurs temps de rétention correspondent à ceux des trisaccharides. Les chromatogrammes des dp3 indiquent trois pics à 16,4 ; 19,6 et 22,1 minutes, d'intensité comparable pour les deux temps de réaction. Plusieurs dp4 et dp5 sont détectés entre 20 et 30 minutes.

La reproductibilité des résultats avec les colonnes Hypercarb est un problème souvent rencontré dans la littérature.⁹ Nous pouvons donc observer des écarts de rétention d'une injection à l'autre.

En comparant les oligosaccharides d'une taille donnée après des temps de réaction différents, nous notons que les espèces observées sont les mêmes (les temps de rétention sont très proches) mais avec des intensités différentes. Comme précisé plus tôt, les plus petits oligosaccharides sont plus intenses lorsque la durée de l'hydrolyse acide est plus longue. La comparaison des temps de rétention des oligosaccharides de différentes tailles montre que les oligosaccharides les plus longs sont retenus plus longtemps dans la colonne.

Dans le chapitre précédent, la méthode de criblage a été développée sur les disaccharides chargés au lithium. Par conséquent, 10% de l'eau de la phase mobile a été remplacée par du chlorure de lithium pour induire des complexes au Li^+ . L'échantillon correspondant aux polysaccharides de *Cetraria islandica* après 24 heures d'hydrolyse acide est injecté. Le chromatogramme des dp2 chargés au lithium (détectés à m/z 349) est représenté dans la Figure 28. Nous pouvons noter la présence de quatre pics chromatographiques entre 10 et 17 minutes. La charge influence donc peu la rétention car les dp2 au Na^+ présentent des temps de rétention proches. En revanche, un pic supplémentaire est observé vers la 11^e minute. Cela indique que les deux composés n'ont pas été séparés sous adduit Na^+ mais qu'ils le sont avec l'adduit Li^+ . Les deux pics suivants vers 15 minutes correspondent à deux pics également présents dans le chromatogramme des $[\text{dp2}+\text{Na}]^+$.

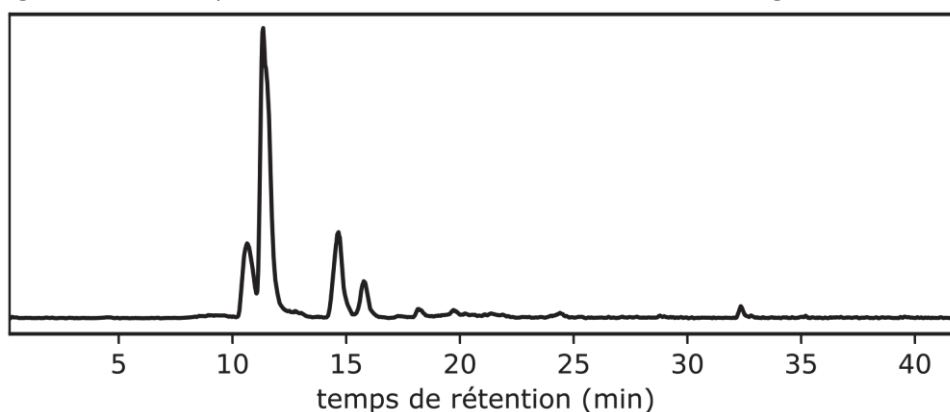


Figure 28 Chromatogrammes des dp2 avec adduit Li^+ (m/z 349) après hydrolyse acide des polysaccharides de *Cetraria Islandica* pendant 24 h.

5.4 Criblage des disaccharides avec le laser Firefly seul

L'échantillon a été passé dans les mêmes conditions chromatographiques (avec du lithium dans la phase mobile) en ajoutant une irradiation laser Firefly à 3650 cm^{-1} . Le chromatogramme se trouve en haut de la Figure 29. Un premier pic est présent à la 13^e minute puis un second beaucoup plus intense est observé à 15 minutes. Enfin, deux pics mal résolus apparaissent vers 19 minutes. Bien que le nombre de pics et leur ordre soient cohérents avec les résultats précédents (Figure 28), nous pouvons

5. Cetraria islandica

voir que les temps de rétention sont significativement différents (différence de 5 min). Cela s'explique par la mauvaise répétabilité des colonnes Hypercarb.

Rappelons que pour la stratégie de criblage, l'ion précurseur correspond au disaccharide observé à m/z 349. Celui-ci est fragmenté par CID et le fragment m/z 169 (correspondant aux fragments b_1 et z_1) est isolé et irradié dans le piège. Des photofragments (de m/z 50 à m/z 165) peuvent alors être observés. Le signal du fragment m/z 169 est représenté par une ligne continue bleue. Le signal des photofragments est affichée avec une aire colorée rouge. L'intensité des photofragments est élevé pour les deux premiers pics, mais paraît faible pour les deux autres pics. Le taux de photofragmentation a été calculé pour chaque composé et est représenté en bas de la Figure 29 en fonction du temps de rétention. Le seuil, fixé à 0,2 dans ces conditions laser, est représenté par une ligne en pointillés. Seul le pic à la 15^e minute présente un taux de photofragmentation supérieur à ce seuil.

En somme, quatre fractions de disaccharides ont été séparées par LC dont l'une répond à notre critère de discrimination choisi avec le laser Firefly. Cette fraction pourrait donc contenir du galactofuranose en extrémité non réductrice.

