



**HAL**  
open science

# Développement d'approches innovantes de marquage isotopique pour la caractérisation par RMN d'anticorps monoclonaux produits en cellules CHO

Arthur Giraud

► **To cite this version:**

Arthur Giraud. Développement d'approches innovantes de marquage isotopique pour la caractérisation par RMN d'anticorps monoclonaux produits en cellules CHO. Biologie structurale [q-bio.BM]. Université Grenoble Alpes [2020-..], 2023. Français. NNT : 2023GRALV060 . tel-04727995

**HAL Id: tel-04727995**

**<https://theses.hal.science/tel-04727995v1>**

Submitted on 9 Oct 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.

THÈSE

Pour obtenir le grade de

**DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES**

École doctorale : CSV- Chimie et Sciences du Vivant

Spécialité : Chimie Biologie

Unité de recherche : Institut de Biologie Structurale

**Développement d'approches innovantes de marquage isotopique pour la caractérisation par RMN d'anticorps monoclonaux produits en cellules CHO**

**Development of innovative isotopic labelling approaches to characterize by NMR mAbs produced in CHO cells**

Présentée par :

**Arthur GIRAUD**

Direction de thèse :

**Jérôme BOISBOUVIER**

DIRECTEUR DE RECHERCHE, Université Grenoble Alpes

Directeur de thèse

**Elodie Crublet**

NMR-Bio

Co-encadrante de thèse

**Oriane FRANCES**

Sanofi

Co-encadrante de thèse

Rapporteurs :

**Christina SIZUN**

DIRECTRICE DE RECHERCHE, CNRS délégation Ile-de-France Sud

**François-Xavier THEILLET**

CHARGE DE RECHERCHE HDR, CNRS délégation Ile-de-France Sud

Thèse soutenue publiquement le **28 septembre 2023**, devant le jury composé de :

**Jérôme BOISBOUVIER**

DIRECTEUR DE RECHERCHE, CNRS délégation Alpes

Directeur de thèse

**Pascal POIGNARD**

PROFESSEUR DES UNIVERSITÉS - PRATICIEN HOSPITALIER,  
Université Grenoble Alpes

Président

**Christina SIZUN**

DIRECTRICE DE RECHERCHE, CNRS délégation Ile-de-France Sud

Rapporteuse

**François-Xavier THEILLET**

CHARGE DE RECHERCHE HDR, CNRS délégation Ile-de-France Sud

Rapporteur

**Maggy HOLOGNE**

MAITRESSE DE CONFÉRENCES HDR, Université Lyon 1

Examinatrice

Invités :

**Elodie Crublet**

DOCTEUR EN SCIENCES, NMR-Bio

**Oriane Frances**

DOCTEUR EN SCIENCES, Sanofi





## Remerciements

Je remercie en premier lieu Jérôme Boisbouvier qui a été mon directeur de thèse et qui a su répondre avec patience (et répétition) à toutes mes questions pour mieux comprendre la RMN. Merci également à Oriane Frances et Elodie Crublet, mes co-directrices de thèse. Oriane pour avoir permis de créer ce projet en réunissant toutes les personnes impliquées à Sanofi et Elodie en tant que CEO de NMR-Bio qui a fournis tous les isotopes utilisés pendant la thèse mais aussi et principalement pour toutes les corrections que tu as apporté au manuscrit. Travailler sous la tutelle de trois responsables en même temps n'est pas chose facile mais on y est arrivé. Vous m'avez félicité pour y être arrivé mais je pense que c'est plutôt à vous de récolter les lauriers pour cette collaboration.

Je remercie Pascal Poignard ayant présidé le jury de thèse, Christina Sizun et François-Xavier Theillet, rapporteur et membres du jury, ainsi que Magalie Hologne, membre du jury de dernière minute, pour avoir évalué mon travail, ainsi que pour la partie discussion de la soutenance qui a été très constructive et intéressante.

Merci à Lionel Imbert, Rida Awad, Adrien Favier et Faustine Henot pour leur grande aide au cours de cette thèse. Lionel pour m'avoir tout appris sur le cell-free, Rida pour l'expression de l'anticorps en CHO, Adrien pour l'aide apporté lors de l'utilisation des logiciels d'attribution et Faustine pour l'attribution des groupements méthyles en fin de thèse. Vous avez été, tous les quatre, les personnes les plus impliquées dans le projet et avez permis de le mener à bien.

Lionel, je te remercie encore une fois, plus personnellement, pour les courses à pied (les marches pour toi) et tout le reste. Tu as réussi l'exploit de me former au cell-free tout en étant absent à cause du COVID. Ça a été long mais on y est arrivé ! Avec tous mes encouragements pour la fin (début) de rédaction ! Ne t'inquiète pas ça fini forcément à un moment^^.

شكرا رضى، رأس الكرنب المفضلة لدي. لقد كان من دواعي سروري سحب الشعر معك في المختبر. لا  
استطيع الإنتظار للعمل معك مرة أخرى لكي تتحمليني صباحا ظهرا و مساء.

Je remercie également Séverine Clavier, Séverine Beauvisage, Solène Bertho, Hélène Le-borgne et Nelly Lechat de Sanofi pour les analyses en spectrométrie de masse qui ont permis de quantifier l'incorporation des isotopes dans l'anticorps. Merci pour les analyses de dernière minute, ainsi que pour vos comptes rendus clairs, détaillés et précis.

Merci également à Peter Güntert de l'ETH de Zurich qui a accepté une collaboration permettant de faire fonctionner le logiciel MethylFlya avec les données que nous avons.

Je tiens à remercier également Ewen Lescop, Thierry Vernet et Nicolas Mouz qui ont été membres de mes comités de suivis de thèse et qui ont suivi son avancement pour son bon déroulement. Vos questions et vos remarque au cours des CSI m'ont permis de me préparer pour les questions du jury de thèse.

Je tiens à remercier aussi Stephanie Seguin Huet, Anaïs Muhr-Naninck et Benoit Mothes de m'avoir formé pendant les deux premiers mois de la thèse à Sanofi pour la production et la purification d'anticorps exprimés en cellules CHO. Merci également à Laetitia Macon de m'avoir aidé à comprendre comment a été produit la souche cellulaire utilisée pendant la thèse.

Le nombre de personne impliqué dans le projet à Sanofi est élevé. Aussi, je vais essayer d'être le plus exhaustif possible en remerciant Maina Kerbrat, Frédéric Herman, Francis Dufieux, Cécile Capdevilla,

Alexey Rak, Magalie Mathieu, Brunot Genet, Jean Michel Menet et Laurent Duhau qui ont tous été impliqués dans le projet. Je remercie plus personnellement François Vallee, Camille Sanson et Oriane Frances qui m'ont bien accueilli lors de mes passages sur le site de Vitry.

Je tiens à remercier également Khalid Mahla qui a été mon neurochirurgien pendant la fin de la thèse. Mon intervention qui a été menée à bien grâce à votre expérience m'a permis de retrouver une santé optimale avant la rédaction de la thèse. Merci à mes amis Nicolas et Léopold d'être venu me voir à l'hôpital pendant cette période. Votre visite m'a fait énormément plaisir. Merci également à Axel Giret, kinésithérapeute et tortionnaire, qui m'a remis sur pieds (au sens propre et figuré). Les séances que j'ai eues avec toi m'ont permis de retrouver une santé physique et mentale qui ont été essentielles pendant la période de rédaction. Plus que ça, tu m'as permis de récupérer un équilibre parfait et j'ai pu, grâce à cela, réaliser des choses incroyables que j'aurais pensé impossible après l'opération.

A big thank you to all the PhD students, post-docs and "youngs" of the NMR team or elsewhere in IBS, I named Iva, Undina, Laura, Ricarda, Nina, Tiago, Faustine, Massilia ABBA, Béatrice (bubulle), Aldo, Arijit, Sonja. Everything we've done together, hikes, parties, afterworks, cooking or PhD hat workshop, and other fun things made this PhD unforgettable. I hope to see you all again all around the glob for futur conferences or even jobs.

Merci à toute l'équipe RMN de l'IBS de m'avoir accueilli pour cette thèse. Je remercie plus personnellement Isabelle, "le cœur" du labo qui permet au wetlab de tenir debout. Merci pour toutes les discussions que nous avons eues sur tout et n'importe quoi, pour ta bonne humeur mais aussi pour tes regards assassins. Promis je n'oublierai plus de remplir les fiches de suivi !!! Merci aussi à Karine et Astrid de m'avoir supporté dans le bureau en cette fin de thèse pendant la rédaction.

Merci à mes parents, grands-parents, mon frère, mes soeurette et Paul d'être venu assister à la soutenance. Merci d'avoir essayé de comprendre (ou de faire semblant) ce que j'ai présenté et merci de m'avoir supporté pendant cette fin de thèse compliquée. Merci pour tout ce que vous avez fait pour moi. Avec un grand merci de la part de tout le monde pour la qualité du buffet qui (je pense) a vraiment plu.

Enfin, je te remercie toi, Annelise. Tu as été avec moi dans les moments les plus difficiles que j'ai eu à affronter dans la vie. Je ne réussirais jamais à te remercier comme tu le mérite avec de simples mots. Nous avons déjà vécu des moments précieux accompagnés de tortues, raies, petits et énormes requins, mais j'espère que nous en aurons pleins d'autres maintenant que les moments difficiles sont passés. Ma thèse terminée, je recherche maintenant une formation spéciale en surf, donc si tu connais quelqu'un de passionné, ça m'intéresse.

## Résumé en français

Les anticorps monoclonaux (mAb) sont des produits bio thérapeutiques ayant apporté de nombreuses solutions pour le traitement de maladies létales ou chroniques. En raison de ce potentiel thérapeutique élevé, le poids des mAbs sur le marché du médicament est en croissance constante. En industrie pharmaceutique, les mAbs sont principalement exprimés en cellules ovariennes de hamster chinois (CHO). Afin de valider l'activité de chaque lot de production, il est impératif de s'assurer de la qualité structurale des mAbs produits. Parmi les nombreuses techniques analytiques utilisées jusqu'à présent, aucune n'offre une résolution suffisante pour apporter une réponse structurale à l'échelle atomique sur les pertes d'activité thérapeutique observées à la suite d'études de stress. Pour cela, la RMN 2D  $^1\text{H}$ - $^{13}\text{C}$  des groupements méthyles, donnant accès à une résolution beaucoup plus élevée, s'avère prometteuse. En effet, les groupements méthyles des anticorps sont distribués uniformément sur la structure des mAbs et leurs caractéristiques physico-chimiques en font des sondes idéales pour suivre par RMN toute perturbation de la structure des mAb à l'échelle atomique. Cependant, l'utilisation de ces spectres RMN 2D  $^1\text{H}$ - $^{13}\text{C}$ , qui reflètent la structure des mAb à l'échelle atomique, est actuellement limitée à la comparaison "d'empreintes spectrales" à un spectre de référence. En effet, l'attribution de chaque signal RMN à son groupement méthyle dans le mAb doit être réalisée en amont pour corrélérer les pertes d'activités avec les perturbations observées sur les spectres RMN. Ce processus d'attribution nécessite l'enrichissement en isotopes stables ( $^{13}\text{C}$ ,  $^{15}\text{N}$ ,  $^2\text{H}$ ) des mAbs. Cependant, aucune méthode de marquage isotopique pour les protéines produites en cellules CHO n'a été proposée à ce jour.

Dans ce contexte, ces travaux de thèse ont eu pour ambition de développer des protocoles robustes de marquages isotopiques permettant l'attribution des résonances RMN des groupements méthyles d'un anticorps exprimé en cellules CHO. Pour cela, une méthode de marquage isotopique des groupements méthyles ( $^{13}\text{CH}_3$ ) de l'anticorps anti-LAMP1 exprimé en cellules CHO a été développée permettant ainsi l'acquisition d'expériences RMN 2D  $^1\text{H}$ - $^{13}\text{C}$  des groupements méthyles de haute qualité. Cependant, la perdeutération des protéines, nécessaire à l'attribution des expériences RMN 3D d'un anticorps, est impossible en cellules de mammifères. Pour résoudre ce problème, une méthode de marquage isotopique des fragments Fab et Fc de l'anticorps a été développée en système d'expression acellulaire (*cell-free*). En effet, contrairement aux cellules CHO, l'utilisation de ce système permet la production de protéines en solvant deutéré.

Ainsi, les fragments uniformément marqués U- $[^2\text{H}, ^{13}\text{C}, ^{15}\text{N}]$  ont été produits pour l'acquisition d'expériences RMN 2D  $^1\text{H}$ - $^{15}\text{N}$  et 3D, permettant de proposer pour la première fois l'attribution des résonances RMN  $^1\text{H}$ - $^{15}\text{N}$  du squelette d'un fragment Fab d'un anticorps. Puis, une méthode de marquage isotopique optimisée  $^{13}\text{CH}_3$  spécifique des acides aminés portant un groupement méthyle a été appliquée sur le fragment Fab perdeutééré et exprimé en système *cell-free*, permettant ainsi d'attribuer les résonances des spectres RMN  $^1\text{H}$ - $^{13}\text{C}$  des groupements méthyles de ce même fragment Fab. La superposition des spectres RMN  $^1\text{H}$ - $^{13}\text{C}$  attribués du fragment Fab exprimé en système *cell-free* avec ceux de l'anticorps marqué  $^{13}\text{CH}_3$  en cellules CHO a ensuite permis de transférer l'attribution des groupements méthyles du fragment Fab vers le spectre de l'anticorps exprimé en cellules CHO. Ces travaux ouvrent la

possibilité d'utiliser la RMN des groupements méthyles pour les caractérisations structurale et dynamique ainsi que pour les études des interactions et des modifications des mAb thérapeutiques à l'échelle atomique.

**Mots clés :** RMN / anticorps / marquage isotopique / groupements méthyles / cellules CHO / système cell-free / attribution

## Summary in english

Monoclonal antibodies (mAbs) are biotherapeutics products that have achieved outstanding success for the treatment of many life-threatening and chronic diseases. Due to their high therapeutic potential towards the development of new treatments to numerous incurable diseases, mAbs market share are steadily expanding on the drug market. In the pharmaceutical industry, mAbs are mainly produced in Chinese Hamster Ovary (CHO) cells. To ensure that each batch retains full therapeutical activities, it is essential to control the structural quality of mAbs. However, none of the many analytical techniques available nowadays offer a sufficient resolution to provide a structural answer at the atomic scale to the loss of therapeutic potency observed during stress studies. To this end, 2D  $^1\text{H}$ - $^{13}\text{C}$  NMR experiments of methyl groups is a promising technique providing access to a much higher resolution. Indeed, methyl groups are very well distributed throughout the mAbs structure, and their physico-chemical characteristics make them ideal probes for monitoring any structural perturbation of the mAb at an atomic scale using NMR. However, the use of 2D  $^1\text{H}$ - $^{13}\text{C}$  NMR spectra, reflecting the mAbs structure at an atomic scale, is currently limited to the comparison of 'spectral fingerprints' with reference spectra. To go further, and in order to correlate the loss of activity with perturbation observed in NMR spectra, each NMR signal must be assigned to its corresponding methyl group in the mAb. This assignment process requires a stable isotope enrichment ( $^{13}\text{C}$ ,  $^{15}\text{N}$ ,  $^2\text{H}$ ) of mAbs. However, no isotopic labeling method for proteins produced in CHO cells has been proposed to date.

In this context, this work aimed to provide robust isotopic labelling protocols enabling NMR resonances assignment of mAbs methyl groups expressed in CHO cells. To this end, an isotopic labelling method of methyl groups ( $^{13}\text{CH}_3$ ) of the anti-LAMP1 antibody expressed in CHO cells has been developed, enabling the acquisition of high-quality 2D  $^1\text{H}$ - $^{13}\text{C}$  NMR experiments of methyl groups. However, protein perdeuteration is impossible in mammalian cells, preventing assignment of 3D NMR signal of the antibody. To overcome this, an isotopic labelling method for the Fab and Fc fragments of the antibody was developed in a cell-free expression system, enabling production of proteins in a perdeuterated solvent (contrary to expression in CHO cells).

Thus, uniformly U- $^2\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$ ]-labelled Fab and Fc fragments were produced for the acquisition of 2D  $^1\text{H}$ - $^{15}\text{N}$  and 3D NMR experiments, enabling for the first time the backbone assignment of  $^1\text{H}$ - $^{15}\text{N}$  NMR resonances of a Fab antibody fragment. Then, a  $^{13}\text{CH}_3$  optimized isotopic labelling method, specific to amino acids bearing a methyl group, was also implemented for the deuterated Fab fragment expressed in cell-free system, providing for the first time the resonances assignment of the  $^1\text{H}$ - $^{13}\text{C}$  NMR spectra of methyl groups this same Fab fragment. Overlay of the assigned  $^1\text{H}$ - $^{13}\text{C}$  NMR spectra of the Fab fragment expressed in cell-free system with those of the  $^{13}\text{CH}_3$ -labelled antibody in CHO cells then enabled to transfer of methyl group assignments of the Fab fragment to the spectrum of the antibody expressed in CHO cells. Paving the way for the use of methyl group NMR for structural and dynamic characterization, this work represents a major advance towards in-depth studies of the interactions and modifications of therapeutic mAbs at the atomic scale resolution.

**Key words :** NMR / antibodies / isotopic labelling / methyl groups / CHO cells / cell-free system / assignment

# Glossaire

ADC: *antibody drug conjugate*  
ADCC: *antibody dependent cells cytotoxicity*  
ADCP: *antibody dependent cellular phagocytosis*  
ADN: acide désoxyribonucléique  
ADP: adénosine diphosphate  
AMPc: adénosine-monophosphate cyclique  
ARN: acide ribonucléotidique  
ARNm: Acide ribonucléique messenger  
ARNt: acide ribonucléique de transfert  
ATP: adénosine triphosphate  
ATU: autorisation temporaire d'utilisation  
BAC: *boronate affinity chromatography*  
BCAT: *branched-chain amino acid amino-transferase*  
BiTE: *bispecific T cell engagers*  
CAM: complexe d'attaque membranaire  
CDC: cytotoxicité dépendante du complément  
CDR: *complementary-determining regions*  
CDRH: *complementary-determining regions* des chaînes lourdes  
CDRL: *complementary-determining regions* des chaînes légères  
CE: *capillary electrophoresis*  
CH1,2,3: domaines constants de la chaîne lourde  
CHO: *chinese hamster ovary (cells)*  
CID: *collision induced dissociation*  
CL: domaine constant de la chaîne légère  
COSY: *correlation spectroscopy*  
CQA: *critical quality attributes*  
cryo-EM: *cryogenic electron microscopy*  
DHFR: *dihydrofolate reductase*  
Dsb: *disulfide bond chaperon protein*  
*E. coli: Escherichia coli*  
EMA: *european medicinal agency*  
ESI: *electro spray ionisation*  
ETD: *electron transfer dissociation*  
Fab: *antigen binding fragment*  
Fc: *crystallisable fragment*  
FcR: récepteur à fragment Fc  
FcRn: récepteur néonatal à fragment Fc  
FDA: *food and drug administration*  
FT-IR: *Fourier transform infrared spectroscopy*  
Fuc: fucose  
Fv: fragment variable  
GalNAc: N-acétylgalactosamine  
GlcNAc: N-acétylglucosamine  
gor: réductase à glutatharédoxine  
GS: glutamine synthétase

GSH: glutathion réduit  
GSSG: glutathion oxydé  
HACA: *human anti-chimeric antibodies*  
HAMA: *human anti-murine antibody*  
HC: chaîne lourde  
HDX: *hydrogen-deuterium exchange*  
HEK 293: *Human embryonic kidney cells*  
HIC: *hyrdophobic interaction chromatography*  
HMQC: *heteronuclear multiple quantum correlation*  
HOS: *higher order structure*  
HSQC: *heteronuclear single quantum cohérence*  
IEC: *ion exchange chromatography*  
IEF: *isoelectric focalisation*  
Ig: immunoglobuline  
LAMP1: *lysosomal associated membrane protein 1*  
LC: chaîne légère  
LC: liquid chromatography  
mAb: *monoclonal antibody*  
MALDI: *matrix assisted laser desorption ionization*  
Man : mannose  
MS: *mass spectrometry*  
MSX: méthionine sulphoximine  
NK: *natural killer (cells)*  
NOE: *nuclear Overhauser effect*  
NOESY: *nuclear Overhauser enhancement spectroscopy*  
NS0: *non-secreting immunoglobuline cells (rat)*  
PCA: *principale component analysis*  
PCS: *pseudocontact shifts*  
PDB: *protein data bank*  
PEG: polyéthylène glycol  
PERC6: cellules de rétines humaines immortalisées  
PLP: pyridoxal 5-phosphate  
PRE: *paramagnetic relaxation enhancement*  
PTM: *posttranslational modifications*  
RMN: résonance magnétique nucléaire  
SC: composant sécrétoire transmembranaire  
scFv: *small scale variable fragment*  
SDS: sodium dodecyl-sulfate  
SEC: *size exclusion chromatography*  
SOFAST: *selective optimized flip angle short transient*  
TOCSY: *total correlation spectroscopy*  
TROSY: *transverse relaxation optimized spectroscopy*  
trxB: réductase à thiorédoxine  
VH: domaine variable de la chaîne lourde  
VL: domaine variable de la chaîne légère  
YB2/0: cellules de myélome de rat

# Table des matières

<b>TABLE DES MATIERES .....</b>	<b>11</b>
<b>LISTE DES FIGURES .....</b>	<b>15</b>
<b>LISTE DES TABLEAUX .....</b>	<b>20</b>
<b>PREAMBULE.....</b>	<b>21</b>
<b>I. INTRODUCTION.....</b>	<b>22</b>
1. LES ANTICORPS PHYSIOLOGIQUES.....	22
1.1. Structure des immunoglobulines de classe G1 : IgG1.....	22
1.1.1. Structure primaire .....	22
a Structure primaire des chaînes légères .....	23
b Structure primaire des chaînes lourdes .....	23
c Nomenclature des différents fragments.....	23
1.1.2. Structure secondaire et tertiaire.....	24
1.1.3. Structure quaternaire .....	26
1.2. Les glycosylations d'IgG.....	27
1.2.1. Structure de base des N-glycosylations et O-glycosylations.....	27
1.2.2. Les différentes N-glycosylations des IgG.....	27
1.3. Structure des cinq classes principales d'immunoglobulines.....	28
1.3.1. Les IgG.....	28
1.3.2. Les IgD.....	29
1.3.3. Les IgE.....	30
1.3.4. Les IgA.....	30
1.3.5. Les IgM.....	31
1.4. Les anticorps dans le système immunitaire.....	33
1.4.1. Rappels sur le système immunitaire .....	33
1.4.2. Les sites actifs des anticorps .....	34
a Sites actifs sur le fragment Fab .....	35
b Sites actifs sur le fragment Fc.....	35
1.4.3. Fonctions effectrices des anticorps.....	37
a Neutralisation, agglomération, et précipitation .....	37
b Opsonisation et phagocytose.....	37
c La cytotoxicité dépendante du complément : CDC .....	38
d Activation des cellules cytotoxiques dépendantes des anticorps : ADCC .....	38
e Inflammation .....	38
2. LES ANTICORPS THERAPEUTIQUES .....	40
2.1. Les premiers anticorps thérapeutiques .....	41
2.1.1. Les anticorps polyclonaux .....	41
2.1.2. Développement des premiers anticorps monoclonaux .....	42
2.1.3. Les anticorps monoclonaux chimériques.....	43
2.1.4. Les anticorps monoclonaux humanisés.....	44
2.1.5. Les anticorps monoclonaux humains.....	44
2.2. Mécanismes d'action des anticorps thérapeutiques .....	45
2.2.1. Les mAbs thérapeutiques neutralisants.....	45
2.2.2. Les mAbs thérapeutiques agonistes et antagonistes .....	46
2.2.3. Les mAbs thérapeutiques cytolytiques .....	47
2.3. Les différentes structures des mAbs thérapeutiques.....	47
2.3.1. IgG entières et fragments.....	47
2.3.2. Anticorps monoclonaux multi spécifiques.....	49
2.3.3. Anticorps conjugués.....	50
2.3.4. Les protéines de fusion au Fc.....	51
2.4. Application des mAbs thérapeutiques face à la COVID-19.....	52
2.5. L'anticorps anti-LAMP1 : modèle de la thèse .....	54
3. LES SYSTEMES D'EXPRESSION DES ANTICORPS THERAPEUTIQUES .....	56

3.1.	Points clés de l'expression d'anticorps .....	56
3.1.1.	Les différentes méthodes de productions.....	56
a	Pour l'industrie ou la recherche.....	56
b	La production en mode <i>batch</i> .....	57
c	La production en mode <i>fed-batch</i> .....	58
d	La production en perfusion .....	58
3.1.2.	Aspects qualitatifs de la production d'anticorps.....	58
a	La monoclonalité des anticorps .....	58
b	Les ponts-disulfures .....	60
c	Les glycosylations .....	62
3.2.	Production d'anticorps en cellules eucaryotes .....	64
3.2.1.	Les cellules murines.....	64
a	Les hybridomes de souris.....	64
b	Les cellules de souris : SP2/0 et NS0 .....	65
3.2.2.	Les cellules ovariennes de hamster chinois : CHO .....	66
a	Les différentes souches de CHO.....	66
b	Production à haute densité cellulaire.....	67
c	Glycosylations des CHO.....	67
3.2.3.	Les autres cellules eucaryotes pour les anticorps thérapeutiques .....	68
a	Cellules humaines : PerC6 et HEK 293 .....	68
b	Les myélomes de rat YB2/0 .....	68
c	Les levures.....	69
3.3.	Production d'anticorps thérapeutiques en système procaryote : les bactéries .....	70
3.3.1.	Expression d'anticorps sous forme soluble dans le périplasma .....	70
3.3.2.	Expression d'anticorps et fragments dans le cytoplasme .....	72
a	Expression de protéines sous forme soluble dans le cytoplasme.....	72
b	Expression des protéines en corps d'inclusion insolubles .....	72
4.	ANALYSES DES CARACTERISTIQUES PHYSICO-CIMIQUES ET STRUCTURALES DES MABS THERAPEUTIQUES .....	74
4.1.	Hétérogénéités physico-chimiques et structurales des mAbs thérapeutiques .....	75
4.1.1.	Hétérogénéité des glycosylations .....	75
4.1.2.	Modifications chimiques des anticorps thérapeutiques.....	76
a	Formation de pyroglutamate en N-terminal des mAbs.....	76
b	Suppression de la lysine ou arginine en C-terminal des mAbs .....	76
c	Désamidation des asparagines et isomérisation des aspartates .....	77
d	Oxydations des méthionines .....	78
e	Glycation .....	79
f	Réduction des ponts disulfures.....	80
4.1.3.	Modifications de la structure hautement ordonnée des mAbs thérapeutiques .....	80
4.2.	Caractérisation de la HOS des anticorps .....	82
4.2.1.	Détermination structurale des anticorps par cryo-microscopie électronique et cristallographie aux rayons X.....	82
4.2.2.	Etude de la HOS par les méthodes de spectroscopie optique.....	83
4.2.3.	La spectrométrie de masse d'échange hydrogène/deutérium .....	83
4.3.	La résonance magnétique nucléaire des anticorps en abondance naturelle.....	85
4.3.1.	La RMN 1D des protons pour la caractérisation des mAbs thérapeutiques .....	85
4.3.2.	La RMN multidimensionnelle en détection <sup>1</sup> H : empreinte digitale de la structure .....	86
a	La RMN <sup>1</sup> H- <sup>15</sup> N des groupements amides des anticorps.....	87
b	La RMN <sup>1</sup> H- <sup>13</sup> C des groupements méthyles des mAbs thérapeutiques .....	88
4.4.	Le marquage isotopique de protéines de haut poids moléculaire pour la RMN.....	92
4.4.1.	Intérêt du marquage isotopique <sup>15</sup> N, <sup>13</sup> C des protéines pour la RMN.....	92
4.4.2.	Nécessité de la deutération des protéines de haut poids moléculaire pour la RMN.....	93
a	La relaxation du signal RMN des protéines de haut poids moléculaire.....	93
b	Optimisation de la relaxation par la variation du τ <sub>c</sub> de la protéine.....	94
c	Optimisation de la relaxation par la deutération des protéines.....	95
5.	STRATEGIE MISE EN PLACE POUR ETUDIER, PAR RMN, LA HOS DES ANTICORPS THERAPEUTIQUES EN SOLUTION ET A L'ECHELLE ATOMIQUE .....	96
5.1.	Méthodes d'attribution des groupements méthyles par RMN .....	96

5.1.1.	Connection des groupements méthyles aux résonances du squelette à l'aide d'expériences HCC .....	96
5.1.2.	Attribution des groupements méthyles par mutagenèse dirigée.....	98
5.1.3.	Utilisation de la structure pour l'attribution des groupements méthyles .....	99
a	Utilisation de la structure et des sondes paramagnétiques pour l'attribution des groupements méthyles .....	99
b	Utilisation de la structure et du réseau de NOE pour l'attribution des groupements méthyles .....	100
5.2.	Stratégie d'attribution des groupements méthyles de l'anticorps anti-LAMP1 exprimé en cellules CHO .....	103
<b>II.</b>	<b>MATERIEL ET METHODES .....</b>	<b>105</b>
1.	MATERIELS ET METHODES POUR LA PRODUCTION DE L'ANTICORPS EN CELLULES CHO .....	105
1.1.	La lignée cellulaire CHO .....	105
1.2.	Culture des cellules CHO et production de l'anticorps anti-LAMP1 .....	106
1.3.	Purification de l'anticorps anti-LAMP1 et quantification.....	106
1.4.	Spectroscopie RMN.....	107
1.5.	Spectrométrie de masse.....	107
2.	MATERIEL ET METHODES POUR L'EXPRESSION ET LE MARQUAGE ISOTOPIQUE DU FRAGMENT FAB EN SYSTEME <i>CELL-FREE</i> .....	109
2.1.	Production du fragment Fab en système <i>cell-free</i> .....	109
2.2.	Purification du fragment Fab en système <i>cell-free</i> .....	114
2.3.	Production et purification de l'isomérase à pont disulfure DsbC .....	115
2.4.	Spectroscopie RMN.....	115
2.5.	Attribution du squelette et des groupements méthyles et dépôt des données.....	117
<b>III.</b>	<b>RESULTATS.....</b>	<b>118</b>
1.	MARQUAGE ISOTOPIQUE DE L'ANTICORPS ANTI-LAMP1 EN CELLULES CHO .....	118
1.1.	Introduction au marquage isotopique en cellules eucaryotes .....	118
1.1.1.	Marquage isotopique <sup>15</sup> N et <sup>13</sup> C de protéines exprimées en cellules eucaryotes pour la RMN .....	118
a	Le marquage isotopique <sup>15</sup> N, <sup>13</sup> C uniforme de protéine en levures .....	118
b	Les stratégies de marquage isotopique uniforme en cellules eucaryotes supérieures .....	119
c	Exemples de marquage isotopique <sup>15</sup> N, <sup>13</sup> C des protéines en cellules eucaryotes supérieures .....	120
1.1.2.	Deutération uniforme des protéines en cellules eucaryotes.....	121
a	Deutération uniforme des protéines en levures.....	121
b	Inconvénient de la deutération uniforme des protéines en milieu perdeutééré.....	121
c	Deutération uniforme des protéines en cellules eucaryotes supérieures .....	122
1.1.3.	Marquage spécifique des groupements méthyles des protéines en cellules eucaryotes.....	123
a	Marquage isotopique des protéines sur les groupements méthyles .....	123
b	Marquage isotopique des groupements méthyles dans les levures .....	123
c	Deutération sélective et marquage isotopique des groupements méthyles dans les cellules d'insectes et de mammifères .....	124
1.1.4.	Etat de l'art sur le marquage isotopique des anticorps exprimés en cellules de mammifères .....	126
a	Production d'un fragment Fv de 25 kDa [U- <sup>13</sup> C, <sup>15</sup> N] en hybridomes de souris pour la RMN.....	127
b	Production d'un fragment Fc de 50 kDa [U- <sup>13</sup> C, <sup>15</sup> N] en cellules CHO pour la RMN.....	128
c	Marquage stéréospécifique <sup>13</sup> C et deutération sélective des leucines d'une IgG2b exprimée en cellules hybridomes de souris .....	129
1.2.	Production de l'IgG1 anti-LAMP1 marquée spécifiquement <sup>13</sup> CH <sub>3</sub> sur les groupements méthyles en cellules CHO .....	131
1.2.1.	Mise en place de la production et de la purification de l'anticorps exprimé par les cellules CHO... ..	131
1.2.2.	Développement d'un milieu de culture adapté au marquage isotopique en cellules de mammifères.....	133
a	Evaluation de trois mélanges d'acides aminés sur le rendement d'expression de l'anticorps .....	133
b	Développement d'un mélange d'acides aminés optimisé.....	134
c	Optimisation du rendement d'expression de l'anticorps anti-LAMP1 en cellules CHO à haute densité cellulaire.....	137
d	Optimisation du taux d'incorporation des acides aminés .....	139
e	Marquage isotopique <sup>13</sup> CH <sub>3</sub> des isoleucines, leucines, méthionines, thréonines et valines.....	141
f	Le marquage isotopique de l'alanine : cas particulier.....	143

1.2.3.	Analyse RMN de l'IgG1 anti-LAMP1 exprimée en cellules CHO et spécifiquement marquée sur les groupements méthyles. ....	145
2.	ATTRIBUTION DU SQUELETTE DU FRAGMENT FAB EXPRIME EN SYSTEME <i>CELL-FREE</i> .....	149
2.1.	Introduction à l'expression d'anticorps en milieu acellulaire .....	149
2.1.1.	Expression de protéines en système <i>cell-free</i> .....	149
a	L'extrait cellulaire.....	150
b	Système de transcription et traduction couplée .....	151
c	Le système énergétique .....	152
2.1.2.	Expression d'anticorps et fragments en système <i>cell-free</i> .....	153
2.2.	Marquage isotopique <sup>15</sup> N, <sup>13</sup> C du squelette des protéines en système <i>cell-free</i> pour la RMN .....	155
2.2.1.	La fuite isotopique des <sup>15</sup> N du squelette des protéines exprimées en système <i>cell-free</i> .....	155
2.2.2.	La fuite isotopique et la deutération des protéines en système <i>cell-free</i> .....	156
2.2.3.	L'avantage du système <i>cell-free</i> face à la fuite isotopique.....	156
2.3.	Article : Enabling Site-Specific NMR Investigations of Therapeutic Fab using a Cell-Free Based Isotopic Labeling approach : Application to anti-LAMP1 Fab. ....	158
3.	ATTRIBUTION DES GROUPEMENTS METHYLES DU FRAGMENT FAB EXPRIME EN SYSTEME <i>CELL-FREE</i> .....	191
3.1.	Optimisations du protocole de production du fragment Fab en système <i>cell-free</i> .....	191
3.1.1.	Etude du rôle de la protéine chaperonne DsbC recombinante ajoutée dans le système <i>cell-free</i> .....	191
3.1.2.	Diminution de la quantité de réactifs du système <i>cell-free</i> .....	193
3.1.3.	Optimisation du protocole de purification du fragment Fab .....	194
3.2.	Marquage isotopique sélectif des groupements méthyles du fragment Fab.....	196
3.2.1.	Production en système <i>cell-free</i> du fragment Fab marqué sélectivement sur les groupements méthyles .....	196
3.2.2.	Acquisition de spectres RMN 2D <sup>1</sup> H- <sup>13</sup> C du fragment Fab marqué sélectivement .....	198
3.2.3.	Validation de la stratégie de division de l'anticorps en fragment pour l'attribution.....	199
3.3.	Attribution des groupements méthyles en utilisant des expériences RMN de type HCC ....	201
3.3.1.	Production du fragment Fab perdeutééré en système <i>cell-free</i> avec un marquage isotopique <sup>13</sup> C linéarisé sur les alanines, isoleucines et valines .....	201
3.3.2.	Acquisition des expériences RMN de type HCC "out and back" .....	203
3.3.3.	Exemple d'attribution pour les valines 132 et 133 .....	205
3.3.4.	Attribution des C $\beta$ des alanines, cas particulier .....	206
3.3.5.	Attribution des groupements méthyles des alanines, isoleucines et valines en utilisant les HCC. ....	206
3.4.	Attribution des groupements méthyles en utilisant des expériences RMN méthyles NOESY et la structure du fragment Fab .....	208
3.4.1.	Production du fragment Fab perdeutééré en système <i>cell-free</i> avec un marquage isotopique <sup>13</sup> CH <sub>3</sub> stéréo et régio spécifique sur les groupements méthyles.....	208
3.4.2.	Attribution du type de résidus pour les résonances du spectre méthyle SOFAST.....	210
3.4.3.	Attribution des groupements méthyles du fragment Fab grâce aux corrélations NOE inter-méthyles .....	212
<b>IV.</b>	<b>CONCLUSION.....</b>	<b>216</b>
1.	CONCLUSION ET DISCUSSION .....	216
1.1.	Marquage isotopique spécifique des groupements méthyles des mAbs produit en cellules CHO .....	216
1.2.	Production et marquage isotopique des Fabs <i>in vitro</i> .....	217
1.3.	Attribution des groupements méthyles du Fab anti-LAMP1 .....	218
2.	PERSPECTIVES.....	220
2.1.	Application des méthodes développées à l'étude du fragment Fc .....	220
2.2.	Etude de la HOS d'un anticorps par RMN .....	221
2.3.	Transfert d'attribution à d'autres fragments Fabs.....	221
<b>V.</b>	<b>BIBLIOGRAPHIE .....</b>	<b>223</b>
<b>VI.</b>	<b>ANNEXES .....</b>	<b>235</b>
1.	COMPOSITION EN ACIDES AMINES DU MELANGE AAOPTICHO .....	235
2.	COMPOSITION EN ACIDES AMINES DU MELANGE AADMEM .....	235

<b>3.</b>	<b>COMPOSITION EN ACIDES AMINES DU MELANGE AA<sub>INSCT</sub></b> .....	<b>236</b>
<b>4.</b>	<b>IDENTIFICATION DES SIGNAUX DES LEUCINES DU FRAGMENT FAB</b> .....	<b>237</b>

## Liste des figures

<b><u>1.1.1 :</u></b>	Nomenclature des différents domaines des IgG et résultats de clivage à la pepsine et à la papaïne .....	23
<b><u>1.1.2 :</u></b>	Structure du fragment Fab de l'anticorps anti-LAMP1 .....	25
<b><u>1.1.3 :</u></b>	Structure des fragments Fc des IgG .....	25
<b><u>1.1.4 :</u></b>	Structure d'une IgG .....	26
<b><u>1.1.5 :</u></b>	Séquences spécifiques et structures principales des N-glycosylations et O-glycosylations .....	27
<b><u>1.1.6 :</u></b>	Représentation schématique des différentes configurations possibles de glycosylation des asparagines 297 des IgG .....	28
<b><u>1.1.7 :</u></b>	Structure des IgG1, IgG2, IgG3, et IgG4 .....	29
<b><u>1.1.8 :</u></b>	Structure des IgD .....	29
<b><u>1.1.9 :</u></b>	Structure des IgE .....	30
<b><u>1.1.10 :</u></b>	Structure des IgA .....	31
<b><u>1.1.11 :</u></b>	Structure des IgM .....	32
<b><u>1.1.12 :</u></b>	Mécanisme simplifié de la réponse humorale du système immunitaire .....	34
<b><u>1.1.13 :</u></b>	Position des sites actifs d'une IgG sur sa structure avec les différents acteurs et leur site de liaison .....	34
<b><u>1.1.14 :</u></b>	Structure cristallographique du fragment Fc d'une IgG1 humaine avec un Fc $\gamma$ R.....	36
<b><u>1.1.15 :</u></b>	Structure cristallographique de la chaîne lourde du fragment Fc d'une IgG1 de rat avec un FcRn .....	36
<b><u>1.1.16 :</u></b>	Neutralisation, précipitation et agglomération des bactéries, virus et toxines par les anticorps .....	37
<b><u>1.1.17 :</u></b>	Oponisation et phagocytose d'un corps pathogène .....	37
<b><u>1.1.18 :</u></b>	Activation de la cascade du complément par la reconnaissance du site C1q d'un anticorps .....	38
<b><u>1.1.19 :</u></b>	Activation d'un ADCC .....	38
<b><u>1.2.1 :</u></b>	Evolution du nombre d'anticorps monoclonaux et de fragments thérapeutiques acceptés par les agences EMA et FDA de 1986 à 2022 .....	40
<b><u>1.2.2 :</u></b>	Principe de l'ADN recombinant .....	43
<b><u>1.2.3 :</u></b>	Les différents types d'anticorps thérapeutiques avec la composition de leur séquence .....	43
<b><u>1.2.4 :</u></b>	Principe d'action du Mogamulizumab .....	46
<b><u>1.2.5 :</u></b>	Représentation schématique des différentes structures d'IgG1 et fragments de mAbs thérapeutiques .....	48

<b><u>1.2.6 :</u></b>	Structure des anticorps thérapeutiques multi spécifiques .....	49
<b><u>1.2.7 :</u></b>	Principe d'action de l'Emicizumab .....	49
<b><u>1.2.8 :</u></b>	Principe d'action des anticorps conjugués à des molécules actives (ADC) .....	50
<b><u>1.2.9 :</u></b>	Schéma structural du Dulaglutide .....	51
<b><u>1.2.10 :</u></b>	Structure obtenue en cryo-EM de la protéine spike seule et en complexe avec des fragments Fab d'IgG .....	52
<b><u>1.2.11 :</u></b>	Modèle structural de la protéine DC-LAMP et LAMP1 à la surface des lysosomes .....	54
<b><u>1.2.12 :</u></b>	Représentation de la structure cristallographique du complexe Fab/LAMP1 et de la structure du squelette du domaine N-terminal de LAMP1 .....	55
<b><u>1.3.1 :</u></b>	Aperçu de la quantité d'anticorps nécessaire pour les différentes phases de développement de la recherche à la commercialisation .....	57
<b><u>1.3.2 :</u></b>	Profil de la croissance cellulaire en fonction des différentes méthodes de culture .....	57
<b><u>1.3.3 :</u></b>	Formation des ponts disulfures grâce à l'activité de la PDI et de l'Ero1 pour les cellules eucaryotes .....	61
<b><u>1.3.4 :</u></b>	Profil de glycosylation humaine et pourcentage de présence des différents glycanes ajoutés sur le motif constant .....	63
<b><u>1.3.5 :</u></b>	Schéma représentant la création d'une cellule hybridome de souris productrice d'anticorps .....	64
<b><u>1.3.6 :</u></b>	Les voies possibles pour la synthèse de protéines à ponts disulfures dans le périplasme ...	71
<b><u>1.4.1 :</u></b>	Format de glycosylations minimales et maximales .....	75
<b><u>1.4.2 :</u></b>	Formation de pyroglutamate en N-terminal des protéines à partir de la glutamine et du glutamate .....	76
<b><u>1.4.3 :</u></b>	Désamidation de l'asparagine, et isomérisation de l'aspartate en passant par l'intermédiaire succinimide .....	77
<b><u>1.4.4 :</u></b>	Les deux étapes de l'oxydation des méthionines .....	78
<b><u>1.4.5 :</u></b>	Schéma représentant la glycation des protéines avec la formation d'un intermédiaire de Schiff, un produit d'Amadori et la formation de AGE .....	79
<b><u>1.4.6 :</u></b>	Structure PDB 1HZH de l'IgG1 B12 .....	82
<b><u>1.4.7 :</u></b>	Schéma représentatif du principe de fonctionnement de la HDX-MS pour l'étude d'interaction anticorps/antigène par cette méthode .....	84
<b><u>1.4.8 :</u></b>	Superposition des spectres RMN 1D <sup>1</sup> H (600MHz) de l'infliximab natif .....	85
<b><u>1.4.9 :</u></b>	SOFAST-HMQC <sup>1</sup> H- <sup>15</sup> N de l'anticorps NIST en abondance naturelle .....	87
<b><u>1.4.10 :</u></b>	Les six acides aminés contenant des groupements méthyles .....	88
<b><u>1.4.11 :</u></b>	Structure des fragments Fab et Fc avec les groupements méthyles représentés en rouge ...	89

<b><u>1.4.12 :</u></b>	Zones des déplacements chimiques théoriques des différents types de groupements méthyles .....	89
<b><u>1.4.13 :</u></b>	Spectre RMN-HSQC 2D $^1\text{H}$ - $^{13}\text{C}$ du mAb NIST .....	90
<b><u>1.4.14 :</u></b>	Spectres RMN-HSQC 2D $^1\text{H}$ - $^{13}\text{C}$ du mAb NIST et de ses fragments .....	91
<b><u>1.5.1 :</u></b>	Schéma représentant les différentes étapes du transfert d'aimantation intra-résidu, illustrées pour une leucine marquée $\text{U}-[{}^2\text{H}, {}^{15}\text{N}, {}^{13}\text{C}]\text{-Leu}-[{}^{13}\text{C}^1\text{H}_3]\delta 2/[{}^{12}\text{C}^2\text{H}_3]\delta 1$ .....	97
<b><u>1.5.2 :</u></b>	Chemin emprunté par l'aimantation au cours d'une expérience de type HCC si le marquage isotopique $^{13}\text{C}$ n'est pas linéaire .....	98
<b><u>1.5.3 :</u></b>	Histogramme résumant les performances de différents programmes d'attribution automatique des groupements méthyles pour cinq protéines .....	102
<b><u>3.1.1 :</u></b>	Spectres HMQC $^1\text{H}$ - $^{13}\text{C}$ des groupements méthyles des méthionines de la protéine $\beta 2\text{AR}$ marquée $[\alpha\beta\text{-}{}^2\text{H}, \text{methyl-}{}^{13}\text{C}\text{-met}]$ .....	125
<b><u>3.1.2 :</u></b>	Intérêt de la deutération des chaînes latérales des résidus marqués sur les groupements méthyles .....	126
<b><u>3.1.3 :</u></b>	Région des groupements amides du spectre HSQC $^1\text{H}$ - $^{15}\text{N}$ du fragment Fv marqué $^{15}\text{N}$ de l'IgG2a anti-dansyl .....	127
<b><u>3.1.4 :</u></b>	Spectre RMN HSQC $^1\text{H}$ - $^{15}\text{N}$ du fragment Fc de l'IgG1 anti-CCR4 marqué $\text{U}-({}^{13}\text{C}, {}^{15}\text{N})$ .....	128
<b><u>3.1.5 :</u></b>	Spectres méthyl TROSY d'une IgG de souris marqué isotopiquement .....	129
<b><u>3.1.6 :</u></b>	Profil d'élution obtenu pour la purification de l'anticorps anti-LAMP1 .....	132
<b><u>3.1.7 :</u></b>	Profil d'élution obtenu pour la purification de l'anticorps anti-LAMP1 .....	132
<b><u>3.1.8 :</u></b>	Rendement d'anticorps exprimé par les cellules CHO en utilisant le milieu commercial dépourvu en acides aminés et les mélanges AADMEM, AAOptiCHO et AAInsect .....	134
<b><u>3.1.9 :</u></b>	Pourcentages de présence des acides aminés dans les mélanges AAInsect et AAOptimAb comparés à ceux dans la séquence de l'anticorps anti-LAMP1 .....	136
<b><u>3.1.10 :</u></b>	Rendement d'anticorps exprimé par les cellules CHO en utilisant le milieu commercial dépourvu en acides aminés et le mélange AAOptimAb développé et optimisé comparé à ceux obtenus pour les trois mélanges précédents .....	136
<b><u>3.1.11 :</u></b>	Profil théorique de croissance des cellules CHO en fonction de la consommation des réactifs du milieu et de la production d'anticorps associée en mode batch en commençant à basse densité cellulaire a) et à haute densité cellulaire b) .....	137
<b><u>3.1.12 :</u></b>	Rendements d'expression de l'IgG1 anti-LAMP1 par les cellules CHO en fonction des milieux de culture utilisés, et de la densité cellulaire initiale .....	138
<b><u>3.1.13 :</u></b>	Impact de l'affamement sur l'intensité des pics RMN $^1\text{H}$ - $^{13}\text{C}$ des isoleucines marquées ${}^{13}\text{CH}_3\text{-}\delta 1$ .....	140
<b><u>3.1.14 :</u></b>	Protocole de culture cellulaire développé pour la production de l'anticorps anti-LAMP1 exprimé et marqué en cellules CHO .....	140
<b><u>3.1.15 :</u></b>	Impact de la quantité d'isoleucine marquée ajoutée lors de la culture sur l'intensité de ses pics sur les spectres RMN 2D $^1\text{H}$ - $^{13}\text{C}$ .....	141

<b><u>3.1.16 :</u></b>	Taux d'incorporation des acides aminés marqués et comportant un groupement méthyle déterminés par spectrométrie de masse .....	142
<b><u>3.1.17 :</u></b>	Impact du type d'inhibiteur et de sa concentration sur l'intensité des pics RMN $^1\text{H}$ - $^{13}\text{C}$ des alanines .....	144
<b><u>3.1.18 :</u></b>	Spectres $^1\text{H}$ - $^{13}\text{C}$ -SOFAST de l'IgG1 anti-LAMP1 marquée U- $[\text{}^1\text{H}]$ , $[\text{}^{13}\text{C}^1\text{H}_3]\epsilon$ -Met et U- $[\text{}^2\text{H}]$ , $[\text{}^{13}\text{C}^1\text{H}_3]\beta$ -alanine et produite en cellules CHO en absence d'inhibiteur de l'alanine transaminase a) et en présence de 5 mg/L de L-cyclosérine b) .....	144
<b><u>3.1.19 :</u></b>	Spectres RMN 2D $^1\text{H}$ - $^{13}\text{C}$ -SOFAST dans la région des groupements méthyles de l'anticorps anti-LAMP1 produit en cellules CHO et marqué $[\text{}^{13}\text{C}^1\text{H}_3]$ .....	146
<b><u>3.2.1 :</u></b>	Principaux composants du système <i>cell-free</i> .....	150
<b><u>3.2.2 :</u></b>	Système de transcription et traduction couplées du système <i>cell-free</i> avec l'exemple de co-expression de deux plasmides comme pour le cas de la chaîne lourde et légère d'un fragment Fab d'anticorps .....	151
<b><u>3.2.3 :</u></b>	Transformation de l'ATP en ADP représentant la source d'énergie utilisée en système <i>cell-free</i> .....	152
<b><u>3.2.4 :</u></b>	Schéma représentant le mécanisme principal du système <i>cell-free</i> .....	153
<b><u>3.2.5 :</u></b>	Fuite isotopique d'un $^{15}\text{N}$ d'un acide aminé à un autre par l'action d'une transaminase dépendante du facteur pyridoxal 5-phosphate (PLP) avec un keto-acide .....	155
<b><u>3.3.1 :</u></b>	Histogramme représentant l'optimisation du rendement d'expression du fragment Fab de l'anticorps anti-ALMP1 en système <i>cell-free</i> grâce à l'ajout de DsbC recombinante .....	192
<b><u>3.3.2 :</u></b>	Rendements d'expression du fragment Fab obtenus en dialyse <i>cell-free</i> en utilisant des ratio "réaction/dialyse" différents .....	193
<b><u>3.3.3 :</u></b>	Chromatogramme obtenu pour la purification du fragment Fab sur une colonne Hi-Trap protein L de 1mL .....	194
<b><u>3.3.4 :</u></b>	Gel SDS-page des fractions de collections contenant les pics d'élution du fragment Fab purifié sur une colonne de streptactine suivie d'une colonne de Ni-NTA a), et purifié en utilisant la colonne Hi-Trap protein L b) .....	194
<b><u>3.3.5 :</u></b>	Chromatogramme obtenu en utilisant une colonne chromatographique d'exclusion de taille Superdex 75 10/300 pour la purification du fragment Fab exprimé en système <i>cell-free</i> et purifié une première fois sur protéine L .....	195
<b><u>3.3.6 :</u></b>	Marquage isotopique des groupements méthyles réalisé pour le marquage sélectif des acides aminés comportant un groupement méthyle .....	196
<b><u>3.3.7 :</u></b>	Synthèse des acides aminés isoleucines et valines à partir de leur précurseur par l'enzyme BCAT (branched-chain amino acid amino-transferase) .....	197
<b><u>3.3.8 :</u></b>	Spectres RMN 2D methyl-SOFAST du fragment Fab marqué sur les groupements méthyles et exprimé en système <i>cell-free</i> .....	198
<b><u>3.3.9 :</u></b>	Superposition des spectres RMN 2D methyl-SOFAST du fragment Fab marqué sur les méthyles et exprimé en système <i>cell-free</i> avec ceux de l'anticorps anti-LAMP1 marqué et exprimé en cellules CHO .....	200

<b><u>3.3.10:</u></b>	Marquage isotopique réalisé sur les résidus alanines, isoleucines et valines pour l'acquisition des expériences RMN "out and back" de type HCC .....	202
<b><u>3.3.11:</u></b>	Spectre RMN 2D méthyl SOFAST du fragment Fab perdeutééré et exprimé en système <i>cell-free</i> avec un marquage isotopique <sup>13</sup> C linéaire pour les alanines, isoleucines- $\delta$ 1 et valines- <i>proR</i> .....	202
<b><u>3.3.12:</u></b>	Chemins empruntés par la magnétisation lors de son transfert pour les expériences RMN "out and back" de type HCC .....	203
<b><u>3.3.13:</u></b>	Exemples d'extraits 2D des expériences RMN 3D "out and back" HCC et HC(C)C et 3D HNCA et HNCACB .....	205
<b><u>3.3.14:</u></b>	Groupements méthyles des alanines, isoleucines et valines attribués en utilisant les expériences RMN HCC "out and back" .....	207
<b><u>3.3.15:</u></b>	Marquage isotopique réalisé sur les acides aminés comportant un groupement méthyle pour l'acquisition des expériences RMN méthyles NOESY .....	208
<b><u>3.3.16:</u></b>	Spectre RMN 2D méthyl SOFAST du fragment Fab de l'anticorps anti-LAMP1 produit en système <i>cell-free</i> perdeutééré .....	209
<b><u>3.3.17:</u></b>	Attribution du type de résidu pour les résonances observées sur le spectre RMN 2D méthyl SOFAST du fragment Fab marqué U-[ <sup>2</sup> H], <sup>13</sup> CH <sub>3</sub> -Ala, Met, Thr, Leu- <i>proS</i> , Val- <i>proR</i> , Ile- $\delta$ 1 .....	211
<b><u>3.3.18:</u></b>	Corrélations NOEs inter-méthyles du fragment Fab de l'anticorps anti-LAMP1 .....	212
<b><u>3.3.19:</u></b>	Représentation des groupements méthyles des thréonines sur la structure du fragment Fab de l'anticorps anti-LAMP1 .....	213
<b><u>3.3.20:</u></b>	Attribution des groupements méthyles des alanines, méthionines, thréonines, leucine- <i>proS</i> , valines- <i>proR</i> et isoleucines- $\delta$ 1 du fragment Fab de l'anticorps anti-LAMP1 exprimé en système <i>cell-free</i> et marqué U-[ <sup>2</sup> H], Ala-[ <sup>13</sup> C <sup>1</sup> H <sub>3</sub> ] $\beta$ , Met-[ <sup>13</sup> C <sup>1</sup> H <sub>3</sub> ] $\epsilon$ , Leu-[ <sup>13</sup> C <sup>1</sup> H <sub>3</sub> ] <i>proS</i> , Val-[ <sup>13</sup> C <sup>1</sup> H <sub>3</sub> ] <i>proR</i> , Ile-[ <sup>13</sup> C <sup>1</sup> H <sub>3</sub> ] $\delta$ 1, Thr-[ <sup>13</sup> C <sup>1</sup> H <sub>3</sub> ] $\gamma$ .....	214
<b><u>4.1.1:</u></b>	Spectres RMN 2D <sup>1</sup> H- <sup>13</sup> C-SOFAST attribués dans la région des groupements méthyles de l'anticorps anti-LAMP1 produit en cellules CHO et marqué [ <sup>13</sup> CH <sub>3</sub> ]-méthionine et : a) [U- <sup>2</sup> H+ <sup>13</sup> CH <sub>3</sub> ]-alanine, b) [U- <sup>2</sup> H+ <sup>13</sup> CH <sub>3</sub> - $\delta$ 1]-isoleucine, c) [U- <sup>2</sup> H+ <sup>13</sup> CH <sub>3</sub> - <i>proS</i> ]-leucine, d) [U- <sup>2</sup> H+ <sup>13</sup> CH <sub>3</sub> ]-thréonine, e) [U- <sup>2</sup> H+ <sup>13</sup> CH <sub>3</sub> - <i>proR</i> ]-valine .....	219

# Liste des tableaux

<b><u>1.1.1 :</u></b>	Tableau récapitulatif des différents isotypes d'anticorps et leurs fonctions .....	39
<b><u>1.2.1 :</u></b>	Dénomination commune internationale des anticorps monoclonaux thérapeutiques .....	41
<b><u>1.5.1 :</u></b>	Mutations suggérées par la matrice BLOSUM pour les acides aminés comportant un groupement méthyle .....	99
<b><u>2.1.1 :</u></b>	Conditions de cultures conseillées par GTP Technology <sup>TM</sup> pour l'expression de l'anticorps anti-LAMP1 en mode <i>fed-batch</i> avec lignée cellulaire CHO développée .....	105
<b><u>2.1.2 :</u></b>	Conditions de cultures mises en place pour la production de l'anticorps anti-LAMP1 en mode batch avec la lignée cellulaire CHO développée par GTP Technology .....	106
<b><u>2.2.1 :</u></b>	Liste des réactifs <i>cell-free</i> et de leur concentration .....	110
<b><u>2.2.2 :</u></b>	Liste des acides aminés non marqués utilisés en système <i>cell-free</i> .....	111
<b><u>2.2.3 :</u></b>	Liste des acides aminés marqués utilisés en système <i>cell-free</i> pour l'expression du fragment Fab perdeutééré en système <i>cell-free</i> deutéré .....	112
<b><u>2.2.4 :</u></b>	Liste des acides aminés et précurseurs marqués sur les groupements méthyles synthétisés à façon par NMR-Bio et utilisés pour la production du fragments Fab marqué U-[ <sup>2</sup> H], <sup>13</sup> CH <sub>3</sub> -Ala, Met, Thr, Leu-proS, Val-proR, Ile- $\delta$ 1 .....	113
<b><u>2.2.5 :</u></b>	Liste des acides aminés et des précurseurs marqués sur les groupements méthyles synthétisés à façon par NMR-Bio pour la production du fragment Fab perdeutééré et avec un marquage <sup>13</sup> C linéaire pour les alanines, isoleucines- $\delta$ 1 et valines-proR .....	113
<b><u>2.2.6 :</u></b>	Liste des expériences RMN acquise avec la fréquence <sup>1</sup> H du spectromètre utilisé et le temps d'acquisition de l'expérience .....	116
<b><u>3.1.1 :</u></b>	Composition finale optimisée pour les 20 acides aminés .....	135
<b><u>3.1.2 :</u></b>	Nombre de pics identifiés sur les spectres 2D <sup>1</sup> H- <sup>13</sup> C SOFAST de l'IgG1 anti-LAMP1 marquée spécifiquement sur les groupements méthyles et produite en cellules CHO .....	147
<b><u>3.2.1 :</u></b>	Liste des principales enzymes présentes dans l'extrait bactérien responsables de fuites isotopiques entre acides aminés marqués, des inhibiteurs pouvant être utilisés et des acides aminés concernés par les fuites .....	157
<b><u>3.3.1 :</u></b>	Nombre de pics identifiés sur les spectres 2D <sup>1</sup> H- <sup>13</sup> C SOFAST du fragment Fab marqué spécifiquement sur les groupements méthyles et exprimé en système <i>cell-free</i> et occurrence de ces résidus dans la séquence de l'anticorps .....	199
<b><u>3.3.2 :</u></b>	Nombre de résonances attribuées pour les C $\alpha$ et C $\beta$ des alanines, isoleucines et valines dans les expériences RMN 3D HNCA et HNCACB avec le nombre de corrélations jusqu'aux groupements méthyles obtenu avec les expériences RMN HCC et l'occurrence de ces acides aminés dans la séquence du fragment Fab de l'anticorps anti-LAMP1 .....	204
<b><u>3.3.3 :</u></b>	Nombre de résonances observées dans les expériences RMN HCC "out and back" correspondant à des C $\alpha$ et des C $\beta$ et l'occurrence de ces acides aminés dans la séquence du fragment Fab de l'anticorps anti-LAMP1 .....	206

# Préambule

La structure hautement ordonnée (HOS) d'une protéine thérapeutique est un paramètre clé à caractériser lors de l'évaluation de la robustesse du processus de bio production pharmaceutique et sa mesure représente un défi analytique important. En effet, les méthodes biophysiques actuellement disponibles ne sont pas suffisantes pour détecter des changements de conformation subtils. De ce fait, il est difficile d'expliquer l'origine structurale de la perte d'activité observée lors d'études de stress de ces molécules biothérapeutiques et cela rend complexe la compréhension de leurs voies de dégradation. Ce projet a pour objectif d'explorer les capacités de la RMN afin de renforcer la panoplie d'outils analytiques permettant de caractériser les produits biothérapeutiques tels que les mAbs et ainsi d'accélérer le développement des produits biothérapeutiques de Sanofi.

Dans ce contexte, l'objectif de cette thèse est de développer une méthode de marquage isotopique des groupements méthyles d'un anticorps exprimé en cellules CHO permettant d'étudier, par RMN en solution, sa structure à l'échelle atomique. Cela requiert une étape primordiale qui est l'attribution. Il a donc été nécessaire de mettre au point des outils et des protocoles permettant d'attribuer cette protéine de 150 kDa.

Cette thèse fait l'objet d'un partenariat entre Sanofi (site de Vitry sur Seine), l'Institut de Biologie Structurale (Grenoble), et NMR-Bio (Grenoble). L'introduction de cette thèse présente le contexte autour des anticorps thérapeutiques qui a mené à ce sujet de thèse. J'ai, pour cela, choisi de rédiger une première partie traitant des anticorps physiologiques et détaillant leur structure et leur fonction dans l'organisme. La deuxième partie décrit les différents anticorps thérapeutiques ainsi que leurs modes d'action dans la lutte contre les maladies. Au cours de la thèse, plusieurs systèmes d'expression de protéines recombinantes ont été utilisés pour la production de l'anticorps anti LAMP1 (anticorps modèle de la thèse) ou de ses fragments. Ainsi, j'ai choisi de rédiger une troisième partie traitant des différents systèmes de production des anticorps thérapeutiques et des problématiques associées. La quatrième partie de l'introduction concerne l'analyse des caractéristiques physico-chimiques et structurales des anticorps thérapeutiques. Cette partie détaille les techniques analytiques actuellement utilisées, ainsi que leurs limitations qui soulignent le besoin de développer une approche permettant d'utiliser la RMN en solution pour la caractérisation structurale des anticorps à l'échelle atomique. Enfin la dernière partie de l'introduction détaille la problématique du sujet de thèse et la stratégie adoptée pour ce projet.

Les résultats sont illustrés dans 3 grandes parties. La première regroupe tous les résultats d'expression, de marquage isotopique et d'analyses RMN obtenus sur l'anticorps entier produit en cellules CHO. La deuxième partie est adossée à un article décrivant la synthèse du fragment Fab de l'anticorps en système acellulaire et l'attribution du squelette de ce fragment. Enfin, la dernière partie traite du marquage des méthyles et de leur attribution sur ce même fragment Fab, toujours synthétisé en système *cell-free*.

# I. Introduction

## 1. Les anticorps physiologiques

Les anticorps sont des immunoglobulines (Ig), des protéines plasmatiques (présentes dans le plasma sanguin) de la famille des globulines et impliquées dans l'immunité. Les anticorps ont été découverts en 1890 lorsque Behring Emil Adolf Von et Shibasaburo Kitasato ont démontré que le sang d'animaux immunisés contre la diphtérie contenait des antitoxines capables d'inactiver spécifiquement la toxine diphtérique [1]. En 1891, la première référence à ces agents neutralisants est faite : ils sont appelés les anticorps. Von Behring reçoit, pour son travail, le prix Nobel de médecine en 1901.

Ensuite, en 1958, Rodney R. Porter a déterminé la structure chimique d'un anticorps en utilisant la papaïne, une enzyme de la papaye, pour le digérer et mieux en comprendre le fonctionnement[2]. Il pensait au départ que les anticorps n'étaient constitués que d'une unique chaîne polypeptidique et a proposé leur structure en « Y » désormais bien connue. Ce sont les travaux de Gerald M. Edelman en 1961 qui ont pu déterminer la présence de ponts disulfures reliant plusieurs chaînes protéiques. Rodney R. Porter et Gerald M. Edelman ont reçu, pour leurs travaux, le prix Nobel de physiologie et de médecine en 1972.

### 1.1. Structure des immunoglobulines de classe G1 : IgG1

Plusieurs isotypes d'anticorps existent (IgA, IgD, IgM, IgE et IgG) et leur structure ainsi que leur fonction varient. Mais la structure en « Y » d'un monomère d'anticorps est commune à toutes les classes. Les IgG sont la classe prédominante des immunoglobulines dans l'organisme d'une personne saine (environ 75% des anticorps présents dans le plasma sanguin) [3]. Il s'agit également de la classe la plus étudiée et la mieux connue et sa structure constitue une base élémentaire idéale pour détailler la structure des différents isotypes d'anticorps. L'organisation typique d'une IgG est détaillée dans les paragraphes suivants.

#### 1.1.1. Structure primaire

Les IgG1 sont des protéines hétérodimériques de 150 kDa composées de deux chaînes lourdes (HC) et de deux chaînes légères (LC), elles-mêmes constituées de différents domaines d'immunoglobuline. Ces domaines ont une longueur d'environ 110 acides aminés et ont une structure spécifique aux immunoglobulines. Tous les isotypes d'immunoglobulines sont composés de domaines constants, dont la séquence en acides aminés est conservée d'un anticorps à un autre, et de domaines variables (VH sur la chaîne lourde ou VL sur la chaîne légère) qui sont spécifiques à chaque anticorps et sont responsables de la spécificité à l'antigène. Ces domaines variables possèdent une partie hautement variable qui représente environ 25% de leur séquence.

### a Structure primaire des chaînes légères

Les chaînes légères de 25 kDa chacune sont composées de deux domaines : un domaine variable (VL) dont 25% de la séquence est variable et spécifique d'une IgG1 à une autre, et un domaine constant (CL). Deux types de chaînes légères existent :  $\kappa$  et  $\lambda$ . Ces chaînes diffèrent au niveau de leur domaine constant et sont exprimées par deux gènes différents dans les cellules immunitaires (lymphocytes B), cependant les deux types de chaînes légères sont capables de s'appareiller à n'importe quel type de chaîne lourde d'anticorps.

### b Structure primaire des chaînes lourdes

L'isotype de l'immunoglobuline est défini selon la chaîne lourde de l'anticorps. Pour les IgG, les chaînes lourdes, de 50 kDa chacune, sont composées de quatre domaines : un domaine variable VH qui, comme la VL de la chaîne légère, comporte 25% de séquences variables et spécifiques, et trois domaines constants (CH1, CH2 et CH3) dont la séquence est très conservée dans toutes les IgG1. Entre les domaines CH1 et CH2 se trouve la région charnière qui comporte une séquence peptidique avec deux cystéines formant des ponts disulfures inter-chaînes, ainsi qu'une séquence d'acides aminés flexibles.

### c Nomenclature des différents fragments

Les domaines des chaînes lourdes et légères se combinent et forment les fragments des anticorps. La découverte de ces fragments a été réalisée par clivage protéolytique à l'aide d'enzymes telles que la pepsine et la papaïne (figure 1.1.1), initialement utilisées pour étudier les fonctions des différentes parties des anticorps.

En effet, la papaïne (enzyme végétale provenant de la papaye) a été utilisée par Rodney R. Porter en 1958 et a permis d'obtenir deux fragments Fab (*antigen binding fragment*) et un fragment Fc (*crystallisable fragment*). Le site d'action de cette enzyme se situe au niveau de la région charnière de l'anticorps, en C-terminal des domaines CH1.

La pepsine clive les domaines CH2 des anticorps en libérant un fragment F(ab')<sub>2</sub> constitué des deux fragments Fab reliés par la région charnière.

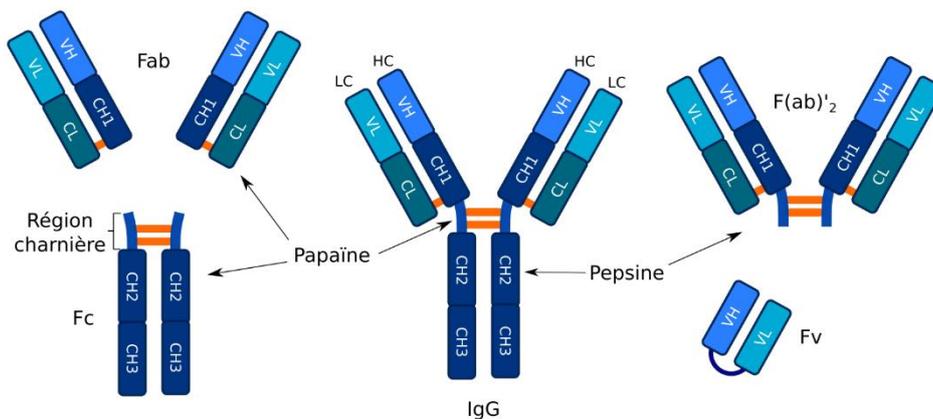


Figure 1.1.1 : Nomenclature des différents domaines des IgG et résultats de clivage à la pepsine et à la papaïne. La chaîne lourde comporte trois domaines constants (CH1-3) et un domaine variable (VH), la chaîne légère présente un domaine constant (CL) et un domaine variable (VL). En orange, les ponts disulfures inter-chaînes. HC : Heavy Chain, chaîne lourde, LC : Light Chain, chaîne légère, Fab : fragment d'anticorps, Fc : fragment constant, Fv : fragment variable.

Le fragment Fc est donc composé des domaines CH2 et CH3 de deux chaînes lourdes reliées par les ponts disulfures de la région charnière. Les fragments Fab sont constitués des domaines VH, CH1 et VL, CL, des chaînes lourdes et légères respectivement, reliés par un pont disulfure au niveau de la partie C-terminale de CL et CH1.

L'association des domaines VH et VL forment le fragment Fv (*variable fragment*) lorsqu'ils sont reliés artificiellement par une chaîne polypeptidique.

### 1.1.2. Structure secondaire et tertiaire

Les domaines d'immunoglobulines sont composés de deux feuillets  $\beta$  antiparallèles rapprochés. Au sein des domaines constants, l'un de ces feuillets  $\beta$  est formé de quatre brins  $\beta$  nommés  $\downarrow A$ ,  $\uparrow B$ ,  $\downarrow E$ ,  $\uparrow D$ , tandis que l'autre en contient trois, nommés  $\downarrow C$ ,  $\uparrow F$ ,  $\downarrow G$  (les flèches correspondent au sens de lecture N-terminal  $\rightarrow$  C-terminal). Ces deux feuillets  $\beta$  sont reliés de manière covalente par la présence d'un pont disulfure intra-domaine reliant deux cystéines positionnées dans les brins  $\beta$   $\uparrow B$  et  $\downarrow F$ , ainsi que par les interactions non-covalentes des chaînes latérales des acides aminés à la surface des feuillets  $\beta$ .

Les domaines variables sont composés de deux brins  $\beta$  supplémentaires nommés  $\downarrow C''$ ,  $\uparrow C'$  qui sont situés entre les brins  $\beta$   $\uparrow C$  et  $\uparrow D$ . Le premier feuillet  $\beta$  possède une composition similaire aux domaines constants, tandis que le deuxième contient cinq brins  $L\beta$  ( $\downarrow C''$ ,  $\uparrow C'$ ,  $\downarrow C$ ,  $\uparrow F$ ,  $\downarrow G$ ). Les feuillets  $\beta$  des domaines variables sont, de la même façon que pour les domaines constants, liés de manière covalente par un pont disulfure intra-domaine entre deux cystéines des brins  $\beta$   $\uparrow B$  et  $\downarrow F$  et interagissent de manière non covalente grâce à l'interaction des chaînes latérales des acides aminés à la surface des feuillets  $\beta$ .

Pour les domaines constants et variables, les différents brins  $\beta$  sont reliés entre eux par des boucles flexibles. Les domaines variables VH et VL des chaînes lourdes et légères comportent chacun trois boucles dont le nombre et la composition en acides aminés varient énormément d'un anticorps à un autre. Il s'agit des 25% de séquences les plus variables dans les anticorps. Ces boucles sont responsables de la reconnaissance spécifique à l'antigène et sont connues sous le nom de CDRs (complementary-determining regions). Les boucles des domaines VH sont nommées CDRH1, CDRH2, CDRH3 et celles des domaines VL sont nommées CDRL1, CDRL2, CDRL3 (figure 1.1.2).

Pour les domaines variables (VH et VL), la conformation des feuillets  $\beta$  est telle que les CDRs sont alignés, formant ainsi le site de reconnaissance aux antigènes dans les fragments Fab des anticorps.

Les domaines variables et constants des chaînes légères et lourdes sont reliés entre eux par une région charnière. L'angle entre les domaines VL et CL est caractéristique pour les chaînes légères  $\kappa$  et  $\lambda$ . Il est situé entre  $116^\circ$  et  $126^\circ$  pour les chaînes légères  $\kappa$ . Pour les chaînes légères  $\lambda$ , la présence d'une glycine supplémentaire dans la séquence implique une variabilité plus importante de cet angle. Pour les domaines VH et CH1 des chaînes lourdes, cet angle est de  $180^\circ$  et est défini par l'interaction des chaînes latérales d'acides aminés hautement conservés dans la séquence des IgG : les acides aminés Phe 148 et Pro 149 du domaine VH avec les Thr110, Leu/Val111 et Ser112 du domaine CH1.

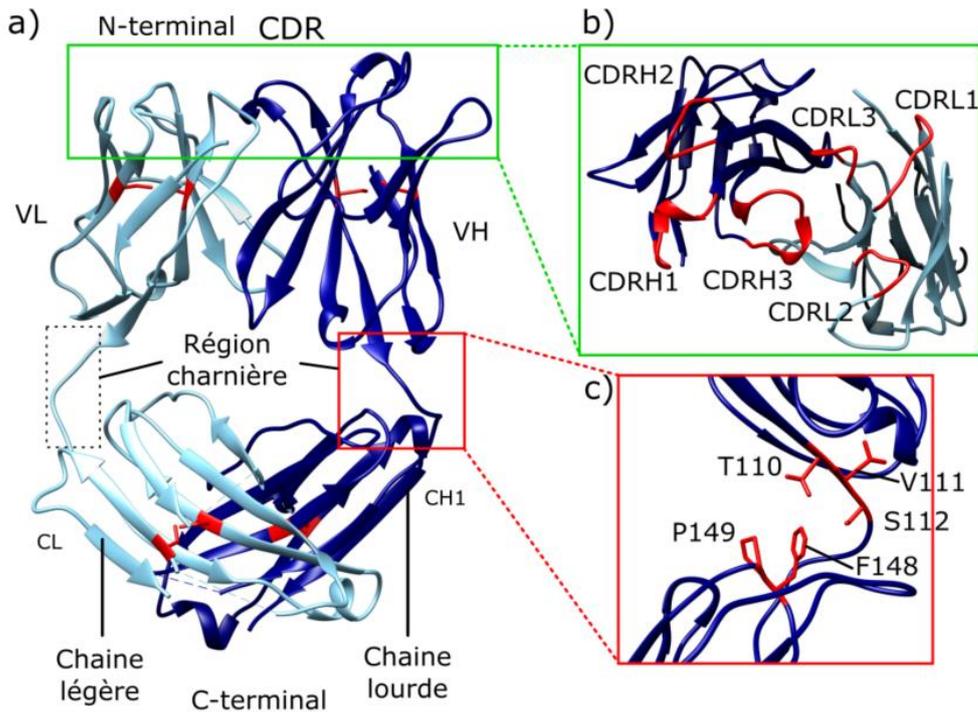


Figure 1.1.2 : a) Structure du fragment Fab de l'anticorps anti-LAMP1 (PDB : 8ATH) les cystéines formant des ponts disulfures intra-domaines sont représentées en rouge. b) Vue du dessus du fragment Fab. Les CDR sont représentés en rouges. c) Zoom sur la région charnière reliant les domaines VH et CH1. Les acides aminés dont l'interaction des chaînes latérales est responsable de l'alignement de ces deux domaines sont représentés en rouge.

La structure des domaines composant le fragment Fc d'une IgG (CH2 et CH3), est semblable à celle des domaines CH1, c'est-à-dire deux feuillets  $\beta$  composés de 4 et 3 brins  $\beta$ . Leur alignement l'un par rapport à l'autre est très conservé dans les différents isotopes d'IgG, et ce, grâce à la conservation des acides aminés Leu 251 (domaine CH2) avec les acides aminés Met 428, His 429, Glu 430 et His 435 (domaine CH3), dont les chaînes latérales interagissent entre elles (figure 1.1.3).

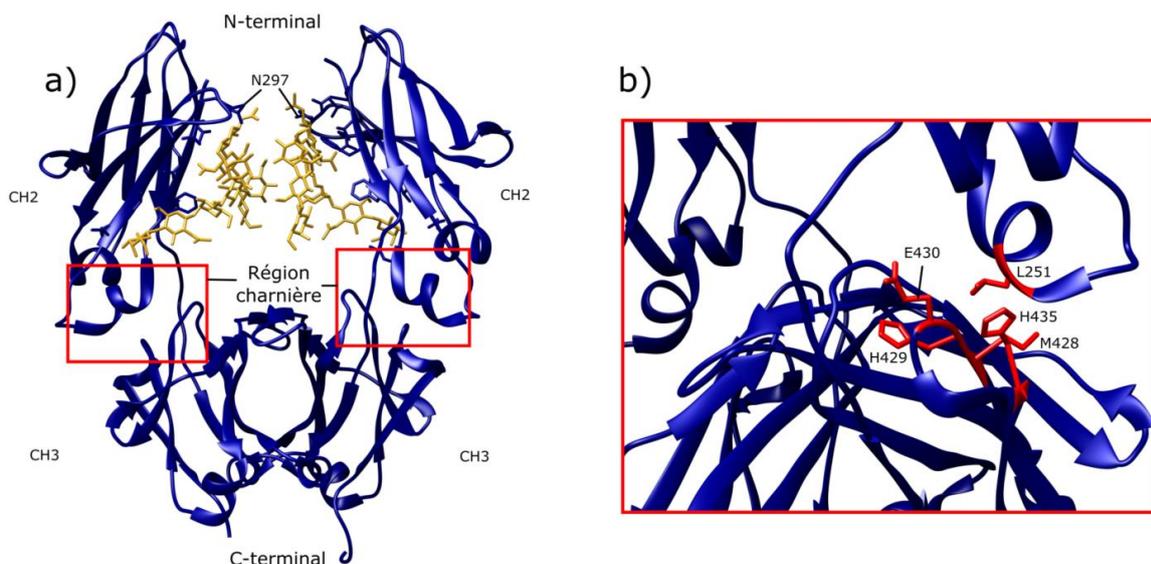


Figure 1.1.3 : a) Structure des fragments Fc des IgG (PDB : 1HZH). Les glycosylations des asparagines 297 sont représentées en jaune. b) zoom sur la région charnière reliant les domaines CH2 et CH3. Les acides aminés dont l'interaction des chaînes latérales est responsable de l'alignement de ces deux domaines sont représentés en rouge.

### 1.1.3. Structure quaternaire

Les chaînes légères et lourdes (pour les fragments Fab) sont reliées entre elles de manière covalente par la présence d'un pont disulfure inter-domaine situé en C-terminal des domaines CL et CH1. Les deux chaînes sont reliées de manière non covalente par la présence de liaisons hydrogènes présentes au cœur de l'interaction des domaines VL/VH et CL/CH1.

Les chaînes lourdes des anticorps sont reliées entre elles de manière covalente par la présence de ponts disulfures inter-domaines reliant des cystéines présentes dans la région charnière entre les domaines CH1 et CH2 (figure 1.1.4). Ces cystéines sont hautement conservées dans les anticorps et leur nombre ainsi que leur position sont spécifiques à chaque isotype d'immunoglobuline [4]. Le détail de leur nombre et de leur conformation pour les autres isotypes d'immunoglobulines est donné dans la partie 1.3.

Pour les fragments Fc des anticorps, les chaînes lourdes ne sont pas liées de manière covalente. Les deux domaines CH3 sont proches dans l'espace et sont liés par des liaisons hydrogènes tandis que les domaines CH2 sont plus éloignés l'un de l'autre [5]. Les domaines CH2 sont cependant liés l'un à l'autre car l'espace qui les sépare est occupé par les glycanes des asparagines 297. Ces glycanes, qui sont hautement conservés dans les IgG, sont cruciaux pour l'activité des anticorps, ainsi que pour leur stabilité. En effet, ils interagissent entre eux et permettent le maintien de la structure des fragments Fc d'anticorps grâce à des liaisons hydrogènes et en piégeant des molécules d'eau (interactions hydrophiles). La flexibilité des domaines CH2 entre eux, rendue possible grâce à ces glycosylations, contribue à l'activité des anticorps avec les différentes cellules du système immunitaire (cf. partie 1.1.4).

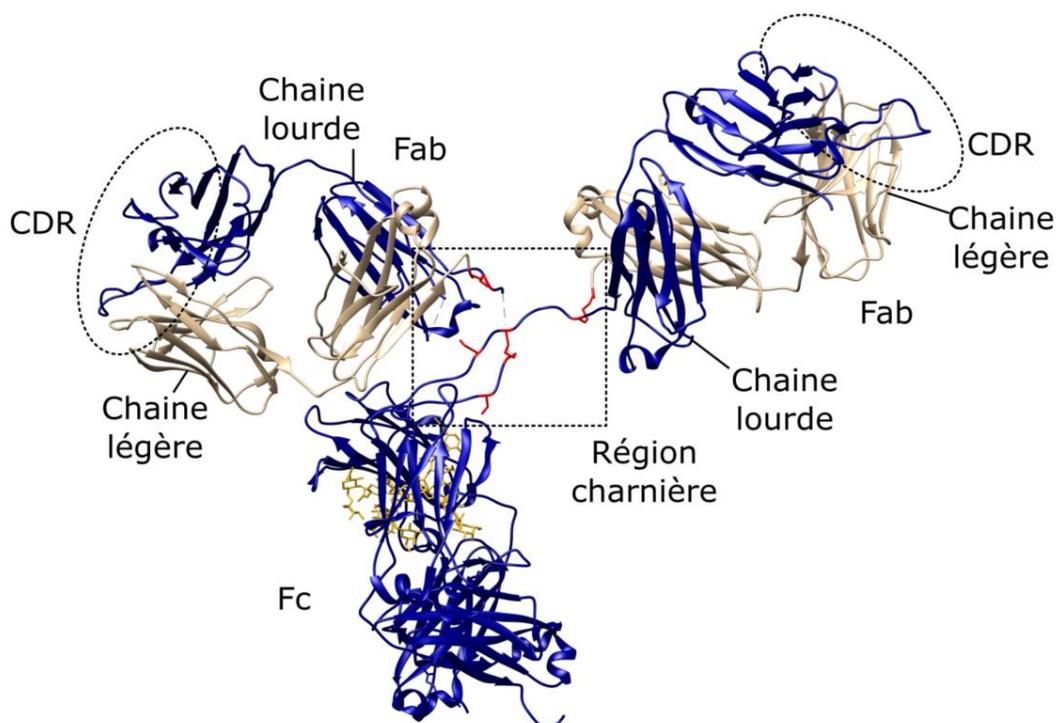


Figure 1.1.4 : Structure d'une IgG (PDB : 1HZH). Les cystéines formant des ponts disulfures inter-domaines sont représentées en rouge au niveau de la région charnière. Dans cette structure, deux ponts disulfures relient les chaînes légères aux chaînes lourdes. La structure cristallographique présente un seul des deux ponts disulfures reliant les chaînes lourdes entre elles, l'absence du second pouvant être expliquée par la flexibilité de la région charnière très dynamique, présentant donc une structure complexe à déterminer par cristallographie.

## 1.2. Les glycosylations d'IgG

Le type de glycosylation impacte la structure et l'activité des immunoglobulines. Plusieurs types de glycosylations existent et, suivant les différents isotopes d'anticorps, leur position et leur nombre varient (cf. partie 1.1.3).

### 1.2.1. Structure de base des N-glycosylations et O-glycosylations

Les N-glycosylations (N-Gly) résultent de l'ajout d'un premier oligosaccharide (ou glycane) sur les chaînes latérales des asparagines si elles sont suivies de la séquence Asn-X-Ser/Thr et X n'étant pas une proline. Les N-glycosylations débutent avec l'ajout d'une N-acétylglucosamine (GlcNAc), puis ont une succession de glycanes communs composés de deux GlcNAc consécutifs, de trois mannoses (Man) (dont deux sont respectivement liés au premier) puis deux GlcNAc à la suite des mannoses terminaux (figure 1.1.5).

Lorsque les glycanes sont liés au groupe hydroxyle de la chaîne latérale des sérines ou thréonines dans des régions riches en séquence Ser/Thr/Pro, on parle de O-glycosylations qui commencent par l'ajout d'une N-acétylgalactosamine (GalNAc).

Les anticorps possèdent de nombreux sites de N-glycosylations très conservés d'une classe d'immunoglobulines à l'autre, mais comportent en revanche peu de O-glycosylations (réservées à deux classes d'immunoglobulines (cf. partie 1.1.3)). Les IgG ne comportant pas de O-glycosylations, ce type de glycosylation ne sera pas abordé plus en détail.

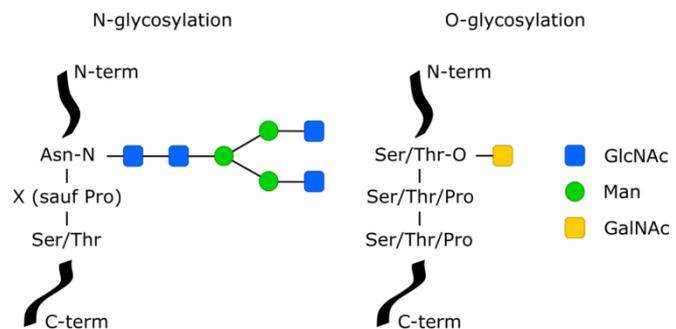


Figure 1.1.5 : Séquences spécifiques et structures principales des N-glycosylations et O-glycosylations.

### 1.2.2. Les différentes N-glycosylations des IgG

Les IgG possèdent un site de N-glycosylation hautement conservé sur leurs Asp 297 [6]. Les profils des glycanes de ces sites peuvent être très hétérogènes. A la séquence de glycanes de base décrite précédemment pour les N-glycosylations, différents autres glycanes peuvent s'ajouter dans des configurations variées pour former des structures plus complexes dont la taille peut varier en fonction du nombre et du type de glycanes (figure 1.1.6). En effet, parmi les différentes associations possibles, un fucose peut s'ajouter au premier GlcNAc, des galactoses peuvent s'ajouter sur les deux GlcNAc du côté terminal de la structure principale, suivis par de l'acide sialique (figure 1.1.6).

Outre les sites de N-glycosylation du fragment Fc, il a été démontré que 15 à 20% des Fab d'IgG étaient glycosylés dans le corps humain [7]. Ces glycosylations se situent au niveau des CDR et peuvent avoir divers effets. Cela peut augmenter ou diminuer leur spécificité aux antigènes, entraver des interactions protéines-protéines, étendre la durée de demi-vie de l'anticorps, ou encore provoquer son aggrégation ou la formation de complexes [6].

L'anticorps anti-Lamp1 utilisé lors de ce projet (cf. partie 1.2.5) ne comporte pas de sites de glycosylation sur son fragment Fab.

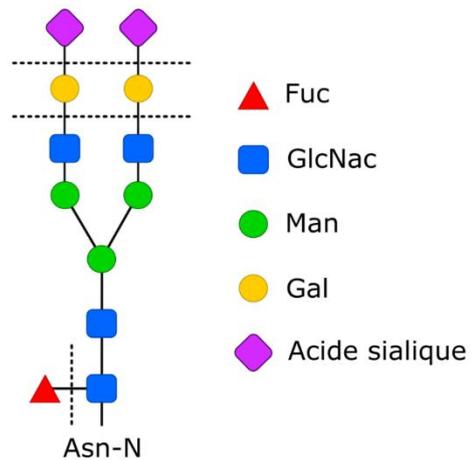


Figure 1.1.6 : Représentation schématique des différentes configurations possibles de glycosylation des asparagines 297 des IgG. Les glycanes pouvant être ajoutés à la structure principale des N-glycosylations sont délimités par des traits en pointillés.

### 1.3. Structure des cinq classes principales d'immunoglobulines

Il existe plusieurs classes d'anticorps : les isotypes IgG, IgE, IgA, IgD et IgM. Leur classification est réalisée selon leurs caractéristiques structurales, telles que le nombre et la position des ponts disulfures, la longueur des chaînes lourdes, le nombre de monomères d'immunoglobulines et le nombre et le type de glycosylation.

#### 1.3.1. Les IgG

Les IgG sont la classe prédominante des immunoglobulines dans l'organisme d'une personne saine (environ 75% des anticorps présents dans le plasma sanguin)[3]. La structure de base d'une IgG1 a été détaillée précédemment dans la partie 1.1.1. Il existe trois autres sous-classes : les IgG2, IgG3 et IgG4. Leur numéro correspond à leur abondance relative dans le plasma sanguin, l'IgG1 étant la plus abondante [8]. Les IgG2 peuvent également être divisées en trois catégories selon la conformation de leurs ponts disulfures : IgG2-A, IgG2 A/B et IgG2-B (figure 1.1.7).

Ces quatre sous-classes d'IgG ont une structure primaire identique à environ 90%, cependant, des différences importantes subsistent, notamment au niveau de la région charnière et des ponts disulfures reliant les deux chaînes lourdes, dont la position et le nombre varient considérablement. L'IgG1 et l'IgG4 en comptent deux, l'IgG2 quatre dont l'arrangement peut varier selon la sous-classe, et l'IgG3 en comporte onze. Cette différence impacte la flexibilité des IgG et donc leur activité.

L'IgG4 naturelle présente quant à elle une particularité supplémentaire par rapport aux autres sous-classes, puisqu'elle est bivalente et correspond à la combinaison de deux parties d'anticorps IgG4 monovalents. Cette sous-classe a la capacité d'échanger des fragments Fab, produisant ainsi des IgG présentant une double spécificité. Sur la figure 1.1.7, les couleurs distinctes indiquent une origine issue d'IgG4 différentes. Le fragment Fc reste inchangé car celui-ci est constant parmi les immunoglobulines, mais il présente deux fragments Fab dont les spécificités sont différentes.

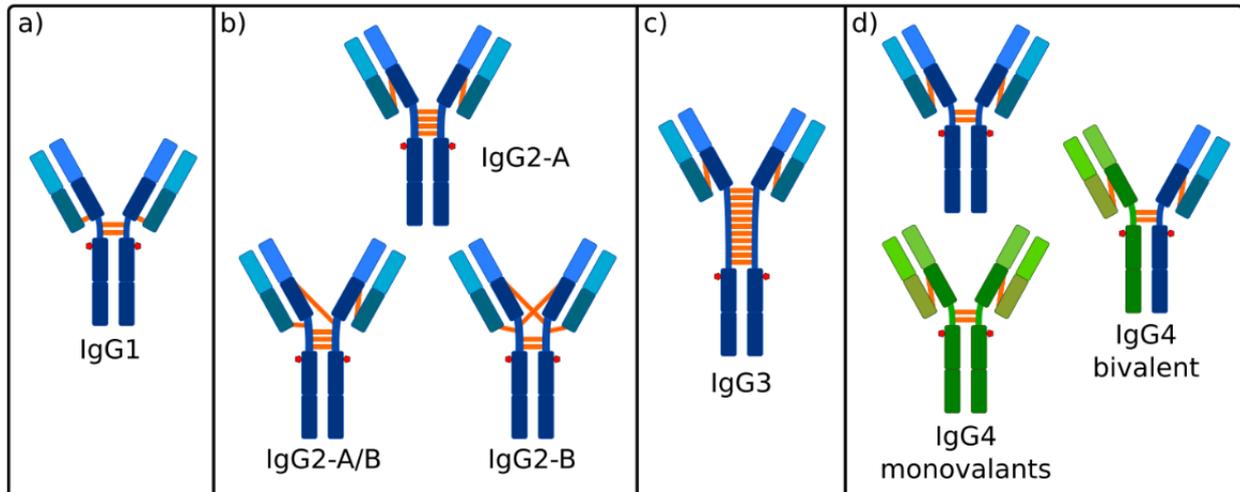


Figure 1.1.7 : Structure des IgG1, IgG2, IgG3, et IgG4 avec représentation des ponts disulfures (orange) et des N-glycosylations (rouge).

### 1.3.2. Les IgD

Les IgD sont très similaires aux IgG. Leurs domaines ainsi que la composition de leurs fragments Fab et Fc est identique. Les différences se situent au niveau des ponts disulfures inter-chaînes. Un seul pont disulfure, situé au niveau de la partie N-terminale des domaines CH2, est présent pour relier les deux chaînes lourdes entre elles. Au niveau des fragments Fab, les ponts disulfures relient la partie N-terminale du domaine CL des chaînes légères avec le C-terminal du domaine CH1 des chaînes lourdes (figure 1.1.8). La présence d'un seul pont disulfure entre les chaînes lourdes rend cette classe d'anticorps plus sensible aux dégradations et impacte sa durée de demi-vie en comparaison des autres classes d'anticorps [9].

Les sites de glycosylation de cette classe sont plus nombreux que chez les IgG. Il y a trois N-glycosylations symétriques : deux sur les domaines CH3 et une sur les domaines CH2. Les IgD constituent l'une des deux classes présentant des sites de O-glycosylation. Ces sites se situent sur la région charnière.

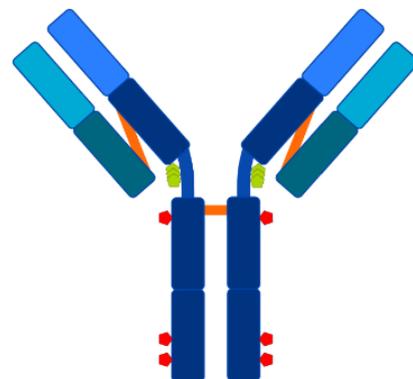


Figure 1.1.8 : Structure des IgD avec représentation des ponts disulfures (orange), des N-glycosylations (rouge) et des O-glycosylations (vert).

### 1.3.3. Les IgE

Les IgE font environ 190 kDa. Leurs chaînes lourdes sont plus longues que celles des IgG et IgD. Elles sont composées d'un 5<sup>ème</sup> domaine en C-terminal : le CH4. Les IgE sont plus rigides car elles ne présentent pas de région charnière, qui est remplacée par un domaine composé de feuillets  $\beta$ . Les ponts disulfures varient. En effet, leur position est identique à celle des IgD dans les fragments Fab, mais ceux reliant les chaînes lourdes, au nombre de deux, se situent en N-terminal et en C-terminal des domaines CH2 (figure 1.1.9).

Les sites de glycosylation sont nombreux et peuvent représenter 12 à 14% de la masse totale de l'anticorps, ce qui en fait la classe d'anticorps la plus glycosylée[10]. Une chaîne lourde d'IgE compte sept sites de N-glycosylation. Trois sont situés sur le domaine CH3, un sur le domaine CH2, et trois très proches les uns des autres sur le domaine CH1. La forte densité de glycanes sur les fragments Fab en plus de l'absence de région charnière des IgE augmentent leur rigidité.

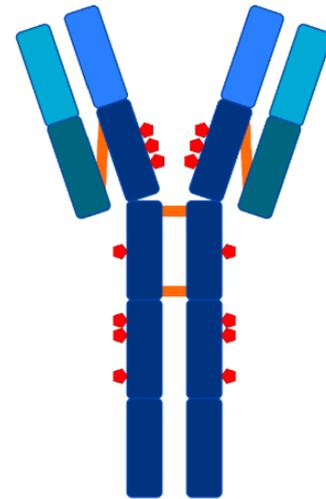


Figure 1.1.9 : Structure des IgE avec représentation des ponts disulfures (orange) et des N-glycosylations (rouge).

### 1.3.4. Les IgA

Les IgA sont composées de deux sous-classes : les IgA1 et IgA2 (figure 1.1.10). La forme monomérique peut servir pour détailler la structure des deux sous-classes. Les chaînes lourdes sont composées du même nombre de domaines que pour les IgG et IgD, et les fragments Fab sont reliés au fragment Fc par une région charnière. Cependant, la région charnière ne comporte pas de pont disulfure, ce qui rend les fragments Fab très mobiles. Les chaînes lourdes sont reliées par deux ponts disulfures au niveau de la partie C-terminale des domaines CH2. Les chaînes légères sont reliées aux chaînes lourdes de la même façon que pour les IgD et IgE.

Les deux sous-classes d'IgA diffèrent au niveau de leurs sites de glycosylation. Les IgA1 ne comportent que deux sites symétriques de N-glycosylation en C-terminal des domaines CH3 et en N-terminal des domaines CH2. Les IgA1 sont la deuxième classe d'anticorps (avec les IgD) à comporter des sites de O-glycosylation, également situés sur leur région charnière. Les IgA2 ne comportent pas de O-glycosylation, mais présentent plus de sites de N-glycosylation. Deux sites existent sur les domaines CH1, deux sur les CH2 et un dernier sur les CH3.

Les IgA existent principalement sous forme de dimères [11]. Dans cette conformation, deux monomères sont reliés par une chaîne polypeptidique appelée J-Chain (*Jonction Chain*) à l'aide de deux ponts disulfures situés en C-terminal des domaines CH3. La J-Chain est composée de 137 acides aminés et est riche en cystéines. Les deux domaines CH3 non reliés à la J-Chain sont reliés entre eux par un disulfure supplémentaire au niveau de leur partie C-terminale.

Les IgA dimériques comportent également un élément supplémentaire, le composant sécrétoire transmembranaire (SC). Ce polypeptide de 75-80 kDa est enroulé autour des dimères

d'IgA et y est attaché par l'intermédiaire de deux ponts disulfures au niveau d'un domaine CH2 des deux monomères.

