

ÉCOLE DOCTORALE DES SCIENCES CHIMIQUES

UMR7178 - IPHC

THÈSE présentée par :

Marie LEY

soutenue le : 6 juillet 2023

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité : Chimie Analytique

**Développements méthodologiques et
combinaisons d'approches de
spectrométrie de masse structurale pour la
caractérisation de complexes multi-
protéiques**

THÈSE dirigée par :

Dr. CIANFERANI Sarah

Directrice de recherche, Université de Strasbourg

RAPPORTEURS :

Pr. BECHARA Chérine

Professeure, Université de Montpellier

Dr. REDEKER Virginie

Directrice de recherche, Université Paris-Saclay

AUTRES MEMBRES DU JURY :

Pr. CAVARELLI Jean

Professeur, Université de Strasbourg

Dr. MARCOUX Julien

Chargé de recherche, Université Toulouse-III-Paul-Sabatier

Dr. MANIVAL Xavier

Directeur de recherche, Université de Lorraine

A mes parents,

A ma famille,

A Jean-Baptiste,

A ma grand-mère,

« Les aventures n'ont-elles jamais une fin ? Je suppose que non. Quelqu'un d'autre doit toujours continuer l'histoire. »

J. R. R. Tolkien, La Communauté de l'Anneau (1954)

Remerciements

Ce travail de thèse a été réalisé au sein du Laboratoire de Spectrométrie de Masse BioOrganique (LSMBO) de l'institut pluri-disciplinaire Hubert Curien de Strasbourg (IPHC, UMR 7178).

Dans un premier temps, je tiens à remercier Sarah Cianférani de m'avoir accordé l'opportunité de réaliser cette thèse au sein du LSMBO, merci de m'avoir guidé, soutenue et encadré ces quatre dernières années (et encore plus en comptant les stages). Mais, je te remercie surtout pour m'avoir fait découvrir la spectrométrie de masse structurale.

Je remercie Virginie Redeker, Chérine Bechara, Jean Cavarelli, Julien Marcoux et Xavier Manival d'avoir accepté d'évaluer mon travail de thèse.

J'adresse également mes remerciements à l'ensemble des collaborateurs, en partie cités dans ce manuscrit, avec lesquels j'ai pu échanger et travailler durant ces années.

Un grand merci à tous les membres, anciens et actuels, du LSMBO qui m'ont accueilli au laboratoire et pour l'ambiance générale :

Merci aux protéomistes, Christine C. et Christine S. (avec le XL), Laurence pour les cours de MS et de m'avoir entretenu pour un stage qui m'a permis de découvrir ce laboratoire, Agnès, Hélène, Valérianne, Delphine, Reiko, Sarahi, Fabrice B., François et Raymond

Merci à l'équipe informatique, Alex et Fabrice V pour vos nombreux dépannages et discussion à propos de driftscope et l'équipe qualité, Véronique, Magali et Stella

Merci à Martine pour les petites discussions et ton aide pour tous l'administratif

Un énorme merci à Jean-Marc pour les réparations machines, les conseils, explications, ton humour, ton pull voyant et les merveilleux beignets.

Merci aux U30, Alexandre (ne reste pas trop au labo), Pauline pour ta bonne humeur, Adrien pour tes vlogs, Noelia, Jeewan et Marie G. pour le Santa secret, Corentin pour ton franc-parler, Cyrielle, Aurélie pour ta gentillesse et ton sourire, Bastien pour les discussions sur SDA et Charline avec HP.

Remerciements

Je tiens à remercier les anciens U30 avec Justine pour le shopping et les longues histoires sur ton appart, Charlotte (ancienne de Zillis) avec le partage de la bibli et notre passion commune de Toopet et enfin Blandine qui a tout le temps été présente pendant six ans et qui a toujours su me faire rire même à des milliers de km.

Je remercie énormément la grande famille de la supramol qui a toujours été un soutien sans faille avec les anciens, Steve pour ta grande confiance en l'humanité et de m'avoir fait découvrir le XL, Antho, Thomas avec les ouvriers en cachette, Elise pour m'avoir pris sous ton aile en M2, Oscar avec tes blagues, Steph et nos discussions sur les pompiers, ta grande famille, COD et la native. Merci à Max qui a été le premier à m'enseigner l'HDX-MS et à m'avoir convaincu de faire une thèse en plus de ton amitié et de grandes discussions scientifiques.

Je tiens à remercier tout particulièrement Evo, mon duo, pour nos pauses, les cancans, les soirées, les cartons, le yodel, et enfin de tous ces moments de discussions, de rires et de doutes. Pendant trois ans, on a su évoluer et bien que tu n'étais plus au labo ma quatrième année, on a su rester en contact alors merci.

Je remercie aussi les nouveaux supras avec les Sarah(s) de reprendre la suite des aventures et bonne continuation.

Les actuels avec Jérôme, mon binôme de l'HDX, désolé pour les petites blagues surtout celle avec de l'eau et merci pour les relectures de mails et toutes les petites discussions. Merci aussi à Rania des petits cafés, repas et de m'avoir aidé pour l'impression de ce manuscrit. Et, Hugo qui est toujours prêt pour faire du bit box (crazy frog) et chanter, bon courage avec le XL-MS.

Enfin je veux remercier ma famille et mes amis, avec mes parents qui m'ont toujours soutenu avec des provisions notamment et qui ont cru en moi. Merci mamie de m'avoir poussé à continuer mes études. Merci à Chacha (d'avoir toujours compris mon état d'esprit) et Léa pour m'avoir toujours dit que je pouvais réussir. A lolo pour nos rendez-vous sur la coursive pendant la pause. Et enfin, merci à Jean-Baptiste pour absolument tout, en plus de m'avoir attendu pendant 5ans, j'en ai rajouté quatre de plus et pourtant tu as toujours su être là, me changer les idées, m'écouter et d'être venu quand je n'avais plus l'énergie de le faire.

Liste des principales abréviations

| | |
|------------|--|
| AAA+ | « ATPases associated with diverse activities » |
| ADN | Acide désoxyribonucléique |
| ADP | Adénosine diphosphate |
| AF | Fonction d'activation |
| AGC target | « Automatic Gain Control target » |
| AMPC | Adénosine monophosphate cyclique |
| ARN | Acide ribonucléique |
| ATD | Distribution des temps d'arrivée des ions |
| ATP | Adénosine triphosphate |
| BN-PAGE | « Blue native polyacrylamide gel electrophoresis » |
| BSA | Albumine de sérum bovin |
| CCS | Section efficace rotationnellement moyennée |
| CID | Dissociation induite par collision |
| CIU | Déploiement induit par collision |
| CNB | Domaine de liaison aux nucléotides cycliques |
| CoA | Co-activateur |
| CoR | Co-répresseur |
| DBD | Domaine de liaison à l'ADN |
| DDA | « Data dependant acquisition » |
| DDM | Dodécyl- β -D-maltoside |
| DEP | « Dishevelled/Egl-10/pleckstrin » |
| DLS | Diffusion dynamique de la lumière |
| DPCD | « Deleted in a mouse of primary cell dyskinesia » |
| DR | Régions désordonnées |
| DSAU | Urée diacétique disuccinimidyl |
| DSBU | Urée dibutyrique disuccinimidyl |
| DT | Tube de dérive |
| DTM | Domaine transmembranaire |
| ECD | Dissociation induite par captures d'électrons |
| EM | Microscopie électronique |
| EPAC1 | Protéine d'échange directement activée par l'AMPC |
| ERR | « Oestrogen related receptor » |
| ESI | Electronébulisation |
| ETD | Dissociation induite par transferts d'électrons |
| FAIMS | « Field asyetric ion mobility » |
| FDR | « False discovery rate » |
| FPOP | Marquage radicalaire |
| FT-ICR | Transformée de fourier – résonance cyclotronique ionique |
| GEF | Facteur d'échange de nucléotides de guanine |

Liste des principales abréviations

| | |
|--------|--|
| GPCR | Récepteur couplé aux protéines G |
| HE | Haute énergie |
| IM | Mobilité ionique |
| IMAC | Chromatographie d'affinité pour les métaux immobilisés |
| H/D | Hydrogène/deutérium |
| Hc | Chaîne lourde des anticorps |
| HCD | Dissociation induite par collision à haute énergie |
| HD | Hélice de dimérisation |
| HDX | Echange hydrogène/deutérium |
| HPLC | Chromatographie liquide à haute performance |
| LBD | Domaine de liaison des ligands |
| Lc | Chaîne légère des anticorps |
| LC | Chromatographie liquide |
| LE | Basse énergie |
| Lip | Protéolyse ménagée |
| LSD1 | Protéine diméthylase 1 spécifique des lysines |
| M-Ph | Photométrie de masse |
| MALDI | Désorption-ionisation assistée par matrice |
| MALS | Diffusion de lumière multi-angle |
| MEMHDX | « Mixed-effects model for HDX experiments » |
| MIBC | Cancer de la vessie envahissant le muscle |
| MS | Spectrométrie de masse |
| N-ter | N-terminal |
| nMS | Spectrométrie de masse native |
| OH· | Hydroxyle radicalaire |
| OMP | « Orientations of proteins in membrane » |
| PA | Purification d'affinité |
| PDB | « Protein data bank » |
| PGC1 | Coactivateur 1-alpha de PPAR γ |
| PhosB | Phosphorylase B |
| Pi | Pression à l'interface |
| PLGS | « Protein lynx global server » |
| PM | Protéine membranaire |
| PPAR | Récepteur activé par les proliférateurs de peroxysomes |
| PRMT | Protéine arginines méthyltransférase |
| PTM | Modification post-traductionnelle |
| PXR | « Pregnane X recepteur » |
| Q | Quadripôle |
| RF | Repliement de Rossmann |
| RFU | Incorporation relative en D |
| RMN | Résonance magnétique nucléaire |
| RN | Récepteur nucléaire |

Liste des principales abréviations

| | |
|----------|---|
| RRM | Motif de reconnaissance de l'ARN |
| RSF1 | Facteur d'épissage répresseur 1 |
| RXR | Récepteur X des rétinoïdes |
| SAM | S-Adénosyl-Méthionine |
| SAXS | Diffusion des rayons X aux petits angles |
| SCX | Chromatographie échangeuses de cations |
| SDS-PAGE | « Sodium dodecyl sulfate polyacrylamide gel electrophoresis » |
| SEC | Chromatographie d'exclusion stérique |
| SH3 | Domaine d'homologie 3 de SRC |
| SID | Collision induite par collision sur une surface |
| SLIM | « Structures for lossless ion manipulations » |
| SnoNRP | Petite ribonucléoprotéine nucléaire |
| SPE | Extraction sur phase solide |
| T | « trapperd » |
| TPR | Répétition tétratricopeptidique |
| TWIMS | « traveling wave ion mobility » |
| Tampon Q | Tampon d'arrêt du marquage |
| TD | « Top-Down » |
| TOF | Temps de vol |
| TW | « traveling wave » |
| UVPD | Photo-dissociation induite par ultraviolet |
| Vc | Voltage de cône |
| VDR | Récepteur à la vitamine D |
| WT | protéine non-mutée |
| X-ray | Cristallographie |
| XL | Pontage chimique covalent |

Liste des principales abréviations

Plan général

| | |
|--|------------|
| INTRODUCTION GENERALE ET OBJECTIFS | 1 |
| PARTIE 1 : DESCRIPTION DES TECHNIQUES HDX-MS ET XL-MS : PRINCIPES, DEVELOPPEMENTS ET OPTIMISATIONS | 13 |
| CHAPITRE 1 : HDX-MS : THEORIE, ETAT DE L'ART, DEVELOPPEMENTS ET OPTIMISATIONS | 15 |
| CHAPITRE 2 : XL-MS : THEORIE, ETAT DE L'ART, DEVELOPPEMENTS ET OPTIMISATIONS | 49 |
| PARTIE 2 : COMBINAISONS D'APPROCHES DE MS STRUCTURALE POUR LA CARACTERISATION DE COMPLEXES MULTI-PROTEIQUES..... | 71 |
| CHAPITRE 1 : CARACTERISATION DE COMPLEXES IMPLIQUANT DES RECEPTEURS NUCLEAIRES PAR NMS ET HDX-MS | 73 |
| CHAPITRE 2 : CARACTERISATION DE PROTEINES MEMBRANAIRES OU DANS UN ENVIRONNEMENT LIPIDIQUE PAR HDX-MS | 99 |
| CHAPITRE 3 : CARACTERISATION DE PROTEINES ARGININES METHYL-TRANSFERASES (PRMTs) PAR MS STRUCTURALE..... | 119 |
| CHAPITRE 4 : APPORT DU XL-MS POUR LA CARACTERISATION DE MACROCOMPLEXES IMPLIQUANT LES ATPASES RUVBL1/2 HUMAINES | 139 |
| CONCLUSIONS GENERALES..... | 157 |
| RÉFÉRENCES BLIBIOPHIQUES..... | 167 |

Plan détaillé

| | |
|---|-----------|
| INTRODUCTION GENERALE ET OBJECTIFS | 1 |
| I. Les approches de MS structurale sans marquage..... | 2 |
| 1. La MS native (nMS)..... | 2 |
| 2. Le couplage spectrométrie de mobilité ionique-MS (IM-MS) | 4 |
| 3. Le séquençage de protéines intactes ou top-down MS (TD-MS) | 4 |
| 4. La protéolyse ménagée suivie par MS (Lip-MS, limited proteolysis MS) | 5 |
| II. Les approches de MS structurale avec marquage | 6 |
| 1. Echange hydrogène/deutérium suivi par MS (HDX-MS) | 6 |
| 2. L'approche de pontage chimique covalent suivi par MS (cross-link MS, XL-MS)..... | 7 |
| 3. Le marquage radicalaire par oxydation (Fast Photochemical Oxidation of Proteins, FPOP)..... | 8 |
| 4. Le marquage chimique covalent ou <i>protein painting</i> | 9 |
| III. La photométrie de masse (M-Ph) : une nouvelle approche en MS structurale | 10 |
| IV. Objectifs de mon travail de thèse | 11 |
| PARTIE 1 : DESCRIPTION DES TECHNIQUES HDX-MS ET XL-MS : PRINCIPES, DEVELOPPEMENTS ET OPTIMISATIONS | 13 |
| CHAPITRE 1 : HDX-MS : THEORIE, ETAT DE L'ART, DEVELOPPEMENTS ET OPTIMISATIONS | 15 |
| I. Théorie et principe | 15 |
| 1. Echange H/D | 15 |
| 2. HDX-MS..... | 18 |
| 3. Traitement des données..... | 21 |
| II. Recommandations et innovations | 22 |
| III. Description de la méthode utilisée dans le cadre de mon travail de thèse..... | 23 |
| 1. Phase pré-analytique | 23 |
| 2. Phase analytique..... | 24 |
| 3. Phase post-analytique | 26 |
| 4. Visualisation des résultats des expériences de HDX-MS..... | 28 |
| IV. Optimisations et développements réalisés dans le contexte de cette thèse | 29 |
| 1. Méthode de deutération | 29 |
| 2. Efficacité des protéases..... | 33 |
| 2.1. Digestion de la phosphorylase B..... | 34 |

| | | |
|---|---|-----------|
| 2.2. | Digestion du pembrolizumab..... | 35 |
| 3. | Apport de la mobilité ionique en HDX-MS | 37 |
| 3.1. | Impact de l'IM sur l'identification des peptides | 38 |
| 3.2. | Impact sur le traitement de données | 40 |
| 4. | Validation statistique des données : quel outil utiliser ? | 42 |
| V. | Conclusions..... | 46 |
| CHAPITRE 2 : XL-MS : THEORIE, ETAT DE L'ART, DEVELOPPEMENTS ET OPTIMISATIONS | | 49 |
| I. | Théorie et principe | 49 |
| 1. | Agent pontant..... | 49 |
| 2. | Types de peptides pontés..... | 51 |
| 3. | Expériences de XL-MS..... | 52 |
| 4. | Traitement de données | 54 |
| II. | Recommandations et innovations | 55 |
| III. | Développements méthodologiques en XL-MS réalisés dans le cadre de mon travail de thèse..... | 57 |
| 1. | Phase pré-analytique..... | 57 |
| 2. | Phase analytique..... | 58 |
| 3. | Phase post-analytique | 60 |
| IV. | Mise au point d'un standard pour évaluer le protocole de XL-MS..... | 61 |
| V. | Optimisations et développements réalisés dans le contexte de cette thèse | 62 |
| 1. | Préparation et analyse..... | 62 |
| 2. | Traitement des données..... | 64 |
| 2.1. | Logiciels de traitements des données avec un cross-linker MS clivable | 64 |
| 2.1.1. | Principe et paramètres de MeroX | 65 |
| 2.1.2. | Principe et paramètres de MaxLynX..... | 66 |
| 2.2. | Evaluation de MeroX et MaxLynX sur trois jeux de données | 66 |
| VI. | Conclusions..... | 68 |
| PARTIE 2 : COMBINAISONS D'APPROCHES DE MS STRUCTURALE POUR LA CARACTERISATION DE COMPLEXES MULTI-PROTEIQUES..... | | 71 |
| CHAPITRE 1 : CARACTERISATION DE COMPLEXES IMPLIQUANT DES RECEPTEURS NUCLEAIRES PAR NMS ET HDX-MS | | 73 |
| I. | Introduction aux récepteurs nucléaires | 73 |

| | |
|---|-----------|
| II. Impact des mutations activatrices et inactivatrices de PPAR γ sur la relation structure/fonction | 74 |
| 1. Contexte analytique | 74 |
| 2. Objectifs en MS structurale | 75 |
| 3. Matériel et méthodes..... | 76 |
| 4. Résultats | 76 |
| 4.1. Impact de la mutation sur la liaison aux co-régulateurs..... | 77 |
| 4.1.1. nMS..... | 77 |
| 4.1.2. HDX-MS..... | 78 |
| 4.1.2.1. Optimisation de la phase pré-analytique..... | 79 |
| 4.1.2.2. Suivi par HDX-MS de l'impact de la fixation du NCoR..... | 79 |
| 4.2. Impact de la mutation ponctuelle de PPAR γ | 80 |
| 5. Valorisation..... | 83 |
| 6. Conclusions | 83 |
| III. Récepteur nucléaire ERR (<i>Oestrogen Related Receptor</i>) | 84 |
| 1. Contexte analytique | 84 |
| 2. Objectifs..... | 85 |
| 3. Matériel et méthodes..... | 85 |
| 4. Résultats | 87 |
| 4.1. Caractérisation du complexe ERR α /LSD1 | 87 |
| 4.1.1. Optimisation des phases pré-analytiques | 87 |
| 4.1.2. HDX-MS..... | 88 |
| 4.1.3. XL-MS..... | 90 |
| 4.2. Caractérisation du complexe ERR α /PGC1 α | 91 |
| 4.2.1. Optimisation des expériences de HDX-MS..... | 91 |
| 4.2.2. HDX-MS..... | 92 |
| 4.2.3. XL-MS..... | 94 |
| IV. Conclusions de l'étude des RNs et perspectives | 96 |
| CHAPITRE 2 : CARACTERISATION DE PROTEINES MEMBRANAIRES OU DANS UN ENVIRONNEMENT LIPIDIQUE | |
| PAR HDX-MS | 99 |
| I. Introduction aux protéines membranaires et HDX-MS | 99 |
| II. HDX-MS pour suivre l'effet de la fixation de ligands sur A _{2A} rant21 | 100 |
| 1. Contexte analytique | 100 |
| 2. Objectifs..... | 101 |

| | | |
|---|---|-----|
| 3. | Matériel, méthodes et optimisations des expériences en HDX-MS..... | 101 |
| 4. | Suivi de l'incorporation en D du récepteur A _{2A} rant 21..... | 102 |
| 5. | Dynamique de fixation des ligands sur A _{2A} rant21 | 103 |
| 6. | Valorisation..... | 106 |
| 7. | Conclusions et perspectives | 106 |
| III. | Etude par HDX-MS de l'interaction de EPAC1 à la membrane | 106 |
| 1. | Contexte analytique | 106 |
| 2. | Matériel, méthodes et optimisations des expériences en HDX-MS..... | 108 |
| 3. | Objectifs du projet..... | 110 |
| 4. | Etude de l'activation de EPAC1..... | 110 |
| 4.1. | Activation de EPAC1 en présence d'AMPC..... | 110 |
| 4.2. | Activation de EPAC1 en présence d'AMPC + liposome | 112 |
| 4.3. | Impact de l'inhibiteur CE3F4-R sur la conformation de EPAC1 | 114 |
| 5. | Valorisation..... | 116 |
| 6. | Conclusions et perspectives | 117 |
| IV. | Conclusions de l'étude des PMs en HDX-MS | 117 |
| CHAPITRE 3 : CARACTERISATION DE PROTEINES ARGININES METHYL-TRANSFERASES (PRMTs) PAR MS | | |
| STRUCTURALE.....119 | | |
| I. | Contexte analytique et biologique | 119 |
| II. | Objectifs | 122 |
| III. | Matériel et méthodes | 122 |
| IV. | Résultats..... | 123 |
| 1. | Caractérisation du complexe PRMT2/RSF1 | 123 |
| 1.1. | nMS..... | 123 |
| 1.2. | HDX-MS..... | 124 |
| 1.2.1. | Optimisation de la phase pré-analytique | 124 |
| 1.2.2. | Impact de la liaison de RSF1 sur PRMT2..... | 124 |
| 1.2.3. | Impact de la liaison de PRMT2 sur RSF1..... | 126 |
| 1.3. | XL-MS | 127 |
| 1.4. | Prédiction de la structure du complexe..... | 129 |
| 1.5. | Valorisation..... | 131 |
| 2. | Caractérisation des complexes impliquant PRMT9 | 132 |
| 2.1. | Complexe ternaire PRMT9/SAP49/SAP145 | 132 |
| 2.2. | Complexe binaire PRMT9/SAP145 ₄₀₁₋₅₅₀ | 132 |

| | |
|--|------------|
| 2.2.1. nMS..... | 133 |
| 2.2.2. HDX-MS..... | 133 |
| 2.2.3. XL-MS..... | 134 |
| V. Conclusions et perspectives | 137 |
| CHAPITRE 4 : APPORT DU XL-MS POUR LA CARACTERISATION DE MACROCOMPLEXES IMPLIQUANT LES ATPASES RUVBL1/2 HUMAINES | 139 |
| I. Contexte analytique et biologique..... | 139 |
| II. Objectifs | 141 |
| III. Matériel et méthodes | 142 |
| IV. Résultats | 142 |
| 1. Caractérisation du complexe R1R2..... | 142 |
| 1.1. M-Ph et nMS | 143 |
| 1.2. XL-MS | 144 |
| 2. Caractérisation du complexe R1R2/DPCD (R2D) | 145 |
| 2.1. M-Ph et nMS | 146 |
| 2.2. XL-MS de DPCD | 146 |
| 2.3. XL-MS du complexe R1R2/DPCD..... | 147 |
| 2.4. Valorisation | 150 |
| 3. Caractérisation du complexe de type R2SP | 151 |
| 3.1. M-Ph et nMS | 151 |
| 3.2. XL-MS de l'hétérodimère SP | 152 |
| 3.3. XL-MS du complexe quaternaire R2SP..... | 153 |
| V. Conclusions et perspectives | 157 |
| CONCLUSIONS GENERALES..... | 157 |
| RÉFÉRENCES BIBLIOGRAPHIQUES | 167 |

Introduction générale et objectifs

Les protéines sont des composantes de toute cellule vivante qui sont nécessaires à son bon fonctionnement. Une grande diversité de protéines et d'assemblages protéiques est déjà répertoriée et impliquée dans des machineries nombreuses et variées. Les protéines adoptent une structure 3D spécifique définie par leur séquence pour remplir une ou plusieurs fonctions biologiques, qui peuvent être modulées lorsqu'elles entrent en interaction avec des métaux, ligands, protéines, ARNs ou ADN (Fig. 1). La caractérisation de la structure et de la fonction des protéines seules ou en complexes est la clé pour comprendre les machineries biologiques¹.

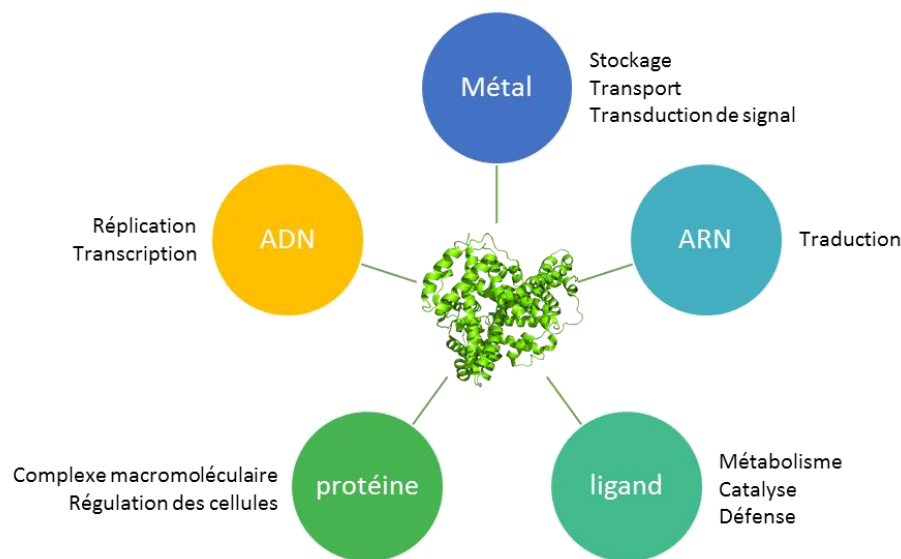


Figure 1 : Différents partenaires d'interaction des protéines ainsi que leurs fonctions.

La caractérisation physico-chimique ou biochimique de ces interactions est indispensable et peut être réalisée à plusieurs niveaux de résolution. Pour cela un large éventail de techniques classiques en biologie structurale est disponible, en particulier : (i) des approches apportant des informations sur la structure tertiaire et quaternaire des protéines avec la filtration sur gel² (taille des protéines) ou encore la diffusion des rayons X aux petits angles³ (SAXS ; conformation et assemblages protéiques); (ii) des techniques comme la fluorescence d'état stable⁴ (ou « Steady state Fluorescence ») ou encore le dichroïsme circulaire⁵ donnant des informations de dynamique et sur la structure secondaire des protéines; (iii) la cristallographie (X-ray), la résonance magnétique nucléaire (RMN) et la microscopie électronique (EM ou cryo-EM) donnant une description au niveau atomique⁶⁻⁷.

Depuis les années 90, la spectrométrie de masse (MS) s'est imposée comme une technique puissante pour la caractérisation de la structure quaternaire et de la dynamique des protéines : on parle de MS structurale⁸⁻¹⁰. Plusieurs méthodologies, établies ou émergentes, ont vu le jour ces trente dernières années, incluant : (i) des approches dites sans marquage comme la MS native (nMS), la MS « top down », la spectrométrie de mobilité ionique (IM), la protéolyse ménagée (Lip-MS) ; et (ii) des approches dites avec marquage telles que l'échange hydrogène/deutérium (HDX), le pontage chimique covalent (XL, aussi appelé « cross-linking »), le marquage radicalaire (FPOP) ou encore le « protein painting ». Chacune de ces méthodes de MS structurale apporte des informations uniques telles que la stœchiométrie, la stabilité, l'affinité, la séquence, les changements conformationnels ou encore les régions en interaction (Fig. 2). Combinées aux techniques biophysiques classiques de biologie structurale (X-ray, RMN, cryo-EM) et à de puissants outils bio-informatiques, les approches de MS structurale permettent de fournir des données complémentaires et participent à guider l'obtention de modèles structuraux.

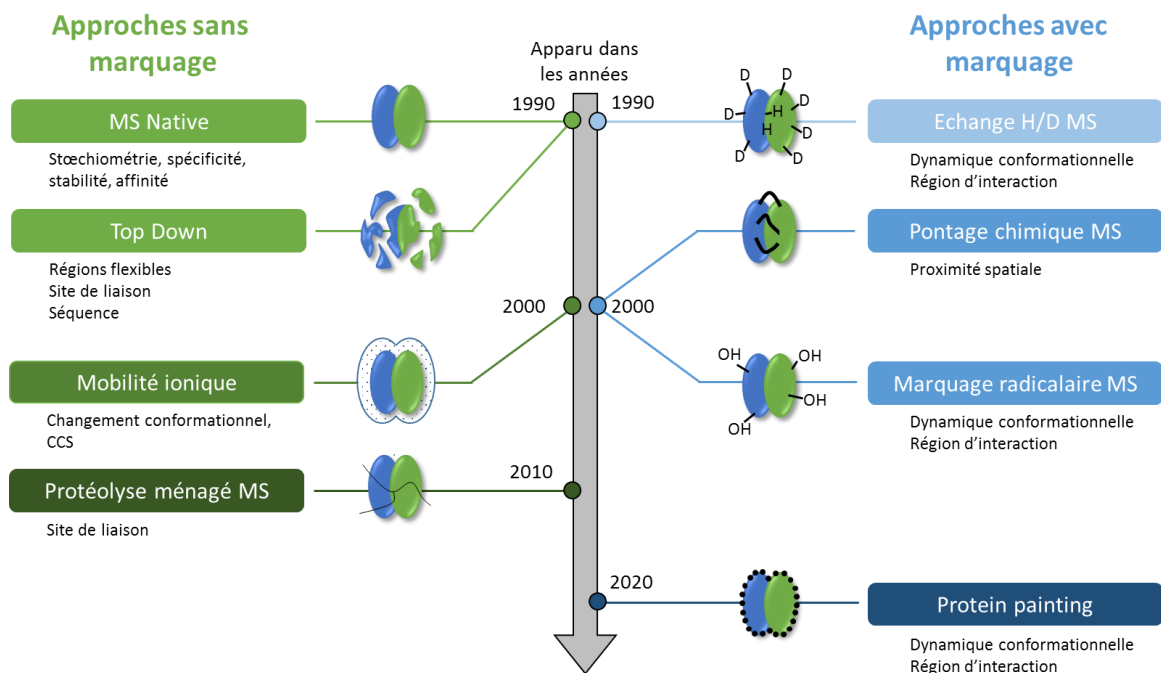


Figure 2 : Chronologie de développement des techniques de MS structurale et informations obtenues avec celles-ci.

I. Les approches de MS structurale sans marquage

1. La MS native (nMS)

Au début des années 90, la MS native a été introduite¹¹⁻¹² grâce au développement de nouvelles sources d'ionisation douces telle que l'électronebulisation (ESI)¹³ permettant

l'analyse de protéines et complexes protéiques en conditions natives. La nMS assure le maintien des interactions non-covalentes et la conservation de la structure quaternaire des protéines^{1, 14}. L'ESI génère des ions multichargés comparé au MALDI qui génère des ions monochargés, ce qui est un avantage pour l'analyse des biomolécules de haut poids moléculaires¹⁵⁻¹⁶. L'évolution de l'instrumentation MS, avec le développement des analyseurs à temps de vol (TOF), Orbitrap et à résonance cyclotronique ionique (FT-ICR), augmente encore les possibilités soit d'analyse rapide (quelques minutes) soit de complexes à haute résolution (7 000 pour le TOF, 25 000 pour l'Orbitrap et 40 000 pour le FT-ICR à 5 000-6 000 m/z)¹⁷. La nMS donne des informations de stœchiométrie d'interaction, de spécificité, stabilité et affinité de complexation (obtention de constantes d'affinité et de dissociation)¹⁸⁻¹⁹. Les instruments de types ESI-Q-TOF sont couramment utilisés pour la nMS du fait de leur résolution (10 000-60 000 FWHM)²⁰⁻²¹. De manière générale, les échantillons sont dessalés par ultracentrifugation, dialyse, filtration sur gel ou encore chromatographie d'exclusion stérique²² (SEC, Annexe 1) permettant d'échanger les tampons de purification ou stockage (composés généralement de sels non-volatils et non-compatibles avec le mode d'ionisation ESI). Le tampon de stockage est échangé avec un autre tampon, permettant de conserver la force ionique et le pH, généralement à base de sels d'ammonium (bicarbonate ou acétate d'ammonium) pour conserver l'intégrité des interactions²³ (Fig. 3). Ces dernières années, la complexité des échantillons analysés a considérablement augmenté avec des échantillons de l'ordre de plusieurs millions de Da (protéines^{15-16, 24} ou virus²⁵⁻²⁷) ou encore l'étude de protéines membranaires dans un environnement mimant les membranes²⁸.

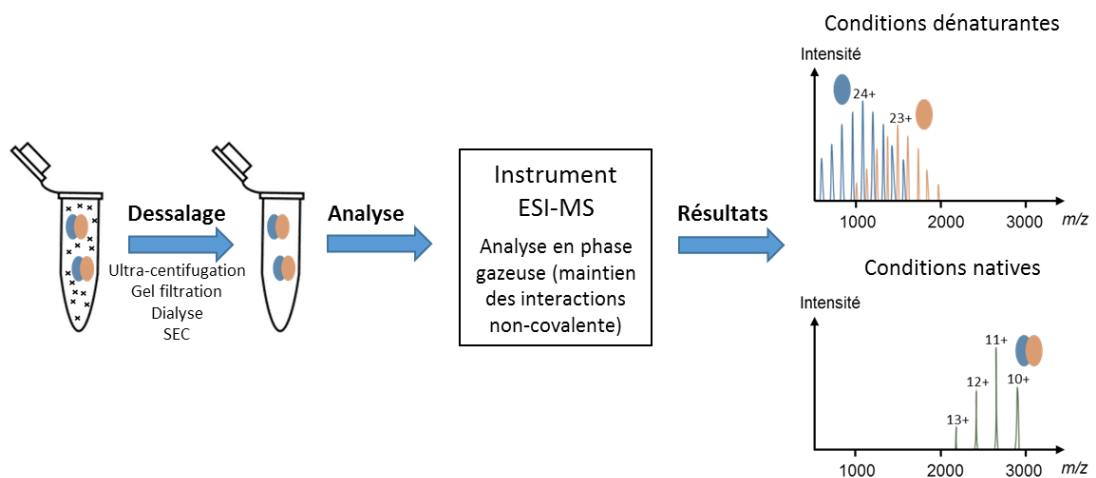


Figure 3 : Représentation des étapes de nMS et des spectres obtenus en conditions dénaturantes ou natives.

2. Le couplage spectrométrie de mobilité ionique-MS (IM-MS)

La nMS peut être couplée à l'IM dans le but d'apporter une méthode de séparation supplémentaire en fonction de la taille, conformation et de la charge de l'analyte au travers d'une cellule sous l'effet d'un gaz et d'un champ électrique²⁹⁻³⁰. Plusieurs cellules de mobilité ionique de géométries différentes sont commercialisées (Fig. 4), comme le tube de dérive (DT), les séparateurs de type « traveling wave » (TW), « trapped » (T), « Field Asymmetric ion mobility » (FAIMS)³¹ et « Structures for Lossless Ion Manipulations » (SLIM)³².

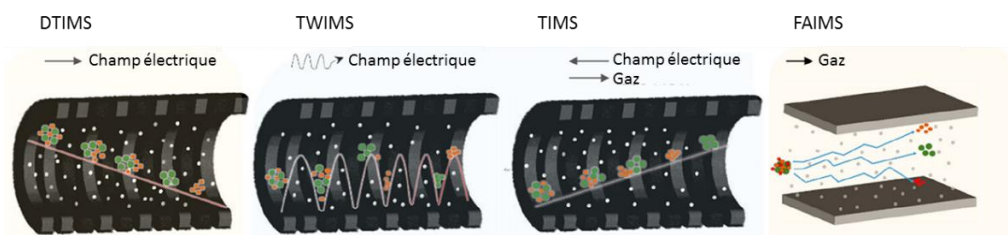


Figure 4 : Représentation des différentes géométries de cellule IM (adapté d'après³³).

La distribution des temps d'arrivée (ATD, arrival time distribution) des ions à la sortie de la cellule IM peut être convertie en section efficace rotationnellement moyennée (CCS, « collision cross section ») qui traduit la conformation des ions en phase gazeuse³⁴. Ainsi, plusieurs conformations d'une même protéine peuvent être différenciées et caractérisées via la comparaison de CCS mesurées à partir de plusieurs états de charge d'une protéine et de CCS théoriques³⁵. Récemment, des expériences de déploiement induit par collision (« collision induced unfolding », CIU) ont été développées pour suivre le déploiement progressif d'une protéine suite à l'activation des ions par collision énergétiques avec un gaz³⁶⁻³⁷. D'autre part, un nouvel instrument, la mobilité ionique cyclique (Waters) de haute résolution, a été développé permettant de caractériser des différences de conformation subtiles sur des protéines de haut poids moléculaires telles que les anticorps monoclonaux thérapeutiques³⁸.

3. Le séquençage de protéines intactes ou top-down MS (TD-MS)

L'approche TD-MS est une technique de MS structurale en plein essor ces dernières années. En effet, elle permet de séquencer les protéines intactes (sans digestion) et d'accéder en une seule analyse à la séquence de la protéine, incluant les modifications post-traductionnelles (PTMs)³⁹, telles que la phosphorylation, l'acétylation et l'ubiquitination⁴⁰. Ces modifications affectent la structure, mais également la fonction des protéines⁴¹⁻⁴². L'étude de ces PTMs est donc primordiale et le TD-MS une approche de choix, car elle donne des informations sur la

séquence et les modifications de séquence des protéines. Le TD-MS permet aussi de caractériser des zones d'interactions protéine-protéine⁴³⁻⁴⁴. Cette technique consiste en la fragmentation des protéines intactes en phase gazeuse (MS/MS) via différentes méthodes d'activation comme la dissociation induite par collision (CID), par collision à haute énergie (HCD), par transfert ou captures d'électrons (ETD/ECD), par collision sur une surface (SID)⁴⁵ ou par photodissociation par ultraviolet (UVPD)⁴⁶⁻⁴⁷. Elle peut être réalisée en conditions natives ou en conditions dénaturantes⁴⁸. Alors que la CID, SID et HCD génèrent des ions fragments du type b/y, l'ETD et l'ECD produisent des ions fragments c/z, tandis que l'UVPD produit des ions a/x, b/y et c/z par UVPD (Fig. 5)⁴⁹. De nouveaux instruments (comme avec Orbitrap Eclipse Tribrid, Thermo Fisher Scientific) combinant plusieurs méthodes de fragmentation permettent d'obtenir encore plus d'informations sur les séquences via l'augmentation de la couverture de séquence.

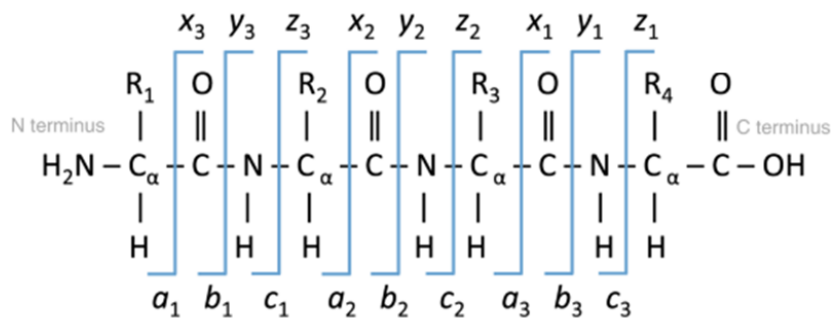


Figure 5 : Description des différents types d'ions fragments : a, b, c, x, y et z (adapté de ⁵⁰).

Le TD-MS permet le séquençage de différents types de protéines de plus en plus complexes, dont les protéines membranaires⁵¹ ou intrinsèquement désordonnées⁵². Le couplage de méthodes de séparation (chromatographie liquide, LC) des protéines en amont des analyses est optimisé, mais aussi le traitement des données⁵³⁻⁵⁴.

4. La protéolyse ménagée suivie par MS (Lip-MS, limited proteolysis MS)

La protéolyse ménagée consiste en la digestion contrôlée et non-complète d'une protéine ou d'un complexe en conditions natives (sans dénaturation) en de larges fragments par une courte exposition à des protéases⁵⁵. Ainsi, les structures tertiaires et quaternaires des protéines sont considérées comme préservées⁵⁶. Le principe de cette approche est de comparer deux états (technique différentielle) d'une même protéine et d'observer au cours du temps la digestion progressive de la protéine en fonction de l'accessibilité au solvant⁵⁷.

Différentes protéases ont déjà été testées pour la digestion ne révélant pas d'enzyme à prioriser pour ce type d'étude⁵⁸. Les fragments obtenus sont ensuite analysés en MS avec des instruments permettant de donner la masse exacte de ces larges fragments (Fig. 6)⁵⁹. Cette approche permet de révéler les zones au cœur des complexes versus les régions exposées à l'enzyme, les interfaces d'interaction, les changements conformationnels et finalement la détermination de la structure des protéines⁶⁰. Récemment, les résultats de la Lip-MS ont pu être intégrés à la génération de modélisation des structures 3D⁶¹.

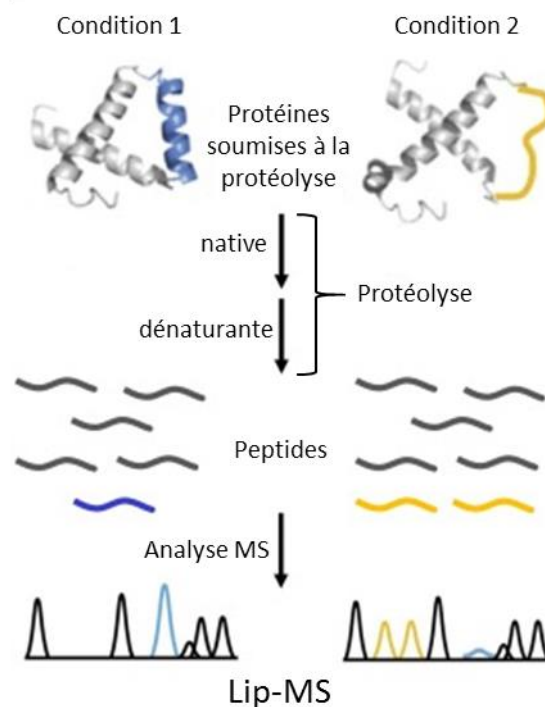


Figure 6 : Représentation du principe de la Lip-MS (adapté de ⁶²)

II. Les approches de MS structurale avec marquage

1. Echange hydrogène/deutérium suivi par MS (HDX-MS)

L'échange H/D consiste en le suivi de l'incorporation en D des protons amidiques de la chaîne principale par MS au cours du temps, pour deux états d'une même protéine⁶³⁻⁶⁴. Cette approche permet de déterminer des régions d'interactions par la comparaison des incorporations en D entre les deux états et d'observer la dynamique conformationnelle des régions exposées au solvant⁶⁵. Globalement, cette technique consiste en la deutération des protéines à différents temps puis au ralentissement (forte limitation) du marquage en abaissant la température et le pH. Les protéines sont ensuite digérées en peptides qui sont enfin analysés en LC-MS/MS (Fig. 7)⁶⁶.

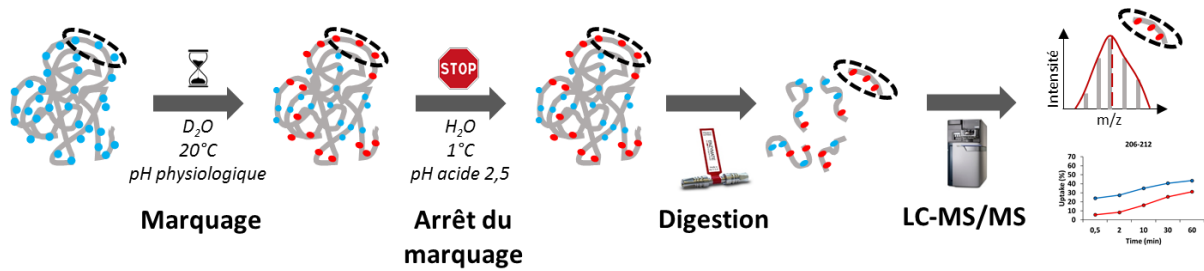


Figure 7 : Représentation schématique des étapes de l'HDX-MS.

Le nombre d'applications est en constante augmentation avec la caractérisation de différents types de complexes impliquant des protéines⁶⁷, mais aussi la comparaison avec la modélisation théorique de l'incorporation en D et des études de stabilité⁶⁸. Au vu de l'augmentation des applications et de la complexité des échantillons analysés, la validation statistique des données est très importante pour qualifier la justesse des différences d'incorporations en D⁶⁹.

2. L'approche de pontage chimique covalent suivi par MS (cross-link MS, XL-MS)

Le pontage chimique suivi par MS est une technique, apparue dans les années 2000⁷⁰, très versatile de par le nombre d'agents pontants (cross-linkers) disponibles. Les cross-linkers possèdent plusieurs caractéristiques dont : (i) deux fonctions réactives situées aux extrémités qui vont réagir avec des groupements spécifiques des chaînes latérales de certains acides aminés, (ii) leurs tailles directement liées à la distance C α -C α entre les deux résidus pontés⁷¹ et (iii) avec/sans une fonction supplémentaire permettant l'enrichissement des peptides pontés ou la fragmentation en MS/MS pour une meilleure identification des peptides pontés⁷². Cette technique consiste en le pontage chimique covalent d'un agent puis de l'arrêt de la réaction, la réduction, alkylation, digestion et enfin l'analyse en LC-MS/MS⁷³ (Fig. 8). Grâce à l'identification des peptides cross-linkés, des régions proches spatialement peuvent être déterminées donnant une idée de la conformation des protéines (cross-link intra-protéines) et/ou des zones de proximité dans le cas de l'étude d'un complexe multi-protéique (cross-link inter-protéines) mais aussi de l'accessibilité au solvant (cross-link *dead end*)⁷⁴. Le XL-MS est une technique puissante pour la caractérisation de l'architecture 3D d'une protéine seule ou en interaction. Cette approche est le plus souvent complémentaire à d'autres techniques plus classiques de biologie structurale pour la confirmation expérimentale d'un

modèle via la mesure des distances Cα-Cα entre deux résidus⁷⁵. Bien que le XL-MS est déjà largement appliqué, de nombreuses questions restent encore posées quant à la bonne identification de peptides cross-linkés, aux intervalles de distances définis pour chaque type de cross-linker et aux statistiques effectuées notamment dans le cas de l'analyse de lysats cellulaires⁷⁶, tissus⁷⁷ ou d'autres systèmes⁷⁸.

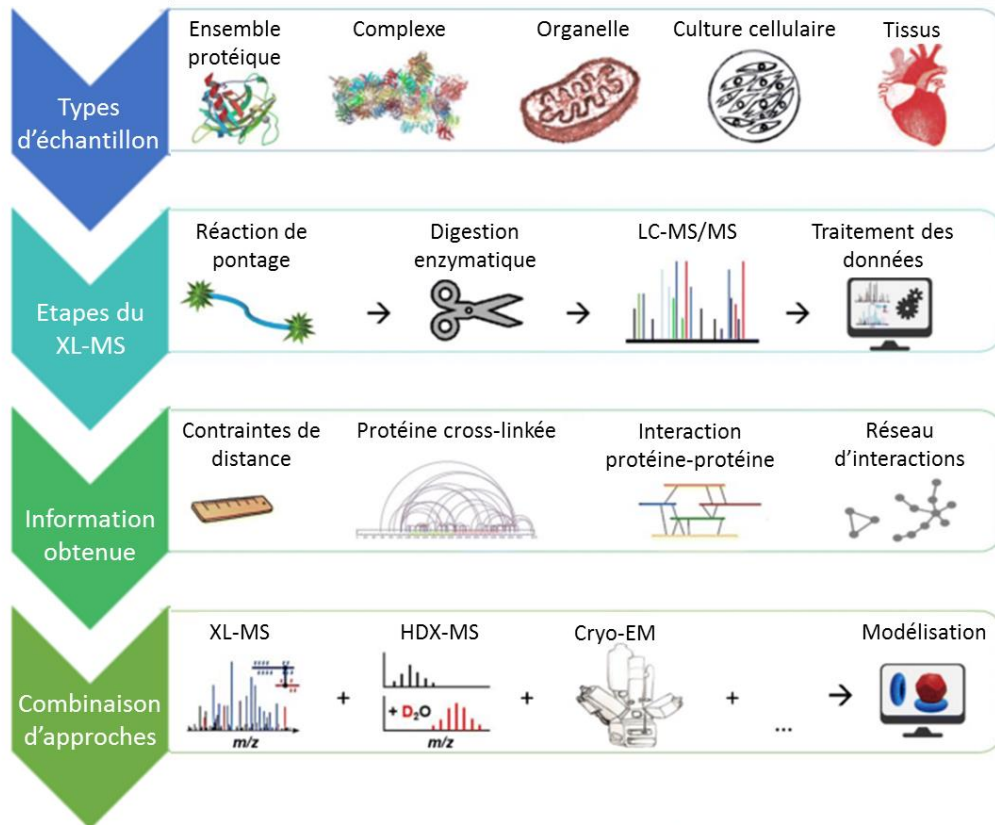


Figure 8 : Descriptif général de la stratégie XL-MS en fonction du type d'échantillon, les étapes expérimentales, les informations obtenues ainsi que l'intégration du XL-MS avec d'autres techniques de biologie structurale pour la modélisation d'une structure 3D (adapté de ⁷⁰).

3. Le marquage radicalaire par oxydation (Fast Photochemical Oxidation of Proteins, FPOP)

Le marquage radicalaire a été introduit en 1999⁷⁹ par l'oxydation des chaînes latérales des acides aminés induite par les radicaux hydroxyles (OH·) en solution aqueuse (FPOP), analysés en LC-MS/MS (Fig. 9). Cette première étude a pu montrer l'intérêt de cette approche pour la détermination de zones accessibles au solvant⁸⁰ connaissant leur taille (proche de celle de l'eau) et l'important taux d'oxydation des chaînes latérales des protéines⁸¹. Cette technique a permis d'apporter des informations de structure, repliement, dynamique des protéines et interactions protéines/protéines et protéines/ADNs⁸² jusqu'aux lysats cellulaires et cellules

intactes⁸³. Les radicaux hydroxyles peuvent être générés de différentes manières incluant la génération chimique dépendante des métaux à partir du peroxyde, la photolyse du peroxyde et la radiolyse de l'eau⁸². La chimie de Fenton ($\text{Fe}^{2+} + \text{H}_2\text{O}_2 \rightarrow \text{Fe}^{3+} + \text{OH}^- + \text{OH}^\cdot$) est la méthode la plus couramment utilisée pour former des hydroxyles⁸⁴.

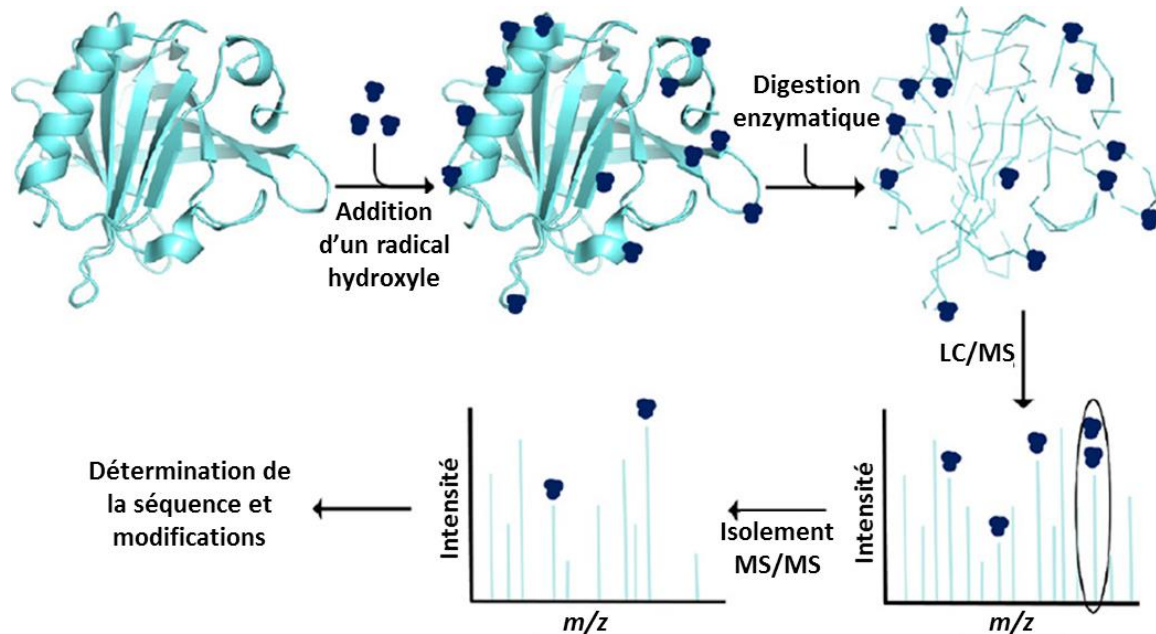


Figure 9 : Description des étapes réalisées lors du marquage radicalaire (adapté de ⁸³).

4. Le marquage chimique covalent ou *protein painting*

Le *protein painting* est une technique citée pour la première fois en 2014 qui consiste en l'étude d'interactions protéines/protéines. De petites molécules colorées (comme par exemple le *fast blue B* aussi appelé FBB ou encore le 8-Anilino-1-naphthalènesulfonic acide ou ANSA⁸⁵) sont fixées de manière covalente à la surface accessible au solvant des protéines⁸⁵⁻⁸⁶. Il s'agit d'une méthode analogue à l'échange H/D avec l'avantage ici d'avoir un marquage covalent des sites accessibles, en évitant les problèmes d'échange inverse D/H (travail en milieu acide et à 0°C). Les protéines sont ensuite dénaturées et digérées avec de la trypsine (seules les régions non marquées par les molécules colorées sont digérées par cette enzyme) puis les peptides sont analysés en MS/MS (Fig. 10). L'avantage de cette technique rapide (5min de coloration) est que les peptides non colorés proviennent uniquement de la zone d'interaction⁸⁷. De plus, cette approche ne nécessite aucun logiciel de traitement spécifique et a déjà permis l'identification de région d'interaction entre le ligand IL1 β , le récepteur IL1RI et la protéine IL1RAcP29⁸⁵, du complexe PD-1/PD-L1 et du complexe YAP2/ZO-1⁸⁸, mais aussi l'interaction de mutants de KIF1A durant l'ATPase⁸⁹.

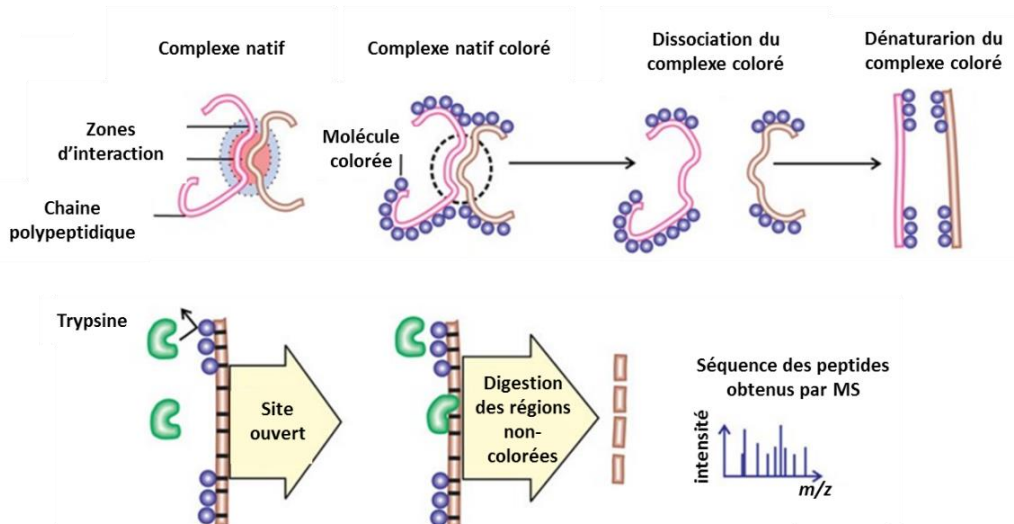


Figure 10 : Représentation schématique des étapes expérimentales du *protein painting* (d'après ⁸⁵).

III. La photométrie de masse (M-Ph) : une nouvelle approche en MS structurale

La photométrie de masse (M-Ph) est fondée sur le principe de la mesure de la diffusion de la lumière des molécules sur une surface solide (interférométrie), qui est directement corrélée à sa masse (Fig. 11)⁹⁰.

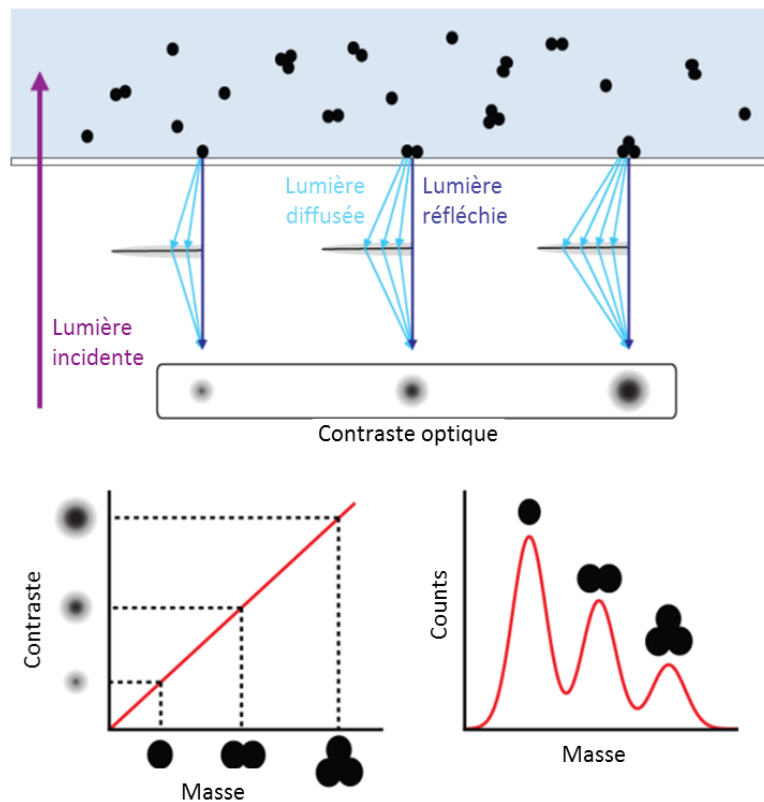


Figure 11 : Principe de la photométrie de masse, détection par la diffusion de la lumière des molécules sur la surface, corrélation entre le contraste et la masse (adapté de ⁹¹).

Un instrument de M-Ph a récemment été développé (Refeyn Two^{MP}) pouvant mesurer la masse des échantillons à une longueur d'onde de 488 nm tout en conservant l'état natif des protéines avec une précision de masse de 2% et seulement des concentrations nano- voire pico-molaires⁹⁰. La corrélation entre la lumière diffusée et réfléchiée et la masse fait que la M-Ph est un outil universel pour l'analyse de tout type de biomolécules (assemblage macromoléculaire⁹²⁻⁹³, protéines associées à la membrane sur une bicouche lipidique⁹⁴ et acides nucléiques⁹⁵). Les expériences de M-Ph donnent des informations sur la pureté, l'homogénéité des échantillons⁹², les états d'oligomérisations⁹⁶ et les affinités⁹⁷. Des analyses avec de faibles concentrations de détergents au sein de l'échantillon sont réalisables, rendant cette technique particulièrement intéressante dans le cas des protéines membranaires⁹⁸. La M-Ph serait toute à fait complémentaire, voire une alternative à la nMS pour la caractérisation de protéines ou de complexes obtenus en faibles concentrations.

IV. Objectifs de mon travail de thèse

Dans ce contexte, mes travaux de thèse se sont concentrés sur le développement, la mise en place ou la combinaison d'approches HDX-MS et XL-MS pour la caractérisation structurale de complexes impliquant différents types de protéines. Ce manuscrit s'articule en deux parties :

Une première partie se focalisera sur l'état de l'art et sur la description des développements méthodologiques réalisés dans le cadre de cette thèse en HDX-MS et XL-MS. Pour chacune des deux techniques, les optimisations réalisées seront décrites aussi bien en phase pré-analytique avec la préparation des échantillons (de l'étape de marquage/pontage à la digestion enzymatique), en phase analytique (LC-MS/MS) ou post-analytique avec le traitement des données, en tenant compte des recommandations de la communauté scientifique dans ces deux domaines.

Une seconde partie regroupe les applications de ces approches HDX-MS et XL-MS pour la caractérisation de plusieurs complexes protéiques impliquant des partenaires de natures différentes (petites molécules, peptides, protéines, ADN), de complexités croissantes (protéine/ligand, protéine/protéine, ou protéine/ADN), allant jusqu'à des masses de plusieurs centaines de kiloDaltons (kDa) mais aussi dans des environnements biologiques particuliers (mimiques de membranes). Différents types de protéines ont été étudiés avec des récepteurs nucléaires, des protéines membranaires, des protéines arginines méthyltransférase et enfin

des complexes d'assemblage impliquant des protéines AAA+ (pour *ATPases Associated with diverse Activities*). Quand pertinentes pour l'obtention d'information de stœchiométries, les approches HDX-MS et/ou XL-MS ont été complétées par des expériences de nMS et/ou de M-Ph, dans le but d'obtenir une caractérisation la plus complète possible.

Partie 1 : Description des techniques HDX-MS et XL-MS : principes, développements et optimisations

Au cours de ma thèse, j'ai principalement réalisé des expériences d'échange H/D et de pontage chimique suivi par MS sur différentes protéines et complexes protéiques de tout type. Lors de ces analyses, plusieurs optimisations ont été réalisées dans le but d'obtenir des informations structurales toujours plus complètes, précises et avérées. Dans chacun des chapitres consacrés à une des deux techniques (HDX-MS et XL-MS), je vais tout d'abord donner la théorie et le principe de l'approche avec les dernières recommandations et innovations. Puis, je vais décrire la mise en place de ces techniques au sein du laboratoire, les différents développements et optimisations réalisés. Je me suis aussi concentrée sur le traitement des données et leurs validations statistiques qui représentent un enjeu majeur pour tous les types d'analyses sachant les nombreux outils mis à notre disposition.

Chapitre 1 : HDX-MS : Théorie, état de l'art, développements et optimisations

Chapitre 2 : XL-MS : Théorie, état de l'art, développements et optimisations

Chapitre 1 : HDX-MS : Théorie, état de l'art, développements et optimisations

L'échange H/D a d'abord été décrit pour l'analyse de l'insuline en solution aqueuse par A. Hvidt et K. Linderstrøm-Lang en 1954⁹⁹. A partir des années 1980, les échanges H/D ont été suivis par RMN pour la caractérisation de petite protéine¹⁰⁰. Les premières références clés de l'approche de HDX-MS sont apparues dans les années 1990, démontrant le potentiel de la MS pour mesurer l'échange H/D au sein d'une protéine ou de complexes multi-protéiques¹⁰¹⁻¹⁰³. Depuis les trois dernières décennies, la technique de HDX-MS a été décrite, développée et appliquée à l'étude de la dynamique des protéines et de leurs interactions¹⁰⁴⁻¹⁰⁵. Le nombre d'études en HDX-MS est en constante augmentation depuis lors¹⁰⁶⁻¹⁰⁷ (Fig. 1) illustrant les bénéfices de cette technique, qui est généralement utilisée de manière différentielle pour comparer de deux états d'une même protéine¹⁰⁸. Des complexes protéine-protéine^{65, 109}, -ligand¹¹⁰ ainsi que les protéines membranaires¹¹¹, pharmaceutiques¹¹² et désordonnées¹¹³⁻¹¹⁴ sont couramment étudiés en HDX-MS en complémentarité avec d'autres techniques classiques de biologie structurale¹⁰⁷.

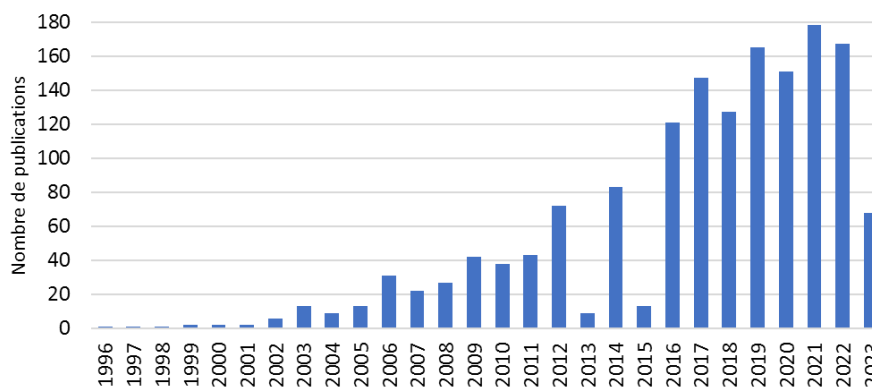


Figure 1 : Représentation du nombre de publications référencées dans PubMed® ayant pour mots clés : HDX-MS.

I. Théorie et principe

1. Echange H/D

Seuls les échanges H/D des H amidiques de la chaîne principale des protéines peuvent être mesurés en MS, les H covalentement liés aux atomes de carbone n'échangeant pas et ceux liés aux hétéroatomes des chaînes latérales échangeant trop rapidement pour être mesurés en

MS ainsi que des liaisons H en N-terminal (N-ter) (Fig. 2)¹⁰⁵. Les prolines ne possédant pas d'H au niveau de la chaîne principale, aucun échange n'est possible.

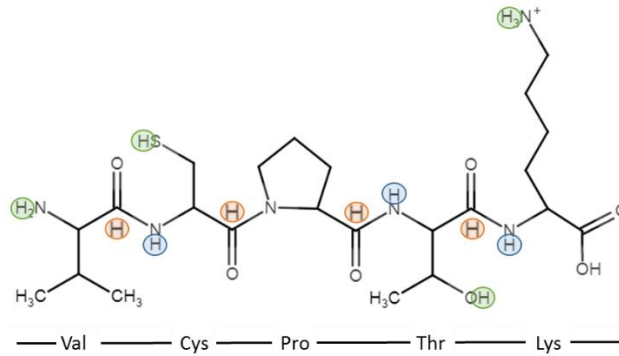


Figure 2 : Représentation des différents types d'atomes H échangeables en HDX-MS du peptide (VCPTK). Les H non-échangeables sont représentés en orange, échangeant trop rapidement pour être mesurés en vert et mesurables en bleu.

L'échange H/D peut être catalysé par un acide, une base ou de l'eau, les H s'échangent alors suivant différents mécanismes¹⁰⁴ (Fig. 3).

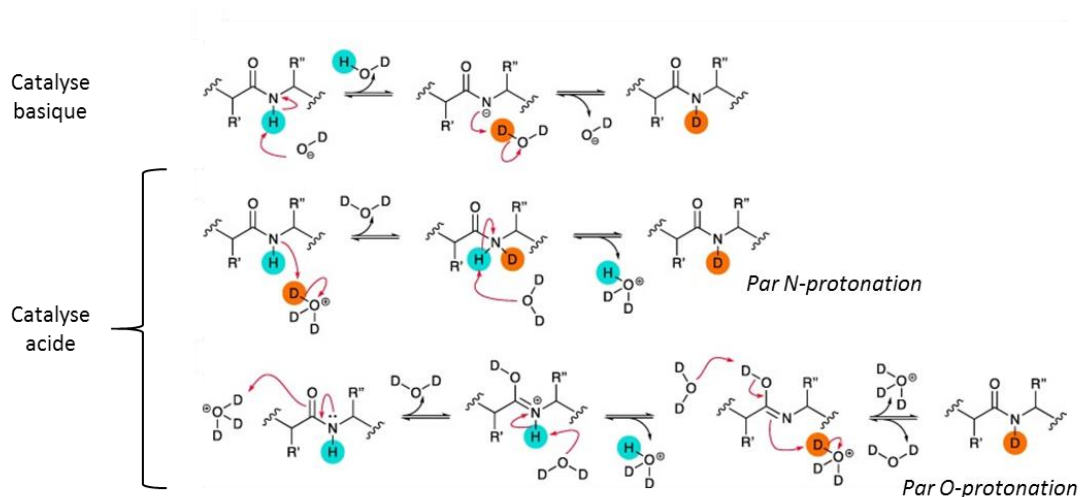


Figure 3 : Mécanismes d'échange H/D dépendant de la catalyse (d'après ¹⁰⁴).

Le taux d'échange est directement dépendant du pH, tandis que la catalyse acide est réalisée à un pH inférieur à 3, un pH supérieur à 3 entraîne une catalyse basique. Ce taux est donné par l'équation 1 où k_{CH} est la constante de vitesse de l'échange, k_{acide} , k_{base} sont les coefficients de vitesse pour l'échange catalysé par l'acide et la base et k_{eau} est le coefficient intrinsèque du transfert de proton catalysé par l'eau¹⁰⁴.

$$\text{Equation 1 : } k_{CH} = k_{acide}[H_3O^+] + k_{base}[OH^-] + k_{eau}[H_2O]$$

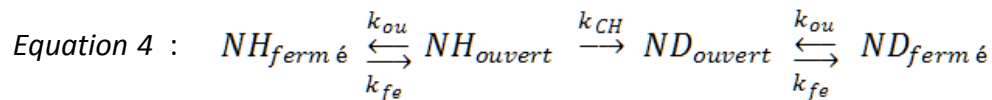
La cinétique d'échange dépend également de la structure des protéines et de leurs repliements suite à la formation de liaison hydrogène notamment. Cette cinétique est décrite

par l'équation 2 où la constante de vitesse observée (k_{obs}) dépend des états repliés (k_f) et déplié (k_u) de la protéine. Les états repliés sont définis par la probabilité de l'échange (β) suivant la constante de vitesse de l'échange (k_{CH}), équation 3¹¹⁵.

$$\text{Equation 2 : } k_{obs} = k_f + k_u$$

$$\text{Equation 3 : } k_f = \beta k_{CH}$$

Plusieurs états sont donc considérés : (i) lorsque l'échange H/D est possible l'hydrogène est dans un état « ouvert » (NH_{ouvert}), accessible au solvant ; et (ii) inversement si l'échange n'est pas possible alors l'état est « fermé » ($NH_{fermé}$). Sachant que les protéines ne sont pas figées en solution, des transitions d'un état ouvert à fermé sont considérées dans le modèle représenté par l'équation 4 décrivant les constantes d'ouverture (k_{ou}) et de fermeture (k_{fe}) ainsi que la constante de vitesse de l'échange (k_{CH}). Ce modèle permet d'évaluer l'échange des H protégés¹¹⁶.



Ainsi une constante traduisant de l'échange globale (k_u) prenant en compte tous ces facteurs a pu être déterminée avec l'équation 5. Sachant que l'échange est réalisé en condition native alors $k_{fe} \gg k_{ou}$ permettant de simplifier la constante globale (k_u) par l'équation 6¹¹⁷.

$$\text{Equation 5 : } k_u = \frac{k_{ou} \times k_{CH}}{k_{ou} + k_{fe} + k_{CH}}$$

$$\text{Equation 6 : } k_u = \frac{k_{ou} \times k_{CH}}{k_{fe} + k_{CH}}$$

À partir de cette dernière équation, deux évènements sont envisageables et décrits comme des cinétiques de types EX1 ou EX2¹¹⁸ (fig. 4):

- Dans le cas de la cinétique EX1, l'échange s'effectue plus rapidement que le repliement de la protéine ($k_{fe} \ll k_{CH}$) signifiant un échange complet à chaque ouverture. En MS, une distribution bimodale sera alors observée.
- Dans le cas de la cinétique EX2, l'échange s'effectue plus lentement que le repliement de la protéine ($k_{fe} \gg k_{CH}$) signifiant que les régions structurées et non accessibles ne seront pas échangées à des temps courts d'exposition, mais au bout de plusieurs minutes, heures ou jours.

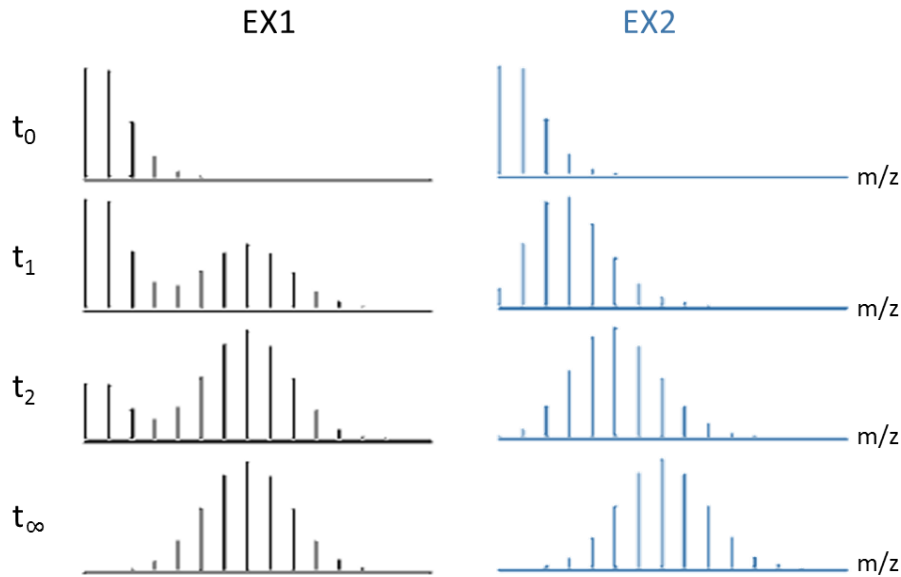


Figure 4 : Distributions isotopiques théoriques correspondant aux cinétiques EX1 et EX2 à différents temps (adapté de ¹¹⁹).

Dans des conditions natives, la cinétique EX2 est généralement observée, mais dans le cas, par exemple, de protéines intrinsèquement désordonnées une cinétique EX1 est également remarquable ainsi que des régimes mixtes EX1/EX2. Il n'est toutefois pas possible de déterminer si les distributions bimodales sont dues à des cinétiques EX1 ou à de l'échange inverse, c'est-à-dire l'échange des D en H ou encore des chevauchements de profils isotopiques de deux ions différents.

La vitesse d'échange (k_{CH}) est, en plus du repliement/dépliement et du pH, directement corrélée à la température. En effet, diminuer la température de 25°C à 0°C abaisse le taux d'échange des hydrogènes d'environ 14 fois¹⁰⁴.

2. HDX-MS

Deux types de deutération sont rapportés dans la littérature : le marquage pulsé et continu (plus couramment utilisé). Le marquage pulsé consiste en l'ajout d'un agent perturbateur (chimique, température, pH) impliquant le déploiement des protéines. Les échantillons sont ensuite marqués par impulsions (de l'ordre de la seconde) et seuls les H accessibles au solvant sont deutérés. Le marquage continu, quant à lui, consiste en la deutération des protéines à l'aide de D₂O (de l'ordre de la minute, heure)¹⁰⁵.

Dans les deux cas, le marquage est ensuite stoppé (*quench*) en baissant la température (de 20°C généralement à 0°C) et le pH (d'un pH physiologique à un pH de 2,5) ce qui permet de limiter l'échange inverse D/H⁶⁸ (Fig. 5).

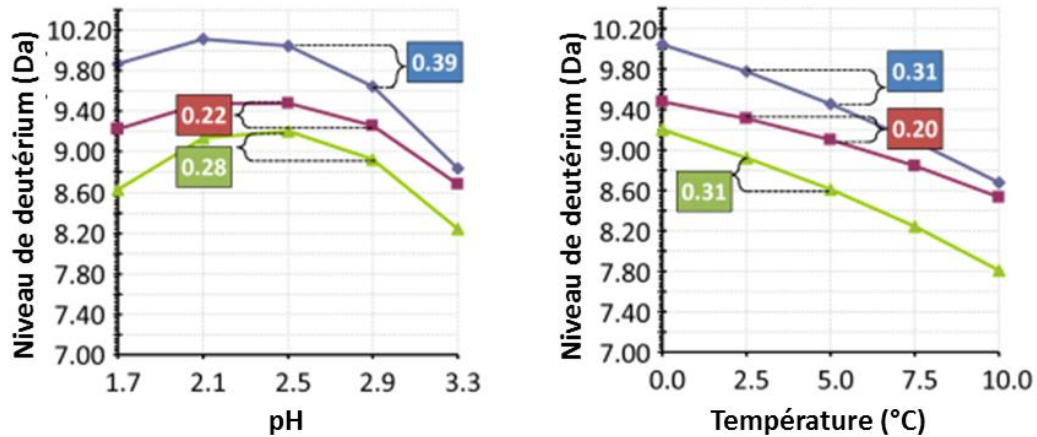


Figure 5 : Niveau de deutérium (Da) en fonction du pH et de la température (°C) de trois peptides, MYSLCEQTVNFK en bleu, QCSVFMTNYEKL en rouge et IHGASDFWVWER en vert (adapté de ⁶⁸).

L'analyse qui suit l'échange H/D peut se faire soit au niveau de la protéine, soit au niveau du peptide. On distingue plusieurs méthodes d'analyses MS : le bottom-up (dénaturation, digestion et fragmentation par CID), top-down (pas de digestion et fragmentation par ETD) et middle-down (digestion et fragmentation par ETD) HDX-MS donnant des informations à différents niveaux de résolution (Fig. 6). La stratégie bottom-up HDX-MS est la méthode la plus couramment utilisée permettant d'obtenir l'incorporation en D au niveau de chaque peptide identifié.

La digestion des protéines doit être effectuée par des protéases actives à pH acide en solution ou immobilisée sur une cartouche. De plus, les enzymes de digestion utilisés en HDX-MS sont non-spécifiques dans le but d'obtenir le plus grand nombre de peptides sur la plus grande proportion de la séquence. La pepsine, l'aspergillopepsine, la zhizopuspepsine, la nepenthesine I et II ont déjà été utilisées en HDX-MS¹²⁰⁻¹²⁴.

De manière générale, des colonnes C18 sont employées pour la séparation des peptides bien que des colonnes C4 ou C8 le sont également pour l'étude de protéines membranaires notamment¹²⁵. Le système chromatographie doit être adapté aux analyses de peptides deutérés avec une température basse (proche de 0°C) pour limiter l'échange inverse D/H. Deux types d'instruments MS sont aussi couramment utilisés, le ESI-Q-TOF ou ESI-Q-Orbitrap.

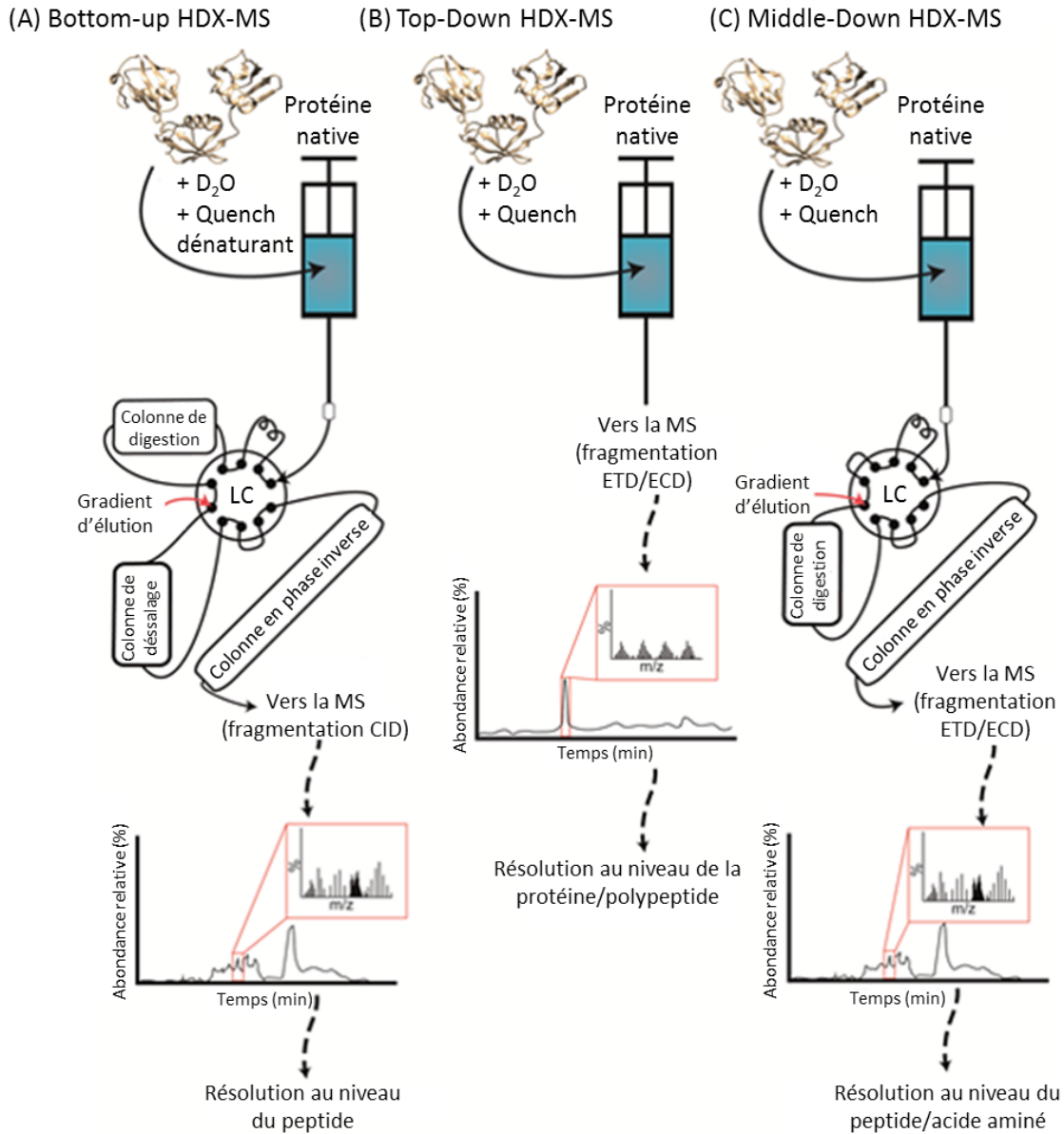


Figure 6 : Représentation schématique des étapes des trois méthodes de HDX-MS, bottom-up, top-down et middle-down (adapté de ¹¹⁶).

Finalement, l'incorporation en D est mesurée à partir du centroïde des intensités des pics (distribution isotopique déterminée à un temps de donné) par rapport au centroïde du temps zéro (t_0) c'est-à-dire sans marquage (distribution isotopique naturelle du peptide). L'incorporation en D relative ($U_{(\%)}$) peut être calculée suivant l'équation 7 prenant en compte l'incorporation absolue en Dalton ($U_{(Da)}$) et l'incorporation maximale du peptide ($maxU$).

$$Equation\ 7 : U_{(\%)} = \frac{U_{(Da)}}{maxU} \times \%D$$

L'incorporation maximale est définie par le nombre de protons amidiques disponibles de la chaîne principale du peptide (en général : nombre de résidus du peptide, excepté la proline,

moins un correspondant au proton N-terminal). Enfin le pourcentage en D ($\%_D$) est déterminé en fonction du taux estimé de D lors du marquage.

3. Traitement des données

Les expériences de HDX-MS deviennent de plus en plus complexes (nombre de protéines étudiées, taille des complexes) et lourdes en termes de taille des données collectées. Voilà pourquoi, le nombre de logiciels a considérablement augmenté ces dernières années. De plus, le traitement des données en HDX-MS peut se diviser en plusieurs étapes avec : (i) l'identification des peptides, (ii) le traitement des profils isotopiques et enfin (iii) l'application de tests statistiques¹¹⁶ (Tableau 1). La différence entre tous ces logiciels est principalement due aux types d'instruments utilisés (Orbitrap, TOF, etc.) pour l'identification et le traitement, aux tests statistiques employés pour la validation ou encore la prise en compte des distributions bimodales (EX1) avec des méthodes de déconvolutions spécialisées. D'autres logiciels sont également disponibles et rapportés dans un article¹²⁶ montrant les différences entre les méthodes de détermination de l'incorporation en D.

| Logiciel | Identification des peptides | Traitement des profils | Statistiques | Visualisation des données | Distribution bimodale |
|--|-----------------------------|------------------------|--------------|---------------------------|-----------------------|
| Protein Lynx Global Server (Waters) | x | | | | |
| PaSER (Brucker) | x | | | | |
| Proteome Discoverer (Thermo Fisher Scientific) | x | | | | |
| MASCOT ¹²⁷ | x | | | | |
| Comet ¹²⁸ | x | | | | |
| HDX Workbench ¹²⁹ | x | x | | x | |
| Mass Spec Studio v2 ¹³⁰ | x | x | | x | |
| ExMS2 ¹³¹ | x | x | x | x | x |
| DynamX (Waters) | | x | | x | |
| HX-Express 2 ^{117, 132} | | x | | x | x |
| HDExaminer (Sierra Analytics) | | x | x | x | |
| deMix ¹³³ | | x | x | x | x |
| MEMHDX ¹³⁴ | | | x | x | |
| Deuteros ¹³⁵⁻¹³⁶ | | | x | x | |
| DECA ¹³⁷ | | | x | x | |
| HaDeX ¹³⁸ | | | | x | |
| HDX-viewer ¹³⁹ | | | | x | |
| MSTools ¹⁴⁰ | | | | x | |

Tableau 1 : tableau décrivant les différents logiciels utilisés en HDX-MS avec leurs spécificités.

II. Recommandations et innovations

De récentes recommandations en HDX-MS ont été publiées dans le but d'harmoniser la description des sets de données HDX-MS, l'interprétation des résultats et d'améliorer la transparence des analyses en HDX-MS pour la communauté scientifique⁶⁶. Ainsi, certaines consignes ont été mises en place, telles que :

- La mise en place d'un contrôle de l'expérience (avec une protéine modèle).
- La réalisation de plusieurs temps de deutération (six comprenant le t_0 , sans marquage).
- Le nombre de réplicats à effectuer pour considérer les résultats comme répétables (3 au minimum).
- L'application d'un seuil pour valider des différences d'incorporation en D significatives sur la base d'une mesure quantitative et répétable (p -value).
- La transparence avec la mise à disposition des données via le référentiel PRIDE¹⁴¹.
- La réalisation de mesures de l'échange inverse D/H du système instrumental ou de la correction de la mesure d'incorporation en D du système protéique étudié.

Les dernières années ont vu émerger de nombreuses optimisations et développements en HDX-MS pour la caractérisation de complexe impliquant des protéines, avec :

- Une nouvelle colonne de digestion avec de la pepsine et de la nepenthesine II co-immobilisée a été commercialisée par Affipro. Cette colonne présente un avantage certain pour l'analyse de protéines membranaires pour améliorer la couverture de séquence et la redondance (nombre moyen de fois où un acide aminé est recouvert par au moins un peptide)¹²⁵. Sachant que le bénéfice de l'utilisation de deux enzymes co-immobilisées pour la digestion des peptides en HDX-MS a déjà été montré avec une colonne contenant de la pepsine et de l'aspergillopepsine¹⁴².
- La séparation des peptides en dessous de 0°C permettant de limiter davantage l'échange inverse D/H. Cette séparation est tout à fait compatible avec le matériel couramment utilisé (colonne C18 HPLC, chromatographie liquide haute performance) et n'entraîne pas de changement de l'efficacité de la séparation (jusqu'à -20°C). Afin d'éviter le gel de la phase mobile (composé d'ACN ou H₂O) à cette température, du méthanol (35%) est ajouté¹⁴³⁻¹⁴⁴.
- La fragmentation par l'ETD, au lieu de la CID ou HCD classiquement utilisées, a aussi été optimisée dans le but d'obtenir une résolution à l'acide aminé près et éviter la

proportion de *scrambling* (migration des D sur d'autres protons amidiques disponibles du peptide causée par la collision avec la phase gazeuse)¹⁴⁵⁻¹⁴⁶.

- Des expériences de HDX-MS ont été réalisées en 2022 sur un nouvel instrument, le SELECT SERIES Cyclic IMS (Waters) montrant les bénéfices de l'implémentation de l'IM cyclique pour l'identification de peptides par la réduction de la saturation des pics des peptides abondants et l'amélioration de la sensibilité. De plus, le constructeur a montré que la durée de la chromatographie peut être réduite, et donc d'augmenter le débit d'analyse et d'éviter davantage les échanges inverses.
- Un nouveau système entièrement automatisé a été développé pour réaliser des expériences de HDX-MS avec une résolution de temps de l'ordre de la milliseconde (50ms - 300s)¹⁴⁷. Ce système permet d'obtenir des cinétiques d'incorporation en D de régions dynamiques (boucles exposées, protéines intrinsèquement désordonnées) qui échangent trop rapidement pour être mesurées avec les méthodes conventionnelles¹⁴⁷⁻¹⁴⁸.
- La reproductibilité des données reste un enjeu majeur en HDX-MS, pour cela de nombreux groupes cherchent à automatiser la préparation des échantillons avec principalement le robot HDX-3 PAL™ (Leap) d'une nouvelle génération, mais aussi avec des étalons internes pour corriger tous décalages au cours des analyses¹⁴⁹⁻¹⁵⁰. La validation statistique des données, encore peu employée, permet également de prendre en compte la reproductibilité pour définir des différences d'incorporation en D statistiquement significative¹³⁴.
- Enfin, l'intégration des expériences de HDX-MS avec la simulation de la dynamique moléculaire est à suivre dans les prochaines années connaissant les difficultés d'interprétations des données expérimentales dans certains cas comme les échanges de type EX1¹⁵¹⁻¹⁵².

III. Description de la méthode utilisée dans le cadre de mon travail de thèse

1. Phase pré-analytique

Classiquement, l'échange H/D s'effectue dans notre cas par le mélange de 3µL de(s) protéine(s) de concentration définie avec 57µL de tampon à base de D₂O grâce au robot PAL LTL (Leap technology) (Fig. 7). Ce dernier est généralement proche du tampon de stockage de

la protéine, composé par exemple de HEPES, K_2HPO_4/KH_2PO_4 , Tris-HCl, etc., mais aussi des sels (NaCl, KCl, $MgCl_2$) avec des détergents et/ou des agents réducteurs en faible concentration (DDM, TCEP, etc.). Le mélange de la protéine avec le tampon deutéré (à base de D_2O) s'effectue à $20^\circ C$ pendant un temps défini de 0,5 à 60min habituellement. Pour stopper le marquage (*quench*) et limiter l'échange inverse D/H, 55 μL de ce mélange est ensuite ajouté à 55 μL de tampon d'arrêt du marquage (Tampon Q) à un pH final acide de 2,5 et une température de $1^\circ C$. En effet, la diminution de la température et du pH diminue les cinétiques d'échanges inverses D/H. Le tampon Q est majoritairement composé d'un agent dénaturant (Guanidine H-Cl ou urée) permettant de faciliter la digestion des protéines, en peptides. Finalement, 80 μL de cette solution sont injectées dans la boucle d'injection de 50 μL . Ces expériences sont répétées au minimum trois fois pour chaque condition (états, temps de deutération) conformément aux recommandations⁶⁶.

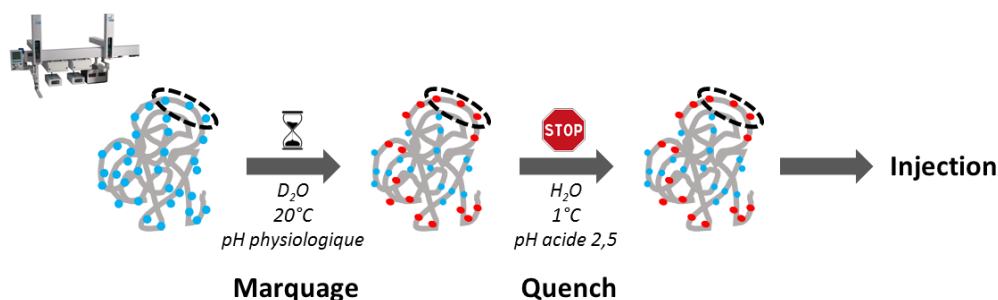


Figure 7 : Description schématique de la phase pré-analytique en HDX-MS avec les étapes de deutération, d'arrêt du marquage/*quench* et d'injection, toutes effectuées par le robot PAL LTL.

2. Phase analytique

Les protéines sont ensuite digérées en ligne (Fig. 8) avec de la pepsine immobilisée sur une cartouche (Enzymate™ BEH Pepsin Column, 2.1x30mm, taille des particules 5 μm , Waters) à $20^\circ C$ dans le module nanoACQUITY UPLC avec HDX technology (Waters)¹⁵³. Le solvant, composé d'eau avec 0,1% d'acide formique (HCOOH), sert au passage des protéines au travers de la cartouche à un débit de 50 à 200 $\mu L/min$. Les peptides sont ensuite concentrés sur une pré-colonne de phase inverse C18 (ACQUITY UPLC BEH C18 VanGuard Pre-column, 2,1x5mm, 1,7 μm , Waters) à $0,1^\circ C$. Ils sont ensuite élués de la pré-colonne et séparés sur une colonne analytique (Colonne ACQUITY UPLC BEH C18, 1x100mm, 1,7 μm , Waters) avec un gradient d'éluion à un débit de 40 $\mu L/min$ composé du solvant A (0.1% de HCOOH, eau) et du solvant B (0.1% de HCOOH, acétonitrile) de 2-40% de B sur 7 min suivi d'une rampe de 0,5 min à 85% de B.

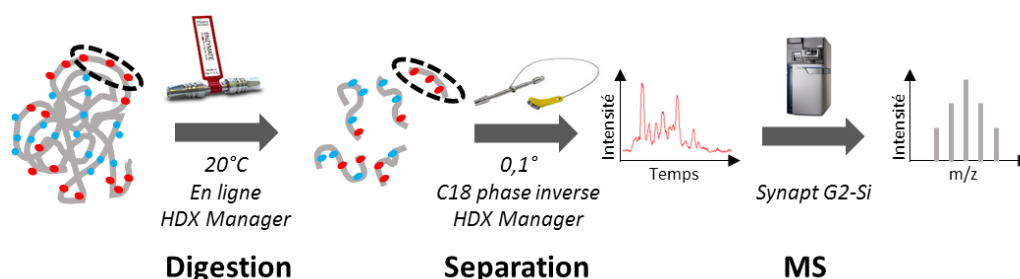


Figure 8 : Description schématique de la phase analytique en HDX-MS avec les étapes de digestion, séparation et analyse MS effectuée à l'aide de la nanoACQUITY UPLC avec HDX technology et du Synapt G2-Si HDMS (Waters).

Les peptides élués sont finalement injectés dans le spectromètre de masse (Fig. 9), le Synapt G2-Si HDMS (Waters), comprenant :

- Une source électrospray de type Z-spray dont la géométrie permet l'élimination des neutres et la transmission des ions avec un débit de N₂ de 90L/h au niveau du cône d'échantillonnage et 600L/h pour la désolvatation.
- Un Stepwave™ composé de deux séries de lentilles superposées avec un flux continu d'azote entraînant les espèces neutres alors que les ions sont guidés au niveau des lentilles supérieures via différents voltages appliqués.
- Un analyseur du type quadripôle utilisé en mode *RF only*, sans sélection des ions.
- Un Triwave™ composé lui-même de trois cellules appelées « Trap, TWIMS, Transfer ». La cellule Trap est primordiale pour l'analyse des ions avec un mode d'acquisition indépendant des données et plus précisément en mode MS^E. Cette méthode consiste en l'application de scans de basses (4 eV) et hautes énergies (15-40 eV) permettant (i) dans le premier cas, l'acquisition des signaux correspondant aux ions précurseurs (MS) où la cellule guide les ions ; et (ii) dans le second cas, la fragmentation des ions parents (MS/MS) par un processus de dissociation induite (CID). La cellule de mobilité ionique (TWIMS) se comporte comme un guide d'ion lorsqu'elle n'est pas activée comme celle de Transfer.
- Un analyseur à temps de vol (TOF) séparant les ions en fonction de leur temps de vol au sein de l'analyseur décrivant une forme de « V » car utilisé en mode de résolution (défini par le constructeur) dans le cadre de nos expériences.
- Un détecteur multiplicateur d'électrons QuanTOF™.