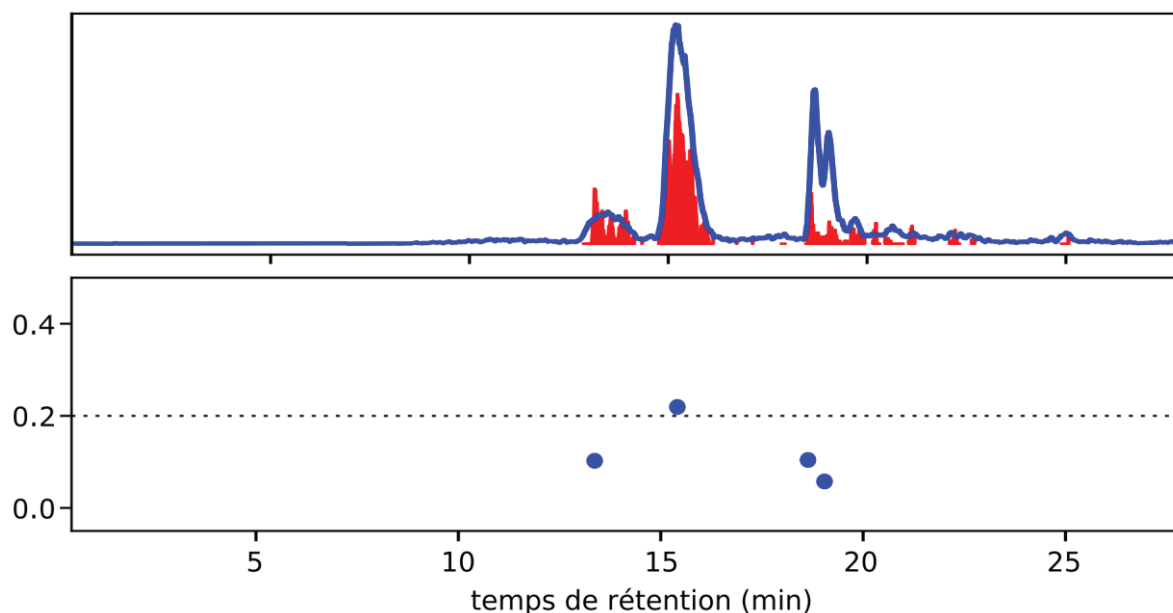


Figure 29 Criblage par laser Firefly à 3650 cm^{-1} des disaccharides issus de l'hydrolyse acide pendant 24h des polysaccharides de *Cetraria islandica*. En haut : la ligne continue correspond au signal de l'ion fragment m/z 169 et l'aire colorée à l'intensité normalisée des photofragments. En bas : ratio de photofragmentation des composés détectés. La valeur seuil fixée à 0,2 apparaît en pointillés.

La stratégie de criblage a été appliquée avec des conditions laser différentes.

5.5 Criblage des disaccharides avec le laser CO2 seul

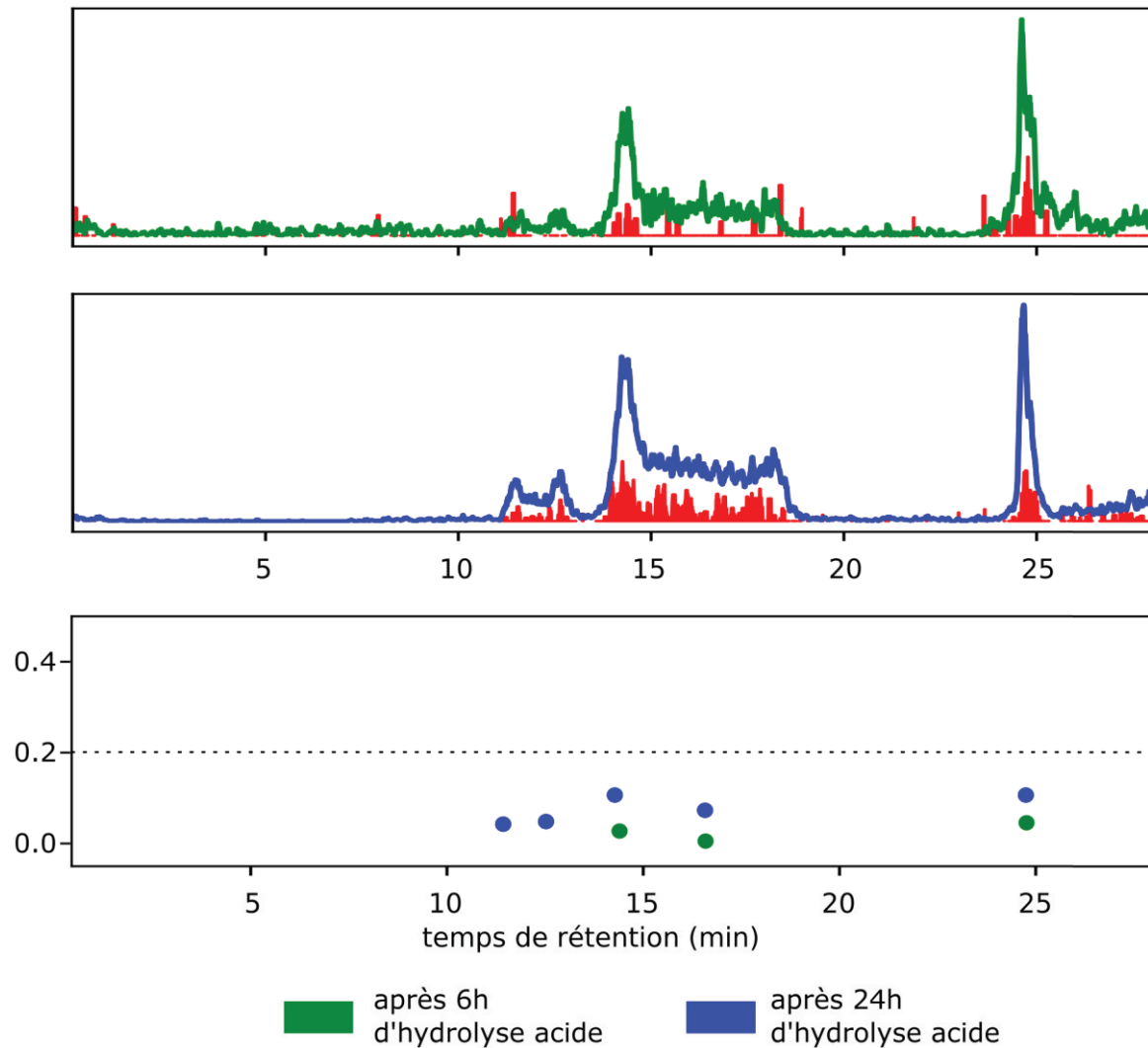
Les échantillons de *Cetraria islandica* dépolymérisés ont été analysés par criblage avec le laser CO2. Nous avons testé deux échantillons avec des durées d'hydrolyse différentes (6 h et 24 h). La séparation LC est effectuée avec une colonne différente par rapport à la partie précédente, les profils chromatographiques sont donc différents.

Les résultats sont en Figure 30, ceux provenant des polysaccharides hydrolysés pendant 6 heures sont représentés en vert et ceux de l'échantillon après 24 heures d'hydrolyse apparaissent en bleu.

Le chromatogramme vert indique un pic d'éluion intense à 13,9-14,7 min suivi d'un long pic d'intensité constante jusqu'à 18,3 minutes. Un autre pic suit à 24,3-25,5 minutes. L'augmentation de la durée d'hydrolyse à 24 h (chromatogramme bleu) révèle deux pics supplémentaires à 11,5 et 12,6 minutes.

L'intensité des photofragments est représentée par une aire rouge sur les deux chromatogrammes. Elle semble visuellement assez faible, les taux de photofragmentation sont calculés pour mieux évaluer. Ces ratios sont représentés en bas de la figure, avec le même code couleur (en vert : après 6 h d'hydrolyse acide, en bleu : après 24 h d'hydrolyse acide). Ils sont tous bien en dessous du seuil.

Par conséquent, cette méthode de criblage ne semble pas détecter des fractions contenant du galactofuranose dans les échantillons de *Cetraria islandica*.



6 *Lasallia pustulata*

6.1 Présentation

L'espèce *Lasallia pustulata* (Figure 31), c'est-à-dire le lichen à pustules, se trouve principalement sur des roches riches en nutriments. Une étude montre la présence de différents polysaccharides dont des pustulanes, glucanes et galactomannanes. Ces derniers possèdent une chaîne principale de α -(1 \rightarrow 6)mannose avec des substitutions de galactofuranose en position O2.¹⁰ Des polysaccharides de grande taille sont attendus (680 000 Da).