Avec l'ajout de la J-Chain et du SC, la masse moléculaire des dimères d'IgA est d'environ 390 kDa sans compter les glycosylations.

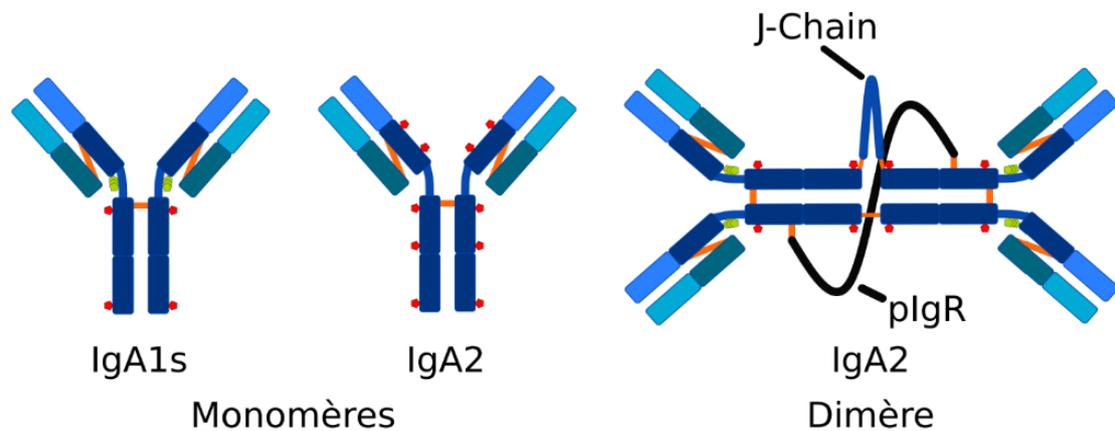


Figure 1.1.10 : Structure des IgA avec représentation des ponts disulfures (orange), des N-glycosylations (rouge) et des O-glycosylations (vert).

### 1.3.5. Les IgM

Les IgM existent sous forme monomérique à la surface des lymphocytes B, mais sont excrétées sous forme multimérique, ce qui leur confère une forme caractéristique de flocon de neige (figure 1.1.11). Cette forme pentamérique en fait l'anticorps le plus imposant avec une masse d'environ 950 kDa, sans compter les nombreuses glycosylations.

Les monomères d'IgM ont une structure très similaire aux IgE et sont très rigides. Ils ont un nombre de domaines identiques, une absence de région charnière, et le même nombre de ponts disulfures aux mêmes positions. En revanche, leurs chaînes lourdes ont cinq sites de glycosylation dont un se situe sur les domaines CH1, et les quatre autres se situent au niveau des domaines CH2, CH3 et CH4 (le domaine CH3 en comporte deux).

La forme pentamérique des IgM est garantie par la présence de nombreux ponts disulfures et la présence de la J-Chain reliée à deux domaines CH4 de deux monomères par des ponts disulfures. Les domaines CH3 de chacun des différents monomères d'IgM sont reliés entre eux par des ponts disulfures inter-immunoglobulines situés au niveau de l'extrémité C-terminale. Les monomères d'IgM comportent également des ponts disulfures internes au niveau de leurs domaines CH3. De même que pour les IgA, la forme multimérique augmente la stabilité de l'anticorps face à un environnement agressif, et augmente ses chances de reconnaissance dans la lutte contre les pathogènes[12].

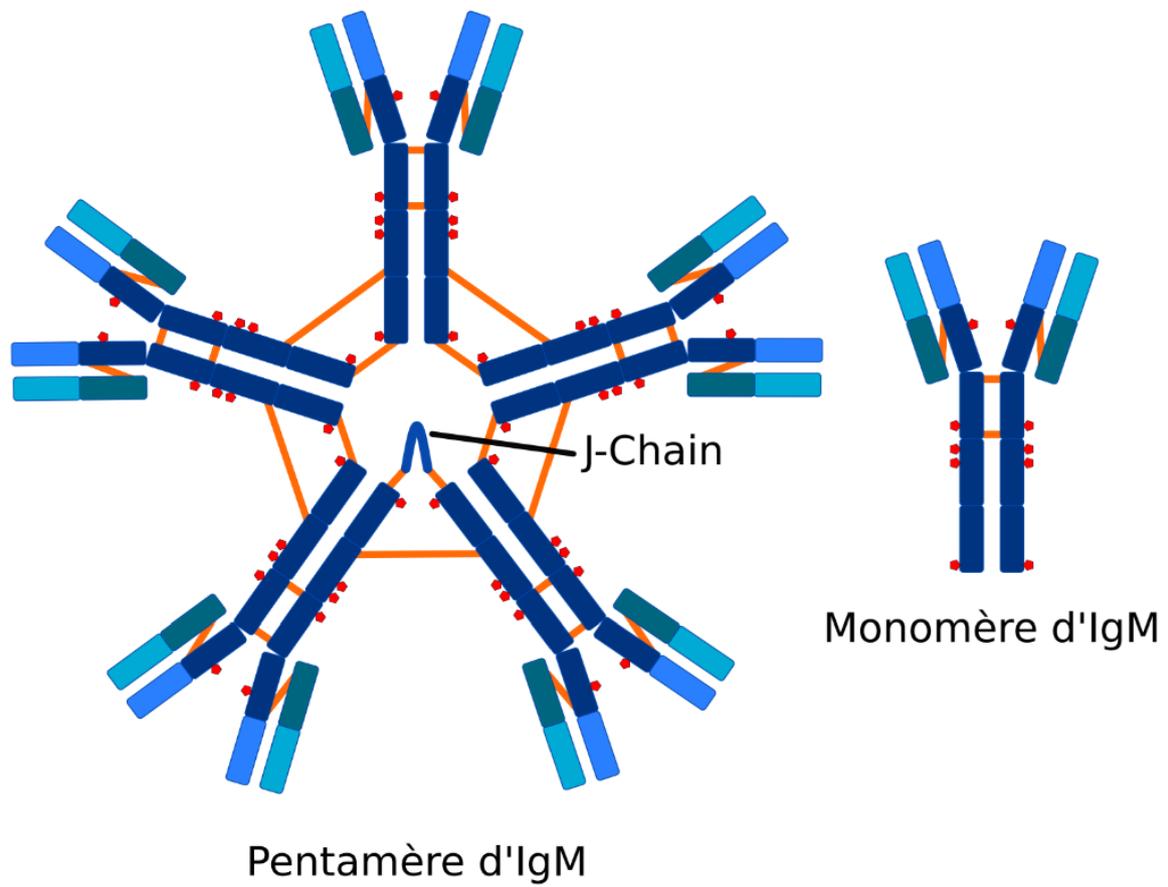


Figure 1.1.11 : Structure des IgM avec représentation des ponts disulfures (orange) et des N-glycosylations (rouge).

## 1.4. Les anticorps dans le système immunitaire

### 1.4.1. Rappels sur le système immunitaire

La réponse immunitaire est la capacité d'un organisme à utiliser des mécanismes de défenses biologiques pour la reconnaissance et l'élimination des agents infectieux extérieurs. Ce système est composé de différents acteurs essentiels à son fonctionnement : les anticorps, les antigènes, les lymphocytes, le système du complément, les cytokines, les cellules dendritiques, les éosinophiles et les macrophages. Ces acteurs réagissent de manière coordonnée, ce qui donne la réponse immunitaire.

Deux types de réponses immunitaires existent : la réponse innée (naturelle), et la réponse acquise (adaptative et spécifique).

La réponse innée est la capacité de tout organisme à répondre face à un nouvel agent infectieux. Les acteurs de cette réponse sont les cellules phagocytaires, les cellules NK (*natural killer*) ou encore des anticorps préformés dit naturels et exprimés par les lymphocytes B immatures. Ces acteurs ne sont pas spécifiques à un antigène mais contiennent des récepteurs spécialisés capables de reconnaître différents composés antigéniques.

La réponse adaptative est beaucoup plus efficace et survient lorsque l'organisme a déjà rencontré et lutté contre un agent infectieux. Les cellules dendritiques (cellules présentatrices des antigènes et initiant cette réponse immunitaire adaptative), les cytokines, les lymphocytes et le système du complément sont des acteurs de cette réponse adaptative. Premièrement, les lymphocytes B et T spécialisés, dans la reconnaissance spécifique d'antigènes, sont activés par la reconnaissance de l'antigène par les anticorps de surface présents sur ces cellules immunitaires. On distingue alors deux mécanismes différents lors de la réponse adaptative : l'immunité cellulaire et l'immunité humorale (figure 1.1.12). La réponse cellulaire est assurée principalement par les lymphocytes T et peut être active même en l'absence d'anticorps. La réponse humorale est assurée par la production d'anticorps en grande quantité par les lymphocytes B. La détection d'un pathogène leur permet d'évoluer en plasmocytes producteurs d'isotypes d'anticorps différents ou en lymphocytes B mémoire permettant une réponse plus rapide lors d'une infection future du même pathogène. C'est lors de la réponse humorale de l'immunité adaptative que les anticorps remplissent leurs fonctions effectrices.

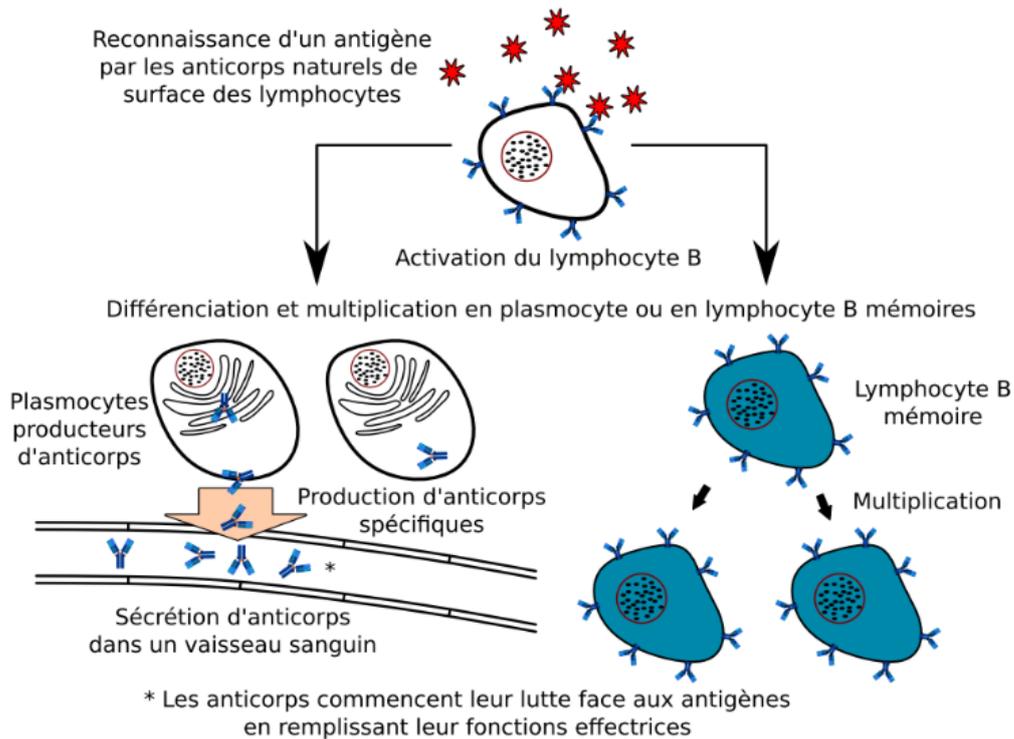


Figure 1.1.12 : Mécanisme simplifié de la réponse humorale du système immunitaire

#### 1.4.2. Les sites actifs des anticorps

Lors de la réponse humorale du système immunitaire, les anticorps remplissent leur rôle à travers des fonctions effectrices rendues possibles grâce à différents sites actifs sur leur structure [13].

Tous les isotypes comportent des sites actifs, mais l'IgG est la classe en possédant le plus. Ces sites peuvent être positionnés en prenant la structure d'une IgG comme référence (figure 1.1.13).

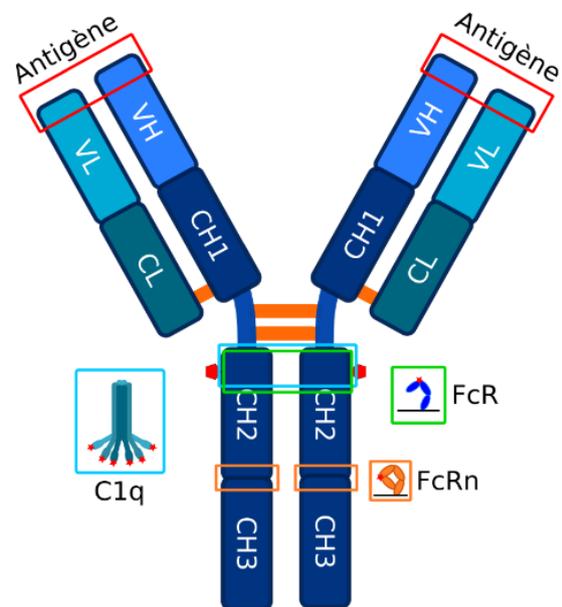


Figure 1.1.13 : Position des sites actifs d'une IgG sur sa structure avec les différents acteurs et leur site de liaison.

### **a Sites actifs sur le fragment Fab**

Les fragments Fab des anticorps portent les sites de reconnaissance des antigènes sur les CDR. Cette spécificité est responsable de plusieurs fonctions effectrices : la neutralisation, l'agglomération ou encore la précipitation de virus, bactéries ou toxines. La reconnaissance d'antigènes permet d'activer les fonctions effectrices dépendantes du fragment Fc des anticorps comme l'opsonisation qui active les phagocytes.

### **b Sites actifs sur le fragment Fc**

Les fragments Fc ont trois sites actifs principaux et distincts remplissant plusieurs fonctions effectrices (figure 1.1.13) : un site de reconnaissance des récepteurs à Fc (FcR), un pour la reconnaissance de l'effecteur C1q du complément, et un autre pour la reconnaissance des récepteurs néonataux de Fc (FcRn).

Le site de reconnaissance au C1q permet d'activer la réaction en cascade des protéines du complément. La zone de reconnaissance est partiellement superposée à celle du site de reconnaissance des FcR. La structure du complexe formé par le fragment Fc et le C1q n'a, à l'heure actuelle, pas été déterminé, mais l'article de S. Schneider *et al.* présente un modèle, qui ne sera pas détaillé ici [14].

Les FcR se situent à la surface de multiples cellules du système immunitaire comme les cellules NK, les phagocytes ou les lymphocytes T. La reconnaissance par l'intermédiaire de ce site actif par un anticorps couplé à un antigène provoque l'activation de différentes voies biologiques. La reconnaissance par une cellule NK active la cytotoxicité à médiation cellulaire dépendante des anticorps (ADCC pour *Antibody dependent cells cytotoxicity*), ou la phagocytose dépendante des anticorps (ADCP pour *antibody-dependent cellular phagocytosis*) lorsque ce site est reconnu par un phagocyte. Il existe plusieurs classes de FcR (FcRI, FcRII, FcRII), dont chacune est spécifique à certains isotypes et/ou sous-classes d'immunoglobulines. Les différentes catégories de FcR [13] ne seront pas développées dans ce manuscrit, mais leur nom varie selon l'isotype reconnu : Fc $\gamma$ R ( IgG), Fc $\mu$ R (IgM), Fc $\epsilon$ R (IgE), et Fc $\alpha$ R (IgA). Ce site actif est localisé en dessous de la région charnière et à proximité des sites de glycosylation du fragment Fc. La reconnaissance des FcR par ce site est médiée par l'interaction avec les glycanes, qui sera impactée selon la nature de ceux-ci. La figure 1.1.14 montre l'interaction du fragment Fc d'une IgG1 à un Fc $\gamma$ R et les acides aminés impliqués dans l'interaction ainsi que la participation des glycanes [15].

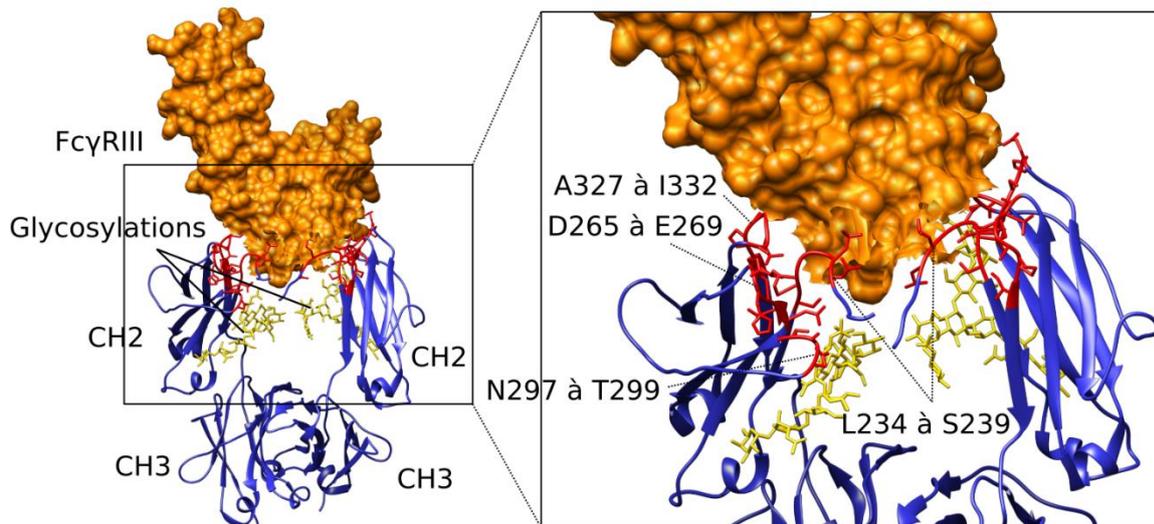


Figure 1.1.14 : Structure cristallographique du fragment Fc (en bleu) d'une IgG1 humaine avec un FcγRIII (en orange). Les acides aminés du fragment Fc impliqués dans l'interaction sont représentés en rouge. Les glycanes sont en beige. PDB : 1E4K DOI: <https://doi.org/10.1038/35018508>

Le reconnaissance des fragments Fc par les récepteurs néonataux (les FcRn) se fait au niveau des régions charnières entre les domaines CH2 et CH3 du fragment Fc (figure 1.1.15). Cette reconnaissance implique plus particulièrement l'isoleucine 253 de la chaîne lourde des anticorps. Des ponts salins avec les résidus His310, Gln311 et His435 renforcent l'interaction avec les FcRn. Ce site actif de reconnaissance ne permet pas de remplir une fonction effectrice mais y participe indirectement car il est impliqué dans les fonctions de bio-distribution et de recyclage des anticorps. Les FcRn des endosomes permettent aux IgG d'échapper au catabolisme endothélial, ce qui augmente leur durée de vie plasmatique et permet aussi leur transcytose à travers différentes membranes. L'étude de ce site est très intéressante lors du développement d'un anticorps thérapeutique car il est associé à la pharmacocinétique de l'anticorps[16], [17].

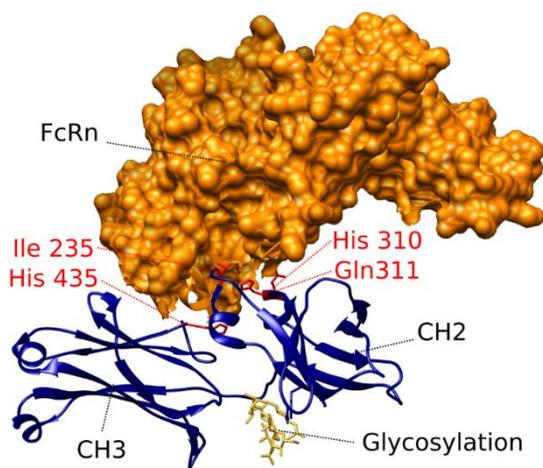


Figure 1.1.15 : Structure cristallographique de la chaîne lourde du fragment Fc d'une IgG1 de rat (bleu) avec un FcRn (en orange). Les glycanes (beige) et les acides aminés du fragment Fc (rouge) impliqués dans l'interaction sont représentés. PDB : 1FRT.

### 1.4.3. Fonctions effectrices des anticorps

#### a Neutralisation, agglomération, et précipitation

Lorsque les anticorps reconnaissent un antigène comme une bactérie ou un virus, ils sont capables d'exercer différentes fonctions effectrices, telles que la neutralisation, l'agglomération ou la précipitation (figure 1.1.16). Ces fonctions sont assurées par la spécificité des fragments Fab des anticorps. Ainsi, les anticorps sont capables de neutraliser les toxines bactériennes avant qu'elles n'interagissent avec d'autres cellules. Ils peuvent également inhiber l'adhérence de bactéries aux surfaces cellulaires, lorsqu'ils présentent une spécificité pour les adhésines de surface, ou encore neutraliser les virus en surchargeant les récepteurs par l'intermédiaire desquels les virus se fixent aux cellules. L'agglomération est le fait de réunir plusieurs antigènes ensemble grâce à la présence des deux fragments Fab ou des multiples isomères d'immunoglobuline. Enfin, la précipitation a lieu lorsque des anticorps forment une chaîne en intercalant des antigènes entre leur fragment Fab, ce qui induit une diminution de leur solubilité et la précipitation du complexe. Ces trois fonctions effectrices sont la base de l'activité des anticorps.

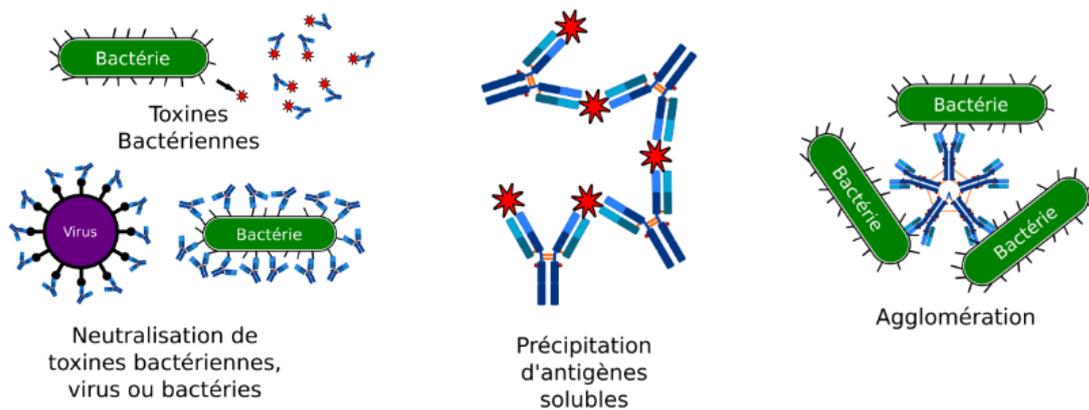


Figure 1.1.16 : Neutralisation, précipitation et agglomération des bactéries, virus et toxines par les anticorps.

#### b Opsonisation et phagocytose

L'opsonisation est le recouvrement de la membrane d'une bactérie par des anticorps pour favoriser sa phagocytose (figure 1.1.17). Les anticorps ainsi adhérents sont appelés des opsonines. En les reconnaissant grâce à leur FcR, les cellules phagocytaires s'activent et le complexe anticorps-antigène permet aux phagocytes de l'internaliser puis de le détruire par phagocytose. Il s'agit de l'activation de la voie l'ADCP.

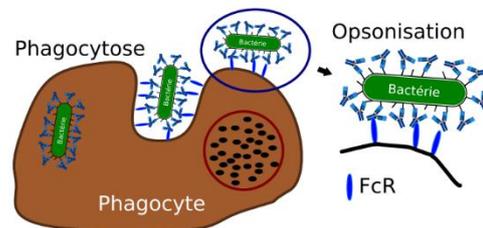


Figure 1.1.17 : Opsonisation et phagocytose d'un corps pathogène

### c La cytotoxicité dépendante du complément : CDC

Le fragment Fc contient un segment spécifique aux effecteurs C1q du complément. L'interaction avec cette molécule active une réaction en cascade avec plusieurs protéines du plasma pour former un complexe d'attaque membranaire (CAM) qui se forme à la surface des pathogènes (figure 1.1.18). Ce complexe a pour effet de former des tunnels à travers les membranes, provoquant ainsi l'éclatement de la cellule cible par sa perforation.

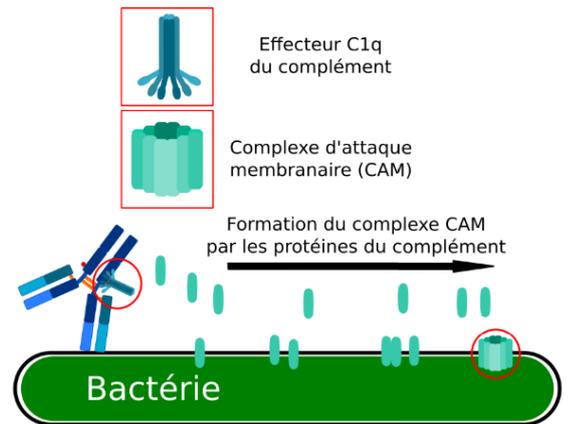


Figure 1.1.18 : Activation de la cascade du complément par la reconnaissance du site C1q d'un anticorps

### d Activation des cellules cytotoxiques dépendantes des anticorps : ADCC

L'ADCC a lieu *via* la reconnaissance des FcR des anticorps couplés aux antigènes par les cellules effectrices comme les cellules NK (*Natural Killer cells* en anglais). Cette reconnaissance permet l'activation des cellules effectrices qui libèrent alors des granules cytotoxiques qui viennent ensuite attaquer la cellule cible (figure 1.1.19).

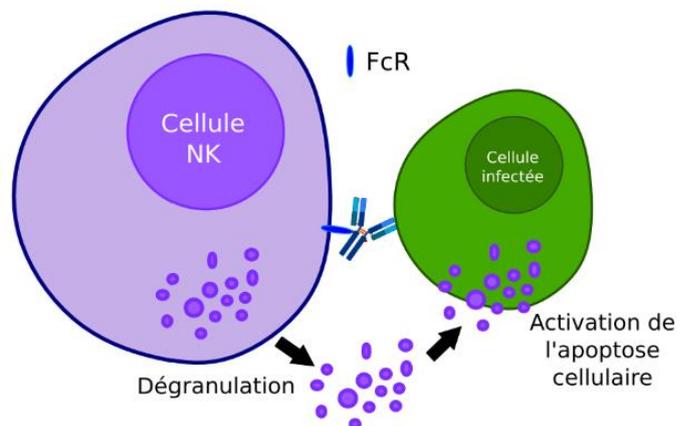


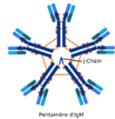
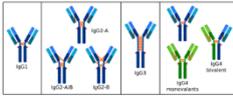
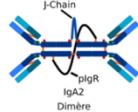
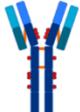
Figure 1.1.19 : Activation d'un ADCC. Le mécanisme d'inflammation est aussi provoqué par la dégranulation

### e Inflammation

Les mastocytes, les éosinophiles ou les basophiles sont activés lors de la reconnaissance d'un FcR d'anticorps couplé. Cette activation provoque leur dégranulation et libère de nombreux médiateurs participant à l'inflammation. Ce mécanisme est notamment présent dans les réactions allergiques.

Le tableau 1.1.1 résume les cinq isotypes d'anticorps physiologiques avec les caractéristiques qui ont été présentées dans ce chapitre.

Tableau 1.1.1 : Tableau récapitulatif des différents isotypes d'anticorps et leurs fonctions

Isotypes	IgM	IgD	IgG	IgA	IgE
Sous-classes	Monomériques Pentamériques	/	IgG1, IgG2a, IgG2b, IgG2a/b, IgG3, IgG4	Monomériques Dimériques IgA1 et IgA2	/
Localisation	Surface des lymphocytes B et muqueuses	Surface des lymphocytes B	Plasma sanguin	Muqueuses	Plasma sanguin
Fonctions effectrices	Activation du complément	/	Neutralisation, opsonisation, activation du complément, ADCC et AIDN	Activation des ADCC et neutralisation	Activation des mastocytes, basophiles et éosinophiles
Rôle dans le système immunitaire	Monomères : Activation des lymphocytes Pentamères : anti-toxines bactériennes des muqueuses	Activation des lymphocytes B	IgG1 : Contre les protéines antigéniques solubles IgG2 : Contre les gènes polysaccharidiques bactériens IgG3 : réponse inflammatoire IgG4 : réponse allergique	Lutte contre les toxines et les bactéries des muqueuses	Réponse allergique et lutte contre les parasites
Schéma structural					

## 2. Les anticorps thérapeutiques

L'utilisation des anticorps en tant que médicament a été rendue possible grâce aux nombreuses avancées technologiques qui ont commencé avec le développement d'anticorps thérapeutiques monoclonaux (mAb) et la découverte de l'immunothérapie par James P. Allison et Tasuku Honjo en 1996 (prix Nobel de physiologie ou de médecine 2018) [18]. Le développement de nouvelles structures aux mécanismes d'action variés a permis de lutter contre des pathologies jusqu'alors incurables. Ainsi, les mAbs thérapeutiques sont devenus la classe de nouveaux médicaments progressant le plus rapidement sur le marché. A eux seuls, ils représentaient près de 50% des biothérapeutiques en cours d'autorisation en 2022, et ce pourcentage augmente tous les ans. Ils représentaient 188 milliards d'Euros en 2022 sur le marché mondial des médicaments et les économistes prévoient 330 milliards pour 2027 [19]. Depuis leur découverte et l'autorisation du premier mAb thérapeutique sur le marché par les agences de sécurité du médicament, 119 ont été acceptés par la FDA, et 113 par l'European Medicines Agency (EMA) en 2022. (chiffres de l'*antibody society* en Mars 2023) <https://www.antibodysociety.org/> (figure 1.2.1).

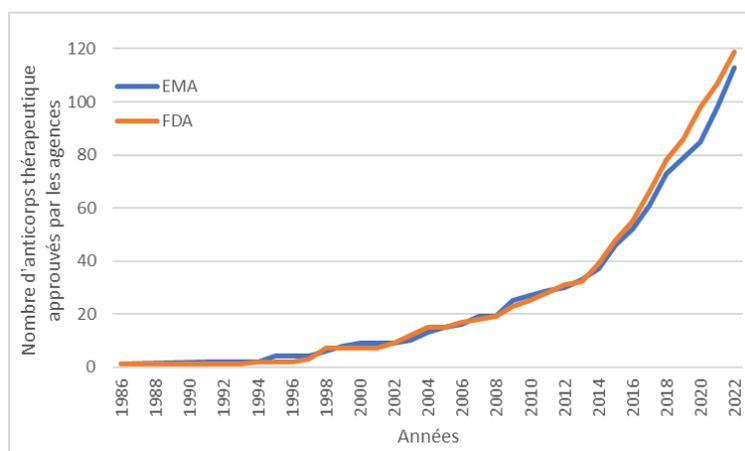


Figure 1.2.1 : Evolution du nombre d'anticorps monoclonaux et de fragments thérapeutiques acceptés par les agences EMA et FDA de 1986 à 2022.

Cette partie présente les avancées majeures de la recherche qui ont mené au développement des anticorps thérapeutiques monoclonaux, ainsi que leurs principaux mécanismes d'action, et les différentes classes des mAbs thérapeutiques actuellement sur le marché.

L'appellation des anticorps thérapeutiques est soumise à une dénomination commune internationale [20]. Cette nomenclature est utilisée dans les paragraphes suivants et se trouve synthétisée dans le tableau 1.2.1. Il faut noter qu'est appelé mAb thérapeutique toute protéine thérapeutique ayant une fonction effectrice de reconnaissance d'antigène spécifique provenant d'un mAb (et plus particulièrement de ses CDR). Ainsi, les différents fragments thérapeutiques qui seront détaillés ultérieurement dans cette partie, les Fab, scFv, F(ab')<sub>2</sub> portent aussi le nom de mAbs thérapeutiques. Les protéines thérapeutiques ne comportant que le fragment Fc d'un anticorps ne sont pas nommées mAbs thérapeutiques, mais protéines de fusion au Fc, car elles sont dépourvues de domaine possédant la partie effectrice des anticorps responsable de la reconnaissance antigénique.

Tableau 1.2.1 : Dénomination commune internationale des anticorps monoclonaux thérapeutiques. Un premier préfixe au choix du producteur est ajouté au début.

Information sur la cible		Information sur la source		Préfixe d'anticorps monoclonal	
Maladies	Préfixe	Source	Préfixe		
Bactériennes	-ba(c)-	Souris	-o-	-mab	
Cardio-vasculaires	-ci(r)-				
Immunologiques	-li(m)-				
Virales	-vi(r)-	Chimère	-xi-		
Tumeurs					
du côlon	-co(l)-	Humanisé	-zu-		
du testicule	-go(t)-				
de l'ovaire	-go(v)-				
mammaires	-ma(r)-				
mélanomes	-me(l)-				
de la prostate	-pr(o)-	Homme	-u-		
diverses	-tu(m)-				
Exemple : Emicizumab		Emi-	-ci-	-zu-	-mab
		Préfixe choisi par le producteur		Anticorps monoclonal humanisé contre une maladie cardiovasculaire	

## 2.1. Les premiers anticorps thérapeutiques

### 2.1.1. Les anticorps polyclonaux

En 1894, après la découverte des anticorps ‘antitoxine’ par le Dr Von Behring, celui-ci invente la sérothérapie et collabore avec le professeur Paul Ehrlich (prix Nobel de physiologie et de médecine 1908) pour la production d’un sérum anti-diphtérie. Face à la pandémie de diphtérie en Allemagne, leurs travaux ont permis de faire baisser la mortalité infantile de 52% à 13% [21]. Le sérum produit à partir du sang de chevaux immunisés contre la maladie est le premier biothérapeutique, contenant des anticorps, administré à des patients.

Ces sérums contiennent un mélange de nombreux anticorps aux spécificités différentes. Ces anticorps sont dits polyclonaux car ils sont produits par différents lymphocytes au sein de l’organisme. L’inconvénient majeur de ce type de biothérapeutique est la présence de protéines animales immunogènes pour l’homme dans le sérum. Les receveurs présentent donc un risque de développer des réactions immunitaires contre ces protéines animales et sont donc susceptibles de s’immuniser contre le traitement en cas d’injections répétées. Le deuxième inconvénient est la multi-spécificité des anticorps produits. Ces biothérapeutiques ne permettent pas la reconnaissance spécifique d’une seule et unique structure antigénique.

Les anticorps polyclonaux d’origine animale sont toutefois encore utilisés dans certaines situations particulières, notamment en l’absence d’alternatives satisfaisantes. Par exemple, il existe des sérums anti-venins ou anti-toxines bactériennes comportant des anticorps polyclonaux produits chez le cheval. Ces médicaments sont utilisés sous ATU (autorisation temporaire d’utilisation). D’autres sont utilisés comme thérapeutiques anti-rejet de greffe comme l’Atgam® ou le Grafalon® (Immunoglobulines polyclonales anti-lymphocyte T humain respectivement d’origine équine ou de lapin).

### 2.1.2. Développement des premiers anticorps monoclonaux

La production d'anticorps monoclonaux a été rendue possible en 1975 par Georges Jean F. Köhler et César Milstein [22] qui ont reçu le prix Nobel de physiologie et de médecine en 1984, avec Niels K. Jerne, pour leurs travaux sur les théories du développement et du contrôle du système immunitaire. Leurs recherches (non brevetées) ont permis de développer la méthode dite des cellules hybridomes à partir de souris immunisées contre les globules rouges de moutons. Une cellule hybridome est alors capable de se multiplier à l'infini (clone) et de produire un anticorps anti-mouton. Les anticorps issus d'une souche de cellules monoclonales sont dits monoclonaux. La production d'anticorps par hybridome est détaillée dans une partie ultérieure traitant des systèmes d'expressions des anticorps (cf. partie 1.3.2.1).

Cette découverte a conduit à l'autorisation et à la mise sur le marché du premier anticorps monoclonal thérapeutique par la FDA en 1986. Il s'agissait du Muromonab-CD3 (OKT<sup>®</sup>3) (immunoglobuline G de classe 2a d'origine murine - IgG2a), dirigé contre les récepteurs CD3 situés sur la membrane des lymphocytes T et responsables du rejet potentiel de greffe d'organe [23]. (Ce biothérapeutique est sorti avant la définition de la nomenclature internationale, expliquant que son nom déroge à la règle de nomenclature internationale établie.)

Le deuxième mAb thérapeutique a été admis par la FDA en 1994. L'autorisation d'un seul anticorps monoclonal murin par la FDA depuis la création de la méthode des hybridomes en 1975 s'explique par le fait que les mAbs ainsi produits comportent une séquence protéique d'origine murine immunogène pour l'homme. Cela provoque la production d'anticorps humains anti-souris, induisant la réponse HAMA (*human anti-murine antibody*) qui a deux conséquences majeures. Premièrement, les anticorps murins n'ont pas d'affinité pour les FcR et FcRn humains et ont une faible durée de vie dans l'organisme humain car ils sont rapidement éliminés par le système immunitaire. Deuxièmement, les patients traités finissent par développer une immunité face au traitement, ce qui rend le biothérapeutique inefficace en cas de deuxième exposition ou, plus grave, mortel lorsque la réponse immunitaire est trop importante (mêmes inconvénients que pour les anticorps polyclonaux cités précédemment). Ces effets secondaires indésirables expliquent la difficulté de développer ces nouveaux biothérapeutiques. Pour contrer la réponse HAMA, les scientifiques ont développé des méthodes permettant de remplacer les séquences immunogènes de souris par des séquences humaines, ce qui a conduit à la création d'anticorps chimériques.

### 2.1.3. Les anticorps monoclonaux chimériques

En 1984, Morisson SL et *al.* ont développé la technique d'ADN recombinant (figure 1.2.2) permettant de remplacer les domaines constants d'anticorps murins par des domaines constants d'anticorps humains [24]. Cette nouvelle méthode a permis l'obtention d'anticorps composés à 70% de séquences humaines (les domaines constants) et 30% de séquences murines spécifiques à un antigène (les domaines variables) (figure 1.2.3). Les anticorps ainsi composés sont dits chimériques. Le remplacement de la partie constante par une séquence humaine a pour conséquence d'éliminer la séquence murine à l'origine de la réponse HAMA. L'immunogénicité de ces anticorps est donc diminuée.

Cette découverte a mené, en 1994, à la seconde autorisation de mise sur le marché d'un mAb thérapeutique qui est, de fait, le premier anticorps monoclonal chimérique : l'Abciximab (ReoPro®). Ce biothérapeutique est également le premier fragment d'anticorps thérapeutique à être autorisé. Il s'agit du fragment Fab de l'anticorps 7E3 dirigé contre le récepteur des glycoprotéines IIb/IIIa des thrombocytes (plaquettes). Son utilisation permet d'éviter la formation de caillots après une opération cardiovasculaire [25]. Le premier mAb chimérique entier à avoir été autorisé (en 1997) est une IgG, il s'agit du Rituximab (Rituxan®), un anticorps anti-CD20 des lymphocytes B, qui permet au système immunitaire de cibler les cellules cancéreuses [26].

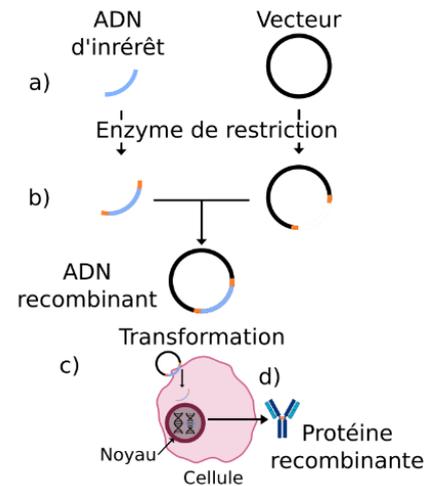


Figure 1.2.2 : Principe de l'ADN recombinant. a) Choix de l'ADN d'intérêt et du vecteur d'intérêt. b) Intégration de l'ADN dans le vecteur grâce à des enzymes de restriction. c) Transformation des cellules et intégration dans leur génome grâce à des enzymes de restriction. d) Production d'une protéine (ou anticorps) recombinante

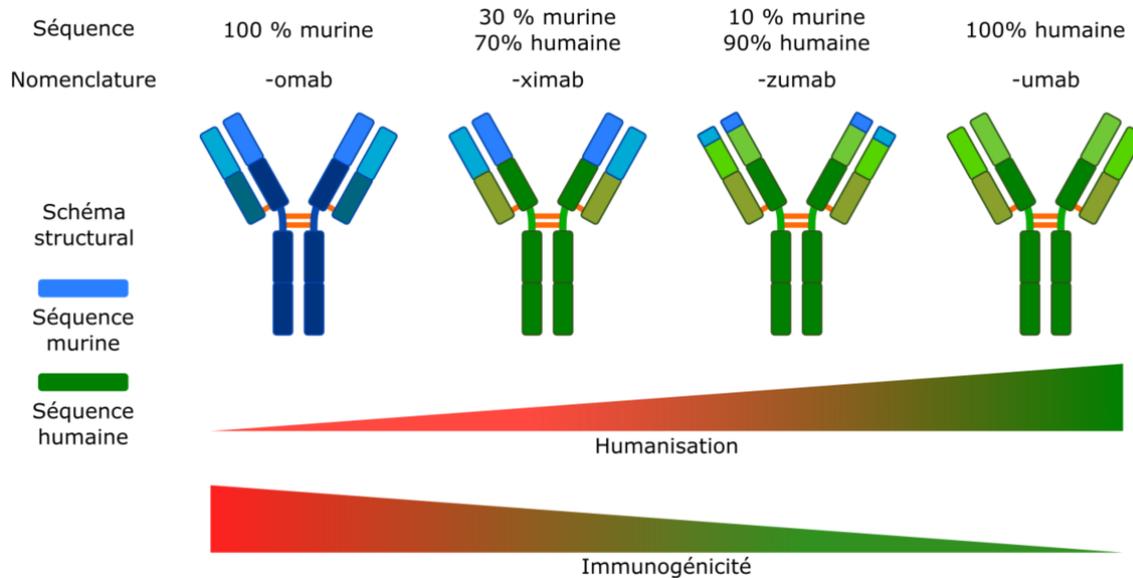


Figure 1.2.3 : Les différents types d'anticorps thérapeutiques avec la composition de leur séquence, leur nomenclature, leur représentation structurale à partir d'une IgG1 et l'évolution de leur immunogénicité en fonction de leur taux d'humanisation

Bien que la réponse HAMA ait été diminuée chez les mAbs chimériques, des réponses immunitaires anti-anticorps chimériques ont été observées chez 40% des patients [27]. Cette réponse HACA (*human anti-chimeric antibodies*) qui provient des 30% de la séquence de l'anticorps qui sont toujours d'origine murine, a conduit les chercheurs à explorer des moyens pour aboutir à une humanisation plus importante de la séquences des mAbs thérapeutiques afin de réduire au maximum leur immunogénicité.

#### 2.1.4. Les anticorps monoclonaux humanisés

Le développement d'anticorps humanisés n'a pas attendu 1994 et l'observation des réponses HACA chez les patients. En 1986, les travaux de Gregory Winter (prix Nobel de chimie 2018) ont permis de développer des anticorps humanisés grâce à l'ingénierie génétique. En fusionnant les gènes des parties constantes d'IgG humaines avec les gènes des CDR de mAbs de souris immunisées, il a obtenu des anticorps à 90% humains [28], dit humanisés (figure 1.2.3).

L'autorisation d'un anticorps humanisé a suivi de très près celle des chimériques. En 1997, la FDA approuve le Daclizumab (Zenapax®) qui est une IgG1 humaine dont on a remplacé les CDR par les CDR d'anticorps murins anti-Tac dirigés contre la chaîne  $\alpha$  des récepteurs humains CD25 des interleukines- 2[29]. Ce mAb a été largement utilisé dans la prévention des rejets de greffe rénale et constitue un exemple d'efficacité des anticorps humanisés avec sûreté.

Le Natalizumab (Tysabri®) représente au contraire un exemple qui prouve que les anticorps humanisés sont toujours susceptibles de provoquer des effets aderses indésirables et imprévisibles. Ce mAb, utilisé contre la sclérose en plaque en ciblant les chaînes des récepteurs  $\alpha 4$  des intégrines, a provoqué des réactions anaphylactiques [30] (réaction immunitaire allergique) et 11% des patients traités présentaient des anticorps humains anti-anticorps humanisés [31].

#### 2.1.5. Les anticorps monoclonaux humains

L'utilisation d'anticorps monoclonaux entièrement humains a débuté en 1990 grâce aux travaux de Grégory Winter et de son équipe [32]. L'obtention de ces anticorps a été possible grâce à l'utilisation de phages filamenteux (bactériophages) présentant des gènes de fragments variables d'anticorps humains provenant d'une grande banque de données. Cette méthode, dénommée *phage display* constitue encore la méthode la plus utilisée aujourd'hui pour la sélection de candidats thérapeutiques. Les séquences de domaines variables sélectionnées sont ensuite extraites et fusionnées par génie génétique à de l'ADN humain d'IgG, qui sera ensuite cloné dans une lignée cellulaire capable d'exprimer les anticorps en quantités industrielles (cf. partie 1.3). L'Adalimumab (Humira®) est le premier mAb thérapeutique entièrement humain et obtenu par *phage display* à avoir été autorisé par la FDA en 2002. Ce mAb se lie au facteur nécrosique tumoral et neutralise sa fonction biologique en empêchant l'interaction avec les récepteurs p55 et p75 situés à la surface cellulaire [33]. Il est indiqué dans les cas de polyarthrites rhumatoïdes, de psoriasis ou encore contre la maladie de Crohn.

Une seconde méthode, apparue en 1994, a permis l'obtention de mAbs entièrement humains à partir de souris transgéniques [34]. Ces souris modifiées génétiquement comportent les gènes humains d'immunoglobulines. L'immunisation d'une souris avec un antigène conduit ainsi celle-ci à produire des anticorps humains. Cette technique a ensuite été couplée à la technique des hybridomes pour la production de mAbs entièrement humains en quantité suffisante pour l'industrie. Le premier mAb mis sur le marché et obtenu par souris transgénique est le Panitumumab (Vectibix<sup>®</sup>) en 2006, une IgG2 ciblant le récepteur du facteur de croissance épidermique et utilisé dans les cas de cancers colorectaux métastatiques exprimant le gène KRAS[35].

## **2.2.Mécanismes d'action des anticorps thérapeutiques**

Les fonctions des mAbs thérapeutiques peuvent rejoindre celles des anticorps naturels (cf. partie 1.1.4). Ils peuvent être neutralisants par la spécificité de leur fragment variable, ou immuno-régulateurs lorsqu'ils régulent des voies de signalisation en activant ou bloquant des liaisons "récepteurs membranaires/ligands". Ils peuvent également être cytolytiques (détruisant les cellules) lorsque les sites effecteurs de leur fragment Fc font intervenir d'autres partenaires du système immunitaire afin de détruire la cellule cible. Enfin, ils présentent parfois des mécanismes d'action non naturels lorsqu'ils comportent des molécules cytolytiques.

### **2.2.1. Les mAbs thérapeutiques neutralisants**

Comme leur nom l'indique, les mAbs thérapeutiques neutralisants sont utilisés pour la grande affinité de leur partie variable à un site biologique d'antigène spécifique, neutralisant ainsi son activité. Ces biothérapeutiques utilisent directement la fonction effectrice neutralisante des fragments variables des anticorps (cf. partie 1.1.4.3). Ce mécanisme d'action est recherché pour le développement de mAbs thérapeutiques dirigés contre les toxines bactériennes et les virus ou contre les auto-antigènes solubles comme les cytokines dans le cas de certaines maladies auto-immunes. Le Palivizumab (Synagis<sup>®</sup>) est le premier mAb thérapeutique neutralisant autorisé par la FDA en 1998. Cette IgG1 humanisée est utilisée dans la prévention contre le virus respiratoire syncytial responsable de la bronchiolite chez le nourrisson [40]. Elle cible la protéine membranaire F du virus permettant la fusion de celui-ci avec la membrane des cellules cibles.

Les mAbs neutralisants ne font pas intervenir de partenaires issus du système immunitaire. Les sites actifs de leur fragment Fc doivent être inactifs. Les IgG2 et IgG4 par exemple, qui présentent peu d'affinité aux différents effecteurs, ainsi que les biothérapeutiques composés uniquement du fragment Fab ou Fv des mAbs constituent de bons candidats pour l'obtention d'une action neutralisante sans activer les acteurs du système immunitaire.

### 2.2.2. Les mAbs thérapeutiques agonistes et antagonistes

Dans certaines pathologies, les cellules immunitaires ne remplissent plus leur rôle de cellules activatrices, inhibitrices ou stimulatrices des récepteurs membranaires. Dans le cas de certaines tumeurs, par exemple, les cellules tumorales sont capables d'éviter les cellules immunitaires, car la mutation de la cellule a modifié l'activité de ses récepteurs. Les mAbs thérapeutiques agonistes ont pour rôle de venir réactiver ces récepteurs, tandis que les mAbs thérapeutiques antagonistes vont agir en bloquant ces récepteurs. Cependant, à ce jour, aucun anticorps thérapeutique agoniste n'a encore été mis sur le marché.

Le Mogamulizumab (Poteligeo<sup>®</sup>) est un exemple de mAb thérapeutique antagoniste (figure 1.2.4). En effet, il bloque les récepteurs CCR4 aux chimiokines impliqués dans le transport des lymphocytes T, empêchant ainsi la prolifération des lymphocytes T vers différents organes tels que la peau, dans le cas de *mycosis fongoïde* ou syndrome de Sézary [37] (lymphomes cutanés provoqués par des lymphocytes T).

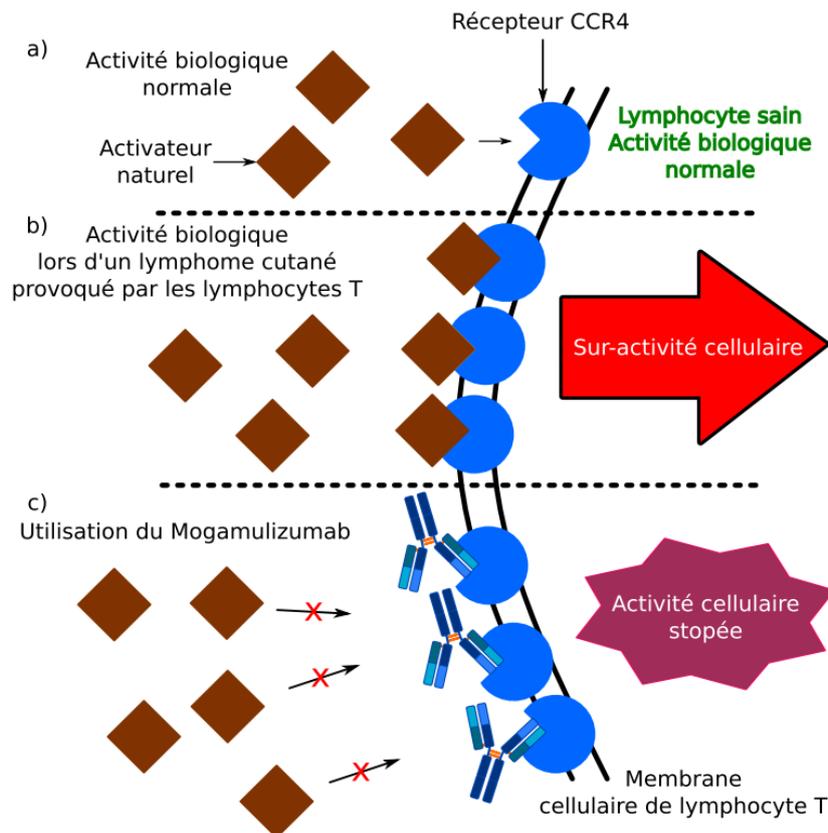


Figure 1.2.4 : Principe d'action du Mogamulizumab. a) Activité biologique chez une personne saine. b) Activité biologique chez une personne atteinte d'un syndrome de Sézary. c) Action antagoniste du Mogamulizumab sur les CCR4

### 2.2.3. Les mAbs thérapeutiques cytolytiques

Les mAbs cytolytiques induisent une lyse des cellules cibles par diverses actions. Certains sont capables de faire intervenir des protéines et cellules du système immunitaire grâce à la présence des sites effecteurs sur le fragment Fc. D'autres, comme les anticorps conjugués (cf. partie 1.2.3.3), portent une molécule pouvant activer de manière spécifique l'apoptose des cellules reconnues par l'anticorps.

L'Isatuximab (Sarclisa<sup>®</sup>) est un exemple de mAb thérapeutique cytolytique (développé par Sanofi). Il s'agit d'une IgG1 spécifique aux récepteurs CD38 des cellules et qui fait intervenir les ADCC, les CDC et les ADCP grâce à la reconnaissance des sites effecteurs du fragment Fc par les différents acteurs du système immunitaire [38]. Il a également pour effet d'immuno-réguler la réponse immunitaire. Son utilisation permet de traiter des patients atteints de myélomes multiples (cancer du sang), dans lequel de nombreux plasmocytes anormaux (comportant des CD38 de surface) s'accumulent. L'Isatuximab permet alors de les éliminer spécifiquement en faisant intervenir les fonctions effectrices du système immunitaire.

## 2.3. Les différentes structures des mAbs thérapeutiques

Les mAbs thérapeutiques sont principalement des IgG car il s'agit de la classe d'immunoglobulines la plus étudiée et donc la plus connue. L'accumulation des connaissances scientifiques et les avancées en termes d'ingénierie génétique, biologique et chimique ont permis de les utiliser de manière très efficace en tant que biothérapeutiques. Parmi l'ensemble des mAbs thérapeutiques acceptés par les agences de sécurité du médicament, environ 67% sont des IgG1 (source : Antibody society au 19/03/2023). On trouve également des IgG4, et des IgG2 et même une IgM. Les paragraphes suivants détaillent les différents types d'IgG thérapeutiques acceptés par les agences de sécurité du médicament.

### 2.3.1. IgG entières et fragments

La construction d'un anticorps thérapeutique dépend du mécanisme que l'on souhaite faire intervenir face à la pathologie considérée. Dans le cas où l'on souhaite faire intervenir des partenaires du système immunitaire, le fragment Fc de l'anticorps portant les différents sites actifs doit être présent. Si l'on souhaite neutraliser l'activité biologique de toxines bactériennes, le fragment Fab de l'anticorps peut suffire. Ainsi, les mAbs thérapeutiques peuvent autant être constitués d'IgG entières que de fragments (figure 1.2.5) (voir exemples précédents pour les IgG entières).

Lors du traitement d'une maladie, la diffusion du thérapeutique à travers les différents tissus cellulaires est un paramètre important à considérer. En effet, l'utilisation d'IgG entières a pour inconvénient leur taille, de 150 kDa, pouvant entraver leur internalisation par certaines cellules tumorales. En raison de leur taille inférieure, les fragments Fab utilisés comme mAb thérapeutique sont internalisés plus efficacement. Par exemple, le Ranibizumab (Lucentis<sup>®</sup>), correspondant au fragment Fab du Bévacizumab (Avastin<sup>®</sup>) (IgG1), s'est révélé plus concluant dans le traitement de la dégénérescence maculaire que l'anticorps entier [39].

L'absence du fragment Fc peut également être bénéfique, car l'absence du site de reconnaissance des FcR contribue à diminuer le potentiel inflammatoire du thérapeutique. En revanche, cela peut, en contrepartie, entraîner une diminution de la durée de vie du biothérapeutique car l'absence du domaine de reconnaissance des FcRn implique que le fragment ne peut pas être recyclé par l'organisme. Pour le Certolizumab pegol (Cimzia®), le problème a été résolu par la conjugaison du fragment avec une molécule de PEG (polyéthylène glycol). Cet anticorps est dirigé contre la cytokine pro-inflammatoire TNF $\alpha$  qui intervient dans les cas de polyarthrites rhumatoïdes ou maladie de Crohn [40].

Les mAbs thérapeutiques peuvent également être des scFv (figure 1.2.5) comme le Brolocizumab ou encore des nanobodies qui ne comportent qu'un seul domaine d'IgG, comme le Caplacizumab. Ces deux catégories présentent une taille encore plus réduite par rapport à celle des fragments Fab (25 kDa et 13 kDa respectivement contre 50 kDa pour les fragments Fab), ce qui améliore largement leur capacité à pénétrer les cellules pour accomplir leur activité biologique.

Beaucoup d'autres fragments d'anticorps sont utilisés en tant que mAbs thérapeutiques ou sont encore en phase d'essais cliniques, tels que les F(ab')<sub>2</sub>, les minibodies (une IgG dépourvue des domaines CH2 et dont les domaines CH3 sont directement rattachés à la région charnière), les diabodies qui correspondent à deux fragments Fabs reliés par des chaînes peptidiques synthétiques, ainsi que ses versions multiples, les triabodies et les tetrabodies (comportant respectivement trois et quatre fragments Fab reliés par des chaînes peptidiques synthétiques). La mise sur le marché de fragments de mAbs thérapeutiques a débuté avec l'Abciximab® en 1994 (cf. partie 1.2.1.3), et le potentiel thérapeutique apporté par ces nouveaux formats en fait des candidats très prometteurs pour le traitement de nombreuses pathologies dont la cible thérapeutique nécessite une pénétration à travers les membranes cellulaires.

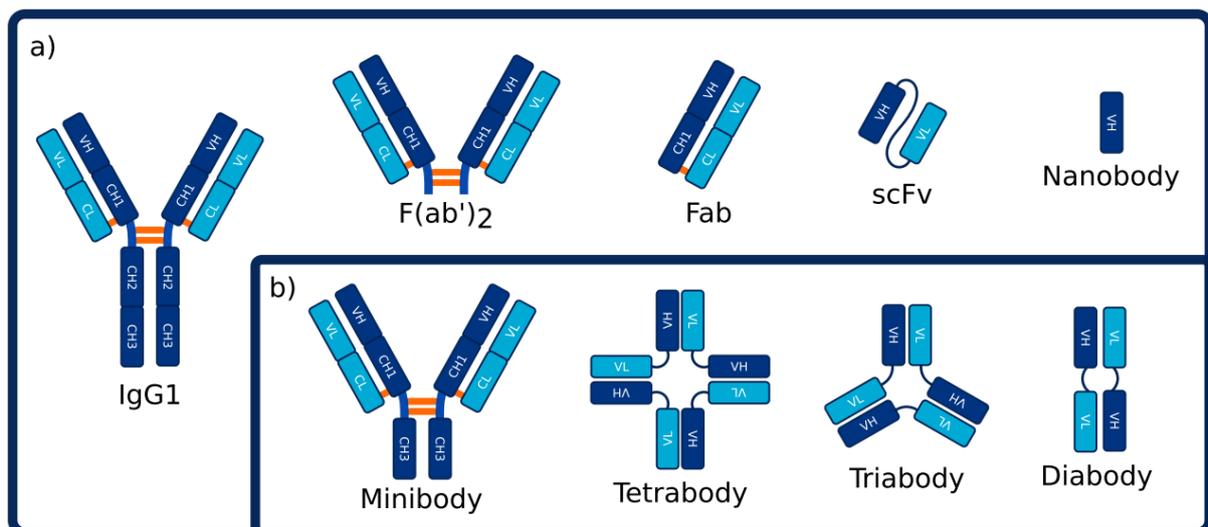


Figure 1.2.5 : Représentation schématique des différentes structures d'IgG1 et fragments de mAbs thérapeutiques. a) Structures de mAbs thérapeutiques ayant été approuvées par la FDA. b) Structures des mAbs thérapeutiques en phase de test

### 2.3.2. Anticorps monoclonaux multi spécifiques

Les mAbs multi-spécifiques sont composés de deux fragments Fab ayant des spécificités différentes, leur permettant de reconnaître deux antigènes différents ou deux épitopes distincts d'un même antigène. Il existe deux types de mAbs multi-spécifiques acceptés par la FDA et l'EMA (figure 1.2.6).

Les IgG asymétriques de type IgG4, ou duobody, sont composées de deux moitiés d'IgG (représentées de couleurs différentes sur la figure 1.2.6). L'Emicizumab (Hemlibra®) est une IgG4 asymétrique bispécifique. Il est utilisé chez les personnes atteintes d'hémophilie A dont le facteur VIII est rendu inactif par la présence d'inhibiteurs induits par lui-même, empêchant ainsi la liaison de celui-ci aux facteurs IX et X. L'absence de liaison provoque l'absence de catalyse du facteur Xa, responsable de la coagulation. L'Emicizumab est capable de se lier aux facteurs IX et X, y compris en présence des inhibiteurs induits par le facteur VIII des personnes hémophiles, car sa structure est différente du facteur VIII et le rend donc non spécifique pour cette interaction [41]. La catalyse du facteur Xa peut ainsi être rétablie, et le mécanisme de coagulation est réactivé (figure 1.2.7).

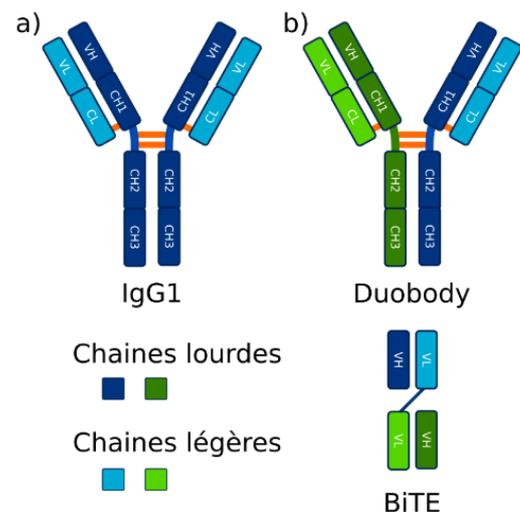


Figure 1.2.6 : Structure des anticorps thérapeutiques multi spécifiques a) Rappel de la structure d'une IgG1. b) Structures des deux types de mAbs thérapeutiques bispécifiques ayant été acceptés par la FDA.

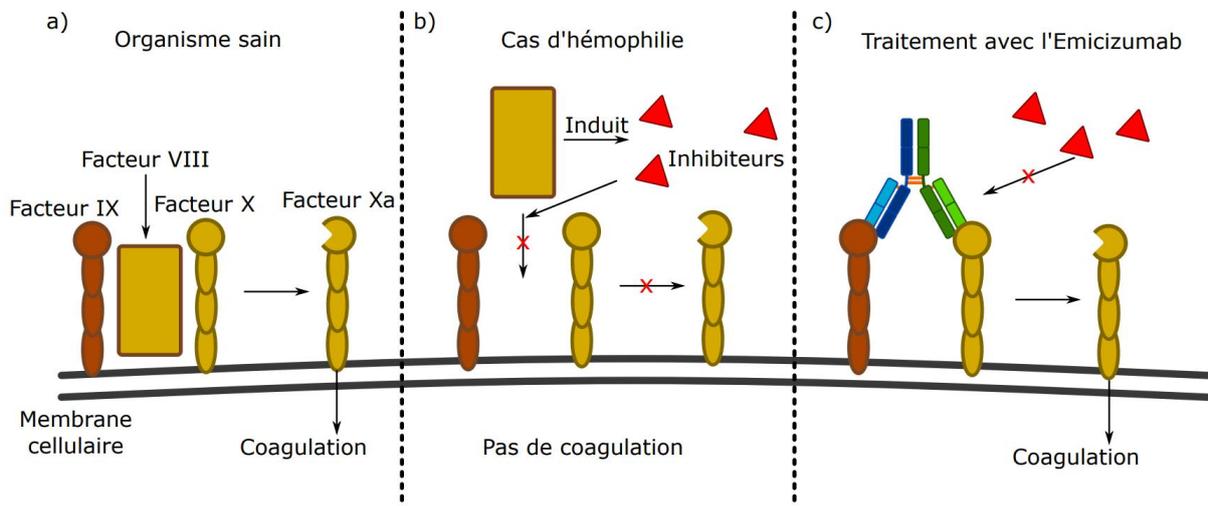


Figure 1.2.7 : Principe d'action de l'Emicizumab. a) Activité biologique d'une personne saine et mécanisme de la coagulation. b) Activité biologique chez une personne atteinte d'hémophilie A, le facteur VIII induit des inhibiteurs l'empêchant d'interagir avec les facteurs IX et X. c) Action du bispécifique Emicizumab sur les facteurs IX et X, ce qui réactive la transformation du facteur X en facteur Xa et donc la coagulation.

Les mAbs multi-spécifiques ne sont pas constitués exclusivement par des IgG entières. Des fragments d'anticorps tels que les mAbs thérapeutiques BiTE (*bispecific T cell engagers*) peuvent également être multi-spécifiques. En effet, les BiTE sont composés de fragments Fv provenant chacun de deux anticorps différents. Ils sont ainsi capables de rapprocher deux antigènes sans créer d'interactions supplémentaires. Le Blinatumomab (Blinicyto®), utilisé pour

certaines leucémies est un exemple de BiTE se liant spécifiquement aux protéines CD19 des lymphocytes B défectueux d'un côté, et aux protéines CD3 des lymphocytes T de l'autre, permettant ainsi l'activation de l'effet cytolytique (dégranulation) [42].

Il s'agit à ce jour des deux seuls formats de mAbs thérapeutiques multi-spécifiques ayant été admis sur le marché. Cependant, au niveau des essais cliniques, les mAbs multi-spécifiques représentent la catégorie de mAbs thérapeutiques comportant le plus de formats différents. La revue de Shijie Jin et al. 2022 présente une liste plus détaillée des différents formats en cours de développement [39].

### 2.3.3. Anticorps conjugués

Les anticorps conjugués, ou les ADC (*Antibody-Drug conjugates*), sont des IgG sur lesquelles une molécule toxique a été greffée au niveau de leur domaine CH3. Il s'agit d'une molécule cytotoxique, une molécule médicamenteuse qui agit sur le métabolisme des cellules et active leur apoptose. Les molécules cytotoxiques sont liées à l'IgG par une chaîne liante (le *linker*) qui est soigneusement construite car sa fonction est de libérer ou non la molécule toxique au bon moment. Le développement de la molécule active ainsi que de la chaîne liante fait appel à une conjugaison d'expertise en chimie organique, biologie moléculaire et de pharmacocinétique. Etant spécifiques, les ADC délivrent ces molécules très précisément à l'endroit où elles sont supposées agir. Le principe des ADC repose sur l'intégration des anticorps conjugués par des endosomes de la cellule cible, qui pourront ensuite relâcher ou non la molécule active (parfois cela n'est pas nécessaire lorsque l'anticorps est dégradé par les lysosomes). Deux catégories de molécules actives existent à l'heure actuelle : celles qui ont pour but d'inhiber l'activité des microtubules de la cellule, et celles qui ont pour but d'endommager son ADN (figure 1.2.8). Parmi les dix ADC sur le marché, sept molécules actives distinctes présentant l'un ou l'autre de ces mécanismes d'action existent.

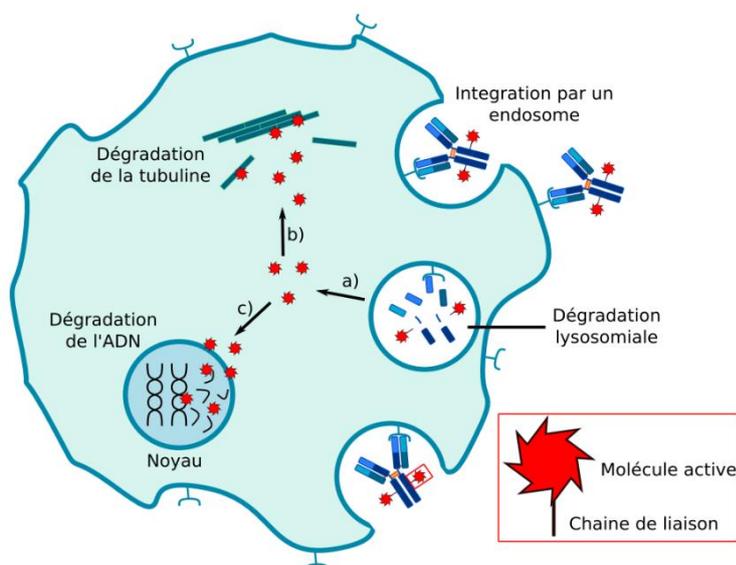


Figure 1.2.8 : Principe d'action des anticorps conjugués à des molécules actives (ADC). a) Libération de la molécule active après la dégradation lysosomiale. b) Molécules actives spécifiques pour la dégradation de la tubuline. c) Molécules actives intégrant le noyau cellulaire et pouvant dégrader l'ADN

Le Brentuximab vedotin (Adcetris®) est le premier ADC à avoir été autorisé par les agences de sécurité du médicament. Il est utilisé dans les traitements de certains lymphomes. Il présente une spécificité aux récepteurs CD30 des cellules tumorales. Son linker est non clivable à l'extérieur de la cellule, mais celui-ci est digéré par protéolyse après transfert du complexe

ADC-CD30 au sein d'un compartiment lysosomal. En se liant à la tubuline, la molécule toxique conjuguée, la monométhylauristatine, va ensuite perturber le cycle cellulaire induisant alors l'apoptose cellulaire [43].

#### 2.3.4. Les protéines de fusion au Fc

Les protéines de fusion au fragment Fc résultent de la conjugaison d'une protéine active couplée au domaine N-terminal d'un fragment Fc d'anticorps par fusion génétique. La présence du fragment Fc permet d'augmenter la durée de vie de la protéine active grâce à la présence du site de reconnaissance des récepteurs néonataux des fragments Fc (FcRn) qui permettent le recyclage du complexe (cf. partie 1.1.4.2). La protéine active possède alors une durée de vie aussi longue que celle d'un anticorps (sans en être un), et l'administration des doses peut intervenir de manière plus espacée. La présence du fragment Fc en fusion peut également constituer un moyen d'améliorer la solubilité de protéines difficiles à exprimer [44].

Le Dulaglutide (Trulicity®) est un exemple de protéine de fusion au Fc. Il est composé du domaine Fc d'une IgG4 dont les fragments Fab ont été remplacés par des peptides analogues du GLP-1 humain (*glucagon-like peptide-1*) attachés en N-terminal des chaînes lourdes de l'anticorps (figure 1.2.9). La protéine GLP-1 humaine est chargée de libérer de l'adénosine monophosphate cyclique en présence de concentrations de glucose élevées dans les cellules  $\beta$ -pancréatiques, ce qui entraîne la libération d'insuline. Chez les personnes diabétiques de type 2, la GLP-1 est inactive. La durée de vie naturelle du GLP-1 humain est de deux minutes au maximum dans l'organisme. La présence du fragment Fc conjugué à l'analogue augmente la durée de vie de la protéine jusqu'à près de cinq jours (avant élimination rénale naturelle), permettant une administration hebdomadaire du médicament. La protéine de fusion est un peptide dans cet exemple, mais peut également être une protéine telle que des cytokines, qui constituent une cible thérapeutique en plein développement [45].

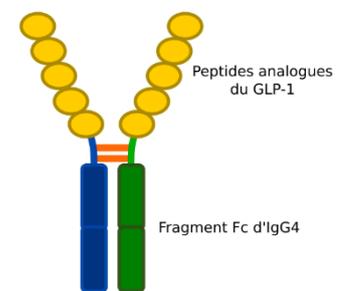


Figure 1.2.9 : Schéma structural du Dulaglutide, protéine de fusion au Fc d'une IgG4

L'activité thérapeutique des protéines de fusion au Fc ne résultant pas d'un anticorps, leur nom ne respecte pas la dénomination commune internationale des mAbs.

## 2.4. Application des mAbs thérapeutiques face à la COVID-19

Cette thèse a commencé au début de l'année 2020, aux prémices de la pandémie mondiale engendrée par le virus COVID-19. Trois ans plus tard, plusieurs millions de personnes en sont décédées. Etant donné l'impact majeur de ce sujet d'actualité, j'ai décidé de rédiger cette partie de manière plus personnelle afin de rendre compte de l'implication des anticorps thérapeutiques dans le traitement de ce virus.

Face à cette crise sanitaire mondiale, la course au vaccin a été très rapide, ainsi que la recherche et le développement d'anticorps thérapeutiques. Contrairement aux vaccins qui ont pour but de fournir une immunisation sur le long terme, les anticorps thérapeutiques sont utilisés dans de tels cas pour diminuer la charge virale de patients déjà malades, et ainsi atténuer les différents symptômes.

Les anticorps thérapeutiques de type neutralisants sont très efficaces pour lutter contre un virus [46]. De nombreux candidats, provenant de patients immunisés (après une première infection) ou de souris humanisées, ont fait leur apparition. Ces anticorps neutralisants ont pour but de limiter la prolifération du virus en ciblant principalement les protéines « spikes » présentes à la surface du virus, et responsables de l'internalisation du virus au sein des cellules [47] (figure 1.2.10). Le but majeur de ces mAbs est de diminuer la charge virale chez les patients « à risque » de manière à les protéger temporairement et éviter ainsi autant que possible les cas d'urgence (réanimation).

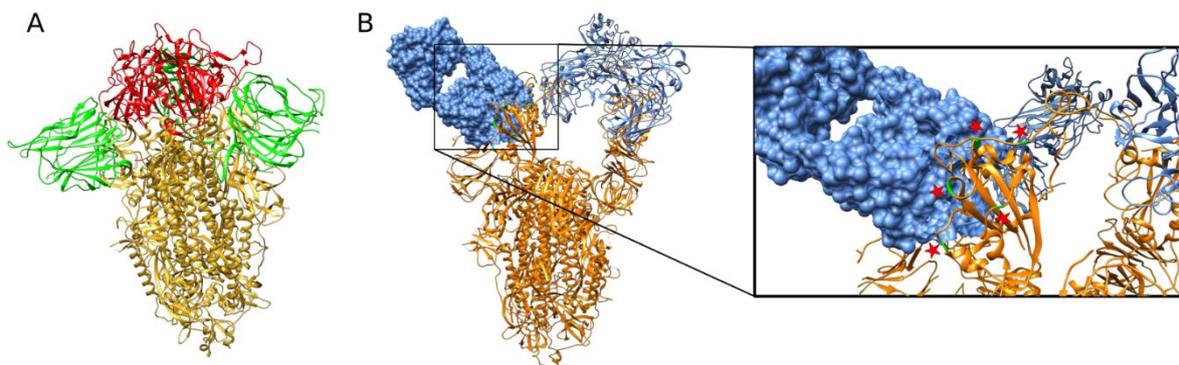


Figure 1.2.10 : Structure obtenue en cryo-EM de la protéine spike seule (A, PDB 6ZGE) et en complexe avec des fragments Fab d'IgG (B, PDB 7LD1). A) En rouge, le receptor binding-domain (RBD) et en vert les domaines N-terminaux (NTD). B) En orange la protéine spike, et en bleu les fragments Fab. Les étoiles rouges représentent les mutations observées sur les RBD pour le variant Omicron.

En Mai 2022, 66 candidats de mAbs thérapeutiques contre la COVID-19 étaient en phases d'essais cliniques, et la moitié d'entre eux faisaient l'objet d'une autorisation temporaire d'urgence de mise sur le marché [48]. Cependant, le développement d'un mAb spécifique au virus est très complexe en raison de la nature mutagène des virus. De nombreux variants de la COVID-19 se sont succédés, rendant alors inactifs de nombreux candidats potentiels (ainsi que certains vaccins), ciblant plus particulièrement les domaines RBD (*receptor-binding domain*) et NTD (*N-terminal domain*) de la protéine spike du virus COVID-19, très propices à la mutation.

Le Bamlanivimab en est un exemple. Il a été utilisé pendant quatre mois en disposant d'une autorisation temporaire d'urgence de mise sur le marché, puis a été abandonné face à l'apparition des variants Beta et Gamma qui ont peu à peu remplacé la souche initiale [49].

L'apparition du variant Omicron constitue également un exemple de la capacité de mutation des virus complexifiant le développement de mAbs thérapeutiques spécifiques sur le long terme. En effet, par rapport aux souches précédentes, ce variant peut présenter les mutations génétiques Q493R, N440K, G446S, S371L ou R436K (parmi d'autres) qui ont été reconnues comme étant responsables de la perte d'efficacité de presque tous les anticorps thérapeutiques alors en développement (figure 1.2.10) [50].

Une des solutions face à cette capacité de mutation est d'essayer de prévoir les futures mutations en identifiant les zones particulièrement variables sur le virus afin d'éviter le choix d'anticorps spécifiques à ces régions pour le développement des futurs anticorps thérapeutiques. Cette stratégie a été utilisée pour le développement d'un mAb neutralisant encore en phase de test. En ciblant les zones de la protéine spike moins soumises aux mutations, il s'est montré efficace contre la plupart des variants du virus, y compris des variants Omicron [51].

L'utilisation de formats spécifiques de mAbs thérapeutiques est aussi envisagée, et notamment les protéines de fusion au Fc. Une stratégie impliquant la protéine ACE2 (*angiotensin-converting enzyme 2*) fusionnée à un Fc a été développée [52]. Cette protéine est un récepteur viral situé sur les protéines humaines et interagissant avec la protéine spike des virus. Ce format d'anticorps neutralisant semble prometteur contre la Covid-19.

Malgré ces difficultés de développement, les anticorps thérapeutiques restent une solution très efficace pour atténuer le risque de forme grave de la pathologie chez certains patients.

## 2.5.L'anticorps anti-LAMP1 : modèle de la thèse

L'anticorps utilisé pendant la thèse est l'IgG1 anti-LAMP1 développé par Sanofi. Il s'agit d'un anticorps à visée thérapeutique de classe neutralisante et spécifique au domaine N-terminal de la protéine membranaire associée aux lysosomes de catégorie 1 (LAMP1 : lysosomal associated membrane protein 1). LAMP1 fait partie de la famille des protéines membranaires des granules cytotoxiques des cellules NK (*natural killer*). Ces granules cytotoxiques sont essentiels pour la fonction immunitaire des cellules NK et sont des lysosomes excrétoires spécialisés contenant des granzymes, de la perforine, de la granulysine, de la cathepsine et des protéines membranaires de la famille LAMP qui contient plusieurs membres : LAMP1/CD107a et LAMP2/CD107b, LAMP3/CD63, BAD-LAMP/UNC-46, ou encore macrosialin/CD68[53]. A la suite de la dégranulation, la protéine LAMP1 se retrouve sur la paroi membranaire des cellules NK (figure 1.2.11). Ainsi, LAMP1 est impliquée dans plusieurs aspects biologiques cellulaires tels que la phagocytose, l'autophagie, les transport lipidiques ou le vieillissement des cellules[54], [55]. Il a été démontré que son absence entraîne une autodestruction des cellules NK lors de leur dégranulation induite par une cellule cible (cancéreuse ou infectée par un virus donc reconnue et désignée comme cible par le système immunitaire)[56]. Il a aussi été démontré que la présence de LAMP1 sur des cellules diminuait l'action de la perforine contre ces cellules. Ainsi, LAMP1 semble avoir une fonction protectrice des cellules NK contre leur propre autodestruction.

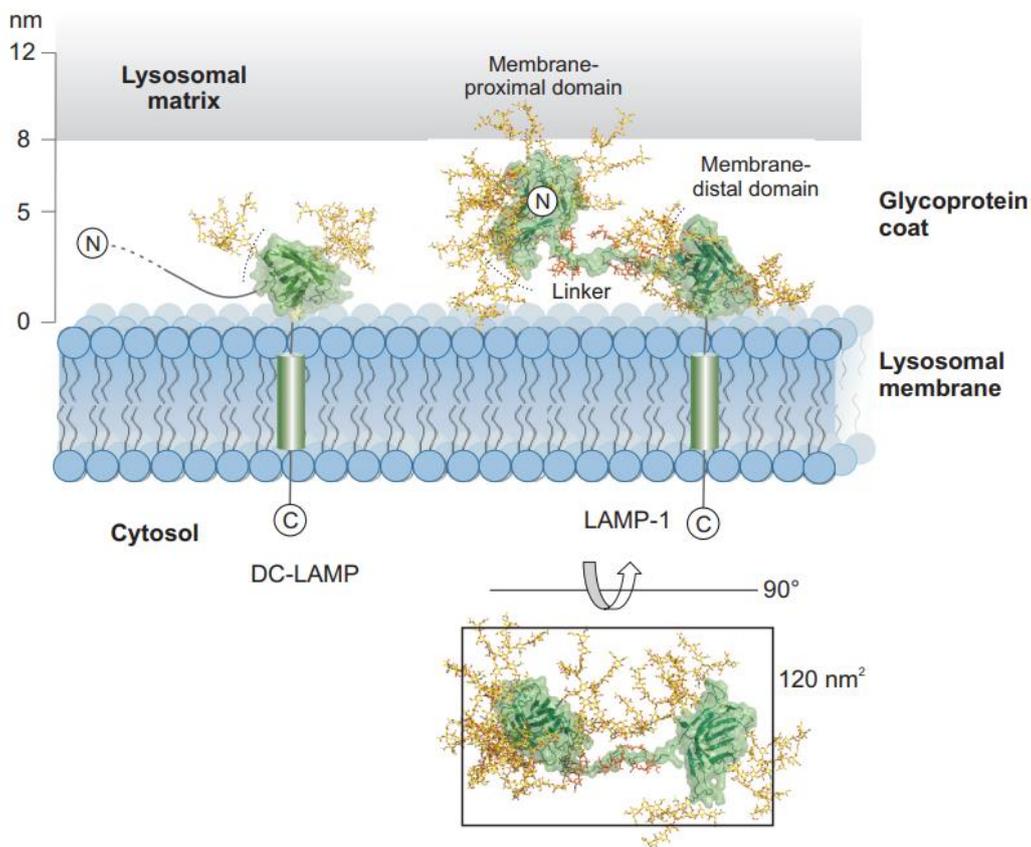


Figure 1.2.11 : Modèle structural de la protéine DC-LAMP (gauche) et LAMP1 (droite) à la surface des lysosomes avec la représentation des nombreuses glycosylations de la protéine (en jaune). Lors de la sécrétion de l'intérieur des lysosomes par les cellules NK, le domaine N-terminal de LAMP1 se retrouve exposé à la surface des cellules. S. Wilke et al. *BMC Biol*, vol. 10, p. 62, Jul. 2012, doi: 10.1186/1741-7007-10-62.

LAMP1 est présente à la surface des cellules NK saines, mais aussi à la surface de certaines cellules tumorales pour les cancers du sein, du colon, du tube digestif, de la prostate, des poumons ou encore des ovaires[55], [57], [58]. Sa présence entraîne une résistance de ces cellules cancéreuses face à la dégranulation des cellules NK et empêche leur élimination par le système immunitaire. Son implication dans ces pathologies en fait donc une cible de choix pour l'immunothérapie. L'anticorps anti-LAMP1 a été développé dans ce contexte et est une IgG1 de type neutralisant dont le but est de bloquer les récepteurs CD107a (LAMP1) des cellules cancéreuses et, ainsi, de les rendre de nouveau sensibles à la dégranulation des cellules NK. De manière spéculative, bloquer ces récepteurs entraînerait également une déplétion sélective des lymphocytes au moment de leur dégranulation[56]. Tout cela n'a pas encore été démontré et l'anticorps anti-LAMP1 n'a pas été mis sur le marché en raison d'un arrêt de son développement (pour des raisons non-divulguées).

LAMP1 est une protéine hautement glycosylée de 40 kDa possédant 17 sites de N-glycosylations et 9 sites de O-glycosylations qui peuvent doubler son poids moléculaire. Cette protéine membranaire est composée principalement de feuilletts  $\beta$  et la structure de son domaine N-terminal (en complexe avec le fragment Fab de l'IgG1 anti-LAMP1) a été déterminée par cristallographie des rayons X : code PDB 8ATH (figure 1.2.12)[59].

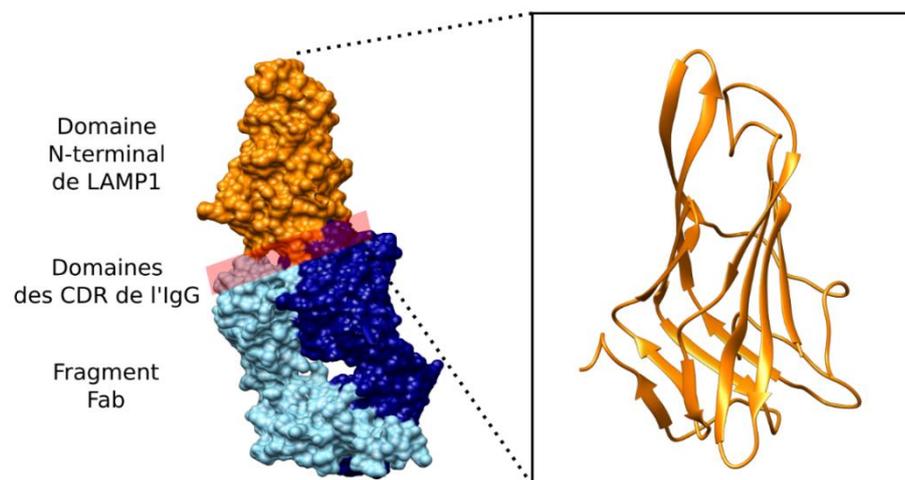


Figure 1.2.12 : Représentation de la structure cristallographique du complexe Fab/LAMP1 (gauche) et de la structure du squelette du domaine N-terminal de LAMP1 (droite). LAMP1 est représenté en orange. Pour le fragment Fab, la chaîne légère est en bleu clair et la chaîne lourde en bleu foncé. Structure PDB 8ATH. T. Pruvost et al., *mAbs*, vol. 15, no. 1, p. 2175311, Dec. 2023, doi: 10.1080/19420862.2023.2175311.

Les avancées en bio-ingénierie pour la production de mAbs thérapeutiques ont permis de développer des formats d'anticorps très variés afin de remplir des fonctions précises et spécifiques. La technologie, pour les développer, est toujours plus novatrice et on estime que pour un mAb thérapeutique mis sur le marché, dix sont en cours de développement ou en attente d'acceptation par les agences de sécurité du médicament.

La durée requise pour le développement des mAbs thérapeutiques est d'environ dix ans, et cela ne concerne pas uniquement le choix de la cible thérapeutique ainsi que le format du mAb. Un développement important est également réalisé sur le système d'expression pour leur production. En effet, le choix du système d'expression est un prérequis déterminant pour de nombreux paramètres tels que le rendement d'expression ou les profils de glycosylations recherchés.

### 3. Les systèmes d'expression des anticorps thérapeutiques

En 1975, la technique permettant de produire des anticorps monoclonaux, en utilisant des cellules hybridomes murines (souris et rat), est mise au point (cf. partie 1.2.1). Cependant, les anticorps issus d'hybridomes présentent certains inconvénients notables, tels que la réponse HAMA, provoquée par la présence de séquences murines immunogènes pour l'homme. Cela a ainsi conduit au développement de mAbs humains grâce aux progrès apportés par les techniques d'ADN recombinant, ainsi que du *screening* de séquences par *phage display*, permettant de transférer n'importe quel système d'expression cellulaire avec la séquence ADN d'anticorps sélectionnée. Le choix du système d'expression est un paramètre crucial qui dépend de plusieurs critères, tels que la capacité de celui-ci à produire des complexes protéiques glycosylés et riches en pont disulfures, l'application à laquelle est destiné l'anticorps produit (thérapeutique ou recherche), la quantité d'anticorps requise, la complexité de mise en œuvre de la production ou encore le coût de la culture. Cette partie détaille les capacités des différents systèmes d'expression pour la production d'anticorps ou de fragments en général (monoclonaux thérapeutiques, ou anticorps à usage clinique ou de recherche). Elle reprend les systèmes d'expression principaux dont, notamment, ceux utilisés pendant la thèse.

#### 3.1. Points clés de l'expression d'anticorps

L'utilisation des anticorps à destination de la recherche ou du médical (thérapeutique) nécessite des qualités et des quantités d'anticorps différentes. Ainsi, le choix du système d'expression, influant sur divers paramètres tels que les glycosylations des anticorps, la quantité de protéine exprimée, ainsi que sur la mise en œuvre des différentes étapes post-production (extraction et purification des protéines excrétées dans le milieu ou non), est principalement déterminé par le type de protéine à produire (IgG aux fonctions effectrices actives ou non, fragment Fab, scFv ou Fc).

##### 3.1.1. Les différentes méthodes de productions

###### a Pour l'industrie ou la recherche

L'obtention d'un anticorps thérapeutique débute par l'étude de son fonctionnement, sa stabilité, son activité ou sa structure, par diverses techniques nécessitant des quantités d'anticorps de l'ordre du milligramme. La progression du produit à travers les différentes phases de validation des thérapeutiques implique de fournir des quantités d'anticorps de plus en plus importantes. Lors des phases pré-cliniques et cliniques, l'injection du produit est testée, nécessitant des quantités pouvant aller du gramme au kilogramme selon les doses thérapeutiques administrées et le nombre de patients impliqués dans les essais. Si le biothérapeutique répond avec succès à l'ensemble ces étapes, il arrive en phase de commercialisation et des centaines de kilogrammes sont alors requis (figure 1.3.1).

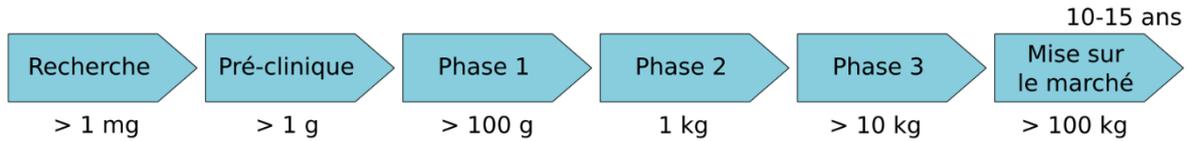


Figure 1.3.1 : Aperçu de la quantité d'anticorps nécessaire pour les différentes phases de développement de la recherche à la commercialisation

Le système d'expression privilégié pour les anticorps monoclonaux sont les cellules de mammifères, mais le développement d'une souche cellulaire suffisamment stable et capable de produire de telles quantités avec des critères de qualité stricts est initié avant les essais pré-cliniques.

Trois méthodes principales de culture cellulaire sont identifiables et vont influencer les rendements de production : le *batch*, le *fed-batch*, et la production en perfusion. Le choix du système se fait en fonction de la quantité de protéine nécessaire, du coût des réactifs (le milieu de culture), et du temps requis pour la production. Ces systèmes ont un impact sur la capacité des cellules à se multiplier et à survivre et vont donc exercer une influence sur la quantité de protéine produite.

### b La production en mode *batch*

Le mode de production en *batch* est la moins complexe. Les cellules sont inoculées à une certaine densité cellulaire (spécifique à chaque espèce et souche cellulaire) dans un milieu de culture pouvant contenir des compléments nutritifs et des pressions de sélection. Les cellules vont se multiplier et produire la protéine d'intérêt jusqu'à épuisement des nutriments du milieu. De manière générale, la production est arrêtée lorsque les cellules initient leur phase d'apoptose (figure 1.3.2) afin d'éviter une accumulation trop importante de débris cellulaires, qui serait délétère pour l'extraction et la purification de la protéine d'intérêt. Cette méthode de production nécessite seulement quelques heures à quelques jours (3 à 24h en bactéries, et jusqu'à environ 4 jours en cellules de mammifères) et permet d'obtenir très rapidement une quantité de protéine suffisante pour des applications de recherche. La mise en place est également très simple et nécessite peu de réactifs.

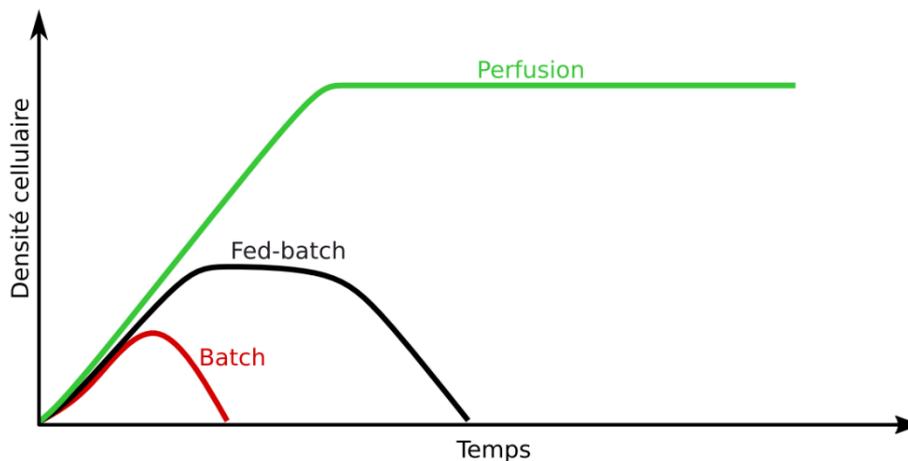


Figure 1.3.2 : Profil de la croissance cellulaire en fonction des différentes méthodes de culture : *batch*, *fed-batch*, et *perfusion* (bioréacteur). Ces profils de croissance cellulaire théoriques sont donnés pour la même densité cellulaire initiale.

### **c La production en mode *fed-batch***

Le mode de production en *fed-batch* est un processus plus long au cours duquel les cellules vont être supplémentées avec un apport de compléments nutritionnels (acides aminés, vitamines, glucose). Le principe de cette méthode est de compléter le milieu de culture cellulaire afin de maintenir les cellules en culture plus longtemps afin de permettre une production en protéines supérieure. Avec cette méthode, les cellules atteignent leur densité maximale après un temps donné et ne dépérissent pas immédiatement. Après ce temps de latence, celles-ci finissent cependant par entrer en phase d'apoptose en raison de l'accumulation de déchets biologiques dans le milieu, rendant celui-ci de plus en plus toxique pour les cellules. Cette méthode de production est souvent utilisée en recherche industrielle afin d'évaluer différents clones producteurs d'anticorps thérapeutiques et permet de maintenir la viabilité cellulaire pendant une dizaine de jours.

### **d La production en perfusion**

La production en perfusion est rendue possible grâce à l'utilisation de bioréacteurs. Cette méthode a été une avancée majeure pour la production des anticorps monoclonaux. L'utilisation d'un bioréacteur permet l'ajout de nutriments ou de compléments pour optimiser l'osmolarité du milieu ainsi que la possibilité de faire des prélèvements de milieu cellulaire pour éliminer les déchets cellulaires qui deviennent, à terme, toxiques pour les cellules. La soustraction du milieu de culture est réalisée en filtrant les cellules afin de maintenir celles-ci en suspension dans le bioréacteur et de conserver la densité cellulaire à la valeur souhaitée. Certains systèmes permettent même de purifier en continu les anticorps qui se trouvent dans le milieu de culture prélevé.

Grâce à ce système, les cellules de mammifères sont capables de survivre jusqu'à 30 jours sans aucune étape de dilution cellulaire. L'ingénierie de ces bioréacteurs a été complexifiée pour la culture en cellules CHO, jusqu'à atteindre des volumes de culture d'environ 20 m<sup>3</sup> et ainsi de produire des centaines de kilogrammes d'anticorps en un seul bioréacteur (cf. partie 1.3.2.2).

## **3.1.2. Aspects qualitatifs de la production d'anticorps**

### **a La monoclonalité des anticorps**

Les anticorps thérapeutiques sont dits monoclonaux car ils sont issus d'une seule cellule capable de se multiplier à l'infini et ayant intégré la séquence codant l'anticorps dans son propre génome. L'obtention d'un clone cellulaire est longue car les cellules transfectées n'intègrent pas toutes l'ADN de la même manière. Après transfection, elles doivent passer par un processus de sélection clonale long et rigoureux. Pour augmenter les chances d'obtention d'un clone stable, plusieurs milliers de clones sont placés dans des plaques 96 puits à la densité d'une cellule par puits. Lors de ce processus, les clones sont initialement sélectionnés sur le critère de leur robustesse. En effet, certaines cellules ne résistent pas à la transfection, et d'autres ne résistent pas à l'agressivité d'une culture à cette densité. La seconde étape, généralement

soumise à l'utilisation d'une pression de sélection, est basée sur la capacité des clones à se multiplier. La durée de leur cycle cellulaire est comparée et les clones se multipliant le plus rapidement sont sélectionnés afin de subvenir à l'objectif final d'obtenir rapidement de la biomasse. Enfin, lorsqu'une ou des dizaines de clones ont été identifiés à la suite de ces deux étapes initiales, les candidats sont alors évalués sur leur productivité. Les clones cellulaires les moins productifs sont éliminés et le candidat retenu présente ainsi les caractéristiques de robustesse, de multiplication rapide, et une haute capacité de production d'anticorps, compatibles avec la production industrielle de biothérapeutiques. Cette lignée cellulaire est appelée lignée cellulaire clonale stable et sert pour la production de tous les lots d'un même biothérapeutique. Cette méthode, longue, coûteuse et complexe à mettre en place est appliquée principalement pour la production d'anticorps (ou protéines) thérapeutiques.

La sélection clonale peut s'avérer moins sélective en définissant une densité cellulaire plus élevée au début du procédé (250 cellules par puits par exemple). L'anticorps produit n'est pas monoclonal et n'est alors pas utilisé en tant qu'anticorps thérapeutique car ces lignées cellulaires sont moins stables que les lignées clonales et sont susceptibles de présenter des mutations génétiques aléatoires et hétérogènes pouvant altérer leur croissance, leur productivité ou leur robustesse sur le long terme. A court terme, ces lignées peuvent cependant être utilisées pour la production d'anticorps à des fins de recherche. La lignée cellulaire utilisée pour la production de l'anticorps pendant la thèse fait partie de cette catégorie, qui peut être dite semi-clonale. La durée d'obtention de ce type de lignée cellulaire est moins longue que pour les lignées clonales et la technique est moins coûteuse, mais aboutit généralement à une productivité inférieure (cela est cependant dépendant de l'espèce cellulaire, de la méthode de production et de la protéine produite).

Pour réduire le temps nécessaire à la production d'un anticorps à partir d'une lignée stable, les anticorps (ou protéines) peuvent être exprimés de façon transitoire. Dans ce cas, les cellules sont transfectées avec l'ADN de l'anticorps, sans que celui-ci ne soit intégré dans le génome de la cellule. La production d'un anticorps transitoire dure environ 5 à 7 jours. Ce type de production n'est pas stable et les rendements sont inférieurs aux lignées semi-clonales, cependant des quantités de l'ordre du milligramme sont suffisantes pour la recherche et le développement en amont de l'usage clinique du mAb, et l'expression transitoire est donc adaptée pour ces étapes. Ainsi, cette méthode constitue une solution adéquate pour initier des études structurales ou d'activité avant l'obtention d'un clone stable producteur de l'anticorps.