L'acquisition est réalisée sur la gamme de m/z de 500 à 2000 et le Glu-fibrino peptide (calibrant interne) est infusé toutes les 30 secondes pour l'étalonnage et la correction de la masse en mode ion positif.

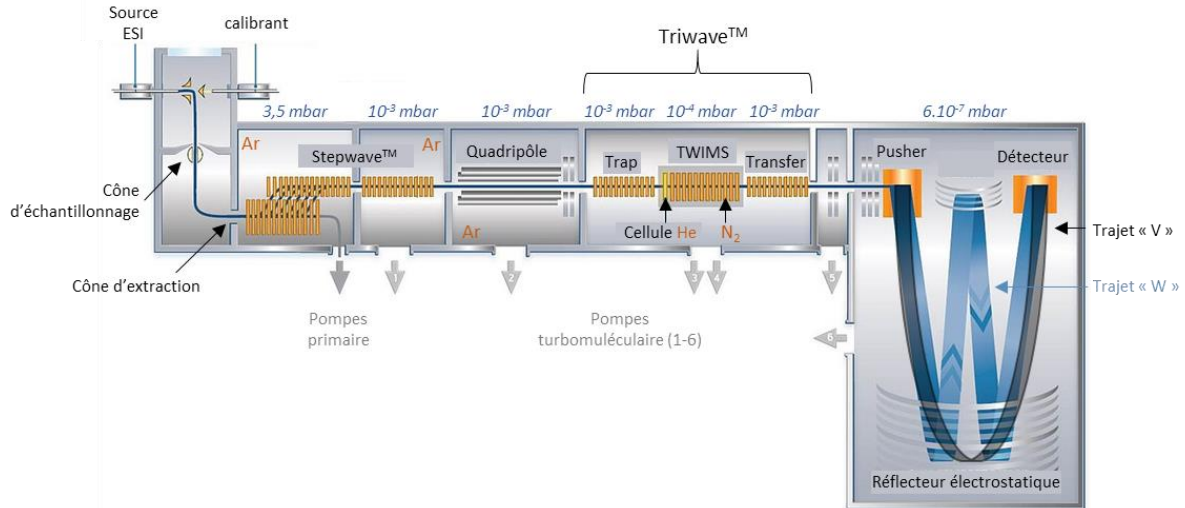


Figure 9 : Représentation schématique du Synapt G2-Si HDMS (Waters) d'après ¹⁵⁴.

3. Phase post-analytique

Une fois les spectres obtenus, le traitement des données s'effectue via l'utilisation de plusieurs logiciels (Fig. 10), dans le but :

D'identifier les peptides avant deutération par le logiciel *Protein Lynx Global Server 2.5.3* (PLGS, Waters). Les valeurs de m/z des scans MS et MS/MS sont comparées à celles obtenues par une digestion *in silico* d'une banque de données composée des séquences des protéines étudiées ainsi que la séquence de la protéase. Plusieurs filtres sont appliqués lors de la recherche avec notamment un FDR de 100% connaissant la pureté de l'échantillon et le nombre insuffisant de peptide pour le calcul d'un FDR, une digestion non-spécifique et une oxydation de la méthionine comme modification variable.

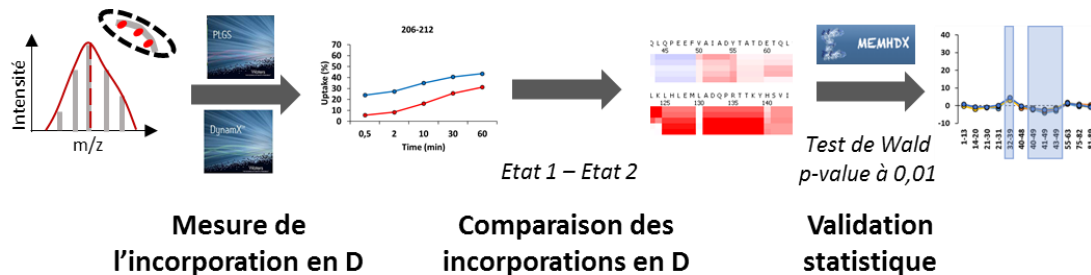


Figure 10 : Description schématique de la phase post-analytique en HDX-MS avec les étapes de mesure de l'incorporation en D et la comparaison de ces incorporations à l'aide des logiciels PLGS et DynamX ainsi que la validation statistique des données effectuée avec MEMHDX.

De déterminer à l'aide de DynamX 3.0 (Waters) les recouvrements de séquences, redondances et nombre de peptides identifiés à partir des données générées par PLGS. Lors de cette étape, différents filtres sont appliqués, dont une intensité minimale de 1000, la présence des peptides dans au moins 2 des 3 réplicats et un minimum de « produit par acide aminé » de 0,3 (pour un peptide de 10 acides aminés au moins trois ions fragments doivent être identifiés). Ce logiciel nous permet plus particulièrement d'effectuer les calculs et les recherches des profils isotopiques de chaque peptide identifié pour chaque réplicat et temps de deutération. Il les assigne dans le but de définir la valeur moyenne d'incorporation en D en fonction de l'intensité de chaque pic du profil isotopique. Ces données sont vérifiées manuellement et un seul état de charge est sélectionné par peptide (requis par le logiciel MEMHDX¹³⁴, selon l'intensité et de l'assignation des profils isotopiques). Les peptides présentant une assignation incomplète, non-gaussienne ou encore montrant un échange inverse sont supprimés.

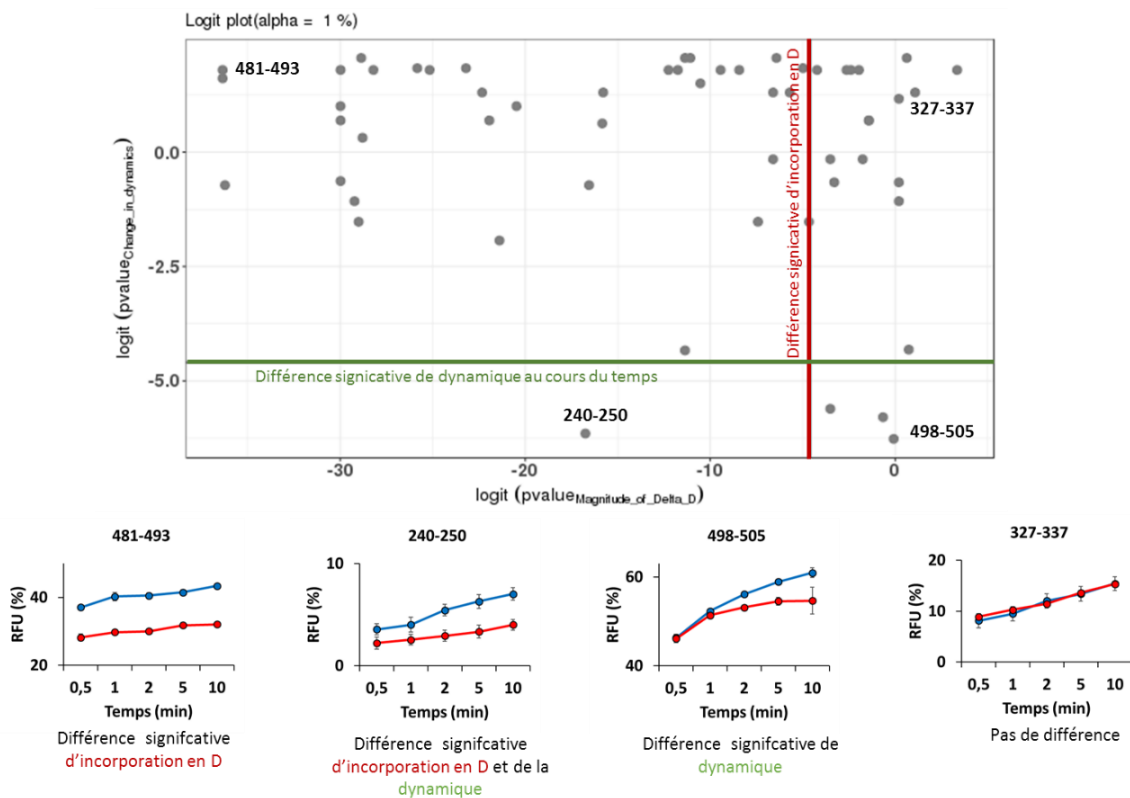


Figure 11 : Exemple de *Logit plot* obtenu grâce au logiciel MEMHDX¹³⁴ présentant les peptides identifiés (points) classés en fonction de leurs différences d'incorporation en D (à gauche de la ligne rouge) et de leurs différences de dynamique au cours du temps (en dessous de la ligne verte). Les incorporations relatives en D de quatre peptides (240-250 ; 327-337 ; 481-493 ; 498-505) sont montrées.

De valider statistiquement des différences d'incorporation en D observables entre deux états d'une même protéine (complexes, mutations, etc.) grâce au logiciel MEMHDX (*Mixed-Effects Model for HDX experiments*, Institut Pasteur)¹³⁴. Un test de Wald est appliqué avec un seuil de significativité statistique de 0,01 (p-value) considérant un pourcentage de deutérium de 95% (déterminé par le pourcentage de D estimé durant la deutération) et un seuil biologique arbitraire généralement fixé à 2%. Ce test révèle, à l'aide du *logit plot* (Fig. 11), les peptides présentant des différences significatives d'incorporation en D entre les deux états (peptide 481-493), des changements significatifs de dynamique au cours du temps (peptide 498-505) ou encore les deux (peptide 240-250).

4. Visualisation des résultats des expériences de HDX-MS

Différentes représentations des données sont disponibles, avec :

L'incorporation relative en D (% , aussi nommé RFU) d'un peptide au cours du temps pour un ou plusieurs états individuellement (Fig. 12A).

Le *difference plot* où la différence d'incorporation relative en D entre deux états est montrée en fonction de chaque peptide identifié et pour tous les temps de deutération (Fig. 12B). Cette représentation sera employée pour montrer les différences d'incorporation en D statistiquement significatives, encadrées en couleur, rouge pour une protection et bleu pour une déprotection.

La *heat map* montrant les différences d'incorporations relatives en D (entre deux états) de chaque acide aminé, via une échelle de couleur, au cours du temps sur la séquence de la protéine (Fig. 12C). Si plusieurs peptides représentent un résidu, les différences représentées sont celles du peptide le plus court chevauchant cet acide aminé. Si les peptides le recouvrant sont de mêmes longueurs, alors les différences du peptide possédant ce résidu le plus proche de son extrémité C-terminale sont conservées. De plus, l'acide aminé en N-terminal de chaque peptide n'est pas représenté. Cette représentation peut également être exportée sur la structure d'une protéine (Fig. 12D) grâce au logiciel Pymol (Schrödinger).

Le *butterfly plot* où les incorporations relatives en D de deux états sont mises en miroir pour chaque peptide identifié et chaque temps de deutération séparément (Fig. 12E).

Toutes les représentations sont mises à disposition par le logiciel DynamX. Toutefois, dans ce manuscrit, la majorité des figures sont réalisées manuellement grâce aux exports du logiciel DynamX (*state data*) et du tableur excel excepté pour les *heat map*.

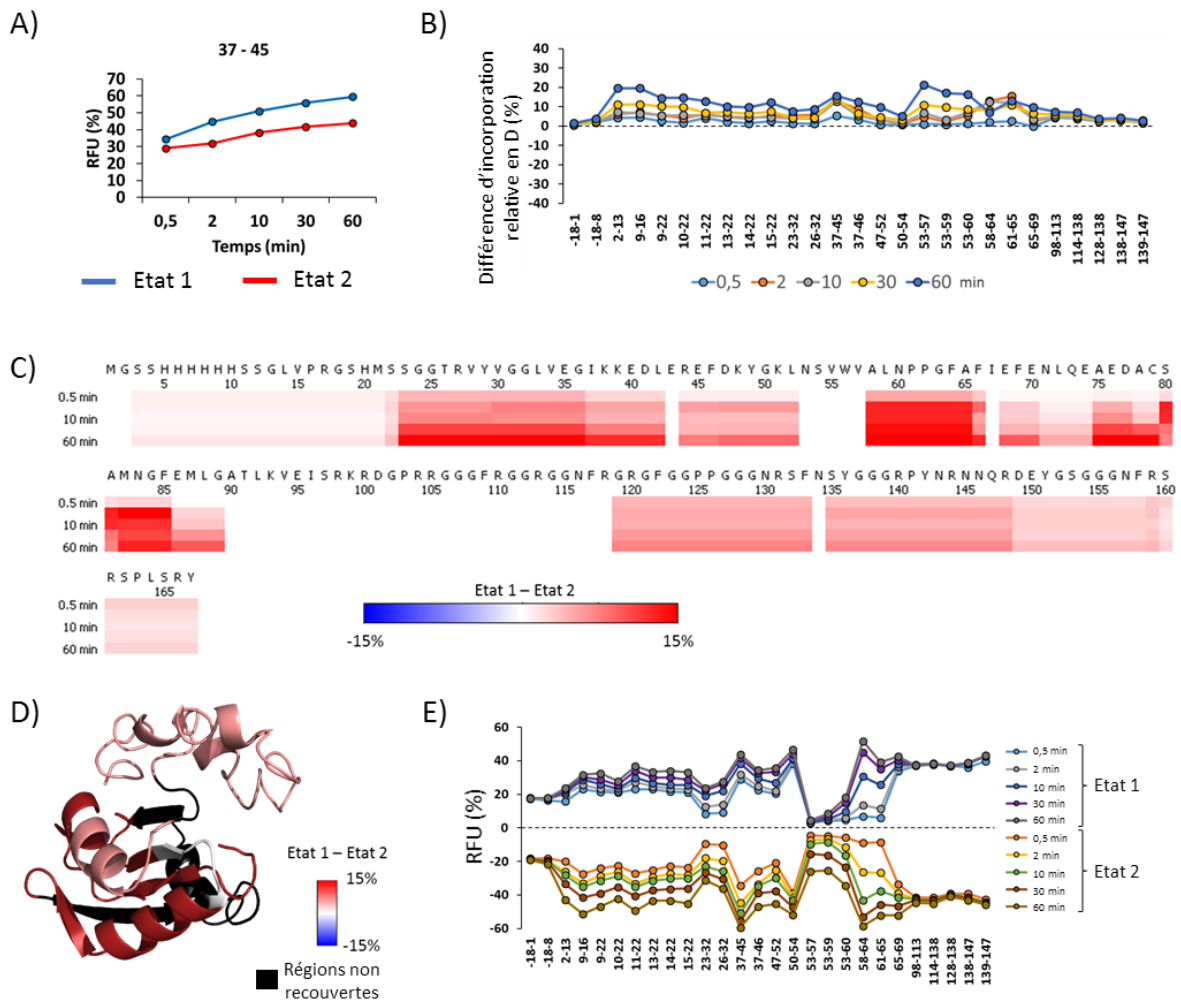


Figure 12 : Différentes représentations des données obtenues en HDX-MS avec (A) l'incorporation en D d'un peptide pour chaque état et temps de deutération, (B) le *difference plot*, (C) la *heat map*, (D) l'export de la *heat map* sur la structure de la protéine et (E) le *butterfly plot*.

IV. Optimisations et développements réalisés dans le contexte de cette thèse

1. Méthode de deutération

Les développements réalisés en HDX-MS permettent d'analyser des systèmes de plus en plus complexes qu'auparavant (instables ou très dynamiques). La méthode de deutération classiquement utilisée au laboratoire (décrite dans le paragraphe III.1.) n'est pas toujours adaptée. C'est dans cette optique que d'autres méthodes de préparation des expériences de HDX-MS ont été réalisées sur une protéine de référence en HDX-MS, la phosphorylase B (phosB, Annexe 2) et évaluées pour trois cinétiques d'incorporation en D (0,5, 15 et 60 minutes).

Trois méthodes de deutération ont donc été testées avec (i) la méthode classique complètement automatisée déjà décrite en amont avec le robot PAL LTL ; (ii) la méthode manuelle ; et (iii) la méthode manuelle après congélation avec une étape supplémentaire de congélation rapide avec de l'azote liquide permettant de conserver des échantillons déjà deutérés pendant un certain laps de temps (Fig. 13). Dans le cas de ces dernières méthodes seule l'injection a été automatisée (une méthode plus rapide). La méthode manuelle après congélation permettrait d'effectuer la deutération directement après l'expression et la purification des protéines et de les analyser plus tard.

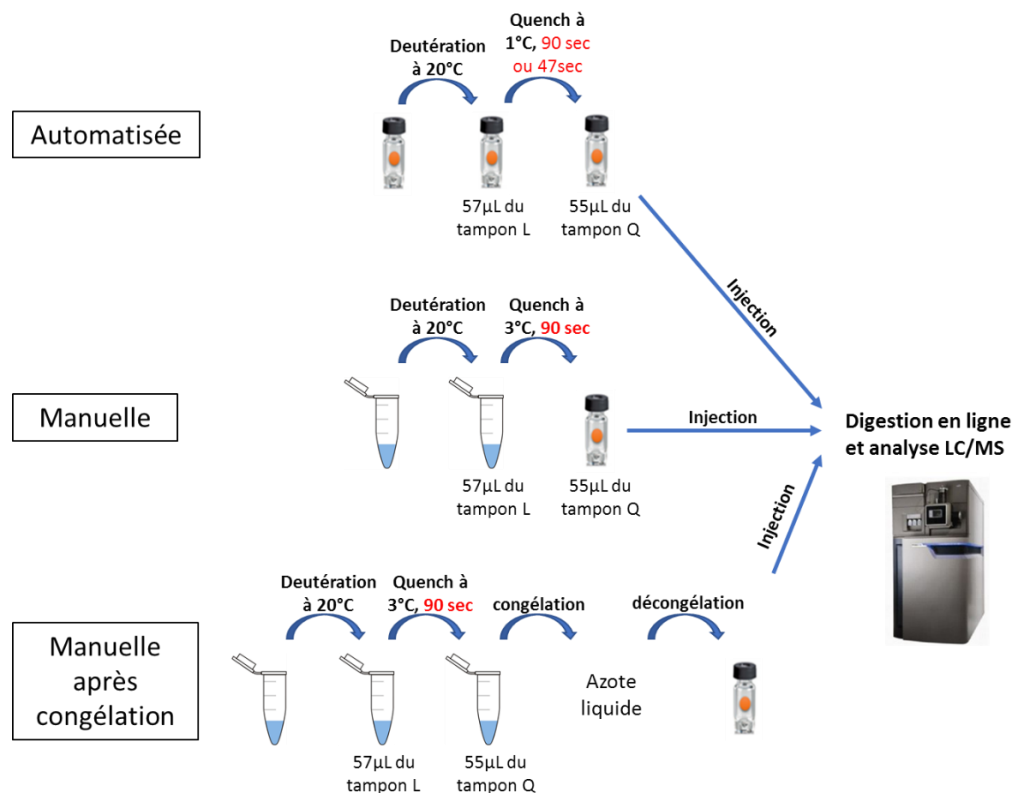


Figure 13 : Description schématique des étapes réalisées pour les différentes méthodes de deutération.

La méthode automatisée a été réalisée avec deux temps différents de *quench* (47 et 90 secondes) :

- 47 secondes est le temps que met le robot pour faire le mélange des deux tampons jusqu'au chargement de la boucle d'injection.
- 90 secondes correspondent au temps minimal pour réaliser cette étape manuellement avec une injection automatisée par le robot.

Dans un premier temps, le nombre de peptides, le pourcentage de recouvrement de séquence et la redondance ont été obtenus pour chacune des méthodes avec le même temps d'arrêt du

marquage de 90 secondes (Tableau 2, Annexe 3). Bien que les recouvrements de séquence soient assez proches (73 à 78%). Le nombre de peptides identifiés et la redondance montrent explicitement que la méthode manuelle permet d'identifier un plus grand nombre de peptides (173).

| Analyses | Automatique | Manuelle | Manuelle après congélation |
|-------------------------------|-------------|------------|----------------------------|
| Couverture de séquence (%) | 73,1 | 78,1 | 73,7 |
| Redondance | 2,8 | 3 | 2,1 |
| Nombre de peptides identifiés | 151 | 173 | 104 |

Tableau 2 : Couverture de séquence, redondance et nombre de peptides identifiés de la phosB relatés pour les trois méthodes.

Dans un second temps, nous avons voulu déterminer si le temps d'arrêt du marquage (47 vs 90 sec) influait sur l'incorporation en D. Pour cela, la différence d'incorporation relative entre les deux temps a été mesurée (Fig. 14). Aucune différence significative d'incorporation en D n'a été remarquée entre les deux temps permettant de continuer à comparer les méthodes pour un temps de *quench* de 90 secondes.

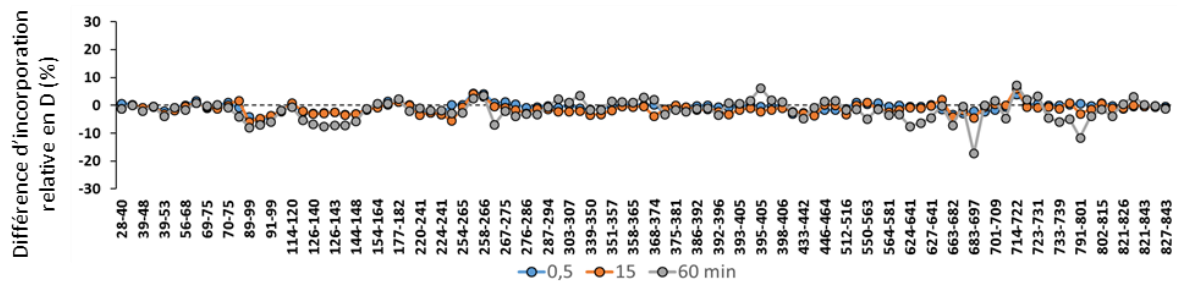


Figure 14 : Différence d'incorporation relative en D pour chaque peptide identifié de la phosB entre les méthodes automatisées avec des temps d'arrêt du marquage de 47 sec et 90 sec.

Enfin, l'incorporation relative en D lors d'une expérience HDX-MS sur la phosphorylase B a été obtenue pour chaque peptide commun (58 peptides) et méthode (Fig. 15).

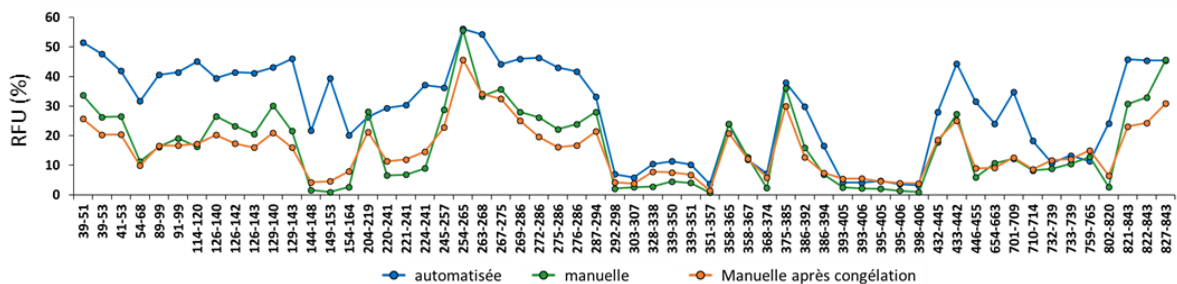


Figure 15 : Incorporation relative en D (RFU) de chaque peptide identifié de la phosB pour chaque méthode testée (automatique en bleu ; manuel direct en vert ; manuel indirect en orange)

Les résultats indiquent une incorporation en D similaire entre les deux méthodes manuelles démontrant que ni la congélation avec l'azote liquide des échantillons après l'arrêt du marquage, ni la décongélation n'ont eu d'impact sur la deutération. Toutefois, l'incorporation en D est majoritairement plus importante pour la méthode automatisée que pour les deux autres méthodes. Une température plus basse est appliquée avec le robot (1°C contre 3°C mesuré pour les méthodes manuelles) lors du *quench* diminuant l'échange inverse ce qui peut expliquer une incorporation en D observée plus importante pour la méthode automatique. Cependant, lors de la vérification des profils isotopiques, des peptides ont été supprimés, car ils présentaient des signes évidents d'échange inverse ou d'un échange coopératif (EX1) dans le cas uniquement de la méthode automatique (Fig. 16). Ces signes ne sont toutefois pas visible avec les méthodes manuelles car l'incorporation en D est globalement plus faible pour ces méthodes.

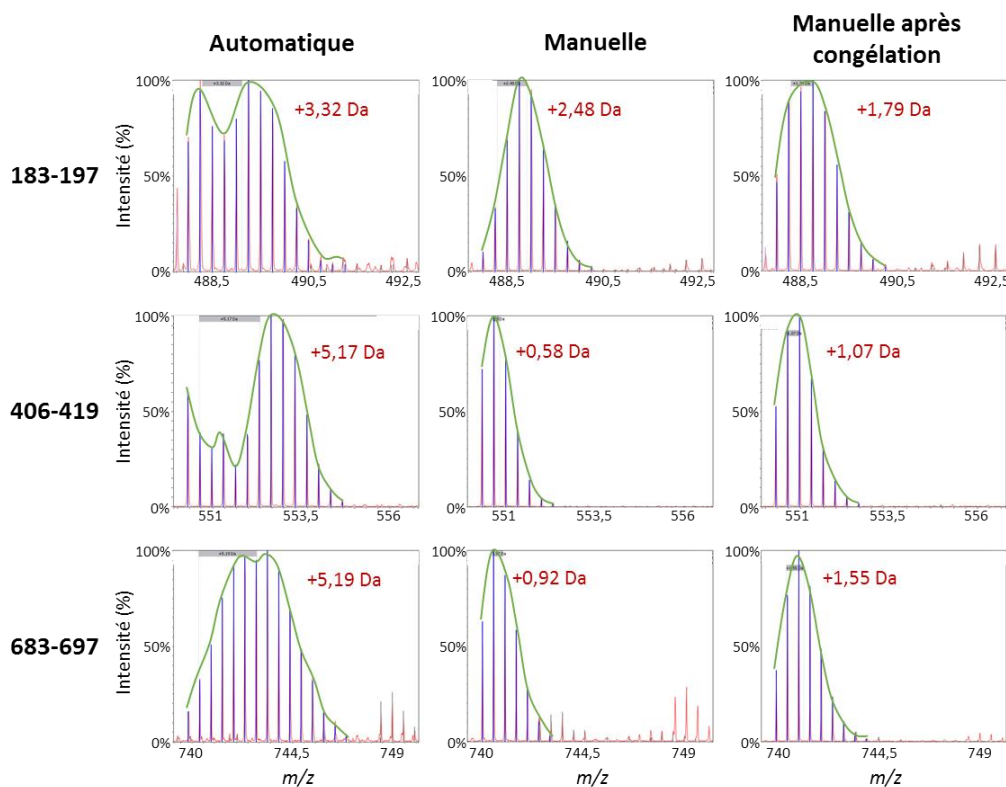


Figure 16 : Profils isotopiques obtenus pour trois peptides de la phosB (183-197 ; 406-419 ; 683-697) et chaque méthode testée (automatique ; manuelle ; manuel après congélation) avec la moyenne d'incorporation en D déterminée pour tous les profils (rouge).

Un avantage mis en avant par l'utilisation du robot, outre le temps, est la répétabilité entre les analyses. La moyenne de l'écart-type relatif pour chaque méthode a alors été déterminée

montrant que les méthodes manuelles (0,34% pour la directe et 0,32% pour l'indirecte) sont plus répétables que la méthode automatisée (0,51%).

En conclusion, les trois méthodes sont utilisables et présentent différents avantages, tels que :

- L'identification de peptides supplémentaires avec la méthode manuelle directe.
- La répétabilité pour les deux méthodes manuelles.
- La possibilité de conserver des échantillons deutérés par une congélation rapide après l'étape de *quench*.

Une perspective de cette étude serait l'utilisation d'un bloc additionnel de dé-/congélation du robot de préparation tel que décrit dans les travaux de Watson, et al.¹⁵⁵.

2. Efficacité des protéases

Avant d'effectuer la deutération des protéines, différentes optimisations sont réalisées comprenant le choix de la protéase. Lors des expériences de HDX-MS, les peptides deutérés doivent être digérés à un pH acide pour limiter l'échange inverse, ce qui diminue le nombre d'enzymes utilisables¹⁵⁶. De plus, la non-spécificité de clivage des enzymes est un prérequis pour obtenir un recouvrement de séquence maximal et une redondance élevée. Il y a encore quelques années, seule la pepsine remplissait tous ces critères et était immobilisée sur une cartouche permettant une digestion efficace, rapide et à une température contrôlée. Récemment, de nouvelles colonnes de digestion ont fait leur apparition avec notamment la rhizopuspepsine ainsi que les nepenthesines I et II (Affipro)¹²⁰⁻¹²². Ces enzymes possèdent une activité optimale à pH acide et sont aspécifiques, mais clivent préférentiellement certains sites qui ne sont pas strictement identiques à ceux de la pepsine (Tableau 3) permettant d'obtenir des peptides différents en fonction de la protéase utilisée.

| Protéase | Organismes | Sites de clivage préférentiels | Gamme de pH pour une activité optimale |
|----------------------|------------------------------|---|--|
| Pepsine | Cochon, Homme, etc. | acides aminés aromatiques, hydrophobes et L | 2-2,5 |
| Rhizopuspepsine | E. coli | acides aminés aromatiques et hydrophobes | 2,9-4,5 |
| Nepenthesine I et II | Nepenthes (plante carnivore) | acides aminés aromatiques, hydrophobes et avant D et entre K-R | 2,6-3 |

Tableau 3 : Tableau récapitulatif des caractéristiques et spécificités des enzymes utilisées^{120-122, 157}.

2.1. Digestion de la phosphorylase B

Dans le but de tester ces enzymes pour réaliser des expériences de HDX-MS, la digestion de la phosphorylase B, une protéine modèle⁶⁶, a été effectuée avec quatre enzymes différentes, la pepsine, la rhizopuspepsine, la nepenthesine I et II. Les tampons utilisés sont composés d'un mélange équimolaire de K_2HPO_4 et KH_2PO_4 de 10mM pour le tampon mimant la deutération de la protéine (H_2O) et 100mM pour le tampon Q. La protéine (40pmoles) a été injectée puis a été transmise vers la colonne de digestion à un débit de 100 μ L/min. Ce débit a été choisi en fonction des limites de pressions applicables au sein des cartouches de digestion de 2000 psi pour la rhizopuspepsine, la nepenthesine I et II (Affipro). Ainsi, les recouvrements de séquences, redondances et nombre de peptides identifiés ont été obtenus pour chacune des analyses (Tableau 4, Annexe 4). Les résultats révèlent un recouvrement de séquence maximale de 93,1% obtenus avec la pepsine. Cependant, l'utilisation de la nepenthesine II permet l'identification d'un nombre plus grand de peptides (277 contre 217 pour la pepsine) augmentant aussi la redondance à 3,5.

| Analyses | Pepsine | Nepenthesine I | Nepenthesine II | Rhizopuspepsine |
|-------------------------------|-------------|----------------|-----------------|-----------------|
| Couverture de séquence (%) | 93,1 | 92,2 | 77,8 | 78,2 |
| Redondance | 3,1 | 3,5 | 2,1 | 2,6 |
| Nombre de peptides identifiés | 217 | 277 | 173 | 207 |

Tableau 4 : Tableau décrivant la couverture de séquence, la redondance et le nombre de peptides identifiés de la phosB avec chaque protéase utilisée.

Au regard des résultats obtenus et des spécificités de clivages connus, les acides aminés en aval ($P1'$) du site de clivage ont été répertoriés pour chacune des enzymes (Fig. 17).

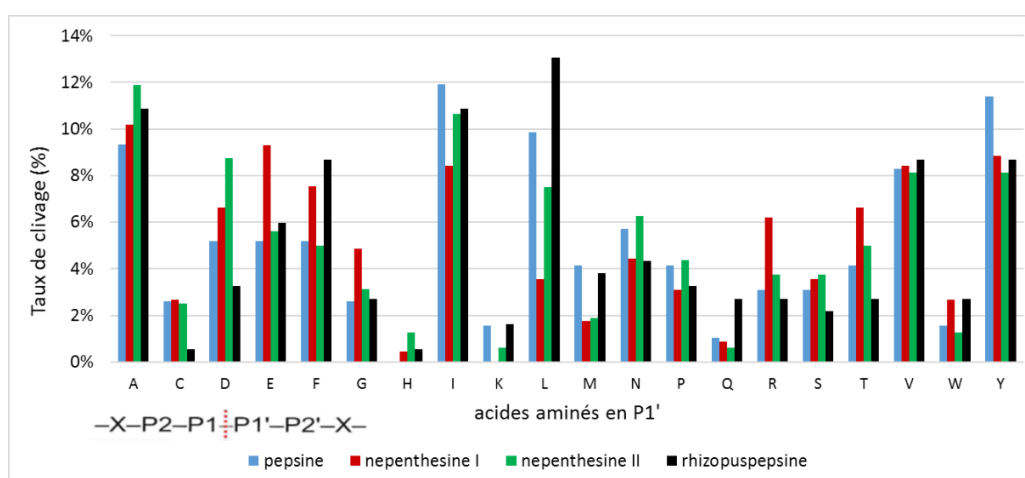


Figure 17 : Histogramme des taux de clivages obtenus en fonction de l'acide aminé en $P1'$ pour chaque protéase, la pepsine (bleu), la nepenthesine I (rouge) et II (vert) et la rhizopuspepsine (noir).

D'après ces résultats la pepsine clive préférentiellement l'isoleucine (12%) sur la phosphorylase B, alors que la nepenthesine I et II clivent l'alanine (10 et 12% respectivement) et la rhizopuspepsine la leucine (13%). De manière générale Les acides aminés ayant les plus forts pourcentages de clivage (A, I et L) sont tous des acides aminés hydrophobes en accord avec les sites connus. Ces résultats montrent l'aspécificité de ces protéases bien qu'elles possèdent des sites préférentiels.

2.2. Digestion du pembrolizumab

Bien que la phosphorylase B soit une protéine modèle en HDX-MS, cette protéine reste facilement digérée par chacune des enzymes. Il est donc intéressant d'augmenter la complexité de la protéine étudiée dans le but d'observer les réels bénéfices et différences des enzymes. Un anticorps, le pembrolizumab (Annexe 2), a ainsi été testé, car il possède des ponts disulfures et de nombreuses régions structurées généralement difficiles à digérer. Les tampons utilisés sont composés de PBS 1x pour le tampon mimant la deutération de la protéine (H₂O) et 2M Guanidine-HCl, 100mM de Glycine et 250mM de TCEP pour le tampon Q. La protéine (50pmoles) a été injectée et un débit de 100µL/min a été appliqué lors de la digestion. Les recouvrements de séquence pour la chaîne légère (Lc) et lourde (Hc) ont été obtenus ainsi que la redondance et le nombre de peptides identifiés (Tableaux 5 et 6, Annexe 5). La Lc du pembrolizumab a été entièrement recouverte par des peptides générés par le nepenthesine I (100% de couverture de séquence). Toutefois, plus de peptides (86) ont été identifiés suite à la digestion avec le nepenthesine II comparée aux autres protéases (Tableau 5).

| Lc | Pepsine | Nepenthesine I | Nepenthesine II | Rhizopuspepsine |
|-------------------------------|---------|----------------|-----------------|-----------------|
| Couverture de séquence (%) | 96,8 | 100 | 96,8 | 94,5 |
| Redondance | 4,2 | 4,8 | 4 | 3,2 |
| Nombre de peptides identifiés | 80 | 77 | 86 | 69 |

Tableau 5 : Tableau décrivant la couverture de séquence, la redondance et le nombre de peptides identifiés de la chaîne légère du pembrolizumab avec chaque protéase utilisée.

Concernant la Hc, bien plus riche en ponts disulfures, la nepenthesine II est l'enzyme la plus efficace avec un plus grand recouvrement de séquence, redondance et nombre de peptides générés (Tableau 6). Ces résultats tendent vers une utilisation préférentielle de la nepenthesine II dans le cadre de l'analyse d'anticorps monoclonaux.

| Hc | Pepsine | Nepenthesine I | Nepenthesine II | Rhizopuspepsine |
|-------------------------------|---------|----------------|-----------------|-----------------|
| Couverture de séquence (%) | 78,7 | 64,3 | 82,4 | 77,8 |
| Redondance | 3,1 | 3,2 | 3,4 | 2,9 |
| Nombre de peptides identifiés | 97 | 70 | 118 | 109 |

Tableau 6 : Tableau décrivant la couverture de séquence, la redondance et le nombre de peptides identifiés de la chaîne lourde du pembrolizumab avec chaque protéase utilisée.

La carte de couverture de séquence de la Hc (Fig. 18) obtenue pour les quatre protéases testées montre que la région N-terminale (résidus 1-113) de cette chaîne n'est que très peu recouverte quelle que soit l'enzyme utilisée. Néanmoins, les zones couvertes de cette région sont différentes suivant l'enzyme utilisée. Cette zone étant la région variable de l'anticorps et le site de reconnaissance des antigènes, il est intéressant d'identifier le plus de peptides différents possibles au niveau de cette extrémité montrant l'intérêt du choix de la protéase.

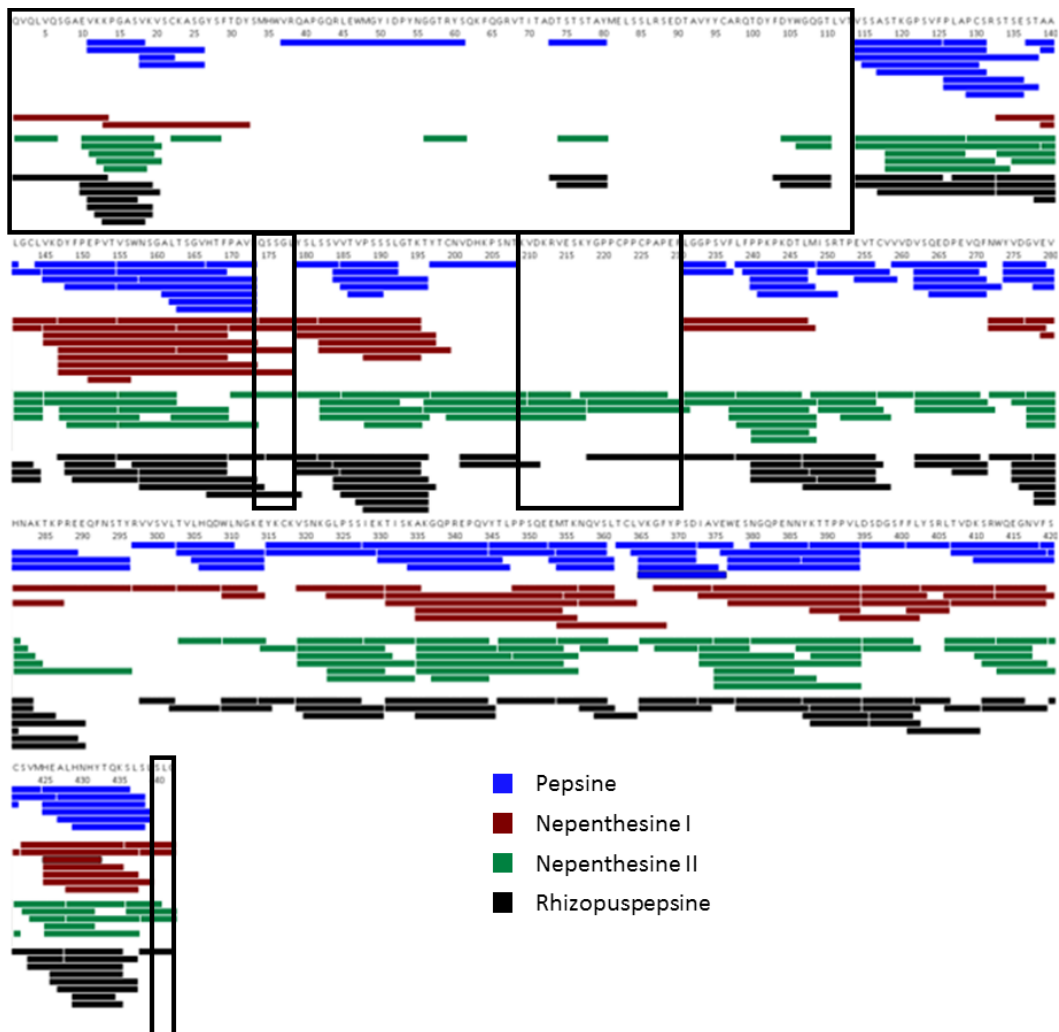


Figure 18 : Carte de couverture de séquence de la chaîne lourde du pembrolizumab par les peptides identifiés avec la pepsine (bleu), la nepenthesine I (rouge), la nepenthesine II (vert) et la rhizopuspepsine (noir) réalisée à l'aide de ¹⁴⁰.

En conclusion, ces enzymes sont efficaces pour la digestion d'anticorps monoclonaux avec un recouvrement de séquence minimal de 77,8% pour la Hc du pembrolizumab. Les peptides générés par les différentes protéases semblent complémentaires, ce qui s'explique par les préférences de clivages des différentes enzymes. Certaines régions sont uniquement recouvertes par des peptides issus de la digestion avec des enzymes différentes de la pepsine. Ces couvertures de séquence sont en accords avec celles référencées (80-100%) dans la littérature pour les anticorps¹⁵⁸⁻¹⁶⁰ et pourraient encore être augmenter avec l'utilisation d'une nouvelle colonne PNGase développée par Affipro¹⁶¹.

3. Apport de la mobilité ionique en HDX-MS

Dans le but d'obtenir toujours plus d'informations sur les protéines/complexes étudiés en HDX-MS, je me suis également intéressée à l'apport de l'IM pour l'analyse HDX-MS. Grâce à la cellule de mobilité ionique (TWIMS) du Triwave™ du Synapt-G2-Si HDMS (Waters), il est possible d'implémenter une séparation supplémentaire des peptides en fonction de leurs charges et conformations¹⁶² (Fig. 19). La cellule TWIMS, lorsqu'elle est activée, fonctionne via (Fig. 19):

- L'application de vagues de potentiels (40V de hauteur de vague et 650 m/s de vitesse de la vague) entraînant les ions et permet de les séparer en fonction de leur charge et plus précisément de leur valeur de m/z .
- L'application d'un flux de N₂ (90 mL/min) dans le sens opposé ralentissant la progression des ions en fonction de leur conformation.

Les pressions au sein de la cellule d'hélium et TWIMS sont de 3 et 5 mbar respectivement au lieu de 10⁻⁴ mbar. L'utilisation de l'IM en HDX-MS permettrait de décomplexifier les spectres (Fig. 19) pouvant affecter directement les résultats. En effet, de nombreux peptides lors de la vérification manuelle des profils isotopiques sont supprimés du fait de chevauchement de plusieurs distributions isotopiques. Ces ions peuvent empêcher l'assignation correcte des profils isotopiques, et donc fausser la valeur moyenne d'incorporation en D des peptides étudiés. Plusieurs expériences de HDX-MS avec l'IM activée ont déjà été effectuées¹⁶³⁻¹⁶⁵ sans toutefois montrer l'impact de celle-ci sur le traitement des données, excepté une publication de 2003 portant sur deux protéines de références la BSA et l'hormone rhGH¹⁶⁶. C'est pourquoi j'ai voulu déterminer l'impact de l'IM sur (i) l'identification des peptides (non deutérés) et (ii) le traitement des données sur cinq types différents de protéines.

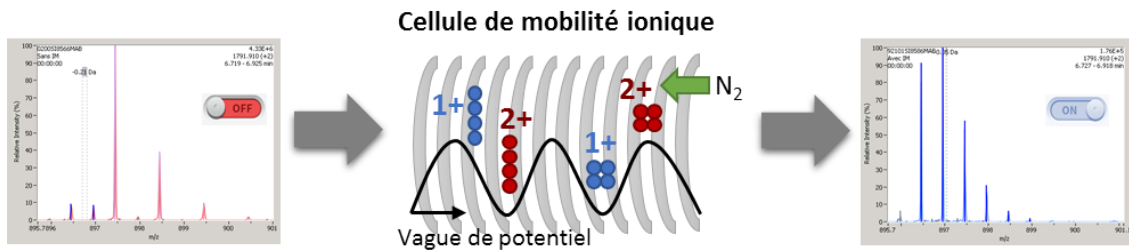


Figure 19 : Représentation schématique de l'impact de l'IM sur des expériences de HDX-MS et de la séparation des ions en fonction leurs charges et conformations au sein de la cellule TWIMS.

3.1. Impact de l'IM sur l'identification des peptides

Dans un premier temps, cinq protéines (PhosB, PXR, PRMT9, Nivolumab et EPAC1, Annexe 2) de différentes complexités (Fig. 20) ont été analysées avec et sans l'IM activée dans le but de définir les avantages de cette méthode de séparation supplémentaire sur l'identification des peptides. Les expériences HDX-MS sur la phosphorylase B (protéine modèle) et PXR ont été réalisées par le Dr. Bourguet Maxime au sein du laboratoire. Toutes les protéines ont été digérées avec de la pepsine dans des conditions optimisées (Annexe 6).

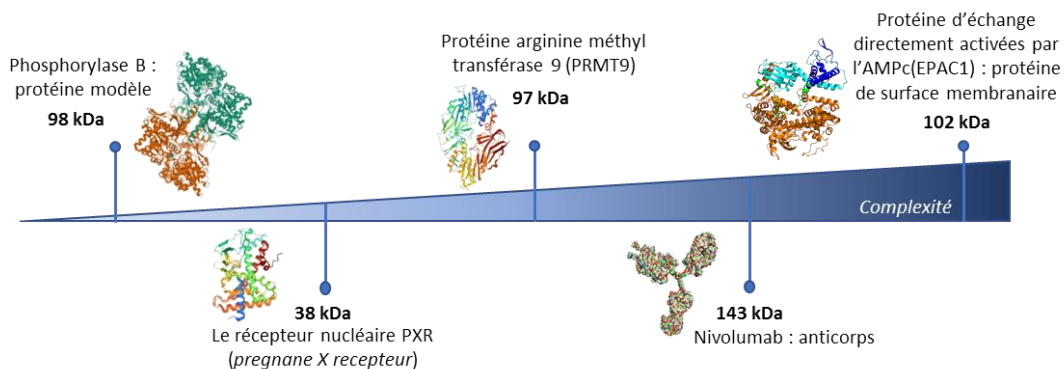


Figure 20 : Classement des protéines étudiées avec et sans IM activée en fonction de leur complexité.

Lors de l'identification des peptides par le logiciel PLGS, les données expérimentales (valeur de m/z , temps de rétention et intensité) sont comparées à des données théoriques issues de la digestion *in sillico* des protéines comprises dans une banque de données (Annexe). L'extraction des données expérimentales et des attributs chromatographiques de chaque ion via un algorithme, nommé Apex3D, permet de produire une liste de tous les ions précurseurs et produits. Pendant cette extraction, des seuils d'intensités par défaut sont appliqués pour les scans de basse énergie (LE) et hautes énergies (HE) qui sont de 250 et 100 respectivement. Avec l'IM activée ces paramètres peuvent être abaissés pour permettre une identification optimale des peptides dans ces conditions, suivant les conseils du constructeur, à 120 (LE) et 20 (HE).

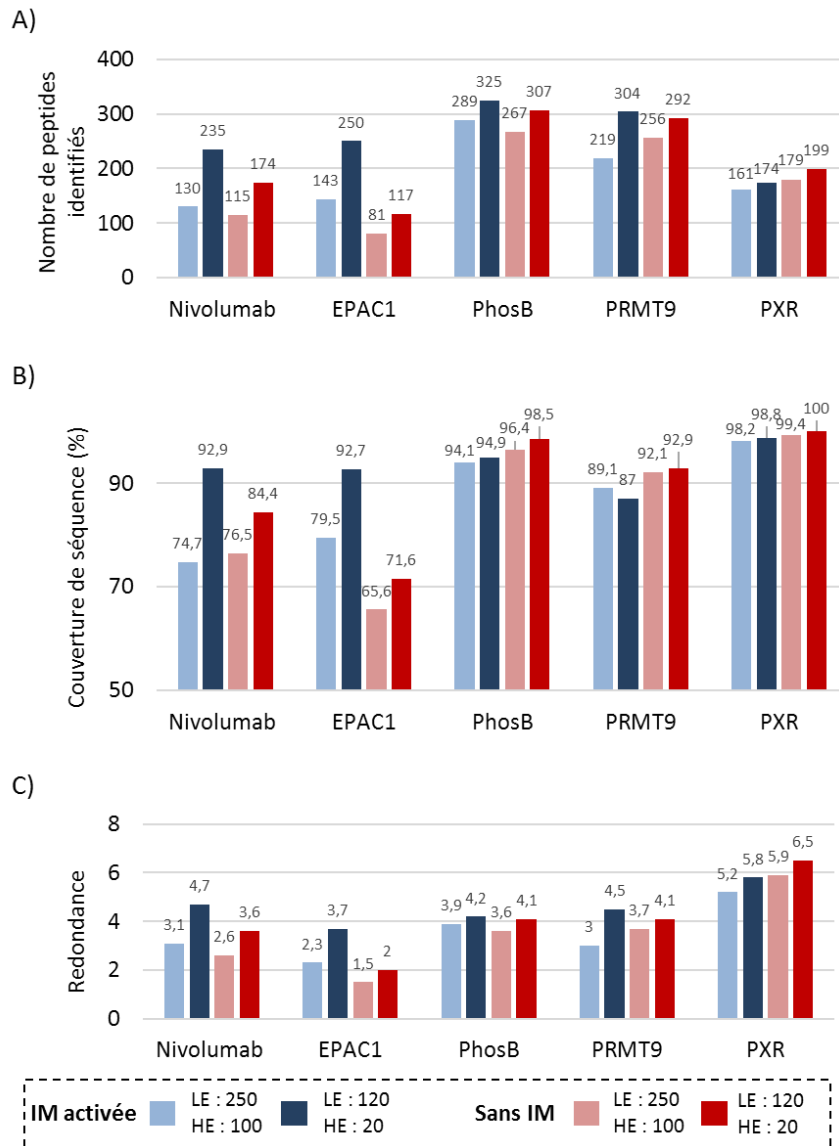


Figure 21 : Histogrammes représentant (A) le nombre de peptides identifiés, (B) la couverture de séquence et (C) la redondance, obtenus lors de la digestion des protéines étudiées (Nivolumab, EPAC1, phosB, PRMT9 et PXR) avec l’IM activée (bleu) et sans IM (rouge) avec les seuils LE et HE par défauts (clair) et diminués (foncé).

Le recouvrement de séquence, la redondance et le nombre de peptides identifiés ont été obtenus avec/sans l’IM activée et suivant les seuils LE et HE testés pour chacune des protéines (Annexe 7). Ces résultats ont révélé que : (i) la diminution des seuils LE et HE augmente considérablement le nombre de peptides identifiés dans le cas de l’IM activée par rapport à sans IM avec un gain de 69 peptides en moyenne avec l’IM contre 39 sans l’IM active (Fig. 21A) ; (ii) un nombre de peptides identifiés généralement plus important avec l’IM activée ; (iii) une couverture de séquence largement augmentée en activant l’IM pour les protéines les plus complexes à analyser en HDX-MS avec 8,5% de plus pour le Nivolumab et 21,1% de plus pour EPAC1 (Fig. 21B) ; et (iv) une redondance également augmentée (Fig. 21C).

Ce critère est très important lors de la vérification manuelle des profils isotopiques, car certains peptides sont supprimés à cause de chevauchement de profils isotopiques de peptides différents.

3.2. Impact sur le traitement de données

Trois des protéines précédemment testées ont été deutérées avec/sans l'IM activée, la phosphorylase B, PXR et le Nivolumab. Les seuils LE et HE par défaut sont appliqués dans le but de pouvoir comparer les deux jeux de données (avec et sans IM activée) avec un procédé d'identification des peptides similaires. Le traitement manuel des données a été effectué par la vérification des assignements des profils isotopiques et la suppression de certains peptides dus à un assignement par le logiciel DynamX jugé incorrect qui fausse la moyenne d'incorporation en D du peptide. Un ré-assignement peut aussi être réalisé, uniquement si le profil isotopique est incomplet (profil gaussien attendu) et après avoir vérifié la correspondance entre la valeur de m/z du pic à assigner et la masse du peptide en fonction de l'état de charge. Les couvertures de séquence, redondances et nombre de peptides identifiés ont été comparés avant et après le traitement manuel des données (Fig. 22, Annexe 8) pour chaque condition (avec/sans IM activée) et protéine.

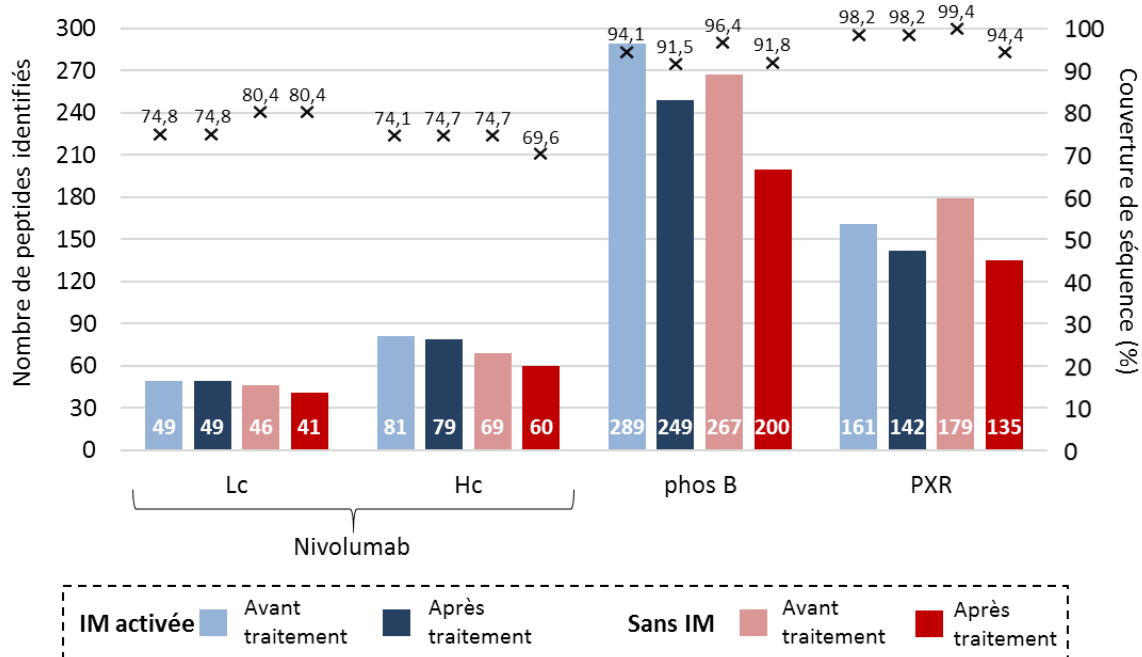


Figure 22 : Représentation du nombre de peptides identifiés (histogrammes) et de la couverture de séquence (croix) des protéines étudiées (Nivolumab Lc et Hc, PXR et PhosB). Les analyses ont été effectuées avec (bleu) et sans (rouge) l'IM activée et les valeurs indiquées ont été obtenues avant (clair) et après (foncé) le traitement manuel des données.

Dans le cas des analyses sans l'IM activée, une perte moyenne de 31 peptides et 3,7% de couverture de séquence alors qu'avec l'IM activée seuls 15 peptides en moyenne sont supprimés soit 0,7% de couverture de séquence. Dans l'ensemble, la mobilité ionique nous permet d'obtenir plus d'informations sur les protéines étudiées. En effet, grâce à cette méthode de séparation supplémentaire, une diminution d'assignation erronée des profils isotopiques est observée par le logiciel DynamX (Fig. 23).

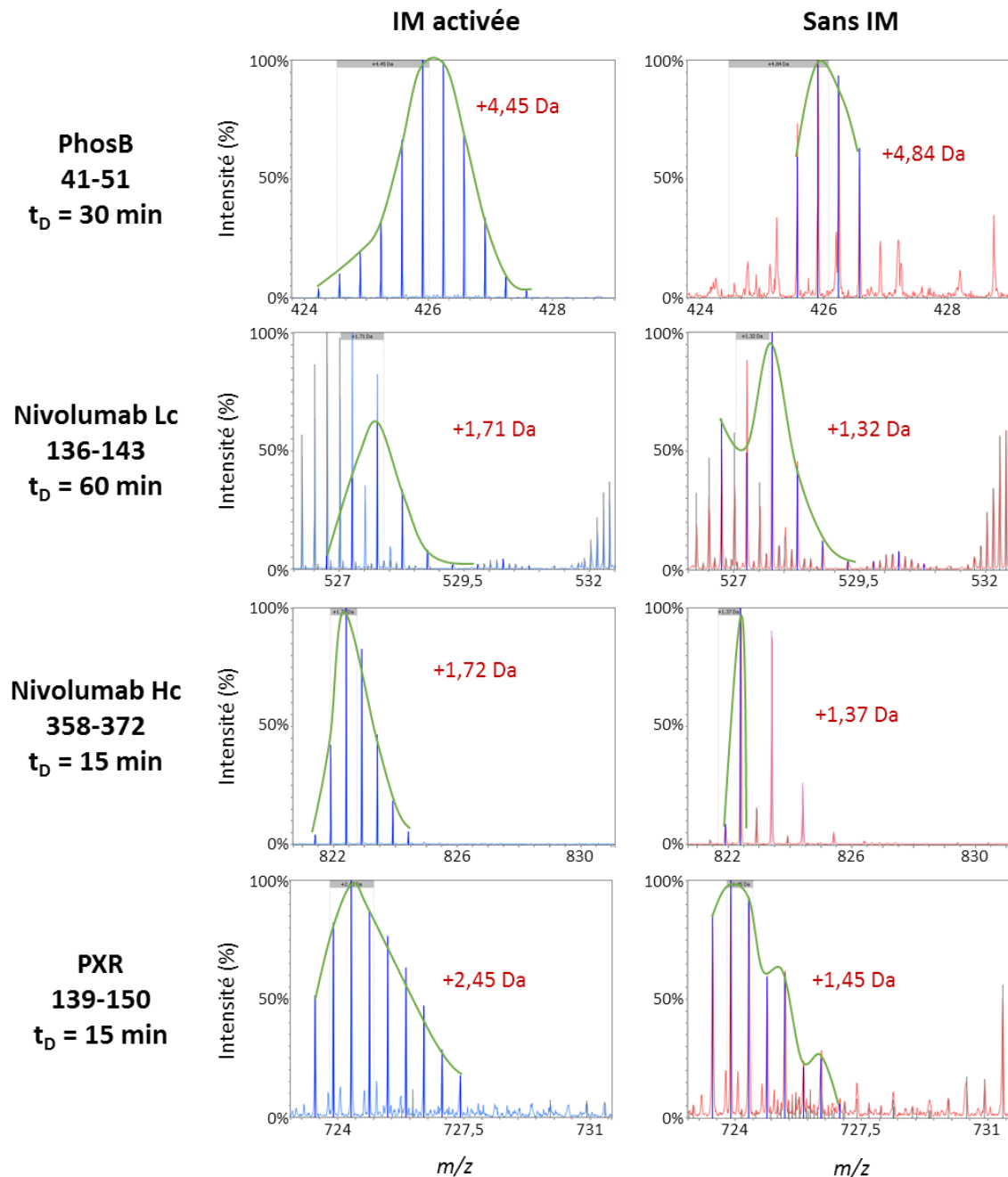


Figure 23 : Profils isotopiques obtenus correspondant à quatre peptides identifiés (41-51 PhosB, 136-143 Nivolumab Lc, 358-372 Nivolumab Hc et 139-150 PXR) avec et sans l'IM activée. Les valeurs moyennes d'incorporation en D de ces profils, déterminées par le logiciel DynamX, sont décrites en rouge. Les ions assignés sont représentés en bleu.

Différents avantages de l'activation de l'IM ont été révélés par les analyses, avec :

- Une augmentation de l'intensité du signal entraînant l'amélioration de la résolution isotopique (peptide 41-51, PhosB)
- Une meilleure séparation des différents ions détectés au sein du même spectre permettant d'éviter des assignations incorrectes (peptide 136-143, Nivolumab Lc) ou incomplètes (peptide 358-372, Nivolumab Hc)
- Une valeur moyenne d'incorporation en D plus juste induite par des gaussiennes assignées de meilleures qualités (peptide 139-150, PXR)

4. Validation statistique des données : quel outil utiliser ?

De nos jours, l'utilisation d'outil statistique est fortement recommandée dans le but de valider les différences d'incorporations en D observé entre deux états⁶⁶. Un grand nombre d'approches statistiques sont actuellement disponibles avec notamment le modèle mixte et le test-t de Student. L'application web MEMHDX¹³⁴ et le logiciel Deuterios 2.0¹³⁵ sont deux outils permettant de mettre en œuvre chacune de ces approches. D'un côté, MEMHDX applique un modèle statistique mixte (fixe et aléatoire) pour chaque peptide, condition expérimentale, temps de deutération et des différents réplicats tout en considérant deux valeurs de probabilité (p-value) pour les changements de dynamique et les différences d'incorporation en D. D'un autre côté, Deuterios emploie le test-t pour comparer les moyennes d'incorporation en D des deux états étudiés pour chaque temps de deutération individuellement. De nombreuses études en HDX-MS ont déjà été validées par chacun de ces logiciels montrant leur pertinence^{164-165, 167-169}.

Ces deux méthodes de validation statistique ont été appliquées à trois jeux de données comparant deux états d'une même protéine, avec (i) l'hétérodimère impliquant le récepteur à la vitamine D (VDR) en complexe avec W1 et avec/sans le ligand ZK12813 ; (ii) le complexe EPAC1 avec/sans la présence de liposome et de l'AMPc ; ainsi que (iii) la protéine non-mutée PPAR γ comparée au mutant F310S de PPAR γ (Annexes 2 et 9). Ces trois expériences permettent d'observer des changements conformationnels distincts et induits par la complexation de petites molécules, par un environnement lipidique ou encore par une mutation. Elles donnent ainsi une vue globale de la pertinence de chacune des approches statistiques pour différentes expériences. Pour les deux tests, une p-value a été fixé à 1% et un pourcentage de deutérium de 95% déterminé par le facteur de dilution des échantillons

dans le tampon deutéré. Les résultats obtenus décrivent les peptides présentant une différence statistiquement significative d'incorporation en D par rapport au temps de deutération (1 ou plusieurs) pour Deuterios et en fonction du pourcentage de différences pour MEMHDX (Fig. 24).

Dans le cas de VDR et PPAR γ (Fig. 24A-B), les résultats obtenus pour les deux méthodes statistiques sont très proches. Ils montrent en plus que les peptides significatifs, donnés par MEMHDX, le sont aussi pour au minimum trois temps de deutération définis par Deuterios. Les recommandations de HDX-MS indiquent la nécessité de réaliser six temps de deutération comprenant le t_0 où les protéines ne sont pas deutérées. Ces recommandations nous amènent à nous demander si un seul temps de deutération significatif est suffisant pour relater d'un changement conformationnel au niveau du peptide. Cependant, le nombre de temps de deutération présentant des différences d'incorporation en D peut amener à des conclusions sur la dynamique des protéines au cours du temps.

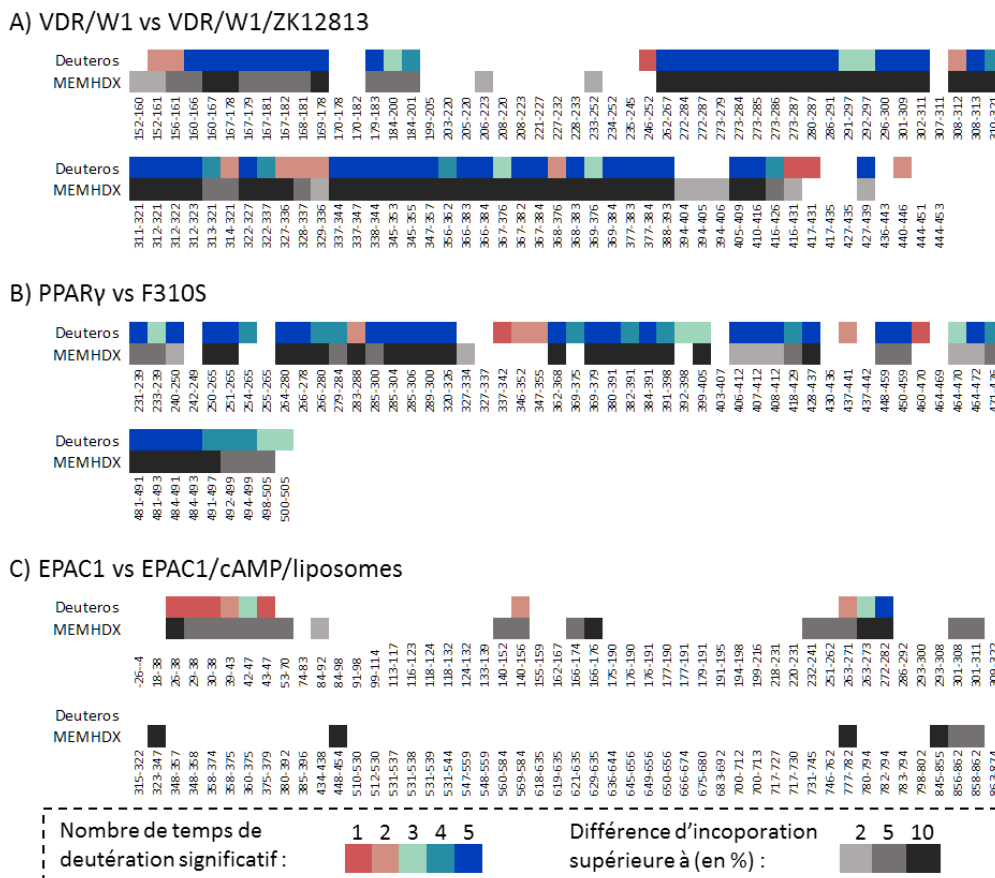


Figure 24 : Représentation des peptides statistiquement significatifs pour les comparaisons (A) VDR/W1 vs VDR/W1/ZK12813, (B) PPAR γ vs F310S et (C) EPAC1 vs EPAC1/AMPC/liposomes déterminés par les outils statistiques MEMHDX et Deuterios. Le nombre de temps de deutération significatif donné par Deuterios sont identifiables par une échelle de couleur et les différences d'incorporation définies par MEMHDX sont montrées par une échelle de gris (allant de 2 à plus de 10%).

Les expériences menées sur EPAC1 (Fig. 24C) révèlent que seulement 10 peptides présentent des différences d'incorporation en D avec Deuterios alors que 25 peptides sont significatifs d'après MEMHDX, dont 5 avec des différences d'incorporation en D supérieures à 10%. Ces 5 peptides présentent une faible répétabilité entre les réplicats de l'ordre de 2,3% contre 1,5% en moyenne pour les autres peptides. Tandis que Deuterios moyenne l'incorporation en D des réplicats, MEMHDX va considérer ces réplicats individuellement révélant ainsi des différences significatives bien que la répétabilité soit faible.

Enfin, une nouvelle méthode pour la validation statistique des données de HDX-MS a été développée par Dr. J. Claesen¹⁷⁰. Il combine le test-t de Student avec la méthode bayésienne empirique améliorant l'estimation du niveau de deutérium en fonction de la répétabilité. Cette approche et les deux précédentes (MEMHDX et Deuterios) ont été testées sur des données comparant l'incorporation en D de la protéine PRMT2 à celle de l'hétérodimère PRMT2/RSF1 (Annexes 2 et 9). Ces données présentent différents avantages pour comparer les trois méthodes statistiques avec des peptides présentant (Fig. 25) d'importantes différences d'incorporation en D supérieures à 10% (256-276), des changements de dynamique au cours du temps (312-319) ainsi qu'une faible répétabilité (12-18 ; 334-340).

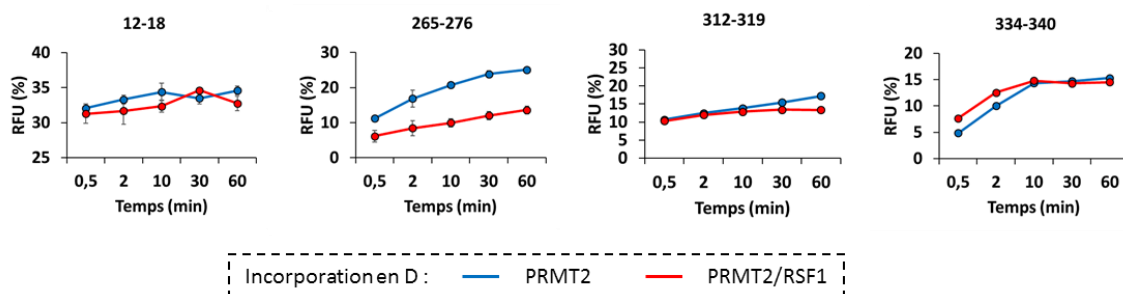


Figure 25 : Incorporation en D (RFU %) en fonction du temps de deutération pour quatre peptides de PRMT2 (12-18 ; 265-276 ; 312-319 ; 334-340). Les deux états représentés sont PRMT2 (bleu) et le complexe PRMT2/RSF1 (rouge).

Ainsi, les trois tests ont été effectués sur les données de HDX-MS montrant une grande similarité entre les peptides significatifs définis par MEMHDX et Deuterios (Fig. 26). Cependant, la quasi-totalité de la région N-terminale (100 premiers acides aminés) présente des différences d'incorporation en D significative avec Deuterios alors que seuls 9 peptides sont validés statistiquement avec MEMHDX. Cette région est le domaine SH3 de PRMT2 et connue comme étant une région flexible¹⁷¹ montrant l'intérêt de l'emploi de Deuterios dans le cas de l'analyse de ce type de protéines. Néanmoins, une minorité des peptides de ce domaine

uniquement validé par Deuterios (8/23) sont significatifs pour au moins 3 temps de deutération laissant un doute sur la réelle significativité de ces peptides. Au regard de ces informations, la nouvelle méthode statistique, appelée dans ce manuscrit *Moderated*, ne valide que peu de différences d'incorporation en D. Bien qu'elle soit décrite comme améliorant l'estimation du niveau de deutérium en fonction de la répétabilité, le peptide 12-18 (Fig. 26) dont les valeurs d'incorporations en D sont peu répétables n'est pas significatif avec cette méthode alors qu'il l'est pour les deux autres tests. Il permet néanmoins de confirmer les différences significatives les plus grandes (> à 5% d'après MEMHDX).

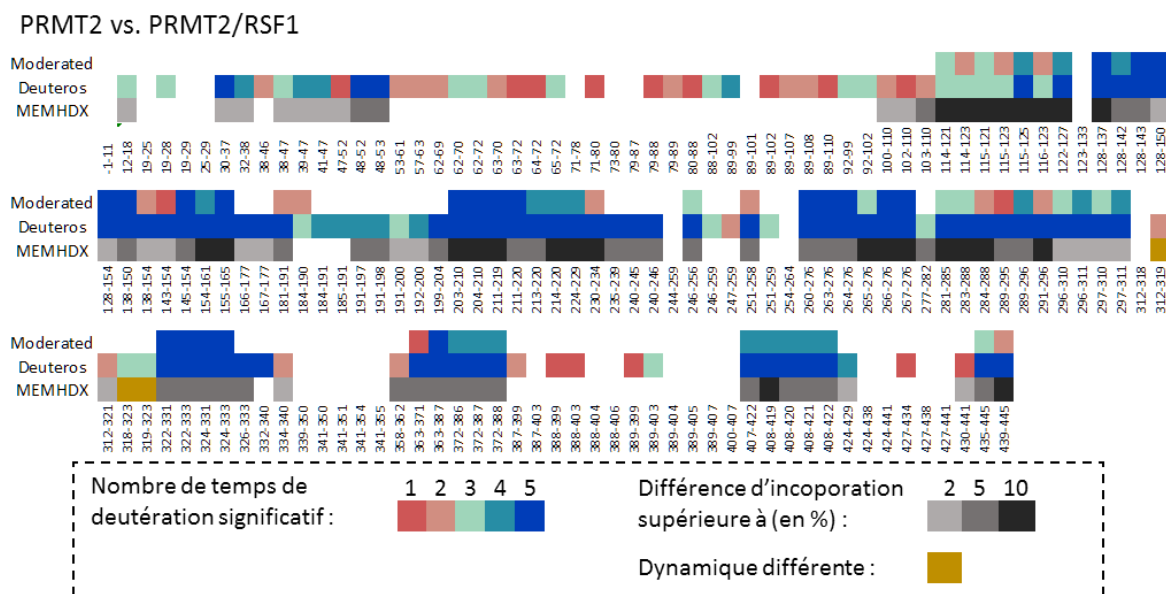


Figure 26 : Représentation des peptides statistiquement significatifs pour la comparaison PRMT2 vs PRMT2/RSF1 déterminé par les outils statistiques MEMHDX, *Moderated* et Deuterios. Le nombre de temps de deutération significatif donné par Deuterios et *Moderated* sont identifiables par une échelle de couleur et les différences d'incorporation définit par MEMHDX sont montrées par une échelle de gris (allant de 2 à plus de 10%) et les différences de dynamique en jaune.

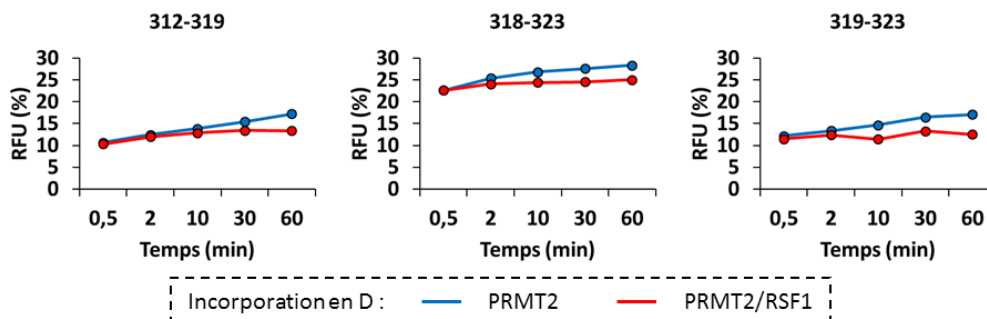


Figure 27 : Incorporation en D (RFU %) en fonction du temps de deutération pour trois peptides de PRMT2 (312-319 ; 318-323 ; 319-323). Les deux états représentés sont PRMT2 (bleu) et le complexe PRMT2/RSF1 (rouge).

Enfin, MEMHDX révèle des différences de dynamiques au cours du temps (Fig. 26-27) pour trois peptides (312-319 ; 318-323 ; 319-323) où la différence d'incorporation en D n'est pas significative traduisant du changement de comportement d'incorporation en D entre les deux conditions prenant en compte les temps de deutération comme Deuterios.

V. Conclusions

Dans ce chapitre, le protocole de HDX-MS, mis en place au laboratoire LSMBO, a d'abord été décrit. Puis, les différents développements et optimisations de cette approche ont été discutés permettant d'apporter une meilleure compréhension de la technique et des différents moyens de la mettre en œuvre. Ces développements ont été réalisés dans le but d'obtenir des informations plus complètes, adaptées et validées statistiquement. En effet, ces études, portant sur chacune des étapes (pré-, post-, analytique) de cette technique, ont montré que :

- Plusieurs méthodes pour effectuer la deutération peuvent être employées et présentent chacune leurs avantages en termes de temps, d'identification de peptide, conservation d'échantillon deutéré et de répétabilité.
- L'efficacité des différentes enzymes de digestion immobilisées sur des colonnes ainsi que leurs complémentarités de par leurs sites de clivage préférentiels. La nepenthesine II et la pepsine semblent être les deux enzymes à privilégier lors des optimisations réalisées en amont de la deutération.
- L'utilisation d'une méthode de séparation supplémentaire des peptides grâce à l'IM présente de nombreux avantages pour tous les types de protéines étudiées avec (i) l'identification d'un plus grand nombre de peptides (augmentation de la couverture de séquence et de la redondance) ainsi que (ii) l'amélioration de la séparation des profils isotopiques obtenus ce qui autorise la conservation de plus de peptides comparés à l'analyse classique des peptides deutérés.
- Les logiciels statistiques proposés emploient différents tests donnant des informations, elles aussi, différentes avec le nombre de temps de deutération significatif (Deuterios et *Moderated*) ou le pourcentage de différences (MEMHDX). Suivant les cas étudiés, MEMHDX et Deuterios sont à considérer pour la validation statistique des données. MEMHDX permet de tenir compte du manque de répétabilité des analyses et met en valeur les différences d'incorporation en D et de dynamique

alors que Deuterios apporte plus d'informations dans le cas de l'étude de protéine flexible pour chaque temps de deutération individuellement.

En perspective, la combinaison de deux protéases, pepsine et nepenthesine II, au sein de la même colonne a été réalisée par Affipro permettant de digérer les protéines par ces deux enzymes simultanément. Cette colonne présenterait de nombreux avantages pour l'étude de protéines et complexes multi-protéiques les plus difficiles à digérer tel que les protéines membranaires¹²⁵.

Il serait également intéressant de déterminer l'impact de deux méthodes de traitement différentes des données de HDX-MS avec la comparaison des logiciels DynamX (Waters) et HDExaminer (Sierra Analytics). HDExaminer est largement employé pour le traitement des données de HDX-MS de par sa versatilité, car il permet de lire des données provenant de différents instruments (Waters ou Thermo Fisher Scientific) en plus de la réalisation directe de test statistique¹⁷²⁻¹⁷⁵.

Chapitre 2 : XL-MS : Théorie, état de l'art, développements et optimisations

En 2000, les premières publications portant sur la combinaison du pontage chimique avec la MS pour la caractérisation des protéines sont apparues avec du BS3 (bis(sulfosuccinimidyle)subérate) et du DTSSP (3,3'-dithiobis(sulfosuccinimidyle propionate)) comme agent pontant. Depuis ce temps, le nombre d'articles faisant référence au XL-MS n'a cessé d'augmenter (Fig. 1) avec l'utilisation d'agents pontants marqués¹⁷⁶, MS-clivables¹⁷⁷ ou encore l'étude d'interactions protéines-protéines¹⁷⁸, de multi-complexes¹⁷⁹, allant jusqu'aux lysats cellulaires¹⁸⁰ et aux tissus⁷⁷. Des études de comparaison conformationnelle ont également été menées¹⁸¹. Au cours de ces dernières années, ces recherches se sont inscrites dans des projets de biologie structurale intégrative en complémentarité d'autres techniques¹⁸².

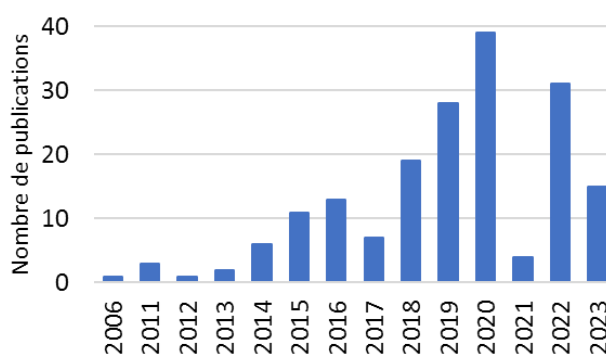


Figure 1 : Représentation du nombre de publications référencés dans PubMed® ayant pour mots clés : XL-MS.

I. Théorie et principe

1. Agent pontant

Bien que le XL-MS soit plutôt récent, le premier pontage a été identifié en 1958 par chromatographie sur papier, depuis lors le nombre de réactifs pour le cross-link n'a cessé d'augmenter avec 268 agents disponibles et commercialisés par Sigma-Aldrich et 104 par Thermo Fisher Scientific en 2021⁷¹ en plus de tous les « fait maison ». Les cross-linker sont divisés en plusieurs catégories en fonction : (i) de leurs fonctions réactives ; (ii) de leurs tailles ; et (iii) d'autres fonctions s'ils en possèdent.

Différentes familles de fonctions réactives sont, à ce jour, répertoriées dont les principales sont représentées ci-dessous (Fig. 2). Les agents pontants homo-bifonctionnels N-hydroxysuccinimide (NHS) sont prédominants pour le cross-link de protéines (avec le DSS, BS3, BS2G, DSG, DSSO, DSBU, CBDPS et PIR), ils réagissent principalement avec les Lys⁷⁸. Il existe cependant d'autres fonctions telles que :

- les thiols-réactifs qui permettent le pontage des Cys, un avantage pour des protéines riches en ponts disulfures
- les photo-réactifs qui peuvent cibler une large gamme de résidus via l'exposition à la lumière UV, avec principalement des cross-linker hétéro-bifonctionnels contenant une partie diazirine ou benzophénone ainsi qu'un groupe réactif amine ou thiol¹⁸²
- les réactifs aux carboxyles (avec le DMTMM) sont très prisés sachant l'abondance de ces groupes au sein des protéines. Les résidus impliqués sont beaucoup moins flexibles que les Lys faisant de l'EDC un agent pontant « zero-length » très utilisé malgré son activité optimale à un bas pH de 5,5¹⁸³
- le formaldéhyde, l'un des plus anciens agents pontants, qui a récemment regagné de l'intérêt pour le XL-MS *in vivo*, car pénètre facilement la membrane cellulaire¹⁸⁴.

| Cross-linker | Structure | Caractéristiques | Cross-linker MS-clivable | Structure | Caractéristiques |
|--------------|-----------|---|--------------------------|-----------|---|
| DSS | | Taille : 12Å Xler marqué disponible | DSSO | | Taille : 10,1Å |
| BS3 | | Taille : 12Å Xler marqué disponible | DSBU | | Taille : 12,5Å |
| EDC | | Zero-length | CBDPS | | Taille : 15Å Enrichissable par purification d'affinité Xler marqué disponible |
| DMTMM | | Zero-length | PIR | | Taille : 43Å Enrichissable par purification d'affinité |
| DSG | | Taille : 7,7Å Xler marqué disponible | | | |
| BS2G | | Taille : 7,7Å Xler marqué disponible | | | |

Figure 2 : Principales familles de cross-linker utilisés avec leurs structures, tailles et autres caractéristiques (adapté de ¹⁸⁵).

Les tailles des agents pontants sont très diverses allant du « zero-length », en réalité du 2Å, jusqu'à plusieurs dizaines d'angström (43Å pour le PIR). Les « zero-length » représentent un atout pour la confirmation d'une structure 3D, car l'intervalle des distances C α -C α n'est alors que définie par la flexibilité de la chaîne carbonée des Lys, comme dans le cas de cross-linker NHS-ester. Des agents pontants de plus grande taille présentent l'avantage de pouvoir aisément posséder une autre fonction permettant l'enrichissement des peptides pontés^{71, 185}.

D'autres fonctions peuvent être présentes au niveau du bras espaceur d'un agent pontant avec : (i) des liaisons de type peptidique, MS-clivables par la fragmentation par CID par exemple¹⁸⁶ ; (ii) des fonctions enrichissables comme pour le PhoX, enrichissable par chromatographie d'affinité pour les métaux immobilisés (IMAC)¹⁸⁷ ou encore le NNP9, enrichissable par purification d'affinité (PA) sur des billes d'agarose de streptavidine¹⁸⁸⁻¹⁸⁹. Le CBDPS possède quant à lui la capacité d'être MS-clivable, mais aussi enrichissable grâce à un groupement biotine par affinité avec de l'avidine¹⁹⁰, comme le NNP9. Cependant et à cause de cette fonction, la taille du CBDPS est plus grande, ce qui peut induire des difficultés pour le pontage de zone non-accessible au solvant.

2. Types de peptides pontés

Les agents pontants sont principalement bifonctionnels et la réaction de pontage entraîne donc plusieurs produits différents avec les peptides pontés (Fig. 3)¹⁹¹ :

- Pontage **inter**, pontage de deux peptides provenant de protéines différentes. Ils peuvent représenter des régions d'interactions.
- Pontage **intra**, pontage de deux peptides provenant de la même protéine permettant de définir des régions proches spatialement ou la formation de multimères. Parmi les pontages-intra deux cas particuliers peuvent être identifiés avec : (i) les pontages **homéotypiques**, entre deux sites identiques de deux peptides de la même protéine ce qui traduit de l'homodimérisation ; (ii) et les pontages « **loop** », cross-link de deux sites au sein du même peptide.
- Pontage **mono ou « deadend »**, pontage d'un peptide, la seconde fonction réactive soit n'a pas réagi et reste libre, ou a réagi avec de l'eau. Ces peptides indiquent les régions accessibles au solvant.

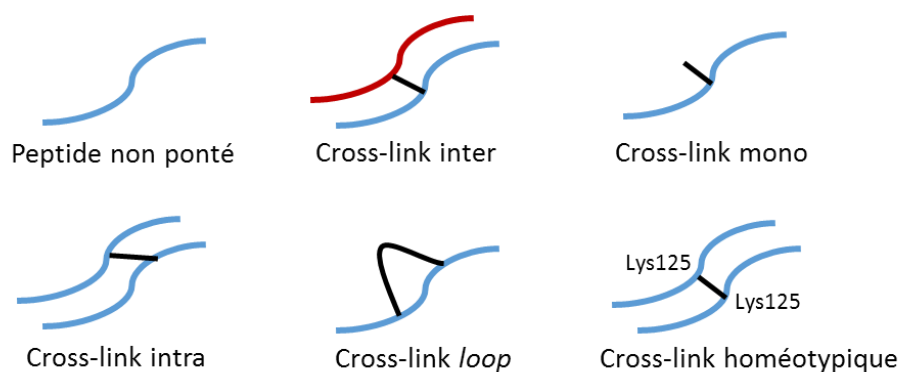


Figure 3 : Représentation schématique des différents types de peptides (bleu ou rouge pour représenter deux protéines différentes) cross-linkés.

3. Expériences de XL-MS

Du fait de la grande diversité d'agents pontants et des conditions de pontage spécifiques au cross-linker (température, temps, pH) et plus particulièrement à son rendement de réaction, il est difficile de décrire toutes les réactions associées (Annexe 10). Toutefois, la réaction de pontage avec le cross-linker le plus utilisé (NHS-ester) est montré avec la figure 4.

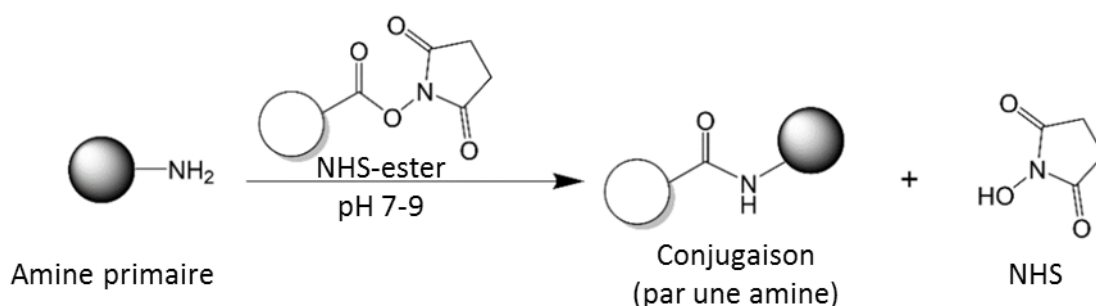


Figure 4 : Chimie du N-hydroxysuccinimide (adapté de ¹⁹²).

Dans le but d'optimiser une réaction de pontage, il est nécessaire de tester plusieurs paramètres tels que ceux susmentionnés, mais également de la quantité d'excès de réactif. À ce jour, l'excès molaire optimal d'agent pontant doit être déterminé expérimentalement pour chaque système étudié, un excès de 100 par rapport à la concentration de la protéine est recommandé pour débiter et peut être évalué par visualisation par électrophorèse sur gel monodimensionnel de polyacrylamide en présence de dodécylsulfate de sodium (gel SDS-PAGE)⁷².

Après le pontage un fractionnement des protéines peut être réalisé à l'aide d'un gel SDS-PAGE ou BN-PAGE (pour *Blue Native Polyacrylamide Gel Electrophoresis*) non dénaturant¹⁹³ sachant que la réaction de cross-link est peu efficace 1-5%¹⁸⁵ (Fig. 5).

La digestion enzymatique des protéines est ensuite réalisée sur gel ou en solution avec en grande majorité de la trypsine, mais aussi des protéases non-spécifiques telles que la pepsine (Fig. 5).

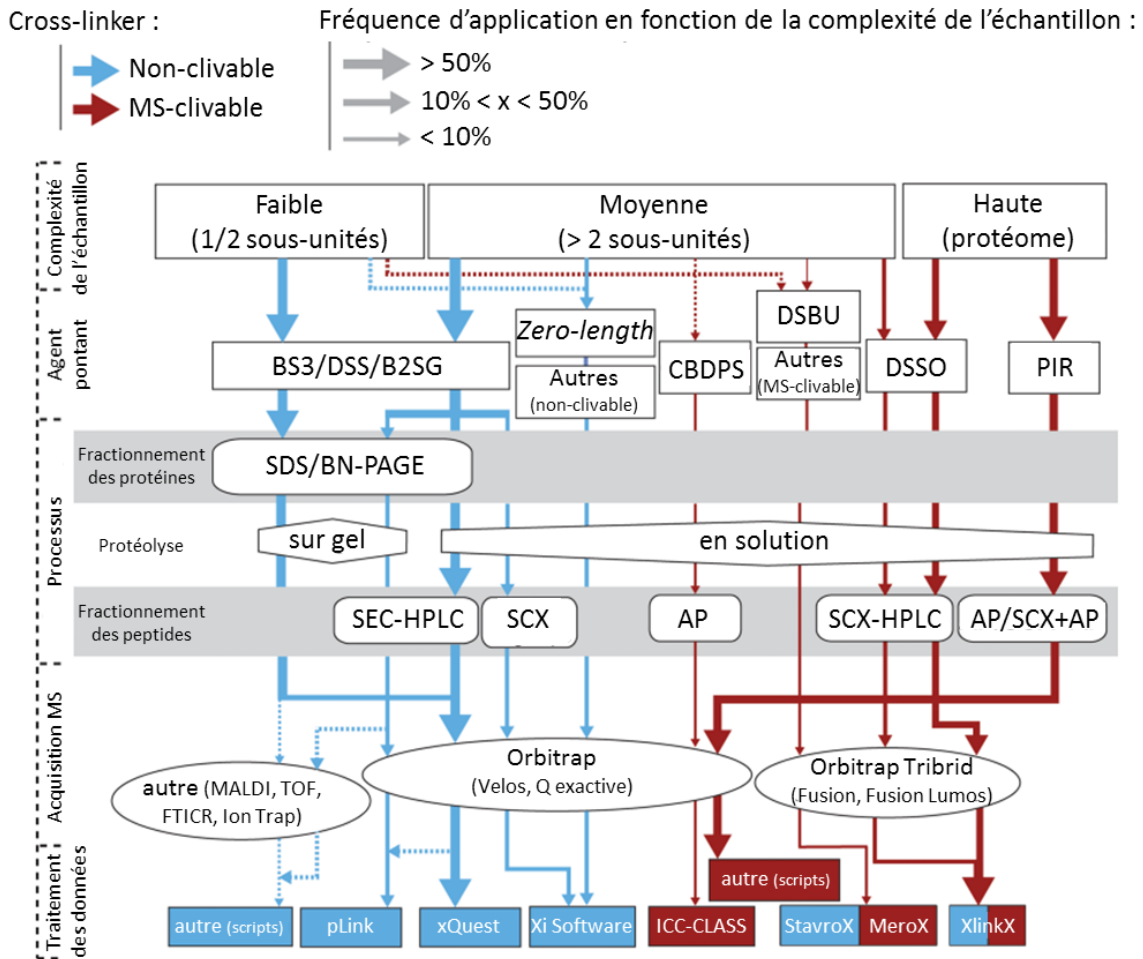


Figure 5 : Étapes du XL-MS déjà répertoriées dans la littérature et fréquences d'applications en fonction de la complexité de l'échantillon, de l'agent pontant, des processus, de l'acquisition MS et du traitement de données (adapté de ¹⁸⁵).

L'enrichissement des peptides pontés est à considérer sachant que l'abondance des peptides pontés est bien inférieure à celles des peptides non pontés. Plusieurs méthodes d'enrichissements sont disponibles avec : (i) la chromatographie d'exclusion stérique (SEC) permettant la séparation des peptides en solution selon leurs « tailles », (ii) la chromatographie échangeuses de cations (SCX) est souvent utilisée, car les peptides pontés portent généralement des charges positives plus nombreuses en raison de leurs deux extrémités N- et C-terminales et (iii) la PA en fonction du cross-linker utilisé, qui peut être employée seule ou combinée avec la SEC et le SCX pour maximiser l'enrichissement¹⁹⁴ (Fig. 5).

La plupart du temps, une source ESI est utilisée pour l'ionisation des ions, mais des expériences avec des instruments possédant une source MALDI (désorption/ionisation laser assistée par matrice) sont aussi réalisées (Fig. 5). Différents instruments peuvent être utilisés pour analyser les expériences de XL-MS tels que le FT-ICR (Solarix et Velos FTICR), les trappes à ion (LTQ XL Linear Ion Trap, Thermo Fisher Scientific) et le Q-TOF (Synapt G2 SI). Cependant, l'acquisition MS est dans la majorité des cas effectuée avec des spectromètres de masses Orbitrap (Q-exactive, Velos, Fusion, Fusion Lumos, Thermo Fisher Scientific) à haute résolution (120 000 à 200 m/z ou 60 000 à 400 m/z en MS ; 15 000 ou 30 000 en MS/MS)⁷⁴. L'acquisition MS peut aussi être optimisée avec les paramètres de fragmentations adéquats¹⁹⁵ ainsi que la sélection des états de charges¹⁷⁹ sachant que les peptides pontés sont des di-peptides, et donc de charges plus élevés que les peptides non-pontés¹⁹⁶.

4. Traitement de données

De nombreux logiciels (Tableau 1) sont à disposition pour l'identification des peptides pontés qu'ils soient MS-clivables ou non, tout en appliquant ou non un FDR (« False Discovery Rate »)¹⁸². La visualisation des résultats peut également s'effectuer avec des logiciels (similaires ou distincts) mettant en avant trois types de représentations¹⁹⁴ : des réseaux d'interactions, des cartes 2D et les pontages sur la structure 3D des protéines étudiées (Fig. 6). Enfin, des modèles de la structure 3D des protéines peuvent être générés en considérant les contraintes de distances apportées par le XL-MS¹⁹⁷.