Figure 31 Photographie d'extraits de *Lasallia pustulata*.

6.2 Résultats de dépolymérisation

Nous obtenons après 6h d'hydrolyse acide des polysaccharides de *Lasallia pustulata* des oligosaccharides de dp2 à dp8, avec une distribution centrée sur dp3. Le spectre de masse montrant la distribution en taille des oligosaccharides est représenté en Figure 32.

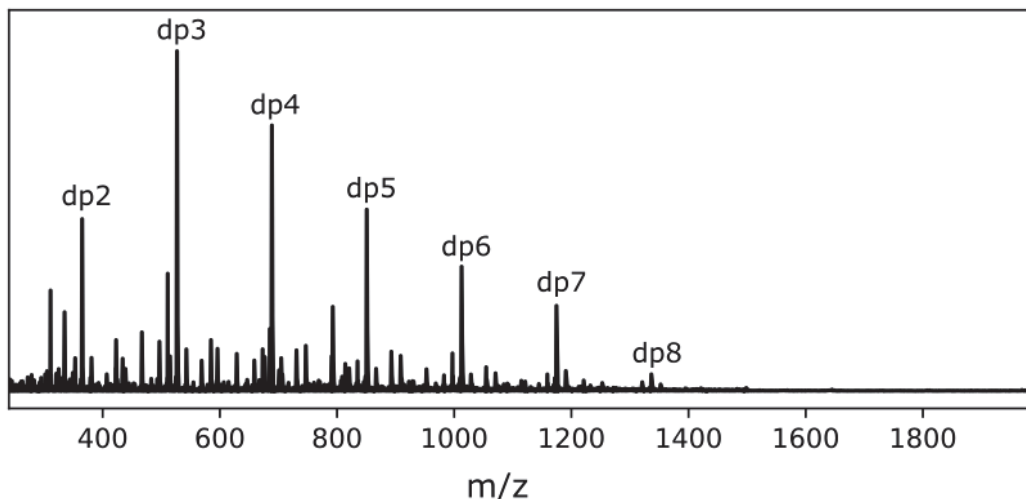


Figure 32 Profil de distribution des oligosaccharides selon leur rapport m/z après 6h d'hydrolyse acide sur les polysaccharides de *Lasallia pustulata*.

La méthode de criblage LC-MS/MS-IR avec le laser Firefly a été appliquée sur les disaccharides issus de *Lasallia pustulata*.

6.3 Criblage des disaccharides avec le laser Firefly seul

Deux échantillons en provenance de *Lasallia pustulata* ont été criblés. Dans le premier échantillon, les polysaccharides ont été dépolymérisés par hydrolyse acide pendant 6 h. Dans le second, la réaction a duré 24 heures. Les résultats du premier échantillon figurent en vert et ceux du deuxième échantillon en bleu (Figure 33).

Nous notons trois pics mal séparés sur le chromatogramme vert avec des temps de rétention de 8,8-9,3 ; 9,7-11 et 11-12,8 minutes. Ces composés sont irradiés à 3650 cm^{-1} . Le signal en photofragments résultant, représenté en rouge, est très intense pour cet échantillon. Les taux de photofragmentation sont calculés pour chaque pic d'élution et sont représentés en bas de la figure.

Le ratio du premier pic est évalué à 0,02, les deux autres pics ont des taux de 0,27 et 0,24 donc supérieurs au seuil (0,2). Il y a ainsi deux fractions sur trois qui répondent au critère de criblage. Des efforts sont nécessaires pour mieux séparer ces composés, ce qui est indispensable pour leur identification.