## **b Les ponts-disulfures**

Les ponts disulfures sont essentiels au maintien de l'intégrité structurale des anticorps (cf. partie 1.1.1). Le système d'expression sélectionné doit donc être capable de produire des protéines dont les ponts disulfures sont correctement formés.

Environ 7 000 protéines d'un protéome eucaryote comportent des ponts disulfures. Cette modification concerne particulièrement les protéines qui sont excrétées hors du cytoplasme par les cellules. Les protéines localisées dans le cytoplasme ne comportent généralement pas (ou peu) de ponts disulfures en raison du caractère réducteur de cet environnement. Le milieu extra-cellulaire est oxydant mais peut cependant contenir de nombreuses protéases capables de dégrader les protéines. Les ponts disulfures participent au maintien des protéines sous forme structurées, dont le cœur hydrophobe est isolé du solvant, contribuant ainsi à les préserver des protéases.

La formation des ponts disulfures dans les cellules n'a pas lieu au même endroit que leur synthèse. Ces différents compartiments sont délimités par des membranes afin de procurer respectivement un milieu plus oxydant propice à la formation des ponts disulfures, ou plus réducteur favorable à la synthèse des protéines. Ces membranes définissent le périplasma des bactéries ou le réticulum endoplasmique et les mitochondries des cellules eucaryotes. Pour les bactéries ou les cellules eucaryotes, les acteurs impliqués peuvent être différents, mais le mécanisme global de formation est similaire.

Un pont disulfure est une modification réversible des protéines dont le résultat est l'oxydation de deux groupements sulfhydriles (SH, ou thiol) des chaînes latérales de deux cystéines, formant une liaison ou disulfure (S-S) covalente entre ces deux résidus. La formation de ponts disulfures au sein d'une protéine est influencée par deux facteurs. Premièrement, les cystéines considérées doivent être proches dans l'espace après le repliement et la formation de la structure secondaire de la protéine. Ce paramètre est protéine-dépendant et est favorisé chez les anticorps car le repliement des feuillets  $\beta$  composant les différents domaines d'immunoglobulines favorise le rapprochement des cystéines formant les ponts disulfures intra-chaînes (cf. partie 1.1.2). La formation de ponts disulfures inter-chaînes est, au contraire, propice à l'obtention d'anticorps non-natifs, c'est-à-dire présentant un mésappariement des chaînes lourdes et légères entre elles, indépendamment de leur structure secondaire. Le second paramètre important est que le milieu doit être suffisamment oxydant pour permettre la formation des ponts disulfures. Toutefois, un environnement trop oxydant, dans lequel les ponts disulfures seraient très stables, est à proscrire car il pourrait contribuer à piéger les protéines non natives présentant des ponts disulfures mal formés en empêchant leur isomérisation vers la forme native. Ce paramètre est ainsi particulièrement essentiel pour les protéines riches en ponts disulfures telles que les anticorps.

*In vivo*, la formation des ponts disulfures est le résultat d'une chaîne de réactions ayant à sa base l'oxygène, qui fournit le pouvoir oxydant nécessaire aux réactions et permet l'oxydation d'enzymes telles que des oxydases à sulfhydriles ou d'oxydoréductases qui agissent de concert pour aboutir à la formation des ponts disulfures. La régulation de la formation des ponts disulfures pour l'obtention de protéines natives et stables nécessite le maintien d'un potentiel rédox au sein des différents compartiments cellulaires. Ce potentiel est régulé par la présence d'agents réducteurs comme le glutathion réduit (GSH). La forte concentration de cette molécule sous forme réduite dans le cytoplasme cellulaire (12 mM) est à l'origine du caractère réducteur de ce milieu. Dans les compartiments propices à la formation de ponts disulfures, la concentration du glutathion sous forme réduite est beaucoup plus faible, au profit de la forme oxydée, la GSSG. Le *ratio* du couple GSH/GSSG assure le maintien du potentiel redox du compartiment à un niveau suffisamment élevé pour permettre la formation des ponts disulfures. Dans ces compartiments, la formation des ponts disulfures natifs est accélérée par la présence d'enzymes pouvant être classées en deux catégories agissant respectivement sur la formation ou l'isomérisation des ponts disulfures.

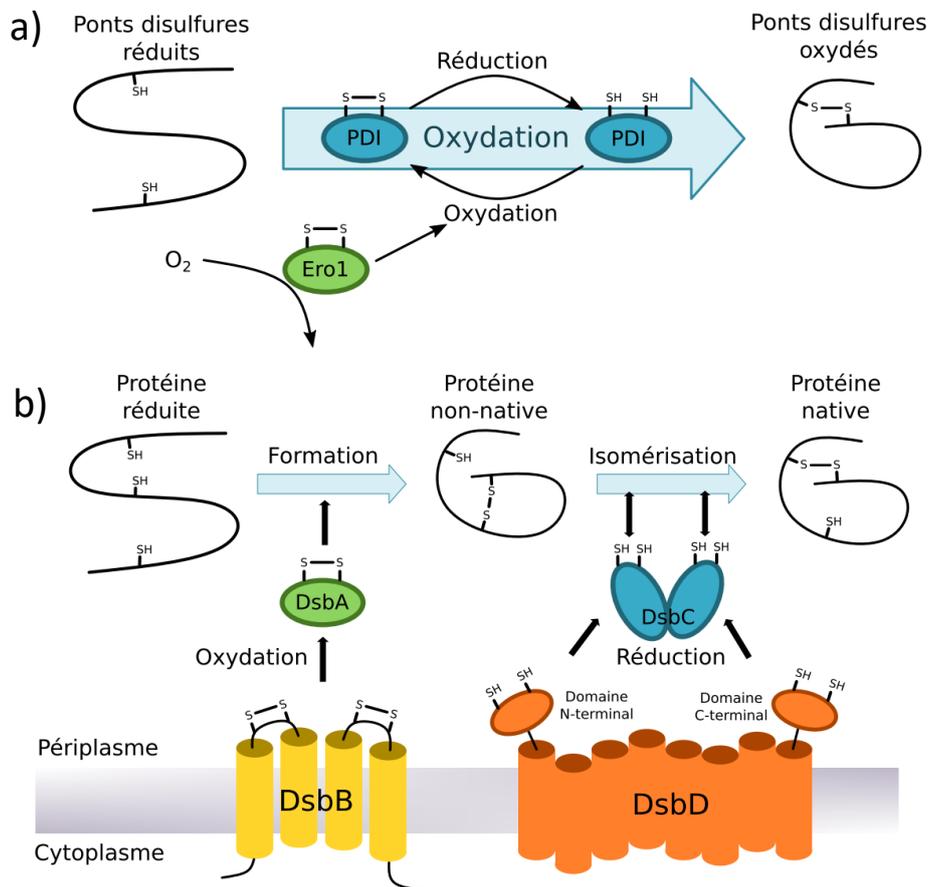


Figure 1.3.3 : Formation des ponts disulfures grâce à l'activité de la PDI et de l'Ero1 pour les cellules eucaryotes a) et des enzymes de la famille des Dsb présentes dans certaines bactéries b).

Une première catégorie d'enzymes permet la catalyse de la formation des ponts disulfures comme la Ero1 pour les cellules eucaryotes (figure 1.3.3 a) ou la DsbA (*Disulfide bond isomerase A*) pour certaines bactéries (figure 1.3.3 b). Ces enzymes sont maintenues sous leur forme oxydée par d'autres enzymes (la DsbB pour la DsbA des bactéries).

La seconde catégorie d'enzymes sont maintenues sous forme réduite et catalysent l'isomérisation des ponts disulfures sur la protéine d'intérêt par échange avec leur propres ponts disulfures jusqu'à ce que la protéine adopte une conformation plus stable et moins réactive. La figure 1.3.3 b représente le mécanisme simplifié du fonctionnement des chaperonnes à ponts disulfures de la famille des Dsb pour certaines bactéries. Le rôle exact du glutathion (réduit et oxydé) dans ce mécanisme est très complexe et certaines voies biologiques impliquant celui-ci sont encore mal comprises. La revue de S. Chakravarthi *et al.* propose une description plus détaillée [60].

### c Les glycosylations

La présence de N-glycosylations sur les asparagines 297 des chaînes lourdes des anticorps implique des contraintes au niveau du choix du système d'expression. En effet, tous les systèmes d'expression ne possèdent pas les enzymes assurant la formation des N-glycosylations avec un profil non immunogène pour l'homme. Lorsque l'anticorps est destiné à des applications thérapeutiques impliquant l'injection chez l'homme, il est nécessaire de s'assurer que le système d'expression choisi ne conduira pas à un anticorps dont le profil de glycosylation est immunogène pour l'individu, ainsi que d'évaluer son impact sur l'activité des sites effecteurs du fragment Fc.

La glycosylation est le résultat d'ajouts de glycanes par le biais d'enzymes, sur différentes structures du vivant afin de former les glycolipides, glycoprotéines, exopolysaccharides ou encore oligosaccharides. La N-glycosylation n'est pas limitée aux cellules de mammifères. Il a en effet été montré que certaines souches de bactéries, telles que *Campylobacter jejuni* (*C. jejuni*), comportaient les enzymes capables de réaliser ce type de modifications post-traductionnelles [61]. Cependant, le profil de glycosylation résultant est très différent par rapport à celui obtenu dans les cellules eucaryotes. En effet, dans les bactéries, la N-glycosylation ne commence pas avec un GlcNac comme pour les eucaryotes (cf. partie 1.1.2), et les glycanes suivants consistent en une succession de GalNac ajoutés les uns à la suite des autres. De manière artificielle, Xuan et son équipe sont parvenus à ajouter un premier GlcNac en utilisant une souche *E. coli* modifiée génétiquement exprimant la glycosyltransférase PgIB de *C. jejuni* [62]. Ils ont ensuite réussi à obtenir des profils de glycosylation humaine par réactions chimiques et enzymatiques, mais les rendements (de glycosylation correcte) obtenus restent faibles. Bien que possible, l'utilisation de bactéries n'est pas encore envisageable pour la production d'anticorps glycosylés.

Dans les cellules eucaryotes, la glycosylation a lieu principalement dans le réticulum endoplasmique et l'appareil de Golgi et fait intervenir de très nombreuses glycosyltransférases (environ 200), glycosidases et autres enzymes [63]. Bien que toutes les cellules eucaryotes soient capables de réaliser la N-glycosylation, le profil d'enchaînement des glycanes est spécifique du système utilisé.

De nombreuses études ont été réalisées concernant l'activité des sites effecteurs du fragment Fc des anticorps en fonction des différents profils de glycosylation. La fucosylation (ajout d'un fucose) sur le premier GlcNac, visible dans 94% des IgG humaines (figure 1.3.4), diminue fortement l'interaction de l'anticorps avec les phagocytes ainsi qu'avec les FcR lors de l'activation des cellules cytotoxiques dépendantes des anticorps (ADCC) [64]. Le système d'expression du rituximab (cellules d'ovaires de hamster chinois : CHO) a par exemple été optimisé de façon à réduire sa fucosylation afin d'augmenter son efficacité thérapeutique [65]. La présence d'un acide sialique en terminaison de la glycosylation a également un impact négatif sur la reconnaissance du Fc par les FcR, car il a été constaté que ce glycane chargé a pour effet de refermer le site de liaison aux FcR [66].

L'organisme humain a évolué de façon à développer des profils de glycosylation spécifiques (figure 1.3.4) [13]. Lors de la production d'un anticorps thérapeutique, une glycosylation non humaine peut entraîner une réponse immunitaire vis-à-vis du biothérapeutique injecté. C'est dans le but d'éviter cette réponse immunitaire que certains systèmes d'expression comme les cellules de mammifères sont préférés à d'autres.

L'ensemble de ces paramètres sont dépendants du type d'anticorps à produire (application thérapeutique ou pour la recherche), et le système d'expression utilisé influe sur les critères de qualité et de quantité d'anticorps produit. Tous les types de systèmes d'expression ont été utilisés pour la biosynthèse d'anticorps ou de fragments et chacun présente des avantages et inconvénients. Ces avantages peuvent se situer au niveau des modifications post-traductionnelles possibles, comme les ponts disulfures et les glycosylations, qui sont propres à chaque espèce. Le système d'expression influe également sur les étapes de post-production telles que l'extraction et la purification, car certaines cellules sont capables d'excréter les anticorps dans le milieu de culture tandis que d'autres les produisent exclusivement de manière intracellulaire, dans leur cytoplasme ou périplasme (pour les bactéries).

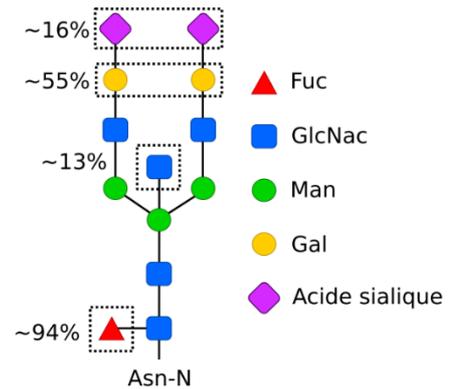


Figure 1.3.4 : Profil de glycosylation humaine et pourcentage de présence des différents glycanes ajoutés sur le motif constant

## 3.2. Production d'anticorps en cellules eucaryotes

### 3.2.1. Les cellules murines

#### a Les hybridomes de souris

Les premiers anticorps monoclonaux ont été produits à partir d'hybridomes de souris (mAbs murins). Cette méthode de production a été développée par Georges Jean F. Köhler et César Milstein (cf. partie 1.2.1.2), en infectant des souris avec des globules rouges de moutons (l'antigène) afin qu'elles développent une immunité et produisent ainsi des anticorps (figure 1.3.5). Leur but étant de produire des anticorps en grande quantité en se dispensant de l'usage de la souris, ils ont extrait les splénocytes (cellules immunitaires de la rate) de la souris. En quittant la rate, ces cellules sont capables de se diversifier en n'importe quel type de globule blanc dont des lymphocytes B sécréteurs d'anticorps. En revanche, ces cellules ne sont initialement pas capables de se reproduire *in vitro*. Ils ont alors l'idée de fusionner les splénocytes producteurs d'anticorps avec des cellules de myélome lymphoïde qui, elles, présentent la capacité de se reproduire à l'infini dans un milieu de culture. Le produit de cette fusion est appelé hybridome et possède les deux caractéristiques principales issues des cellules mères respectives. L'hybridome est capable de produire un anticorps identique à celui produits par les splénocytes originaux, et possède la capacité de se reproduire à l'infini comme les cellules de myélome.

Bien que cette méthode révolutionnaire basée sur les clones d'hybridomes ait permis la mise sur le marché du premier anticorps thérapeutique (Muromonab-CD3), celle-ci présente cependant le défaut de produire des anticorps dont les séquences murines sont immunogènes pour l'homme (réponse HAMA) conduisant finalement à privilégier la production d'anticorps recombinants.

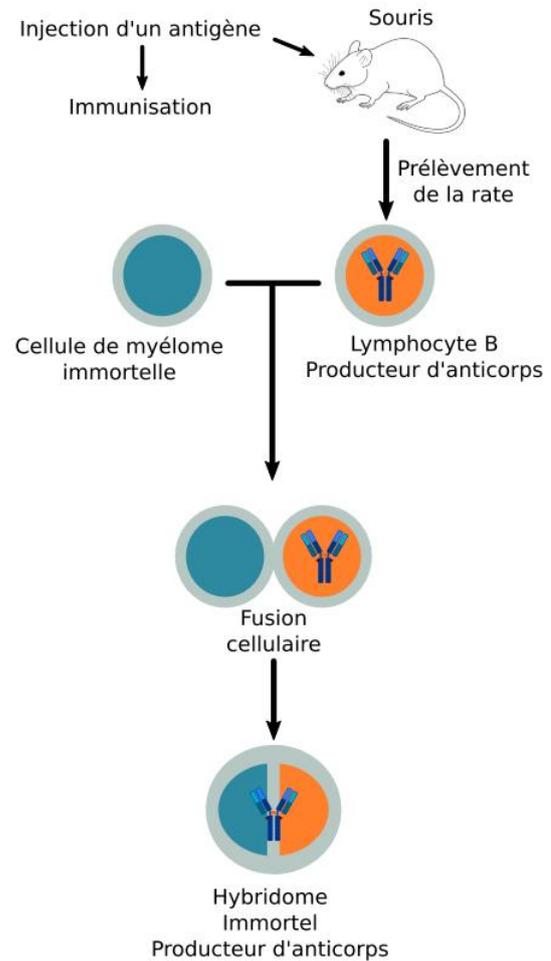


Figure 1.3.5 : Schéma représentant la création d'une cellule hybridome de souris productrice d'anticorps

## **b Les cellules de souris : SP2/0 et NS0**

Les avancés scientifiques en ingénierie génétique, avec l'apparition de l'ADN recombinant, ont permis de résoudre partiellement le problème de la réponse HAMA. Il a alors été possible de réaliser des transfections dans divers systèmes d'expression avec l'ADN codant la séquence d'anticorps souhaitée. Les cellules hybridomes SP2/0 et de myélomes murins NS0 ont constitué les souches initialement privilégiées pour la production des anticorps chimériques et humanisés recombinants.

Les cellules murines SP2/0 et NS0 ont été historiquement utilisées pour la production d'anticorps issus d'hybridomes. Les cellules SP2/0-Ag14 (de leur nom complet) constituent une lignée cellulaire clonale d'hybridomes de souris non sécrétrice d'immunoglobuline endogène. Cette lignée n'est pas résistante aux milieux de culture contenant de l'hypoxanthine, aminoptérine et thymidine (HAT). L'hybridation de cette lignée avec un partenaire possédant un gène codant pour l'anticorps et un gène de résistance aux HAT permet de sélectionner les clones qui expriment l'anticorps. Bien que les anticorps soient issus de l'hybridation de ces cellules avec un autre partenaire, le nom SP2/0 est conservé pour la lignée cellulaire productrice.

Les cellules NS0 (NS0/1 de leur nom complet) sont issues de tumeurs à plasmocytome (les MOPC21) générées chez une souris après injection d'une huile minérale, qui produisaient initialement des immunoglobulines endogènes. De multiples sélections clonales des MOPC21 ont été réalisées afin d'obtenir un clone non producteur d'immunoglobuline : les NS0/1 (*non-secreting immunoglobuline cells*). De même que la majorité des cellules de myélome, ces cellules ont d'abord été utilisées en tant que partenaires pour la création d'hybridomes (comme pour les SP2/0), puis la même lignée a été utilisée pour la production d'anticorps recombinants en utilisant comme pression de sélection la glutamine synthétase (GS). En effet, les NS0 ont une déficience naturelle pour la GS et, en transfectant les cellules avec le gène codant pour l'anticorps et la glutamine synthétase, il est possible de sélectionner les clones producteurs qui survivent dans un milieu de culture déplété en glutamine. Cette pression de sélection constitue une méthode très utilisée pour les cellules de mammifères. La production de l'anticorps par les cellules peut ensuite être induite/augmentée en inhibant l'activité de la GS par l'utilisation de méthionine sulphoximine (MSX), un inhibiteur de la glutamine synthétase. L'activité de la GS étant réduite, les cellules sont poussées à surexprimer le gène de la GS et donc de l'anticorps.

L'utilisation des cellules SP2/0 et NS0 a constitué les premiers pas de la production des mAbs thérapeutiques, permettant d'obtenir des rendements atteignant jusqu'à 2 g/L. Cependant, ces cellules présentent l'inconvénient de produire des profils de glycosylation pouvant être immunogènes pour l'homme. En effet, les cellules de souris peuvent conduire à la formation de galactose-alpha-1,3-galactose ( $\alpha$ -gal) et d'acide N-glycolylneuraminique (acide sialique) (figure 1.3.4), qui induisent une immunogénicité humaine [67] ( $\alpha$ -gal : antigène de Galili ; présent chez la plupart des mammifères sauf chez les primates qui ont perdu, au cours de l'évolution, le gène à l'origine de sa formation (de même que l'acide sialique spécifiquement absent chez les humains) ; ils sont donc immunogènes). La présence de ces profils de glycosylation peut aussi entraîner une élimination plus rapide de l'anticorps par l'organisme comparé aux composés non-immunogènes, ce qui diminue le temps durant lequel le biothérapeutique est présent dans le sang et diminue donc son activité, constituant un facteur

limitant pour l'utilisation de ces lignées cellulaires pour la production d'anticorps monoclonaux thérapeutiques. Ces lignées cellulaires sont donc limitées à la production des mAbs thérapeutiques les plus anciens et représentaient environ 23 % des systèmes d'expression d'anticorps thérapeutiques en 2020 (source : FDA et EMA).

### 3.2.2. Les cellules ovariennes de hamster chinois : CHO

L'utilisation des cellules ovariennes de hamster chinois, les CHO, a marqué un tournant dans l'histoire de la bio-production des anticorps monoclonaux. Plusieurs souches modifiées ont ainsi été développées dans le but d'augmenter la production d'anticorps. En effet, ce sont des cellules très adaptées à la culture en suspension et donc à la production en bioréacteurs à haute densité cellulaire (mieux que les SP2/0 et NS0), et les anticorps produits sont porteurs de glycosylations proches de celles observées chez les humains. Les CHO sont aussi plébiscitées par l'industrie pharmaceutique et les agences de sécurité du médicament car l'accumulation des connaissances sur ce système est considérable, en faisant le système le plus contrôlé pour la bio-production des mAbs thérapeutiques. En 2022, plus de 71% des mAbs thérapeutiques étaient produits grâce à ce type de cellules et ce pourcentage tend encore à augmenter.

#### a Les différentes souches de CHO

Aujourd'hui, deux souches de cellules CHO sont principalement utilisées pour la production des mAbs thérapeutiques : les CHO-DG44 et les CHO-GS. Ces souches sont issues de la même souche initiale CHO-K1, modifiée différemment afin d'atteindre un potentiel producteur de mAb le plus élevé possible [68]. Les cellules CHO-GS sont issues d'une souche déficiente en glutamine synthétase dont le principe de fonctionnement de la pression de sélection a été expliqué dans la partie 3.2.1. Les CHO-DG44 sont des cellules déficientes en réductase à dihydrofolate (DHFR : *dihydrofolate reductase*), une enzyme essentielle pour la prolifération cellulaire en permettant la biosynthèse de tétrahydrofolate à partir du dihydrofolate. Cette souche est transfectée avec le gène codant la DHFR et le gène de l'anticorps sur un même plasmide. Une concentration élevée de méthotrexate, un inhibiteur de la DHFR, impacte l'activité de celle-ci, conduisant alors les cellules à surexprimer le gène de la DHFR et donc de l'anticorps [69]. Ces souches constituent les souches les plus utilisées en industrie pharmaceutique pour leur capacité d'expression de gènes exogènes.

La production d'anticorps recombinants transitoires pour la recherche peut, par ailleurs, être basée sur l'utilisation de la souche initiale des CHO (la CHO-K1) qui n'est pas déficiente en DHFR ou GS. Ces pressions de sélection drastiques sont alors être remplacées par des gènes de résistance à des antibiotiques. La lignée cellulaire utilisée pendant la thèse pour la production de l'anticorps est issue d'une souche CHO-K1 transfectée avec des plasmides comportant les gènes de l'anticorps anti-LAMP1 et de la résistance à la généticine et l'hygromycine.

## **b Production à haute densité cellulaire**

La production de mAbs thérapeutiques implique l'utilisation de bioréacteurs dont l'une des difficultés principales réside dans la nécessité d'un apport d'oxygène en quantité suffisante et de manière homogène à toutes les cellules. Plus les niveaux de densités cellulaires augmentent, et plus l'oxygénation de la culture devient critique, rendant le maintien de la viabilité cellulaire difficile. En 1987, Miller et *al.* ont montré que les cellules CHO peuvent consommer 50% d'oxygène en moins en conservant une viabilité et productivité stable lorsque la quantité de glucose est suivie et optimisée au cours de la culture[70]. Cela a notamment contribué à augmenter la taille des bioréacteurs, ainsi que la densité cellulaire utilisée. Les cellules CHO sont capables d'atteindre une densité cellulaire maximale de  $8.10^6$  à  $10.10^6$  cellules/mL de culture en flasque. En considérant l'avancée apportée par Miller et *al.* ainsi que l'optimisation des milieux de culture en bioréacteur, les cellules CHO sont désormais capables de croître jusqu'à atteindre des densité de  $25-30.10^6$  cellules/mL[71].

En 1990, les rendements de production d'anticorps monoclonaux en CHO étaient d'environ 100 mg/L en flasques. Aujourd'hui, grâce à toutes les avancées en génétique cellulaire, bio-ingénierie des bioréacteurs, et optimisation des milieux de culture, les rendements sont en moyenne de 3 g/L mais peuvent atteindre jusqu'à 10 g/L avec des volumes de culture de l'ordre de 2 000 à 20 000 L en bioréacteurs. Cette nouvelle technologie de culture cellulaire à très haute échelle permet le maintien des cellules en culture pendant une vingtaine de jours, permettant d'atteindre, en un seul bioréacteur, une production finale de centaines de kilogrammes de mAbs thérapeutiques.

En raison du prix de certains réactifs, pour les applications de recherche, la production en *batch*, dont les rendements sont plutôt de quelques dizaines de mg/L, est privilégiée. Cependant, le rendement de production d'anticorps en cellules CHO est généralement modulé selon la lignée cellulaire considérée ou le type de production utilisé (bioréacteur, *fed-batch*, *batch*).

## **c Glycosylations des CHO**

Par rapport aux cellules murines, les cellules CHO ont, pour la production des anticorps thérapeutiques, l'avantage de ne pas posséder la glycosaminetransférase responsable de la formation de galactose- $\alpha$ -1,3-galactose ( $\alpha$ -gal) et d'acide N-glycolylneuraminique, immunogènes pour l'homme.

De plus, l'absence de fucosylation sur la glycosylation du fragment Fc (en position N-297), a pour effet de renforcer la liaison de l'anticorps aux FcR (notamment avec ceux des cellules NK). L'absence de fucosylation contribue donc à augmenter l'activité thérapeutique de l'anticorps. Les connaissances approfondies accumulées sur ce type de cellules ont permis l'identification du gène codant l'enzyme responsable de la fucosylation des protéines, et sa suppression dans les souches CHO afin de renforcer l'activité des mAbs thérapeutiques produits. Ces cellules modifiées sont appelées CHO-FUT8-KO [72], correspondant à la lignée cellulaire CHO-DG44 ou -K1 dont le gène FUT8 a été supprimé (KO : *knock-out*). A titre d'exemple, les anticorps rituximab, Benralizumab et Mogamulizumab sont issus de souches CHO modifiées produisant des anticorps faiblement fucosylés.

### **3.2.3. Les autres cellules eucaryotes pour les anticorps thérapeutiques**

Les anticorps thérapeutiques sont principalement produits par les trois types de cellules décrites précédemment : CHO, NS0 et SP2/0. Cependant, d'autres systèmes existent et ceux ayant été utilisés dans la production d'anticorps thérapeutiques sont listés dans cette partie.

#### **a Cellules humaines : PerC6 et HEK 293**

Les cellules humaines présentent l'avantage de disposer de toute la machinerie nécessaire pour produire des anticorps de qualité compatible à un usage thérapeutique chez l'homme. Les modifications post-traductionnelles et les glycosylations qu'elles sont capables d'effectuer sont représentatives de celles réalisées dans un organisme humain. Le risque d'immunogénicité est donc très réduit.

Cependant l'origine humaine de ces cellules leur confèrent l'inconvénient d'être susceptibles de se trouver infectées par des virus humains, potentiellement portés par le manipulateur. Elles sont donc plus fragiles que les cellules CHO ou murines qui ne peuvent pas être soumises à ce genre d'infection.

Les cellules PerC6 (cellules de rétine humaines immortalisées) ont été utilisées pour la production d'un anticorps thérapeutique qui n'a pas passé toutes les phases de tests et dont la production a été abandonnée, en raison d'une activité de l'anticorps inférieure aux attentes, et non d'un problème lié aux cellules. Il s'agissait de l'Otolimab qui ne s'est pas montré suffisamment efficace pour traiter les cas d'arthrite rhumatoïde sévère.

Les cellules HEK 293 (souvent simplement appelées HEK) sont des cellules embryonnaires de rein humain et correspondent à l'une des classes de cellules de mammifères les plus utilisées en laboratoire. Elles n'ont pas été utilisées pour la production de mAbs thérapeutiques, mais pour la production de protéines thérapeutiques de fusion au Fc. Le Dulaglutide (cf. partie 1.2.3.4) a été produit à partir des HEK.

Les cellules humaines sont considérées comme étant très prometteuses pour la production d'anticorps thérapeutiques [73]. Elles sont également souvent utilisées en laboratoire pour la production d'anticorps recombinants destinés à la recherche.

#### **b Les myélomes de rat YB2/0**

Les cellules de myélomes de rat YB2/0 sont utilisées pour la production de l'Ublituximab (Briumvi<sup>®</sup>), dont la mise sur le marché a été admise en 2022. Cet anticorps, dirigé contre les récepteurs CD20 des cellules tumorales [74], est indiqué dans les cas de leucémie lymphoïde chronique et de sclérose en plaques de type récurrente-rémittente. Le gène FUT8 codant l'enzyme responsable de la fucosylation est faiblement exprimé dans les cellules YB2/0, ce qui rend ce système d'expression très intéressant pour la production d'anticorps faiblement fucosylés [75]. De plus, pour ce type de cellules, l'expression de ce gène est aussi dépendant de l'osmolarité du milieu de culture, qui peut être optimisée grâce au contrôle des divers paramètres de culture, [76] (idéale lors d'une production en bioréacteurs).

### c Les levures

Les levures présentent l'avantage de se multiplier plus rapidement que les cellules de mammifères (doublement de la population toutes les heures contre 10 à 24 h pour les cellules de mammifères). De plus, le milieu de culture de ces cellules est simple et peu coûteux. La culture de *Pichia pastoris* en bioréacteur a permis d'obtenir des rendements de 1 g/L de mAb [77]. Il s'agit donc d'un système de production plus économique que les cellules de mammifères. Cependant, les profils de glycosylation naturellement issus de cette espèce sont immunogènes pour l'homme. Ceci représente un inconvénient pour la production d'IgG entières, dont les sites effecteurs du fragment Fc doivent être actifs, mais pas pour les fragments ne contenant que les parties variables. Ce système est ainsi considéré comme étant très intéressant pour la biosynthèse des fragments de mAbs thérapeutiques (Fab, scFv, F(ab')<sub>2</sub>, nanobodies et autres catégories).

L'Eptinézumab (Vyepi<sup>®</sup>) est un anticorps thérapeutique antagoniste et inhibiteur du peptide CGRP impliqué dans les cas de migraines épisodiques et chroniques. Cet anticorps ne fait pas intervenir de fonctions effectrices issues de son fragment Fc et il n'est donc pas nécessaire qu'il soit glycosylé. Ainsi, celui-ci a été muté sur le site de glycosylation (N297A), et sa production a été réalisée dans une souche de *Pichia pastoris* [78].

La modification génétique des levures pour introduire les gènes codant les enzymes responsables des glycosylations humaines est en cours de développement et pourrait se révéler très prometteuse [79].

### **3.3. Production d'anticorps thérapeutiques en système procaryote : les bactéries**

L'utilisation de bactéries pour la production de protéines recombinantes présente l'avantage d'être moins coûteuse que l'utilisation des systèmes eucaryotes. Les bactéries se multiplient rapidement et nécessitent des milieux de culture peu coûteux. Ce système d'expression est cependant limité pour la production de protéines riches en ponts disulfures et glycosylées, telles que les IgG et/ou certains de ses fragments. Certaines bactéries possèdent les enzymes nécessaires à la réalisation de glycosylations, mais le profil de glycanes formé n'est pas viable pour des protéines humaines (cf. partie 1.3.1.2 c). Ainsi, l'utilisation des bactéries est restreinte à la production d'IgG et de fragments non glycosylés. Plusieurs solutions existent pour la formation des ponts disulfures dans les bactéries. Ceux-ci peuvent être formés dans le cytoplasme ou dans le périplasma en utilisant des souches modifiées. En revanche, leur formation intracellulaire constitue un inconvénient pour l'extraction et la purification des anticorps qui nécessite de lyser les bactéries, libérant une quantité importante de débris cellulaires pouvant conduire à une pureté initiale moindre.

#### **3.3.1. Expression d'anticorps sous forme soluble dans le périplasma**

Le trajet auquel sont soumises les protéines, du cytoplasme, où celles-ci sont synthétisées, jusqu'au périplasma, permettant la formation des ponts disulfures, n'est pas idéal pour la plupart des protéines en raison du caractère réducteur et agressif du cytoplasme bactérien, qui peut entraîner la formation de corps d'inclusion (insolubles). Cet environnement réducteur est maintenu par la présence de diverses enzymes telles que la réductase à thiorédoxine (*trxB*) et la réductase à glutathion (*gor*) chez *E. coli*. Pour éviter la formation des corps d'inclusion, de nombreuses souches modifiées, présentant une déficience en réductases et/ou coexprimant des oxydases ou isomérases, ont été développées [80]. Afin de rendre leur cytoplasme oxydant, les souches FA113 et Rosetta-gami 2 de *E. coli* ont par exemple été modifiées, les rendant déficientes pour les réductases *trxB* et *gor*. Parfois, la présence de protéases peut également faire diminuer les rendements de production. Ainsi, les bactéries *E. coli* BL21 (DE3) ont été rendues déficientes en la protéase Lon, responsable de la dégradation des protéines exogènes. La co-expression de protéines chaperonnes, aidant au bon repliement des protéines et à la formation des ponts disulfures, est une autre solution souvent utilisée pour la production de protéines à ponts disulfure dans les bactéries. Ainsi, la souche initiale K12 de *E. coli* a été modifiée pour créer la souche SHuffle, capable de coexprimer l'isomérase à pont disulfures DsbC dans le cytoplasme, en plus d'être déficiente pour les gènes *trxB* et *gor*.

Si le cytoplasme des bactéries est rendu suffisamment oxydant, les protéines sous forme solubles, correctement repliées, peuvent emprunter la voie de translocation de l'arginine jumelle TAT (*TAT pathway* pour *twin-arginine translocation pathway*) afin de se rendre dans le périplasma, dont le caractère plus oxydant que le cytoplasme permet la formation des ponts disulfures dans les bactéries (figure 1.3.6). Pour améliorer l'adressage au périplasma, une technique consiste en l'utilisation de peptides signaux. Plusieurs types de peptides signaux existent et font intervenir deux mécanismes de translocation distincts dans le périplasma de *E. coli*. Les peptides signaux PelB, OmpA, PhoA et pIII permettent d'utiliser la voie SEC (*secretary pathway*), tandis que les peptides DsbA, TorT et TolB permettent d'emprunter la voie dépendante de SRP (*signal recognition particle*) [81].

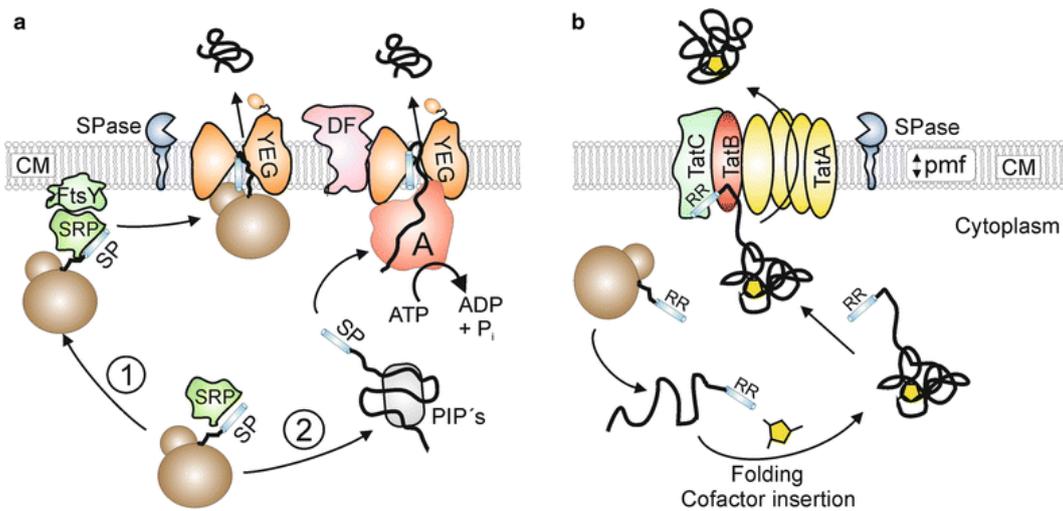


Figure 1.3.6 : Les voies possibles pour la synthèse de protéines à ponts disulfures dans le périplasma : a) SRP et SEC pour les protéines non repliées contenant des peptides signaux. b) TAT pour les protéines repliées. R. Freudl, et al. 2018, doi: 10.1186/s12934-018-0901-3.

Une fois la protéine néosynthétisée délivrée au périplasma, la formation de ponts disulfures est réalisée grâce à la présence dans ce compartiment de DsbA, une isomérase à pont disulfures maintenue sous une forme oxydée par DsbB. La famille des isomérases à pont disulfures DsbA, DsbB, DsbC, DsbD regroupe des enzymes impliquées dans la catalyse des ponts disulfures et localisées dans le périplasma de certaines souches bactériennes. La DsbA et la DsbC sont utilisées pour isomériser les ponts disulfures, tandis que la DsbB et la DsbD ont pour rôle de maintenir l'état oxydé de la DsbA et la DsbC. D'autres enzymes périplasmiques sont présentes et travaillent en coordination avec ces deux premières catégories : la FkpA et la SurA, qui sont naturellement chargées du maintien structural de la membrane externe et du repliement des protéines extra-membranaires.

L'utilisation d'une souche *E. coli* non modifiée peut également être envisagée si le plasmide codant la protéine d'intérêt contient des gènes pour l'expression d'enzymes spécifiques ou de chaperonnes. Cette méthode a permis l'obtention d'un rendement de production d'un fragment Fab atteignant 18 mg/L de culture, grâce à l'utilisation du plasmide pTUM4 qui code pour l'expression de quatre oxydoréductases et chaperonnes périplasmiques [80].

### 3.3.2. Expression d'anticorps et fragments dans le cytoplasme

L'expression dans le cytoplasme des IgG ou de ses fragments est possible selon deux méthodes : la production de protéines sous formes solubles, ou la production en tant que corps d'inclusion, insolubles.

#### a Expression de protéines sous forme soluble dans le cytoplasme

L'expression de protéine à ponts disulfures sous forme soluble dans le cytoplasme est largement dépendante du type de souche bactérienne utilisée. En effet, le cytoplasme doit présenter un caractère oxydant pour que cela soit possible, nécessitant ainsi l'utilisation de souches modifiées déficientes en protéases et réductases, et éventuellement des techniques de coexpression de chaperonnes en complément. Le système CyDisCo (*cytoplasmic disulfide bond formation*) permet l'expression de fragments Fab solubles dans le cytoplasme en utilisant la coexpression de l'oxydase à sulfhydriles Ery1P ainsi qu'une isomérase à pont disulfures (DsbC ou PDI) [82]. Le système CyDisCo a permis l'obtention de rendements d'expression de 42 mg/L de culture pour des fragments Fab et de 240 mg/L pour des fragments scFv.

L'ajout de partenaires de fusion comme la protéine de liaison au maltose (MBP : *maltose-binding protein*) ou le peptide modificateur lié à l'ubiquitine (SUMO : small ubiquitin-related modifier) peut contribuer à augmenter la solubilité de la protéine produite dans le cytoplasme et favoriser ainsi l'action des oxydases et isomérases assurant la formation des ponts disulfures [83]. Ce système a cependant conduit à des rendements plus faibles (entre 10 et 20 mg/L de fragments Fab) en comparaison au système CyDisCo décrit précédemment.

Certaines protéines, telles que l'enzyme DnaK (HSP70) dépendante de l'ATP, ainsi que les co-chaperonnes Dnaj et GrpE, peuvent également être coexprimées avec la protéine d'intérêt[84]. Celles-ci agissent de manière coordonnée pour améliorer le repliement correct de la protéine cible dans des conditions agressives.

#### b Expression des protéines en corps d'inclusion insolubles

Une technique d'expression souvent utilisée pour la production de fragments Fab en bactérie consiste en la production de la protéine d'intérêt sous forme de corps d'inclusion insolubles, suivie d'une étape de resolubilisation de ceux-ci pour aboutir à la protéine sous forme soluble. Les souches *E. coli* W3110 et BL21 (DE3) sont très utilisées pour la production de fragments Fab selon cette technique.

Cette méthode offre plusieurs avantages, tels que la pureté de la protéine d'intérêt dans la fraction contenant les agrégats, et le fait que, sous cette forme, les protéines sont moins sensibles à l'activité des protéases. Cependant, cette méthode implique de devoir extraire la protéine sous forme de corps d'inclusion, la solubiliser puis assurer son repliement avec la formation correcte de ses ponts disulfures. La solubilisation est réalisée en utilisant des dénaturants tels que l'urée et le chlorhydrate de guanidine, ou encore des détergents, des solvants organiques et des agents chaotropes [85]. Le repliement et la formation des ponts disulfures sont essentiels pour les IgG et leurs fragments. Pour cette étape, des concentrations optimisées de couples redox peuvent être utilisées comme le glutathion réduit/oxydé (GSH/GSSG). L'ajout d'additifs, tels que la L-arginine ou des faibles concentrations d'urée,

peut également améliorer le processus en évitant le mauvais repliement de la protéine et en empêchant la formation d'agrégats [86].

Cette méthode de production est la seule utilisée à ce jour pour la production de fragments Fab de mAbs thérapeutiques (pour le Ranibizumab (Lucentis<sup>®</sup>), et le Certolizumab pegol (Cimzia<sup>®</sup>). Avec un système optimisé (souche *E. coli*, extraction, repliement, solubilisation, formation des ponts disulfures et purification), le Ranibizumab est obtenu avec des rendements d'expression de 400 mg/L en tant que corps d'inclusion avec un rendement de renaturation de 20%, soit un rendement de production de protéine active de 80 mg/L [87]. Bien que cette méthode soit complexe et nécessite beaucoup d'étapes post-production, il s'agit de la plus efficace pour l'obtention d'anticorps ou fragments en bactéries.

Ainsi, parmi les différents systèmes d'expression disponibles pour la production des anticorps, le choix d'un système adapté est essentiel pour garantir l'activité des biothérapeutiques. Des caractéristiques physicochimiques déterminantes, telles que les glycosylations, sont spécifiques à chaque système d'expression. En raison de l'impact direct de ces propriétés physicochimiques sur l'activité des anticorps, l'utilisation de analytiques permet de contrôler la qualité des anticorps produits.

## 4. Analyses des caractéristiques physico-chimiques et structurales des mAbs thérapeutiques

La mise sur le marché d'anticorps monoclonaux est soumise à réglementation par les agences de sécurité du médicament. La FDA (*Food and Drug agency*) en Amérique du Nord, l'EMA (European Medicinal Agency) pour l'Europe, et les autorités réglementaires du Japon ont fondé, en 1990, un conseil international pour l'harmonisation des exigences techniques pour l'enregistrement des produits pharmaceutiques à usage humain, ou ICH (*international council for harmonisation*). Les critères d'acceptation des mAbs thérapeutiques sont détaillés précisément par l'ICH, qui définit les caractéristiques des mAbs thérapeutiques à surveiller. L'efficacité, la sûreté, la stabilité et l'activité des mAbs thérapeutiques produits doivent être acceptables pour un nouveau biothérapeutique, et répondre à des normes d'acceptabilité en cas de production d'un nouveau lot ou d'un biosimilaire<sup>1</sup> ou lors de changements dans les méthodes de production, purification, stockage ou injection [88]. Ces critères d'acceptabilité sont liés aux caractéristiques qualitatives critiques du produit (CQA : *critical quality attributes*) qui sont définis par les propriétés ou caractéristiques physiques, chimiques, biologiques ou microbiologiques devant se situer dans une limite, une plage ou une distribution appropriée pour assurer la qualité du produit [89]. Il est également défini que les entreprises pharmaceutiques doivent apporter la preuve de la maîtrise de ces paramètres à travers l'utilisation de plusieurs techniques analytiques orthogonales de façon à renforcer le contrôle de leur produit. Nombreuses de ces CQA, telles que la pureté du produit, les excipients ou les intermédiaires de production (particules issues des cellules ou des méthodes de purifications) ne concernent pas directement les caractéristiques de la protéine thérapeutique. Cette partie se concentre sur les CQA concernant les caractéristiques physico-chimiques et structurales des anticorps, ainsi que leur analyse.

Les caractéristiques physico-chimiques et structurales des mAbs thérapeutiques peuvent être impactées à différentes étapes de la production et du stockage du produit. Ces hétérogénéités peuvent survenir dès les étapes de production de la protéine par les cellules, avec les glycosylations qui ont une nature hétérogène, ou encore les différentes modifications post-traductionnelles (PTM), mais peut également apparaître au cours des étapes de purification, stockage ou même d'injection du produit. Lors de ces étapes, certaines conditions (température, pH, solvant, excipients, durée de stockage) peuvent induire diverses modifications chimiques différentes telles que la désamidation, l'oxydation, la formation de pyroglutamate en N-terminal, la suppression de la lysine ou de l'arginine en C-terminal, l'isomérisation ou encore la glycation, qui constituent les modifications chimiques principalement surveillées pour les anticorps thérapeutiques [90]. Ces modifications chimiques ont un impact sur la charge, la masse et la structure des anticorps, ce qui peut modifier leur activité, leur efficacité ou leur sûreté.

---

<sup>1</sup> Biosimilaire : Médicament biologique cliniquement équivalent et comparable à un médicament biologique de référence qui a déjà une autorisation de mise sur le marché et dont le brevet est tombé dans le domaine public. (définition EMA)

## 4.1. Hétérogénéités physico-chimiques et structurales des mAbs thérapeutiques

L'hétérogénéité des caractéristiques physico-chimiques et structurales des mAbs thérapeutiques est due à la formation de modifications chimiques sur leur séquence d'acides aminés. Les modifications post-traductionnelles (PTM) désignent, en biologie, les modifications chimiques réalisées sur la protéine après l'étape de traduction au sein des cellules. Dans le contexte pharmaceutique, des modifications chimiques, parfois identiques aux PTM, peuvent apparaître sur les protéines lors des étapes de purification ou stockage. Dans cette partie, les glycosylations, particulièrement importantes pour les anticorps, sont traitées distinctement des autres modifications chimiques (PTM ou non) pouvant survenir sur les anticorps.

### 4.1.1. Hétérogénéité des glycosylations

Les glycosylations ont déjà été abordées en raison de leur impact sur l'activité des anticorps, notamment à travers l'activation des fonctions effectrices ou d'une réponse immunogène. Il faut retenir que l'hétérogénéité des glycosylations d'un anticorps, et notamment celles positionnées sur les asparagines 297 du fragment Fc, provoquent des effets ayant un impact sur la sécurité (caractère immunogène de certains profils de glycosylation), l'efficacité (activation des mécanismes du système immunitaire comme l'ADCC, ADCP ou la CDC), et la clairance du thérapeutique (qui peut être trop rapide en cas d'absence de glycosylation).

Etant donné l'impact et le caractère hétérogène de cette modification, il est important de fournir des analyses détaillées sur leur composition. Les profils de glycanes principalement surveillés en raison de leur influence sur la stabilité, l'efficacité et l'activité des mAbs sont la présence d'acide sialique, la quantité de mannose, et l'absence de fucose sur la glycosylation de la N-297.

Les glycosylations peuvent avoir un impact significatif sur la masse de l'anticorps et la rendre hétérogène. La masse moyenne minimale d'un profil de glycosylation (avec les premiers glycanes obligatoires) est de 1.4 kDa et peut atteindre 2.8 kDa pour les formes les plus complexes, ce qui représente 1.5% et 4% de la masse d'une IgG entière (figure 1.4.1).

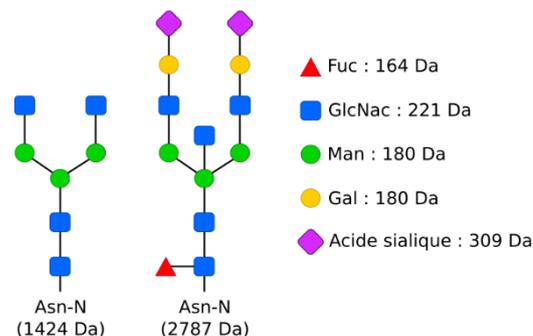


Figure 1.4.1 : Format de glycosylations minimales (gauche) et maximales (droite) avec leur masse respective et la masse des différents oligosaccharides

La méthode de choix pour la caractérisation des différentes glycosylations est la chromatographie liquide couplée à la spectrométrie de masse (LC/MS) [91]. L'anticorps est d'abord dénaturé puis les glycosylations sont clivées grâce à la PNGase F (peptide N-glycosidase F) qui est une enzyme capable de déglycosyler les protéines. Les glycosylations sont ensuite séparées de l'anticorps en utilisant une colonne de chromatographie HILIC (*hydrophilic interaction chromatography*). Ensuite, la chromatographie peut être couplée à la spectrométrie de masse en utilisant une source d'ionisation ESI (*electro-spray ionisation*). La

détermination de la position des glycosylations dans la séquence se fait également par LC/MS, mais en utilisant la trypsine pour digérer la protéine en peptides dans le but de réaliser du *peptide mapping*. Les peptides présentant une ou plusieurs glycosylations sont identifiés sur le critère de leur masse beaucoup plus élevée que leur forme non glycosylée.

#### 4.1.2. Modifications chimiques des anticorps thérapeutiques

Comme expliqué précédemment, les modifications chimiques des anticorps peuvent s'apparenter aux PTM des protéines. Toutes les modifications chimiques possibles ne seront pas détaillées ici, car plus de 400 PTM sont référencées pour les protéines. Cette partie détaille les principales modifications surveillées dans le cas des mAbs thérapeutiques.

##### a Formation de pyroglutamate en N-terminal des mAbs

La formation de pyroglutamate se fait sur les acides aminés glutamate (Glu) ou glutamine (Gln) en N-terminal et résulte en leur réarrangement cyclique (figure 1.4.2). La cyclisation des résidus Gln et Glu peut être réalisée de manière enzymatique au cours de la production par la glutaminyle cyclase, présente dans divers organismes, y compris l'humain [92]. Cependant, ce réarrangement se fait principalement de manière spontanée (non catalysée par l'intermédiaire d'enzymes) et peut survenir pendant les étapes de purification, de stockage ou après injection chez le patient en raison d'une variation de pH.

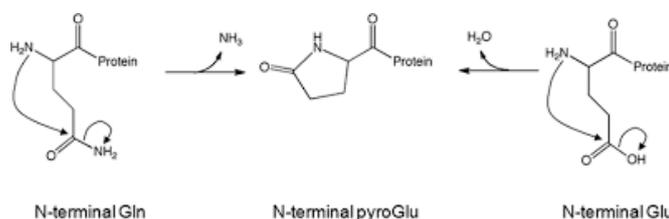


Figure 1.4.2 : Formation de pyroglutamate en N-terminal des protéines à partir de la glutamine et du glutamate

La formation de pyroglutamate en N-terminal diminue le pI des mAbs et entraîne des difficultés lors des étapes de purification du produit en raison de l'hétérogénéité de charge. Dans la séquence d'un anticorps, le domaine N-terminal est proche du site de reconnaissance des antigènes (CDR), ce qui peut altérer la fonction effectrice des fragments Fab. L'impact de cette modification sur la sécurité et l'efficacité des mAbs thérapeutiques n'est cependant pas démontrée et semble dépendant de chaque anticorps [92].

##### b Suppression de la lysine ou arginine en C-terminal des mAbs

La lysine pouvant se trouver en C-terminal des anticorps est susceptible d'être supprimée par l'activité de carboxypeptidases basiques présentes dans les cellules [93]. Cette modification n'a pas d'impact sur la reconnaissance des Fc $\gamma$ r par les parties des Fc d'IgG1 ayant perdu leur lysine en C-terminal [94]. Cette modification change la masse et la charge de l'anticorps et peut entraîner (comme pour la formation de pyroglutamate), une hétérogénéité du thérapeutique. Ces deux premières modifications n'ont pas d'impact sur la sûreté et l'efficacité

des anticorps, mais sont utilisées pour prouver le contrôle du procédé de fabrication du biothérapeutique.

La suppression de l'acide aminé en C-terminal ou la formation de pyroglutamate en N-terminal peuvent être analysées par spectrométrie de masse en mode "masse intacte" en utilisant une source d'ionisation ESI ou MALDI.

### c Désamidation des asparagines et isomérisation des aspartates

La désamidation des asparagines (Asn) (de charge neutre) conduit à la formation d'aspartate (Asp) ou d'iso-aspartate (IsoAsp) (de charge négative) en passant par sa cyclisation en un intermédiaire succinimide (de charge neutre) (figure 1.4.3). Cette modification implique un changement structural local et une modification de la charge de la protéine. Elle est sensible au pH de l'environnement et peut survenir aussi bien pendant la production par les cellules que pendant les étapes de purification et stockage [95]. Elle a lieu préférentiellement sur les Asn suivies d'une glycine et dépend de la structure locale. En effet, la formation de l'intermédiaire cyclique succinimide ne peut survenir que si les azotes de la chaîne latérale de l'Asn et du squelette de la protéine sont suffisamment proches dans l'espace, ce qui est plus souvent le cas dans les zones flexibles et donc au niveau des séquences comportant une ou plusieurs glycines. La dégradation des asparagines est une modification courante dans les anticorps et certains sites sont connus pour être plus propices à cette modification, tels que l'Asn 54 suivie de la glycine 55 dans la chaîne lourde [96]. Cette asparagine est située à proximité d'un des CDR et sa désamidation peut affecter la reconnaissance de l'anticorps à l'antigène et, ainsi, entraîner une diminution de l'activité du biothérapeutique [97].

L'isomérisation de l'aspartate en iso-aspartate est similaire à la désamidation des asparagines car elle passe également par l'intermédiaire succinimide (figure 1.4.3) et se trouve impactée par les mêmes facteurs environnementaux. L'isomérisation des résidus Asp est favorisée pour les motifs de séquence de type 'Asp-Gly' ou 'Asp-Ser' et des études ont montré que des sites se trouvant dans la région des CDR (donc flexibles) étaient plus susceptibles de subir cette modification.

Contrairement à la désamidation, cette modification ne provoque pas de changement de poids moléculaire, sauf s'il y a formation d'un intermédiaire succinimide stable (perte de 17 Da lors de la transformation de l'Asp et l'IsoAsp en succinimide).

L'aspartate et l'iso-aspartate sont tous les deux chargés négativement. L'hétérogénéité de charge de l'anticorps peut se produire lors de la formation d'un intermédiaire succinimide stable, dont la charge est neutre. L'isomérisation entraîne une modification structurale locale

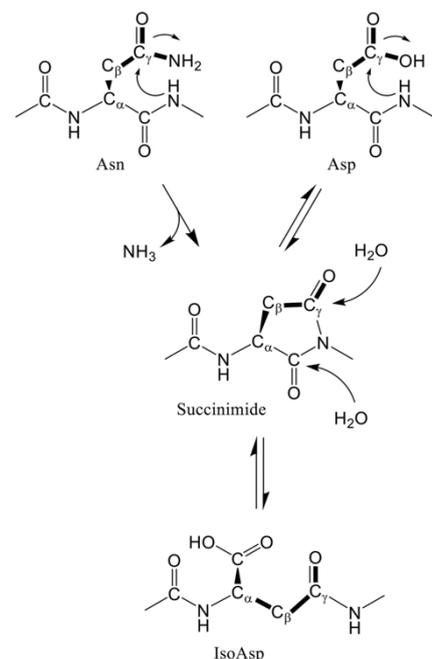


Figure 1.4.3 : Désamidation de l'asparagine, et isomérisation de l'aspartate en passant par l'intermédiaire succinimide.

qui peut provoquer le clivage de l'anticorps, ou mener à la formation d'agrégats, et donc altérer la stabilité et l'activité de l'anticorps [98].

Les produits de dégradation de la désamidation d'asparagine en Asp ou IsoAsp ainsi que de la formation de succinimide stable peuvent être séparés par chromatographie en phase inverse, échangeuse d'ions (IEC) ou par focalisation isoélectrique capillaire (IEF) [99]. La séparation du produit de l'isomérisation des Asp en Iso-Asp peut être réalisée en utilisant la chromatographie d'interaction hydrophobe (HIC) [100]. Ces méthodes de séparation utilisées en HPLC sont ensuite couplées à la spectrométrie de masse en mode peptide mapping et notamment en mode LC/MS/MS qui permet de situer précisément les PTM et de les quantifier [99].

#### d Oxydations des méthionines

L'oxydation des méthionines résulte en leur transformation en méthionine sulfoxyde, et plus rarement en méthionine sulfone lorsqu'une oxydation supplémentaire a lieu (figure 1.4.4). Cette modification entraîne la formation d'un groupement plus polaire, qui a pour effet de modifier la structure locale de la protéine. Elle est non enzymatique, et a lieu principalement pendant les étapes de purification et de stockage des biothérapeutiques car elle est provoquée par des changements de température ou la composition des différents tampons utilisés.

Cette modification peut impacter l'activité biologique des anticorps et même provoquer une réponse immunitaire du patient face au mAb oxydé. Il a été démontré que l'oxydation des méthionines 33 et 209 des

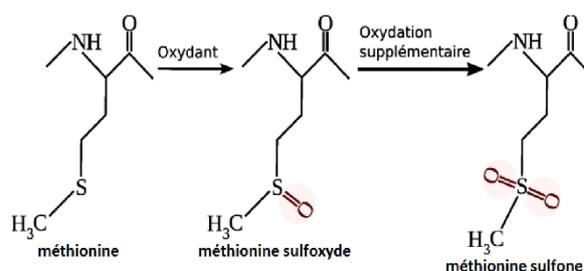


Figure 1.4.4: Les deux étapes de l'oxydation des méthionines

chaînes lourdes d'un mAb provoquait une perte de stabilité face à des variations de température, et favorisait la désamidation locale des asparagines [101]. Des études de l'oxydation des méthionines ont également été réalisées par HDX-MS (*hydrogen-deuterium exchange mass spectrometry*). Elles ont montré que l'oxydation des méthionines 257 et 433 des chaînes lourdes des anticorps (situées sur le domaine CH<sub>2</sub>) a un impact sur la reconnaissance des récepteurs néonataux de Fc (FcRn) et diminue l'interaction des fragments Fc avec les récepteurs. L'oxydation de ces méthionines a donc un impact sur la durée de vie de l'anticorps, car sa capacité à être recyclé dans l'organisme est diminuée (cf. partie 1.1.4.2).

Le contrôle qualité de l'oxydation des mAbs est réalisé par spectrométrie de masse en mode peptide mapping pour l'identification des méthionines oxydées, et la méthode est couplée à de l'HPLC échangeuse d'ions pour la séparation des peptides contenant les méthionines oxydées et non-oxydées.

## e Glycation

La glycation (à ne pas confondre avec la glycosylation) est une modification non enzymatique (à l'inverse des glycosylations) qui survient lorsqu'un sucre réducteur peut être issu du glucose, galactose ou fructose réagit avec l'amine primaire en N-terminal, ou l'amine de la chaîne latérale d'une lysine (préférentiellement avec les lysines de surface) [102]. Dans un premier temps, le groupement amine réagit de manière réversible avec le groupe aldéhyde d'un sucre réducteur et forme un intermédiaire de Schiff instable. Puis, cet intermédiaire peut subir des réarrangements (réarrangements d'Amadori) pour aboutir à la formation de kétoamine stable (figure 1.4.5). L'oxydation d'une glycation d'Amadori (réaction de Hodge) peut conduire à la formation de carboxyméthyllysine ou pentosidine qui constituent des produits de glycation avancés (AGEs: *advanced glycation end products*) (figure 1.4.5)[103].

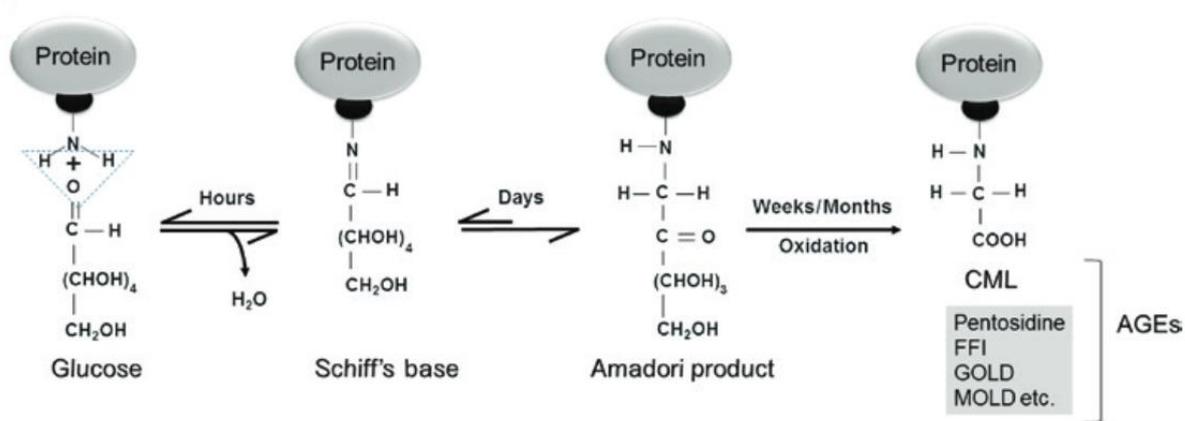


Figure 1.4.5 : Schéma représentant la glycation des protéines avec la formation d'un intermédiaire de Schiff, un produit d'Amadori et la formation de AGE. A. K. Pasupulati, et al. 2016, doi: 10.1515/bmc-2016-0021.

Les sucres réduits peuvent provenir de la formulation du milieu de culture utilisé, et notamment du glucose utilisé comme source énergétique pour les cellules et dont la réduction peut être accélérée sous l'effet de la température, de la force ionique ou du pH [104]. Elle peut aussi survenir lors du stockage, lorsque des sucres comme le sucrose disaccharide sont utilisés en formulation [105].

La formation des AGEs lors de la bio-fabrication des mAbs thérapeutiques est très surveillée car ceux-ci sont très réactifs et peuvent être à l'origine de réactions immunitaires, altérer certaines fonctions cellulaires, et sont toxiques pour certaines cellules telles que les cellules endothéliales, les neurones, ou encore les leucocytes [106].

L'analyse de la glycation des anticorps est réalisée en utilisant la chromatographie d'affinité au borate (BAC : *boronate affinity chromatography*) qui constitue une technique d'isolation et d'enrichissement de composés cis-diols tels que les sucres [107]. La focalisation isoélectrique capillaire (IEF) et la chromatographie échangeuse d'ion (IEC) peuvent également être utilisées en raison de la perte de charge positive des amines primaires transformées en produit d'Amadori, ce qui a pour effet de modifier la charge de l'anticorps [108]. Ces méthodes peuvent être couplées à la spectrométrie de masse en mode masse intacte ou peptide mapping pour déterminer le pourcentage de glycation ainsi que les sites modifiés. La présence de l'intermédiaire kétoamine peut aussi être contrôlée par un test colorimétrique en utilisant du

nitrobleu tétrazolium (NBT : *nitroblue tetrazolium*), dont la couleur se modifie lors de son oxydation par la kétoamine[109].

## f Réduction des ponts disulfures

Les ponts disulfures peuvent faire l'objet de deux types de modifications chimiques : ils peuvent s'isomériser en donnant une conformation non-native à la protéine, ou alors ne pas se former du tout. La mauvaise isomérisation peut être favorisée à pH basique ou en présence d'autres cystéines réduites. Les ponts disulfures ayant déjà été évoqués en détail dans la partie 3, ils ne seront pas plus détaillés ici.

La méthode utilisée afin de déterminer une mauvaise isomérisation des ponts disulfures consiste à combiner l'utilisation de l'ETD (*electron transfer dissociation*) et de la CID (*collision induced dissociation*) comme source d'ionisation en LC/MS. Wang *et al.* ont utilisé cette technique pour déterminer la conformation de ponts disulfures mal isomérisés pour trois mAbs différents [110].

### 4.1.3. Modifications de la structure hautement ordonnée des mAbs thérapeutiques

Que ce soit à une échelle globale ou locale, toutes les modifications chimiques peuvent avoir un impact sur la structure 3D des anticorps. De ce fait, l'analyse des caractéristiques physico-chimiques des anticorps est corrélée aux analyses structurales. A une échelle globale (du polypeptide à l'échelle macroscopique), les modifications chimiques peuvent provoquer la formation d'agrégats ou la fragmentation des anticorps. Les modifications structurales à cette échelle sont très importantes mais faciles à analyser par chromatographie d'exclusion de taille (SEC) ou par électrophorèse capillaire en dodécyl-sulfate de sodium (CE-SDS). Ces méthodes peuvent être couplées à la spectrométrie de masse pour la détermination du type de fragments ou d'agrégats. A une échelle plus locale (*i.e.* acide aminé), les modifications structurales sont plus compliquées à analyser, et il est difficile de donner une explication structurale à la perte d'activité ou d'efficacité due à une modification chimique [111].

L'activité liée à la structure des petites molécules thérapeutiques peut être définie de manière simple en utilisant les règles de stéréochimie. A l'inverse, l'activité des anticorps thérapeutiques nécessite une formation structurale complexe appelée structure hautement ordonnée (HOS : *higher order structure*). Cette HOS est le résultat du repliement de la protéine, inclut l'ensemble des niveaux structuraux des protéines (de la structure primaire à la structure quaternaire) et est impactée par toutes les modifications physico-chimiques pouvant avoir lieu. Ce terme, décrivant la structure des protéines, est principalement utilisé dans les domaines de l'industrie pharmaceutique et de la sécurité des biothérapeutiques.

La HOS des anticorps est directement impactée par les glycosylations ou les modifications chimiques détaillées précédemment, qui peuvent entraîner des variations d'activité, de stabilité, de sécurité ou d'efficacité des biothérapeutiques. Il est donc important de comprendre comment la HOS influence les caractéristiques qualitatives critiques des mAbs. En plus du contrôle de la qualité, le suivi de la HOS lors du développement du biothérapeutique permet une compréhension approfondie de l'impact des conditions de production sur la qualité

du produit. Les attentes des agences de sécurité du médicament sont très élevées sur la caractérisation de la HOS des anticorps monoclonaux à usages thérapeutiques. C'est pourquoi la caractérisation de la HOS doit être réalisée par différentes méthodes analytiques orthogonales et complémentaires afin de pouvoir mieux développer les procédés de bio-fabrication des mAbs thérapeutiques.

## 4.2. Caractérisation de la HOS des anticorps

### 4.2.1. Détermination structurale des anticorps par cryo-microscopie électronique et cristallographie aux rayons X.

La première structure à haute résolution d'un fragment Fab d'anticorps a été obtenue dans les années 1970 par cristallographie aux rayons-X [112]. Depuis, la cristallographie des protéines et la cryo-microscopie électronique (cryo-EM) ont fourni plus de 9 000 structures PDB de fragments Fab ou Fc des anticorps. Peu de structures d'anticorps entiers ont été déterminées en raison de la flexibilité des fragments d'anticorps les uns par rapport aux autres qui complexifie la formation de cristaux. On en trouve cependant quelques-unes, telles que la structure de l'IgG1 B12 développée pour la lutte contre le VIH [113] (figure 1.4.6). La plupart des structures cristallographiques disponibles concernent des études de la structure des fragments Fab et Fc en complexe avec un antigène ou un partenaire biologique. Bien que la cristallographie aux rayons-X et la cryo-EM puissent fournir des données structurales de haute résolution sur la HOS des anticorps, ces méthodes capturent des intermédiaires correspondant à des états figés, et ne suffisent pas pour être représentative de la dynamique moléculaire des anticorps en solution. De plus, les conditions favorisant l'obtention des cristaux de protéines sont loin de celles de la formulation pour la biosynthèse des anticorps.

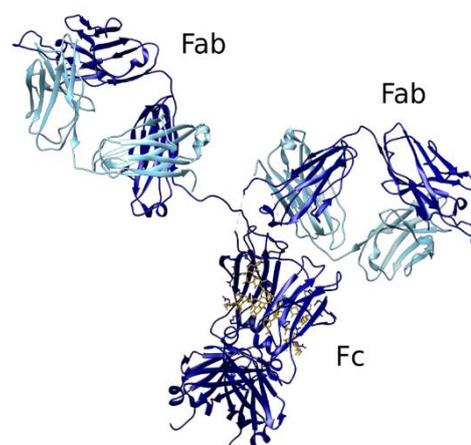


Figure 1.4.6 : Structure PDB 1HZH de l'IgG1 B12. Les chaînes lourdes sont en bleu clair et les chaînes légères en bleu foncé. PDB DOI: 10.2210/pdb1HZH/pdb

Malgré cela, la détermination de la structure des anticorps est essentielle à la meilleure compréhension de leur fonctionnement, notamment à travers la détermination de la structure des différents fragments en complexe avec des partenaires (antigènes ou récepteurs). Les structures ainsi déterminées peuvent également servir de base pour l'utilisation complémentaire d'autres techniques structurales, telles que la spectrométrie RMN.

Cependant, l'analyse de la HOS des mAbs thérapeutiques concerne plus spécifiquement la structure adoptée par les anticorps lorsqu'ils sont en solution et l'impact structural exercé par les différents paramètres de production. Ainsi, l'emploi de méthodes analytiques complémentaires est nécessaire pour évaluer la structure en solution des anticorps.

#### 4.2.2. Etude de la HOS par les méthodes de spectroscopie optique

L'étude de la HOS des anticorps peut se faire à différentes échelles. La présence de fragments ou d'agrégats implique des modifications importantes de la HOS qui peuvent être suivies par des méthodes de spectroscopie. De nombreuses méthodes spectroscopiques existent avec une basse ou moyenne résolution, telles que la fluorescence intrinsèque, le dichroïsme circulaire, la spectroscopie infra-rouge ou Raman à transformée de Fourier (FT-IR). Ces techniques analytiques fournissent des informations sur la structure secondaire et globale de l'anticorps et permettent de connaître l'état de la protéine (repliée/dépliée, proportion en feuillets  $\beta$  et hélices  $\alpha$ , dégradée/intacte). Elles sont très utilisées en industrie pour la détection de variations dans la qualité du produit lors de changements de formulation ou de développement d'un bio-similaire. Leur avantage est la rapidité de mise en œuvre et la sensibilité élevée de ces techniques, dont la réalisation des analyses nécessite de faibles quantités de protéine (quelques micro-grammes). Cependant, elles sont limitées par leur résolution, qui ne permet pas d'attribuer les modifications du signal à des changements spécifiques de la structure secondaire. De ce fait, il est difficile d'identifier le changement structural subtil à l'origine d'une perte d'activité ou d'efficacité en se dispensant de l'emploi d'une autre méthode analytique orthogonale.

#### 4.2.3. La spectrométrie de masse d'échange hydrogène/deutérium

De manière plus subtile, des PTM impactant l'activité et l'efficacité des mAbs sans modification de leur structure globale peuvent apporter des modifications structurales locales à la HOS. A ce niveau de résolution, les méthodes analytiques décrites précédemment ne suffisent plus et des méthodes comme la HDX-MS et la RMN sont utilisées pour offrir une résolution à l'échelle atomique de l'analyse de la HOS des mAbs thérapeutiques.

Le principe de la HDX-MS est basé sur la capacité des protons des groupements amides du squelette carboné des protéines à s'échanger avec les protons du solvant à des vitesses variées, et dépendantes de la structure globale et locale de la protéine en fonction de leur accessibilité. En effet, lorsqu'ils sont impliqués dans des liaisons hydrogènes, les protons du cœur de la protéine sont, par exemple, protégés de l'échange. La protéine est initialement placée dans un tampon contenant du  $D_2O$  pendant des temps d'incubation différents afin de permettre aux protéines l'incorporation de différentes quantités de deutérium, puis l'échange hydrogène/deutérium est stoppé à basse température et à pH acide. Ensuite, les échantillons sont digérés par réaction enzymatique (avec la pepsine ou la trypsine par exemple) en peptides, analysés par LC/MS (*peptide mapping*). L'analyse de ces échantillons fournit des spectres de masses spécifiques à chaque peptide et à chaque durée d'incubation dans le  $D_2O$  (figure 1.4.7). Ainsi, il est possible de positionner le site de deutération et de déterminer sa cinétique d'échange, communiquant une information sur l'accessibilité au solvant des sites de la protéine, et donc une information sur la structure de la protéine en solution. En appliquant la HDX-MS lors d'une interaction avec un partenaire tel qu'un antigène, il est possible de déterminer quels résidus des CDR sont impliqués, car ceux-ci sont protégés de l'échange lors de l'interaction avec l'épitope (figure 1.4.7).

L'avantage de la spectrométrie de masse est sa sensibilité, ainsi de faibles quantités de protéines sont requises (microgrammes). De plus, la résolution de la spectrométrie de masse est

très élevée (de l'ordre du Da en fonction de la méthode utilisée) et permet de situer avec précision des changements de masse, ce qui permet d'étudier les zones de la protéine accessibles au solvant qui peuvent être impactées par la dynamique de la protéine.

La HDX-MS est utilisée en industrie pharmaceutique pour l'étude des changements conformationnels dynamiques des mAbs en solution causés par les glycosylations, l'interaction à l'antigène, des mutations, ou des modifications de l'environnement (tampon, pH, température, force ionique). Elle peut apporter des informations sur les changements conformationnels subis par le squelette carboné des protéines en fonction des conditions de stockage et de stress [114].

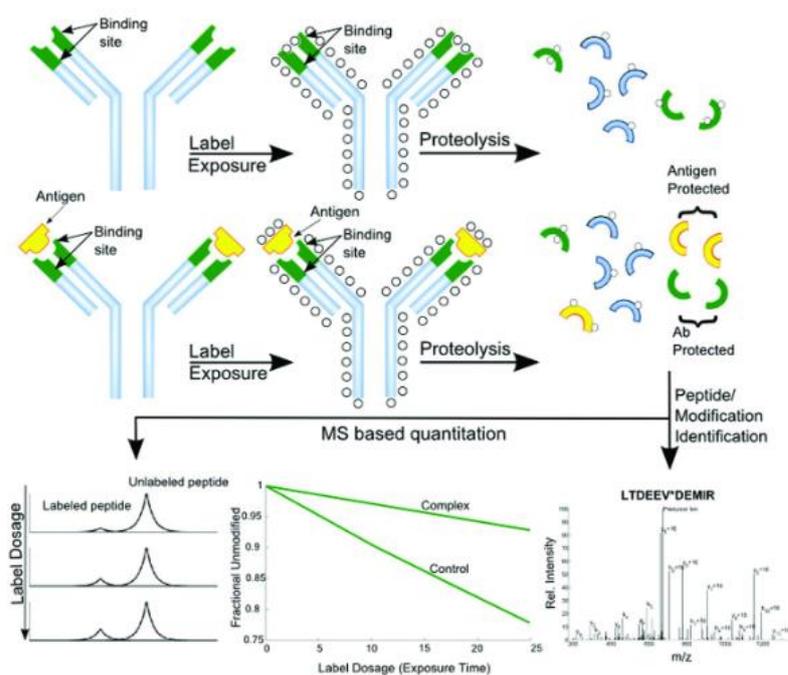


Figure 1.4.7 : Schéma représentatif du principe de fonctionnement de la HDX-MS pour l'étude d'interaction anticorps/antigène par cette méthode. P. Kaur et al. 2015 doi: 10.1021/bk-2015-1202.ch003.

Cependant, cette méthode présente des limitations, la rendant insuffisante pour la caractérisation précise de la HOS des mAbs en solution. En effet, elle nécessite la digestion de la protéine en peptides et il est impossible d'ioniser de manière satisfaisante (en spectrométrie de masse) l'ensemble des peptides issus d'une digestion enzymatique, empêchant ainsi d'observer certains acides aminés. De plus, la rapidité d'échange de certains protons avec le solvant utilisé pendant l'étape de chromatographie liquide constitue une autre limitation. En effet, ce phénomène de *back-exchange* du deutérium empêche l'analyse en HDX-MS de certains acides aminés dont le proton amide s'échange trop rapidement.

### 4.3. La résonance magnétique nucléaire des anticorps en abondance naturelle

De nombreuses techniques analytiques de biochimie et biophysique existent pour la caractérisation de la HOS des biothérapeutiques. Cependant, ces techniques manquent de résolution et ne permettent pas d'expliquer les modifications subtiles de la structure des anticorps impactant leur activité, efficacité et stabilité. La structure des anticorps peut être obtenue à une résolution atomique en utilisant la cristallographie des rayons X et la cryo-microscopie électronique mais ces méthodes n'offrent qu'une structure d'un état le plus représenté, parmi les multiples conformations que peuvent présenter les anticorps en solution.

La résonance magnétique nucléaire (RMN) des mAbs thérapeutiques en solution peut fournir des informations à l'échelle atomique sur leur structure. Les signaux observés en RMN représentent tous un atome ou une liaison spécifique (atome d'hydrogène ou liaison hydrogène-azote ou hydrogène-carbone par exemple) au sein de la protéine, et la position de ces signaux dans le spectre RMN dépend de l'environnement local de chaque atome. Selon cette définition, la RMN est la méthode permettant de combler les limitations des autres techniques analytiques l'évaluation non-invasive de la HOS des mAbs thérapeutiques à une résolution atomique et permet d'analyser la structure de leurs différentes conformations en solution.

Cette partie décrit les méthodes analytiques de la RMN en abondance naturelle (sans enrichissement isotopique) utilisées dans la caractérisation de la HOS des mAbs thérapeutiques.

#### 4.3.1. La RMN 1D des protons pour la caractérisation des mAbs thérapeutiques

En RMN, la position des raies de résonances d'un noyau est impactée par la distribution des électrons proches du noyau qui encombrant le champ magnétique. Cela signifie que le déplacement chimique des protons d'une protéine sur un spectre 1D  $^1\text{H}$  est dépendant de la structure de la protéine. Ainsi, la HOS des anticorps peut être analysée par des expériences RMN des protons en une dimension. Cependant, les anticorps sont des grosses protéines comportant un grand nombre de protons, fournissant un spectre  $^1\text{H}$  de plus de 9 900 raies se superposant (figure 1.4.8). De surcroît, la taille et la lente réorientation des anticorps sur eux-mêmes impacte le phénomène de relaxation transversale (cf. partie 1.4.4). Plus la masse d'une protéine est élevée, plus le temps de corrélation rotationnel ( $\tau_c$ ) est élevé, et plus la relaxation transversale se fait rapidement, diminuant ainsi le temps disponible pour l'acquisition du signal. Cela se traduit par une diminution de l'intensité du signal et un élargissement des raies spectrales. La superposition des résonances, l'intensité et la résolution réduites pour les anticorps constituent les problèmes majeurs pour l'utilisation de la RMN 1D  $^1\text{H}$  pour l'étude de leur HOS. Cependant, plusieurs astuces permettent d'utiliser la RMN 1D des protons et il est tout de même

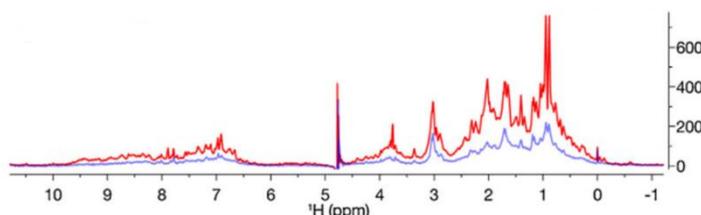


Figure 1.4.8 : Superposition des spectres RMN 1D  $^1\text{H}$  (600MHz) de l'inliximab natif (en bleu) et clivé avec l'enzyme IdeS (en rouge). IdeS est une enzyme qui clive les anticorps en  $F(ab')_2$ . K. Chen et al. 2016, doi: 10.1016/j.jpba.2016.06.007.

possible, malgré ces inconvénients, d'obtenir des informations utiles pour la HOS des mAbs grâce à cette technique.

La formulation des milieux et tampons utilisés pour la synthèse, la purification et le stockage des anticorps contiennent des additifs et excipients dont il est intéressant d'étudier l'impact sur la HOS des mAbs. En RMN, la taille de ces petites molécules par rapport aux anticorps est un problème car l'intensité des résonances de leurs protons est très nettement supérieure à celle de l'anticorps. Lors de l'étude de l'impact de ces composés sur la HOS des anticorps par RMN, cela se traduit par un recouvrement des raies spectrales de l'anticorps par les raies spectrales plus intenses des petits composés. La méthode PROFILE est une technique basée sur l'utilisation de séquences d'impulsion spécifiques et d'un algorithme de traitement du signal permettant de retrouver le spectre 1D  $^1\text{H}$  d'une protéine en solution dans un milieu complexe. La méthode permet de séparer les signaux des excipients du solvant grâce à leur coefficient de diffusion translationnel, différent de ceux des grosses molécules, telles que les anticorps. Ainsi le spectre reflète l'état de la protéine impactée par la présence de nombreux petits composés, en faisant abstraction de leur signal sur le spectre [115]. En couplant cette méthode avec le phénomène de diffusion des spins entre protons d'une protéine, il est alors possible de différencier les régions plus structurées des autres, permettant ainsi de réaliser des analyses plus sélectives de la HOS des mAbs par RMN [116].

La RMN 1D  $^1\text{H}$  peut également être utilisée pour l'analyse des interactions non spécifiques et réversibles des mAbs entre eux, qui peut survenir en amont de la formation d'agrégats. L'intensité du signal d'un spectre  $^1\text{H}$  est proportionnelle à la taille de la protéine analysée et donc aux oligomères qu'elles peuvent former. Ce type d'étude a été réalisé avec l'Infliximab qui est susceptible de former des oligomères réversibles et qui fournit des spectres ayant un *ratio* signal/bruit plus faible que les spectres du rituximab, qui ne forme pas d'oligomères [117].

La RMN 1D  $^1\text{H}$  des anticorps en abondance naturelle reste cependant très limitée. Elle est limitée à une analyse globale de la structure, comme les méthodes de spectroscopie optique (cf. partie 1.4.2.2). L'utilisation de méthode statistique (comme la PCA : *principale component analysis*) pour la compréhension des modifications observées sur les spectres RMN est souvent nécessaire, mais les données sont, en général, difficiles à interpréter car les spectres présentent un recouvrement de près de 10 000 signaux  $^1\text{H}$ , et les modifications subtiles de la HOS ne peuvent pas être étudiées par cette méthode.

#### 4.3.2. La RMN multidimensionnelle en détection $^1\text{H}$ : empreinte digitale de la structure

La RMN multidimensionnelle en détection  $^1\text{H}$  consiste à corrélérer deux spin  $\frac{1}{2}$  dans deux dimensions grâce au couplage scalaire  $^1\text{J}_{\text{H}-^{13}\text{C}}$  ou  $^1\text{J}_{\text{H}-^{15}\text{N}}$ . L'information ainsi obtenue comporte la fréquence  $^1\text{H}$  d'un proton et celle d'un hétéroatome ( $^{13}\text{C}$  ou  $^{15}\text{N}$ ) auquel celui-ci est covalentement lié. Cette méthode est moins sensible que la RMN 1D, mais beaucoup plus spécifique car les signaux observés ne concernent que certaines liaisons atomiques ( $^1\text{H}-^{13}\text{C}$  ou  $^1\text{H}-^{15}\text{N}$ ). Les spectres 2D de protéines permettent ainsi d'obtenir une véritable empreinte digitale de leur structure. Cependant, la taille des anticorps reste un défi pour la RMN 2D car

leur mobilité moléculaire est lente et impacte la relaxation transversale, ce qui a pour effet de diminuer la sensibilité.

L'abondance naturelle relative des noyaux  $^{13}\text{C}$  et  $^{15}\text{N}$  est respectivement de 1.11% et 0.37%. Cela signifie que, pour un spectre d'une protéine non enrichie en isotopes, seuls 1.11% des carbones, et 0.37% des azotes des protéines en solution sont visibles. En abondance naturelle, le  $^{12}\text{C}$  ne fournit pas de signal RMN et le  $^{14}\text{N}$  est un noyau quadripolaire dont la relaxation très efficace implique des largeurs de raies importantes. Leur remplacement par leur isotope stable ( $^{13}\text{C}$  et  $^{15}\text{N}$  respectivement), lors de la production de la protéine, permet d'augmenter le signal. L'augmentation du signal est très significative lorsque le taux d'incorporation des isotopes est élevé. Par exemple, si 100% des carbones des groupements méthyles sont remplacés par des  $^{13}\text{C}$ , le signal est augmenté d'un facteur 100. Cependant, le marquage isotopique des anticorps thérapeutiques pour le contrôle de leur HOS par RMN n'est pas envisageable. En effet, la finalité des mAbs thérapeutique impliquant l'injection chez le patient exclut un marquage isotopique de l'anticorps thérapeutique. Cette technique est donc utilisée pour l'étude structurale des protéines par RMN mais pas pour le suivi de la qualité des lots de production. Cette partie détaille seulement la RMN des anticorps en abondance naturelle telle qu'utilisée pour le contrôle de leur HOS. Le marquage isotopique sera détaillé ultérieurement dans les résultats (cf. partie 3.1.1).

#### a La RMN $^1\text{H}$ - $^{15}\text{N}$ des groupements amides des anticorps

La RMN 2D hétéronucléaire  $^1\text{H}$ - $^{15}\text{N}$  fournit un spectre très représentatif de la structure du squelette carboné de la protéine analysée. Son profil indique si la protéine est repliée ou non, et la réalisation des études statistiques de la modification des spectres  $^1\text{H}$ - $^{15}\text{N}$  permet de surveiller des modifications de la HOS des anticorps dues à des glycosylations ou modifications chimiques ou en conséquence de l'utilisation d'additifs dans les différents tampons [118]. J. P.

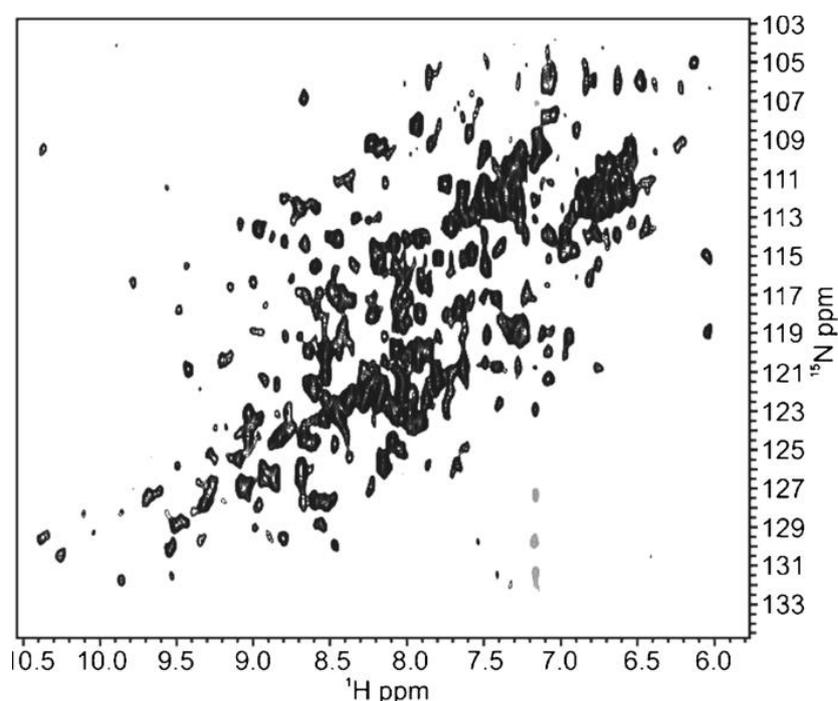


Figure 1.4.9 : SOFAST-HMQC  $^1\text{H}$ - $^{15}\text{N}$  de l'anticorps NIST en abondance naturelle. Acquisition à 900 MHz et 50°C. J. P. Marino et al., 2015, doi: 10.1021/bk-2015-1202.ch002.

Marino *et al.* ont montré la possibilité d'obtenir des spectres RMN  $^1\text{H}$ - $^{15}\text{N}$  d'anticorps en abondance naturelle de qualité (figure 1.4.9) pour un échantillon à 300  $\mu\text{M}$ , en acquérant des expériences RMN SOFAST-HMQC sur un spectromètre équipé d'une sonde cryogénique, opérant à une fréquence  $^1\text{H}$  de 900 MHz pendant 68 h et à 50°C [119]. En raison du nombre d'acides aminés composant l'anticorps, l'utilisation des spectres RMN des amides du squelette carboné est limitée. Le nombre de groupements amides différents dans un anticorps est d'environ 700 (1 400 acides aminés mais les anticorps sont des hétéro-dimères symétriques), ce qui provoque de nombreuses superpositions des pics sur ces spectres, entraînant donc une interprétation des données difficile.

Les anticorps peuvent être digérés en fragment Fab, F(ab')<sub>2</sub> ou Fc par la papaïne ou la pepsine. Cela n'altère pas la structure locale des différents domaines d'IgG et diminue la taille de la protéine analysée. Cela permet l'obtention de spectres RMN de meilleure qualité, car la superposition de pic est restreinte comparée à la protéine entière, et la taille réduite des fragments permet une meilleure sensibilité, contribuant ainsi à améliorer tous les types d'analyses par RMN.

Concernant les paramètres intrinsèques au spectromètre lors de l'acquisition, plusieurs techniques, telles que l'augmentation de la température d'acquisition et l'utilisation d'une sonde cryogénique (valables pour tous les types d'acquisition), ou encore l'utilisation de séquences d'impulsion optimisées (SOFAST, TROSY), peuvent être utilisées afin de s'affranchir d'une partie de l'effet de relaxation.

## b La RMN $^1\text{H}$ - $^{13}\text{C}$ des groupements méthyles des mAbs thérapeutiques

La RMN des groupements méthyles des protéines a représenté une avancée majeure pour l'étude de gros complexes protéiques comme les anticorps. Parmi les vingt acides aminés, six possèdent des groupements méthyles (alanine, isoleucine, leucine, méthionine, thréonine, valine) (figure 1.4.10).

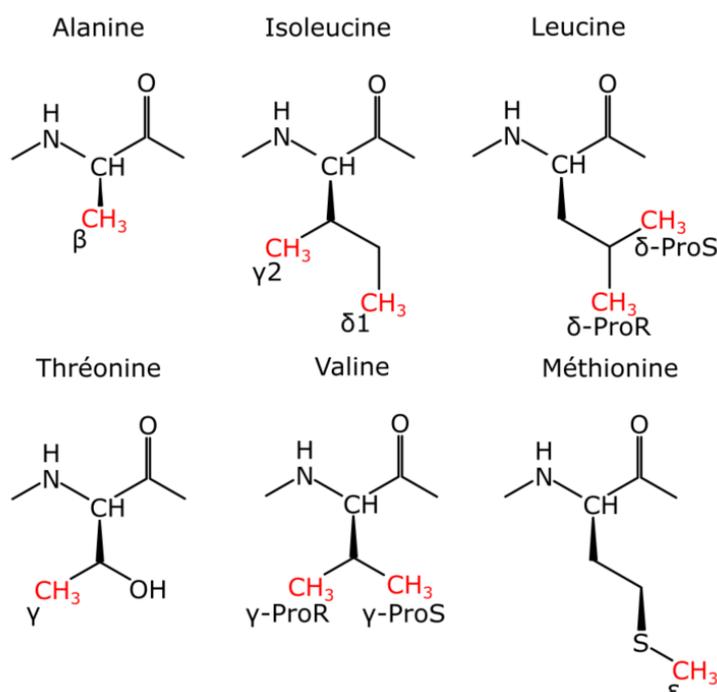


Figure 1.4.10 : Les six acides aminés contenant des groupements méthyles (en rouge)

Ils représentent environ 40% de la séquence des protéines, résultant en un spectre RMN  $^1\text{H}^{13}\text{C}$  moins chargé comparé à un spectre  $^1\text{H}^{15}\text{N}$  (environ 340 groupements méthyles contre 700 groupements amides ; les isoleucines, leucines et valines en contenant deux chacune). Lorsque l'on représente la position des groupements méthyles sur la structure d'un anticorps (figure 1.4.11), on remarque qu'ils sont très bien répartis et couvrent l'ensemble de la structure de façon homogène. Ainsi, un spectre RMN de ces groupements méthyles est très représentatif de la structure d'un anticorps et ceux-ci constituent de très bonnes sondes, car quel que soit le lieu où survient une modification chimique, un ou plusieurs groupements méthyles se situeront à proximité de celle-ci, et son/leur déplacement chimique sur le spectre sera impacté. Les groupements méthyles sont des "sondes" permettant de suivre la HOS des anticorps thérapeutiques. Les déplacements chimiques des carbones sont impactés par la conformation locale de la chaîne latérale des acides aminés, tandis que ceux des protons sont très sensibles à l'environnement de ceux-ci sans nécessairement être sensibles aux modifications locales.

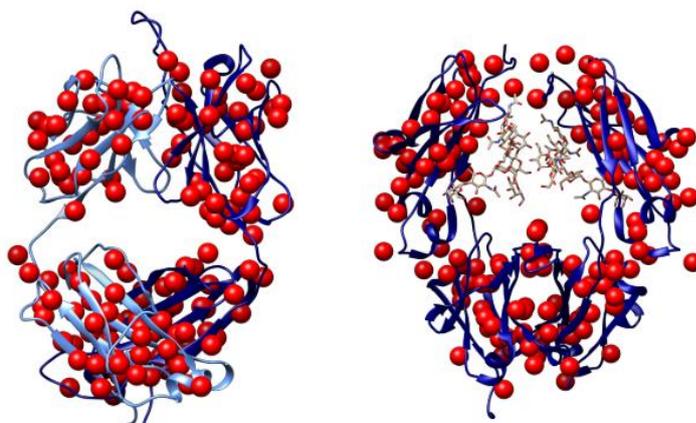


Figure 1.4.11 : Structure des fragments Fab (gauche) et Fc (droite) avec les groupements méthyles représentés en rouge. Les glycosylations sont représentées en beige sur le fragment Fc. Le squelette carboné de la chaîne légère est en bleu clair et celui de la chaîne lourde est en bleu foncé. PDB : 1HZH. DOI: 10.2210/pdb1HZH/pdb

De plus, la zone de résonance des groupements méthyles sur un spectre RMN n'est pas surchargée, et les différents groupements méthyles (des différents acides aminés concernés) résonnent à des déplacements chimiques spécifiques, limitant ainsi le nombre de pics superposés et offrant une meilleure interprétation des données (figure 1.4.12), même si de nombreuses superpositions restent cependant possibles.

Le pic RMN  $^1\text{H}^{13}\text{C}$  d'un groupement méthyle correspond à l'intensité de 3 protons aux propriétés chimiques et physiques identiques, donc d'intensité trois fois supérieure à un pic de proton seul comme pour un groupement amide N-H du squelette carboné des protéines. De plus, le groupement méthyle a un mouvement de rotation rapide et, lorsqu'il est situé en extrémité de la chaîne latérale des acides aminés, fait partie d'une zone flexible. Cela procure à ces

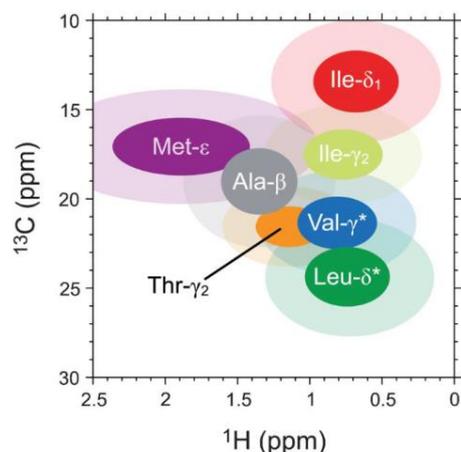


Figure 1.4.12 : Zones des déplacements chimiques théoriques des différents types de groupements méthyles. Plevin et Boisbouvier 2012

groupements de meilleures propriétés pour la relaxation transversale et longitudinale que pour les autres protons de la protéine et procure un gain d'un facteur deux en moyenne sur l'intensité du signal par rapport à un signal  $^1\text{H}$ - $^{15}\text{N}$ . Un autre avantage de la RMN des groupements méthyles d'une protéine en abondance naturelle, est l'abondance naturelle relative du carbone 13 qui est de 1,1% par rapport à l'azote 15 qui est de 0,37%. En comparaison des  $^{15}\text{N}$  et en abondance naturelle, il y a donc statistiquement trois fois plus de  $^{13}\text{C}$  observables. Grâce à ces avantages physiques, la RMN  $^1\text{H}$ - $^{13}\text{C}$  des groupements méthyles offre des pics 15 à 20 fois plus intenses que ceux de la RMN  $^1\text{H}$ - $^{15}\text{N}$  des groupements amides d'une protéine en abondance naturelle.