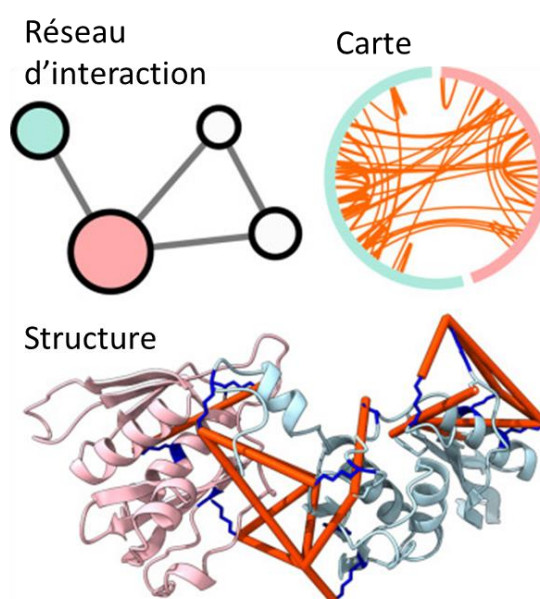


Figure 6 : Types de représentations des résultats de XL-MS avec les réseaux d'interactions, cartes et structure (adapté de ¹⁹⁸).

| logiciel | cross-linkers | | Identification des peptides | FDR | Visualisation | | Modèle | Référence |
|--------------------|---------------|--------------|-----------------------------|-----|---------------|-------|--------|-----------|
| | non clivable | MS clivables | | | réseau | carte | | |
| xQuest/xProphet | x | x | x | x | | | | 196 |
| pLink | x | x | x | x | | | | 199 |
| Protein Prospector | x | | x | x | | | | 200 |
| Kojak | x | | x | x | | | | 201 |
| SIM-XL | x | | x | x | x | x | | 202 |
| Xilmass | x | | x | x | | | | 203 |
| Xlsearch | x | x | x | x | | | | 204 |
| ECL2.0 | x | | x | x | | | | 205 |
| StavroX | x | | x | x | | | | 206 |
| MaxQuant | x | x | x | x | | | | 207 |
| xTract | x | | | | | | | 208 |
| pQuant | x | | | | | | | 209 |
| mMass | x | | | | | | | 210 |
| XiQ | x | | | | | | | 211 |
| CLMSVault | x | | | x | x | x | | 212 |
| xVis | x | | | | x | | | 213 |
| MeroX | | x | x | x | | | | 214 |
| XlinkX | | x | x | | | | | 215 |
| xiNET | | | | | x | | | 216 |
| ProXL | | | | | x | x | | 217 |
| XlinkDB 2.0 | | | | | x | x | x | 218 |
| Xwalk | | | | | | x | | 219 |
| Xlmap | | | | | | x | | 220 |
| Xlink analyzer | | | | | | x | x | 221 |
| XL-MOD | | | | | | | x | 222 |
| DynaXL | | | | | | | x | 223 |
| ROSETTA | | | | | | | x | 224 |
| HADDOCK | | | | | | | x | 225 |
| IMP | | | | | | | x | 226 |

Tableau 1 : Tableau récapitulatif des logiciels accessibles au public pour l'identification, la visualisation et/ou la modélisation des pontages (d'après ¹⁹⁷, où les références et sites web sont disponibles).

II. Recommandations et innovations

Une étude inter-laboratoires (70 auteurs et 30 laboratoires) en XL-MS portant sur l'analyse d'albumine de sérum bovin (BSA) a permis de mettre en place les premières recommandations⁷⁴. Certaines lignes directrices ont été posées suite à cette étude, telles que :

- L'application d'un FDR pour que la validation des peptides pontés soit satisfaisante.
- Une précision de masse doit être définie dans le but de réduire les fausses attributions des ions correspondants aux peptides pontés.

- Les conditions expérimentales du XL-MS, les paramètres instrumentaux et filtres appliqués lors du traitement de données doivent être reportés en détail dans un tableur excel type mis à disposition en information supplémentaire de l'article⁷⁴.
- Les données doivent être partagées dans le référentiel PRIDE¹⁴¹.
- La réaction de pontage chimique doit être évaluée grâce au gel SDS-PAGE par exemple.
- La sélectivité des cross-linkers (les sites possibles de XL) est à prendre en compte en fonction de la complexité des échantillons étudiés. Dans le cas des NHS-ester, le pontage des Lys est uniquement recherché pour l'analyse d'échantillons complexes alors que la Ser, Thr et Tyr sont aussi considérées comme sites potentiels de pontage pour l'analyse d'échantillons peu complexe (< 10 protéines).
- Pour l'analyse de la BSA, 80 peptides pontés doivent être identifiés pour considérer la réaction comme efficace, suffisante et optimale.

En 2020, une autre publication a permis de mettre en place d'autres recommandations complémentaires avec notamment la réalisation de triplicatas expérimentaux pour tenir compte de la reproductibilité de la réaction²²⁷.

L'approche de XL-MS est en pleine expansion avec l'analyse d'échantillons complexes et la modélisation de structures 3D. Plusieurs avancées ont vu le jour ces dernières années, avec :

- Un agent pontant enrichissable sur IMAC. Le PhoX a été développé et commercialisé, qui permet d'enrichir 300 fois les peptides pontés au sein des échantillons et ayant une spécificité de 97% pour l'enrichissement¹⁸⁷. De plus, la réaction de pontage et le protocole d'enrichissement peuvent être automatisés via l'utilisation du robot de préparation AssayMAP Bravo (Agilent) apportant une plus grande confiance quant à la répétabilité des étapes pré-analytiques.

- Des instruments de nouvelles générations ont déjà montré leurs efficacités pour l'analyse en XL-MS avec plus particulièrement le TimsTOF Pro (Bruker) comprenant une cellule de mobilité ionique Tims. Cet instrument a permis d'augmenter considérablement le nombre de peptides pontés identifiés (623 sites uniques pour la BSA contre 80 en moyenne, 670 pour le ribosome E. coli et 1623 uniques pour les lysats nucléaires, correspondant à 1088 interactions intra- et 535 interactions inter-protéines)²²⁸.

- L'étude des interactions protéine-acides nucléiques en XL-MS avec par exemple l'agent pontant 1,2:3,4-diepoxybutane (DEB). Sachant que de nombreux processus biologiques

importants dans les cellules sont réalisés par des complexes à plusieurs sous-unités constitués de protéines et d'ARN, où les techniques à haute résolution plus établies pourraient ne pas être en mesure de produire des données concluantes, il est intéressant de pouvoir réaliser un pontage des protéines avec de l'ADN ou de l'ARN²²⁹⁻²³².

- Enfin, la modélisation de structure 3D de protéines ou complexes impliquant des protéines, en considérant en données d'entrée les contraintes de distances apportées par les expériences de XL-MS, restent difficiles à mettre en œuvre sans avoir de connaissance préalable des logiciels Rosetta, HADDOCK ou IMP. Cependant, un nouvel outil, appelé AlphaLink vient d'être développé pour la prédiction de structures à partir des algorithmes AlphaFold2 et des contraintes de distances apportées par les peptides photo-pontés uniquement²³³. Il ne fait aucun doute que cet outil sera encore optimisé pour la prise en compte de produits d'autres types de cross-linkers dans les prochaines années.

III. Développements méthodologiques en XL-MS réalisés dans le cadre de mon travail de thèse

1. Phase pré-analytique

Le protocole de pontage chimique mis en place au LSMBO est réalisé suivant les étapes décrites dans la publication de C. Iacobucci⁷² avec de l'urée dibutyrique disuccinimidyl (DSBU) ou de l'urée diacétique disuccinimidyl (DSAU). Ces deux agents pontants (cross-linker) MS-clivables contiennent deux extrémités NHS-ester réagissant avec les amines primaires principalement, mais aussi avec les groupes hydroxyles²³⁴. Le DSBU (Fig. 7) possède un bras espaceur de 11 atomes avec deux chaînes carbonées (C4) pour une longueur de 12,5Å et une masse de 426,38 Da quand le DSAU lui est plus court (C2 ; 7,7Å ; 370,27 Da).

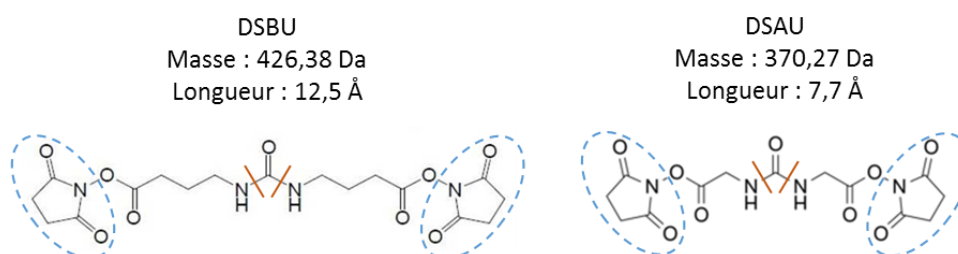


Figure 7 : Formule topologique du DSBU et DSAU avec leurs masses et longueurs. Les extrémités NHS ester sont entourées en bleu et les liaisons MS-clivables montrées en orange.

La réaction de pontage des protéines (environ 20µg de protéine) s'effectue pendant 45min à température ambiante (T_{amb}) avec un excès molaire de cross-linker de 50 à 200 au préalable

dilué dans du DMSO (Fig. 8). La quantité de DMSO finale n'excède pas les 10% (v/v). Cette réaction est réalisée à un pH entre 7,2-8,0 dans un tampon de même composition que celui de stockage sauf s'il contient du Tris-HCl ou des agents réducteurs à base de thiols (DTT, 2-Mercaptoéthanol, etc.) car ils réagissent avec l'agent pontant. Une fois le pontage effectué, la réaction est stoppée (*quench*) par l'ajout de bicarbonate d'ammonium à une concentration de 20mM pendant 20min à T_{amb} dans le but d'hydrolyser les agents pontants libres en solution. L'échantillon est ensuite coupé en deux, d'abord pour continuer les étapes pour l'analyse ($\approx 15\mu\text{g}$) et, d'une autre part, pour faire un contrôle visuel de la réaction de cross-link ($\approx 5\mu\text{g}$) comparé à un échantillon non-ponté par gel SDS-PAGE à 10% de SDS (Annexe 11). Les protéines pontées sont ensuite réduites et alkylées avec 5mM de DTT pendant 30min à 37°C et 15mM d'iodoacétamide durant 1h dans le noir et à T_{amb} . La digestion des protéines est par la suite réalisée avec un mix de trypsine/Lys-C (Promega, V5073) à un ratio de 1:50 (enzyme:substrat, en g/L) pendant la nuit à 37°C. Les peptides sont enfin concentrés par extraction sur phase solide (SPE, Sep-Pak C18 1cc, 50mg phase) puis évaporés grâce à un concentrateur à vide (Savant SC100 SpeedVac, Thermo Fisher Scientific) avant la remise en suspension dans une solution de $\text{H}_2\text{O}/\text{ACN}/\text{HCOOH}$ (98/2/0,1). Le pontage est entièrement réalisé en triplicata pour chaque expérience suivant les recommandations²²⁷.

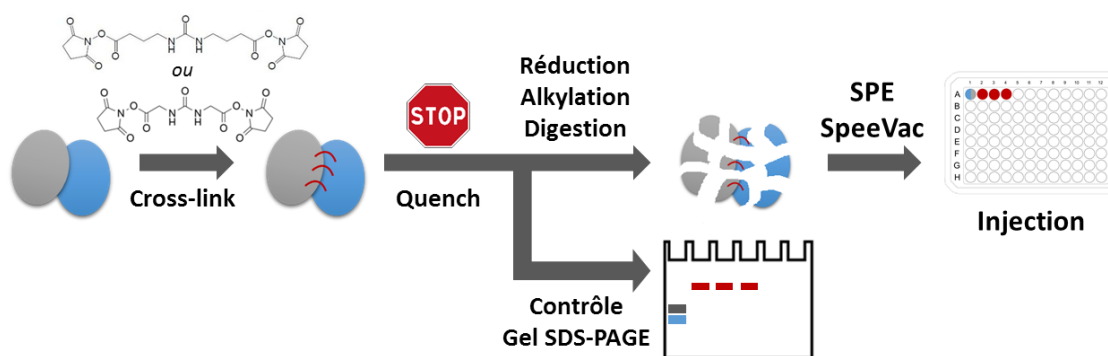


Figure 8 : Description schématique de la phase pré-analytique en XL-MS : pontage chimique, *quench*, réduction, alkylation et digestion puis concentration via la SPE et le SpeeVac. Après l'étape de quench, une partie de l'échantillon est réservée au contrôle par gel SDS-PAGE.

2. Phase analytique

L'injection de 50-200ng peptides est ensuite réalisée en LC-MS/MS (Q Exactive HF-X Orbitrap ou Q-Exactive Plus Orbitrap, Thermo Fisher Scientific, Annexe 12). Chacun des deux systèmes sont couplés à une nanoAcquity Ultra-Haute Performance LC (Waters) permettant l'injection, la concentration des peptides sur une pré-colonne (C18, 180 μm x 20mm, 5 μm , Waters) et la séparation des peptides sur une colonne de phase inverse (C18, 75 μm x 250mm, 1,7 μm ,

Waters). Les solvants d'élutions sont composés d'eau avec 0,1% de HCOOH (A) et d'ACN avec 0,1% de HCOOH (B) à un débit de 400nL/min et une température de colonne de 60°C.

Les peptides élués sont finalement transmis au spectromètre de masse (source nano-ESI) comprenant une géométrie Q-orbitrap. La température de source a été fixée à 250°C et l'acquisition est réalisée en mode positif sur une gamme de m/z de 300 à 1800 avec une valeur *AGC target* (« Automatic Gain Control target ») fixée à 3×10^6 (correspondant à la population dans l'analyseur). La calibration interne s'effectue à l'aide de l'ion polysiloxane de l'air ambiant à 445.12 m/z . Une méthode DDA (*Data Dependent Acquisition*) de type top 10 est utilisée : les 10 ions les plus intenses sont sélectionnés et isolés par fenêtre de 2 m/z et sont fragmentés avec une cellule HCD (énergie de collision normalisée de $30 \pm 3\%$). En plus de la fragmentation des peptides, les cross-links MS clivables seront également fragmentés au niveau de leurs liaisons peptidiques (Fig. 8). Les spectres MS/MS ont été acquis avec une valeur *AGC target* fixée à 1×10^5 et d'exclusion dynamique de 60s. Tous les paramètres LC, MS et MS/MS qui diffèrent suivant l'instrument utilisé sont répertoriés dans le Tableau 2. Les données ont été converties au format .mgf avec le logiciel MS Convert GUI 3.0.21295 (ProteoWizard).

| spectromètre de masse | Q-Exactive Plus Orbitrap | Q Exactive HF-X Orbitrap |
|---------------------------------------|---|--|
| LC | | |
| pré-colonne | nanoACQUITY UPLC precolumn (Waters) | nanoEase M/Z Symmetry C18 (Waters) |
| colonne | nanoACQUITY UPLC column (Waters) | nanoEase M/Z Peptide BEH C18 (Waters) |
| Gradient | 1-8 % B (2 min), 8-35 % B (77 min), 35-90 % B (1 min), 90 % B (5 min), 90-1 % B (2 min), 1 % B (2 min) | 3 % B (15 min), 3-40 % B (90 min), 40-99 % B (2 min), 99 % B (5 min), 99-3 % B (2 min) |
| Temps (min) | 105 | 135 |
| MS | | |
| Voltage du capillaire (kV) | 1,8 | 1,9 |
| Temps d'injection maximal (ms) | 50 | 100 |
| Résolution (200 m/z) | 140000 | 120000 |
| MS/MS | | |
| Résolution (200 m/z) | 35000 | 30000 |
| Etat de charges | 2-8 | 2 minimum |
| Temps d'injection maximal (ms) | 100 | 200 |

Tableau 2 : Description des paramètres LC, MS et MS/MS appliqués suivant le spectromètre de masse utilisé.

3. Phase post-analytique

Les données au format .mgf sont ensuite exportées dans le logiciel MeroX²¹⁴ 2.0.1.4. Les peptides ont été identifiés à partir d'une banque de séquence « faite maison » contenant les séquences d'une centaine de contaminants connus (kératines, protéases, etc.) et des protéines analysées, ainsi que ces mêmes séquences inversées. Suite à la digestion par la Trypsine/Lys-C, les sites de clivage sont les K et R en acceptant 3 « *missed cleavages* ». Les longueurs minimales et maximales des peptides sont de 5-30 résidus dont la charge minimale est de 2. Les modifications fixes et variables prises en compte sont respectivement la carbamidométhylation des Cys et l'oxydation des Met. Les di-peptides pontés sont déterminés considérant les Lys et l'acide aminé N-terminal comme premier site ponté ainsi que les Ser, Thr et Tyr en plus des précédents comme site secondaire. La précision de mesure de masse a été fixée à 5 ppm pour l'ion précurseur et à 10 ppm pour les ions produits MS/MS. Seuls les types d'ions a, b et y ont été recherchés. Un FDR de 5 % a été choisi, une intensité relative de 10 %, un seuil de score de 30 (score déterminé par le logiciel), un rapport signal/bruit (S/N) de 2 et le mode RISEUP a été choisi comme mode d'analyse considérant le nombre d'entrées dans la banque de séquence et recommandé par MeroX⁷². Les données ont finalement été validées manuellement en fonction des fragments caractéristiques attendus par ce type de cross-linker (Fig. 9).

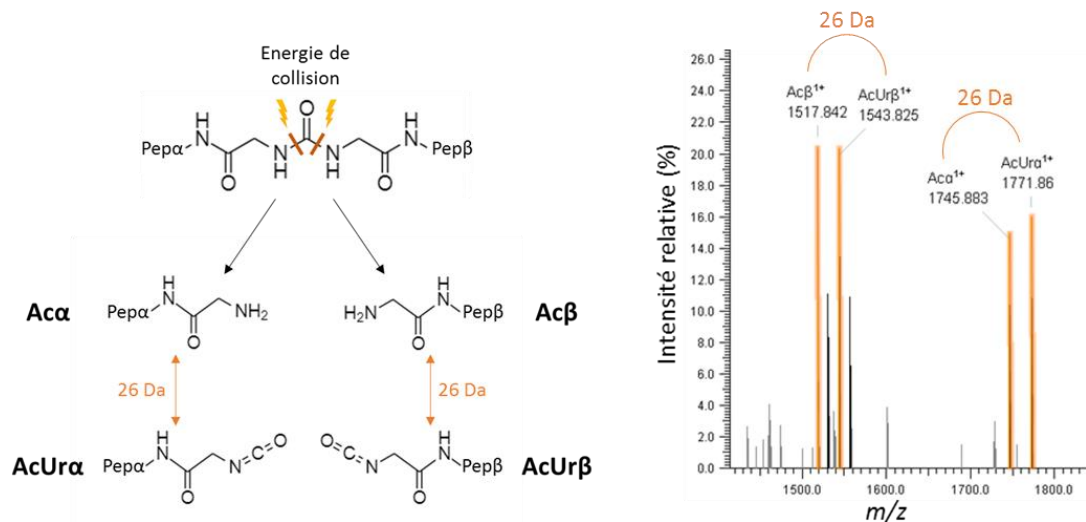


Figure 9: Produits de la fragmentation du DSAU, un cross-linker MS clivable et spectre MS/MS des fragments doubles caractéristiques.

Seuls les di-peptides pontés, possédant au moins un des deux fragments caractéristiques pour chaque peptide, sont conservés ainsi que ceux présents dans au moins 2 des 3 répiclats. Quand

une structure 3D est disponible, les distances $\text{C}\alpha\text{-C}\alpha$ sont mesurées entre les acides aminés cross-linkés grâce au logiciel Pymol et doivent être comprises entre 6-30Å pour le DSBU et 6-26Å pour le DSAU.

IV. Mise au point d'un standard pour évaluer le protocole de XL-MS

Dans le but de vérifier l'efficacité de notre protocole de pontage et d'implémenter un contrôle qualité de la réaction, des expériences de XL-MS sur de la BSA (Annexe 2) ont été menées avec du DSBU. Sachant qu'une étude inter-laboratoire a été réalisée sur cette protéine suivant différents protocoles menant à la conclusion qu'environ 80 peptides pontés doivent être identifiés pour que le cross-link soit jugé efficace⁷⁴. Un triplicata expérimental de BSA a donc été analysé sur chacun des deux instruments (Q Exactive HF-X Orbitrap ou Q-Exactive Plus Orbitrap). La BSA (30µg) a été pontée dans un tampon (25mM K_2HPO_4 , 25mM KH_2PO_4 et 0,5mM TCEP) à un pH de 7,8 avec 200 excès molaire de DSBU. Ces analyses ont permis l'identification de 122 et 97 peptides pontés (>80) présents dans au moins 2 des 3 réplicats (Fig. 10). Une structure du dimère de la BSA est disponible (PDB : 6QS9) et bien que le dimère est présent en faible abondance (15% déterminé en M-Ph, Annexe 13), cette structure permet de mesurer les distances $\text{C}\alpha\text{-C}\alpha$ entre les résidus impliqués dans le cross-link. Parmi ces mesures, 71% entrent dans la gamme de distance (6-30Å) montrant que les peptides pontés sont en accords avec la structure. La reproductibilité de la réaction est bonne (66-67%) connaissant le faible rendement de réaction ($\approx 5\%$) et la faible abondance des peptides pontés par rapport aux peptides non-pontés¹⁸⁷.

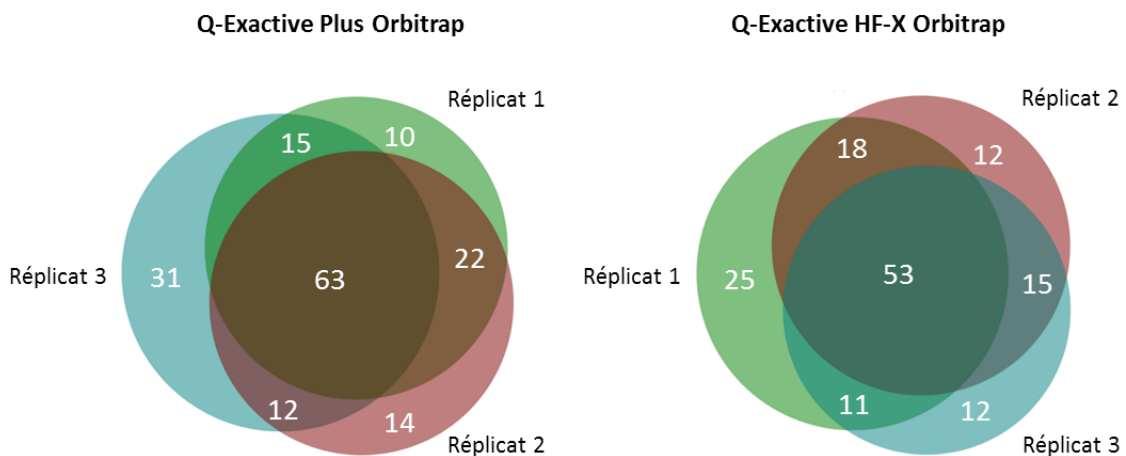


Figure 10 : Diagrammes de Venn (DeepVenn) des peptides cross-linkés pour chaque réplicat déterminés avec le Q-Exactive Plus Orbitrap (gauche) et le Q Exactive HF-X Orbitrap (droite).

V. Optimisations et développements réalisés dans le contexte de cette thèse

1. Préparation et analyse

Dans un premier temps, l'optimisation de la réaction de pontage est estimée visuellement sur gel SDS-PAGE (dénaturant) nécessitant une petite quantité de protéine (1-5 μ g) pour observer le cross-link par rapport à un échantillon contrôle (non-ponté). La diminution de l'intensité des bandes correspondantes aux protéines non pontées est un indicateur du pontage covalent des protéines. En parallèle, l'apparition d'une nouvelle bande du complexe covalent cross-linké est un indicateur du bon déroulement de la réaction.

La réaction de pontage doit-elle-même être optimisée pour chaque système étudié. Pour illustrer l'importance de cette étape de pontage chimique, je vais décrire l'optimisation de la réaction dans le cadre de l'analyse du complexe PRMT9/SAP145₄₀₁₋₅₅₀ (Annexe 2). Plusieurs temps (15 ou 45min), tampons (20mM K₂HPO₄, 20mM KH₂PO₄ et 0,5mM TCEP ou 20mM HEPES, 100mM NaCl et 0,5mM TCEP) et températures (4°C, T_{amb} et 37°C) de la réaction de pontage avec du DSBU (1:200) ont donc été testés sur gel SDS-PAGE avec le complexe de PRMT9 (95 kDa) et de SAP145₄₀₁₋₅₅₀ (18 kDa). Les résultats obtenus (Fig. 11A) révèlent bien la diminution de l'intensité des bandes des deux protéines par rapport à celles du contrôle et l'apparition de bandes de masses supérieures indiquant le pontage. Bien que le gel ne présente pas de différences majeures entre toutes les conditions testées, il est tout de même possible de remarquer une disparition plus importante des bandes correspondantes à SAP145₄₀₁₋₅₅₀ pour le pontage avec le tampon composé de HEPES ainsi qu'une plus forte intensité pour les bandes de masses supérieures à celle de PRMT9 pour le XL à T_{amb} pendant 45min. Différents excès molaires de DSBU ont été également testés (25, 50, 100 et 200) pour chacun des tampons (Fig. 11B) montrant que plus il y a d'agent pontant plus le cross-link est visible comme attendu. Finalement, les optimisations sur gel ont permis de sélectionner la température (T_{amb}) et le temps de réaction (45min) avec un cross-linker en excès molaire de 200 (Fig. 11C). Néanmoins, les expériences sur gel n'ont pas permis de déterminer un agent pontant ou tampon préférentiels.

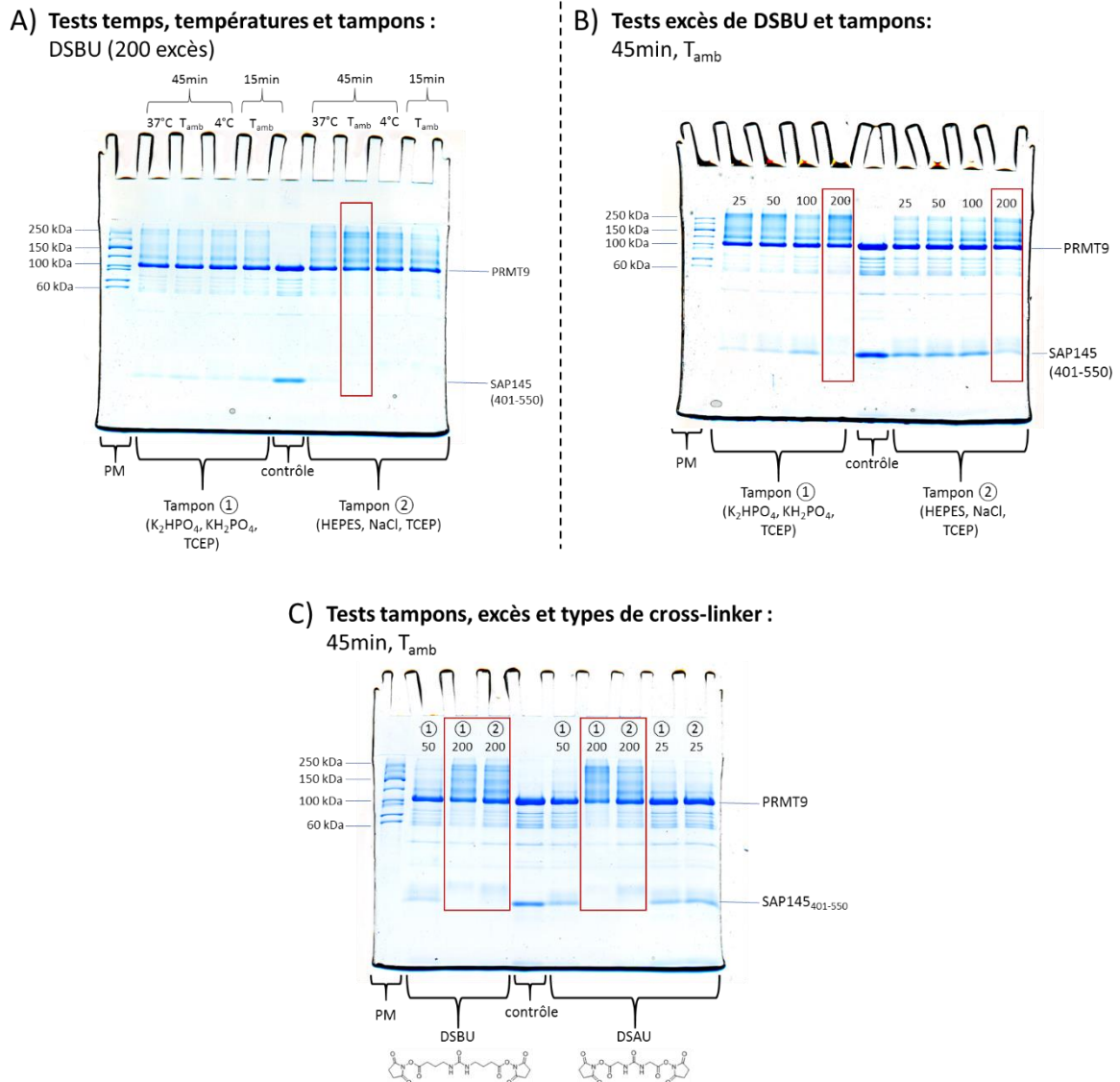


Figure 11 : Gels SDS-PAGE montrant les différentes optimisations réalisées de la réaction de pontage chimique sur le complexe PRMT9/SAP145₄₀₁₋₅₅₀. Les conditions de pontage sont décrites pour chacun des puits du gel sachant que le premier puits contient un marqueur de poids moléculaires (PM).

Les protéines pontées à T_{amb} pendant 45min avec un excès de 200 avec du DSAU ou DSBU dans chacun des tampons décrits ci-dessus ont été analysées en triplicat sur le Q Exactive HF-X Orbitrap (Tableau 3). Les analyses ont permis de montrer que pour chacun des cross-linkers un autre tampon, HEPES pour le DSBU et K_2HPO_4 , KH_2PO_4 pour le DSAU, permet d'obtenir le plus grand nombre de peptides pontés identifiés. La variabilité du nombre de peptides pontés identifiés pour le même cross-linker est importante (de 31 à 96) montrant que le rendement de la réaction de pontage n'est pas dépendant du type de fonction réactive (NHS ester) de l'agent pontant. L'ensemble de ces résultats montre qu'hormis la température et le temps de pontage, toutes les autres conditions (tampon, type et excès de cross-linker) doivent être

testées et optimisées indépendamment dans le but d'obtenir le plus grand nombre de peptides pontés identifiés. Le choix du tampon pour effectuer les expériences de XL-MS est un paramètre important et permet dans certains cas de tripler le nombre de peptides pontés identifiés (31 à 93 peptides pour le pontage du complexe PRMT9/SAP145₄₀₁₋₅₅₀ avec le DSBU).

| Cross-linker | Réplikat | Tampon | Nombre de XL inter-PRMT9/SAP145 | Nombre de XL intra-protéine | Nombre de XL total |
|--------------|----------|---|---------------------------------|-----------------------------|--------------------|
| DSBU | 1 | HEPES | 50 | 43 | 93 |
| DSBU | 2 | HEPES | 52 | 44 | 96 |
| DSBU | 3 | HEPES | 45 | 40 | 85 |
| DSBU | 1 | K ₂ HPO ₄ , KH ₂ PO ₄ | 18 | 14 | 32 |
| DSBU | 2 | K ₂ HPO ₄ , KH ₂ PO ₄ | 24 | 17 | 41 |
| DSBU | 3 | K ₂ HPO ₄ , KH ₂ PO ₄ | 15 | 16 | 31 |
| DSAU | 1 | HEPES | 27 | 35 | 62 |
| DSAU | 2 | HEPES | 12 | 22 | 34 |
| DSAU | 3 | HEPES | 29 | 33 | 62 |
| DSAU | 1 | K ₂ HPO ₄ , KH ₂ PO ₄ | 42 | 34 | 76 |
| DSAU | 2 | K ₂ HPO ₄ , KH ₂ PO ₄ | 42 | 22 | 64 |
| DSAU | 3 | K ₂ HPO ₄ , KH ₂ PO ₄ | 44 | 23 | 67 |

Tableau 3 : Nombre de peptides cross-linkés (inter, intra et total) identifiés pour chaque condition (cross-linkers et tampons) et réplikat.

Dans l'idée de diminuer le temps d'analyse (105min/réplikat) et sachant que les échantillons sont purs et peu complexes, seul 1-4 protéines différentes, un gradient LC plus court a été testé de 70min au lieu de 90min (Tableau 2). Un échantillon de BSA a été analysé avec chaque gradient, avec celui de 70min seuls 51 peptides pontés ont été identifiés contre 93 avec le gradient de 90min. Dans le but d'optimiser l'identification des peptides pour le gradient de 70min, les 20 ions les plus intenses sont sélectionnés et isolés au lieu de 10 initialement. Dans ces conditions, 47 peptides pontés ont été identifiés. Ces nouveaux paramètres n'améliorant pas l'identification de ces peptides, les analyses ont été réalisées avec les conditions initiales (paragraphe III.2.).

2. Traitement des données

2.1. Logiciels de traitements des données avec un cross-linker MS clivable

Au cours de ces dernières années, de nombreux logiciels ont été développés pour identifier les peptides pontés rendant les possibilités de traitement de données variées et nombreuses. L'emploi de cross-linkers MS-clivables réduit considérablement le nombre de logiciels

utilisables pour l'identification du type de peptides générés¹⁸⁵ avec : xQuest²³⁵, pLink²³⁶, Xi¹⁸⁶, MeroX⁷² et XlinkX²³⁷. Le logiciel MeroX a été choisi pour traiter les données de XL-MS, car il permet de vérifier manuellement les profils de fragmentations obtenus pour chaque peptide ponté identifié et la présence des fragments caractéristiques MS-clivables. En 2022, un module du logiciel MaxQuant appelé MaxLynX²³⁸ a été développé et comparé à d'autres logiciels (MeroX et XlinkX) pour l'identification des peptides pontés synthétiques ou des embryons de *Drosophila melanogaster* (mouche des fruits). Alors qu'un nombre très proche de peptides synthétiques cross-linkés a été identifié pour MaxLynX (242) et MeroX (237), une différence plus importante est observée pour le traitement des données de XL-MS des embryons avec 21% de moins d'identification avec MeroX²³⁸.

2.1.1. Principe et paramètres de MeroX

MeroX est un logiciel simple d'utilisation et permet, à partir des données au format .mgf, des paramètres définis et d'un fichier fasta, d'identifier les peptides pontés. Les paramètres utilisés sont décrits dans le paragraphe III.3. de ce chapitre. MeroX calcule les masses de deux peptides cross-linkés à partir des fragments doubles correspondants dans les spectres MS/MS et compare les masses des peptides aux protéines digérées *in silico*. Un score est défini en fonction de l'intensité et du nombre de fragments identifiés des peptides pontés par rapport au nombre d'ions identifiés au sein du spectre, le score minimal est placé à 30 par défaut. La qualité de l'identification peut être contrôlée grâce à l'application d'un taux de faux positif (FDR) placé à 5% suivant les recommandations⁷⁴. Cependant, si ce taux est appliqué avec une petite banque de données (<10 protéines) l'estimation ne sera plus pertinente. C'est pourquoi, les séquences de contaminants connus (\approx 100 entrées) ont été ajoutées à notre banque de données. MeroX calcule également la probabilité pour chaque combinaison possible de sites pontés et sites secondaires au sein du même peptide en fonction de l'intensité de tous les ions fragments identifiés pour chaque site possible⁷². Ce calcul de probabilité est un réel avantage par rapport aux autres logiciels, mais peut impliquer une vérification manuelle de tous les spectres et probabilités de sites. Cette vérification manuelle devient complexe à réaliser dans le cas de l'étude de protéome, cellules et tissus.

2.1.2. Principe et paramètres de MaxLynX

MaxLynX est un outil de traitement de données intégré dans l'environnement MaxQuant, largement utilisé en protéomique. Les paramètres sont identiques à ceux de MeroX. MaxLynX permet d'identifier les peptides pontés à partir des données brutes. Les peptides non pontés sont identifiés grâce au moteur de recherche Andromeda²³⁹ alors que pour identifier les peptides cross-linkés un autre moteur spécialisé est utilisé. De la même manière, un FDR est applicable sur la base d'un calcul de probabilité d'erreur entre les fausses correspondances peptide ponté/spectres divisé par le nombre de cibles. Deux FDR distincts peuvent être appliqués pour les peptides pontés intra et inter contrairement à MeroX. MaxLynX permet de visualiser tous les peptides identifiés dont ceux non cross-linkés (MeroX ne donne pas d'informations sur les peptides non pontés). Cependant et au contraire de MeroX, cet outil ne calcule pas de probabilité de sites pontés. Pour détecter les peptides pontés avec un cross-linker MS-clivable, trois étapes sont appliquées et filtres successivement : (i) la différence des masses des fragments doubles sur le spectre MS/MS ; (ii) la vérification de l'intensité de ces fragments (au moins un fragment double doit correspondre à un des trois pics les plus intenses) ; (iii) une approche plus « souple » où seul un fragment double est nécessaire pour identifier le peptide comme étant cross-linké. À l'inverse de MeroX, MaxLynX ne reconnaît pas les peptides homéotypiques pour l'instant²³⁸.

2.2. Evaluation de MeroX et MaxLynX sur trois jeux de données

Au vu des gains en identifications apportés par MaxLynx sur les embryons²³⁸, cette extension a donc été testée sur trois jeux de données incluant la protéine de référence (BSA) et deux complexes PRMT2/RSF1 (120 kDa) et RUVBL1-2/DPCD (R2D, 383 kDa, Annexes 2 et 14). Les résultats ont été aussi comparés à ceux obtenus avec le logiciel MeroX. Ainsi, environ deux fois plus de peptides pontés ont été identifiés avec Maxlynx dans au moins deux des trois réplicats pour chacune des protéines testées (Fig. 12, Annexe 15). Les taux de peptides pontés communs aux deux traitements sont globalement assez faibles (31, 15 et 24%). Cette disparité peut s'expliquer par l'identification des sites les plus probables de cross-link au sein d'un même di-peptide. En effet, pour un seul peptide cross-linké plusieurs sites de pontage sont disponibles et définis par MeroX et non par MaxLynX. Dans le cas du XL des deux peptides (400-412 et 460-468, Tableau 4), les sites K401-S466 (3 réplicats) et K401-K463 (1 réplicat) ont été identifiés avec Maxlynx alors que seul le site K401-K463 (2 réplicats) a été identifié avec

MeroX. Cette différence est certainement due aux probabilités des sites donnés par MeroX et le choix manuel des sites en fonction des profils de fragmentation.

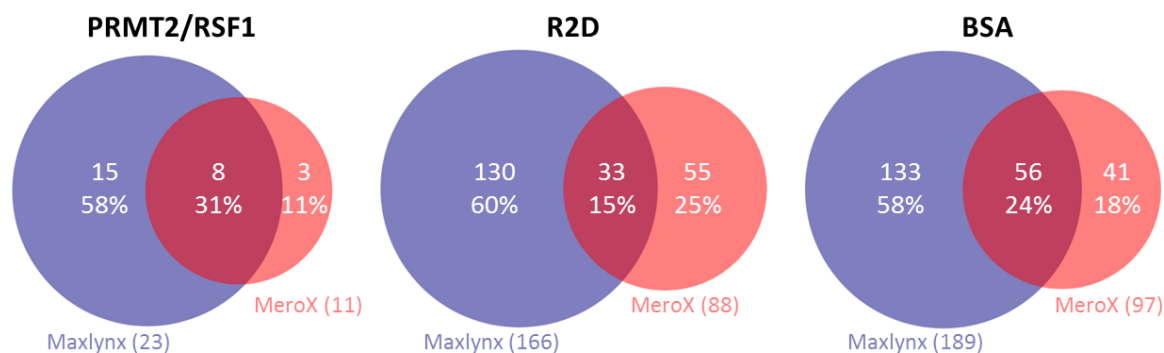


Figure 12 : Diagrammes de Venn (DeepVenn) des peptides cross-linkés déterminés pour chaque protéine (BSA) ou complexe protéique (PRMT2/RSF1 ou R2D) avec les deux logiciels de traitement, Maxlynx (violet) et MeroX (rouge).

| Logiciel | Sites | Séquence 1 | Début 1 | Fin 1 | Séquence 2 | Début 2 | Fin 2 | Réplicats | Charge | m/z | Masse |
|----------|-----------|-----------------|---------|-------|-------------|---------|-------|-----------|--------|--------|---------|
| MaxLynx | K401-S466 | [LKHLVDEPQNLIK] | 400 | 412 | [CCTKPESER] | 460 | 468 | 1 | 4 | 727,87 | 2907,46 |
| | K401-S466 | [LKHLVDEPQNLIK] | 400 | 412 | [CCTKPESER] | 460 | 468 | 2 | 4 | 727,87 | 2907,46 |
| | K401-S466 | [LKHLVDEPQNLIK] | 400 | 412 | [CCTKPESER] | 460 | 468 | 3 | 4 | 727,87 | 2907,46 |
| | K401-K463 | [LKHLVDEPQNLIK] | 400 | 412 | [CCTKPESER] | 460 | 468 | 3 | 3 | 970,16 | 2907,46 |
| MeroX | K401-K463 | [LKHLVDEPQNLIK] | 400 | 412 | [BBTKPESER] | 460 | 468 | 1 | 5 | 582,50 | 2908,47 |
| | K401-K463 | [LKHLVDEPQNLIK] | 400 | 412 | [BBTKPESER] | 460 | 468 | 1 | 4 | 727,87 | 2908,47 |
| | K401-K463 | [LKHLVDEPQNLIK] | 400 | 412 | [BBTKPESER] | 460 | 468 | 3 | 5 | 582,50 | 2908,47 |
| | K401-K463 | [LKHLVDEPQNLIK] | 400 | 412 | [BBTKPESER] | 460 | 468 | 3 | 3 | 970,16 | 2908,47 |
| | K401-K463 | [LKHLVDEPQNLIK] | 400 | 412 | [BBTKPESER] | 460 | 468 | 3 | 4 | 727,87 | 2908,47 |

Tableau 4 : Résidus (sites) impliqués dans le cross-link des peptides 400-412 et 460-468 déterminés par les logiciels Maxlynx et MeroX pour chaque ion avec leur masse, valeur de m/z, charge et séquence.

La vérification manuelle des profils de fragmentation des peptides pontés n'est réalisable qu'avec MeroX et seuls les di-peptides pontés possédant au moins un des deux fragments caractéristiques pour chaque peptide sont conservés. Ce filtrage manuel des données n'étant pas possible avec Maxlynx, mesurer les distances entre les résidus pontés pourrait indiquer la pertinence des XL identifiés avec chaque logiciel. Pour la BSA, une structure est déjà disponible (PDB : 6QS9) alors que, pour les autres complexes, des modélisations proches des structures réelles ont été obtenues par certains de nos collaborateurs ce qui permet de mesurer les distances Ca-Ca . De manière générale des taux plus importants (Tableau 5, Annexe 15), de distances mesurées comprises dans l'intervalle défini (6-30Å), sont obtenus dans le cas des XL identifiés avec MeroX qu'avec Maxlynx.

| protéine/complexe | structure | Taux de distances validées (%) | |
|-------------------|-----------------------------------|--------------------------------|---------|
| | | MeroX | Maxlynx |
| BSA | BSA (PDB : 6QS9) | 70 | 66 |
| PRMT2/RSF1 | modèle PRMT2/RSF1 | 95 | 78 |
| R2D (R1R2/DPCD) | modèle DPCD | 85 | 85 |
| | R1R2 (PDB : 5OAF ²⁴⁰) | 75 | 68 |

Tableau 5 : Taux de distances validées (comprises entre 6 et 30Å) des résidus cross-linkés identifiés avec les logiciels Maxlynx et MeroX pour chaque structure et protéine/complexe étudiés.

Ces résultats montrent que malgré une identification d'un plus grand nombre de peptides pontés avec Maxlynx, les taux de distances concordantes sont plus importants avec MeroX donnant des avantages à chacun de ces logiciels.

VI. Conclusions

J'ai d'abord mis en place un protocole de référence (BSA) de pontage chimique basé sur l'utilisation d'agents MS-clivables adapté en fonction de nos instruments et des recommandations²²⁷. Ces analyses ont montré l'efficacité de la méthode avec plus de 80 peptides pontés identifiés. Ensuite, j'ai pu montrer l'importance de l'étape d'optimisation de la réaction de pontage pour sélectionner les meilleures conditions expérimentales grâce au gel SDS-PAGE. Différents paramètres instrumentaux ont également été testés sans montrer d'amélioration majeure du nombre de peptides cross-linkés identifiés. Enfin, au vu du nombre de logiciel permettant de traiter des données de XL-MS, MeroX et Maxlynx ont été choisis, comparés et ont montré différents avantages à leur utilisation pour l'identification de peptides pontés. Dans le cadre de l'analyse de complexe protéique avec moins de cinq protéine, il semblerait que MeroX est à privilégier pour le traitement de donnée de XL-MS et est employé au sein du LSMBO.

En perspective, le gel SDS-PAGE, bien que largement utilisé dans tous les protocoles de XL-MS, reste une technique peu précise, car uniquement visuelle (intensité des bandes). Les optimisations de la réaction de pontage par gel sont donc imprécises et qualitatives. Une nouvelle technique disponible au sein du laboratoire, la M-Ph, pourrait constituer une alternative intéressante. Des expériences de M-Ph en conditions dénaturantes permettraient d'observer le pontage covalent de deux protéines par rapport à un échantillon contrôle non-ponté donnant des informations sur leurs masses (de l'ordre du kDa) et proportions (intensité proportionnelle au signal). La M-Ph donnerait des informations plus précises pour effectuer les optimisations exclusivement dans le cas d'analyse d'oligomères ou de complexes.

L'utilisation d'un nouvel agent pontant NHS-ester récemment commercialisé par Bruker, le PhoX¹⁸⁷, pourrait permettre d'enrichir les peptides cross-linkés grâce à un groupement phosphate fixé sur l'agent pontant via la IMAC. L'utilisation de cet agent pontant pourrait considérablement augmenter le nombre de peptides cross-linkés identifiés après enrichissement. Au sein du laboratoire, un protocole d'enrichissement des peptides phosphorylés en phospho-protéomique a déjà été mis en place avec une cartouche IMAC avec le robot de préparation AssayMAP Bravo (Agilent). Ce protocole peut être transposable à l'enrichissement des peptides pontés avec le PhoX et l'automatisée rendant celui-ci reproductible.

Enfin, l'utilisation du TimsTOF Pro (Bruker) possédant une cellule de mobilité ionique TIMS²²⁸ apporterait une séparation supplémentaire des peptides en fonction de leur masse et charge ce qui permettrait de séparer les di-peptides pontés (peu abondant) des peptides non cross-linké (en large abondance). Cet instrument a notamment permis d'identifier plus de 623 peptides pontés de la BSA avec du DSB²²⁸.

Partie 2 : Combinaisons d'approches de MS structurale pour la caractérisation de complexes multi-protéiques

L'objectif de cette partie est de caractériser plusieurs complexes protéiques avec différents partenaires par l'application de plusieurs méthodes de MS structurale pour chaque système. Le premier chapitre s'intéresse à l'analyse, en nMS et HDX-MS, de récepteurs nucléaires et l'impact des co-régulateurs sur leurs conformations. Le second chapitre se concentre sur la caractérisation par HDX-MS de protéines membranaires en présence de petites molécules dans un environnement mimant la membrane. Le troisième chapitre concerne les protéines arginines méthyltransférases et leur caractérisation structurale en présence de leurs substrats par nMS, HDX-MS et XL-MS. Enfin, le dernier chapitre est dédié à l'étude de complexes d'assemblages de hauts poids moléculaires (300-400 kDa) impliquant des protéines AAA+ par XL-MS.

Chapitre 1 : Caractérisation de complexes impliquant des récepteurs nucléaires par nMS et HDX-MS

Chapitre 2 : Caractérisation de protéines membranaires ou dans un environnement lipidique par HDX-MS

Chapitre 3 : Caractérisation de protéines arginines méthyltransférases (PRMTs) par MS structurale

Chapitre 4 : Apport du XL-MS pour la caractérisation de macrocomplexes impliquant les ATPases RUVBL1/2 humaines

Chapitre 1 : Caractérisation de complexes impliquant des récepteurs nucléaires par nMS et HDX-MS

I. Introduction aux récepteurs nucléaires

Les récepteurs nucléaires (RN) sont une famille de facteurs de transcription qui régulent de nombreux processus physiologiques tels que le métabolisme, la reproduction, l'inflammation, ainsi que le rythme circadien²⁴¹. Les RN peuvent former des monomères, des homodimères ou encore des hétérodimères avec RXR (Récepteur X des rétinoïdes)²⁴². Ils possèdent une architecture commune avec plusieurs domaines²⁴³⁻²⁴⁴ (Fig. 1), dont :

- Un domaine N-terminal (A/B), une région désordonnée contenant la fonction d'activation 1 (AF-1) qui interagit avec des co-régulateurs.
- Un domaine de liaison à l'ADN (DBD, C) qui contient deux doigts de zinc coordonnés chacun par 4 cystéines et l'hélice de reconnaissance de l'ADN.
- Une petite région charnière (D) qui est un lien plus ou moins flexible entre le DBD et le domaine de liaison des ligands (LBD).
- Un LBD (E) qui est le domaine de fixation des ligands. Il contient, au sein de l'hélice H12, une fonction d'activation ligand-dépendante (AF-2), permettant la liaison de co-régulateurs. Le LBD porte aussi l'interface principale de dimérisation des RNs.
- Un domaine F est un domaine hautement variable situé à l'extrémité C-terminale des RNs, qui est sujet à de fortes PTMs et permet le recrutement d'ubiquitine ligases et de protéines. Il est présent dans une minorité des RNs.

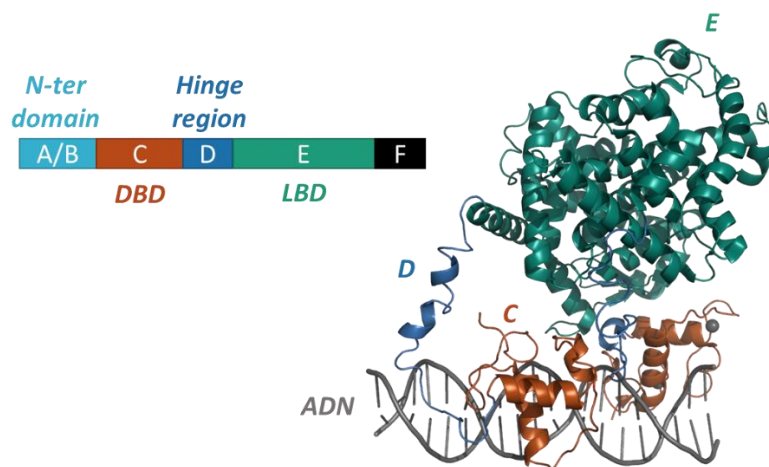


Figure 1 : Représentation schématique des domaines communs aux RNs avec une visualisation de ces domaines sur la structure d'un hétérodimère RXR/LXR (PDB : 4NQA).

La liaison de ligands (hormones, vitamines, métabolites) permet le recrutement de divers co-régulateurs transcriptionnels et conduit à la transactivation ou la transrépression des gènes cibles²⁴⁵⁻²⁴⁷. Pour beaucoup de RNs, la liaison des ligands fonctionnels induit un changement conformationnel du récepteur qui implique un repositionnement de l'hélice H12²⁴⁸. Quand les RNs sont dans une conformation active pour la transcription, une poche hydrophobe et une pince électrostatique chargée au niveau du LBD participent à l'interface d'interaction et connectent les motifs LXXLL des co-activateurs²⁴⁹. Les complexes RNs/co-activateurs vont acétyler des histones entraînant la décompaction de la chromatine et une augmentation de la transcription des gènes cibles. Alors que l'association avec des corépresseurs conduit à la désacétylation des histones et donc à la compaction et la répression de l'expression des gènes²⁵⁰.

Une dysfonction de la signalisation du RN peut être à l'origine de certaines maladies telles que des cancers, l'infertilité, l'obésité ou le diabète²⁵¹. De ce fait, les RNs font l'objet de recherches, de plus en plus nombreuses, comme cibles médicamenteuses pour différentes pathologies²⁴⁵.

II. Impact des mutations activatrices et inactivatrices de PPAR γ sur la relation structure/fonction

1. Contexte analytique

Parmi ces RNs, le récepteur γ activé par les proliférateurs de peroxyosomes (PPAR γ) joue des rôles clés dans le métabolisme humain et l'homéostasie du glucose²⁵²⁻²⁵³. La régulation du PPAR γ est, par exemple, critique dans la tumorigenèse du cancer de la vessie envahissant le muscle (MIBC)²⁵⁴⁻²⁵⁵. Un criblage génétique a permis d'identifier des mutations de perte de fonction dans le LBD de PPAR γ dans les MIBC basaux, associées à une sous-expression/inactivation de PPAR γ soutenant un rôle de suppresseur de tumeur pour PPAR γ . PPAR γ^{LBD} est composé de 13 hélices α (1-12 et 2') et de trois petits brins β ²⁵⁶. La surface de liaison des co-régulateurs AF-2 du LBD (H3-H5) est essentielle pour déterminer la fonction transcriptionnelle de PPAR γ par le recrutement différentiel des co-régulateurs (Fig. 2)²⁵⁷⁻²⁵⁸. Plusieurs ligands naturels et synthétiques (GW1929, rosiglitazone, GW9662 et le T0070907) et co-régulateurs (MED1, PGC1 α et NCoR) sont déjà caractérisés comme entrant en interaction avec PPAR γ ²⁵⁹. La structure de PPAR γ^{LBD} a été aussi largement étudiée en présence de différents ligands et/ou de fragments de co-régulateurs²⁶⁰, révélant les différentes

conformations de ce RN et plus particulièrement le repositionnement de l'hélice H12 du LBD (Fig. 2). Récemment, différents mutants de PPAR γ ^{LBD} ont été découverts par nos collaborateurs dans le cadre de l'étude du MIBC avec 8 mutations ponctuelles favorisant la croissance des tumeurs (mutants activateurs)²⁵⁴ et 12 favorisant la suppression de ces tumeurs (mutants inactivateurs).

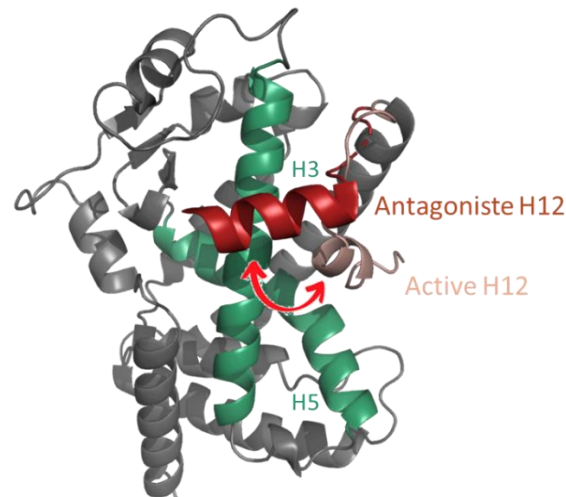


Figure 2 : Combinaison des structures de PPAR γ ^{LBD} (PDB : 1PRG, 6CST) présentant les hélices H3 et H5 (vert) et l'hélice H12 dans une conformation active (rose) ou antagoniste (rouge).

Ce projet a été réalisé dans le cadre d'un projet ANR (Nurescape) en collaboration avec l'Institut de Génétique et de Biologie Moléculaire et Cellulaire avec le Dr. N. Rochel et son équipe (IGBMC, UMR 7104, Illkirch). Cette étude aspire à tendre vers une meilleure caractérisation structurale des RNs pour élucider les conséquences des mutations de PPAR γ ^{LBD}. Pour cela, différentes techniques ont été employées comme des tests fonctionnels, de la résonance plasmonique de surface, de la thermophorèse à micro-échelle, de la dynamique moléculaire, de la cristallographie aux rayons X et des approches de MS structurale, ces dernières faisant l'objet de mon travail.

2. Objectifs en MS structurale

L'objectif de cette étude est d'utiliser différentes approches de MS structurale pour étudier l'impact des mutations de PPAR γ sur sa structure et son activité. Des expériences de nMS en présence de co-régulateurs et de HDX-MS ont été effectuées sur 4 mutants différents de PPAR γ ^{LBD} : les mutants PPAR γ les plus inactifs (F310S et H494Y), un inactivateur partiel (Y505H), le mutant activateur T475M et PPAR γ non muté (WT). Ce projet a été mené dans le but de (i) définir l'impact de ces mutations sur la liaison des ligands et co-régulateurs et (ii)

d'observer les changements conformationnels induits par ces mutations sur la structure de PPAR γ ^{LBD}.

3. Matériel et méthodes

La nMS a été effectuée par infusion directe sur le synapt-G2 (Waters), les voltages de cônes et pression à l'interface sont donnés dans les légendes des spectres nMS. Les données ont été exploitées via le logiciel MassLynx 4.1. (Waters). En amont de l'analyse, les échantillons ont été dessalés sur VivaspineTM (10 kDa, Fisher Scientific) avec 200mM d'acétate d'ammonium à un pH de 7,0 pour toutes les analyses, 7 cycles de 6 min à 15 000xg et à 4°C ont été réalisés. Les ligands et/ou co-régulateurs ont été ajoutés en léger excès molaire (1:3, protéine:co-régulateur ou 1:3:3, protéine:ligand:co-régulateur). Ainsi, le co-activateur PGC1 α (CoA) a été analysé en présence de PPAR γ ^{LBD} puis avec le ligand agoniste GW1929. De même que le co-répresseur (CoR) NCoR a été étudié en présence de PPAR γ ^{LBD} puis avec deux ligands SR16832 (antagoniste) et T0070907 (inverse agoniste).

Les expériences d'échange H/D ont été effectuées en triplicat en présence d'un tampon à base de D₂O (20mM Tris-HCl, 200mM NaCl à un pH de 8) à différents temps de deutération (0,5, 2, 10, 30 et 60 minutes). L'arrêt du marquage a été réalisé en présence de 2M de Guanidine-HCl et 100mM de Glycine à un pH final de 2,4. Les protéines (15pmoles) avec/sans NCoR (30pmoles) ont été digérées avec de la pepsine avec un débit de 200 μ L/min. Après digestion, les peptides ont été séparés sur une colonne C18 et analysés sur un Synapt-G2 Si avec la mobilité ionique activée (Waters). Les peptides ont été identifiés avec le logiciel PLGS 2.5.3 (Waters) et filtrés avec DynamX 3.0 (Waters). Les profils isotopiques sont vérifiés manuellement sans correction de l'échange inverse (D/H). Les résultats ont été validés statistiquement avec une p-value de 0,01 avec le logiciel MEMHDX¹³⁴.

4. Résultats

Le LBD de PPAR γ a été étudié sous sa forme « sauvage » (non mutée, WT) ainsi que mutée (4 mutants) : 1 mutation « gain de fonction/activatrice » (T475M) et 3 mutations « perte de fonction/inactivatrice » (F310S, Y505H et H494Y).

4.1. Impact de la mutation sur la liaison aux co-régulateurs

4.1.1. nMS

Les expériences de nMS des protéines seules (Fig. 3) ont montré des préparations protéiques pures et homogènes avec des masses en accord avec la séquence en acide aminé (Annexe 2). Masse du WT (33575 ± 1 Da) : +26 Da pour le H494Y (33601 ± 1 Da) ; -26 Da pour le Y505H (33548 ± 1 Da) ; -60 Da pour le F310S (33515 ± 1 Da) ; +30 Da pour le T475M (33605 ± 1 Da).

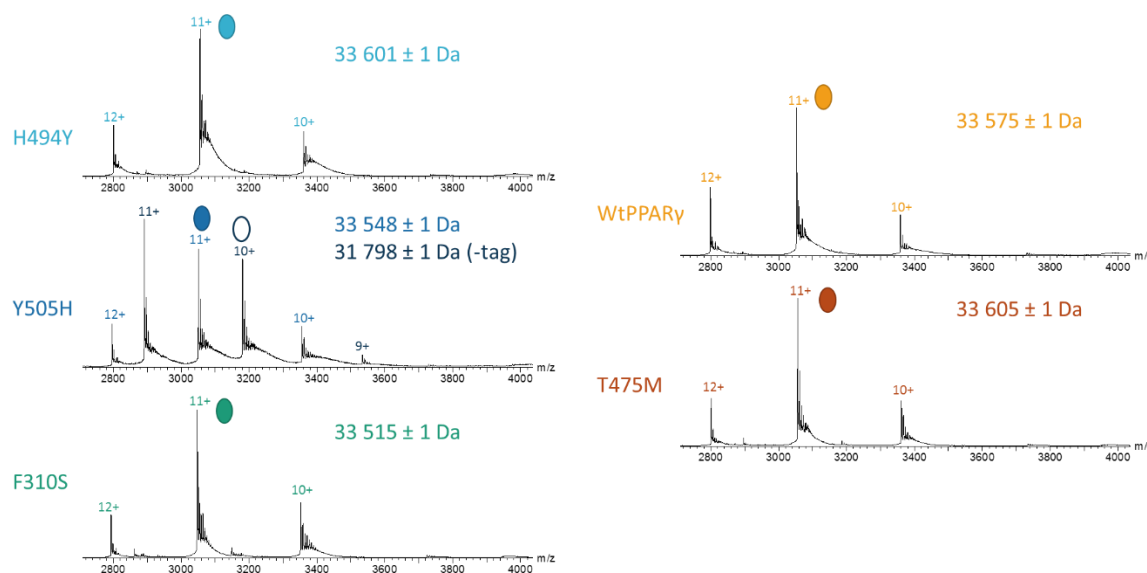


Figure 3 : Spectres de masse obtenus par nMS de H494Y, Y505H, F310S, WT et T475M avec un voltage de cône fixé à 20 V et une pression à l'interface de 6 mbar.

D'autres analyses en présence de différents co-régulateurs et ligands ont été effectuées dans le but de déterminer des taux de complexation des co-régulateurs (Fig. 4, Annexe 16). Les expériences ont été réalisées en triplicat et les taux ont été mesurés à partir des intensités de chaque espèce détectée (protéine seule et en complexe) pour cinq états de charges (+9, +10, +11, +12 et +13). Une stœchiométrie de 1:1 a été observée pour tous les complexes protéines/co-régulateurs et 1:1:1 pour les complexes protéines:ligands:co-régulateurs impliquant les ligands GW1929 et SR16832. La liaison du ligand T0070907 (277 Da) de stœchiométrie 1:1:1 (protéines:ligands:co-régulateurs) a été observée avec une masse de 242 Da, ce qui correspond à la perte d'un chlore (-35 Da) traduisant la formation d'une liaison covalente du ligand avec le RN (orthostérique). Une stœchiométrie de 1:2:1 (protéines:ligands:co-régulateurs) a également été observée avec d'une part, cette même perte d'un chlore (+242 Da) et d'autre part, la liaison non-covalente d'un deuxième ligand (+277 Da) montrant la fixation allostérique de ce second T0070907.

Ainsi, les analyses de nMS des protéines en présence du co-activateur PGC1 α révèlent une liaison de celui-ci d'environ 30% pour le T475M, le WT et le Y505H alors qu'un plus faible taux (<20%) est observé pour H494Y et F310S (Fig. 4A), suggérant une diminution de l'interaction des mutants inactivateurs avec le CoA. L'ajout du ligand agoniste GW1929 augmente considérablement la liaison du CoA à \approx 70% pour T475M, WT et Y505H et 54% pour les deux autres mutants (Fig. 4A). La liaison de PGC1 α est donc plus faible pour le H494Y et le F310S même en présence du GW1929 comparé à PPAR γ WT, confirmant la diminution de l'interaction protéine/CoA en présence des mutants inactivateurs. Le taux de complexation du mutant Y505H sont plus proches du WT que d'autres mutants inactivateurs, en accord avec son pouvoir d'inactivateur partiel.

Des expériences similaires ont été effectuées en présence d'un peptide co-répresseur NCoR montrant une liaison plus importante de celui-ci avec les mutants H494Y (45%) et F310S (63%) tandis que le WT, le T475M et le Y505H sont plus proches des 40% (Fig. 4B). Ces résultats montrent l'augmentation de la liaison du CoR avec les mutants inactivateurs. L'addition d'un ligand antagoniste SR16832 n'augmente pas de manière significative de la liaison du NCoR quelle que soit la protéine (Fig. 4B), suggérant que la liaison du ligand ne dispose pas d'influence sur la fixation du CoR. En présence de l'inverse agoniste T0070907, l'affinité pour T475M, WT et H494Y du NCoR est augmentée (Fig. 4B).

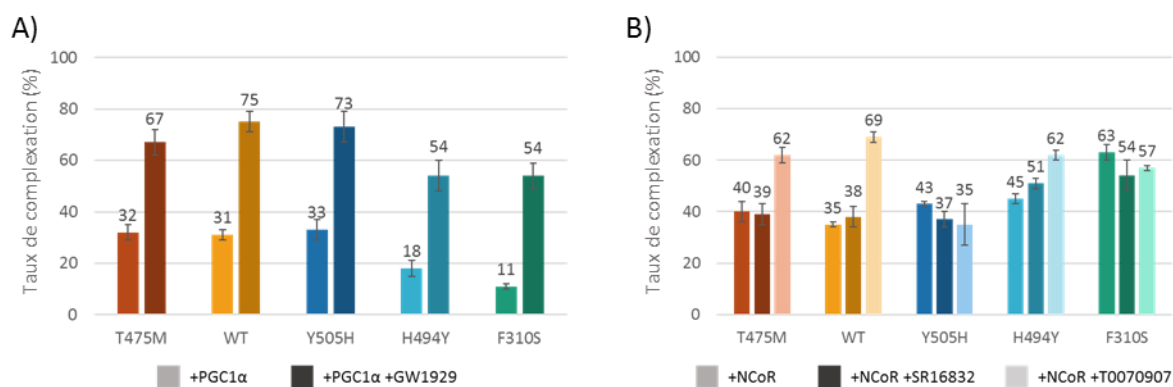


Figure 4 : Histogrammes des taux de complexations observés entre les protéines et (A) PGC1 α avec/sans le GW1929 et (B) NCoR avec/sans le SR16832 ou le T0070907.

4.1.2. HDX-MS

Pour mieux comprendre la dynamique de l'interaction du CoR avec PPAR γ WT et le mutant F310S. Des expériences de HDX-MS ont été réalisées sur ces deux protéines en absence et en présence du co-répresseur NCoR.

4.1.2.1. Optimisation de la phase pré-analytique

Dans les conditions optimales et avant les échanges H/D, la couverture de séquence (Annexe 17) obtenue pour chacune des protéines (Tableau 1) est très satisfaisante (plus de 98%) de même que la redondance (> 3,6). Les échanges H/D de PPAR γ ont été effectués à des temps courts d'incubation (0,5, 1, 2, 5 et 10 min) du fait d'une instabilité du récepteur pour des temps longs de deutération (>10 min) induisant un manque de répétabilité en HDX-MS. Après vérification des profils isotopiques, des informations ont été obtenues sur plus de 86% de la séquence avec une redondance à 2,3 pour les analyses du F310S et du WT (Tableau 1).

| Analyses | PPAR γ | F310S après HDX-MS | WT après HDX-MS |
|-------------------------------|---------------|--------------------|-----------------|
| Couverture de séquence (%) | 98,2 | 86,9 | 90,2 |
| Redondance | 3,6 | 2,3 | 2,3 |
| Nombre de peptides identifiés | 105 | 57 | 61 |

Tableau 1 : Couverture de séquence, redondance et nombre de peptides identifiés relatés pour les analyses de PPAR γ .

4.1.2.2. Suivi par HDX-MS de l'impact de la fixation du NCoR

Les expériences de HDX-MS ont révélé plusieurs différences d'incorporation en D statistiquement significatives (MEMHDX avec une p-value à 0,01¹³⁴) entre les protéines seules ou en complexes avec NCoR (Fig. 5).

Une plus grande incorporation en D est mesurée pour toutes les hélices de PPAR γ WT en présence du NCoR indiquant une plus grande accessibilité au solvant et donc une déprotection de ces hélices suite à la liaison (Annexe 18). La partie C-terminale de PPAR γ WT, comprenant son hélice H12, est déprotégée en accord avec le repositionnement de H12 dû à la liaison de NCoR. La seule protection observée est localisée au niveau du peptide 320-326 de l'hélice H3 qui est aussi le site de liaison des co-régulateurs.

Le mutant inactivateur F310S est globalement moins impacté par la liaison du CoR révélant un changement conformationnel moins important pour le mutant que pour la protéine WT (Annexe 19). La même protection du peptide 320-326 a été remarquée pour le F310S démontrant l'interaction du NCoR. Cependant, l'hélice H12 de F310S ne présente pas de différence d'incorporation en D suite à la fixation du CoR. L'ensemble de ces observations en HDX-MS laisse supposer que F310S serait déjà dans une conformation antagoniste (favorable à l'interaction de NCoR) contrairement à la protéine non mutée.

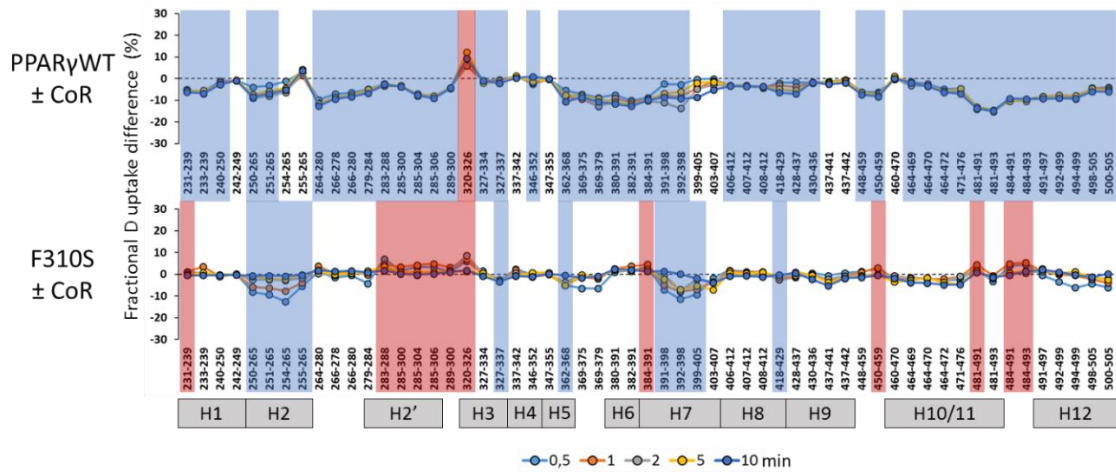


Figure 5 : Représentation des différences d'incorporations relatives en D entre PPAR γ WT ou F310S et le complexe avec NCoR pour chaque peptide identifié de PPAR γ aux différents temps de deutération. Les peptides encadrés présentent des différences statistiquement significatives d'incorporation en D (MEMHDX, p-value à 0,01¹³⁴) supérieures à 2% (rouge) montrant une protection des protéines en présence de NCoR et inférieures à -2% (bleu) révélant une déprotection.

4.2. Impact de la mutation ponctuelle de PPAR γ

Afin d'avoir une visualisation directe de l'impact des mutations sur la conformation de PPAR γ , des expériences de HDX-MS ont été menées sur les protéines seules et l'incorporation en D de chaque mutant a été comparée à celle du WT. L'échange H/D a révélé des différences d'incorporation en D statistiquement significatives (MEMHDX avec un p-value à 0,01¹³⁴) spécifiques à chaque mutant (Fig. 6 et 7, Annexe 20).

Dans un premier temps, le mutant activateur T475M est globalement protégé comparé à PPAR γ WT alors que les autres mutants inactivateurs sont principalement déprotégés montrant une première tendance liant l'activité transcriptionnelle (in-/activatrice) et les dynamiques d'incorporation en D (dé-/protection).

Dans un second temps, une augmentation du nombre de régions présentant des différences d'incorporation en D significatives est observable suivant la mutation comparée à PPAR γ WT (6 peptides pour le Y505H ; 12 pour le H494Y ; 44 pour le F310S). En effet, le mutant Y505H est déprotégé au niveau des régions comprenant les résidus 380-405 et 477-481 (soit 5,9% des acides aminés) identique aux deux autres mutants alors que H494Y présente des déprotections supplémentaires des résidus 264-280 et 359-379 (soit 11%). F310S est quant à lui encore plus impacté avec 87,4% des résidus montrant des déprotections par rapport à PPAR γ WT. Ces résultats de HDX-MS peuvent être mis en lien avec la relation structure/fonction des mutants indiquant que plus le nombre de régions présentant des

différences d'incorporations en D est grand plus le pouvoir inactivateur est important. Ces résultats sont en accords avec les tests fonctionnels réalisés par nos collaborateurs.

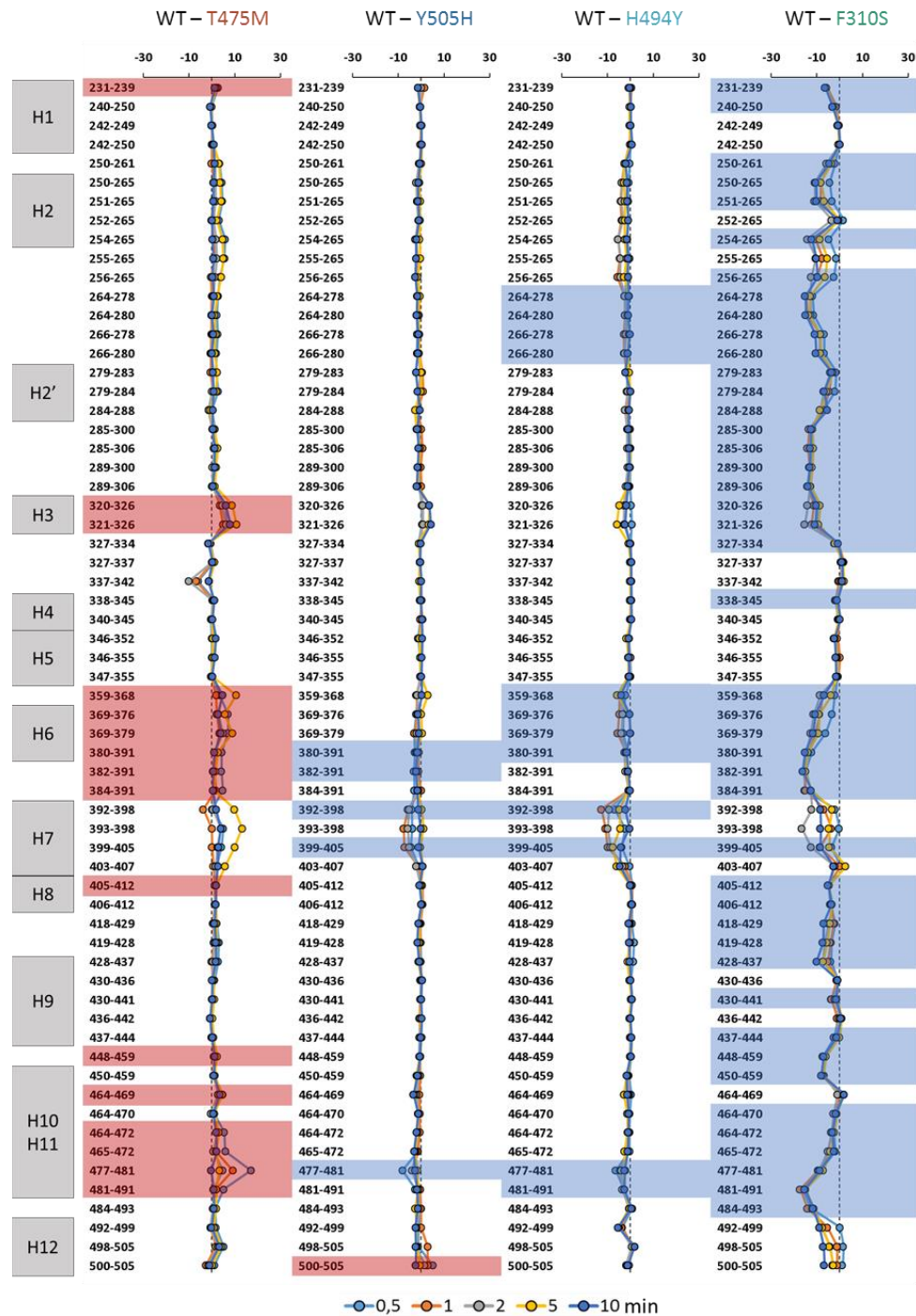


Figure 6 : Représentation des différences d'incorporations relatives en D entre PPAR γ WT et les mutants (T475M, Y505H, H494Y et F310S) pour chaque peptide identifié de PPAR γ aux différents temps de deutération. Les peptides encadrés présentent des différences statistiquement significatives d'incorporation en D (MEMHDX, p-value à 0,01¹³⁴) supérieures à 2% (rouge) montrant une protection du mutant par rapport au WT et inférieures à -2% (bleu) relatant d'une déprotection.

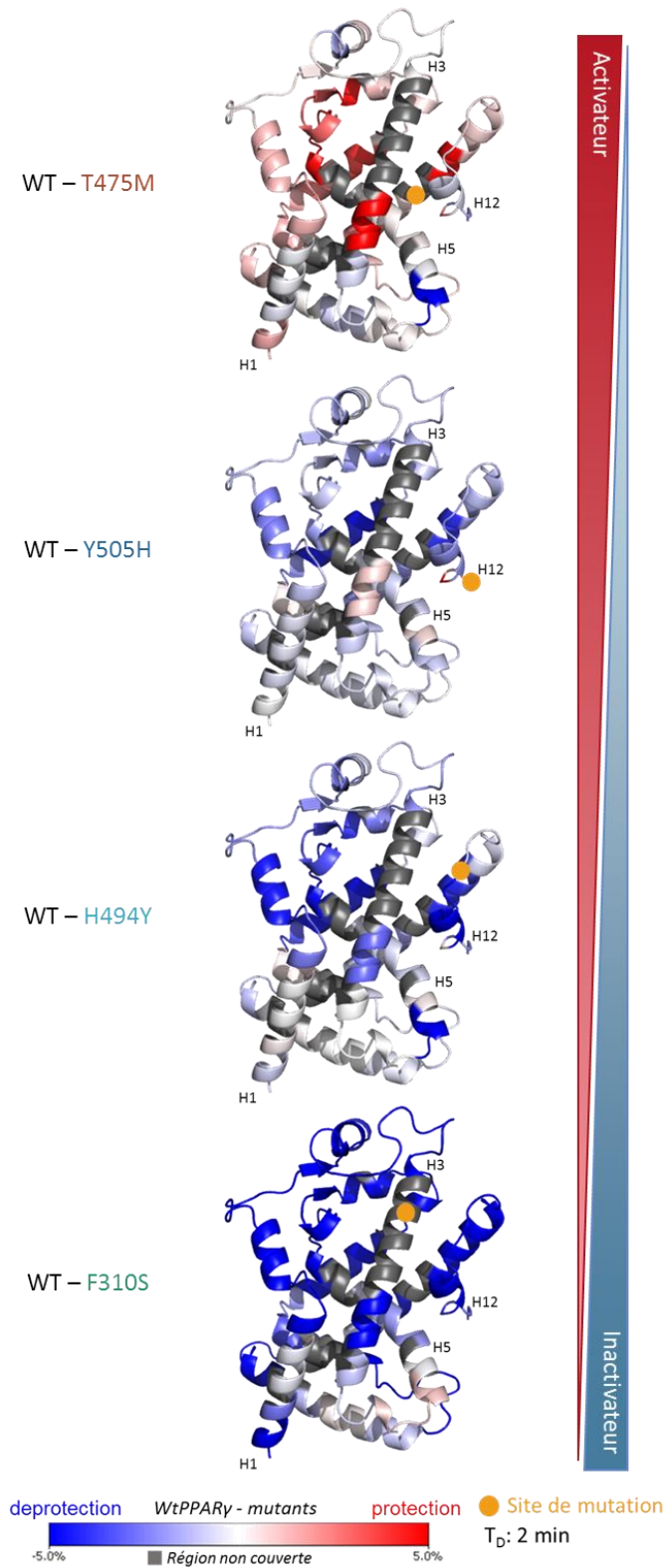


Figure 7 : Représentation des différences d'incorporation en D (temps de deutération : 2min) entre PPAR γ WT et les mutants (T475M, Y505H, H494Y et F310S) sur la structure de PPAR γ WT (PDB : 1PRG). Les régions présentant des différences d'incorporation en D sont colorées suivant l'échelle donnée (-5% ; 5%), celles pour lesquelles aucune information n'est disponible sont en grise.

5. Valorisation

J'ai pu valoriser ces travaux par une communication par affiche : ***Deciphering structure-function relationship of loss-of-function PPAR γ mutants involved in bladder cancer using native MS and H/D exchange mass spectrometry.*** Marie Ley, Judit Osz, Syrine Beji, Natacha Rochel-Guiberteau, Sarah Cianférani. Congrès Analytics 2022, 05-08/09/22, Nantes.

Deux articles sont en préparation :

- ***Basal subtype-specific tumor suppressor role for PPAR γ in bladder cancer.*** Coutos-Thévenot, L.; et al. DOI : 10.1101/868190
- ***Structure-function relationships of PPAR γ mutants identified in bladder cancer.*** Ley, M.; et al.

6. Conclusions

Les résultats obtenus grâce aux techniques de MS structurale nous indiquent que :

- Y505H est un mutant inactivateur partiel montré en nMS par les taux de complexation du co-activateur en présence/absence du ligand agoniste proche de ceux de PPAR γ WT. Les résultats non concluant obtenus en nMS pour Y505H et H494Y en présence de CoR et du T0070907 peuvent être dus à la mutation au niveau de l'hélice H12, car ce ligand induirait le repositionnement de l'hélice H12.
- La fixation des ligands antagonistes SR12813 et T0070907 n'influe pas sur la liaison du NCoR de manière significative pour les mutants inactivateurs (Y505H, H494Y et F310S).
- L'hélice H12 du F310S ne présente pas en HDX-MS de réarrangement structural suite à la liaison du NCoR contrairement à PPAR γ WT montrant que cette hélice serait dans une conformation antagoniste pour les mutants inactivateurs. Ces résultats confirment que le mutant F310S est bien inactivateur.
- L'activateur T475M est protégé alors que les inactivateurs sont déprotégés comparé au WT montrant que l'activité transcriptionnelle est lié à l'incorporation en D (relation structure/fonction).
- Le nombre de régions subissant un réarrangement structural est en lien avec le pouvoir d'activation transcriptionnel des mutants permettant de les classer par HDX-MS. Le F310S serait plus inactivateur que le H494Y et le Y505H ce qui est en accord avec les taux de complexations observés en nMS.

III. Récepteur nucléaire ERR (*Oestrogen Related Receptor*)

1. Contexte analytique

Le récepteur nucléaire orphelin (ERR) appartient à la sous-famille des RNs stéroïdiens et a été identifié en 1998 du fait de l'homologie de son DBD (68% du domaine C et 36% du domaine E) avec celui du récepteur des œstrogènes (ER)²⁶¹⁻²⁶³. Il existe trois isoformes de ERR (α , β , γ) possédant tous des caractéristiques spécifiques des RNs en plus de ceux des ER (Fig. 8), qui sont tous les trois orphelins (pas de ligand naturel identifié à ce jour). Ils sont exprimés dans plusieurs tissus durant l'embryogénèse et chez l'adulte. Ils ont une activité transcriptionnelle ligand-indépendante, qui est nécessairement médiée par le recrutement de co-régulateurs. ERR α est impliqué dans de nombreuses fonctions reliées au métabolisme énergétique dans les tissus métaboliquement très actifs, tels que le foie, les muscles, les tissus adipeux ou encore le cœur²⁶⁴⁻²⁶⁵.

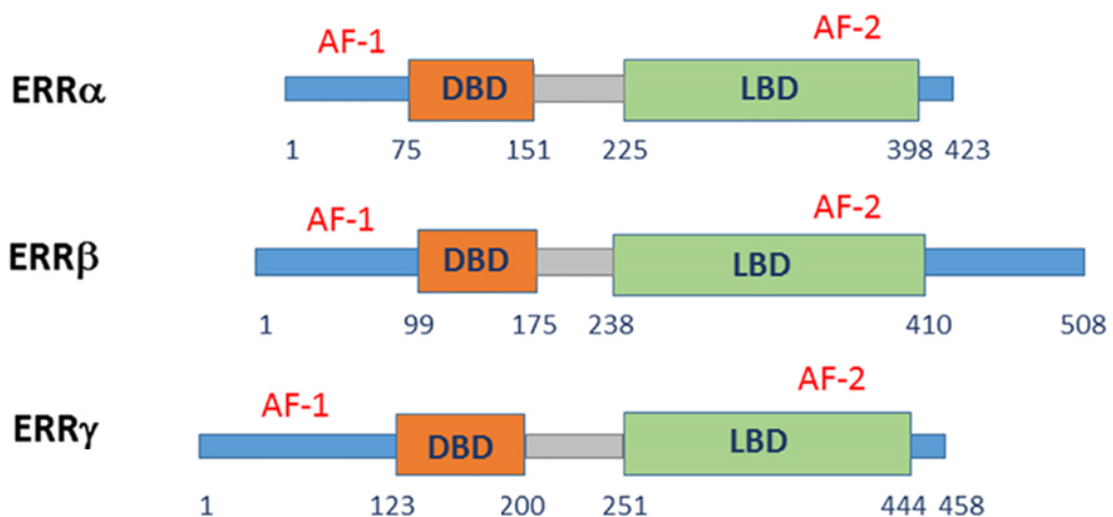


Figure 8 : Représentation schématique des trois isoformes de ERR avec les extrémités N- et C-terminales (bleu) et la région charnière (gris) entre le DBD (orange) et le LBD (vert) (d'après²⁶⁶).

ERR α est le seul de ces isoformes qui est présent dans tous les tissus décrits précédemment. Il est composé comme tous les autres RNs, d'une fonction d'activation 1 et 2, d'un DBD et d'un LBD (avec 12 hélices)²⁶⁶. ERR α a largement été étudié en interaction avec le co-activateur PGC1 α ²⁶⁷ (Coactivateur 1-alpha du récepteur gamma activé par les proliférateurs de peroxysomes) à l'origine découvert en interaction avec PPAR γ . L'interaction de ERR α (homodimère) et du CoA PGC1 α est nécessaire pour l'activité transcriptionnelle du RN²⁶⁸. Ce co-facteur PGC1 α contient le motif LXXLL connu comme étant le site de reconnaissance des RNs²⁶⁴ mais possède aussi de nombreuses régions intrinsèquement désordonnées prédites

par Interpro²⁶⁹ sur UniProt (Q9UBK2). Bien que le complexe ERR α /PGC1 α a été étudié révélant des zones d'interactions probables, la structure de ce complexe n'a pas encore été obtenue. Un autre partenaire de ERR α a été découvert plus récemment, la protéine diméthylase 1 spécifique des lysines (LSD1)²⁷⁰. LSD1 est la première enzyme identifiée comme responsable du retrait des groupes mono-/diméthylés des lysines 4 (K4) et 9 (K9) de l'histone H3 menant à une répression (H3K4) ou une activation (H3K9) de la transcription²⁷⁰⁻²⁷¹. Deux études ont montré par une approche de séquençage d'ARN que ERR α régule l'activité de LSD1^{270, 272}. En effet, une inactivation de ce complexe entraîne une augmentation de la diméthylation de H3K9me2. ERR α et LSD1 constituent des facteurs prédictifs de différents cancers²⁷³⁻²⁷⁵. Aucune information sur les zones d'interactions ces deux protéines ni aucune structure du complexe ERR α /LSD1, même partielle n'ont encore été obtenues. La structure de LSD1 est disponible²⁷⁶, révélant deux superhélices²⁷⁷, un domaine SWIRM pouvant servir comme site d'ancrage pour les histones²⁷⁸ ainsi que le domaine d'activité de diméthylase.

Ce projet a été réalisé dans le cadre d'une collaboration avec l'Institut de Génétique et de Biologie Moléculaire et Cellulaire avec le Dr. I. Billas et son équipe (IGBMC, UMR 7104, Illkirch). Dans le but de tendre vers une meilleure compréhension structurale des différents complexes impliquant ERR α , différentes techniques ont été employées comme la migration sur gel natif, la cryo-EM et des approches de MS structurale, ces dernières faisant l'objet de mon travail.

2. Objectifs

Ce chapitre décrit les résultats obtenus grâce aux expériences de HDX-MS et XL-MS sur les complexes ERR α /PGC1 α et ERR α /LSD1 dans le but d'obtenir des informations structurales inédites sur ces complexes.

3. Matériel et méthodes

Les expériences d'échange H/D ont été effectuées en triplicats en présence d'un tampon à base de D₂O à un pH de 6,5 (10mM HEPES, 120mM KCl, 1mM TCEP et 0,2mM MgCl₂) à différents temps de deutération (0,5, 2, 10, 30 et 60 minutes) pour ERR α /PGC1 α . L'arrêt du marquage a été réalisé en présence de 2M de Guanidine-HCl, 100mM de Glycine, 1mM TCEP et 0,2mM MgCl₂ à un pH final de 2,4. Les protéines (20pmoles de ERR α et 40pmoles de PGC1 α)

ont été digérées avec de la pepsine immobilisée sur une colonne (Waters) avec un débit de 200 μ L/min.

Dans le cas de l'analyse du complexe ERR α /LSD1 un tampon de marquage à un pH de 8,5 (15mM Tris-HCl, 120mM KCl, 1mM TCEP et 0,2mM MgCl₂) a été utilisé, alors que le tampon d'arrêt du marquage est composé de 2M de Guanidine-HCl et 100mM de Glycine à un pH final de 2,4. ERR α (10pmoles) et LSD1 (20pmoles) ont été digérées avec de la rizhopuspepsine immobilisée sur une colonne (Affipro) avec un débit de 100 μ L/min.

Après digestion, les peptides ont été séparés sur une colonne C18 et analysés sur un Synapt-G2 Si (Waters). Les peptides ont été identifiés avec le logiciel PLGS 2.5.3 (Waters) et filtrés avec DynamX 3.0 (Waters). Les profils isotopiques sont vérifiés manuellement sans correction de l'échange inverse. Les résultats ont été validés statistiquement avec une p-value de 0,01 avec le logiciel MEMHDX (institut Pasteur).

Pour les expériences de XL-MS, les tampons de stockage des protéines ont été échangés par un tampon composé de 10mM de K₂HPO₄ et 10mM KH₂PO₄, 100mM de NaCl, 0,2 MgCl₂ et 1mM de TCEP à un pH de 7,2 (ERR α /PGC1 α) ou 8,0 (ERR α /LSD1) sur Zeba (7 kDa, Thermo Fisher Scientific). Le XL-MS a été réalisé avec du DSBU à un excès molaire de 200 à 20°C pendant 45 min pour le complexe ERR α /LSD1 et aussi du DSAU pour le complexe ERR α /PGC1 α . Dans le cas du XL-MS du complexe ERR α /PGC1 α , les protéines ont été pontées en présence d'ADN. Dans les mêmes conditions, le pontage a été effectué avec du DSAU sur le complexe ERR α /PGC1 α . Tous les pontages ont été réalisés en triplicat puis réduits, alkylés et digérés par un mix de Trypsine/Lys-C. Les excès de réactifs ont été éliminés par passage sur SPE C18, séparés sur une colonne C18 et analysés sur un Orbitrap Q-Exactive HF-X (Thermo Fisher Scientific). MeroX 2.0.1.4 a été utilisé pour l'identification des peptides pontés avec un FDR fixé à 5%. La présence des fragments caractéristiques attendus pour ce type d'agent pontant a été vérifiée et seuls les cross-links validés dans au moins 2 des 3 réplicats ont été conservés. L'intervalle de distance est fixé de 6 à 30Å pour le DSBU et 6 à 26Å pour le DSAU.

4. Résultats

4.1. Caractérisation du complexe ERR α /LSD1

4.1.1. Optimisation des phases pré-analytiques

Pour mieux comprendre l'interaction entre les protéines ERR α et LSD1 (Annexe 2), des expériences de HDX-MS ont été réalisées sur les protéines seules et en complexe. Mais, avant d'effectuer l'échange H/D, la digestion des protéines est optimisée. Ainsi la rhizopuspepsine a été choisie pour effectuer la digestion permettant d'obtenir un maximum de couverture de séquence (>73%, Tableau 2, Annexe 21).

| Analyses | ERR α | LSD1 | Complexe | | Après HDX-MS | |
|-------------------------------|--------------|------|--------------|------|--------------|------|
| | | | ERR α | LSD1 | ERR α | LSD1 |
| Couverture de séquence (%) | 82,8 | 95,2 | 73,4 | 85,8 | 68,5 | 82,7 |
| Redondance | 4,3 | 5,7 | 2,6 | 4,2 | 2,5 | 3,4 |
| Nombre de peptides identifiés | 142 | 421 | 80 | 297 | 66 | 220 |

Tableau 2 : Couverture de séquence, redondance et nombre de peptides identifiés relatés pour les analyses de ERR α , LSD1 et du complexe ERR α /LSD1.

Après les expériences d'échange H/D et la vérification manuelle des profils isotopiques, seuls 66 peptides ont été conservés pour ERR α . Néanmoins, les peptides sont assez bien répartis tout le long de la séquence laissant au moins quelques peptides représentant tous les domaines/hélices (Fig. 9).

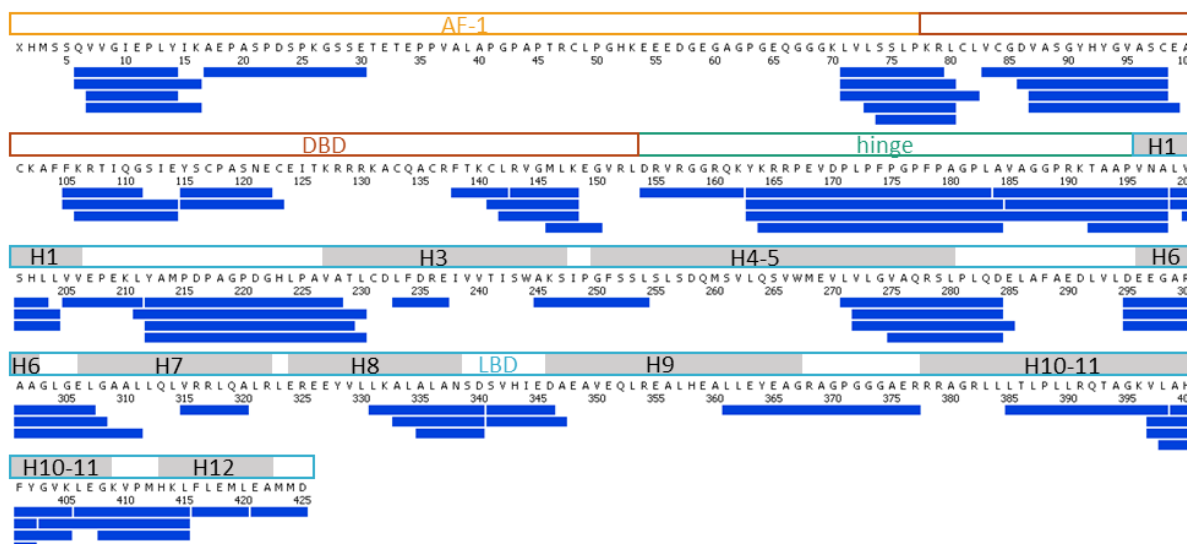


Figure 9 : Carte de recouvrement de séquence des peptides identifiés de ERR α en HDX-MS pour la comparaison entre ERR α seul et en complexe avec LSD1. Les différents domaines de la protéine sont annotés comme la fonction AF-1, le DBD, la région charnière entre le DBD et le LBD (hinge) et le LBD.