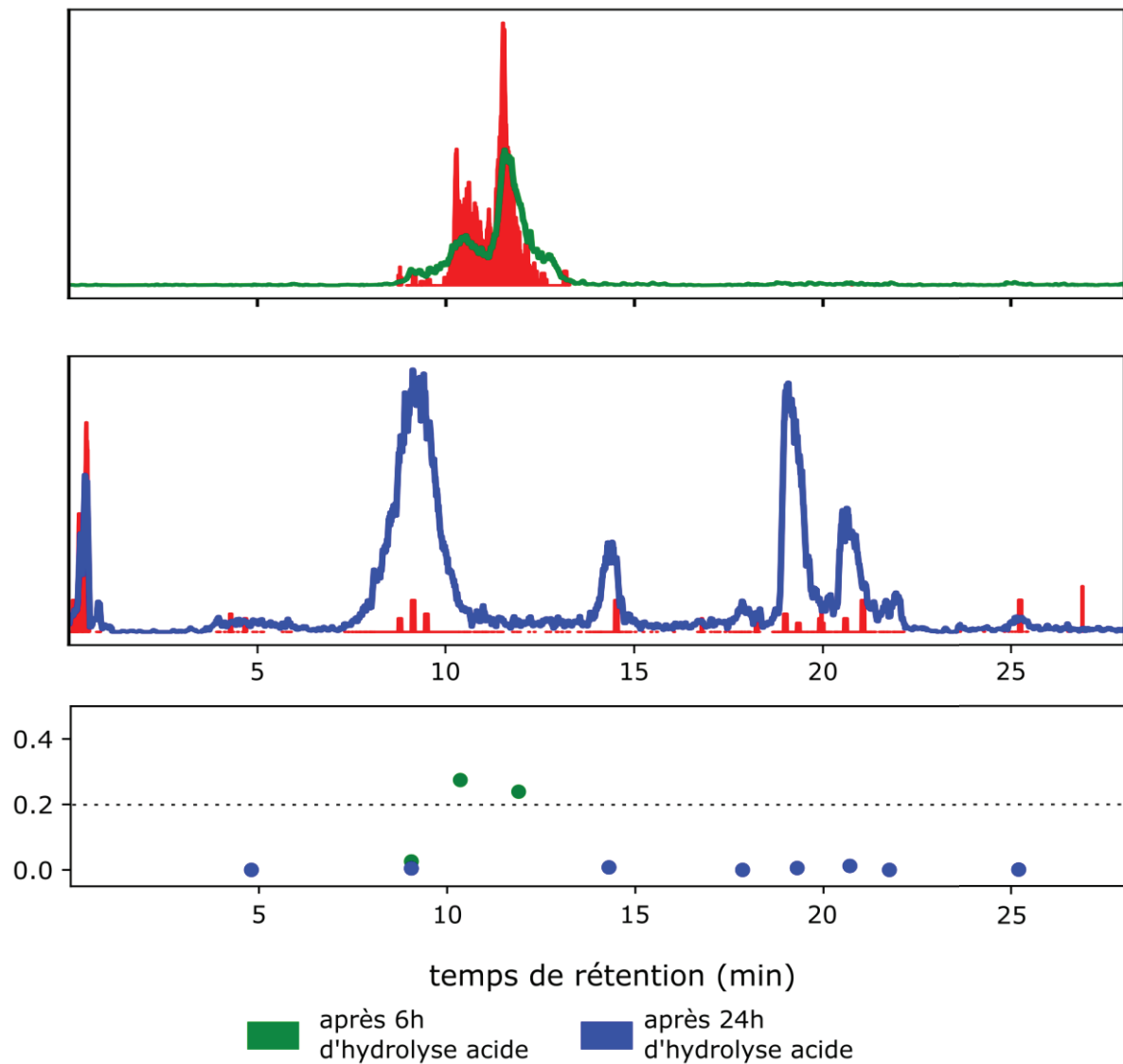


Figure 33 Criblage par laser Firefly des disaccharides issus de l'hydrolyse acide pendant 6 h (en vert) ou 24 h (en bleu) des polysaccharides de *Lasallia pustulata*. En haut : la ligne continue correspond au signal de l'ion fragment m/z 169 et l'aire colorée à l'intensité normalisée des photofragments. En bas : ratio de photofragmentation des composés détectés. La valeur seuil fixée à 0,2 apparait en pointillés.

6. *Lasallia pustulata*

Les disaccharides obtenus après 24 h d'hydrolyse acide des polysaccharides de *Lasallia pustulata* ont également été analysés. Les pics du chromatogramme sont plus abondants et une meilleure séparation est observée. Il y a huit pics dont les périodes de rétention sont : [3,8;5,8], [7,9;10,2], [13,9;14,7], [17,6;18,1], [18,8;19,8], [20,3;21,1], [21,5;22] et [25;25,4] minutes. L'intensité des photofragments figure en rouge, une absence de photofragmentation peut être notée tout au long du chromatogramme (le peu de signal correspond à du bruit de fond). Les taux calculés valident cette constatation, les huit composés ont un taux de photofragmentation proche de 0.

7 Roccella fuciformis

7.1 Présentation

La troisième espèce étudiée est *Roccella fuciformis* (Figure 34) et se trouve très souvent en bord de mer. Les lichens *Roccella* possèdent des galactomannanes avec une chaîne principale de α -(1 \rightarrow 4)mannose. Des substitutions d'unités β -galactofuranose en O6 ont été observées.¹¹ Les polysaccharides attendus sont de faible taille (3300 Da).



Figure 34 Photographie d'extraits de *Roccella fuciformis*.

7.2 Résultats de dépolymérisation

Les temps de réaction pour l'hydrolyse acide ont été raccourci pour cette espèce (15 min, 45 min, 2 h ou 3 h) en raison du faible poids moléculaire des polysaccharides. Après 3 heures d'hydrolyse, des oligosaccharides de dp2 à dp10 sont observés, avec un très fort signal en dp3 (Figure 35).

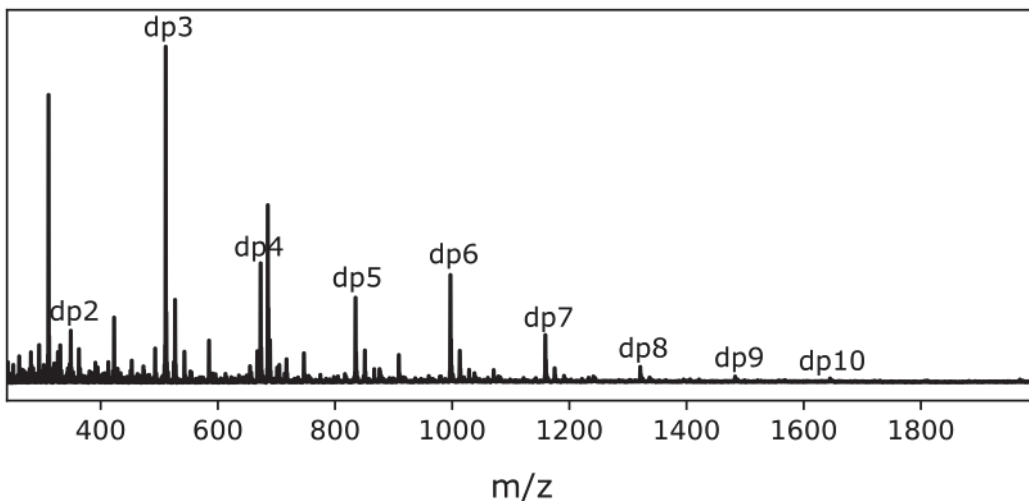


Figure 35 Profil de distribution des oligosaccharides selon leur rapport m/z après 3 h d'hydrolyse acide sur les polysaccharides de *Roccella fuciformis*.

7.3 Criblage des disaccharides avec laser Firefly seul

L'échantillon contenant les polysaccharides de *Roccella fuciformis* hydrolysés pendant deux heures a été injecté. Le chromatogramme obtenu est représenté en Figure 36. Deux fractions sont distinguées : un premier pic large et mal résolu entre 11,1 et 13,6 minutes puis un second entre 18,6 et 20 minutes. Ces composés sont irradiés avec le laser Firefly émettant à 3650 cm^{-1} . L'intensité des photofragments peut être visualisée par le signal en rouge. Elle est plutôt élevée pour le premier composé et plus faible pour le second. Le calcul des taux de photofragmentation indique que le premier pic peut contenir du galactofuranose car le ratio est au-dessus de 0,2. En revanche, le taux du second composé ne dépasse pas la valeur spécifiée.