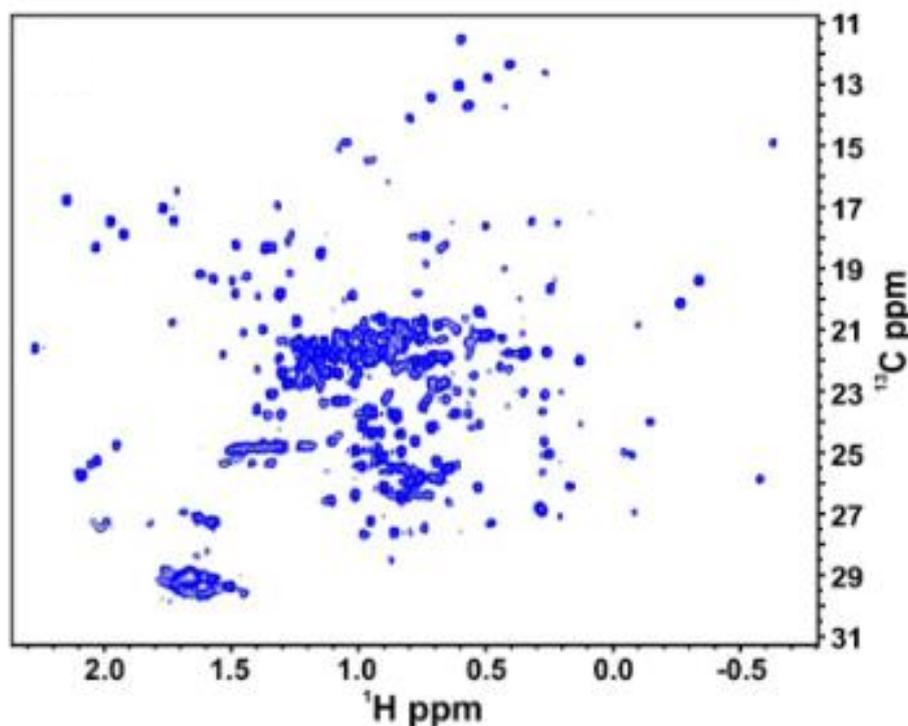


Figure 1.4.13 : Spectre RMN-HSQC 2D  $^1\text{H}$ - $^{13}\text{C}$  du mAb NIST obtenu avec un spectromètre opérant à une fréquence  $^1\text{H}$  de 900 MHz équipé d'une sonde cryogénique. L'acquisition a durée 12 h à 50°C. L. W. Arbogast et al. *Anal. Chem.*, vol. 87, no 7, p. 3556-3561, avr. 2015, doi: 10.1021/ac504804m.

Ces avantages font que la RMN  $^1\text{H}$ - $^{13}\text{C}$  des groupements méthyles des anticorps est très utilisée pour la caractérisation de leur HOS. Pour un anticorps en abondance naturelle, il est possible d'obtenir un spectre  $^1\text{H}$ - $^{13}\text{C}$  des groupements méthyles de bonne qualité en utilisant un spectromètre RMN opérant à très haut champ équipé d'une sonde cryogénique et une acquisition de 12 h à 50°C (figure 1.4.13)[120]. Actuellement, la RMN des groupements méthyles des anticorps est utilisée pour l'obtention 'd'empreintes digitales' de la structure des anticorps et permet de contrôler leur qualité. Cependant, lorsqu'une modification a lieu sur les spectres, cette méthode a besoin d'être analysée en combinaison avec des études statistiques (par PCA), et de LC/MS pour positionner et quantifier les PTM des mAbs thérapeutiques[121].

La RMN des groupements méthyles des fragments Fab et Fc a montré que les structures des domaines d'IgG sont hautement conservées lors du clivage de l'anticorps à la papaïne, et les spectres obtenus sont plus simples à interpréter. De plus, la taille réduite des fragments Fab et Fc par rapport à une IgG entière permet d'obtenir des spectres 2D  $^1\text{H}$ - $^{13}\text{C}$  de qualité, de manière plus rapide (4.75h contre 12h) (figure 1.4.14).

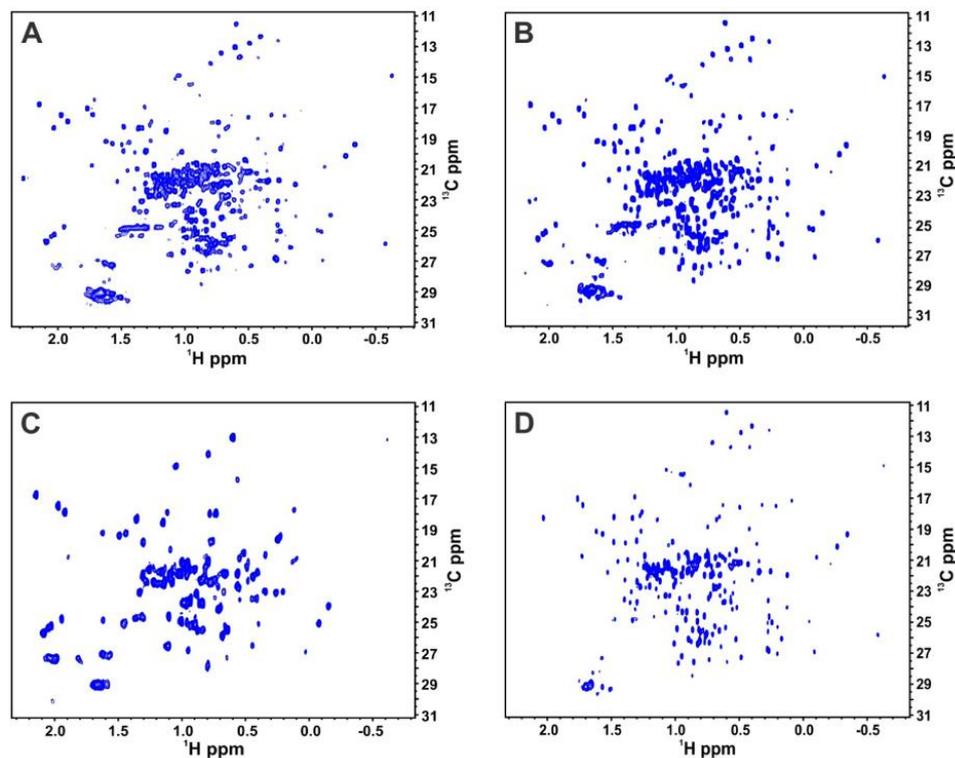


Figure 1.4.14 : Spectres RMN-HSQC 2D  $^1\text{H}$ - $^{13}\text{C}$  des groupements méthyles à 900MHz et 50°C. Spectre du mAb NIST entier à environ 260  $\mu\text{M}$  (A), du mix de Fab et Fc (500  $\mu\text{M}$  et 250  $\mu\text{M}$ ) après clivage à la papaïne(B), et des fragments Fc à 250  $\mu\text{M}$  (C) et Fab à 500  $\mu\text{M}$  (D). L'acquisition du spectre A a été de 12 h, contre 4.75 h pour les autres. L. W. Arbogast et al. *Anal. Chem.*, vol. 87, no 7, p. 3556-3561, avr. 2015, doi: 10.1021/ac504804m.

En résumé, la caractérisation de la HOS des mAbs thérapeutiques par RMN est confrontée à la taille importante des anticorps. Bien que la RMN des anticorps en abondance naturelle permette d'obtenir une empreinte digitale précise de la HOS, celle-ci présente des limitations et doit être utilisée avec des méthodes de statistiques pour évaluer si les changements observés sur les spectres sont significatifs d'une modification de la HOS. Cependant, lorsqu'un déplacement chimique est modifié à cause d'une altération de la HOS, il est difficile, voire impossible, de savoir quel acide aminé est concerné sans avoir recours à des méthodes orthogonales comme la LC/MS. Pour apporter plus de précision à l'analyse de la HOS des mAbs thérapeutiques par RMN, les spectres 2D  $^1\text{H}$ - $^{15}\text{N}$  et surtout  $^1\text{H}$ - $^{13}\text{C}$  doivent être attribués. Pour cela, un marquage isotopique des anticorps doit être réalisé afin d'acquies les expériences RMN 3D nécessaires, mais une telle méthode n'existe pas encore pour les anticorps produits en cellules CHO. La mise en place d'une méthode de marquage isotopique pour l'attribution des spectres  $^1\text{H}$ - $^{13}\text{C}$  des groupements méthyles des anticorps est l'objet principal de cette thèse, et la stratégie adoptée pour résoudre cette problématique sera détaillée dans la partie 1.5.

## 4.4. Le marquage isotopique de protéines de haut poids moléculaire pour la RMN

### 4.4.1. Intérêt du marquage isotopique $^{15}\text{N}$ , $^{13}\text{C}$ des protéines pour la RMN

Les principaux atomes observés en RMN des protéines sont les hydrogènes, carbones et azotes. Cependant, en RMN, seuls les noyaux possédant un moment angulaire intrinsèque, ou spin, d'une valeur de  $\frac{1}{2}$  fournissent un signal exploitable. Or, ce n'est pas le cas pour les isotopes  $^{12}\text{C}$  et  $^{14}\text{N}$  des noyaux de carbone et d'azote qui sont les plus abondants parmi leurs isotopes naturels. En RMN des protéines, les isotopes de spin  $\frac{1}{2}$  qui fournissent un signal sont l'hydrogène  $^1\text{H}$ , le carbone  $^{13}\text{C}$ , et l'azote  $^{15}\text{N}$ . L'hydrogène  $^1\text{H}$  possède une abondance naturelle proche de 100% (99.98%), ce qui veut dire que la quasi-totalité des protons d'une protéine non marquée vont fournir un signal. En revanche, les abondances naturelles du carbone  $^{13}\text{C}$  et de l'azote  $^{15}\text{N}$  sont respectivement de 1.11% et de 0.37%. Cela signifie que pour une protéine non marquée, seuls 1.11% des noyaux de carbones et 0.37% des noyaux d'azotes participent au signal RMN. Cette abondance naturelle est suffisante pour l'acquisition de spectres RMN 2D de protéines, mais le temps d'acquisition va dépendre de la taille de la protéine. Par exemple, l'obtention d'un spectre RMN 2D des groupements méthyles de bonne qualité pour une protéine de 50 kDa non marquée en utilisant un spectromètre RMN équipé d'une sonde cryogénique nécessite environ 16 h d'acquisition, tandis qu'il faut environ deux jours pour un spectre RMN des groupements amide. Ces temps d'acquisition sont longs et coûteux. Pour acquérir des spectres plus rapidement, le marquage isotopique est nécessaire.

Pour le carbone et l'azote, le marquage isotopique d'une protéine consiste à remplacer ces noyaux par leurs isotopes de spin  $\frac{1}{2}$  ( $^{13}\text{C}$  et  $^{15}\text{N}$ ), ce qui a pour effet d'augmenter le signal. Cette augmentation est très significative si le taux d'incorporation des isotopes est élevé. Par exemple, si 100% des carbones et des azotes sont marqués par des  $^{13}\text{C}$  ou  $^{15}\text{N}$ , leur signal est augmenté respectivement d'un facteur 90 et 270. Cette augmentation du signal a pour effet de permettre une diminution de la durée d'acquisition des spectres RMN. Pour la même qualité de signal (même *ratio* signal sur bruit des résonances), un gain d'intensité d'un facteur 100 équivaut à une durée d'acquisition divisée par un facteur 10 000. Ainsi, l'acquisition d'un spectre RMN 2D des groupements méthyles d'un anticorps marqué sur 100% des groupements carbones nécessite environ une minute (comparé à environ 16 h).

Le marquage isotopique  $^{15}\text{N}$ ,  $^{13}\text{C}$  des protéines est également indispensable car l'étude des protéines par RMN a une résolution atomique nécessite l'attribution des spectres RMN (cf. partie 1.4.3). Pour cela, des expériences RMN 3D doivent être acquises. Pour ces expériences, le transfert de magnétisation peut se faire entre deux atomes de carbones ou d'azotes. Dès lors, si la protéine n'est pas marquée, il y a très peu de probabilité que les deux atomes soient des isotopes donnant un signal RMN exploitable. Dans ce contexte, le marquage isotopique  $^{15}\text{N}$ ,  $^{13}\text{C}$  des protéines devient essentiel.

#### 4.4.2. Nécessité de la deutération des protéines de haut poids moléculaire pour la RMN

##### a La relaxation du signal RMN des protéines de haut poids moléculaire

Lors d'une expérience RMN, des impulsions radiofréquences sont appliquées pour amener l'aimantation dans le plan transversal et éditor les fréquences de chaque spin. L'aimantation dans le plan transversal a une durée de vie limitée. Sa disparition au cours du temps est appelée relaxation transversale ( $R_2$ ). Cette relaxation transversale est définie principalement par les interactions dipolaires qui ont lieu à travers l'espace lorsque le champ magnétique d'un (ou des) spins agit sur celui d'un (ou des) autre(s) et *vice-versa*. Pour un proton donné, appartenant à un groupement méthyle, la relaxation transversale est définie par l'équation 1.

$$\text{Equation 1 : } R_2(^1\text{Hi}) = R_{\text{Hi-Ci}}^{\text{DD,DD}} + R_{\text{Hi-Hi}_{\text{intra}}}^{\text{DD,DD}} + \sum_j R_{\text{Hi-Hj}_{\text{inter}}}^{\text{DD,DD}}$$

Les deux premiers termes de la formule concernent la relaxation due aux interactions dipolaires entre les spins du groupement méthyle. Le proton interagit avec l'hétéroatome (ici un carbone pour un groupement méthyle) qui lui est lié de manière covalente (1<sup>er</sup> terme), et avec les autres protons liés de manière covalente à cet hétéroatome (2<sup>ème</sup> terme) (toujours dans le cas d'un groupement méthyle, les deux autres protons du groupement par exemple). Le troisième terme concerne l'interaction dipolaire du proton avec les protons situés à l'extérieur du groupement chimique étudié.

Pour l'acquisition d'expériences RMN basées sur les groupements amides ou méthyles, les isotopes  $^1\text{H}$ ,  $^{13}\text{C}$  du groupement méthyle (ou  $^{15}\text{N}$  pour le squelette) doivent être présents, et l'impact de leurs interactions dipolaires sur la relaxation sera toujours le même. Pour minorer ces deux termes, des expériences RMN TROSY ont été développées (*Transverse Relaxation Optimized Spectroscopy*). Les séquences de pulses de ces expériences ont été optimisées pour favoriser la relaxation des atomes d'un même système de spin par optimisation des phénomènes d'interférence. Les articles de Tugarinov *et al.* et de Pervushin *et al.* détaillent l'avantage des expériences RMN TROSY respectivement pour les groupements méthyles et amides[122], [123].

En revanche, on peut agir sur le troisième terme grâce au marquage isotopique. Cela peut s'expliquer en considérant son équation qui définit les différents facteurs impactant la relaxation due aux interactions dipolaires entre les protons proches dans l'espace (équation 2).

$$\text{Equation 2 : } R_{\text{Hi,Hj}}^{\text{DD,DD}} = K \frac{\gamma_i^2 \gamma_j^2}{r_{ij}^6} \times (5J(0) + 9J(\omega_H) + 6J(2\omega_H))$$

K est composé uniquement de constantes ( $\pi$ , la constante de Planck  $h$  et la constante magnétique  $\mu_0$ ) et est défini par l'équation 3:

$$\text{Equation 3 : } K = \frac{1}{8} \times \frac{h^2}{2\pi} \times \frac{\mu_0^2}{(4\pi)^2}$$

Pour les protéines de haut poids moléculaire (>40 kDa), les termes  $J(\omega_H)$  et  $J(2\omega_H)$  qui sont les fonction de densités spectrales pour les protons sont négligeables en comparaison du terme  $J(0)$ . Et  $J(0)$  est définie par l'équation 4 :

$$\text{Equation 4 : } J(0) = \frac{2}{5} \tau_c S^2$$

Ainsi, l'équation 4 devient :

$$\text{Equation 5 : } R_{Hi,Hj}^{DD,DD} = K \frac{\gamma_i^2 \gamma_j^2}{r_{ij}^6} \times 2 \tau_c S^2$$

En dehors de la constante  $K$ , trois facteurs peuvent être différenciés dans cette équation. En effet, la relaxation d'un proton (ou d'un atome) est proportionnelle aux moments gyromagnétiques  $\gamma$  des spins impliqués dans l'interaction dipolaire, à la distance les séparant à la puissance six ( $r_{ij}$ ), et à la taille de la protéine représentée par le  $\tau_c$  avec  $S$  qui est le paramètre d'ordre du système de spin concerné.

Le facteur  $r^6$  représente la distance entre les deux noyaux impliqués dans l'interaction dipolaire. Le fait que la vitesse de relaxation soit inversement proportionnelle à cette distance à la puissance six explique que les interactions dipolaires l'impactant le plus sont celles ayant lieu avec les protons les plus proches dans l'espace. Ainsi, les interactions dipolaires sont fortes entre les protons d'un groupement chimique et ceux de la chaîne latérale de l'acide aminé par exemple. Cela explique aussi que la relaxation est défavorisée dans les parties de la protéine où la densité de protons est élevée.

## **b Optimisation de la relaxation par la variation du $\tau_c$ de la protéine**

Le facteur  $\tau_c$  est le temps de corrélation de la molécule étudiée et est proportionnel à sa taille. Il s'agit du temps moyen nécessaire pour qu'une molécule se réoriente d'un radian. Plus une molécule possède une masse importante, plus le facteur  $\tau_c$  est important. Ce facteur est également impacté par la viscosité du milieu dans lequel la protéine est solubilisée en RMN du liquide.

Un moyen d'améliorer le signal RMN est de réduire le  $\tau_c$  de la protéine. Cela peut être réalisé par augmentation de la température d'acquisition. En effet, l'augmentation de la température a pour effet d'augmenter l'agitation thermique, ce qui entraîne une diminution du  $\tau_c$  environ d'un facteur deux tous les 20°C. C'est pour cela que les expériences RMN sont acquises à une température la plus élevée possible (sans dégrader la protéine). Cette augmentation de température a également pour effet de diminuer le facteur  $S^2$  ce qui est favorable pour la relaxation.

De manière détournée, il est possible de modifier le  $\tau_c$  en divisant la taille de la protéine. Un anticorps d'une masse de 150 kDa a un  $\tau_c$  plus élevé que celui de ses fragments de masse plus faible. Ainsi, pour les mêmes conditions d'acquisition, les signaux d'un fragment

d'anticorps sont plus intenses et moins larges que ceux de l'anticorps entier. Cependant, cette stratégie n'est valable que si le fragment analysé possède la même structure seul et en présence du reste de la protéine.

### c Optimisation de la relaxation par la deutération des protéines

La deutération des protéines est un moyen d'améliorer la relaxation des signaux qui est due aux interactions dipolaires entre protons. En remplaçant la plupart des protons que l'on ne souhaite pas observer par du deutérium ( $^2\text{H}$ ), les interactions dipolaires diminuent fortement. En effet, le deutérium possède un facteur gyromagnétique faible en comparaison des protons ( $\gamma[^2\text{H}]/\gamma[^1\text{H}] \approx 0.15$ ) et cela se traduit par une vitesse de relaxation plus lente des noyaux qui sont à proximité, et donc un signal plus intense et des raies de résonance moins larges. Pour deux hydrogènes proches dans l'espace, si l'un est remplacé par un deutérium, l'interaction dipolaire est divisée par un facteur 44 ( $1/0.15^2$ ), d'où l'affinement des raies.

Cette deutération est d'autant plus importante pour l'acquisition d'expériences RMN 3D de type HNCA. Pour ces expériences, l'azote est protoné et sert de point de départ pour transférer l'aimantation jusqu'au  $\text{C}\alpha$ . L'interaction dipolaire entre l'azote et son proton ne peut donc pas être évitée. Cependant, le  $\text{C}\alpha$  possède également un proton qui n'est pas utilisé dans le transfert et qui crée une forte interaction dipolaire ( $^{13}\text{C}\alpha\text{-}^1\text{H}$ ). Cette interaction est donc divisée par un facteur 44 lors de la deutération des protéines et permet de récupérer un signal pour les expériences RMN 3D qui sont des expériences peu sensibles.

De plus, en deutérant ou en marquant de manière spécifique une protéine, c'est-à-dire seulement certains groupements chimiques ou acides aminés, les spectres RMN ne contiennent que les signaux correspondant au marquage effectué. Ainsi, il est possible de simplifier l'interprétation des spectres RMN par cette méthode car ils contiennent moins de signaux qui sont plus intenses et moins larges.

## 5. Stratégie mise en place pour étudier, par RMN, la HOS des anticorps thérapeutiques en solution et à l'échelle atomique

La RMN en abondance naturelle des anticorps est limitée en raison de leur taille et du fait que les spectres  $^1\text{H}$ - $^{13}\text{C}$  des groupements méthyles ne sont pas attribués. Jusqu'à présent, la méthode est ainsi limitée à l'utilisation d'outils statistiques pour évaluer les modifications observées sur les spectres. L'attribution des spectres RMN  $^1\text{H}$ - $^{13}\text{C}$  des groupements méthyles de l'anticorps apporterait des informations supplémentaires sur l'origine structurale des pertes d'activité observées lors d'études de stress.

### 5.1. Méthodes d'attribution des groupements méthyles par RMN

L'attribution des groupements méthyles par RMN est possible en utilisant différentes méthodes. Le choix se fait en tenant compte du système d'expression de la protéine, du nombre de groupements méthyles à attribuer et de la taille de la protéine. L'attribution peut être réalisée (i) par transfert d'attribution en partant du squelette de la protéine déjà attribué, (ii) par mutagenèse dirigée de chacun des résidus portant un groupement méthyle (iii) ou en utilisant la structure et des sondes paramagnétiques ou l'effet NOE (*nuclear overhauser effect*).

#### 5.1.1. Connection des groupements méthyles aux résonances du squelette à l'aide d'expériences HCC

Pour cette méthode, les signaux RMN des  $\text{C}\alpha$  et  $\text{C}\beta$  du squelette attribués au préalable sont utilisés pour attribuer les groupements méthyles. Historiquement, il a été proposé d'attribuer les protons des chaînes latérales des acides aminés à l'aide d'expériences RMN basées sur le couplage scalaire  $J_{\text{H-H}}$  [124] (COSY, *CORrelation Spectroscopy* ou TOCSY, *Total Correlation Spectroscopy*). L'introduction du marquage isotopique uniforme  $^{15}\text{N}$  et  $^{13}\text{C}$  a permis l'acquisition d'expériences RMN 3D HCCH-TOCSY permettant de décrire des chaînes latérales entières de l'interleukin- $1\beta$  [125]. Cependant, cette protéine ne fait que 17 kDa, et la problématique n'est pas la même pour des protéines de masse plus élevée ( $> 30$  kDa) qui peuvent limiter l'utilisation de ces expériences en raison d'une relaxation transversale rapide et d'une résolution plus faible pour les protéines de masse moléculaire plus importante.

Le développement du marquage isotopique en  $^{13}\text{C}$  et  $^2\text{H}$  combiné à de nouvelles expériences RMN 3D basées sur le couplage  $J_{\text{C-C}}$  a permis de dépasser cette limitation. En effet, ces expériences 3D exploitent les couplages scalaires  $J_{\text{C-C}}$  ( $=35\text{Hz}$ ) et maintiennent la relaxation transversale à une vitesse réduite pendant ces transferts grâce à la perdeutération des chaînes latérales.

Deux types de séquences d'impulsion menant à deux chemins de cohérence différents pour la magnétisation existent (figure 1.5.1).

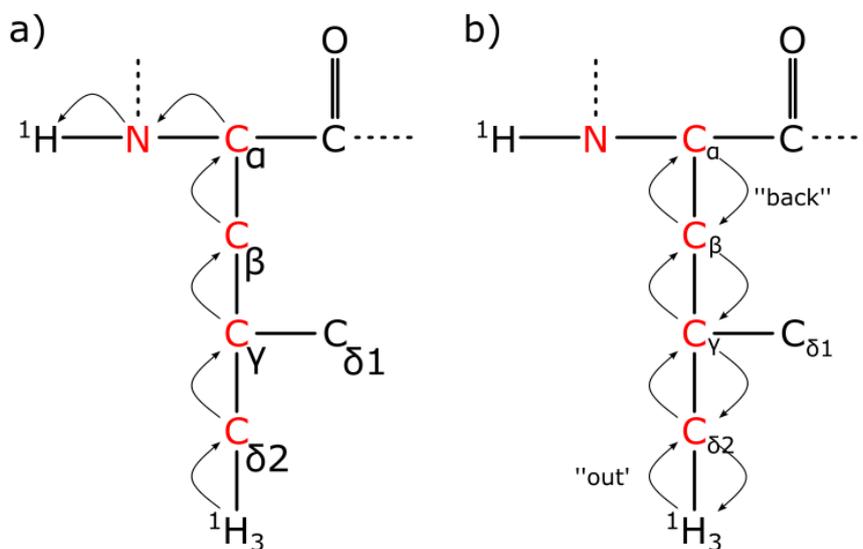


Figure 1.5.1 : Schéma représentant les différentes étapes du transfert d'aimantation intra-résidu, illustrées pour une leucine marquée  $U\text{-}[^2\text{H}, ^{15}\text{N}, ^{13}\text{C}]\text{-Leu-}[^{13}\text{C}^1\text{H}_3]\delta 2/[^{12}\text{C}^2\text{H}_3]\delta 1$  a) pour la séquence d'impulsion  $(\text{HM})\text{CM}(\text{CGCBCA})\text{NH}$ . b) pour la séquence d'impulsion « out and back HCC-COSY ». Figure adaptée de Tugarinov and Kay 2003. Les azotes et les carbones marqués en  $^{15}\text{N}$  et  $^{13}\text{C}$  sont représentés en rouge.

Le transfert d'aimantation des groupements méthyles jusqu'aux groupements amides du squelette peut être réalisé avec des expériences de type  $\text{H}_\text{M}\text{C}_\text{M}\text{C}_\text{G}\text{C}_\text{B}\text{C}_\text{A}\text{NH}$  (figure 1.5.1 a). Cette méthode est avantageuse car elle permet d'obtenir des spectres avec une bonne dispersion des déplacements chimiques dans les dimensions  $^1\text{H}$  et  $^{15}\text{N}$ , mais peut s'avérer limitée pour les protéines de haut poids moléculaire.

Le deuxième type d'expérience consiste en un transfert d'aimantation des groupements méthyles jusqu'aux carbones du squelette ( $\text{C}_\alpha$  et/ou  $\text{C}_\beta$ ), puis un retour de l'aimantation sur le groupement méthyle de départ pour la détection (figure 1.5.1 b). Ces expériences sont appelées « out and back » en raison de l'aller-retour effectué par l'aimantation au cours de l'expérience. Cette méthode est plus sensible que la précédente car il n'est pas nécessaire d'aller jusqu'au groupement amide.

Cette stratégie d'attribution des groupements méthyles nécessite un marquage isotopique spécifique des acides aminés. Le transfert de l'aimantation est optimal si les acides aminés sont porteurs d'une chaîne linéaire de  $^{13}\text{C}$  en partant de leur groupement méthyle jusqu'aux carbones (et  $^{15}\text{N}$  si l'aimantation va jusqu'à l'azote) du squelette. Il en résulte une limitation à l'utilisation de cette stratégie pour l'attribution des groupements méthyles de certains acides aminés. Par exemple, les groupements méthyles des méthionines ne peuvent être attribués de cette façon en raison de la présence de l'atome de soufre dans leur chaîne latérale qui ne permet pas les transferts d'aimantation de ce type. Les acides aminés leucine, isoleucine et valine présentent aussi une complication car la présence de leurs deux groupements méthyles offre deux chemins de cohérence différents pour l'aimantation. Si on prend l'exemple de l'isoleucine, et si l'aimantation débute sur le carbone  $\delta 1$ , l'aimantation va être transférée sur le  $\text{C}_\gamma$ , puis sur le  $\text{C}_\beta$ , puis à la fois sur le  $\text{C}_\gamma 2$  et le  $\text{C}_\alpha$  (figure 1.5.2). Cela a pour effet de réduire l'efficacité du transfert jusqu'au  $\text{C}_\alpha$  et donc de diminuer la sensibilité de l'expérience. Ce

problème a été surmonté avec le développement du marquage isotopique permettant d'obtenir des acides aminés marqués de manière régiospécifique (isoleucine) et stéréospécifique (valine et leucine) sur le groupement méthyle et connecté au squelette par une chaîne de  $^{13}\text{C}$  linéaire [126], [127].

Cette stratégie a permis l'attribution des signaux  $\delta 1$  des isoleucines et proR des valines de TET2 (468 kDa) et MSG (82 kDa) [126].

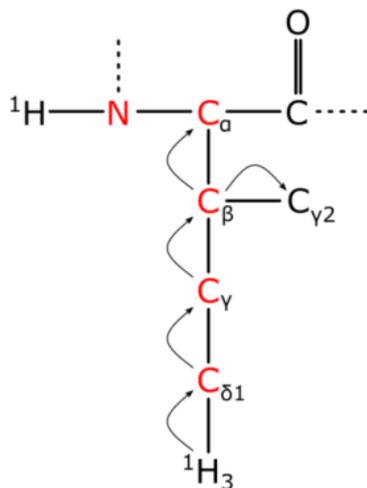


Figure 1.5.2 : Chemin emprunté par l'aimantation au cours d'une expérience de type HCC si le marquage isotopique  $^{13}\text{C}$  n'est pas linéaire. L'exemple est donné pour une isoleucine.

### 5.1.2. Attribution des groupements méthyles par mutagenèse dirigée.

La méthode d'attribution des groupements méthyles par mutagenèse consiste à produire des mutants ponctuels de la protéine à attribuer. Le mutant est produit avec un marquage spécifique sur les résidus du même type que celui muté (par exemple, marquage de toutes les isoleucines si le résidu muté est une isoleucine) de façon à obtenir des spectres RMN 2D  $^1\text{H}-^{13}\text{C}$  pour chacun des mutants. L'objectif est ensuite de comparer chacun de ces spectres avec le spectre de référence de la protéine non mutée pour identifier le pic manquant (du résidu muté) et donc attribuer ce résidu sur le spectre de référence [128].

Néanmoins, en pratique, la mutation d'un acide aminé peut entraîner une modification structurale locale dans la protéine, ce qui peut impliquer, outre la disparition du pic de l'acide aminé muté, la modification du déplacement chimique de certains pics. Ces déplacements chimiques modifiés, dit secondaires, compliquent l'attribution. Cette méthode peut également être très chronophage car si l'on souhaite attribuer les groupements méthyles d'une protéine qui en possède 250, alors il faut produire 250 mutants différents et acquérir autant, de spectres RMN. Les étapes de production et purification des mutants peuvent, dans ce cas, devenir fastidieuses, et, par ailleurs, la manipulation d'un grand nombre d'échantillons peut présenter un risque d'erreur (mélange d'échantillons par le manipulateur). De plus, certains mutants, impactant la nature de la protéine, peuvent ne pas être exprimés correctement (perte de rendement ou d'expression, perte de stabilité).

Afin de limiter les déplacements chimiques secondaires il faut remplacer l'acide aminé muté par un acide aminé homologue pour limiter les perturbations locales. Pour ce faire, on peut se baser sur les matrices de BLOSUM récapitulant les acides aminés les plus susceptibles d'être interchangés dans une séquence de protéine (tableau 1.5.1).

La préparation des échantillons a ensuite été optimisée par Crublet *et al.* pour une production en bactérie[129]. Le milieu de culture M9 classique a été remplacé par un milieu de culture riche, permettant d'augmenter les rendements et donc de diminuer les volumes de culture et le coût de préparation des échantillons marqués. Les flasques de culture ont ainsi été remplacés par des plaques de culture DeepWell 24 puits. Cela a permis de paralléliser la production, mais aussi la purification puisque la diminution des volumes à traiter a permis une purification des mutants, simultanément, sur des faibles volumes de résine dans des plaques de purification de 96 puits. Cette optimisation leur a permis d'attribuer, en deux semaines (dont 4 jours d'acquisition de spectres RMN 2D  $^1\text{H}$ - $^{13}\text{C}$ ), une centaine de groupements méthyles pour les acides aminés alanine, isoleucine et valine de la protéine TET2.

Bien que permettant d'attribuer les groupements méthyles de gros complexes protéiques comme TET2, cette méthode est cependant limitée lorsque la protéine est difficile à exprimer en bactérie, ou si elle doit être exprimée avec un autre système. Si les rendements d'expression sont faibles, les volumes de milieu de culture riche, et d'acides aminés marqués isotopiquement sont trop importants et cette méthode devient trop coûteuse. De plus, l'utilisation de cette méthode se complexifie en fonction du nombre de groupements méthyles que contient la protéine.

### 5.1.3. Utilisation de la structure pour l'attribution des groupements méthyles

L'utilisation de la structure 3D (préalablement déterminée par Cryo-EM ou cristallographie par exemple) pour l'attribution des groupements méthyles peut se faire de deux manières différentes. Une des méthodes consiste à insérer des sondes paramagnétiques dans la protéine et utiliser leur effet sur le signal des protons proches dans l'espace, tandis que l'autre utilise des expériences RMN spécifiques exploitant l'effet NOE pour observer les interactions des protons entre eux et dans l'espace.

#### a Utilisation de la structure et des sondes paramagnétiques pour l'attribution des groupements méthyles

Cette méthode d'attribution est basée sur l'incorporation d'un centre paramagnétique et l'étude des effets de relaxation et de déplacements chimique induits dans les spectres RMN. Ce centre paramagnétique est une espèce chimique possédant un électron apparié comme le groupement nitroxyde, le manganèse ( $\text{Mn}^{2+}$ ) ou les noyaux de la famille des lanthanides. Leur effet paramagnétique peut aller jusqu'à 50Å de distance en fonction du centre paramagnétique choisi et élargit les raies des résonances. Tous ont un impact sur le phénomène de relaxation dit

Tableau 1.5.1 : Mutations suggérées par la matrice BLOSUM pour les acides aminés comportant un groupement méthyle

Acides aminés comportant un groupement méthyle	Mutation suggérée par la matrice BLOSUM
Alanine	Sérine
Isoleucine	Valine
Leucine	Isoleucine
Valine	Isoleucine
Thréonine	Sérine
Méthionine	Leucine

effet PRE (*paramagnetic relaxation enhancement*), et certains ont également un impact sur les déplacements chimiques, dit effet PCS (*pseudocontact shifts*). L'intensité de l'effet paramagnétique est inversement proportionnelle à la distance au centre paramagnétique à la puissance 6 pour l'effet PRE, et dépend de l'orientation du vecteur "centre paramagnétique-spin nucléaire" dans le tenseur de susceptibilité magnétique du centre paramagnétique et de la norme de ce vecteur à la puissance 3 pour l'effet PCS. Connaissant la structure de la protéine et les caractéristiques du centre paramagnétique (i.e. tenseur de susceptibilité magnétique) les effets de l'insertion du centre paramagnétique peuvent être prédits et comparés aux valeurs expérimentales en comparant les effets observables obtenus en présence d'un centre paramagnétique ou d'un centre diamagnétique à la même position dans deux échantillons différents. La comparaison entre les valeurs PRE et PCS expérimentales et prédites peut être utilisée pour proposer une attribution des signaux RMN.

Cette méthode nécessite la production de plusieurs échantillons avec des incorporations d'un centre paramagnétique à des endroits spécifiques. Les cystéines réduites (donc non réalisable pour un anticorps ou l'un de ses fragments) ou les lysines sont généralement utilisées pour insérer un centre paramagnétique. La production de ces échantillons peut être chronophage car l'incorporation d'un centre paramagnétique n'est pas facile et peut entraîner des pertes de rendements ou une absence d'expression de la protéine.

En plaçant le centre paramagnétique judicieusement, cette méthode est avantageuse pour obtenir des informations sur un groupement méthyle se trouvant isolé sur la structure (sans autres groupements méthyles à proximité), et donc non attribuable en utilisant l'effet NOE (cf. partie 5.1.3 b). Cependant, l'attribution de la totalité des groupements méthyles d'une protéine est compliquée voire impossible en utilisant uniquement cette méthode. F. Henot *et al.* ont par exemple montré que l'attribution des groupements méthyles (des valines et leucines) du domaine N-terminal de la protéine HSP90 (25 kDa) réalisée en utilisant des sondes paramagnétiques n'était juste que dans 80% des cas [130], [131]. Pour cette raison, cette méthode ne sera pas détaillée plus que cela.

## **b Utilisation de la structure et du réseau de NOE pour l'attribution des groupements méthyles**

L'effet nucléaire Overhauser (NOE) est créé par l'interaction dipolaire des moments magnétiques de deux noyaux (ou plus) proches dans l'espace. En interagissant ensemble, leur moment magnétique a un impact sur le phénomène de relaxation de l'autre et *vice-versa* ; on parle alors de relaxation croisée (cross-relaxation). En utilisant des séquences de *pulse* spécifiques appelées NOESY qui sont des expériences RMN 2D, il est possible d'obtenir un spectre représentant l'effet NOE entre deux spins (ou groupes de spins équivalents comme les groupements méthyles) où chaque signal correspond à l'interaction d'un groupement méthyle avec un autre dans l'espace. La relaxation croisée est proportionnelle à  $1/r^6$  où  $r$  est la distance moyenne entre les deux (groupes de) spins dans l'espace. Cette interaction diminue donc fortement en fonction de la distance et est observable entre des groupements méthyles séparés par une distance allant jusqu'à 6-8Å. Ainsi, si un effet NOE est observé pour deux systèmes de spin, alors les deux noyaux concernés sont proches dans l'espace, et l'intensité du signal fournit

une information sur la distance les séparant (pour une présentation plus détaillé de l'effet NOE, je recommande le livre du Dr. James Keeler [132]).

Cette technique est très utilisée pour l'étude structurale des protéines par RMN et elle permet l'attribution des groupements méthyles en se basant sur la connaissance de la structure de la protéine à haute résolution et l'acquisition d'expériences RMN fournissant un réseau de NOE. Par comparaison des groupements méthyles proches dans l'espace sur la structure, et des effets NOE observés, il est possible d'attribuer les signaux des spectres aux acides aminés concernés.

Contrairement aux méthodes présentées précédemment, celle-ci ne nécessite pas de produire des mutants et un seul échantillon peut suffire pour la collection des données RMN. Cependant, l'effet NOE n'est observable que sur de courtes distances et la diffusion des spins à temps de mélanges élevés (dans les parties hautement structurées) peut compliquer l'interprétation des spectres et mener à des attributions ambiguës [133]. L'isolement de certains groupements méthyles (à plus de 8 Å de distance d'un autre groupement méthyle) peut également empêcher leur attribution car ils ne présenteront aucun effet NOE. Ces complications peuvent être surmontées en partie par la deutération complète de la protéine et la protonation spécifique des groupements méthyles [134]. Cela a permis d'améliorer la relaxation et ainsi l'observation d'effet NOE entre spins séparés de 12 Å de distance pour la protéine MSG (malate synthase G) de 82k Da [127].

La relation d'une structure tridimensionnelle et d'un réseau de corrélations NOE peut s'avérer très difficile à interpréter pour les grosses protéines hautement structurées (> 40 kDa). Ce travail peut mener à des ambiguïtés d'attribution. Pour aider, de nombreux programmes d'attribution employant l'effet NOE ont été développés. MAP-XSII a été le premier développé en 2009 par Xu *et al.* [135], [136] suivi rapidement par cinq autres : le programme FLAMEnGO 2.0 développé par Chao *et al.* en 2014 [137], MAGMA développé par Pritišanac *et al.* en 2017 [138], MAGIC développé par Monneau *et al.* en 2017 [139], MethylFLYA développé par Pritišanac *et al.* en 2019 [140] et enfin MAUS développé par Nerli *et al.* en 2021 [141].

Ces programmes sont basés sur l'utilisation de différents algorithmes permettant de comparer un réseau de corrélations NOE théoriques (calculées à partir de la structure 3D) au réseau NOE obtenu grâce à des expériences RMN 3D NOESY. Chaque programme a ses avantages et inconvénients et permet d'obtenir des attributions plus ou moins complètes et/ou justes. On peut distinguer trois types d'attributions qui permettent de comparer ces logiciels : les attributions correctes et non-ambiguës, les attributions correctes et ambiguës, et les attributions incorrectes. Pour des protéines de 45 kDa (HLA-A01 et HLA-A02) et l'attribution d'environ 90 de leurs groupements méthyles (pour les résidus AILV), le programme MAUS a permis d'obtenir plus d'attributions correctes (ambiguës ou non) que les autres. Mais, le pourcentage d'attributions correctes et non ambiguës est plus élevé pour le programme Methyl-Flya et MAP-XSII. Enfin, MAP-XSII a fourni plus d'attributions incorrectes. La figure 1.5.3 représente un histogramme résumant les performances des différents programmes d'attribution automatiques et réalisé par le docteur Faustine Henot à partir des tableaux de Nerli *et al.* [128], [141]. Cet histogramme donne un bon aperçu de leur performance en fonction de la taille de la protéine et de la quantité de groupements méthyles à attribuer.

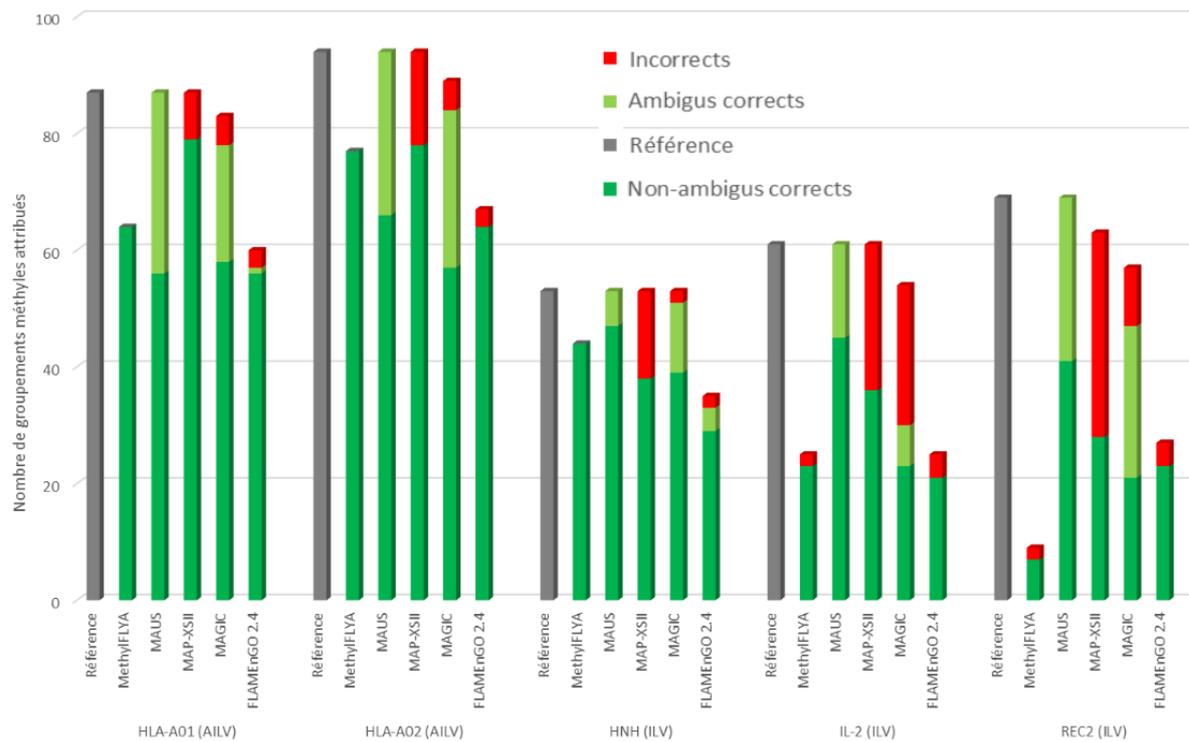


Figure 1.5.3 : Histogramme résumant les performances de différents programmes d'attribution automatique des groupements méthyles pour cinq protéines : HLA-A01, HLA-A02, HNH, IL-2 et REC2. Les données proviennent du tableau de Nerli et al. (Nerli et al. 2021), et l'histogramme provient du manuscrit de thèse du docteur Faustine Henot. F. Henot, "HSP90 une cible thérapeutique dynamique : modulation des conformations peuplées par la liaison de ligands," 2022.

De manière générale, la performance de ces logiciels est améliorée lorsque des informations sur le type d'acide aminé sont fournies, et les marquages isotopiques stéréospécifiques et régiosélectifs l'améliorent également. Mais leur emploi reste difficile et, par expérience, aucun de ces programmes n'a permis l'attribution complète des groupements méthyles d'une protéine de 29 kDa étudiée au laboratoire[131] (Henot et al 2022).

## 5.2.Stratégie d'attribution des groupements méthyles de l'anticorps anti-LAMP1 exprimé en cellules CHO

L'objectif de cette thèse est d'utiliser la RMN des groupements méthyles pour des études structurales ou de dynamique des anticorps en solution et à l'échelle atomique. Pour cela, les signaux RMN des spectres  $^1\text{H}$ - $^{13}\text{C}$  doivent être attribués aux acides aminés portant un groupement méthyle de l'anticorps exprimé en cellules CHO. En effet, les anticorps thérapeutiques sont principalement exprimés en cellules CHO dans les entreprises pharmaceutiques et la production de l'anticorps anti-LAMP1 a été développée dans ces cellules par Sanofi. Il a été expliqué précédemment que les caractéristiques physico-chimiques telles que les glycosylations des anticorps étaient fortement impactées par le système d'expression utilisé. Il est donc important que la méthode de marquage isotopique  $^{13}\text{CH}_3$  des groupements méthyles de l'anticorps soit développée dans le même système d'expression que celui utilisé en industrie.

Pour l'attribution des signaux RMN d'une protéine de la taille d'un anticorps, celui-ci doit être perdeutééré et marqué spécifiquement sur les groupements méthyles pour l'acquisition d'expériences RMN 3D. Cependant, bien que le marquage isotopique des groupements méthyles soit possible en cellules CHO grâce au protocole développé au cours de la thèse, la perdeutération n'est pas applicable en raison de la toxicité du deutérium pour les cellules de mammifère. Il n'est donc pas possible de produire l'anticorps avec le marquage isotopique nécessaire à l'attribution de ses groupements méthyles en cellules CHO. Cependant, afin d'améliorer la qualité des signaux RMN des spectres  $^1\text{H}$ - $^{13}\text{C}$  des groupements méthyles de l'anticorps, une méthode de marquage isotopique  $^{13}\text{CH}_3$  des alanines, méthionines, thréonines, isoleucines- $\delta 1$ , leucines-proS et valines-proR a été développée pour ce système d'expression.

Le système acellulaire (*cell-free*) basé sur un extrait bactérien qui est plus adapté pour la production de protéines perdeutéérées avec un marquage isotopique des groupements méthyles a ensuite été utilisé. Cependant, même avec une perdeutération et un marquage isotopique optimal, un anticorps possède une masse moléculaire qui reste élevée pour la RMN (150 kDa). Ainsi, à la place de l'anticorps entier, les fragments Fab et Fc de l'anticorps ont été produits en système *cell-free*. De cette manière, la masse moléculaire de la protéine à attribuer est divisée par trois et les spectres RMN correspondants ont un rapport signal/bruit de meilleure qualité, ce qui facilite le travail d'attribution. Cette stratégie de "division" est encore plus avantageuse pour le fragment Fc de l'anticorps qui est un homodimère symétrique. Le spectre RMN des groupements méthyles d'un fragment Fc possède un nombre de signaux correspondant à une protéine de 25 kDa et non de 50 kDa comme pour le fragment Fab. Cependant, chaque NOE peut être interprété comme provenant d'une interaction entre deux groupements méthyles appartenant à la même chaîne ou appartenant aux deux chaînes du fragment Fc. Cette ambiguïté peut être résolue par un marquage isotopique différent des deux chaînes [142], mais ne sera pas explicité dans cette thèse focalisée sur l'attribution des signaux du domaine Fab.

L'expression de protéines riches en ponts disulfures en système *cell-free* bactérien est possible (cf. partie 3.2.1). Cependant, exprimer un fragment Fab d'anticorps en système acellulaire consiste à produire un complexe de deux chaînes protéiques riches en ponts disulfures (chaîne légère et chaîne lourde). Les premiers rendements d'expression obtenus au

début de la thèse sur le fragment Fab de l'anticorps anti-LAMP1 ont révélé un besoin d'optimisation de la méthode d'expression afin de diminuer les coûts de production d'un échantillon marqué isotopiquement. Ainsi, la méthode d'expression en système *cell-free* a été optimisée pour l'expression de fragments d'anticorps tout au long de la thèse.

La première étape pour l'attribution des groupements méthyles du fragment Fab a consisté à attribuer son squelette grâce à un échantillon U-[<sup>2</sup>H,<sup>13</sup>C,<sup>15</sup>N] et à l'acquisition d'expériences RMN 3D classiques (HNCO, HNCACO, HNCA, HNCOCA, HNCACB, HNCOCACB)[143].

A la suite de cette attribution, le fragment Fab perdeutééré avec un marquage isotopique <sup>13</sup>C linéaire et <sup>13</sup>CH<sub>3</sub> spécifique sur les groupements méthyles des alanines, isoleucines- $\delta$ 1 et valines-proR a été produit dans le but d'acquérir des expériences RMN de type HCC "out and back" pour l'attribution des groupements méthyles de ces trois types de résidus. L'attribution des groupements méthyles ne peut pas être complète avec cette méthode seule car le marquage isotopique <sup>13</sup>C linéaire n'est pas possible pour les méthionines (atome de soufre dans la chaîne latérale), et les thréonines et leucines U-[<sup>2</sup>H,<sup>13</sup>C] + <sup>13</sup>CH<sub>3</sub> (proS pour les leucines) n'existent pas commercialement.

Ainsi, pour compléter l'attribution obtenue avec les expériences de type HCC, le fragment Fab perdeutééré avec un marquage isotopique spécifique <sup>13</sup>CH<sub>3</sub> des alanines, méthionines, thréonines, isoleucines- $\delta$ 1, leucines-proS et valines-proR a été produit. Cet échantillon a servi à acquérir une expérience RMN NOESY inter-méthyles dont le réseau de NOE a été utilisé en combinaison avec la structure cristallographique du fragment Fab pour étendre l'attribution des groupements méthyles. Pour cette attribution, le logiciel Methyl-Flya a été utilisé.

Enfin, les attributions ont été transférés à l'anticorps anti-LAMP1 par comparaison des spectres RMN de l'anticorps exprimé et marqué <sup>13</sup>CH<sub>3</sub> en cellules CHO au spectre des groupements méthyles attribués du fragment Fab exprimé en système *cell-free*.

## II. Matériel et méthodes

### 1. Matériels et méthodes pour la production de l'anticorps en cellules CHO

#### 1.1. La lignée cellulaire CHO

La lignée cellulaire utilisée pour la production de l'anticorps anti-LAMP a été produite par GTP Technology™ : lignée cellulaire CHOvolution™. Les cellules ont été transfectées avec deux vecteurs d'expression linéarisés pour exprimer la chaîne légère et la chaîne lourde de l'IgG. Un vecteur comportait le gène de résistance à la généticine et l'autre le gène de résistance à l'hygromycine. L'utilisation de deux vecteurs exprimant la même protéine avec deux gènes de résistance différents a été réalisée pour obtenir une lignée cellulaire hyper productrice de l'anticorps.

Il ne s'agissait pas d'une lignée cellulaire clonale qui nécessite de démarrer la sélection clonale à une cellule par puits, donc l'anticorps utilisé pendant la thèse n'était pas un anticorps monoclonal. GTP Technology™ a sélectionné cette lignée cellulaire en démarrant l'étape de sélection à environ 250 cellules par puits. Bien que non monoclonale, la lignée cellulaire a été certifiée stable et productrice de l'anticorps à un rendement d'expression de 400 mg/L dans les conditions standards de culture fournies par GTP Technology™. Ce protocole de production est un mode *fed-batch* avec un apport régulier de nutriments jusqu'à la phase d'apoptose des cellules. Leur protocole conseille l'utilisation du milieu CDM4PERMAb (Cytiva) supplémenté avec les milieux Cell Boost 7a et Cell Boost 7b (Cytiva), 4 mM de glutamine (Merck), 10% de poloxamer 188, 400 µg/mL d'hygromycine B Gold et 400 µg/mL de généticine (Invitrogen). Les conditions de culture utilisées pour un mode *fed-batch* sont résumées dans le tableau 2.1.1.

Tableau 2.1.1 : Conditions de cultures conseillées par GTP Technology™ pour l'expression de l'anticorps anti-LAMP1 en mode *fed-batch* avec lignée cellulaire CHO développée

Paramètres	Valeurs
Température	36,5°C
Vitesse d'agitation	150 rpm (Ø 25mm) ou équivalent
Pourcentage de CO <sub>2</sub>	5%
Densité cellulaire de départ	0,5.10 <sup>6</sup> cells/mL
Durée de culture maximale	13 jours
Conditions d'arrêt de la culture	Si viabilité cellulaire < 60%
Concentration de glucose	Ajustée à 5,5 g/L de J5 à J7 Puis à 2g/L de J10 à J13
Concentration de glutamine	Ajustée à 2 % (v/v) de J7 à J10
Ajout des cell-boost 7a et 7b	À J0, 2% et 0,2% À J3, 4% et 0,4% À J5, 6% et 0,6% À J7, 9% et 0,9% À J10, 5% et 0,5%

## 1.2. Culture des cellules CHO et production de l'anticorps anti-LAMP1

Pour la maintenance des cellules en laboratoire, le milieu de culture CDM4PERMAb complétement comme décrit précédemment a été utilisé. Les cellules ont été cultivées à 37°C et 5% de CO<sub>2</sub> avec une vitesse d'agitation de 200 rpm (Ø 25mm). Les dilutions cellulaires étaient effectuées tous les trois ou quatre jours avec une densité cellulaire visée de 0,5.10<sup>6</sup> cellules/mL. Les cellules étaient comptées sur cellule de Mallassez avec une coloration au bleu de trypan.

Les productions de l'anticorps anti-LAMP1 en abondance naturelle ont été réalisées en flasque et en mode *batch* en utilisant le milieu de culture CDM4PERMAb complétement comme décrit précédemment. Les productions ont été réalisées avec des densités cellulaires initiales de 1.10<sup>6</sup> cellules/mL pendant des durées d'incubation de 4 jours à 37°C, 5% de CO<sub>2</sub> et 200 rpm (Ø = 25mm). Le tableau 2.1.2 résume les paramètres de culture cellulaire mis en place.

Tableau 2.1.2 : Conditions de cultures mises en place pour la production de l'anticorps anti-LAMP1 en mode batch avec la lignée cellulaire CHO développée par GTP Technology

Paramètres	Valeurs
Température	37°C
Vitesse d'agitation	200 rpm (Ø 25mm)
Pourcentage de CO <sub>2</sub>	5%
Densité cellulaire de départ	1.10 <sup>6</sup> cells/mL
Durée de culture maximale	4 jours
Conditions d'arrêt de la culture	Si viabilité cellulaire < 80%

## 1.3. Purification de l'anticorps anti-LAMP1 et quantification

Le surnageant des cultures, qui contient l'anticorps sécrété par les cellules CHO, a été prélevé juste avant la phase d'apoptose des cellules afin de limiter la quantité de débris cellulaires à purifier par la suite. Cela a été réalisé par centrifugation des cellules à 200 g pendant 15 minutes. Le surnageant ainsi obtenu a ensuite été dilué trois fois dans un tampon 20mM Tris base, 20 mM NaCl pH 7.4 (tampon A).

Pour la purification de l'anticorps, une colonne Hiscreen MabSelect Sure (protéine A) de 5 mL a été utilisée et connectée à un système de chromatographie NGC (Bio-Rad). Le surnageant dilué a été injecté sur la colonne équilibrée en tampon A à un débit de 1 mL/min. Un premier lavage avec 5 CV de tampon A a été réalisé, suivi d'un lavage avec 2 CV de tampon B (20 mM Tris Base, 1 M NaCl, pH 7.4). Pour réduire la quantité de sel présente avant l'élution, un troisième lavage a été appliqué avec 5 CV de tampon A. L'élution de l'anticorps a été réalisée avec 6 CV de tampon C (100 mM glycine, 10 mM NaCl, pH 3). Le suivi du profil d'élution a été effectué par lecture de l'absorbance à 280 nm. Avant de collecter les fractions d'élution (1 mL), 30 µL de tampon D (1 M Tris Base, pH 8) ont été ajoutés dans les puits de la plaque de collection pour tamponner les fractions d'élution.

Les fractions contenant l'anticorps ont ensuite été rassemblées et concentrées jusqu'à un volume de 0.5 mL en utilisant un Amicon possédant un *cut-off* de 50 kDa.

À la suite de cette concentration, les échantillons ont été injectés sur une colonne de filtration sur gel pour effectuer une deuxième étape de purification par chromatographie d'exclusion de taille. Cela a été réalisé sur une colonne de gel superose 6 10/300 (Cytiva) à un débit de 0.5 mL/min. Les fractions contenant l'anticorps ont été analysées sur gel SDS-PAGE à 12.5% de polyacrylamide. Les fractions contenant l'anticorps ont été réunies puis concentrées sur Amicon avec *cut-off* de 50 kDa jusqu'à un volume de 0.5 mL.

Le calcul de la concentration d'anticorps en solution a été réalisé par absorbance à 280 nm ( $\epsilon = 223\,400\text{ M}^{-1}\cdot\text{cm}^{-1}$  ;  $A = 1.544$  ; MW=144 724 Da PI : 7.9) en utilisant un Nanodrop 2000 (Thermo Fisher).

### 1.4.Spectroscopie RMN

L'anticorps produit en cellules CHO pour les analyses RMN l'a été en utilisant le milieu de culture CDM4PERMAb dépourvu d'acides aminés et complété avec le mélange d'acides aminés développé, 400  $\mu\text{g/mL}$  d'hygromycine B Gold et 400  $\mu\text{g/mL}$  de généticine. Les acides aminés comportant des groupements méthyles étaient U-[ $^1\text{H}$ ], [ $^{13}\text{C}^1\text{H}_3$ ] $\epsilon$ -Met (299146, Sigma Aldrich), 2-( $^2\text{H}$ )-3-( $^{13}\text{C}$ )-L-Alanine (NMR-Bio),  $^{15}\text{N}$ -2,3,3,4,4,4-(D5)-proS- $^{13}\text{CH}_3$ -L-Val (NMR-Bio),  $^{15}\text{N}$ -2,3,4,4-(D4)-3-méthyl-(D3) 5- $^{13}\text{CH}_3$ -L- Ile $\delta$ 1 (NMR-Bio), 2,3-(D2) 4-( $^{13}\text{CH}_3$ )-L-Thr (NMR-Bio), 2,3,3,4,5,5,5-(D7)-pro-S-méthyl- $^{13}\text{C}$ -Leu (900795, Sigma Aldrich). De la L-cyclosérine (10 mg/L) a été utilisée en tant qu'inhibiteur de transaminase lors du marquage isotopique de l'alanine. Les concentrations d'acides aminés marqués ajoutés pendant la culture ont varié de 100 à 350 mg/L au cours du développement.

Après purification, le tampon des échantillons à analyser par RMN a été échangé avec un tampon MES (acide 2-(N-morpholino)éthanosulfonique) 50 mM, 100 mM NaCl, pH 6.5, 10% ou 100% de  $^2\text{H}_2\text{O}$  (en fonction du type d'échantillon).

Les expériences RMN ont été acquises à 35°C en utilisant des spectromètres RMN équipés d'une sonde cryogénique et opérant à une fréquence  $^1\text{H}$  de 950, 850 et 600 MHz. Les expériences RMN acquises étaient des 2D  $^1\text{H}$ - $^{13}\text{C}$  SOFAST centrées sur la région des groupements méthyles. Le logiciel utilisé pour l'exploitation des spectres était la version 4.1 de TopSpin (Bruker). Tous les échantillons ont été analysés dans le même tampon RMN dans des tubes RMN de 3 mm.

### 1.5.Spectrométrie de masse

Pour la spectrométrie de masse, un marquage isotopique différent de celui réalisé pour la RMN a été effectué. L'anticorps a été marqué isotopiquement sur un acide aminé comportant un groupement méthyle à la fois. Les acides aminés utilisés pour cela étaient U-[ $^1\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$ ] pour la valine, la thréonine, la leucine et l'isoleucine. La méthionine était U-[ $^1\text{H}$ ], [ $^{13}\text{C}^1\text{H}_3$ ] $\epsilon$  (ref 299146, Sigma Aldrich) et l'alanine U-[ $^1\text{H}$ ], [ $^{13}\text{C}^1\text{H}_3$ ] $\beta$  (ref CLM-117-pk).

Ce marquage isotopique a été utilisé afin d'obtenir une masse homogène pour les résidus marqués. Les échantillons ont été produits en suivant le protocole développé et en ajoutant 200 ou 350 mg/L du résidu marqué pour les isoleucines, leucines, méthionines, thréonines et

## Matériel et méthodes

valines. Un seul échantillon avec un marquage isotopique des alanines a été produit avec un ajout de 300 mg/L d'alanine marquée et 10 mg/L de L-cyclosérine lors de la culture cellulaire.

Avant l'analyse par spectrométrie de masse, l'anticorps a été réduit et déglycosylé en utilisant 12.5 µL de *Rapid PNGase F reducing buffer* (ROCHE)(10 minutes à 75°C) et 1µL de PNGase F (ROCHE) (15 minutes à 50°C) pour des volumes d'échantillon de 50 µL à 0.5 mg/mL d'anticorps.

La détermination de l'incorporation des acides aminés marqués a été effectuée en utilisant un spectromètre de masse LC EXION / MS : X500B (AB Sciex) en infusant directement les échantillons purifiés. Le calcul des taux d'incorporation a été effectué avec la formule suivante :

$$1 - \left( \frac{\text{Masse observée de l'anticorps marqué} - \text{Masse théorique de l'anticorps}}{\text{incrément de masse par résidu marqué} \times \text{nombre de résidus marqués}} \right) \times 100$$

## 2. Matériel et méthodes pour l'expression et le marquage isotopique du fragment Fab en système *cell-free*

### 2.1. Production du fragment Fab en système *cell-free*

La chaîne légère (LC – résidus 1 à 214) et la chaîne lourde (HC – résidus 1 à 227) de l'anticorps anti-LAMP1 (développé par Sanofi) ont été clonées séparément dans des plasmides pIVEX 2.4d et pIVEX 2.3d contenant respectivement, en N-terminal, un tag 6-His (LC) et un tag Twin-Strep (HC) suivis d'un site de clivage spécifique à la protéase TEV. Ces deux chaînes ont été exprimées *in vitro* en système *cell-free* en mode dialyse[144]. Les plasmides ont été ajoutés à une concentration totale de 16 µg/mL (*ratio* LC :HC = 1 :2) dans des tubes de dialyse GeBaflex de 3 mL (Euromedex). L'extrait *E. coli* S30 a été préparé selon le protocole de Imbert et al. [145] Les concentrations de tous les réactifs *cell-free* dans le Gebaflex et dans le tampon de dialyse *cell-free* sont résumées dans le tableau 2.2.1. Pour permettre la formation et l'isomérisation des ponts disulfures, l'extrait S30 initialement utilisé a été préparé à partir de la souche *E. coli* SHuffle exprimant l'isomérase à pont disulfures DsbC. Par la suite, la souche BL21 (DE3) a été utilisée et de la DsbC recombinante purifiée a été ajoutée dans le mélange réactionnel à une concentration de 15µM, ce qui s'est révélé plus efficace. Les réactions *cell-free* ont été réalisées en présence de glutathion réduite (GSH) et oxydée (GSSG) à des concentrations respectives de 0.9 et 3.6 mM. Les GeBaflex ont été placés dans des tubes de centrifugation de 50 mL contenant 30 mL de tampon de dialyse *cell-free*, puis dans des tubes de 25 mL avec 15 mL de tampon de dialyse *cell-free* après optimisation du *ratio* du volume d'échantillon par rapport au volume de tampon de dialyse. Ce tampon de dialyse contenait, entre autres (tableau 2.2.1), les 20 acides aminés à une concentration de 1mM chacun. Le tableau 2.2.2 liste les acides aminés non marqués qui ont été utilisés. Les réactions *cell-free* ont été incubées 16h à 30°C en agitation douce (20 rpm dans un four d'hybridation (Techne)).

## Matériel et méthodes

Tableau 2.2.1 : Liste des réactifs cell-free et de leur concentration. Les réactifs en noir sont présent dans le tube et dans le tampon de dialyse. Les réactifs en rouge ne sont présent que dans le tube de dialyse.

Réactifs	Concentrations	Fournisseur	Référence
rCTP	0,8 mM	Jena	NU-1011-CSTM
rGTP	0,8 mM	Jena	NU-1011-CSTM
rUTP	0,8 mM	Jena	NU-1011-CSTM
rATP	1,2 mM	Jena	NU-1011-CSTM
HEPES	55 mM	Sigma	H4034
Acide folinique	68 µM	Sigma	47612
AMP cyclique	0,64 mM	Sigma	A6885
Spermidine	2mM	Sigma	S2626
Créatine phosphate	80 mM	Sigma	27920
Acétate d'ammonium	177,5 mM	Sigma	09689
Bétaïne	15 mM	Sigma	B3501
Glutathionne réduite	0,9 mM	Sigma	3541
Glutathionne oxydée	3,6 mM	Sigma	G4376
Acétate de magnésium	14 µM	Sigma	M5661
Acides aminés	1 mM	Voir tableaux 2.2.2 ; 2.2.3 ; 2.2.4 ; 2.2.5	
<b>ARNt</b>	<b>0,175 mg/mL</b>	Roche Sigma	10109541001
<b>Créatine kinase</b>	<b>0,25 mg/mL</b>	Roche Sigma	10127566001
<b>ARN T7 polymérase</b>	<b>1% (v/v)</b>	Fait maison	/
<b>Plasmide chaine légère</b>	<b>4,97 µg/mL</b>	ROCHE	pIVEX2.4d
<b>Plasmide chaine lourde</b>	<b>11,03 µg/mL</b>	ROCHE	pIVEX2.3d
<b>DsbC recombinante</b>	<b>15 µM</b>	Fait maison	/
<b>Extrait bactérien S30</b>	<b>40% (v/v)</b>	Fait maison	/

Tableau 2.2.2: Liste des acides aminés non marqués utilisés en système cell-free

Acides aminés	Fournisseur	Référence
L-Alanine	Sigma	05130-25G
L-Arginine Hydrochloride	Sigma	A5131-10G
L-Asparagine	Sigma	11149-25G-F
L-Aspartic Acid	Sigma	11190-50G
L-Cystéine	Sigma	168149-25G
L-Glutamic Acid	Sigma	G1251-100G
L-Glutamine	Sigma	49419-25G
Glycine	Sigma	G7126-100G
L-Histidine hydrochloride	Sigma	H8125-25G
L-Isoleucine	Sigma	I2752-10G
L-Leucine	Sigma	L8000-50G
L-Lysine hydrochloride	Sigma	L5626-100G
L-Methionine	Sigma	M9625-25G
L-Phenylalanine	Sigma	78019-25G
L-Proline	Sigma	81710-10G
L-Serine	Sigma	84959-25G
L-Threonine	Sigma	T8625-10G
L-Tryptophan	Sigma	T0254-25G
L-Tyrosine	Sigma	T3754-50G
L-Valine	Sigma	V0500-25G

## Matériel et méthodes

Le fragment Fab U-[<sup>2</sup>H, <sup>13</sup>C, <sup>15</sup>N] ayant servi à l'attribution de son squelette a été marqué en utilisant des acides aminés U-[<sup>2</sup>H, <sup>13</sup>C, <sup>15</sup>N] (CIL). Pour éviter la reprotonation des C $\alpha$  et C $\beta$  par les transaminases lors de la production du fragment en milieu cell-free protoné (H<sub>2</sub>O), l'extrait S30 a été traité avec du NaBH<sub>4</sub>, du AOA et du D-malate en suivant le protocole de Imbert et al. [145]

Lors du marquage isotopique des groupements méthyles, les acides aminés non marqués sur les groupement méthyles étaient deutéré. Le tableau 2.2.3 résume ceux ayant été utilisé.

Tableau 2.2.3 : Liste des acides aminés marqués utilisés en système cell-free pour l'expression du fragment Fab perdeutééré en système cell-free deutéré.

Acides aminés	Marquage isotopique	Fournisseur	Référence
L-Arginine	U- <sup>2</sup> H, <sup>15</sup> N	Martek isotopes	DNARG
L-Asparagine	2,3,3- <sup>2</sup> H <sub>3</sub>	Martek isotopes	D3ASP
L-Aspartic Acid	U- <sup>2</sup> H, <sup>15</sup> N	Martek isotopes	DNASP
L-Cystéine	U- <sup>2</sup> H, <sup>15</sup> N	Martek isotopes	DNCYS
L-Glutamic Acid	U- <sup>2</sup> H, <sup>15</sup> N	Martek isotopes	DNGLA
L-Glutamine	U- <sup>2</sup> H	Martek isotopes	DGLN
L-Glycine	2,2- <sup>2</sup> H <sub>2</sub>	Eurisotop	DLM-1674-5
L-Histidine	U- <sup>2</sup> H	Martek isotopes	DHIS
L-leucine	U- <sup>2</sup> H, <sup>15</sup> N	Martek isotopes	CNLeu
L-thréonine	U- <sup>2</sup> H, <sup>15</sup> N, <sup>13</sup> C	Martek isotopes	3BTHR
L-Lysine hydrochloride	U- <sup>2</sup> H	Martek isotopes	DLys
L-Phenylalanine	U- <sup>2</sup> H, <sup>15</sup> N	Martek isotopes	DNPHE
L-Proline	50% <sup>2</sup> H, U- <sup>13</sup> C, <sup>15</sup> N	Martek isotopes	CND(50) Pro
L-Serine	U- <sup>2</sup> H	Martek isotopes	DNSER
L-Threonine	U- <sup>2</sup> H	Martek isotopes	DTHR
L-Tryptophan	U- <sup>2</sup> H	Martek isotopes	DTRY
L-Tyrosine	<sup>13</sup> C <sub>2</sub> , <sup>15</sup> N, <sup>2</sup> H	Martek isotopes	3BTYR

## Matériel et méthodes

Le fragment Fab marqué sur les groupements méthyles de manière spécifique ou sur les six résidus portant un groupement méthyle a été produit en utilisant les acides aminés et précurseurs K-Ile et KIV résumés dans le tableau 2.2.4

Tableau 2.2.4 : Liste des acides aminés et précurseurs marqués sur les groupements méthyles synthétisés à façon par NMR-Bio et utilisés pour la production du fragments Fab marqué U-[<sup>2</sup>H], <sup>13</sup>CH<sub>3</sub>-Ala, Met, Thr, Leu-proS, Val-proR, Ile- $\delta$ 1.

Acides aminés Et précurseurs	Marquage isotopique	Fournisseur
Alanine	U-[ <sup>2</sup> H, <sup>15</sup> N]- <sup>13</sup> CH <sub>3</sub>	NMR-Bio
Méthionine	U-[ <sup>2</sup> H, <sup>15</sup> N]- <sup>13</sup> CH <sub>3</sub>	NMR-Bio
Leucine	U-[ <sup>2</sup> H, <sup>15</sup> N]- <sup>13</sup> CH <sub>3</sub> -proS	Aldrich
K-Ile	2-kéto-3-( <sup>12</sup> C <sup>2</sup> H <sub>3</sub> -méthyle)-3,4-( <sup>2</sup> H)-5( <sup>13</sup> C)-pentanoate	NMR-Bio
Thréonine	U-[ <sup>2</sup> H, <sup>15</sup> N]- <sup>13</sup> CH <sub>3</sub>	NMR-Bio
KIV	2-Keto-3-( <sup>2</sup> H)-3-( <sup>12</sup> C <sup>2</sup> H <sub>3</sub> -méthyle)-2,4- <sup>13</sup> (C <sub>2</sub> ) butyrate	NMR-Bio

Le fragment Fab perdeutééré avec un marquage isotopique <sup>13</sup>C linéaire pour les alanines, isoleucines et valines a été produit à partir de l'alanine et des précurseurs K-Ile et KIV marqués qui sont résumés dans le tableau 2.2.5.

Tableau 2.2.5 : Liste des acides aminés et des précurseurs marqués sur les groupements méthyles synthétisés à façon par NMR-Bio pour la production du fragment Fab perdeutééré et avec un marquage <sup>13</sup>C linéaire pour les alanines, isoleucines- $\delta$ 1 et valines-proR.

Acides aminés Et précurseurs	Marquage isotopique	Fournisseur
Alanine	U-[ <sup>2</sup> H, <sup>13</sup> C]- <sup>13</sup> CH <sub>3</sub>	NMR-Bio
K-Ile	2-kéto-3-( <sup>12</sup> C <sup>2</sup> H <sub>3</sub> -méthyle)-3,4-( <sup>2</sup> H <sub>2</sub> )-1,2,3,4,5( <sup>13</sup> C)-pentanoate	NMR-Bio
KIV	2-Keto-3-( <sup>2</sup> H)-3-( <sup>12</sup> C <sup>2</sup> H <sub>3</sub> -méthyle)-1,2,3,4- <sup>13</sup> C <sub>2</sub> butyrate	NMR-Bio

## 2.2. Purification du fragment Fab en système *cell-free*

Les réactions *cell-free* ont ensuite été centrifugées 20 min à 25 000 g puis le fragment Fab anti-LAMP1 a été purifié par deux étapes de chromatographie d'affinité successives, basées sur l'affinité pour le 6-His tag et le Twin-Strep tag. Le protocole est détaillé dans l'article de la partie 3.2.3.

Pour l'échantillon de Fab U-[<sup>2</sup>H,<sup>13</sup>C,<sup>15</sup>N] seulement (ayant servi à l'attribution du squelette), les tags Strep et His6 ont été clivés à la TEV. Le protocole est détaillé dans la publication de la partie 3.2.3.

Ces deux méthodes de purification ont ensuite été remplacées par une purification sur résine de protéine L. Pour cela, après centrifugation de la réaction *cell-free* (20 min à 25000 g), une colonne Hi-Trap protein L de 1 mL a été utilisée sur une chaîne chromatographique NGC (Bio-Rad). La réaction *cell-free* a été diluée par 4 dans un tampon phosphate 20 mM, NaCl 150 mM pH 6.8 (tampon d'équilibrage) et injectée dans la colonne équilibrée avec ce même tampon et à un débit de 1 mL/min. Un premier lavage a été effectué avec 5 CV du tampon d'équilibration, suivi d'un deuxième lavage avec 5 CV de tampon phosphate 20 mM à pH 6.8. L'élution du fragment Fab a été effectuée avec un gradient variant de 0 à 100% en 10 CV de tampon d'élution (glycine 100 mM à pH 3) suivi de 100% de tampon d'élution pendant 10 CV. Le suivi du profil d'élution a été réalisé par lecture de l'absorbance à 280 nm. Des fractions de 1 mL ont été recueillies dans des plaques 96 puits dans lesquels 30 µL de tampon de neutralisation TrisBase 1M à pH 8 ont été ajoutés au préalable.

Les fractions contenant le fragment Fab ont ensuite été déposées sur gel SDS-PAGE à 12.5% de polyacrylamide. Des western blot anti-Fab ont également été réalisés (anti-human IgG (Fab specific)-peroxidase antibody ref A0293, Merck).

Les fractions contenant le Fab ont ensuite été rassemblées et concentrées jusqu'à un volume de 0.5 mL en utilisant un Amicon possédant un *cut-off* de 10 kDa. Puis, les échantillons ont été injectés à 0.5 mL/min sur une colonne de filtration sur gel (Superdex 75 10/300 (Cytiva)) en utilisant une chaîne chromatographique NGC (Bio-Rad). Le tampon utilisé était un tampon MES 50 mM, NaCl 100 mM à pH 6.8. Les fractions contenant l'anticorps ont ensuite été réunies pour être concentrées sur Amicon avec un *cut-off* de 10 kDa jusqu'à un volume de 0.5 mL.

Le calcul de la concentration pour la détermination des rendements de production du fragment Fab a été réalisé par lecture de l'absorbance à 280 nm en utilisant un Nanodrop 2000 (Thermo Fisher) ( $\epsilon = 75\,915\text{ M}^{-1}\cdot\text{cm}^{-1}$  ;  $A = 1.596$  ;  $MW = 47\,552\text{ Da}$  ;  $PI = 8.21$ ).

### 2.3. Production et purification de l'isomérase à pont disulfure DsbC

Des bactéries BL21 (DE3) ont été transformées avec le vecteur pET-28a-c codant pour la DsbC (fourni par NMR-Bio) possédant, en N-terminal, un tag 6-His suivi d'un site de clivage à la TEV. La surexpression de la DsbC a été initiée par induction avec 1 mM d'IPTG pendant 16 h à 20°C. Puis, les bactéries ont été récoltées par centrifugation, lysées par sonication (4°C deux fois 2 minutes à 50%) dans un tampon 50 mM Tris, 500 mM NaCl, pH 7.5, 40 µg/ml DNase, contenant du complete-EDTA-free (ref) et centrifugées pour différencier la fraction soluble des corps d'inclusion. La protéine DsbC a été purifiée à partir de la fraction soluble, en utilisant une colonne d'affinité Ni-NTA sur une chaîne chromatographique NGC (Bio-Rad) préalablement nettoyée de tout composant à caractère RNase avec de 0.5 mM NaOH. Le surnageant du lysat bactérien a été injecté sur la colonne Ni-NTA (Ni-NTA XK home made) équilibrée avec un tampon 50 mM Tris, 500 mM NaCl, 20 mM imidazole, pH 7.5 à un débit de 1 mL/min. Un lavage de 5 CV a été effectué avec le tampon d'équilibration suivi de l'élution de la DsbC par un gradient variant de 20 à 500 mM d'imidazole sur 10 CV.

A la suite de cette purification, les fractions contenant la DsbC ont été rassemblées et concentrées sur AMICON avec un cut-off de 10 kDa jusqu'à un volume de 8 mL. Le tag 6-His de la DsbC a ensuite été clivé avec la TEV avec un ratio de 1/10 (masse de TEV/masse de DsbC) (16 h, 4°C, en dialyse et sous agitation magnétique).

La DsbC clivée a ensuite été purifiée par chromatographie sur colonne NiNTA inversée (récupération de la fraction clivée et du premier lavage de la colonne) puis par chromatographie d'exclusion de taille en utilisant une colonne de filtration sur gel HiLoad 26/600 Superdex 75 (Cytiva) équilibrée dans un tampon 50 mM Hepes, 150 mM NaCl, 5% glycérol, pH 7.5.

Les fractions d'élution contenant la DsbC ont ensuite été concentrées sur AMICON avec un cut-off de 10 kDa jusqu'à une concentration de 200 µM. La DsbC a été conservée à -80°C.

### 2.4. Spectroscopie RMN

Toutes les expériences RMN acquises sur le fragment Fab de l'anticorps anti-LAMP1 ont été enregistrées à 35°C sur des spectromètres équipés d'une sonde cryogénique.

Six expériences 3D BEST-TROSY (HNCA, HN(CO)CA, HN(CA)CB, HN(COCA)CB, HNCO, HN(CA)CO) [143] ont été acquises sur des spectromètres Bruker Avance III HD opérants à des fréquences  $^1\text{H}$  de 600 à 950 MHz pour une durée totale d'acquisition de 18 jours sur un échantillon concentré à 300 µM de fragment Fab U- $^{2}\text{H}$ ,  $^{15}\text{N}$ ,  $^{13}\text{C}$ .

L'attribution des groupements méthyles  $\delta 1$  des isoleucines, des alanines et proR des valines a été réalisée en utilisant des expériences RMN 3D "out and back" HCC, HC(C)C et HC(CC)C [122]. Ces expériences ont été acquises sur un spectromètre RMN équipé d'une sonde cryogénique et opérant à une fréquence  $^1\text{H}$  de 950 MHz sur un échantillon de fragment Fab de 260 µM U- $^{2}\text{H}$ ,  $^{14}\text{N}$ ,  $^{12}\text{C}$ , Ala- $^{2}\text{H}$ ; 1,2- $^{13}\text{C}_2$ ; [ $^{13}\text{C}^1\text{H}_3$ ] $\beta$ ], Ile- $^{2}\text{H}_4$ ; 1,2,3,4- $^{13}\text{C}_4$ ; [ $^{13}\text{C}^1\text{H}_3$ ] $\delta 1$ /[ $^{12}\text{C}^2\text{H}_3$ ] $\gamma 2$ ], Val- $^{2}\text{H}_2$ ; 1,2,3- $^{13}\text{C}_3$ ; [ $^{13}\text{C}^1\text{H}_3$ Pro-R]/[ $^{12}\text{C}^2\text{H}_3$ Pro-S].

L'attribution des groupements méthyles en utilisant les NOEs inter-méthyles a été réalisée en utilisant une expériences RMN 3D CCH HMQC-NOESY-HMQC [142], [146]. Cette expérience a été acquise sur un spectromètre RMN équipé d'une sonde cryogénique et

opérant à une fréquence  $^1\text{H}$  de 950 MHz sur un échantillon de fragment Fab de 330  $\mu\text{M}$  U- $^{2}\text{H}$ ,  $^{14}\text{N}$ ,  $^{12}\text{C}$ ], Ala- $^{13}\text{C}^1\text{H}_3$ ] $\beta$ , Ile- $^{13}\text{C}^1\text{H}_3$ ] $\delta 1$ , Val- $^{13}\text{C}^1\text{H}_3$ ] $\text{proR}$ , Thr- $^{13}\text{C}^1\text{H}_3$ ] $\gamma$ , Met- $^{13}\text{C}^1\text{H}_3$ ] $\xi$ , Leu- $^{13}\text{C}^1\text{H}_3$ ] $\text{proS}$ .

Les données RMN ont été processées en utilisant Topspin 4.1.4 (Bruker) ou nmrPipe/nmrDraw [147], CcpNmr v3.0.4 [148], et MethylFLYA[149].

Les expériences RMN acquises pendant la thèse sont résumées dans le tableau 2.2.6.

Tableau 2.2.6 : Liste des expériences RMN acquise avec la fréquence  $^1\text{H}$  du spectromètre utilisé et le temps d'acquisition de l'expérience

Expériences RMN		Champs $B_0$	Durée d'acquisition
Best-TROSY	2D $^1\text{H}$ - $^{15}\text{N}$	950 MHz	30 min
	HNCO	700 MHz	1 jour
	HN(CA)CO	700 MHz	3,75 jours
	HNCA	950 MHz	2 jours
	HN(CO)CA	600 MHz	2 jours
	HNCACB	950 MHz	4 jours
	HN(COCA)CB	600 MHz	5,5 jours
HCC "out and back"	HCC	950 MHz	2 jours
	HC(C)C	950 MHz	4 jours
	HC(CC)C	950 MHz	6 jours
Methyl-SOFAST		950 MHz	1 heure
3D CCH HMQC-NOESY-HMQC		950 MHz	9 jours

## 2.5. Attribution du squelette et des groupements méthyles et dépôt des données

L'attribution séquentielle du squelette de 89 % de  $^1\text{H}_\text{N}$ , 89 % de  $^{15}\text{N}_\text{H}$ , 92 % de  $^{13}\text{CO}$ , 93 % de  $^{13}\text{C}_\alpha$  et 89 % de  $^{13}\text{C}_\beta$  fait partie des résultats présentés dans l'article de la partie 3.2.3. Des expériences « out and back » HCC, HC(C)C et HC(CC)C ont permis de connecter des résonances du squelette précédemment attribuées aux alanines- $\beta$ , valines-proR et isoleucines- $\delta 1$ . 53 %, 43 % et 33 % des alanines- $\beta$ , des valines-pro-R et des isoleucines- $\delta 1$  ont pu être reliées sans ambiguïté aux résonances C $\alpha$  et C $\beta$  attribuées au préalable.

Ces premières résonances des groupes méthyles attribuées, ainsi que le type d'acide aminé correspondant à chaque signal méthyle, la structure 3D du fragment fab Anti-Lamp1 et les 434 connectivités inter-méthyles NOE enregistrées sur le fragment Fab anti-LAMP1 U- $[\text{}^2\text{H}, \text{}^{14}\text{N}, \text{}^{12}\text{C}]$ , Ala- $[\text{}^{13}\text{C}^1\text{H}_3]_\beta$ , Ile- $[\text{}^{13}\text{C}^1\text{H}_3]_{\delta 1}$ , Val- $[\text{}^{13}\text{C}^1\text{H}_3]_{\text{Pro-R}}$ , Thr- $[\text{}^{13}\text{C}^1\text{H}_3]_\gamma$ , Met- $[\text{}^{13}\text{C}^1\text{H}_3]_\xi$ , Leu- $[\text{}^{13}\text{C}^1\text{H}_3]_{\text{Pro-S}}$  ont été donnés comme entrée pour le logiciel MethylFLYA[149]. Le logiciel automatisé d'attribution des groupes méthyles a été exécuté avec des seuils de distance de 5,5/6,0 et 6,5 Å et 100 passages par seuil de distance. 10 groupes méthyles supplémentaires ont été attribués avec une grande confiance grâce à un score d'attribution élevé donné par MethylFLYA et, en outre, la résonance attribuée correspondait à des résidus précédemment attribués de manière ambiguë à l'aide d'expériences HCC « out and back ».

Le programme MethylFLYA a été relancé avec les groupements méthyles nouvellement attribués et combiné à une analyse manuelle minutieuse du spectre NOESY pour attribuer 100 % des isoleucines- $\delta 1$ , 100 % des alanines- $\beta$ , 100 % des valines-proR, 100 % des méthionines- $\xi$ , 96 % des leucines-proS et 40 % des thréonines- $\gamma$ .

### III. Résultats

## 1. Marquage isotopique de l'anticorps anti-LAMP1 en cellules CHO

### 1.1. Introduction au marquage isotopique en cellules eucaryotes

#### 1.1.1. Marquage isotopique $^{15}\text{N}$ et $^{13}\text{C}$ de protéines exprimées en cellules eucaryotes pour la RMN

Le marquage isotopique de protéines est réalisé principalement en bactérie. Ce système d'expression est en effet bien adapté puisque les bactéries sont capables de croître en milieu de culture minimum. De plus, les bactéries sont capables de synthétiser leurs propres acides aminés à partir de composés simples comme le glucose et un sel d'ammonium. Cela est un avantage pour le marquage isotopique uniforme en bactérie car l'utilisation de  $^{13}\text{C}_6$ -glucose et de  $^{15}\text{NH}_4\text{Cl}$  permet d'obtenir un taux d'incorporation des isotopes proches de 100% pour tous les acides aminés à coût réduit. En effet, les composés simples comme le glucose ou le sel d'ammonium marqués sont peu coûteux.

Cependant, comme expliqué dans la partie 1.3.3.3, les bactéries sont limitées pour la production de certaines protéines. Pour les protéines thérapeutiques comme les anticorps qui sont des protéines riches en ponts disulfures et dont la présence de glycosylations est nécessaire, les cellules eucaryotes sont plus appropriées. Ainsi, il s'avère parfois nécessaire d'utiliser les cellules eucaryotes pour produire des protéines plus complexes (ponts disulfures, PTM, glycosylations) marquées isotopiquement, pour pouvoir étudier leur structure en RMN.

Le marquage isotopique d'une protéine peut être uniforme ou spécifique, sur le squelette ou les chaînes latérales des acides aminés (groupements méthyles et autres), avec ou sans deutération. Le choix du type de marquage dépend de la protéine et du projet d'étude et du type de spectre RMN que l'on souhaite acquérir. Chaque système d'expression présente ainsi ses avantages et inconvénients dans le choix du type de marquage isotopique.

#### a Le marquage isotopique $^{15}\text{N}$ , $^{13}\text{C}$ uniforme de protéine en levures

Pour obtenir des protéines correctement repliées avec des modifications post-traductionnelles ou des profils de glycosylations proches d'un organisme humain, les cellules eucaryotes supérieures (mammifères ou insectes) sont plus appropriées. Cependant, ces cellules ne sont pas aussi adaptées pour le marquage isotopique que les cellules eucaryotes inférieures comme les levures. En effet, les levures sont proches des bactéries dans le sens où elles sont capables de croître dans des milieux de culture dits "minimum" qui sont adaptés pour la production de protéines marquées isotopiquement. Ces milieux ont une composition définie en sources de carbone et d'azote pouvant servir à la biosynthèse de tous les acides aminés par les levures.

Ainsi, le marquage isotopique uniforme  $^{15}\text{N}$ ,  $^{13}\text{C}$  de protéines en levures est simplement réalisé en utilisant du  $^{13}\text{C}$ -méthanol,  $^{13}\text{C}$ -glycérol ou  $^{13}\text{C}$ -glucose comme source de carbone marqué, et du  $^{15}\text{NH}_4\text{Cl}$  ou du  $(^{15}\text{NH}_4)_2\text{SO}_4$  comme source d'azote marqué. Cela en fait le système eucaryote le plus économique pour le marquage isotopique uniforme de protéines (55€/g de  $^{15}\text{NH}_4\text{Cl}$  et 119€/g de  $^{13}\text{C}$ -glucose<sup>2</sup>).

Le marquage isotopique  $^{15}\text{N}$ ,  $^{13}\text{C}$  uniforme de protéines en levure a été introduit par Laroche *et al.* en 1994[150]. Ils ont produit, en *Pichia Pastoris*, le peptide anticoagulant des tiques (TAP) uniformément marqué  $^{15}\text{N}$ ,  $^{13}\text{C}$  et  $^{15}\text{N}/^{13}\text{C}$  grâce à l'utilisation de sulfate d'ammonium marqué ( $(^{15}\text{NH}_4)_2\text{SO}_4$ ), de D-glucose- $^{13}\text{C}_6$  et de méthanol- $^{13}\text{C}$ . Cela leur a permis de déterminer la structure du peptide TAP en utilisant la RMN en solution.

## **b Les stratégies de marquage isotopique uniforme en cellules eucaryotes supérieures**

Bien que le marquage isotopique soit plus efficace pour les cellules eucaryotes inférieures, il est parfois plus intéressant ou nécessaire de synthétiser certaines protéines, comme les protéines membranaires ou les protéines glycosylées, dans des systèmes eucaryotes supérieurs. Ces cellules, contrairement aux levures et bactéries, ont perdu la capacité de synthétiser certains acides aminés, dits dans ce cas essentiels, qui doivent alors être ajoutés dans les milieux de culture.

Deux stratégies existent pour le marquage isotopique uniforme  $^{15}\text{N}$ ,  $^{13}\text{C}$  des protéines en cellules eucaryotes supérieures. La méthode la plus directe est d'ajouter chacun des acides aminés marqués sous leur forme pure. Cependant, le coût de cette méthode est très élevé et non envisageable pour un marquage isotopique uniforme, elle est plus adéquate lors d'un marquage isotopique spécifique. Certains milieux commerciaux isotopiquement marqués, basés sur cette stratégie, existent, tels que le milieu Bioexpress-2000 pour les cellules d'insectes ou le milieu Bioexpress-6000 pour les cellules de mammifères (CIL), mais à des prix très élevés. Pour exemple, le milieu Bioexpress-2000 est vendu à 10 k\$/L ou 20 k\$/L respectivement pour un marquage  $^{15}\text{N}$  ou  $^{15}\text{N}$  et  $^{13}\text{C}$ [151].

La deuxième méthode, qui permet de réduire le coût d'expression de protéines marquées  $^{15}\text{N}$  et  $^{13}\text{C}$ , est d'utiliser des extraits d'algues ou de levures marqués, qui permettent d'atteindre des taux d'incorporation élevés à des coûts compatibles avec la recherche[151], [152].

En 1992, Hansen *et al.* ont développé, une méthode permettant d'obtenir des extraits d'algues marqués uniformément  $^{15}\text{N}$  et  $^{13}\text{C}/^{15}\text{N}$ [153]. Cet extrait, incorporé ensuite dans des cultures de cellules SP2/0, leur a permis d'obtenir de l'urokinase uniformément marquée  $^{15}\text{N}$  et  $^{13}\text{C}/^{15}\text{N}$ . Cette méthode consiste à produire un extrait d'algues ayant été cultivées en présence de  $^{15}\text{NH}_4\text{Cl}$  pour la source d'azote marquée, et en  $^{13}\text{CO}_2$  pour la source de carbone marqué. Ces algues, peu coûteuses à produire, sont ensuite hydrolysées à pH acide et à chaud (4 M d'acide méthanesulfonique à 115°C pour le protocole de Hansen *et al.*) pour obtenir des extraits d'algues contenant, entre autres, des acides aminés marqués isotopiquement. Cette hydrolyse entraîne la dégradation de certains acides aminés et le milieu de culture final doit être complété avec de l'Asn, du Trp, de la Gln et de la Cys marqué  $^{15}\text{N}$ . Cette méthode leur a permis d'obtenir

---

<sup>2</sup> Prix catalogue au 04/07/2023 pour Cortec Net

la première protéine marquée uniformément en  $^{15}\text{N}$  et exprimée en cellules de mammifères pour des analyses RMN[153].

Les extraits d'algues sont aujourd'hui commercialement disponibles (sous les noms de Celtone<sup>®</sup> pour CIL (350 €/g) et Isogro<sup>®</sup> pour Merck (827€/g)<sup>3</sup>) avec plusieurs types de marquages isotopiques :  $^{15}\text{N}$  ;  $^{15}\text{N}$ - $^2\text{H}$  ;  $^{13}\text{C}$  ;  $^{15}\text{N}$ - $^{13}\text{C}$  ;  $^{15}\text{N}$ - $^{13}\text{C}$ - $^2\text{H}$ . Cela rend l'utilisation d'extraits d'algues marqués très accessible pour le marquage isotopique uniforme de protéines exprimées en cellules eucaryotes. Cependant, l'utilisation de ces extraits ne suffit pas pour obtenir un marquage uniforme en raison des acides aminés non-essentiels des cellules. C'est pour cela que le marquage isotopique uniforme  $^{15}\text{N}$ ,  $^{13}\text{C}$  est réalisé en utilisant un extrait marqué, complété avec du glucose  $^{13}\text{C}$  et du  $^{15}\text{NH}_4\text{Cl}$  qui sont utilisés par les cellules pour la biosynthèse des acides aminés marqués. Une autre solution est d'utiliser les acides aminés non-essentiels marqués avec des inhibiteurs de transaminases afin de stopper la biosynthèse de l'acide aminé non marqué et ainsi d'éviter la dilution isotopique de l'acide aminé marqué.

Les extraits de levure sont une autre source d'acides aminés marqués. En 2009, Egorova-Zachernyuk *et al.* ont développé un protocole de production d'un extrait de levures (*Pichia pastoris*) marquées  $^{15}\text{N}$  et  $^{13}\text{C}$ [154]. Les levures sont cultivées dans des milieux comprenant du  $^{15}\text{NH}_4\text{Cl}$  comme source d'azote et du glucose- $^{13}\text{C}$  comme source de carbone. L'hydrolyse est ensuite réalisée de façon enzymatique, moins agressive que l'hydrolyse acide, et, contrairement aux extraits d'algues, seule la glutamine doit être ajoutée au moment de l'expression. Ces extraits de levures sont cependant moins disponibles commercialement et plus coûteux que les extraits d'algues et les rendements d'expression, moins bons[152].

### c Exemples de marquage isotopique $^{15}\text{N}$ , $^{13}\text{C}$ des protéines en cellules eucaryotes supérieures

Un exemple de l'utilisation d'un extrait d'algues marqué pour une expression en cellules CHO a été illustré par Lustbader *et al.* en 1996 pour la production de l'hormone chorionique gonadotrope humaine (hCG) marquée U-( $^{13}\text{C}$  ;  $^{15}\text{N}$ ) [155]. La protéine hCG est une protéine glycosylée présente dans le placenta des femmes enceintes qui peut également être présente dans certaines cellules cancéreuses. Pour produire la protéine marquée, un milieu de culture pour cellules CHO (CHO-S-SFM-I ; Life Technologies) déplété en sérum et en acides aminés et un mix d'acides aminés U-( $^{13}\text{C}$ ,  $^{15}\text{N}$ ) (Celtone ; Martek) supplémenté avec les acides aminés Cys, Arg et Gln ont été utilisés. De plus, ont été ajoutés du glucose  $^{13}\text{C}$  et du pyruvate de sodium  $^{13}\text{C}$  pour éviter la dilution isotopique de l'alanine, du glutamate, et de l'aspartate. En effet, ces acides aminés sont non essentiels pour les cellules de mammifères et peuvent être synthétisés à partir du glucose qui se transforme en pyruvate (qui conduit à l'alanine), puis en  $\alpha$ -keto-glutarate (glutamate), et à l'oxoacetate (aspartate). L'objectif de Lustbader *et al.* a été d'utiliser la RMN pour étudier la structure de la protéine en présence des glycosylations. La détermination du taux d'incorporation des isotopes dans la séquence de la protéine réalisée par spectrométrie de masse a révélé un marquage uniforme avec un rendement d'incorporation de

---

<sup>3</sup> \*les prix sont donnés pour du Celtone<sup>®</sup> et de l'Isogro<sup>®</sup> marqué U-[ $^{15}\text{N}$ ,  $^{13}\text{C}$ ]. Prix catalogues de CIL et Sigma au 20/05/2023.

90% qui leur a permis d'acquérir des expériences RMN 3D de type HNCA présentant 95 % des résonances attendues pour les connections inter-résidus[155].

Pour les extraits de levures, seule la glutamine est dégradée pendant l'hydrolyse et doit être ajoutée pendant la culture[153]. Pour déterminer les coûts de production, Meola *et al.* ont démontré, en 2014, qu'il était possible de remplacer la glutamine marquée par un sel d'ammonium marqué pour le marquage isotopique  $^{15}\text{N}$  des protéines en cellules d'insectes[156]. Leur travail est basé sur le principe que la glutamine n'est pas un acide aminé essentiel pour les cellules d'insectes, et qu'un ajout de  $^{15}\text{NH}_4\text{Cl}$  est suffisant pour que les cellules synthétisent de la glutamine  $^{15}\text{N}$ . En utilisant un milieu de culture pour cellules d'insectes déplété en acides aminés et complété en extrait de levure marqué  $^{15}\text{N}$  ainsi que du  $^{15}\text{NH}_4\text{Cl}$ , ils ont pu obtenir des ectodomains de protéine membranaire uniformément marquée  $^{15}\text{N}$  en cellules d'insectes D. S2.

### **1.1.2. Deutération uniforme des protéines en cellules eucaryotes**

#### **a Deutération uniforme des protéines en levures**

Comme les bactéries, les levures peuvent s'adapter à la présence de  $\text{D}_2\text{O}$  après une étape d'adaptation en milieu deutéré à 70% ou même 90%, ce qui en fait un organisme de choix pour la deutération. Clark *et al.* l'ont illustré en produisant la protéine membranaire GPCRs (*G protein-coupled receptors*) dans *Pichia pastoris*[157]. Après une adaptation des cellules dans un milieu de culture préparé avec 90 % de  $\text{D}_2\text{O}$ , les levures sont transférées dans un milieu constitué de 100 % de  $\text{D}_2\text{O}$ . Le pourcentage de deutération peut être augmenté en utilisant du glucose ou du glycérol deutéré dans le milieu, mais cela a pour effet d'augmenter le coût de production. Ce protocole (en utilisant du  $^2\text{H}_8$ -glycérol) leur a permis d'obtenir la GPCRs deutérée à 90%.