Du point de vue de LSD1, encore plus de régions sont significativement impactées par la fixation de $ERR\alpha$, avec (Fig. 11, Annexe 23) :

- ~77% de LSD1 est impacté par la présence de $ERR\alpha$ révélant une déprotection de tous les domaines de LSD1.
- Seulement un peptide comprenant les acides aminés 303-307 du domaine d'activation de la diméthylase est protégé suite à l'interaction avec $ERR\alpha$.

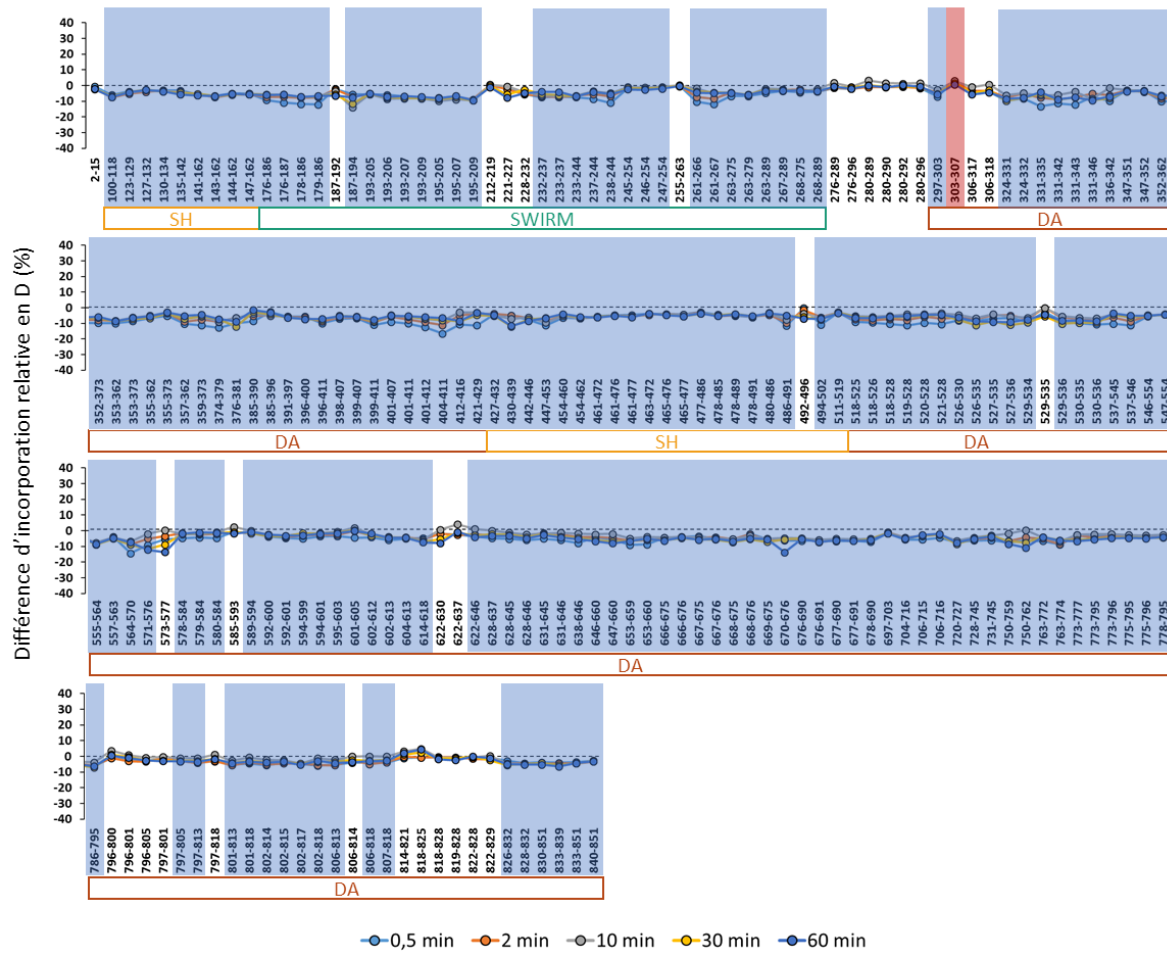


Figure 11 : Représentation des différences d'incorporations relatives en D entre LSD1 et le complexe avec $ERR\alpha$ pour chaque peptide identifié de LSD1 aux différents temps de deutération. Les peptides encadrés présentent des différences statistiquement significatives d'incorporation de deutérium (MEMHDX, p-value à 0,01¹³⁴) supérieures à 2% (rouge) relatant d'une protection du complexe comparé à la protéine seule et inférieures à -2% (bleu) montrant une déprotection. Les différents domaines de la protéine sont annotés comme les superhélices (SH), le SWIRM et le domaine d'activité (DA).

Dans l'ensemble, les résultats d'HDX-MS montrent que (i) le complexe $ERR\alpha$ /LSD1 est plus accessible au solvant que les protéines seules et (ii) que la présence des deux protéines $ERR\alpha$ et LSD1 impacte énormément la conformation globale de chacune d'entre elles. Les données HDX-MS sont en accord avec une grande flexibilité des protéines en lien avec les nombreuses régions désordonnées de LSD1 (1-172) et les boucles flexibles connectant les hélices

correspondant aux peptides 130-134, 347-352, 461-476 et 828-832 qui révèlent tous une saturation (pas d'augmentation de l'incorporation en D au cours du temps) de l'incorporation en D même aux temps courts de deutération (Fig. 12).

Cependant, l'interaction des protéines n'est pas réellement distinguable par le manque de régions protégées. Ces résultats montrent les limitations des expériences de HDX-MS malgré les optimisations et l'obtention d'information de dynamique et changement conformationnel sur 68,5% (ERR α) et 82,7% (LSD1) des séquences.

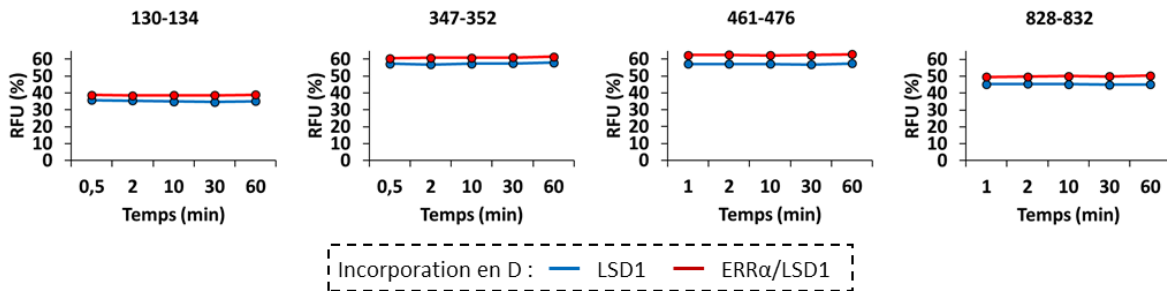


Figure 12 : Représentation de l'incorporation en D relative (RFU) de LSD1 (bleu) et du complexe ERR α /LSD1 (rouge) en fonction du temps de deutération pour différents peptides de LSD1 identifiés en HDX-MS.

4.1.3. XL-MS

Des expériences de XL-MS, avec du DBSU, ont alors été menées dans le but de confirmer l'interaction de ERR α et LSD1 et de définir des régions proches spatialement. Malheureusement, aucun cross-link inter-ERR α /LSD1 n'a été détecté alors que 3 peptides pontés intra-ERR α et 21 intra-LSD1 ont été identifiés. Les trois cross-links intra-ERR α sont localisés sur le DBD et la partie C-terminale de la fonction AF-1 (Fig. 13, Annexe 24). Les peptides pontés intra-LSD1 sont répartis au niveau des différents domaines de cette protéine.

De manière générale, ces résultats ne nous permettent pas de conclure quant à l'interaction de ERR α et LSD1. Cependant, les données obtenues en HDX-MS montrent tout de même un impact de la présence des deux protéines sur la conformation globale de ERR α et de LSD1.

Les expériences de XL-MS tout comme celles de HDX-MS n'ont pas permis de confirmation la liaison de LSD1 à ERR α , bien qu'il y ait suffisamment de résidus cross-linkables et que, comme nous le verrons dans le paragraphe 4.2.3. de ce chapitre, les conditions de XL-MS sont adaptées à l'analyse de ce type de protéines. Enfin, ce projet est alors mis en pause, de la M-Ph est envisagée pour tenter de détecter l'interaction directe entre ERR α et LSD1.

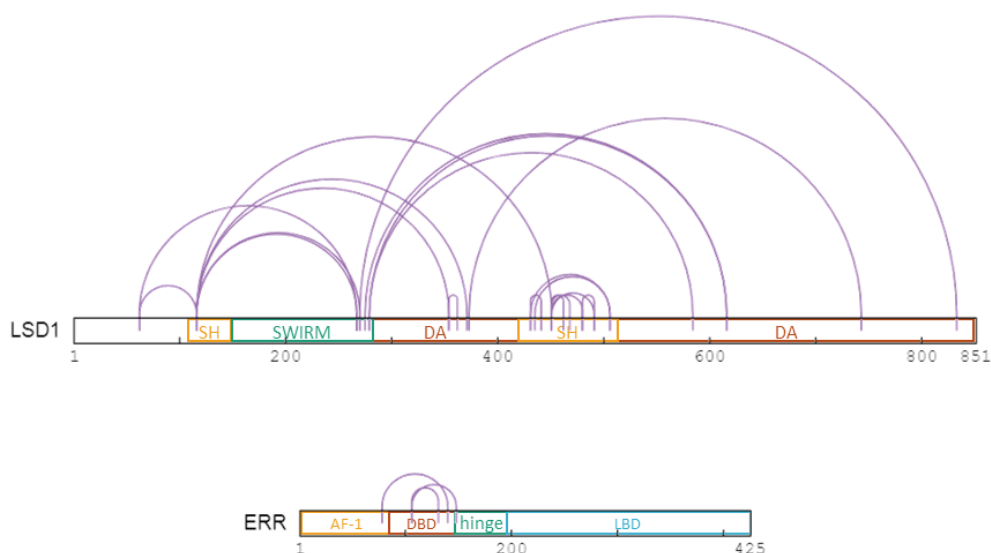


Figure 13 : Représentation schématique des peptides pontés intra-protéine (violet) du complexe $ERR\alpha/LSD1$ (obtenus avec Xinet²¹⁶) avec les superhélices (SH), le SWIRM et le domaine d'activité (DA) de LSD1 ainsi que la fonction AF-1, le DBD, la région charnière entre le DBD et le LBD (hinge) et enfin le LBD de $ERR\alpha$.

4.2. Caractérisation du complexe $ERR\alpha/PGC1\alpha$

De la même manière que pour $ERR\alpha/LSD1$, des expériences de HDX-MS et XL-MS ont été effectuées sur $ERR\alpha/PGC1\alpha$ (Annexe 2). Ces analyses ont été réalisées dans le but d'obtenir de nouvelles informations sur l'interaction entre ces deux protéines et de montrer l'impact de la liaison de $PGC1\alpha$ sur $ERR\alpha$ pleine taille, d'autant que de nombreuses données biophysiques et structurales ont été obtenues par nos collaborateurs (Dr. I. Billas et son équipe, IGBMC).

4.2.1. Optimisation des expériences de HDX-MS

Une première étude en HDX-MS a été menée par Elise DEL NERO au sein du laboratoire, révélant qu'une grande partie de $PGC1\alpha$ est désordonnée comme attendu (confirmé sur UniProt et par InterPro²⁶⁹). Ces régions présentant une forte dynamique d'échange H/D, sont difficiles à caractériser, car n'adoptent pas une conformation stable et unique. La saturation rapide de l'incorporation en D de ces régions, de par leur dynamique, complique l'interprétation des données et la comparaison de deux états. Abaisser le pH d'une unité permet de diminuer la vitesse d'échange par 10^{280} et ainsi d'éviter une saturation complète des domaines désordonnés (DR). Cette méthode ayant déjà été testée sur $PGC1\alpha$ sans montrer de changement conformationnel²⁸¹, nous avons voulu le tester pour l'étude du complexe $ERR\alpha/PGC1\alpha$.

Après optimisation des conditions expérimentales, le recouvrement de séquence, la redondance et le nombre de peptides identifiés ont été obtenus pour chacune des protéines et pour le complexe ERR α /PGC1 α (Tableau 3, Annexe 25).

| Analyses | ERR α | PGC1 α | Complexe | | Après HDX-MS | |
|-------------------------------|--------------|---------------|--------------|---------------|--------------|---------------|
| | | | ERR α | PGC1 α | ERR α | PGC1 α |
| Couverture de séquence (%) | 92,9 | 87,7 | 83,1 | 87,7 | 60,7 | 77,2 |
| Redondance | 4,4 | 4,2 | 2,1 | 6,2 | 1,7 | 4,3 |
| Nombre de peptides identifiés | 136 | 88 | 56 | 136 | 36 | 107 |

Tableau 3 : Couverture de séquence, redondance et nombre de peptides identifiés pour les analyses de ERR α , PGC1 α et du complexe ERR α /PGC1 α .

Après les expériences d'échange H/D et la vérification manuelle des profils isotopiques, seuls 36 peptides ont été conservés pour ERR α laissant tout de même chaque domaine représenté par au moins un peptide excepté la région charnière entre le DBD et le LBD. Bien que plusieurs temps de deutération aient été effectués (0,5, 2, 10, 30 et 60 minutes), un manque de reproductibilité est observé pour les temps longs de deutération (30 et 60 minutes). Ces cinétiques ont donc été retirées.

4.2.2. HDX-MS

Les résultats de HDX-MS montrent plusieurs différences d'incorporation en D statistiquement significatives entre ERR α et le complexe ERR α /PGC1 α , avec (Fig. 14, Annexe 26) :

- La protection de ERR α en complexe avec PGC1 α au niveau des résidus 15-38 du domaine AF-1, 104-114 et 142-153 du DBD et 295-307 de l'hélice H6 du LBD.
- La déprotection des résidus 6-14 de la partie N-terminale de ERR α en interaction avec PGC1 α , mais aussi le peptide 147-152 du DBD et la partie C-terminale de LBD avec les hélices H10-12 (384-419).

L'hélice H4 et la boucle H8-H9 de ERR α , proches de PGC1 α dans la structure cristallographique (PDB : 3D24)²⁷⁹, ne sont pas significativement affectées (pas de différence d'incorporation en D) par la présence de PGC1 α . Néanmoins, la déprotection de l'hélice H12, traduisant d'un changement conformationnel montré par trois peptides recouvrant les mêmes résidus, est caractéristique des RNs en interaction avec des co-régulateurs. Il est également intéressant de noter la forte protection (jusqu'à 35%) du DBD de ERR α en présence de PGC1 α , une information nouvelle et importante apportée par nos données de HDX-MS.

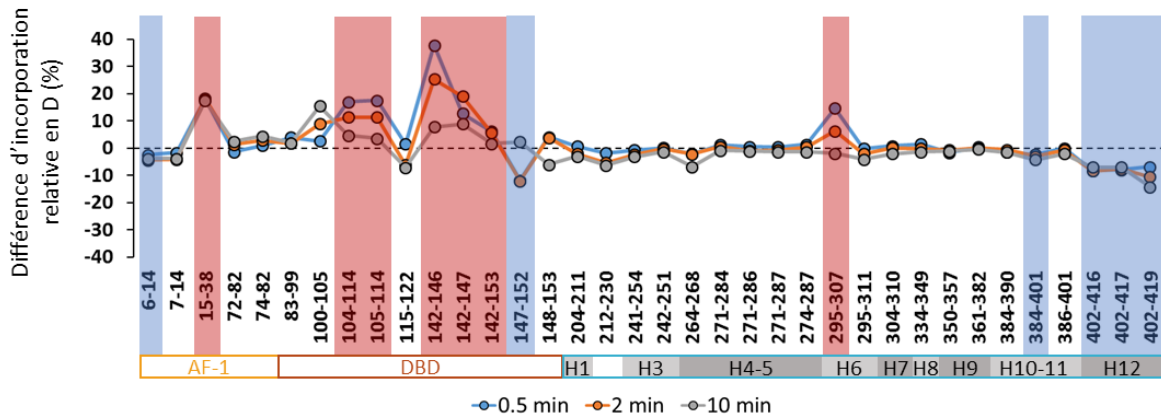


Figure 14 : Représentation des différences d'incorporations relatives en D entre ERRα et le complexe avec PGC1α pour chaque peptide identifié de ERRα pour aux différents temps de deutération. Les peptides encadrés présentent des différences statistiquement significatives d'incorporation de deutérium (MEMHDX, p-value à 0,01¹³⁴) supérieures à 2% (rouge) relatant d'une protection du complexe comparé à la protéine seule et inférieures à -2% (bleu) montrant une déprotection. Les différents domaines de la protéine sont annotés comme la fonction AF-1, le DBD, la région charnière entre le DBD et le LBD (hinge) et le LBD.

Dans le cas de PGC1α (Annexe 27), aucune différence d'incorporation en D n'a été identifiée entre la protéine seule et le complexe ERRα/PGC1α. Une saturation de l'incorporation en D est remarquée pour les deux états (Fig. 15) bien que la deutération a été réalisée à un pH de 6,5. Seul le peptide 96-101 présente une augmentation au cours du temps de l'incorporation en D de plus de 5% pendant les deux états restent similaires.

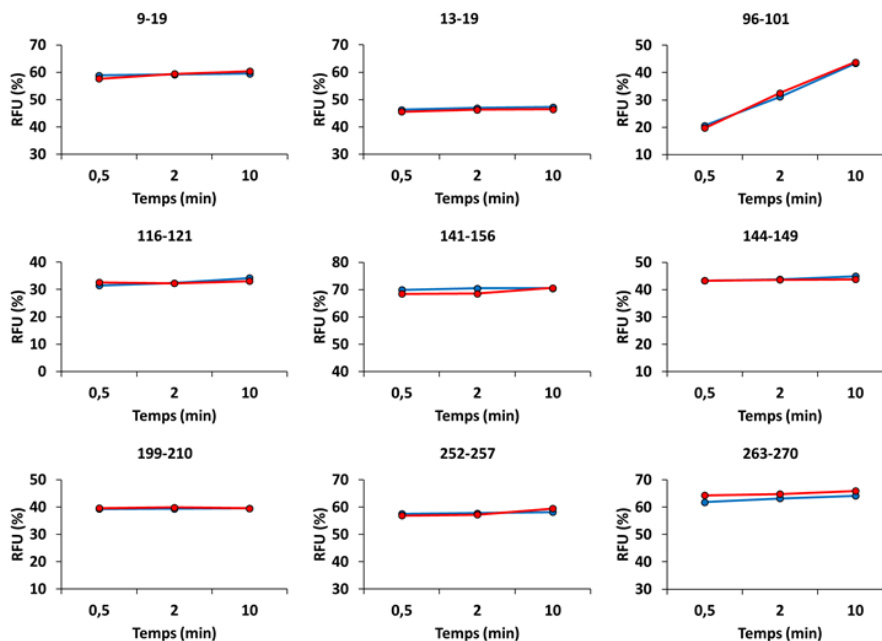


Figure 15 : Représentation de l'incorporation en D relative (RFU) de PGC1α (rouge) et du complexe ERRα/PGC1α (bleu) en fonction du temps de deutération pour différents peptides de PGC1α identifiés en HDX-MS.

4.2.3. XL-MS

Des expériences de XL-MS ont alors été menées dans le but d'obtenir des informations structurales supplémentaires pour chacune des protéines en interaction. Ainsi, 11 cross-links intra-ERR α , 1 intra-PGC1 α (Annexe 28) et 13 inter-ERR α /PGC1 α (Tableau 4) ont été identifiés suite à l'analyse du complexe.

| Protéine 1 | Protéine 2 | Site 1 | Site 2 | Agents pontants | | Domaine site 1 | Domaine site 2 | Distance C α -C α (Å) ERR α /PGC1 α |
|------------|------------|--------|--------|-----------------|------|----------------|----------------|--|
| | | | | DSAU | DSBU | | | |
| PGC1 | ERR | K184 | K189 | X | X | - | Hinge | - |
| PGC1 | ERR | K184 | K403 | X | | - | H10-11 | 25,6 |
| PGC1 | ERR | K184 | K407 | X | | - | LBD | 26,1 |
| PGC1 | ERR | K192 | K403 | X | | - | H10-11 | 26,4 |
| PGC1 | ERR | K194 | K407 | X | | - | LBD | 26,9 |
| PGC1 | ERR | S195 | K403 | X | | - | H10-11 | 24,8 |
| PGC1 | ERR | S195 | K407 | X | | - | LBD | 29,7 |
| PGC1 | ERR | K201 | K394 | X | | - | H10-11 | 29,4 |
| PGC1 | ERR | K201 | K403 | X | X | - | H10-11 | 25,5 |
| PGC1 | ERR | K201 | K407 | X | | - | LBD | 31,9 |
| PGC1 | ERR | K224 | K76 | X | X | DR | DBD | 71,1 |
| PGC1 | ERR | K224 | K189 | X | | DR | Hinge | - |
| PGC1 | ERR | K235 | K403 | X | | DR | H10-11 | 56,1 |

Tableau 4 : Cross-links inter-protéines identifiés pour l'analyse du complexe ERR α /PGC1 α avec une description des sites, de l'agent pontant utilisé, des domaines et des distances mesurées sur la structure du complexe obtenue par nos collaborateurs.

Ces peptides pontés sont localisés au niveau de plusieurs domaines (Fig. 16), dont :

- La deuxième région désordonnée de PGC1 α ainsi que les résidus K184, K192, K194, S195 et K201 (Tableau 4) précédant le motif inversé LxxLL de PGC1 α (₂₁₀LLKYL₂₁₅). Ce motif est connu comme étant le site de reconnaissance spécifique pour ERR α ²⁸².
- Les domaines AF-1 et DBD de même que la région charnière entre le DBD et le LBD de ERR α sont les régions comprenant les résidus pontés intra-ERR α . Parmi ces peptides, deux peptides pontés homéotypiques (pontage de deux acides aminés identiques de deux peptides de même séquence) des résidus K138 et K189 ont été identifiés confirmant la dimérisation de ERR α .
- La partie N-terminale du DBD de ERR α est proche du motif LxxLL inversé de PGC1 α (K224-K76) ainsi que la région charnière entre le DBD et le LBD avec le résidus K189.
- Les hélices H10-11 avec les résidus K394 et K403 et la boucle entre les hélices H11-12 représentée par la lysine 407 de ERR α sont proches de PGC1 α . Ces résultats sont en

accord avec les domaines d'interaction des RNs avec les sites de reconnaissance des co-régulateurs²⁴⁹.

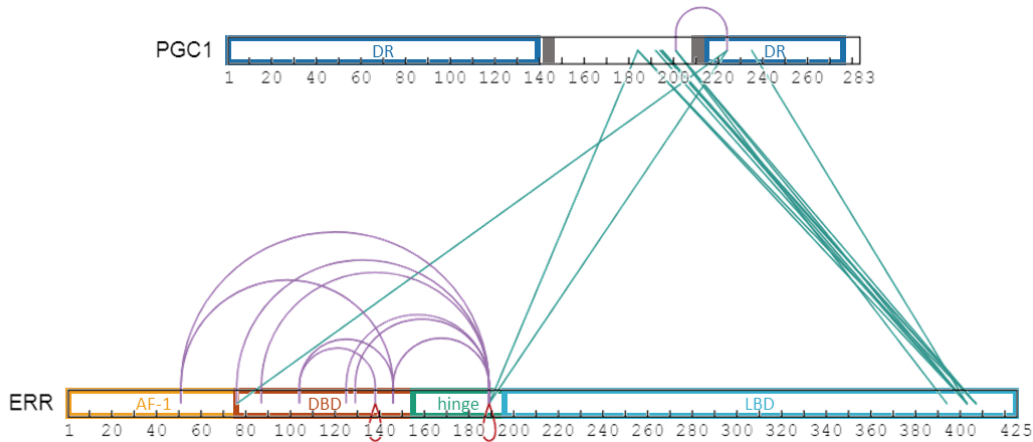


Figure 16 : Représentation schématique des peptides pontés intra-protéine (violet) et inter-protéines (vert) du complexe ERRα/PGC1α (obtenus avec Xinet²¹⁶) avec les régions désordonnées (DR), et les motifs LXXLL ou LKYLL (gris) de PGC1α ainsi que la fonction AF-1, le DBD, la région charnière entre le DBD et le LBD (hinge) et enfin le LBD de ERRα.

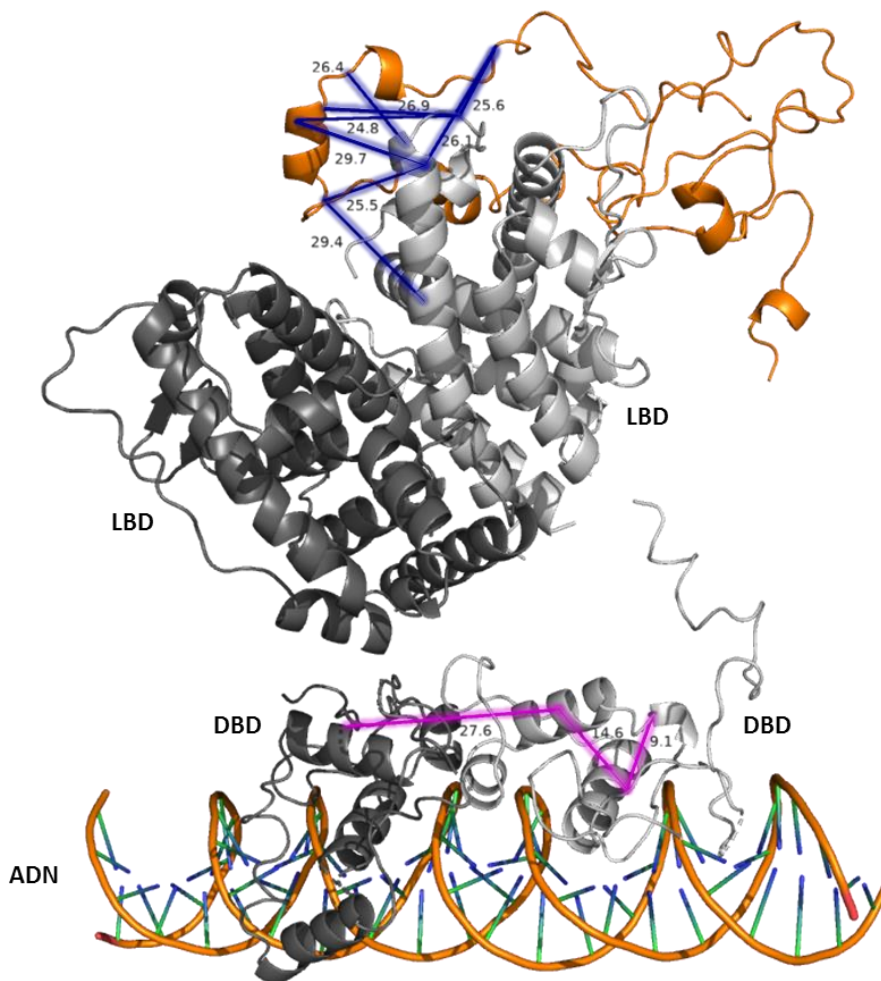


Figure 17 : Représentation des cross-links validés intra (violet) et inter (bleu) de ERRα/PGC1α sur la structure cryo-EM du complexe. Les deux monomères de ERRα sont différenciables par leurs couleurs (gris clair et foncé) et PGC1α (orange).

Une structure tronquée du complexe $ERR\alpha/PGC1\alpha$ a été obtenue par nos collaborateurs en cryo-EM avec $ERR\alpha^{DBD-LBD}$ et $PGC1\alpha_{122-235}$ (Fig. 17). Cette structure nous permet de comparer les mesures de distances C α -C α entre les acides aminés pontés et les intervalles de distance fixés pour chaque agent pontant (6-30Å pour le DSBU et 6-26Å pour le DSAU).

Toutefois, en considérant la forte dynamique de $PGC1\alpha$ et différentes références bibliographiques^{70, 73, 283-284} portant sur les intervalles de distance C α -C α pour plusieurs cross-linkers, des valeurs plus élevées de distances maximales ont été choisies avec 30Å pour le DSAU et 32Å pour le DSBU. Ainsi, 72% des distances mesurables entre les résidus pontés inter- $ERR\alpha/PGC1\alpha$ sont comprises dans les intervalles (Fig. 17). De la même manière, les trois distances mesurables correspondant aux peptides intra- $ERR\alpha$ sont inférieures à 30Å. Les données expérimentales de XL-MS sont donc en accord avec la structure cryo-EM du complexe $ERR\alpha/PGC1\alpha$.

Finalement, les résultats obtenus en HDX-MS ont permis de mettre en lumière des zones du DBD de $ERR\alpha$ impacté par l'interaction avec $PGC1\alpha$. De la même manière, un changement conformationnel de l'hélice H12 de $ERR\alpha$ a pu être montré. Les expériences de XL-MS ont :

- confirmé la présence du dimère de $ERR\alpha$ grâce aux cross-links homéotypiques,
- indiqué la proximité du DBD de $ERR\alpha$ avec $PGC1\alpha$ complémentaire aux données obtenues en HDX-MS,
- révélé que la partie C-terminale du LBD de $ERR\alpha$ entre en interaction avec des régions proches du motif inversé de $PGC1\alpha$,
- permis de valider la structure cryo-EM du complexe.

IV. Conclusions de l'étude des RNs et perspectives

Les RNs, de par leur diversité et la grande variété de mécanismes associés, font l'objet de plus en plus d'études structurales. Une meilleure compréhension de cette importante famille de protéines est effectivement recherchée et grâce à ces travaux des réponses ont pu être apportées. La combinaison d'approches de MS structurale (HDX-MS et XL-MS) a pu montrer :

- L'impact de mutation ponctuelle de $PPAR\gamma^{LBD}$ sur la liaison de co-régulateur en nMS et HDX-MS avec les taux de complexation avec le co-régulateurs mesurés pour chaque protéine mutée et l'étude de l'interaction du NCoR sur le WT et F310S.

- La possibilité de classier les mutants de PPAR γ ^{LBD} en HDX-MS en fonction de leurs activités transcriptionnelles (activatrices ou inhibitrices).
- Les limitations des techniques de HDX-MS et XL-MS dans l'étude du complexe ERR α /LSD1 ne permettant pas pour le moment d'identifier des régions d'interactions.
- L'apport de données de MS structurale pour la caractérisation du complexe ERR α /PGC1 α et la validation de la structure cryo-EM.

Des analyses supplémentaires en nMS avec d'autres co-régulateurs et PPAR γ ^{LBD} pourraient permettre de classier de nouveaux co-régulateurs en fonction de leurs types (CoA ou CoR). De la même manière, d'autres mutants dont l'activité transcriptionnelle n'est pas déterminée pourront être étudiés en HDX-MS. Les profils obtenus en HDX-MS permettent la classification des mutants en fonction de leurs activités transcriptionnelles.

Des expériences de M-Ph sont envisageables pour confirmer l'interaction entre ERR α et LSD1. Des analyses de HDX-MS à un pH de 6 avec des temps court de deutération pourront être menées dans le but d'observer l'impact de la liaison de ERR α sur PGC1 α . La mesure de l'échange inverse D/H de PGC1 α peut également être intéressante sachant que cette protéine est intrinsèquement désordonnée et permettrait d'obtenir une deutération totale plus aisément donnant des indications sur les peptides échangeant plus rapidement.

Chapitre 2 : Caractérisation de protéines membranaires dans un environnement lipidique par HDX-MS

I. Introduction aux protéines membranaires et HDX-MS

Les protéines membranaires (PM) jouent un rôle majeur dans les cellules et consistent 30% du protéome humain²⁸⁵⁻²⁸⁶. Elles sont nécessaires à la production d'énergie, la transmission de signaux, l'activité enzymatique et le transport de petites molécules²⁸⁷⁻²⁸⁸. Compte tenu de leurs nombreuses fonctions, les PMs sont subdivisées en plusieurs familles incluant les canaux ioniques, transporteurs, récepteurs et enzymes²⁸⁹ (Fig. 1). Les PMs contiennent des domaines transmembranaires (DTMs) hydrophobes très organisés.

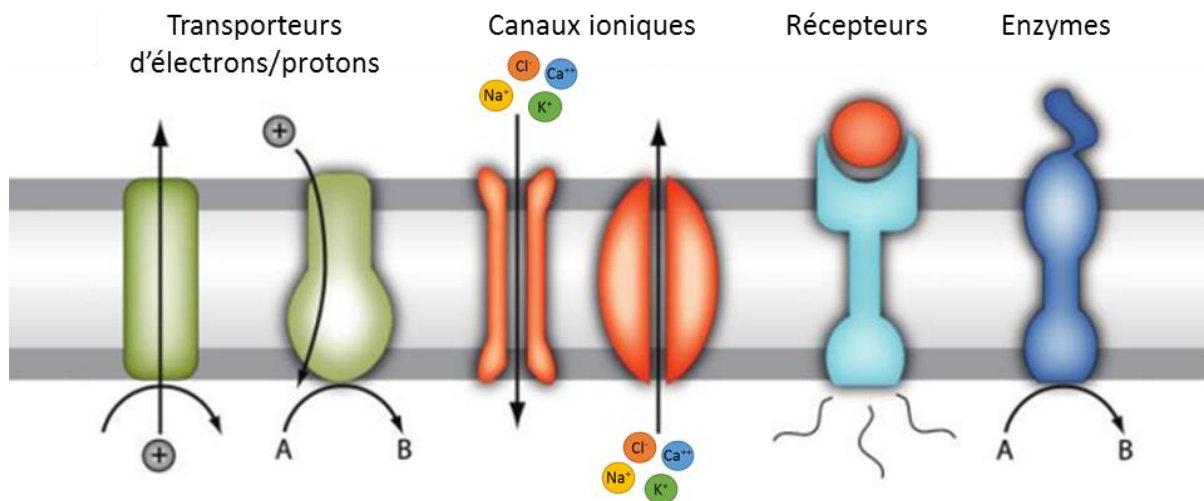


Figure 1 : Représentation de différentes familles de protéines membranaires (d'après ²⁹⁰).

Les PMs constituent 50% des cibles thérapeutiques et font l'objet de beaucoup d'études structurales pour mieux comprendre leurs relations structure-fonction²⁹¹. Néanmoins, la caractérisation structurale des PMs reste difficile et compliquée et ceci pour les raisons suivantes¹²⁵ :

- Leur faible rendement de production
- Leur instabilité en solution (protéine amphiphile)
- L'emploi de détergents, souvent en forte concentration, dans le but de permettre leur solubilité et de conserver leur conformation native

Du fait de ces enjeux, un nombre limité de techniques biophysiques est utilisé pour l'étude des PMs au niveau moléculaire. Parmi celles-ci, l'approche HDX-MS a sa place, car elle fournit

des informations sur les changements conformationnels et la dynamique d'échange sans nécessiter de cristallisation (qui fige les PMs dans un état conformationnel thermodynamiquement stable)^{9, 292}. De nombreuses études ont d'ores et déjà été menées sur les protéines membranaires en HDX-MS apportant ainsi des informations toujours plus complètes sur leurs structures^{111, 293}. Plusieurs revues scientifiques ont été publiées donnant des recommandations pour optimiser la préparation et les analyses en HDX-MS^{125, 294}, telles que :

- L'utilisation de protéases telles que la nepenthesine II au lieu de la pepsine, classiquement utilisée¹²¹.
- La stabilisation des PMs par des détergents compatibles avec la MS comme par exemple le dodécyl- β -D-maltoside (DDM)²⁹⁵.
- La présence de systèmes mimant la membrane comme les bicelles, nanodisques, liposomes, amphipols et polymères de type SMALP²⁹⁶ permettant d'analyser les PMs au plus proche de leur état natif.

II. HDX-MS pour suivre l'effet de la fixation de ligands sur A_{2A} rant21

1. Contexte analytique

Les récepteurs couplés aux protéines G (GPCRs) sont une des familles les plus importantes des PMs et tiennent le rôle de médiateurs dans le cadre de la majorité des réponses cellulaires hormonales et de la neurotransmission²⁹⁷. Ils incluent le récepteur à l'adénosine A_{2A} qui comporte 7 hélices transmembranaires et un domaine de fixation des ligands connu comprenant les résidus F168, N253 et I274²⁹⁸. Bien que le ligand endogène d'A_{2A} soit l'adénosine, d'autres ligands ont été découverts comme entrant en interaction avec le récepteur. Ces ligands agonistes (NECA et CGS21680) ou antagonistes (ZM241385, caféine et théophylline) initient des réponses biologiques différentes et sont étudiés en tant que médicaments anti-inflammatoires ou pour soigner les maladies neurodégénératives (Fig. 2)²⁹⁹⁻³⁰⁴. Il est primordial de comprendre l'impact de ces ligands sur la structure des protéines. Des mutations du récepteur A_{2A} (Annexe 2 et 29) ont déjà été rapportées comme favorisant la liaison de ligand agoniste (mutation rag23) ou antagoniste (mutation rant21) car elles stabiliseraient A_{2A} dans une conformation active ou inactive, respectivement³⁰⁵.

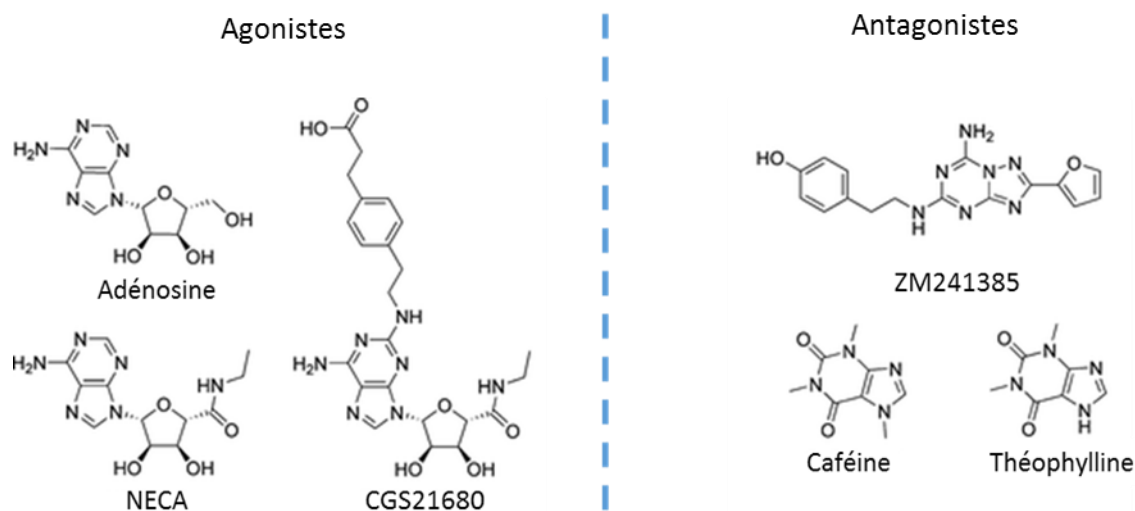


Figure 2 : Structure chimique des ligands (agonistes et antagonistes) connus de A_{2A} (d'après ²⁹⁹).

2. Objectifs

Le but de ce projet est, dans un premier temps, d'étudier en HDX-MS l'impact de la liaison des deux ligands ZM241385 (antagoniste) et CGS21680 (agoniste) sur la structure de A_{2A} rant21. Les changements conformationnels observés suite à l'interaction avec les ligands permettraient de déterminer des régions spécifiques à la liaison de chaque type de ligand (agoniste versus antagoniste). La connaissance de ces régions permettrait finalement la classification d'autres ligands par HDX-MS.

Le projet a été réalisé dans le cadre d'une collaboration avec la plateforme IMPReSS (avec le Dr. R. Wagner, Strasbourg).

3. Matériel, méthodes et optimisations des expériences en HDX-MS

Le mutant de A_{2A} (rant21) a été exprimé et purifié par nos collaborateurs avec les acides aminés L85A, A119T, A153K et A270V mutés (Annexe 2 et 29).

Pour mieux comprendre la dynamique de l'interaction des ligands avec A_{2A} rant21, des expériences de HDX-MS ont été réalisées sur la récepteur A_{2A} rant21 seul ou en présence de ligands. Avant d'effectuer l'échange H/D, la digestion des protéines a été optimisée. Dans le cas de l'analyse d'A_{2A} rant21, un tampon mimant la deutération (50mM HEPES, 200mM NaCl et 0,05% DDM à un pH de 7,4) a été utilisé ainsi qu'un tampon d'arrêt du marquage (100mM Glycine et 2M guanidine-HCl à un pH de 1,5). La digestion a été effectuée sur une colonne contenant de la nepenthesine II immobilisée (Affipro, République Tchèque) avec un débit de

100 $\mu\text{L}/\text{min}$. 52 pmol de protéine A_{2A} rant21 ont été injectées. Après digestion, les peptides ont été séparés sur une colonne C18 et analysés sur un Synapt-G2 Si (Waters). Les peptides ont été identifiés avec le logiciel PLGS 2.5.3 (Waters) et filtrés avec DynamX 3.0 (Waters). Les filtres appliqués sont un minimum d'intensité de 1000 et les peptides doivent être présents dans au moins 2 des 3 réplicats pour chaque état. Un minimum de « produit par acide aminé » a été fixé à 0,2 au lieu de 0,3 permettant d'augmenter considérablement le nombre de peptides identifiés (plus de 70 peptides) (Tableau 1). Dans ces conditions, la couverture de séquence (Annexe 30) obtenue pour la protéine A_{2A} (Tableau 1) est très satisfaisante (> 89%) de même que la redondance ($\geq 6,9$). De plus, la couverture de séquence des DTMs est plus importante (83,8%) avec un « produit par acide aminé » de 0,2.

| Analyse | A _{2A} rant21 | |
|---------------------------------|------------------------|------|
| | 0,3 | 0,2 |
| « produit par acide aminé » | 0,3 | 0,2 |
| Couverture de séquence (%) | 82,5 | 89,9 |
| Redondance | 4,6 | 6,9 |
| Nombre de peptides identifiés | 163 | 233 |
| Couverture de séquence des DTMs | 77,2 | 83,8 |

Tableau 1 : Couverture de séquence, redondance et nombre de peptides identifiés pour l'analyse de A_{2A} rant21 selon le minimum de produit par acide aminé appliqué (0,2 ou 0,3).

Lors de l'expérience d'échange H/D, seul 70 peptides ont été conservés donnant une couverture de séquence de 70,8 % et une redondance de 2,5 pour la comparaison en absence/présence du ZM241385. Néanmoins, 6 des 7 DTMs sont couverts par au moins 2 peptides et les régions connues de liaison des ligands sont également recouvertes par des peptides.

Les profils isotopiques ont été vérifiés manuellement sans correction de l'échange inverse. Les résultats ont été validés statistiquement avec une p-value de 0,01 avec le logiciel MEMHDX¹³⁴.

4. Suivi de l'incorporation en D du récepteur A_{2A} rant 21

Le suivi de l'incorporation en D de A_{2A} rant21 à différents temps (0,5, 2, 10, 30 et 60 minutes) des peptides des DTMs nous permet d'observer la dynamique de ces régions en l'absence de ligand. En effet, les peptides de ces domaines (résidus 20-31 ; 43-65 ; 113-126 ; 132-141 ; 234-241 ; 245-253 ; 271-288) présentent une incorporation en D généralement plus faible que les autres régions de A_{2A} rant21, majoritairement en dessous de 30% (Fig. 3) en accord avec leurs

caractères hydrophobes. Ces résultats sont en accord avec la conservation de la conformation native de A_{2A} rant21.

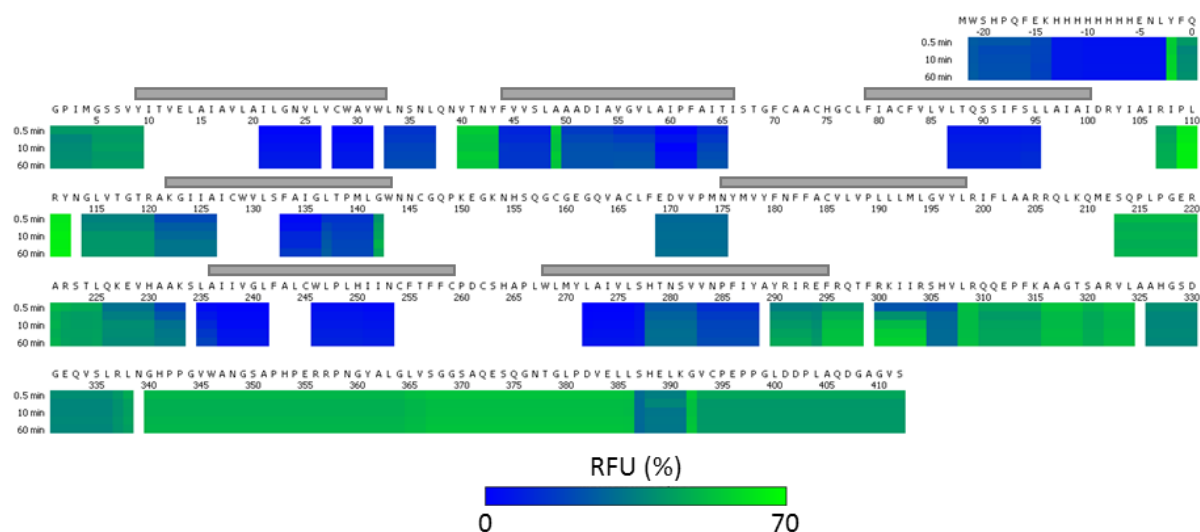


Figure 3 : Incorporation en D relative (%) représentée sur la séquence d'A_{2A} rant21 en fonction des temps d'incubation. DTMs en gris.

5. Dynamique de fixation des ligands sur A_{2A} rant21

Des expériences HDX-MS ont ensuite été réalisées en présence de ligands (2 excès molaire). Les incorporations en D ont été comparées entre A_{2A} rant21 seul et le complexe A_{2A} rant21/ligand (Annexe 31). De manière générale, les comparaisons montrent des différences statistiquement significatives positives (MEMHDX avec une p-value à 0,01¹³⁴, Fig. 4), traduisant de la protection de certaines régions de A_{2A} rant21 en présence des ligands.

Nous avons commencé par suivre l'effet de la fixation de ZM241385, un ligand antagoniste de référence. Plusieurs régions de A_{2A} rant21 sont protégées en présence du ligand antagoniste ZM241385 (Fig. 4A) :

- Le domaine connu de liaison au ligand (résidus F168 et N253) représenté par les peptides 168-175, 245-250 et 245-253. Cependant, l'isoleucine 274 également montrée comme importante pour la fixation du ligand et recouvert par trois peptides (271-276 ; 272-276 ; 272-280), ne présente pas de différence significative. Cet acide aminé étant proche d'un site de mutation de la protéine (A270V) de A_{2A} rant21, l'absence de différence significative des peptides de cette zone pourrait être due à la mutation.
- Les peptides 59-65, 276-288, 277-282 et 277-288 : cette région forme une poche pour accueillir le groupe phénol de ZM241385²⁹⁸. Celle-ci implique aussi la partie C-

terminale de l'hélice transmembranaire 7 (DTM7), les peptides 281-288, 289-293, 289-294 et 294-298.

- Les DTM3 et DTM4 ainsi que la boucle entre les deux sont également fortement impactés.

Nous avons ensuite évalué la liaison d'un autre type de ligand, le CGS21680 (agoniste). La liaison de ce ligand entraîne aussi la protection statistiquement significative de nombreuses régions de A_{2A} rant21 (Fig. 4B) telles que :

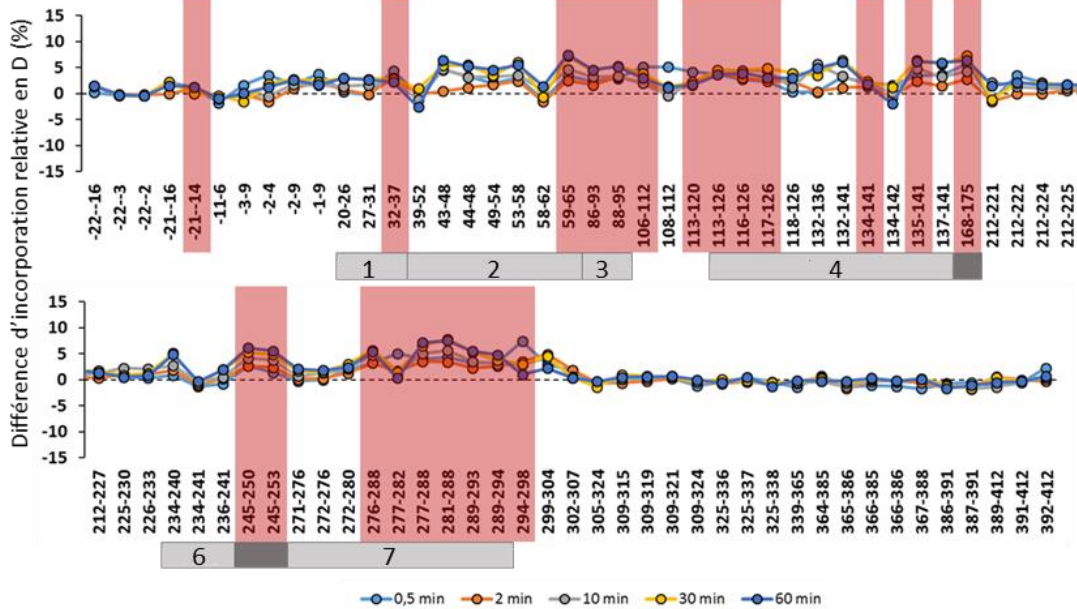
- Les hélices transmembranaires 2 à 7 sont statistiquement protégés traduisant d'un réarrangement structural important de la quasi-totalité (84% des peptides identifiés) de A_{2A} rant21 en présence de l'agoniste CGS21680.
- Les peptides (245-250 ; 245-253 ; 273-280) qui contiennent les résidus (N253 et I274) impliqués dans la fixation du ligand sont protégés, ce qui confirme expérimentalement l'interaction entre A_{2A} rant21 et le CGS21680. La phénylalanine 168 n'étant pas recouverte par un peptide en HDX, nous ne disposons pas d'information sur ce résidu.
- La partie C-terminale est aussi impactée par la liaison du CGS21680 avec cinq peptides (299-304 ; 309-315 ; 309-319 ; 325-336 ; 325-338) qui montrent une diminution du taux d'incorporation en D.

Les comparaisons de l'incorporation en D de A_{2A} rant21 seul et en présence de chacun des ligands ont permis de confirmer la fixation des ligands par la protection des zones connues de liaisons des ligands. Elles ont aussi révélé que la liaison du ligand agoniste CGS21680 avec A_{2A} rant21 induit une protection de son domaine C-terminal, ce qui n'est pas le cas lors de la liaison de l'antagoniste ZM241385. De la même manière, l'hélice entre le DTM5 et DTM6 recouverte par les peptides (212-222 ; 212-224 ; 212-225 ; 212-227) est protégée suite à la fixation du CGS21680 à A_{2A} rant21 et ne l'est pas pour le complexe A_{2A} rant21/ZM241385. Le changement conformationnel de ces deux zones est donc caractéristique de la liaison d'un ligand agoniste au récepteur A_{2A} rant21.

L'étude de complexes avec d'autres ligands dont le type (antagoniste ou agoniste) est connu pourrait être réalisée pour confirmer ces résultats et ces spécificités de réarrangements structuraux. Néanmoins, le récepteur étudié est le mutant A_{2A} rant21, qui est stabilisé dans une conformation inactive. Par conséquent, il est attendu que la liaison d'un ligand agoniste

provoquera plus de réarrangements structuraux que la liaison du ligand antagoniste ce qui correspond à nos résultats HDX-MS. Ces résultats pourraient alors être extrapolés pour la classification en HDX-MS de ligands dont le type n'est pas connu.

A) Impact du ZM241385 sur A_{2A}



B) Impact du CGS21680 sur A_{2A}

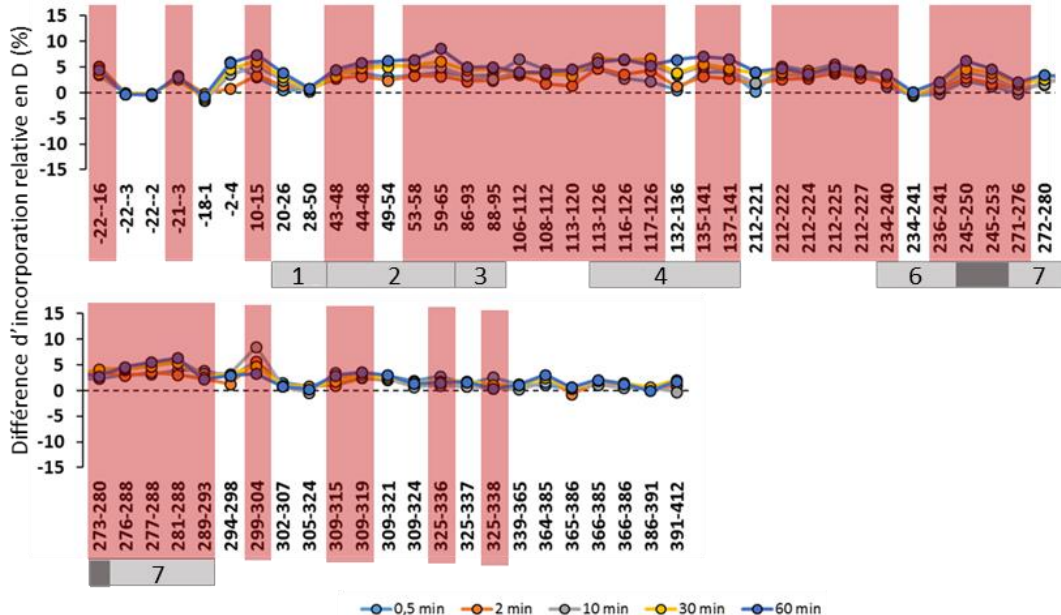


Figure 4 : Représentation des différences d'incorporations relatives en D entre A_{2A} rant21 et les complexes (A) A_{2A} rant21/ZM241385 et (B) A_{2A} rant21/CGS21680 pour chaque peptide identifié de A_{2A} rant21 pour les différents points de deutération. Les peptides encadrés présentent des différences statistiquement significatives d'incorporation en deutérium (MEMHDX, p-value à 0,01¹³⁴) supérieures à 2% (rouge) et montrent une protection globale du complexe A_{2A} rant21/ligand comparé à A_{2A} rant21 seul. Les DTMs sont numérotés de 1 à 7 et les peptides contenant les trois résidus impliqués dans la fixation des ligands (F168, N253 et I274) sont représenté par un encadré gris foncé.

6. Valorisation

L'étude en HDX-MS de la fixation du ligand ZM241385 sur le récepteur à l'adénosine A_{2A} a été valorisée dans la revue suivante (Annexe 32) :

Banères, J. L.; Botzanowski, T.; Boutin, J. A.; Calamini, B.; Castel, J.; Catoire, L. J.; Cianférani, S.; Demesmay, C.; Ferguson, G.; Ferry, G.; Kniazeff, J.; Krimm, I.; Langer, T.; Lebon, G.; Ley, M.; Nyerges, M.; Schwob, M.; Venien-Bryan, C.; Wagner, R.; Zeder-Lutz, G.; Zilian-Stohrer, C., **Biophysical Dissection of Isolated GPCRs: The Adenosine A_{2A} Receptor under the Bistouries.** *Receptors* **2023**, 2, 47-92.

7. Conclusions et perspectives

Les expériences de HDX-MS menées sur A_{2A} ont permis de montrer que :

- Une couverture de séquence satisfaisante (>70%) peut être obtenue pour l'analyse de A_{2A} suite à l'optimisation de la phase pré-analytique (utilisation de différentes protéases) et post-analytique (filtres appliqués lors de l'identification des peptides).
- L'impact de la liaison de peptides molécules (ligands) sur la conformation de A_{2A} peut être défini. De cette manière, différentes zones ont pu être montrées comme spécifiques à la liaison d'un ligand suivant son type (agoniste/antagoniste).
- Des informations de dynamiques même sur des DTMs hydrophobes.

Finalement, il serait envisageable de continuer les expériences de HDX-MS sur A_{2A} avec d'autres ligands dont le type (agoniste ou antagoniste) n'est pas connu dans le but de les classifier, si possible, par la détermination des changements conformationnels significatifs de A_{2A} induits par ces ligands observés en HDX-MS.

III. Etude par HDX-MS de l'interaction de EPAC1 à la membrane

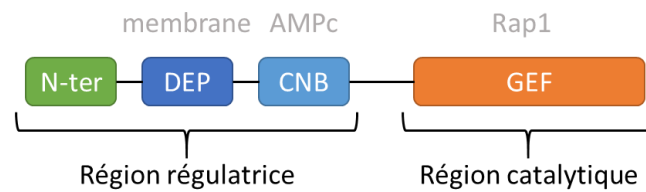
1. Contexte analytique

Les protéines périphériques à la membrane font également partie intégrante des protéines membranaires. Ce sont des protéines qui interagissent avec la membrane de façon non-covalente, comme les protéines d'échange directement activées par l'AMPC (adénosine monophosphate cyclique) aussi appelées EPAC³⁰⁶. En effet, ces protéines ont été identifiées comme étant partiellement activées par la membrane³⁰⁷. Elles sont membres de la famille des facteurs d'échange de nucléotides de guanine (GEFs) pour les Rap GTPases³⁰⁸. EPAC1 et EPAC2

sont les deux isoformes principales d'EPAC et sont encodés par différents gènes³⁰⁹. Elles possèdent (Fig. 5A) :

- Une région régulatrice N-terminale avec un domaine DEP (Dishevelled/Egl-10/pleckstrin) interagissant avec la membrane et un domaine de liaison aux nucléotides cycliques (CNB) répondant à l'AMPc.
- Une région catalytique C-terminale avec un domaine GEF se liant avec les protéines Rap1.

A)



B)

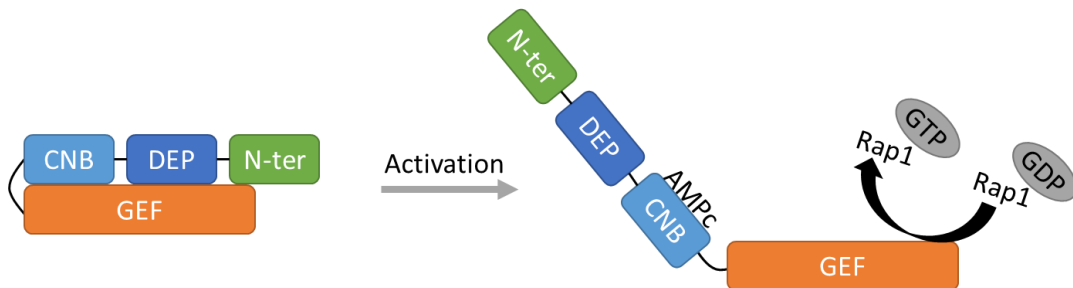


Figure 5 : (A) Représentation schématique des domaines de la protéine EPAC1 avec leurs partenaires connus (gris). (B) Illustration du changement conformationnel de EPAC1 induit par l'activation de la protéine (d'après ³¹⁰).

De nombreuses études ont déjà été menées sur EPAC2 alors que EPAC1 reste peu caractérisée structuralement. Ainsi, la manière dont EPAC1 entre en interaction avec la membrane reste encore peu comprise. Néanmoins, il a déjà été montré que l'activation de EPAC1 entraîne un réarrangement de sa structure avec une ouverture du domaine régulateur par rapport à la région catalytique (Fig. 5B)³¹⁰. EPAC1 est exprimée dans le cœur et est à l'origine de certains troubles cardiaques faisant d'elle une cible pour la recherche d'inhibiteurs. L'activité inhibitrice du CE3F4 a été découverte par criblage à haut débit³¹¹. Il a été montré que le CE3F4 n'entre pas en compétition avec l'AMPc et stabilise une conformation d'EPAC1 qui restreint l'accès au site catalytique de Rap1^{310, 312}. L'étude de l'interaction du CE3F4 avec EPAC1 a montré que l'énantiomère R du ligand CE3F4 (CE3F4-R) se fixe avec 10 fois plus d'affinités à EPAC1 que l'énantiomère S³¹³.

Le projet a été réalisé dans le cadre d'une collaboration avec le Laboratoire de Biologie et Pharmacologie Appliquée (avec le Dr. J. Cherfils, ENS Paris-Saclay). Ce projet a eu pour objectif d'affiner les connaissances structurales des PMs. Pour cela, différentes techniques ont été employées pour l'analyse d'EPAC1 comme la SAXS, la X-ray et des essais de cinétique d'échange de nucléotides.

2. Matériel, méthodes et optimisations des expériences en HDX-MS

Dans le but de mieux comprendre la dynamique de l'activation de EPAC1 et l'effet de la fixation de l'inhibiteur CE3F4, des expériences de HDX-MS ont été réalisées sur la protéine en présence/absence d'AMPc et/ou de liposomes. Deux jeux de données ont été réalisés sur EPAC1, un premier avec liposomes et l'AMPc et un deuxième en présence du CE3F4-R et de l'AMPc avec plus d'un an d'écart. Dans le cas de l'analyse de EPAC1 dans un environnement lipidique un mélange de liposomes contenant 48% de phosphatidylcholine, 20% de phosphatidyléthanolamine, 10% de phosphatidylsérine, 20% de cholestérol et 2% de phosphatidylinositol-4,5-bisphosphate ont été utilisés. Les liposomes ont été extrudés le jour même de l'analyse à travers un filtre de porosité 0,2 µm. Avant d'effectuer l'échange H/D, la digestion des protéines est optimisée. Dans le cas de l'analyse d'EPAC1, la combinaison du tampon mimant la deutération (50mM HEPES, 150mM NaCl, 2mM de MgCl₂ et 2mM de β-mercaptoéthanol à un pH de 8) à celle du tampon d'arrêt du marquage (150mM Glycine et 4M guanidine-HCl à un pH de 2,2) est utilisée. Dans le cadre de toutes les analyses avec de l'AMPc, 4,5mM de cette molécule a été ajoutée au tampon mimant le marquage. La digestion (Fig. 6), à la pepsine (comparaisons EPAC1 vs EPAC1/cAMP et EPAC1 vs EPAC1/cAMP/liposomes) ou à la nepenthesine II (comparaison EPAC1 vs EPAC1/CE3F4 et EPAC1 vs EPAC1/cAMP/CE3F4), a été effectuée avec une colonne contenant l'enzyme immobilisée (pepsine, Waters ; nepenthesine II, Affipro) et un débit appliqué de 100µL/min. Les protéines (avec/sans partenaires) ont été préparées fraîchement tous les trois réplicats, les liposomes ont été extrudés et ajoutés à une concentration, estimée suivant la concentration des liposomes initiale, à 3,125mM. 19pmoles de protéine ont été injectées en MS. Après digestion, les peptides ont été séparés sur une colonne C18 et analysés sur un Synapt-G2 Si (Waters).

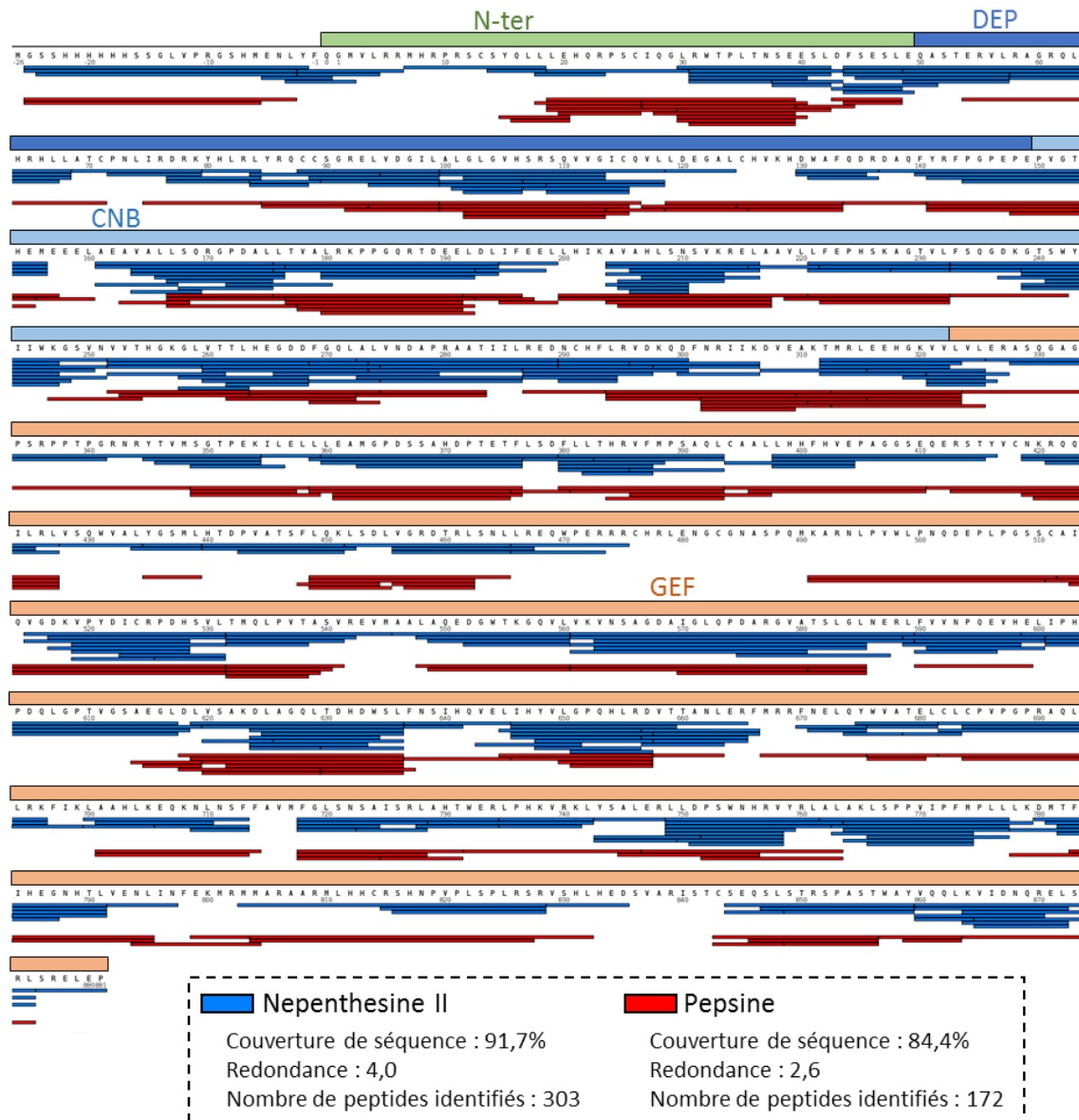


Figure 6 : Carte de recouvrement de séquence des peptides identifiés de EPAC1 obtenus suite à la digestion à la nepenthesine II (bleu) et à la pepsine (rouge) sur la séquence de EPAC1 réalisée à l'aide de ¹⁴⁰. Les domaines sont représentés en vert pour le N-terminale, en bleu foncé pour le DEP, en bleu clair pour le CNB et en orange pour le GEF.

Le second jeu d'analyses d'EPAC1 a été réalisé avec l'IM activée. Les peptides ont été identifiés avec le logiciel PLGS 2.5.3 (Waters) et filtré avec DynamX 3.0 (Waters).

Dans ces conditions, pour le premier jeu de données, une digestion à la pepsine a permis d'obtenir la plus grande couverture de séquence avec 84,4 % (Fig. 6). Pour le second jeu de données, la nepenthesine II a été utilisée et a permis d'identifier 303 peptides avec une redondance de 4 et une couverture de séquence de 91,7% (Fig. 6).

Après avoir réalisé les expériences de HDX-MS, un recouvrement de séquence de 70% minimum a été obtenu pour chaque comparaison. Les profils isotopiques ont été vérifiés

manuellement sans correction de l'échange inverse. Les résultats ont été validés statistiquement avec une p-value de 0,01 avec le logiciel MEMHDX¹³⁴ (institut Pasteur).

3. Objectifs du projet

Dans le but d'étudier l'impact de l'activation de EPAC1 par l'AMPc, des expériences de HDX-MS ont été effectuées en absence puis en présence de liposomes pour évaluer l'impact de l'environnement lipidique mimant la membrane sur EPAC1/AMPc. Une seconde étude a été menée pour déterminer l'effet du CE3F4-R sur EPAC1 en présence/absence d'AMPc (Fig. 7).

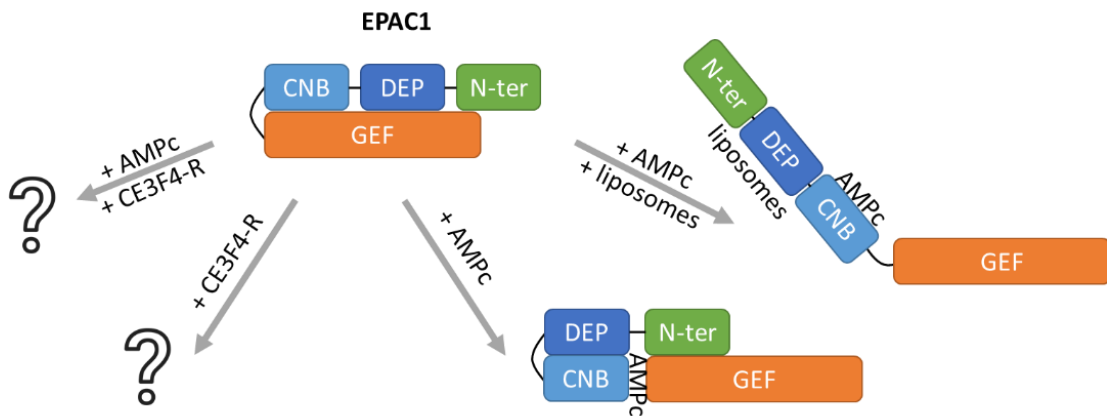


Figure 7 : Représentation schématique des comparaisons menées en HDX-MS et des différentes conformations attendues.

4. Etude de l'activation de EPAC1

4.1. Activation de EPAC1 en présence d'AMPc

L'activation de EPAC1 (tronquée) par l'AMPc a déjà été décrite et entraîne un changement conformationnel important de la protéine (Fig. 8). Des expériences de HDX-MS sur la protéine entière en présence d'AMPc ont donc été menées dans le but d'observer l'impact de la liaison de l'AMPc sur EPAC1 (entière, Annexe 2).

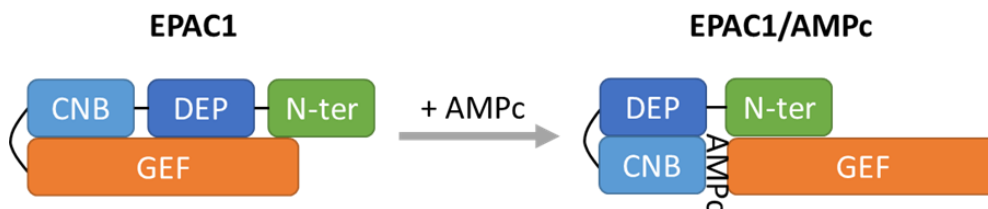


Figure 8 : Représentation schématique des différents changements de conformations attendus suite à l'ajout de l'AMPc sur la structure d'EPAC1.

La différence d'incorporation entre EPAC1 et EPAC1/AMPC est globalement positive, ce qui traduit d'une forte protection (> 5%) de tous les domaines de EPAC1 suite à la liaison de l'AMPC (Fig. 9, Annexe 33A), à savoir :

- Le domaine N-terminal représenté par un peptide (18-38)
- Le domaine DEP (résidus 99-115 ; 118-132 ; 140-156)
- Le domaine CNB et plus particulièrement le site connu de fixation de l'AMPC présentant une différence d'incorporation en D jusqu'à 25% pour le peptide (272-282) ce qui confirme l'interaction.
- La quasi-totalité de la région catalytique avec le domaine GEF (avec les résidus 323-347 ; 358-379 ; 448-454 ; 510-538 ; 547-584 ; 618-644 ; 675-680 ; 717-730 ; 780-794).

Toutes ces régions protégées confirment un réarrangement global de la structure d'EPAC1 en présence d'AMPC.

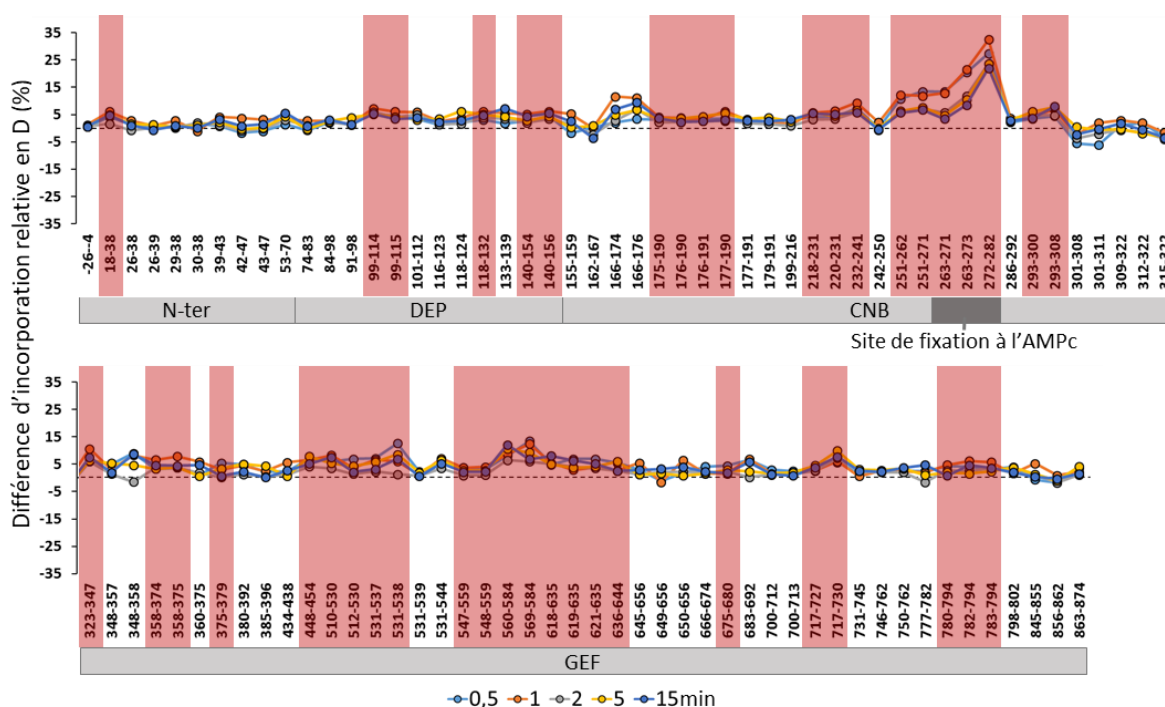


Figure 9 : Représentation des différences d'incorporations relatives en D entre EPAC1 seule et en complexe avec de l'AMPC pour chaque peptide identifié de EPAC1 aux différents temps de deutération. Les peptides encadrés présentent des différences statistiquement significatives d'incorporation de deutérium (MEMHDX, p-value à 0,01¹³⁴) supérieures à 5% (rouge). Les différents domaines de EPAC1 sont indiqués en dessous de chaque peptide ainsi que le site de fixation à l'AMPC.

4.2. Activation de EPAC1 en présence d'AMPC + liposome

L'activation complète de EPAC1 avec l'AMPC et les liposomes mimant la membrane entraînerait un changement conformationnel important³¹⁰ (Fig. 10) pour permettre l'échange de GEFs pour les Rap GTPases³⁰⁸.

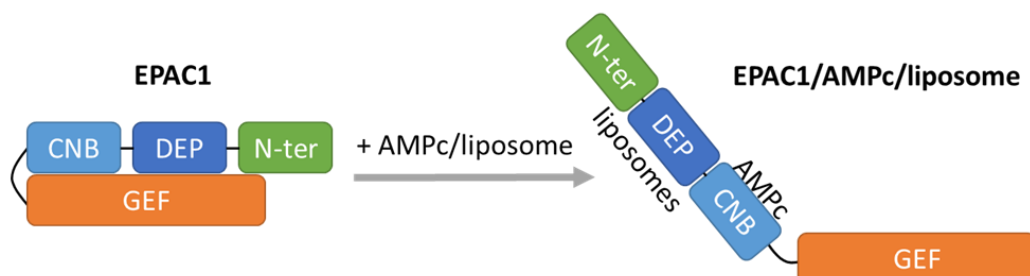


Figure 10 : Représentation schématique du changement de conformation attendu suite à l'ajout de AMPC/liposome sur la structure d'EPAC1.

Le profil HDX-MS du complexe EPAC1/AMPC/liposomes a été comparé à celui de EPAC1 seule (Fig. 11, Annexe 33B). Cette comparaison montre des différences d'incorporation en D statistiquement significatives (supérieures à 5%, MEMHDX¹³⁴) de EPAC1, telles que :

- Le domaine N-terminal est globalement protégé en présence d'AMPC et de liposomes, ce qui est montré par les peptides comprenant les résidus 26 à 47. Cependant, la boucle liant les domaines DEP de EPAC1 (peptide 53-70) présente une augmentation de l'incorporation en D. La déprotection de cette boucle traduit d'un réarrangement des domaines l'un par rapport à l'autre rendant cette boucle plus accessible au solvant.

- Le domaine DEP présente des diminutions de l'incorporation en D suite à l'interaction avec AMPC/liposome, visibles par les peptides 99-114, 140-152 et 140-156. Ces deux derniers peptides sont communs avec ceux observés pour l'analyse précédente sans liposomes (EPAC1 vs EPAC1/AMPC) et sont voisins au domaine CNB. Au contraire, la protection du peptide 99-114 est spécifique à cette comparaison (EPAC1 vs EPAC1/AMPC/liposomes) et témoigne de l'interaction de EPAC1 avec les liposomes.

- Les différences significatives au niveau du site connu de fixation de l'AMPC et de quelques résidus en amont de ce site révèlent une protection de cette région (résidus 232-241 ; 251-282). Elles prouvent expérimentalement la fixation de l'AMPC avec EPAC1 en présence de liposomes. Cependant, le domaine CNB présente aussi deux régions plus accessibles au solvant (166-176 et 301-311).

- Le domaine GEF (région catalytique de EPAC1) dévoile uniquement deux régions protégées suite à la liaison de l'AMPC et des liposomes. Ces régions sont situées en N-ter et C-ter du domaine et représentées par les peptides 323-347 et 845-862. Le domaine GEF incorpore globalement plus de D en présence de AMPC/liposome en accord avec une ouverture de la protéine. Toutefois, seulement deux peptides (448-454 et 777-782) de cette zone montrent une augmentation de l'incorporation en D significative en présence de liposomes et d'AMPC. Il est à noter qu'une importante différence d'incorporation en D mais non-significative entre les réplicats a été constatée au niveau de ce domaine, pouvant tout de même signifier que la région est dynamique.

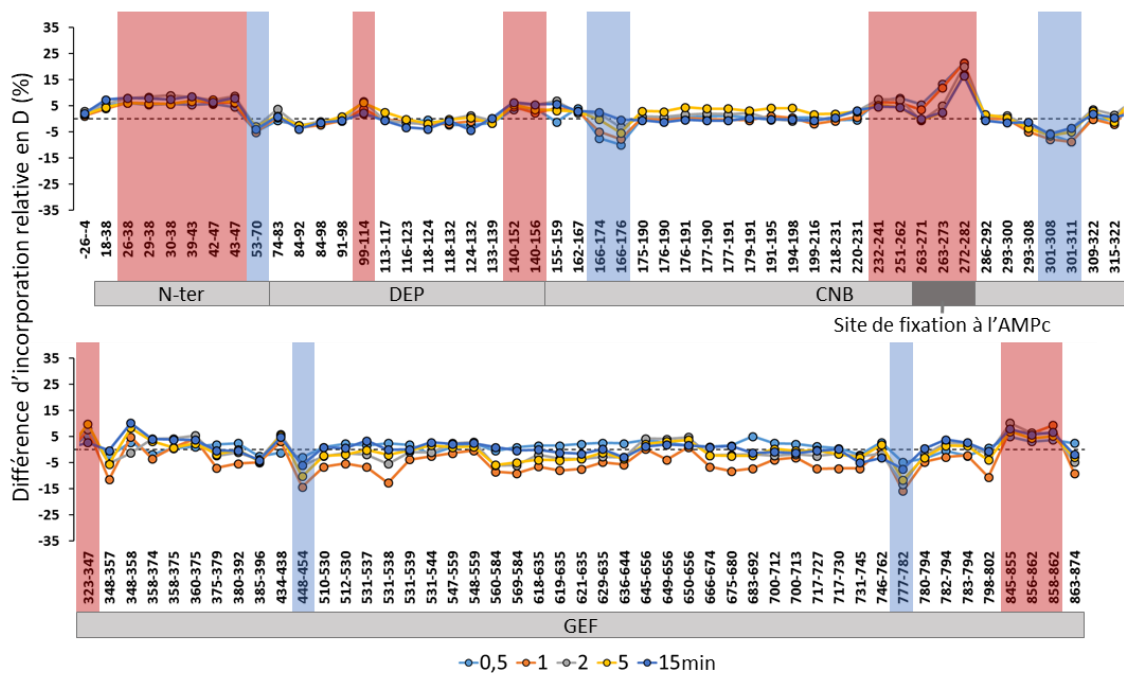


Figure 11 : Représentation des différences d'incorporations relatives en D entre EPAC1 seule et en complexe avec AMPC/liposomes, pour chaque peptide identifié de EPAC1 aux différents temps de deutération. Les peptides encadrés présentent des différences statistiquement significatives d'incorporation de deutérium (MEMHDX, p-value à 0,01¹³⁴) supérieures à 5% (rouge) relatant d'une protection de cette région et inférieures à -5% (bleu) soit une déprotection. Les différents domaines de EPAC1 sont indiqués en dessous de chaque peptide ainsi que le site de fixation à l'AMPC.

Un modèle de la structure de EPAC1 en conformation active en présence d'AMPC et d'une membrane lipidique a été obtenu par nos collaborateurs avec le serveur OMP (*orientations of proteins in membrane*)³¹⁴. Ce modèle a été réalisé à partir de la structure cristalline de EPAC1^{DEP-CNB-AMPC} (également obtenue par nos collaborateurs) et de la structure homologue de EPAC2^{CNB-GEF-AMPC-Rap1} (PDB : 3CF6). Ce modèle de EPAC1 présente une ouverture de la région régulatrice, qui est en accord avec les résultats expérimentaux de HDX-MS (Fig. 12).

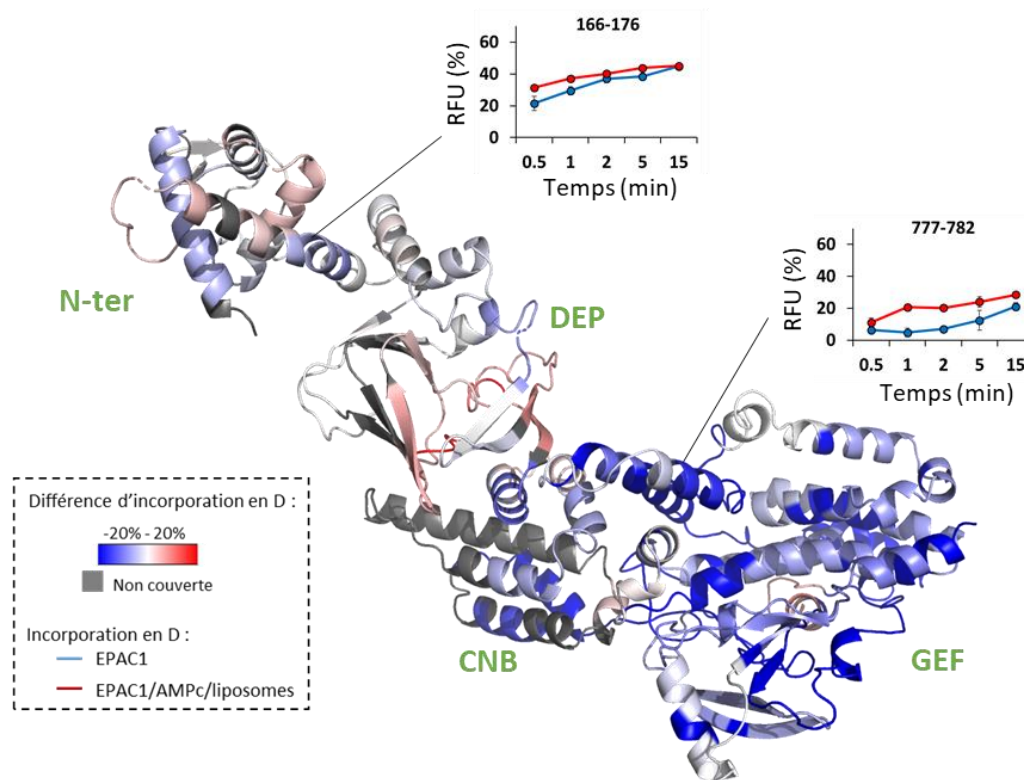


Figure 12 : Représentation des différences d'incorporation en D (temps de deutération : 1min) entre EPAC1 et EPAC1/AMPC/liposomes sur le modèle d'EPAC1. Les régions présentant des différences d'incorporation en D sont colorées suivant l'échelle donnée (-20% ; 20%), celles pour lesquelles aucune information n'est disponible sont en gris. L'incorporation fractionnelle relative en D (RFU) de EPAC1 (bleu) et du complexe (rouge) des deux peptides (166-176 ; 777-782) est représentée en fonction du temps d'incubation.

Effectivement, les déprotections de l'hélice coudée entre les domaines DEP et CNB (166-176) et celle plus globale du domaine GEF montré par le peptide 777-782 (Fig. 12) sont en accord avec le modèle proposé.

4.3. Impact de l'inhibiteur CE3F4-R sur la conformation de EPAC1

Suite à l'obtention de nouvelles informations sur l'impact structural de l'activation d'EPAC1, des expériences de HDX-MS ont été réalisées en présence de l'inhibiteur CE3F4-R afin d'étudier l'impact de l'inhibition de EPAC1 sur sa conformation.

Dans un premier temps, la différence d'incorporation en D entre EPAC1 et EPAC1/CE3F4-R a été étudiée. Cependant, la réalisation de tests statistiques avec le logiciel MEMHDX¹³⁴ sur ces expériences de HDX-MS n'a révélé aucune différence statistiquement significative (Annexe 34). Deux hypothèses peuvent être énoncées suite à l'absence de différence d'incorporation en D entre EPAC1 seule et EPAC1/CE3F4-R : (i) le CE3F4-R ne s'est pas lié à EPAC1 ; ou (ii)

l'inhibiteur n'induit pas de changement de conformation majeur de EPAC1 ; ou (iii) EPAC1 n'est pas activée ; ou encore (iv) l'interaction est perdue au moment de la dilution car de faible affinité

Afin de s'assurer de l'activation de EPAC1, une nouvelle série d'expériences de HDX-MS ont été effectuées. Deux comparaisons ont été réalisées : (i) EPAC1 seule versus EPAC1/AMPC/CE3F4-R et (ii) EPAC1/AMPC versus EPAC1/AMPC/CE3F4-R sachant que la présence d'AMPC active EPAC1 (Fig. 13, Annexe 35)³¹³.

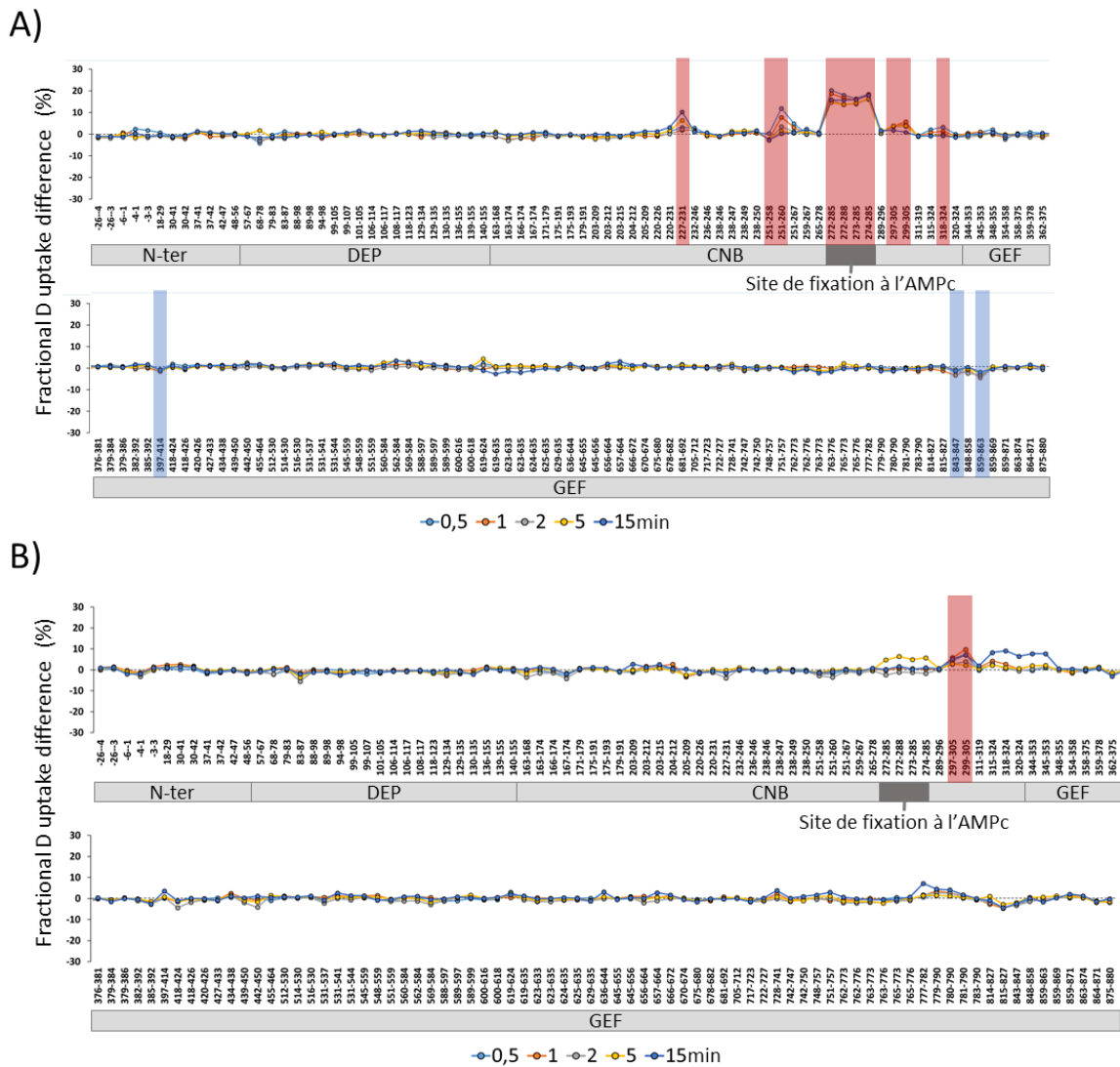


Figure 13 : Représentation des différences d'incorporations relatives en D entre (A) EPAC1 et EPAC1/AMPC/CE3F4-R ainsi qu'entre (B) EPAC1/AMPC et EPAC1/AMPC/CE3F4-R pour chaque peptide identifié de EPAC1 aux différents temps de deutération. Les peptides encadrés présentent des différences statistiquement significatives d'incorporation de deutérium (MEMHDX, p-value à 0,01¹³⁴) supérieures à 5% (rouge) relatant d'une protection de cette région et inférieures à -5% (bleu), soit une déprotection. Les différents domaines de EPAC1 sont indiqués en dessous de chaque peptide et ainsi que le site de fixation à l'AMPC.

La comparaison (i) révèle plusieurs régions du domaine CNB présentant une diminution de l'incorporation en D (protection) suite à la liaison de l'AMPc et du CE3F4-R comparé à EPAC1 seule (Fig. 13A). Certains des peptides protégés (232-246 ; 251-267 ; 272-285) sont communs aux expériences menées précédemment avec l'AMPc (EPAC1 vs EPAC1/AMPc) dont ceux du site de fixation connue de celui-ci. De manière intéressante, de nouvelles régions (résidus 297-305 ; 318-324) sont aussi identifiées au niveau de la partie C-terminale du domaine CNB. Cette région charnière est ici protégée suite à la fixation du CE3F4-R en présence d'AMPc suggérant que cette zone correspond au site de liaison de l'inhibiteur. Trois peptides (397-414 ; 843-847 ; 859-863) du domaine GEF présentent des déprotections suite à l'interaction du CE3F4-R. Ce site de fixation possible du CE3F4-R est confirmé par les deux peptides protégés (297-305 ; 299-305) de la comparaison (ii) entre EPAC1/AMPc et EPAC1/CE3F4-R/AMPc (Fig. 13B). Cette comparaison permet d'observer l'impact de la liaison du CE3F4-R uniquement sur la conformation de EPAC1 activée par l'AMPc.

En conclusion, ces résultats confirment que le CE3F4 et l'AMPc n'entrent pas en compétition et se fixent au niveau de deux sites différents (allostérie). De plus, ils confirment que l'inhibiteur CE3F4-R interagit préférentiellement avec EPAC1 lorsque la protéine est sous sa conformation activée (AMPc) (Fig. 14).

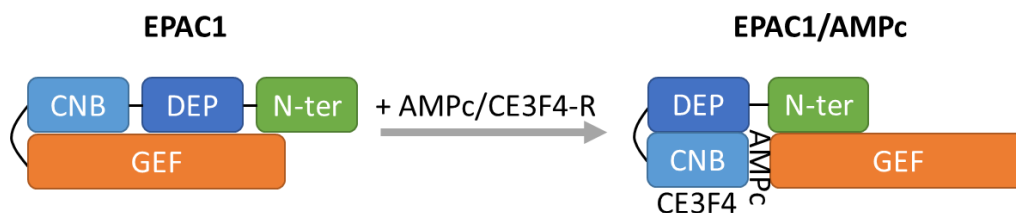


Figure 14 : Représentation schématique de changement de conformation attendu suite à l'ajout de AMPc/CE3F4-R sur la structure d'EPAC1.

5. Valorisation

L'étude en HDX-MS de l'impact de l'AMPc et des liposomes sur EPAC1 a été valorisée dans un article accepté dans le journal *Nature communications* :

Membranes prime the RapGEF EPAC1 to transduce cAMP signaling. Sartre, C.; Ferrandez, Y.; Peurois, F.; Ley, M.; Zhang, W.; Kryszke, M. H.; Courilleau, D.; Fischmeister, R.; Ambroise, Y.; Zeghouf, M.; Cianferani, S.; Peyroche, G.; Cherfils, J. DOI : 10.21203/rs.3.rs-2608569/v1

6. Conclusions et perspectives

Les expériences de HDX-MS menées sur EPAC1 et ces différents complexes ont permis de montrer que :

- Une couverture de séquence satisfaisante (>70%) peut être obtenue pour l'analyse de EPAC1 suite à l'optimisation de la phase pré-analytique (utilisation de différentes protéases).
- Les différentes comparaisons d'incorporation en D menées sur EPAC1 ont permis de montrer (i) l'impact de l'activation de EPAC1 par l'AMPc et les liposomes sur sa conformation avec l'ouverture de celle-ci ; ainsi que (ii) l'impact de la fixation de l'inhibiteur CE3F4-R sur la structure de EPAC1 qui nécessite d'être activée par l'AMPc au préalable.