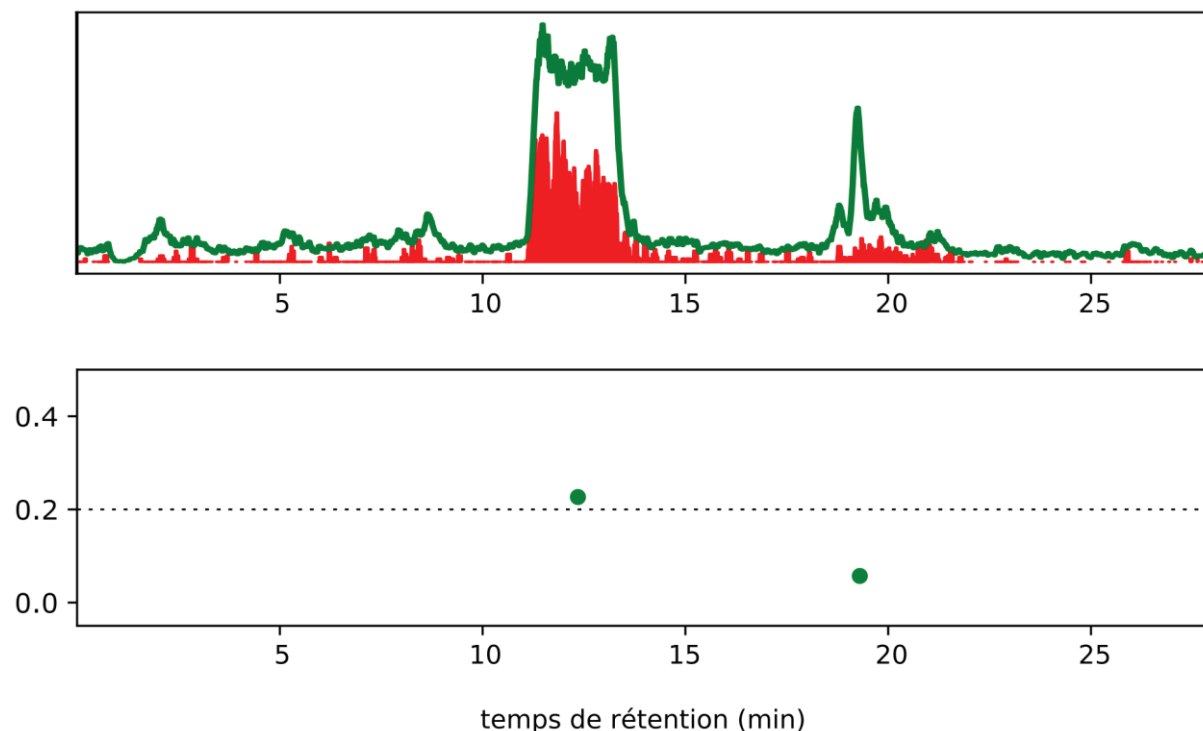


Figure 36 Criblage par laser Firefly des disaccharides issus après 2h d'hydrolyse des polysaccharides de *Roccella fuciformis*. En haut : la ligne continue correspond au signal de l'ion fragment m/z 169 et l'aire colorée à l'intensité normalisée des photofragments. En bas : ratio de photofragmentation des composés détectés. La valeur seuil fixée à 0,2 apparaît en pointillés.

7.4 Criblage des disaccharides avec laser CO2 seul

La stratégie de criblage avec le laser CO₂ a été appliquée sur l'échantillon *Roccella fuciformis* après 3 h d'hydrolyse. Sur le chromatogramme (en haut de la Figure 37), deux pics mal séparés sont représentés à 11-12,5 minutes et à 13,6-14,2 minutes. Un pic intense est ensuite observé à 24,3-25,6 minutes, suivi d'un pic faible jusqu'à la fin du chromatogramme. L'intensité absolue des photofragments (en rouge) est élevée pour les trois premiers pics. Nous devons en revanche corréliser ce signal avec celui de leur ion parent (dont l'intensité est représentée en bleu). Le taux de photofragmentation permet de prendre en compte l'intensité de l'ion parent. Par ordre de rétention, il a été calculé à 0,16 ; 0,14 ; 0,16 et 0,09 (voir les marqueurs ronds en bas de la figure). À l'exception du quatrième pic, les taux de photofragmentation sont très proches de 0,2. Pour cette raison, nous avons effectué la même analyse avec une seconde injection. Le chromatogramme est représenté au milieu de la figure. Le chromatogramme de l'ion fragment (en bleu) montre les quatre pics aux mêmes positions et avec une intensité comparable aux résultats précédents. L'intensité des photofragments (en rouge) semble également similaire. Les taux de photofragmentation ont été calculés et sont représentés par des croix

en bas de la figure. Ils sont par ordre d'apparition des pics de 0,24 ; 0,2 ; 0,14 et 0,11. Les deux premiers pics ont donc un taux dépassant le seuil pour cette seconde analyse.

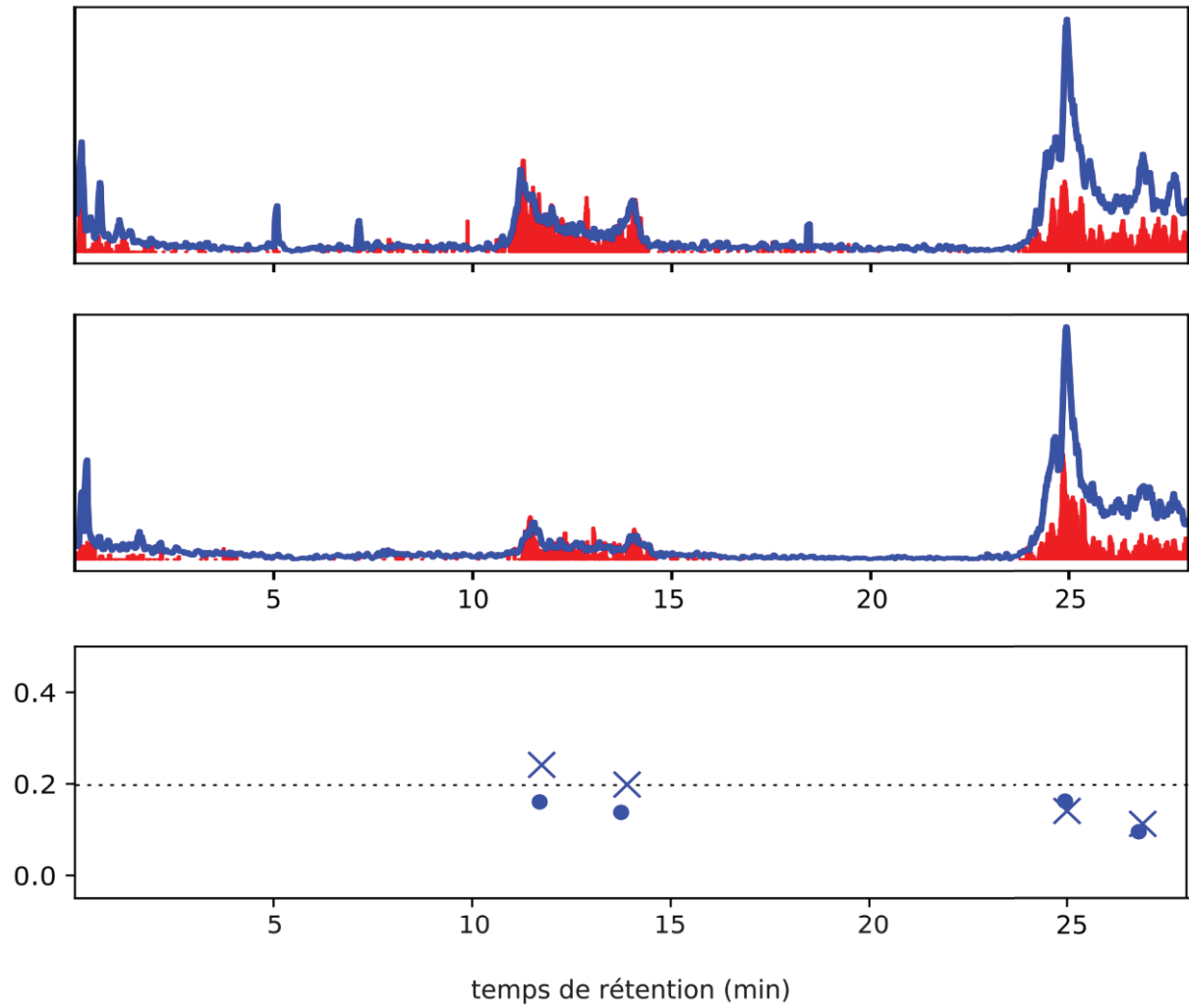


Figure 37 Criblage par laser CO₂ des disaccharides issus après 3 h d'hydrolyse des polysaccharides de *Rocella fuciformis*. En haut : la ligne continue correspond au signal de l'ion fragment m/z 169 et l'aire colorée à l'intensité normalisée des photofragments. En bas : ratio de photofragmentation des composés détectés (des ronds pour la 1^{ère} analyse et des croix pour la 2^{nde} analyse). La valeur seuil fixée 0,2 apparait en pointillés.