Bien qu'il soit possible d'obtenir le même résultat en cellules d'algues, ce système est moins utilisé pour la production de protéines. En effet, ce système est très différent des organismes eucaryotes supérieurs, il est donc plus intéressant d'utiliser les levures ou les cellules eucaryotes supérieures pour cela. Cependant, tout comme pour le marquage  $^{15}\text{N}$  et  $^{13}\text{C}$ , la biomasse d'algue marquée est utilisée pour la production d'extraits d'acides aminés deutérés pour les cellules eucaryotes supérieures (partie suivante).

#### **b Inconvénient de la deutération uniforme des protéines en milieu perdeutééré**

La production de protéines en système perdeutééré, comme il est possible de faire en bactéries ou en levures, présente un inconvénient lors d'un marquage isotopique  $^{15}\text{N}$ ,  $^{13}\text{C}$ ,  $^2\text{H}$ . En effet, la production d'une protéine en milieu perdeutééré mène à une protéine dont tous les groupements amides sont deutérés. Or, pour l'acquisition d'expériences RMN basées sur l'analyse du squelette des protéines, les groupements amides ont besoin d'être protonés. Cela est possible dans un solvant protoné (100% de  $^1\text{H}_2\text{O}$ ), mais seulement pour les groupements amides situés à la surface de la protéine. En effet, ceux situés au cœur des protéines ont des

cinétiques d'échange beaucoup plus lente et certains peuvent ne pas s'échanger du tout. La solution est alors de dénaturer la protéine pour procéder à l'échange de tous les groupements amides dans un tampon protoné. Cependant, cette étape implique que la protéine puisse ensuite être repliée, et, quand c'est possible, le rendement s'en trouve souvent affecté.

### c Deutération uniforme des protéines en cellules eucaryotes supérieures

Contrairement aux levures et aux algues, les cellules eucaryotes supérieures sont incapables de croître dans des milieux de culture perdeutérés. En effet, pour les cellules d'insectes et de mammifères, le D<sub>2</sub>O est toxique et conduit à la mort cellulaire. Cependant, il est possible de deutérer partiellement des protéines en cellules d'insectes cultivées en solvant protoné, en utilisant des extraits d'algues ou de levures perdeutérés [151], [152].

Opitz *et al.* et Sitarska *et al.* ont développé des protocoles de culture de cellules d'insectes en présence d'extraits de levures et d'algues perdeutérés [151], [152]. Le D<sub>2</sub>O est toxique pour les cellules d'insectes, mais l'utilisation d'un mix d'acides aminés perdeutérés est possible. Cependant, la perdeutération est impossible de cette manière pour plusieurs raisons. En effet, certaines voies métaboliques des cellules d'insectes permettent d'échanger les protons en position C $\alpha$  des acides aminés non essentiels et notamment le glutamate, l'aspartate et la glycine qui sont les plus concernés. Étant non-essentiels, ces acides aminés sont aussi impactés par un phénomène de dilution isotopique car les cellules synthétisent la forme non-marquée de ces acides aminés à partir du glucose et du solvant protoné. Ces échanges sont réalisés par des transaminases capables d'échanger les protons en position C $\alpha$  des acides aminés avec les protons du solvant. Deuxièmement, la deutération des acides aminés alanine, glutamine et aspartate peut être diluée si le glucose utilisé lors de la culture est protoné [151], [152]. Ces sources de dilutions isotopiques entraînent une hétérogénéité de la deutération en position C $\alpha$  des acides aminés, et cela peut provoquer l'apparition d'un deuxième pic sur les spectres RMN. En effet, le déplacement chimique du carbone n'est pas le même s'il est lié à un atome d'hydrogène ou de deutérium, on parle de déplacement chimique isotopique (différence d'environ 0.3 ppm par deutérium lié au carbone). Cela limite l'utilisation d'expériences RMN 3D.

Malgré ces sources de dilution isotopique, les extraits de levures et d'algues perdeutérés permettent d'obtenir des taux de deutération respectifs d'environ 60% et 76% en cellules d'insectes [151], [152]. Cette méthode présente un avantage par rapport aux systèmes permettant de produire des protéines perdeutérées en solvant 100 % deutéré comme pour les levures (cf. partie 6.3.2) puisque lors d'un marquage avec des acides aminés triplement marqués (<sup>15</sup>N, <sup>13</sup>C, <sup>2</sup>H), la production en solvant protoné permet d'obtenir une protéine 100% protonée sur les groupements amide, même si les acides aminés sont deutérés. A. Grahl *et al.* ont illustré cela en produisant le récepteur  $\beta_1$ -adrenergique marqué <sup>15</sup>N à 90% et <sup>2</sup>H à 60% en cellules d'insectes (Sf9) [158]. Pour cela, ils ont exprimé la protéine en cellules d'insectes cultivées dans un milieu de culture dépourvu d'acides aminés et complété avec un extrait de levure marqué <sup>15</sup>N, <sup>2</sup>H, et de la glutamine <sup>15</sup>N. Ce marquage isotopique leur a permis d'étudier la dynamique de la protéine seule et en interaction avec des ligands par RMN des groupements amides.

Malheureusement, pour les protéines de haut poids moléculaire, ces taux de deutération ne sont pas suffisants pour l'acquisition de triples résonances nécessaires à l'attribution des spectres RMN  $^1\text{H}$ - $^{15}\text{N}$ . Pour cela, un système d'expression permettant la perdeutération de la protéine doit être utilisé, ou alors, la RMN des groupements méthyles doit être envisagée.

### **1.1.3. Marquage spécifique des groupements méthyles des protéines en cellules eucaryotes**

#### **a Marquage isotopique des protéines sur les groupements méthyles**

Comme expliqué dans la partie 4.3.2 b, la RMN des groupements méthyles offre un meilleur rapport signal/bruit que la RMN des groupements amides du squelette. Cependant, pour les protéines de haut poids moléculaire devant être exprimées en cellules eucaryotes, la difficulté est d'obtenir un rendement de marquage isotopique élevé sur les groupements méthyles en même temps qu'un pourcentage de deutération élevé de la protéine. Le marquage isotopique des groupements méthyles peut être réalisé en utilisant les six acides aminés marqués. Cependant, le coût de ces isotopes est très élevé et rend ce type de marquage, selon le rendement de la protéine à étudier, souvent inabordable. En bactérie, l'utilisation de précurseurs d'acides aminés marqués a été développée afin de rendre le marquage isotopique des groupements méthyles financièrement plus accessible. En effet, les bactéries peuvent synthétiser tous les acides aminés. Ainsi, il est possible de marquer les groupements méthyles des isoleucines, valines et leucines en utilisant de l' $\alpha$ -ketobutyrate (Ile) et de l'acétolactate ou de l' $\alpha$ -ketoisovalerate (Leu et Val). Les groupements méthyles des alanines, méthionines et thréonines doivent être marqués en utilisant directement l'acide aminé.

Cette méthode est possible en bactérie grâce à l'absence de compartimentation du métabolisme. En effet, tout a lieu dans le cytoplasme et les précurseurs sont facilement accessibles aux enzymes de biosynthèse des acides aminés. Pour les cellules eucaryotes qui sont des organismes plus complexes, la compartimentation du métabolisme est un problème car les précurseurs doivent atteindre les mitochondries où se situent les enzymes de biosynthèse des acides aminés[159]. De plus, pour les cellules eucaryotes supérieures, la plupart des acides aminés sont essentiels, ce qui signifie que les enzymes permettant leur biosynthèse n'existent pas et l'utilisation de précurseurs est alors impossible. Ainsi, en cellules eucaryotes, il est plus difficile voire impossible de marquer isotopiquement les groupements méthyles des protéines en utilisant des précurseurs marqués.

#### **b Marquage isotopique des groupements méthyles dans les levures**

En levure, l'utilisation d'un précurseur marqué pour le marquage des groupements méthyles n'a été montré que pour l'isoleucine. En produisant l'actine (protéine du cytosquelette des cellules eucaryotes) en levures (*P. pastoris*), qui ne peut pas être exprimée en bactérie, dans un milieu fortement deutéré (90%), Clark *et al.* ont démontré qu'il n'était possible de marquer que 50 % des groupements méthyles  $\delta 1$  des isoleucines en utilisant de l' $\alpha$ -ketobutyrate marqué  $^{13}\text{CH}_3$ . Leurs travaux ont consisté à optimiser le protocole de culture cellulaire afin d'augmenter

les rendements d'expression de la protéine d'intérêt dans le but d'utiliser moins de précurseurs pour le marquage de l'isoleucine en position  $\delta 1$ . Cela leur a permis de réduire de 50% le coût de production de l'actine et du récepteur A<sub>2A</sub> marqués sur les isoleucines- $\delta 1$  et exprimés en levure.

Pour les valines et les leucines, l' $\alpha$ -ketoisovalérate, précurseur situé en amont dans la voie de biosynthèse, a été utilisé pour le marquage, mais cela a conduit à des taux d'incorporation de seulement 3% dans les levures [159]. L'hypothèse avancée a été que les levures ne possèdent pas les transporteurs pour ce précurseur et qu'il ne peut donc pas intégrer le cytosol des cellules [160]. Pour un marquage isotopique optimal de ces groupements méthyles, l'acide aminé final doit dans ce cas être utilisé.

Par la suite, Suzuki *et al.* ont montré, en 2018, qu'il était possible d'augmenter le transport de l' $\alpha$ -ketoisovalérate dans le cytosol des levures en optimisant le pH de la culture pour la production de la  $\beta$ -lactoglobuline bovine [161]. En diminuant le pH de la culture de 6 à 3.5, et en utilisant de l' $\alpha$ -ketoisovalérate U- $^{2}\text{H}$ , [ $^{13}\text{C}^1\text{H}_3$ ], ils ont montré que le marquage isotopiques des valines pouvait atteindre 74%. Cette nette augmentation n'a cependant pas été observées pour les leucines qui n'ont été marquées qu'à 10%, ceci pouvant s'expliquer par le fait que la biosynthèse de la leucine fait intervenir plus d'étapes que celle de la valine à partir du même précurseur (l' $\alpha$ -ketoisovalérate).

### **c Deutération sélective et marquage isotopique des groupements méthyles dans les cellules d'insectes et de mammifères**

Pour les cellules d'insectes et de mammifères, le marquage spécifique des groupements méthyles ne peut pas être réalisé en utilisant des précurseurs. Ces systèmes d'expression hétérotrophes nécessitent l'ajout des acides aminés marqués  $^{13}\text{CH}_3$  lors de l'expression. C'est pour cela que les cellules d'insectes et de mammifères sont les systèmes les plus coûteux pour le marquage isotopique des groupements méthyles. Un autre inconvénient pour ces systèmes est que la perdeutération est impossible en raison de la toxicité du D<sub>2</sub>O pour ces cellules. Pour permettre l'étude de protéines ayant des masses moléculaires toujours plus élevées en utilisant la RMN des groupements méthyles, une autre méthode de deutération a été développée.

En RMN, l'impact des interactions dipolaires entre protons sur le phénomène de relaxation est proportionnel à  $1/r^6$  ou  $r$  est la distance entre les deux noyaux (d'hydrogènes  $^1\text{H}$  dans notre cas). D'après cette définition, on déduit que les protons impactant le plus la relaxation sont ceux situés le plus près du groupement chimique étudié. Ainsi, pour un groupement méthyle, les protons les plus proches dans l'espace sont ceux situés sur l'acide aminé portant le groupement méthyle ou sur des acides aminés proches dans l'espace (au sein des cœurs hydrophobes des protéines). Ainsi, les méthodes développées pour contrer la toxicité du D<sub>2</sub>O ont consisté à ne deutérer que les protons les plus proches des groupements méthyles étudiés, c'est-à-dire ceux des chaînes latérales des acides aminés portant le groupement méthyle, ou ceux des autres acides aminés proches dans l'espace.

L'avantage du marquage spécifique des groupements méthyles et de la deutération sélective des résidus proches dans l'espace a été démontré par Kofuku *et al.* en 2014 [162] par le marquage isotopique des groupements méthyles des méthionines. Pour cet acide aminé, le groupement méthyle est coupé du reste des protons de la chaîne latérale par la présence de l'atome de soufre. Ainsi, les protons les plus proches voisins du groupement méthyle ne sont pas forcément ceux de la chaîne latérale mais ceux des autres résidus à proximité. Ils ont déterminé la nature des résidus proches dans l'espace des méthionines d'intérêt en étudiant la structure cristallographique de leur protéine, le récepteur  $\beta$ 2-adrenergic ( $\beta$ 2AR), pour sélectionner ceux devant être deutérés. En deutérant sélectivement dix types d'acides aminés (ala, cys, ile, leu, met (excepté sur le méthyl), phe, thr, val, tyr et trp), le signal des méthionines étudiées a été multiplié par un facteur cinq (figure 3.1.1). Ils ont ensuite appliqué la même stratégie pour le marquage spécifique des groupements méthyles des alanines en utilisant de la L-cyclosérine pour inhiber la synthèse d'alanine endogène, qui aurait pour effet de diluer le marquage isotopique [163].

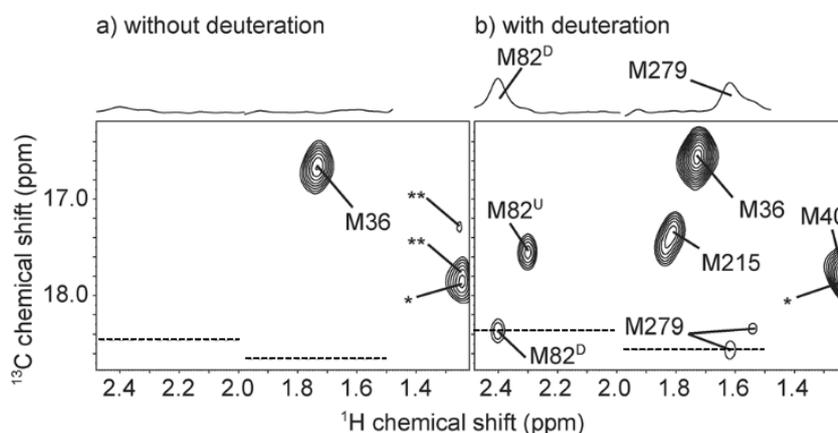


Figure 3.1.1 : Spectres HMQC  $^1\text{H}$ - $^{13}\text{C}$  des groupements méthyles des méthionines de la protéine  $\beta$ 2AR marquée [ $\alpha\beta$ - $^2\text{H}$ , methyl- $^{13}\text{C}$ -met] a), et marquée [ $^2\text{H}$ -9AA,  $\alpha\beta\gamma$ - $^2\text{H}$ , methyl- $^{13}\text{C}$ -met]. Extrait de Y. Kofuku *et al.*, *Angewandte Chemie International Edition*, vol. 53, no. 49

Dubey *et al.* ont ensuite cherché à calculer le gain de résolution et d'intensité des signaux RMN pour la deutération sélective des chaînes latérales des leucines marquées sur les groupements méthyles [164]. Pour cela, ils ont exprimé la protéine RBM39 humaine en cellules d'insectes (Sf9) avec un marquage stéréospécifique  $^{13}\text{C}$  des carbones  $\delta$ 1 des leucines et avec ou sans deutération de la chaîne latérale et des  $\text{H}\delta$ 2. Ainsi, ils ont calculé que la deutération de l'acide aminé comportant le groupement méthyle apporte un gain de résolution de 30 à 50%, et une intensité des pics deux fois supérieures (figure 3.1.2). En effet, pour les leucines (ainsi que les valines) les protons les plus proches dans l'espace d'un de leurs groupements méthyles sont situés sur leur deuxième groupement méthyle et sur le  $\text{C}\gamma$  (leucine) ou  $\text{C}\beta$  (valine). Leur travail démontre donc également l'un des avantages du marquage stéréospécifique des groupements méthyles de ces résidus. En effet, cela permet de supprimer les interactions dipolaires entre les protons des groupements méthyles d'un même acide aminé. L'autre avantage du marquage stéréospécifique des groupements méthyles est la simplification des spectres RMN qui ne

montrent alors plus qu'un seul pic par résidu marqué au lieu de deux (valable aussi pour les valines). Les spectres RMN sont donc moins chargés et présentent moins de pics superposés.

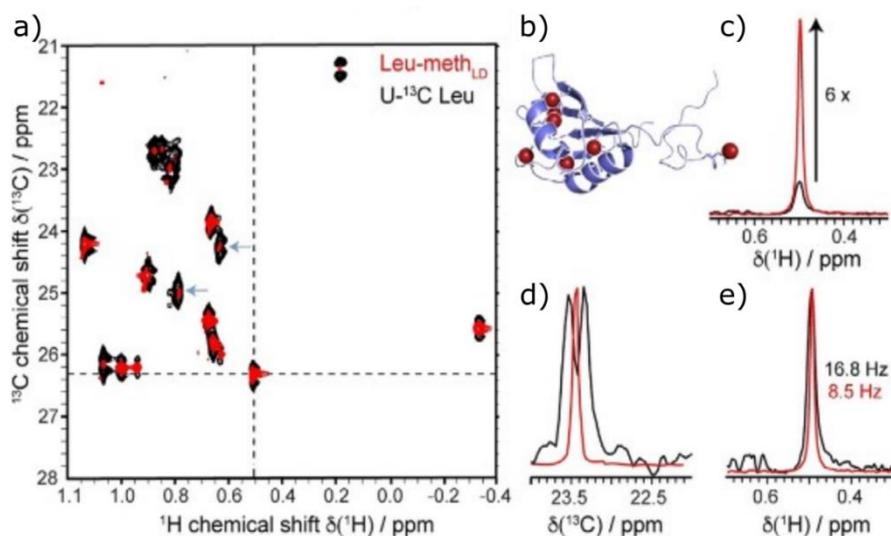


Figure 3.1.2 : Intérêt de la deutération des chaînes latérales des résidus marqués sur les groupements méthyles a) Spectre HMQC- $^{13}\text{C}$ , $^1\text{H}$  du domaine RRM2 de la RBM39 humaine exprimée en cellules d'insectes avec un marquage  $[\text{U}-^{13}\text{C}$ ,  $^1\text{H}$ ]-Leu (noir) ou  $[\text{U}-^2\text{H}]-\delta^{13}\text{CH}_3$ -Leu (Leu-meth $_{\text{LD}}$ ) (rouge). b) Structure du domaine RRM2 (PDB : 2JRS) avec les leucines marquées utilisées pour l'étude en sphères rouges. c) à e), extraction de plans 1D des dimensions  $^1\text{H}/^{13}\text{C}$ . c) le signal est augmenté par un facteur 6. d) un singulet est obtenu à la place d'un doublet ce qui facilite l'acquisition à haute résolution dans la dimension  $^{13}\text{C}$ . e) la largeur à mi-hauteur des pics est réduite de 49% dans la dimension  $^1\text{H}$ . Extrait et adapté de A. Dubey et al. *Angewandte Chemie International Edition*, vol. 60, no. 25, pp. 13783–13787, 2021, doi: [10.1002/anie.202016070](https://doi.org/10.1002/anie.202016070).

#### 1.1.4. Etat de l'art sur le marquage isotopique des anticorps exprimés en cellules de mammifères

Etant donnée l'importance des anticorps thérapeutiques, des méthodes de marquage isotopique en cellules de mammifères ont déjà été mises en place pour leur étude structurale, ainsi que pour l'attribution de certains domaines. Cette partie détaille les trois travaux majeurs publiés pour la RMN des anticorps exprimés et marqués isotopiquement en cellules de mammifères. Le marquage isotopique en cellules de mammifères est une méthode difficile à mettre en place et son application à des anticorps est très limitée. Il est intéressant de noter que les trois papiers cités proviennent du même groupe de chercheurs coordonné par le professeur Koichi Kato (Université Japonaise de Nagoya).





également produit d'autres échantillons avec des marquages sélectifs d'acides aminés, comme Shindo *et al.*, pour procéder à l'identification des types de résidus sur les spectres RMN.

Il faut noter qu'un fragment Fc d'anticorps est un hétérodimère. Le spectre RMN possède donc un nombre de résonances équivalent à une protéine de 25 kDa, ce qui facilite le travail d'attribution. L'effet de la taille de la protéine sur la relaxation à travers le  $\tau_c$  est tout de même celui d'une protéine de 50 kDa, mais l'acquisition des spectres à 52°C leur a permis d'obtenir des spectres de bonne résolution par une diminution du  $\tau_c$  de la protéine analysée (cf. partie 6.1.2).

### c Marquage stéréospécifique $^{13}\text{C}$ et deutération sélective des leucines d'une IgG2b exprimée en cellules hybridomes de souris

Yanaka *et al.* ont démontré l'intérêt d'utiliser un marquage stéréospécifique des groupements méthyles, en plus de la deutération sélective du reste des protons de l'acide aminé marqué, pour la suppression des interactions dipolaires  $^1\text{H}$ - $^1\text{H}$  intra-résidus[167]. Pour cela, une IgG2b a été exprimée en cellules hybridomes de souris en utilisant de la leucine marquée [ $\text{C}\delta_2$ - $^{13}\text{C}$  ;  $\text{H}\alpha$ ,  $\text{H}\beta$ ,  $\text{H}\gamma$ ,  $\text{H}\delta_1$ - $^2\text{H}_7$ ] ou [ $\text{C}\delta_2$ - $^{13}\text{C}$ ] pour la comparaison des spectres TROSY-méthyles. Les résultats ont montré une nette amélioration de la qualité des spectres obtenus et ont permis d'observer la totalité des pics pour les 42 groupements  $\delta_2$  de leucines du fragment Fab de l'anticorps (figure 3.1.5). Ils ont ensuite essayé d'optimiser la deutération de l'anticorps en utilisant du glucose- $^2\text{H}_7$ , mais cela n'a abouti qu'à une faible amélioration des spectres. Ils en ont conclu que la deutération des protéines à travers les voix métaboliques lors d'une expression en milieu de culture protoné était faible et de l'ordre de 10%.

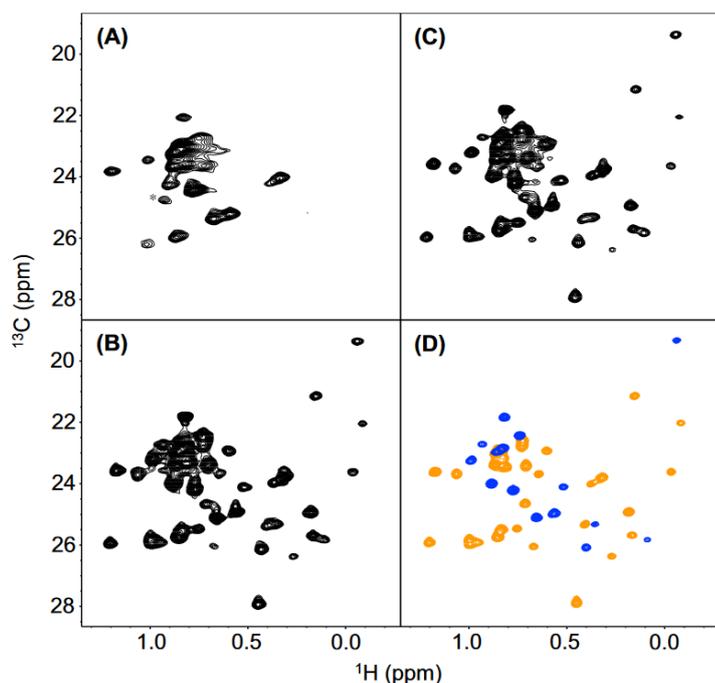


Figure 3.1.5 : Spectres méthyl TROSY d'une IgG de souris marquée isotopiquement avec : A, de la leucine [ $\text{C}\delta_2$ - $^{13}\text{C}$ ] ; B de la leucine [ $\text{C}\delta_2$ - $^{13}\text{C}$  ;  $\text{H}\alpha$ ,  $\text{H}\beta$ ,  $\text{H}\gamma$ ,  $\text{H}\delta_1$ - $^2\text{H}_7$ ] ; C de la leucine [ $\text{C}\delta_2$ - $^{13}\text{C}$  ;  $\text{H}\alpha$ ,  $\text{H}\beta$ ,  $\text{H}\gamma$ ,  $\text{H}\delta_1$ - $^2\text{H}_7$ ] + du glucose- $^2\text{H}_7$ . D, Superposition des spectres des fragments Fab (orange) et Fc (bleu) obtenus après clivage de l'IgG à la papaine. La concentration des échantillons était d'environ 65  $\mu\text{M}$  pour l'IgG entière et 200  $\mu\text{M}$  pour les fragments. Les acquisitions ont été réalisées à 37°C et pH 7.4 en utilisant un spectromètre RMN opérant à une fréquence du proton de 800 MHz. S. Yanaka et al. *J Biomol NMR*, vol. 71, no. 3, pp. 193–202, Jul. 2018, doi: 10.1007/s10858-018-0169-2.

Ils ont également démontré le maintien de la structure des fragments Fab et Fc lors d'un clivage de l'anticorps à la papaïne. Une fois superposés, les spectres des fragments Fab et Fc, obtenus après clivage et qui contiennent 28 et 14 leucines non-équivalentes, ont reproduit le spectre de l'IgG2b intacte. Cependant, aucune attribution n'a encore été réalisée pour les groupements méthyles d'un anticorps produit en cellules de mammifères.

Malgré toutes les avancées dans le domaine, le marquage isotopique d'anticorps exprimés en cellules de mammifères n'a permis l'attribution que du squelette des fragments Fc et Fv d'un anticorps. Les exemples donnés sont ceux d'un fragment Fv de 25 kDa et d'un fragment Fc de 50 kDa. Or, un fragment Fc d'anticorps est un homodimère et son spectre RMN représente en réalité celui de deux fois la même protéine de 25 kDa (avec malgré tout, les conséquences d'une protéine d'une taille de 50 kDa sur la relaxation du signal RMN). Les deux attributions de squelette de fragments d'anticorps n'ont donc jamais concerné plus de 200 acides aminés à la fois (environ). Le marquage isotopique des groupements méthyles d'un anticorps exprimé en cellules de mammifère n'a été réalisé que pour les leucines en position  $\delta 2$ , et aucune attribution n'a été réalisée. Cela souligne les défis que représentent, d'un côté, le marquage isotopique des groupements méthyles de protéines en cellules de mammifères et, de l'autre, la taille d'un anticorps de 150 kDa et plus de 1300 acides aminés (plus de 700 non équivalents) pour la RMN.

## 1.2. Production de l'IgG1 anti-LAMP1 marquée spécifiquement $^{13}\text{C}_3$ sur les groupements méthyles en cellules CHO

A L'heure actuelle, il n'existe aucune attribution des groupements méthyles d'un anticorps exprimé en cellules de mammifères. Or, la caractérisation de la HOS de l'anticorps anti-LAMP1 par la RMN nécessite cette attribution préalable. Dans ce projet, nous proposons de réaliser l'attribution des fragments Fab et Fc de l'anticorps exprimés en système *cell-free*. La taille plus faible de ces fragments, comparées à celle de l'anticorps entier, est, en effet, un avantage pour la RMN. Puis, l'attribution de ces fragments sera transférée sur les spectres  $^1\text{H}$ - $^{13}\text{C}$  de l'anticorps anti-LAMP1 produit en cellules CHO. Cependant, le spectre RMN en abondance naturelle d'un anticorps présente un très grand nombre de résonances pour les groupements méthyles (au nombre de 436 pour l'anticorps anti-LAMP1, en comptant les deux groupements méthyles des résidus Ile, Leu, Val). Leur superposition peut rendre le transfert d'attribution compliqué.

Pour simplifier les spectres RMN de l'anticorps entier et pouvoir, à terme, transférer l'attribution plus facilement, nous avons choisi de marquer les groupements méthyles de manière stéréospécifique pour les valines et leucines et régiosélective pour les isoleucines.

Cette stratégie a nécessité le développement et l'optimisation de plusieurs paramètres de la production pour obtenir le meilleur rendement de production possible lors d'un marquage isotopique, ainsi que le développement d'un mélange d'acides aminés optimisé pour l'incorporation des acides aminés marqués.

### 1.2.1. Mise en place de la production et de la purification de l'anticorps exprimé par les cellules CHO

La première étape du développement a consisté à mettre en place un protocole de culture cellulaire adapté à la production et à la purification de l'anticorps en laboratoire. En effet, la lignée cellulaire utilisée pour ce projet permet d'exprimer l'anticorps anti-LAMP1 à un rendement de 400 mg/L en utilisant une méthode de culture en mode *fed-batch* et en utilisant des milieux de cultures riches. Pour rappel, le mode de culture en *fed-batch* consiste à nourrir les cellules quotidiennement pour maintenir stables les concentrations de certains composés comme le glucose ou les acides aminés. Pour la production de protéines marquées isotopiquement, cette méthode de culture cellulaire n'est pas judicieuse en raison d'une consommation trop importante de matériel marqué. Les cellules CHO ont donc été cultivées en mode *batch* (en flasque) dans des milieux de culture commerciaux enrichis dans un premier temps (le milieu CDM4PERMAb (SH30871.02) et les compléments de milieux Cell-Boost 7a (référence SH3119.01) et 7b (référence SH31120.01) de Cytiva). La composition exacte de ces milieux commerciaux riches est inconnue, mais ils sont optimisés pour augmenter la croissance cellulaire et l'expression d'anticorps par les cellules.

Les conditions standards de culture de cellules CHO ont été appliquées pour ce protocole. La densité cellulaire pour le début de la production d'anticorps a été fixée à  $1.10^6$  cellules/mL en utilisant les milieux de culture enrichis. Le début de la phase d'apoptose des

cellules dans ces conditions a été observé après quatre jours dans l'incubateur et le surnageant de la culture a été récupéré par centrifugation, l'anticorps étant excrété par les cellules.

La purification de l'anticorps a été réalisée en utilisant la spécificité de la protéine A pour les anticorps. La figure 3.1.6 présente le profil d'élution obtenu pour cette première étape de purification.

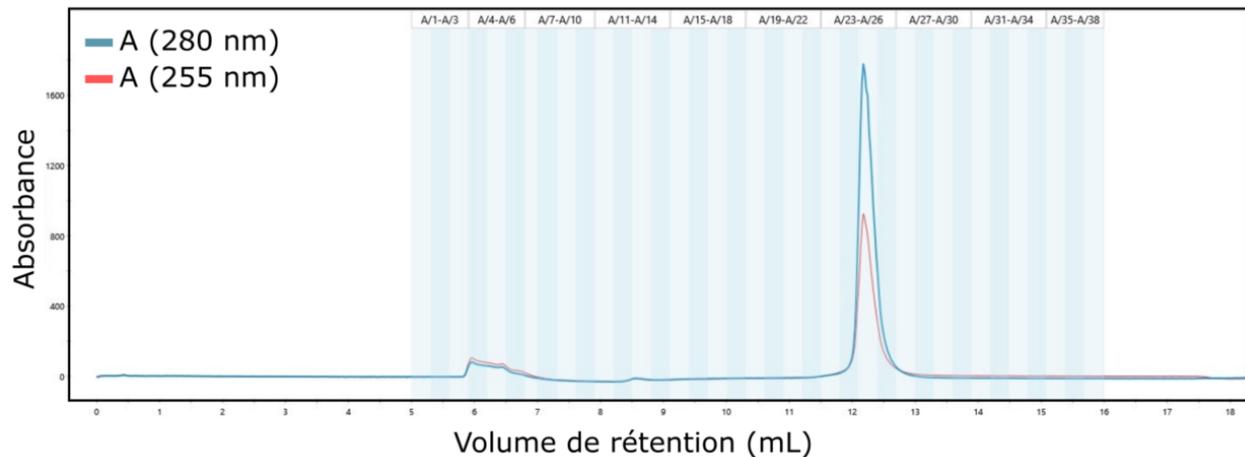


Figure 3.1.6 : Profil d'élution obtenu pour la purification de l'anticorps anti-LAMP1 par chromatographie d'affinité sur une colonne Hiscreen MabSelect sure de 5mL. Débit de 1 mL/min.

Dans un second temps, une étape de purification par chromatographie d'exclusion a été mise en place afin d'éliminer les contaminants potentiellement présents. Le profil d'élution obtenu (figure 3.1.7) a démontré l'absence d'agrégats ou de fragments et le temps de rétention du pic visible correspond à une protéine de 150 kDa d'après les standards de la colonne.

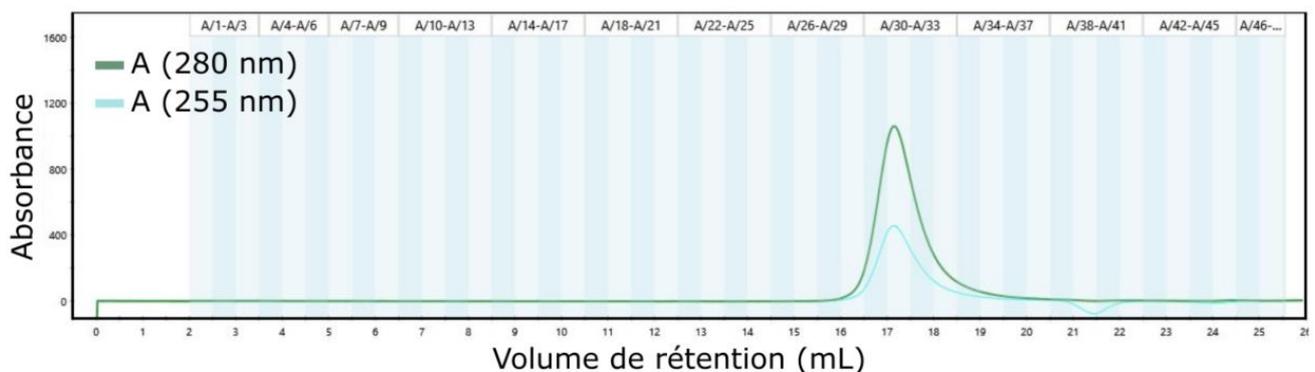


Figure 3.1.7 : Profil d'élution obtenu pour la purification de l'anticorps anti-LAMP1 par chromatographie d'exclusion de taille sur une colonne de gel superose 6 10/300. Débit de 0.5 mL/min.

Le rendement pour une production utilisant les milieux de culture riches et commerciaux avec une densité cellulaire initiale de  $1.10^6$  cellules/mL en flasque a ainsi été estimé à 76 mg d'anticorps produit par litre de culture.

### **1.2.2. Développement d'un milieu de culture adapté au marquage isotopique en cellules de mammifères**

Les milieux de culture commerciaux riches comme le CDM4PERMAb (et les compléments de milieu Cell Boost 7a et 7b) favorisent la croissance cellulaire et la production d'anticorps pour obtenir des hauts rendements d'expression. En revanche, pour le marquage isotopique, il n'est pas envisageable d'utiliser un milieu de culture contenant des acides aminés non marqués (CDM4PERMAb), ainsi que des compléments de milieu (Cell Boost 7a et Cell Boost 7b) qui contiennent également et entre autres des acides aminés. L'objectif du marquage isotopique est d'obtenir un taux d'incorporation des acides aminés marqués dans la protéine d'intérêt le plus élevé possible tout en utilisant une quantité minimale d'acides aminés marqués. La présence d'acides aminés non marqués dans le milieu de culture lors du marquage isotopique entraînerait une dilution des acides aminés marqués et une diminution de leur taux d'incorporation (on parle de dilution isotopique). La solution choisie a donc été de supprimer les sources d'acides aminés non marqués en utilisant un milieu de culture commercial dépourvu en acides aminés, et d'ajouter un mélange d'acides aminés marqué optimisé à la fois pour les rendements et le taux d'incorporation. Ce mélange a tout d'abord été développé avec des acides aminés non marqués pour vérifier que le rendement par rapport au milieu commercial n'était pas affecté.

#### **a Evaluation de trois mélanges d'acides aminés sur le rendement d'expression de l'anticorps**

Le milieu de culture CDM4PERMAb a été commandé à façon, dépourvu d'acides aminés et les compléments de milieu n'ont plus été utilisés. Ensuite, un mélange d'acides aminés a été mis au point avec pour objectif d'obtenir un rendement de production d'anticorps le plus élevé possible. Pour cela, des mélanges d'acides aminés issus de milieux commerciaux de compositions connues et d'une publication ont été testés au préalable : le milieu DMEM et le milieu OptiCHO commerciaux et un mélange d'acides aminés pour cellules d'insectes développé par Alvar Gossert *et al.* [168] et optimisé par NMR-Bio. Les mélanges créés seront appelés AADMEM, AAOptiCHO et AAInsect par la suite et leur composition (en concentration d'acides aminés dans le milieu final) est détaillée en annexe 6.1, 6.2 et 6.3.

La figure 3.1.8 représente les rendements d'expression de l'anticorps obtenus en utilisant chacun de ces mélanges d'acides aminés ajouté dans le milieu de culture commercial CDM4PERMAb dépourvu d'acides aminés. On remarque que les rendements obtenus sont quasiment similaires pour les deux mélanges AADMEM et AAOptiCHO. Cependant, lorsque l'on somme la masse des acides aminés présents dans ces mélanges, on remarque que le mélange AADMEM en contient 1.4 fois plus en masse que le AAOptiCHO : 1.62 g/L pour le mélange AADMEM contre 1.17 g/L pour le mélange AAOptiCHO. Cela signifie que la proportion de chacun des acides aminés dans le mélange AAOptiCHO compense la quantité d'acides aminés totale plus basse. Ainsi, l'hypothèse est que la proportion des acides aminés dans ces mélanges est importante pour obtenir un bon rendement d'expression.

Pour le mélange AAInscet, outre la différence des proportions d'acides aminés dans ce mélange par rapport aux deux autres, la masse totale d'acides aminés y est 2 à 3 fois supérieure. En effet, ce mélange contient 3.07 g/L d'acides aminés. En utilisant le mélange AAInscet dans le milieu de culture commercial dépourvu en acides aminés, le rendement a été estimé à **25 mg/L** (figure 3.1.8). Cela correspond à une augmentation du rendement par un facteur 3. On en conclut que la quantité d'acides aminés est également importante pour obtenir un rendement d'expression d'anticorps plus élevé.

### b Développement d'un mélange d'acides aminés optimisé

Ainsi, une quantité totale de 3 g/L d'acides aminés dans le milieu de culture est optimale et l'optimisation de leur proportion peut améliorer le rendement d'expression.

Pour développer un mélange d'acides aminés optimisé, la proportion de chaque acide aminé a été optimisée sur la base du mélange AAInscet et sur les travaux de Carrillo-Cocom *et al.* qui ont démontré que la consommation en acides aminés des cellules CHO fluctue en fonction de la composition en acides aminés de la protéine sur exprimée par ces cellules[169].

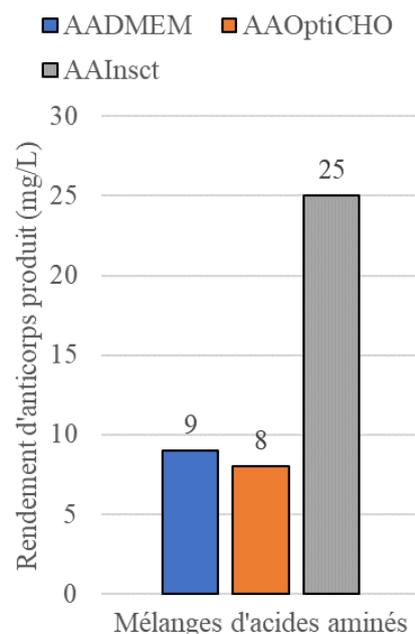


Figure 3.1.8 : Rendement d'anticorps exprimé par les cellules CHO en utilisant le milieu commercial dépourvu en acides aminés et les mélanges AADMEM, AAOptiCHO et AAInscet.

Ainsi, le mélange AAInscet a été utilisé comme point de départ car c'est avec celui-ci que le rendement le plus élevé a été obtenu. L'optimisation de la proportion des acides aminés dans ce mélange a été réalisée en déterminant la proportion de chaque acide aminé dans la séquence de l'anticorps. L'optimisation de la concentration des acides aminés dans le mélange n'a été réalisée que sur les acides aminés essentiels pour les cellules CHO puisque les autres sont synthétisés par la cellule et n'ont, en théorie, pas besoin d'être ajoutés au milieu de culture. La composition du mélange optimisé d'acides aminés (AAOptimAb) est détaillée dans le tableau 3.1.1.

Tableau 3.1.1: Composition finale optimisée pour les 20 acides aminés. Cette composition correspond au mélange AAOptimAb

Acides aminés	Concentrations (mg/L)	Acides aminés	Concentrations (mg/L)
Alanine	93,3	Leucine	270
Arginine	83,5	Lysine	238,8
Asparagine	260	Méthionine	35
Aspartate	300	Phénylalanine	126,6
Cystéine	46	Proline	186,8
Glutamate	356	Sérine	283,2
Glutamine	400	Thréonine	246,3
Glycine	100,32	Tryptophane	24
Histidine	71,6	Tyrosine	200
Isoleucine	60	Valine	236,8

La figure 3.1.9 présente la proportion de chaque acide aminé dans le mélange AAInscet et le AAOptimAb et dans la séquence de l'anticorps anti-LAMP1. Cette figure montre que la composition du mélange AAOptimAb est plus proche de la composition de l'anticorps en comparaison de celle du AAInscet.

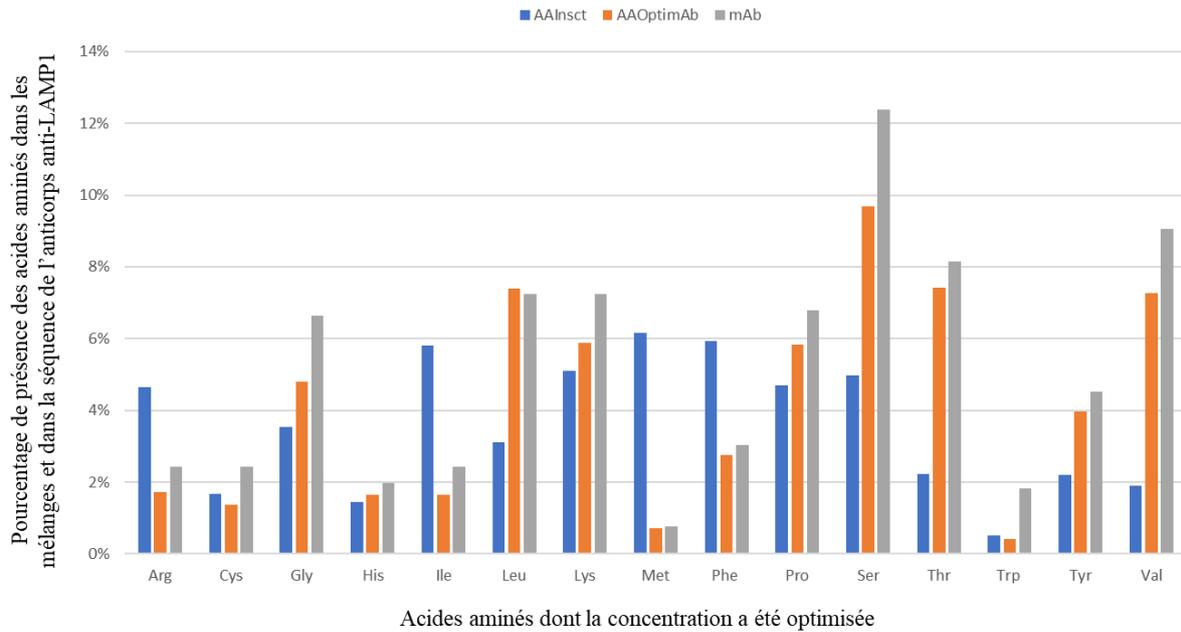


Figure 3.1.9 : Pourcentages de présence des acides aminés dans les mélanges AAInsct et AAOptimAb comparés à ceux dans la séquence de l'anticorps anti-LAMP1. Les valeurs ne sont pas données pour les acides aminés non-essentiels dont les concentrations n'ont pas été optimisées par rapport au mélange AAInsct.

La production de l'anticorps en utilisant le milieu commercial dépourvu d'acides aminés et le mélange AAOptimAb a été évaluée à 40 mg/L, soit 1.6 fois supérieure au rendement avec le mélange AAInsct non optimisé (figure 3.1.10).

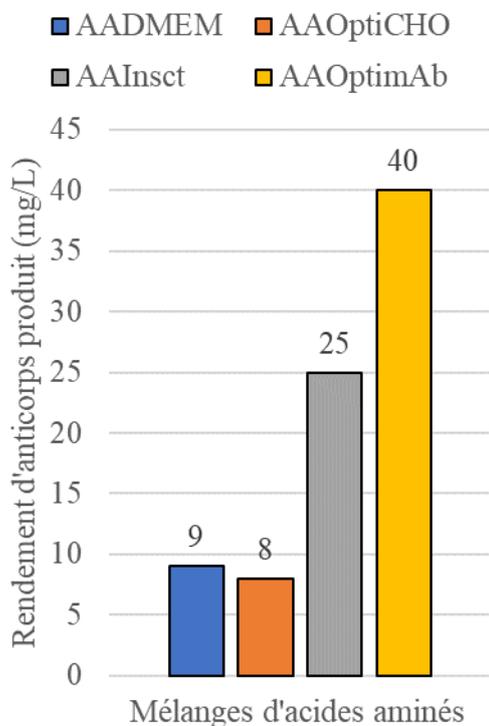


Figure 3.1.10 : Rendement d'anticorps exprimé par les cellules CHO en utilisant le milieu commercial dépourvu en acides aminés et le mélange AAOptimAb développé et optimisé comparé à ceux obtenus pour les trois mélanges précédents.

Le développement d'un mélange d'acides aminés optimisé selon la séquence de l'anticorps a permis d'augmenter les rendements et d'atteindre **40 mg/L**, soit encore deux fois moins que le rendement obtenu en utilisant les milieux commerciaux enrichis (76 mg/L). Pour essayer d'atteindre les rendements obtenus avec les milieux commerciaux, nous avons joué avec les densités cellulaires pendant la phase de production.

### c Optimisation du rendement d'expression de l'anticorps anti-LAMP1 en cellules CHO à haute densité cellulaire

Au début d'une culture, les cellules s'adaptent au nouveau milieu de culture. En débutant à une densité cellulaire de  $0.5 \cdot 10^6$  cellules/mL et pour des cellules CHO cultivées en flasque, cette adaptation peut durer un à deux jours. Lors de cette phase, les cellules se multiplient lentement et produisent peu d'anticorps. Ensuite, les cellules entrent dans une phase de croissance exponentielle. Durant cette étape qui dure de 2 à 4 jours, les cellules consomment les nutriments du milieu pour se multiplier de manière exponentielle jusqu'à atteindre une densité cellulaire limite ( $8 \cdot 10^6$  cellules/mL pour la lignée cellulaire de la thèse cultivée en flasque). Lors de cette phase, les cellules expriment tout de même de l'anticorps, mais pas en quantité importante car les cellules sont peu nombreuses. Une fois que la densité cellulaire limite est atteinte, les cellules ne consomment plus les nutriments du milieu pour se multiplier, mais pour produire des protéines. L'expression d'anticorps est donc optimale à cette densité cellulaire maximale (la quantité d'anticorps produit par cellule est identique, mais la quantité de cellules est beaucoup plus importante, la production d'anticorps est donc supérieure). Enfin, après deux à trois jours, les cellules ont consommé la plupart des nutriments du milieu et les concentrations restantes ne sont plus suffisantes pour les maintenir en vie. Les cellules commencent alors leur phase d'apoptose et la culture s'arrête.

L'optimisation du marquage isotopique en cellules de mammifères consiste à incorporer les acides aminés marqués au bon moment de la culture. Lors des phases d'adaptation et de croissance exponentielle, la majeure partie des nutriments n'est pas consommée pour la production de l'anticorps. C'est pourquoi la culture à haute densité cellulaire a été mise en place pour le développement de ce marquage isotopique en cellules CHO. En flasque, les cellules CHO ne peuvent dépasser une densité de  $10 \cdot 10^6$  cellules/mL. En débutant la culture à cette densité, on s'assure alors que les cellules ne consomment pas les nutriments du milieu pour croître, mais pour produire des protéines, et donc l'anticorps (figure 3.1.11).

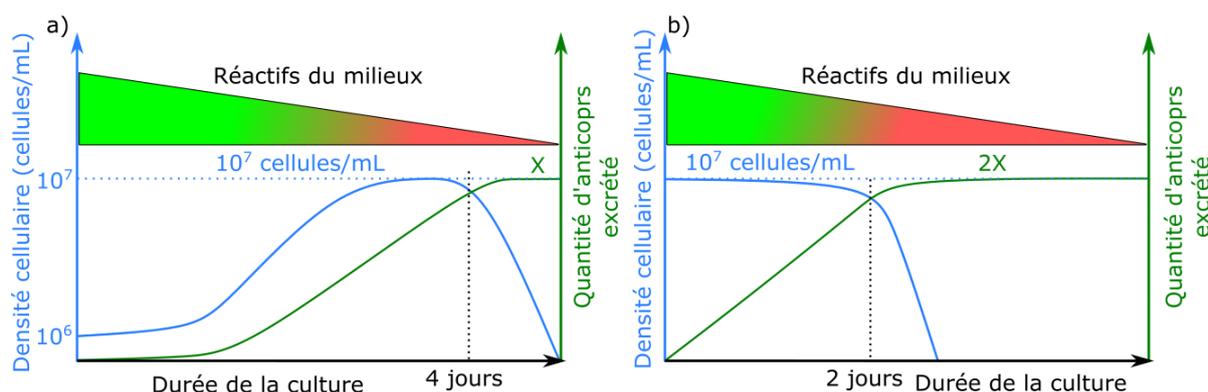


Figure 3.1.11 : Profil théorique de croissance des cellules CHO en fonction de la consommation des réactifs du milieu et de la production d'anticorps associée en mode batch en commençant à basse densité cellulaire a) et à haute densité cellulaire b).

Pour la production de l'anticorps à haute densité cellulaire, les cellules ont été cultivées au préalable jusqu'à atteindre une densité cellulaire de 2 à  $4 \cdot 10^6$  cellules/mL dans des flasques de cultures de 500 mL (200 rpm, 5 % de  $\text{CO}_2$  et  $37^\circ\text{C}$ ). Cette étape consiste à obtenir de la biomasse (des cellules) en quantité suffisante pour démarrer des cultures à haute densité cellulaire. Les cellules ont ensuite été centrifugées dans des flacons de centrifugation stériles à

200 g pendant 10 minutes et re-suspendues dans du milieu de culture frais à une densité de  $10 \cdot 10^6$  cellules/mL. La production d'anticorps non marqué dans ces conditions dure 48 h avant que les cellules ne commencent leur phase d'apoptose.

Le rendement de production à haute densité cellulaire a été déterminé après purification et comparé à ceux obtenus lors des étapes de développement précédentes : production avec les milieux riches, puis avec le mix d'acides aminés développé à une densité cellulaire initiale de  $1 \cdot 10^6$  cellules/mL (figure 3.1.12). Les résultats ont montré que le rendement de production de l'anticorps à basse densité cellulaire et en utilisant les milieux riches a été restauré grâce à cette méthode de production à haute densité cellulaire avec le mix d'acides aminés développé (**80 mg/L**).

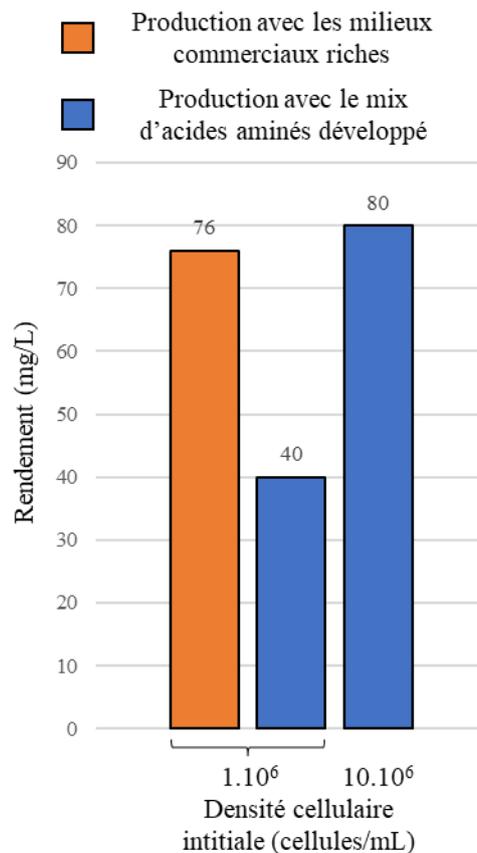


Figure 3.1.12 : Rendements d'expression de l'IgG1 anti-LAMP1 par les cellules CHO en fonction des milieux de culture utilisés, et de la densité cellulaire initiale.

#### **d Optimisation du taux d'incorporation des acides aminés**

L'optimisation du taux d'incorporation des acides aminés marqués passe par une élimination des acides aminés non marqués présents dans la culture. En effet, lors de la phase de croissance cellulaire, les cellules sont cultivées dans des milieux de culture commerciaux enrichis qui contiennent des acides aminés non marqués. Pour les éliminer, les cellules sont centrifugées et rincées deux fois avec du milieu de culture déplété en acides aminés.

Cette étape de lavage permet d'éliminer les acides aminés présents dans le surnageant de culture mais ne suffit pas à s'affranchir des acides aminés non marqués présents dans les cellules. En effet, puisque les cellules sont initialement cultivées dans un milieu de culture contenant des acides aminés non marqués, on peut faire l'hypothèse que, malgré l'étape de lavage, les cellules contiennent toujours un *pool* interne d'acides aminés qui peuvent être utilisés par les cellules lors de l'expression de la protéine et provoquer une dilution isotopique des acides aminés marqués. Pour éliminer un maximum de ces acides aminés présents dans les cellules, celles-ci ont été affamées en les cultivant à haute densité et dans le milieu commercial déplété en acides aminés pendant quelques heures. Durant cette période, les cellules ne se multiplient pas, car la densité cellulaire est trop haute, et leur *pool* d'acides aminés interne est utilisé uniquement pour leur survie ou la production d'anticorps non marqué excrété dans le milieu de culture. L'évaluation du bénéfice de cette étape d'affamement a été réalisée, dans un premier temps, sur la capacité des cellules à survivre dans de telles conditions. Des durées d'affamement de 4, 8 et 16 h ont été testées et la viabilité cellulaire a été vérifiée pour chaque durée d'incubation. Les résultats ont montré que même après 16 h à  $10 \cdot 10^6$  cellules/mL dans un milieu de culture déplété en acides aminés, les cellules présentaient une viabilité cellulaire acceptable et supérieure à **90%**. Cette durée a donc été retenue pour essayer d'éliminer le *pool* interne d'acides aminés des cellules. Après 16 h, les cellules ont été centrifugées et lavées deux fois avec le milieu de culture déplété en acides aminés en supprimant le surnageant qui ne contient normalement pas d'acides aminés en abondance naturelle, mais potentiellement de l'anticorps non marqué produit pendant cette étape d'affamement.

Une fois que la durée d'affamement permettant de maintenir une bonne viabilité a été sélectionnée, il a été nécessaire de vérifier son impact sur l'incorporation des acides aminés marqués. Cela a été réalisé en comparant la qualité des spectres RMN  $^1\text{H}$ - $^{13}\text{C}$  obtenus pour l'anticorps produit avec et sans étape d'affamement en marquant, dans les deux cas, le groupement méthyle en position  $\delta 1$  des isoleucines. L'intensité des pics RMN d'un groupement chimique étant proportionnelle à l'abondance isotopique des noyaux de ce groupement, il est possible d'utiliser la RMN pour évaluer l'efficacité de la méthode utilisée pour l'incorporation des acides aminés marqués. L'échantillon a été produit avec de l'isoleucine marquée  $^{13}\text{CH}_3$  en position  $\delta 1$  (le marquage isotopique de l'isoleucine était régiosélectif (en position  $\delta 1$ ) de façon à n'obtenir qu'un signal par isoleucine sur le spectre RMN) à 100 mg/L et également de la méthionine marquée  $^{13}\text{CH}_3$  à 1 g/L (en large excès pour servir d'étalon interne). L'utilisation de la méthionine en excès comme étalon interne permet de s'affranchir de l'impact de la concentration de l'échantillon sur l'intensité des résonances RMN, en supposant qu'à cette concentration, les méthionines sont marquées à 100% dans les anticorps produits. Deux échantillons ont ainsi été obtenus :

- 1 : Production **sans** affaînement,  $^{13}\text{CH}_3\text{-Met}$  à 1 g/L et  $^{13}\text{CH}_3\text{-Ile}\delta 1$  à 100 mg/L
- 2 : Production **avec** affaînement pendant 16 h,  $^{13}\text{CH}_3\text{-Met}$  à 1 g/L et  $^{13}\text{CH}_3\text{-Ile}\delta 1$  à 100 mg/L

Pour les deux échantillons, les productions ont été réalisées dans le milieu CDM4PERMAb déplété en acides aminés, supplémenté avec le mix d'acides aminés non marqués développé (sans méthionine et isoleucine), de la méthionine et de l'isoleucine marquées  $^{13}\text{CH}_3$ .

La figure 3.1.13 présente les résultats obtenus. L'échantillon produit avec l'étape d'affaînement présente un ratio d'intensité des pics (Ile/Met) de 0.76 contre 0.57 pour l'échantillon produit sans étape d'affaînement. On peut donc conclure que l'étape d'affaînement améliore d'un **facteur 1.3** la qualité des spectres RMN en améliorant l'incorporation des acides aminés marqués lors de la production de la protéine. De plus, le rendement de production obtenu en utilisant l'étape d'affaînement des cellules a été estimé à 79 mg/L. La méthode d'affaînement permet donc d'obtenir une meilleure incorporation des acides aminés marqués sans impacter le rendement d'expression.

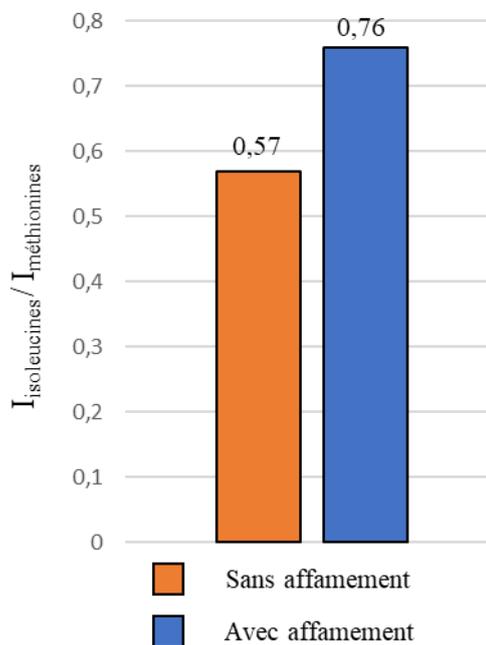


Figure 3.1.13 : Impact de l'affaînement sur l'intensité des pics RMN  $^1\text{H-}^{13}\text{C}$  des isoleucines marquées  $^{13}\text{CH}_3\text{-}\delta 1$

On en conclut donc que l'étape d'affaînement des cellules améliore significativement l'incorporation des acides aminés marqués et que la méthode de culture mise en place permet bien de diminuer le *pool* interne d'acides aminés non marqués des cellules. Le protocole développé pour le marquage isotopique de l'anticorps anti-LAMP1 produit en cellules CHO est présenté dans la figure 3.1.14.

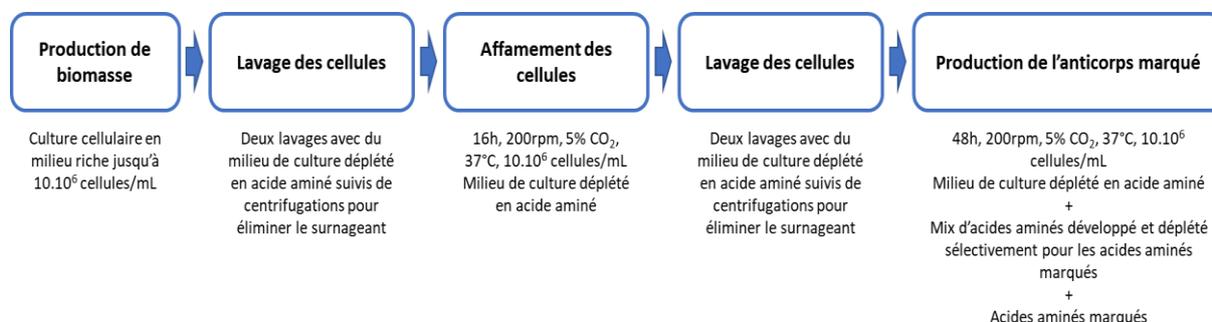


Figure 3.1.14 : Protocole de culture cellulaire développé pour la production de l'anticorps anti-LAMP1 exprimé et marqué en cellules CHO.

### e Marquage isotopique $^{13}\text{CH}_3$ des isoleucines, leucines, méthionines, thréonines et valines

#### a) Suivi qualitatif de l'incorporation des acides aminés par RMN :

Les acides aminés isoleucines, leucines, méthionines, thréonines et valines font partie des acides aminés essentiels pour les cellules de mammifères. Pour évaluer la quantité de chacun de ces acides aminés à ajouter dans le milieu, nous avons émis l'hypothèse que la capacité des cellules à incorporer ces acides aminés était identique. En effet, L. M. Carrillo-Cocom *et al.*[169] ont démontré que les cellules CHO présentent un profil de consommation identique pour ces cinq acides aminés en cas de surexpression d'une protéine recombinante ou non par les cellules avec des concentrations initiales d'acides aminés toutes inférieures à 50 mg/L. L'isoleucine a été choisie comme marquage de référence puis les résultats ont été extrapolés aux autres résidus.

L'incorporation de l'isoleucine a d'abord été évaluée sur l'amélioration de la qualité des spectres RMN  $^1\text{H}$ - $^{13}\text{C}$  pour le marquage isotopique des groupements méthyles en position  $\delta 1$  des isoleucines.

Les cellules CHO exprimant l'anticorps anti-LAMP1 ont été cultivées à haute densité cellulaire en présence d'isoleucine marquée  $^{13}\text{CH}_3$  en position  $\delta 1$  à quatre concentrations différentes (100, 150, 200 et 250 mg/L) et de méthionine (1 g/L) pour utiliser ses résonances RMN comme référence. L'anticorps ainsi produit a été purifié et des spectres HMQC  $^1\text{H}$ ,  $^{13}\text{C}$ -SOFAST ont été acquis. Le ratio de la moyenne des intensités des pics des isoleucines sur celle des pics des méthionines a été calculé pour chaque échantillon. Les résultats obtenus sont présentés dans la figure 3.1.15. La moyenne des intensités des pics des isoleucines a augmenté proportionnellement à la quantité d'isoleucine ajoutée. En revanche, le plateau représentant le maximum d'incorporation de l'acide aminé n'a pas été atteint. Sans échantillon supplémentaire, il n'a pas été possible de déterminer si l'incorporation de l'acide aminé marqué était maximale à une concentration de 250 mg/L.

La spectrométrie de masse étant quantitative et plus appropriée aux besoins du projet, des échantillons supplémentaires avec un marquage dédié ont été produits.

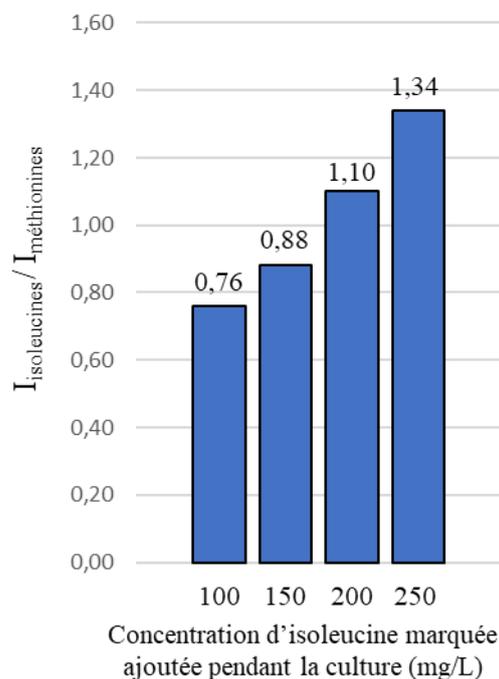


Figure 3.1.15 : Impact de la quantité d'isoleucine marquée ajoutée lors de la culture sur l'intensité de ses pics sur les spectres RMN 2D  $^1\text{H}$ - $^{13}\text{C}$ .

## b) Suivi quantitatif de l'incorporation des acides aminés par spectrométrie de masse :

Pour quantifier le taux d'incorporation des acides aminés marqués lors de la production de l'anticorps en cellules CHO, la spectrométrie de masse en mode masse intacte a été utilisée. Pour cela, des échantillons ont été produits avec un seul type d'acide aminé marqué à la fois. Le type de marquage isotopique a été différent de celui utilisé pour la RMN. En effet, pour la RMN, certains acides aminés étaient deutérés sur les chaînes latérales (et marqués stéréospécifiquement pour la valine et la leucine, et régiosélectivement pour l'isoleucine). Or, le deutérium en position C $\alpha$  des acides aminés peut être échangé avec les protons du solvant à cause de transaminases présentes dans les cellules. Cela implique qu'un même résidu deutéré pourrait avoir des masses différentes. Par spectrométrie de masse, cette hétérogénéité de masse entraîne une perte de précision sur l'analyse. C'est pourquoi des acides aminés non deutérés ont été utilisés. Les isoleucines, leucines, thréonines et valines ont été marquées [U- $^1\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$ ], les méthionines U-[ $^1\text{H}$ ], [ $^{13}\text{C}^1\text{H}_3$ ] $\epsilon$  et les alanines U-[ $^1\text{H}$ ], [ $^{13}\text{C}^1\text{H}_3$ ] $\beta$ .

Basés sur les estimations préliminaires obtenues lors des analyses RMN pour l'incorporation de l'isoleucine, les taux d'incorporation des acides aminés marqués ont été évalués pour l'isoleucine, la leucine, la méthionine, la thréonine et la valine ajoutées à des concentrations de 200 et 350 mg/L lors de la culture. Les résultats sont détaillés dans la figure 3.1.16.

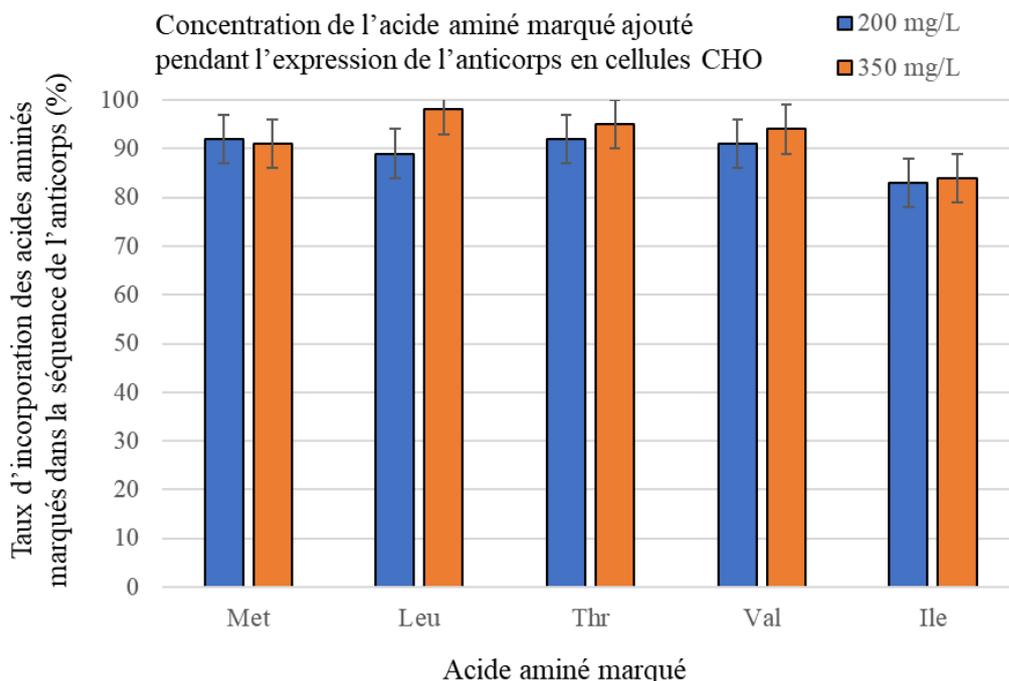


Figure 3.1.16 : Taux d'incorporation des acides aminés marqués et comportant un groupement méthyle déterminés par spectrométrie de masse.

L'incorporation des méthionines, leucines, valines et des thréonines a été évaluée à **90 %** pour un ajout de **200 mg/L** de l'acide aminé marqué pendant la culture. L'isoleucine est incorporée au maximum à 83% avec un ajout de 200 mg/L lors de la culture. Les taux d'incorporation n'ont pas été significativement améliorés par ajout d'une quantité supérieure d'acides aminés (300 mg/L).

Le résultat obtenu pour l'isoleucine contredit l'estimation préliminaire qui a été faite en RMN. En effet, précédemment, l'amélioration de l'intensité des signaux des isoleucines sur les spectres RMN indiquait une amélioration significative du taux d'incorporation de l'isoleucine pour une concentration de 250 mg/L en comparaison à 200 mg/L. Or, l'analyse par spectrométrie de masse montre un maximum d'incorporation pour 200 mg/L. Cela montre que l'évaluation du taux d'incorporation d'un acide aminé sur l'amélioration de la qualité des spectres RMN est très relative.

#### **f Le marquage isotopique de l'alanine : cas particulier**

Parmi les six acides aminés comportant un groupement méthyle dans les cellules eucaryotes, l'alanine est le seul acide aminé non-essentiel synthétisé à partir du glucose (non marqué) présent dans le milieu de culture. Cela a pour effet de provoquer une dilution isotopique de l'acide aminé marqué lors de la production. Pour éviter cela, la solution adoptée a été d'utiliser un inhibiteur de l'alanine transaminase qui est responsable de la biosynthèse d'alanine à partir du pyruvate (produit du glucose). Deux inhibiteurs ont été testés à des concentrations différentes en évaluant leur impact sur le taux d'incorporation d'alanine marquée par RMN : la L-cyclosérine et la  $\beta$ -chloroalanine [151], [163]. Les échantillons ont été produits avec le protocole mis en place précédemment (30 mL de culture), et avec la méthionine marquée  $^{13}\text{CH}_3$  ajoutée à 1 g/L dans le milieu de culture. Tous les échantillons ont été produits en utilisant 100 mg/L d'alanine marquée  $^{13}\text{CH}_3$ . Cinq échantillons ont été produits :

- 1** : Pas d'inhibiteurs de l'alanine transaminase
- 2** : Ajout de  $\beta$ -chloroalanine à 20 mg/L
- 3** : Ajout de  $\beta$ -chloroalanine à 40 mg/L
- 4** : Ajout de L-cyclosérine à 5 mg/L
- 5** : Ajout de L-cyclosérine à 10 mg/L

Les résultats obtenus (figure 3.1.17) montrent que les intensités des pics RMN  $^1\text{H}$ - $^{13}\text{C}$  des alanines sont nettement supérieures lors de l'utilisation d'inhibiteur. En moyenne les pics des alanines sont 5.8 fois plus intenses lors de l'expression en présence d'inhibiteurs, indépendamment du type d'inhibiteur et de la concentration utilisée. On en conclut donc que pour ces deux inhibiteurs et aux concentrations testées, l'alanine transaminase est inhibée de la même façon. La figure 3.1.18 présente une comparaison des spectres  $^1\text{H}$ - $^{13}\text{C}$  obtenus sans (a) et avec (b) ajout de L-cyclosérine à 5 mg/L et permet d'illustrer l'amélioration de la qualité des spectres RMN.

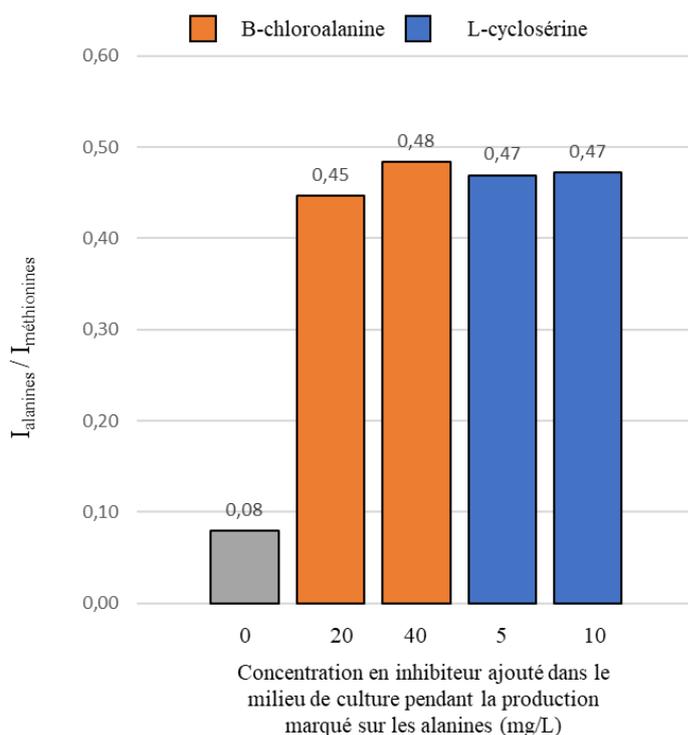


Figure 3.1.17 : Impact du type d'inhibiteur et de sa concentration sur l'intensité des pics RMN  $^1\text{H}$ - $^{13}\text{C}$  des alanines

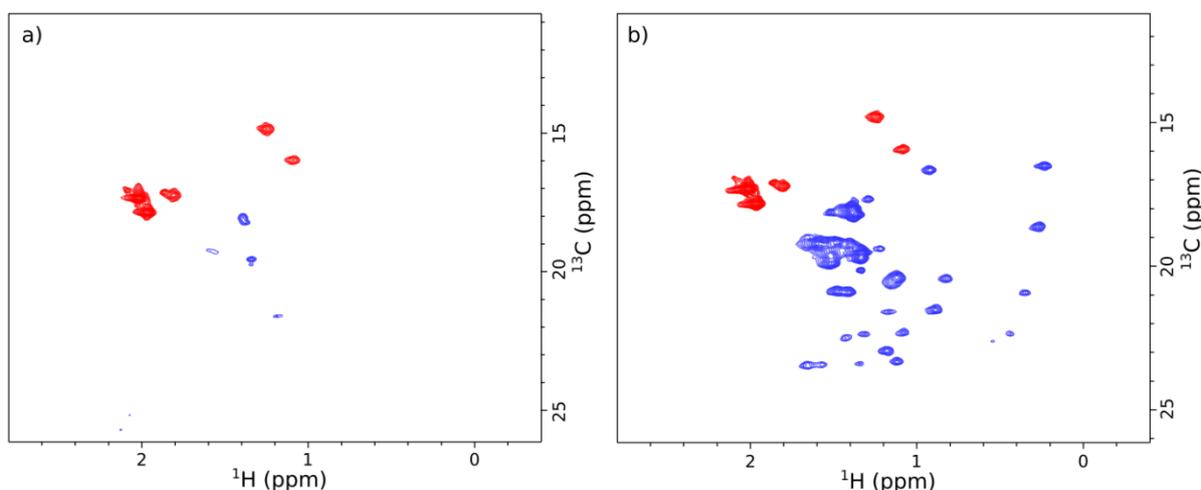


Figure 3.1.18 : Spectres  $^1\text{H}$ - $^{13}\text{C}$ -SOFAST de l'IgG1 anti-LAMP1 marquée U- $^{13}\text{C}$ ,  $^{13}\text{C}$ -Met et U- $^{13}\text{C}$ - $\beta$ -alanine et produite en cellules CHO en absence d'inhibiteur de l'alanine transaminase a) et en présence de 5 mg/L de L-cyclosérine b). Les pics correspondant aux méthionines sont représentés en rouge sur les deux spectres et les alanines sont en bleu. Les échantillons étaient à une concentration d'anticorps de 40  $\mu\text{M}$  a) et 55  $\mu\text{M}$  b). Les spectres ont été acquis à 35°C et à une fréquence  $^1\text{H}$  de 600 MHz avec un spectromètre RMN équipé d'une sonde cryogénique.

La L-cyclosérine et la  $\beta$ -chloroalanine peuvent être utilisées toutes les deux aux concentrations testées avec la même efficacité. La L-cyclosérine a été arbitrairement sélectionnée (à une concentration de 5 mg/L) et deux autres échantillons produits à partir de 100 et 300 mg/L d'alanine (et 1g/L de méthionine) ont été préparés. L'intensité des pics des alanines par rapport à celle des pics des méthionines a été identique dans les deux cas.

Le taux d'incorporation de l'alanine a ensuite été déterminé par spectrométrie de masse en mode masse intacte pour un ajout d'alanine marquée  $^{13}\text{CH}_3$  à 300 mg/L et de L-cyclosérine à 10 mg/L pendant la culture cellulaire. L'incorporation de l'alanine marquée dans ces conditions a été estimée à 85% par spectrométrie de masse.

### 1.2.3. Analyse RMN de l'IgG1 anti-LAMP1 exprimée en cellules CHO et spécifiquement marquée sur les groupements méthyles.

Pour faciliter le transfert d'attribution *in fine*, un marquage spécifique des résidus comportant un groupement méthyle a été réalisé à raison d'un ou deux types d'acides aminés marqués à la fois. De cette manière, le transfert d'attribution des fragments produits en *cell-free* jusqu'au mAb produit en cellules CHO, pourra être réalisé par type d'acide aminé sur des spectres RMN moins encombrés. Ainsi, cinq échantillons ont été produits en utilisant le protocole et le mélange d'acides aminés développés. Les acides aminés isoleucine, leucine et valine ont été marqués sur un seul groupement méthyl par acide aminé (Ile- $\delta_1$ , Leu-proS, Val-proR). Ce marquage stéréospécifique permet à la fois de simplifier les spectres RMN pour n'obtenir qu'un signal par résidu marqué, mais également de réduire les interactions dipolaires entre les protons des groupements méthyles qui sont très proches dans l'espace (cf. partie 3.1.1.3). En effet, lors du marquage stéréospécifique de ces résidus, un des groupements méthyles est marqué  $^{13}\text{CH}_3$  tandis que l'autre groupement méthyle est marqué  $^{12}\text{C}^2\text{H}_3$ .

Les échantillons produits ont été les suivants :

- 1 Anticorps marqué U- $^1\text{H}$ ,  $^{13}\text{C}^1\text{H}_3$ ε-Met et U- $^2\text{H}$ ,  $^{13}\text{C}^1\text{H}_3$ β-alanine
- 2 Anticorps marqué U- $^1\text{H}$ ,  $^{13}\text{C}^1\text{H}_3$ ε-Met et U- $^2\text{H},^{15}\text{N}$ ,  $^{13}\text{C}^1\text{H}_3$ δ1-isoleucine
- 3 Anticorps marqué U- $^1\text{H}$ ,  $^{13}\text{C}^1\text{H}_3$ ε-Met et [U- $^2\text{H}$ ,  $^{13}\text{C}^1\text{H}_3$ ]proS-leucine
- 4 Anticorps marqué U- $^1\text{H}$ ,  $^{13}\text{C}^1\text{H}_3$ ε-Met et U- $^2\text{H}$ ,  $^{13}\text{C}^1\text{H}_3$ γ-thréonine
- 5 Anticorps marqué U- $^1\text{H}$ ,  $^{13}\text{C}^1\text{H}_3$ ε-Met et U- $^2\text{H},^{15}\text{N}$ ,  $^{13}\text{C}^1\text{H}_3$ proR-valine

La méthionine a été marquée dans tous les échantillons de façon à pouvoir référencer correctement la position des pics des autres résidus les uns par rapport aux autres lors de la reconstruction du spectre de l'anticorps entier par superposition spectrale. L'acquisition des spectres RMN  $^1\text{H}$ - $^{13}\text{C}$  SOFAST pour ces échantillons a donné les spectres présentés en figure 3.1.19.

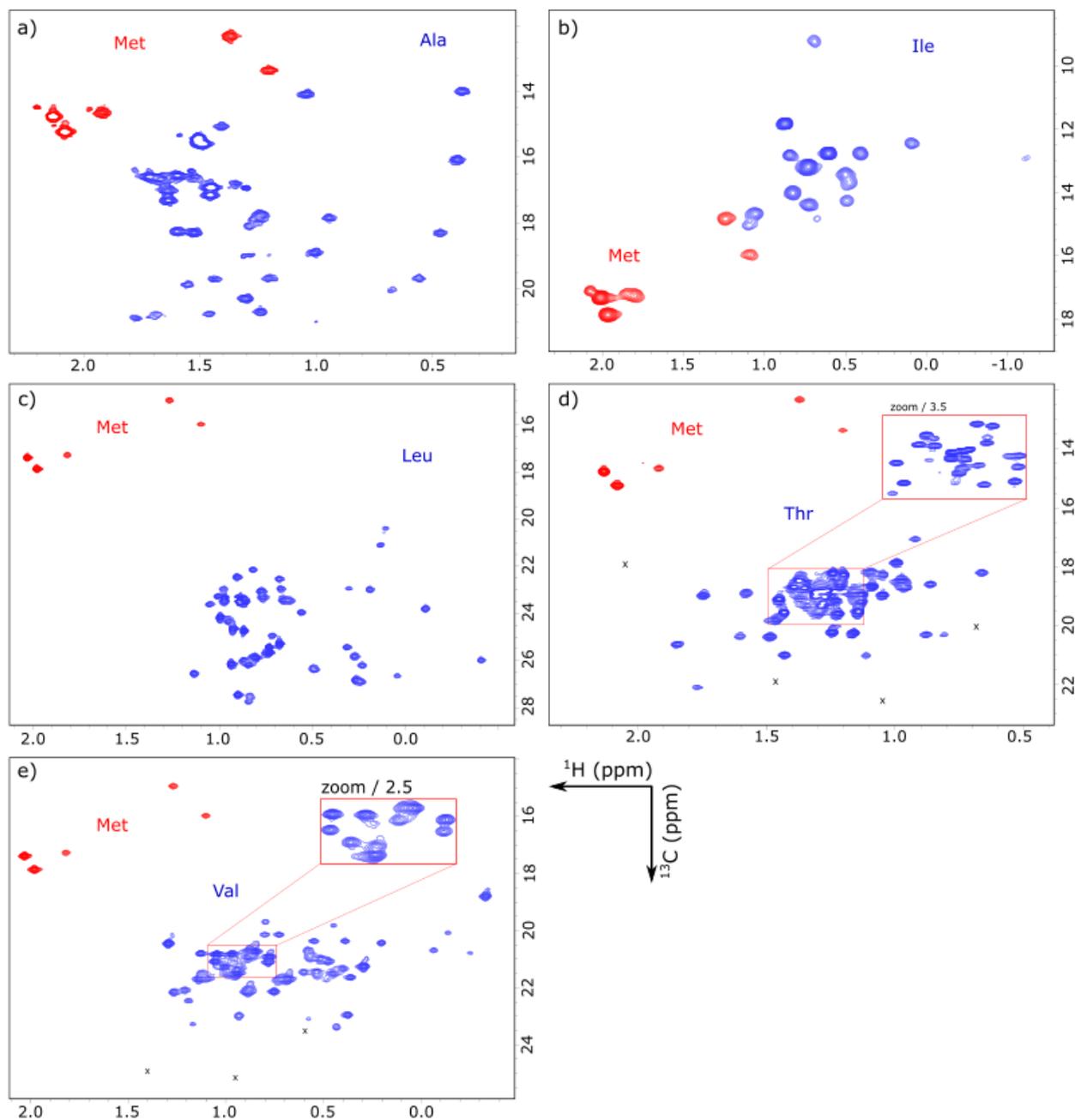


Figure 3.1.19 : Spectres RMN 2D  $^1\text{H}$ - $^{13}\text{C}$ -SOFAST dans la région des groupements méthyles de l'anticorps anti-LAMP1 produit en cellules CHO et marqué [ $^{13}\text{C}_3$ ]-méthionine et : a) [U- $^2\text{H}+^{13}\text{C}_3$ ]-alanine, b) [U- $^2\text{H}+^{13}\text{C}_3$ ]-isoleucine, c) [U- $^2\text{H}+^{13}\text{C}_3$ ]-proS]-leucine, d) [U- $^2\text{H}+^{13}\text{C}_3$ ]-thréonine, e) [U- $^2\text{H}+^{13}\text{C}_3$ ]-proR]-valine. Les expériences ont été acquises à 35°C sur des spectromètres équipés d'une sonde cryogénique opérant à une fréquence  $^1\text{H}$  de 850 MHz a) d), 950 MHz c) e) et 600 MHz b). Pour une meilleure représentation des pics, plusieurs niveaux de zoom ont été effectués pour les spectres d) et e) (carrés rouges). Les pics correspondant aux groupements méthyles des méthionines sont représentés en rouge sur chaque spectre.

Ces spectres ont permis d'identifier et de dénombrer les signaux correspondant à chaque type de résidu comportant un groupement méthyle (tableau 3.1.2). Toutes les alanines, isoleucines, leucines et méthionines sont observées sur ces spectres. Les pics manquant pour la thréonine et la valine s'expliquent par une superposition des pics.

Tableau 3.1.2 : Nombre de pics identifiés sur les spectres 2D  $^1\text{H}$ - $^{13}\text{C}$  SOFAST de l'IgG1 anti-LAMP1 marquée spécifiquement sur les groupements méthyles et produite en cellules CHO et quantité des résidus marqués dans la séquence de l'anticorps. \* /2 pour avoir le nombre de résidus non équivalent car l'anticorps est symétrique.

Acide aminé marqué	Nombre de pics dans le spectre $^1\text{H}$ - $^{13}\text{C}$ -SOFAST correspondant	Nombre de l'acide aminé dans la séquence de l'IgG1 anti-LAMP1 (/2*)
Alanine	35	35
Isoleucine	16	16
Leucine	48	48
Méthionine	5	5
Thréonine	53	54
Valine	53	60

Le développement du protocole de culture cellulaire a permis de mettre en place une méthode de culture adaptée à la production d'anticorps au laboratoire pour l'analyse RMN. Le développement d'un mélange d'acides aminés aux concentrations connues a été un succès. La connaissance de la composition de ce mélange d'acides aminés permet de le reproduire à façon selon le marquage isotopique souhaité et d'ajouter ensuite les acides aminés marqués voulus. Ce mélange d'acides aminés combiné à l'utilisation d'une haute densité cellulaire initiale ( $10 \cdot 10^6$  cellules/mL) a permis de restaurer les rendements d'expression d'anticorps non marqué par les cellules CHO obtenus avec des milieux commerciaux riches. Le rendement d'expression dans les conditions de culture développées a été estimé à 80 mg d'anticorps exprimé par litre de culture, ce qui permet d'obtenir 4 mg d'anticorps pour un volume de culture cellulaire de 50 mL, volume suffisant pour la production d'un échantillon marqué pour la RMN à des fins de recherche.

L'incorporation des acides aminés marqués en cellules CHO a pu être optimisée grâce à la mise en place d'une méthode d'affamement des cellules pour une élimination maximale des acides aminés non marqués contenu dans les cellules. Cette étape, réalisée juste avant de procéder à la production de l'anticorps marqué, a permis d'améliorer la qualité spectrale des signaux des isoleucines par un facteur 1.3.

Pour l'ajout de 300 mg/L d'acides aminés marqués et pour un ajout de 10 mg/L de L-cyclosérine lors du marquage de l'alanine, les taux d'incorporation des résidus comportant un groupement méthyle sont tous supérieurs à 84% voire 90% pour certains.

D'après Arbogast *et al.*, l'acquisition de spectres RMN 2D  $^1\text{H}$ - $^{13}\text{C}$  des groupements méthyles d'un anticorps non marqué (abondance naturelle) nécessite 12 h d'acquisition en utilisant un spectromètre RMN opérant à haut champs et haute température (900 MHz et 50°C)

et avec environ 250  $\mu\text{M}$  d'anticorps [120] (cf. partie 1.4.3.2). En prenant en compte le gain d'intensité obtenu lors du marquage isotopique en  $^{13}\text{C}$ , on peut faire l'approximation qu'une incorporation des acides aminés marqués sur les groupements méthyles de 80% diminue cette durée à moins d'une minute pour les mêmes conditions d'acquisition.

La méthode mise en place permet donc d'obtenir des spectres RMN 2D  $^1\text{H}$ - $^{13}\text{C}$  des groupements méthyles dans un temps réduit en utilisant 4 mg d'acides aminés marqués par échantillon produit à partir de 50 mL de culture cellulaire.

Le marquage isotopique réalisé (deutération des chaînes latérales et marquage stéréospécifique des groupements méthyles) a permis d'obtenir des spectres RMN 2D  $^1\text{H}$ - $^{13}\text{C}$  de très bonne qualité et il a été possible d'observer tous les pics correspondant aux alanines, isoleucines, leucines et méthionines. Une superposition de pics reste cependant présente pour les thréonines et plusieurs pour les valines.

Le développement de cette méthode de marquage isotopique d'anticorps produit en cellules CHO a permis d'obtenir des spectres RMN 2D d'une qualité suffisante pour envisager le transfert d'attribution.

La méthode développée pour le marquage isotopique des groupements méthyles de l'anticorps en cellules CHO a permis d'obtenir la première identification des types de résidus pour les signaux des spectres RMN 2D  $^1\text{H}$ - $^{13}\text{C}$  d'un anticorps. Cependant, cette identification est insuffisante pour pouvoir suivre la structure des anticorps en solution et à l'échelle atomique. Pour cela, l'attribution spécifique des signaux à chaque acide aminé portant un groupement méthyle doit être réalisée. Comme expliqué dans la partie 1.5.2, la stratégie qui a été adoptée pour répondre à cette problématique et de passer par l'utilisation d'expériences RMN de type HCC, puis de compléter l'attribution en acquérant des expériences NOESY inter-méthyles. Pour les expériences de type HCC, l'attribution du squelette doit être réalisée. Pour ces trois types méthodes d'attribution, et pour une protéine de la taille d'un anticorps doit être perdeutééré afin d'acquérir les expériences RMN 3D nécessaires aux attributions. Cela étant impossible en cellules CHO à cause de la toxicité du  $\text{D}_2\text{O}$  pour les cellules de mammifères, le système *cell-free* a été utilisé. De plus, afin de réduire la taille des protéines analysées pour améliorer la qualité des spectres RMN et faciliter les attributions, seuls les fragments Fab et Fc de l'anticorps anti-LAMP1 ont été synthétisés en système *cell-free*.

En raison d'un manque de temps, les groupements méthyles du fragment Fc n'ont pas été attribués lors de la thèse. La stratégie d'attribution pour ce fragment est différente de celle adoptée pour le fragment Fab en raison de la présence des glycosylations sur les asparagines 297. Cette stratégie est détaillée dans les perspectives de la thèse et la partie suivante traite du travail réalisé en système *cell-free* pour l'attribution des groupements méthyles du fragment Fab de l'anticorps anti-LAMP1.

## 2. Attribution du squelette du fragment Fab exprimé en système *cell-free*

### 2.1. Introduction à l'expression d'anticorps en milieu acellulaire

Le système acellulaire (*cell-free*), ou production de protéines *in vitro*, utilise un extrait cellulaire (eucaryote ou procaryote) comportant la machinerie cellulaire responsable de la transcription et de la traduction des protéines. Le gène codant la protéine d'intérêt est ajouté à cet extrait avec de nombreux réactifs (acides aminés, tampon, enzymes, chaperonnes et autres co-facteurs) pour faire fonctionner la machinerie cellulaire. La biosynthèse de composés en système *cell-free* était utilisée, dans les années 1960, pour comprendre le fonctionnement du code génétique[170]. Pour la fabrication des anticorps et/ou de fragments, les systèmes de production *in vivo* étaient préférés en raison des limitations du système *cell-free* : forte consommation de réactifs coûteux, faibles rendements, difficulté à augmenter l'échelle des réactions. De nos jours, ces limitations ont été surpassées grâce à de multiples avancées comme une meilleure compréhension de la formation des ponts disulfures, une meilleure connaissance des protéines chaperonnes ou alors l'obtention d'extrait cellulaire provenant de nouvelles souches cellulaires plus compétentes pour la synthèse de protéines à pont disulfures. Et aujourd'hui, grâce à ces avancées, des extraits cellulaires de diverses origines (eucaryotes ou procaryotes) ont déjà été utilisés pour la synthèse d'anticorps ou de fragments en milieu *cell-free*[171]–[174].

#### 2.1.1. Expression de protéines en système *cell-free*

Il existe deux méthodes de production en système *cell-free* : un système utilisant un extrait cellulaire et le système PURE (*protein synthesis using recombinant elements*). Le système PURE consiste à utiliser toutes les protéines et enzymes responsables de la transcription et de la traduction sans utiliser d'extrait cellulaire. Les protéines du système sont produites et purifiées au préalable. Cela a l'avantage de n'apporter que ce qui est nécessaire à la réaction, et d'éviter la présence de nombreuses protéases ou autres déchets cellulaires toujours présents dans les extraits. Mais l'inconvénient est le coût d'une telle procédure qui en fait un système plutôt utilisé pour des expressions à petites échelles (50 µL). Ainsi, le système *cell-free* se fait préférentiellement à partir d'extrait cellulaire, auquel sont ajoutés l'ADN de la protéine d'intérêt, des cofacteurs de transcription et traduction, un tampon, d'autres composés pour imiter le milieu intra-cellulaire, et un système énergétique permettant de faire fonctionner la machinerie cellulaire plus longtemps (figure 3.2.1). Pour la synthèse de protéines complexes à ponts disulfures comme les anticorps ou leurs fragments, des enzymes ou chaperonnes peuvent être ajoutées dans ce milieu de synthèse ouvert.

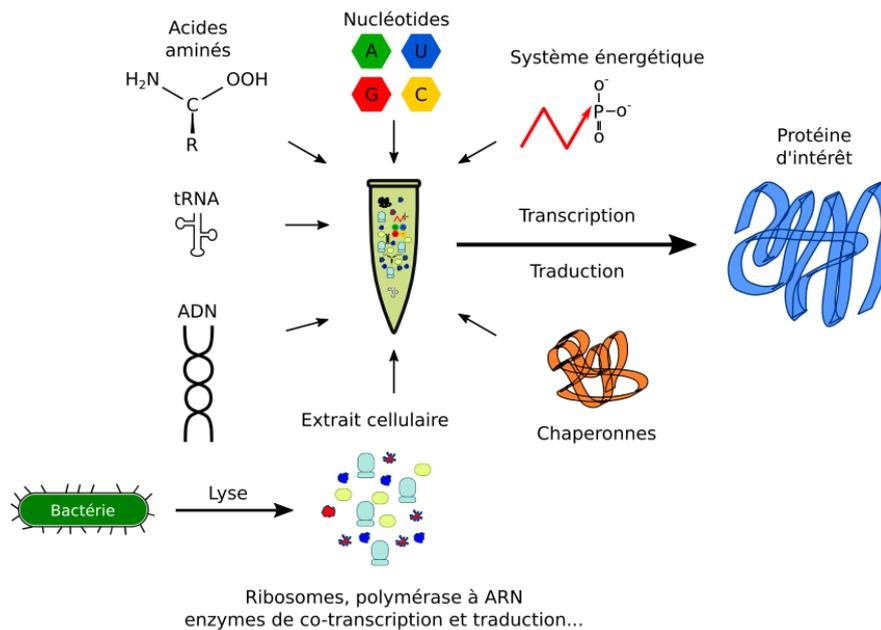


Figure 3.2.1 : Principaux composants du système cell-free

### a L'extrait cellulaire

La production *in vitro* de protéines consiste à réaliser la transcription de l'ADN en ARNm, puis la traduction de l'ARNm en protéine dans un tube. Pour cela, un extrait cellulaire contenant les ribosomes, des cofacteurs de traduction, des ARNt et l'ARN polymérase est utilisé. L'extrait cellulaire est la composante principale du système *cell-free*. Cet extrait peut être d'origine eucaryote ou procaryote. Les systèmes *cell-free* basés sur un extrait cellulaire procaryote utilisent principalement un extrait provenant de souches de *E. coli*. Ceux d'origine eucaryote utilisent un extrait cellulaire provenant de plusieurs types de cellules comme les cellules de germe de blé et les réticulocytes de lapin qui sont les plus utilisés.

En dehors des enzymes nécessaires à la réaction, l'extrait peut également apporter des nucléases, protéases ou réductases ayant un impact négatif sur la production de la protéine. C'est pour cela que plusieurs souches cellulaires présentant des modifications génétiques spécifiques sont utilisées pour la préparation de ces extraits. Par exemple, la souche SHuffle de *E. coli* peut être utilisée car elle surexprime la DsbC, une isomérase jouant un rôle majeure dans la formation et l'isomérisation des ponts disulfures (cf. partie 1.3.1.2.b) et est déficiente en réductases *trxB* et *gor*. Bien que l'utilisation de ces souches soit un avantage pour faciliter le protocole de production en évitant l'ajout de protéines chaperonnes (qu'il faut produire et purifier au préalable), les rendements d'expression sont souvent inférieurs à ceux obtenus pour les extraits de souches standards complétés avec les protéines chaperonnes[175].

Pour la synthèse d'anticorps et de leurs fragments, des extraits obtenus à partir de CHO, de levures (*Saccharomyces cerevisiae*), cellules de tabac BY-2 ou encore cellules d'insectes Sf21 (*Spodoptera frugiperda 21*) ont été utilisés. L'utilisation d'un extrait eucaryote permet d'apporter les enzymes nécessaires à la formation de modifications inexistantes en bactéries (comme les glycosylations). Cependant, les rendements restent nettement inférieurs à ceux

obtenus avec un système *cell-free* procaryote. La préparation de l'extrait cellulaire est donc une étape déterminante, et le choix du système d'origine est dépendant de la protéine à exprimer et des modifications que l'on souhaite obtenir (principalement les glycosylations et ponts disulfures pour les anticorps et/ou leurs fragments).