L'utilisation de liposomes, dans le cadre des analyses de EPAC1, a permis de conserver un état natif de la protéine EPAC1. Ils pourraient être employés dans l'étude de l'impact du CE3F4-R sur la conformation de EPAC1 complètement activée.

IV. Conclusions de l'étude des PMs en HDX-MS

Les différentes études menées sur les PMs en HDX-MS démontrent encore une fois la pertinence de cette technique dans ce contexte. Certaines optimisations sont toutefois nécessaires notamment au niveau de la phase pré-analytique avec l'utilisation de différentes enzymes pour la digestion, de la phase analytique avec l'utilisation de la mobilité ionique (pour les comparaisons EPAC1 vs EPAC1/CE3FA et EPAC1 vs EPAC1/CE3FA/AMPc) et enfin de la phase post-analytique avec l'ajustement d'un filtre appliqué lors de l'identification des peptides (« produit par acide aminé »). Finalement, ces optimisations ont permis d'identifier de nombreux peptides permettant ainsi d'apporter des informations de changements conformationnels et de dynamique importants sur une grande partie de la séquence (couverture de séquence > 70%) et d'obtenir des informations sur les peptides des DTMs bien qu'hydrophobes. Il est aussi notable que les effets des liposomes sur la conformation de EPAC1 ont pu être mis en évidence grâce à cette technique.

En perspective, l'utilisation d'une nouvelle cartouche combinant deux enzymes de digestion co-immobilisées (pepsine et nepenthesine II) développée et commercialisée par Affipro pourrait permettre d'augmenter encore le nombre de peptides identifiés et le pourcentage de couverture de séquence des PMs¹²⁵.

Chapitre 3 : Caractérisation de protéines arginines méthyl-transférases (PRMTs) par MS structurale

I. Contexte analytique et biologique

La méthylation des arginines est une PTM courante et de plus en plus étudiée, car essentielle pour la régulation de la transcription des gènes suite à la méthylation des histones³¹⁵⁻³¹⁷. Cette PTM est aussi impliquée dans différents processus cellulaires tels que la réparation de l'ADN, la signalisation ou encore le métabolisme de l'ARN³¹⁸. Elles affectent les protéines par la modification de leurs propriétés physico-chimiques comme l'augmentation de leur hydrophobicité suite à la diminution des sites potentiels de liaison hydrogène³¹⁹. Elle s'effectue à l'aide d'un groupe donneur, la S-Adénosylméthionine (SAM) et est catalysée par les protéines arginine méthyl-transférases (PRMTs)^{316, 320}. Plusieurs PRMTs existent et sont classées chez les mammifères en trois types différents en fonction de leurs produits, un quatrième type existe uniquement chez les levures (*Saccharomyces cerevisiae*)³²¹. PRMT7 est l'unique enzyme produisant une mono-méthylation des arginines (MMA, type III) alors que PRMT1-4, 6 et 8 vont engendrer des di-méthylations asymétriques (ADMA, type I) et enfin PRMT5 et 9 catalysent une di-méthylation symétrique (SDMA, type II, Fig. 1A)³²²⁻³²⁴.

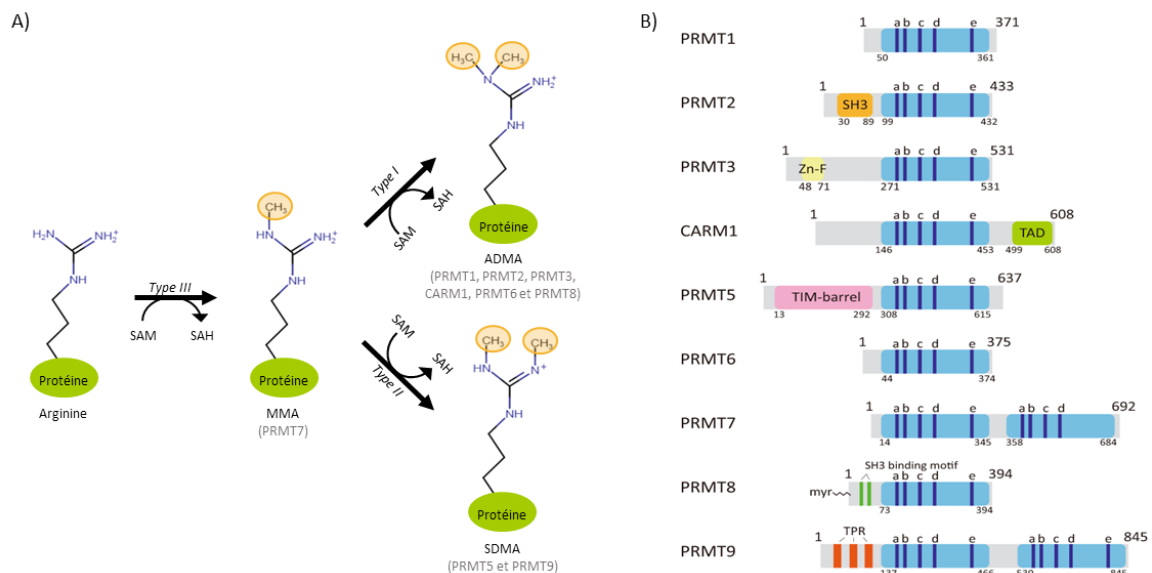


Figure 1 : (A) Représentations des différents produits de la méthylation des arginines suivant les types de PRMTs. (B) Description des PRMTs avec leurs domaines (repliement de Rossmann et tonneau β en bleu clair) et motifs uniques (a, Motif I: VLD/EVGxGxG; b, Post-I: V/IxG/AxD/E; c, Motif II: F/I/VDI/L/K; d, Motif III: LR/KxxG; e, boucle THW)(adapté de Hwang et al, 2021³²⁵).

Ces différentes protéines sont des homologues et présentent des similarités de séquence d'environ 37%³¹⁹, avec en commun différents domaines comme le repliement de Rossmann (RF), un module catalytique interagissant avec les groupements donneurs, le tonneau β ainsi que des motifs uniques (Fig. 1B)³²⁶. Des motifs sont également conservés³²⁶ et composent le noyau catalytique des PRMTs, avec :

- Le motif I (VLD/EVGxGxG) et le post-motif I (V/lxG/AxD/E) formant la poche de liaison du donneur de méthyle (SAM) avec formation de liaison H.
- Le motif II (F/I/VDI/L/K) et le motif III forment des brins β et stabilisent le motif I.
- Le motif double E (SExMGxxLxxExM) positionne l'arginine du substrat.
- La boucle THW importante pour la liaison du substrat.

De nombreuses maladies cardiovasculaires, pulmonaires et certains cancers sont connus comme étant directement liés à la dérégulation des PRMTs³¹⁸.

Parmi ces PRMTs, PRMT2 (50 kDa) reste peu caractérisée, car aucun substrat de PRMT2 n'était connu ne permettant pas de détecter son activité enzymatique jusqu'à l'identification de RSF1 (le facteur d'épissage répresseur 1) en 2017 par l'équipe du Pr. J. Cavarelli. RSF1 (18 kDa) qui copurifie avec PRMT2 dans les cellules de l'insecte *Spodoptera frugiperda*¹⁷¹. RSF1 est un répresseur de l'assemblage des spliceosomes et de l'épissage des pré-ARNm³²⁷ spécifique aux insectes et fait partie de la famille des régulateurs d'épissages riches en sérine et arginine chez la drosophile³²⁸⁻³³¹. Ce facteur est composé d'un domaine RRM (motif de reconnaissance de l'ARN) comprenant deux motifs SxGGxRxY et une partie C-terminale riche en arginine, glycine et sérine nécessaires pour une activité complète³³². Quatre mono-méthylations (sur R24, R100, R111 et R139) et deux di-méthylations (sur R7 et R120 sur les motifs SxGGxRxY) ont été identifiées en présence de PRMT2 et du donneur SAM¹⁷¹. PRMT2 se distingue d'autres enzymes par sa région N-terminale composée de 45 résidus spécifiques et son domaine d'homologie 3 de SRC (SH3) de 50 acides aminés³¹⁹. Il est également connu pour former un homodimère asymétrique dont la structure a été résolue (Fig. 2), sans toutefois permettre la visualisation du domaine SH3 car très flexible (PDB : 5FUL)¹⁷¹.

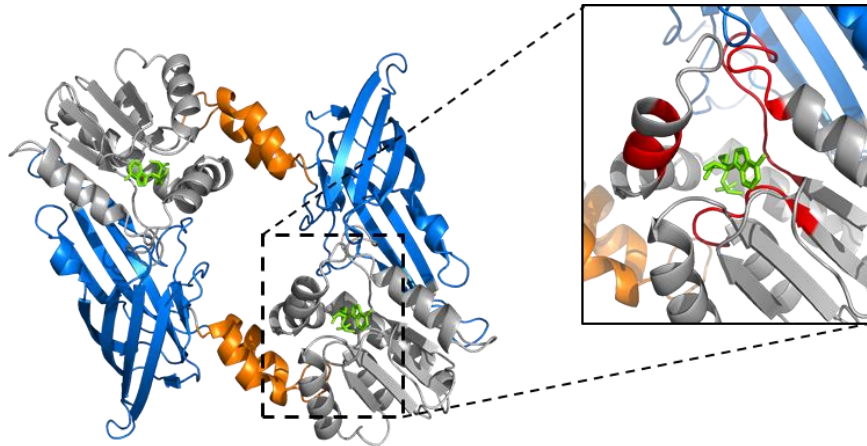


Figure 2 : Structure du dimère de PRMT2 (PDB : 5FUL) avec le domaine RF (gris), le tonneau β (bleu), les hélices de dimérisation (orange) et les trois motifs YFxxY, DVGxGxG et SExMGxxLxxExM (rouge) nécessaires pour la liaison des donneurs SAM (vert).

PRMT9 (95 kDa) a déjà été identifiée en interaction avec deux partenaires SAP49 (44k Da) et SAP145 (100 kDa) par la méthylation de R508 de SAP145^{319, 333-334}, mais le complexe n'a pas été caractérisé structuralement. SAP145 et SAP49 sont des protéines associées au spliceosome et des composés de U2 snoNRP (petite ribonucléoprotéine nucléaire) impliquée dans les premières étapes de l'épissage³³⁴. SAP49 contient deux domaines RRM (1 et 2) et une partie C-terminale riche en proline et glycine³³⁵ et SAP145 comprend un domaine CUS1 (382-563) qui interagit avec le RRM1 de SAP49³³⁶⁻³³⁷. PRMT9 présente différentes caractéristiques avec deux domaines catalytiques (domaine de liaison des SAM) comme PRMT7 (type III) rendant sa classification de type II discutable³³⁸ et son domaine N-terminal (140 acides aminés) comprenant trois répétitions tétratricopeptidiques (TPR, *Tetratrico Peptide Repeat*) spécifiques à PRMT9 et participe à la reconnaissance d'autres protéines^{325, 339}.

Ce projet a été réalisé dans le cadre d'un projet ANR (InSiPRMTs) en collaboration avec deux laboratoires partenaires (Pr. J. Cavarelli et son équipe, département de Biologie Structurale Intégrative, IGBMC, UMR 7104, Illkirch ; Dr. A. Wagner, BFC, UMR 7199, Strasbourg). Cette ANR a pour but de comprendre comment les PRMTs reconnaissent leurs substrats et de fournir des informations structurales sur les complexes PRMT/substrats. Pour cela, différentes techniques ont été employées comme la diffusion dynamique de la lumière (DLS), la SAXS, la SEC couplée à la diffusion de lumière multi-angle (SEC-MALS), la cryo-EM, la X-ray et des approches de MS structurale, ces dernières faisant l'objet de mon travail.

II. Objectifs

Ce chapitre décrit les résultats obtenus grâce aux différentes approches de MS structurale sur les complexes PRMT2/RSF1 et PRMT9/SAP145/SAP49. Dans un premier temps, PRMT2/RSF1 a été analysé en nMS, HDX-MS et XL-MS dans le but de caractériser ce complexe en obtenant des informations sur la stœchiométrie, la dynamique conformationnelle et les proximités spatiales. Dans un second temps, le complexe ternaire PRMT9/SAP145/SAP49 a été analysé en nMS tandis que le sous-complexe PRMT9/SAP145₄₀₁₋₅₅₀ a fait l'objet d'analyses en nMS, HDX-MS et XL-MS.

III. Matériel et méthodes

La nMS a été effectuée par infusion directe sur le synapt-G2 (Waters), les voltages de cônes et pression dans l'interface sont données dans les légendes des spectres montrés. Les données ont été exploitées avec le logiciel MassLynx 4.1. (Waters). En amont de l'analyse, les échantillons ont été dessalés deux fois sur Zeba (7 kDa, Thermo Fisher Scientific) à 1 500xg avec 500mM d'acétate d'ammonium pour les analyses concernant PRMT2 et RSF1 et 200mM pour PRMT9 à un pH de 6,9.

Les expériences d'échange H/D ont été effectuées en triplicat en présence d'un tampon à base de D₂O (20mM HEPES, 100mM NaCl et 0,5mM DTT, pH 7,0) à différents temps de deutération (0,5, 2, 10, 30 et 60 minutes) pour PRMT2. PRMT9 a été mis en présence d'un tampon composé de 20mM de Tris-HCl, 100mM de NaCl et 2mM de TCEP à un pH de 8,0 pendant 0,5, 1, 10, 30 et 60 min. L'arrêt du marquage a été réalisé en présence de 2M de Guanidine-HCl et 100mM de Glycine à un pH final de 2,4. Les protéines (20pmoles) ont été digérées avec de la pepsine avec un débit de 100 (PRMT2) et 200µL/min (PRMT9). Après digestion, les peptides ont été séparés sur une colonne C18 et analysés sur un Synapt-G2 Si (Waters). Les peptides ont été identifiés avec le logiciel PLGS 2.5.3 (Waters) et filtré avec DynamX 3.0 (Waters). Les résultats ont été validés statistiquement avec une p-value de 0,01 avec le logiciel MEMHDX (institut Pasteur).

Le XL-MS a été réalisé avec du DSBU à un excès molaire de 100 dans un tampon composé de 20mM de HEPES, 100mM de NaCl et 0,5mM de TCEP à un pH de 7,5 à 20°C pendant 45 min pour le complexe PRMT2/RSF1. Concernant PRMT9, il a été effectué avec du DSAU et DSBU séparément à un excès molaire de 200 dans 20mM de HEPES, 100mM de NaCl et 2mM de

TCEP à 20°C pendant 45 min. Tous les pontages ont été réalisés en triplicat (1h à 20°C) puis réduits, alkylés et digérés par un mix de Trypsine/Lys-C. Les excès de réactifs sont éliminés par passage sur SPE C18, séparés sur une colonne C18 et analysés sur un Q-Exacte HF-X Orbitrap (Thermo Fisher Scientific). MeroX 2.0.1.4 a été utilisé pour l'identification des peptides pontés avec un FDR fixé à 5%. La présence des fragments caractéristiques attendus pour ce type d'agent pontant a été vérifiée et seuls les cross-links validés dans au moins 2 des 3 réplicats ont été conservés. Les intervalles de distance sont fixés de 6 à 30Å pour le DSBU et 6 à 26Å pour le DSAU.

IV. Résultats

1. Caractérisation du complexe PRMT2/RSF1

1.1. nMS

Les expériences de nMS des protéines seules ont montré la présence de RSF1 majoritairement sous forme monomérique ($17\,989 \pm 1$ Da, Fig. 3A) et PRMT2 présent en équilibre entre monomères ($50\,721 \pm 1$ Da, Fig. 3B) et dimères ($101\,450 \pm 2$ Da) pour PRMT2 (Annexe 2).

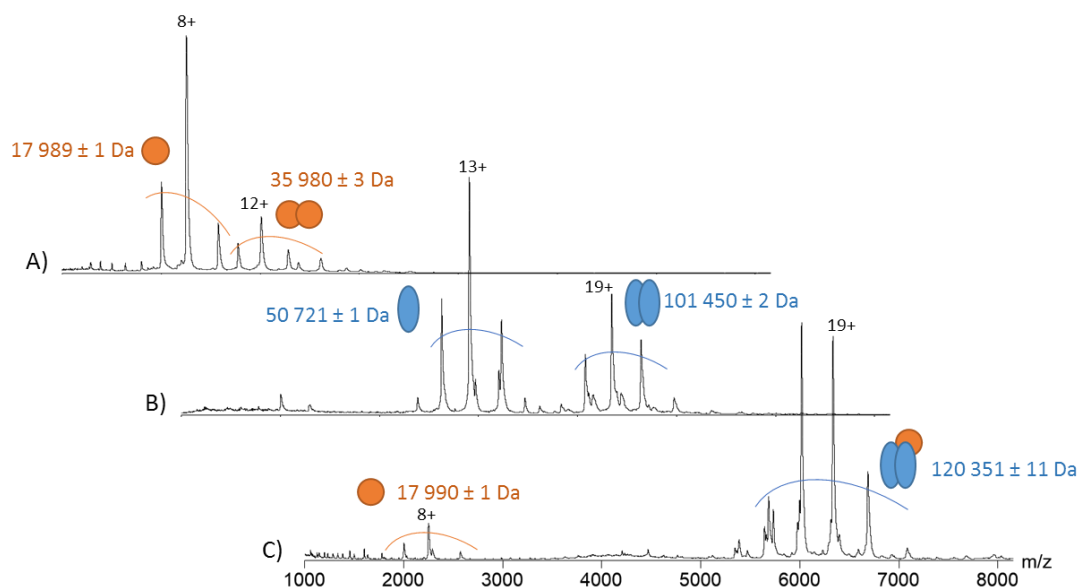


Figure 3 : Spectres de masse obtenus par nMS de RSF1 (A), PRMT2 (B) et PRMT2/RSF1 (C) avec un voltage de cône (Vc) fixé à 120 V et une pression dans l'interface (Pi) de 6 mbar.

L'analyse de PRMT2/RSF1 a permis d'identifier la formation d'un complexe entre un dimère de PRMT2 et un RSF1 ($120\,351 \pm 11$ Da, Fig. 3C) avec deux groupes donneurs SAM (+398 Da chacun).

Ces résultats confirment que :

- Le complexe PRMT2/RSF1 est asymétrique avec la fixation d'un RSF1 pour un dimère de PRMT2. Aucune autre stœchiométrie n'a été observée pour ce complexe.
- La présence des donneurs de méthyle SAM avec le complexe ce qui est en accord avec le mécanisme enzymatique.
- RSF1 est bien un substrat de PRMT2 par la stabilisation du dimère de PRMT2.

1.2. HDX-MS

1.2.1. Optimisation de la phase pré-analytique

Pour mieux comprendre la dynamique de l'interaction PRMT2/RSF1 des expériences de HDX-MS ont été réalisées sur PRMT2, RSF1 et le complexe. Dans les conditions optimales, la couverture de séquence (Annexe 36) obtenu pour chacune des protéines (Tableau 1) est très satisfaisante (> 93%) de même que la redondance ($\geq 2,9$).

| Analyses | PRMT2 | RSF1 | Complexe | |
|-------------------------------|-------|------|----------|------|
| | | | PRMT2 | RSF1 |
| Couverture de séquence (%) | 93,3 | 97 | 95,5 | 97 |
| Redondance | 3,2 | 3 | 3,7 | 2,9 |
| Nombre de peptides identifiés | 127 | 36 | 144 | 37 |

Tableau 1 : Couverture de séquence, redondance et nombre de peptides identifiés pour les analyses de PRMT2, RSF1 et du complexe PRMT2/RSF1.

1.2.2. Impact de la liaison de RSF1 sur PRMT2

Les échanges H/D de PRMT2 et du complexe PRMT2/RSF1 ont été effectués à différents temps d'incubation (0,5, 2, 10, 30 et 60 min) (Annexe 37). Les incorporations en D ont été comparées entre les deux états de la protéine (PRMT2 vs PRMT2/RSF1) et ont révélé des différences d'incorporation en D importantes (jusqu'à 35%), nombreuses (70% des résidus) et statistiquement significatives (MEMHDX avec une p-value à 0,01¹³⁴, Fig. 4).

Généralement, une plus faible incorporation en D suite à la liaison d'un partenaire d'interaction par rapport à la protéine seule traduit d'une moins grande accessibilité au solvant de cette protéine en complexe et donc d'une protection suite à l'interaction (zone d'interaction, changement conformationnel, liaison H, etc).

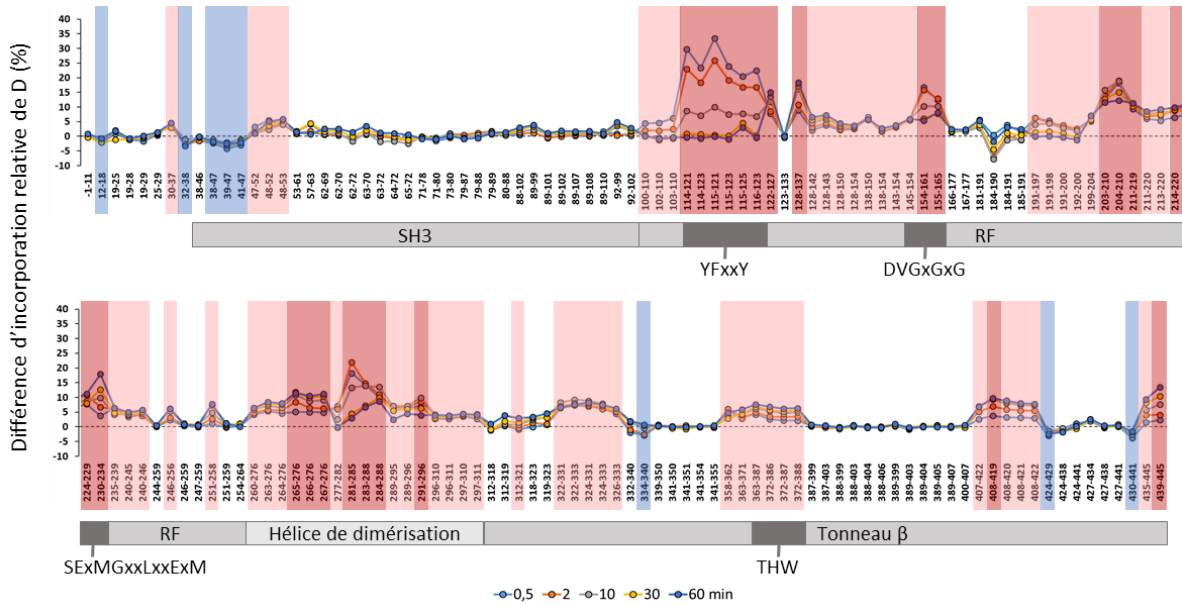


Figure 4 : Représentation des différences d'incorporations relatives en D en PRMT2 et le complexe PRMT2/RSF1 pour chaque peptide identifié de PRMT2 aux différents points de deutérotation. Les peptides encadrés présentent des différences statistiquement significatives d'incorporation de deutérium (MEMHDX, p-value à 0,01¹³⁴) supérieures à 10% (rouge), entre 2% et 10% (rose) montrant la protection du complexe par rapport à la protéine seule et inférieures à -2% (bleu) relatant d'une déprotection. Les différents domaines de PRMT2 sont indiqués en gris.

Dans le cas de la partie N-terminale du domaine SH3 de PRMT2, une déprotection est observée (acides aminés 38-47) puis une protection (acides aminés 47-53) mettant en lumière la flexibilité de cette région³¹⁹.

Le domaine catalytique, caractérisé par le repliement de Rossmann, présente quant à lui une importante protection suite à la liaison de RSF1. Les trois motifs de cette région de PRMT2 (YFxxY spécifique aux PRMTs de type I, DVGxGxG le motif I et SEXMGxxLxxExM le motif double E), aidant à la reconnaissance des donneurs de méthyle, sont complètement protégés avec des différences supérieures à 10% (résidus 114-123, 145-165 et 224-234, Fig. 4) en présence de RSF1, en accord avec l'activité enzymatique de PRMT2.

L'hélice de dimérisation (aussi notée HD, résidus 260-311) est entièrement protégée avec RSF1 ce qui est en accord avec les expériences de nMS par la stabilisation du dimère de PRMT2 en complexe.

Enfin, le tonneau β est principalement protégé au niveau de sa partie N-terminale, de la boucle THW important pour la liaison du substrat, de la région 407-422 et du C-terminale. Toutefois, trois peptides sont déprotégés et exposés au solvant (334-340 ; 424-429 ; 430-441). Ces déprotections sont néanmoins localisées aux extrémités du tonneau et seraient plus accessibles au solvant (Fig. 5).

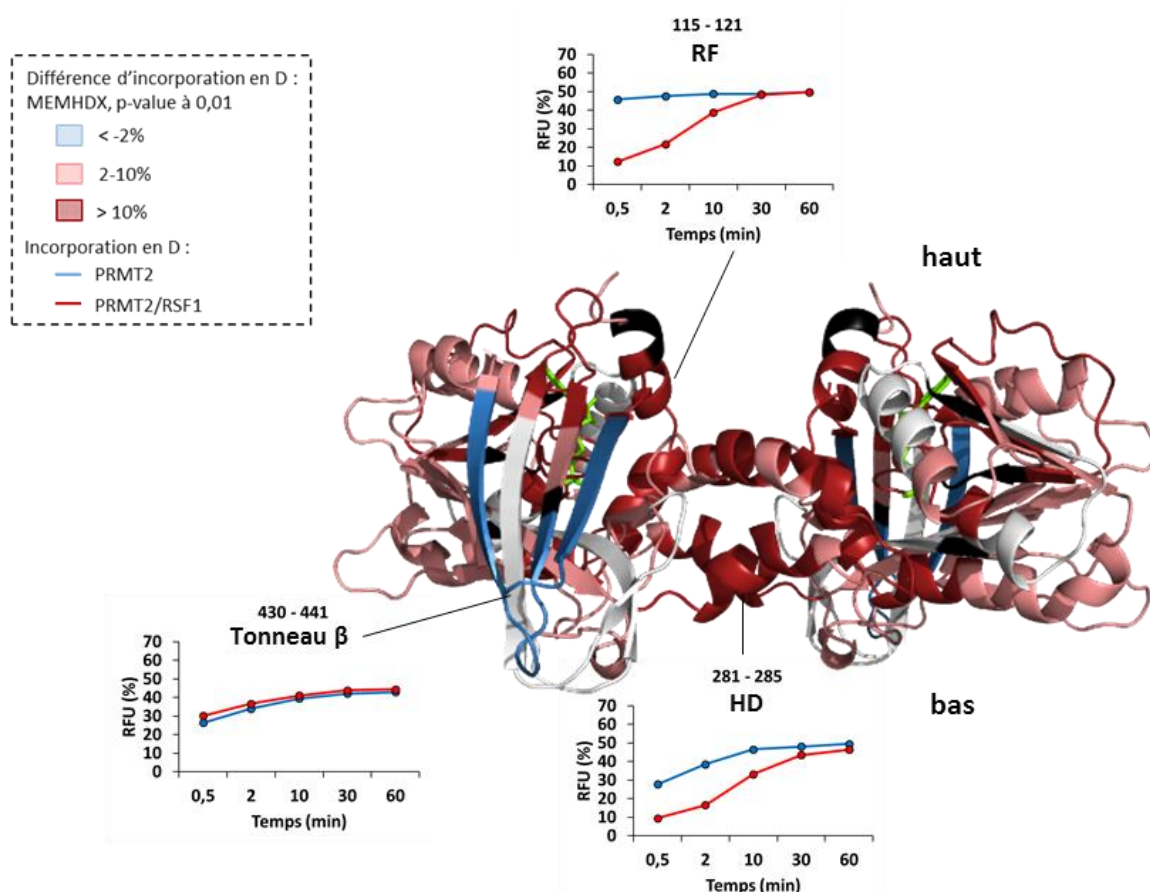


Figure 5 : Représentation des différences d'incorporation statistiquement significatives en deutérium (MEMHDX, p-value à 0,01¹³⁴) entre PRMT2 et le complexe PRMT2/RSF1 sur la structure du dimère de PRMT2 (PDB : 5FUL). Donneur de méthyle (vert) ; régions présentant des différences d'incorporation en D supérieures à 10% (rouge) ; entre 2% et 10% (rose) ; inférieures à -2% (bleu) ; régions non-couvertes (noir). L'incorporation fractionnelle relative en D (RFU) de PRMT2 (bleu) et du complexe PRMT2/RSF1 (rouge) des trois peptides (115-121 ; 281-285 ; 430-441) est représentée en fonction du temps d'incubation.

De manière plus générale, la quasi-totalité de PRMT2 est protégée en présence de RSF1 comme l'export des données sur la structure le montre (Fig. 5) et les différences les plus grandes (supérieures à 10%) sont situées sur le « haut » de la protéine (au niveau de l'interface avec les domaines SH3). L'interaction entre PRMT2 et RSF1 pourrait donc s'effectuer au niveau du « haut » du dimère proche des sites de liaison des donneurs SAM.

1.2.3. Impact de la liaison de PRMT2 sur RSF1

Des différences d'incorporations en D significatives ont été mesurées sur 75% de la séquence de RSF1 montrant un énorme changement conformationnel sur l'ensemble de cette petite protéine de 18 kDa. Les résultats de l'échange H/D révèlent une protection du domaine RRM et des motifs SxGGxRxY (Fig. 6, Annexe 38). De plus, les arginines R24, R100, R111 et R139

précédemment identifiées comme étant mono-méthylées¹⁷¹ sont, elles aussi, protégées (peptides 23-32 ; 26-32 ; 98-113 ; 138-147 ; 139-147), en accord avec une interaction enzyme/substrat entre PRMT2 et RSF1.

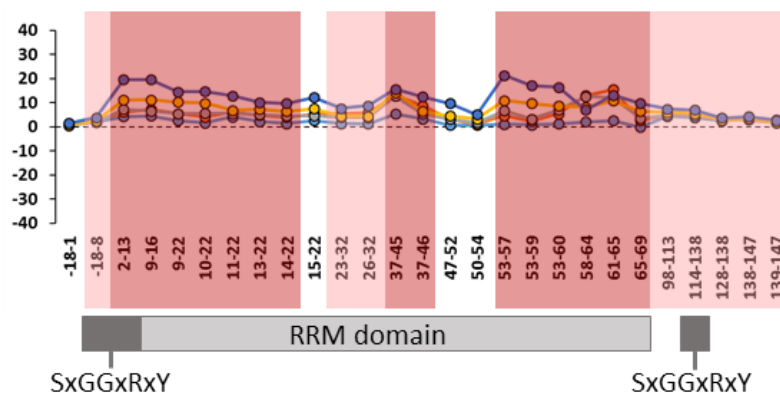


Figure 6 : Représentation des différences d'incorporations relatives en D entre RSF1 et le complexe PRMT2/RSF1 pour chaque peptide identifié de RSF1 aux différents points de deutération. Les peptides encadrés présentent des différences statistiquement significatives d'incorporation de deutérium (MEMHDX, p-value à 0,01¹³⁴) supérieures à 10% (rouge) et entre 2% et 10% (rose).

Finalement, l'approche HDX-MS a mis en évidence des protections globales de PRMT2 et RSF1 lors de la formation du complexe PRMT2/RSF1. Le domaine de liaison des SAM de PRMT2 est protégé en présence de RSF1 ainsi que la région des R méthylées de RSF1. Ces résultats traduisent de l'activité enzymatique de PRMT2 en interaction avec RSF1 et permettent de partiellement localiser les régions d'interaction au niveau du « haut » du dimère.

1.3. XL-MS

Afin de préciser les zones d'interaction, des expériences de XL-MS ont été effectuées sur le complexe. Le pontage chimique de PRMT2/RSF1 a été réalisé avec du DSBU (excès molaire de 1:100). Il a permis d'identifier 11 peptides pontés, dont :

- 4 cross-links intra-PRMT2 engageant le domaine SH3 (K68), RF (K136) et les hélices de dimérisation (Y270, S290, K294) (Tableau 2).
- 1 intra-RSF1 impliquant le domaine RRM (K19) avec le premier motif SxGGxRxY (S2) (Tableau 2).
- 6 inter-PRMT2/RSF1 (Fig. 7) rendant compte de la proximité spatiale du domaine SH3 et de l'hélice de dimérisation de PRMT2 avec le domaine RRM et les deux motifs de RSF1.

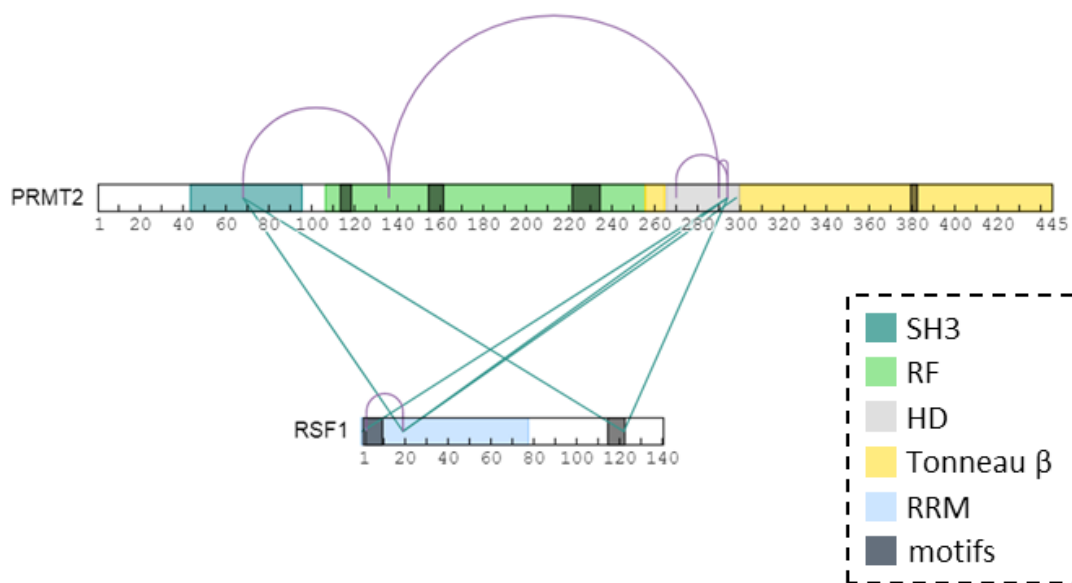


Figure 7 : Représentation schématique des peptides pontés inter (vert) et intra-protéine (violet) du complexe PRMT2/RSF1 (obtenus avec Xinet²¹⁶) avec les domaines SH3 (turquoise), RF (vert), RRM (bleu clair) ainsi que le tonneau β (jaune), l'hélice de dimérisation (gris) et les motifs (noir).

| Protéine 1 | Protéine 2 | Site 1 | Site 2 | Domaine site 1 | Domaine site 2 | Distance C α -C α (Å) | |
|------------|------------|--------|--------|----------------|----------------|-------------------------------------|-------------|
| | | | | | | PDB : 5FUL | modèle RSF1 |
| PRMT2 | PRMT2 | K136 | K68 | RF | SH3 | N | N |
| PRMT2 | PRMT2 | K136 | S290 | RF | HD | 28,4 | N |
| PRMT2 | PRMT2 | K294 | Y270 | HD | HD | 12,4 | N |
| PRMT2 | PRMT2 | K294 | S290 | HD | HD | 6,2 | N |
| RSF1 | RSF1 | S2 | K19 | SxGGxRxY | RRM | N | 20 |

Tableau 2 : Cross-links intra-protéines identifiés pour l'analyse du complexe PRMT2/RSF1.

Les distances C α -C α entre les résidus pontés de PRMT2 ont été mesurées sur la structure du dimère de PRMT2 (PDB : 5FUL, Tableau 2), excepté pour le domaine SH3 non visible sur la structure. Les 3 distances mesurées correspondant aux peptides intra-PRMT2 (K136-S290 ; K294-S290 ; K294-Y270) sont comprises dans l'intervalle de 6 à 30Å admis pour le DSBU (28,4 ; 6,2 ; 12,4 respectivement). Le pontage entre les résidus K136-S290 a été observé entre les deux monomères de PRMT2 permettant de valider l'interface dimérique. Un modèle de RSF1, généré par SWISS-MODEL³⁴⁰ a été obtenu par nos collaborateurs. La distance mesurée sur ce modèle (S2-K19 ; 20Å) est en accord avec la gamme de distance atteignable par l'agent pontant (Tableau 2).

1.4. Prédiction de la structure du complexe

Ne disposant pas d'une structure 3D de haute résolution du complexe, un modèle a été généré via les logiciels Robetta³⁴¹ puis AlphaFold³⁴² par nos collaborateurs. Cette structure représente le dimère asymétrique de PRMT2 (comprenant le domaine SH3) en interaction avec la partie N-terminale de RSF1 (résidus 1-95).

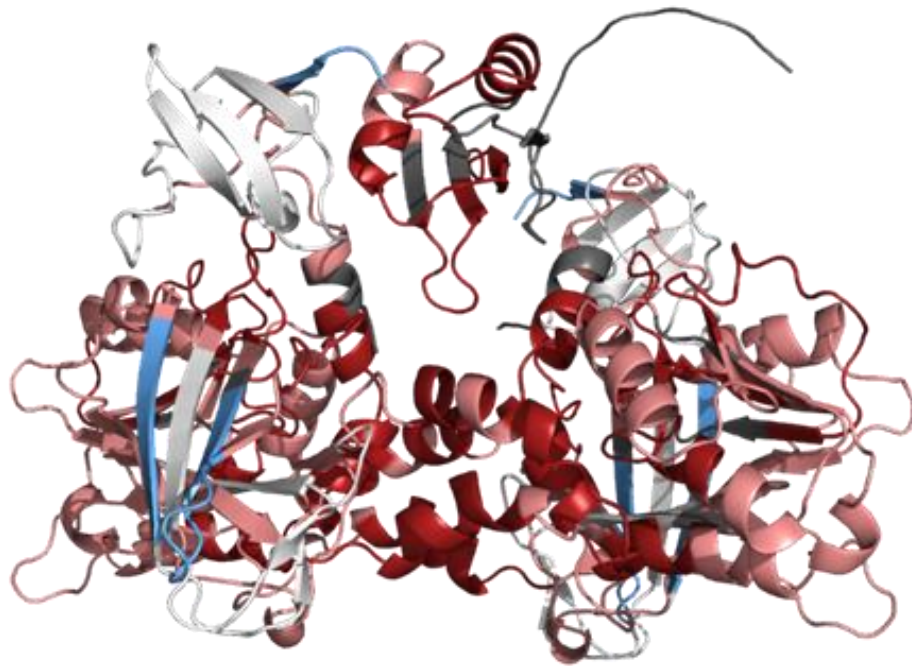
D'après ce modèle de PRMT2/RSF1 (Fig. 8A), l'interaction entre les protéines s'effectue par le « haut » du dimère en accord avec les résultats obtenus en HDX-MS (paragraphe 1.2.2.).

| Protéine 1 | Protéine 2 | Site 1 | Site 2 | Domaine site 1 | Domaine site 2 | Distance Cα-Cα (Å) modèle PRMT2/RSF1 |
|------------|------------|--------|--------|----------------|----------------|--------------------------------------|
| PRMT2 | PRMT2 | K136 | K68 | RF | SH3 | 50,3 |
| PRMT2 | PRMT2 | K136 | S290 | RF | HD | 28,6 |
| PRMT2 | PRMT2 | K294 | Y270 | HD | HD | 12,3 |
| PRMT2 | PRMT2 | K294 | S290 | HD | HD | 6,3 |
| RSF1 | RSF1 | S2 | K19 | SxGGxRxY | RRM | 24 |
| PRMT2 | RSF1 | K68 | K19 | SH3 | RRM | 22,3 |
| PRMT2 | RSF1 | K68 | Y122 | SH3 | SxGGxRxY | - |
| PRMT2 | RSF1 | K294 | S-1 | HD | - | - |
| PRMT2 | RSF1 | K294 | K19 | HD | RRM | 26 |
| PRMT2 | RSF1 | K294 | Y122 | HD | SxGGxRxY | - |
| PRMT2 | RSF1 | S298 | K19 | HD | RRM | 23,5 |

Tableau 3 : Cross-links intra- et inter-protéines identifiés pour l'analyse du complexe PRMT2/RSF1.

Les distances mesurées entre les acides aminés impliqués dans le cross-links intra-PRMT2 et intra-RSF1 sont comprises dans notre gamme de 6 à 30Å (K136-S290, 28,6Å ; K294-S290, 6,3Å ; K294-Y270, 12,4Å ; S2-K19, 24Å) exceptés pour la distance de 50,3Å entre les résidus (K136-K68) (Tableau 3). La position du domaine SH3 de PRMT2 (comprenant la K68) semblerait alors inexacte, ceci pouvant s'expliquer par la flexibilité/dynamique de ces domaines SH3 ou par l'asymétrie du complexe (qui n'est pas prise en compte lors de la modélisation). Finalement, 3 des 6 distances inter-PRMT2/RSF1 sont mesurables (K294-K19, 26Å ; S298-K19, 23,5Å ; K68-K19, 22,3Å) et sont en accord avec les distances atteignables en XL-MS (Figure 8B).

A)



B)

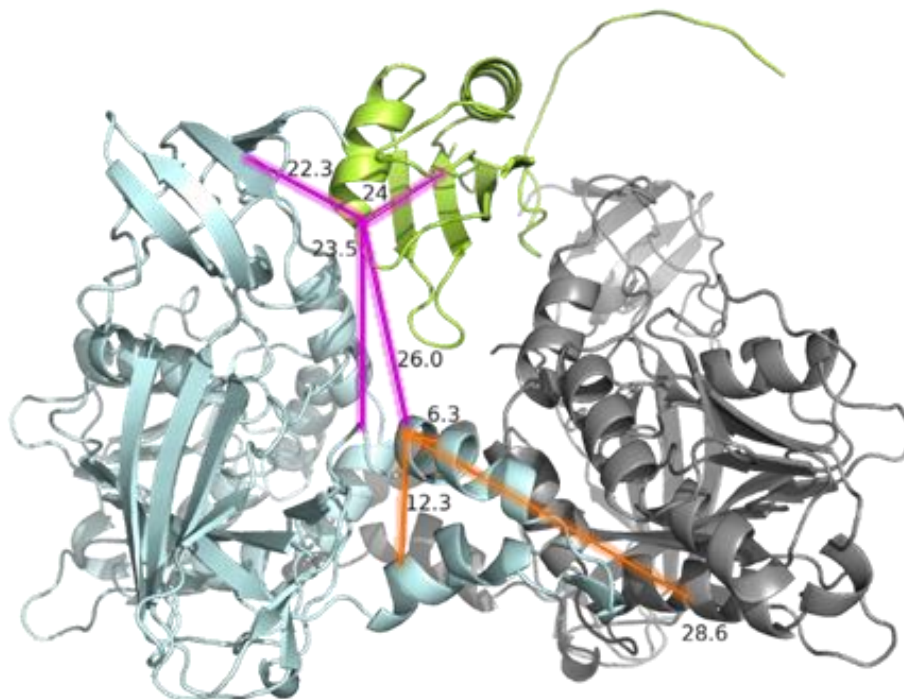


Figure 8 : (A) Représentation des différences d'incorporation statistiquement significatives en deutérium (MEMHDX, p-value à 0,01¹³⁴) entre PRMT2 et le complexe PRMT2/RSF1 sur le modèle du complexe PRMT2/RSF1. Régions présentant des différences d'incorporation en D supérieures à 10% (rouge) ; entre 2% et 10% (rose) ; inférieures à -2% (bleu) ; régions non-couvertes (noir). (B) Représentation des cross-links validés intra (orange) et inter (violet) de PRMT2/RSF1 sur le modèle du complexe. Les deux monomères de PRMT2 sont différenciables par leurs couleurs, cyan et gris et RSF1 en vert.

Les expériences de MS structurale ont pu révéler les premières informations structurales sur ce complexe PRMT2/RSF1, qui sont :

- La formation d'un complexe asymétrique d'un monomère de RSF1 pour un dimère de PRMT2 avec deux donneurs de méthyles (SAM) montré par nMS.
- La confirmation de l'activité enzymatique révélée en HDX-MS par la protection des domaines de liaison des SAM de PRMT2 et des arginines de RSF1, précédemment identifiées comme méthylées.
- La localisation partielle de la zone d'interaction des deux protéines a été démontré en HDX-MS (vers le « haut » du dimère).
- La bonne formation du dimère prouvée par le cross-link inter-monomère de PRMT2 en accord avec la structure.
- La proximité spatiale de PRMT2 et RSF1 représenté par les cinq peptides pontés inter-PRMT2/RSF1.
- L'obtention de résultats expérimentaux (HDX-MS et XL-MS) permettant de confirmer le modèle du complexe, bien que la position des domaines SH3 serait encore à préciser.

1.5. Valorisation

Ces résultats ont été valorisés par une communication orale :

Integration of mass spectrometry data to structural biology for the characterization of PRMT2/RSF1 interaction. Marie Ley, Vincent Cura, Nils Marechal, Nathalie Troffer-Charlier, Jean Cavarelli, Sarah Cianférani. BSI du 29 novembre au 3 décembre 2021

Une communication par affiche :

Structural mass spectrometry approaches for the characterization of the interaction between protein arginine methyl transferase 2 and its binding partner RSF1. Marie Ley, Vincent Cura, Nathalie Troffer-Charlier, Jean Cavarelli and Sarah Cianférani. ASMS du 31 octobre au 4 novembre 2021

Un manuscrit est en cours de préparation :

Structural insights into the arginine methyltransferase 2 protein complex with an arginine-rich partner

2. Caractérisation des complexes impliquant PRMT9

De la même manière que pour le complexe PRMT2/RSF1, PRMT9 a été étudié en interaction avec différents partenaires par MS structurale. PRMT9 possède, à ce jour, deux substrats connus SAP145 et SAP49³³³⁻³³⁴. Toutefois aucune structure de ces complexes n'a encore été obtenue.

2.1. Complexe ternaire PRMT9/SAP49/SAP145

Avant de réaliser des expériences de XL-MS ou HDX-MS pour apporter des précisions sur les proximités spatiales et la dynamique d'interaction, les complexes sont étudiés en nMS. Différentes optimisations des conditions de dessalage (100 ou 200mM d'acétate d'ammonium) et d'analyses (voltage de cône 60, 120, 180 et 200 V) ont été effectuées pour l'analyse du complexe ternaire PRMT9/SAP145/SAP49 (Annexe 2). Cependant, seul SAP49 (44 513 ± 13 Da) et PRMT9 (94 914 ± 35 Da) libres sont détectées (Fig. 9). La protéine SAP145 n'a pas été révélée par nMS. L'absence de complexe peut être expliquée par la faible concentration de l'échantillon après dessalage (6,2 µM maximum) qui peut soit être insuffisante en nMS, ou ne pas permettre le maintien du complexe en solution.

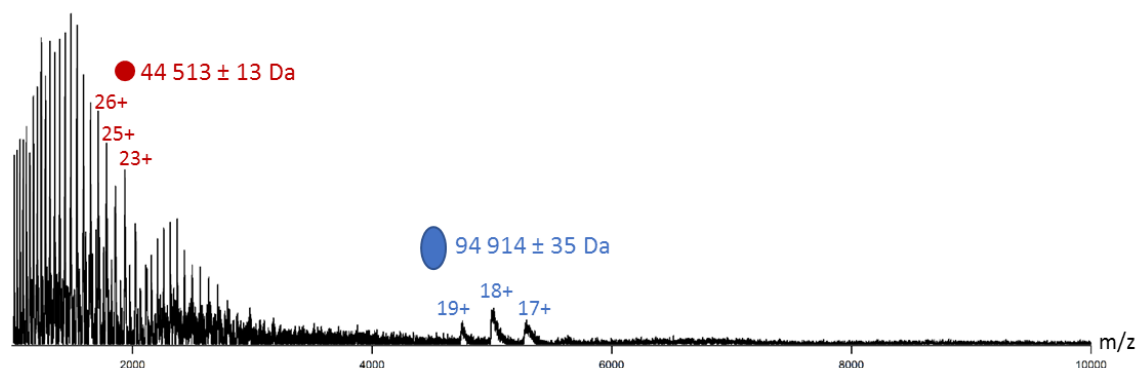


Figure 9 : Spectre de masse obtenu par nMS de PRMT9/SAP145/SAP49 avec un Vc de 60 V et une Pi de 6 mbar.

2.2. Complexe binaire PRMT9/SAP145₄₀₁₋₅₅₀

Malgré ces premières analyses peu concluantes, l'étude a été poursuivie en s'intéressant cette fois au sous-complexe binaire entre PRMT9 (94 804 Da) et un fragment de 150 acides aminés de SAP145 (SAP145₄₀₁₋₅₅₀, 18 067 Da) dont l'interaction a déjà été décrite³³⁴ (Annexe 2). Ce sous-complexe est de masse inférieure au complexe ternaire PRMT9/SAP145/SAP49 (113 kDa au lieu de 241 kDa) et obtenu à une plus grande concentration molaire, laquelle facilite l'analyse en nMS.

2.2.1. nMS

L'analyse en nMS du sous-complexe PRMT9/SAP145₄₀₁₋₅₅₀ a mis en évidence les deux protéines monomériques ($18\,362 \pm 1$ Da pour SAP145₄₀₁₋₅₅₀ ; $94\,844 \pm 11$ Da pour PRMT9, Fig. 10). Le fragment de SAP145₄₀₁₋₅₅₀ possède un incrément de masse non expliqué de 295 Da par rapport à la masse théorique. Toutefois, la présence d'un signal intense correspondant au complexe PRMT9/SAP145₄₀₁₋₅₅₀ ($113\,192 \pm 8$ Da) de stœchiométrie 1:1 atteste de la formation du complexe. De plus, ce complexe est identifiable même à haut voltage (200V) suggérant une forte interaction entre ces deux protéines et confirmant que l'espèce détectée de 18kDa est bien le fragment de SAP145.

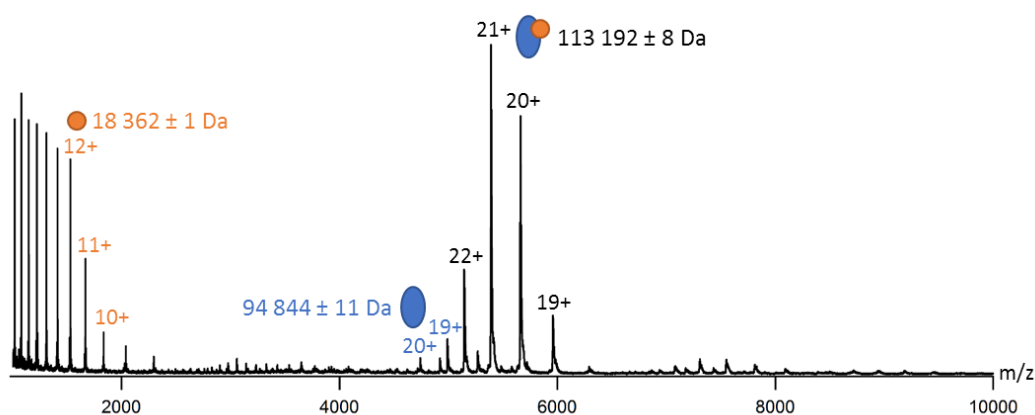


Figure 10 : Spectre de masse obtenu par nMS de PRMT9/SAP145₄₀₁₋₅₅₀ avec un Vc de 200 V et une Pi de 6 mbar.

2.2.2. HDX-MS

Suite à l'observation en nMS du complexe PRMT9/SAP145₄₀₁₋₅₅₀ en tant qu'espèce majoritaire, des expériences manuelles d'échange H/D ont été menées sur le complexe permettant d'obtenir une couverture de séquence de 92,4% de PRMT9 avec une redondance de 2,9 et 209 peptides identifiés (Annexe 39). Après l'échange et la vérification des profils isotopiques, seuls 130 peptides ont pu être utilisés, correspondant à 78,1% de la séquence de PRMT9. Un écart type important des incorporations en D des triplicats d'analyses a été observé, allant jusqu'à 3 Da de différence entre les réplicats pour un même temps de deutération. Ce manque de répétabilité empêche la validation statistique de différence d'incorporation en D (Annexe 40). Ces variations sont imputables soit à une instabilité des protéines, soit à une forte dynamique du complexe avec des échanges rapides de D en H et inversement. Par conséquent, malgré plusieurs tentatives, l'approche HDX-MS n'a pas pu être mise en œuvre sur ce complexe.

2.2.3. XL-MS

Le XL-MS est une technique de choix dans ce cas, car elle peut permettre de stabiliser le complexe (ou un état structural du complexe) par le pontage des acides aminés au sein et entre les deux protéines. Des optimisations des conditions du cross-link ont été effectuées sur gel SDS-PAGE (temps, température, excès d'agent pontant) puis par MS (tampon de dilution). Ces optimisations sont décrites dans ce manuscrit (Partie I, Chapitre 2, V.1.). Finalement, les résultats obtenus ont permis l'identification de 75 cross-links unique après le pontage du complexe en triplicata avec du DSBU et du DSAU.

Parmi ces peptides pontés, 23 peptides intra-SAP145₄₀₁₋₅₅₀ ont été identifiés dans au moins deux des trois réplicats (Fig. 11, Annexe 41). Les résidus pontés ne sont pas localisés dans une région précise de la protéine, mais indique que de manière générale les régions sont proches les unes des autres, suggérant une conformation relativement compacte de SAP145₄₀₁₋₅₅₀.

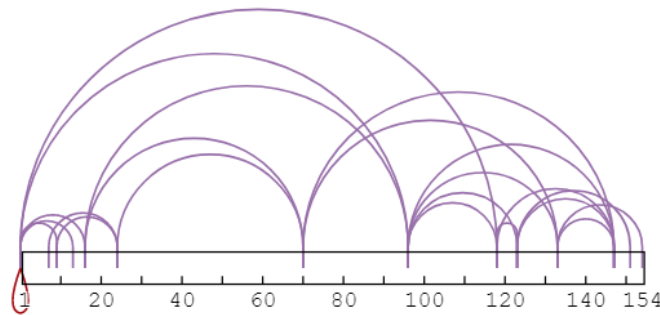


Figure 11 : Représentation schématique des peptides pontés (violet) intra-SAP145₄₀₁₋₅₅₀ (obtenus avec Xinet²¹⁶).

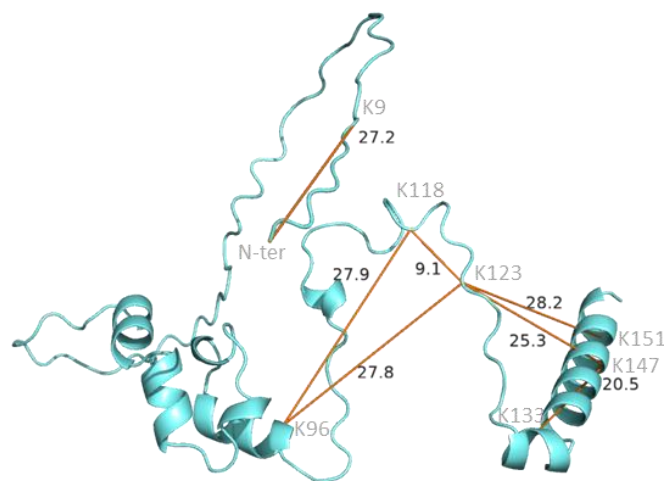


Figure 12 : Représentation des cross-links validés intra-SAP145₄₀₁₋₅₅₀ (orange) sur le modèle AlphaFold³⁴³.

J'ai pu obtenir un modèle de SAP145₄₀₁₋₅₅₀ via le logiciel Alphafold³⁴³, avec une faible confiance sur les 50 premiers acides aminés (Annexe 42). Les distances C α -C α ont été mesurées sur ce modèle (Fig. 12) et seulement 7 distances (soit 30%) sont comprises dans nos gammes allant de 6 à 26/30Å pour le DSAU et le DSBU. La fiabilité de ce modèle est donc remise en cause.

Les 13 cross-links intra-PRMT9 sont aussi répartis sur l'ensemble des domaines de la protéine (Annexe 43). Des distances C α -C α ont pu être mesurées sur la structure de PRMT9 (PDB : 6PDM) entre les résidus impliqués dans le pontage (Figure 13, Tableau 4). Cette structure reste toutefois incomplète avec la partie N-terminale comprenant les trois répétitions tétratricopeptidiques (TRP) non visibles. Toutes les distances mesurables sont en accord avec la gamme de distance des deux agents pontants (de 6 à 26/30Å) confirmant la structure du domaine RF1, du tonneau β 2 de PRMT9 (Tableau 4) et validant la qualité de nos données XL-MS.

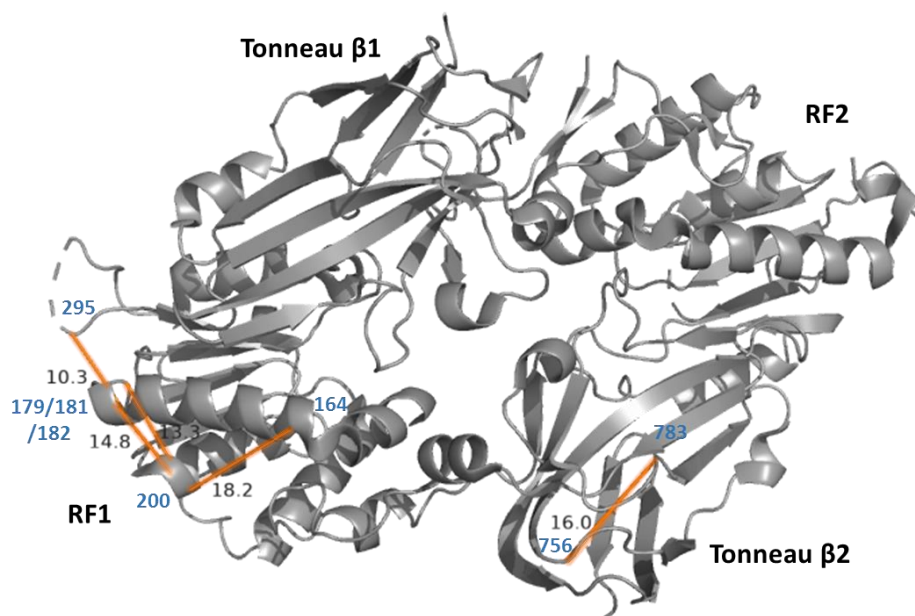


Figure 13 : Représentation des cross-links validés intra-PRMT9 (orange) sur sa structure (PDB : 6PDM). Les numérotations des acides aminés impliqués dans le pontage sont indiquées en bleu.

| Protéine 1 | Protéine 2 | Site 1 | Site 2 | Domaine site 1 | Domaine site 2 | Distance C α -C α (Å) PDB : 6PDM |
|------------|------------|--------|--------|-------------------|-------------------|---|
| PRMT9 | PRMT9 | K164 | K200 | - | RF1 | 18,2 |
| PRMT9 | PRMT9 | S179 | K200 | - | RF1 | 14,8 |
| PRMT9 | PRMT9 | S181 | K200 | - | RF1 | 13,3 |
| PRMT9 | PRMT9 | K182 | K295 | - | RF1 | 10,3 |
| PRMT9 | PRMT9 | K756 | K783 | Tonneau β 2 | Tonneau β 2 | 16 |

Tableau 4 : Cross-links intra-protéines identifiés pour l'analyse du complexe PRMT9/SAP145₄₀₁₋₅₅₀.

La plupart des peptides pontés (39 sur 75) sont inter-PRMT9/SAP145₄₀₁₋₅₅₀ cependant seul 13 résidus de PRMT9 sont cross-linkés. Ces acides aminés permettent de localiser cinq régions de PRMT9 proches de SAP145₄₀₁₋₅₅₀ (Fig. 14), dont :

- La partie N-terminale de PRMT9 avec les résidus pontés N-ter, S17 et S29.
- Le TRP3 avec les lysines 129, 132 et 141.
- Le domaine RF1 (K200, K238, K243 et K246) et RF2 (K655).
- Le motif III (S691) qui stabilise la poche de liaison du donneur de méthyle (SAM).
- Le tonneau β par l'acide aminé K778.

Le pontage de SAP145₄₀₁₋₅₅₀ a permis l'identification de 12 résidus cross-linkés principalement situés au niveau des régions N-terminale et C-terminale (Fig. 14, Annexe 1). L'arginine 508 déjà identifiée comme étant méthylée lorsque SAP145 interagit avec PRMT9³³⁴ correspond dans notre cas au résidu 112. La lysine 118, proche de l'arginine 112, est impliqué dans trois pontages avec les K238, K243 et K246 (du domaine catalytique RF1 près du motif II). Cette proximité de l'arginine 112 et de la poche de liaison des SAM confirme la relation substrat/enzyme de ce complexe PRMT9/SAP145₄₀₁₋₅₅₀.

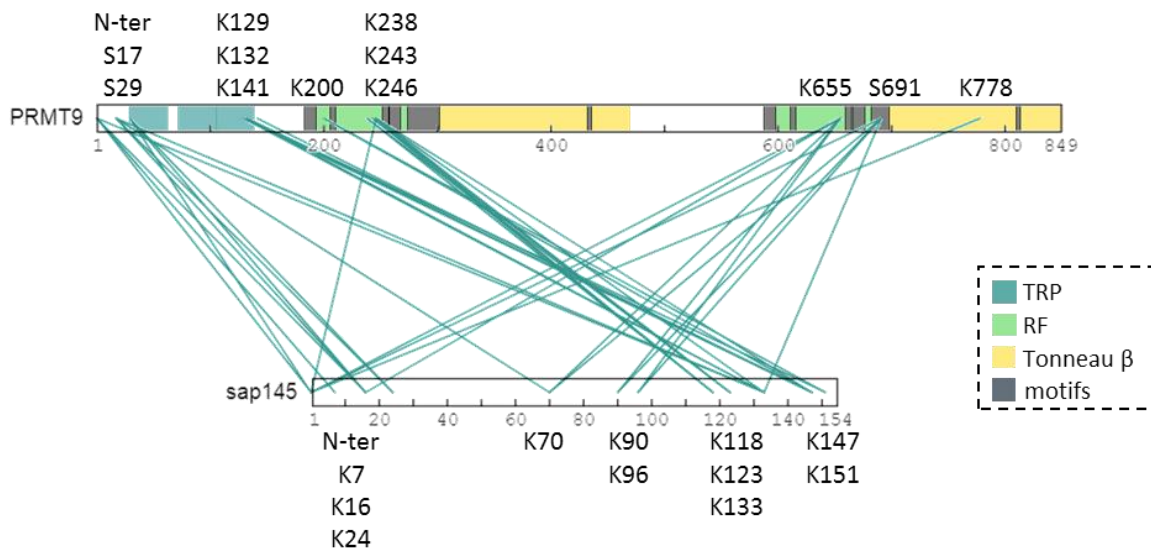


Figure 14 : Représentation schématique des peptides pontés (vert) inter-PRMT9/SAP145₄₀₁₋₅₅₀ (obtenus avec Xinet²¹⁶). Les répétitions TRP sont identifiables en vert-bleu, les repliements de Rossmann en vert, les tonneaux β en jaune et les motifs en gris. SAP145₄₀₁₋₅₅₀ numéroté de 1 à 154 avec les 4 premiers résidus composant le tag.

Finalement, bien qu'aucune structure du complexe ternaire PRMT9/SAP145/SAP49 ou le sous-complexe binaire PRMT9/SAP145₄₀₁₋₅₅₀ n'ait été obtenue pour le moment, les expériences de MS structurale ont pu mettre en lumière différentes informations telles que :

- La formation du sous-complexe binaire de stœchiométrie 1:1 montré par nMS confirmant l'interaction entre le fragment de SAP145 et l'enzyme PRMT9.
- Une forte dynamique conformationnelle du complexe PRMT9/SAP145₄₀₁₋₅₅₀ révélé par les expériences d'échange H/D.
- Un nombre important de peptides pontés intra-SAP145₄₀₁₋₅₅₀ (23) répartis tout au long de la séquence, qui pourront être utilisées pour générer ou confirmer un futur modèle.
- Une proximité du domaine catalytique de PRMT9 avec l'arginine 508 (numéroté R112 sur la figure 14) de SAP145₄₀₁₋₅₅₀, identifiée comme étant méthylé en présence de l'enzyme.
- Les régions N-terminale et C-terminale de PRMT9 proches du fragment de SAP145.

V. Conclusions et perspectives

La combinaison des expériences de nMS, HDX-MS et XL-MS a fourni des informations structurales sur les complexes impliquant les PRMTs telles que :

- La présence/absence des complexes ainsi que leurs stœchiométries par nMS.
- La confirmation de l'activité enzymatique de PRMT2 en présence de RSF1 par la protection des régions jouant un rôle dans le processus de méthylation observée en HDX-MS.
- La validation expérimentale d'un modèle du complexe PRMT2/RSF1 par XL-MS.
- La proximité spatiale de régions de PRMT9 et SAP145₄₀₁₋₅₅₀ pour la première fois définie par XL-MS.

L'ensemble de ces résultats sont complémentaires et tendent à converger vers la compréhension des mécanismes de reconnaissance des substrats et de la structure de ces complexes.

Bien que l'étude du complexe PRMT2/RSF1 permette de faire la lumière sur l'interaction enzyme/substrat, RSF1 est une protéine spécifique aux insectes et ne peut donc pas être un substrat naturel de PRMT2 humaine. Il serait de ce fait intéressant de chercher une protéine appartenant à la même famille de protéine riche en sérine/arginine (SR) qui serait un potentiel

substrat naturel de PRMT2. Chez l'Humain, deux protéines SRSF7 et SRSF3, qui présentent une forte similarité de séquence avec RSF1, sont des facteurs d'épissages SR et pourraient être étudiés de la même manière que RSF1³¹⁹.

D'un autre côté, les informations obtenues suite à l'analyse du sous-complexe PRMT9/SAP145₄₀₁₋₅₅₀ pourraient aider à la résolution structurale du complexe ternaire PRMT9/SAP145/SAP49. Des expériences de M-Ph peuvent être réalisées sur ce complexe, car cette technique ne nécessite pas une concentration de l'ordre du μM mais du nM. Cette technique permettrait également de nous affranchir de la difficulté que présente l'ionisation de composés de hautes masses comme peut l'être ce complexe (environ 241 kDa) ainsi que le maintien des interactions. Le pontage chimique de PRMT9/SAP145/SAP49 pourra être effectué à partir des optimisations réalisées sur le sous-complexe. Enfin, des expériences de HDX-MS du sous-complexe PRMT9/SAP145₄₀₁₋₅₅₀ pourrait être réitérées avec une deutération manuelle dans le but de pallier à la faible répétabilité obtenu précédemment.

Chapitre 4 : Apport du XL-MS pour la caractérisation de macrocomplexes impliquant les ATPases RUVBL1/2 humaines

I. Contexte analytique et biologique

Les protéines RUVBL1 (R1) et RUVBL2 (R2) sont des protéines de ≈ 50 kDa qui font partie de la superfamille des protéines AAA+ (pour *ATPases Associated with cellular Activities*) et présentent une similarité importante de séquence de 65% chez l'Humain³⁴⁴. En fonction de leurs interactions avec différents partenaires protéiques ou nucléiques, elles jouent un rôle essentiel dans de nombreux processus cellulaires comme la réplication de l'ADN, la dégradation des protéines, la fusion des membranes, la transduction des signaux ou encore la régulation de l'expression des gènes³⁴⁵⁻³⁴⁸. Les ATPases R1 et R2 s'assemblent en hétéro-dodécamères (stœchiométrie 6:6 R1:R2, 790 kDa), qui sont ensuite dissociés en hétéro-hexamères (3:3 R1:R2, 344 kDa) lors de la liaison à leurs partenaires et fixation/hydrolyse de l'ATP pour assurer leur activité de protéine chaperon^{240, 349}. Les RUVBLs sont conservées chez les eucaryotes et composées de trois domaines (DI, DII et DIII), DI et DIII formant l'anneau hexamérique AAA+ tandis que le DII est le principal domaine d'interaction avec d'autres protéines/co-facteurs (Fig. 1)³⁵⁰⁻³⁵¹. Plusieurs structures de l'hexamère R1R2 sont disponibles avec différentes conformations du domaine DII, ouverte (PDB : 7OLE³⁵⁰) et fermée (PDB : 5OAF²⁴⁰).

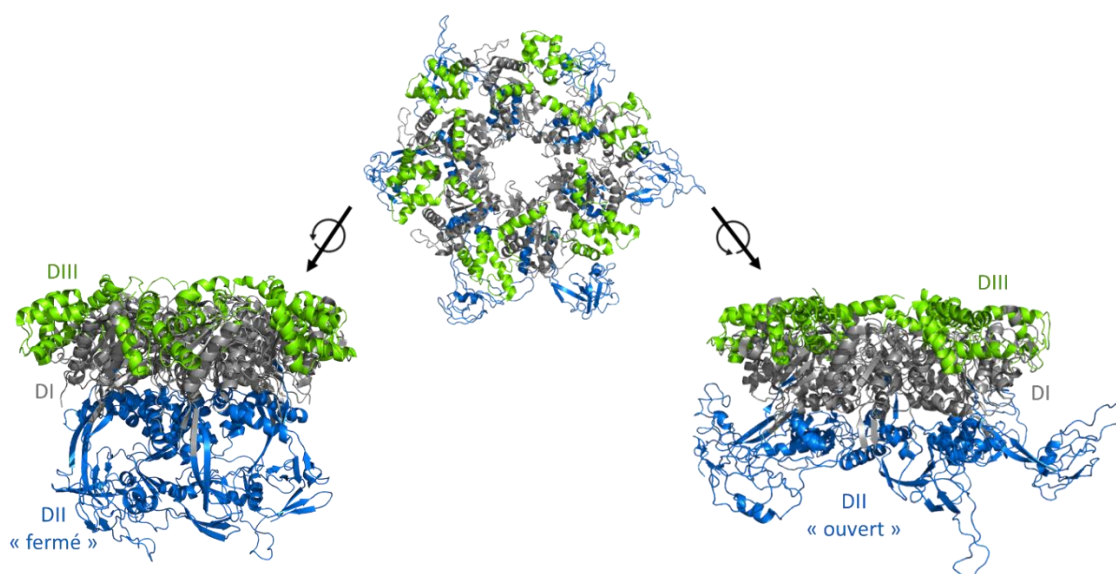


Figure 1. Structures de l'hexamère R1R2 (PDB : 5OAF gauche, PDB : 7OLE droite et centre) avec les domaines DI en gris, DII en bleu et DIII en vert.

Les RUVBLs sont notamment impliquées dans les complexes R2TP avec l'hétérodimère RPAP3/PIH1D1 (≈ 360 kDa) et R2TP-like³⁵²⁻³⁵⁵ (Fig. 2). RPAP3 est un adaptateur entre les HSP70/90 et les R1/R2 tout en stabilisant PIH1D1 qui recrute lui-même les clients via notamment son domaine N-terminal de fixation au phospho-peptide³⁵⁶⁻³⁵⁷. Le R2TP assure une fonction de co-protéine chaperon avec HPS70/90 dont une des cibles est la petite particule ribonucléoprotéique nucléolaire à boîtes C/D (snoRNP C/D) et H/ACA (snoRNP H/ACA)³⁵⁸. En 2018, un complexe de type R2TP, nommé R2SP, a été mis en évidence par nos collaborateurs, dans lequel RPAP3 et PIH1D1 sont respectivement remplacées par deux protéines homologues, SPAG1 et PIH1D2³⁵⁹. Ce macrocomplexe est enrichi dans les testicules et sa spécificité est modifiée pour reconnaître d'autres partenaires tels que des complexes de liprines³⁶⁰ ou encore les bras de dynéines³⁶¹. Ces protéines possèdent des domaines en commun avec leurs homologues telles que pour SPAG1, un domaine C-terminal (RPAP3-C) dont la fonction est de recruter les R1/R2, une région d'interaction avec PIH1D2 et trois domaines TPR consécutifs (2 pour RPAP3) servant notamment comme région d'interaction avec les HSP70/90³⁶²⁻³⁶⁴. Dans le cas de PIH1D2, les domaines similaires à PIH1D1 sont : (i) PIH en N-terminal possédant une poche de liaison phospho-peptidique responsable de la reconnaissance de certains substrats ; et (ii) CS (*CHORD-containing proteins and SGT1*) en C-terminal assurant l'hétérodimérisation avec SPAG1³⁶⁵.

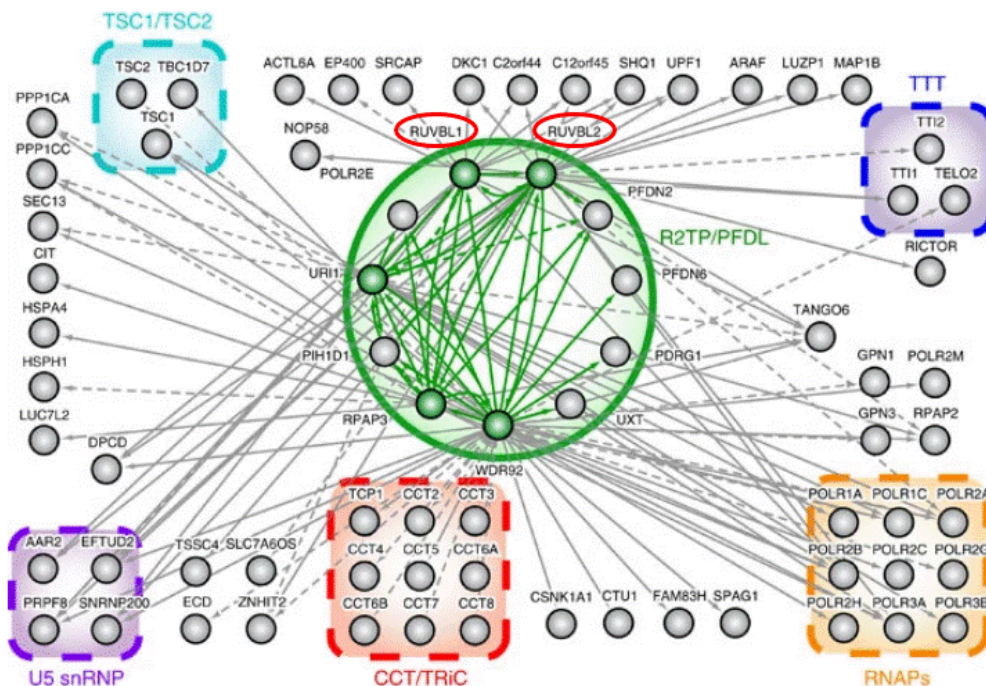


Figure 2. Diagramme du réseau d'interactions formé autour du complexe R2TP + PFDL (*PreFoldin-Like*) ou encore PAQosome (adapté de Cloutier et al, 2017³⁵⁴).

Un autre partenaire d'interaction de R1R2 a été découvert par analyse protéomique : DPCD (pour *deleted in a mouse model of primary ciliary dyskinesia*)^{354, 359}. Cette protéine, de 23 kDa, joue un rôle clé dans la formation des cils, son absence entraînant une dyskinésie ciliaire primaire³⁶⁶. Ce rôle est en accord avec la description de R1 comme étant une protéine associée à la ciliopathie et au pré-assemblage des complexes protéiques ciliaires^{354, 367}. La structure 3D de DPCD était encore inconnue au début de cette étude, car DPCD n'a aucun homologue fonctionnel ou de domaines fonctionnels conservés à ce jour. DPCD contient néanmoins plusieurs sites potentiels de phosphorylation, d'amidation et de glycosylation³⁶⁸.

Ce projet a été réalisé dans le cadre d'un projet ANR collaboratif (ANR snoRNPassembly) impliquant cinq équipes partenaires (Dr. E. Bertrand IGH -UMR 9002, Montpellier ; Dr. C. Plisson-Chastang CBI, Toulouse ; Pr. B. Charpentier/ Dr. X. Manival IMoPA - UMR 7365, Nancy ; Dr. P. Meyer LBMCE - UMR 8226, Paris et Dr. T. Bandejas iBET - Lisbon, Portugal). Le but de ce consortium est la caractérisation fonctionnelle et structurale des assemblages impliquant les RUVBLs en combinant différentes techniques comme la nMS, le XL-MS, la RMN, la diffraction des rayons X, la cryo-EM, le SAXS et d'autres techniques biophysiques pour la partie structurale. J'ai contribué à ce projet en mettant en œuvre les approches de MS structurales décrites précédemment.

II. Objectifs

Dans le cadre de ma thèse, j'ai combiné les résultats obtenus en nMS à la M-Ph et au XL-MS pour obtenir des informations de stœchiométries d'interactions et de régions de proximités spatiales, ceci afin de guider la résolution de structures 3D obtenues par nos collaborateurs. Nous avons sélectionné le XL-MS plutôt que l'HDX-MS du fait de la faible quantité de matériel biologique disponible (200 µg, ne suffisent pas pour l'étude différentielle en HDX-MS), la stabilité de certains complexes dans le temps ainsi que l'hétérogénéité des systèmes. Dans cette étude, je me suis focalisée sur la M-Ph et le XL-MS, les spectres de nMS ayant été obtenus par le Dr. Evolène Deslignière. L'ensemble de ces approches a été mis en œuvre pour l'analyse des protéines/complexes suivants (Annexe 2) :

- R1R2
- DPCD
- R1R2/DPCD (R2D)
- SPAG1₆₂₂₋₉₂₆/PIH1D2 (SP)
- R1R2/SPAG1₆₂₂₋₉₂₆/PIH1D2 (R2SP)

III. Matériel et méthodes

Les expériences de M-Ph ont été réalisées sur les complexes suivants : R1R2, R1R2/DPCD et R1R2/SPAG1₆₂₂₋₉₂₆/PIH1D2. Avant l'analyse, les échantillons ont été dilués dans du PBS (1x) à un pH de 7,0 et sont déposés sur une lame de verre pour atteindre une concentration finale d'analyse de 10nM. Une calibration du photomètre de masse (Refeyn Two^{MP}) a été effectuée avec trois protéines, la BSA (66 kDa), un anticorps (150 kDa) et la GLDH (330 kDa). L'enregistrement du signal a été effectué pendant une minute et traité avec le logiciel Discover^{MP} 2.3.0. permettant l'obtention des histogrammes de distribution de contraste.

La nMS a été effectuée par infusion directe sur l'orbitrap EMR (Thermo Fisher Scientific). Les paramètres instrumentaux sont donnés dans les légendes des spectres montrés. Les données ont été exploitées avec BioPharma Finder (Thermo Fisher Scientific). En amont de l'analyse, les échantillons ont été dessalés deux fois sur Zeba (7 kDa, Thermo Fisher Scientific) à 1 500xg avec 200mM d'acétate d'ammonium à un pH de 7,0.

Le XL-MS a été réalisé avec du DSBU à un excès molaire de 100 dans un tampon composé de 20mM de HEPES, 100mM de NaCl et 0,5mM de TCEP à un pH de 7,5 à 20°C pendant 45 min pour tous les complexes. Tous les pontages ont été réalisés en triplicat puis réduits (DTT), alkylés (IAM) et digérés par de la Trypsine/Lys-C. Les excès de réactifs sont éliminés par passage sur SPE C18, séparés sur une colonne C18 et analysés sur un Orbitrap Q-Exactive Plus (pour R1R2/DPCP et DPCD) ou un Orbitrap Q-Exactive HF-X (pour R1R2, R2SP et SP) (Thermo Fisher Scientific). MeroX 2.0.1.4 a été utilisé pour l'identification des peptides pontés avec un FDR fixé à 5%. La présence des fragments caractéristiques attendu pour ce type d'agent pontant a été vérifiée et seuls les cross-links validés dans au moins 2 des 3 répliquats ont été conservés. L'intervalle de distance est fixé de 6 à 30Å pour le DSBU.

IV. Résultats

1. Caractérisation du complexe R1R2

Dans un premier temps le complexe R1R2, bien que déjà étudié structuralement, a été analysé en M-Ph, nMS et XL-MS.

1.1. M-Ph et nMS

Afin d'obtenir des informations sur la stœchiométrie des espèces ainsi que l'homogénéité des échantillons, des analyses en nMS et en M-Ph ont été effectuées. La M-Ph a permis de révéler la présence de ces trois espèces de : (i) 52 kDa (monomère) ; (ii) 331 kDa (hexamère de stœchiométrie 3:3 R1:R2) ; et (iii) 649 kDa (dodécamère de stœchiométrie 6:6 R1:R2) en proportions différentes (48% de monomère, 33% d'hexamère et 19% de dodécamère, Fig. 3A). Ces résultats sont en accord avec ceux obtenus en nMS, une technique de plus haute résolution. Le monomère (non montré, $\approx 52\%$), l'hexamère ($313\,086 \pm 59$ Da, 3:3 R1:R2) et le dodécamère ($625\,789 \pm 169$ Da, 6:6 R1:R2) ont été identifiés pour l'analyse en nMS du complexe R1R2 (Fig 3B). La M-Ph présente l'avantage, par rapport à la nMS, de pouvoir analyser une faible quantité de matériel de l'ordre du nano molaire sans effectuer de préparation préalable à l'analyse et d'être effectué en solution (sans compétition à l'ionisation)^{90, 369}. Ces avantages sont primordiaux dans le cas de l'étude d'échantillons hétérogènes et de hauts poids moléculaires comme le sont ces protéines. Ils nous permettent aussi de réaliser une vérification rapide (10 min) des échantillons en amont des expériences de XL-MS. Néanmoins du fait de la résolution plus élevée, la nMS est la seule approche permettant de valider par exemple la présence d'ATP/ADP dans les échantillons ce qui est crucial pour étudier la mécanistique des R1R2.

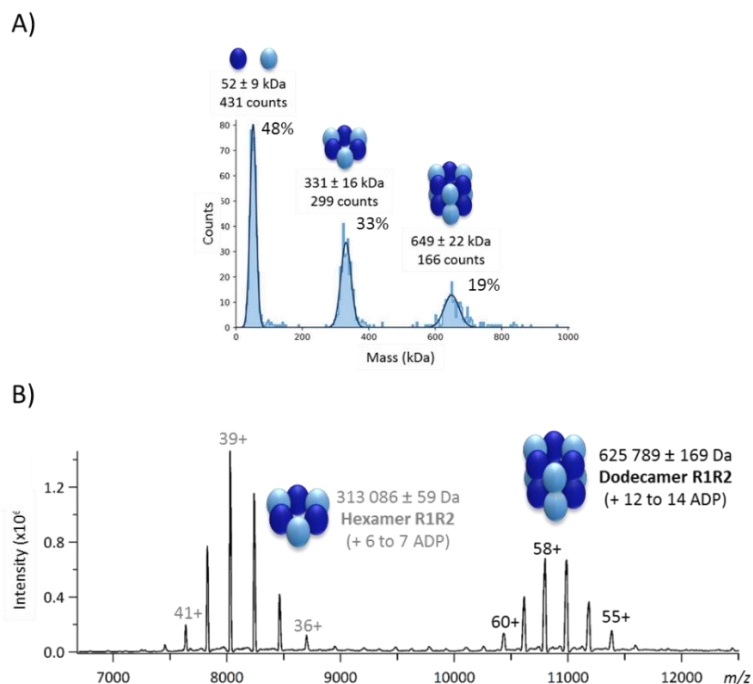


Figure 3. (A) Histogramme de distribution de contraste obtenu en M-Ph et (B) spectre de masse native du complexes R1R2.

1.2. XL-MS

Les deux ATPases forment donc différents complexes (hexamères et dodécamères), avec deux conformations connues pour l'hexamère (ouverte PDB : 7OLE³⁵⁰ et fermée PDB : 5OAF²⁴⁰). Dans le but de déterminer la conformation de notre complexe en solution, des expériences de XL-MS ont été réalisées avec du DSBU. L'analyse XL-MS a permis d'identifier 25 peptides pontés inter-R1/R2, 10 intra-R2 et 42 intra-R1 (Annexe 45). Ce nombre important de cross-links identifiés (77) permet d'obtenir des informations sur les proximités spatiales des acides aminés de tous les domaines des RUVBLs. Toutefois, une grande proportion des peptides cross-linkés implique le domaine DII des deux ATPases (90% des cross-links intra-R2, 52% des intra-R1 et 56% des inter-R1/R2). Le domaine DII en contact avec le solvant, prend différentes conformations pour former l'hexamère ouvert ou fermé et intervient aussi dans l'interface de formation du dodécamère³⁷⁰. Sachant, que deux structures de l'hexamère ont déjà été obtenues, les distances entre les acides aminés pontés peuvent être mesurées. Ainsi, considérant l'intervalle de distance allant de 6 à 30Å fixé par le cross-linker (DSBU), 46% des distances mesurables sur la structure « ouverte » 7OLE sont comprises dans cet intervalle (Fig. 4) et 42% sur la structure « fermé » 5OAF.

Les distances mesurées qui entrent dans la gamme sont inférieures à 50%¹⁸², nous pouvons donc en conclure que les structures ne sont pas en accord avec les cross-links. Ces résultats peuvent être expliqués par :

- La génération d'artéfacts dus aux conditions de cross-links, ce qui est peu probable du fait du faible excès molaire de DSBU ajoutés (x100)
- La présence de mélanges d'oligomères (hexamère/dodécamère), plus probable au vu des résultats obtenus en M-Ph et nMS
- La présence de différents conformères des RUVBLs en équilibre avec la conformation ouverte ou fermée de leurs domaines DII, sachant qu'aucun partenaire est présent pour interagir avec les RUVBLs et que l'interaction entraîne souvent la stabilisation d'un conformère

Finalement, les expériences de M-Ph, nMS et XL-MS du complexe R1R2 ont conforté la coexistence de plusieurs populations (oligomères) de ces protéines. Effectivement, bien que ce complexe soit relativement bien décrit structurellement aucune des deux structures disponibles dans la PDB, n'est en accord total avec nos données de XL-MS.

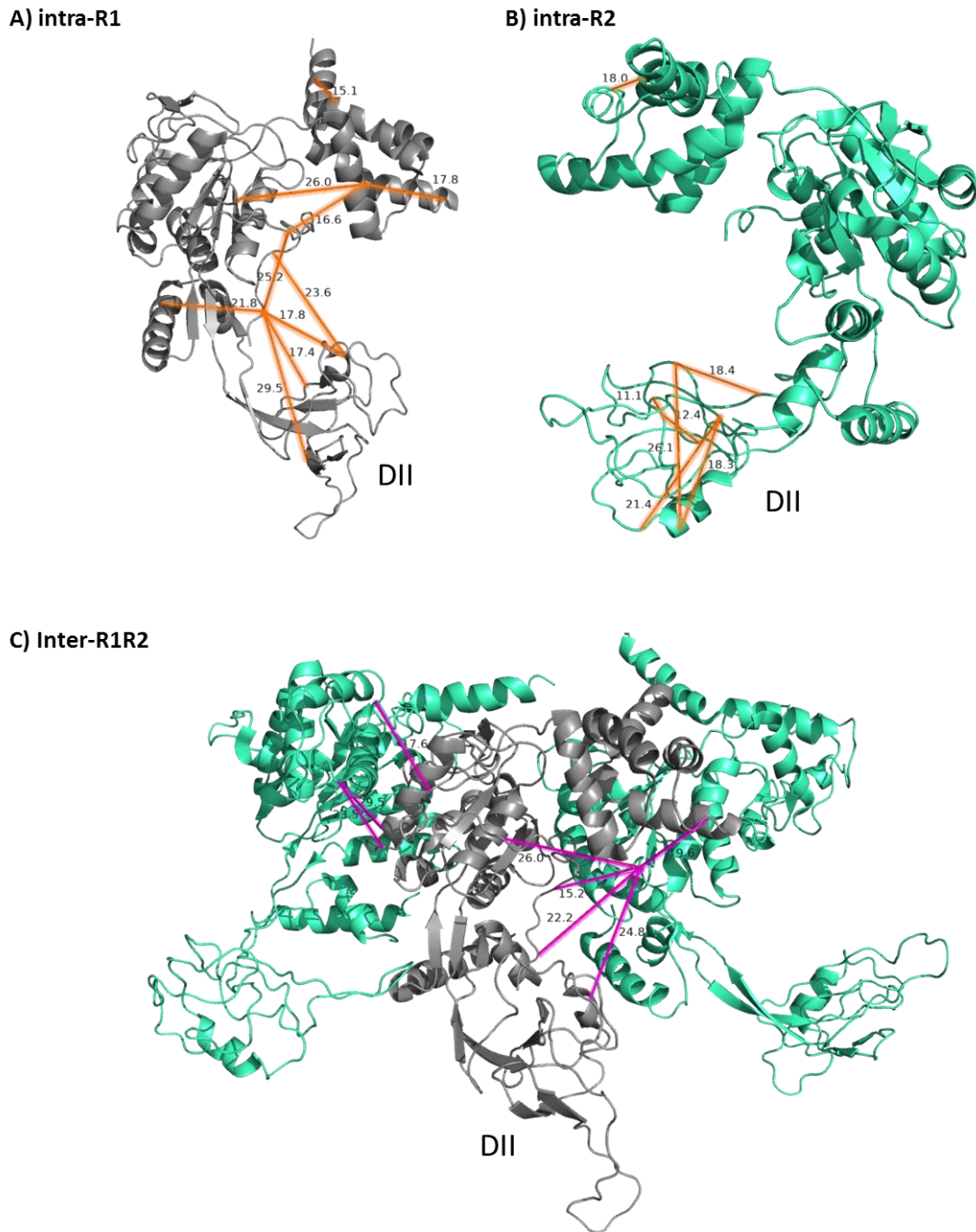


Figure 4. Représentation des distances mesurables et comprises dans l'intervalle de distance (A) intra-R1, (B) intra-R2 et (C) inter-R1/R2 du complexe R1R2 sur la structure (PDB : 7OLE). R1 et R2 sont représentés en gris et cyan respectivement, les cross-links intra-protéine (orange) et inter-R1/R2 (violet).

2. Caractérisation du complexe R1R2/DPCD (R2D)

Après avoir obtenu les premiers résultats sur le complexe impliquant les RUVBLs sous forme libre, l'étude s'est ensuite focalisée sur le complexe R1R2/DPCD (R2D) dans le but d'obtenir des informations structurales.

2.1. M-Ph et nMS

La M-Ph du complexe R2D (Fig. 5A) a révélé trois distributions : (i) monomères R1 ou R2 (84 ± 37 kDa) ; (ii) le complexes R2D de stœchiométries 3:3:2 ou 3:3:3 (380 ± 43 kDa) largement majoritaire (69%) ; et (iii) le complexe dodécamérique R1R2 (717 ± 105 kDa) de trop faible intensité pour déterminer sa masse exacte. Comparée à l'analyse de R1R2, l'abondance des monomères est plus faible en présence de DPCD (seulement 19% contre 48% pour R1R2) de même que pour le dodécamère (12% contre 19%). La présence de DPCD favorise donc la formation de l'hexamère, ce qui est en accord avec la fonction de protéine chaperon des RUVBLs dont la transition dodécamère/hexamère est la base mécanistique de leurs activités biologiques³⁴⁵. La nMS du complexe confirme les résultats de M-Ph avec la présence des deux complexes R2D de stœchiométries différentes 3:3:2 ($360\,082 \pm 77$ Da) et 3:3:3 ($383\,611 \pm 66$ Da) avec 6-7 molécules d'ADP (Fig. 5B).

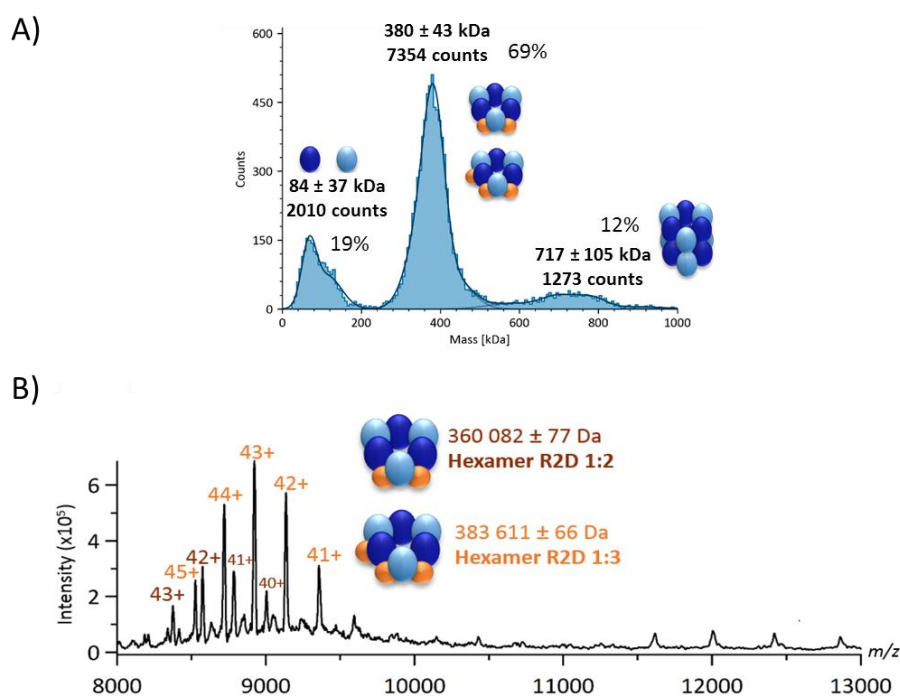


Figure 5. (A) Histogramme de distribution de contraste obtenue en M-Ph et (B) spectre de masse native du complexes R2D.

2.2. XL-MS de DPCD

DPCD est encore très peu décrit structuralement, mais un modèle basé sur la prédiction *de novo* de la structure a été obtenu par nos collaborateurs, suggérant la présence d'un domaine CS putatif (K99-K172)³⁷¹. Dans le but de corroborer cette hypothèse, des expériences de XL-MS ont été réalisées sur DPCD libre. Le cross-link avec le DSBU a montré 10 peptides pontés

dont 7 impliquant le groupement amine en N-terminal (Annexe 46). La distance entre les acides aminés pontés a été mesurée au niveau du modèle et 8 sont comprises dans l'intervalle de 6 à 30Å (Fig 6). Environ 80% des distances mesurées sur le modèle sont en accords avec le XL-MS. Ces analyses ont donc permis dans le cas de DPCD, de confirmer l'existence en solution du modèle prédit *ab-initio*.

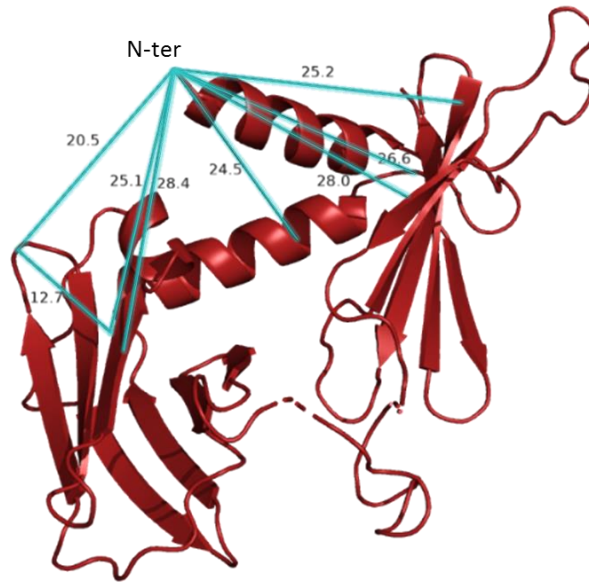


Figure 6. Représentation des distances mesurées, inférieures à 30Å, intra-DPCD (cyan) sur le modèle.