8 Discussion et perspectives

En somme, nous avons trois sources d'information pour connaître les espèces de lichen synthétisant du galactofuranose : la littérature, les résultats RMN de nos collaborateurs et nos analyses en IRMPD. Nous pouvons affirmer dans un premier temps que la littérature sur les polysaccharides de lichen est faible et obsolète. Les données que nous avons trouvées sur *Cetraria islandica* datent par exemple de 1984 et ne signalent pas la présence de galactofuranose.⁸ Cette espèce avait donc été sélectionnée comme témoin négatif mais les résultats RMN de nos collaborateurs ainsi que les nôtres valident la présence de galactofuranose. À l'inverse, *Lichina pygmaea* contient du galactofuranose d'après la bibliographie¹² mais les résultats RMN n'en ont pas détecté (les polysaccharides de cette espèce n'ont pas été fournis et donc n'ont pas été analysés par IRMPD). Les différentes méthodes peuvent être comparées pour valider ou non la présence de galactofuranose. Par ailleurs, les données acquises durant le projet ont permis d'enrichir la base d'informations dans la littérature.

Tableau 6 Récapitulatif des résultats de criblage selon l'espèce de lichen, la durée de l'hydrolyse acide et les conditions laser.

lichen	durée de l'hydrolyse acide (en heures)	détection de furanose par criblage avec le laser Firefly		détection de furanose par criblage avec le laser CO2	
		oui	non	oui	non
<i>Cetraria islandica</i>	6	-	-	0	0
	24	1	3	0	0
<i>Lasallia pustulata</i>	6	2	0	-	-
	24	0	8	-	-
<i>Rocella fuciformis</i>	2	1	1	-	-
	3	-	-	2	2

La présence de galactofuranose dépend ainsi de l'espèce de lichen selon sa capacité ou non à synthétiser cette forme. Puis nous avons vu qu'elle était également liée à la préparation d'échantillon. La durée de l'hydrolyse acide altère la structure du galactofuranose. Par exemple, nous en avons détecté pour l'espèce *Lasallia pustulata* après 6 h de réaction et plus après 24 h de réaction (voir Tableau 6). Les analyses par RMN sur cette espèce suggèrent des résultats similaires : le signal en galactofuranose disparaît si la durée d'hydrolyse acide est trop longue.

Enfin, notre méthode de criblage révèle pour un échantillon donné (même espèce et même durée d'hydrolyse acide) certaines fractions contenant du galactofuranose et d'autres non. Par exemple, comme l'indique le Tableau 6, quatre fractions de *Cetraria islandica* après 24 h de réaction dont une seule contient du galactofuranose. Cela s'explique par les différentes combinaisons de disaccharides obtenues après dépolymérisation. Ceux qui contiennent du galactofuranose en extrémité non réductrice (les différents régioisomères Gal β -Man) répondent au criblage tandis les autres (les isomères Man-Man, Gal p -Man, etc.) n'y répondent pas.

Par ailleurs, des fractions contenant du galactofuranose sont détectées avec le Firefly à 3650 cm⁻¹ mais pas avec le laser CO2. Lors du développement de la méthode de criblage, nous avons également testé le couplage des deux lasers (à 3410 cm⁻¹) qui a donné le meilleur contraste entre disaccharides contenant ou non du galactofuranose. Il est donc intéressant et plus robuste d'effectuer le criblage dans ces conditions avec lesquelles les meilleures performances sont obtenues. Cela n'a pas pu être réalisé dans le cadre du projet.

Une autre perspective est le travail à réaliser en chromatographie liquide. La mauvaise répétabilité des colonnes Hypercarb ainsi que notre manque d'expérience sur le développement de méthodes LC sont des points faibles. Malgré cela, une bonne résolution a été obtenue pour certains chromatogrammes. Des efforts doivent être faits pour séparer les composés mal résolus. Cette optimisation est indispensable pour mener à bien la dernière étape de la démarche, c'est-à-dire l'identification des composés discriminés.

En effet, une fois les séparations optimales effectuées, nous proposons d'identifier les composés pouvant retenir du galactofuranose. En première intention, nous pouvons acquérir le spectre IRMPD de l'ion précurseur (le disaccharide) pour le comparer à notre base de données (étape 7 du schéma page 124). Si l'identification n'est pas possible, une stratégie de séquençage pourra être appliquée. Les deux monosaccharides identifiés séparément permettront de déterminer la structure du disaccharide.

Les développements instrumentaux effectués ces dernières années permettent l'acquisition de spectres IRMPD en moins d'une minute.¹³ Il est envisageable d'effectuer cette dernière étape en LC-MS-IR. Si la durée du pic est trop courte, les fractions d'intérêt peuvent être collectées (voir page 21) puis injectées en infusion directe pour une analyse en MS-IR.