Les extraits cellulaires d'origines eucaryotes font de plus en plus leur apparition pour la production d'anticorps et de fragments [172], [173], [176], mais les rendements d'expression sont plus faibles. Par exemple, pour des anticorps ou leurs fragments, les rendements d'expression sont de 0.9 à 1.6  $\mu\text{g/mL}$  pour des extraits de cellules de mammifères [176], ou encore 20  $\mu\text{g/mL}$  pour les extraits de cellules d'insectes [172]. Les extraits bactériens peuvent de leur côté, permettre d'obtenir des rendements d'expression de l'ordre du g/L dans certains cas [177], [178]. Ainsi, le détail du système *cell-free* se concentrera sur le système bactérien car il s'agit du plus performant en terme de rendement de production, et de celui utilisé pendant la thèse.

### b Système de transcription et traduction couplée

La transcription et la traduction des protéines en système *in vitro* peuvent se faire de deux façons : soit les deux étapes sont faites séparément, soit elles sont réalisées lors de la même réaction. Il s'agit alors du système *cell-free* à transcription et traduction couplées qui est le système principalement utilisé (figure 3.2.2).

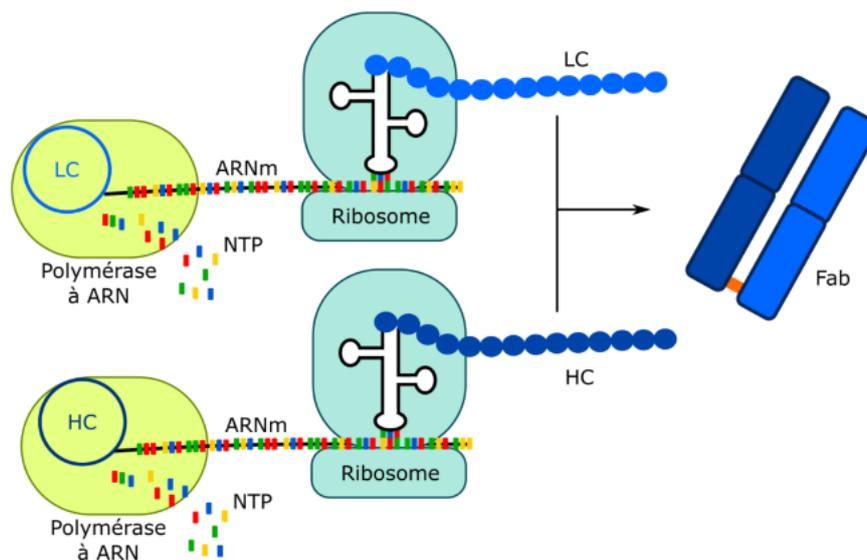


Figure 3.2.2 : Système de transcription et traduction couplées du système *cell-free* avec l'exemple de co-expression de deux plasmides comme pour le cas de la chaîne lourde et légère d'un fragment Fab d'anticorps. La synthèse des deux chaînes est réalisée au cours d'une même expression.

L'extrait cellulaire seul ne suffit pas car lors de sa préparation, de nombreux cofacteurs de transcription et traduction peuvent être perdus. Ainsi, en plus des acides aminés et de l'ADN portant le gène de la protéine d'intérêt, le système est complété avec de l'ARN polymérase (T7-ARN polymérase) et des ARN de transfert. Les nucléotides (ATP, GTP, UTP, CTP) sont ajoutés pour composer l'ARNm et faire partie du système énergétique.

L'absence d'environnement cellulaire peut être agressive pour certaines enzymes provenant de l'extrait. L'assemblage des différents sous-domaines des ribosomes peut être garanti par l'ajout de spermidine qui a également un effet stabilisateur de l'ARNm. La spermidine est un polycation et agit sur la viscosité du milieu qui est importante pour le maintien structural de cette machinerie. Dans la même optique, des agents chimiques comme le glutamate de potassium ou l'acétate d'ammonium peuvent être utilisés afin d'offrir au système une viscosité proche de celle présente dans les cellules[144].

Un tampon (de pH optimisé pour chaque extrait cellulaire et protéine exprimée) est utilisé de manière à être le plus proche des conditions d'un milieu cellulaire propice à l'activité des différentes enzymes.

Un autre composant essentiel à la réaction en système *cell-free* est le magnésium. *In vivo*, les ions  $Mg^{2+}$  jouent un rôle dans de nombreuses réactions et sont l'un des cations divalents les plus abondants dans les cellules[179]. Ils sont notamment impliqués dans l'assemblage des ribosomes et le fonctionnement de nombreuses enzymes et cofacteurs intervenant lors de la transcription et la traduction des protéines. La concentration de  $Mg^{2+}$  est réglée avec l'ajout de glutamate ou acétate de magnésium et est optimisée pour chaque extrait cellulaire et protéine d'intérêt à exprimer.

### c Le système énergétique

L'ARN polymérase, les ribosomes, et les ARNt synthétases ont besoin d'une source d'énergie pour fonctionner. Cette source d'énergie est apportée par un système énergétique basé sur la libération de groupements phosphates de l'ATP lorsqu'il se transforme en ADP (figure 3.2.3). En raison de la consommation d'ATP comme source d'énergie, sa concentration est plus importante que pour les autres nucléotides dans le système (1.2 mM contre 0.8 mM). L'énergie ainsi libérée est consommée, par exemple, lors de l'ajout des acides aminés sur les différents ARN de transfert ou lors de la régénération des nucléotides.

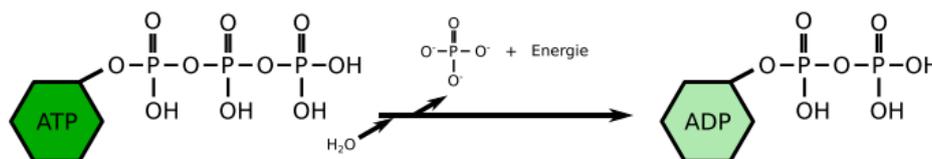


Figure 3.2.3: Transformation de l'ATP en ADP représentant la source d'énergie utilisée en système *cell-free*

Un bon rendement d'expression en système *cell-free* n'est possible qu'avec un système énergétique suffisamment robuste, et pouvant se régénérer pour faire perdurer la synthèse de protéine. Un système de régénération est alors utilisé avec l'ajout de créatine phosphate et de créatine kinase. La première apporte un groupement phosphate pouvant être transféré sur l'ADP pour le régénérer en ATP, et la deuxième catalyse la réaction (figure 3.2.4).

La molécule d'ADP contient deux autres groupements phosphate dont un pourrait encore être consommé comme source d'énergie. Cela entraînerait la production d'AMP qui ne peut être régénéré par la créatine kinase et la créatine phosphate. Pour éviter l'utilisation du phosphate de l'ADP, on peut ajouter de l'AMP cyclique (AMPC) qui va avoir pour effet de

bloquer l'équilibre de la réaction et éviter la consommation de l'ADP et favoriser ainsi sa régénération en ATP (figure 3.2.4).

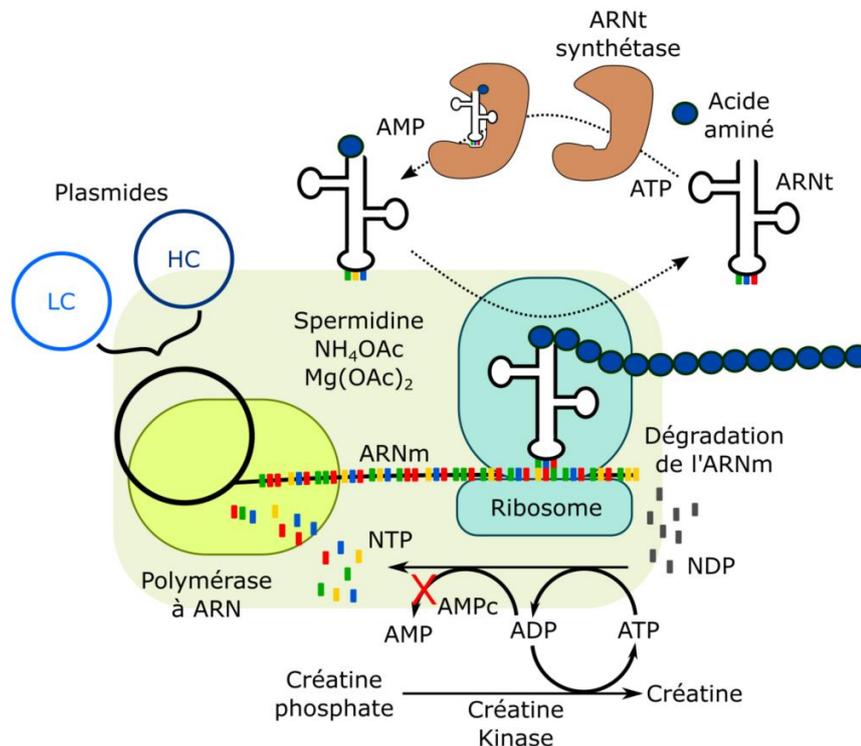


Figure 3.2.4 : Schéma représentant le mécanisme principal du système cell-free. Le schéma est donné comme exemple pour la co-expression de protéines comme c'est le cas pour l'expression à partir des plasmides de chaîne légères et lourdes des anticorps ou de leurs fragments.

### 2.1.2. Expression d'anticorps et fragments en système *cell-free*

Le *cell-free* n'offre pas de compartiments spéciaux pour le repliement des protéines et la formation des ponts disulfures comme c'est le cas dans les cellules eucaryotes (réticulum endoplasmique) et procaryotes (périplasma). Tout a lieu au même endroit, et pour des protéines à pont disulfures comme les anticorps (ou les fragments), un milieu oxydant est nécessaire. Cependant, lors de la préparation standard de l'extrait (bactérien), des agents réducteurs, comme le DTT (dithiothréitol), sont utilisés pour préserver l'activité de la machinerie cellulaire. L'extrait provenant de certaines souches bactériennes peut également contenir des réductases endogènes comme la réductase à thiorédoxine (trxB) et la réductase à glutathion (gor) chez *E. coli* (cf. partie 3.3). Plusieurs solutions ont été développées pour permettre la formation de ponts disulfures en système *cell-free*. Tout d'abord, les agents réducteurs sont évités voire supprimés du système, et remplacés par des couples oxydo-réducteurs comme le glutathion réduit et oxydé (GSH/GSSG) qui ajustent un potentiel redox approprié à la formation des ponts disulfures [180]. L'extrait peut être traité avec des inhibiteurs enzymatiques comme l'iodoacétamide (IAM) afin de neutraliser l'activité des réductases. Etant donnée leur nature non spécifique, certaines enzymes nécessaires au système peuvent aussi être inhibées. On peut alors éviter l'utilisation d'inhibiteur en produisant l'extrait à partir de souches modifiées comme la souche SHuffle de *E. coli* qui n'exprime pas la trxB et la gor (en plus de surexprimer l'isomérase DsbC) [175]. Le système *cell-free* est très utilisé pour l'expression de fragments

Fab ou scFv d'anticorps à des fins de recherche et permet d'obtenir rapidement (quelques heures) des rendements de production de l'ordre de 0.3 mg/mL de réaction *cell-free* (pour les systèmes utilisant des extraits bactériens) [178].

Le système *cell-free* est un système ouvert. L'absence des différents compartiments cellulaires n'offre pas la même protection aux protéines, mais cela à l'avantage de permettre l'ajout de n'importe quelle enzyme ou chaperonne purifiées. De nombreuses chaperonnes existent et ont permis d'améliorer la production de fragments scFv ou Fab solubles comme la DnaK, la DnaJ, la GroEL, ou la GroES [180]. L'ajout d'isomérases à pont disulfures permet également d'augmenter nettement le rendement de production de protéines à ponts disulfures. L'utilisation d'une souche *E. coli* modifiée, l'adaptation du milieu et l'ajout de DsbC et de FkpA exogènes ont, par exemple, permis la production d'une IgG entière à un rendement de 1 g/L de réaction [177].

Une approche différente a consisté à utiliser un extrait provenant des cellules d'insectes Sf21 apportant des vésicules remplissant le rôle du réticulum endoplasmique. L'ajout d'un peptide signal a permis le transfert dans les vésicules de fragments scFv d'IgG [172]. Les rendements de production restent cependant faibles (20 µg/mL de réaction).

Le système *cell-free* offre un avantage supplémentaire par rapport aux autres systèmes. Le fait que le système soit ouvert et la bonne connaissance de la machinerie cellulaire de *E. coli* peuvent permettre l'incorporation d'acides aminés non naturels dans la protéine. Pour cela, des ARN de transfert (ARNt) orthogonaux au système et spécifiques à un acide aminé non naturel sont utilisés [181], [182]. L'incorporation de l'acide aminé non naturel à un endroit spécifique de la séquence se fait ensuite par l'utilisation de codons spécifiques à ces ARNt orthogonaux. Les acides aminés non naturels peuvent être très intéressants pour de multiples techniques analytiques pour l'étude structurale des protéines, ou encore pour la production d'anticorps couplés à des molécules actives. Ainsi, l'utilisation du *cell-free* avec cette technique d'incorporation d'acides aminés non naturels commence à faire son apparition pour la production d'ADC thérapeutiques[183]. Le principe est d'incorporer, à un ou des endroits spécifiques de l'anticorps, un (ou des) acides aminés non-naturels portant des groupements chimiques réactifs favorisant la fixation de la molécule active. Cette technologie est encore en phase de développement mais semble prometteuse.

## 2.2. Marquage isotopique $^{15}\text{N}$ , $^{13}\text{C}$ du squelette des protéines en système *cell-free* pour la RMN

Le marquage isotopique  $^{15}\text{N}$ ,  $^{13}\text{C}$  en système *cell-free* est réalisé par ajout des acides aminés marqués lors de la réaction. Tout comme les systèmes *in vivo*, le système *cell-free* est lui aussi impacté par le phénomène de fuites isotopique auquel il faut prêter attention pour obtenir le meilleur taux d'incorporation des isotopes.

### 2.2.1. La fuite isotopique des $^{15}\text{N}$ du squelette des protéines exprimées en système *cell-free*

Les systèmes *in vivo* sont fortement impactés par les phénomènes de dilution et de fuite isotopique. La dilution a lieu lors de la biosynthèse d'acides aminés par les cellules à partir de composés du milieu ou à cause de la présence d'acides aminés non marqués dans le milieu. Pour le système *cell-free*, la préparation des extraits bactériens est réalisée à partir de bactéries mises en culture dans un milieu contenant des acides aminés non marqués. De ce fait, les extraits contiennent des acides aminés non marqués, mais il est possible de les éliminer par purification de l'extrait sur une colonne de dessalage [184]. Ce système *cell-free* est également concerné par le phénomène de fuite isotopique et ce, notamment, lors du marquage isotopique spécifique  $^{15}\text{N}$  sur le squelette des protéines. En effet, la préparation de l'extrait bactérien supprime la plupart des enzymes impliquées dans les voies de biosynthèse des acides aminés en bactéries [185], [186]. Cependant, certaines transaminases restent présentes et actives, notamment celles dépendantes du facteur pyridoxal 5-phosphate (PLP) qui sont responsables de la transamination des acides aminés avec un keto-acide. Cette fuite isotopique concerne tous les acides aminés à plus ou moins grande échelle, impactant en particulier l'asparagine, l'aspartate, la glutamine et le glutamate [188].

La fuite isotopique lors d'un marquage isotopique spécifique d'un acide aminé entraîne le marquage isotopique d'un ou d'autres acides aminés par transfert des isotopes par les transaminases (figure 3.2.5). Le spectre RMN obtenu dans ce cas est inexploitable en raison d'un nombre de résonances non identifiées et plus élevé que prévu.

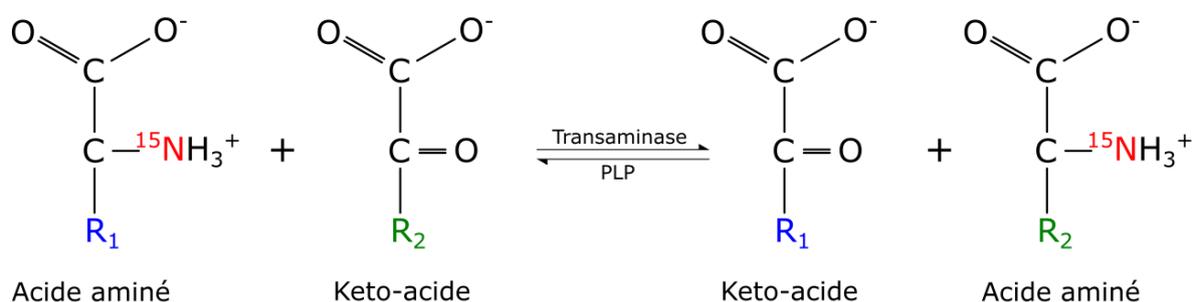


Figure 3.2.5 : Fuite isotopique d'un  $^{15}\text{N}$  d'un acide aminé à un autre par l'action d'une transaminase dépendante du facteur pyridoxal 5-phosphate (PLP) avec un keto-acide.

Pour limiter l'action de ces enzymes, des souches de bactéries auxotrophes ont été développées [186]. Cependant, le développement de nouvelles souches peut être compliqué et toutes les enzymes ne peuvent être éliminées de cette façon.

### 2.2.2. La fuite isotopique et la deutération des protéines en système *cell-free*

La fuite isotopique causée par ces enzymes ne concerne pas uniquement le marquage  $^{15}\text{N}$  sur le squelette des protéines. Pour l'attribution du squelette d'une grosse protéine par RMN, des acides aminés  $^{15}\text{N}$ ,  $^{13}\text{C}$  et deutérés doivent être utilisés. Si la réaction de synthèse est réalisée en solvant protoné, alors les atomes de deutérium liés aux  $\text{C}\alpha$  et  $\text{C}\beta$  des acides aminés sont susceptibles d'être échangés avec les protons du solvant par l'action des transaminases. La fuite isotopique concernant les  $^{15}\text{N}$  du squelette n'a un impact sur les spectres RMN qu'en cas de marquage spécifique (seulement un ou quelques acides aminés à la fois). L'échange des protons en position  $\text{C}\alpha$  et  $\text{C}\beta$  concerne les marquages spécifique et uniforme. En effet, tout dépend de la composition du milieu. Il y aura un échange si la deutération entre les acides aminés et le milieu n'est pas la même (ex : résidus deutérés et milieu 100%  $^1\text{H}_2\text{O}$  ou résidus protonés et milieu 100%  $^2\text{H}_2\text{O}$ ).

*In vivo*, il a été proposé d'utiliser des acides aminés marqués en excès lors de la production de la protéine (3 g/L d'extrait d'algue marqué U- $^{2}\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$ ) en milieu 100% protoné. Cela permet d'obtenir des protéines complètement protonées ( $^1\text{H}$ ) sur les groupements amides du squelette tandis que 85 % des autres positions de protons seraient théoriquement deutérées [189]. Cependant, cette deutération est impactée par l'action des transaminases et peut être responsable de la reprotonation de 30 à 80% des  $\text{C}\alpha$  dans ce genre de cas [190]. En système *cell-free*, même si la plupart des enzymes sont supprimées lors de la préparation de l'extrait, cet échange s'élève tout de même à 25-50% (en fonction des acides aminés) [191].

Une autre méthode pour éviter la reprotonation des  $\text{C}\alpha$  et  $\text{C}\beta$  par les enzymes est d'exprimer la protéine en milieu *cell-free* 100% deutéré. Ainsi, les enzymes sont toujours actives mais échangeront les atomes de deutérium des acides aminés avec les atomes de deutérium du milieu. Cependant, comme expliqué dans la partie 6, cette méthode présente un inconvénient pour l'acquisition de spectres RMN basés sur les groupements amides du squelette lors du marquage isotopique  $^{15}\text{N}$ ,  $^{13}\text{C}$ . En effet, pour les protéines de haut poids moléculaire, les  $\text{H}_\text{N}$  du cœur de la protéine s'échangent lentement avec le solvant. La protonation des groupements amides doit alors se faire par une étape de dépliement et de repliement qui est une source de perte de protéine marquée et coûteuse. De plus, la production de protéines en système *cell-free* 100% deutéré a pour inconvénient de diminuer les rendements d'expression d'un facteur deux en moyenne [191].

### 2.2.3. L'avantage du système *cell-free* face à la fuite isotopique

Comparé aux systèmes *in vivo*, le système *cell-free* a un réel avantage face à la fuite isotopique. Il s'agit d'un système ouvert qui permet l'ajout de nombreux composés comme des inhibiteurs enzymatiques ne pouvant être utilisés en système *in vivo* car ceux-ci seraient toxiques pour les cellules.

Plusieurs inhibiteurs compatibles avec la production de protéines en système *cell-free* existent. Xun-Cheng Su *et al.* ont rapporté que le tétrahydroborate de sodium ( $\text{NaBH}_4$ ) a un effet inhibiteur sur l'activité des enzymes PLP [192]. Le traitement de l'extrait bactérien au  $\text{NaBH}_4$  avant son utilisation permet en effet de réduire de manière irréversible les intermédiaires réactionnels nommés bases de Schiff qui se forment entre la chaîne latérale des lysines des enzymes et le cofacteur PLP. Cet inhibiteur a l'avantage de ne pas influencer sur le rendement

d'expression de la protéine car ces enzymes ne sont pas nécessaires lors des étapes de transcription et traduction, mais l'activité de toutes les enzymes PLP ne sont pas inhibées.

Pour éviter la plupart des autres sources de fuite isotopique, l'activité d'autres enzymes peut être inhibée par l'ajout d'aminooxyacétate (AOA) et d'acide D-malique [188]. Les deux enzymes ciblées par ces inhibiteurs permettent d'éviter la reprotonation des C $\alpha$  des alanines, glutamates, phénylalanines, tyrosines, tryptophanes et des C $\beta$  des asparagines et aspartates. La 5-diazo-4-oxo-L-norvaline (DONV) et la 6-diazo-5-oxo-L-norleucine (DON) permettent d'inhiber les enzymes responsables de la fuite isotopique vers les glutamines, prolines, glycines arginines, alanines et spécifiquement vers les aspartates (DONV) et glutamates (DON). La glutamine synthétase peut être inhibée par l'ajout de L-méthionine sulfoximine, la thréonine déshydrogénase par l'ajout de L-allothréonine et la Sérine hydroxyméthyltransférase par l'ajout de 5-formyl-tetrahydrofolate pyroglutamate. Ces exemples sont issus de la revue de B. Hoffmann *et al.* et le tableau 3.2.1 résume la liste des principales enzymes responsables de la fuite isotopique en système *cell-free*, ainsi que les inhibiteurs correspondant pouvant être utilisés et les acides aminés concernés par ces fuites [186] [185].

Tableau 3.2.1 : Synthèse des principales enzymes présentes dans l'extrait bactérien responsables de fuites isotopiques entre acides aminés marqués, des inhibiteurs pouvant être utilisés et des acides aminés concernés par les fuites. Cette liste est tirée de l'article de B. Hoffmann *et al.* Progress in Nuclear Magnetic Resonance Spectroscopy, vol. 105, pp. 1–22, Apr. 2018, doi: 10.1016/j.pnmrs.2017.11.004.

Enzymes responsables de fuites isotopiques	Inhibiteurs possibles	Acides aminés concernés
Aspartate transaminase	AOA, NaBH <sub>4</sub>	Glu, Gln, Pro, Gly, Arg, Ala
Alanine transaminase	AOA, NaBH <sub>4</sub>	Ala, Gln, Pro, Gly, Arg
Glutamine synthétase	L-méthionine sulfoximine	Glu
Aspartate ammonia lyase	Acide D-malique	Gln, Glu, Pro, Gly, Arg, Ala
Asparaginase	DONV	Asp, Gln, Pro, Gly, Arg, Ala
Glutaminase	DON	Glu, Gln, Pro, Gly, Arg, Ala
Threonine dehydrogenase	L-allothreonine	Gly, Ser, Glu, Gln, Pro, Arg, Ala
2-amino-3-ketobutyrate CoA ligase	NaBH <sub>4</sub>	Gly, Glu, Gln, Pro, Arg, Ala, (Ser)
Sérine hydroxyméthyltransférase	5-formyl-tetrahydrofolate pyroglutamate	Ser (suite à l'action de l'enzyme précédente)
Thréonine deamidase	NaBH <sub>4</sub>	Gln, Ile, Glu, Gln, Pro, Arg, Ala, Gly

Inhiber l'intégralité des enzymes présentes dans l'extrait bactérien n'est pas envisageable en raison de leur nombre et du fait qu'elles ne soient pas toutes sensibles aux mêmes inhibiteurs car certains inhibiteurs ont un effet négatif sur le rendement d'expression de la protéine d'intérêt. Cependant, lorsqu'ils sont utilisés judicieusement, ils permettent une nette diminution des fuites isotopiques et donc, une amélioration du signal RMN. Ainsi, Aisha L.G. *et al.* ont montré que l'ajout de seulement trois inhibiteurs (AOA, acide D-malique, D-cyclosérine) a permis d'améliorer le signal RMN tout en ayant une perte de rendement limitée à 15% [185].

Un autre exemple de l'utilisation d'inhibiteurs est présenté par Imbert *et al.* qui utilisent un extrait de bactéries préalablement traité au NaBH<sub>4</sub>, et ajoutent des inhibiteurs AOA et acide D-malique lors de la réaction de synthèse. Leur protocole a permis de diminuer la reprotonation en position C $\alpha$  des résidus jusqu'à 10% pour certains résidus et même 5% pour la plupart, au lieu de 25 à 50% sans inhibiteur [191].

### **2.3. Article : Enabling Site-Specific NMR Investigations of Therapeutic Fab using a Cell-Free Based Isotopic Labeling approach : Application to anti-LAMP1 Fab.**

L'article présenté ci-dessous décrit les résultats obtenus pour l'attribution du squelette du fragment Fab de l'anticorps anti-LAMP1. La mise en place et les optimisations du protocole d'expression du fragment Fab en système cell-free détaillées ont tout d'abord permis d'améliorer les rendements de production et de diminuer les coûts de production des échantillons marqués. Le fragment Fab U-[<sup>2</sup>H,<sup>13</sup>C,<sup>15</sup>N] a été produit avec le protocole optimisé, permettant d'obtenir 2.8 mg de protéine.

Puis, la méthode utilisée pour l'attribution des résonances RMN des H<sub>N</sub>, N<sub>H</sub>, C<sub>O</sub>, C<sub>α</sub>, C<sub>β</sub> est abordée. L'attribution des signaux RMN pour le squelette a été réalisée en utilisant CCP NMR3 et un logiciel d'aide à l'attribution, développé par le Dr Adrien Favier, ingénieur de recherche dans le groupe de RMN biomoléculaire de l'IBS. Enfin, j'ai également étudié, par RMN <sup>1</sup>H-<sup>15</sup>N en solution, l'interaction du fragment Fab avec son antigène LAMP1.

Ces travaux nous ont permis de proposer la première attribution du squelette du fragment Fab anti-LAMP1, complète à 90%.

Cet article a été envoyé au service de propriété intellectuelle de Sanofi le 21 Juillet 2023 et nous sommes actuellement en attente de leur validation avant de le soumettre à un journal scientifique pour qu'ils soit revu.

Résultats Article : Enabling site-specific NMR investigation of therapeutic Fab using a cell-free based isotopic labeling approach : application to anti-LAMP1 Fab.

## **Enabling Site-Specific NMR Investigations of Therapeutic Fab**

**using a Cell-Free Based Isotopic Labeling Approach:**

**Application to anti-LAMP1 Fab.**

Arthur Giraud<sup>a,b,c#</sup>, Lionel Imbert<sup>a#</sup>, Adrien Favier<sup>a</sup>, Faustine Henot<sup>a,b</sup>, Francis Duffieux<sup>b</sup>, Camille Samson<sup>b</sup>, Oriane Frances<sup>b\*</sup>, Elodie Crublet<sup>c\*</sup>, Jérôme Boisbouvier<sup>a\*</sup>

<sup>a</sup> *Univ. Grenoble Alpes, CNRS, CEA, Institut de Biologie Structurale (IBS), 71, Avenue des Martyrs, F-38044 Grenoble, France.*

<sup>b</sup> *Sanofi Research & Development, 94403 Vitry-sur-Seine, France.*

<sup>c</sup> *NMR-Bio, 5 place Robert Schuman, F-38025 Grenoble, France.*

*# These authors contributed equally to this work*

*\* Correspondence to be addressed to [Oriane.Frances@sanofi.com](mailto:Oriane.Frances@sanofi.com) or [crublet@nmr-bio.com](mailto:crublet@nmr-bio.com) or [jerome.boisbouvier@ibs.fr](mailto:jerome.boisbouvier@ibs.fr)*

*Number of characters: 49 371*

*Number of words: 7295*

## **Enabling Site-Specific NMR Investigations of Therapeutic mAb Fragments using a Cell-Free Based Isotopic Labeling Approach:**

### **Application to anti-LAMP1 Fab.**

#### **Abstract:**

Monoclonal antibodies (mAbs) are biotherapeutics that have achieved outstanding success in treating many life-threatening and chronic diseases. The recognition of an antigen is mediated by the fragment antigen binding (Fab) regions composed by four different disulfide bridge-linked immunoglobulin domains. NMR is a powerful method to assess the integrity, the structure and interaction of Fabs, but site specific analysis has been so far hampered by the size of the Fabs and the lack of approaches to produce isotopically labeled samples. We proposed here an efficient *in vitro* method to produce [<sup>15</sup>N, <sup>13</sup>C, <sup>2</sup>H]-labeled Fab enabling high resolution NMR investigations of these powerful therapeutics. As an open system, the cell-free expression mode enables fine-tuned control of the redox potential in the presence of disulfide bond isomerase to enhance the formation of native disulfide bonds. Moreover, inhibition of transaminases in the S30 cell-free extract offers the opportunity to produce perdeuterated Fab sample directly in <sup>1</sup>H<sub>2</sub>O medium, without the need for a time-consuming and inefficient refolding process. This specific protocol was applied to produce optimally labeled sample of a therapeutic Fab, enabling the first sequential assignment of <sup>1</sup>H<sub>N</sub>, <sup>15</sup>N, <sup>13</sup>C', <sup>13</sup>C<sub>α</sub>, <sup>13</sup>C<sub>β</sub> resonances of a full-length Fab. 90 % of the backbone resonances of Fab domain directed against the human LAMP1 glycoprotein were assigned successfully, opening new opportunities to study at atomic resolution Fabs high order structures, dynamics and interactions, using solution-state NMR.

**Keywords:** Fab, mAb, LAMP1, Cell-Free, Isotope Labeling, perdeuteration, NMR, Backbone, Assignment.

## Introduction

Monoclonal antibodies (mAbs) have become the fastest growing class of therapeutics owing to their beneficial impacts in treating many life-threatening diseases.<sup>1</sup> The development of protein-based pharmaceuticals is challenging, and requires overcoming various manufacturing hurdles such as purity, stability and characterization of product structure. The 3D structure is unique to biological products and is directly linked to their function. It is a key parameter to consider when assessing the robustness of pharmaceutical biomanufacturing. Chemical modifications commonly encountered during bioproduction and long-term storage of mAbs can potentially affect structure and alter therapeutic activities. The biophysical tools currently available in the pharmaceutical industry (fluorescence, DSC, CD) for characterizing therapeutic mAbs under development do not have sufficient resolution to highlight subtle changes in mAb structure that nevertheless have an impact on potency. NMR methods have recently been introduced to address this challenge, enabling the generation of a well-resolved spectral map offering a complete atomic fingerprint of a mAb's primary, secondary, tertiary and quaternary structures.<sup>2,3</sup> Structural modifications, such as methionine oxidation, induced by oxidative stresses applied to a mAb batch, can be detected using <sup>13</sup>C-natural abundance 2D-NMR spectra of mAb.<sup>4</sup> However, from these 2D NMR fingerprint spectra, it is not yet possible to precisely localize the corresponding modifications in the 3D structure, and thus to highlight the direct impact of the structural modification on biological activity. To gain insight at the atomic level, it is necessary to assign each detected NMR signal to the corresponding atom in the 3D structure of the mAb. Assigning NMR frequencies would make it possible to decipher all the structural information encoded in 2D NMR spectra, hitherto unexploited in 2D NMR fingerprinting studies of mAbs. Prior knowledge of a mAb's NMR signals would enable in-depth characterization of any structural modifications using NMR spectra acquired directly on clinical batches of mAbs. However, the assignment of a 150 kDa mAb requires the production of samples uniformly enriched in <sup>2</sup>H, <sup>15</sup>N and <sup>13</sup>C isotopes, which has been the main bottleneck for atomic-resolution NMR studies of mAbs to date.

In this article, we present a general method for the cell-free production of optimally labeled Fabs, enabling high-resolution NMR studies. Our strategy is based on the in vitro co-expression of Fab light chain (LC) and heavy chain (HC), in the presence of the disulfide isomerase DsbC and GSH/GSSG to enable native pairing of the Fab disulfide bonds produced using the S30 extract of the *E. coli* SHuffle strain (Fig. 1). Transaminases in the cell-free extract were inactivated prior to protein production, to allow incorporation of unmodified U-[<sup>2</sup>H, <sup>13</sup>C, <sup>15</sup>N] amino acids. This approach was applied to characterize the NMR frequencies of the therapeutic Fab directed against lysosome-associated membrane protein 1 (LAMP1<sup>5</sup>). LAMP1 is a glycosylated protein found mainly on the membrane of lysosomes, which are involved in many aspects of cell biology and can influence cellular processes such as phagocytosis, autophagy, lipid transport and aging.<sup>6,7</sup> Furthermore, LAMP1 has been shown to be involved in tumor growth and metastatic spread.<sup>7-9</sup> The anti-LAMP1 mAb was developed to interact

Résultats Article : Enabling site-specific NMR investigation of therapeutic Fab using a cell-free based isotopic labeling approach : application to anti-LAMP1 Fab.

with LAMP1 overexpressed on the surface of cancer cells. Its importance underlines the need to better understand the link between the structure of anti-LAMP1 and its activity.

Our cell-free protocol was applied to produce 3 mg of U-[ $^2\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$ ]-labeled anti-LAMP1 Fab NMR sample and has led us, for the first time, to assign more than 90% of a Fab backbone NMR signals. This result paves the way for a specific study of dynamics and interactions, and enables us to better track changes in the higher-order structure (HOS) of this Fab that have an impact on its potency.

## Results

The site-specific analysis of a protein using NMR requires the assignment of each NMR signal to the corresponding atom in the protein. The identification of NMR frequencies is usually achieved using a series of triple resonance experiments on a sample uniformly enriched in  $^{13}\text{C}$  and  $^{15}\text{N}$  isotopes in order to sequentially link the backbone resonances of successive amino acids.<sup>10</sup> However, these 3D NMR experiments are limited by spin relaxation, and the perdeuteration of non-exchangeable protons is necessary to sequentially assign backbone NMR frequencies for proteins over 30 kDa.<sup>11,12</sup> With a total molecular weight of 47 kDa, the sequence-specific NMR study of a Fab requires the large-scale synthesis (> 1 mg) of a sample uniformly enriched in  $^2\text{H}$ ,  $^{13}\text{C}$  and  $^{15}\text{N}$ , with all backbone amino protons exchanged for the  $^1\text{H}$  isotope. This is usually achieved by overexpression of the target protein in *E. coli* in a minimal medium using  $^2\text{H}_2\text{O}$  as solvent and [ $^2\text{H}_7$ ,  $^{13}\text{C}_6$ ]-glucose and  $^{15}\text{N}^2\text{H}_4^+$ -salt as carbon and nitrogen sources, respectively.<sup>13</sup>  $^2\text{H}/^1\text{H}$  back-exchange of the backbone  $\text{H}_\text{N}$  protons is then carried out by purifying the protein in a  $^1\text{H}_2\text{O}$ -based buffer. But in large globular proteins, the  $\text{H}_\text{N}$  located in the hydrophobic core of the proteins are protected from solvent exchange, preventing the insertion of NMR-active  $^1\text{H}$  probes on the backbone of a large number of residues located in the folded part of the Fab. Efficient  $^2\text{H}/^1\text{H}$  exchange generally requires unfolding/refolding of the target protein in a  $^1\text{H}_2\text{O}$ -based buffer,<sup>14</sup> which is a challenge for large proteins containing disulfide-bonds and, if successful, such protocol is usually characterized by a low protein yield. To solve this problem, we decided to take advantage of cell-free expression system to set up an efficient protocol enabling the production of optimally-labeled anti-LAMP1 Fab without requiring time-consuming and inefficient unfolding/refolding steps.

***Optimization of Cell-Free production of a correctly folded disulfide-rich Fab fragment.***

*E. coli*-based cell-free (CF) expression of proteins exploits transcription and translation machineries extracted from these prokaryotic cells.<sup>15,16</sup> The cell extract, containing the required protein synthesis machineries, is recovered after removal of the DNA/RNA of the original organism. These machineries are supplemented with T7 RNA polymerase, amino acids and energy sources, in order to express the target protein from provided DNA template. Cell-free protein synthesis has been shown to be a promising technology for the production of proteins whose synthesis is challenging in conventional *in vivo* protein production approach.<sup>17-21</sup> We therefore decided to co-express the Fab heavy (HC) and light (LC) chains *in vitro* by inserting the corresponding plasmids into a CF reaction mixture<sup>22</sup> and to assess the presence of folded Fab in the soluble fraction by performing a western blot, in non-reducing conditions, using an anti-His tag antibody. Analysis of the first small-scale expression tests performed using a standard S30 extract from *E. coli* BL21(DE3) strain did not reveal the presence of soluble folded Fab (Fig. 2a lane A). In the standard *E. coli* cell extract, reducing agents such as DTT (dithiothreitol) are added to preserve the activity of the transcription/translation machineries, but they are not suitable for the folding of disulfide-rich proteins. In addition, standard *E. coli* extracts contain endogenous reductases such as thioredoxin reductase (trxB) and glutaredoxin reductase (gor), contributing to maintain cysteine in a reduced state. Although the reducing agents can be excluded from S30 extract, the presence of reductases in S30 extract still hinders the formation of disulfide bonds. We therefore decided to prepare S30 extracts from *E. coli* SHuffle, a strain modified to suppress the gor and thrx reductases in the cytoplasm and to overexpress the disulfide isomerase DsbC.<sup>23,24</sup> Co-expression of Fab HC and LC using this optimized S30 extract enabled detection of folded, soluble Fabs in the cell-free reaction mixture (Fig. 2a lane B). To further promote the formation of these disulfide bonds and rearrangement by DsbC, we decided to add an excess of reduced and oxidized glutathione in a 1:4 ratio.<sup>25</sup> Maintaining the redox potential of the reaction mixture under optimal conditions using GSSG/GSH increased the concentration of soluble Fab in the cell-free reaction by a factor of *ca.* 3 (Fig. 2a lane C).

Preliminary cell-free expression tests for light or heavy chains separately indicated that the light chain (LC) is well expressed in the soluble fraction, while the heavy chain (HC) alone is expressed at a lower level and mainly in the insoluble fraction (data not shown). We therefore decided to optimize the plasmid ratio, in order to maximize the level of expression of folded anti-LAMP1 Fab in the soluble fraction. The total plasmid concentration was maintained constant in the cell-free reaction mixture, whilst the LC/HC plasmid ratio was modified from 0.45 to 0.05. The soluble fraction of the reaction mixture was analyzed by western blot using an anti-His-tag antibody to detect the His-tagged light chain alone or in the Fab complex, whereas the unpaired heavy chain was not detected (insoluble strep-tag construct). Western blot detecting strep-tag confirmed that the band at 50 kDa corresponded to the heterodimeric Fab (supplementary figure S1). The intensity of western-blot lanes was plotted on figure

Résultats Article : Enabling site-specific NMR investigation of therapeutic Fab using a cell-free based isotopic labeling approach : application to anti-LAMP1 Fab.

2b. At high LC/HC plasmid ratios, we observed both Fab and monomeric LCs, indicating that LCs (in excess) are paired with HCs to form a soluble Fab complex. By decreasing the LC/HC plasmid *ratio*, we observed a concomitant increase in the amount of Fab and a decrease in the unpaired LC chain up to a ratio of 0.22, where the unpaired LC chain disappears almost completely and the highest amount of soluble Fab is detected. This result indicates that the LC expression level is about 3.5 to 4 times higher than that of HC. As expected, for a lower value of the LC/HC plasmid ratio, the amount of Fab decreased significantly due to the defect in light chain production.

We then tested the influence of polyions on the yield of Fab production. Betaine, ectoine and choline can be used *in vivo* and *in vitro* to stabilize monomeric proteins in solution.<sup>26-28</sup> We therefore compared the total amounts of Fab recovered in the soluble fraction for four cell-free syntheses performed in the absence of polyion, or with the addition of choline, ectoin or betaine (figure 2c). To enable reliable quantification of Fab produced in the soluble fraction, cell-free reactions were carried out in continuous exchange mode (CECF<sup>21</sup>) with a reaction mixture volume of 1 mL and a dialysis/reaction volume ratio of 10. While the addition of ectoin or choline (10 mM each) did not improve yield, a 2.7-fold increase in soluble Fab production was observed in the presence of betaine. After optimization, the total yield of anti-LAMP1 Fab recovered from the CECF synthesis using SHuffle S30 extract in the presence of GSH/GSSG and betaine was estimated to 125 µg per mL of cell-free reaction mix.

### ***In vitro production of [<sup>15</sup>N, <sup>13</sup>C, <sup>2</sup>H]-Fab in <sup>1</sup>H<sub>2</sub>O-based buffer***

This optimized protocol was applied for the production of perdeuterated proteins by replacing unlabeled amino acids with a 20- amino acid mixture uniformly enriched in [<sup>2</sup>H,<sup>13</sup>C,<sup>15</sup>N]. By carrying out protein synthesis in a <sup>1</sup>H<sub>2</sub>O-based buffer, all exchangeable protons such as backbone H<sub>N</sub> sites are fully occupied by NMR active <sup>1</sup>H isotope, without requiring an unfolding/refolding step in <sup>1</sup>H<sub>2</sub>O as is the case for large perdeuterated proteins expressed *in vivo* in M9-<sup>2</sup>H<sub>2</sub>O medium.<sup>14</sup> Nevertheless, *E. coli* transaminases are still present in the S30 extract and can catalyze the exchange of the <sup>2</sup>H located in the α- and β-positions of amino acids by the <sup>1</sup>H isotope of the solvent.<sup>29-31</sup> To inactivate the main transaminases, the S30 extract can be treated with NaBH<sub>4</sub>, D-malate and aminooxyacetate, reducing the <sup>2</sup>H/<sup>1</sup>H exchange by a factor of 5.31 As a side-effect, this transaminase inactivation also reduces the efficiency of the S30 extract and an average decrease in protein production yield of a factor of 2 is usually observed, but such decrease remains acceptable compared to the low recovery yield of typical urea-unfolding/refolding protocol required to reprotonate all H<sub>N</sub> sites in a perdeuterated protein produced *in vivo* using <sup>2</sup>H<sub>2</sub>O buffer.

After purification by affinity chromatography using successively the HC-strep tag and the LC-His tag (Fig 3A), this optimized in vitro production protocol enabled to recover 125 µg of pure [<sup>15</sup>N, <sup>13</sup>C, <sup>2</sup>H]-Fab per mL of cell-free reaction mixture (Fig. 3a). As the two tag sequences are subject to proteolysis and may give rise to intense NMR signals compared with the core Fab signals, they were cleaved using TEV cleavage sites inserted between each tag and target protein sequence (supplementary figure S2). Note that the amount of TEV-protease was optimized to compensate the loss of activity in absence of DTT. After reverse Ni-NTA purification to remove the His-tagged TEV and to recover the cleaved Fab (Fig 3b), the sample was concentrated to 0.3 mM and loaded into a 4 mm shigemi tube. The 2D <sup>15</sup>N-TROSY spectrum of the labeled Fab (Fig. 3c) shows well-dispersed NMR signals, characteristic of a folded 50 kDa protein rich in β-sheets. The interaction assay carried out in the presence of LAMP1 antigen confirmed that the U-[<sup>15</sup>N, <sup>13</sup>C, <sup>2</sup>H]- Fab fragment of the anti-LAMP1 mAb was active and conserved its ability to form a complex with its antigen (supplementary figure S3).

### ***Backbone assignment of anti-LAMP1 Fab NMR signals***

Sequential assignment of anti-LAMP1 Fab backbone was performed using 3D-BEST-TROSY HNCA, HN(CA)CB, HNC(O), HN(CA)CO, HN(CO)CA and HN(COCA)CB triple resonance experiments.<sup>32,33</sup> Figure 4 presents examples of 2D extracts for four consecutive amino acids, which were the most difficult ones to assign in the protein sequence as no connection to other amino acids was observed. Once all the remaining identified spin-systems were linked together and assigned to their corresponding amino acids in the Fab sequence, this four amino acid long segment could be potentially placed only in two remaining positions of the light chain: Y<sub>37</sub>Q<sub>38</sub>D<sub>39</sub>K<sub>40</sub> or D<sub>93</sub>N<sub>94</sub>L<sub>95</sub>W<sub>96</sub>. Considering on one hand that K<sub>40</sub> is followed by a proline and therefore cannot be connected, and on the other hand that sequential signals observed for assigned T<sub>97</sub> do no match with signals observed for the fourth segment residue, the segment could be assigned to position <sup>37-40</sup>. Both W<sub>36</sub> and Y<sub>92</sub> were not assigned and cannot be used to confirm connection with weak sequential signals observed for first residues of the segment. We therefore used structure-based prediction of chemical shift using ShiftX<sup>2</sup><sup>34</sup> to confirm that only Y<sub>37</sub>Q<sub>38</sub>D<sub>39</sub>K<sub>40</sub> is compatible with observed resonances for the four-residue long segment.

Extensive analysis of triple resonance experiments enabled the detection and assignment of 89 % of <sup>1</sup>H<sub>N</sub>, 89% of <sup>15</sup>N<sub>H</sub>, 92% of <sup>13</sup>CO, 93% of <sup>13</sup>C<sub>α</sub> and 89% of <sup>13</sup>C<sub>β</sub> (Fig. 5). The completeness of assignment reaches 97% for constant domains of light and heavy chains (C<sub>L</sub> and C<sub>H1</sub>) and drops to 53% for the 6 loops forming the antigen Complementarity Determining Region (CDR) of Fab variable domain. The level of assignment for the 6 CDR loops correlates with an average 3-fold lower intensities for the CDR signals in the 3D experiments compared with the remaining Fab signals. The weaker intensities observed for the CDR signals are presumably due to µs-ms conformational exchange occurring in these six loops in the absence of antigenic target. Analysis of all the signals observable in

Résultats Article : Enabling site-specific NMR investigation of therapeutic Fab using a cell-free based isotopic labeling approach : application to anti-LAMP1 Fab.

the high resolution 2D  $^{15}\text{N}$ -TROSY (Figure 5), indicates that only 13 peaks with a signal/noise *ratio* higher than 10 were not sequentially assigned. 5 of them correspond to tryptophan side chain iminos, 2 can be ambiguously assigned to 3 remaining Serines (S<sub>33</sub>, S<sub>34</sub> or S<sub>78</sub> from the HC), and 6 others cannot be connected in the Fab sequence. Secondary structure predictions of both light and heavy chains (supplementary figure 4), based on backbone chemical assignment using TALOS+ <sup>35</sup>, displayed secondary structure elements in accordance with the crystal structure (PDB code: 8ATH <sup>5</sup>).

## Discussion

Therapeutic mAbs or Fab fragments are efficiently produced in optimized media using eukaryotic cells such as Chinese hamster ovary cells (CHO). Although production yields can reach 1 g/L in batch and 1-10 g/L in fed-batch processes,<sup>36</sup> this expression system is not compatible with the production of perdeuterated proteins due to the toxicity of <sup>2</sup>H<sub>2</sub>O (D<sub>2</sub>O) for mammalian cells.<sup>37</sup> A bacterial expression system can also be used for Fab preparation, as this mAb fragment generally lacks glycosylation or other specific post-transcriptional modifications. Even if this system usually remains advantageous because it is affordable and allows easy handling, the production of correctly folded eukaryotic proteins is often difficult, particularly when the protein contains multiple disulfide bonds, as Fabs. Indeed, even if some Fabs have been produced in *E. coli*, bacterial expression is not considered optimal as most Fabs are prone to accumulation as insoluble.<sup>38</sup> Complex unfolding/refolding steps are then required to recover active folded Fabs, but the association of light and heavy chains via disulfide bridges is poorly efficient in *E. coli* due to the large number of combinations that can occur in making and breaking the disulfide bonds. As an example, a recent patent covers the production process of the Fab ranibizumab, in which protein refolding was performed by diluting solubilised inclusion bodies into a refolding buffer, and both pH and temperature shifts were performed at suitable time intervals to obtain the refolded protein, with a refolding yield of 13.6% (PCT No.: PCT / IB2016 / 054926).

Alternatively, as the cytoplasm of *E. coli* is a reducing environment preventing the formation of disulfide bonds, disulfide-rich proteins can be addressed to the more oxidative periplasmic compartment, where the proteins should adopt their native topologies with the help of isomerases, including the disulfide isomerase DsbC.<sup>39-41</sup> Inefficient translocation is one of the main obstacles to periplasmic expression, and different signal sequences must be tried to effectively increase expression of correctly folded soluble Fabs. Secreted expression to the periplasm, then, usually results in low yields of soluble Fabs. For example, previous attempts to express Fabs in the periplasm were characterized by moderate yields (from 1 to 20 mg/L), although expression was carried out in high cell density cultures.<sup>42</sup> Fabs can also be expressed as soluble proteins in the cytoplasm, using modified *E. coli* strains lacking thioredoxin and glutaredoxin reductases and expressing the disulfide isomerase DsbC to allow formation and native pairing of cysteines.<sup>24,43</sup> This allowed to obtain yields varying between 8 and 20 mg/L of correctly folded Fab.<sup>44</sup> To achieve higher yields, Gaciarz et al. developed the CyDisCo system, based on coexpression of a catalyst of disulfide bond formation and a catalyst of disulfide bond isomerization (usually DsbC or PDI). Yields of purified Fab fragments up to 40 mg/L were obtained using rich media.<sup>43</sup>

Even if recent advances in cloning technologies and host cell engineering have resulted in a significant improvement in the expression of Fabs in *E. coli*, it should be kept in mind that the production of isotopically labeled proteins requires the use of M9 minimal medium, which reduces the average

Résultats Article : Enabling site-specific NMR investigation of therapeutic Fab using a cell-free based isotopic labeling approach : application to anti-LAMP1 Fab.

protein yield by an order of magnitude compared with ultra-rich *E. coli* culture media used in previously published Fab synthesis protocols. In addition, the production of perdeuterated proteins requires the adaptation of *E. coli* in  $^2\text{H}_2\text{O}$  solvent,<sup>13</sup> which further significantly reduces the total soluble proteins recovered. Finally, observation of  $\text{H}_\text{N}$  probes from the backbone at the core of the protein requires unfolding and urea refolding in  $^1\text{H}_2\text{O}$  buffer to allow reprotonation of all  $\text{H}_\text{N}$  sites. Optimization of these folding protocols is complex and time-consuming and results in significant losses of valuable labeled proteins.

The high flexibility of the cell-free system, which allows direct handling of the reaction medium, makes it suitable for the preparation of disulfide-containing proteins.<sup>45-47</sup>

In this study, by carefully controlling the oxidative environment to enhance the formation of suitable disulfide bonds, we used the acellular system to produce optimally labeled Fab directly in  $^1\text{H}_2\text{O}$  medium, without the need for a time-consuming refolding process. *In vivo*, expression of a protein in  $^1\text{H}_2\text{O}$  using deuterated glucose and a mixture of deuterated amino acids results in significant incorporation of the  $^1\text{H}$  isotope in the  $\alpha$  and  $\beta$  positions due to endogenous enzymatic activities. This incorporation of protons into the aliphatic position of the backbone results in signal loss due to relaxation, NMR signal broadening and duplication of  $^{13}\text{C}$  NMR signals due to  $^1\text{H}/^2\text{H}$  isotopic shifts,<sup>11,13</sup> preventing the use of such sample for sequential NMR signal assignment in large proteins. On the contrary, in the open cell-free system, these transaminases can be inactivated by reducing the pyridoxal-5-phosphate (PLP) cofactors or by adding various inhibitors.<sup>29-31</sup> Although it reduces the yield of protein synthesis by around 50% on average, the loss is comparable to the drop in protein yield observed during cell-free protein synthesis in  $^2\text{H}_2\text{O}$  buffer. In addition, transaminase inactivation enables *in vitro* production of perdeuterated proteins in  $^1\text{H}_2\text{O}$ , with all  $\text{H}_\text{N}$  sites fully protonated, a prerequisite for avoiding inefficient unfolding/refolding protocols.

The production of an optimally isotopically labeled Fab fragment has enabled the sequential assignment of the NMR signal from the backbone of a therapeutic Fab. Previous attempts using  $^{15}\text{N}$ ,  $^{13}\text{C}$ -labeled fragments (without perdeuteration) only allowed the assignment of the homodimeric Fc fragment.<sup>48</sup> As Fabs are hetero-oligomers, their NMR assignment is more challenging and has been, until now, and was limited to 25 kDa chimeric ScFv constructs with an artificial linker connecting the two variable VH and VL domains.<sup>49</sup> This first sequential assignment of therapeutic Fab NMR signals offers the possibility of studying, at atomic resolution, the interactions, dynamics or localization of changes in the Fab high-order structure (HOS) of a Fab during the production and storage processes.<sup>2</sup> Although the aim of this study was not to carry out an in-depth NMR study of the interaction between the anti-LAMP1 Fab (47 kDa) and its antigen, figure 6 shows an example of observed perturbations of NMR signals in the presence of the glycosylated soluble LAMP1 construct (80 kDa<sup>5</sup>) in a well-resolved area of 2D  $^{15}\text{N}$ -TROSY spectrum.

Résultats Article : Enabling site-specific NMR investigation of therapeutic Fab using a cell-free based isotopic labeling approach : application to anti-LAMP1 Fab.

The developed method presented in this article is general and allows the *in vitro* production of isotopically labeled Fabs suitable for site-specific NMR studies of these important therapeutic mAb fragments. Here, we showed that this cell-free system could be a useful Fab production platform for structural analysis of potential therapeutic antibodies by NMR as production in this system is rapid and scalable. These results represent an important step towards the general use of NMR to enable atomic resolution investigation of therapeutic mAb fragments in solution.

## **Materials and Methods**

### ***Cell-Free S30 extracts preparation***

The S30 lysates were prepared based on the protocol previously published<sup>30</sup> using *E. coli* BL21 (DE3) strain, commonly used for cell-free protein synthesis,<sup>50</sup> or SHuffle strain to promote formation of native disulfide bonds topology.<sup>51</sup> To limit the back protonation of the H<sub>α</sub> and H<sub>β</sub> positions of the supplied perdeuterated amino acids by transaminases for the production of U-[<sup>15</sup>N, <sup>13</sup>C, <sup>2</sup>H]-Fab sample, the S30 extract was previously treated with NaBH<sub>4</sub>, AOA, and D-malate.<sup>30</sup>

### ***Protein plasmids preparation***

The light (LC - residues 1 to 214) and heavy (HC - residues 1 to 227) chain sequences of the anti-LAMP1 mAb (Fab B sequence)<sup>5</sup> were cloned separately in pIVEX 2.4d and pIVEX 2.3d plasmids (respectively) containing either an N-terminal 6-His tag (LC) or a Twin-Strep-tag (HC) followed by a TEV protease cleavage site. Each plasmid was transformed into TOP10 *E. coli* strain grown in LB medium (500 mL) with 100 µg/mL of ampicillin. Plasmids were then purified using NucleoBond Xtra Maxi Plus kit (Macherey-Nagel) and final elution concentration was defined to 1 mg/mL in RNase-free water.

### ***Optimisation of Anti-LAMP1 Fab in vitro production***

For optimization purpose, small scale cell-free Fab syntheses were performed in 50 µL batch mode or in 1 mL continuous exchange mode (CECF) with dialysis/reaction volume ratio of 1:10. All reactions were performed at 30°C with a gentle shaking (20 rpm) using an hybridization oven (Techne), in 1.5 mL Eppendorf tubes for batches and in 3 mL GeBAflex dialysis devices (Euromedex) for CECF. Optimisations were performed using BL21(DE3) or SHuffle S30 extracts added at 40 % of the final reaction volume. In batch mode, LC and HC plasmids were added at a total final concentration of 10 µg/mL (with different ratios), and 16 µg/mL for CECF reactions. To form disulfide bridges, reduced and oxidized glutathiones were used at 1 and 4 mM respectively, according to the previously published protocol.<sup>25</sup> The benefits of polyions was tested using a 10 mM final concentration of betaine, ectoine or choline in the reaction and the cell-free dialysis buffer.

Other standard cell-free components (NTPs, HEPES, MgOAc<sub>2</sub>, creatine phosphate, NH<sub>4</sub>OAc, spermidine, AMPc, folic acid) were added as described in the previously published protocol.<sup>30</sup> The cell-free reactions were stopped after 3 h for batch and 16 h for CECF and reaction mixtures were centrifuged for 20 min at 14 000 g before next steps. Batch reactions were analysed with Western Blot using an anti-His-tag or anti-strep-tag antibodies. CECF reactions were purified using the purification

Résultats Article : Enabling site-specific NMR investigation of therapeutic Fab using a cell-free based isotopic labeling approach : application to anti-LAMP1 Fab.

protocol described below. Yields of purified protein were determined by absorbance measurement at 280 nm on a Nanodrop 2000 (Thermo Fisher) ( $\epsilon = 75\,915\text{ M}^{-1}\cdot\text{cm}^{-1}$  ;  $A = 1.596$  ;  $MW = 47\,552\text{ Da}$ ).

### ***Production of U-[<sup>15</sup>N, <sup>13</sup>C, <sup>2</sup>H]-Fab anti-LAMP1***

Both chains were co-expressed *in vitro* using cell-free protein synthesis performed CECF.<sup>52</sup> LC and HC plasmids were added at a total concentration of 16  $\mu\text{g/mL}$  (LC/HC plasmids ratio of 0,22) to 3 mL GeBAflex dialysis devices (Euromedex) containing creatine kinase (Merck - 250  $\mu\text{g/mL}$ ), T7 RNA polymerase (produced in house - 100  $\mu\text{g/mL}$ ), tRNA (Merck – 175  $\mu\text{g/mL}$ ) and S30 *E. coli* SHuffle bacterial extract. Cell free synthesis was performed with oxidized (GSSG) and reduced (GSH) glutathione at a concentration of 3.6 and 0.9 mM, respectively. The 20 U-[<sup>2</sup>H, <sup>13</sup>C, <sup>15</sup>N]-labeled amino acids (CIL) were added as a mix at a final concentration of 3 mg/mL and betaine at 15 mM. Other standard cell-free components (NTPs, HEPES, MgOAc<sub>2</sub>, creatine phosphate, NH<sub>4</sub>OAc, spermidine, AMPc, Folinic acid) were added as described previously.<sup>30</sup> Each GeBAflex was placed in a 50 mL centrifugation tube containing 30 mL of dialysis mixture (same as described above, without T7, tRNA, S30-extract, creatine kinase and DNA-plamids - see Imbert et al. 2021 or detailed protocol<sup>30</sup>). Protein synthesis was performed during 16 h in 16 GeBAflex units in parallel at 30°C with gentle agitation (20 rpm in an hybridization oven, (Techne)). To ensure full protonation of all HN positions of the perdeuterated Fab, *in vitro* synthesis was performed with H<sub>2</sub>O-based HEPES buffer at pH 7.4 using S30 extract treated to limit activities of transaminases.<sup>30</sup>

### ***Fab purification***

Cell-free reaction mixtures were centrifuged for 20 min at 14 000 g and supernatants were pooled and diluted by a factor 4 in buffer A, before two successive affinity chromatography steps using streptactin (StrepTrap HP 5 mL column – Cytiva) and Ni-NTA (HisTrap FF 1 mL column – Cytiva) resins (See supplementary table S1 for detailed purification protocols and buffers composition). After removal of imidazole using a centrifugal filtration unit (Amicon® 10 kDa MWCO), the two purification tags were cleaved using His-tagged TEV protease. Cleavages were performed overnight at 4°C, with the Fab fragment concentrated at 1mg/mL and the TEV at 0.6 mg/mL in a buffer containing 50 mM Tris-HCl pH 8, 250 mM NaCl, 0.5 mM EDTA, 3 mM GSH, 0.3 mM GSSG and 10% of glycerol (DTT usually used for TEV cleavage was avoided to prevent disulfide bonds reduction). His-tagged TEV was removed by reverse Ni-NTA chromatography (HisTrap FF 1 mL column – Cytiva). The flow through containing the cleaved Fab fragment was buffer exchanged and concentrated in a NMR buffer containing 10% D<sub>2</sub>O, 50 mM MES at pH 6.5, 100 mM NaCl and cOmplete protease inhibitors (1 tablet for 50 mL buffer solution) using a centrifugal filtration unit (Amicon® 10 kDa MWCO). The final U-[<sup>15</sup>N,<sup>13</sup>C,<sup>2</sup>H] Fab sample (200  $\mu\text{L}$ ) was placed in a 4 mm Shigemi NMR tube at a final concentration of 0.3 mM.

### ***Liquid-state NMR spectroscopy***

2D-<sup>15</sup>N-TROSY experiments and 6 triple resonance experiments (HNCA, HN(CO)CA, HN(CA)CB, HN(COCA)CB, HNCO, HN(CA)CO)<sup>32,33</sup> were acquired at 35°C on Bruker Avance III HD spectrometers equipped with cryogenic probes and operating at <sup>1</sup>H frequency of 600 to 950 MHz for a total duration of 18 days using a 0.3 mM sample of U-[<sup>2</sup>H, <sup>15</sup>N, <sup>13</sup>C] anti-LAMP1-Fab (See details on supplementary table S2). Data were processed using nmrPipe/nmrDraw package,<sup>53</sup> and analysed using FLYA,<sup>54</sup> TALOS+,<sup>35</sup> ShiftX2,<sup>34</sup> CcpNMR 3<sup>55</sup> with the help of an in-house developed software.

The assignment of chemical shifts of the anti-LAMP1-Fab, together with all the NMR data sets used, will be available upon publication into the BMRB (<https://www.bmrwisc.edu/>).

### ***Interaction of anti-LAMP1 Fab with LAMP1***

LAMP1 was produced using HEK293 cells as described previously.<sup>5</sup> The interaction of the Fab fragment with its antigen LAMP1 was assessed by size exclusion chromatography using a superose 6 increase 3.2/300 column. For the anti-LAMP1 Fab, LAMP1, or LAMP1/anti-LAMP1 Fab complex, 20 µL sample in PBS 2X pH 7.4 buffer was loaded at a flow rate of 0.15 mL/min and a temperature of 22°C. The anti-LAMP1 Fab fragment was loaded at 1 mg/mL and LAMP1 at 0.4 mg/mL. The complex was formed by mixing LAMP1 and Fab (both prepared in PBS buffer) at a molar *ratio* of 8:10 for 20 min at 37°C. The complex was then loaded on the column at a concentration of 1.17 mg/mL.

For the NMR interaction study, the complex was formed by mixing LAMP1 (280 µM) and Fab (140 µM), both prepared in PBS buffer, and incubated at 37°C during 20 min. The complex was then concentrated up to 200 µL on an Amicon with a cut-off of 50 kDa before acquisition of 2D <sup>15</sup>N-TROSY experiment.

Résultats Article : Enabling site-specific NMR investigation of therapeutic Fab using a cell-free based isotopic labeling approach : application to anti-LAMP1 Fab.

### **Acknowledgments.**

The authors thank Isabel Ayala and Rida Awad for advice and stimulating discussions. This work is supported by the French National Research Agency in the framework of the "Investissements d'avenir" program (ANR-15-IDEX-02), NMR4mAbs project n° ANR-22-CE29-0024-01 and by the CEA/SANOFI collaborative research program C36478/191073. This work used the high field NMR and Cell-Free facilities at the Grenoble Instruct-ERIC Center (ISBG; UAR 3518 CNRS-CEA-UGA-EMBL) within the Grenoble Partnership for Structural Biology (PSB). Platform access was supported by FRISBI (ANR-10-INBS-05-02) and GRAL, a project of the University Grenoble Alpes graduate school (Ecoles Universitaires de Recherche) CBH-EUR-GS (ANR-17-EURE-0003) and IR INFRANALYTICS FR2054. IBS acknowledges integration into the Interdisciplinary Research Institute of Grenoble (IRIG, CEA).

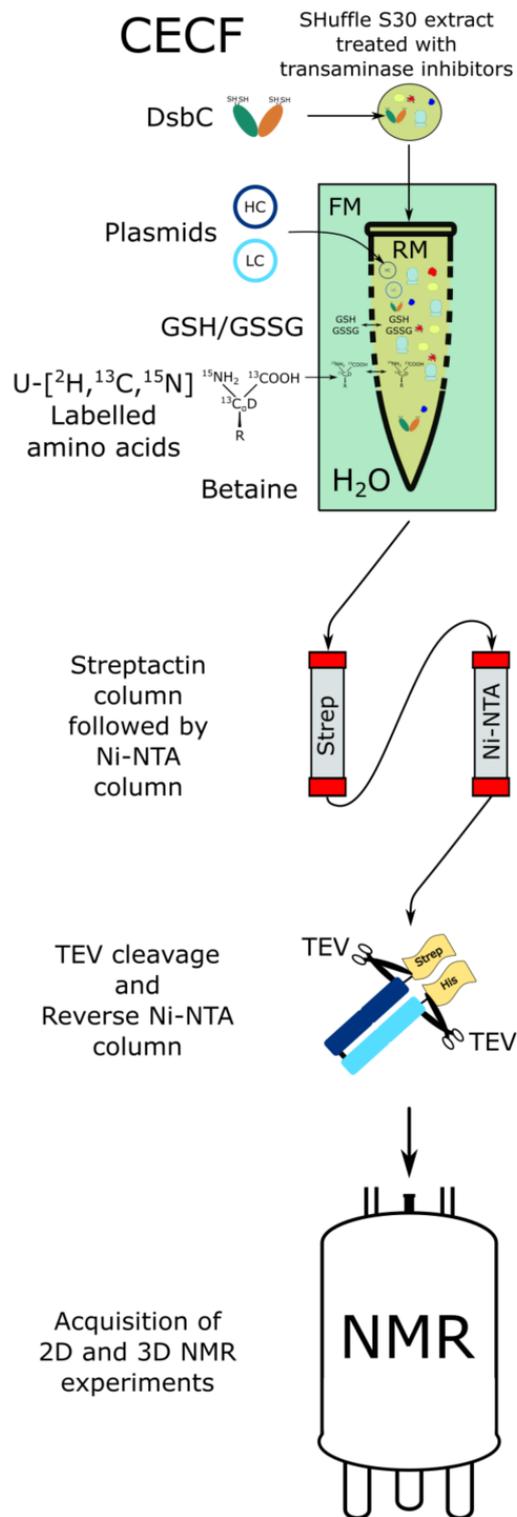
## References :

1. Kaplon H, Crescioli S, Chenoweth A, Visweswaraiah J, Reichert JM. Antibodies to watch in 2023. *mAbs*. 2023;15(1):2153410. doi:10.1080/19420862.2022.2153410
2. Brinson RG, Marino JP, Delaglio F, et al. Enabling adoption of 2D-NMR for the higher order structure assessment of monoclonal antibody therapeutics. *mAbs*. 2018;11(1):94-105. doi:10.1080/19420862.2018.1544454
3. Marino JP, Brinson RG, Hudgens JW, et al. Emerging Technologies To Assess the Higher Order Structure of Monoclonal Antibodies. In: *State-of-the-Art and Emerging Technologies for Therapeutic Monoclonal Antibody Characterization Volume 3. Defining the Next Generation of Analytical and Biophysical Techniques*. Vol 1202. ACS Symposium Series. American Chemical Society; 2015:17-43. doi:10.1021/bk-2015-1202.ch002
4. Cerofolini L, Ravera E, Fischer C, et al. Integration of NMR Spectroscopy in an Analytical Workflow to Evaluate the Effects of Oxidative Stress on Abituzumab: Beyond the Fingerprint of mAbs. *Anal Chem*. 2023;95(24):9199-9206. doi:10.1021/acs.analchem.3c00317
5. Pruvost T, Mathieu M, Dubois S, Maillère B, Vigne E, Nozach H. Deciphering cross-species reactivity of LAMP-1 antibodies using deep mutational epitope mapping and AlphaFold. *mAbs*. 2023;15(1):2175311. doi:10.1080/19420862.2023.2175311
6. Terasawa K, Tomabechi Y, Ikeda M, et al. Lysosome-associated membrane proteins-1 and -2 (LAMP-1 and LAMP-2) assemble via distinct modes. *Biochem Biophys Res Commun*. 2016;479(3):489-495. doi:10.1016/j.bbrc.2016.09.093
7. Cameron B, Dabdoubi T, Berthou-Soulié L, et al. Complementary epitopes and favorable developability of monoclonal anti-LAMP1 antibodies generated using two transgenic animal platforms. *PLOS ONE*. 2020;15(7):e0235815. doi:10.1371/journal.pone.0235815
8. Alessandrini F, Pezzè L, Ciribilli Y. LAMPs: Shedding light on cancer biology. *Semin Oncol*. 2017;44(4):239-253. doi:10.1053/j.seminoncol.2017.10.013
9. Agarwal AK, Gude RP, Kalraiya RD. Regulation of melanoma metastasis to lungs by cell surface Lysosome Associated Membrane Protein-1 (LAMP1) via galectin-3. *Biochem Biophys Res Commun*. 2014;449(3):332-337. doi:10.1016/j.bbrc.2014.05.028
10. Ikura M, Kay LE, Bax A. A novel approach for sequential assignment of proton, carbon-13, and nitrogen-15 spectra of larger proteins: heteronuclear triple-resonance three-dimensional NMR spectroscopy. Application to calmodulin. *Biochemistry*. 1990;29(19):4659-4667. doi:10.1021/bi00471a022
11. Venters RA, Farmer II BT, Fierke CA, Spicer LD. Characterizing the Use of Perdeuteration in NMR Studies of Large Proteins:  $^{13}\text{C}$ ,  $^{15}\text{N}$  and  $^1\text{H}$  Assignments of Human Carbonic Anhydrase II. *J Mol Biol*. 1996;264(5):1101-1116. doi:10.1006/jmbi.1996.0699
12. Garrett DS, Seok YJ, Liao DI, Peterkofsky A, Gronenborn AM, Clore GM. Solution Structure of the 30 kDa N-Terminal Domain of Enzyme I of the Escherichia coli Phosphoenolpyruvate: Sugar Phosphotransferase System by Multidimensional NMR. *Biochemistry*. 1997;36(9):2517-2530. doi:10.1021/bi962924y
13. Gardner KH, Kay LE. The Use of  $^2\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$  Multidimensional Nmr Gto Study the Structure and Dynamics of Proteins. *Annu Rev Biophys Biomol Struct*. 1998;27(1):357-406. doi:10.1146/annurev.biophys.27.1.357
14. Tugarinov V, Kanelis V, Kay LE. Isotope labeling strategies for the study of high-molecular-weight proteins by solution NMR spectroscopy. *Nat Protoc*. 2006;1(2):749-754. doi:10.1038/nprot.2006.101
15. Zubay G. In Vitro Synthesis of Protein in Microbial Systems. *Annu Rev Genet*. 1973;7(1):267-287. doi:10.1146/annurev.ge.07.120173.001411
16. Pratt C. Kinetics and regulation of cell-free alkaline phosphatase synthesis. *J Bacteriol*. 1980;143(3):1265-1274. doi:10.1128/jb.143.3.1265-1274.1980
17. Segelke BW, Schafer J, Coleman MA, et al. Laboratory scale structural genomics. *J Struct Funct Genomics*. 2004;5(1):147-157. doi:10.1023/B:JSFG.0000029193.82120.d1
18. Kigawa T. Cell-Free Protein Preparation Through Prokaryotic Transcription-Translation Methods. In: Endo Y, Takai K, Ueda T, eds. *Cell-Free Protein Production: Methods and*

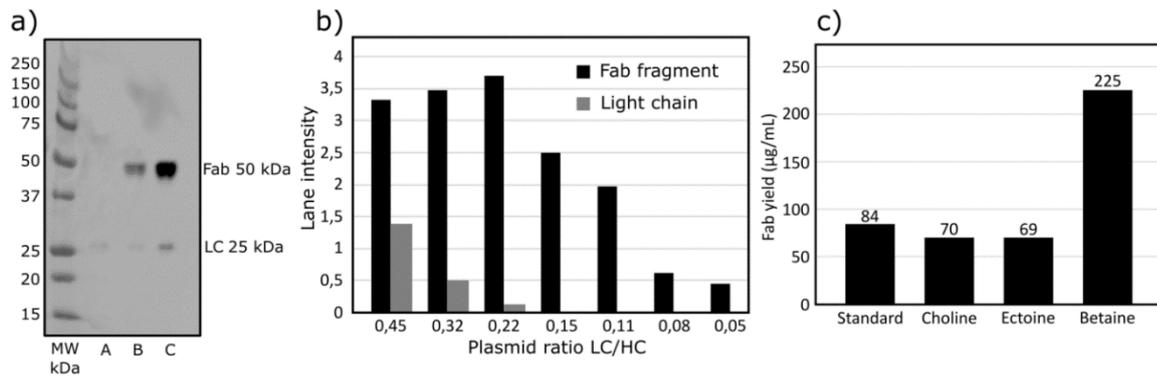
- Protocols*. Methods in Molecular Biology. Humana Press; 2010:1-10. doi:10.1007/978-1-60327-331-2\_1
19. Schwarz D, Klammt C, Koglin A, et al. Preparative scale cell-free expression systems: New tools for the large scale preparation of integral membrane proteins for functional and structural studies. *Methods*. 2007;41(4):355-369. doi:10.1016/j.ymeth.2006.07.001
  20. Sawasaki T, Hasegawa Y, Tsuchimochi M, et al. A bilayer cell-free protein synthesis system for high-throughput screening of gene products. *FEBS Lett*. 2002;514(1):102-105. doi:10.1016/S0014-5793(02)02329-3
  21. Spirin AS, Baranov VI, Ryabova LA, Ovodov S, Alakhov YB. A Continuous Cell-Free Translation System Capable of Producing Polypeptides in High Yield. *Science*. 1988;242(4882):1162-1164. doi:10.1126/science.3055301
  22. Matsuda T, Ito T, Takemoto C, et al. Cell-free synthesis of functional antibody fragments to provide a structural basis for antibody–antigen interaction. Pizzo SV, ed. *PLOS ONE*. 2018;13(2):e0193158. doi:10.1371/journal.pone.0193158
  23. Dopp JL, Reuel NF. Simple, functional, inexpensive cell extract for in vitro prototyping of proteins with disulfide bonds. *Biochem Eng J*. 2020;164:107790. doi:10.1016/j.bej.2020.107790
  24. Lobstein J, Emrich CA, Jeans C, Faulkner M, Riggs P, Berkmen M. SHuffle, a novel Escherichia coli protein expression strain capable of correctly folding disulfide bonded proteins in its cytoplasm. *Microb Cell Factories*. 2012;11(1):753. doi:10.1186/1475-2859-11-56
  25. Yin G, Swartz JR. Enhancing multiple disulfide bonded protein folding in a cell-free system. *Biotechnol Bioeng*. 2004;86(2):188-195. doi:10.1002/bit.10827
  26. Kai L, Dötsch V, Kaldenhoff R, Bernhard F. Artificial Environments for the Co-Translational Stabilization of Cell-Free Expressed Proteins. *PLoS ONE*. 2013;8(2). doi:10.1371/journal.pone.0056637
  27. Welsh DT, Herbert RA. Osmotically induced intracellular trehalose, but not glycine betaine accumulation promotes desiccation tolerance in Escherichia coli. *FEMS Microbiol Lett*. 1999;174(1):57-63. doi:10.1111/j.1574-6968.1999.tb13549.x
  28. Arora A, Ha C, Park CB. Inhibition of insulin amyloid formation by small stress molecules. *FEBS Lett*. 2004;564(1-2):121-125. doi:10.1016/S0014-5793(04)00326-6
  29. Etezady-Esfarjani T, Hiller S, Villalba C, Wüthrich K. Cell-free protein synthesis of perdeuterated proteins for NMR studies. *J Biomol NMR*. 2007;39(3):229-238. doi:10.1007/s10858-007-9188-0
  30. Imbert L, Lenoir-Capello R, Crublet E, et al. In Vitro Production of Perdeuterated Proteins in H<sub>2</sub>O for Biomolecular NMR Studies. In: Chen YW, Yiu CPB, eds. *Structural Genomics: General Applications*. Methods in Molecular Biology. Springer US; 2021:127-149. doi:10.1007/978-1-0716-0892-0\_8
  31. Su XC, Loh CT, Qi R, Otting G. Suppression of isotope scrambling in cell-free protein synthesis by broadband inhibition of PLP enzymes for selective <sup>15</sup>N-labelling and production of perdeuterated proteins in H<sub>2</sub>O. *J Biomol NMR*. 2011;50(1):35-42. doi:10.1007/s10858-011-9477-5
  32. Favier A, Brutscher B. NMRlib: user-friendly pulse sequence tools for Bruker NMR spectrometers. *J Biomol NMR*. 2019;73(5):199-211. doi:10.1007/s10858-019-00249-1
  33. Lescop E, Schanda P, Brutscher B. A set of BEST triple-resonance experiments for time-optimized protein resonance assignment. *J Magn Reson*. 2007;187(1):163-169. doi:10.1016/j.jmr.2007.04.002
  34. Han B, Liu Y, Ginzinger SW, Wishart DS. SHIFTX2: significantly improved protein chemical shift prediction. *J Biomol NMR*. 2011;50(1):43-57. doi:10.1007/s10858-011-9478-4
  35. Shen Y, Delaglio F, Cornilescu G, Bax A. TALOS+: A hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol NMR*. 2009;44(4):213-223. doi:10.1007/s10858-009-9333-z
  36. Kunert R, Reinhart D. Advances in recombinant antibody manufacturing. *Appl Microbiol Biotechnol*. 2016;100(8):3451-3461. doi:10.1007/s00253-016-7388-9
  37. Katz JJ, Crespi HL. Deuterated Organisms: Cultivation and Uses. *Science*. 1966;151(3715):1187-1194. doi:10.1126/science.151.3715.1187

38. Gupta SK, Shukla P. Microbial platform technology for recombinant antibody fragment production: A review. *Crit Rev Microbiol.* 2017;43(1):31-42. doi:10.3109/1040841X.2016.1150959
39. Rietsch A, Belin D, Martin N, Beckwith J. An in vivo pathway for disulfide bond isomerization in Escherichia coli. *Proc Natl Acad Sci U S A.* 1996;93(23):13048-13053. doi:10.1073/pnas.93.23.13048
40. Missiakas D, Georgopoulos C, Raina S. The Escherichia coli dsbC (xprA) gene encodes a periplasmic protein involved in disulfide bond formation. *EMBO J.* 1994;13(8):2013-2020.
41. Nakamoto H, Bardwell JCA. Catalysis of disulfide bond formation and isomerization in the Escherichia coli periplasm. *Biochim Biophys Acta BBA - Mol Cell Res.* 2004;1694(1):111-119. doi:10.1016/j.bbamcr.2004.02.012
42. Hsu CC, Thomas OR, Overton TW. Periplasmic expression in and release of Fab fragments from Escherichia coli using stress minimization. *J Chem Technol Biotechnol.* 2016;91(3):815-822. doi:10.1002/jctb.4672
43. Gaciarz A, Veijola J, Uchida Y, et al. Systematic screening of soluble expression of antibody fragments in the cytoplasm of E. coli. *Microb Cell Factories.* 2016;15(1):22. doi:10.1186/s12934-016-0419-5
44. Venturi M, Seifert C, Hunte C. High level production of functional antibody fab fragments in an oxidizing bacterial cytoplasm. Edited by J. Karn. *J Mol Biol.* 2002;315(1):1-8. doi:10.1006/jmbi.2001.5221
45. Michel E, Wüthrich K. Cell-free expression of disulfide-containing eukaryotic proteins for structural biology: Cell-free expression of disulfide-containing proteins. *FEBS J.* 2012;279(17):3176-3184. doi:10.1111/j.1742-4658.2012.08697.x
46. Matsuda T, Watanabe S, Kigawa T. Cell-free synthesis system suitable for disulfide-containing proteins. *Biochem Biophys Res Commun.* 2013;431(2):296-301. doi:10.1016/j.bbrc.2012.12.107
47. Wael Gad. Cell-free expression systems for disulfide rich proteins. In: *Cell-Free Protein Expression Systems.* KS Omniscriptum Publishing; 2014.
48. Yagi H, Zhang Y, Yagi-Utsumi M, et al. Backbone <sup>1</sup>H, <sup>13</sup>C, and <sup>15</sup>N resonance assignments of the Fc fragment of human immunoglobulin G glycoprotein. *Biomol Nmr Assign.* 2015;9(2):257-260. doi:10.1007/s12104-014-9586-7
49. Ghasriani H, Ahmadi S, Hodgson DJ, Aubin Y. Backbone and side-chain resonance assignments of the NISTmAb-scFv and antigen-binding study. *Biomol NMR Assign.* 2022;16(2):391-398. doi:10.1007/s12104-022-10109-z
50. Kigawa T, Yabuki T, Matsuda N, et al. Preparation of Escherichia coli cell extract for highly productive cell-free protein expression. *J Struct Funct Genomics.* 2004;5(1):63-68. doi:10.1023/B:JSFG.0000029204.57846.7d
51. de Marco A. Strategies for successful recombinant expression of disulfide bond-dependent proteins in Escherichia coli. *Microb Cell Factories.* 2009;8(1):26. doi:10.1186/1475-2859-8-26
52. Kim DM, Choi CY. A semicontinuous prokaryotic coupled transcription/translation system using a dialysis membrane. *Biotechnol Prog.* 1996;12(5):645-649. doi:10.1021/bp960052l
53. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR.* 1995;6(3):277-293. doi:10.1007/BF00197809
54. Schmidt E, Güntert P. A New Algorithm for Reliable and General NMR Resonance Assignment. *J Am Chem Soc.* 2012;134(30):12817-12829. doi:10.1021/ja305091n
55. Skinner SP, Fogh RH, Boucher W, Ragan TJ, Mureddu LG, Vuister GW. CcpNmr AnalysisAssign: a flexible platform for integrated NMR analysis. *J Biomol Nmr.* 2016;66(2):111-124. doi:10.1007/s10858-016-0060-y

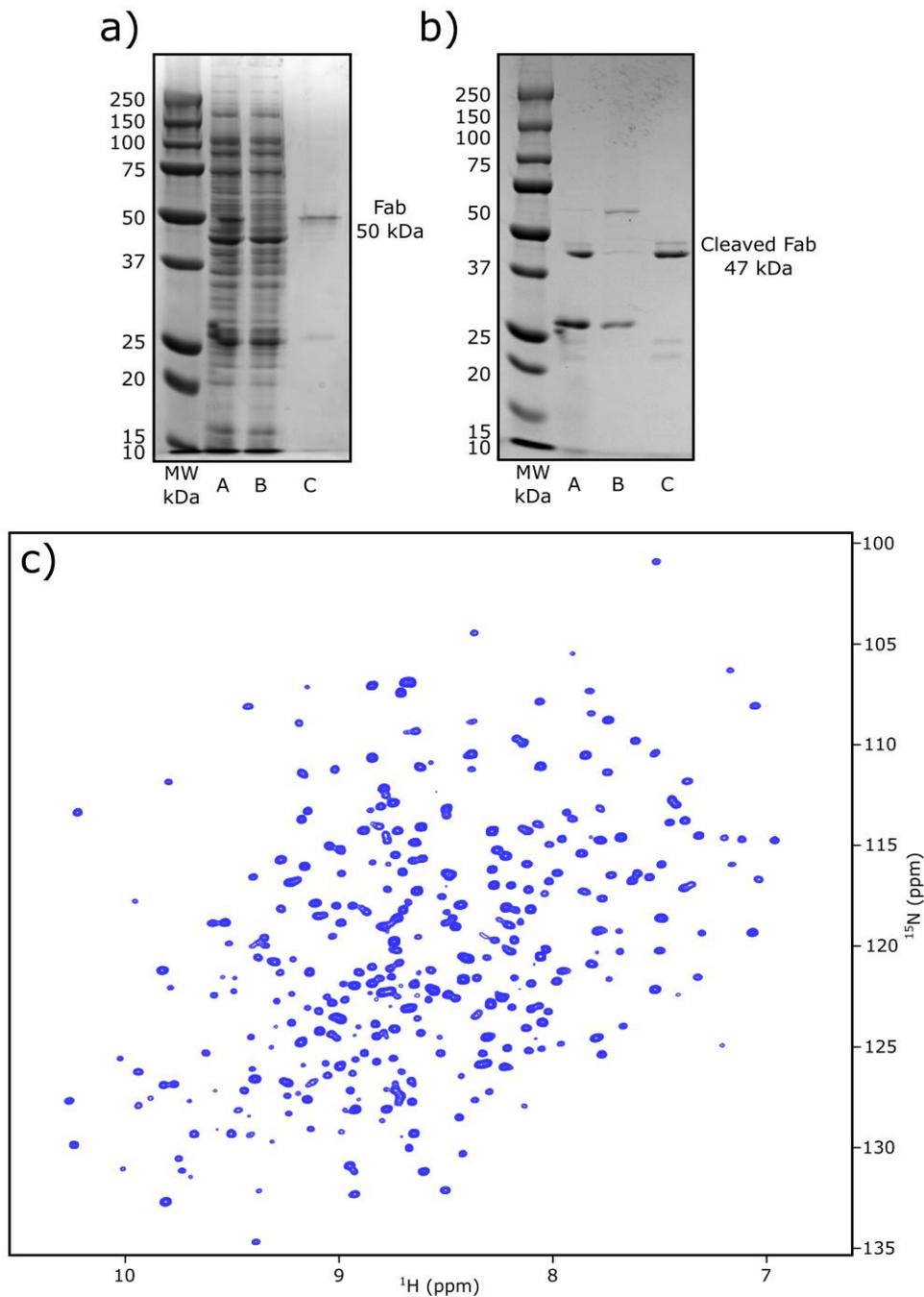
## Figures



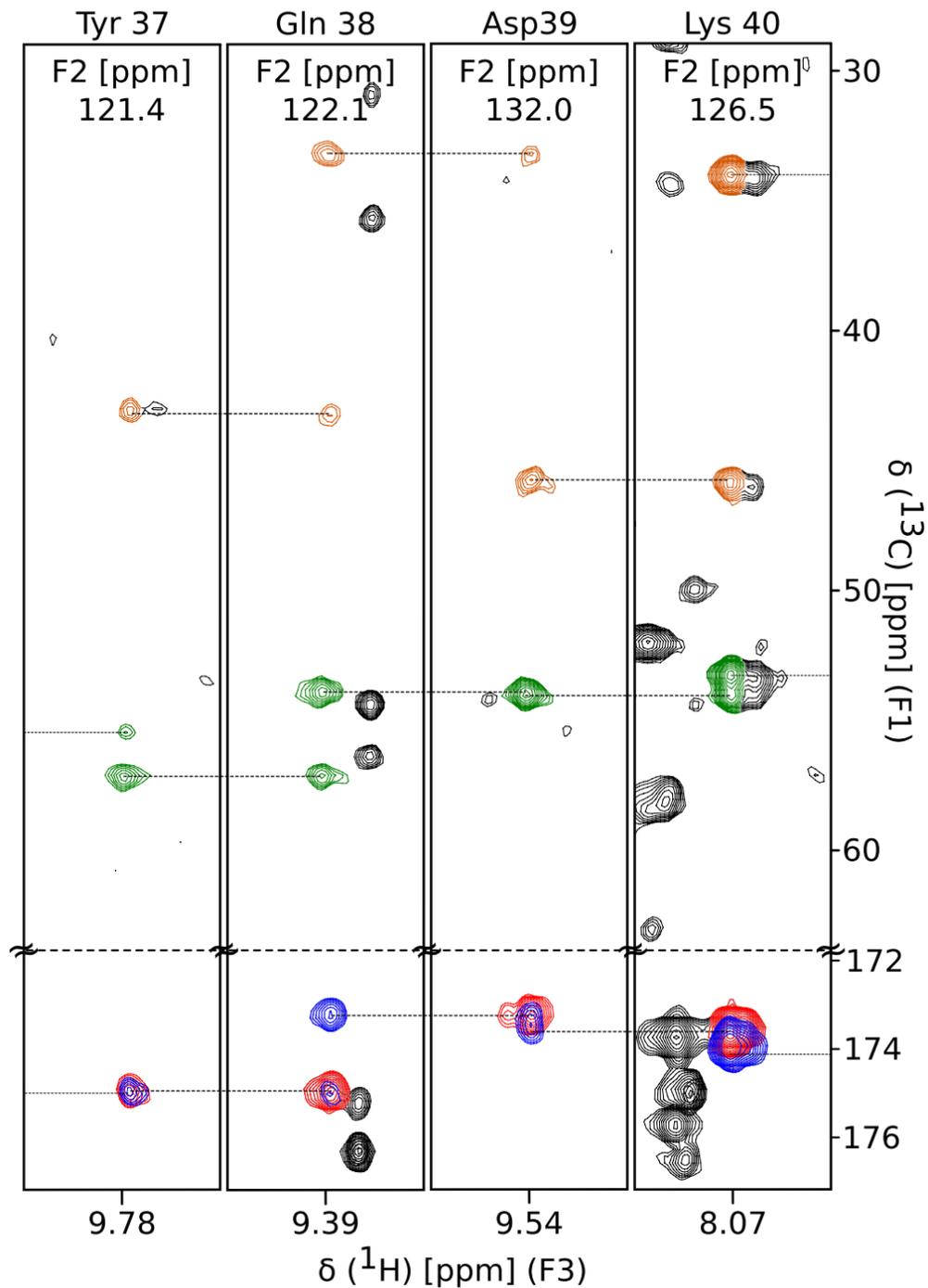
**Figure 1 : Schematic representation of the optimized cell-free protocol for the production of U-[<sup>2</sup>H, <sup>13</sup>C, <sup>15</sup>N] in H<sub>2</sub>O buffer for site specific NMR investigation of Fabs.**



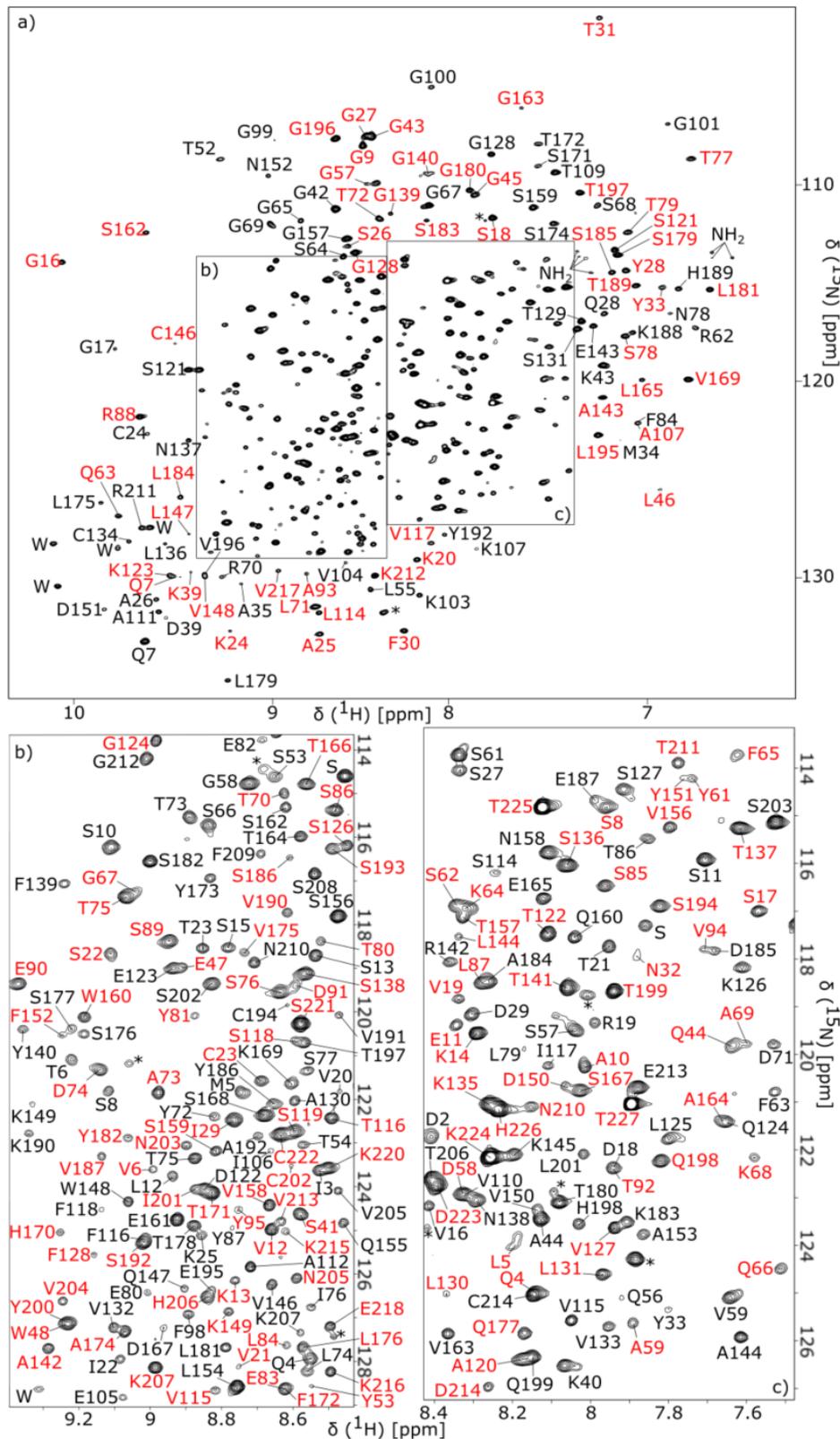
**Figure 2: Optimization of Cell-free production of Fab.** **a)** Detection of soluble Fab produced *in vitro* (50 µL batch production) using western blot performed with an anti-His-tag antibody. The band at 50 kDa stands for the Fab fragment and the one at 25 kDa for the His-tagged light chain produced in excess. Cell-free reactions were performed using S30 extract prepared using (lane A) *E. coli* BL21(DE3) strain; (lane B) *E. coli* SHuffle strain; or (lane C) *E. coli* SHuffle strain with the addition of reduced (1 mM) and oxidized (4 mM) glutathione. **b)** Screening of the LC/HC plasmid ratio on cell free reaction yields. The histogram represents the intensity of Fab (black) and LC (grey) bands quantified on a western blot performed with an anti-HIS-tag antibody. The ratio is the concentration of the light chain plasmid over the concentration of the heavy chain plasmid for a total plasmid concentration of 10 µg/mL in the cell-free reaction. **c)** Screening of the effect of addition of choline, ectoïne and betaine (added at 10 mM concentration each) on cell free production of anti-LAMP1 Fab. The cell free production was performed in CECF mode (dialysis) in a 3 mL GeBAflex and the histogram reports the amount of Fab purified per mL of CECF reaction. The “standard” stands for to the reference cell-free reaction performed without polyions.



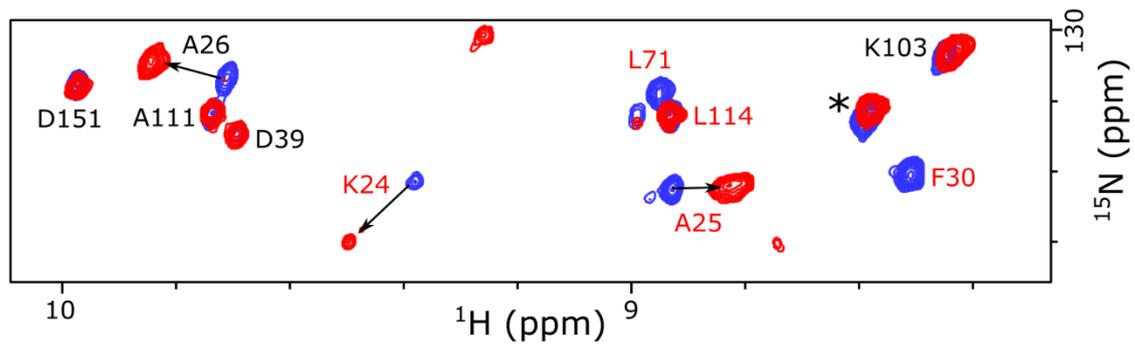
**Figure 3. Large Scale production U-[<sup>2</sup>H,<sup>13</sup>C,<sup>15</sup>N] anti-LAMP1 Fab.** a) SDS-PAGE gel analysis of the Fab purification steps. Lane A: total soluble extract after centrifugation of the cell-free reaction mix; Lane B: Flow through recovered at outlet of HiTrap Streptavidin HP column; Lane C: Sample collected from the elution pic of HiTrap Streptavidin HP column. (b) Fab sample was cleaved by TEV protease to remove purification tags and cleaved Fab was purified by Ni-NTA reverse purification. Cleavage was monitored by SDS-PAGE gel analysis. Lane A: Fab sample after TEV cleavage showing cleaved Fab fragment (47 kDa) and His-TEV (30 kDa); Lane B: purified TEV. Lane C : flow through collected at the outlet of the Ni-NTA column performed after the cleavage that contains the cleaved Fab fragment alone (47 kDa). c) 2D <sup>15</sup>N-TROSY NMR spectrum of the U-[<sup>2</sup>H,<sup>13</sup>C,<sup>15</sup>N] Fab sample.



**Figure 4: Sequential Assignment of anti-LAMP1 Fab.** 2D-strips presenting 2D extracts from 3D BEST-HNCO, -HN(CA)CO, -HNCA, -HN(CA)CB experiments. The selected 2D-strips illustrated assignment of Tyr-37, Gln38, Asn-39, Lys-40 segment from Fab light chain. In blue  $\text{C}'(i-1)$  resonances, in red  $\text{C}'(i)$  resonances, in green  $\text{C}_\alpha(i-1)$  and  $\text{C}_\alpha(i)$  resonances and in orange  $\text{C}_\beta(i-1)$  and  $\text{C}_\beta(i)$  resonances.