2.3. XL-MS du complexe R1R2/DPCD

Nous nous sommes finalement intéressés au pontage chimique du complexe R2D. Il a permis d'identifier 13 peptides intra-DPCD, 21 intra-R1, 9 intra-R2, 13 inter-DPCD/R1, 14 inter-DPCD/R2 et 18 inter-R1/R2 (Fig. 7, Annexe 47). Certains peptides pontés (42%) sont communs avec ceux identifiés au préalable pour DPCD seul et le complexe R1R2 libre (6 des 13 intra-DPCD, 9 des 21 intra-R1, 2 des 9 intra-R2 et 8 des 18 inter-R1/R2). Ils apportent une certaine confiance quant au jeu de données de XL-MS.

Trois régions de DPCD sont principalement impliquées dans l'interaction avec les RUVBLs :

- Les extrémités N- et C-ter comprenant des hélices α proches les unes des autres
- La région du domaine CS putatif

De manière intéressante, plus de la moitié des cross-links inter-RUVBLs/DPCD (8 des 13 pour R1 et 7 des 14 pour R2) se situent au niveau des domaines DII des RUVBLs connus comme étant le domaine d'interaction majeur avec d'autres protéines³⁵⁰.

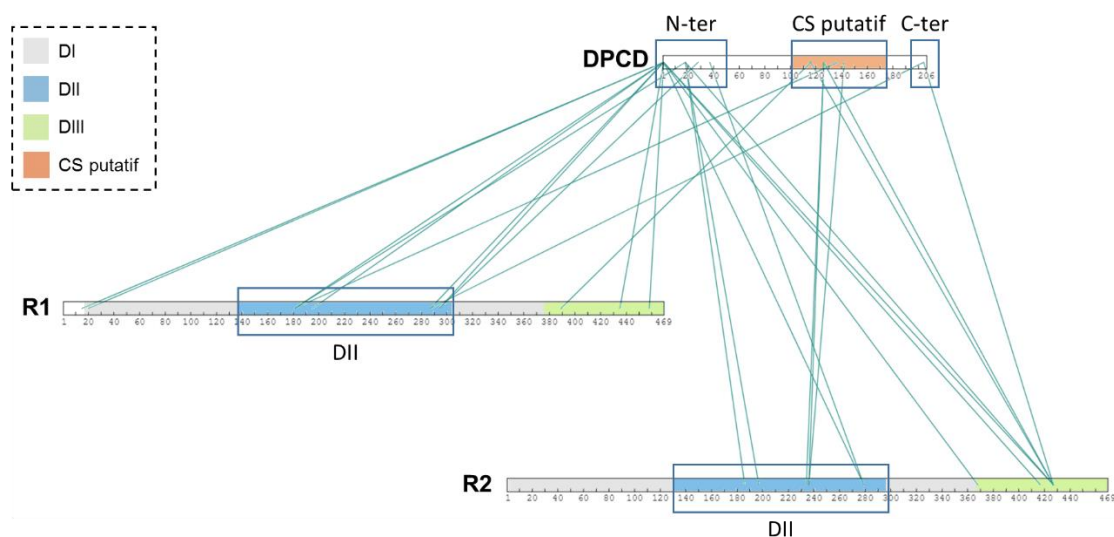


Figure 7. Représentation schématique des peptides pontés inter-DPCD/R1 et inter-DPCD/R2 (vert) du complexe R2D (obtenue avec Xinet²¹⁶).

Un modèle préliminaire du complexe R1R2/DPCD issu d'une carte Cryo-EM (à 6Å de résolution) a été obtenu par nos collaborateurs où DPCD a été placé en fonction des densités additionnelles peu précises, présentes sur les cartes EM, à côté de 4 des 6 domaines DII de RUVBLs qui sont dans une configuration ouverte (Fig. 8). Parmi les distances mesurables de nombreuses sont comprises dans l'intervalle de tolérance des expériences de XL-MS (6-30Å) :

- 92% des peptides intra-DPCD
- 100% des peptides intra-R1
- 63% des peptides intra-R2

Les expériences de XL-MS ont donc permis de valider les structures de R1 et R2 et le modèle de DPCD libre. Concernant la conformation de l'hexamère R1R2, 58% des distances correspondant au cross-links inter-R1/R2 sont inférieures à 30Å. Ce pourcentage est légèrement supérieur à celui obtenu pour les analyses de l'hexamère R1R2 lié en conformation ouverte (7OLE³⁵⁰, 40%) ou fermé (5OAF²⁴⁰, 46%). Il est intéressant de noter que toutes les distances mesurées des cross-links inter-R1/R2 lié entre le domaine DII et le DI ou le DIII (K181-K427, K287-K427 et K435-K160) ne sont pas comprises dans l'intervalle de 6-30Å (Tableau 1), ce qui ne permet pas de confirmer la conformation ouverte ou fermé des DII. Ces résultats sont en accords avec les tentatives d'ajustement des modèles de R2D sur des données SAXS présentant un χ^2 élevé et dont la courbe montre l'existence de plusieurs états conformationnels des RUVBLs en raison de la flexibilité probable leurs domaines DII même sous forme lié.

| Protéine 1 | Protéine 2 | Site 1 | Site 2 | Domaine site 1 | Domaine site 2 | Distance Ca-Ca |
|------------|------------|-----------|--------|----------------|----------------|----------------|
| RuvBL1 | RuvBL2 | K20 | K67 | DI | DI | 24,3 |
| RuvBL1 | RuvBL2 | K24 | K67 | DI | DI | 18,3 |
| RuvBL1 | RuvBL2 | K46 | K417 | DI | DIII | 19,4 |
| RuvBL1 | RuvBL2 | K46 | K427 | DI | DIII | 20,6 |
| RuvBL1 | RuvBL2 | K70/S71 | K417 | DI | DIII | 28,4 |
| RuvBL1 | RuvBL2 | T180/K181 | K427 | DII | DIII | 65,8 |
| RuvBL1 | RuvBL2 | K281 | K269 | DII | DII | N |
| RuvBL1 | RuvBL2 | K281 | K279 | DII | DII | N |
| RuvBL1 | RuvBL2 | K281 | K365 | DII | DI | N |
| RuvBL1 | RuvBL2 | K287 | K269 | DII | DII | 16,4 |
| RuvBL1 | RuvBL2 | K287 | K427 | DII | DIII | 43,2 |
| RuvBL1 | RuvBL2 | K294 | K269 | DII | DII | 23,1 |
| RuvBL1 | RuvBL2 | K389 | K427 | DIII | DIII | 35,1 |
| RuvBL1 | RuvBL2 | K435 | K160 | DIII | DII | 64,5 |
| RuvBL1 | RuvBL2 | K458 | K427 | DIII | DIII | 36,8 |
| RuvBL1 | RuvBL2 | K466 | K427 | DIII | DIII | N |
| RuvBL1 | RuvBL2 | K466 | K444 | DIII | DIII | N |
| RuvBL1 | RuvBL2 | K469 | K427 | DIII | DIII | N |

Tableau 1 : Tableau des différents cross-links inter-R1/R2 identifiés suite à l'analyse du complexe impliquant de R1 et R2. Les protéines, sites et domaines impliqués sont explicités ainsi que les distances Ca-Ca mesurées sur la structure de R1R2 (PDB : 7OLE³⁵⁰).

Du point de vue des peptides inter-R1/DPCD, seulement 3 cross-links, parmi les 11, ont une distance inférieure à 30 Å (27,5 Å pour K18-K181 ; 20,2 Å pour K28-K294 ; 25,4 Å pour K204-K287) et aucun des peptides inter-R2/DPCD ne correspond aux distances mesurées sur le modèle de R2D (Fig. 8, Annexe 47). La position de DPCD dans le modèle est donc probablement inexacte, en accord avec une moins bonne résolution dans ces zones et une structure probablement non optimale. Elle pourra être précisée grâce aux contraintes de distances données par le XL-MS.

En conclusion, les expériences de M-Ph et de nMS ont permis de confirmer l'existence du complexe 1:3 entre l'hexamère R1R2 et DPCD. La conformation ouverte des domaines DII des RUVBLs a été démontrée comme plus probable par la mesure de distance des résidus pontés sur le modèle du complexe R2D. Tandis que, les résultats obtenus en XL-MS ont aussi montré que la position de DPCD par rapport à l'hexamère R1R2 est encore à ajuster.

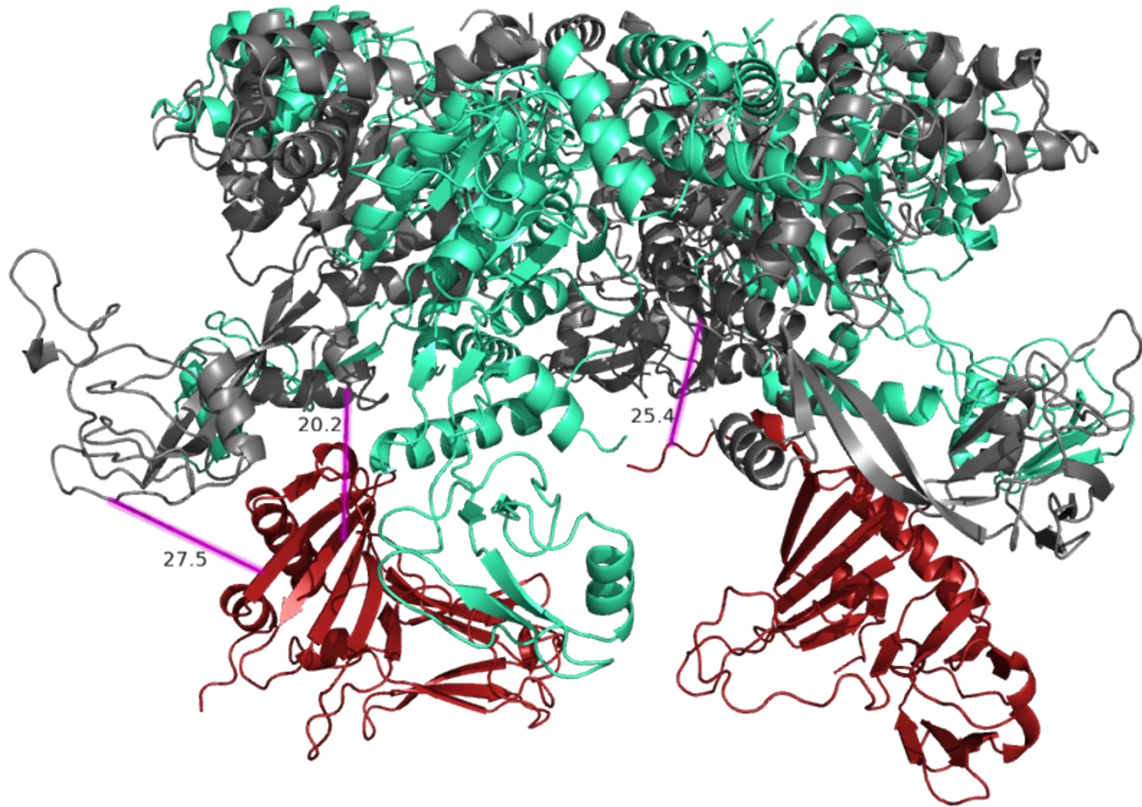


Figure 8. Représentation des distances mesurées comprises dans l'intervalle (6-30Å) correspondant aux cross-links inter-R1/DPCD (violet) sur la structure du complexe R2D. DPCD, R1 et R2 sont colorés en rouge, gris et cyan respectivement.

2.4. Valorisation

Une partie des résultats présentés en amont a pu être valorisée dans la publication suivante dans laquelle je suis co-première auteure (Annexe 48):

Dos Santos Morais*, R.; Santo, P. E.*; Ley, M.*; et al., **Deciphering cellular and molecular determinants of human DPCD protein in complex with RUVBL1/RUVBL2 AAA-ATPases.** *Journal of Molecular Biology* **2022**, 434, 167760.

Et dans une communication par affiche :

Structural Mass Spectrometry approaches to decipher interactions between DPCD and RUVBL1/RUVBL2, a 380 kDa ribonucleoprotein subcomplex. Marie Ley, Evolène Deslignière, Paulo E. Santo, Steve Hessmann, Cédric Schelcher, Marie-Eve Chagot, Ana C.F. Paiva, Pedro M. F. Sousa, Bruno Charpentier, Tiago M. Bandejas, Xavier Manival, Raphael D. S. Morais and Sarah Cianférani. ASMS du 31 octobre au 4 novembre 2021

3. Caractérisation du complexe de type R2SP

Une étude similaire à celle réalisée sur le complexe R2D a été effectuée avec le complexe R1R2/SPAG1₆₂₂₋₉₂₆/PIH1D2 (ou R2SP) afin de déterminer la stœchiométrie et les zones de contacts des différents partenaires au sein du complexe.

3.1. M-Ph et nMS

R2SP (Annexe 2) a été analysé en M-Ph permettant d'identifier plusieurs complexes (Fig 9A). Tout d'abord, les monomères de R1 ou R2 (40 ± 9 kDa) ont été révélés en plus forte abondance puis l'hétérodimère SPAG1₆₂₂₋₉₂₆/PIH1D2 (70 ± 8 kDa, aussi noté SP). Trois distributions ont été observées correspondant à l'hexamère R1R2 (319 ± 8 kDa) et aux complexes de l'hexamère R1R2 et de l'hétérodimère SP de stœchiométries (hexamère/hétérodimère) 1:1 (387 ± 8 kDa) et 1:2 (460 ± 12 kDa). Le dodécamère de R1R2 a aussi été identifié (631 ± 8 kDa) en plus faible abondance.

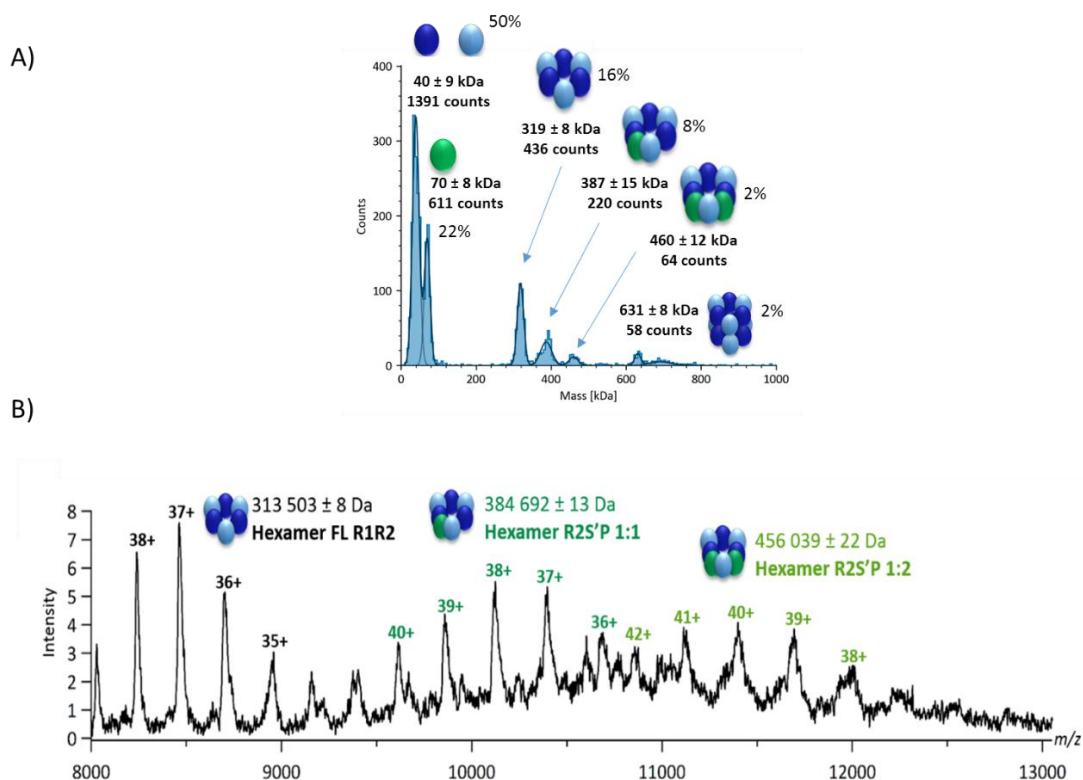


Figure 9. (A) Histogramme de distribution de contraste obtenue en M-Ph et (B) spectre de masse native du complexe R2SP.

La nMS (Fig. 9B) a confirmé la présence des complexes R2SP de différentes stœchiométries ($384\,692 \pm 13$ Da pour le complexe 1:1 ; $456\,039 \pm 22$ Da pour le complexe 1:2 avec 6-7 ADP) comme précédemment montré par la M-Ph.

3.2. XL-MS de l'hétérodimère SP

Le cross-link du complexe SP a d'abord été réalisé avec pour objectif de confirmer l'interaction des deux partenaires et les régions en interaction. Ainsi, 27 peptides pontés intra-SPAG1, 10 intra-PIH1D2 et 21 inter-SPAG1/PH1D2 ont pu être identifiés (Annexe 49), dont :

- 7 peptides inter-S/P liés au domaine TPR de SPAG1 connu comme région d'interaction avec les HSP70/90.
- 7 cross-links inter-S/P concernant les résidus K741 et K743 en aval du domaine TPR3 de SPAG1, une région également impliqué dans la liaison de PIH1D2³⁵⁹ (Fig. 10).
- 1 cross-links inter-S/P du domaine d'hétérodimérisation (CS) de PIH1D2
- 11 peptides inter-S/P sont situés au niveau de la poche de liaison phospho-peptidique (PIH).

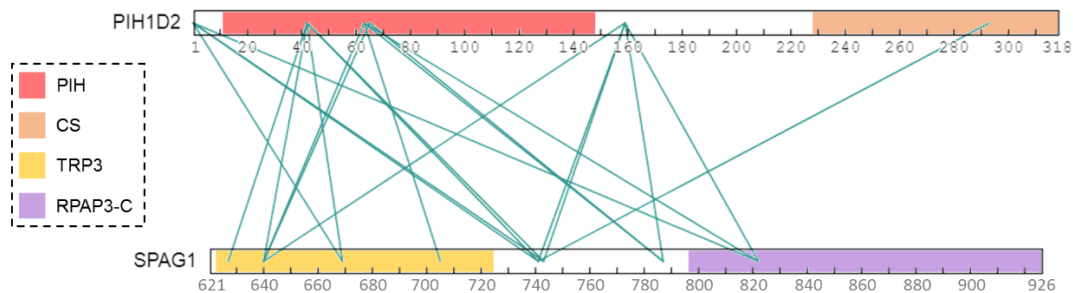


Figure 10 : Représentation schématique des peptides pontés inter-PIH1D2/SPAG1 (vert) du complexe SP (obtenus avec Xinet²¹⁶).

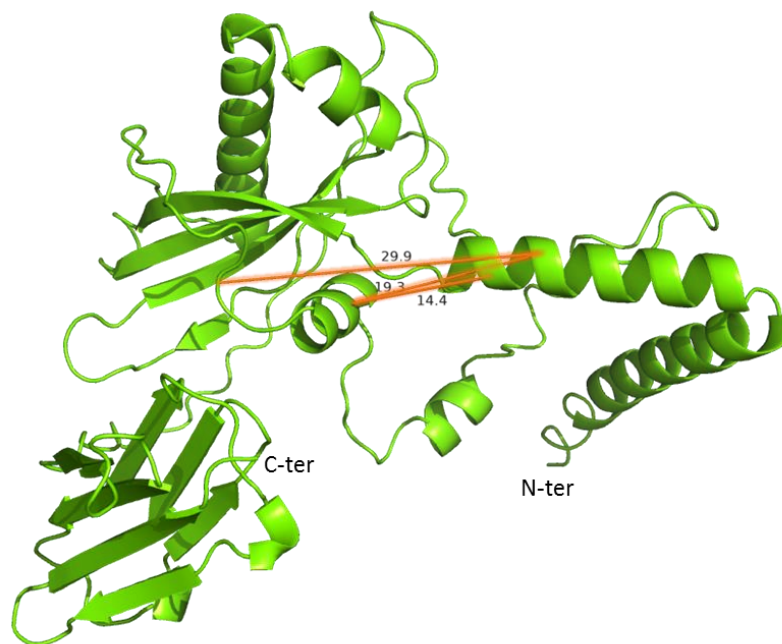


Figure 11 : Représentation des distances mesurées comprises dans l'intervalle (6-30Å) correspondant aux cross-links intra-PIH1D2 (orange) sur la structure de PIH1D2 obtenue avec AlphaFold³⁴³.

Encore aucune structure 3D PIH1D2 et du complexe avec SPAG1 n'existe à l'heure actuelle. Sachant que PIH1D1 et PIH1D2 sont des protéines homologues³⁵⁹ et qu'il existe des structures 3D de PIH1D1³⁵⁶⁻³⁵⁷, j'ai réalisé la modélisation de la structure de PIH1D2 avec Alphafold³⁴³ (Annexe 50). Toutefois, seules 3 des 10 distances mesurables correspondant aux peptides pontés intra-PIH1D2 sont dans la gamme de distances mesurables par XL-MS. Les 7 autres distances concernent le résidu N-terminal de PIH1D2 indiquant que cette région est probablement flexible (Fig. 11).

Le domaine TPR3 de SPAG1 a, quant à lui, été résolu par nos collaborateurs (PDB : 6I57³⁶⁴) nous donnant la possibilité de mesurer la distance $\text{C}\alpha\text{-C}\alpha$ entre les résidus pontés. 11 des 12 distances mesurables entre dans la gamme allant de 6 à 30Å (Fig. 12, Annexe 49).

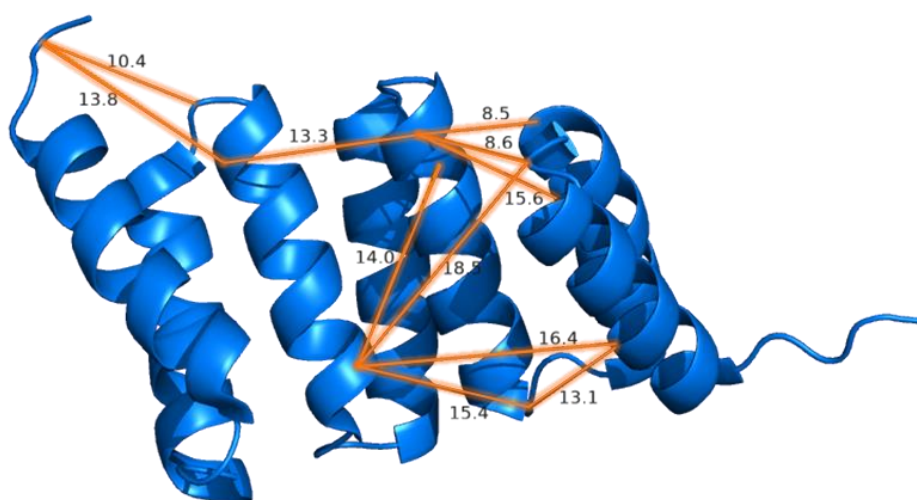
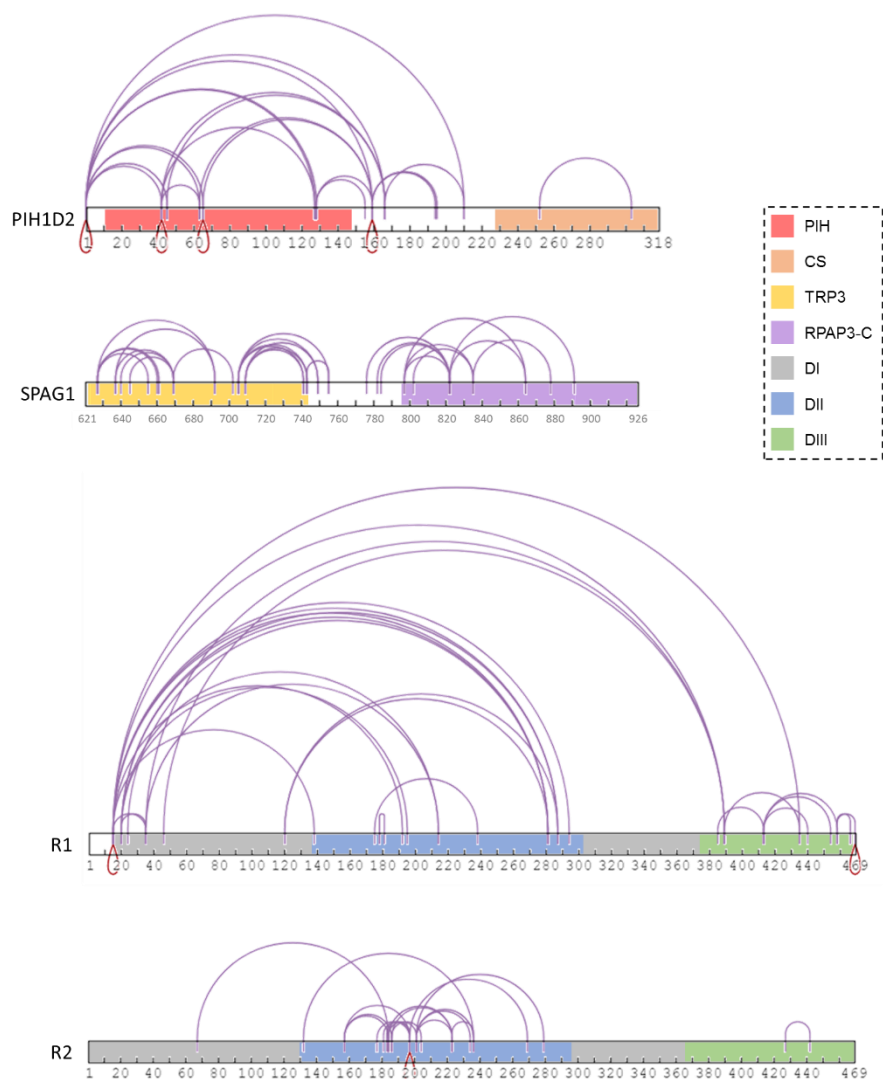


Figure 12 : Représentation des distances mesurées comprises dans l'intervalle (6-30Å) correspondant aux cross-links intra-SPAG1 (orange) sur la structure de SPAG1 (PDB : 6I57³⁶⁴).

3.3. XL-MS du complexe quaternaire R2SP

Un grand nombre de cross-links (208) a été obtenu par l'analyse du complexe R2SP, avec de nombreux peptides pontés intra-protéines (Fig. 13, Annexe 51) :

- 24 intra-PIH1D2 dont 16 au niveau de la poche de liaison PIH ce qui est en accord avec l'analyse XL-MS de l'hétérodimère SP seul,
- 26 intra-SPAG1 avec 15 situés sur les trois TRP et 10 du côté C-terminal du domaine RPAP3-C,
- 30 intra-R1 avec 10 cross-links inter-domaines et 19 intra-R2 dont seulement 2 inter-domaines, indiquant une conformation fermée plus probable des domaines DII des RUVBLs par la mesure des distances entre les résidus des structures 7OLE³⁵⁰ (82%) et 5OAF²⁴⁰ (84%).



621

Figure 13 : Représentation schématique des peptides pontés intra-protéines (violet) identifiés suite au XL-MS du complexe R2SP (obtenus avec Xinet²¹⁶).

Outre les peptides intra-protéines identifiés, des cross-links inter-protéines ont aussi permis de révéler les domaines d'interactions des différentes protéines entre elles (Fig. 14), tels que :

- 26 inter-PIH1D2/R1 comprenant 10 peptides au niveau de PIH et 11 concernant le DII des RUVBLs,
- 14 inter-PIH1D2/R2 majoritairement identifiés sur le domaine DII des RUVBLs (12), le principal domaine de liaison avec d'autres protéines,
- 17 inter-PIH1D2/SPAG1 provenant principalement du domaine CS de PIH1D2 et du TRP3 de SPAG1 contrairement à l'analyse de l'hétérodimère SP prouvant que le domaine CS de PIH1D2 est fortement impliqué dans la formation de R2SP,
- 25 inter-R2/R1 impliquant tous les domaines permettant aussi de caractériser la configuration des domaines DII qui serait cette fois-ci dans une configuration fermée,

- 19 inter-R1/SPAG1 qui incluent également l'ensemble des domaines de R1,
- 8 inter-R2/SPAG1 avec 7 peptides pontés entre le domaine DIII et SPAG1 indiquant la proximité de cette protéine avec ce domaine.

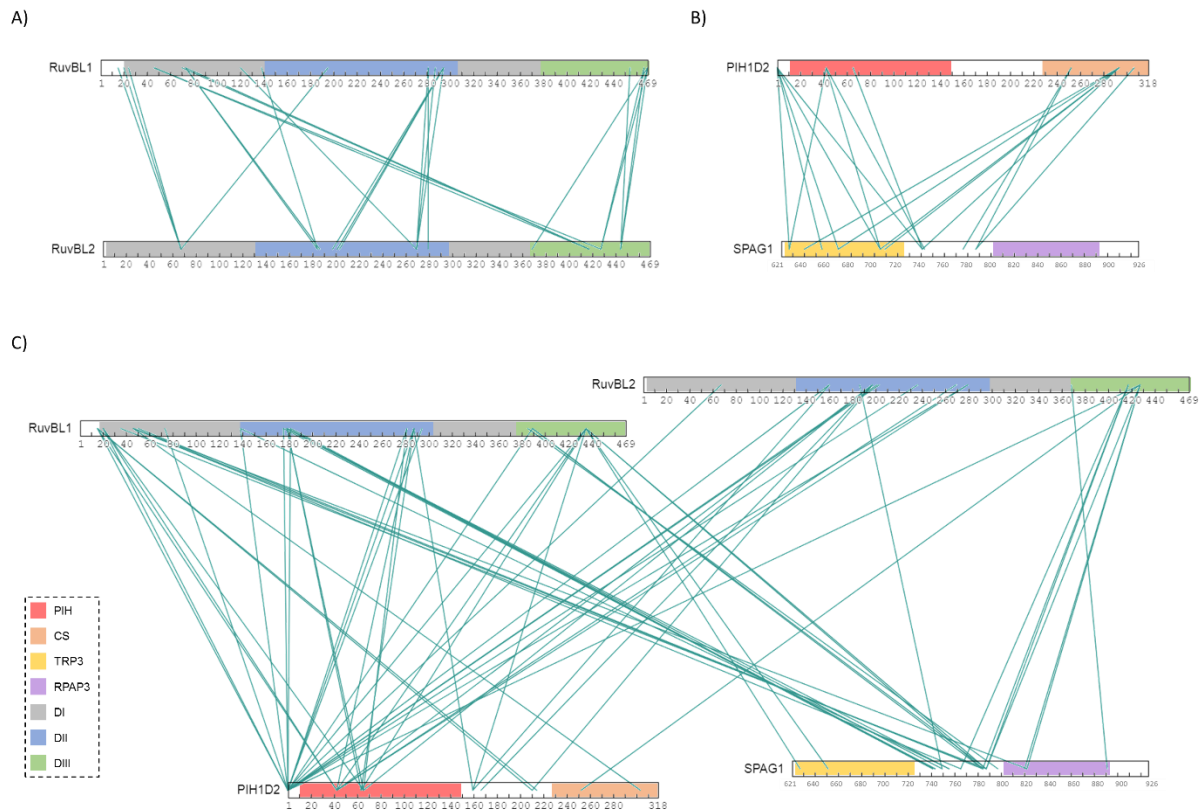


Figure 14. Représentation schématique des peptides pontés (vert) inter-R1/R2 (A), inter-SPAG1/PIH1D2 (B) ainsi que les inter-SPAG1/RUVBLs et inter-PIH1D2/RUVBLs (C) identifiés suite au XL-MS du complexe R2SP (obtenus avec Xinet²¹⁶).

De manière générale, 46% des cross-links intra-SPAG1, intra-PIH1D2 et inter-PIH1D2/SPAG1 sont communs à ceux identifiés pour l'analyse du complexe SP. La totalité des distances mesurées pour les intra-SPAG1 sur la structure du TPR3 (PDB : 6I57) a été validée apportant une confiance quant à ces résultats. Concernant les RUVBLs, 49% des intra-R1, intra-R2 et inter-R2/R1 sont similaires à l'analyse de l'hexamère R1R2. Les distances C α -C α mesurées ont permis de valider 73% et 84% des peptides pontés pour les structures 7OLE et 5OAF respectivement, suggérant que les domaines DII de R1R2 soient dans une configuration fermée.

Finalement, une structure 3D de l'hexamère R1R2 en interaction avec la partie C-terminale de SPAG1 (134 résidus) a été obtenue par nos collaborateurs en Cryo-EM avec le domaine DII des RUVBLs tronqué ne permettant pas de définir leurs conformations (Fig. 15). Toutes les distances mesurables (33) sur cette structure sont comprises dans l'intervalle allant de 6 à 30

À de l'expérience de XL-MS, montrant la concordance entre les deux types d'analyses. En accord avec la proximité supposée précédemment entre le domaine DIII des RUVBLs et RPAP3-C de SPAG1, la structure a montré l'interaction de ces domaines confortée par les distances (20,9 ; 21,5 ; 17,4 ; 21,4) mesurées des pontages impliquant ces deux domaines (Fig. 15, Annexe 51).

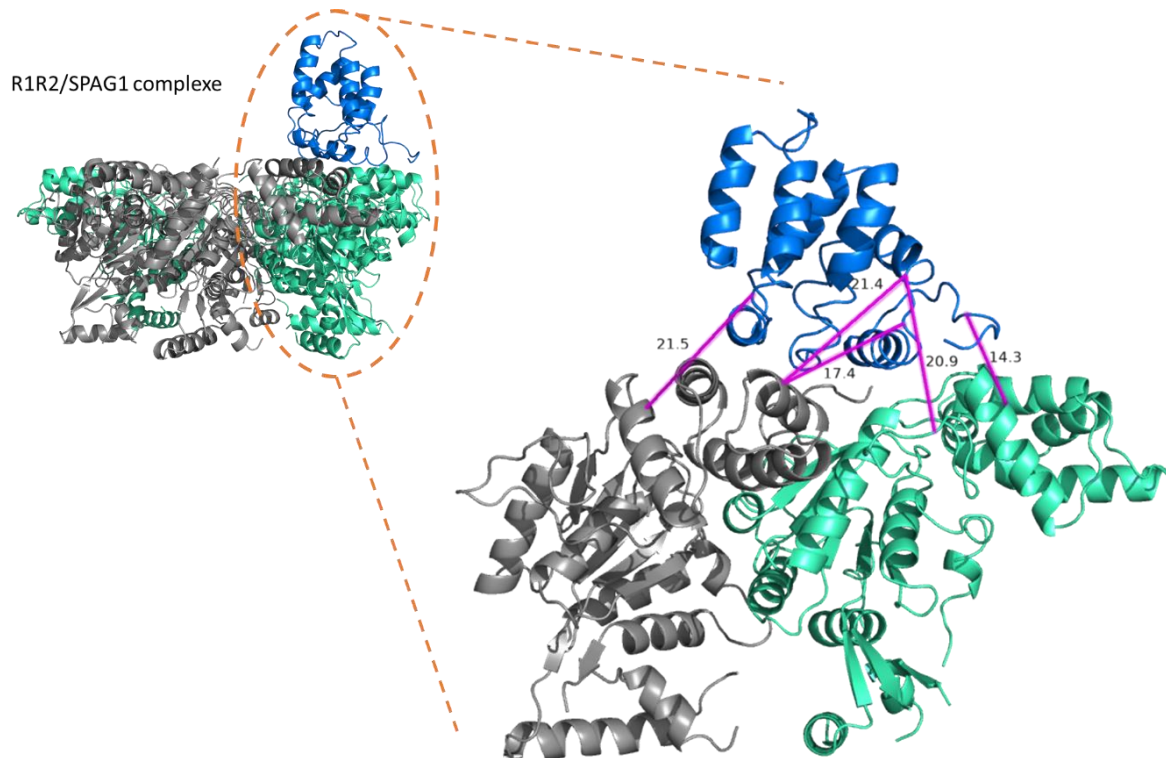


Figure 15. Représentation du complexe et des cross-links inter-R1/R2/S₆₂₂₋₉₂₆ (violet) du complexe R2SP avec R1 (gris), R2 (cyan) et SPAG1₇₉₂₋₉₂₆ (bleu).

Le grand nombre de peptides pontés identifiés suite au XL-MS du complexe binaire et quaternaire impliquant R1R2, SPAG1 et PIH1D2 nous ont apporté des premières informations structurales sur le complexe R2SP. Ils ont aussi permis de démontrer le domaine d'interaction de PIH1D2 par la proximité de la poche liaison phospho-peptidique de PIH1D2 avec les autres protéines du complexe. La concordance des résultats de XL-MS du complexe R2SP et de cryo-EM du complexe R2S a pu être mise en avant par la proximité du domaine DIII avec SPAG1. Il est également possible de supposer, grâce à l'identification des peptides pontés, (i) que les domaines DII des RUVBLs sont dans une configuration « fermée » ; et (ii) que PIH1D2 est proche spatialement des DII des RUVBLs avec 23/40 cross-links inter-RUVBLs/PIH1D2 impliquant ce domaine.

V. Conclusions et perspectives

Dans le cadre de ce projet, les différents complexes impliquant les RUVBLs ont été étudiés par l'utilisation des approches de nMS, M-Ph et XL-MS permettant ainsi la caractérisation structurale de l'interaction de nouveaux partenaires de R1R2.

Les expériences de M-Ph ont mis en évidence la pertinence de cette technique en complémentarité de la nMS. En effet, les résultats ont permis de montrer les abondances relatives des différents oligomères au sein des échantillons avec la présence de monomères, hexamères (majoritaire) et dodécamères. De plus, la faible concentration nécessaire à l'analyse (de l'ordre du nM) et la rapidité de la mise en œuvre font de cette approche, une méthode de choix à utiliser en première ligne malgré sa résolution en masse plus faible que la nMS.

Les résultats de XL-MS des différents complexes ont permis de mettre en évidence expérimentalement les conformations des domaines DII (ouvertes ou fermées) des RUVBLs et l'état oligomérique de ces protéines (hexamérique ou dodécamérique), mais aussi les régions d'interaction des différents partenaires.

En perspectives, ces résultats pourront encore permettre de confirmer/réfuter des modèles ou structures de haute résolution des différents protéines/complexes étudiés. Les pontages nous donnent aussi la possibilité de modéliser les complexes en imposant des contraintes de distances avec par exemple le logiciel DisVis³⁷². Le « rigid docking » peut être effectué à partir de données SAXS et XL-MS tout en prenant en compte la présence d'états conformationnels multiples et la distinction des peptides intra- et inter-protéines³⁷³. L'utilisation d'agents pontants plus courts pourrait aider la modélisation grâce à la réduction de l'intervalle de distance bien que la flexibilité de la chaîne carbonée des lysines est toujours à prendre en considération. En dernier lieu, ces protocoles pourront être employés pour réaliser d'autres pontages impliquant les RUVBLs et leurs divers partenaires pour gagner en précision à propos de leurs différentes structures.

Conclusions générales

Mes travaux de thèse se sont focalisés sur le développement, l'optimisation et la combinaison d'approches de MS structurale et plus particulièrement les techniques de HDX-MS et XL-MS pour la caractérisation de différents complexes protéiques. Dans un premier temps, je me suis concentrée sur la description des développements méthodologiques et des optimisations réalisées pour toutes les phases des deux approches HDX-MS et XL-MS en tenant compte des différentes recommandations^{66, 74, 227}. Dans un second temps, j'ai pu caractériser différents types de complexes impliquant des protéines telles que les RNs, les PMs, les PRMTs et les complexes d'assemblages avec des protéines AAA+. En plus, des deux techniques sus-mentionnées, des analyses de nMS et/ou M-Ph ont été réalisées dans le but d'obtenir des informations complémentaires de stœchiométries ou de taux de complexations.

Mes travaux de thèse ont principalement contribué à faire évoluer des méthodologies analytiques de MS structurale pour la caractérisation de complexes biologiques. Les principaux résultats méthodologiques concernent :

- **L'apport de la mobilité ionique en HDX-MS.** Mon travail a surtout permis d'évaluer la contribution de l'IM en HDX-MS. J'ai pu réaliser une étude complète de l'apport de la mobilité ionique en HDX-MS, grâce aux différents types de protéines étudiées allant du récepteur nucléaire à la protéine membranaire. La séparation supplémentaire via l'utilisation de l'IM a permis, non seulement d'augmenter considérablement le nombre de peptides identifiés et par la même occasion la couverture de séquence et la redondance, mais aussi de faciliter le traitement des données de HDX-MS. En effet, la séparation des profils isotopiques est largement améliorée, évitant l'élimination de certains peptides à cause d'assignations erronées (chevauchement de distributions isotopiques). Dans un souci d'amélioration, les paramètres de hauteur et vitesse de vague, au sein de la cellule de IM, pourraient être encore optimisés, mais restent en accord avec les paramètres utilisés par d'autres groupes¹⁶⁶. Parmi les points non abordés lors de mon travail, reste la question de l'impact de l'IM sur l'échange inverse D/H qui pourrait être évalué⁶⁶ par exemple sur une protéine modèle telle que la phosphorylase B.

- **L'importance des outils de traitement et de validation statistique des données HDX-MS.** Le traitement des données inclue une première phase d'assignation/reconnaissance « automatique » des profils isotopiques qui détermine ensuite les peptides à considérer dans l'étude. Malheureusement, actuellement, cette assignation « automatique » doit être corrigée manuellement pour éviter la prise en compte de mauvaises assignations peptidiques, et reste donc assez subjective et dépendante de l'utilisateur. Cette correction est tout à fait possible avec le logiciel DynamX (Waters) alors que son concurrent HDEaminer (Sierra Analytics) largement utilisé dans la communauté, ne permet pas la correction manuelle des assignations. Une harmonisation ainsi qu'une amélioration des logiciels de reconnaissance des profils isotopiques reste toujours souhaitable dans le but d'éliminer l'intervention de l'expérimentateur pour la validation des données. Après le traitement des données, la validation statistique des données HDX-MS qui permet de faire ressortir les différences significatives entre deux conditions, est également perfectible. Tout au long de mon travail de thèse, de nouveaux outils pour la validation statistique des données de HDX-MS ont été développés (MEMHDX¹³⁴, Deuterios¹³⁵⁻¹³⁶ et *Moderated*¹⁷⁰), ce qui m'a permis de manière constante et continue de les évaluer dans le cadre de mes projets en cours et de montrer ainsi leurs avantages et inconvénients. De mon point de vue, il n'existe pas encore « d'outil de validation idéal ». Certains logiciels (MEMHDX et *Moderated*) tiennent compte de la répétabilité (prise en compte des réplicats techniques) alors que d'autres (Deuterios) ne les considèrent pas. Cependant, les logiciels Deuterios et *Moderated* rendent compte des différences significatives par temps de deutération alors que MEMHDX non. L'utilisation de ces outils de validation statistique est donc à adapter en fonction des systèmes biologiques étudiés, dépendant de la stabilité des protéines, de leurs dynamiques et des différences fines d'incorporation en D entre les deux états. Finalement, j'ai pu optimiser chacune des phases de la technique de HDX-MS et évaluer la pertinence des tests statistiques appliqués aux données sur différentes protéines (modèle ou issus de projets collaboratifs) dans le but d'effectuer des expériences optimales et spécifiques à chaque projet biologique étudié.

- **Le développement d'un « test de performance » pour évaluer l'ensemble de la stratégie XL-MS.** Au début de ma thèse en 2019, l'approche XL-MS n'était utilisée que de manière ponctuelle, non systématique et sans test de performance au LSMBO. Durant ma deuxième année de thèse, j'ai mis en place un protocole de cross-link tenant compte des

recommandations parues en 2019 pour la réalisation d'expériences HDX-MS⁶⁶. Je me suis d'abord attelée à mettre en place un « test de performance » en utilisant la BSA (66 kDa) qui permet, non seulement de valider l'étape de pontage chimique elle-même (SDS-PAGE) mais aussi la méthode LC-MS/MS et l'identification des peptides pontés (> 80 pour la BSA) avec deux spectromètres de masse (Q Exactive HF-X Orbitrap et Q-Exactive Plus Orbitrap, Thermo Fisher Scientific). Les distances C α -C α mesurées sur la structure de la BSA ont été comparées aux peptides identifiés, ce qui a permis de valider que 71% des peptides identifiés sont dans la gamme des distances C α -C α ciblée. Néanmoins, de mon point de vue, la BSA n'est pas un système adapté pour l'évaluation de la stratégie XL-MS. En effet, plusieurs points sont critiquables : i) la BSA est une protéine qui n'homodimèrise qu'en faible abondance (132 kDa, 15% d'après la M-Ph), ce qui rend la détection des peptides pontés faiblement abondants encore plus difficile ; ii) la BSA est une protéine recombinante qui ne reflète pas la complexité de la plupart des échantillons biologiques (complexes multi-protéiques de plusieurs centaines de kDa, pas toujours purs, obtenus en faibles quantités) et qui peut donner des résultats positifs alors qu'aucun peptide ponté n'est identifié dans le cas du complexe d'intérêt. Il serait alors intéressant de trouver un autre modèle tel que des lysats cellulaires par exemple¹⁸², dans le but de déterminer des interactions protéine-protéine caractéristiques d'une réaction de pontage efficace. Malgré la description récente d'une autre stratégie utilisant l'addition par ajouts dosés de peptides pontés synthétiques dans l'échantillon biologique ponté qui pourrait constituer un étalon interne pour valider la partie LC-MS/MS, nous pensons que l'utilisation de standard protéiques plus complexes serait une meilleure alternative qui permettrait de valider en même temps les étapes de pontage chimique, d'analyse LC-MS/MS et de traitement bio-informatique des données.

- **L'intérêt des agents pontants MS-clivables en XL-MS et le développement d'une stratégie analytique XL-MS adaptée.** Au début de mon travail de thèse, les principaux agents pontants utilisés au laboratoire étaient des réactifs homo-bifonctionnels ciblant les Lys de types BS3 ou DSS qui présentent d'inconvénients, notamment pour la validation des peptides pontés par les outils bio-informatiques existants. Je me suis ainsi focalisée sur l'optimisation de chacune des étapes de l'approche XL-MS avec un cross-linker MS-clivable, des conditions de pontage (excès de cross-linker, pH, température et temps de réaction), à l'adaptation des

méthodes LC-MS/MS (gradient LC et mode d'acquisition) et l'évaluation de plusieurs outils d'identification des données avec MaxLynX²³⁸ et MeroX⁷².

Les deux logiciels MeroX et MaxLynX de traitement des données de XL-MS, pouvant identifier les fragments caractéristiques (cross-linker MS-clivable), ont donc été testés. Les avantages de chaque logiciel ont été montrés avec une identification plus importante de peptides pontés avec MaxLynX et des taux de distances concordantes plus importants avec MeroX traduisant d'une identification plus fiable des cross-links. Parmi les points spécifiques à MeroX, on peut citer la possibilité de vérifier « manuellement » chaque correspondance (fragments – pics) et les probabilités de sites de pontage, ce qui rend cet outil particulièrement intéressant pour la génération de sets de données plus « fiables et vérifiés ». Cependant, cette vérification peut s'avérer être un inconvénient pour l'analyse de jeux de données plus importants (complexes impliquant un grand nombre de partenaires ou XL *in situ*). Un autre point limitant de MeroX est la nécessité de transformer les données brutes en .mgf avant leur traitement via l'utilisation d'un logiciel tiers, ce qui amène potentiellement de nombreux biais si la conversion ne se fait pas correctement. Concernant MaxLynX, l'étape de conversion des données brutes en format .mgf n'est pas nécessaire, ce qui présente un avantage pour l'harmonisation des données. Cependant, les procédés d'identification des peptides pontés par le logiciel MaxLynX ne sont pas dévoilés explicitement par les développeurs, empêchant une comparaison fiable des procédés d'identification. De plus, aucune vérification des fragments caractéristiques (nombre et identification) des cross-linkers MS-clivables n'est réalisable ne permettant pas d'apporter une fiabilité identique à celle apportée par MeroX.

En conclusion, j'ai mis en place un protocole de XL-MS avec des agents pontants MS-clivables, les optimisations à réaliser lors de l'application à un nouveau projet biologique, un test de performance dans le but de traduire de l'efficacité du pontage chimique ainsi qu'une évaluation de l'identification des peptides pontés. En perspectives, au vu du faible rendement de la faible abondance (0,15% en termes d'intensité¹⁸⁷) des peptides pontés par rapport à ceux non-pontés, deux possibilités s'offrent à nous avec : (i) l'utilisation d'instruments de nouvelles générations avec le TimsTOF Pro (Bruker) avec une cellule d'IM TIMS permettant de séparer les peptides pontés (de charges plus élevées) de ceux non-pontés¹⁹⁶ ou avec l'Orbitrap Eclipse Tribid (Thermo Fisher Scientific) avec une cellule FAIMS en entrée du spectromètre de masse ; (ii) l'emploi d'un agent pontant enrichissable comme le PhoX¹⁸⁷ (Bruker) sachant

que cet enrichissement peut en plus être automatisé avec le robot de préparation AssayMAP Bravo (Agilent) rendant celui-ci plus reproductible.

- **La photométrie de masse comme alternative à la MS native ou apport de la photométrie de masse dans les projets de biologie structurale.**

Un des challenges les plus récurrents en chimie analytique est la quantité et la concentration d'échantillon disponibles pour effectuer des analyses. Dans certains cas, cette concentration n'est pas suffisante pour réaliser des expériences de nMS avec une étape de dessalage en amont des analyses.

Dans cette optique, un nouvel instrument a été mis en place au sein du laboratoire, la M-Ph (Refeyn Two^{MP}). La M-Ph permet notamment de réaliser des analyses avec des concentrations de l'ordre du pico-nano molaire. J'ai pu ainsi effectué une comparaison des analyses de nMS et M-Ph sur des complexes de haut poids moléculaire (R2D, 330 kDa) qui apportent des informations similaires de stœchiométries et d'abondances relative des différentes espèces bien que de plus faible résolution de masse (kDa en M-Ph). Cette technique est alors complémentaire et une bonne alternative à la nMS.

De mon point de vue, des analyses de M-Ph devraient être réalisées en amont de toutes autres expériences de MS structurale dans le but de déterminer la pureté, l'homogénéité des échantillons⁹², les états d'oligomérisations⁹⁶ et les affinités⁹⁷.

En perspective, ces expériences de M-Ph pourraient être adaptées au contrôle de la réaction de pontage chimique (à la place du gel SDS-PAGE) qui serait assurément plus rapide (quelques minutes) que le gel (plus d'un jour) mais aussi à l'évaluation de l'efficacité de la réaction et pour l'analyse des échantillons en conditions dénaturantes (dissociation des complexes non-pontés).

La deuxième partie de mon travail de thèse a consisté à appliquer les développements méthodologiques décrits en amont à différents projets biologiques qui ont été réalisés dans le cadre de collaborations scientifiques. Ces applications m'ont notamment permis d'approfondir les différentes optimisations ou les développements des techniques de XL-MS et HDX-MS. L'ensemble des études des complexes protéiques, décrits dans cette deuxième partie, montre l'apport de la combinaison d'approches de MS structurale pour la

caractérisation de différents types de complexes (protéine-protéine, protéine-liposomes, protéine-peptide, protéine-ligand), avec :

- **Dans un premier temps, la caractérisation de deux RNs qui ont montré la pertinence de l'emploi de plusieurs approches de MS structurale.**

En effet, la MS native et l'HDX-MS ont été utilisés pour montrer l'impact de la liaison des co-régulateur avec les mutants de PPAR γ ^{LBD} et de cette mutation ponctuelle montrant la similarité des informations obtenues. La poursuite de l'étude de ces mutations (un seul résidu muté) pourrait être effectué en HDX-MS en utilisant la stratégie « middle down » comprenant la fragmentation ETD en MS/MS et donnant des informations à l'acide aminé près.

L'avantage de la multiplicité des différentes expériences a aussi été montré par les complexes impliquant ERR α et des protéines intrinsèquement désordonnées où les résultats obtenus ont pu être cross-validés. Il serait aussi intéressant d'effectuer un pontage différentiel protéine-ADN de ERR α (en présence versus en absence de PGC1 α) dans le but d'observer l'impact de PGC1 α sur la liaison de ERR α avec l'ADN grâce à l'agent pontant DEB³⁷⁴ par exemple.

- **Dans un second temps, j'ai pu montrer l'importance de l'HDX-MS pour l'étude de protéines membranaires** avec la large possibilité d'optimisation de la technique permettant l'identification de peptides provenant des DTMs difficile à caractériser. Cette approche autorise aussi l'analyse des PMs dans des environnements de plus en plus proches des conditions physiologiques.

L'étude menée sur A_{2A} en présence de différents types de ligands (agoniste ou antagoniste) a révélé des différences d'incorporation en D significative et spécifique aux types de ligands. Ces changements conformationnels spécifiques pourraient permettre *in fine* de classifier de nouveaux ligands dont le type n'est pas encore connu. *Ces optimisations ont été intégrées à la rédaction d'une revue sur la caractérisation biophysique de A2A (Banères, J. L. ; et al. 2023, Receptors).*

L'étude de EPAC1 en HDX-MS a montré les avantages de la technique pour l'analyse de PMs dans un environnement favorable à la conservation de l'état natif des protéines (liposome, détergents) et de déterminer l'impact de la liaison de liposome et de ligands (AMPc et/ou CE3F4-R) sur EPAC1. Ces analyses ont mis en évidence la capacité d'optimisation de la digestion permettant, avant deutération, d'obtenir un pourcentage de couverture de

séquence tout à fait correct (> 84%). Néanmoins, après avoir réalisé le traitement de données, cette couverture de séquence n'est plus que de 70% montrant que les optimisations de départ doivent être plus poussées avec, par exemple, l'utilisation d'une colonne de digestion contenant deux enzymes co-immobilisées (pepsine et nepenthesine II, Affipro)¹²⁵. Ces deux enzymes ont montré leurs potentiels respectifs pour la digestion de PMs¹²⁵. *Une première publication a été acceptée dans un journal de biologie à haut impact facteur (Sartre, C. ; et al. 2023, Nature communications) et une deuxième publication est en cours de rédaction.*

- **Dans un troisième temps, les expériences effectuées en MS native, HDX-MS et XL-MS sur les PRMTs (PRMT2 et PRMT9) ont permis de déterminer les premières informations structurales sur les complexes enzymes/substrats.** Tandis que la MS native nous révèle la stœchiométrie du complexe, l'HDX-MS donne des informations conformationnelles à l'échelle du peptide et le XL-MS révèle les résidus spatialement proches.

La combinaison de ces approches a finalement mis en évidence des régions probables d'interaction entre PRMT2 et RSF1, de confirmer un modèle de la structure 3D de ce complexe et d'aller vers la compréhension des mécanismes de reconnaissance des substrats. Cette étude pourra être poursuivie par l'analyse de l'interaction entre PRMT2 et SRSF7 présent chez l'Humain et de séquences similaires à RSF1.

Les expériences en nMS et XL-MS menées sur le complexe PRMT9/SAP145₍₄₀₁₋₅₅₀₎ ont mis en place les conditions optimales pour l'analyse de ce complexe et ont révélé l'interaction de ces deux protéines ainsi que les proximités spatiales. Ces optimisations pourront être adaptées à la caractérisation du complexe ternaire PRMT9/SAP145/SAP49. Des expériences de M-Ph pourront être réalisées sur le complexe ternaire, car cette méthode ne nécessite qu'une faible quantité d'échantillons et permet de déterminer la pureté et l'homogénéité de l'échantillon, un pré-requis dans le cas d'une nouvelle étude.

Deux articles concernant ces deux projets sont en cours de préparation suite à l'obtention de nouveaux jeux de données haute résolution en cryo-EM de nos collaborateurs.

- **Dans un dernier temps, la caractérisation des complexes multimériques impliquant les RUVBLs a été réalisée grâce à la M-Ph, la MS native et le XL-MS.** Dans le cas des deux types de complexes étudiés impliquant les RUVBLs (R2D et R2SP), les expériences de nMS et M-Ph ont apporté des informations de stœchiométrie et abondances relatives des différents

oligomères. Tandis que, les analyses de XL-MS ont mis en évidence les différentes conformations des domaines DII des RUVBLs en plus des régions d'interactions avec les autres partenaires (DPCD ou SPAG1/PIH1D2).

Une première publication concernant l'étude du complexe R2D, en co-première auteure, a été acceptée (Dos Santos Morais, R. ; et al. 2022, Journal of Molecular Biology) et une deuxième publication portant sur le complexe R2SP est en cours de rédaction.

Tout au long de mes travaux de thèse, il m'a paru évident de me tenir informée des derniers développements concernant la modélisation de structures 3D et plus particulièrement de l'intégration des données de MS structurale pour la modélisation de structures. Actuellement, plusieurs logiciels permettant cette intégration des données sont disponibles, tels que Rosetta²²⁴, HADDOCK²²⁵ ou IMP²²⁶. Cependant, ces logiciels nécessitent des connaissances approfondies en modélisation et ne sont donc pas facilement accessibles aux débutants ou non-initiés.

Durant ma thèse, un nouvel outil de prédiction de la structure 3D des protéines, AlphaFold³⁴², a émergé et est très utilisé par la communauté scientifique. Plus accessible que les autres logiciels, j'ai pu l'utiliser au cours de ma thèse dans le but de prédire des structures et de les confirmer grâce aux contraintes de distances obtenues en XL-MS. Néanmoins, ces prédictions réalisées à partir des structures déjà référencées dans la PDB « Protein Data Bank » ne sont pas précises et possèdent de nombreux domaines désordonnés, car non référencés, ce qui est montré par les mesures de distances entre résidus apportées par le pontage. L'intégration directe de ces contraintes de distances au sein de AlphaFold serait alors un réel atout pour la prédiction de structures 3D. À ce jour, AlphaLink a été développé pour la prédiction de structures à partir des algorithmes AlphaFold2 et des contraintes de distances apportées par les peptides pontés par un cross-linker hétéro-bifonctionnel du type photo/acide aminé uniquement²³³. Il est donc primordial de suivre l'évolution de ces outils/logiciels, permettant d'aller toujours plus loin dans la caractérisation complète et précise de la structure des protéines.

En conclusion, ces travaux de thèse ont montré que la combinaison des approches de MS structurale est bénéfique pour la caractérisation de tous types de protéines/complexes (PRMTs, PMs, RNs et RUVBLs). De nombreuses pistes d'amélioration ont été données dans le

Conclusions générales

but de poursuivre les études concernant ces projets biologiques ainsi que de continuer d'optimiser et d'harmoniser les expériences et obtenus en HDX-MS et XL-MS.

Références bibliographiques

1. Heck, A. J., Native mass spectrometry: a bridge between interactomics and structural biology. *Nat Methods* **2008**, *5* (11), 927-33.
2. Hagel, L., Gel-filtration chromatography. *Curr Protoc Mol Biol* **2001**, *Chapter 10*, Unit 10 9.
3. Da Vela, S.; Svergun, D. I., Methods, development and applications of small-angle X-ray scattering to characterize biological macromolecules in solution. *Curr Res Struct Biol* **2020**, *2*, 164-170.
4. Matyus, L.; Szollosi, J.; Jenei, A., Steady-state fluorescence quenching applications for studying protein structure and dynamics. *J Photochem Photobiol B* **2006**, *83* (3), 223-36.
5. Miles, A. J.; Janes, R. W.; Wallace, B. A., Tools and methods for circular dichroism spectroscopy of proteins: a tutorial review. *Chem Soc Rev* **2021**, *50* (15), 8400-8413.
6. Walport, L. J.; Low, J. K. K.; Matthews, J. M.; Mackay, J. P., The characterization of protein interactions - what, how and how much? *Chem Soc Rev* **2021**, *50* (22), 12292-12307.
7. Kaur, U.; Johnson, D. T.; Chea, E. E.; Deredge, D. J.; Espino, J. A.; Jones, L. M., Evolution of Structural Biology through the Lens of Mass Spectrometry. *Anal Chem* **2019**, *91* (1), 142-155.
8. Allison, T. M.; Bechara, C., Structural mass spectrometry comes of age: new insight into protein structure, function and interactions. *Biochem Soc Trans* **2019**, *47* (1), 317-327.
9. Marcoux, J.; Cianferani, S., Towards integrative structural mass spectrometry: Benefits from hybrid approaches. *Methods* **2015**, *89*, 4-12.
10. Benesch, J. L.; Ruotolo, B. T., Mass spectrometry: come of age for structural and dynamical biology. *Curr Opin Struct Biol* **2011**, *21* (5), 641-9.
11. Katta, V.; Chait, B. T., Observation of the heme-globin complex in native myoglobin by electrospray-ionization mass spectrometry. *J Am Chem Soc* **1991**, *113* (22), 8534-8535.
12. Ganem, B.; Li, Y. T.; Henion, J. D., Detection of noncovalent receptor-ligand complexes by mass spectrometry. *Journal of the American Chemical Society* **1991**, *113* (22), 8534-8535.
13. Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M., Electrospray ionization for mass spectrometry of large biomolecules. *Science* **1989**, *246* (4926), 64-71.
14. Hall, Z.; Politis, A.; Bush, M. F.; Smith, L. J.; Robinson, C. V., Charge-state dependent compaction and dissociation of protein complexes: insights from ion mobility and molecular dynamics. *J Am Chem Soc* **2012**, *134* (7), 3429-38.
15. Sanglier, S.; Leize, E.; Van Dorsselaer, A.; Zal, F., Comparative ESI-MS study of approximately 2.2 MDa native hemocyanins from deep-sea and shore crabs: from protein oligomeric state to biotope. *J Am Soc Mass Spectrom* **2003**, *14* (5), 419-29.
16. Rostom, A. A.; Fucini, P.; Benjamin, D. R.; Juenemann, R.; Nierhaus, K. H.; Hartl, F. U.; Dobson, C. M.; Robinson, C. V., Detection and selective dissociation of intact ribosomes in a mass spectrometer. *Proc Natl Acad Sci U S A* **2000**, *97* (10), 5185-90.
17. Tamara, S.; den Boer, M. A.; Heck, A. J. R., High-Resolution Native Mass Spectrometry. *Chem Rev* **2022**, *122* (8), 7269-7326.
18. Rogniaux, H.; Van Dorsselaer, A.; Barth, P.; Biellmann, J. F.; Barbanton, J.; van Zandt, M.; Chevrier, B.; Howard, E.; Mitschler, A.; Potier, N.; Urzhumtseva, L.; Moras, D.; Podjarny, A., Binding of aldose reductase inhibitors: correlation of crystallographic and mass spectrometric studies. *J Am Soc Mass Spectrom* **1999**, *10* (7), 635-47.

19. Stojko, J.; Fieulaine, S.; Petiot-Becard, S.; Van Dorselaer, A.; Meinel, T.; Giglione, C.; Cianferani, S., Ion mobility coupled to native mass spectrometry as a relevant tool to investigate extremely small ligand-induced conformational changes. *Analyst* **2015**, *140* (21), 7234-45.
20. Holcapek, M.; Jirasko, R.; Lisa, M., Recent developments in liquid chromatography-mass spectrometry and related techniques. *J Chromatogr A* **2012**, *1259*, 3-15.
21. Leney, A. C.; Heck, A. J. R., Native Mass Spectrometry: What is in the Name? *J Am Soc Mass Spectrom* **2017**, *28* (1), 5-13.
22. Deslignière, E.; Ley, M.; Bourguet, M.; Ehkirch, A.; Botzanowski, T.; Erb, S.; Hernandez, O. A.; Cianferani, S., Pushing the limits of native MS: Online SEC-native MS for structural biology applications. *Int. J. Mass Spectrom.* **2021**, *461*.
23. Beck, J. L.; Colgrave, M. L.; Ralph, S. F.; Sheil, M. M., Electrospray ionization mass spectrometry of oligonucleotide complexes with drugs, metals, and proteins. *Mass Spectrom Rev* **2001**, *20* (2), 61-87.
24. van de Waterbeemd, M.; Tamara, S.; Fort, K. L.; Damoc, E.; Franc, V.; Bieri, P.; Itten, M.; Makarov, A.; Ban, N.; Heck, A. J. R., Dissecting ribosomal particles throughout the kingdoms of life using advanced hybrid mass spectrometry methods. *Nat Commun* **2018**, *9* (1), 2493.
25. Yang, Y.; Du, Y.; Kaltashov, I. A., The Utility of Native MS for Understanding the Mechanism of Action of Repurposed Therapeutics in COVID-19: Heparin as a Disruptor of the SARS-CoV-2 Interaction with Its Host Cell Receptor. *Anal Chem* **2020**, *92* (16), 10930-10934.
26. Yang, Y.; Ivanov, D. G.; Kaltashov, I. A., The challenge of structural heterogeneity in the native mass spectrometry studies of the SARS-CoV-2 spike protein interactions with its host cell-surface receptor. *Anal Bioanal Chem* **2021**, *413* (29), 7205-7214.
27. Snijder, J.; van de Waterbeemd, M.; Damoc, E.; Denisov, E.; Grinfeld, D.; Bennett, A.; Agbandje-McKenna, M.; Makarov, A.; Heck, A. J., Defining the stoichiometry and cargo load of viral and bacterial nanoparticles by Orbitrap mass spectrometry. *J Am Chem Soc* **2014**, *136* (20), 7295-9.
28. Karch, K. R.; Snyder, D. T.; Harvey, S. R.; Wysocki, V. H., Native Mass Spectrometry: Recent Progress and Remaining Challenges. *Annu Rev Biophys* **2022**, *51*, 157-179.
29. Shvartsburg, A. A.; Smith, R. D., Fundamentals of traveling wave ion mobility spectrometry. *Anal Chem* **2008**, *80* (24), 9689-99.
30. Ruotolo, B. T.; Benesch, J. L.; Sandercock, A. M.; Hyung, S. J.; Robinson, C. V., Ion mobility-mass spectrometry analysis of large protein complexes. *Nat Protoc* **2008**, *3* (7), 1139-52.
31. Dodds, J. N.; Baker, E. S., Ion Mobility Spectrometry: Fundamental Concepts, Instrumentation, Applications, and the Road Ahead. *J Am Soc Mass Spectrom* **2019**, *30* (11), 2185-2195.
32. Hollerbach, A. L.; Conant, C. R.; Nagy, G.; Ibrahim, Y. M., Implementation of Ion Mobility Spectrometry-Based Separations in Structures for Lossless Ion Manipulations (SLIM). *Methods Mol Biol* **2022**, *2394*, 453-469.
33. Christofi, E.; Barran, P., Ion Mobility Mass Spectrometry (IM-MS) for Structural Biology: Insights Gained by Measuring Mass, Charge, and Collision Cross Section. *Chem Rev* **2023**, *123* (6), 2902-2949.
34. Bush, M. F.; Hall, Z.; Giles, K.; Hoyes, J.; Robinson, C. V.; Ruotolo, B. T., Collision cross sections of proteins and their complexes: a calibration framework and database for gas-phase structural biology. *Anal Chem* **2010**, *82* (22), 9557-65.

35. May, J. C.; Leaptrot, K. L.; Rose, B. S.; Moser, K. L. W.; Deng, L.; Maxon, L.; DeBord, D.; McLean, J. A., Resolving Power and Collision Cross Section Measurement Accuracy of a Prototype High-Resolution Ion Mobility Platform Incorporating Structures for Lossless Ion Manipulation. *J Am Soc Mass Spectrom* **2021**, *32* (4), 1126-1137.
36. Rabuck, J. N.; Hyung, S. J.; Ko, K. S.; Fox, C. C.; Soellner, M. B.; Ruotolo, B. T., Activation state-selective kinase inhibitor assay based on ion mobility-mass spectrometry. *Anal Chem* **2013**, *85* (15), 6995-7002.
37. Botzanowski, T.; Hernandez-Alba, O.; Malissard, M.; Wagner-Rousset, E.; Desligniere, E.; Colas, O.; Haeuw, J. F.; Beck, A.; Cianferani, S., Middle Level IM-MS and CIU Experiments for Improved Therapeutic Immunoglobulin Subclass Fingerprinting. *Anal Chem* **2020**, *92* (13), 8827-8835.
38. Desligniere, E.; Ollivier, S.; Beck, A.; Ropartz, D.; Rogniaux, H.; Cianferani, S., Benefits and Limitations of High-Resolution Cyclic IM-MS for Conformational Characterization of Native Therapeutic Monoclonal Antibodies. *Anal Chem* **2023**, *95* (8), 4162-4171.
39. Donnelly, D. P.; Rawlins, C. M.; DeHart, C. J.; Fornelli, L.; Schachner, L. F.; Lin, Z.; Lippens, J. L.; Aluri, K. C.; Sarin, R.; Chen, B.; Lantz, C.; Jung, W.; Johnson, K. R.; Koller, A.; Wolff, J. J.; Campuzano, I. D. G.; Auclair, J. R.; Ivanov, A. R.; Whitelegge, J. P.; Pasa-Tolic, L.; Chamot-Rooke, J.; Danis, P. O.; Smith, L. M.; Tsybin, Y. O.; Loo, J. A.; Ge, Y.; Kelleher, N. L.; Agar, J. N., Best practices and benchmarks for intact protein analysis for top-down mass spectrometry. *Nat Methods* **2019**, *16* (7), 587-594.
40. Ramazi, S.; Zahiri, J., Post-translational modifications in proteins: resources, tools and prediction methods. *Database* **2021**, *2021*, 1-20.
41. Walsh, G.; Jefferis, R., Post-translational modifications in the context of therapeutic proteins. *Nat Biotechnol* **2006**, *24* (10), 1241-52.
42. Jensen, O. N., Modification-specific proteomics: characterization of post-translational modifications by mass spectrometry. *Curr Opin Chem Biol* **2004**, *8* (1), 33-41.
43. Habeck, T.; Lermyte, F., Seeing the complete picture: proteins in top-down mass spectrometry. *Essays Biochem* **2023**, *67* (2), 283-300.
44. Belov, M. E.; Damoc, E.; Denisov, E.; Compton, P. D.; Horning, S.; Makarov, A. A.; Kelleher, N. L., From protein complexes to subunit backbone fragments: a multi-stage approach to native mass spectrometry. *Anal Chem* **2013**, *85* (23), 11163-73.
45. Stiving, A. Q.; VanAernum, Z. L.; Busch, F.; Harvey, S. R.; Sarni, S. H.; Wysocki, V. H., Surface-Induced Dissociation: An Effective Method for Characterization of Protein Quaternary Structure. *Anal Chem* **2019**, *91* (1), 190-209.
46. Fornelli, L.; Srzentic, K.; Huguët, R.; Mullen, C.; Sharma, S.; Zabrouskov, V.; Fellers, R. T.; Durbin, K. R.; Compton, P. D.; Kelleher, N. L., Accurate Sequence Analysis of a Monoclonal Antibody by Top-Down and Middle-Down Orbitrap Mass Spectrometry Applying Multiple Ion Activation Techniques. *Anal Chem* **2018**, *90* (14), 8421-8429.
47. Zenaidee, M. A.; Wei, B.; Lantz, C.; Wu, H. T.; Lambeth, T. R.; Diedrich, J. K.; Ogorzalek Loo, R. R.; Julian, R. R.; Loo, J. A., Internal Fragments Generated from Different Top-Down Mass Spectrometry Fragmentation Methods Extend Protein Sequence Coverage. *J Am Soc Mass Spectrom* **2021**, *32* (7), 1752-1758.
48. Catherman, A. D.; Skinner, O. S.; Kelleher, N. L., Top Down proteomics: facts and perspectives. *Biochem Biophys Res Commun* **2014**, *445* (4), 683-93.
49. Brodbelt, J. S., Ion Activation Methods for Peptides and Proteins. *Anal Chem* **2016**, *88* (1), 30-51.

50. Cox, J., Prediction of peptide mass spectral libraries with machine learning. *Nat Biotechnol* **2023**, *41* (1), 33-43.
51. Brown, K. A.; Tucholski, T.; Alpert, A. J.; Eken, C.; Wesemann, L.; Kyrvasilis, A.; Jin, S.; Ge, Y., Top-Down Proteomics of Endogenous Membrane Proteins Enabled by Cloud Point Enrichment and Multidimensional Liquid Chromatography-Mass Spectrometry. *Anal Chem* **2020**, *92* (24), 15726-15735.
52. Beveridge, R.; Calabrese, A. N., Structural Proteomics Methods to Interrogate the Conformations and Dynamics of Intrinsically Disordered Proteins. *Front Chem* **2021**, *9*, 603639.
53. Melby, J. A.; Roberts, D. S.; Larson, E. J.; Brown, K. A.; Bayne, E. F.; Jin, S.; Ge, Y., Novel Strategies to Address the Challenges in Top-Down Proteomics. *J Am Soc Mass Spectrom* **2021**, *32* (6), 1278-1294.
54. Brown, K. A.; Melby, J. A.; Roberts, D. S.; Ge, Y., Top-down proteomics: challenges, innovations, and applications in basic and clinical research. *Expert Rev Proteomics* **2020**, *17* (10), 719-733.
55. Schopper, S.; Kahraman, A.; Leuenberger, P.; Feng, Y.; Piazza, I.; Muller, O.; Boersema, P. J.; Picotti, P., Measuring protein structural changes on a proteome-wide scale using limited proteolysis-coupled mass spectrometry. *Nat Protoc* **2017**, *12* (11), 2391-2410.
56. Fontana, F.; Gelain, F., Probing mechanical properties and failure mechanisms of fibrils of self-assembling peptides. *Nanoscale Adv* **2020**, *2* (1), 190-198.
57. Serpa, J. J.; Patterson, A. P.; Pan, J.; Han, J.; Wishart, D. S.; Petrotchenko, E. V.; Borchers, C. H., Using multiple structural proteomics approaches for the characterization of prion proteins. *J Proteomics* **2013**, *81*, 31-42.
58. Picotti, P.; Marabotti, A.; Negro, A.; Musi, V.; Spolaore, B.; Zambonin, M.; Fontana, A., Modulation of the structural integrity of helix F in apomyoglobin by single amino acid replacements. *Protein Sci* **2004**, *13* (6), 1572-85.
59. Holfeld, A.; Quast, J. P.; Bruderer, R.; Reiter, L.; de Souza, N.; Picotti, P., Limited Proteolysis-Mass Spectrometry to Identify Metabolite-Protein Interactions. *Methods Mol Biol* **2023**, *2554*, 69-89.
60. Petrotchenko, E. V.; Borchers, C. H., Protein Chemistry Combined with Mass Spectrometry for Protein Structure Determination. *Chem Rev* **2022**, *122* (8), 7488-7499.
61. Proctor, E. A.; Fee, L.; Tao, Y.; Redler, R. L.; Fay, J. M.; Zhang, Y.; Lv, Z.; Mercer, I. P.; Deshmukh, M.; Lyubchenko, Y. L.; Dokholyan, N. V., Nonnative SOD1 trimer is toxic to motor neurons in a model of amyotrophic lateral sclerosis. *Proc Natl Acad Sci U S A* **2016**, *113* (3), 614-9.
62. de Souza, N.; Picotti, P., Mass spectrometry analysis of the structural proteome. *Curr Opin Struct Biol* **2020**, *60*, 57-65.
63. Konermann, L.; Pan, J.; Liu, Y. H., Hydrogen exchange mass spectrometry for studying protein structure and dynamics. *Chem Soc Rev* **2011**, *40* (3), 1224-34.
64. Smith, D. L.; Deng, Y.; Zhang, Z., Probing the non-covalent structure of proteins by amide hydrogen exchange and mass spectrometry. *J Mass Spectrom* **1997**, *32* (2), 135-46.
65. Harrison, R. A.; Engen, J. R., Conformational insight into multi-protein signaling assemblies by hydrogen-deuterium exchange mass spectrometry. *Curr Opin Struct Biol* **2016**, *41*, 187-193.
66. Masson, G. R.; Burke, J. E.; Ahn, N. G.; Anand, G. S.; Borchers, C.; Brier, S.; Bou-Assaf, G. M.; Engen, J. R.; Englander, S. W.; Faber, J.; Garlish, R.; Griffin, P. R.; Gross, M. L.; Guttman, M.; Hamuro, Y.; Heck, A. J. R.; Houde, D.; Iacob, R. E.; Jorgensen, T. J. D.; Kaltashov, I. A.; Klinman, J. P.; Konermann, L.; Man, P.; Mayne, L.; Pascal, B. D.; Reichmann, D.; Skehel, M.;

- Snijder, J.; Strutzenberg, T. S.; Underbakke, E. S.; Wagner, C.; Wales, T. E.; Walters, B. T.; Weis, D. D.; Wilson, D. J.; Wintrode, P. L.; Zhang, Z.; Zheng, J.; Schriemer, D. C.; Rand, K. D., Recommendations for performing, interpreting and reporting hydrogen deuterium exchange mass spectrometry (HDX-MS) experiments. *Nat Methods* **2019**, *16* (7), 595-602.
67. Yen, H. Y.; Jazayeri, A.; Robinson, C. V., G Protein-Coupled Receptor Pharmacology-Insights from Mass Spectrometry. *Pharmacol Rev* **2023**, *75* (3), 397-415.
68. Engen, J. R.; Wales, T. E., Analytical Aspects of Hydrogen Exchange Mass Spectrometry. *Annu Rev Anal Chem (Palo Alto Calif)* **2015**, *8*, 127-48.
69. Chalmers, M. J.; Busby, S. A.; Pascal, B. D.; West, G. M.; Griffin, P. R., Differential hydrogen/deuterium exchange mass spectrometry analysis of protein-ligand interactions. *Expert Rev Proteomics* **2011**, *8* (1), 43-59.
70. Piersimoni, L.; Sinz, A., Cross-linking/mass spectrometry at the crossroads. *Anal Bioanal Chem* **2020**, *412* (24), 5981-5987.
71. Belsom, A.; Rappsilber, J., Anatomy of a crosslinker. *Curr Opin Chem Biol* **2021**, *60*, 39-46.
72. Iacobucci, C.; Gotze, M.; Ihling, C. H.; Piotrowski, C.; Arlt, C.; Schafer, M.; Hage, C.; Schmidt, R.; Sinz, A., A cross-linking/mass spectrometry workflow based on MS-cleavable cross-linkers and the MeroX software for studying protein structures and protein-protein interactions. *Nat Protoc* **2018**, *13* (12), 2864-2889.
73. Gotze, M.; Iacobucci, C.; Ihling, C. H.; Sinz, A., A Simple Cross-Linking/Mass Spectrometry Workflow for Studying System-wide Protein Interactions. *Anal Chem* **2019**, *91* (15), 10236-10244.
74. Iacobucci, C.; Piotrowski, C.; Aebersold, R.; Amaral, B. C.; Andrews, P.; Bernfur, K.; Borchers, C.; Brodie, N. I.; Bruce, J. E.; Cao, Y.; Chaignepain, S.; Chavez, J. D.; Claverol, S.; Cox, J.; Davis, T.; Degliesposti, G.; Dong, M. Q.; Edinger, N.; Emanuelsson, C.; Gay, M.; Gotze, M.; Gomes-Neto, F.; Gozzo, F. C.; Gutierrez, C.; Haupt, C.; Heck, A. J. R.; Herzog, F.; Huang, L.; Hoopmann, M. R.; Kalisman, N.; Klykov, O.; Kukacka, Z.; Liu, F.; MacCoss, M. J.; Mechtler, K.; Mesika, R.; Moritz, R. L.; Nagaraj, N.; Nesati, V.; Neves-Ferreira, A. G. C.; Ninnis, R.; Novak, P.; O'Reilly, F. J.; Pelzing, M.; Petrotchenko, E.; Piersimoni, L.; Plasencia, M.; Pukala, T.; Rand, K. D.; Rappsilber, J.; Reichmann, D.; Sailer, C.; Sarnowski, C. P.; Scheltema, R. A.; Schmidt, C.; Schriemer, D. C.; Shi, Y.; Skehel, J. M.; Slavin, M.; Sobott, F.; Solis-Mezarino, V.; Stephanowitz, H.; Stengel, F.; Stieger, C. E.; Trabjerg, E.; Trnka, M.; Vilaseca, M.; Viner, R.; Xiang, Y.; Yilmaz, S.; Zelter, A.; Ziemianowicz, D.; Leitner, A.; Sinz, A., First Community-Wide, Comparative Cross-Linking Mass Spectrometry Study. *Anal Chem* **2019**, *91* (11), 6953-6961.
75. Sinz, A., The advancement of chemical cross-linking and mass spectrometry for structural proteomics: from single proteins to protein interaction networks. *Expert Rev Proteomics* **2014**, *11* (6), 733-43.
76. Yang, B.; Tang, S.; Ma, C.; Li, S. T.; Shao, G. C.; Dang, B.; DeGrado, W. F.; Dong, M. Q.; Wang, P. G.; Ding, S.; Wang, L., Spontaneous and specific chemical cross-linking in live cells to capture and identify protein interactions. *Nat Commun* **2017**, *8* (1), 2240.
77. Sinz, A., Crosslinking Mass Spectrometry Goes In-Tissue. *Cell Syst* **2018**, *6* (1), 10-12.
78. Chavez, J. D.; Bruce, J. E., Chemical cross-linking with mass spectrometry: a tool for systems structural biology. *Curr Opin Chem Biol* **2019**, *48*, 8-18.
79. Maleknia, S. D.; Brenowitz, M.; Chance, M. R., Millisecond radiolytic modification of peptides by synchrotron X-rays identified by mass spectrometry. *Anal Chem* **1999**, *71* (18), 3965-73.

80. Maleknia, S. D.; Downard, K. M., Protein Footprinting with Radical Probe Mass Spectrometry- Two Decades of Achievement. *Protein Pept Lett* **2019**, *26* (1), 4-15.
81. Maleknia, S. D.; Downard, K. M., Advances in radical probe mass spectrometry for protein footprinting in chemical biology applications. *Chem Soc Rev* **2014**, *43* (10), 3244-58.
82. Kiselar, J. G.; Chance, M. R., Future directions of structural mass spectrometry using hydroxyl radical footprinting. *J Mass Spectrom* **2010**, *45* (12), 1373-82.
83. McKenzie-Coe, A.; Montes, N. S.; Jones, L. M., Hydroxyl Radical Protein Footprinting: A Mass Spectrometry-Based Structural Method for Studying the Higher Order Structure of Proteins. *Chem Rev* **2022**, *122* (8), 7532-7561.
84. Tullius, T. D.; Dombroski, B. A., Hydroxyl radical "footprinting": high-resolution information about DNA-protein contacts and application to lambda repressor and Cro protein. *Proc Natl Acad Sci U S A* **1986**, *83* (15), 5469-73.
85. Luchini, A.; Espina, V.; Liotta, L. A., Protein painting reveals solvent-excluded drug targets hidden within native protein-protein interfaces. *Nat Commun* **2014**, *5*, 4413.
86. Todorovic, V., Painting a picture of protein interaction. *Nat Methods* **2014**, *11* (10), 993.
87. Gunther, S.; Deredge, D.; Bowers, A. L.; Luchini, A.; Bonsor, D. A.; Beadenkopf, R.; Liotta, L.; Wintrode, P. L.; Sundberg, E. J., IL-1 Family Cytokines Use Distinct Molecular Mechanisms to Signal through Their Shared Co-receptor. *Immunity* **2017**, *47* (3), 510-523 e4.
88. Haymond, A.; Dey, D.; Carter, R.; Dailing, A.; Nara, V.; Nara, P.; Venkatayogi, S.; Paige, M.; Liotta, L.; Luchini, A., Protein painting, an optimized MS-based technique, reveals functionally relevant interfaces of the PD-1/PD-L1 complex and the YAP2/ZO-1 complex. *J Biol Chem* **2019**, *294* (29), 11180-11198.
89. Morikawa, M.; Jerath, N. U.; Ogawa, T.; Morikawa, M.; Tanaka, Y.; Shy, M. E.; Zuchner, S.; Hirokawa, N., A neuropathy-associated kinesin KIF1A mutation hyper-stabilizes the motor-neck interaction during the ATPase cycle. *EMBO J* **2022**, *41* (5), e108899.
90. Young, G.; Hundt, N.; Cole, D.; Fineberg, A.; Andrecka, J.; Tyler, A.; Olerinyova, A.; Ansari, A.; Marklund, E. G.; Collier, M. P.; Chandler, S. A.; Tkachenko, O.; Allen, J.; Crispin, M.; Billington, N.; Takagi, Y.; Sellers, J. R.; Eichmann, C.; Selenko, P.; Frey, L.; Riek, R.; Galpin, M. R.; Struwe, W. B.; Benesch, J. L. P.; Kukura, P., Quantitative mass imaging of single biological macromolecules. *Science* **2018**, *360* (6387), 423-427.
91. Paul, S. S.; Lyons, A.; Kirchner, R.; Woodside, M. T., Quantifying Oligomer Populations in Real Time during Protein Aggregation Using Single-Molecule Mass Photometry. *ACS Nano* **2022**, *16* (10), 16462-16470.
92. Sonn-Segev, A.; Belacic, K.; Bodrug, T.; Young, G.; VanderLinden, R. T.; Schulman, B. A.; Schimpf, J.; Friedrich, T.; Dip, P. V.; Schwartz, T. U.; Bauer, B.; Peters, J. M.; Struwe, W. B.; Benesch, J. L. P.; Brown, N. G.; Haselbach, D.; Kukura, P., Quantifying the heterogeneity of macromolecular machines by mass photometry. *Nat Commun* **2020**, *11* (1), 1772.
93. Malay, A. D.; Miyazaki, N.; Biela, A.; Chakraborti, S.; Majsterkiewicz, K.; Stupka, I.; Kaplan, C. S.; Kowalczyk, A.; Piette, B.; Hochberg, G. K. A.; Wu, D.; Wrobel, T. P.; Fineberg, A.; Kushwah, M. S.; Kelemen, M.; Vavpetic, P.; Pelicon, P.; Kukura, P.; Benesch, J. L. P.; Iwasaki, K.; Heddle, J. G., An ultra-stable gold-coordinated protein cage displaying reversible assembly. *Nature* **2019**, *569* (7756), 438-442.
94. Foley, E. D. B.; Kushwah, M. S.; Young, G.; Kukura, P., Mass photometry enables label-free tracking and mass measurement of single proteins on lipid bilayers. *Nat Methods* **2021**, *18* (10), 1247-1252.
95. Li, Y.; Struwe, W. B.; Kukura, P., Single molecule mass photometry of nucleic acids. *Nucleic Acids Res* **2020**, *48* (17), e97.

96. Haussermann, K.; Young, G.; Kukura, P.; Dietz, H., Dissecting FOXP2 Oligomerization and DNA Binding. *Angew Chem Int Ed Engl* **2019**, *58* (23), 7662-7667.
97. Wu, D.; Piszczek, G., Measuring the affinity of protein-protein interactions on a single-molecule level by mass photometry. *Anal Biochem* **2020**, *592*, 113575.
98. Lebedeva, M. A.; Palmieri, E.; Kukura, P.; Fletcher, S. P., Emergence and Rearrangement of Dynamic Supramolecular Aggregates Visualized by Interferometric Scattering Microscopy. *ACS Nano* **2020**, *14* (9), 11160-11168.
99. Hvidt, A.; Linderstrom-Lang, K., Exchange of hydrogen atoms in insulin with deuterium atoms in aqueous solutions. *Biochim Biophys Acta* **1954**, *14* (4), 574-5.
100. Wagner, G.; Wuthrich, K., Amide proton exchange and surface conformation of the basic pancreatic trypsin inhibitor in solution. Studies with two-dimensional nuclear magnetic resonance. *J Mol Biol* **1982**, *160* (2), 343-61.
101. Miranker, A.; Robinson, C. V.; Radford, S. E.; Dobson, C. M., Investigation of protein folding by mass spectrometry. *FASEB J* **1996**, *10* (1), 93-101.
102. Zhang, Z.; Post, C. B.; Smith, D. L., Amide hydrogen exchange determined by mass spectrometry: application to rabbit muscle aldolase. *Biochemistry* **1996**, *35* (3), 779-91.
103. Katta, V.; Chait, B. T., Conformational changes in proteins probed by hydrogen-exchange electrospray-ionization mass spectrometry. *Rapid Commun Mass Spectrom* **1991**, *5* (4), 214-7.
104. Oganessian, I.; Lento, C.; Wilson, D. J., Contemporary hydrogen deuterium exchange mass spectrometry. *Methods* **2018**, *144*, 27-42.
105. Wales, T. E.; Engen, J. R., Hydrogen exchange mass spectrometry for the analysis of protein dynamics. *Mass Spectrom Rev* **2006**, *25* (1), 158-70.
106. Engen, J. R., Analysis of protein conformation and dynamics by hydrogen/deuterium exchange MS. *Anal Chem* **2009**, *81* (19), 7870-5.
107. Engen, J. R.; Botzanowski, T.; Peterle, D.; Georgescauld, F.; Wales, T. E., Developments in Hydrogen/Deuterium Exchange Mass Spectrometry. *Anal Chem* **2021**, *93* (1), 567-582.
108. Kochert, B. A.; Jacob, R. E.; Wales, T. E.; Makriyannis, A.; Engen, J. R., Hydrogen-Deuterium Exchange Mass Spectrometry to Study Protein Complexes. *Methods Mol Biol* **2018**, *1764*, 153-171.
109. Belorossova, A. Y.; Bourguet, M.; Hessmann, S.; Chalhoub, S.; Kieffer, B.; Cianferani, S.; Rochel, N., Molecular determinants of MED1 interaction with the DNA bound VDR-RXR heterodimer. *Nucleic Acids Res* **2020**, *48* (19), 11199-11213.
110. Marciano, D. P.; Kuruvilla, D. S.; Pascal, B. D.; Griffin, P. R., Identification of Bexarotene as a PPARgamma Antagonist with HDX. *PPAR Res* **2015**, *2015*, 254560.
111. Vadas, O.; Jenkins, M. L.; Dornan, G. L.; Burke, J. E., Using Hydrogen-Deuterium Exchange Mass Spectrometry to Examine Protein-Membrane Interactions. *Methods Enzymol* **2017**, *583*, 143-172.
112. Deng, B.; Lento, C.; Wilson, D. J., Hydrogen deuterium exchange mass spectrometry in biopharmaceutical discovery and development - A review. *Anal Chim Acta* **2016**, *940*, 8-20.
113. Devarakonda, S.; Gupta, K.; Chalmers, M. J.; Hunt, J. F.; Griffin, P. R.; Van Duyne, G. D.; Spiegelman, B. M., Disorder-to-order transition underlies the structural basis for the assembly of a transcriptionally active PGC-1alpha/ERRgamma complex. *Proc Natl Acad Sci U S A* **2011**, *108* (46), 18678-83.
114. Balasubramaniam, D.; Komives, E. A., Hydrogen-exchange mass spectrometry for the study of intrinsic disorder in proteins. *Biochim Biophys Acta* **2013**, *1834* (6), 1202-9.

115. Engen, J. R.; Smith, D. L., Investigating protein structure and dynamics by hydrogen exchange MS. *Anal Chem* **2001**, *73* (9), 256A-265A.
116. Vinciauskaite, V.; Masson, G. R., Fundamentals of HDX-MS. *Essays Biochem* **2023**, *67* (2), 301-314.
117. Guttman, M.; Weis, D. D.; Engen, J. R.; Lee, K. K., Analysis of overlapped and noisy hydrogen/deuterium exchange mass spectra. *J Am Soc Mass Spectrom* **2013**, *24* (12), 1906-12.
118. Xiao, H.; Hoerner, J. K.; Eyles, S. J.; Dobo, A.; Voigtman, E.; Mel'cuk, A. I.; Kaltashov, I. A., Mapping protein energy landscapes with amide hydrogen exchange and mass spectrometry: I. A generalized model for a two-state protein and comparison with experiment. *Protein Sci* **2005**, *14* (2), 543-57.
119. Jensen, P. F.; Rand, K. D., Hydrogen Exchange: A Sensitive Analytical Window into Protein Conformation and Dynamics. *Wiley, Editor* **2016**, 1-17.
120. Kadek, A.; Mrazek, H.; Halada, P.; Rey, M.; Schriemer, D. C.; Man, P., Aspartic protease nepenthesin-1 as a tool for digestion in hydrogen/deuterium exchange mass spectrometry. *Anal Chem* **2014**, *86* (9), 4287-94.
121. Yang, M.; Hoepfner, M.; Rey, M.; Kadek, A.; Man, P.; Schriemer, D. C., Recombinant Nepenthesin II for Hydrogen/Deuterium Exchange Mass Spectrometry. *Anal Chem* **2015**, *87* (13), 6681-7.
122. Rey, M.; Man, P.; Brandolin, G.; Forest, E.; Pelosi, L., Recombinant immobilized rhizopuspepsin as a new tool for protein digestion in hydrogen/deuterium exchange mass spectrometry. *Rapid Commun Mass Spectrom* **2009**, *23* (21), 3431-8.
123. Brier, S.; Maria, G.; Carginale, V.; Capasso, A.; Wu, Y.; Taylor, R. M.; Borotto, N. B.; Capasso, C.; Engen, J. R., Purification and characterization of pepsins A1 and A2 from the Antarctic rock cod *Trematomus bernacchii*. *FEBS J* **2007**, *274* (23), 6152-66.
124. Ahn, J.; Cao, M. J.; Yu, Y. Q.; Engen, J. R., Accessing the reproducibility and specificity of pepsin and other aspartic proteases. *Biochim Biophys Acta* **2013**, *1834* (6), 1222-9.
125. Moller, I. R.; Slivacka, M.; Hausner, J.; Nielsen, A. K.; Pospisilova, E.; Merkle, P. S.; Liskova, R.; Polak, M.; Loland, C. J.; Kadek, A.; Man, P.; Rand, K. D., Improving the Sequence Coverage of Integral Membrane Proteins during Hydrogen/Deuterium Exchange Mass Spectrometry Experiments. *Anal Chem* **2019**, *91* (17), 10970-10978.
126. Claesen, J.; Burzykowski, T., Computational methods and challenges in hydrogen/deuterium exchange mass spectrometry. *Mass Spectrom Rev* **2017**, *36* (5), 649-667.
127. Perkins, D. N.; Pappin, D. J.; Creasy, D. M.; Cottrell, J. S., Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **1999**, *20* (18), 3551-67.
128. Eng, J. K.; Jahan, T. A.; Hoopmann, M. R., Comet: an open-source MS/MS sequence database search tool. *Proteomics* **2013**, *13* (1), 22-4.
129. Pascal, B. D.; Willis, S.; Lauer, J. L.; Landgraf, R. R.; West, G. M.; Marciano, D.; Novick, S.; Goswami, D.; Chalmers, M. J.; Griffin, P. R., HDX workbench: software for the analysis of H/D exchange MS data. *J Am Soc Mass Spectrom* **2012**, *23* (9), 1512-21.
130. Raval, S.; Sarpe, V.; Hepburn, M.; Crowder, D. A.; Zhang, T.; Viner, R.; Schriemer, D. C., Improving Spectral Validation Rates in Hydrogen-Deuterium Exchange Data Analysis. *Anal Chem* **2021**, *93* (9), 4246-4254.
131. Kan, Z. Y.; Ye, X.; Skinner, J. J.; Mayne, L.; Englander, S. W., ExMS2: An Integrated Solution for Hydrogen-Deuterium Exchange Mass Spectrometry Data Analysis. *Anal Chem* **2019**, *91* (11), 7474-7481.