9 Bibliographie

- 1 F. Cuskin, E. C. Lowe, M. J. Temple, Y. Zhu, E. Cameron, N. A. Pudlo, N. T. Porter, K. Urs, A. J. Thompson, A. Cartmell, A. Rogowski, B. S. Hamilton, R. Chen, T. J. Tolbert, K. Piens, D. Bracke, W. Vervecken, Z. Hakki, G. Speciale, J. L. Munõz-Munõz, A. Day, M. J. Peña, R. McLean, M. D. Suits, A. B. Boraston, T. Atherly, C. J. Ziemer, S. J. Williams, G. J. Davies, D. W. Abbott, E. C. Martens and H. J. Gilbert, *Nature*, 2015, **517**, 165–169.
- 2 F. A. de Moura, F. T. Macagnan and L. P. da Silva, *International Journal of Food Science & Technology*, 2015, **50**, 275–281.
- 3 M. J. Amicucci, E. Nandita, A. G. Galermo, J. J. Castillo, S. Chen, D. Park, J. T. Smilowitz, J. B. German, D. A. Mills and C. B. Lebrilla, *Nat Commun*, 2020, **11**, 3963.
- 4 J. Boustie *et al.*, Depolymerisation of polysaccharide with conservation of galactofuranosyl units (en rédaction).
- 5 O. Yeni *et al.*, Screening using LC-MS/MS-IR to detect galactofuranose (en rédaction).
- 6 L. Barnes, A.-R. Allouche, S. Chambert, B. Schindler and I. Compagnon, *International Journal of Mass Spectrometry*, 2020, **447**, 116235.
- 7 Strategy for Structural Elucidation of Polysaccharides: Elucidation of a Maize Mucilage that Harbors Diazotrophic Bacteria | Analytical Chemistry, <https://pubs.acs.org/doi/10.1021/acs.analchem.9b00789>, (accessed September 28, 2022).
- 8 P. A. J. Gorin and M. Iacomini, *Carbohydrate Research*, 1984, **128**, 119–132.
- 9 T. E. Bapiro, F. M. Richards and D. I. Jodrell, *Anal. Chem.*, 2016, **88**, 6190–6194.
- 10 M. T. Pereyra, A. Prieto, M. Bernabé and J. A. Leal, *The Lichenologist*, 2003, **35**, 177–185.
- 11 E. R. Carbonero, L. M. C. Cordeiro, C. G. Mellinger, G. L. Sassaki, E. Stocker-Wörgötter, P. A. J. Gorin and M. Iacomini, *Carbohydr Res*, 2005, **340**, 1699–1705.
- 12 A. Prieto, J. A. Leal, M. Bernabé and D. L. Hawksworth, *Mycological Research*, 2008, **112**, 381–388.
- 13 O. Yeni, B. Schindler, B. Moge and I. Compagnon, *Analyst*, 2022, **147**, 312–317.

Conclusion et perspectives

L'activité de spectroscopie IRMPD sur les sucres a été instaurée à l'Institut Lumière Matière en 2013 par Isabelle Compagnon. Un travail conséquent a alors été effectué, notamment durant les thèses de Baptiste Schindler et Gina Renois Predelus. D'abord, l'IRMPD n'étant pas une méthode commerciale, il a fallu installer le dispositif expérimental et développer un logiciel pour l'acquisition et l'interprétation des données. Ensuite, des travaux de recherche fondamentale ont été effectués pour démontrer l'intérêt d'utiliser l'IRMPD en glycomique. La complexité des échantillons a progressivement augmenté avec d'abord des monosaccharides standards (glucosamine...), puis des standards de plus grande taille (disaccharides, trisaccharides...), jusqu'à l'analyse de produits naturels commerciaux (glycopeptide de poulet, amidon...). Cela a nécessité le développement d'une méthode innovante : le séquençage des sucres par MS/MS-IR. Le dispositif a sans cesse été amélioré dans le but de s'étendre à de la recherche appliquée, notamment par l'ajout d'une chromatographie liquide au dispositif instrumental. Dans le même but, des premiers travaux ont été effectués pour augmenter la rapidité d'acquisition de spectre IRMPD.

Cette thèse s'instaure dans la continuité de ces travaux. Des efforts ont été fournis pour le développement de nouveaux outils. Les méthodes existantes et nouvellement développées ont été valorisées durant cette thèse dont l'objectif est de satisfaire un besoin analytique. À partir de ce qui est disponible dans la nature (des lichens), nous répondons à une question biologique par IRMPD (détection et caractérisation d'oligosaccharides contenant du galactofuranose). Cela requiert une approche globale et interdisciplinaire. Le projet ALGAIMS réunit des experts de différents domaines (chimistes, lichénologues, massistes) et se divise en deux parties.

La première partie du projet a été menée sur des produits synthétiques afin de développer les méthodes expérimentales. Chaque molécule synthétisée par nos collaborateurs chimistes a été analysée par IRMPD. D'une part, ces données nous ont permis de vérifier la préservation de l'information sur la taille du cycle en spectrométrie de masse (en modes ESI-MS et ESI-MS/MS). D'autre part, les signatures IRMPD acquises sur les standards (en MS-IR) ont servi à créer une base de données afin de pouvoir identifier les produits naturels. La conservation de l'information après fragmentation CID nous a permis de développer trois stratégies pour l'analyse d'oligosaccharides. La première concerne l'analyse en fragment monosaccharidique (en MS/MS-IR) pour révéler la taille du cycle. La seconde est l'étude des fragments disaccharidiques et renseigne, en plus de l'information sur la taille du cycle, la régiochimie de la liaison. Enfin, le criblage en LC-MS/MS-IR permet de discriminer les oligosaccharides contenant du galactofuranose. Nous pouvons ainsi choisir la stratégie adéquate selon l'information désirée.

La seconde partie du projet consiste en l'analyse de différentes espèces de lichen sélectionnées par nos collaborateurs lichénologues : *Lasallia pustulata*, *Cetraria islandica* et *Roccella fuciformis*. Ils ont effectué les étapes de collecte et d'extraction des polysaccharides. L'IRMPD a été utile dans le choix de la méthode de dépolymérisation. Les oligosaccharides issus de la dépolymérisation ont été analysés en MS. Nous avons appliqué la stratégie de criblage sur ces échantillons naturels. Ainsi notre démarche en plusieurs étapes a été effectuée sur différentes espèces de lichen pour identifier les oligosaccharides contenant du galactofuranose. Les résultats illustrent le potentiel de notre démarche.

Conclusion

Pour pousser plus loin l'approche analytique, il serait intéressant d'identifier ces oligosaccharides discriminés. Cela nécessite de développer une stratégie basée sur l'acquisition de données LC-MS-IRMPD. Nos récents développements instrumentaux permettent l'acquisition d'un spectre IRMPD en moins d'une minute, ouvrant accès à la stratégie LC-MS-IRMPD. Une alternative est de collecter les fractions d'intérêt pour une analyse ultérieure en MS-IR (stratégie plus longue). Dès lors cette étape franchie, on pourra alors proposer aux biologistes un protocole complet pour discriminer et identifier au détail près les oligosaccharides avec du galactofuranose.

En parallèle, nous avons analysé le galactofuranose par une seconde approche qui consiste à caractériser, avec l'aide d'outils théoriques, ses préférences conformationnelles en phase gazeuse. Il en résulte que la phase gazeuse réduit l'espace conformationnel du galactofuranose et stabilise sa structure. Nous avons détaillé les différentes conformations à l'échelle de l'ångström et ce dans différentes conditions de micro-environnement.

Pour approfondir l'approche théorique, il serait pertinent d'étudier des oligosaccharides de longueur croissante qui contiennent un, plusieurs ou pas d'unité(s) galactofuranoside. Nous pourrions ainsi étudier leur contribution sur la flexibilité d'oligosaccharides présents dans la nature.