**Figure 5: Assigned 2D-<sup>15</sup>N-TROSY spectra of anti-LAMP1 Fab.** a) <sup>15</sup>N BEST-TROSY spectrum of Fab acquired at 308K and at a <sup>1</sup>H frequency of 950 MHz. On panels b) and c), zooms from panel a) are presented. Each assigned signal (in a, b, c) is annotated with the corresponding residue number of the light chain (black) or heavy chain (red). W annotations indicate tryptophan indole signals, NH<sub>2</sub> signals of side chains, S unassigned serine signals, \* unassigned signals characterized by signal/noise > 10.



**Figure 6. Example of chemical shift perturbations on anti-LAMP1 NMR spectra induced by the formation of the complex with LAMP1 antigen.** Overlays of a region of 2D  $^{15}\text{N}$ -TROSY spectra of free Fab (47 kDa) in blue and LAMP1/anti-LAMP1 Fab in red (130 kDa). Fab assigned NMR signals are annotated on Fab blue spectrum and putative corresponding signals in the spectrum of the complex are indicated by arrows. Unassigned Fab NMR signals are identified by an asterisk. Full 2D  $^{15}\text{N}$ -TROSY spectrum of anti-LAMP1/LAMP1 complex is presented on supplementary Figure S5.

Résultats Article : Enabling site-specific NMR investigation of therapeutic Fab using a cell-free based isotopic labeling approach : application to anti-LAMP1 Fab.

## **Enabling Site-Specific NMR Investigations of Therapeutic Fab**

**using a Cell-Free Based Isotopic Labelling Approach:**

### **Application to anti-LAMP1 Fab.**

Arthur Giraud<sup>a,b,c,#</sup>, Lionel Imbert<sup>a,#</sup>, Adrien Favier<sup>a</sup>, Faustine Henot<sup>a,b</sup>, Francis Duffieux<sup>b</sup>, Camille Samson<sup>b</sup>, Oriane Frances<sup>b,\*</sup>, Elodie Crublet<sup>c,\*</sup>, Jérôme Boisbouvier<sup>a,\*</sup>

<sup>a</sup> *Univ. Grenoble Alpes, CNRS, CEA, Institut de Biologie Structurale (IBS), 71, Avenue des Martyrs, F-38044 Grenoble, France.*

<sup>b</sup> *Sanofi Research & Development, 94403 Vitry-sur-Seine, France.*

<sup>c</sup> *NMR-Bio, 5 place Robert Schuman, F-38025 Grenoble, France.*

*# These authors contributed equally to this work*

*\* Correspondence to be addressed [Oriane.Frances@sanofi.com](mailto:Oriane.Frances@sanofi.com) or [crublet@nmr-bio.com](mailto:crublet@nmr-bio.com) or [jerome.boisbouvier@ibs.fr](mailto:jerome.boisbouvier@ibs.fr)*

## **Supplemental material**

Résultats Article : Enabling site-specific NMR investigation of therapeutic Fab using a cell-free based isotopic labeling approach : application to anti-LAMP1 Fab.

**Supplementary Table S1: Experimental details for the affinity chromatography purification of anti-LAMP1 Fab.**

	StrepTrap HP 5 mL	HisTrap FF 1 mL
Device	NGC chromatography system (Bio-Rad)	
Flow rate	4 mL/min	1 mL/min
Equilibration	5 CV of buffer A	5 CV of buffer C and 6 % of buffer D
Sample loading	Sample diluted by 4 in buffer A and loaded at 2 mL/min	Elution from the StrepTrap column loaded at 0,5 mL/min
Wash	6 CV of buffer A	6 CV of buffer C
Elution	Buffer A and 0 % to 100 % of buffer B in 6 CV, and 5 CV of buffer B	Buffer C and 6 % to 100 % of buffer D in 25 CV, and 5 CV of buffer D
Buffer A	100 mM Tris HCl, 1 mM EDTA, 150 mM NaCl pH 7,4	
Buffer B	100 mM Tris HCl, 1 mM EDTA, 150 mM NaCl, 2,5 mM D-desthiobiotine pH 7,4	
Buffer C	2X PBS pH 7,4	
Buffer D	2X PBS, 250 mM Imidazole pH 7,4	

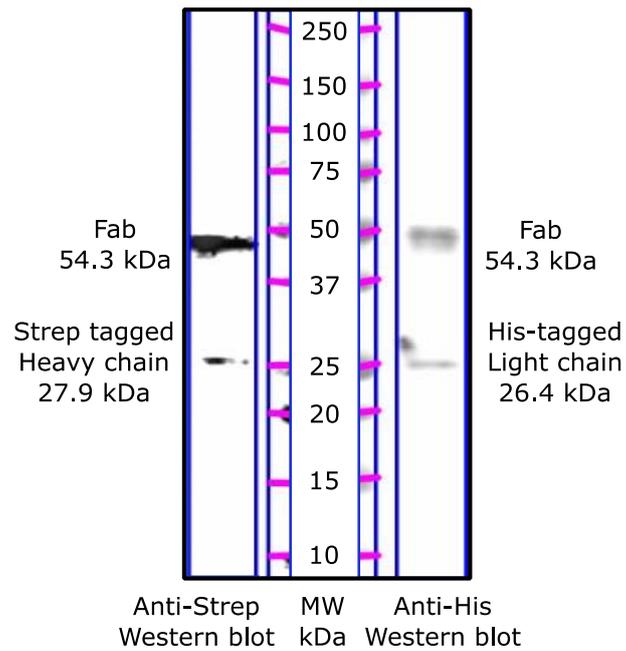
CV : column volum

Résultats Article : Enabling site-specific NMR investigation of therapeutic Fab using a cell-free based isotopic labeling approach : application to anti-LAMP1 Fab.

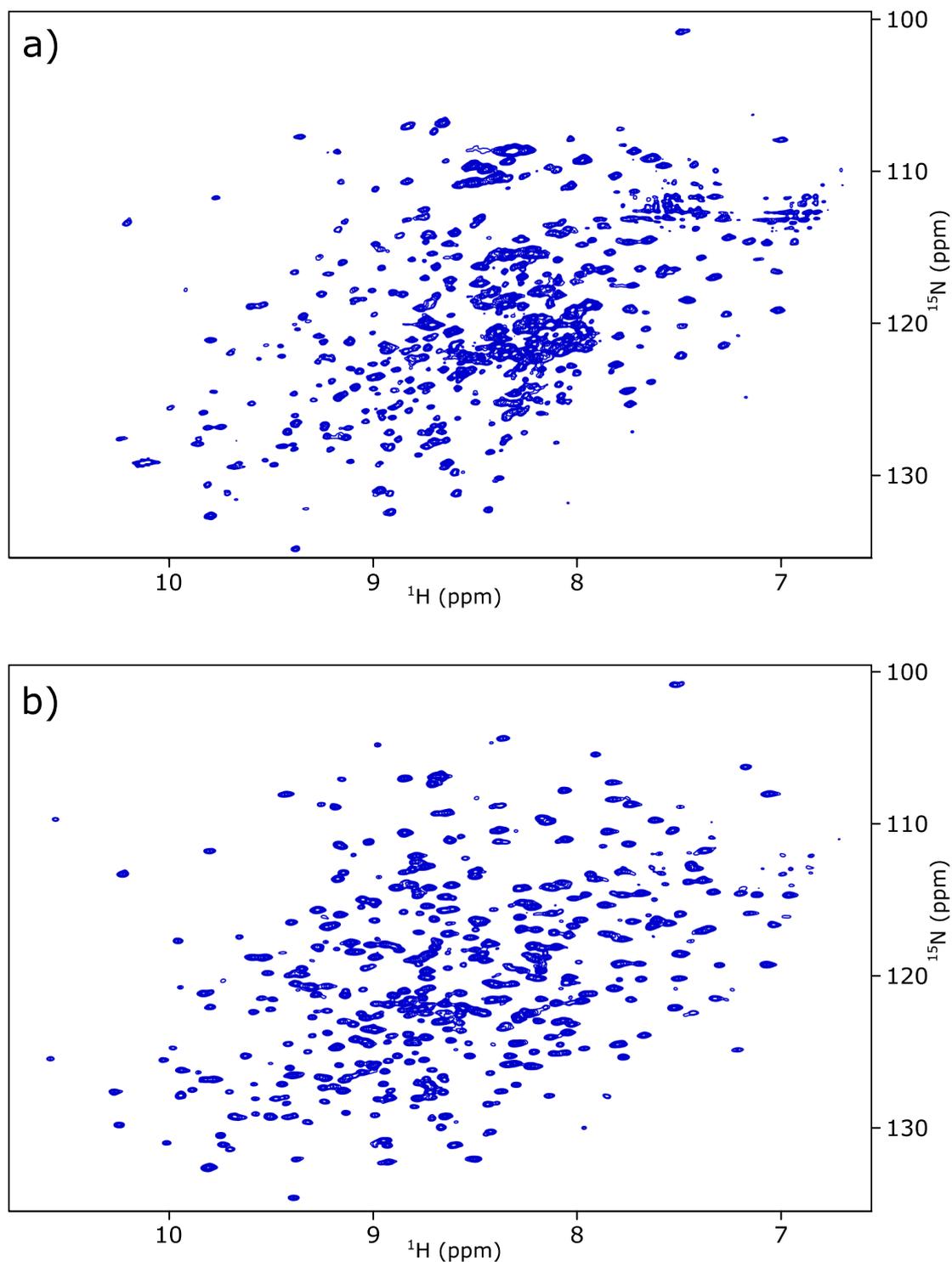
**Supporting Table S2: Summary of NMR experiments acquired for backbone assignment of U- $^2\text{H}$ ,  $^{15}\text{N}$ ,  $^{13}\text{C}$ ]-Fab sample.**

2D-BEST-TROSY <sup>a</sup>	1 h	950 MHz
3D-BEST-HNCO <sup>a</sup>	1 day	700 MHz
3D-BEST-HNCA <sup>a</sup>	2 days	950 MHz
3D-BEST-HN(CA)CB <sup>a</sup>	4 days	950 MHz
3D-BEST-HN(CA)CO <sup>a</sup>	4 days	700 MHz
3D-BEST-HN(CO)CA <sup>a</sup>	2 days	600 MHz
3D-BEST-HN(COCA)CB <sup>a</sup>	5 days	600 MHz

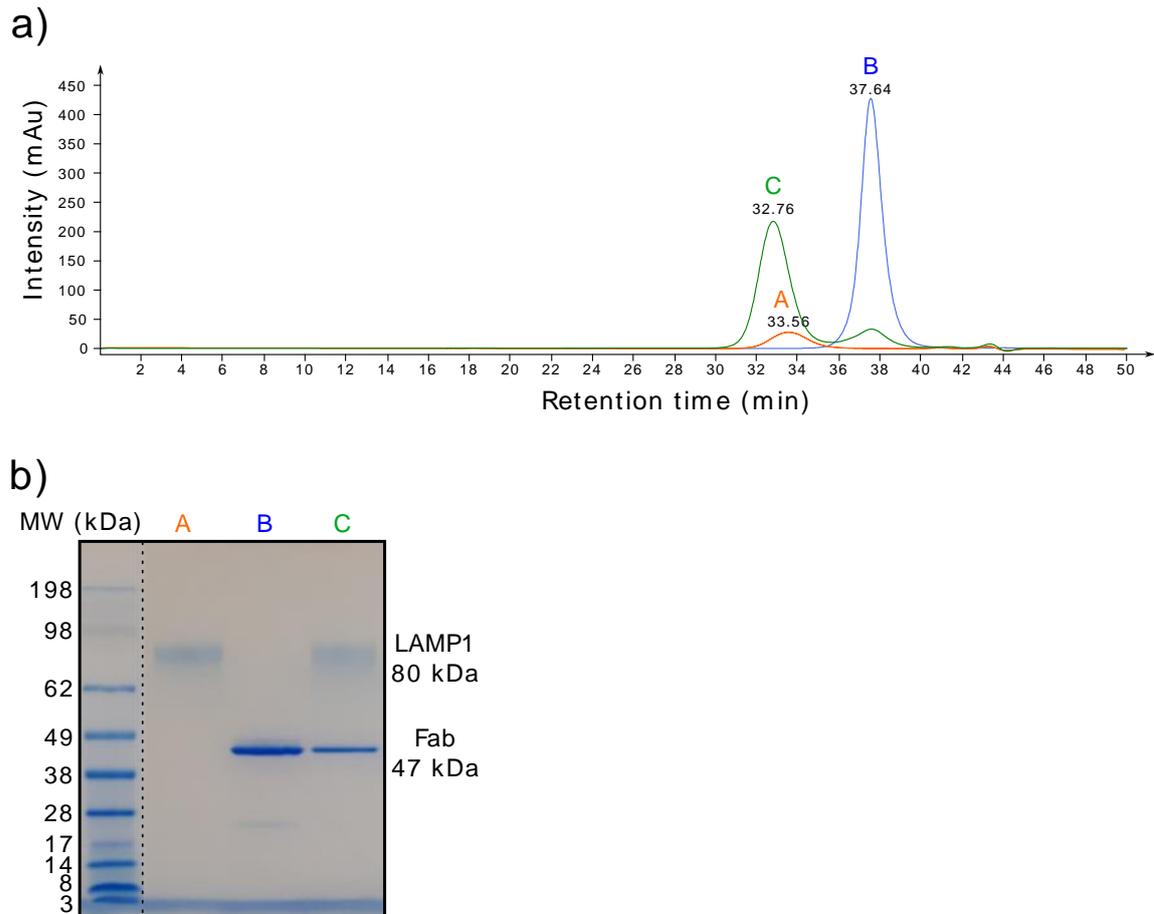
*a. (Lescop et al. J. Magn. Res. 2007; Favier and Brutscher J. Biomol NMR 2019)*



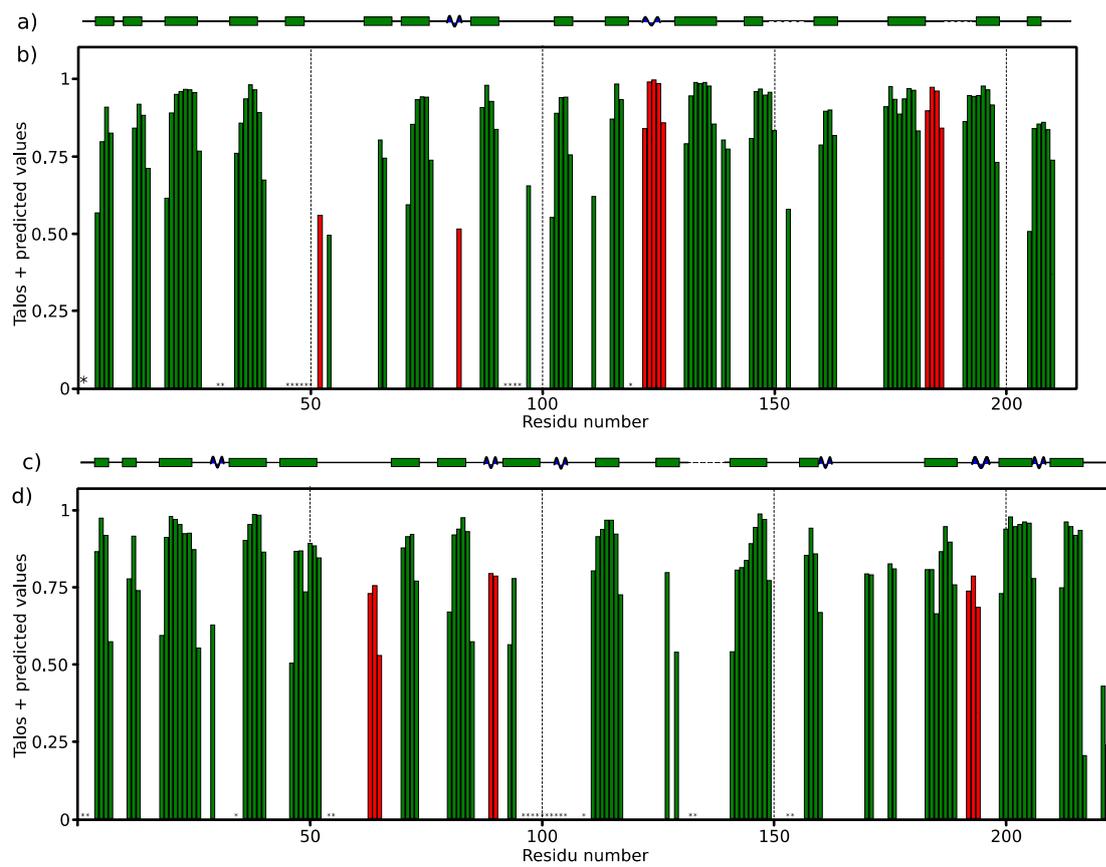
**Supplementary Figure S1:** Detection of soluble Fab produced *in vitro* (batch production of 50  $\mu$ L) by western blot with an anti-strep-tag antibody (left) or an anti-His-tag antibody (right). The light chain has a 6-His tag, while the heavy chain has a strep tag.



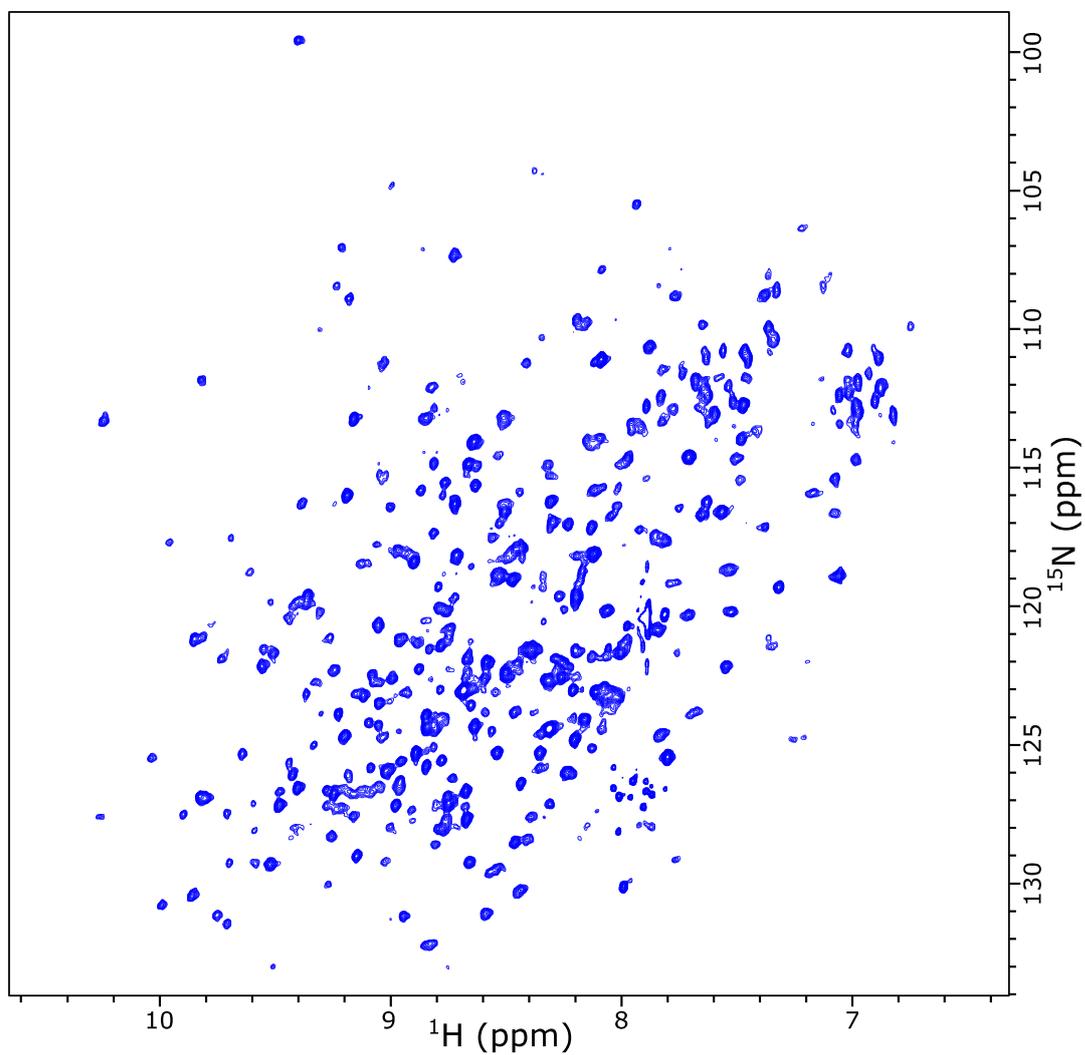
**Supplementary Figure S2:** 2D  $^{15}\text{N}$ -TROSY NMR spectra of U- $[\text{}^2\text{H}, \text{}^{15}\text{N}]$  anti-LAMP1 Fab sample before a) and after b) tags cleavage. Experiments were acquired at 35°C on an NMR spectrometer operating at a  $^1\text{H}$  frequency of 600 MHz and equipped with a cryogenic probe head.



**Supplementary Figure S3: Interaction test between anti-LAMP1 fab and Lamp1 antigen.** a) Size exclusion chromatography characterization of LAMP1 (A), anti-LAMP1 Fab (B) and complex between anti-LAMP1 Fab and LAMP1 antigen (C). This chromatogram was obtained by injecting samples on a superose 6 increase 3.2/300 SEC column. b) SDS-PAGE gel analysis in non-reducing conditions of the three elution pics annotated in a).



**Supplementary Figure S4: Secondary structure predictions – comparison to crystallographic data.** Panels **a** and **c** display the secondary structure elements as reported on the crystallographic structure for Fab Light (a) and Heavy (c) chains in complex with LAMP1 antigen (PDB code: 8ATH) . Panels **b** and **d** present the secondary structure predictions based on backbone assignment of Fab Light (b) and Heavy (d) chains using TALOS +. In red and green are predicted  $\alpha$ -helix and  $\beta$ -sheets secondary structure elements, respectively. TALOS + predicted loops segments are not represented. \* indicate residues for which none of the backbone resonances were assigned.



**Supplementary Figure S5:** 2D  $^{15}\text{N}$ -TROSY NMR spectrum of U- $^{2}\text{H}$ ,  $^{15}\text{N}$ ] anti-LAMP1 Fab sample in complex with LAMP1 antigen. Experiments were acquired at 35°C on an NMR spectrometer operating at a  $^1\text{H}$  frequency of 950 MHz and equipped with a cryogenic probe head.

### **3. Attribution des groupements méthyles du fragment Fab exprimé en système *cell-free***

L'interaction du fragment Fab avec son antigène LAMP1 a été présentée dans l'article de la partie 3.2.3 avec la superposition des spectres RMN  $^1\text{H}$ - $^{15}\text{N}$  du fragment seul et en complexe. Bien que des modifications de déplacement chimique soient observées, il est impossible de transférer l'attribution du squelette du fragment Fab seul au fragment en complexe avec LAMP1 en raison de la complexité du spectre RMN  $^1\text{H}$ - $^{15}\text{N}$  du fragment Fab, dû au nombre élevé de signaux (plus de 440) et à la taille importante du complexe Fab/LAMP1 (environ 130 kDa) qui impacte la relaxation et donc la qualité des signaux. Cela souligne le besoin d'utiliser la RMN des groupements méthyles pour étudier la structure des fragments d'anticorps en solution.

Avant de procéder à l'attribution des groupements méthyles du fragment Fab exprimé en système *cell-free*, le protocole de production a dû être optimisé afin d'augmenter les rendements.

#### **3.1. Optimisations du protocole de production du fragment Fab en système *cell-free***

La stratégie choisie pour l'attribution des groupements méthyles du fragment Fab est d'acquérir des expériences RMN 3D de type HCC pour corrélérer les résonances des  $\text{C}\alpha$  et  $\text{C}\beta$  attribuées du squelette aux résonances des spectres RMN 2D  $^1\text{H}$ - $^{13}\text{C}$ . Il a été décidé d'acquérir, en complément, des expériences RMN NOESY inter-méthyles afin d'attribuer les groupements méthyles grâce à la connaissance de la structure du fragment Fab. Pour ces deux types d'expériences RMN, une protéine de la taille d'un fragment Fab doit être marquée spécifiquement sur les groupements méthyles et perdeutérée. Ce marquage isotopique, réalisé à partir d'acides aminés ou de précurseurs marqués est plus coûteux que le marquage U- $[\text{}^2\text{H}, \text{}^{13}\text{C}, \text{}^{15}\text{N}]$  réalisé pour l'attribution du squelette de fragment. Ainsi, une optimisation du protocole de production du fragment Fab en système *cell-free* a été réalisée afin d'augmenter les rendements d'expression et de réduire les coûts de production des échantillons nécessaires à l'attribution des groupements méthyles.

##### **3.1.1. Etude du rôle de la protéine chaperonne DsbC recombinante ajoutée dans le système *cell-free***

Pour l'expression du fragment Fab U- $[\text{}^{13}\text{C}, \text{}^{15}\text{N}, \text{}^2\text{H}]$  de l'anticorps en système *cell-free*, le potentiel redox avait déjà été adapté grâce à l'ajout du couple glutathion oxydé/glutathion réduit (*ratio* GSH : GSSG = 1 : 4 mM). Afin de favoriser la bonne isomérisation des ponts disulfures, l'extrait cellulaire utilisé provenait de la souche SHuffle de *E. coli*. Cette souche bactérienne ne possède pas les gènes des réductases *trx*B et *gor* et surexprime la protéine DsbC qui est une isomérase à pont disulfures (cf. partie 1.3.3.1). Elle peut, de ce fait, surexprimer des protéines à pont disulfures bien repliées, comme les anticorps [193]. L'inconvénient d'utiliser cette souche pour les extraits bactériens utilisés en *cell-free* est qu'il n'est pas possible de quantifier la DsbC présente dans les extraits et de vérifier qu'elle est toujours aussi active après

la lyse des bactéries. Par exemple, Jared L. Dopp et N. Reuel ont montré que l'utilisation d'un extrait de la souche SHuffle était moins efficace que l'utilisation d'un extrait provenant d'une souche bactérienne standard comme les KGK10, complémenté avec la protéine chaperonne DsbC (100 µg/mL ou 4,3 µM) [175].

Nous avons donc choisi de comparer les deux protocoles. Pour cela, nous avons, au préalable, produit et purifié de la DsbC recombinante en système d'expression bactérien. L'extrait SHuffle (utilisé pour la préparation du fragment Fab U-[<sup>13</sup>C, <sup>15</sup>N, <sup>2</sup>H]) a été comparé à un extrait préparé à partir d'une souche de BL21 (DE3) puis complémenté avec 15 µM de DsbC purifiée dans le mélange final de la réaction *cell-free*. Ces productions ont été réalisées dans un milieu réactionnel protoné en suivant les protocoles d'expression et de purification du fragment Fab détaillés dans la partie 3.2.3.

Les résultats obtenus rejoignent ceux de Jared L. Dopp et N. Reuel (figure 3.3.1a). Le rendement d'expression obtenu en utilisant l'extrait des BL21 complémenté avec 15 µM de DsbC purifiée est de 0,56 µg/mL, soit 2,2 fois supérieur à celui obtenu en utilisant l'extrait des bactéries SHuffle (0,25 µg/mL).

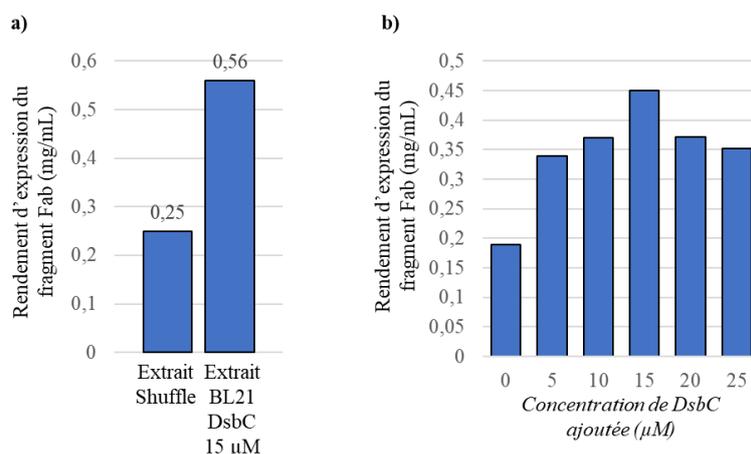


Figure 3.3.1 : Histogramme représentant l'optimisation du rendement d'expression du fragment Fab de l'anticorps anti-ALMP1 en système cell-free grâce à l'ajout de DsbC recombinante. a) Rendements d'expression du fragment Fab en système cell-free en utilisant un extrait SHuffle contenant de la DsbC endogène et un extrait de BL21 supplémenté avec 15 µM de DsbC exogène. b) Rendements d'expression du fragment Fab en système cell-free en fonction de la quantité de DsbC exogène ajoutée dans l'extrait de BL21. Les expériences ont été réalisées en H<sub>2</sub>O pour a) et en D<sub>2</sub>O pour b).

La concentration de DsbC ajoutée a ensuite été optimisée en utilisant l'extrait provenant des bactéries BL21 complémenté avec des concentrations de DsbC recombinante variant de 0 à 25 µM. Ces expressions ont été réalisées dans un milieu réactionnel deutéré (solvant D<sub>2</sub>O) pour évaluer, simultanément, le rendement d'expression lors de la production du fragment Fab perdeutééré. Les résultats sont présentés dans la figure 3.3.1b. Le rendement d'expression le plus élevé (0.4 mg de fragment Fab par mL de réaction *cell-free*) a été obtenu pour 15 µM de DsbC ajoutée. La différence de rendement entre les 2 figures (a et b) pour 15 µM de DsbC s'explique par le fait que la réaction a été réalisée en milieu protoné dans un cas (a) et deutéré dans l'autre cas (b). Il est connu que le D<sub>2</sub>O affecte les rendements d'expression de protéines, notamment dans le système *cell-free* [145].

L'ajout de 15  $\mu\text{M}$  de DsbC dans la réaction *cell-free* a permis de multiplier le rendement de production du fragment Fab par un facteur 2, ce qui permet, pour une même quantité finale, de fragment Fab, de diviser les volumes réactionnels par 2 et donc d'utiliser deux fois moins d'acides aminés marqués.

Lors d'expériences préliminaires, il a été observé que la chaîne lourde est moins soluble que la chaîne légère. Ainsi, pour obtenir un maximum d'appariement des deux chaînes, la chaîne lourde doit être plus exprimée que la chaîne légère. Cela se traduit par l'ajout du plasmide de la chaîne lourde en excès par rapport à celui de la chaîne légère. Cependant, l'appariement des deux chaînes entre elles est favorisé lors de l'ajout de DsbC dans la réaction *cell-free*. Ainsi, le rapport des concentrations des deux plasmides ajoutés (pour la chaîne légère et lourde) a été optimisé de nouveau en suivant le protocole détaillé dans la partie 3.2.3. Le rapport optimal pour les expériences suivantes a été défini comme tel (en quantité massique) : 2 chaînes lourdes pour 1 chaîne légère sur une base totale de 16  $\mu\text{g/mL}$  de plasmide. Ce nouveau rapport contient plus de chaîne légère par rapport à la chaîne lourde que le précédent (5 chaînes lourdes pour 1 chaîne légère).

### 3.1.2. Diminution de la quantité de réactifs du système *cell-free*

Dans le système *cell-free*, les réactions de synthèse peuvent être réalisées en « batch », où tous les composés sont ajoutés dans le tube de synthèse au début de la réaction, ou en système CECF. (pour *continuous exchange cell-free*) appelé également « dialyse ». Dans ce cas, la réaction est réalisée dans un tube de dialyse, alimenté en continu en énergie et en substrats, dont des acides aminés. Cela permet d'augmenter la durée de vie de la réaction et donc les rendements de protéines et d'éliminer les déchets de synthèse produits pendant la réaction. Le *ratio* volumétrique réaction : tampon de dialyse est en général fixé à 1 :10.

Pour des échantillons marqués, les acides aminés (ou précurseurs d'acides aminés) marqués doivent être ajoutés dans le tube de dialyse et dans le tampon réactionnel. Un échantillon réalisé dans ces conditions contient alors 11 fois plus d'acides aminés marqués qu'un échantillon produit en mode « batch ».

Nous avons voulu vérifier s'il était possible de diminuer ce *ratio* de 1 :10 sans impacter le rendement d'expression. Ainsi, plusieurs réactions *cell-free* ont été effectuées avec des *ratio* "réaction/dialyse" de 1 :10, 1 :6, 1 :5, 1 :4 et 1 :3.

Les résultats obtenus (figure 3.3.2) montrent que le rendement d'expression du fragment Fab ne varie pas lorsque le *ratio* diminue de 1 :10 à 1 :6 et 1 :5. Cela signifie que le rendement d'expression n'est pas impacté par une diminution d'un facteur deux de la quantité de réactifs utilisés.

Ces résultats nous ont permis de valider le *ratio* 1 :5 pour les synthèses de Fab marqué et donc de diminuer d'un facteur deux la quantité de réactifs marqués.

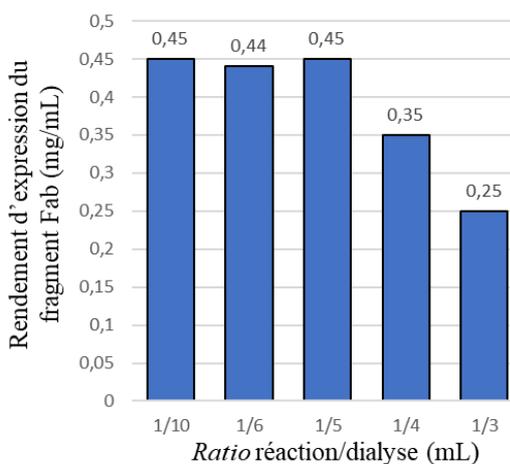


Figure 3.3.2: Rendements d'expression du fragment Fab obtenus en dialyse *cell-free* en utilisant des *ratio* "réaction/dialyse" différents

### 3.1.3. Optimisation du protocole de purification du fragment Fab

Le protocole de purification du fragment Fab mis en place précédemment consistait à utiliser deux colonnes de purification successives basées sur l'affinité pour les deux tags (His Tag et Strep-Tag) ajoutés au niveau des domaines N-terminaux des chaînes légères et lourdes. Or, le rendement diminue généralement avec le nombre d'étapes de purification. Une nouvelle méthode de purification a donc été mise en place en utilisant, à la place des 2 purifications successives, une colonne de résine protéine L. Ce type de résine se lie spécifiquement à la région variable des chaînes  $\kappa$  des chaînes légères des anticorps. Cette colonne de chromatographie d'affinité est basée sur une rétention des fragments d'anticorps à pH neutre (supérieur à 6) et leur élution à pH acide (inférieur à 3,5).

La purification a été menée en parallèle en utilisant l'ancien et le nouveau protocole. La figure 3.3.3 présente le profil d'élution obtenu sur la nouvelle colonne, et la figure 3.3.4 présente

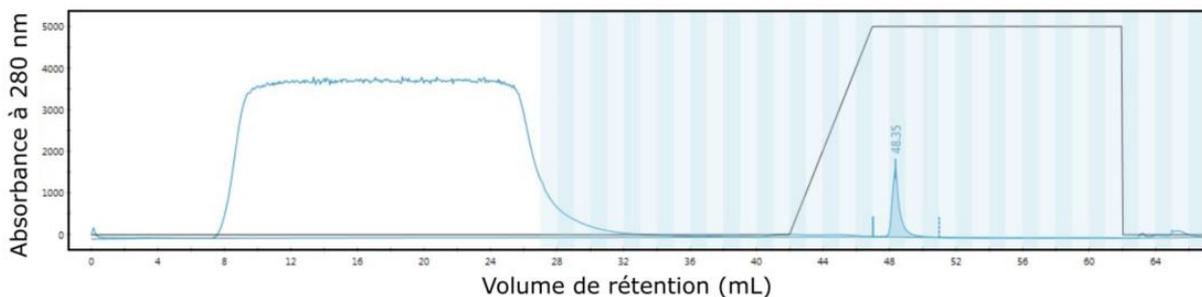


Figure 3.3.3 : Chromatogramme obtenu pour la purification du fragment Fab sur une colonne Hi-Trap protein L de 1mL. Débit de 1 mL/min. Les rayures représentent les fractions de collection. Le pourcentage de tampon éluant est représenté en noir sur le chromatogramme.

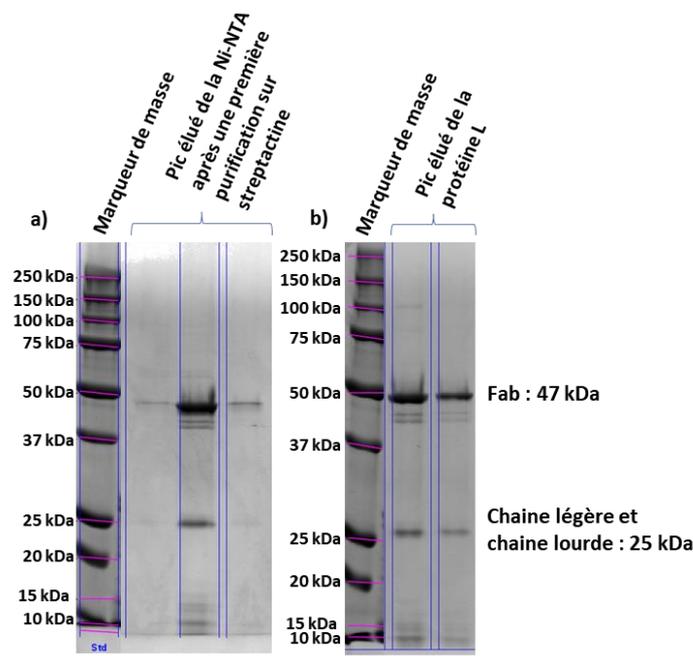


Figure 3.3.4 : Gel SDS-page des fractions de collections contenant les pics d'élution du fragment Fab purifié sur une colonne de streptactine suivie d'une colonne de Ni-NTA a), et purifié en utilisant la colonne Hi-Trap protein L b).

une comparaison du dépôt sur gel SDS-PAGE des pics d'élution obtenu pour les deux méthodes.

Les résultats montrent que le rendement obtenu avec une purification par affinité sur la protéine L est 1.4 fois supérieur à celui obtenu avec la purification par résine streptactine suivie d'une résine Ni-NTA. Cela s'explique par le fait que chaque étape de traitement de l'échantillon est une source de perte et qu'il est donc plus avantageux de n'utiliser qu'une seule méthode de purification si celle-ci est aussi efficace que deux successives.

Nous avons donc choisi de poursuivre avec la protéine L. Par ailleurs, à terme, cette méthode nous permettra de nous affranchir des tag qui peuvent venir perturber le signal RMN par leur flexibilité.

Le fragment Fab purifié par affinité sur la protéine L a ensuite été chargé sur une colonne de chromatographie par exclusion de taille pour vérifier l'absence d'agrégats. Le profil d'élution est présenté en figure 3.3.5.

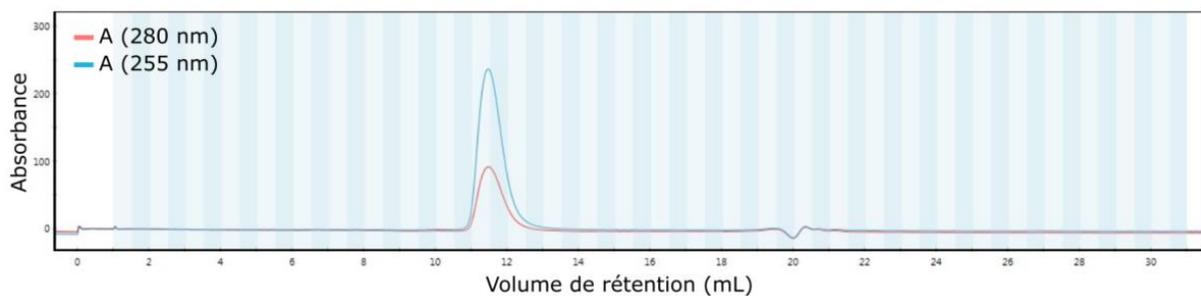


Figure 3.3.5 : Chromatogramme obtenu en utilisant une colonne chromatographique d'exclusion de taille Superdex 75 10/300 pour la purification du fragment Fab exprimé en système cell-free et purifié une première fois sur protéine L. Le volume de rétention du pic correspond à une protéine de 50 kDa comme le fragment Fab.

L'optimisation du protocole de production du fragment Fab en système *cell-free* a consisté à utiliser de la DsbC recombinante à 15  $\mu$ M, à diminuer le volume du tampon de dialyse et donc la quantité de réactifs par un facteur deux, et à mettre en place une méthode de purification plus efficace que la précédente. Cela a été réalisé afin de réduire les coûts de production d'échantillons marqués isotopiquement en raison du coût des acides aminés marqués. On peut donc parler de rendement par rapport à la quantité d'acides aminés dans la réaction *cell-free*. Ainsi, l'ensemble de ces optimisations a permis d'augmenter le rendement par un facteur 5.4 par rapport à la quantité d'acides aminés utilisés.

## 3.2. Marquage isotopique sélectif des groupements méthyles du fragment Fab

### 3.2.1. Production en système *cell-free* du fragment Fab marqué sélectivement sur les groupements méthyles

Dans un premier temps, le marquage spécifique des groupements méthyles du fragment Fab a été réalisé sur un type de résidu à la fois. Parmi les six acides aminés comportant un groupement méthyle, seule la leucine n'a pas fait l'objet d'un marquage sélectif en raison d'une disponibilité commerciale limitée pour ce résidu marqué. Les méthionines et les isoleucines ont été marquées dans le même échantillon. En effet, les déplacements chimiques respectifs de leurs résonances sont suffisamment différents pour pouvoir les différencier sur les spectres. Pour les expériences RMN 2D des groupements méthyles d'une protéine de 50 kDa, la perdeutération de la protéine n'est pas nécessaire. En effet, la deutération des chaînes latérales des acides aminés marqués sur les groupements méthyles est suffisante pour obtenir un compromis entre le gain de sensibilité et de résolution par rapport au coût de production de ces échantillons (cf. partie 3.1.1.3 c). Pour éviter l'échange du deutérium des groupements amides de ces acides aminés avec l'hydrogène du solvant, les échantillons (Volume 5 mL- *ratio* de dialyse 1 :5) ont été produits dans un milieu réactionnel préparé en D<sub>2</sub>O. Cela a également pour effet d'éviter l'échange du deutérium en position C $\alpha$  des acides aminés par les transaminases. Le marquage isotopique des groupements méthyles a été réalisé de façon stéréospécifique sur les valines en position *proR* et régiospécifique sur les isoleucines en position  $\delta 1$ . Ce type de marquage isotopique concorde avec celui réalisé sur l'anticorps produit en cellules CHO (cf. partie 3.1.2) et permet de faciliter la compréhension des spectres (une résonance par acide aminé marqué) ainsi que d'éliminer les interactions dipolaires créés par les protons d'un groupement méthyle et favoriser le phénomène de relaxation. Ainsi quatre échantillons ont été produits avec le marquage isotopique présenté en figure 3.3.6.

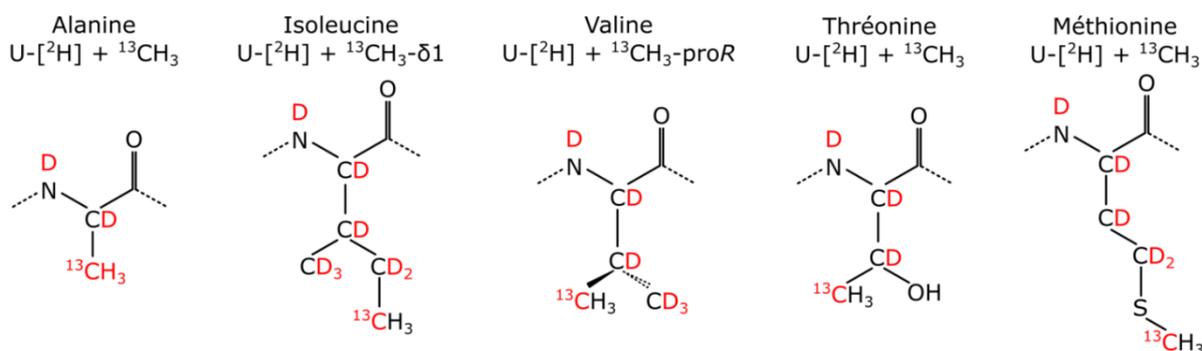


Figure 3.3.6 : Marquage isotopique des groupements méthyles réalisé pour le marquage sélectif des acides aminés comportant un groupement méthyle. Les carbones marqués en <sup>13</sup>C et la deutération sont représentés en rouge.

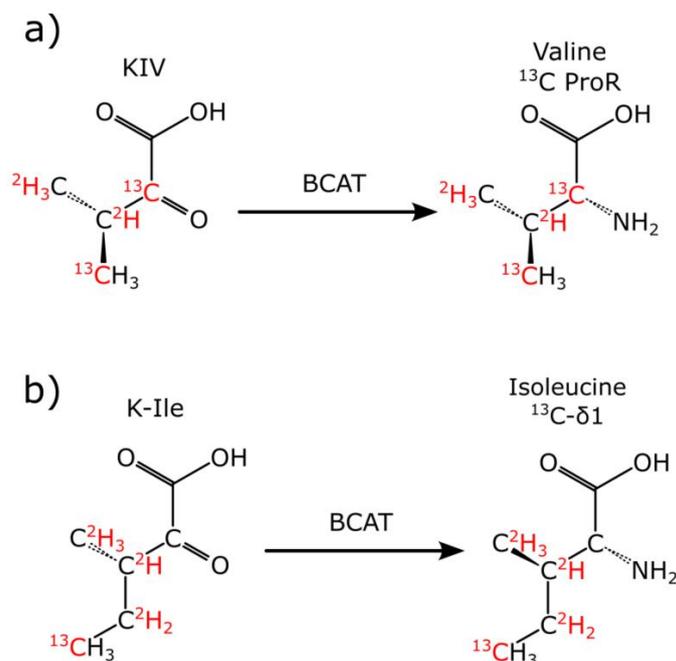


Figure 3.3.7 : Synthèse des acides aminés isoleucines et valines à partir de leur précurseur par l'enzyme BCAT (*branched-chain amino acid amino-transferase*). a) Synthèse de la valine marquée  $^{13}\text{C}$  en position proR en utilisant le 2-Keto-3-( $^2\text{H}$ )-3-( $^{13}\text{C}^2\text{H}_3$ -méthyle)-2,4- $^{13}\text{C}_2$  butyrate (KIV) marqué  $^{13}\text{C}$  en position proR. b) Synthèse de l'isoleucine marquée en position  $\delta 1$  en utilisant le 2-kéto-3-( $^{12}\text{C}^2\text{H}_3$ -méthyle)-3,4-( $^2\text{H}_2$ )-5( $^{13}\text{C}$ )-pentanoate marqué en position  $\delta 1$ .

Les alanines, méthionines et thréonines ont été marquées en ajoutant directement l'acide aminé marqué isotopiquement. Pour marquer l'isoleucine et la valine, le (S)-3 -méthyle-2-ketopentanoate (appelé K-Ile) marqué (2-kéto-3-( $^{12}\text{C}^2\text{H}_3$ -méthyle)-3,4-( $^2\text{H}_2$ )-5( $^{13}\text{C}$ )-pentanoate) et le 2-ketoisovalerate (appelé KIV) marqué (2-Keto-3-( $^2\text{H}$ )-3-( $^{13}\text{C}^2\text{H}_3$ -méthyle)-2,4- $^{13}\text{C}_2$  butyrate) ont respectivement été utilisés. Ces précurseurs d'acides aminés sont transformés en isoleucine et en valine par l'enzyme BCAT (*branched-chain amino acid aminotransferase*)[194] naturellement contenue dans les extraits bactériens (figure 3.3.7).

### 3.2.2. Acquisition de spectres RMN 2D $^1\text{H}$ - $^{13}\text{C}$ du fragment Fab marqué sélectivement

Les 4 échantillons produits ont été analysés par RMN par acquisition d'expériences méthyles-SOFAST (figure 3.3.8).

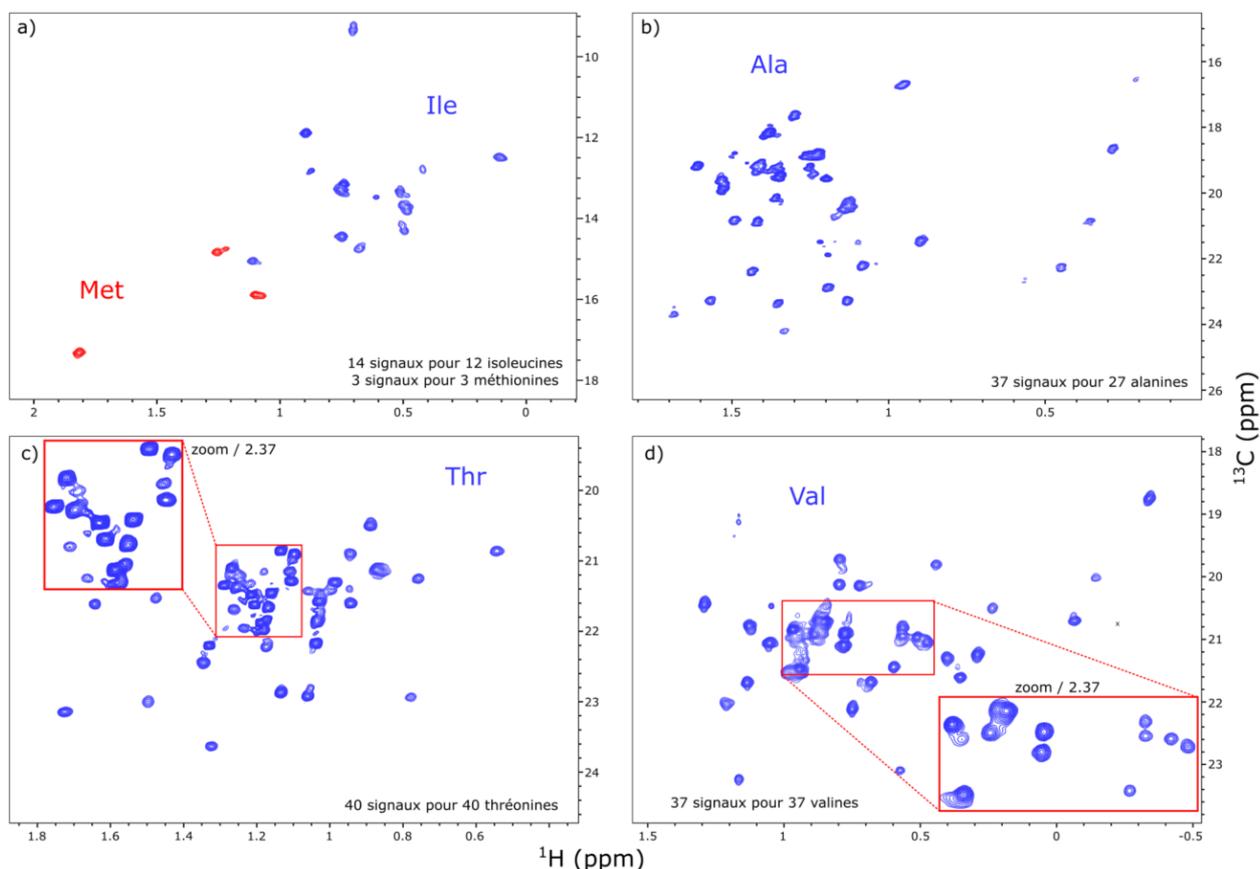


Figure 3.3.8 : Spectres RMN 2D methyl-SOFAST du fragment Fab marqué sur les groupements méthyles et exprimé en système cell-free. a) Spectre du fragment Fab marqué  $^{13}\text{CH}_3$ -Met,  $^{13}\text{CH}_3$ -Ile- $\delta$ 1. Les résonances appartenant à des méthionines sont indiquées en rouge. b) Spectre du fragment Fab marqué  $^{13}\text{CH}_3$ -Ala. c) Spectre du fragment Fab marqué  $^{13}\text{CH}_3$ -Thr. d) Spectre du fragment Fab marqué  $^{13}\text{CH}_3$ -Val-proR. Les septres ont été acquis sur un spectromètre RMN opérant à une fréquence de proton de 950 MHz équipé d'une sonde cryogénique. Les acquisitions ont duré 30 min pour des échantillons concentrés de 20 à 70  $\mu\text{M}$  en fragment Fab dans des tubes Shigemi de 4 mm.

L'analyse de ces spectres a permis d'identifier et de compter le nombre de résonances appartenant à chaque type d'acide aminé portant un groupement méthyle (tableau 3.3.1). Le nombre de résonances correspond exactement au nombre de méthionines, thréonines et valines dans le fragment Fab. Pour les alanines et les isoleucines, la qualité des spectres n'a pas permis de compter les résonances avec précision et plus de résonances qu'attendues sont visibles. Cela peut s'expliquer à cause d'une dégradation de l'échantillon. Pour éviter cela, de l'inhibiteur de protéase a été ajouté dans les échantillons par la suite.

Tableau 3.3.1 : Nombre de pics identifiés sur les spectres 2D  $^1\text{H}$ - $^{13}\text{C}$  SOFAST du fragment Fab marqué spécifiquement sur les groupements méthyles et exprimé en système cell-free et occurrence de ces résidus dans la séquence de l'anticorps.

Acide aminé marqué	Nombre de résonances dans le spectre 1H-13C-SOFAST correspondant	Occurrence de l'acide aminé dans la séquence du fragment Fab de l'anticorps anti-LAMP1 $\beta$
Alanine	37	28
Isoleucine	14	12
Méthionine	3	3
Thréonine	40	40
Valine	37	37
Leucine*	30	30

### 3.2.3. Validation de la stratégie de division de l'anticorps en fragment pour l'attribution

Ces spectres ont ensuite été superposés à ceux obtenus pour l'échantillon de mAb produit en cellules CHO (cf. partie X) afin de comparer leur profils et donc la structure de leurs domaines Fab (figure 3.3.9).

On peut voir que les résonances des groupements méthyles du fragment Fab exprimé en système *cell-free* ont le même déplacement chimique que celles du fragment Fab provenant de l'anticorps exprimé en cellules CHO. En effet, tous les pics sont superposés pour les méthionines, alanines, isoleucines, thréonines et valines, ce qui indique que la structure du fragment Fab isolé exprimé en système *cell-free* est correcte.

Il est à noter que les résonances des groupements méthyles de l'anticorps qui ne sont pas superposées à celles du fragment Fab exprimé en système *cell-free* appartiennent au fragment Fc de l'anticorps.

Les leucines marquées dans l'échantillon produit en système *cell-free*, ayant servi à acquérir les expériences RMN NOE (cf. partie 3.3.4), ont pu être déterminées par déduction après avoir identifié les résonances des cinq autres résidus sur le spectre. La superposition des résonances des leucines du fragment Fab exprimé en système *cell-free* et venant de l'anticorps exprimé en cellules CHO est montré en annexe 6.4.

## Résultats Attribution des groupements méthyles du fragment Fab exprimé en système cell-free

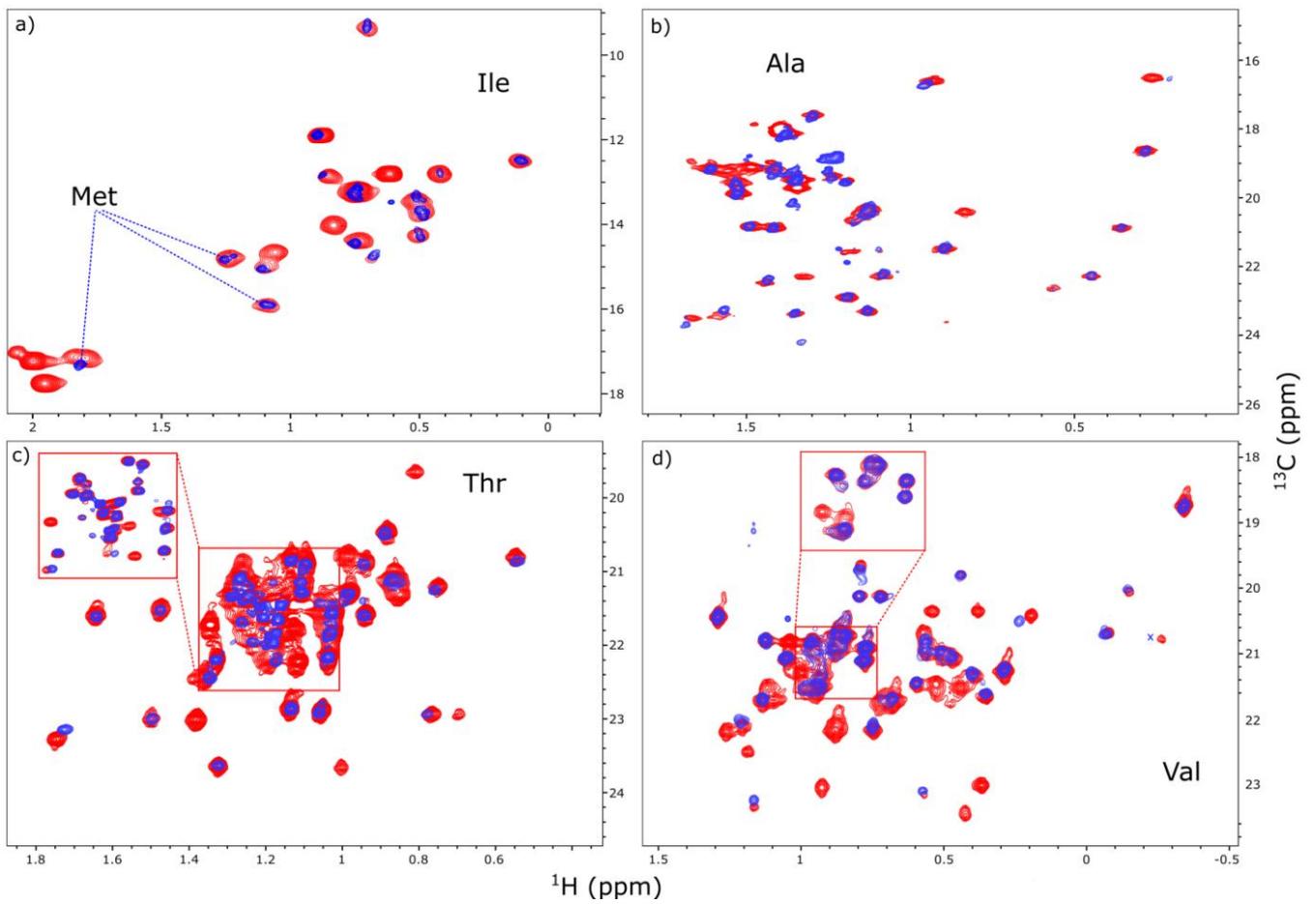


Figure 3.3.9 : Superposition des spectres RMN 2D methyl-SOFAST du fragment Fab marqué sur les méthyles et exprimé en système cell-free (pics bleus) avec ceux de l'anticorps anti-LAMP1 marqué et exprimé en cellules CHO (pics rouges, cf. partie 3.1.2). Il s'agit d'une superposition des spectres présentés dans les figures 3.1.19 et 3.3.8 a) Superposition des spectres pour les méthionines et isoleucines, les pics des méthionines appartenant au fragment Fab sont indiquées par les traits en pointillés. b) Superposition des spectres pour les alanines. c) Superposition des spectres pour les thréonines. d) Superposition des spectres pour les valines.

Cette superposition de spectres valide donc la stratégie d'attribution qui a été présentée en partie 1.5.2. L'expression du fragment Fab de l'anticorps peut être réalisée en système *cell-free* pour attribuer les résonances des groupements méthyles du fragment Fab de l'anticorps exprimé en cellules CHO.

### 3.3. Attribution des groupements méthyles en utilisant des expériences

#### RMN de type HCC

##### 3.3.1. Production du fragment Fab perdeutééré en système *cell-free* avec un marquage isotopique $^{13}\text{C}$ linéarisé sur les alanines, isoleucines et valines

Pour l'attribution des groupements méthyles d'une protéine, l'attribution des résonances du squelette de cette protéine peut être utilisée comme point d'ancrage (cf. partie 3.2.3). Pour le fragment Fab, cela a été réalisé grâce à l'acquisition d'expériences RMN de type HCC "out and back" qui permettent de transférer l'aimantation en partant des groupements méthyles pour aller jusqu'aux atomes du squelette ( $\text{C}\alpha$  ou  $\text{C}\beta$ )[122]. Pour cela, les protéines de haut poids moléculaire comme un fragment Fab d'anticorps (50 kDa) doivent être perdeutéérées pour améliorer les propriétés de la protéine face à la relaxation du signal RMN. Le marquage isotopique  $^{13}\text{C}$  doit être linéaire pour permettre le transfert de magnétisation et corrélérer ainsi les  $\text{C}\alpha$  ou  $\text{C}\beta$  aux carbones des groupements méthyles. Pour la simplification des spectres RMN, le marquage isotopique des groupements méthyles doit également être stéréospécifique. Le marquage isotopique idéal pour ces expériences est donc une perdeutération de la protéine avec un marquage  $^{13}\text{C}$  linéaire,  $^{13}\text{C}^1\text{H}_3$  sur un groupement méthyle par acide aminé et  $^{12}\text{C}^2\text{H}_3$  pour le groupement méthyle non marqué (pour les valines, les leucines et les isoleucines).

Le marquage isotopique  $^{13}\text{C}$  linéaire n'est pas possible pour les méthionines en raison de l'atome de soufre qui sépare le groupement méthyle du reste de la chaîne latérale. Il n'est donc pas possible d'attribuer les groupements méthyles des méthionines avec la méthode des expériences RMN HCC. Cela ne peut pas non plus être réalisé pour les leucines et les thréonines. Bien que ces acides aminés soient commercialement disponibles avec un marquage isotopique  $^{13}\text{C}$  uniforme (qui permettrait de transférer la magnétisation pendant les expériences RMN HCC), ils ne sont pas disponibles avec de la deutération uniforme et une protonation spécifique de leur groupement méthyle (proS pour les leucines).

Un échantillon de fragment Fab a donc été préparé, dans lequel les alanines ont été marquées  $\text{U-}^2\text{H}$ ,  $^{13}\text{C}\alpha$ ,  $^{13}\text{C}\beta$ ,  $^1\text{H}\beta$ , les isoleucines  $\text{U-}^2\text{H-}^{13}\text{C}\alpha^1$ ,  $^{13}\text{C}\beta$ ,  $^{13}\text{C}\gamma_1$ ,  $^{12}\text{C}\gamma_2$ ,  $^{13}\text{C}\delta_1$ ,  $^1\text{H}_3\delta_1$  et les valines  $\text{U-}^2\text{H-}^{13}\text{C}\alpha$ ,  $^{13}\text{C}\beta$ ,  $^{12}\text{C}\gamma\text{proS}$ ,  $^{13}\text{C}\gamma\text{proR}$ ,  $^1\text{H}_3\gamma\text{proR}$  (figure 3.3.10). Ce marquage isotopique a été réalisé en utilisant de l'alanine marquée  $\text{U-}[^2\text{H}, ^{13}\text{C}]-^{13}\text{C}^1\text{H}_3$  et les précurseurs 2-kéto-3-( $^{12}\text{C}^2\text{H}_3$ -méthyle)-3,4-( $^2\text{H}_2$ )-1,2,3,4,5( $^{13}\text{C}$ )-pentanoate (K-Ile) et 2-Keto-3-( $^2\text{H}$ )-3-( $^{12}\text{C}^2\text{H}_3$ -méthyle)-1,2,3,4- $^{13}\text{C}_2$  butyrate (KIV) marqués. Les 17 autres acides aminés ont été ajoutés sous la forme  $\text{U-}[^2\text{H}, ^{12}\text{C}, ^{14}\text{N}]$ . Cet échantillon (25 mL en mode dialyse) a été préparé dans un milieu réactionnel deutéré et concentré à 260  $\mu\text{M}$  puis transféré dans un tube shigemi de 4 mm. Une expérience méthyle-SOFAST a été acquise pour contrôler la qualité de l'échantillon (figure 3.3.11) suivie de l'acquisition des expériences RMN HCC. Notons que le marquage  $^{13}\text{C}$

linéarisé induit un couplage  $^1J_{CC}$  (35Hz) pour le carbone du groupement méthyle, d'où le dédoublement des résonances sur le spectre présenté en figure 3.3.11.

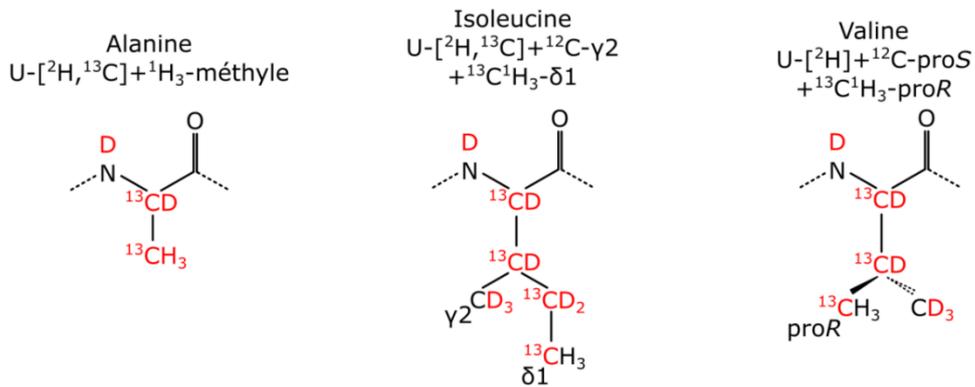


Figure 3.3.10 : Marquage isotopique réalisé sur les résidus alanines, isoleucines et valines pour l'acquisition des expériences RMN "out and back" de type HCC.

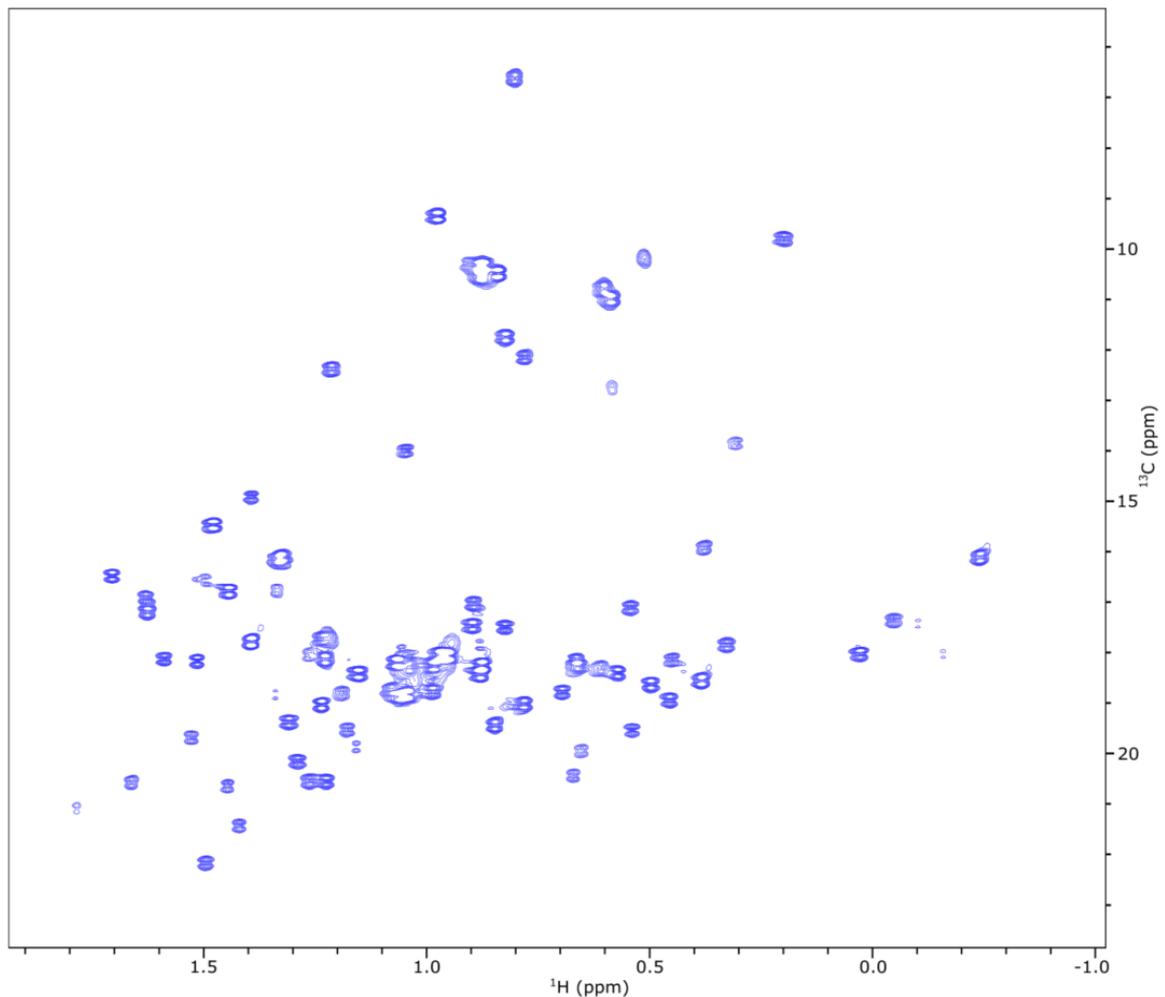


Figure 3.3.11 : Spectre RMN 2D méthyl SOFAST du fragment Fab perdeutééré et exprimé en système cell-free avec un marquage isotopique  $^{13}C$  linéaire pour les alanines, isoleucines- $\delta 1$  et valines-proR. Cette expérience RMN a été acquise à 35°C pour un échantillon de fragment Fab concentré à 260  $\mu M$  sur un spectromètre équipé d'une sonde cryogénique et opérant à une fréquence du proton de 950 MHz.

### 3.3.2. Acquisition des expériences RMN de type HCC ‘out and back’

Les expériences RMN de type HCC permettent de transférer l’aimantation des groupements méthyles à attribuer jusqu’aux atomes du squelette de la protéine (qui ont été attribués au préalable). Pour ce projet, il a été choisi de réaliser des expériences RMN HCC dites ‘out and back’ qui sont plus sensibles que les expériences de type HCC allant jusqu’aux groupements NH du squelette [122] (cf. partie 1.5.1). Les expériences ‘out and back’ sont donc plus adaptées pour les protéines de haut poids moléculaire.

Le nombre de liaisons atomiques qui séparent les six résidus portant un groupement méthyle des  $C\alpha$  et  $C\beta$  n’est pas le même. En effet, le groupement méthyle de l’alanine, par exemple, correspond à son  $C\beta$ . La fréquence  $^{13}C$  correspondant au groupement méthyle des alanines est donc directement éditée dans les expériences de type HN(CA)CB qui ont été utilisées pour l’attribution du squelette et sur le spectre 2D  $^1H$ - $^{13}C$  des groupements méthyles. En revanche, le groupement méthyle en position  $\delta 1$  des isoleucines se situe à deux liaisons atomiques de distance de leur  $C\beta$ . Ainsi, la magnétisation doit être transférée sur cette distance pour corrélérer les groupements méthyles des isoleucines à leur  $C\beta$  (et trois liaisons pour corrélérer le  $C\alpha$ ). Il est donc nécessaire d’acquérir plusieurs expériences RMN de type HCC pour pouvoir corrélérer les différents groupements méthyles aux  $C\alpha$  et  $C\beta$ .

Ainsi, trois expériences HCC ‘out and back’ différentes ont été nécessaires[122]. De la plus à la moins sensible, une expérience HCC simple a été acquise pour corrélérer les groupements méthyles des alanines à leur  $C\alpha$ , et ceux des valines à leur  $C\beta$  (figure 3.3.12a). Puis, une expérience HCC *relay* (ou HC(C)C) a été acquise pour corrélérer les groupements méthyles des valines à leur  $C\alpha$ , et le groupement méthyle  $\delta 1$  des isoleucines à leur  $C\beta$  (figure 3.3.12b). Enfin, une troisième expérience HCC double *relay* (ou HC(CC)C) a été acquise pour corrélérer les groupements méthyles en position  $\delta 1$  des isoleucines avec leur  $C\alpha$  (figure 3.3.12c).

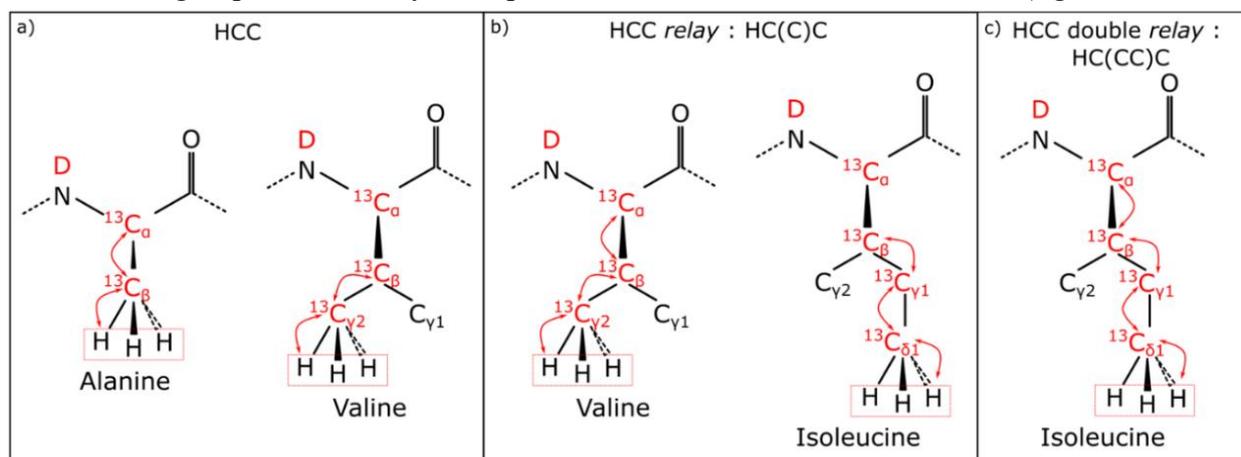


Figure 3.3.12 : Chemins empruntés par la magnétisation lors de son transfert pour les expériences RMN ‘out and back’ de type HCC a), HCC relay b) et HCC double relay c). Le chemin est donné pour la corrélation des groupements méthyles des résidus alanines, isoleucines et valines avec leur  $C\alpha$  et  $C\beta$ .

Les déplacements chimiques des résonances obtenues avec ces expériences ont ensuite été comparés à celles des expériences HNCA et HN(CA)CB pour être attribués car le C $\alpha$  d'un résidu a la même valeur de déplacement chimique dans la dimension  $^{13}\text{C}$  dans une expérience de type HCC et dans une HNCA (la logique est la même pour les C $\beta$ , sauf exception pour les alanines en fonction de leur deutération, cf. partie 3.3.3.4).

Il faut noter que l'attribution des groupements méthyles en utilisant les expériences HCC est possible uniquement si les résonances des C $\alpha$  et C $\beta$  du résidu ont été attribuées dans les expériences HNCA et HNCACB. Ainsi, toutes les résonances non attribuées pour le squelette ne peuvent l'être en utilisant les HCC. Le tableau 3.3.2 résume le nombre de signaux attribués pour les C $\alpha$  et C $\beta$  des alanines, isoleucines et valines lors de l'attribution du squelette du fragment Fab, et transférés aux groupements méthyles grâce aux expériences RMN HCC.

Tableau 3.3.2 : Nombre de résonances attribuées pour les C $\alpha$  et C $\beta$  des alanines, isoleucines et valines dans les expériences RMN 3D HNCA et HNCACB avec le nombre de corrélations jusqu'aux groupements méthyles obtenu avec les expériences RMN HCC et l'occurrence de ces acides aminés dans la séquence du fragment Fab de l'anticorps anti-LAMP1.

Acide aminé marqué	C $\alpha$ attribués dans la HNCA	Nombre de C $\alpha$ transférés aux méthyles grâce aux HCC	C $\beta$ attribués dans la HNCACB	Nombre de C $\beta$ transférés aux méthyles grâce aux HCC	Occurrence de l'acide aminé dans la séquence du fragment Fab de l'anticorps anti-LAMP1
Alanines	26	15 (58 % des C $\alpha$ )	26	15 (58 % des C $\beta$ )	28
Isoleucines	8	4 (50 % des C $\alpha$ )	9	4 (50 % des C $\beta$ )	12
Valines	34	16 (47 % des C $\alpha$ )	34	16 (47 % des C $\beta$ )	37

### 3.3.3. Exemple d'attribution pour les valines 132 et 133

Les valines 132 et 133 sont deux valines qui se succèdent dans la séquence de la chaîne légère du fragment Fab. Les résonances des C $\alpha$  et des C $\beta$  de ces deux valines ont été attribuées dans les expériences RMN 3D HNCA et HN(CA)CB. Cela a permis d'attribuer les résonances de leurs groupements méthyles respectifs grâce aux expériences RMN HCC et HC(C)C. La figure 3.3.13 représente des extraits 2D des expériences RMN 3D corrélant les C $\alpha$  de ces deux valines dans les expériences HNCA et HC(C)C et les C $\beta$  dans les expériences HN(CA)CB et HCC.

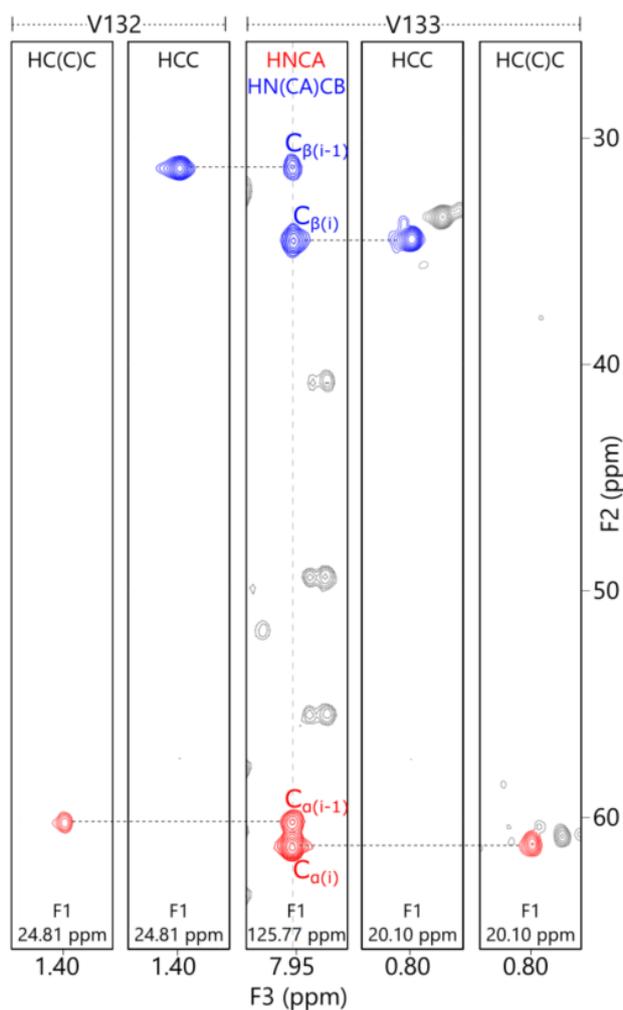


Figure 3.3.13 : Exemples d'extraits 2D des expériences RMN 3D "out and back" HCC et HC(C)C et 3D HNCA et HNCACB. L'exemple est donné pour la corrélation des C $\alpha$  (rouges) et C $\beta$  (bleus) des valines 132 et 133 dont les résonances ont été attribuées lors de l'attribution du squelette du fragment Fab (cf. partie 3.2.3). Les spectres ont été acquis à 35°C sur un échantillon de fragment Fab concentré à 300 $\mu$ M pour les HNCACB et HNCA, et à 330 $\mu$ M pour les HCC et HC(C)C. Le spectromètre RMN utilisé était équipé d'une sonde cryogénique et opérait à une fréquence du proton de 950 MHz pour toutes les expériences.

### 3.3.4. Attribution des C $\beta$ des alanines, cas particulier

Il a été expliqué précédemment que le groupement méthyle des alanines était également leur C $\beta$  et que la résonance correspondante était celle observée sur les spectres RMN 2D  $^1\text{H}$ - $^{13}\text{C}$  des groupements méthyles et sur les expériences 3D HN(CA)CB utilisées pour l'attribution du squelette. Cependant, la deutération du résidu est différente pour les échantillons utilisés dans les deux cas. En effet, les C $\beta$  des alanines sont deutérés lors de l'acquisition des expériences 3D HN(CA)CB et sont protonés lors de l'acquisition des 2D  $^1\text{H}$ - $^{13}\text{C}$  et 3D HCC pour les groupements méthyles. Cette différence de deutération entraîne une variation du déplacement chimique du carbone du groupement méthyle dans les deux expériences. En effet, le déplacement chimique isotopique n'est pas le même pour un carbone protoné ou deutéré. Ainsi, une différence de 0.8 ppm est observée dans la dimension carbone des spectres entre un carbone 100% protoné et un 100 % deutéré. Pour les isoleucines et les valines, cette différence de déplacement chimique entre les expériences HCC et les HNCA et HN(CA)CB existe mais est négligeable.

### 3.3.5. Attribution des groupements méthyles des alanines, isoleucines et valines en utilisant les HCC.

L'acquisition des expériences RMN HCC n'a pas permis d'observer le nombre exact de résonances pour les C $\alpha$  et C $\beta$  correspondant au nombre d'alanines, isoleucines et valines dans la séquence du fragment Fab (tableau 3.3.3).

Tableau 3.3.3 : Nombre de résonances observées dans les expériences RMN HCC "out and back" correspondant à des C $\alpha$  et des C $\beta$  et l'occurrence de ces acides aminés dans la séquence du fragment Fab de l'anticorps anti-LAMP1.

Acide aminé marqué	Nombre de résonances pour les C $\alpha$ dans les expériences HCC	Nombre de résonances pour les C $\beta$ dans les expériences HCC	Occurrence de l'acide aminé dans la séquence du fragment Fab de l'anticorps anti-LAMP1 $\beta$	Pourcentage d'attribution
Alanine	30	30	28	53 %
Isoleucine	8	12	12	33 %
Valine	34	34	37	43 %

L'attribution des résonances observées a été effectuée comme détaillé précédemment et de manière systématique. Seules les corrélations non ambiguës ont été attribuées. Les corrélations ambiguës, c'est-à-dire celles avec un manque d'information (manque de signal pour le C $\alpha$  et/ou le C $\beta$  dans les HNCA, HNCACB ou HCC) et/ou une superposition de signal, n'ont pas été utilisées pour l'attribution. Cependant, ces données ont été conservées dans l'optique de pouvoir lever certaines ambiguïtés d'attribution grâce aux expériences RMN NOESY inter-méthyles.

De cette manière, 53%, 43% et 33% des groupements méthyles des alanines (C $\beta$ ), valines (proR) et isoleucines ( $\delta$ 1) ont été attribués. Pour plus de visibilité (résonances non dédoublées), l'attribution des groupements méthyles en utilisant les expériences RMN HCC a été reportée sur les spectres RMN 2D  $^{13}\text{CH}_3$  SOFAST du fragment Fab marqué  $^{13}\text{CH}_3$  sur les groupements méthyles de chaque type de résidu (figure 3.3.14) (cf. partie 3.1.2.3).

L'attribution des groupements méthyles en utilisant des expériences RMN de type HCC est loin d'être complète et n'a permis d'attribuer que 45% des alanines, isoleucines et valines marquées. De plus, il reste les groupements méthyles des leucines, méthionines et thréonines à attribuer. Cela signifie que parmi les 150 groupements méthyles à attribuer pour le fragment Fab (en ne comptant qu'un seul groupement méthyle pour les résidus Ile, Leu et Val), seuls 23% ont été attribués en utilisant cette méthode. Ainsi, pour compléter cette attribution, un autre échantillon de fragment Fab a été produit en système *cell-free* avec un marquage isotopique permettant l'acquisition d'expériences RMN de type NOESY  $^{13}\text{CH}_3$ - $^{13}\text{CH}_3$ .

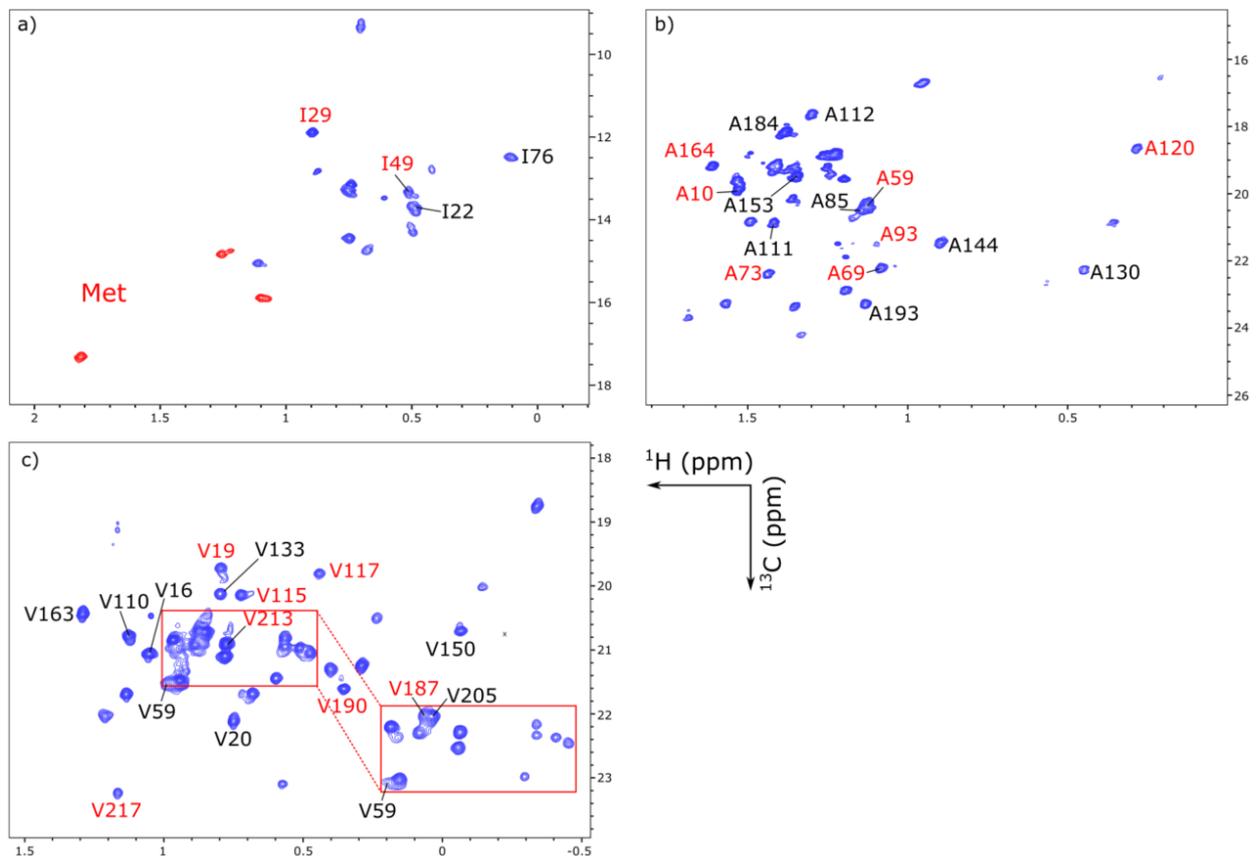


Figure 3.3.14 : Groupements méthyles des alanines, isoleucines et valines attribués en utilisant les expériences RMN HCC "out and back". L'attribution est représentée sur les spectres RMN 2D methyl-SOFAST du fragment Fab marqué sur les groupements méthyles et exprimé en système *cell-free*. L'attribution des résidus de la chaîne légère est en noir, et celle des résidus de la chaîne lourde est en rouge. Les spectres correspondent au fragment Fab marqué a)  $^{13}\text{CH}_3$ -Met,  $^{13}\text{CH}_3$ -Ile- $\delta$ 1 (Les résonances appartenant à des méthionines sont indiquées en rouge), b)  $^{13}\text{CH}_3$ -Ala, c)  $^{13}\text{CH}_3$ -Val-proR. Les spectres ont été acquis sur un spectromètre RMN opérant à une fréquence du proton de 950 MHz équipé d'une sonde cryogénique. Les acquisitions ont duré 30 min pour des échantillons concentrés de 20 à 70  $\mu\text{M}$  en fragment Fab dans des tubes Shigemi de 4 mm.

### 3.4. Attribution des groupements méthyles en utilisant des expériences RMN méthyles NOESY et la structure du fragment Fab

Afin de compléter l'attribution réalisée par connexion des  $C\alpha$  et des  $C\beta$  aux groupements méthyles en utilisant des expériences HCC, un nouvel échantillon de fragment Fab marqué sur les groupements méthyles des résidus alanines, leucines-proS, isoleucines- $\delta 1$ , méthionines, thréonines et valines-proR a été produit. Cet échantillon a servi à acquérir des expériences RMN 3D NOESY inter-méthyles. Ce réseau de NOE combiné à la structure connue de la protéine permet d'attribuer les méthionines, thréonines et leucines qui n'ont pas pu être marquées de manière linéaire en  $^{13}C$  pour les expériences HCC, ainsi que les alanines, isoleucines et valines qui n'ont pas pu être attribuées précédemment.

#### 3.4.1. Production du fragment Fab perdeutééré en système *cell-free* avec un marquage isotopique $^{13}CH_3$ stéréo et régio spécifique sur les groupements méthyles

Les expériences RMN basées sur l'effet NOE sont fortement impactées par les interactions dipolaires entre les protons de la protéine. Ainsi, pour détecter un réseau de NOEs le plus complet possible pour le fragment Fab de l'anticorps, son expression en système *cell-free* a été réalisée en milieu réactionnel 100% deutéré avec des acides aminés perdeutéérés. L'échantillon a été produit à partir de 25 mL de réaction *cell-free* en mode dialyse, puis purifié et concentré jusqu'à 330  $\mu M$  (tube Shigemi de 4 mm).

Les acides aminés marqués sur les groupements méthyles étaient également perdeutéérés avec une protonation spécifique sur leur groupement méthyle marqué  $^{13}C$  (figure 3.3.15).

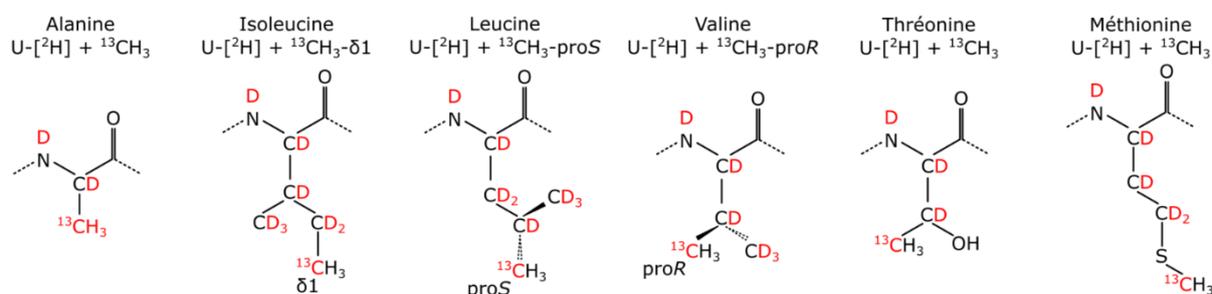


Figure 3.3.15 : Marquage isotopique réalisé sur les acides aminés comportant un groupement méthyle pour l'acquisition des expériences RMN méthyles NOESY. Les  $^{13}C$  et  $^2D$  sont représentés en rouge.

Comme précédemment, les isoleucines et valines ont été marquées en utilisant leur précurseur respectif marqué isotopiquement, c'est-à-dire le 2-keto-3-( $^{13}\text{C}$  methyl)-4-( $^2\text{H}_2$ )-5-( $^2\text{H}_3$ methyl)- 1,2,3,4- $^{13}\text{C}_4$  pentanoate (K-Ile) et le 2-Keto-3-( $^2\text{H}_3$  methyl)-1,2,3,4- $^{13}\text{C}_4$  butyrate (KIV). Les alanines, méthionines, thréonines et leucines ont été marquées en utilisant directement les acides aminés marqués.

Pour contrôler la qualité de l'échantillon de fragment Fab marqué  $^{13}\text{CH}_3$  sur les six acides aminés comportant un groupement méthyle, une expérience RMN 2D méthyle SOFAST a été acquise (figure 3.3.16).

A la suite de l'acquisition de ce spectre, une expérience RMN 3D HMQC-NOESY-HMQC a été acquise pour obtenir le réseau de NOE inter-méthyles nécessaire à l'attribution.

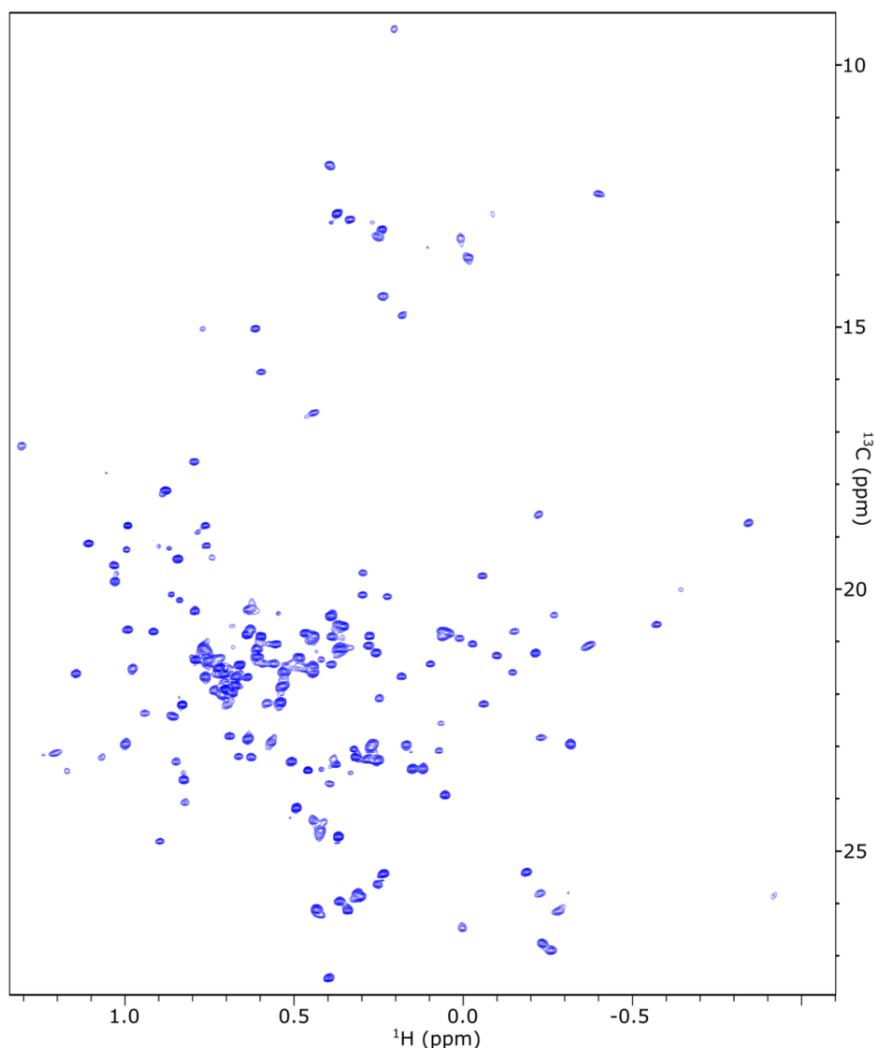


Figure 3.3.16 : Spectre RMN 2D methyl SOFAST du fragment Fab de l'anticorps anti-LAMP1 produit en système cell-free perdeutééré. Le fragment Fab est U-[ $^2\text{H}$ ],  $^{13}\text{CH}_3$ -Ala, Met, Thr, Leu-proS, Val-proR, Ile- $\delta 1$ . L'acquisition a été réalisée à 35°C pendant 30 min sur un échantillon concentré à 330  $\mu\text{M}$  de fragment Fab, placé dans un tube RMN Shigemi de 4 mm. Le spectromètre RMN utilisé est équipé d'une sonde cryogénique et opère à une fréquence  $^1\text{H}$  de 950 MHz.

### **3.4.2. Attribution du type de résidu pour les résonances du spectre méthyle SOFAST**

Une étape essentielle à l'attribution des groupements méthyles en utilisant les expériences RMN méthyles NOESY corrélées à la structure connue de la protéine est l'identification du type de résidu pour les résonances observées sur les spectres. Cela a été réalisé en superposant les spectres 2D méthyles du fragment Fab marqué  $^{13}\text{CH}_3$  sélectivement sur un groupement méthyle (spectres présentés en figure 3.3.8) et marqué sur les six acides aminés à la fois. Cette superposition de spectres nous a permis d'obtenir une information sur l'appartenance de chaque résonance du spectre RMN des groupements méthyles. Cette identification est représentée par le code couleur en figure 3.3.17. Les résonances des groupements méthyles proS des leucines ont été identifiées par déduction logique après avoir identifié les résonances des cinq autres résidus.

L'identification du type de résidu pour les résonances du spectre RMN 2D  $^1\text{H}$ - $^{13}\text{C}$  indique que des résonances supplémentaires sont observées (signaux en noir dans la figure 3.3.17). Ces pics supplémentaires pourraient être le résultat d'une dégradation ou d'une forme minoritaire du fragment Fab présente dans le tube RMN, mais n'ont pas impacté l'attribution car ils ne présentent pas de signaux de corrélations NOEs.