132. Weis, D. D.; Engen, J. R.; Kass, I. J., Semi-automated data processing of hydrogen exchange mass spectra using HX-Express. *J Am Soc Mass Spectrom* **2006**, *17* (12), 1700-3.
133. Na, S.; Lee, J. J.; Joo, J. W. J.; Lee, K. J.; Paek, E., deMix: Decoding Deuterated Distributions from Heterogeneous Protein States via HDX-MS. *Sci Rep* **2019**, *9* (1), 3176.
134. Hourdel, V.; Volant, S.; O'Brien, D. P.; Chenal, A.; Chamot-Rooke, J.; Dillies, M. A.; Brier, S., MEMHDX: an interactive tool to expedite the statistical validation and visualization of large HDX-MS datasets. *Bioinformatics* **2016**, *32* (22), 3413-3419.
135. Lau, A. M.; Claesen, J.; Hansen, K.; Politis, A., Deuterios 2.0: peptide-level significance testing of data from hydrogen deuterium exchange mass spectrometry. *Bioinformatics* **2021**, *37* (2), 270-272.
136. Lau, A. M. C.; Ahdash, Z.; Martens, C.; Politis, A., Deuterios: software for rapid analysis and visualization of data from differential hydrogen deuterium exchange-mass spectrometry. *Bioinformatics* **2019**, *35* (17), 3171-3173.
137. Lumpkin, R. J.; Komives, E. A., DECA, A Comprehensive, Automatic Post-processing Program for HDX-MS Data. *Mol Cell Proteomics* **2019**, *18* (12), 2516-2523.
138. Puchala, W.; Burdukiewicz, M.; Kistowski, M.; Dabrowska, K. A.; Badaczewska-Dawid, A. E.; Cysewski, D.; Dadlez, M., HaDeX: an R package and web-server for analysis of data from hydrogen-deuterium exchange mass spectrometry experiments. *Bioinformatics* **2020**, *36* (16), 4516-4518.
139. Bouyssie, D.; Lesne, J.; Locard-Paulet, M.; Albigot, R.; Burlet-Schiltz, O.; Marcoux, J., HDX-Viewer: interactive 3D visualization of hydrogen-deuterium exchange data. *Bioinformatics* **2019**, *35* (24), 5331-5333.
140. Kavan, D.; Man, P., MSTools — Web based application for visualization and presentation of HXMS data. *Int. J. Mass Spectrom.* **2011**, *302*, 53-58.
141. Perez-Riverol, Y.; Bai, J.; Bandla, C.; Garcia-Seisdedos, D.; Hewapathirana, S.; Kamatchinathan, S.; Kundu, D. J.; Prakash, A.; Frericks-Zipper, A.; Eisenacher, M.; Walzer, M.; Wang, S.; Brazma, A.; Vizcaino, J. A., The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res* **2022**, *50* (D1), D543-D552.
142. Hamuro, Y.; Zhang, T., High-Resolution HDX-MS of Cytochrome c Using Pepsin/Fungal Protease Type XIII Mixed Bed Column. *J Am Soc Mass Spectrom* **2019**, *30* (2), 227-234.
143. Fang, M.; Wang, Z.; Cupp-Sutton, K. A.; Welborn, T.; Smith, K.; Wu, S., High-throughput hydrogen deuterium exchange mass spectrometry (HDX-MS) coupled with subzero-temperature ultrahigh pressure liquid chromatography (UPLC) separation for complex sample analysis. *Anal Chim Acta* **2021**, *1143*, 65-72.
144. Wales, T. E.; Fadgen, K. E.; Eggertson, M. J.; Engen, J. R., Subzero Celsius separations in three-zone temperature controlled hydrogen deuterium exchange mass spectrometry. *J Chromatogr A* **2017**, *1523*, 275-282.
145. Wollenberg, D. T. W.; Pengelley, S.; Mouritsen, J. C.; Suckau, D.; Jorgensen, C. I.; Jorgensen, T. J. D., Avoiding H/D Scrambling with Minimal Ion Transmission Loss for HDX-MS/MS-ETD Analysis on a High-Resolution Q-TOF Mass Spectrometer. *Anal Chem* **2020**, *92* (11), 7453-7461.
146. Landgraf, R. R.; Chalmers, M. J.; Griffin, P. R., Automated hydrogen/deuterium exchange electron transfer dissociation high resolution mass spectrometry measured at single-amide resolution. *J Am Soc Mass Spectrom* **2012**, *23* (2), 301-9.
147. Kish, M.; Smith, V.; Lethbridge, N.; Cole, L.; Bond, N. J.; Phillips, J. J., Online Fully Automated System for Hydrogen/Deuterium-Exchange Mass Spectrometry with Millisecond Time Resolution. *Anal Chem* **2023**, *95* (11), 5000-5008.

148. Walters, B. T.; Mayne, L.; Hinshaw, J. R.; Sosnick, T. R.; Englander, S. W., Folding of a large protein at high structural resolution. *Proc Natl Acad Sci U S A* **2013**, *110* (47), 18898-903.
149. Burkitt, W.; O'Connor, G., Assessment of the repeatability and reproducibility of hydrogen/deuterium exchange mass spectrometry measurements. *Rapid Commun Mass Spectrom* **2008**, *22* (23), 3893-901.
150. Liyanage, O. T.; Seneviratne, C. A.; Gallagher, E. S., Applying an Internal Standard to Improve the Repeatability of In-electrospray H/D Exchange of Carbohydrate-Metal Adducts. *J Am Soc Mass Spectrom* **2019**, *30* (8), 1368-1372.
151. Park, I. H.; Venable, J. D.; Steckler, C.; Cellitti, S. E.; Lesley, S. A.; Spraggon, G.; Brock, A., Estimation of Hydrogen-Exchange Protection Factors from MD Simulation Based on Amide Hydrogen Bonding Analysis. *J Chem Inf Model* **2015**, *55* (9), 1914-25.
152. Huang, L.; So, P. K.; Yao, Z. P., Protein dynamics revealed by hydrogen/deuterium exchange mass spectrometry: Correlation between experiments and simulation. *Rapid Commun Mass Spectrom* **2019**, *33* Suppl 3, 83-89.
153. Wales, T. E.; Fadgen, K. E.; Gerhardt, G. C.; Engen, J. R., High-speed and high-resolution UPLC separation at zero degrees Celsius. *Anal Chem* **2008**, *80* (17), 6815-20.
154. Gray, C. J.; Thomas, B.; Upton, R.; Migas, L. G.; Evers, C. E.; Barran, P. E.; Flitsch, S. L., Applications of ion mobility mass spectrometry for high throughput, high resolution glycan analysis. *Biochim Biophys Acta* **2016**, *1860* (8), 1688-709.
155. Watson, M. J.; Harkewicz, R.; Hodge, E. A.; Vorauer, C.; Palmer, J.; Lee, K. K.; Guttman, M., Simple Platform for Automating Decoupled LC-MS Analysis of Hydrogen/Deuterium Exchange Samples. *J Am Soc Mass Spectrom* **2021**, *32* (2), 597-600.
156. Iacob, R. E.; Engen, J. R., Hydrogen exchange mass spectrometry: are we out of the quicksand? *J Am Soc Mass Spectrom* **2012**, *23* (6), 1003-10.
157. Hamuro, Y.; Coales, S. J.; Morrow, J. A.; Molnar, K. S.; Tuske, S. J.; Southern, M. R.; Griffin, P. R., Hydrogen/deuterium-exchange (H/D-Ex) of PPARgamma LBD in the presence of various modulators. *Protein Sci* **2006**, *15* (8), 1883-92.
158. Houde, D.; Arndt, J.; Domeier, W.; Berkowitz, S.; Engen, J. R., Characterization of IgG1 Conformation and Conformational Dynamics by Hydrogen/Deuterium Exchange Mass Spectrometry. *Anal Chem* **2009**, *81* (14), 5966.
159. Zhang, A.; Fang, J.; Chou, R. Y.; Bondarenko, P. V.; Zhang, Z., Conformational difference in human IgG2 disulfide isoforms revealed by hydrogen/deuterium exchange mass spectrometry. *Biochemistry* **2015**, *54* (10), 1956-62.
160. Grauslund, L. R.; Calvaresi, V.; Pansegrau, W.; Norais, N.; Rand, K. D., Epitope and Paratope Mapping by HDX-MS Combined with SPR Elucidates the Difference in Bactericidal Activity of Two Anti-NadA Monoclonal Antibodies. *J Am Soc Mass Spectrom* **2021**, *32* (7), 1575-1582.
161. Gramlich, M.; Maier, S.; Kaiser, P. D.; Traenkle, B.; Wagner, T. R.; Voglmeir, J.; Stoll, D.; Rothbauer, U.; Zeck, A., A Novel PNGase Rc for Improved Protein N-Deglycosylation in Bioanalytics and Hydrogen-Deuterium Exchange Coupled With Mass Spectrometry Epitope Mapping under Challenging Conditions. *Anal Chem* **2022**, *94* (27), 9863-9871.
162. May, J. C.; McLean, J. A., The Influence of Drift Gas Composition on the Separation Mechanism in Traveling Wave Ion Mobility Spectrometry: Insight from Electrodynamic Simulations. *Int J Ion Mobil Spectrom* **2003**, *16* (2), 85-94.
163. Paoletti, F.; Covaceuszach, S.; Cassetta, A.; Calabrese, A. N.; Novak, U.; Konarev, P.; Grdadolnik, J.; Lamba, D.; Golic Grdadolnik, S., Distinct conformational changes occur within

the intrinsically unstructured pro-domain of pro-Nerve Growth Factor in the presence of ATP and Mg(2). *Protein Sci* **2023**, *32* (2), e4563.

164. Braet, S. M.; Buckley, T. S. C.; Venkatakrishnan, V.; Dam, K. A.; Bjorkman, P. J.; Anand, G. S., Timeline of changes in spike conformational dynamics in emergent SARS-CoV-2 variants reveal progressive stabilization of trimer stalk with altered NTD dynamics. *Elife* **2023**, *12*.

165. Calvaresi, V.; Wrobel, A. G.; Toporowska, J.; Hammerschmid, D.; Doores, K. J.; Bradshaw, R. T.; Parsons, R. B.; Benton, D. J.; Roustan, C.; Reading, E.; Malim, M. H.; Gamblin, S. J.; Politis, A., Structural dynamics in the evolution of SARS-CoV-2 spike glycoprotein. *Nat Commun* **2023**, *14* (1), 1421.

166. Cryar, A.; Groves, K.; Quaglia, M., Online Hydrogen-Deuterium Exchange Traveling Wave Ion Mobility Mass Spectrometry (HDX-IM-MS): a Systematic Evaluation. *J Am Soc Mass Spectrom* **2017**, *28* (6), 1192-1202.

167. Leger, C.; Pitard, I.; Sadi, M.; Carvalho, N.; Brier, S.; Mechaly, A.; Raoux-Barbot, D.; Davi, M.; Hoos, S.; Weber, P.; Vachette, P.; Durand, D.; Haouz, A.; Guijarro, J. I.; Ladant, D.; Chenal, A., Dynamics and structural changes of calmodulin upon interaction with the antagonist calmidazolium. *BMC Biol* **2022**, *20* (1), 176.

168. Khan, S. H.; Braet, S. M.; Koehler, S. J.; Elacqua, E.; Anand, G. S.; Okafor, C. D., Ligand-induced shifts in conformational ensembles that describe transcriptional activation. *Elife* **2022**, *11*.

169. Javed, W.; Vallet, S.; Clement, M. P.; Le Roy, A.; Moulin, M.; Hartlein, M.; Breyton, C.; Burette-Schiltz, O.; Marcoux, J.; Orelle, C.; Ebel, C.; Martel, A.; Jault, J. M., Structural Insights into the Catalytic Cycle of a Bacterial Multidrug ABC Efflux Pump. *J Mol Biol* **2022**, *434* (9), 167541.

170. Claesen, J.; Krishnamurthy, S.; Lau, A. M.; Economou, A., Moderated Test Statistics to Detect Differential Deuteration in Hydrogen/Deuterium Exchange Mass Spectrometry Experiments. *Anal Chem* **2021**, *93* (49), 16341-16349.

171. Cura, V.; Marechal, N.; Troffer-Charlier, N.; Strub, J. M.; van Haren, M. J.; Martin, N. I.; Cianferani, S.; Bonnefond, L.; Cavarelli, J., Structural studies of protein arginine methyltransferase 2 reveal its interactions with potential substrates and inhibitors. *FEBS J* **2017**, *284* (1), 77-96.

172. Hageman, T. S.; Weis, D. D., Reliable Identification of Significant Differences in Differential Hydrogen Exchange-Mass Spectrometry Measurements Using a Hybrid Significance Testing Approach. *Anal Chem* **2019**, *91* (13), 8008-8016.

173. Bar Barroeta, A.; van Galen, J.; Stroo, I.; Marquart, J. A.; Meijer, A. B.; Meijers, J. C. M., Hydrogen-deuterium exchange mass spectrometry highlights conformational changes induced by factor XI activation and binding of factor IX to factor XIa. *J Thromb Haemost* **2019**, *17* (12), 2047-2055.

174. Glasgow, A.; Hobbs, H. T.; Perry, Z. R.; Wells, M. L.; Marqusee, S.; Kortemme, T., Ligand-specific changes in conformational flexibility mediate long-range allostery in the lac repressor. *Nat Commun* **2023**, *14* (1), 1179.

175. Wales, T. E.; Eggertson, M. J.; Engen, J. R., Considerations in the analysis of hydrogen exchange mass spectrometry data. *Methods Mol Biol* **2013**, *1007*, 263-88.

176. Muller, D. R.; Schindler, P.; Towbin, H.; Wirth, U.; Voshol, H.; Hoving, S.; Steinmetz, M. O., Isotope-tagged cross-linking reagents. A new tool in mass spectrometric protein interaction analysis. *Anal Chem* **2001**, *73* (9), 1927-34.

177. Soderblom, E. J.; Goshe, M. B., Collision-induced dissociative chemical cross-linking reagents and methodology: Applications to protein structural characterization using tandem mass spectrometry analysis. *Anal Chem* **2006**, *78* (23), 8059-68.
178. Tang, X.; Munske, G. R.; Siems, W. F.; Bruce, J. E., Mass spectrometry identifiable cross-linking strategy for studying protein-protein interactions. *Anal Chem* **2005**, *77* (1), 311-8.
179. Chen, Z. A.; Jawhari, A.; Fischer, L.; Buchen, C.; Tahir, S.; Kamenski, T.; Rasmussen, M.; Lariviere, L.; Bukowski-Wills, J. C.; Nilges, M.; Cramer, P.; Rappsilber, J., Architecture of the RNA polymerase II-TFIIF complex revealed by cross-linking and mass spectrometry. *EMBO J* **2010**, *29* (4), 717-26.
180. Bursomanno, S.; Beli, P.; Khan, A. M.; Minocherhomji, S.; Wagner, S. A.; Bekker-Jensen, S.; Mailand, N.; Choudhary, C.; Hickson, I. D.; Liu, Y., Proteome-wide analysis of SUMO2 targets in response to pathological DNA replication stress in human cells. *DNA Repair (Amst)* **2015**, *25*, 84-96.
181. Schmidt, C.; Zhou, M.; Marriott, H.; Morgner, N.; Politis, A.; Robinson, C. V., Comparative cross-linking and mass spectrometry of an intact F-type ATPase suggest a role for phosphorylation. *Nat Commun* **2013**, *4*, 1985.
182. Piersimoni, L.; Kastiris, P. L.; Arlt, C.; Sinz, A., Cross-Linking Mass Spectrometry for Investigating Protein Conformations and Protein-Protein Interactions horizontal line A Method for All Seasons. *Chem Rev* **2022**, *122* (8), 7500-7531.
183. Leitner, A.; Joachimiak, L. A.; Unverdorben, P.; Walzthoeni, T.; Frydman, J.; Forster, F.; Aebersold, R., Chemical cross-linking/mass spectrometry targeting acidic residues in proteins and protein complexes. *Proc Natl Acad Sci U S A* **2014**, *111* (26), 9455-60.
184. Sutherland, B. W.; Toews, J.; Kast, J., Utility of formaldehyde cross-linking and mass spectrometry in the study of protein-protein interactions. *J Mass Spectrom* **2008**, *43* (6), 699-715.
185. Steigenberger, B.; Albanese, P.; Heck, A. J. R.; Scheltema, R. A., To Cleave or Not To Cleave in XL-MS? *J Am Soc Mass Spectrom* **2020**, *31* (2), 196-206.
186. Giese, S. H.; Fischer, L.; Rappsilber, J., A Study into the Collision-induced Dissociation (CID) Behavior of Cross-Linked Peptides. *Mol Cell Proteomics* **2016**, *15* (3), 1094-104.
187. Steigenberger, B.; Pieters, R. J.; Heck, A. J. R.; Scheltema, R. A., PhoX: An IMAC-Enrichable Cross-Linking Reagent. *ACS Cent Sci* **2019**, *5* (9), 1514-1522.
188. Nury, C.; Redeker, V.; Dautrey, S.; Romieu, A.; van der Rest, G.; Renard, P. Y.; Melki, R.; Chamot-Rooke, J., A novel bio-orthogonal cross-linker for improved protein/protein interaction analysis. *Anal Chem* **2015**, *87* (3), 1853-60.
189. Rey, M.; Dupre, M.; Lopez-Neira, I.; Duchateau, M.; Chamot-Rooke, J., eXL-MS: An Enhanced Cross-Linking Mass Spectrometry Workflow To Study Protein Complexes. *Anal Chem* **2018**, *90* (18), 10707-10714.
190. Petrotchenko, E. V.; Serpa, J. J.; Borchers, C. H., An isotopically coded CID-cleavable biotinylated cross-linker for structural proteomics. *Mol Cell Proteomics* **2011**, *10* (2), M110001420.
191. Yilmaz, S.; Shiferaw, G. A.; Rayo, J.; Economou, A.; Martens, L.; Vandermarliere, E., Cross-linked peptide identification: A computational forest of algorithms. *Mass Spectrom Rev* **2018**, *37* (6), 738-749.
192. Lu, L.; Duong, V. T.; Shalash, A. O.; Skwarczynski, M.; Toth, I., Chemical Conjugation Strategies for the Development of Protein-Based Subunit Nanovaccines. *Vaccines (Basel)* **2021**, *9* (6).

193. Fiala, G. J.; Schamel, W. W.; Blumenthal, B., Blue native polyacrylamide gel electrophoresis (BN-PAGE) for analysis of multiprotein complexes from cellular lysates. *J Vis Exp* **2011**, (48).
194. Tang, X.; Wippel, H. H.; Chavez, J. D.; Bruce, J. E., Crosslinking mass spectrometry: A link between structural biology and systems biology. *Protein Sci* **2021**, *30* (4), 773-784.
195. Kolbowski, L.; Mendes, M. L.; Rappsilber, J., Optimizing the Parameters Governing the Fragmentation of Cross-Linked Peptides in a Tribrid Mass Spectrometer. *Anal Chem* **2017**, *89* (10), 5311-5318.
196. Rinner, O.; Seebacher, J.; Walzthoeni, T.; Mueller, L. N.; Beck, M.; Schmidt, A.; Mueller, M.; Aebersold, R., Identification of cross-linked peptides from large sequence databases. *Nat Methods* **2008**, *5* (4), 315-8.
197. Yu, C.; Huang, L., Cross-Linking Mass Spectrometry: An Emerging Technology for Interactomics and Structural Biology. *Anal Chem* **2018**, *90* (1), 144-165.
198. Graziadei, A.; Rappsilber, J., Leveraging crosslinking mass spectrometry in structural and cell biology. *Structure* **2022**, *30* (1), 37-54.
199. Yang, B.; Wu, Y. J.; Zhu, M.; Fan, S. B.; Lin, J.; Zhang, K.; Li, S.; Chi, H.; Li, Y. X.; Chen, H. F.; Luo, S. K.; Ding, Y. H.; Wang, L. H.; Hao, Z.; Xiu, L. Y.; Chen, S.; Ye, K.; He, S. M.; Dong, M. Q., Identification of cross-linked peptides from complex samples. *Nat Methods* **2012**, *9* (9), 904-6.
200. Trnka, M. J.; Baker, P. R.; Robinson, P. J.; Burlingame, A. L.; Chalkley, R. J., Matching cross-linked peptide spectra: only as good as the worse identification. *Mol Cell Proteomics* **2014**, *13* (2), 420-34.
201. Hoopmann, M. R.; Zelter, A.; Johnson, R. S.; Riffle, M.; MacCoss, M. J.; Davis, T. N.; Moritz, R. L., Kojak: efficient analysis of chemically cross-linked protein complexes. *J Proteome Res* **2015**, *14* (5), 2190-8.
202. Lima, D. B.; de Lima, T. B.; Balbuena, T. S.; Neves-Ferreira, A. G. C.; Barbosa, V. C.; Gozzo, F. C.; Carvalho, P. C., SIM-XL: A powerful and user-friendly tool for peptide cross-linking analysis. *J Proteomics* **2015**, *129*, 51-55.
203. Yilmaz, S.; Drepper, F.; Hulstaert, N.; Cernic, M.; Gevaert, K.; Economou, A.; Warscheid, B.; Martens, L.; Vandermarliere, E., Xilmass: A New Approach toward the Identification of Cross-Linked Peptides. *Anal Chem* **2016**, *88* (20), 9949-9957.
204. Ji, C.; Li, S.; Reilly, J. P.; Radivojac, P.; Tang, H., XLSearch: a Probabilistic Database Search Algorithm for Identifying Cross-Linked Peptides. *J Proteome Res* **2016**, *15* (6), 1830-41.
205. Yu, F.; Li, N.; Yu, W., ECL: an exhaustive search tool for the identification of cross-linked peptides using whole database. *BMC Bioinformatics* **2016**, *17* (1), 217.
206. Gotze, M.; Pettelkau, J.; Schaks, S.; Bosse, K.; Ihling, C. H.; Krauth, F.; Fritzsche, R.; Kuhn, U.; Sinz, A., StavroX--a software for analyzing crosslinked products in protein interaction studies. *J Am Soc Mass Spectrom* **2012**, *23* (1), 76-87.
207. Chen, Z. A.; Fischer, L.; Cox, J.; Rappsilber, J., Quantitative Cross-linking/Mass Spectrometry Using Isotope-labeled Cross-linkers and MaxQuant. *Mol Cell Proteomics* **2016**, *15* (8), 2769-78.
208. Walzthoeni, T.; Joachimiak, L. A.; Rosenberger, G.; Rost, H. L.; Malmstrom, L.; Leitner, A.; Frydman, J.; Aebersold, R., xTract: software for characterizing conformational changes of protein complexes by quantitative cross-linking mass spectrometry. *Nat Methods* **2015**, *12* (12), 1185-90.
209. Tan, D.; Li, Q.; Zhang, M. J.; Liu, C.; Ma, C.; Zhang, P.; Ding, Y. H.; Fan, S. B.; Tao, L.; Yang, B.; Li, X.; Ma, S.; Liu, J.; Feng, B.; Liu, X.; Wang, H. W.; He, S. M.; Gao, N.; Ye, K.; Dong, M. Q.;

- Lei, X., Trifunctional cross-linker for mapping protein-protein interaction networks and comparing protein conformational states. *Elife* **2016**, *5*.
210. Kukacka, Z.; Rosulek, M.; Strohm, M.; Kavan, D.; Novak, P., Mapping protein structural changes by quantitative cross-linking. *Methods* **2015**, *89*, 112-20.
211. Fischer, L.; Chen, Z. A.; Rappsilber, J., Quantitative cross-linking/mass spectrometry using isotope-labelled cross-linkers. *J Proteomics* **2013**, *88*, 120-8.
212. Courcelles, M.; Coulombe-Huntington, J.; Cossette, E.; Gingras, A. C.; Thibault, P.; Tyers, M., CLMSVault: A Software Suite for Protein Cross-Linking Mass-Spectrometry Data Analysis and Visualization. *J Proteome Res* **2017**, *16* (7), 2645-2652.
213. Grimm, M.; Zimniak, T.; Kahraman, A.; Herzog, F., xVis: a web server for the schematic visualization and interpretation of crosslink-derived spatial restraints. *Nucleic Acids Res* **2015**, *43* (W1), W362-9.
214. Gotze, M.; Pettelkau, J.; Fritzsche, R.; Ihling, C. H.; Schafer, M.; Sinz, A., Automated assignment of MS/MS cleavable cross-links in protein 3D-structure analysis. *J Am Soc Mass Spectrom* **2015**, *26* (1), 83-97.
215. Liu, F.; Rijkers, D. T.; Post, H.; Heck, A. J., Proteome-wide profiling of protein assemblies by cross-linking mass spectrometry. *Nat Methods* **2015**, *12* (12), 1179-84.
216. Combe, C. W.; Fischer, L.; Rappsilber, J., xiNET: cross-link network maps with residue resolution. *Mol Cell Proteomics* **2015**, *14* (4), 1137-47.
217. Riffle, M.; Jaschob, D.; Zelter, A.; Davis, T. N., ProXL (Protein Cross-Linking Database): A Platform for Analysis, Visualization, and Sharing of Protein Cross-Linking Mass Spectrometry Data. *J Proteome Res* **2016**, *15* (8), 2863-70.
218. Schweppe, D. K.; Zheng, C.; Chavez, J. D.; Navare, A. T.; Wu, X.; Eng, J. K.; Bruce, J. E., XLinkDB 2.0: integrated, large-scale structural analysis of protein crosslinking data. *Bioinformatics* **2016**, *32* (17), 2716-8.
219. Kahraman, A.; Malmstrom, L.; Aebersold, R., Xwalk: computing and visualizing distances in cross-linking experiments. *Bioinformatics* **2011**, *27* (15), 2163-4.
220. Schweppe, D. K.; Chavez, J. D.; Bruce, J. E., XLmap: an R package to visualize and score protein structure models based on sites of protein cross-linking. *Bioinformatics* **2016**, *32* (2), 306-8.
221. Kosinski, J.; von Appen, A.; Ori, A.; Karius, K.; Muller, C. W.; Beck, M., Xlink Analyzer: software for analysis and visualization of cross-linking data in the context of three-dimensional structures. *J Struct Biol* **2015**, *189* (3), 177-83.
222. Ferber, M.; Kosinski, J.; Ori, A.; Rashid, U. J.; Moreno-Morcillo, M.; Simon, B.; Bouvier, G.; Batista, P. R.; Muller, C. W.; Beck, M.; Nilges, M., Automated structure modeling of large protein assemblies using crosslinks as distance restraints. *Nat Methods* **2016**, *13* (6), 515-20.
223. Ding, Y. H.; Gong, Z.; Dong, X.; Liu, K.; Liu, Z.; Liu, C.; He, S. M.; Dong, M. Q.; Tang, C., Modeling Protein Excited-state Structures from "Over-length" Chemical Cross-links. *J Biol Chem* **2017**, *292* (4), 1187-1196.
224. Kahraman, A.; Herzog, F.; Leitner, A.; Rosenberger, G.; Aebersold, R.; Malmstrom, L., Cross-link guided molecular modeling with ROSETTA. *PLoS One* **2013**, *8* (9), e73411.
225. Trahan, C.; Oeffinger, M., Targeted cross-linking-mass spectrometry determines vicinal interactomes within heterogeneous RNP complexes. *Nucleic Acids Res* **2016**, *44* (3), 1354-69.
226. Russel, D.; Lasker, K.; Webb, B.; Velazquez-Muriel, J.; Tjioe, E.; Schneidman-Duhovny, D.; Peterson, B.; Sali, A., Putting the pieces together: integrative modeling platform software for structure determination of macromolecular assemblies. *PLoS Biol* **2012**, *10* (1), e1001244.

227. Leitner, A.; Bonvin, A.; Borchers, C. H.; Chalkley, R. J.; Chamot-Rooke, J.; Combe, C. W.; Cox, J.; Dong, M. Q.; Fischer, L.; Gotze, M.; Gozzo, F. C.; Heck, A. J. R.; Hoopmann, M. R.; Huang, L.; Ishihama, Y.; Jones, A. R.; Kalisman, N.; Kohlbacher, O.; Mechtler, K.; Moritz, R. L.; Netz, E.; Novak, P.; Petrotchenko, E.; Sali, A.; Scheltema, R. A.; Schmidt, C.; Schriemer, D.; Sinz, A.; Sobott, F.; Stengel, F.; Thalassinos, K.; Urlaub, H.; Viner, R.; Vizcaino, J. A.; Wilkins, M. R.; Rappsilber, J., Toward Increased Reliability, Transparency, and Accessibility in Cross-linking Mass Spectrometry. *Structure* **2020**, *28* (11), 1259-1268.
228. Ihling, C. H.; Piersimoni, L.; Kipping, M.; Sinz, A., Cross-Linking/Mass Spectrometry Combined with Ion Mobility on a timsTOF Pro Instrument for Structural Proteomics. *Anal Chem* **2021**, *93* (33), 11442-11450.
229. Sarnowski, C. P.; Bikaki, M.; Leitner, A., Cross-linking and mass spectrometry as a tool for studying the structural biology of ribonucleoproteins. *Structure* **2022**, *30* (4), 441-461.
230. Kramer, K.; Sachsenberg, T.; Beckmann, B. M.; Qamar, S.; Boon, K. L.; Hentze, M. W.; Kohlbacher, O.; Urlaub, H., Photo-cross-linking and high-resolution mass spectrometry for assignment of RNA-binding sites in RNA-binding proteins. *Nat Methods* **2014**, *11* (10), 1064-70.
231. Knorlein, A.; Sarnowski, C. P.; de Vries, T.; Stoltz, M.; Gotze, M.; Aebersold, R.; Allain, F. H.; Leitner, A.; Hall, J., Nucleotide-amino acid pi-stacking interactions initiate photo cross-linking in RNA-protein complexes. *Nat Commun* **2022**, *13* (1), 2719.
232. Gotze, M.; Sarnowski, C. P.; de Vries, T.; Knorlein, A.; Allain, F. H.; Hall, J.; Aebersold, R.; Leitner, A., Single Nucleotide Resolution RNA-Protein Cross-Linking Mass Spectrometry: A Simple Extension of the CLIR-MS Workflow. *Anal Chem* **2021**, *93* (44), 14626-14634.
233. Stahl, K.; Graziadei, A.; Dau, T.; Brock, O.; Rappsilber, J., Protein structure prediction with in-cell photo-crosslinking mass spectrometry and deep learning. *Nat Biotechnol* **2023**.
234. Ward, C. C.; Kleinman, J. I.; Nomura, D. K., NHS-Esters As Versatile Reactivity-Based Probes for Mapping Proteome-Wide Ligandable Hotspots. *ACS Chem Biol* **2017**, *12* (6), 1478-1483.
235. Leitner, A.; Walzthoeni, T.; Aebersold, R., Lysine-specific chemical cross-linking of protein complexes and identification of cross-linking sites using LC-MS/MS and the xQuest/xProphet software pipeline. *Nat Protoc* **2014**, *9* (1), 120-37.
236. Chen, Z. L.; Meng, J. M.; Cao, Y.; Yin, J. L.; Fang, R. Q.; Fan, S. B.; Liu, C.; Zeng, W. F.; Ding, Y. H.; Tan, D.; Wu, L.; Zhou, W. J.; Chi, H.; Sun, R. X.; Dong, M. Q.; He, S. M., A high-speed search engine pLink 2 with systematic evaluation for proteome-scale identification of cross-linked peptides. *Nat Commun* **2019**, *10* (1), 3404.
237. Klykov, O.; Steigenberger, B.; Pektas, S.; Fasci, D.; Heck, A. J. R.; Scheltema, R. A., Efficient and robust proteome-wide approaches for cross-linking mass spectrometry. *Nat Protoc* **2018**, *13* (12), 2964-2990.
238. Yilmaz, S.; Busch, F.; Nagaraj, N.; Cox, J., Accurate and Automated High-Coverage Identification of Chemically Cross-Linked Peptides with MaxLynx. *Anal Chem* **2022**, *94* (3), 1608-1617.
239. Cox, J.; Neuhauser, N.; Michalski, A.; Scheltema, R. A.; Olsen, J. V.; Mann, M., Andromeda: a peptide search engine integrated into the MaxQuant environment. *J Proteome Res* **2011**, *10* (4), 1794-805.
240. Aramayo, R. J.; Willhoft, O.; Ayala, R.; Bythell-Douglas, R.; Wigley, D. B.; Zhang, X., Cryo-EM structures of the human INO80 chromatin-remodeling complex. *Nat Struct Mol Biol* **2018**, *25* (1), 37-44.

241. Gronemeyer, H.; Bourguet, W., Allosteric effects govern nuclear receptor action: DNA appears as a player. *Sci Signal* **2009**, *2* (73), pe34.
242. Germain, P.; Staels, B.; Dacquet, C.; Spedding, M.; Laudet, V., Overview of nomenclature of nuclear receptors. *Pharmacol Rev* **2006**, *58* (4), 685-704.
243. Weikum, E. R.; Liu, X.; Ortlund, E. A., The nuclear receptor superfamily: A structural perspective. *Protein Sci* **2018**, *27* (11), 1876-1892.
244. Warnmark, A.; Treuter, E.; Wright, A. P.; Gustafsson, J. A., Activation functions 1 and 2 of nuclear receptors: molecular strategies for transcriptional activation. *Mol Endocrinol* **2003**, *17* (10), 1901-9.
245. Gadaleta, R. M.; Magnani, L., Nuclear receptors and chromatin: an inducible couple. *J Mol Endocrinol* **2014**, *52* (2), R137-49.
246. Renaud, J. P.; Moras, D., Structural studies on nuclear receptors. *Cell Mol Life Sci* **2000**, *57* (12), 1748-69.
247. Moras, D.; Gronemeyer, H., The nuclear receptor ligand-binding domain: structure and function. *Curr Opin Cell Biol* **1998**, *10* (3), 384-91.
248. Batista, M. R.; Martinez, L., Dynamics of nuclear receptor Helix-12 switch of transcription activation by modeling time-resolved fluorescence anisotropy decays. *Biophys J* **2013**, *105* (7), 1670-80.
249. Porter, B. A.; Ortiz, M. A.; Bratslavsky, G.; Kotula, L., Structure and Function of the Nuclear Receptor Superfamily and Current Targeted Therapies of Prostate Cancer. *Cancers (Basel)* **2019**, *11* (12).
250. Lonard, D. M.; O'Malley, B. W., Nuclear receptor coregulators: modulators of pathology and therapeutic targets. *Nat Rev Endocrinol* **2012**, *8* (10), 598-604.
251. Gronemeyer, H.; Gustafsson, J. A.; Laudet, V., Principles for modulation of the nuclear receptor superfamily. *Nat Rev Drug Discov* **2004**, *3* (11), 950-64.
252. Shang, J.; Kojetin, D. J., Structural mechanism underlying ligand binding and activation of PPARgamma. *Structure* **2021**, *29* (9), 940-950 e4.
253. Johnson, B. A.; Wilson, E. M.; Li, Y.; Moller, D. E.; Smith, R. G.; Zhou, G., Ligand-induced stabilization of PPARgamma monitored by NMR spectroscopy: implications for nuclear receptor activation. *J Mol Biol* **2000**, *298* (2), 187-94.
254. Rochel, N.; Krucker, C.; Coutos-Thevenot, L.; Osz, J.; Zhang, R.; Guyon, E.; Zita, W.; Vanthong, S.; Hernandez, O. A.; Bourguet, M.; Badawy, K. A.; Dufour, F.; Peluso-Iltis, C.; Heckler-Beji, S.; Dejaegere, A.; Kamoun, A.; de Reynies, A.; Neuzillet, Y.; Rebouissou, S.; Beraud, C.; Lang, H.; Massfelder, T.; Allory, Y.; Cianferani, S.; Stote, R. H.; Radvanyi, F.; Bernard-Pierrot, I., Recurrent activating mutations of PPARgamma associated with luminal bladder tumors. *Nat Commun* **2019**, *10* (1), 253.
255. Tate, T.; Xiang, T.; Wobker, S. E.; Zhou, M.; Chen, X.; Kim, H.; Batourina, E.; Lin, C. S.; Kim, W. Y.; Lu, C.; McKiernan, J. M.; Mendelsohn, C. L., Pparg signaling controls bladder cancer subtype and immune exclusion. *Nat Commun* **2021**, *12* (1), 6160.
256. Miyamae, Y., Insights into Dynamic Mechanism of Ligand Binding to Peroxisome Proliferator-Activated Receptor gamma toward Potential Pharmacological Applications. *Biol Pharm Bull* **2021**, *44* (9), 1185-1195.
257. Shang, J.; Mosure, S. A.; Zheng, J.; Brust, R.; Bass, J.; Nichols, A.; Solt, L. A.; Griffin, P. R.; Kojetin, D. J., A molecular switch regulating transcriptional repression and activation of PPARgamma. *Nat Commun* **2020**, *11* (1), 956.
258. Hughes, T. S.; Chalmers, M. J.; Novick, S.; Kuruvilla, D. S.; Chang, M. R.; Kamenecka, T. M.; Rance, M.; Johnson, B. A.; Burris, T. P.; Griffin, P. R.; Kojetin, D. J., Ligand and receptor

- dynamics contribute to the mechanism of graded PPARgamma agonism. *Structure* **2012**, *20* (1), 139-50.
259. Chrisman, I. M.; Nemetcheck, M. D.; de Vera, I. M. S.; Shang, J.; Heidari, Z.; Long, Y.; Reyes-Caballero, H.; Galindo-Murillo, R.; Cheatham, T. E., 3rd; Blayo, A. L.; Shin, Y.; Fuhrmann, J.; Griffin, P. R.; Kamenecka, T. M.; Kojetin, D. J.; Hughes, T. S., Defining a conformational ensemble that directs activation of PPARgamma. *Nat Commun* **2018**, *9* (1), 1794.
260. Sanchez, D. J.; Missiaen, R.; Skuli, N.; Steger, D. J.; Simon, M. C., Cell-Intrinsic Tumorigenic Functions of PPARgamma in Bladder Urothelial Carcinoma. *Mol Cancer Res* **2021**, *19* (4), 598-611.
261. Giguere, V.; Yang, N.; Segui, P.; Evans, R. M., Identification of a new class of steroid hormone receptors. *Nature* **1988**, *331* (6151), 91-4.
262. Vanacker, J. M.; Bonnelye, E.; Chopin-Delannoy, S.; Delmarre, C.; Cavailles, V.; Laudet, V., Transcriptional activities of the orphan nuclear receptor ERR alpha (estrogen receptor-related receptor-alpha). *Mol Endocrinol* **1999**, *13* (5), 764-73.
263. Mazaira, G. I.; Zgajnar, N. R.; Lotufo, C. M.; Daneri-Becerra, C.; Sivils, J. C.; Soto, O. B.; Cox, M. B.; Galigniana, M. D., The Nuclear Receptor Field: A Historical Overview and Future Challenges. *Nucl Receptor Res* **2018**, *5*.
264. Deblois, G.; St-Pierre, J.; Giguere, V., The PGC-1/ERR signaling axis in cancer. *Oncogene* **2013**, *32* (30), 3483-90.
265. Bookout, A. L.; Jeong, Y.; Downes, M.; Yu, R. T.; Evans, R. M.; Mangelsdorf, D. J., Anatomical profiling of nuclear receptor expression reveals a hierarchical transcriptional network. *Cell* **2006**, *126* (4), 789-99.
266. Misawa, A.; Inoue, S., Estrogen-Related Receptors in Breast Cancer and Prostate Cancer. *Front Endocrinol (Lausanne)* **2015**, *6*, 83.
267. Shinozuka, T.; Ito, S.; Kimura, T.; Izumi, M.; Wakabayashi, K., Discovery of a Novel Class of ERRalpha Agonists. *ACS Med Chem Lett* **2021**, *12* (5), 817-821.
268. Schreiber, S. N.; Knutti, D.; Brogli, K.; Uhlmann, T.; Kralli, A., The transcriptional coactivator PGC-1 regulates the expression and activity of the orphan nuclear receptor estrogen-related receptor alpha (ERRalpha). *J Biol Chem* **2003**, *278* (11), 9013-8.
269. Paysan-Lafosse, T.; Blum, M.; Chuguransky, S.; Grego, T.; Pinto, B. L.; Salazar, G. A.; Bileschi, M. L.; Bork, P.; Bridge, A.; Colwell, L.; Gough, J.; Haft, D. H.; Letunic, I.; Marchler-Bauer, A.; Mi, H.; Natale, D. A.; Orengo, C. A.; Pandurangan, A. P.; Rivoire, C.; Sigrist, C. J. A.; Sillitoe, I.; Thanki, N.; Thomas, P. D.; Tosatto, S. C. E.; Wu, C. H.; Bateman, A., InterPro in 2022. *Nucleic Acids Res* **2023**, *51* (D1), D418-D427.
270. Carnesecchi, J.; Forcet, C.; Zhang, L.; Tribollet, V.; Barenton, B.; Boudra, R.; Cerutti, C.; Billas, I. M.; Serandour, A. A.; Carroll, J. S.; Beaudoin, C.; Vanacker, J. M., ERRalpha induces H3K9 demethylation by LSD1 to promote cell invasion. *Proc Natl Acad Sci U S A* **2017**, *114* (15), 3909-3914.
271. Lim, S.; Janzer, A.; Becker, A.; Zimmer, A.; Schule, R.; Buettner, R.; Kirfel, J., Lysine-specific demethylase 1 (LSD1) is highly expressed in ER-negative breast cancers and a biomarker predicting aggressive biology. *Carcinogenesis* **2010**, *31* (3), 512-20.
272. Carnesecchi, J.; Cerutti, C.; Vanacker, J. M.; Forcet, C., ERRalpha protein is stabilized by LSD1 in a demethylation-independent manner. *PLoS One* **2017**, *12* (11), e0188871.
273. Ariazi, E. A.; Jordan, V. C., Estrogen-related receptors as emerging targets in cancer and metabolic disorders. *Curr Top Med Chem* **2006**, *6* (3), 203-15.

274. Shao, G.; Wang, J.; Li, Y.; Liu, X.; Xie, X.; Wan, X.; Yan, M.; Jin, J.; Lin, Q.; Zhu, H.; Zhang, L.; Gong, A.; Shao, Q.; Wu, C., Lysine-specific demethylase 1 mediates epidermal growth factor signaling to promote cell migration in ovarian cancer cells. *Sci Rep* **2015**, *5*, 15344.
275. Nagasawa, S.; Sedukhina, A. S.; Nakagawa, Y.; Maeda, I.; Kubota, M.; Ohnuma, S.; Tsugawa, K.; Ohta, T.; Roche-Molina, M.; Bernal, J. A.; Narvaez, A. J.; Jeyasekharan, A. D.; Sato, K., LSD1 overexpression is associated with poor prognosis in basal-like breast cancer, and sensitivity to PARP inhibition. *PLoS One* **2015**, *10* (2), e0118002.
276. Mimasu, S.; Sengoku, T.; Fukuzawa, S.; Umehara, T.; Yokoyama, S., Crystal structure of histone demethylase LSD1 and tranylcypromine at 2.25 Å. *Biochem Biophys Res Commun* **2008**, *366* (1), 15-22.
277. Lupas, A. N.; Bassler, J.; Dunin-Horkawicz, S., The Structure and Topology of alpha-Helical Coiled Coils. *Subcell Biochem* **2017**, *82*, 95-129.
278. Tochio, N.; Umehara, T.; Koshiba, S.; Inoue, M.; Yabuki, T.; Aoki, M.; Seki, E.; Watanabe, S.; Tomo, Y.; Hanada, M.; Ikari, M.; Sato, M.; Terada, T.; Nagase, T.; Ohara, O.; Shirouzu, M.; Tanaka, A.; Kigawa, T.; Yokoyama, S., Solution structure of the SWIRM domain of human histone demethylase LSD1. *Structure* **2006**, *14* (3), 457-68.
279. Greschik, H.; Althage, M.; Flaig, R.; Sato, Y.; Chavant, V.; Peluso-Iltis, C.; Choulier, L.; Cronet, P.; Rochel, N.; Schule, R.; Stromstedt, P. E.; Moras, D., Communication between the ERRalpha homodimer interface and the PGC-1alpha binding surface via the helix 8-9 loop. *J Biol Chem* **2008**, *283* (29), 20220-30.
280. Hodge, E. A.; Benhaim, M. A.; Lee, K. K., Bridging protein structure, dynamics, and function using hydrogen/deuterium-exchange mass spectrometry. *Protein Sci* **2020**, *29* (4), 843-855.
281. Goswami, D.; Devarakonda, S.; Chalmers, M. J.; Pascal, B. D.; Spiegelman, B. M.; Griffin, P. R., Time window expansion for HDX analysis of an intrinsically disordered protein. *J Am Soc Mass Spectrom* **2013**, *24* (10), 1584-92.
282. Li, Y.; Kovach, A.; Suino-Powell, K.; Martynowski, D.; Xu, H. E., Structural and biochemical basis for the binding selectivity of peroxisome proliferator-activated receptor gamma to PGC-1alpha. *J Biol Chem* **2008**, *283* (27), 19132-9.
283. Tuting, C.; Iacobucci, C.; Ihling, C. H.; Kastiris, P. L.; Sinz, A., Structural analysis of 70S ribosomes by cross-linking/mass spectrometry reveals conformational plasticity. *Sci Rep* **2020**, *10* (1), 12618.
284. Merkley, E. D.; Rysavy, S.; Kahraman, A.; Hafen, R. P.; Daggett, V.; Adkins, J. N., Distance restraints from crosslinking mass spectrometry: mining a molecular dynamics simulation database to evaluate lysine-lysine distances. *Protein Sci* **2014**, *23* (6), 747-59.
285. Marty, M. T.; Hoi, K. K.; Robinson, C. V., Interfacing Membrane Mimetics with Mass Spectrometry. *Acc Chem Res* **2016**, *49* (11), 2459-2467.
286. Keener, J. E.; Zhang, G.; Marty, M. T., Native Mass Spectrometry of Membrane Proteins. *Anal Chem* **2021**, *93* (1), 583-597.
287. Bender, J.; Schmidt, C., Mass spectrometry of membrane protein complexes. *Biol Chem* **2019**, *400* (7), 813-829.
288. Bolla, J. R.; Agasid, M. T.; Mehmood, S.; Robinson, C. V., Membrane Protein-Lipid Interactions Probed Using Mass Spectrometry. *Annu Rev Biochem* **2019**, *88*, 85-111.
289. Oberai, A.; Ihm, Y.; Kim, S.; Bowie, J. U., A limited universe of membrane protein families and folds. *Protein Sci* **2006**, *15* (7), 1723-34.
290. Hedin, L. E.; Illergard, K.; Elofsson, A., An introduction to membrane proteins. *J Proteome Res* **2011**, *10* (8), 3324-31.

291. Almeida, J. G.; Preto, A. J.; Koukos, P. I.; Bonvin, A.; Moreira, I. S., Membrane proteins structures: A review on computational modeling tools. *Biochim Biophys Acta Biomembr* **2017**, *1859* (10), 2021-2039.
292. Dafun, A. S.; Marcoux, J., Structural mass spectrometry of membrane proteins. *Biochim Biophys Acta Proteins Proteom* **2022**, *1870* (8), 140813.
293. Calabrese, A. N.; Radford, S. E., Mass spectrometry-enabled structural biology of membrane proteins. *Methods* **2018**, *147*, 187-205.
294. O'Brien, D. P.; Hourdel, V.; Chenal, A.; Brier, S., Hydrogen/Deuterium Exchange Mass Spectrometry for the Structural Analysis of Detergent-Solubilized Membrane Proteins. *Methods Mol Biol* **2020**, *2127*, 339-358.
295. Coppieters 't Wallant, K.; Martens, C., Hydrogen-deuterium exchange coupled to mass spectrometry: A multifaceted tool to decipher the molecular mechanism of transporters. *Biochimie* **2022**.
296. Jodaitis, L.; van Oene, T.; Martens, C., Assessing the Role of Lipids in the Molecular Mechanism of Membrane Proteins. *Int J Mol Sci* **2021**, *22* (14).
297. Rosenbaum, D. M.; Rasmussen, S. G.; Kobilka, B. K., The structure and function of G-protein-coupled receptors. *Nature* **2009**, *459* (7245), 356-63.
298. Hino, T.; Arakawa, T.; Iwanari, H.; Yurugi-Kobayashi, T.; Ikeda-Suno, C.; Nakada-Nakura, Y.; Kusano-Arai, O.; Weyand, S.; Shimamura, T.; Nomura, N.; Cameron, A. D.; Kobayashi, T.; Hamakubo, T.; Iwata, S.; Murata, T., G-protein-coupled receptor inactivation by an allosteric inverse-agonist antibody. *Nature* **2012**, *482* (7384), 237-40.
299. Carlsson, J.; Yoo, L.; Gao, Z. G.; Irwin, J. J.; Shoichet, B. K.; Jacobson, K. A., Structure-based discovery of A2A adenosine receptor ligands. *J Med Chem* **2010**, *53* (9), 3748-55.
300. Lebon, G.; Warne, T.; Edwards, P. C.; Bennett, K.; Langmead, C. J.; Leslie, A. G.; Tate, C. G., Agonist-bound adenosine A2A receptor structures reveal common features of GPCR activation. *Nature* **2011**, *474* (7352), 521-5.
301. Carpenter, B.; Lebon, G., Human Adenosine A2A Receptor: Molecular Mechanism of Ligand Binding and Activation. *Front Pharmacol* **2017**, *8*, 898.
302. Lebon, G.; Edwards, P. C.; Leslie, A. G.; Tate, C. G., Molecular Determinants of CGS21680 Binding to the Human Adenosine A2A Receptor. *Mol Pharmacol* **2015**, *87* (6), 907-15.
303. Liu, W.; Chun, E.; Thompson, A. A.; Chubukov, P.; Xu, F.; Katritch, V.; Han, G. W.; Roth, C. B.; Heitman, L. H.; AP, I. J.; Cherezov, V.; Stevens, R. C., Structural basis for allosteric regulation of GPCRs by sodium ions. *Science* **2012**, *337* (6091), 232-6.
304. Cheng, R. K. Y.; Segala, E.; Robertson, N.; Deflorian, F.; Dore, A. S.; Errey, J. C.; Fiez-Vandal, C.; Marshall, F. H.; Cooke, R. M., Structures of Human A(1) and A(2A) Adenosine Receptors with Xanthines Reveal Determinants of Selectivity. *Structure* **2017**, *25* (8), 1275-1285 e4.
305. Magnani, F.; Shibata, Y.; Serrano-Vega, M. J.; Tate, C. G., Co-evolving stability and conformational homogeneity of the human adenosine A2a receptor. *Proc Natl Acad Sci U S A* **2008**, *105* (31), 10744-9.
306. Musheshe, N.; Schmidt, M.; Zaccolo, M., cAMP: From Long-Range Second Messenger to Nanodomain Signalling. *Trends Pharmacol Sci* **2018**, *39* (2), 209-222.
307. Consonni, S. V.; Maurice, M. M.; Bos, J. L., DEP domains: structurally similar but functionally different. *Nat Rev Mol Cell Biol* **2014**, *15* (5), 357-62.
308. Rehmann, H., Interaction of Epac with Non-canonical Cyclic Nucleotides. *Handb Exp Pharmacol* **2017**, *238*, 135-147.

309. Lezoualc'h, F.; Fazal, L.; Laudette, M.; Conte, C., Cyclic AMP Sensor EPAC Proteins and Their Role in Cardiovascular Function and Disease. *Circ Res* **2016**, *118* (5), 881-97.
310. Bouvet, M.; Blondeau, J. P.; Lezoualc'h, F., The Epac1 Protein: Pharmacological Modulators, Cardiac Signalosome and Pathophysiology. *Cells* **2019**, *8* (12).
311. Courilleau, D.; Bisselier, M.; Jullian, J. C.; Lucas, A.; Bouyssou, P.; Fischmeister, R.; Blondeau, J. P.; Lezoualc'h, F., Identification of a tetrahydroquinoline analog as a pharmacological inhibitor of the cAMP-binding protein Epac. *J Biol Chem* **2012**, *287* (53), 44192-202.
312. Boulton, S.; Selvaratnam, R.; Ahmed, R.; Van, K.; Cheng, X.; Melacini, G., Mechanisms of Specific versus Nonspecific Interactions of Aggregation-Prone Inhibitors and Attenuators. *J Med Chem* **2019**, *62* (10), 5063-5079.
313. Courilleau, D.; Bouyssou, P.; Fischmeister, R.; Lezoualc'h, F.; Blondeau, J. P., The (R)-enantiomer of CE3F4 is a preferential inhibitor of human exchange protein directly activated by cyclic AMP isoform 1 (Epac1). *Biochem Biophys Res Commun* **2013**, *440* (3), 443-8.
314. Lomize, M. A.; Pogozheva, I. D.; Joo, H.; Mosberg, H. I.; Lomize, A. L., OPM database and PPM web server: resources for positioning of proteins in membranes. *Nucleic Acids Res* **2012**, *40* (Database issue), D370-6.
315. Guccione, E.; Richard, S., The regulation, functions and clinical relevance of arginine methylation. *Nat Rev Mol Cell Biol* **2019**, *20* (10), 642-657.
316. Jarrold, J.; Davies, C. C., PRMTs and Arginine Methylation: Cancer's Best-Kept Secret? *Trends Mol Med* **2019**, *25* (11), 993-1009.
317. Dai, W.; Zhang, J.; Li, S.; He, F.; Liu, Q.; Gong, J.; Yang, Z.; Gong, Y.; Tang, F.; Wang, Z.; Xie, C., Protein Arginine Methylation: An Emerging Modification in Cancer Immunity and Immunotherapy. *Front Immunol* **2022**, *13*, 865964.
318. Wu, Q.; Schapira, M.; Arrowsmith, C. H.; Barsyte-Lovejoy, D., Protein arginine methylation: from enigmatic functions to therapeutic targeting. *Nat Rev Drug Discov* **2021**, *20* (7), 509-530.
319. Cura, V.; Cavarelli, J., Structure, Activity and Function of the PRMT2 Protein Arginine Methyltransferase. *Life (Basel)* **2021**, *11* (11).
320. Fuhrmann, J.; Clancy, K. W.; Thompson, P. R., Chemical biology of protein arginine modifications in epigenetic regulation. *Chem Rev* **2015**, *115* (11), 5413-61.
321. Niewmierzycka, A.; Clarke, S., S-Adenosylmethionine-dependent methylation in *Saccharomyces cerevisiae*. Identification of a novel protein arginine methyltransferase. *J Biol Chem* **1999**, *274* (2), 814-24.
322. Blanc, R. S.; Richard, S., Arginine Methylation: The Coming of Age. *Mol Cell* **2017**, *65* (1), 8-24.
323. Lorton, B. M.; Shechter, D., Cellular consequences of arginine methylation. *Cell Mol Life Sci* **2019**, *76* (15), 2933-2956.
324. Bedford, M. T.; Richard, S., Arginine methylation an emerging regulator of protein function. *Mol Cell* **2005**, *18* (3), 263-72.
325. Hwang, J. W.; Cho, Y.; Bae, G. U.; Kim, S. N.; Kim, Y. K., Protein arginine methyltransferases: promising targets for cancer therapy. *Exp Mol Med* **2021**, *53* (5), 788-808.
326. Tewary, S. K.; Zheng, Y. G.; Ho, M. C., Protein arginine methyltransferases: insights into the enzyme structure and mechanism at the atomic level. *Cell Mol Life Sci* **2019**, *76* (15), 2917-2932.
327. Sertznig, H.; Roesmann, F.; Wilhelm, A.; Heininger, D.; Bleekmann, B.; Elsner, C.; Santiago, M.; Schuhenn, J.; Karakoese, Z.; Benatzky, Y.; Snodgrass, R.; Esser, S.; Sutter, K.;

- Dittmer, U.; Widera, M., SRSF1 acts as an IFN-I-regulated cellular dependency factor decisively affecting HIV-1 post-integration steps. *Front Immunol* **2022**, *13*, 935800.
328. Bradley, T.; Cook, M. E.; Blanchette, M., SR proteins control a complex network of RNA-processing events. *RNA* **2015**, *21* (1), 75-92.
329. Labourier, E.; Bourbon, H. M.; Gallouzi, I. E.; Fostier, M.; Allemand, E.; Tazi, J., Antagonism between RSF1 and SR proteins for both splice-site recognition in vitro and Drosophila development. *Genes Dev* **1999**, *13* (6), 740-53.
330. Yu, B.; Li, P.; Zhang, Q. C.; Hou, L., Differential analysis of RNA structure probing experiments at nucleotide resolution: uncovering regulatory functions of RNA structure. *Nat Commun* **2022**, *13* (1), 4227.
331. Salifu, S. P.; Doughan, A., New Clues to Prognostic Biomarkers of Four Hematological Malignancies. *J Cancer* **2022**, *13* (8), 2490-2503.
332. Anko, M. L., Regulation of gene expression programmes by serine-arginine rich splicing factors. *Semin Cell Dev Biol* **2014**, *32*, 11-21.
333. Li, A. S. M.; Li, F.; Eram, M. S.; Bolotokova, A.; Dela Sena, C. C.; Vedadi, M., Chemical probes for protein arginine methyltransferases. *Methods* **2020**, *175*, 30-43.
334. Yang, Y.; Hadjikyriacou, A.; Xia, Z.; Gayatri, S.; Kim, D.; Zurita-Lopez, C.; Kelly, R.; Guo, A.; Li, W.; Clarke, S. G.; Bedford, M. T., PRMT9 is a type II methyltransferase that methylates the splicing factor SAP145. *Nat Commun* **2015**, *6*, 6428.
335. Champion-Arnaud, P.; Reed, R., The prespliceosome components SAP 49 and SAP 145 interact in a complex implicated in tethering U2 snRNP to the branch site. *Genes Dev* **1994**, *8* (16), 1974-83.
336. Terada, Y.; Yasuda, Y., Human immunodeficiency virus type 1 Vpr induces G2 checkpoint activation by interacting with the splicing factor SAP145. *Mol Cell Biol* **2006**, *26* (21), 8149-58.
337. Pauling, M. H.; McPheeters, D. S.; Ares, M., Jr., Functional Cus1p is found with Hsh155p in a multiprotein splicing factor associated with U2 snRNA. *Mol Cell Biol* **2000**, *20* (6), 2176-85.
338. Yang, Y.; Bedford, M. T., Protein arginine methyltransferases and cancer. *Nat Rev Cancer* **2013**, *13* (1), 37-50.
339. Smith, E.; Zhou, W.; Shindiapina, P.; Sif, S.; Li, C.; Baiocchi, R. A., Recent advances in targeting protein arginine methyltransferase enzymes in cancer therapy. *Expert Opin Ther Targets* **2018**, *22* (6), 527-545.
340. Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F. T.; de Beer, T. A. P.; Rempfer, C.; Bordoli, L.; Lepore, R.; Schwede, T., SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* **2018**, *46* (W1), W296-W303.
341. Kim, D. E.; Chivian, D.; Baker, D., Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res* **2004**, *32* (Web Server issue), W526-31.
342. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Zidek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D., Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596* (7873), 583-589.
343. Mirdita, M.; Schutze, K.; Moriwaki, Y.; Heo, L.; Ovchinnikov, S.; Steinegger, M., ColabFold: making protein folding accessible to all. *Nat Methods* **2022**, *19* (6), 679-682.

344. Silva, S. T. N.; Brito, J. A.; Arranz, R.; Sorzano, C. O. S.; Ebel, C.; Douth, J.; Tully, M. D.; Carazo, J. M.; Carrascosa, J. L.; Matias, P. M.; Bandejas, T. M., X-ray structure of full-length human RuvB-Like 2 - mechanistic insights into coupling between ATP binding and mechanical action. *Sci Rep* **2018**, *8* (1), 13726.
345. Zhou, C. Y.; Stoddard, C. I.; Johnston, J. B.; Trnka, M. J.; Echeverria, I.; Palovcak, E.; Sali, A.; Burlingame, A. L.; Cheng, Y.; Narlikar, G. J., Regulation of Rvb1/Rvb2 by a Domain within the INO80 Chromatin Remodeling Complex Implicates the Yeast Rvbs as Protein Assembly Chaperones. *Cell Rep* **2017**, *19* (10), 2033-2044.
346. Maimoni Campanella, J. E.; Ramos Junior, S. L.; Rodrigues Kiraly, V. T.; Severo Gomes, A. A.; de Barros, A. C.; Mateos, P. A.; Freitas, F. Z.; de Mattos Fontes, M. R.; Borges, J. C.; Bertolini, M. C., Biochemical and biophysical characterization of the RVB-1/RVB-2 protein complex, the RuvBL/RVB homologues in *Neurospora crassa*. *Biochimie* **2021**, *191*, 11-26.
347. Matias, P. M.; Baek, S. H.; Bandejas, T. M.; Dutta, A.; Houry, W. A.; Llorca, O.; Rosenbaum, J., The AAA+ proteins Pontin and Reptin enter adult age: from understanding their basic biology to the identification of selective inhibitors. *Front Mol Biosci* **2015**, *2*, 17.
348. Jha, S.; Dutta, A., RVB1/RVB2: running rings around molecular biology. *Mol Cell* **2009**, *34* (5), 521-33.
349. Dauden, M. I.; Lopez-Perrote, A.; Llorca, O., RUVBL1-RUVBL2 AAA-ATPase: a versatile scaffold for multiple complexes and functions. *Curr Opin Struct Biol* **2021**, *67*, 78-85.
350. Pal, M.; Munoz-Hernandez, H.; Bjorklund, D.; Zhou, L.; Degliesposti, G.; Skehel, J. M.; Hesketh, E. L.; Thompson, R. F.; Pearl, L. H.; Llorca, O.; Prodromou, C., Structure of the TLO2-TTI1-TTI2 complex and its function in TOR recruitment to the R2TP chaperone. *Cell Rep* **2021**, *36* (1), 109317.
351. Zhang, G.; Li, S.; Cheng, K. W.; Chou, T. F., AAA ATPases as therapeutic targets: Structure, functions, and small-molecule inhibitors. *Eur J Med Chem* **2021**, *219*, 113446.
352. Houry, W. A.; Bertrand, E.; Coulombe, B., The PAQosome, an R2TP-Based Chaperone for Quaternary Structure Formation. *Trends Biochem Sci* **2018**, *43* (1), 4-9.
353. Rivera-Calzada, A.; Pal, M.; Munoz-Hernandez, H.; Luque-Ortega, J. R.; Gil-Carton, D.; Degliesposti, G.; Skehel, J. M.; Prodromou, C.; Pearl, L. H.; Llorca, O., The Structure of the R2TP Complex Defines a Platform for Recruiting Diverse Client Proteins to the HSP90 Molecular Chaperone System. *Structure* **2017**, *25* (7), 1145-1152 e4.
354. Cloutier, P.; Poitras, C.; Durand, M.; Hekmat, O.; Fiola-Masson, E.; Bouchard, A.; Faubert, D.; Chabot, B.; Coulombe, B., R2TP/Prefoldin-like component RUVBL1/RUVBL2 directly interacts with ZNHIT2 to regulate assembly of U5 small nuclear ribonucleoprotein. *Nat Commun* **2017**, *8*, 15615.
355. Kakihara, Y.; Houry, W. A., The R2TP complex: discovery and functions. *Biochim Biophys Acta* **2012**, *1823* (1), 101-7.
356. Henri, J.; Chagot, M. E.; Bourguet, M.; Abel, Y.; Terral, G.; Maurizy, C.; Aigueperse, C.; Georgescauld, F.; Vandermoere, F.; Saint-Fort, R.; Behm-Ansmant, I.; Charpentier, B.; Pradet-Balade, B.; Verheggen, C.; Bertrand, E.; Meyer, P.; Cianferani, S.; Manival, X.; Quinternet, M., Deep Structural Analysis of RPAP3 and PIH1D1, Two Components of the HSP90 Co-chaperone R2TP Complex. *Structure* **2018**, *26* (9), 1196-1209 e8.
357. Horejsi, Z.; Stach, L.; Flower, T. G.; Joshi, D.; Flynn, H.; Skehel, J. M.; O'Reilly, N. J.; Ogradowicz, R. W.; Smerdon, S. J.; Boulton, S. J., Phosphorylation-dependent PIH1D1 interactions define substrate specificity of the R2TP cochaperone complex. *Cell Rep* **2014**, *7* (1), 19-26.

358. Boulon, S.; Marmier-Gourrier, N.; Pradet-Balade, B.; Wurth, L.; Verheggen, C.; Jady, B. E.; Rothe, B.; Pescia, C.; Robert, M. C.; Kiss, T.; Bardoni, B.; Krol, A.; Branlant, C.; Allmang, C.; Bertrand, E.; Charpentier, B., The Hsp90 chaperone controls the biogenesis of L7Ae RNPs through conserved machinery. *J Cell Biol* **2008**, *180* (3), 579-95.
359. Maurizy, C.; Quinteret, M.; Abel, Y.; Verheggen, C.; Santo, P. E.; Bourguet, M.; A, C. F. P.; Bragantini, B.; Chagot, M. E.; Robert, M. C.; Abeza, C.; Fabre, P.; Fort, P.; Vandermoere, F.; P, M. F. S.; Rain, J. C.; Charpentier, B.; Cianferani, S.; Bandejas, T. M.; Pradet-Balade, B.; Manival, X.; Bertrand, E., The RPAP3-Cterminal domain identifies R2TP-like quaternary chaperones. *Nat Commun* **2018**, *9* (1), 2093.
360. Coulombe, B.; Cloutier, P.; Gauthier, M. S., How do our cells build their protein interactome? *Nat Commun* **2018**, *9* (1), 2955.
361. Knowles, M. R.; Ostrowski, L. E.; Loges, N. T.; Hurd, T.; Leigh, M. W.; Huang, L.; Wolf, W. E.; Carson, J. L.; Hazucha, M. J.; Yin, W.; Davis, S. D.; Dell, S. D.; Ferkol, T. W.; Sagel, S. D.; Olivier, K. N.; Jahnke, C.; Olbrich, H.; Werner, C.; Raidt, J.; Wallmeier, J.; Pennekamp, P.; Dougherty, G. W.; Hjeij, R.; Gee, H. Y.; Otto, E. A.; Halbritter, J.; Chaki, M.; Diaz, K. A.; Braun, D. A.; Porath, J. D.; Schueler, M.; Baktai, G.; Griese, M.; Turner, E. H.; Lewis, A. P.; Bamshad, M. J.; Nickerson, D. A.; Hildebrandt, F.; Shendure, J.; Omran, H.; Zariwala, M. A., Mutations in SPAG1 cause primary ciliary dyskinesia associated with defective outer and inner dynein arms. *Am J Hum Genet* **2013**, *93* (4), 711-20.
362. Zeytuni, N.; Zarivach, R., Structural and functional discussion of the tetra-trico-peptide repeat, a protein interaction module. *Structure* **2012**, *20* (3), 397-405.
363. Dermouche, S.; Chagot, M. E.; Manival, X.; Quinteret, M., Optimizing the First TPR Domain of the Human SPAG1 Protein Provides Insight into the HSP70 and HSP90 Binding Properties. *Biochemistry* **2021**, *60* (30), 2349-2363.
364. Chagot, M. E.; Dos Santos Morais, R.; Dermouche, S.; Lefebvre, D.; Manival, X.; Chipot, C.; Dehez, F.; Quinteret, M., Binding properties of the quaternary assembly protein SPAG1. *Biochem J* **2019**, *476* (11), 1679-1694.
365. Pal, M.; Morgan, M.; Phelps, S. E.; Roe, S. M.; Parry-Morris, S.; Downs, J. A.; Polier, S.; Pearl, L. H.; Prodromou, C., Structural basis for phosphorylation-dependent recruitment of Tel2 to Hsp90 by Pih1. *Structure* **2014**, *22* (6), 805-18.
366. Mirvis, M.; Stearns, T.; James Nelson, W., Cilium structure, assembly, and disassembly regulated by the cytoskeleton. *Biochem J* **2018**, *475* (14), 2329-2353.
367. Dafinger, C.; Rinschen, M. M.; Borgal, L.; Ehrenberg, C.; Basten, S. G.; Franke, M.; Hohne, M.; Rauh, M.; Gobel, H.; Bloch, W.; Wunderlich, F. T.; Peters, D. J. M.; Tasche, D.; Mishra, T.; Habbig, S.; Dotsch, J.; Muller, R. U.; Bruning, J. C.; Persigehl, T.; Giles, R. H.; Benzing, T.; Schermer, B.; Liebau, M. C., Targeted deletion of the AAA-ATPase Ruvbl1 in mice disrupts ciliary integrity and causes renal disease and hydrocephalus. *Exp Mol Med* **2018**, *50* (6), 1-17.
368. Zariwala, M.; O'Neal, W. K.; Noone, P. G.; Leigh, M. W.; Knowles, M. R.; Ostrowski, L. E., Investigation of the possible role of a novel gene, DPCD, in primary ciliary dyskinesia. *Am J Respir Cell Mol Biol* **2004**, *30* (4), 428-34.
369. Olerinyova, A.; Sonn-Segev, A.; Gault, J.; Eichmann, C.; Schimpf, J.; Kopf, A. H.; Rudden, L. S. P.; Ashkinadze, D.; Bomba, R.; Frey, L.; Greenwald, J.; Degiacomi, M. T.; Steinhilper, R.; Killian, J. A.; Friedrich, T.; Riek, R.; Struwe, W. B.; Kukura, P., Mass Photometry of Membrane Proteins. *Chem* **2021**, *7* (1), 224-236.
370. Gorynia, S.; Bandejas, T. M.; Pinho, F. G.; McVey, C. E.; Vonrhein, C.; Round, A.; Svergun, D. I.; Donner, P.; Matias, P. M.; Carrondo, M. A., Structural and functional insights

into a dodecameric molecular machine - the RuvBL1/RuvBL2 complex. *J Struct Biol* **2011**, *176* (3), 279-91.

371. Dos Santos Morais, R.; Santo, P. E.; Ley, M.; Schelcher, C.; Abel, Y.; Plassart, L.; Desligniere, E.; Chagot, M. E.; Quinternet, M.; Paiva, A. C. F.; Hessmann, S.; Morellet, N.; P, M. F. S.; Vandermoere, F.; Bertrand, E.; Charpentier, B.; Bandejas, T. M.; Plisson-Chastang, C.; Verheggen, C.; Cianferani, S.; Manival, X., Deciphering cellular and molecular determinants of human DPCD protein in complex with RUVBL1/RUVBL2 AAA-ATPases. *J Mol Biol* **2022**, *434* (19), 167760.

372. van Zundert, G. C.; Trellet, M.; Schaarschmidt, J.; Kurkcuoglu, Z.; David, M.; Verlato, M.; Rosato, A.; Bonvin, A. M., The DisVis and PowerFit Web Servers: Explorative and Integrative Modeling of Biomolecular Complexes. *J Mol Biol* **2017**, *429* (3), 399-407.

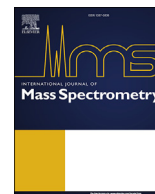
373. Koukos, P. I.; Bonvin, A., Integrative Modelling of Biomolecular Complexes. *J Mol Biol* **2020**, *432* (9), 2861-2881.

374. Michaelson-Richie, E. D.; Loeber, R. L.; Codreanu, S. G.; Ming, X.; Liebler, D. C.; Campbell, C.; Tretyakova, N. Y., DNA-protein cross-linking by 1,2,3,4-diepoxybutane. *J Proteome Res* **2010**, *9* (9), 4356-67.

Annexes

Annexe 1

Deslignière, E.; Ley, M.; Bourguet, M.; Ehkirch, A.; Botzanowski, T.; Erb, S.; Hernandez-Alba, O.; Cianférani, S., **Pushing the limits of native MS: Online SEC-native MS for structural biology applications.** *International Journal of Mass Spectrometry* 2021, 461.



Pushing the limits of native MS: Online SEC-native MS for structural biology applications



Evolène Deslignière, Marie Ley, Maxime Bourguet, Anthony Ehkirch, Thomas Botzanowski, Stéphane Erb, Oscar Hernandez-Alba, Sarah Cianférani*

Laboratoire de Spectrométrie de Masse BioOrganique, Université de Strasbourg, CNRS, IPHC UMR 7178, 67000, Strasbourg, France

ARTICLE INFO

Article history:

Received 23 October 2020
Received in revised form
17 December 2020
Accepted 18 December 2020
Available online 30 December 2020

Keywords:

Size exclusion chromatography
Native mass spectrometry
Ion mobility
LC-MS

ABSTRACT

Native mass spectrometry (nMS) is now widely applied to investigate non-covalently assembled biomolecule complexes. nMS requires the use of near-neutral pH and volatile buffers to preserve the native state of proteins. However, buffer exchange into nMS-compatible solvent is usually performed manually, which results in a time-consuming and tedious process, thus appearing as a major drawback for nMS analysis. Conversely, online coupling of size exclusion chromatography (SEC) to nMS affords a fast-automated and improved desalting, but also provides an additional dimension of separation for complex protein mixtures. We illustrate here the benefits of SEC-nMS compared to manual offline desalting for the characterization of a wide variety of biological systems, ranging from multiprotein assemblies, protein–ligand and protein–nucleic acid complexes, to proteins in a detergent environment. We then highlight the potential of the coupling to further integrate ion mobility while preserving the native conformations of proteins, allowing for rapid collision cross section measurement and even collision-induced unfolding experiments. Finally, we show that online SEC coupling can also serve as the basis for multidimensional non-denaturing liquid chromatography (LC) workflows, with the SEC acting as a fast desalting device, helping to achieve first dimension LC separation in optimal chromatographic conditions while being compatible with further nMS analysis.

© 2020 Elsevier B.V. All rights reserved.

1. Introduction

Native MS (nMS) remained for a long time the prerogative of few academic expert laboratories. The recent growing interest of biopharma companies for this methodology, especially for the characterization of therapeutic proteins like monoclonal antibodies (mAbs), prompted manufacturers to develop new instrumentations adapted to biopharma specific requirements, from sample preparation to data acquisition and treatment.

nMS was first introduced in the early 90s by the groups of Katta&Chait and Henion for the analyses of myoglobin [1] and kinase/substrate [2] complexes. The validity of gas phase nMS to conclude about solution binding stoichiometries and even more for affinities was for a long time a question of heated debates: “Do gas phase native MS spectra really reflect solution phase behavior?”; “Are gas phase data reliable to extrapolate solution phase

behavior?”. Major advances in studying macromolecular interaction by nMS have been performed by the groups of Carol Robinson and Albert Heck, and helped converting early skepticism into competence and consistency of practical use of nMS. nMS provides unambiguous determination of subunit stoichiometry, the most straightforward and important application of nMS, on a variety of biological assemblies ranging from multiprotein, e.g. transthyretin, retinol binding [3], GroEL [4], or HSP [5], to protein–ligand and protein–nucleic acid assemblies, including RNA editing complexes like the CRISPR-Cas systems [6], RNA polymerase II [7], and ribonucleoparticles [8]. Applications of nMS also encompass highly complex membrane proteins such as TRP channels [9], GPCR [10] or ABC transporters [11], and binding of lipids [12] or other small molecules [13] to membranes for which high resolution instrumentation is a real breakthrough. Even MegaDaltons ribosomes [14], hemocyanines [15], COP9 signalosome [16], dynactin complex [17] and virus particles with masses up to 18 MDa [18] can be measured by nMS.

From 2005, hyphenation of ion mobility (IM) spectrometry to nMS added a new dimension of gas-phase data interpretation,

* Corresponding author.
E-mail address: sarah.cianferani@unistra.fr (S. Cianférani).

affording conformational characterization along with stoichiometry determination. Ruotolo et al. first described a new type of instrumentation combining nMS and ion mobility (nIM-MS) and demonstrated that the overall topology of a ring-shaped protein complex could be retained in the gas phase [19]. More advanced strategies combining gas phase activation to nIM-MS, termed collision-induced unfolding (CIU), allow assessing differential gas phase unfolding behavior of proteins. Overall, nIM-MS demonstrated that most solution structures are retained into the gas phase [19,20].

However, despite all those technological breakthroughs, nMS was still poorly implemented in industrial environments in 2015, mostly due to its lack of automation. Indeed, prior to nMS, sample preparation consisting of buffer exchange has to be performed manually, hampering throughput necessary in companies. nMS analyses have to be performed in a volatile buffer, compatible with electrospray ionization and able to preserve weak non-covalent complex assemblies in solution [21]. Ammonium buffers are thus classically used for nMS analyses. Sample preparation usually consists of manual buffer exchange, or desalting, into ammonium acetate using a variety of devices such as gel filtration devices, microconcentrators or dialysis units, a step which is quite time-consuming, labor intensive and in some cases delicate. It can also lead to alteration of the sample, including aggregation, precipitation, or chemical modifications. Hence, this manual buffer exchange process appears as a major drawback for nMS analysis automation and throughput increase. The size exclusion chromatography (SEC), for which separation is based on the differences in hydrodynamic volumes, is achieved through a column packed with particles having precise pore sizes, which makes SEC interesting for fast buffer exchange, that is, separation of small nonvolatile salt molecules from protein species, and/or separation of high versus low mass proteins contained in a mixture.

Even if already suggested in early 2003 by Cavanagh et al. using self-packed gel filtration cartridges [22] and few years later (2008) by Waitt et al. with commercial SEC columns [23], the online coupling of SEC to nMS only became reality five years later for reference proteins [24,25], soluble protein/protein complexes [26] and biotherapeutics [27–32] analysis. Several reasons may account for the delay between first experiments and its effective wider application, among which: (i) the fact that older generation of MS instruments were not sufficiently tolerant to high salt contents necessary for nMS analysis, even volatile salts such as ammonium acetate, (ii) the lack of chemical inertness of SEC columns leading to poor peak shapes with volatile salts, and (iii) the amount of starting biomaterial required for the analysis. SEC afforded first a way to automate nMS experiments providing rapid assessment of proteins and complexes integrity of large numbers of samples, using small sample quantities (between 2 and 5 µg).

The benefit of SEC-nMS has been broadly documented by our group [27,28,33] and others [29–32] for therapeutic proteins characterization, more particularly for mAb-based product analysis. In the present work, we aim at widening the scope of SEC-nMS use for different systems of interest we had in the lab. After first presenting the SEC-nMS technical parameters to be optimized for each system, several examples will serve to discuss benefits but also limitations of online SEC-nMS. We then show the efficiency of the coupling to further integrate IM for fast collision cross section measurement and CIU experiments. Lastly, we demonstrate how online SEC coupling can also be used as the basis to develop multidimensional non-denaturing liquid chromatography (LC) setups, with the SEC column acting as a fast desalting device, helping to achieve first dimension LC separation in optimal chromatographic conditions while being compatible with nMS analysis.

2. Practical aspects, advantages and drawbacks of online SEC-nMS

Columns – Several types of SEC columns are currently available on the market from different manufacturers for the analysis of proteins and macromolecular complexes. They offer a wide variety in terms of column dimensions, stationary phase chemistry, particle sizes and porosities. More particularly, the development of sub-3 µm (from 1.7 to 2.7 µm) particle size SEC columns, also considered as the new generation SEC columns, allowed to improve separation and column efficiency while significantly reducing the analysis time between three- and five-fold with runs performed in dozens of minutes [34] (Figure S1). As expected, the highest resolution is achieved with the lowest SEC particle size (<2 µm). SEC columns are also available with different lengths, from 30 to 300 mm, that will affect the separation of the different protein populations and their nMS signal.

Mobile phase – Among the different parameters that have to be optimized to perform online SEC-nMS, the ionic strength and the pH of the mobile phase are critical to maintain the native conformation of the proteins. Those parameters mainly depend on the nature of the analyte and are the same as for offline nMS. SEC-nMS experiments can be performed with ammonium acetate, bicarbonate or formate mobile phases. Ventouri et al. conducted in-depth investigation on the ability of SEC-nMS to preserve the native fold of reference proteins depending on the nature, pH and ionic strength of the solvent [35]. Ammonium acetate was more effective to retain the native protein conformation under near-physiological pH conditions, although adsorption and peak tailing could be observed especially at lower ionic strength, which was also reported for mAb analysis [36]. Ammonium bicarbonate and formate may lead to higher fractions of denatured populations [35,37]; however, the presence and level of denatured species depend on the analyte, and need to be evaluated for other types of proteins.

Online SEC-nMS for high throughput online buffer exchange – For online desalting purposes, the focus will be on using short SEC columns (30 – 50 mm), as increased runtimes will be obtained with longer columns [24]. A compromise between short time of analysis, keeping efficient separation of low molecular weight non-volatile salts and proteins, limited protein adsorption on the stationary phase and acceptable MS intensities has to be found. Several short new-generation SEC columns between 30 and 50 mm are currently available from different manufacturers (ThermoFisher, Waters, Agilent, Phenomenex, etc.). Online SEC-nMS analysis of therapeutic mAbs could be successfully achieved with any of those columns, but in our hands, Waters BEH200 30 mm and Agilent AdvanceBio SEC 50 mm afford the best compromise for therapeutic protein analysis and are used at first line. However, protein adsorption could be different for other types of proteins. The flowrate also plays a critical role on SEC-nMS data: even if flowrates ranging between 100 and 250 µL/min can be used, lower elution flowrates (100 µL/min) usually provide increased nMS signal-to-noise ratio (Figure S2) and are thus better suited [28]. To conclude, for high throughput purposes resulting in <5 min runs, the shortest columns (30 – 50 mm) with smallest particle size (<3 µm) should be preferentially selected (Figure S1).

SEC is particularly well-adapted to perform online buffer exchange for proteins for which manual desalting may induce the precipitation of the sample or does not completely remove the non-volatile salts of the original buffer. The benefits of online SEC-nMS is clearly exemplified with a mAb-RNA conjugate (Fig. 1A). After 6 cycles of Vivaspin, only minor noisy MS signal that could correspond to the mAb-RNA conjugate is observed in the m/z range 6000 – 7500 (Fig. 1 A1), avoiding any mass calculation and

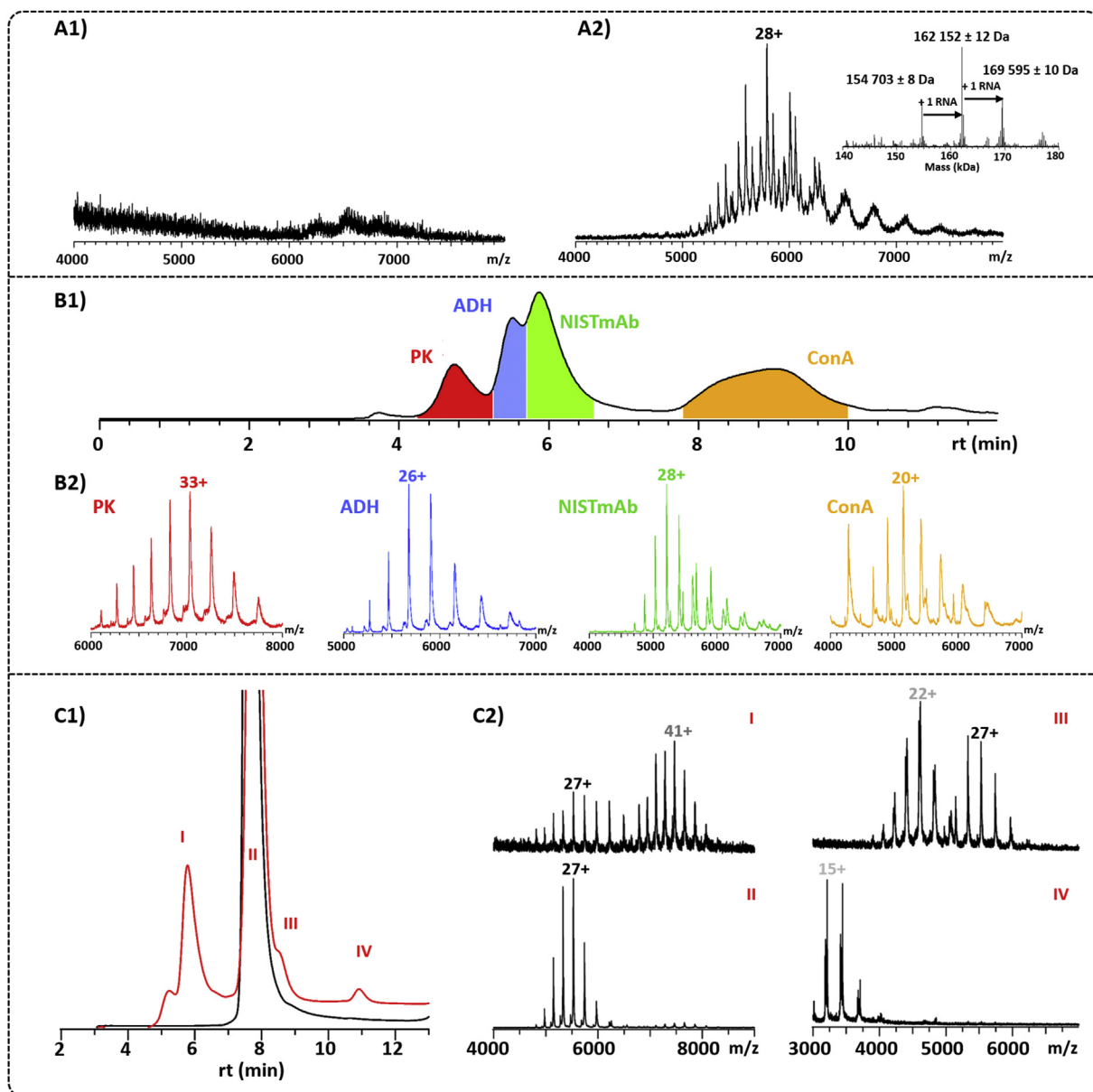


Fig. 1. SEC-nMS for efficient desalting and improved protein separation. Native mass spectrum of mAb-RNA conjugate obtained with (A1) manual desalting and (A2) online SEC-nMS. (B1) Chromatographic separation and (B2) corresponding native mass spectra of a mixture of four reference proteins (PK, ADH, ConA, and deglycosylated NISTmAb) obtained with a 300 mm SEC column (Waters BEH125). (C1) Chromatogram of nonstressed (black line) and thermally-stressed (red line) bevacizumab. (C2) Native mass spectra corresponding to the high molecular weight species (HMWS) and low molecular weight species (LMWS) of bevacizumab after thermal stress. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

determination of RNA distribution and binding stoichiometry. Conversely, SEC-nMS analysis affords efficient sample desalting and detection of several species with accurate mass measurements (Fig. 1 A2).

Online SEC-nMS for improved protein separation – One of the main issues in analyzing complex protein mixtures is the difference in ionization efficiencies of the constituents that might lead to an overwhelming MS signal of the protein with the highest ionization efficiency, and very minor signals detected in the background of the native mass spectrum. In such cases where several species may hamper the detection of each other, SEC-nMS might be an interesting alternative to simultaneously afford sample desalting and separation of proteins based on their hydrodynamic volume, leading to an easier and straightforward native mass spectra

interpretation. For separation purposes, longer columns (150 or 300 mm) with enhanced chromatographic separation capabilities but also longer run times will be preferred [28]. A mixture composed of four reference proteins, ConA, ADH, NISTmAb and PK, was analyzed with an Acquity BEH200 (4.6×30 mm, $1.7 \mu\text{m}$) column (Fig. 1B). Reference proteins were efficiently separated in the chromatographic dimension as depicted in Fig. 1 B1. Only NISTmAb and ADH show two overlapping peaks since they exhibit very similar hydrodynamic volumes. In this case, the mass spectra associated with both peaks allowed the identification of both proteins (Fig. 1 B2). Thereby proteins with different sizes can be separated with long SEC columns (300 mm) avoiding ionization competitions between the analytes, which enables a more reliable identification and quantification of different populations.

Another application of SEC-based protein separation is the separation and quantification of LMWS and HMWS [38] in forced degraded studies. When using nMS as a standalone technique, the relative quantification of protein size variants can be misleading due to the modification of the intensity of the different oligomers that may arise at the interface of the mass spectrometer as a consequence of the ionization process (disruption of noncovalent bonds, formation of nonspecific interactions, ionization competition between HMWS and LMWS). Upon thermal stress, the intensity of HMWS and LMWS increases compared to the non-stressed mAb sample (Fig. 1 C1), and the relative quantitation of each population can be assessed based on the areas of chromatographic peaks. However, the relative quantification of HMWS and LMWS significantly varies when using nMS as standalone technique (Figure S3). In this case, the relative intensity of the dimer is estimated to 21.5% (Figure S3A) compared to 14.5% (Figure S3B) with online SEC-nMS data, which corresponds to a variation of 50%. These differences stem on the formation of nonspecific HMWS in the interface of the mass spectrometer during the ionization process compared to nanoESI-nMS. For this reason, online SEC-nMS is particularly well-suited to assess protein stability through forced degraded studies by comparing the relative intensity of the size variant populations. Online SEC-nMS enables an efficient separation, a more reliable quantification and a simultaneous identification of high- and low-molecular species to afford a precise characterization of protein degradation.

Sensitivity of SEC-nMS – Overall, SEC-nMS coupling often requires larger amount of sample compared to nanoESI-nMS. Even though the sensitivity of the technique depends on several factors such as the ionization efficiency of the analyte or the type of mass spectrometer used, injected material can vary from 2 to 5 μg . However, mass spectra with a suitable S/N ratio can be obtained with less material. The limit of detection of our online SEC coupled to a Synapt-G2 HDMS mass spectrometer with glycosylated pertuzumab was established to 1 μg (Figure S4), where the S/N ratio (80) and the resolution of the MS peaks ($R_s @m/z\ 5290 = 587$) allow an accurate mass measurement and assignment of all mAb populations. Lower amounts could be injected in other SEC-nMS setups [26], nevertheless 1 μg of loaded mAb offers an adequate trade-off between the amount of sample and mass spectrum quality.

Main SEC-nMS drawbacks – The use of volatile salts in the mobile phase of SEC-nMS systems may induce further interactions with the stationary phase of the column, leading to the coelution of different protein populations [36] or providing some discrepancies between the elution volumes and the hydrodynamic volume of the oligomers [24]. Since SEC separation is based on the hydrodynamic volume, only proteins with different sizes will be efficiently separated, otherwise, only coeluting or partially resolved chromatographic peaks will be obtained with this kind of chromatography. In addition to possible interactions with the stationary phase, sample dilution in the mobile phase will occur due to diffusion in the column, which may also dissociate unstable and/or low affinity complexes.

Normally, nMS is performed in the nanoESI regime to reduce the amount of injected sample. Conversely, SEC-nMS experiments are performed in ESI ionization mode and thus require higher amounts of starting material. Upon addition of SEC dimension, the flowrate is increased significantly (between 100 and 300 $\mu\text{L}/\text{min}$) implying the use of higher desolvation gas and source temperatures to improve solvent evaporation and hence, enhance nMS data. The latter parameter has been recently reported in the literature to have an impact on mAb gas-phase energetics by increasing the internal energy of the ions [33]. These results suggest that source temperatures, and desolvation gas temperatures, need to be carefully adjusted to avoid further source activation and/or fragmentation. In

spite of these potential drawbacks, this review illustrates numerous examples where SEC in combination with nMS has been used to provide further insights into the characterization of therapeutic proteins and macromolecular complexes, and hence, showing the suitability of this coupling for structural biology.

3. SEC-nMS is suitable for a variety of non-covalent complexes analyses

Protein/ligand and protein/metal interactions – SEC-nMS can help to probe non-covalent interactions between proteins and small molecules. nMS has been reported for high-throughput screening of ligands libraries against enzymes or receptors, the latter being extensively applied in drug discovery. nMS has already been reported as a straightforward, fast and reproducible method to detect ligand binding to specific targets, but also to determine binding affinity, stoichiometry and specificity [39], thus appearing as an attractive technique to complement more common approaches including nuclear magnetic resonance, or surface plasmon resonance [40]. However, the use of SEC coupled to nMS has been scarcely reported in literature.

Quinn's group has presented an online SEC-ESI-FTICR-MS approach to detect protein-ligand non-covalent complexes and to screen natural product extracts [41]. A second study details the technical developments and optimizations of SEC-nMS as a robust, quantitative, and automated platform to measure affinities of non-covalent protein–small molecule interactions on an Orbitrap instrument using indoleamine 2,3-dioxygenase 1 (IDO1), a catabolic enzyme, and inhibitory ligands as case study [42]. The scarce use of SEC coupled to nMS for protein–ligand screening may be explained by the risk of cross contamination, as high affinity ligand adsorption on the SEC column may prevent further detection of binding of a lower affinity molecule to the target protein. Alternative nanoESI microfluidic devices, e.g. the Triversa Nanomate from Advion, have been developed, combining capabilities of sample preparation such as ligand addition to the protein and temperature-controlled incubation to reproducible nanoESI injection and nMS analysis [43]. Benefits of such systems is that using individual nozzles for each ligand/mixture of ligand prevents from cross contamination of capillaries or columns.

SEC-nMS can also be of interest to investigate protein–metal interactions, as reported by Jia et al. for [2Fe–2S] cluster-bridged complexes, for which offline nanoESI-nMS failed to detect cluster-bound species [44]. As iron-sulfur clusters are sensitive to oxygen, SEC-nMS helps to minimize the possibility to oxidize unstable analytes due to reduced sample preparation and shorter time of analysis. Here, SEC-nMS performed under inert atmosphere (injection valve, sample syringe and samples flushed with argon) revealed the presence of a GLRX5 homodimer with one [2Fe–2S] cluster. This setup also allowed to monitor cluster transfer reactions to get better insight into intermediate [2Fe–2S] species, proving that SEC-nMS can be used as a robust and fast technique to elucidate both the cluster and protein components.