En somme, nous avons employé l'IRMPD et ses deux approches pour contribuer au peu de connaissance disponible sur le galactofuranose. Parmi les espèces de lichen analysées, nous avons discriminé celles pouvant posséder du galactofuranose. Notre développement méthodologique est clairement applicable sur les autres espèces de lichen. Il est même possible d'analyser séparément les différents composants du lichen. De nombreuses questions sont soulevées sur le rôle des organismes symbiotiques et leur échange mutuel. Des études basées sur notre démarche pourraient éclairer sur la biosynthèse et l'échange des polysaccharides entre les différents organismes chez le lichen. La méthodologie proposée peut ainsi fournir de nombreuses informations aux lichénologues. Par ailleurs, le protocole a été développé sur des espèces de lichen dans ce projet car ils sont inoffensifs pour l'homme et présents en abondance. Notre méthode convient à l'étude des polysaccharides de lichen, mais est également transférable à d'autres organismes. En effet, seules les premières étapes de la démarche sont spécifiques au lichen et peuvent être remplacées par des méthodes adaptées au type d'échantillon. Il est ainsi possible d'étudier des espèces de virus, de bactéries et de champignons qui synthétisent aussi du galactofuranose et dont certaines sont pathogènes pour l'homme (*Mycobacterium tuberculosis*, *Aspergillus fumigatus*...). Cela ouvre ainsi des perspectives en santé.

J'ai pu tester dans le cadre de cette thèse les performances de l'IRMPD à des fins analytiques. Grâce à cette méthode, notre équipe a non seulement pu apporter des éléments de réponse dans la partie du projet qui nous était destinée (identifier les oligosaccharides avec du galactofuranose) mais elle a en plus été utile à nos collaborateurs responsables d'autres étapes du projet (tester les méthodes de dépolymérisation). L'IRMPD semble aujourd'hui avoir atteint la maturité nécessaire pour être employé dans des projets analytiques.

Il s'agit pour l'équipe du premier projet avec une démarche aussi complète. Nous avons ainsi été confrontés pour la première fois à autant de données. Un travail conséquent a été réalisé pour traiter, visualiser et interpréter toutes ces données. La base de données créée dans le cadre du projet était de taille importante. Alors que les comparaisons étaient principalement effectuées visuellement, nous avons introduit un système de comparaison par score que nous avons publié pour la première fois. Les calculs théoriques ont également abouti à une quantité considérable de données. Nous avons ainsi introduit une nouvelle façon de décrire les préférences conformationnelles des furanoses en visualisant d'abord les différentes familles sur les roues de conformation puis en analysant

Conclusion

spécifiquement chaque conformation. Pour les échantillons provenant de source naturelle, un nombre élevé de composés à analyser était attendu (corrélé au nombre d'espèces de lichen, des modes de préparation d'échantillon et des différentes fractions LC). Nous avons ici également pu traiter ces données en introduisant la méthode phare du projet qui est le criblage. L'utilisation d'un laser à longueur d'onde fixe est très avantageuse car elle facilite grandement la configuration du couplage avec la LC.

L'IRMPD a une place parfaitement légitime dans des projets de recherche appliquée en glycomique. Les futurs projets s'appuieront sur les outils développés par mes prédécesseurs (dispositif expérimental, logiciel, séquençage, etc.), ceux développés dans le cadre de cette thèse (criblage, comparaison par scores, etc.) et demanderont nécessairement de nouveaux développements.

Liste des publications

En rédaction :

1. **O. Yeni**, S. Ollivier, D. Ropartz, H. Rogniaux, C. Guitteny, F. Le Dévéhat, J. Boustie, L. Legentil, V. Ferrières, B. Moge and I. Compagnon. Screening using LC-MS/MS-IR to detect galactofuranose.
2. **O. Yeni**, A. Camiruaga, A. Tillet, A-R Allouche, L. Legentil, V. Ferrières, P. Carçabal and I. Compagnon. Galactofuranose flexibility in different micro-environment.
3. J. Boustie, I. Compagnon, V. Ferrières, C. Guitteny, **O. Yeni**, L. Legentil, S. Ollivier, H. Rogniaux, D. Ropartz and F. Le Dévéhat. Depolymerisation of polysaccharide with conservation of galactofuranosyl units.

Soumis :

4. **O. Yeni**, A-R Allouche, L.L. Kiessling, L. Legentil, V. Ferrières and I. Compagnon. Conformational preferences of the flexible galactofuranose sugar in gas-phase.
5. **O. Yeni**, S. Ollivier, D. Ropartz, H. Rogniaux, F. Le Dévéhat, J. Boustie, L. Legentil, V. Ferrières, B. Moge and I. Compagnon. Ring-size memory of galactose-containing MS/MS fragment and application to the detection of galactofuranose in oligosaccharides, and their sequencing.
6. L-P David, S. Ferron, B. Favreau, D. Ropartz, **O. Yeni**, I. Compagnon, Vincent Ferrières, F. Le Dévéhat and L. Legentil. Lichenic oligosaccharides: synthesis of galactomannans fragments to help NMR assignment for extracted polysaccharides.

Publiés :

7. B. Moge, O. Yeni, A. Infantino and I. Compagnon, CO₂ laser enhanced rapid IRMPD spectroscopy for glycan analysis, *International Journal of Mass Spectrometry*, 2023, **490**, 117071.
8. B. Moge, B. Schindler, **O. Yeni** and I. Compagnon, Fucose Migration Pathways Identified Using Infrared Spectroscopy**, *Angewandte Chemie International Edition*, 2023, **62**, e202300538.
9. S. Ollivier, L. Legentil, **O. Yeni**, L.-P. David, V. Ferrières, I. Compagnon, H. Rogniaux and D. Ropartz, Gas-Phase Behavior of Galactofuranosides upon Collisional Fragmentation: A Multistage High-Resolution Ion Mobility Study, *J. Am. Soc. Mass Spectrom.*, 2023, **34**, 627–639.
10. **O. Yeni**, B. Schindler, B. Moge and I. Compagnon, Rapid IRMPD (InfraRed multiple photon dissociation) analysis for glycomics, *Analyst*, 2022, **147**, 312–317.
11. **O. Yeni**, A. Gharbi, S. Chambert, J. Rouillon, A.-R. Allouche, B. Schindler and I. Compagnon, Acetylated sugars in the gas phase: stability, migration, positional isomers and conformation, *Phys. Chem. Chem. Phys.*, 2022, **24**, 1016–1022.
12. J. S. Ho, A. Gharbi, B. Schindler, **O. Yeni**, R. Brédy, L. Legentil, V. Ferrières, L. L. Kiessling and I. Compagnon, Distinguishing Galactoside Isomers with Mass Spectrometry and Gas-Phase Infrared Spectroscopy, *J. Am. Chem. Soc.*, 2021, **143**, 10509–10513.
13. B. Favreau, **O. Yeni**, S. Ollivier, J. Boustie, F. L. Dévéhat, J.-P. Guégan, M. Fanuel, H. Rogniaux, R. Brédy, I. Compagnon, D. Ropartz, L. Legentil and V. Ferrières, Synthesis of an Exhaustive Library of Naturally Occurring Galp-Manp and Galp-Manp Disaccharides. Toward Fingerprinting According to Ring Size by Advanced Mass Spectrometry-Based IM-MS and IRMPD, *J. Org. Chem.*, 2021, **86**, 6390–6405.