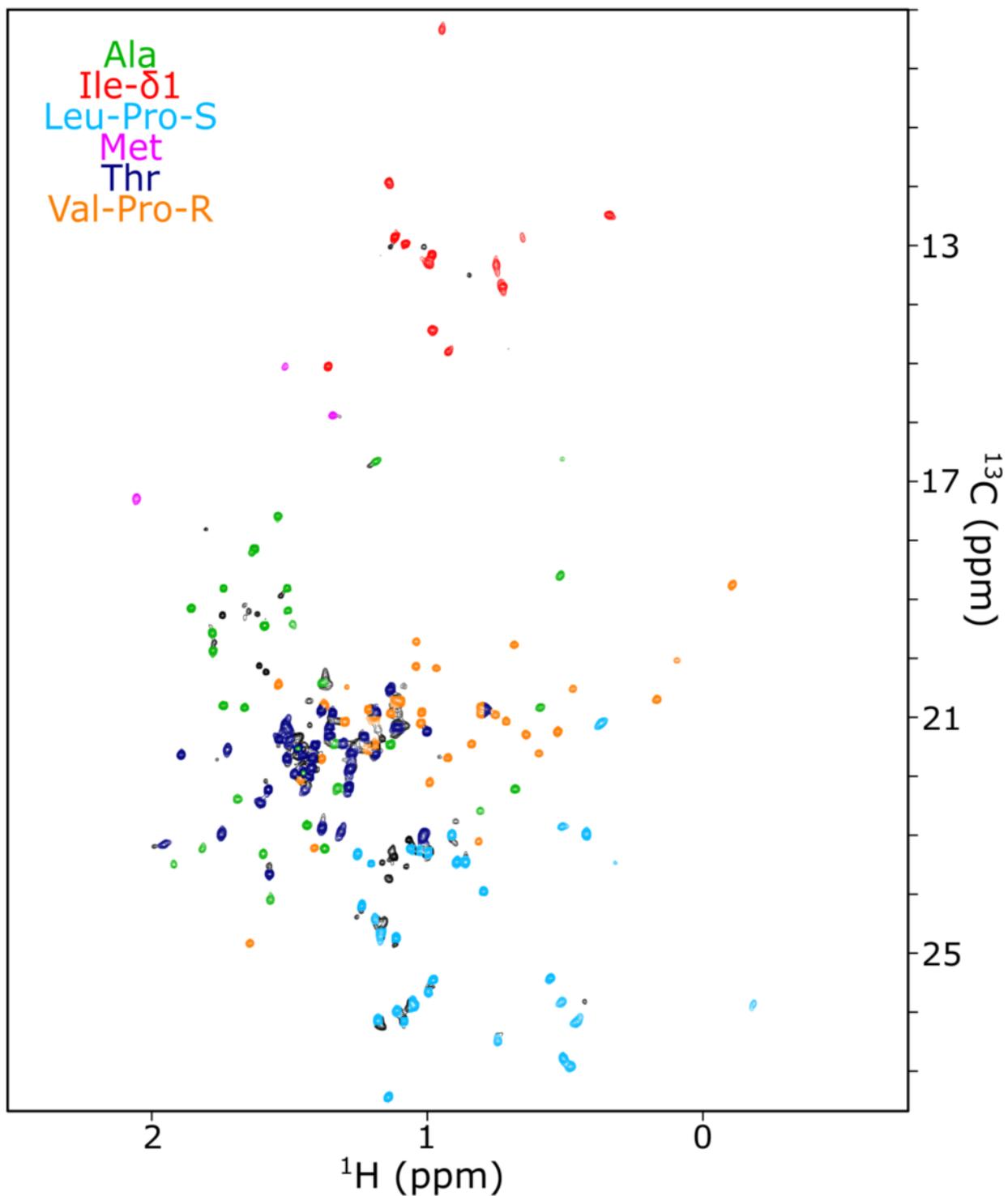
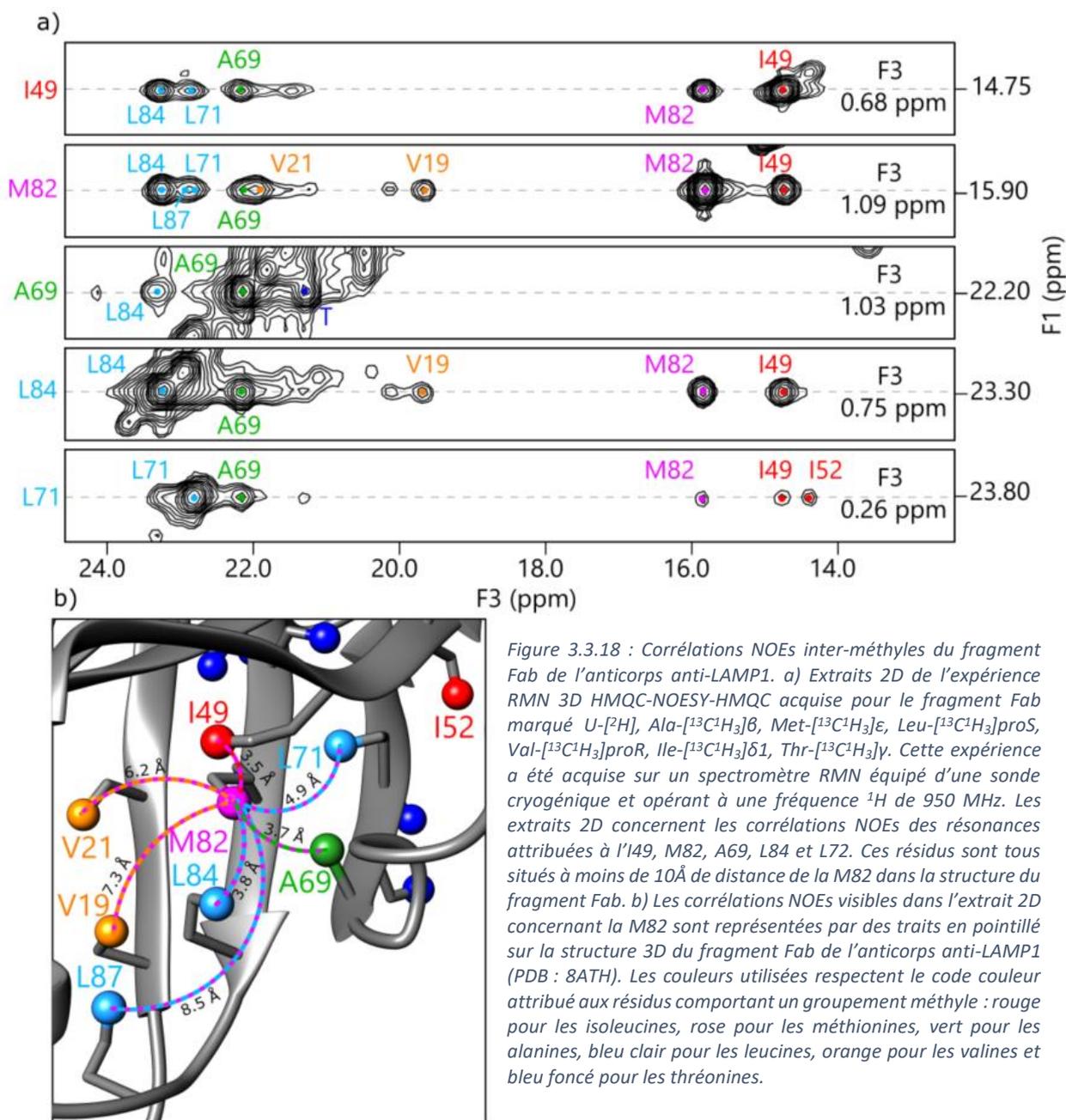


Figure 3.3.17 : Attribution du type de résidu pour les résonances observées sur le spectre RMN 2D methyl SOFAST du fragment Fab marqué U- $^{2}\text{H}$ ,  $^{13}\text{CH}_3$ -Ala, Met, Thr, Leu-proS, Val-proR, Ile- $\delta$ 1. Les résonances des groupements méthyles correspondant à des alanines en vert, des isoleucines ( $\delta$ 1) en rouge, des méthionines en rose, des valines (ProR) en orange, des thréonines en bleu marine, des leucine (ProS) en bleu clair.

### 3.4.3. Attribution des groupements méthyles du fragment Fab grâce aux corrélations NOE inter-méthyles

L'acquisition de l'expérience RMN 3D HMQC-NOESY-HMQC sur le fragment Fab marqué U- $^2\text{H}$ , Ala- $^{13}\text{C}^1\text{H}_3$ ] $\beta$ , Met- $^{13}\text{C}^1\text{H}_3$ ] $\epsilon$ , Leu- $^{13}\text{C}^1\text{H}_3$ ] $\text{proS}$ , Val- $^{13}\text{C}^1\text{H}_3$ ] $\text{proR}$ , Ile- $^{13}\text{C}^1\text{H}_3$ ] $\delta 1$ , Thr- $^{13}\text{C}^1\text{H}_3$ ] $\gamma$  a permis d'identifier 434 pics de corrélations NOE inter-méthyles avec un rapport signal sur bruit  $S/N \geq 3$ . La figure 3.3.18 présente un exemple de corrélations NOE inter-méthyles et l'attribution proposée pour les résonances observées.



Pour attribuer les groupements méthyles grâce aux corrélations NOEs interméthyles obtenues, le programme MethylFLYA a été utilisé [149]. Afin d'augmenter la fiabilité de l'attribution réalisée par le programme, les alanines, isoleucines et valines ayant été attribuées de manière non ambiguë en utilisant les expériences RMN HCC lui ont été renseignées (35 au total, cf. partie...). L'utilisation de ce programme a permis d'attribuer de nouvelles résonances

de groupements méthyles, ainsi que de lever des ambiguïtés d'attribution obtenues grâce aux expériences RMN HCC.

Ce programme d'attribution automatique des groupements méthyles a été développé pour l'attribution des résidus ILV avec un marquage isotopique non stéréospécifique. Le fragment Fab produit pour l'attribution des groupements méthyles en utilisant les NOE a été marqué les six résidus portant un groupement méthyle dont les groupements proR des valines, proS des leucines et  $\delta 1$  des isoleucines. De plus, le programme de MethylFLYA a été basé sur l'utilisation d'expériences RMN 4D à faible temps de mélange pour l'identification des groupements méthyles proS et proR des leucines et valines [149]. Or, cela a été réalisé grâce au marquage isotopique stéréospécifique des groupements méthyles du fragment Fab. Ainsi, les données fournies au programme pour réaliser l'attribution présentaient quelques différences et notamment au niveau du marquage isotopique réalisé.

Afin de faire corréler le programme de MethylFLYA avec le marquage isotopique réalisé et avec les données en notre possession, une collaboration a été mise en place avec le professeur Peter Gunter (ETH Zurich), responsable de l'équipe ayant développé MethylFLYA. Cependant, le programme ainsi modifié n'avait jamais été utilisé pour l'attribution des groupements méthyles des alanines, méthionines et thréonines. Ainsi, une vérification manuelle systématique a été réalisée pour les attributions fournies par MethylFLYA.

Puis, une attribution manuelle des groupements méthyles a été réalisée en prenant en compte les données fournies par les expériences HCC, et les NOEs restantes non attribuées. De cette manière, l'attribution des groupements méthyles a été réalisée pour 100% des alanines, des méthionines, des isoleucines- $\delta 1$  et des valines-proR. Sur les 30 leucines du fragment Fab, seule la leucine 95 n'a pas pu être attribuée à une résonance du spectre RMN 2D  $^1\text{H}$ - $^{13}\text{C}$  en raison d'une superposition des pics supplémentaires observés dans la région des leucines. Les thréonines ont le pourcentage d'attribution le plus faible parmi les six types de résidus comportant un groupement méthyle. Seuls 40% de leurs groupements méthyles ont pu être attribués. Cela s'explique par le fait que les thréonines sont principalement situées en surface du fragment Fab et sont donc éloignées du reste des autres groupements méthyles du cœur hydrophobe (à plus de 10Å de distance) (figure 3.3.19). Ainsi, la plupart des thréonines ne possèdent pas de pics de corrélation NOE inter-méthyles autres qu'avec d'autres thréonines non attribuées. De plus, le fait qu'elles soient toutes isolées à la surface du fragment implique que leur déplacement chimique est presque identique, ce qui est visible sur le spectre 2D  $^1\text{H}$ - $^{13}\text{C}$  des groupements méthyles ou les résonances des

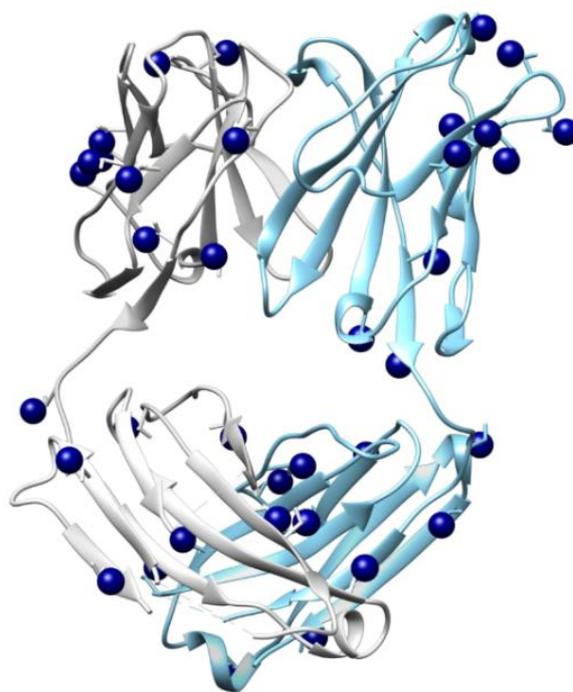
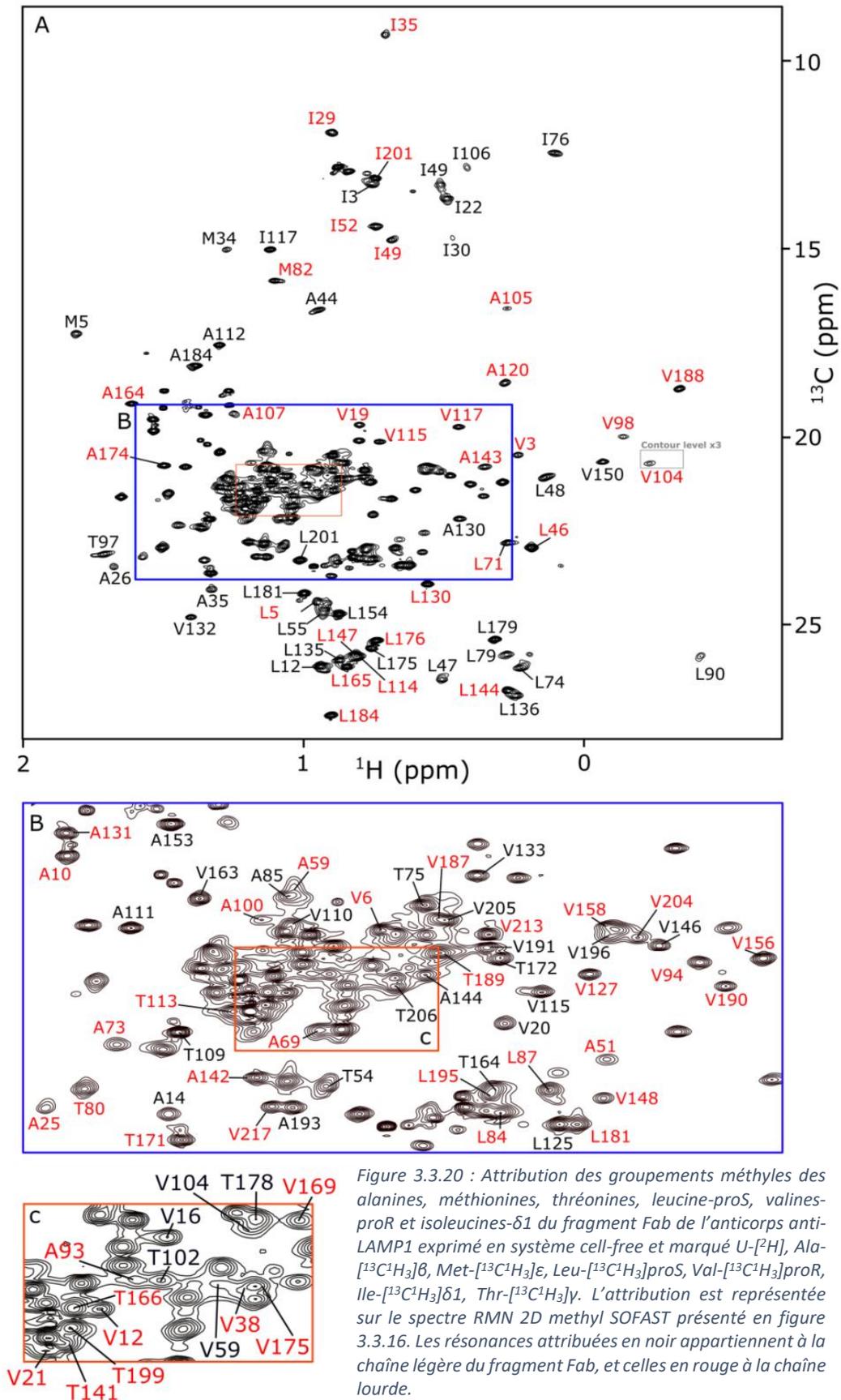


Figure 3.3.19 : Représentation des groupements méthyles des thréonines (sphères bleu foncé) sur la structure du fragment Fab de l'anticorps anti-LAMP1. (PDB : 8ATH). La chaîne légère est représentée en gris clair et la chaîne lourde en bleu clair.

thréonines sont superposées (figure 3.3.17). L'attribution finale des groupements méthyles est présentée sur le spectre 2D méthyle SOFAST en figure 3.3.20.



En combinant les données fournies par les expériences RMN HCC et les NOESY inter-méthyles, 100 % des groupements méthyles des alanines, méthionines, isoleucines- $\delta 1$  et valines *proR* ont été attribués. Un groupement *proS* de leucine n'a pas pu être attribué en raison d'une superposition de résonances. Pour les thréonines dont les groupements méthyles ont été attribués à 40% seulement, une méthode complémentaire devra être utilisée. Etant situées à la surface du fragment Fab, il n'est pas possible de les attribuer en utilisant des NOE inter-méthyles. Les  $C\alpha$  et  $C\beta$  des thréonines ont cependant tous été attribués sur les expériences HNCA et HN(CA)CB utilisées pour l'attribution du squelette. Une solution serait d'utiliser un marquage linéaire en  $^{13}C$  sur la chaîne latérale des thréonines et des expériences RMN HCC. Le fragment Fab a été marqué avec de la thréonine marquée U- $^{13}C$  disponible commercialement. Ce type de marquage isotopique permet en théorie d'acquérir des expériences RMN HCC et de transférer la magnétisation du groupement méthyle aux  $C\alpha$  et  $C\beta$ . Cependant, cela n'a pas permis d'acquérir des spectres exploitables car la thréonine marquée U- $^{13}C$  n'est pas deutérée. Pour aller plus loin avec des expériences RMN HCC, il faudrait utiliser de la thréonine U- $^{2}H,^{13}C$  +  $^{13}C^1H_3$  qui n'est pas disponible commercialement.

## **IV. Conclusion**

### **1. Conclusion et discussion**

Ce projet de thèse s'est inscrit dans le cadre d'une collaboration entre Sanofi, l'IBS et NMR-Bio afin de développer une technique innovante de marquage isotopique pour mieux caractériser, par RMN, les anticorps monoclonaux exprimés en cellules CHO. La place des anticorps thérapeutiques sur le marché mondial des médicaments est en augmentation croissante depuis leur apparition en 1986. Ces biothérapeutiques permettent d'apporter des solutions thérapeutiques à de nombreuses pathologies mortelles comme les cancers et les maladies chroniques. Pour les agences de sécurité du médicament telles que la FDA et l'EMA, il est d'une importance capitale de pouvoir relier l'activité de ces anticorps à leur structure, à l'échelle atomique et en solution. Pour cela, les nombreuses techniques analytiques actuellement utilisées dans les entreprises pharmaceutiques ne suffisent pas en raison d'un manque de résolution. La RMN des groupements méthyles peut être la solution à cette problématique. Cependant, en absence d'attribution des spectres RMN, cette méthode est limitée à l'utilisation d'outils statistiques pour la compréhension des modifications observés sur les spectres à la suite de pertes d'activité des mAbs.

Mes travaux de thèse ont visé à développer des techniques de marquages isotopiques optimisé pour la RMN afin de permettre l'étude des mAbs et de leur fragments à résolution atomique en solution.

#### **1.1. Marquage isotopique spécifique des groupements méthyles des mAbs produit en cellules CHO**

La première étape a consisté à développer une méthode de marquage isotopique des groupements méthyles pour l'anticorps anti-LAMP1 exprimé en cellules CHO. L'optimisation de ce mélange d'acides aminés, en combinaison avec une méthode de culture à haute densité cellulaire ( $10^7$  cellules/mL), a permis d'atteindre les rendements obtenus avec les milieux commerciaux riches. L'incorporation des acides aminés marqués dans l'anticorps a été optimisée en mettant en place un protocole pour affamer les cellules, ce qui a permis d'augmenter le taux d'incorporation des acides aminés marqués et d'atteindre des valeurs supérieures à 84 % pour les alanines et isoleucines et supérieures à 90 % pour les méthionines, leucines, thréonines et valines. Pour l'alanine, qui constitue un acide aminé non-essentiel pour les cellules de mammifère, ce taux d'incorporation a été atteint en inhibant la voie de biosynthèse de cet acide aminé par les cellules par ajout de 10 mg/L de L-cyclosérine pendant la culture.

Le développement de cette méthode de marquage isotopique en cellules CHO a permis d'obtenir, pour la première fois dans la communauté scientifique, des spectres RMN  $^1\text{H}$ - $^{13}\text{C}$  des groupements méthyles d'un anticorps à haute résolution et, ainsi, d'identifier les signaux appartenant aux six acides aminés comportant des groupements méthyles.

Cette méthode, développée sur l'anticorps anti-LAMP1, peut être étendue à toutes les protéines exprimées en cellules CHO, système de choix pour la production de protéines

thérapeutiques. Pour cela, le mélange d'acides aminés développé devra être optimisé en adaptant la proportion des acides aminés par rapport à la séquence de la protéine étudiée.

## 1.2. Production et marquage isotopique des Fabs *in vitro*

La deutération d'une protéine de la taille d'un anticorps est essentielle pour attribuer ses signaux RMN. Cependant, les cellules CHO ne sont pas adaptées pour la deutération des protéines. Le système *cell-free* a donc été utilisé pour produire des échantillons marqués de façon à pouvoir attribuer les signaux RMN du squelette puis des groupements méthyles. La stratégie a également été de diviser l'anticorps anti-LAMP1 en exprimant ses fragments Fab et Fc à la place d'exprimer l'anticorps entier de façon à améliorer la qualité des spectres RMN.

Au préalable, les conditions d'expression ont été optimisées afin d'augmenter les rendements et de diminuer les coûts de production des fragments marqués. Par exemple, pour l'expression du fragment Fab, l'ajout de bétaïne à 15 mM a permis d'augmenter le rendement d'un facteur deux. Ce résultat est très différent de ceux obtenus par Kai *et al.* qui ont étudié l'utilisation de trois polyions, dont la bétaïne, en système *cell-free* pour améliorer le rendement [195]. Ils ont déterminé que l'ajout de bétaïne n'augmentait pas le rendement de production de la sGFP. Ainsi, l'amélioration du rendement d'expression grâce à l'ajout d'un polyion est un paramètre protéine-dépendant, à optimiser pour chaque protéine exprimée en système *cell-free*.

Le rendement d'expression a également été augmenté avec l'ajout de DsbC recombinante à 14  $\mu\text{M}$  dans un extrait bactérien commun en remplacement de l'extrait issu de la souche SHuffle de *E. coli*. Cette augmentation du rendement d'expression rejoint les résultats de Dopp *et al.*, mais la concentration utilisée est plus élevée (4.3  $\mu\text{M}$  pour Dopp *et al.*) [196]. Il semble en effet que la concentration en protéine chaperonne utilisée lors de la réaction soit protéine dépendant. D'autres protéines chaperonnes pour la formation et l'isomérisation des protéines existent et leur utilisation a été décrite pour l'augmentation des rendements d'expression d'immunoglobulines ou de fragments (Fab, scFv) en système *cell-free* (cf. partie 3.2.1.2) [177], [178], [197]. Pour les futures expressions de protéines à ponts disulfures en système *cell-free*, il serait donc intéressant d'étudier leur effet (seules ou en combinaison) afin d'augmenter encore les rendements.

Le volume de réactifs utilisé pour le tampon de dialyse du système *cell-free* a également pu être diminué par deux sans impacter le rendement d'expression. Cela correspond à une augmentation du rendement par un facteur deux par rapport à la quantité de réactifs utilisés, et donc une diminution d'un facteur deux du prix de production de protéine avec des acides aminés marqués.

Enfin, la méthode de purification du fragment Fab a été optimisée en mettant en place l'utilisation d'une colonne de protéine L spécifique aux fragments Fab d'anticorps.

L'ensemble de ces optimisations a permis d'atteindre un rendement de 400  $\mu\text{g/mL}$  (pour les échantillons produits en mode dialyse), ce qui rejoint les rendements d'expression obtenus dans la littérature pour des fragments Fab exprimés en système *cell-free* bactérien. Cela représente une augmentation d'un facteur 10 par rapport au rendement de production obtenu au

début du projet et donc une diminution d'un facteur 10 des coûts de production pour la préparation des échantillons marqués.

### 1.3. Attribution des groupements méthyles du Fab anti-LAMP1

Une fois le protocole d'expression optimisé, l'attribution du squelette du fragment Fab a pu être entreprise grâce à la production du fragment marqué U-[<sup>2</sup>H, <sup>13</sup>C, <sup>15</sup>N]. La qualité de l'échantillon produit a permis d'acquérir les expériences RMN 3D nécessaires à l'attribution du squelette. Ce travail m'a permis de proposer la première attribution du spectre RMN <sup>1</sup>H-<sup>15</sup>N du squelette d'un fragment Fab entier.

L'attribution du spectre RMN <sup>1</sup>H-<sup>15</sup>N du fragment Fab de l'anticorps anti-LAMP1 a ensuite été utilisée pour observer l'interaction du Fab anti-LAMP1 avec son antigène LAMP1. Des modifications de déplacements chimiques ont été observées pour des résidus situés dans les domaines CDR du fragment Fab. Cependant, en raison du nombre élevé de signaux RMN sur le spectre <sup>1</sup>H-<sup>15</sup>N d'un fragment Fab d'anticorps (environ 450), beaucoup de signaux sont superposés, ce qui a empêché le transfert d'attribution du squelette du spectre du fragment Fab seul au spectre du fragment Fab en complexe avec LAMP1. Ainsi, il s'est avéré impossible de suivre l'interaction du fragment Fab (et encore moins celle de l'anticorps) avec son antigène par RMN des groupements amides du squelette. Pour cela, la RMN des groupements méthyles peut permettre de décomplexifier les spectres tout en fournissant une empreinte globale du fragment Fab.

L'attribution des groupements méthyles du fragment Fab de l'anticorps anti-LAMP1 a ensuite été réalisée en utilisant deux méthodes d'attribution complémentaires : l'utilisation d'expériences RMN de type HCC 'out and back' pour transférer l'attribution du squelette réalisée au préalable sur les groupements méthyles des alanines, isoleucines ( $\delta 1$ ) et valines (*proR*), et l'utilisation d'un réseau de NOEs inter-méthyles corrélé à la structure connue du fragment Fab de l'anticorps anti-LAMP1. Grâce à cela, 100 % des groupements méthyles des alanines, méthionines, isoleucines- $\delta 1$  et valines-*proR* ont été attribués. Tous les groupements méthyles des leucines *proS* ont été attribués, à l'exception de celui de la leucine 95 de la chaîne légère en raison d'une superposition de signaux. Le groupement amide de cette leucine n'a pas non plus été attribué lors de l'attribution du squelette de la protéine. Cette leucine est située sur le CDRL3 du fragment Fab, et donc, sur une partie très flexible de la protéine.

Les groupements méthyles des thréonines ont été attribués à 40 %. En effet, la thréonine U-[<sup>2</sup>H, <sup>13</sup>C] + <sup>13</sup>CH<sub>3</sub> n'existe pas commercialement, empêchant d'utiliser la méthode des expériences RMN HCC pour attribuer leur groupements méthyles pour une protéine de 50 kDa. De plus, les thréonines étant principalement situées à la surface du fragment Fab, leurs groupements méthyles ne présentent pas de NOEs et possèdent un déplacement chimique presque identique, pouvant expliquer la superposition des signaux RMN sur le spectre <sup>1</sup>H-<sup>13</sup>C. L'intégralité des  $C\alpha$  et  $C\beta$  des thréonines ont été attribués lors de l'attribution du squelette du fragment Fab. Pour compléter l'attribution, il serait nécessaire de réaliser un marquage isotopique du fragment Fab avec de la thréonine U-[<sup>2</sup>H, <sup>13</sup>C] + <sup>13</sup>CH<sub>3</sub> pour permettre l'utilisation d'expériences RMN HCC, ou alors, utiliser la méthode de la mutagénèse dirigée.

Enfin, l'attribution des résonances RMN des groupements méthyles du fragment Fab exprimé en système *cell-free* a été transférée sur les spectres RMN <sup>1</sup>H-<sup>13</sup>C de l'anticorps entier



## 2. Perspectives

### 2.1. Application des méthodes développées à l'étude du fragment Fc

D'une manière similaire à la stratégie développée pour le fragment Fab, consistant à utiliser le système *cell-free* pour l'expression et le marquage isotopique du fragment en vue d'attribuer ses groupements méthyles, puis de transférer cette attribution sur les spectres de l'anticorps produit et marqué en cellules CHO, une approche semblable pourrait être envisagée pour le fragment Fc. Par faute de temps, l'attribution des groupements méthyles du fragment Fc de l'anticorps anti-LAMP1 n'a pas été réalisée pendant la thèse. Des études de faisabilité pour l'expression du fragment Fc en système *cell-free* ont néanmoins été réalisées par une étudiante en stage de Master 1 en biologie, sous ma supervision. Les résultats obtenus ont permis de mettre au point le protocole de production du fragment Fc en système *cell-free*, avec des rendements de 0.4 mg/mL de réaction.

Le squelette carboné du fragment Fc d'une IgG1 a déjà été attribué par Yagi *et al.* et la séquence attribuée est identique à celle de l'anticorps anti-LAMP1[198]. Par ailleurs, le fragment Fc est un homodimère constitué de deux fois 25 kDa. Cela implique des spectres RMN deux fois moins chargés que ceux du fragment Fab. De plus, à concentration égale avec le fragment Fab, les signaux seront deux fois supérieurs pour le fragment Fc en raison de la symétrie de celui-ci. L'ensemble de ces éléments constitue un avantage considérable qui facilitera l'attribution des groupements méthyles du fragment Fc.

En revanche, la présence de glycosylations sur les asparagines 297 des domaines CH2 des chaînes lourdes du fragment Fc pourrait entraîner des difficultés. Yagi *et al.* ont attribué le squelette d'un fragment Fc d'IgG1 glycosylée[198], ce qui signifie qu'utiliser la méthode des expériences RMN de type HCC pour le fragment Fc exprimé en système *cell-free* (absence de glycosylation) n'est pas possible pour le domaine CH2 du fragment, dont la structure est fortement impactée par les glycosylations. En effet, les déplacements chimiques des C $\alpha$  et C $\beta$  du squelette carboné du domaine CH2 seraient trop différents pour des fragments Fc glycosylés et non glycosylés. Dans ces conditions, la méthode est applicable uniquement pour le domaine CH3, dont la structure n'est pas impactée par la présence de glycosylations.

Le profil de glycosylation du fragment Fc dont le squelette a été attribué n'est pas connu. Ainsi, dans un premier temps, il serait nécessaire d'attribuer le squelette carboné du fragment Fc déglycosylé. Puis, la méthode des expériences RMN HCC pourrait être appliquée pour initier l'attribution des groupements méthyles du domaine CH2. Cependant, cela resterait insuffisant pour transférer cette attribution sur le spectre du fragment Fc de l'IgG1 entière exprimée en cellules CHO car celle-ci différerait en raison des glycosylations présentes. Deux approches pourraient alors être envisagées pour surmonter cette difficulté :

- Déglycosyler l'anticorps produit en cellules CHO pour permettre le transfert d'attribution en utilisant les HCC pour l'attribution des groupements méthyles. Il serait ensuite nécessaire d'étudier l'impact des différents profils de glycosylation sur les signaux RMN des groupements méthyles du domaine CH2 afin de pouvoir transférer l'attribution selon les profils de glycosylation du fragment Fc.
- Développer une méthode permettant de produire un fragment Fc d'anticorps en système *cell-free* avec un profil de glycosylation contrôlé. Une fois la méthode développée, il serait nécessaire d'attribuer le spectre RMN du fragment Fc avec un

profil de glycosylation identique à celui de l'anticorps exprimé en cellules CHO. Cette méthode permettrait directement d'étudier l'impact des différents profils de glycosylation sur les spectres RMN  $^1\text{H}$ - $^{13}\text{C}$  des groupements méthyles.

La deuxième approche a été envisagée dans le cadre d'une nouvelle collaboration entre Sanofi et l'IBS pour un sujet de thèse CIFRE qui commencera fin 2023.

## 2.2. Etude de la HOS d'un anticorps par RMN

En attribuant indépendamment les fragments Fab et Fc puis en les transférant à l'anticorps entier, il sera possible d'étudier l'anticorps anti-LAMP1 produit dans les cellules CHO, système d'expression utilisé en routine pour la production industrielle, et de caractériser sa HOS ainsi que d'expliquer l'origine structurale d'une potentielle perte d'activité.

Il est important de noter que l'attribution qui a été réalisée pour les isoleucines, leucines et valines ne concerne qu'un seul de leurs deux groupements méthyles. Les spectres RMN de l'anticorps non marqué présenteront donc des signaux non attribués qui correspondront aux groupements méthyles  $\gamma_2$  des isoleucines, *proR* des leucines et *proS* des valines. Ainsi, pour le contrôle de la qualité des lots de production de l'anticorps anti-LAMP1, l'attribution de ces groupements méthyles devra être réalisée. Cela peut être réalisé en utilisant des acides aminés marqués  $^{13}\text{C}$  linéaires et  $^{13}\text{CH}_3$  stéréospécifiques sur les  $\gamma_2$  des isoleucines, *proR* des leucines et *proS* des valines puis en acquérant des expériences RMN de type HCC pour corréler ces groupements méthyles aux  $\text{C}\alpha$  et  $\text{C}\beta$  des résidus [126]. Une fois cette attribution complétée, il sera possible, grâce au spectre RMN  $^1\text{H}$ - $^{13}\text{C}$  des groupements méthyles de référence de l'anticorps anti-LAMP1 attribué, de caractériser tous les lots de ce même anticorps sortant des chaînes de production, en acquérant simplement un spectre RMN des méthyles en abondance naturelle et en le comparant au spectre de référence, et, ainsi, pouvoir suivre la HOS de ces anticorps non marqués pendant toutes les étapes de développement et pour tous les lots de production. Cela apportera une solution plus robuste et complémentaire aux méthodes statistiques actuellement utilisées pour la compréhension des spectres RMN d'anticorps.

Lors de la mise au point d'un nouvel anticorps thérapeutique potentiel, dans un premier temps, la méthode développée au cours de cette thèse sur l'anticorps anti-LAMP1 pourra être appliquée de façon systématique, à savoir : production de l'anticorps marqué ou de ces fragments pendant la phase de R&D, attribution des signaux RMN des méthyles, puis utilisation de ce spectre de référence pour valider les lots de production (non marqués).

Dans un second temps, et vue la nature conservée à 75% des anticorps appartenant à une même classe, il serait envisageable de transférer directement l'attribution obtenue pour l'anticorps anti-LAMP1 à un nouvel anticorps sans devoir réaliser son attribution *de novo*.

## 2.3. Transfert d'attribution à d'autres fragments Fabs

Le développement de la méthode permettant l'attribution des groupements méthyles d'un fragment Fab d'anticorps a nécessité trois ans de travail. Maintenant que cette méthode est développée et compte tenu de la conservation de séquence pour les domaines constants des anticorps, l'attribution des groupements méthyles d'un autre fragment Fab d'IgG1 ne nécessiterait que quelques mois. En effet, on peut espérer que 75% de l'attribution déjà réalisée

## Conclusion

sera transférable aux autres fragments Fab d'IgG1 humaines et que seules les séquences des parties variables, et notamment celles des CDR, devront être attribuées. Il serait donc intéressant d'évaluer la possibilité de n'exprimer que les fragments scFv (25 kDa) des anticorps en système *cell-free*, de les attribuer et de vérifier que cette attribution est transférable au fragment Fab dans un premier temps, puis à l'anticorps entier.

L'application de cette méthode de manière systématique à chaque anticorps thérapeutique pourrait permettre d'acquérir une base de données de l'empreinte spectrale de leurs groupements méthyles conséquente. Il faut s'attendre à ce que les premiers transferts d'attribution de l'anticorps anti-LAMP1 à un autre soient compliqués à mettre en œuvre en raison de la complexité des spectres RMN de protéines de 150 kDa. Cependant, l'accumulation des données pourrait permettre, à terme, de faciliter l'attribution des groupements méthyles des anticorps thérapeutiques, ce qui ouvrira la porte à la RMN en solution pour accéder à des informations, à l'échelle atomique, sur les propriétés conformationnelles des anticorps, les mouvements locaux, les interactions avec les partenaires ainsi qu'à la cartographie des épitopes.

## V. Bibliographie

- [1] « EMIL ADOLF VON BEHRING (1854-1917) - Encyclopædia Universalis ». <https://www.universalis.fr/encyclopedie/emil-adolf-von-behring/> (consulté le 13 mars 2023).
- [2] « RODNEY ROBERT PORTER (1917-1985) - Encyclopædia Universalis ». <https://www.universalis.fr/encyclopedie/rodney-robert-porter/> (consulté le 13 mars 2023).
- [3] G. Vidarsson, G. Dekkers, et T. Rispens, « IgG Subclasses and Allotypes: From Structure to Effector Functions », *Front. Immunol.*, vol. 5, p. 520, oct. 2014, doi: 10.3389/fimmu.2014.00520.
- [4] A. F. Williams et A. N. Barclay, « The Immunoglobulin Superfamily—Domains for Cell Surface Recognition », vol. 6, p. 381-405, 1988, doi: 0732-0582/88/0410-0381.
- [5] E. Lobner, M. W. Traxlmayr, C. Obinger, et C. Hasenhindl, « Engineered IgG1-Fc – one fragment to bind them all », *Immunol. Rev.*, vol. 270, n° 1, p. 113-131, mars 2016, doi: 10.1111/imr.12385.
- [6] J. N. Arnold, M. R. Wormald, R. B. Sim, P. M. Rudd, et R. A. Dwek, « The Impact of Glycosylation on the Biological Function and Structure of Human Immunoglobulins », *Annu. Rev. Immunol.*, vol. 25, n° 1, p. 21-50, 2007, doi: 10.1146/annurev.immunol.25.022106.141702.
- [7] M. Holland *et al.*, « Differential glycosylation of polyclonal IgG, IgG-Fc and IgG-Fab isolated from the sera of patients with ANCA-associated systemic vasculitis », *Biochim. Biophys. Acta BBA - Gen. Subj.*, vol. 1760, n° 4, p. 669-677, avr. 2006, doi: 10.1016/j.bbagen.2005.11.021.
- [8] A. Ferrante, L. Beard, et R. Feldman, « IgG subclass distribution of antibodies to bacterial and viral antigens », *Pediatr. Infect. Dis. J.*, vol. 9(8), p. p 516-524, août 1990, doi: 10.1097/00006454-199008001-00004.
- [9] J.-L. Preud'homme, I. Petit, A. Barra, F. Morel, J.-C. Lecron, et E. Lelievre, « Structural and functional properties of membrane and secreted IgD », *Mol. Immunol.*, 2000.
- [10] R. Plomp *et al.*, « Site-Specific N-Glycosylation Analysis of Human Immunoglobulin E », *J. Proteome Res.*, vol. 13, n° 2, p. 536-546, févr. 2014, doi: 10.1021/pr400714w.
- [11] J. M. Woof et M. W. Russell, « Structure and function relationships in IgA », *Mucosal Immunol.*, vol. 4, n° 6, Art. n° 6, nov. 2011, doi: 10.1038/mi.2011.39.
- [12] V. B. Klimovich, « IgM and Its Receptors: Structural and Functional Aspects », vol. 76, n° 5, 2011.
- [13] S. W. de Taeye, T. Rispens, et G. Vidarsson, « The Ligands for Human IgG and Their Effector Functions », *Antibodies*, vol. 8, n° 2, Art. n° 2, juin 2019, doi: 10.3390/antib8020030.
- [14] S. Schneider et M. Zacharias, « Atomic resolution model of the antibody Fc interaction with the complement C1q component », *Mol. Immunol.*, vol. 51, n° 1, p. 66-72, mai 2012, doi: 10.1016/j.molimm.2012.02.111.
- [15] P. Sonderrmann, R. Huber, V. Oosthuizen, et U. Jacob, « The 3.2-Å crystal structure of the human IgG1 Fc fragment–FcγRIII complex », *Nature*, vol. 406, n° 6793, Art. n° 6793, juill. 2000, doi: 10.1038/35018508.
- [16] M.-B. Charlotte, O. Marc, et W. Hervé, « FcRn, un récepteur d'IgG aux multiples facettes », vol. 25, 2009.
- [17] W. P. Burmeister, A. H. Huber, et P. J. Bjorkman, « Crystal structure of the complex of rat neonatal Fc receptor with Fc », *Nature*, vol. 372, n° 6504, Art. n° 6504, nov. 1994, doi: 10.1038/372379a0.
- [18] D. R. Leach, M. F. Krummel, et J. P. Allison, « Enhancement of Antitumor Immunity by CTLA-4 Blockade », vol. 271, 1996.
- [19] R.-M. Lu *et al.*, « Development of therapeutic antibodies for the treatment of diseases », *J. Biomed. Sci.*, vol. 27, n° 1, p. 1, janv. 2020, doi: 10.1186/s12929-019-0592-z.
- [20] S. Kopp-Kubel, « Dénominations communes internationales (DCI) pour les substances pharmaceutiques », *Bull. World Health Organ.*, vol. 73, n° 4, p. 425-429, 1995.

- [21] G. Gachelin, Y. Bottineau-Fuchs, A. Opinel, P. Rivoirard, et P. Bourdelais, « La naissance des biotechnologies pharmaceutiques en France (1887-1914) ».
- [22] C. M. G. KÖHLER, « Continuous cultures of fused cells secreting antibody of predefined specificity », *Nature*, vol. 256, p. 495-497, 1975, doi: 256(5517):495-7.
- [23] P. A. Todd et R. N. Brogden, « Muromonab CD3. A review of its pharmacology and therapeutic potential », *Drugs*, vol. 37, n° 6, p. 871-899, juin 1989, doi: 10.2165/00003495-198937060-00004.
- [24] S. L. Morrison, M. J. Johnson, L. A. Herzenberg, et V. T. Oi, « Chimeric human antibody molecules: mouse antigen-binding domains with human constant region domains. », *Proc. Natl. Acad. Sci. U. S. A.*, vol. 81, n° 21, p. 6851-6855, nov. 1984.
- [25] C. Griffiths et M. Fisher, « Abciximab-induced thrombocytopenia », *QJM Int. J. Med.*, vol. 95, n° 9, p. 635-636, sept. 2002, doi: 10.1093/qjmed/95.9.635.
- [26] C. Luo *et al.*, « Efficacy and safety of new anti-CD20 monoclonal antibodies versus rituximab for induction therapy of CD20+ B-cell non-Hodgkin lymphomas: a systematic review and meta-analysis », *Sci. Rep.*, vol. 11, p. 3255, févr. 2021, doi: 10.1038/s41598-021-82841-w.
- [27] W. Y. K. Hwang et J. Foote, « Immunogenicity of engineered antibodies », *Methods*, vol. 36, n° 1, p. 3-10, mai 2005, doi: 10.1016/j.ymeth.2005.01.001.
- [28] P. T. Jones, P. H. Dear, J. Foote, M. S. Neuberger, et G. Winter, « Replacing the complementarity- determining regions in a human antibody with those from a mouse », *Nature*, 1986.
- [29] N. Pfender et R. Martin, « Daclizumab (anti-CD25) in multiple sclerosis », *Exp. Neurol.*, vol. 262 Pt A, p. 44-51, déc. 2014, doi: 10.1016/j.expneurol.2014.04.015.
- [30] O. Stüve, R. Gold, A. Chan, E. Mix, U. Zettl, et B. C. Kieseier, «  $\alpha$ 4-Integrin antagonism with natalizumab », *J. Neurol.*, vol. 255, n° 6, p. 58-65, déc. 2008, doi: 10.1007/s00415-008-6011-0.
- [31] D. H. Miller *et al.*, « A Controlled Trial of Natalizumab for Relapsing Multiple Sclerosis », *N. Engl. J. Med.*, vol. 348, n° 1, p. 15-23, janv. 2003, doi: 10.1056/NEJMoa020696.
- [32] J. McCafferty, A. D. Griffiths, G. Winter, et D. J. Chiswell, « Phage antibodies: filamentous phage displaying antibody variable domains », *Nature*, vol. 348, n° 6301, Art. n° 6301, déc. 1990, doi: 10.1038/348552a0.
- [33] J. Askling, K. Fahrback, B. Nordstrom, S. Ross, C. H. Schmid, et D. Symmons, « Cancer risk with tumor necrosis factor alpha (TNF) inhibitors: meta-analysis of randomized controlled trials of adalimumab, etanercept, and infliximab using patient level data », *Pharmacoepidemiol. Drug Saf.*, vol. 20, n° 2, p. 119-130, févr. 2011, doi: 10.1002/pds.2046.
- [34] N. Lonberg *et al.*, « Antigen-specific human antibodies from mice comprising four distinct genetic modifications », *Nature*, vol. 368, n° 6474, Art. n° 6474, avr. 1994, doi: 10.1038/368856a0.
- [35] F. Battaglin, A. Puccini, S. Ahcene Djaballah, et H.-J. Lenz, « The impact of panitumumab treatment on survival and quality of life in patients with RAS wild-type metastatic colorectal cancer », *Cancer Manag. Res.*, vol. 11, p. 5911-5924, juin 2019, doi: 10.2147/CMAR.S186042.
- [36] « Le palivizumab et l'immunoglobuline spécifique contre le virus respiratoire syncytial par voie intraveineuse en prophylaxie chez les nourrissons à haut risque », *Paediatr. Child Health*, vol. 4, n° 7, p. 483-489, oct. 1999.
- [37] S. Makita et K. Tobinai, « Mogamulizumab for the treatment of T-cell lymphoma », *Expert Opin. Biol. Ther.*, vol. 17, n° 9, p. 1145-1153, sept. 2017, doi: 10.1080/14712598.2017.1347634.
- [38] J. Richter, L. Sanchez, et S. Thibaud, « Therapeutic potential of isatuximab in the treatment of multiple myeloma: Evidence to date », *Semin. Oncol.*, vol. 47, n° 2, p. 155-164, avr. 2020, doi: 10.1053/j.seminoncol.2020.04.004.
- [39] S. Jin *et al.*, « Emerging new therapeutic antibody derivatives for cancer treatment », *Signal Transduct. Target. Ther.*, vol. 7, n° 1, Art. n° 1, févr. 2022, doi: 10.1038/s41392-021-00868-x.

- [40] H. Yamazaki *et al.*, « Certolizumab pegol for induction of remission in Crohn's disease », *Cochrane Database Syst. Rev.*, vol. 2019, n° 8, p. CD012893, août 2019, doi: 10.1002/14651858.CD012893.pub2.
- [41] P. J. Lenting, C. V. Denis, et O. D. Christophe, « Emicizumab, a bispecific antibody recognizing coagulation factors IX and X: how does it actually compare to factor VIII? », *Blood*, vol. 130, n° 23, p. 2463-2468, déc. 2017, doi: 10.1182/blood-2017-08-801662.
- [42] C. A. Portell, C. M. Wenzell, et A. S. Advani, « Clinical and pharmacologic aspects of blinatumomab in the treatment of B-cell acute lymphoblastic leukemia », *Clin. Pharmacol. Adv. Appl.*, vol. 5, n° Suppl 1, p. 5-11, avr. 2013, doi: 10.2147/CPAA.S42689.
- [43] N. L. Bartlett *et al.*, « Brentuximab vedotin activity in diffuse large B-cell lymphoma with CD30 undetectable by visual assessment of conventional immunohistochemistry », *Leuk. Lymphoma*, vol. 58, n° 7, p. 1607-1616, juill. 2017, doi: 10.1080/10428194.2016.1256481.
- [44] B. L. Duivelshof *et al.*, « Therapeutic Fc-fusion proteins: Current analytical strategies », *J. Sep. Sci.*, vol. 44, n° 1, p. 35-62, 2021, doi: 10.1002/jssc.202000765.
- [45] P. Murer et D. Neri, « Antibody-cytokine fusion proteins: A novel class of biopharmaceuticals for the therapy of cancer and of chronic inflammation », *New Biotechnol.*, vol. 52, p. 42-53, sept. 2019, doi: 10.1016/j.nbt.2019.04.002.
- [46] J. Favresse *et al.*, « Neutralizing Antibodies in COVID-19 Patients and Vaccine Recipients after Two Doses of BNT162b2 », *Viruses*, vol. 13, n° 7, Art. n° 7, juill. 2021, doi: 10.3390/v13071364.
- [47] P. C. Taylor, A. C. Adams, M. M. Hufford, I. de la Torre, K. Winthrop, et R. L. Gottlieb, « Neutralizing monoclonal antibodies for treatment of COVID-19 », *Nat. Rev. Immunol.*, vol. 21, n° 6, Art. n° 6, juin 2021, doi: 10.1038/s41577-021-00542-x.
- [48] V. P. Chavda *et al.*, « Therapeutic monoclonal antibodies for COVID-19 management: an update », *Expert Opin. Biol. Ther.*, vol. 22, n° 6, p. 763-780, juin 2022, doi: 10.1080/14712598.2022.2078160.
- [49] V. P. Chavda, A. B. Patel, et D. D. Vaghasiya, « SARS-CoV-2 variants and vulnerability at the global level », *J. Med. Virol.*, vol. 94, n° 7, p. 2986-3005, 2022, doi: 10.1002/jmv.27717.
- [50] L. Liu *et al.*, « Striking antibody evasion manifested by the Omicron variant of SARS-CoV-2 », *Nature*, vol. 602, n° 7898, Art. n° 7898, févr. 2022, doi: 10.1038/s41586-021-04388-0.
- [51] L. Liu *et al.*, « Antibodies that neutralize all current SARS-CoV-2 variants of concern by conformational locking ». bioRxiv, p. 2023.04.08.536123, 10 avril 2023. doi: 10.1101/2023.04.08.536123.
- [52] H. L. Svilenov *et al.*, « Extrinsic stabilization of antiviral ACE2-Fc fusion proteins targeting SARS-CoV-2 », *Commun. Biol.*, vol. 6, n° 1, Art. n° 1, avr. 2023, doi: 10.1038/s42003-023-04762-w.
- [53] S. Wilke, J. Krausze, et K. Büssow, « Crystal structure of the conserved domain of the DC lysosomal associated membrane protein: implications for the lysosomal glycocalyx », *BMC Biol.*, vol. 10, p. 62, juill. 2012, doi: 10.1186/1741-7007-10-62.
- [54] K. Terasawa *et al.*, « Lysosome-associated membrane proteins-1 and -2 (LAMP-1 and LAMP-2) assemble via distinct modes », *Biochem. Biophys. Res. Commun.*, vol. 479, n° 3, p. 489-495, oct. 2016, doi: 10.1016/j.bbrc.2016.09.093.
- [55] B. Cameron *et al.*, « Complementary epitopes and favorable developability of monoclonal anti-LAMP1 antibodies generated using two transgenic animal platforms », *PLOS ONE*, vol. 15, n° 7, p. e0235815, juill. 2020, doi: 10.1371/journal.pone.0235815.
- [56] A. Cohnen *et al.*, « Surface CD107a/LAMP-1 protects natural killer cells from degranulation-associated damage », *Blood*, vol. 122, n° 8, p. 1411-1418, août 2013, doi: 10.1182/blood-2012-07-441832.
- [57] F. Alessandrini, L. Pezzè, et Y. Ciribilli, « LAMPs: Shedding light on cancer biology », *Semin. Oncol.*, vol. 44, n° 4, p. 239-253, août 2017, doi: 10.1053/j.seminoncol.2017.10.013.
- [58] A. K. Agarwal, R. P. Gude, et R. D. Kalraiya, « Regulation of melanoma metastasis to lungs by cell surface Lysosome Associated Membrane Protein-1 (LAMP1) via galectin-3 », *Biochem. Biophys. Res. Commun.*, vol. 449, n° 3, p. 332-337, juill. 2014, doi: 10.1016/j.bbrc.2014.05.028.

- [59] T. Pruvost, M. Mathieu, S. Dubois, B. Maillère, E. Vigne, et H. Nozach, « Deciphering cross-species reactivity of LAMP-1 antibodies using deep mutational epitope mapping and AlphaFold », *mAbs*, vol. 15, 2023, doi: 10.1080/19420862.2023.2175311.
- [60] S. Chakravarthi, C. E. Jessop, et N. J. Bulleid, « The role of glutathione in disulphide bond formation and endoplasmic-reticulum-generated oxidative stress », *EMBO Rep.*, vol. 7, n° 3, p. 271-275, mars 2006, doi: 10.1038/sj.embor.7400645.
- [61] H. Nothhaft et C. M. Szymanski, « Protein glycosylation in bacteria: sweeter than ever », *Nat. Rev. Microbiol.*, vol. 8, n° 11, Art. n° 11, nov. 2010, doi: 10.1038/nrmicro2383.
- [62] « N-Glycan Engineering: Constructing the N-GlcNAc Stump », p. 7, 2022.
- [63] S. Groux-Degroote, F. Foulquier, S. Cavdarli, et P. Delannoy, « Les mécanismes de régulation de la glycosylation - Exemples d'altérations des chaînes glycaniques dans les cancers », *médecine/sciences*, vol. 37, n° 6-7, Art. n° 6-7, juin 2021, doi: 10.1051/medsci/2021082.
- [64] R. L. Shields *et al.*, « Lack of Fucose on Human IgG1 N-Linked Oligosaccharide Improves Binding to Human FcγRIII and Antibody-dependent Cellular Toxicity\* », *J. Biol. Chem.*, vol. 277, n° 30, p. 26733-26740, juill. 2002, doi: 10.1074/jbc.M202069200.
- [65] R. Jefferis, « Recombinant antibody therapeutics: the impact of glycosylation on mechanisms of action », *Trends Pharmacol. Sci.*, vol. 30, n° 7, p. 356-362, juill. 2009, doi: 10.1016/j.tips.2009.04.007.
- [66] P. Sondermann, A. Pincetic, J. Maamary, K. Lammens, et J. V. Ravetch, « General mechanism for modulating immunoglobulin effector function », *Proc. Natl. Acad. Sci. U. S. A.*, vol. 110, n° 24, p. 9868-9872, juin 2013, doi: 10.1073/pnas.1307864110.
- [67] J. Kim, H. Luo, W. White, W. Rees, R. Venkat, et M. Albarghouthi, « Impact of Fc N-linked glycans on in vivo clearance of an immunoglobulin G1 antibody produced by NS0 cell line », *mAbs*, vol. 12, n° 1, p. 1844928, doi: 10.1080/19420862.2020.1844928.
- [68] V. G. Dhara, H. M. Naik, N. I. Majewska, et M. J. Betenbaugh, « Recombinant Antibody Production in CHO and NS0 Cells: Differences and Similarities », *BioDrugs*, vol. 32, n° 6, p. 571-584, déc. 2018, doi: 10.1007/s40259-018-0319-9.
- [69] G. Urlaub et L. A. Chasin, « Isolation of Chinese hamster cell mutants deficient in dihydrofolate reductase activity. », *Proc. Natl. Acad. Sci. U. S. A.*, vol. 77, n° 7, p. 4216-4220, juill. 1980.
- [70] A. A. Lin, R. Kimura, et W. M. Miller, « Production of tPA in recombinant CHO cells under oxygen-limited conditions », *Biotechnol. Bioeng.*, vol. 42, n° 3, p. 339-350, 1993, doi: 10.1002/bit.260420311.
- [71] Y. Takagi, T. Kikuchi, R. Wada, et T. Omasa, « The enhancement of antibody concentration and achievement of high cell density CHO cell cultivation by adding nucleoside », *Cytotechnology*, vol. 69, n° 3, p. 511-521, juin 2017, doi: 10.1007/s10616-017-0066-7.
- [72] N. Yamane-Ohnuki *et al.*, « Establishment of FUT8 knockout Chinese hamster ovary cells: An ideal host cell line for producing completely defucosylated antibodies with enhanced antibody-dependent cellular cytotoxicity », *Biotechnol. Bioeng.*, vol. 87, n° 5, p. 614-622, 2004, doi: 10.1002/bit.20151.
- [73] Z. S.-L. Heng, J. Y. Yeo, D. W.-S. Koh, S. K.-E. Gan, et W.-L. Ling, « Augmenting recombinant antibody production in HEK293E cells: optimizing transfection and culture parameters », *Antib. Ther.*, vol. 5, n° 1, p. 30, janv. 2022, doi: 10.1093/abt/tbac003.
- [74] E. Fox *et al.*, « A phase 2 multicenter study of ublituximab, a novel glycoengineered anti-CD20 monoclonal antibody, in patients with relapsing forms of multiple sclerosis », *Mult. Scler. Houndmills Basingstoke Engl.*, vol. 27, n° 3, p. 420-429, mars 2021, doi: 10.1177/1352458520918375.
- [75] Y. Kanda *et al.*, « Comparison of cell lines for stable production of fucose-negative antibodies with enhanced ADCC », *Biotechnol. Bioeng.*, vol. 94, n° 4, p. 680-688, 2006, doi: 10.1002/bit.20880.
- [76] Y. Konno *et al.*, « Fucose content of monoclonal antibodies can be controlled by culture medium osmolality for high antibody-dependent cellular cytotoxicity », *Cytotechnology*, vol. 64, n° 3, p. 249-265, mai 2012, doi: 10.1007/s10616-011-9377-2.

- [77] K. R. Love, N. C. Dalvie, et J. C. Love, « The yeast stands alone: the future of protein biologic production », *Curr. Opin. Biotechnol.*, vol. 53, p. 50-58, oct. 2018, doi: 10.1016/j.copbio.2017.12.010.
- [78] L. Van Landuyt, C. Lonigro, L. Meuris, et N. Callewaert, « Customized protein glycosylation to improve biopharmaceutical function and targeting », *Curr. Opin. Biotechnol.*, vol. 60, p. 17-28, déc. 2019, doi: 10.1016/j.copbio.2018.11.017.
- [79] M. Berdichevsky, M. d'Anjou, M. R. Mallem, S. S. Shaikh, et T. I. Potgieter, « Improved production of monoclonal antibodies through oxygen-limited cultivation of glycoengineered yeast », *J. Biotechnol.*, vol. 155, n° 2, p. 217-224, sept. 2011, doi: 10.1016/j.jbiotec.2011.06.021.
- [80] L. Friedrich *et al.*, « Bacterial production and functional characterization of the Fab fragment of the murine IgG1/ $\lambda$  monoclonal antibody cmHsp70.1, a reagent for tumour diagnostics », *Protein Eng. Des. Sel.*, vol. 23, n° 4, p. 161-168, avr. 2010, doi: 10.1093/protein/gzp095.
- [81] R. Freudl, « Signal peptides for recombinant protein secretion in bacterial expression systems », *Microb. Cell Factories*, vol. 17, n° 1, p. 52, mars 2018, doi: 10.1186/s12934-018-0901-3.
- [82] A. Gaciarz *et al.*, « Systematic screening of soluble expression of antibody fragments in the cytoplasm of *E. coli* », *Microb. Cell Factories*, vol. 15, n° 1, p. 22, janv. 2016, doi: 10.1186/s12934-016-0419-5.
- [83] F. Rezaie *et al.*, « Cytosolic expression of functional Fab fragments in *Escherichia coli* using a novel combination of dual SUMO expression cassette and EnBase® cultivation mode », *J. Appl. Microbiol.*, vol. 123, n° 1, p. 134-144, 2017, doi: 10.1111/jam.13483.
- [84] Chung, K.T, Lee, T.H, et Kang, G.S, « Isolation of proteins that specifically interact with the ATPase domain of mammalian ER chaperone, BiP. », *Biotechnol Bioprocess Eng*, vol. 8, p. 192-198, 2003, doi: <https://doi.org/10.1007/BF02935896>.
- [85] A. Singh, V. Upadhyay, A. K. Upadhyay, S. M. Singh, et A. K. Panda, « Protein recovery from inclusion bodies of *Escherichia coli* using mild solubilization process », *Microb. Cell Factories*, vol. 14, n° 1, p. 41, mars 2015, doi: 10.1186/s12934-015-0222-8.
- [86] D. Arora et N. Khanna, « Method for increasing the yield of properly folded recombinant human gamma interferon from inclusion bodies », *J. Biotechnol.*, vol. 52, n° 2, p. 127-133, déc. 1996, doi: 10.1016/S0168-1656(96)01636-7.
- [87] P. Priyanka et A. S. Rathore, « A novel strategy for efficient expression of an antibody fragment in *Escherichia coli*: ranibizumab as a case study », *J. Chem. Technol. Biotechnol.*, vol. 97, n° 1, p. 42-54, 2022, doi: 10.1002/jctb.6883.
- [88] European Medicines Agency, « ICH Q5E Comparability of Biotechnological\_Biological Products.pdf ». juin 2005.
- [89] « Q8R(2) ICH harmonised tripartite guideline of Pharmaceutical development ». août 2009. Consulté le: 6 avril 2023. [En ligne]. Disponible sur: [https://database.ich.org/sites/default/files/Q8\\_R2\\_Guideline.pdf](https://database.ich.org/sites/default/files/Q8_R2_Guideline.pdf)
- [90] H. Liu, G. Gaza-Bulsecu, D. Faldu, C. Chumsae, et J. Sun, « Heterogeneity of monoclonal antibodies », *J. Pharm. Sci.*, vol. 97, n° 7, p. 2426-2447, 2008, doi: 10.1002/jps.21180.
- [91] H. J. An, J. W. Froehlich, et C. B. Lebrilla, « Determination of Glycosylation Sites and Site-specific Heterogeneity in Glycoproteins », *Curr. Opin. Chem. Biol.*, vol. 13, n° 4, p. 421-426, oct. 2009, doi: 10.1016/j.cbpa.2009.07.022.
- [92] Y. D. Liu, A. M. Goetze, R. B. Bass, et G. C. Flynn, « N-terminal Glutamate to Pyroglutamate Conversion in Vivo for Human IgG2 Antibodies », *J. Biol. Chem.*, vol. 286, n° 13, p. 11211-11217, avr. 2011, doi: 10.1074/jbc.M110.185041.
- [93] R. J. Harris, « Processing of C-terminal lysine and arginine residues of proteins isolated from mammalian cell culture », *J. Chromatogr. A*, vol. 705, n° 1, p. 129-134, juin 1995, doi: 10.1016/0021-9673(94)01255-D.

- [94] V. Faid, Y. Leblanc, M. Berger, A. Seifert, N. Bihoreau, et G. Chevreux, « C-terminal lysine clipping of IgG1: impact on binding to human FcγRIIIa and neonatal Fc receptors », *Eur. J. Pharm. Sci.*, vol. 159, 2021, doi: 10.1016/j.ejps.2021.105730.
- [95] K. Kato, T. Nakayoshi, E. Kurimoto, et A. Oda, « Mechanisms of Deamidation of Asparagine Residues and Effects of Main-Chain Conformation on Activation Energy », *Int. J. Mol. Sci.*, vol. 21, n° 19, p. 7035, sept. 2020, doi: 10.3390/ijms21197035.
- [96] J. J. Phillips *et al.*, « Rate of Asparagine Deamidation in a Monoclonal Antibody Correlating with Hydrogen Exchange Rate at Adjacent Downstream Residues », *Anal. Chem.*, vol. 89, n° 4, p. 2361-2368, févr. 2017, doi: 10.1021/acs.analchem.6b04158.
- [97] M. Habegger *et al.*, « Assessment of chemical modifications of sites in the CDRs of recombinant antibodies », *mAbs*, vol. 6, n° 2, p. 327-339, mars 2014, doi: 10.4161/mabs.27876.
- [98] J. Kuang *et al.*, « Understanding the pathway and kinetics of aspartic acid isomerization in peptide mapping methods for monoclonal antibodies », *Anal. Bioanal. Chem.*, vol. 413, n° 8, p. 2113-2123, mars 2021, doi: 10.1007/s00216-021-03176-z.
- [99] B. Spanov, N. Govorukhina, N. C. van de Merbel, et R. Bischoff, « Analytical tools for the characterization of deamidation in monoclonal antibodies », *J. Chromatogr. Open*, vol. 2, p. 100025, nov. 2022, doi: 10.1016/j.jcoa.2021.100025.
- [100] A. A. Wakankar *et al.*, « Aspartate Isomerization in the Complementarity-Determining Regions of Two Closely Related Monoclonal Antibodies », *Biochemistry*, vol. 46, n° 6, p. 1534-1544, févr. 2007, doi: 10.1021/bi061500t.
- [101] D. Liu *et al.*, « Structure and Stability Changes of Human IgG1 Fc as a Consequence of Methionine Oxidation », *Biochemistry*, vol. 47, n° 18, p. 5088-5100, mai 2008, doi: 10.1021/bi702238b.
- [102] B. Wei, K. Berning, C. Quan, et Y. T. Zhang, « Glycation of antibodies: Modification, methods and potential effects on biological functions », *mAbs*, vol. 9, n° 4, p. 586-594, 2017, doi: 10.1080/19420862.2017.1300214.
- [103] M. A. Grillo et S. Colombatto, « Advanced glycation end-products (AGEs): involvement in aging and in neurodegenerative diseases », *Amino Acids*, vol. 35, n° 1, p. 29-36, juin 2008, doi: 10.1007/s00726-007-0606-0.
- [104] I. H. Yuk *et al.*, « Controlling glycation of recombinant antibody in fed-batch cell cultures », *Biotechnol. Bioeng.*, vol. 108, n° 11, p. 2600-2610, nov. 2011, doi: 10.1002/bit.23218.
- [105] D. D. Banks, D. M. Hambly, J. L. Scavezze, C. C. Siska, N. L. Stackhouse, et H. S. Gadgil, « The effect of sucrose hydrolysis on the stability of protein therapeutics during accelerated formulation studies », *J. Pharm. Sci.*, vol. 98, n° 12, p. 4501-4510, déc. 2009, doi: 10.1002/jps.21749.
- [106] G. Basta *et al.*, « Advanced Glycation End Products Activate Endothelium Through Signal-Transduction Receptor RAGE », *Circulation*, vol. 105, n° 7, p. 816-822, févr. 2002, doi: 10.1161/hc0702.104183.
- [107] R. A. Saleem *et al.*, « A chemical and computational approach to comprehensive glycation characterization on antibodies », *mAbs*, vol. 7, n° 4, p. 719-731, 2015, doi: 10.1080/19420862.2015.1046663.
- [108] C. Quan *et al.*, « A study in glycation of a therapeutic recombinant humanized monoclonal antibody: Where it is, how it got there, and how it affects charge-based behavior », *Anal. Biochem.*, vol. 373, n° 2, p. 179-191, févr. 2008, doi: 10.1016/j.ab.2007.09.027.
- [109] S. Ahmad, Moinuddin, et A. Ali, « Immunological studies on glycated human IgG », *Life Sci.*, vol. 90, n° 25, p. 980-987, juin 2012, doi: 10.1016/j.lfs.2012.05.002.
- [110] Y. Wang, Q. Lu, S.-L. Wu, B. L. Karger, et W. S. Hancock, « Characterization and Comparison of Disulfide Linkages and Scrambling Patterns in Therapeutic Monoclonal Antibodies - Using LC-MS with Electron Transfer Dissociation », *Anal. Chem.*, vol. 83, n° 8, p. 3133-3140, avr. 2011, doi: 10.1021/ac200128d.

- [111] M. Habberger *et al.*, « Rapid characterization of biotherapeutic proteins by size-exclusion chromatography coupled to native mass spectrometry », *mAbs*, vol. 8, n° 2, p. 331-339, févr. 2016, doi: 10.1080/19420862.2015.1122150.
- [112] E. A. Padlan, D. M. Segal, T. F. Spande, D. R. Davies, S. Rudikoff, et M. Potter, « Structure at 4.5 Å; Resolution of a Phosphorylcholine-binding Fab », *Nature. New Biol.*, vol. 245, p. 165-167, 1973, doi: <https://doi.org/10.1038/newbio245165a0>.
- [113] E. O. Saphire *et al.*, « Crystal Structure of a Neutralizing Human IgG Against HIV-1: A Template for Vaccine Design », *Science*, vol. 293, n° 5532, p. 1155-1159, août 2001, doi: 10.1126/science.1061692.
- [114] C. Y. Tremblay, P. Limpikirati, et R. W. Vachet, « Complementary Structural Information for Stressed Antibodies from Hydrogen-Deuterium Exchange and Covalent Labeling Mass Spectrometry », *J. Am. Soc. Mass Spectrom.*, vol. 32, n° 5, p. 1237-1248, mai 2021, doi: 10.1021/jasms.1c00072.
- [115] L. Poppe, J. B. Jordan, K. Lawson, M. Jerums, I. Apostol, et P. D. Schnier, « Profiling Formulated Monoclonal Antibodies by <sup>1</sup>H NMR Spectroscopy », *Anal. Chem.*, vol. 85, n° 20, p. 9623-9629, oct. 2013, doi: 10.1021/ac401867f.
- [116] J. Franks, J. N. Glushka, M. T. Jones, D. H. Live, Q. Zou, et J. H. Prestegard, « Spin Diffusion Editing for Structural Fingerprints of Therapeutic Antibodies », *Anal. Chem.*, vol. 88, n° 2, p. 1320-1327, janv. 2016, doi: 10.1021/acs.analchem.5b03777.
- [117] K. Chen, D. S. Long, S. C. Lute, M. J. Levy, K. A. Brorson, et D. A. Keire, « Simple NMR methods for evaluating higher order structures of monoclonal antibody therapeutics with quinary structure », *J. Pharm. Biomed. Anal.*, vol. 128, p. 398-407, sept. 2016, doi: 10.1016/j.jpba.2016.06.007.
- [118] S. Majumder *et al.*, « Utility of High Resolution NMR Methods to Probe the Impact of Chemical Modifications on Higher Order Structure of Monoclonal Antibodies in Relation to Antigen Binding », *Pharm. Res.*, vol. 36, n° 9, p. 130, juill. 2019, doi: 10.1007/s11095-019-2652-1.
- [119] J. P. Marino *et al.*, « Emerging Technologies To Assess the Higher Order Structure of Monoclonal Antibodies », in *State-of-the-Art and Emerging Technologies for Therapeutic Monoclonal Antibody Characterization Volume 3. Defining the Next Generation of Analytical and Biophysical Techniques*, in ACS Symposium Series, no. 1202, vol. 1202. American Chemical Society, 2015, p. 17-43. doi: 10.1021/bk-2015-1202.ch002.
- [120] L. W. Arbogast, R. G. Brinson, et J. P. Marino, « Mapping Monoclonal Antibody Structure by 2D 13C NMR at Natural Abundance », *Anal. Chem.*, vol. 87, n° 7, p. 3556-3561, avr. 2015, doi: 10.1021/ac504804m.
- [121] R. G. Brinson *et al.*, « Enabling adoption of 2D-NMR for the higher order structure assessment of monoclonal antibody therapeutics », *mAbs*, vol. 11, n° 1, p. 94-105, déc. 2018, doi: 10.1080/19420862.2018.1544454.
- [122] V. Tugarinov, P. M. Hwang, J. E. Ollerenshaw, et L. E. Kay, « Cross-Correlated Relaxation Enhanced 1H-13C NMR Spectroscopy of Methyl Groups in Very High Molecular Weight Proteins and Protein Complexes », *J. Am. Chem. Soc.*, vol. 125, n° 34, p. 10420-10428, août 2003, doi: 10.1021/ja030153x.
- [123] K. Pervushin, R. Riek, G. Wider, et K. Wüthrich, « Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution », *Proc. Natl. Acad. Sci.*, vol. 94, n° 23, p. 12366-12371, nov. 1997, doi: 10.1073/pnas.94.23.12366.
- [124] K. Wüthrich, « NMR with Proteins and Nucleic Acids », *Europhys. News*, vol. 17, n° 1, p. 11-13, 1986, doi: 10.1051/epn/19861701011.
- [125] G. M. Clore, A. Bax, P. C. Driscoll, P. T. Wingfield, et A. M. Gronenborn, « Assignment of the side-chain proton and carbon-13 resonances of interleukin-1.β. using double- and triple-resonance heteronuclear three-dimensional NMR spectroscopy », *Biochemistry*, vol. 29, n° 35, p. 8172-8184, sept. 1990, doi: 10.1021/bi00487a027.

- [126] G. Mas, E. Crublet, O. Hamelin, P. Gans, et J. Boisbouvier, « Specific labeling and assignment strategies of valine methyl groups for NMR studies of high molecular weight proteins », *J. Biomol. NMR*, vol. 57, n° 3, p. 251-262, nov. 2013, doi: 10.1007/s10858-013-9785-z.
- [127] I. Ayala, L. Chiari, R. Kerfah, J. Boisbouvier, P. Gans, et O. Hamelin, « Asymmetric Synthesis of Methyl Specifically Labelled L-Threonine and Application to the NMR Studies of High Molecular Weight Proteins », *ChemistrySelect*, vol. 5, n° 17, p. 5092-5098, 2020, doi: 10.1002/slct.202000827.
- [128] F. Henot, « HSP90 une cible thérapeutique dynamique: modulation des conformations peuplées par la liaison de ligands », 2022.
- [129] E. Crublet *et al.*, « A Cost-Effective Protocol for the Parallel Production of Libraries of <sup>13</sup>CH<sub>3</sub>-Specifically Labeled Mutants for NMR Studies of High Molecular Weight Proteins », in *Structural Genomics: General Applications*, Y. W. Chen, Éd., in *Methods in Molecular Biology*. Totowa, NJ: Humana Press, 2014, p. 229-244. doi: 10.1007/978-1-62703-691-7\_17.
- [130] M. Lescanne, P. Ahuja, A. Blok, M. Timmer, T. Akerud, et M. Ubbink, « Methyl group reorientation under ligand binding probed by pseudocontact shifts », *J. Biomol. NMR*, vol. 71, n° 4, p. 275-285, août 2018, doi: 10.1007/s10858-018-0190-5.
- [131] F. Henot *et al.*, « Optimized precursor to simplify assignment transfer between backbone resonances and stereospecifically labelled valine and leucine methyl groups: application to human Hsp90 N-terminal domain », *J. Biomol. NMR*, vol. 75, n° 6, p. 221-232, juill. 2021, doi: 10.1007/s10858-021-00370-0.
- [132] James Keeler, *Understanding NMR spectroscopy*, Wiley.
- [133] I. Pritišanac, « Automated assignment of methyl NMR spectra from large proteins », *Prog. Nucl. Magn. Reson. Spectrosc.*, p. 20, 2020.
- [134] R. Sounier, L. Blanchard, Z. Wu, et J. Boisbouvier, « High-Accuracy Distance Measurement between Remote Methyls in Specifically Protonated Proteins », *J. Am. Chem. Soc.*, vol. 129, n° 3, p. 472-473, janv. 2007, doi: 10.1021/ja067260m.
- [135] Y. Xu et S. Matthews, « MAP-XSII: an improved program for the automatic assignment of methyl resonances in large proteins », *J. Biomol. NMR*, vol. 55, n° 2, p. 179-187, févr. 2013, doi: 10.1007/s10858-012-9700-z.
- [136] Y. Xu *et al.*, « Automated Assignment in Selectively Methyl-Labeled Proteins », *J. Am. Chem. Soc.*, vol. 131, n° 27, p. 9480-9481, juill. 2009, doi: 10.1021/ja9020233.
- [137] F.-A. Chao, J. Kim, Y. Xia, M. Milligan, N. Rowe, et G. Veglia, « FLAMEnGO 2.0: An enhanced fuzzy logic algorithm for structure-based assignment of methyl group resonances », *J. Magn. Reson.*, vol. 245, p. 17-23, août 2014, doi: 10.1016/j.jmr.2014.04.012.
- [138] I. Pritišanac *et al.*, « Automatic Assignment of Methyl-NMR Spectra of Supramolecular Machines Using Graph Theory », *J. Am. Chem. Soc.*, vol. 139, n° 28, p. 9523-9533, juill. 2017, doi: 10.1021/jacs.6b11358.
- [139] Y. R. Monneau *et al.*, « Automatic methyl assignment in large proteins by the MAGIC algorithm », *J. Biomol. NMR*, vol. 69, n° 4, p. 215-227, déc. 2017, doi: 10.1007/s10858-017-0149-y.
- [140] I. Pritišanac, J. M. Würz, T. R. Alderson, et P. Güntert, « Automatic structure-based NMR methyl resonance assignment in large proteins », *Nat. Commun.*, vol. 10, n° 1, Art. n° 1, oct. 2019, doi: 10.1038/s41467-019-12837-8.
- [141] S. Nerli, V. S. De Paula, A. C. McShan, et N. G. Sgourakis, « Backbone-independent NMR resonance assignments of methyl probes in large proteins », *Nat. Commun.*, vol. 12, n° 1, p. 691, janv. 2021, doi: 10.1038/s41467-021-20984-0.
- [142] R. Törner, R. Awad, P. Gans, B. Brutscher, et J. Boisbouvier, « Spectral editing of intra- and inter-chain methyl-methyl NOEs in protein complexes », *J. Biomol. NMR*, vol. 74, n° 1, p. 83-94, janv. 2020, doi: 10.1007/s10858-019-00293-x.
- [143] A. Favier et B. Brutscher, « NMRlib: user-friendly pulse sequence tools for Bruker NMR spectrometers », *J. Biomol. NMR*, vol. 73, n° 5, p. 199-211, mai 2019, doi: 10.1007/s10858-019-00249-1.

- [144] A. S. Spirin et J. R. Swartz, « Cell-Free Protein Synthesis Systems: Historical Landmarks, Classification, and General Methods », in *Cell-Free Protein Synthesis*, John Wiley & Sons, Ltd, 2007, p. 1-34. doi: 10.1002/9783527622702.ch1.
- [145] L. Imbert *et al.*, « In Vitro Production of Perdeuterated Proteins in H<sub>2</sub>O for Biomolecular NMR Studies », in *Structural Genomics: General Applications*, Y. W. Chen et C.-P. B. Yiu, Éd., in *Methods in Molecular Biology*. New York, NY: Springer US, 2021, p. 127-149. doi: 10.1007/978-1-0716-0892-0\_8.
- [146] V. Tugarinov et L. E. Kay, « Methyl Groups as Probes of Structure and Dynamics in NMR Studies of High-Molecular-Weight Proteins », *ChemBioChem*, vol. 6, n° 9, p. 1567-1577, 2005, doi: 10.1002/cbic.200500110.
- [147] F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, et A. Bax, « NMRPipe: A multidimensional spectral processing system based on UNIX pipes », *J. Biomol. NMR*, vol. 6, n° 3, p. 277-293, nov. 1995, doi: 10.1007/BF00197809.
- [148] S. P. Skinner, R. H. Fogh, W. Boucher, T. J. Ragan, L. G. Mureddu, et G. W. Vuister, « CcpNmr AnalysisAssign: a flexible platform for integrated NMR analysis », *J. Biomol. Nmr*, vol. 66, n° 2, p. 111-124, 2016, doi: 10.1007/s10858-016-0060-y.
- [149] I. Pritišanac, J. M. Würz, T. R. Alderson, et P. Güntert, « Automatic structure-based NMR methyl resonance assignment in large proteins », *Nat. Commun.*, vol. 10, n° 1, Art. n° 1, oct. 2019, doi: 10.1038/s41467-019-12837-8.
- [150] Y. Laroche, V. Storrne, J. De Meutter, J. Messens, et M. Lauwereys, « High-Level Secretion and Very Efficient Isotopic Labeling of Tick Anticoagulant Peptide (TAP) Expressed in the Methylotrophic Yeast, *Pichia pastoris* », *Biotechnol. N*, vol. 12, p. 1119-24, nov. 1994, doi: 10.1038/nbt1194-1119.
- [151] A. Sitarska *et al.*, « Affordable uniform isotope labeling with 2H, 13C and 15N in insect cells », *J. Biomol. NMR*, vol. 62, n° 2, p. 191-197, juin 2015, doi: 10.1007/s10858-015-9935-6.
- [152] C. Opitz, S. Isogai, et S. Grzesiek, « An economic approach to efficient isotope labeling in insect cells using homemade 15N-, 13C- and 2H-labeled yeast extracts », *J. Biomol. NMR*, vol. 62, n° 3, p. 373-385, juill. 2015, doi: 10.1007/s10858-015-9954-3.
- [153] A. P. Hansen, A. M. Petros, A. P. Mazar, T. M. Pederson, A. Rueter, et S. W. Fesik, « A practical method for uniform isotopic labeling of recombinant proteins in mammalian cells », *Biochemistry*, vol. 31, n° 51, p. 12713-12718, déc. 1992, doi: 10.1021/bi00166a001.
- [154] T. A. Egorova-Zachernyuk, G. J. C. G. M. Bosman, A. M. A. Pistorius, et W. J. DeGrip, « Production of yeastolates for uniform stable isotope labelling in eukaryotic cell culture », *Appl. Microbiol. Biotechnol.*, vol. 84, n° 3, p. 575-581, sept. 2009, doi: 10.1007/s00253-009-2063-z.
- [155] JoyceW. Lustbader *et al.*, « Expression of human chorionic gonadotropin uniformly labeled with NMR isotopes in Chinese hamster ovary cells: An advance toward rapid determination of glycoprotein structures », *J. Biomol. NMR*, vol. 7, n° 4, juin 1996, doi: 10.1007/BF00200431.
- [156] A. Meola *et al.*, « Robust and low cost uniform 15N-labeling of proteins expressed in *Drosophila* S2 cells and *Spodoptera frugiperda* Sf9 cells for NMR applications », *J. Struct. Biol.*, vol. 188, n° 1, p. 71-78, oct. 2014, doi: 10.1016/j.jsb.2014.08.002.
- [157] L. Clark, I. Dikiy, D. M. Rosenbaum, et K. H. Gardner, « On the use of *Pichia pastoris* for isotopic labeling of human GPCRs for NMR studies », *J. Biomol. NMR*, vol. 71, n° 4, p. 203-211, août 2018, doi: 10.1007/s10858-018-0204-3.
- [158] A. Grahl, L. A. Abiko, S. Isogai, T. Sharpe, et S. Grzesiek, « A high-resolution description of  $\beta$ 1-adrenergic receptor functional dynamics and allosteric coupling from backbone NMR », *Nat. Commun.*, vol. 11, n° 1, Art. n° 1, mai 2020, doi: 10.1038/s41467-020-15864-y.
- [159] M. Miyazawa-Onami, K. Takeuchi, T. Takano, T. Sugiki, I. Shimada, et H. Takahashi, « Perdeuteration and methyl-selective 1H, 13C-labeling by using a *Kluyveromyces lactis* expression system », *J. Biomol. NMR*, vol. 57, n° 3, p. 297-304, nov. 2013, doi: 10.1007/s10858-013-9789-8.

- [160] M. Zhang *et al.*, «  $^{13}\text{C}$  metabolic flux profiling of *Pichia pastoris* grown in aerobic batch cultures on glucose revealed high relative anabolic use of TCA cycle and limited incorporation of provided precursors of branched-chain amino acids », *FEBS J.*, vol. 284, n° 18, p. 3100-3113, 2017, doi: 10.1111/febs.14180.
- [161] R. Suzuki, M. Sakakura, M. Mori, M. Fujii, S. Akashi, et H. Takahashi, « Methyl-selective isotope labeling using  $\alpha$ -ketoisovalerate for the yeast *Pichia pastoris* recombinant protein expression system », *J. Biomol. NMR*, vol. 71, n° 4, p. 213-223, août 2018, doi: 10.1007/s10858-018-0192-3.
- [162] Y. Kofuku *et al.*, « Functional Dynamics of Deuterated  $\beta$ 2-Adrenergic Receptor in Lipid Bilayers Revealed by NMR Spectroscopy », *Angew. Chem. Int. Ed.*, vol. 53, n° 49, p. 13376-13379, 2014, doi: 10.1002/anie.201406603.
- [163] Y. Kofuku *et al.*, « Deuteration and selective labeling of alanine methyl groups of  $\beta$ 2-adrenergic receptor expressed in a baculovirus-insect cell expression system », *J. Biomol. NMR*, vol. 71, n° 3, p. 185-192, juill. 2018, doi: 10.1007/s10858-018-0174-5.
- [164] A. Dubey *et al.*, « Local Deuteration Enables NMR Observation of Methyl Groups in Proteins from Eukaryotic and Cell-Free Expression Systems », *Angew. Chem. Int. Ed.*, vol. 60, n° 25, p. 13783-13787, 2021, doi: 10.1002/anie.202016070.
- [165] K. Shindo, K. Masuda, H. Takahashi, Y. Arata, et I. Shimada, « Letter to the Editor: Backbone  $^1\text{H}$ ,  $^{13}\text{C}$ , and  $^{15}\text{N}$  resonance assignments of the anti-dansyl antibody Fv fragment », *J. Biomol. NMR*, vol. 17, p. 357-358, 2000, doi: <https://doi.org/10.1023/A:1008331100127>.
- [166] H. Yagi *et al.*, « Backbone  $^1\text{H}$ ,  $^{13}\text{C}$ , and  $^{15}\text{N}$  resonance assignments of the Fc fragment of human immunoglobulin G glycoprotein », *Biomol. Nmr Assign.*, vol. 9, n° 2, p. 257-260, 2015, doi: 10.1007/s12104-014-9586-7.
- [167] S. Yanaka, H. Yagi, R. Yogo, M. Yagi-Utsumi, et K. Kato, « Stable isotope labeling approaches for NMR characterization of glycoproteins using eukaryotic expression systems », *J. Biomol. NMR*, vol. 71, n° 3, p. 193-202, juill. 2018, doi: 10.1007/s10858-018-0169-2.
- [168] A. D. Gossert, A. Hinniger, S. Gutmann, W. Jahnke, A. Strauss, et C. Fernández, « A simple protocol for amino acid type selective isotope labeling in insect cells with improved yields and high reproducibility », *J. Biomol. NMR*, vol. 51, n° 4, p. 449-456, déc. 2011, doi: 10.1007/s10858-011-9570-9.
- [169] L. M. Carrillo-Cocom *et al.*, « Amino acid consumption in naïve and recombinant CHO cell cultures: producers of a monoclonal antibody », *Cytotechnology*, vol. 67, n° 5, p. 809-820, oct. 2015, doi: 10.1007/s10616-014-9720-5.
- [170] M. W. Nirenberg et J. H. Matthaei, « The dependence of cell-free protein synthesis in *E. coli* upon naturally occurring or synthetic polyribonucleotides », *Proc. Natl. Acad. Sci.*, vol. 47, n° 10, p. 1588-1602, oct. 1961, doi: 10.1073/pnas.47.10.1588.
- [171] X. Jiang, Y. Ookubo, I. Fujii, H. Nakano, et T. Yamane, « Expression of Fab fragment of catalytic antibody 6D9 in an *Escherichia coli* in vitro coupled transcription/translation system », *FEBS Lett.*, vol. 514, n° 2-3, p. 290-294, 2002, doi: 10.1016/S0014-5793(02)02383-9.
- [172] M. Stech, « Production of functional antibody fragments in a vesicle-based eukaryotic cell-free translation system », *J. Biotechnol.*, 2012.
- [173] A. Ryo, « Development of Monoclonal Antibody and Diagnostic Test for Middle East Respiratory Syndrome Coronavirus Using Cell-free Synthesized Nucleocapsid Antigen », *Front. Microbiol.*, vol. 7, p. 15, 2016.
- [174] S. Murakami, R. Matsumoto, et T. Kanamori, « Constructive approach for synthesis of a functional IgG using a reconstituted cell-free protein synthesis system », *Sci. Rep.*, vol. 9, n° 1, Art. n° 1, janv. 2019, doi: 10.1038/s41598-018-36691-8.
- [175] J. L. Dopp et N. F. Reuel, « Simple, functional, inexpensive cell extract for in vitro prototyping of proteins with disulfide bonds », *Biochem. Eng. J.*, vol. 164, p. 107790, déc. 2020, doi: 10.1016/j.bej.2020.107790.

- [176] M. Stech *et al.*, « Cell-free synthesis of functional antibodies using a coupled in vitro transcription-translation system based on CHO cell lysates », *Sci. Rep.*, vol. 7, n° 1, Art. n° 1, sept. 2017, doi: 10.1038/s41598-017-12364-w.
- [177] D. Groff *et al.*, « Engineering toward a bacterial “endoplasmic reticulum” for the rapid expression of immunoglobulin proteins », *mAbs*, vol. 6, n° 3, p. 671-678, mai 2014, doi: 10.4161/mabs.28172.
- [178] J. F. Zawada *et al.*, « Microscale to manufacturing scale-up of cell-free cytokine production--a new approach for shortening protein production development timelines », *Biotechnol. Bioeng.*, vol. 108, n° 7, p. 1570-1578, juill. 2011, doi: 10.1002/bit.23103.
- [179] X. Guo, Y. Zhu, L. Bai, et D. Yang, « The Protection Role of Magnesium Ions on Coupled Transcription and Translation in Lyophilized Cell-Free System », *ACS Synth. Biol.*, vol. 9, n° 4, p. 856-863, avr. 2020, doi: 10.1021/acssynbio.9b00508.
- [180] L. A. Ryabova, D. Desplancq, A. S. Spirin, et A. Pluckthun, « Functional antibody production using cell-free translation: Effects of protein disulfide isoinerase and chaperones », vol. 15, 1997.
- [181] C. J. Noren, S. J. Anthony-Cahill, M. C. Griffith, et P. G. Schultz, « A General Method for Site-specific Incorporation of Unnatural Amino Acids into Proteins », *Science*, vol. 244, n° 4901, p. 182-188, avr. 1989, doi: 10.1126/science.2649980.
- [182] D. Groff *et al.*, « Development of an E. coli strain for cell-free ADC manufacturing », *Biotechnol. Bioeng.*, vol. 119, n° 1, p. 162-175, janv. 2022, doi: 10.1002/bit.27961.
- [183] « Sutro Biopharma ». [https://www.sutrobio.com/wp-content/uploads/2020/03/AACR\\_NCI\\_EORTC\\_Poster-scaled.jpg](https://www.sutrobio.com/wp-content/uploads/2020/03/AACR_NCI_EORTC_Poster-scaled.jpg) (consulté le 6 avril 2023).
- [184] T. Torizawa, M. Shimizu, M. Taoka, H. Miyano, et M. Kainosho, « Efficient production of isotopically labeled proteins by cell-free synthesis: A practical protocol », *J. Biomol. NMR*, vol. 30, n° 3, p. 311-325, nov. 2004, doi: 10.1007/s10858-004-3534-2.
- [185] A. LaGuerre, F. Löhr, F. Bernhard, et V. Dötsch, « Chapter Fifteen - Labeling of Membrane Proteins by Cell-Free Expression », in *Methods in Enzymology*, Z. Kelman, Éd., in Isotope Labeling of Biomolecules - Labeling Methods, vol. 565. Academic Press, 2015, p. 367-388. doi: 10.1016/bs.mie.2015.06.001.
- [186] B. Hoffmann, F. Löhr, A. Laguerre, F. Bernhard, et V. Dötsch, « Protein labeling strategies for liquid-state NMR spectroscopy using cell-free synthesis », *Prog. Nucl. Magn. Reson. Spectrosc.*, vol. 105, p. 1-22, avr. 2018, doi: 10.1016/j.pnmrs.2017.11.004.
- [187] R. Percudani et A. Peracchi, « A genomic overview of pyridoxal-phosphate-dependent enzymes », *EMBO Rep.*, vol. 4, n° 9, p. 850-854, sept. 2003, doi: 10.1038/sj.embor.embor914.
- [188] J. Yokoyama, T. Matsuda, S. Koshiba, N. Tochio, et T. Kigawa, « A practical method for cell-free protein synthesis to avoid stable isotope scrambling and dilution », *Anal. Biochem.*, vol. 411, n° 2, p. 223-229, avr. 2011, doi: 10.1016/j.ab.2011.01.017.
- [189] J. Fiaux, E. B. Bertelsen, A. L. Horwich, et K. Wüthrich, « Uniform and Residue-specific <sup>15</sup>N-labeling of Proteins on a Highly Deuterated Background », *J. Biomol. NMR*, vol. 29, n° 3, p. 289-297, juill. 2004, doi: 10.1023/B:JNMR.0000032523.00554.38.
- [190] K. H. Gardner et L. E. Kay, « The Use of <sup>2</sup>H, <sup>13</sup>C, <sup>15</sup>N Multidimensional Nmr Gto Study the Structure and Dynamics of Proteins », *Annu. Rev. Biophys. Biomol. Struct.*, vol. 27, n° 1, p. 357-406, 1998, doi: 10.1146/annurev.biophys.27.1.357.
- [191] L. Imbert *et al.*, « In Vitro Production of Perdeuterated Proteins in H<sub>2</sub>O for Biomolecular NMR studies », p. 37.
- [192] X.-C. Su, C.-T. Loh, R. Qi, et G. Otting, « Suppression of isotope scrambling in cell-free protein synthesis by broadband inhibition of PLP enymes for selective <sup>15</sup>N-labeling and production of perdeuterated proteins in H<sub>2</sub>O », *J. Biomol. NMR*, vol. 50, n° 1, p. 35-42, mai 2011, doi: 10.1007/s10858-011-9477-5.
- [193] J. B. Eaglesham, A. Garcia, et M. Berkmen, « Chapter Five - Production of antibodies in SHuffle Escherichia coli strains », in *Methods in Enzymology*, Z. Kelman et W. B. O'Dell, Éd., in

- Recombinant Protein Expression: Prokaryotic Hosts and Cell-Free Systems, vol. 659. Academic Press, 2021, p. 105-144. doi: 10.1016/bs.mie.2021.06.040.
- [194] D. Foshag *et al.*, « The E. coli S30 lysate proteome: A prototype for cell-free protein production », *New Biotechnol.*, vol. 40, p. 245-260, janv. 2018, doi: 10.1016/j.nbt.2017.09.005.
- [195] L. Kai, V. Dötsch, R. Kaldenhoff, et F. Bernhard, « Artificial Environments for the Co-Translational Stabilization of Cell-Free Expressed Proteins », *PLoS ONE*, vol. 8, n° 2, févr. 2013, doi: 10.1371/journal.pone.0056637.
- [196] J. L. Dopp et N. F. Reuel, « Simple, functional, inexpensive cell extract for in vitro prototyping of proteins with disulfide bonds », *Biochem. Eng. J.*, vol. 164, p. 107790, déc. 2020, doi: 10.1016/j.bej.2020.107790.
- [197] L. A. Ryabova, D. Desplancq, A. S. Spirin, et A. Plückthun, « Functional antibody production using cell-free translation: Effects of protein disulfide isomerase and chaperones », *Nat. Biotechnol.*, vol. 15, n° 1, Art. n° 1, janv. 1997, doi: 10.1038/nbt0197-79.
- [198] H. Yagi *et al.*, « Backbone 1H, 13C, and 15N resonance assignments of the Fc fragment of human immunoglobulin G glycoprotein », *Biomol. Nmr Assign.*, vol. 9, n° 2, p. 257-260, 2015, doi: 10.1007/s12104-014-9586-7.

## VI. Annexes

### 1. Composition en acides aminés du mélange AAOptiCHO

Cette composition d'acides aminés a été reproduite à partir de l'article de L. M. Carrillo Cocom et al.[169]. Il s'agit de la composition en acides aminés du milieu de culture commercial pour cellules CHO nommé CD OptiCHO™.

Acides aminés	Concentrations (mg/L)	Acides aminés	Concentrations (mg/L)
Alanine	5	Leucine	42,5
Arginine	40	Lysine	30
Asparagine	12	Méthionine	20
Aspartate	/	Phénylalanine	30
Cystéine	7	Proline	40
Glutamate	17	Sérine	80
Glutamine	/	Thréonine	30
Glycine	5	Tryptophane	/
Histidine	125	Tyrosine	20
Isoleucine	38	Valine	30

### 2. Composition en acides aminés du mélange AADMEM

Cette composition d'acides aminés a été reproduite à partir de la composition du milieu de culture commercial DMEM utilisé pour les cellules HEK et qui est disponible sur le web.

Acides aminés	Concentrations (mg/L)	Acides aminés	Concentrations (mg/L)
Alanine	/	Leucine	105
Arginine	84	Lysine	146
Asparagine	/	Méthionine	30
Aspartate	/	Phénylalanine	66
Cystéine	62,6	Proline	0
Glutamate	/	Sérine	42
Glutamine	584	Thréonine	95
Glycine	30	Tryptophane	16
Histidine	42	Tyrosine	120
Isoleucine	105	Valine	95

### 3. Composition en acides aminés du mélange AAInsect

Cette composition d'acides aminés a été optimisée par NMR-Bio à partir de l'article d'Alvar Gosser et al. qui ont utilisé un mélange d'acides aminés optimisé l'incorporation d'acides aminés marqués en cellules d'insectes[168].

<b>Acides aminés</b>	<b>Concentrations (mg/L)</b>	<b>Acides aminés</b>	<b>Concentrations (mg/L)</b>
Alanine	93,4	Leucine	92
Arginine	183,4	Lysine	169
Asparagine	260	Méthionine	208
Aspartate	300	Phénylalanine	222
Cystéine	46	Proline	122,4
Glutamate	356	Sérine	118,4
Glutamine	400	Thréonine	60
Glycine	60	Tryptophane	24
Histidine	50,4	Tyrosine	90
Isoleucine	172	Valine	50,4

#### 4. Identification des signaux des leucines du fragment Fab

Superposition des spectres RMN 2D methyl-SOFAST du fragment Fab marqué U-[ $^2\text{H}$ ],  $^{13}\text{CH}_3$ -Ala, Met, Thr, Leu-proS, Val-proR, Ile- $\delta 1$  et exprimé en système cell-free (pics bleus) avec ceux de l'anticorps anti-LAMP1 marqué sur les groupements méthyles des leucines-proS et exprimé en cellules CHO (pics rouges).