Protein/protein interactions – SEC is commonly used to investigate protein oligomerization state or more generally the distribution of size variants of a sample [45]. It is also often the last step of protein/complex purification. We thus performed systematic comparison of manual desalting followed by nMS analysis versus online SEC-nMS for different systems of increasing complexity. The first example relies on the oligomeric state assessment of a homo-oligomeric soluble protein PRMT2, a member of the protein arginine methyltransferase family that has diverse roles in transcriptional regulation [46]. Fig. 2 A1–2 presents the nMS mass spectra obtained after manual desalting (two cycles of Zeba 7 kDa, AcONH₄ 500 mM, pH 7.0) or SEC-nMS analysis of

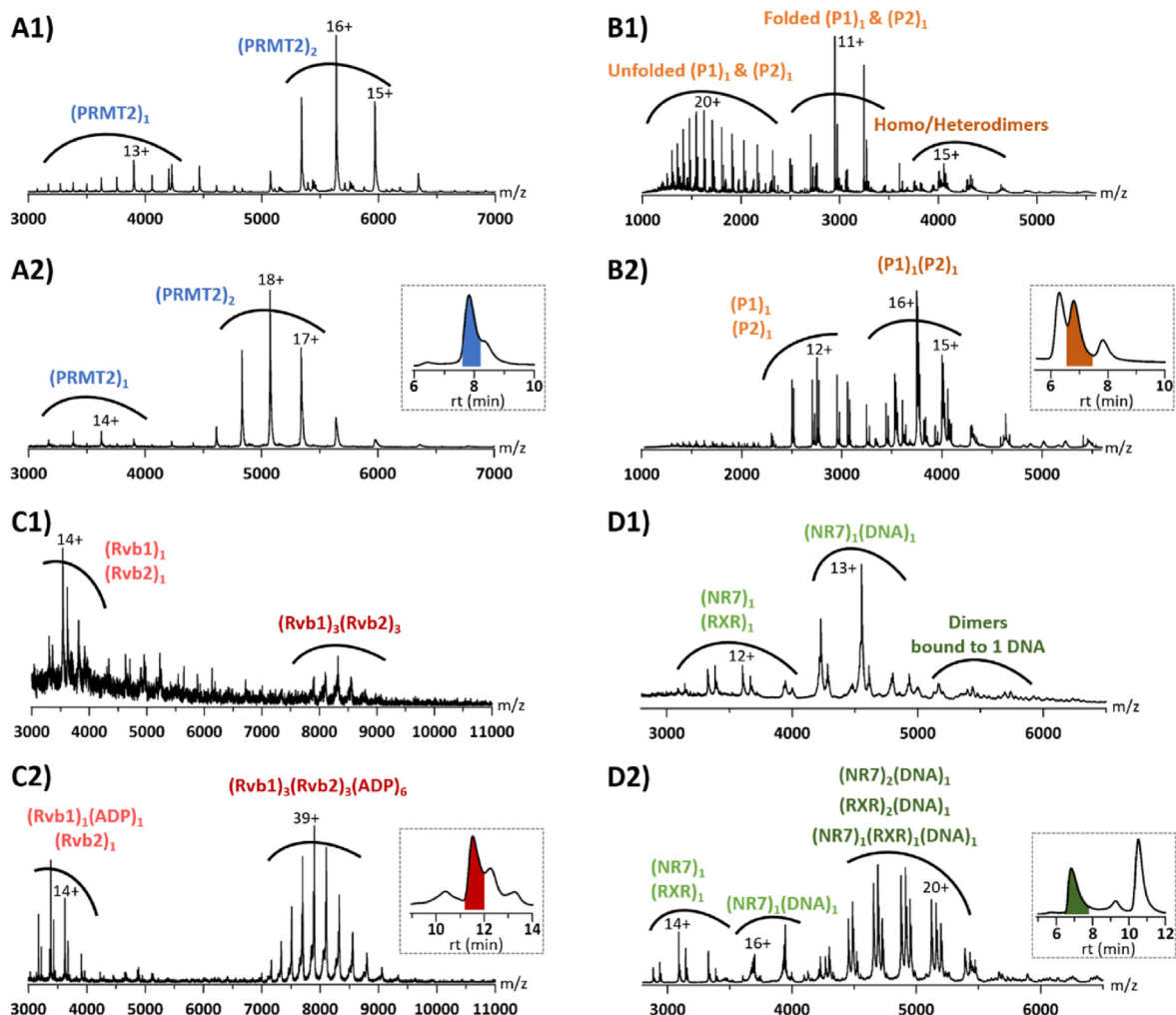


Fig. 2. Offline nanoESI-nMS (A1, B1, C1, D1) versus online SEC-nMS experiments (A2, B2, C2, D2) with corresponding SEC-UV chromatograms at 280 nm depicted in insets. Native mass spectra of (A) PRMT2, (B) an equimolar 1:1 P1:P2 mixture, (C) HMW Rvbs complexes, and (D) nuclear receptors RXR and NR7 bound to DNA.

mouse PRMT2. Except differences in the charge state distributions (CSD) due to the use of nanoESI versus ESI, similar mass spectra were obtained, highlighting the presence of two CSDs, a minor one with a mass of $50\,726 \pm 1$ Da corresponding to monomeric (PRMT2)₁ and a major one of $101\,517 \pm 3$ Da corresponding to dimeric (PRMT2)₂. The PRMT2 example shows the potential of SEC-nMS to preserve the oligomeric state of a protein known as dimeric. Additionally, SEC-nMS ensures the spray stability at high concentration of AcONH₄ (500 mM) contrary to nanoESI-nMS.

Benefits of SEC is further illustrated on the multiprotein P1–P2 complex system. The nMS spectrum (Fig. 2 B1) obtained after manual desalting is highly complex and presents several CSDs. The first one between m/z 1000 and 2500 corresponds to partially unfolded monomeric P1 and P2; from m/z 2500 to 3500, the most intense CSDs correspond to folded monomeric P1 and P2 species; finally, in the m/z range 3700 – 5000, minor signals of homo/heterodimers are detected. Conversely, online SEC-nMS analysis of an equimolar mixture of P1 and P2 (100 μ M of each, 1-h room temperature incubation) revealed three partially resolved chromatographic peaks (inset, Fig. 2 B2). The most intense chromatographic peak corresponds to (P1)₂ homodimers ($64\,914 \pm 1$ Da), followed by a second peak with a mass of $59\,973 \pm 2$ Da in agreement with a 1:1 stoichiometric (P1)₁(P2)₁ heterodimer (Fig. 2 B2) and finally a peak at 7.8 min corresponding to (P2)₂ homodimers ($55\,030 \pm 1$ Da). This

example illustrates the benefits of using SEC separation capabilities (even if partial) to compensate for overlapping CSDs, “polishing” the differences in ionization efficiencies of monomers and dimers and resulting in better quality nMS spectra.

Finally, SEC-nMS is also adapted to the characterization of HMW complexes (>100 kDa). Yeast Rvb1 and Rvb2 are involved in various cellular processes ranging from ribonucleoprotein complex biogenesis to chromatin remodeling [47]. Rvbs predominantly assemble into hexameric rings >300 kDa [47], and are generally purified with buffers containing ADP, which helps to maintain the integrity of the hexamers [48]. The nMS spectrum obtained after manual desalting (Zeba 7 kDa, AcONH₄ 150 mM, pH 7.5) shows that monomers are mainly detected, and only low intensity signals with $S/N = 5$ are observed for hexamers, suggesting that complex stabilization *via* ADP was not retained with manual buffer exchange (Fig. 2 C1). Online SEC-nMS experiments performed using 150 mM AcONH₄ as mobile phase demonstrate the potential of the coupling to preserve non-covalent interactions within the hexameric ring of Rvbs, while improving the resolution ($R_s @m/z$ 8106 = 811) and the intensity of the signal (71 S/N) (Fig. 2 C2). SEC-nMS highlights the presence of two populations of hexamers bearing 5 and 6 ADP respectively, the predominant one being bound to 6 ADP ($307\,973 \pm 8$ Da).

Protein/nucleic acid interactions – For protein-nucleic acid

complexes with either RNA or DNA, an additional challenge has to be faced related to the presence of cations associated to oligonucleotide binding which in turn often leads to low quality nMS signals with broad peaks, low mass spectral resolution, and difficult data interpretation. Interactions of nuclear receptors with their cognate DNA response elements serve to illustrate this point. NR7 and RXR are transcription factors that can bind a variety of ligands which can further interact with specific DNA sequences, known as response elements, either as monomers or hetero/homo-dimers. nMS is often used to assess the binding stoichiometry of complexes involving nuclear receptors [49,50]. In this example, nMS was employed to uncover the dimerization properties of NR7 and RXR upon DNA binding. The nMS spectrum obtained after manual buffer exchange (Zeba 7 kDa, ACONH₄ 150 mM, pH 8.0) followed by complex reconstitution of NR7, RXR and DNA (1 h at 4 °C) reveals that the most intense CSD corresponds to the binding of one monomer of NR7 to a DNA fragment of 15 940 Da (Fig. 2 D1). Low intensity signals corresponding to HMWS are also observed in the 5300 – 5800 *m/z* region, but do not allow accurate mass measurements and proper mass-based identification. Conversely, the online SEC-nMS analysis (150 mM ACONH₄ pH 8.0 after the complex reconstitution) exhibits very different results. The first chromatographic peak at 6.8 min reveals the coexistence of different dimers each bound to one DNA fragment: (NR7)₂ homodimers (102 411 ± 1 Da), (RXR)₂ homodimers (103 944 ± 2 Da) and 1:1 (NR7)₁(RXR)₁ heterodimers (103 177 ± 2 Da) (Fig. 2 D2). The second chromatographic peak at 9.3 min corresponds to the binding of NR7 to DNA as a monomer (59 178 ± 2 Da), while the last peak at 10.5 min shows the presence of isolated DNA species. Thus, online SEC-nMS allowed to unambiguously characterize the binding combinations between NR7, RXR and the DNA, highlighting the fact that these nuclear receptors predominantly bind DNA as dimers. Here, these results emphasize the potential of SEC-nMS to preserve and transmit heterogeneous protein–nucleic acids complexes.

SEC-nMS for proteins in a detergent environment – Membrane protein analysis by nMS is still challenging as it needs to be performed in a detergent environment. Online SEC-nMS can be applied also in presence of detergents, as depicted with the determination of the oligomeric species constituting amyloid-beta (Aβ) pores that form in the membrane of neurons to explain neurotoxicity in Alzheimer's disease [51]. Fig. 3A presents the nMS spectrum of the β-sheet pore-forming (βPFO) Aβ(1–42) after manual desalting in C8E5 detergent. Several species ranging from monomers of high intensity to possible octamers of very low intensity were detected, precluding to conclude about the main species present in solution. To circumvent these limitations, SEC-nMS was performed in 200 mM ammonium bicarbonate pH 9.0 with 14.2 mM C8E5 detergent as mobile phase to benefit from SEC separation capabilities. SEC-nMS analysis of βPFOs_{Aβ(1–42)} resulted in two partially separated chromatographic peaks: the peak at 6.3 min corresponds to octamers (36 114 ± 1 Da) due to the presence of intense 9+ and 7+ ions of the octamer, while the peak at 7.0 min could be attributed to tetramers (18 056 ± 1 Da) (Fig. 3B). Comparison of SEC-nMS analysis of low versus high concentrations of Aβ(1–42) revealed that the former was enriched in Aβ(1–42) tetramers and the latter in octamers (inset, Fig. 3B) [51]. SEC-nMS presented a unique opportunity to establish the stoichiometry of the potentially distinct oligomer species of βPFOs_{Aβ(1–42)}, revealing that the main species present in the βPFOs_{Aβ(1–42)} sample were Aβ(1–42) tetramers and octamers [51]. In addition, since no charge states specific for other oligomer stoichiometries between tetramers and octamers were detected, these results suggested that tetramers were the building block for octamer formation. Overall, SEC-nMS appears promising to characterize membrane proteins in detergent environments. However, further developments are still

necessary, as higher amounts of non-volatile detergents are required compared to nanoESI-nMS. Besides, higher flowrates used in online SEC-nMS compared to nanoESI-nMS may reduce MS sensitivity, and dissociation of unstable membrane protein complexes may occur [52].

4. Hyphenation of ion mobility to SEC-nMS for conformational characterization of protein and protein complexes

nIM-MS is now broadly used in structural biology to study the gas-phase conformations of protein and their non-covalent complexes [53,54]. IM separates ions based on their size, shape and charge under the influence of an electric field as they drift through an inert buffer gas in the mobility cell. Drift times can be further converted into rotationally averaged collision cross sections (CCS), which correspond to projected areas of ions. IM affords an additional level of gas-phase characterization, and can help not only to separate isomers but also to examine conformational changes upon complex formation. Hence, the hyphenation of SEC to nIM-MS, which was first reported by Van der Rest et al. [24], is of main interest to rapidly gain further insight into proteins conformations.

SEC-nIM-MS for spectrum cleaning and unambiguous CSDs attribution of oligomers – When several oligomerization states of a protein coexist, unambiguous peak assignment may not be achieved with SEC-nMS due to both low intensities or lack of isotopic resolution for higher oligomeric states. The benefit of SEC-nIM-MS for the separation of isobaric species along with unambiguous peak attribution of oligomers is clearly evidenced in the case of Aβ(1–42)s, which form monomers, dimers, trimers, tetramers and octamers [51]. For instance, mass spectra obtained with SEC-nMS (Fig. 3B) exhibit a peak at *m/z* 3010 that could correspond to either dimeric (*z* = 3+), tetrameric (*z* = 6+) or octameric (*z* = 12+) species. With the addition of IM, the arrival time distribution (ATD) corresponding to *m/z* 3010 highlights the presence of two different populations at 8.55 and 12.01 ms (inset, Fig. 3C) which were identified as tetramers (*z* = 6+) and dimers (*z* = 3+), respectively. Since ions belonging to the same series are diagonally aligned on bidimensional drift time vs *m/z* plots, tetrameric 5+ ions and octameric 8+ ions could also be assigned (Fig. 3C). This example demonstrates the potential of SEC-nIM-MS to assess the coexistence of oligomeric states for identical *m/z* while ensuring unambiguous CSDs attributions.

SEC-nIM-MS for high-throughput CCS measurements of biotherapeutics – While the use of nIM-MS is commonplace in academic research laboratories, it remains scarce in biopharmaceutical companies. Although nIM-MS has already proven valuable to characterize biotherapeutic proteins such as mAbs, bispecific antibodies (bsAb) and antibody–drug conjugates (ADC) [55], bottlenecks hampering its routine use in biopharma include (i) relatively manual experiments that require highly skilled operators performed with nanoelectrospray capillaries of microfluidic devices and (ii) lack of automation of nIM-MS data processing.

Van der Rest et al. first demonstrated that CCS values of reference proteins obtained using SEC-nIM-MS were in good agreement with nanoESI-nIM-MS values [24]. With therapeutic mAbs as model proteins, we compared ^{TW}CCS_{N₂} obtained with nanoESI-nIM-MS after manual desalting and SEC-nIM-MS analysis [28,33]. Different points should be considered when performing SEC-nIM-MS for CCS calculation. As SEC-nIM-MS experiments are performed in ESI mode, CSDs in SEC-ESI are centered on higher charge states than in nanoESI. Dipole-dipole interactions and coulombic repulsions increasing with charge state, higher charge states might be more activated, which might result in slight protein unfolding. As shown in Fig. 4A, similar IM calibration equations were obtained

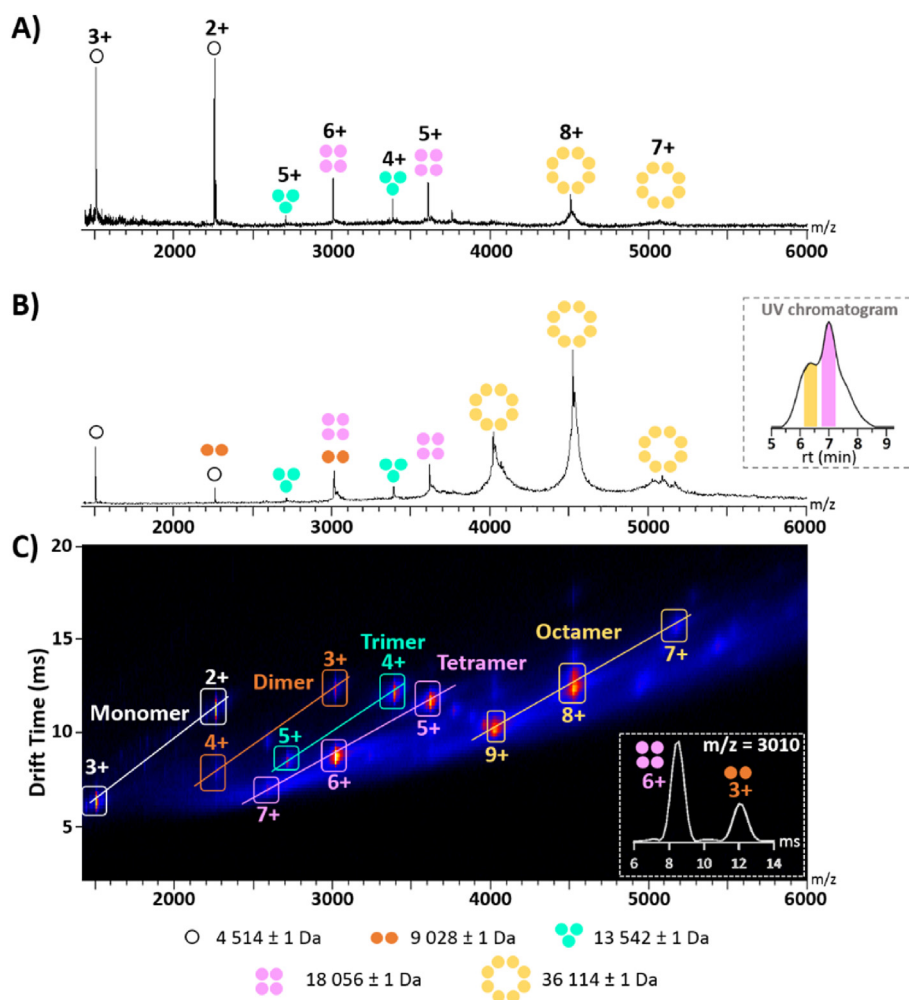


Fig. 3. NanoESI- and SEC-ESI- nIM-MS experiments of Aβ(1–42). (A) Native nanoESI mass spectrum obtained after manual desalting. (B) Native SEC-ESI mass spectrum corresponding to the chromatographic peak of the octamer (yellow slice of UV chromatogram shown in inset). Species were identified as monomers (white), dimers (orange), trimers (green), tetramers (pink) and octamers (yellow). (C) 2D SEC-nIM-MS Driftscope plot with charge state distributions of the different oligomers. The extracted ATD at m/z 3010 is shown in inset. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

for $^{TW}CCS_{N_2}$ measurements performed in SEC-nIM-MS and offline injections of manually desalted proteins. Keeping that in mind, CCS were measured for a series of mAbs and show good correlation with $^{TW}CCS_{N_2}$ measured in nanoESI-nIM-MS (Fig. 4B). In this context, SEC-nIM-MS appears as an attractive setup to automate sample preparation while widening the information content available in a single few-minutes experiment as illustrated by Ehkirch et al. for the characterization of different therapeutic mAbs formats, including IgGs, ADCs and bsAbs [28]. Additionally, SEC-nIM-MS affords further characterization of HMWS in stressed mAbs samples, revealing a significant conformational compaction upon mAb aggregation [24].

Altogether, our results highlight that online coupling of SEC to nIM-MS does not significantly influence the global conformation of mAbs, which of course needs to be checked for other types of proteins.

SEC coupled to Collision-Induced Unfolding – CIU is an IM-based approach used to probe ion gas-phase unfolding. In travelling-wave IM spectrometry (TWIMS) instruments, CIU experiments are performed by raising collision voltages (CVs) in the trap cell before IM separation, subsequently activating ions that can undergo conformational changes [57]. While CIU appears as an

elegant alternative to circumvent low IM resolution of linear TWIMS cells, lack of CIU automation in both buffer exchange and data acquisition has precluded its wide adoption. In order to fully automate CIU experiments from sample preparation to data interpretation, we have developed a fast-online SEC-CIU coupling using short SEC columns, where CVs are automatically raised along sample elution (Fig. 4C) [33]. CIU fingerprints are generated with the CIUSuite 2 software to better visualize ATDs variations corresponding to conformational transitions (Fig. 4D). This SEC-CIU setup is particularly interesting to rapidly distinguish mAbs subclasses, as the latter exhibit different unfolding patterns at both intact [58] and middle [59] levels. Diagnostic CVs regions allow to identify mAbs subclasses with scores >80% (Fig. 4D). Targeted scheduled SEC-CIU, that is, acquisition of solely the most diagnostic CVs, helps to further reduce the data collection time while retaining clear-cut mAbs classifications.

Altogether, the combination of improved high-throughput desalting and automated data collection afforded by SEC drastically shortens the overall time process of CIU experiments, from 3 h for classical CIU experiments with manual buffer exchange to 15 min for targeted scheduled SEC-CIU.

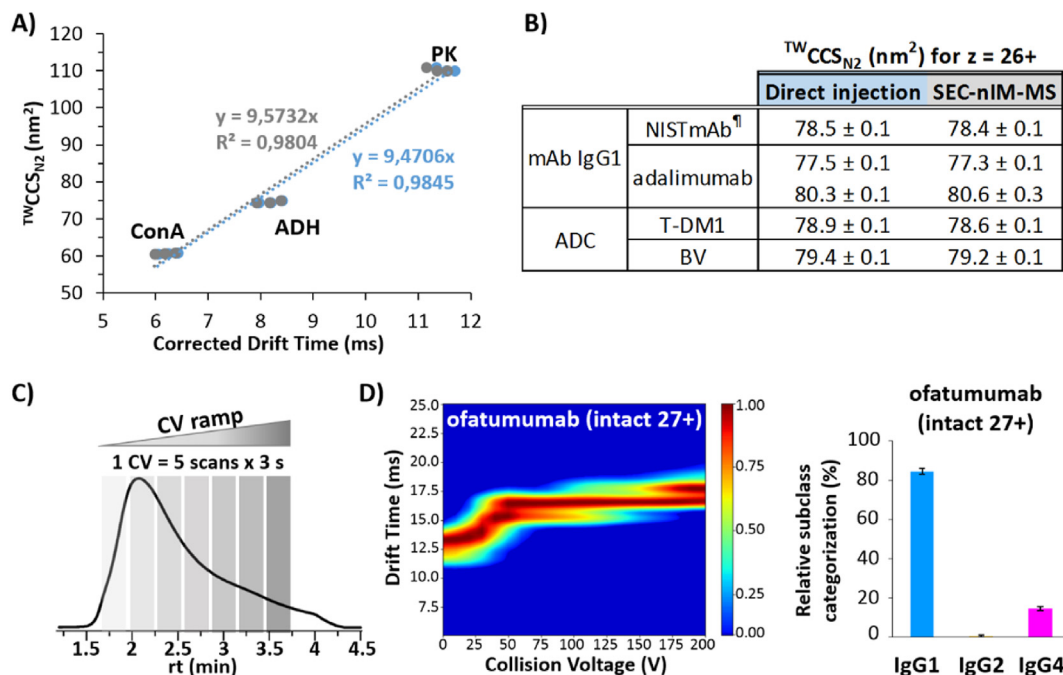


Fig. 4. SEC-nIM-MS experiments for CCS measurements and CIU approaches. (A) Calibration curves for SEC-nIM-MS (grey) and offline injection (blue). ConA, ADH and PK were used as external calibrants as reported by Bush et al. [56]. (B) Table summarizing $^{TW}CCS_{N_2}$ measurements at intact level obtained with direct nanoESI versus SEC injections. [†]All products were deglycosylated except for NISTmAb. (C) Schematic representation of SEC-CIU acquisitions. (D) SEC-CIU fingerprint and subclass classification obtained for intact ofatumumab. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

5. Hyphenation of non-denaturing LC through a SEC-based multidimensional methodology

SEC can be used for two main purposes: (i) separation and relative quantification of protein oligomers and (ii) to perform online buffer exchange prior to the analysis of proteins by nMS. However, as a consequence of the latter functionality, SEC can pave the way to allow the hyphenation of non-denaturing LC with nMS.

Coupling HIC to nMS: Development of HICxSEC-nMS strategy – Hydrophobic interaction chromatography (HIC) is the reference technique in quality control laboratories to separate and quantify the different drug-antibody ratio (DAR) populations of ADCs [60,61], especially in the case of cysteine-linked ADCs for which maintaining non-covalent interactions is critical for the analysis of the whole ADC scaffold. However, the analysis of the different populations of highly heterogeneous ADCs can lead to ambiguous interpretations when this analysis is based solely on the HIC dimension. To this end, nMS can afford additional information through the mass measurement of the intact species in their native state to enable the identification of the DAR populations. Since conventional HIC solvents are not compatible with nMS, several approaches have been proposed in the literature allowing the straight hyphenation of HIC with nMS. According to the main drawbacks associated with this coupling, the different experimental strategies were based on the modification of the composition of HIC mobile phase [62], the adoption of new HIC stationary phase materials [63] or the addition of a makeup flow followed by a splitting flow [64]. However, these strategies require the use of nMS-compatible mobile phases, that is, ammonium acetate, along with the addition of organic modifiers, which leads to the reduction of the peak separation and the potential denaturation of proteins. To this end, Ekhkirch et al. designed and adapted a 2D LC system where the SEC cartridge

(AdvanceBio SEC, 4.6 × 50 mm, 2.7 μm, 300 Å) is implemented within the interface of the HIC-MS coupling enabling online buffer exchange in the front end of the mass spectrometer [65] (Fig. 5A). This experimental setup was used to characterize the different populations of a conjugated mAb (Fig. 5B). In this case, only five ADC populations were expected (D0, D2, D4, D6, and D8), however one additional peak is observed at 40 min in the HIC profile (Fig. 5B). The mass measurement enables the assignment of the additional HIC feature as a DAR4 positional isomer (Fig. 5C). The all-in-one combination of HICxSEC-nMS provided a comprehensive and streamlined characterization of all the species observed within the first LC dimension without compromising neither the chromatography separation nor the native structure of the proteins.

2D SECxSEC-nMS setup for improved SEC performances – The first non-denaturing HICxSEC-nMS setup can be considered as the basis for the conception of new MS-based multidimensional LC coupling strategies [65]. The idea is to benefit from synergic effects of performing first dimension non-denaturing LC without compromising chromatographic performances while being compatible with nMS. The implementation of SEC for fast desalting between the first analytical dimension of separation and the mass spectrometer has thus a great potential to overcome the limitations of these techniques when used as standalone methods. Indeed, SEC-UV is usually performed with high concentration of non-volatile salts (typically around 100–500 mM) to reduce the non-specific interactions between the stationary phase and the size variant species [66]. When SEC is coupled to nMS and hence, classical SEC mobile phase is replaced by MS-compatible solvents, proteins may undergo more interaction with the stationary phase (ionic and hydrophobic interactions), especially those with a pI greater than 7.0, giving rise to broader peaks with larger retention times, and leading to an underestimation of the relative intensity of size variants [36]. This is illustrated by the example of the

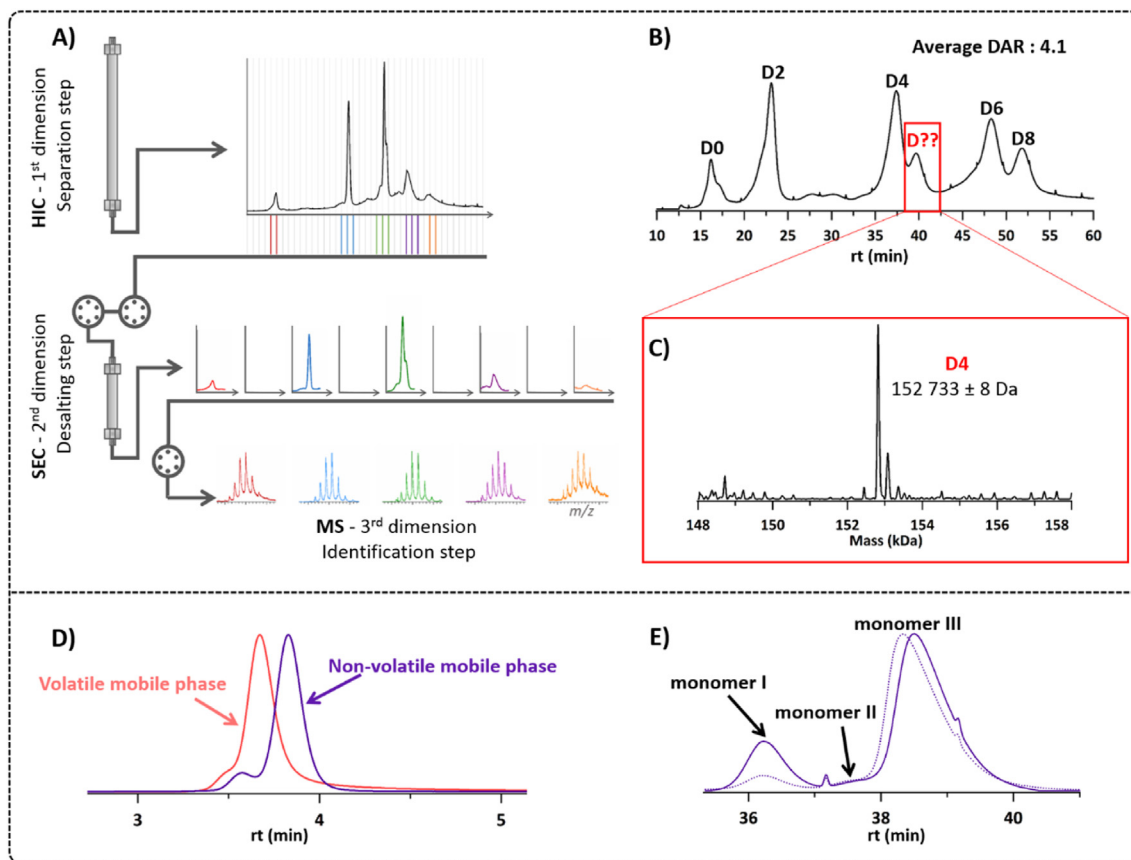


Fig. 5. 2D HICxSEC-nMS and SECxSEC-nMS experiments. (A) HICxSEC-nMS setup. (B) HIC profile of an in-house investigational ADC. (C) Deconvoluted mass of the highlighted HIC peak. (D) SEC profile of intact pembrolizumab with non-volatile and volatile salts in the mobile phase. (E) Comparison of SEC profile of intact pembrolizumab (solid line) and thermally-stressed pembrolizumab (dotted line).

characterization of HMWS and LMWS in the case of the therapeutic mAb pembrolizumab (Fig. 5D). Ehkirch et al. thus developed an experimental setup based on comprehensive 2D SECxSEC coupled to nMS to determine HMWS and LMWS of non-stressed and thermally-stressed mAbs [67]. The particular case of pembrolizumab pinpoints that the use of phosphate buffer in the first SEC dimension and the online combination of nMS are essential to provide baseline chromatography resolution and unambiguously assign mAb populations, respectively (Fig. 5D–E). The co-elution of the two chromatographic peaks when performing SEC with ammonium acetate mobile phase hinders the relative quantitation of both species. Furthermore, three distributions can be observed in pembrolizumab's chromatographic profile upon thermal stress (Fig. 5E), which may suggest the presence of different oligomer populations, that is, trimer, dimer, and monomer. However, nMS not only revealed that different chromatographic peaks corresponded to different types of monomeric populations (Fig. 5E), but also that those monomers had different degrees of oxidation.

The results obtained with these multidimensional LCxLC-IMxMS for therapeutics analysis under non-denaturing conditions pinpoints the complementarity of these techniques to expand the capabilities of non-denaturing LC and nMS techniques. In one hand, LC dimensions allow the separation and quantification of the different mAb populations based on the apparent hydrophobicity or size. On the other hand, nMS plays a pivotal role to unravel complex or unexpected chromatographic profiles affording a precise mass measurement of each individual chromatographic distribution.

6. Concluding remarks

In the present study, we exemplify the use of online SEC-nMS for a broader range of biomolecules ranging from homo-oligomeric proteins to proteins with nucleic acid, but also protein-small molecule complexes or even larger assemblies up to hundreds of kDa. SEC-nMS is suitable for most soluble protein-protein complexes and can also be adapted for a selected series of detergents for membrane protein analysis. Online SEC-nMS in most cases not only affords rapid and improved desalting efficiency but also separation of co-eluting/overlapping species. Online nMS is quite versatile as many types of commercial columns, but also manually packed ones, and LC systems are available that can be coupled to any nMS compliant mass spectrometer. In addition, the SEC-nMS coupling provides enough sensitivity to analyze low amounts of material (few micrograms) with accurate mass measurements. This increase in sensitivity could be of major interest to tackle low affinity ($K_a > 100 \mu\text{M}$) assemblies or complexes maintained by weak interactions.

SEC-nMS was also the basis for the development of more complicated multidimensional 2D LCxLC workflows in which the SEC acts as fast desalting device, allowing to achieve first dimension LC separation in optimal chromatographic conditions while being compatible with downstream nMS analysis. The combination of multidimensional LC with nMS provides a synergic effect for the comprehensive characterization of mAb-based therapeutics in one single run without limiting the technical capabilities of these techniques when used as standalone techniques.

Online SEC does not hamper further IM measurements affording

conformational characterization of samples either through automated CCS measurements or even for CIU experiments. In addition, further MS/MS experiments (reported for the moment only with collision-induced dissociation (CID) fragmentation) are possible. We expect ultraviolet photodissociation (UVPD) and electron-transfer dissociation (ETD) being possible also, opening the way for automatic top-down nMS.

CRedit authorship contribution statement

Evolène Deslignière: Formal analysis, Writing - original draft, Writing - review & editing. **Marie Ley:** Formal analysis. **Maxime Bourguet:** Formal analysis. **Anthony Ehkirch:** Formal analysis. **Thomas Botzanowski:** Formal analysis. **Stéphane Erb:** Formal analysis. **Oscar Hernandez-Alba:** Formal analysis, Writing - original draft, Writing - review & editing. **Sarah Cianféroni:** Supervision, Writing - original draft, Writing - review & editing, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

The authors declare no competing financial interest.

Acknowledgments

This work was supported by the CNRS, the University of Strasbourg and its IdeX program, the "Agence Nationale de la Recherche" (ANR grants ANR-16-CE11-0032 and ANR-19-CE11-0010), the French Proteomic Infrastructure (ProFI; ANR-10-INBS-08-03) and the Medalis LabEx. The authors would like to thank GIS IBISA and Région Alsace for financial support in purchasing a Synapt G2 HDMS instrument. E.D., A.E and T.B. acknowledge the French Ministry for Education and Research, the "Association Nationale de la Recherche et de la Technologie" (ANRT) and Syndivia, and the Institut de Recherche Servier, for funding of their PhD, respectively. M. B. was supported by a fellowship from the Région Alsace.

The authors would like to thank their collaborators for kindly providing samples: Isabelle Billas Massobrio (RXR/NR7/DNA complexes) and Jean Cavarelli (PRMT2) from the IGBMC (Strasbourg, France), Alain Wagner (mAb-oligonucleotide complex, BFC, Strasbourg, France), Florence Vincent (P1/P2) from University of Aix-Marseille (Marseille, France), Xavier Manival (Rvb) from IMoPA (Nancy, France), Natàlia Carulla from the IECB (Bordeaux, France) and Alain Beck from the Institut de Recherche Pierre Fabre (Saint-Julien-en-Genevois, France).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ijms.2020.116502>.

References

- [1] V. Katta, B.T. Chait, Observation of the heme-globin complex in native myoglobin by electrospray-ionization mass spectrometry, *J. Am. Chem. Soc.* 113 (22) (1991) 8534–8535, <https://doi.org/10.1021/ja00022a058>.
- [2] B. Ganem, Y.T. Li, J.D. Henion, Detection of noncovalent receptor-ligand complexes by mass spectrometry, *J. Am. Chem. Soc.* 113 (16) (1991) 6294–6296, <https://doi.org/10.1021/ja00016a069>.
- [3] A.A. Rostom, M. Sunde, S.J. Richardson, G. Schreiber, S. Jarvis, R. Bateman, C.M. Dobson, C.V. Robinson, Dissection of multi-protein complexes using mass spectrometry: subunit interactions in transthyretin and retinol-binding protein complexes, *Proteins Suppl* 2 (1998) 3–11, [https://doi.org/10.1002/\(sici\)1097-0134\(1998\)33:2<3:aid-prot2>3.3.co;2-8](https://doi.org/10.1002/(sici)1097-0134(1998)33:2<3:aid-prot2>3.3.co;2-8).

- [4] A.A. Rostom, C.V. Robinson, Detection of the intact GroEL chaperonin assembly by mass spectrometry, *J. Am. Chem. Soc.* 121 (19) (1999) 4718–4719, <https://doi.org/10.1021/ja990238r>.
- [5] G.K.A. Hochberg, J.L.P. Benesch, Dynamical structure of α B-crystallin, *Prog. Biophys. Mol. Biol.* 115 (1) (2014) 11–20, <https://doi.org/10.1016/j.pbiomolbio.2014.03.003>.
- [6] M.M. Jore, M. Lundgren, E. van Duijn, J.B. Bultema, E.R. Westra, S.P. Waghmare, B. Wiedenheft, Ü. Pul, R. Wurm, R. Wagner, M.R. Beijer, A. Barendregt, K. Zhou, A.P.L. Snijders, M.J. Dickman, J.A. Doudna, E.J. Boekema, A.J.R. Heck, J. van der Oost, S.J.J. Brouns, Structural basis for CRISPR RNA-guided DNA recognition by cascade, *Nat. Struct. Mol. Biol.* 18 (5) (2011) 529–536, <https://doi.org/10.1038/nsmb.2019>.
- [7] Martínez-Rucobo, W. Fuensanta, R. Kohler, M. van de Waterbeemd, Albert J.R. Heck, M. Hemann, F. Herzog, H. Stark, P. Cramer, Molecular basis of transcription-coupled pre-mRNA capping, *Mol. Cell.* 58 (6) (2015) 1079–1089, <https://doi.org/10.1016/j.molcel.2015.04.004>.
- [8] J.-M. Saliou, X. Manival, A.-S. Tillault, C. Atmanene, C. Bobo, C. Branlant, A. Van Dorsselaer, B. Charpentier, S. Cianféroni, Combining native MS approaches to decipher archaeal box H/ACA ribonucleoprotein particle structure and activity, *Proteomics* 15 (16) (2015) 2851–2861, <https://doi.org/10.1002/pmic.201400529>.
- [9] K.K. Viet, A. Wagner, K. Schwickert, N. Hellwig, M. Brennich, N. Bader, T. Schirmeister, N. Morgner, H. Schindelin, U.A. Hellmich, Structure of the human TRPML2 ion channel extracytosolic/luminal domain, *Structure* 27 (8) (2019) 1246–1257, <https://doi.org/10.1016/j.str.2019.04.016>, e5.
- [10] V.A. Jackson, S. Mehmood, M. Chavent, P. Roversi, M. Carrasquero, D. del Toro, G. Seyit-Bremer, F.M. Ranaivoson, D. Comolletti, M.S.P. Sansom, C.V. Robinson, R. Klein, E. Seiradake, Super-complexes of adhesion GPCRs and neural guidance receptors, *Nat. Commun.* 7 (1) (2016), <https://doi.org/10.1038/ncomms11184>.
- [11] C. Bechara, A. Nöll, N. Morgner, M.T. Degiacomi, R. Tampé, C.V. Robinson, A subset of annular lipids is linked to the flippase activity of an ABC transporter, *Nat. Chem.* 7 (3) (2015) 255–262, <https://doi.org/10.1038/nchem.2172>.
- [12] J. Marcoux, S.C. Wang, A. Politis, E. Reading, J. Ma, P.C. Biggin, M. Zhou, H. Tao, Q. Zhang, G. Chang, N. Morgner, C.V. Robinson, Mass spectrometry reveals synergistic effects of nucleotides, lipids, and drugs binding to a multidrug resistance efflux pump, *Proc. Natl. Acad. Sci. U.S.A.* 110 (24) (2013) 9704–9709, <https://doi.org/10.1073/pnas.1303888110>.
- [13] J. Gault, J.A.C. Donlan, I. Liko, J.T.S. Hopper, K. Gupta, N.G. Housden, W.B. Struwe, M.T. Marty, T. Mize, C. Bechara, Y. Zhu, B. Wu, C. Kleanthous, M. Belov, E. Damoc, A. Makarov, C.V. Robinson, High-resolution mass spectrometry of small molecules bound to membrane proteins, *Nat. Methods* 13 (4) (2016) 333–336, <https://doi.org/10.1038/nmeth.3771>.
- [14] L.L. Ilag, H. Videler, A.R. McKay, F. Sobott, P. Fucini, K.H. Nierhaus, C.V. Robinson, Heptameric (L12)6/L10 rather than canonical pentameric complexes are found by tandem MS of intact ribosomes from thermophilic bacteria, *Proc. Natl. Acad. Sci. U.S.A.* 102 (23) (2005) 8192–8197, <https://doi.org/10.1073/pnas.0502193102>.
- [15] S. Sanglier, E. Leize, A. Dorsselaer, F. Zal, Comparative ESI-MS study of ~2.2 MDa native hemocyanins from deep-sea and shore crabs: from protein oligomeric state to biotope, *J. Am. Soc. Mass Spectrom.* 14 (5) (2003) 419–429, [https://doi.org/10.1016/s1044-0305\(03\)00131-4](https://doi.org/10.1016/s1044-0305(03)00131-4).
- [16] S. Rozen, Füzesi-Levi, G. Maria, G. Ben-Nissan, L. Mizrahi, A. Gabashvili, Y. Levin, S. Ben-Dor, M. Eisenstein, M. Sharon, CSNAP is a stoichiometric subunit of the COP9 signalosome, *Cell Rep.* 13 (3) (2015) 585–598, <https://doi.org/10.1016/j.celrep.2015.09.021>.
- [17] L. Urnavicius, K. Zhang, A.G. Diamant, C. Motz, M.A. Schlager, M. Yu, N.A. Patel, C.V. Robinson, A.P. Carter, The structure of the dynactin complex and its interaction with dynein, *Science* 347 (6229) (2015) 1441–1446, <https://doi.org/10.1126/science.aaa4080>.
- [18] J. Snijder, R.J. Rose, D. Veesler, J.E. Johnson, A.J.R. Heck, Studying 18 MDa virus assemblies with native mass spectrometry, *Angew. Chem. Int. Ed.* 52 (14) (2013) 4020–4023, <https://doi.org/10.1002/anie.201210197>.
- [19] B.T. Ruotolo, G. Kevin, I. Campuzano, A.M. Sandercock, R.H. Bateman, C.V. Robinson, Evidence for macromolecular protein rings in the absence of bulk water, *Science* 310 (5754) (2005) 1658–1661, <https://doi.org/10.1126/science.1120177>.
- [20] Z. Hall, C.V. Robinson, Do charge state signatures guarantee protein conformations? *J. Am. Soc. Mass Spectrom.* 23 (7) (2012) 1161–1168, <https://doi.org/10.1007/s13361-012-0393-z>.
- [21] E. Duijn, Current limitations in native mass spectrometry based structural biology, *J. Am. Soc. Mass Spectrom.* 21 (6) (2010) 971–978, <https://doi.org/10.1016/j.jasms.2009.12.010>.
- [22] J. Cavanagh, L.M. Benson, R. Thompson, S. Naylor, In-line desalting mass spectrometry for the study of noncovalent biological complexes, *Anal. Chem.* 75 (14) (2003) 3281–3286, <https://doi.org/10.1021/ac030182q>.
- [23] G.M. Waitt, R. Xu, G.B. Wisely, J.D. Williams, Automated in-line gel filtration for native state mass spectrometry, *J. Am. Soc. Mass Spectrom.* 19 (2) (2008) 239–245, <https://doi.org/10.1016/j.jasms.2007.05.008>.
- [24] G. Van der Rest, F. Halgand, Size exclusion chromatography-ion mobility-mass spectrometry coupling: a step toward structural biology, *J. Am. Soc. Mass Spectrom.* 28 (11) (2017) 2519–2522, <https://doi.org/10.1007/s13361-017-1810-0>.
- [25] K. Muneeruddin, J.J. Thomas, P.A. Salinas, I.A. Kaltashov, Characterization of

- small protein aggregates and oligomers using size exclusion chromatography with online detection by native electrospray ionization mass spectrometry, *Anal. Chem.* 86 (21) (2014) 10692–10699, <https://doi.org/10.1021/ac502590h>.
- [26] Z.L. VanAernum, F. Busch, B.J. Jones, M. Jia, Z. Chen, S.E. Boyken, A. Sahasrabudde, D. Baker, V.H. Wysocki, Rapid online buffer exchange for screening of proteins, protein complexes and cell lysates by native mass spectrometry, *Nat. Protoc.* 15 (3) (2020) 1132–1157, <https://doi.org/10.1038/s41596-019-0281-0>.
- [27] T. Botzanowski, S. Erb, O. Hernandez-Alba, A. Ekhkirch, O. Colas, E. Wagner-Rousset, D. Rabuka, A. Beck, P.M. Drake, S. Cianferani, Insights from native mass spectrometry approaches for top- and middle- level characterization of site-specific antibody-drug conjugates, *mAbs* 9 (5) (2017) 801–811, <https://doi.org/10.1080/19420862.2017.1316914>.
- [28] A. Ekhkirch, O. Hernandez-Alba, O. Colas, A. Beck, D. Guillaume, S. Cianferani, Hyphenation of size exclusion chromatography to native ion mobility mass spectrometry for the analytical characterization of therapeutic antibodies and related products, *J. Chromatogr. B: Anal Technol Biomed Life Sci* 1086 (2018) 176–183, <https://doi.org/10.1016/j.jchromb.2018.04.010>.
- [29] J. Woodard, H. Lau, R.F. Latypov, Nondenaturing size-exclusion chromatography-mass spectrometry to measure stress-induced aggregation in a complex mixture of monoclonal antibodies, *Anal. Chem.* 85 (13) (2013) 6429–6436, <https://doi.org/10.1021/ac401455f>.
- [30] S.M. Hengel, R. Sanderson, J. Valliere-Douglass, N. Nicholas, C. Leiske, S.C. Alley, Measurement of in vivo drug load distribution of cysteine-linked antibody–drug conjugates using microscale liquid chromatography mass spectrometry, *Anal. Chem.* 86 (7) (2014) 3420–3425, <https://doi.org/10.1021/ac403860c>.
- [31] M. Habegger, M. Leiss, A.-K. Heidenreich, O. Pester, G. Hafenmair, M. Hook, L. Bonnington, H. Wegele, M. Haindl, D. Reusch, P. Bulau, Rapid characterization of biotherapeutic proteins by size-exclusion chromatography coupled to native mass spectrometry, *mAbs* 8 (2) (2015) 331–339, <https://doi.org/10.1080/19420862.2015.1122150>.
- [32] J. Jones, L. Pack, J.H. Hunter, J.F. Valliere-Douglass, Native size-exclusion chromatography-mass spectrometry: suitability for antibody–drug conjugate drug-to-antibody ratio quantitation across a range of chemotypes and drug-loading levels, *mAbs* 12 (1) (2019) 1682895, <https://doi.org/10.1080/19420862.2019.1682895>.
- [33] E. Deslignière, A. Ekhkirch, T. Botzanowski, A. Beck, O. Hernandez-Alba, S. Cianferani, Toward automation of collision-induced unfolding experiments through online size exclusion chromatography coupled to native mass spectrometry, *Anal. Chem.* 92 (19) (2020) 12900–12908, <https://doi.org/10.1021/acs.analchem.0c01426>.
- [34] A. Goyon, A. Beck, O. Colas, K. Sandra, D. Guillaume, S. Fekete, Evaluation of size exclusion chromatography columns packed with sub-3 μm particles for the analysis of biopharmaceutical proteins, *J. Chromatogr. A* 1498 (2017) 80–89, <https://doi.org/10.1016/j.chroma.2016.11.056>.
- [35] I.K. Ventouri, D.B.A. Malheiro, R.L.C. Voeten, S. Kok, M. Honing, G.W. Somsen, R. Haselberg, Probing protein denaturation during size-exclusion chromatography using native mass spectrometry, *Anal. Chem.* 92 (6) (2020) 4292–4300, <https://doi.org/10.1021/acs.analchem.9b04961>.
- [36] A. Goyon, V. D'Atri, O. Colas, S. Fekete, A. Beck, D. Guillaume, Characterization of 30 therapeutic antibodies and related products by size exclusion chromatography: feasibility assessment for future mass spectrometry hyphenation, *J. Chromatogr. B: Anal Technol Biomed Life Sci* 1065–1066 (2017) 35–43, <https://doi.org/10.1016/j.jchromb.2017.09.027>.
- [37] L. Koneremann, Addressing a common misconception: ammonium acetate as neutral pH “buffer” for native electrospray mass spectrometry, *J. Am. Soc. Mass Spectrom.* 28 (9) (2017) 1827–1835, <https://doi.org/10.1007/s13361-017-1739-3>.
- [38] S. Fekete, A. Beck, J.-L. Veuthey, D. Guillaume, Theory and practice of size exclusion chromatography for the analysis of protein aggregates, *J. Pharmaceut. Biomed. Anal.* 101 (2014) 161–173, <https://doi.org/10.1016/j.jpba.2014.04.011>.
- [39] V. Vivat Hannah, C. Atmanene, D. Zeyer, A. Van Dorsselaer, S. Sanglier-Cianferani, Native MS: an ‘ESL’ way to support structure- and fragment-based drug discovery, *Future Med. Chem.* 2 (1) (2010) 35–50, <https://doi.org/10.4155/fmc.09.141>.
- [40] L. Pedro, R. Quinn, Native mass spectrometry in fragment-based drug discovery, *Molecules* 21 (8) (2016) 984, <https://doi.org/10.3390/molecules21080984>.
- [41] H. Vu, N.B. Pham, R.J. Quinn, Direct screening of natural product extracts using mass spectrometry, *J. Biomol. Screen* 13 (4) (2008) 265–275, <https://doi.org/10.1177/1087057108315739>.
- [42] C. Ren, A.O. Bailey, E. VanderPorten, A. Oh, W. Phung, M.M. Mulvihill, S.F. Harris, Y. Liu, G. Han, W. Sandoval, Quantitative determination of protein–ligand affinity by size exclusion chromatography directly coupled to high-resolution native mass spectrometry, *Anal. Chem.* 91 (1) (2018) 903–911, <https://doi.org/10.1021/acs.analchem.8b03829>.
- [43] H.J. Maple, O. Scheibner, M. Baumert, M. Allen, R.J. Taylor, R.A. Garlish, M. Bromirski, R.J. Burnley, Application of the Exact Plus EMR for automated protein–ligand screening by non-covalent mass spectrometry, *Rapid Commun. Mass Spectrom.* 28 (13) (2014) 1561–1568, <https://doi.org/10.1002/rcm.6925>.
- [44] M. Jia, S. Sen, C. Wachnowsky, I. Fidai, J.A. Cowan, V.H. Wysocki, Characterization of [2Fe–2S]-Cluster-Bridged protein complexes and reaction intermediates by use of native mass spectrometric methods, *Angew. Chem. Int. Ed.* 59 (17) (2020) 6724–6728, <https://doi.org/10.1002/anie.201915615>.
- [45] D. Lowe, K. Dudgeon, R. Rouet, P. Schofield, L. Jeremutus, D. Christ, Aggregation, stability, and formulation of human antibody therapeutics, *Adv. Protein Chem. Struct. Biol.* 84 (2011) 41–61, <https://doi.org/10.1016/b978-0-12-386483-3.00004-5>.
- [46] V. Cura, N. Marechal, N. Troffer-Charlier, J.M. Strub, M.J. Haren, N.I. Martin, S. Cianferani, L. Bonnefond, J. Cavarelli, Structural studies of protein arginine methyltransferase 2 reveal its interactions with potential substrates and inhibitors, *FEBS J.* 284 (1) (2016) 77–96, <https://doi.org/10.1111/febs.13953>.
- [47] A. Jeganathan, V. Leong, L. Zhao, J. Huen, N. Nano, W.A. Houry, J. Ortega, Yeast Rvb1 and Rvb2 proteins oligomerize as a conformationally variable dodecamer with low frequency, *J. Mol. Biol.* 427 (10) (2015) 1875–1886, <https://doi.org/10.1016/j.jmb.2015.01.010>.
- [48] K.L.Y. Cheung, J. Huen, Y. Kakiyama, W.A. Houry, J. Ortega, Alternative oligomeric states of the yeast rvb1/rvb2 complex induced by histidine tags, *J. Mol. Biol.* 404 (3) (2010) 478–492, <https://doi.org/10.1016/j.jmb.2010.10.003>.
- [49] K. Mohideen-Abdul, K. Tazibt, M. Bourguet, I. Hazemann, I. Lebars, M. Takacs, S. Cianferani, B.P. Klaholz, D. Moras, I.M.L. Billas, Importance of the sequence-directed DNA shape for specific binding site recognition by the estrogen-related receptor, *Front. Endocrinol.* 8 (2017), <https://doi.org/10.3389/fendo.2017.00140>.
- [50] N. Rochel, C. Crucker, L. Coutos-Thevenot, J. Osz, R. Zhang, E. Guyon, W. Zita, S. Vanthong, O.A. Hernandez, M. Bourguet, K.A. Badawy, F. Dufour, C. Peluso-Itlis, S. Heckler-Beji, A. Dejaegere, A. Kamoun, A. de Reynies, Y. Neuzillet, S. Rebouissou, C. Beraud, H. Lang, T. Massfelder, Y. Allory, S. Cianferani, R.H. Stote, F. Radvanyi, I. Bernard-Pierrot, Recurrent activating mutations of PPARgamma associated with luminal bladder tumors, *Nat. Commun.* 10 (1) (2019) 253, <https://doi.org/10.1038/s41467-018-08157-y>.
- [51] S. Ciudad, E. Puig, T. Botzanowski, M. Meigouni, A.S. Arango, J. Do, M. Mayzel, M. Bayoumi, S. Chaignepain, G. Maglia, S. Cianferani, V. Orekhov, E. Tajkhorshid, B. Bardiaux, N. Carulla, A β (1–42) tetramer and octamer structures reveal edge conductivity pores as a mechanism for membrane damage, *Nat. Commun.* 11 (1) (2020), <https://doi.org/10.1038/s41467-020-16566-1>.
- [52] J.E. Keener, G. Zhang, M.T. Marty, Native mass spectrometry of membrane proteins, *Anal. Chem.* (2020), <https://doi.org/10.1021/acs.analchem.0c04342>.
- [53] A.C. Leney, A.J. Heck, Native mass spectrometry: what is in the name? *J. Am. Soc. Mass Spectrom.* 28 (1) (2017) 5–13, <https://doi.org/10.1007/s13361-016-1545-3>.
- [54] A. Konijnenberg, A. Butterer, F. Sobott, Native ion mobility-mass spectrometry and related methods in structural biology, *Biochim. Biophys. Acta* 1834 (6) (2013) 1239–1256, <https://doi.org/10.1016/j.bbapap.2012.11.013>.
- [55] G. Terral, A. Beck, S. Cianferani, Insights from native mass spectrometry and ion mobility-mass spectrometry for antibody and antibody-based product characterization, *J. Chromatogr. B: Anal Technol Biomed Life Sci* 1032 (2016) 79–90, <https://doi.org/10.1016/j.jchromb.2016.03.044>.
- [56] M.F. Bush, Z. Hall, K. Giles, J. Hoyes, C.V. Robinson, B.T. Ruotolo, Collision cross sections of proteins and their complexes: a calibration framework and database for gas-phase structural biology, *Anal. Chem.* 82 (22) (2010) 9557–9565, <https://doi.org/10.1021/ac102295j>.
- [57] S.M. Dixit, D.A. Polasky, B.T. Ruotolo, Collision induced unfolding of isolated proteins in the gas phase: past, present, and future, *Curr. Opin. Chem. Biol.* 42 (2018) 93–100, <https://doi.org/10.1016/j.cbpa.2017.11.010>.
- [58] Y. Tian, L. Han, A.C. Buckner, B.T. Ruotolo, Collision induced unfolding of intact antibodies: rapid characterization of disulfide bonding patterns, glycosylation, and structures, *Anal. Chem.* 87 (22) (2015) 11509–11515, <https://doi.org/10.1021/acs.analchem.5b03291>.
- [59] T. Botzanowski, O. Hernandez-Alba, M. Malissard, E. Wagner-Rousset, E. Deslignière, O. Colas, J.-F. Haeuw, A. Beck, S. Cianferani, Middle level IM-MS and CIU experiments for improved therapeutic immunoglobulin subclass fingerprinting, *Anal. Chem.* 92 (13) (2020) 8827–8835, <https://doi.org/10.1021/acs.analchem.0c00293>.
- [60] M. Rodriguez-Aller, D. Guillaume, A. Beck, S. Fekete, Practical method development for the separation of monoclonal antibodies and antibody-drug-conjugate species in hydrophobic interaction chromatography, part 1: optimization of the mobile phase, *J. Pharmaceut. Biomed. Anal.* 118 (2016) 393–403, <https://doi.org/10.1016/j.jpba.2015.11.011>.
- [61] A. Cusumano, D. Guillaume, A. Beck, S. Fekete, Practical method development for the separation of monoclonal antibodies and antibody-drug-conjugate species in hydrophobic interaction chromatography, part 2: optimization of the phase system, *J. Pharmaceut. Biomed. Anal.* 121 (2016) 161–173, <https://doi.org/10.1016/j.jpba.2016.01.037>.
- [62] B. Wei, G. Han, J. Tang, W. Sandoval, Y.T. Zhang, Native hydrophobic interaction chromatography hyphenated to mass spectrometry for characterization of monoclonal antibody minor variants, *Anal. Chem.* 91 (24) (2019) 15360–15364, <https://doi.org/10.1021/acs.analchem.9b04467>.
- [63] B. Chen, Y. Peng, S.G. Valeja, L. Xiu, A.J. Alpert, Y. Ge, Online hydrophobic interaction chromatography–mass spectrometry for top-down proteomics, *Anal. Chem.* 88 (3) (2016) 1885–1891, <https://doi.org/10.1021/acs.analchem.5b04285>.
- [64] Y. Yan, T. Xing, S. Wang, T.J. Daly, N. Li, Online coupling of analytical hydrophobic interaction chromatography with native mass spectrometry for the characterization of monoclonal antibodies and related products, *J. Pharmaceut. Biomed. Anal.* 186 (2020) 113313, <https://doi.org/10.1016/>

- [jjpba.2020.113313](https://doi.org/10.1016/j.jpba.2020.113313).
- [65] A. Ehkirch, V. D'Atri, F. Rouviere, O. Hernandez-Alba, A. Goyon, O. Colas, M. Sarrut, A. Beck, D. Guillaume, S. Heinisch, S. Cianferani, An online four-dimensional HIC×SEC-IM×MS methodology for proof-of-concept characterization of antibody drug conjugates, *Anal. Chem.* 90 (3) (2018) 1578–1586, <https://doi.org/10.1021/acs.analchem.7b02110>.
- [66] A. Goyon, A. Beck, J.-L. Veuthey, D. Guillaume, S. Fekete, Comprehensive study on the effects of sodium and potassium additives in size exclusion chromatographic separations of protein biopharmaceuticals, *J. Pharmaceut. Biomed. Anal.* 144 (2017) 242–251, <https://doi.org/10.1016/j.jpba.2016.09.031>.
- [67] A. Ehkirch, A. Goyon, O. Hernandez-Alba, F. Rouviere, V. D'Atri, C. Dreyfus, J.F. Haeuw, H. Diemer, A. Beck, S. Heinisch, D. Guillaume, S. Cianferani, A novel online four-dimensional SEC×SEC-IM×MS methodology for characterization of monoclonal antibody size variants, *Anal. Chem.* 90 (23) (2018) 13929–13937, <https://doi.org/10.1021/acs.analchem.8b03333>.

Annexe 2

Tableau récapitulatif des séquences des protéines étudiées au cours de ces travaux de thèse.

| | |
|------------------------------------|--|
| A _{2A} rant21 | MWSHPQFEKHHHHHHHENLYFQGPIMGSSVYITVELAIAVLAILGNVLCWAVWLNSNLQNV TNYFVVSLLAAADIAGVGLAIPFAITISTGFCAACHGCLFIACFVLVLTQSSIFSLIAIAIDRY IAIRIPLRYNGLVTGTRAKGIIAICWVLSFAIGLTPMLGWNNCGQPKEGKNHSQGCQVAC LFEDVPMNYMVFNFACVLVPLLLMLGVYLRIFLAARRQLKQMESQPLPGERARSTLQKEV HAAKSLAIIIVGLFALCWLPLHIINCFTEFCPDCSHAPLWMLYLAIVLSHTNSVVPFIYAYRI REFRQTFRKIIRSHVLRQQEPFKAAGTSARVLAHGSQVSLRLNGHPPGVWANGSAPHPE RRPNGYALGLVSGGSAQESQNTGLPDVELLSHELKGVCPPEPGLDDPLAQDGAGVS |
| ADN (ERR α /PGC1 α) | AGTCCCATGTCAAGGTCACCGTGACCTTTACGAATTTAATAA (+brin complémentaire) |
| BSA | DTHKSEIAHRFKDLGEEHFGLVLIASFQYLQCPDFEHVKLVNELTEFAKTCVADESHAGCE KSLHTLFGDELCKVASLRETYGDMADCCCEKQEPERNECFLSHKDDSPDLPLKLPDPNTLCDEF KADEKFFWGKYLIEIARRHPYFYAPELLYYANKYNGVFECCQAEDKGACLLPKIETMREKVL TSSARQLRCASIQKFERALKAWSVARLSQKFPKAEFVEVTKLVTDLTKVHKECCHGDLLEC ADDRADLAKYICDNQDTISSKLECCDKPLLEKSHCIAEVEKDAIPENLPLTADFAEDKDVC KNYQEAQDAFLGSLYYSRRHPEYAVSVLLRLAKEYEATLECCAKDDPHACYSTVFDKLVK LVDEPQNLKQNCQDFEKLGEYGFQNALIVRYTRKVPQVSTPTLVEVSRSLGKVGTRCCTKPE SERMPCTEDYLSLILNRLCVLHEKTPVSEKVTKCCTESLVNRRPCFSALTPDETYVPKAFDEK LFTFHADICTLPDTEKQIKKQALVELLKHKPKATEEQKKTVMENFVAFVDKCAADDKEACF AVEGPKLVVSTQTALA |
| DPCD | GPHMAVTGWLESLRTAQKTALLQDGRRKVHYLFPDGKEMAEYDEKTSSELLVRKWRVKSALGA MGQWQLEVGDPAPLGAGNLGPELIKESNANPIFMRKDTKMSFQWRIRNLPYPKDVSVSVDQK ERICIIVRTTNKKYKKSIPDLDRHQLPLDDASLSFAHANCTLIISYQKPKVEVVAESELQKE LKKVKTAHSNDGDCKTQ |
| EPAC1 | MGSSHHHHHSSGLVPRGSHMENLYFQGMVLRMRHRPRSCSYQLLLEHQRPSCIQGLRWTPLT NSEESLDFSESLEQASTERVLRAGRQLHRHLLATCPNLIRDRKYHLRLYRQCCSGRELVDGIL ALGLGVHSRSQVVGICQVLLDEGALCHVKHDWAFQDRDAQFYRFPGPPEPEPVGTHEMEEELAE AVALLSQRGPDALLTVALRKPPGQRTDEELDLIFEELLHIKAVAHLSNSVKRELAAVLLFEPH SKAGTVLFSQGDKGTSWYIIWKGSVNVVTHGKGLVTTLHEGDDFGQLALVNDAPRAATIIILRE DNCHFRLVDKQDFNRIIKDVEAKTMRL EEHGKVVLLVERASQGAGPSRPPTGRNRYTVMSGT PEKILELLL EAMGPDSSAHDPTEFTLSDFLLTHRVMPSAQLCAALLHHFHVEPAGGSEQERS TYVCNKRQQILRLVSQWVALYGSMLHTDPVATSF LQKLSDLVGRDTRLNLLREQWPERRCH RLENGCGNASPQMKARNLPVWLPNQDEPLPGSSCAIQVGDKVPYDICRPDHSVLTMQPLPTAS VREVMAALAEQEDGWTGQVLVKVN SAGDAIGLQPDARGVATSLGLNERLFVWNPQEVHELIPH PDQLGPTVGS AEGDLVSAKDLAQQLTDHDWSL FNSIHQVELIHYVLGPQHLRDVTTANLERF MRRFNEQLQYVATELCLCPVGPRAQLLRKFIKLAHLKEQKNLNSFFAVMFGLSNSAISRLA HTWERLPHKVRKLYSALERLLDPSWNHRVYRLALAKLSPPVIPFMPLLLKDMTFIHEGNHTLV ENLINF EKMRMARAARMLHCHRSHNPVPLSPLRSRVSHLHEDSVARISTCSEQSLSTRSPAS TWAYVQQLKVIDNQRELSRLSRELEP |
| ERR α | GHMSSQVVGIEPLYIKAEPASDPSPKGSSETETEPVALAPGPAPTRCLPGHKEEEDGEGAGP GEQGGGKLVLSL PKRLCLVCGDVASGYHYGVASCEACKAFFKRTIQGSI EYSCPASNECEIT KRRRKACQACRFKTLRVGMLKEGVRLDRVRGGRQYKRRPEVDPLFPFPFPAGPLAVAGGP RKTAAPVNALVSHLLVVEPEKLYAMPDPAGPDGHLPAVATLCDLDFREIVVTISWAKSIPGFS SLSLSDQMSVLQSVWMEVLVLGVAQRSLPLQDELAF AEDLVLDEEGARAAGLGELGAALLQLV RRLQALRLEREYVLLKALALANSDSVHIEDAEAVEQLREALHEALLEYEAGRAGPGGGAERR RAGRLLLTLP LLRQTAGKVL AHFYGVKLEGVPMHKLFL EMLEAMMD |
| LSD1 | LSGKKA AAAAAAAAAAATGTEAGPGTAGGSENGSEVAAQ PAGLSGPAEVGPGAVGERTPRKKE PPRASPPGGLAEPPGSAGPQAGPTVVP GSATPMETGIAETPEGRRTSRRKRAKVEYREMDSE ANLSEDEYYSEEERNAKAEKEKLP P P P P P QAPPEEENESEPEEPSGVEGA AFQSRLPHDRMTS QEAACFPDIISGPQQTQKVF L FIRNRTLQLWLDNPKIQLTFEATLQQL EAPYNSD TVLVHRVH SYLERHGLINFGIYKRIKPLPTKKTGKVIIIGSGVSLAAARQLQSF GMDVTLLEARDRVGGR VATFRKGNVYADLGAMVVTGLGPNMAVVSQVNMELAKIKQKCLYEANGQAVPKEKDEMVE QEFNRLLEATSYLSHQLD FNVLNKPVSLGQALEVVIQLQEKHVKDEQIEHWKKIVKTQEE LK ELLNKMVNLKEKIKELHQYKEASEVKPPRDITAEFLVKS KHRDLTALCKEYDELAETQ GKLE EKLQELEANPPSDVYLSRRDRQILDWHFANLEFANATPLSTLSLKHWDQDDDFEFTGSHLTVR NGYSCVPVALAEGLDIKLNTAVRQVRYTASGCEVIAVNTRSTSQTFFIYKDAVLC TLPLGV LK |

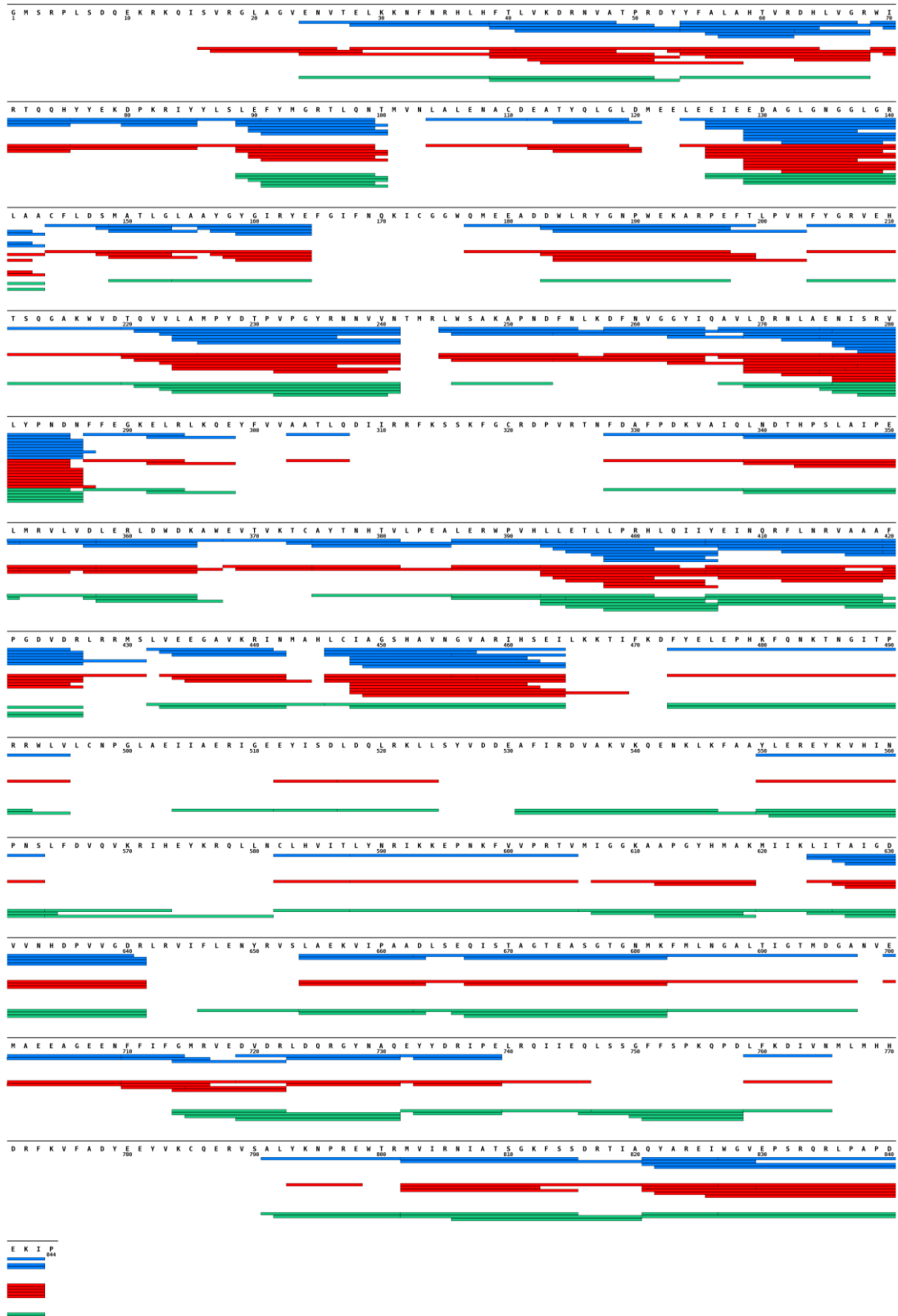
| | |
|-----------------|---|
| | <p>QQPPAVQFVPLPEWKTSAVQRMFGNLNKVVLFCFDRVFWDPVSNLFGHVGSTTASRGELFLF WNLYKAPILLALVAGEAAGIMENISDDVIVGRCLAILKGFSSAVPQPKEVTVSRWRADPWA RGSYSYVAAGSSGNDYDLMAQPIITPGPSIPGAPQPIPRLFFAGEHTIRNYPATVHGALLSGLR EAGRIADQFLGAMYTLRQATPGVPAQQSPSM</p> |
| Nivolumab | <p>Lc :EIVLTQSPATLSLSPGERATLSCRASQSVSSYLAWYQQKPGQAPRLLIYDASNRATGIP ARFSGSGSGTDFTLTISSELEPEDFAVYYCQQSSNWPRTFGQGTKEIKRTVAAPSVFIFPPSD EQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDYSLSTLTLSKAD YEKHKVYACEVTHQGLSSPVTKSFNRGEC Hc :VQLVESGGGVVQPGRSLRLDCKASGITFSNSGMHWVRQAPGKGLEWVAVIWYDGSKRYY ADSVKGRFTISRDNKNTLFLQMNSLRAEDTAVYYCATNDDYWGQGLTIVSSASTKGPSVFP LAPCSRSTSESTAALGCLVKDYFPEPVTVSWNSGAL TSGVHTFPAVLQSSGLYSLSSVTVPS SSLGKTYTCNVDHKPSNTKVDKRVESKYGPPCPPCPAPEFLGGPSVFLFPPKPKDTLMISRT PEVTCVVVDVSDQEDPEVQFNWYVDGVEVHNAKTKPREEQFNSTYRVVSVLTVLHQDNLNGKEY KCKVSNKGLPSSIEKTIKAKGQPREPQVYTLPPSQEEMTKNQVSLTCLVKGFYPSDIAVEWE SNGQPENNYKTTPPVLDSDGSFFLYSRLTVDKSRWQEGNVFSCSVMHEALHNHYTQKSLSL G</p> |
| Pembrolizumab | <p>Lc :EIVLTQSPATLSLSPGERATLSCRASKGVSTSGYSYLHWYQQKPGQAPRLLIYLASYLE SGVPARFSGSGSGTDFTLTISSELEPEDFAVYYCQHSRDLPLTFGGGTKEIKRTVAAPSVFIF PPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDYSLSTLT SKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC Hc :VQLVESGGGVVQPGRSLRLDCKASGITFSNSGMHWVRQAPGKGLEWVAVIWYDGSKRYY ADSVKGRFTISRDNKNTLFLQMNSLRAEDTAVYYCATNDDYWGQGLTIVSSASTKGPSVFP LAPCSRSTSESTAALGCLVKDYFPEPVTVSWNSGAL TSGVHTFPAVLQSSGLYSLSSVTVPS SSLGKTYTCNVDHKPSNTKVDKRVESKYGPPCPPCPAPEFLGGPSVFLFPPKPKDTLMISRT PEVTCVVVDVSDQEDPEVQFNWYVDGVEVHNAKTKPREEQFNSTYRVVSVLTVLHQDNLNGKEY KCKVSNKGLPSSIEKTIKAKGQPREPQVYTLPPSQEEMTKNQVSLTCLVKGFYPSDIAVEWE SNGQPENNYKTTPPVLDSDGSFFLYSRLTVDKSRWQEGNVFSCSVMHEALHNHYTQKSLSL G</p> |
| PGC1α | <p>MRSHHHHHHPGLVPRGSMAMDCNQDSESVWSDIECAALVGEDQPLCPDLPELDLSELDVND LDTDSFLGGLKWCSDQSEIISNQYNNEPSNIFEKIDEENEANLLAVLTETLDSLVPDEDGLPS FDALTDGDVTTDNEASPSSMPDGTTPPPQEAEPSELLKLLAPANTQLSYNECSGLSTQNHAN HNHRIRTNPAIVKTENSWSNKAKSICQQKQRRRPSSELLKYLTTNDPPHTKPTENRNSRD KCTSKKKSHTQSQSQHLQAKPTTSLPLTPESPNDPKGSPFENKTIERTL</p> |
| Phosphorylase B | <p>GMSRPLSDQEKRKQISVRGLAGVENVTELKKNFNRLHFTLVKDRNVATPRDYFALAHTVRD HLVGRWIRTQQHYEKPDKRIYLSLEFYMGRTLQNTMVNLALACDEATYQLGLDMEEL IEEDAGLNGGLGRLAACFLDSMATLGLAAYGYGIRYEFGIFNQKICGGWQMEEADDWLRYGN PWEKARPEFTLPVHFYGRVEHTSQGAKWVDTQVVLAMPYDTPVPGYRNNVNTMRLWSAKAPN DFNLKDFNVGGYIQAVALDRNLAENISRVLYPNDNFFEGKELRLKQYFVVAATLQDIIRRFKS SKFGCRDPVRTNFDAFPDKVAIQLNDTHPSLAPELMRVLDLERLDWDKAWVTVKTCAYTN HTVLPALERWVHLLLETLLPRHLQIYEINQRF LNRVAAAFPGDVRRLRRMSLVEEGAVKRI NMAHLCIAGSHAVNGVARIHSEILKKTIFKDFYELEPHKFNKNTNGITPRRWLVLCPGLAEI IAERIGEEYISDLQRLKLLSYVDDEAFIRDVAKVKQENKLFKFAAYLREYKVHINPNSLFDV QVKRIHEYKRQLLNCLHVITLYNRIKKEPNKFVVPRTVMIGGKAAPGYHMAKMIKLIITAIGD VNVHDPVVGDRLRVIFLENYRVSLAEKVIPAADLSEQISTAGTEASGTGNMFKMLNGALTIGT MDGANVEMAEAGEENFFIFGMRVEDVDRLDQRGYNAQEYYDRIPELRQIIEQLSSGFFSPKQ PDLFKDIVNMLMHHDRFKVFADYEEYVKCQERVSALYKNPREWTRMVIRNIATS GKFSSDRTI AQYAREIWGVEPSRQRLPAPDEKIP</p> |
| PIH1D2 | <p>GPHMETSSKGLLTQVTQFWNLDDLAQSDPEGYEKF IQQQKKEGKQLCAAPEPQLCLQTRILK PKEKILFINLCQWTRIPAPQSTTHPVPLTVGKPEDTTEISDAYTVIDVAYNPVLAHAAEKDQV KKNQLIQMAMKCIEEKFQFTLSHSYHITKFRIKGSIQRMKQNL MGIQTDSDILREKMRRELT GQIRSSMTSNPDHFPQLLLPKDQVSGKAVCLIEEISSTEIQVEMKMPAYELKIVHDHSEKPLK IELKVELPGINSVSLCDLSVSEDDLLIEVSEKYRLHLNLPKLIIDEMTTAKFIKEKSTLIITM PLV</p> |
| PPARγ (F310S) | <p>MGSSHHHHHSSGLVPRGSHMQLNPESADLRALAKHLYDSYIKSFPLTKAKARAILTGKTTDK SPFVIYDMNSLMMGEDKIKFKHITPLQEQSKEVAIRISQGCQFRSVEAVQEI TEYAKSIPGFV NLDLNDQVTLKYGVHEIYTMASL MNKDGVLISEGQGFMTREFLKSRLKPFQDFMEPKFEF AVKFNALDSDLAIFIAVIIISGDRPGLLNKPIEDIQDNLLQALELQLKLNHPSSQLFA KLLQKMTDLRQIVTEHVQLLQVIKKTETDMSLHPLLQEIYKDLY</p> |

| | |
|-----------------------|--|
| PPAR γ (H494Y) | MGSSHHHHHSSGLVPRGSHMQLNPESADLRALAKHLYDSYIKSFPLTKAKARAILTGKTTDK SPFVIYDMNSLMMGEDKIKFKHITPLQEQSKEVAIRIFQGCQFRSVEAVQEITEYAKSIPGFV NLDLNDQVTLKYGVEHIIYTMLASLMNKDGVLISEGQGFMTREFLKSRLKPFQDFMEPKFEF AVKFNALELDDSDLAIFIAVIIISGDRPGLLNKPIEDIQDNLQALELQLKLNHPESQLFA KLLQKMTDLRQIVTEHVQLLQVIKKTETDMSLYPLLQEIKDLY |
| PPAR γ (T475M) | MGSSHHHHHSSGLVPRGSHMQLNPESADLRALAKHLYDSYIKSFPLTKAKARAILTGKTTDK SPFVIYDMNSLMMGEDKIKFKHITPLQEQSKEVAIRIFQGCQFRSVEAVQEITEYAKSIPGFV NLDLNDQVTLKYGVEHIIYTMLASLMNKDGVLISEGQGFMTREFLKSRLKPFQDFMEPKFEF AVKFNALELDDSDLAIFIAVIIISGDRPGLLNKPIEDIQDNLQALELQLKLNHPESQLFA KLLQKMTDLRQIVMEHVQLLQVIKKTETDMSLHPLLQEIKDLY |
| PPAR γ (WT) | MGSSHHHHHSSGLVPRGSHMQLNPESADLRALAKHLYDSYIKSFPLTKAKARAILTGKTTDK SPFVIYDMNSLMMGEDKIKFKHITPLQEQSKEVAIRIFQGCQFRSVEAVQEITEYAKSIPGFV NLDLNDQVTLKYGVEHIIYTMLASLMNKDGVLISEGQGFMTREFLKSRLKPFQDFMEPKFEF AVKFNALELDDSDLAIFIAVIIISGDRPGLLNKPIEDIQDNLQALELQLKLNHPESQLFA KLLQKMTDLRQIVTEHVQLLQVIKKTETDMSLHPLLQEIKDLY |
| PPAR γ (Y505H) | MGSSHHHHHSSGLVPRGSHMQLNPESADLRALAKHLYDSYIKSFPLTKAKARAILTGKTTDK SPFVIYDMNSLMMGEDKIKFKHITPLQEQSKEVAIRIFQGCQFRSVEAVQEITEYAKSIPGFV NLDLNDQVTLKYGVEHIIYTMLASLMNKDGVLISEGQGFMTREFLKSRLKPFQDFMEPKFEF AVKFNALELDDSDLAIFIAVIIISGDRPGLLNKPIEDIQDNLQALELQLKLNHPESQLFA KLLQKMTDLRQIVTEHVQLLQVIKKTETDMSLHPLLQEIKDLH |
| PRMT2 | GHMEAPGEGPCSESVIPVLEEDPVDYGCQMQLLQDGAQLQLQLQPEEFVAIADYTATDETQL SFLRGEKILILRQTTADWWGERAGCCGYIPANHLGKQLEEDPEDTQDEEYFDSYGTGLKLH LEMLADQPRTTYHVSILQNKESLKDQVILDVGCCTGIISLFCAHHARPAVYAVEASDMAQH TSQLVLQNGFADTITVFQKVEDVVLPEKVDVLVSEWMTCLLFEFMIESILYARDTWLKGDG IIWPTTAALHLVPCSAEKDYHSKVLFDNAYEFNLSALKSLAIKEFFSRPKSNHILKPEDCLS EPCITLQLDMRTVQVPDLETMRGELRFDIQKAGTLHGFTAWFSVYFQSLEEGQPQVLSGTL HPTTHWKQTLFMMDDPVPVHTGDVVTGSVVLQRNPVWRRHMSVLSWVVTALDPTSQRVGEK VFPIWW |
| PRMT9 | GPHMPNSRPRRRGAGSGAGAPGRDELVSRLQSAEHCLGSQDFGTAYAHYLLVLSLAPLKD DVKETFQYTLFRWAEELDALSRQDLLGCYEQALELFPDDEVICNSMGEHLFRMGFRDEAAGY FHKAVKLNPFDNDAKENFYRVANWLVERWHFIMLNDTKRNTIYNAAIQKAVCSGSKSVLDIGA GTGILSMFARKAGASVYACELSKTMYELACDVVAANEMEEGKLLHMKSLDIKIPKHIPERV SLVVTETVDAGLFGEGIVESLIHAWEHLLQPKTKGENGDCEKYGKVIPASAVIFGIAVECAE IRRHHRVGKDVAGISLPTNVKQSPAYSSVDGEETIEPYTTEKMSRVPGGYLALTECFEIMT VDFNNLQELKSLATKPKDKIGVPIKEGVLDAIVVWVFLQLDDEHSLSTSPSEETCWEQAVYP VHDLADYWIKLGDDEVIMEVSCQDCYLRIQSISVFHSEHEMEVGKSFTRNKDLSFGNEAELCS ALANLQTSKPDVVEQLCVLESTEIALLNIPYHEGFKMAMKVLSSLTPEKLGQAMDTQCQNN EMSYESGPSNSDQSTPEPLYLDVSEGSVLPVIAGTLGQVKPYSSVEKDQHRATLDIICEAN HFPKDTLEFWLRHVEDESAVLQRPKSKLWSIIILDVIEPSGLIQQEIMEKAAISRCLLQSGG KIFPQYVLMFGLLVESQTLMEESAVQGAERTLGFNIAPFINQFQVPRVFLDLASLPCIPLSK PVELLRDLMTPLYLNTSNREVKVRICKSGQVTAIPFWYHMYLDDEIRLDTSSSEASHWKQAAVV LDNPIQVEMGDELVLISIQHHSNVSITIKQ |
| PXR | GSHMSERTGTQPLGVQGLTEEQRMIRELMDAQMKTFDITFSHFKNFRLPGVLSGCELPEL QAPSREEAAKWSQVRKDLCSLKVSLQLRGEDGSVWNYKPPADSGGKEIFSLPHMADMSTYMF KGIISFAKVISYFRDLPIEDQISLLKGAAFELCQLRFNTVFNAETGTWECGRLSYCLEDTAGG FQQLLEPMLKFHYMLKKLQLHEEEYVLMQAISLFSPPDRPVLQHRVVDQLQEFAITLKS YIECNRPQPAHRFLFLKIMAMLTLSINAQHTQRLRLRIQDIHPFATPLMQELFGITGSGGSGGS SHSSLTERHKILHRLLEQEGSPS |
| RSF1 | MGSSHHHHHSSGLVPRGSHMSSGGTRVYVGGVLEGIKKEDLEREFDKYGLNSVWVALNPPG FAFIEFENLQEAEDACSAMNGFEMLGATLKVEISRKRDPRRGGGFRGGRGNFRGRGFGGPP GGGNRSFNYSYGGGRPYNRNNQRDEYGSGGGNFRS |
| RUVBL1 (R1) | MVHHHHHHLVPRGSKIEEVKSTTKTQRIASHSHVKGLGLDESLAKQAASGLVGQENAREACG VIVELIKSKKMAGRAVLLAGPPGTGKTALALAIQELGSKVFPFPMVGSEVYSTEIKKTEVLM ENFRRAIGLRIKETKEVYEGETELTPCETENPMGGYKTIHVIIIGLKTAKGKQLKLDPSI FESLQKERVEAGDVIYIEANSQAVKRGRCDTYATEFDLEAEYVPLPKGDVHKKKEIIQDVT LHDLVANARPQGGQDILSMMQMLMKPKKTEITDKLRGEINKVVNKYIDQGAELVPGVLFVD EVHMLDIECFYLRHALESSIAPVIVIFASNRGNCVIRGTEDITSPHGIPLDLLDRVMIIRTML YTPQEMKQIIKIRAQTEGINISEALNHLGEIGTKTTLRYSVQLLTPANLLAKINGKDSIEKE HVEEISELFYDAKSSAKILADQQDKYMK |

| | |
|---------------------------|---|
| RUVBL2 (R2) | MATVTATTKVPEIRDVTRIERIGAHSHIRGLGLDDALEPRQASQGMVQQLAARRAAGVVLEMI REGKIAGRAVL IAGQPGTGKTAIAMGMAQALGPDTPFTAIGSEIFSLMSKTEALTQAFRRS IGVRIKEETEIEIEGEVVEIQIDRPATGTGSKVGKLTLLKTTMETIYDLGTMIESLTKDKVQA GDVITIDKATGKISKLGRSFTRRARDYDAMGSQTKFVQCPDGELQKRKEVVHTVSLHEIDVINS RTQGFLALFSGDTGEIKSEVREQINAKVAEWREEGKAEIIPGVLFIDEVHMLDIESFSFLNRA LESDMAPVLIMATNRGITRIRGTSYQSPHGIPIDLLDRLLIVSTTPYSEKDTKQILRIRCEEE DVEMSEDAYTVLTRIGLETSRVAIQQLITAASLVCRKRKGTVEVQVDDIKRVYSLFLDESSTQ YMKEYQDAFLFNELKGETMDSLEVLQ |
| SAP49 | MAAGPISERNQDATVYVGGLEKVESEPLLWELFLQAGPVVNTTHMPKDRVTGQHQQGYGFVEFLS EEDADYAIKIMNMIKLYGKPIRVNKASAHNKNLDVGANIFIGNLDPEIDEKLLYDTFSAFGVI LQTPKIMRDPDTGNSKGYAFINFASFDAASDAIEAMNGQYLCNRPITVSYAFKKDSKGERHGS AAERLLAAQNPLSQADRPQLFADAPPPSAPNPVSSLSGSLPPPMPPPGSPFPVPPPGA LPPGIPPAMPPPPMPGAAGHGPPSAGTPGAGHPGHGSHPHFPFPPGMPHPGMSQMLAHHG PHGLGHPHAGPPGSGGQPPRPPGMPHPGPPPMGMPRGPFPFSPMGHPGMPPHGMRGPPP LMPPHGYTGPPRPPPYGYQRGLPPPRPTPRPPVPPRGLRGLPQ |
| SAP145 | GPATEHPEPPKAELQLPPPPPPGHYGAWAQELQAKLAEIGAPIQGNREELVERLQSYTRQTG IVLNRPVLRGEDGDKAAPPPMSAQLPGIPMPPPPGLPPLQPPPPPPPPPLGLGLGFMAHPP NLGPPPLRVGEPVALSEERLKLAAQQAAALMQQEERAKQQGDHSLKEHELLEQQKRAVLL EQRQEQEIAKMGTPVPRPPQDMGQIGVRTPLGPRVAAPVGPVGPPTPTVLPAGAPVPRGPP PPGDENREDDPSVGPKIPQALEKILQLKESRQEEMNSQQEEMETDARSSLGQSASETEED TVSVSKKEKNRKRNRKRRKRRKRRKRRKRRKRRKRRKRRKRRKRRKRRKRRKRRKRRKRRKRR IYEPNFIFFKRIFEAFKLTDDVKEKEKEKEPEKLDKLENSAAPKKKGFEEHKSDDDDSSDDEQ EKKPEAPKLSKKKLRMRNRFTAELKQLVARPDVEMHDVTAQDPKLLVHLKATRNVPVPRH WCFKRKYLQGKRGIEKPPFELPDFIKRTGIQEMREALQEKEEQTKMKSMMREKVRPKMGKIDI DYQKLHDAFFKWQTKPKLTIHGDLYYEGKEFETRLEEKKPGDLSDELRLISLGMPPVGNNAHVP PPWLIAMQRYGPPPSYPNLKIPGLNSPIESCFSFGYHAGGWGKPPVDETKPLYGDVFGTNA EFQTKTEEEEIDRTPWGELEPSDEESSEEEEESEDEKPDETGFITPAGSLITPGGFSSVP AGMETPELIGLRKKKIEEAMDGETPQLFTVLPEKRTATVGGAMMGSTHIYDMSTVMSRKGPA PELQGEVALAPEELELDPAMTQKYEHEVREQAQVEKEDFSMVAEHAQKQKRRKAQPQ DSRGGSKKYKEFKF |
| SAP145 ₄₀₁₋₅₅₀ | GPHMKEKEKEPEKLDKLENSAAPKKKGFEEHKSDDDDSSDDEQEKPEAPKLSKKKLRMRNR FTVAELKQLVARPDVEMHDVTAQDPKLLVHLKATRNVPVPRHWCFKRKYLQGKRGIEKPPF ELPDFIKRTGIQMREALQEKEEQTKM |
| SPAG ₁₆₂₂₋₉₂₆ | MTFKALKEEGNQCVDNKYKDALSKYSECLKINNKECAIYTNRALCYLKLQCQFEAKQDCDQA LQLADGNVKAIFYRRALAHKGLKNYQKSLIDLNKVILLDPSIIIEAKMELEEVTLLNLKDKTAP FNKEKERRKIEIQEVNEGKEEGRPAGEVSMGCLASEKGGKSSRSPEDPEKLPKPNAYEF GQIINALSTRKDEACAHLLAITAPKDLPMFLSNKLEGDFTLLLIQSLKNNLIEKDPFLVYQH LLYLSKAERFKMMLTLISKGQKELIEQLFEDLSDTPNNHFTLEDIQALKRQYEL |
| VDR | MGSSHHHHHSSGLVPRGSHMPKLSEEQRRIIAILLDAHHTYDPTYSDFCQFRPPVRVNDGG GSHPSRPNRSRHTPSFGDSSSSSDHCITSSDMMDSSSFNLDLSEEDSDPSVTLELSQLSM LPHLADLVSYIQKVIQFAKMIIPGFRDLTSEDQIVLLKSSAIEVIMLRNESFTMDMSWTG NQDYKYRVSDVTKAGHSLELIEPLIKFQVGLKLNHHEEHVLLMAICIVSPDRPGVQDAALI EAIQDRLSNTLQTYIRCRHPPPGSHLLYAKMIQKLADLRSLNEEHSKQYRCLSFQPECSMMLT PLVLEVFGNEIS |
| W1 | GSHMPKGRWVEGITSEGYHYYYDLISGASQWEKPEGFQGDLLKKTAVKTVWVEGLSEDFGFTYY YNTETGESRWEKPPD |

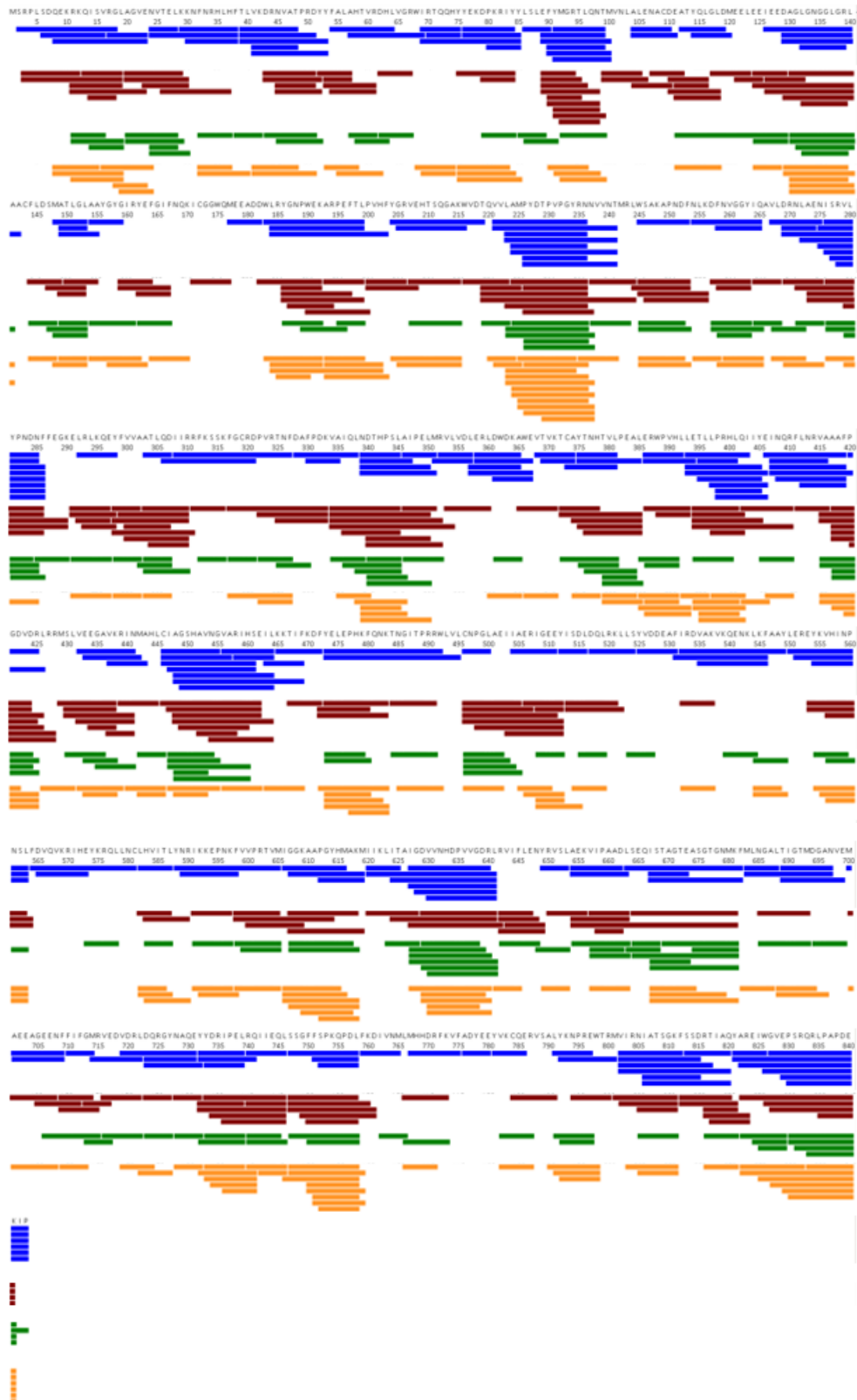
Annexe 3

Carte de couverture de séquence obtenue en HDX-MS pour la phosphorylase B avec les différentes méthodes de deutération (automatisée, bleu ; manuelle, rouge ; manuelle après congélation, vert) réalisé grâce à ¹.



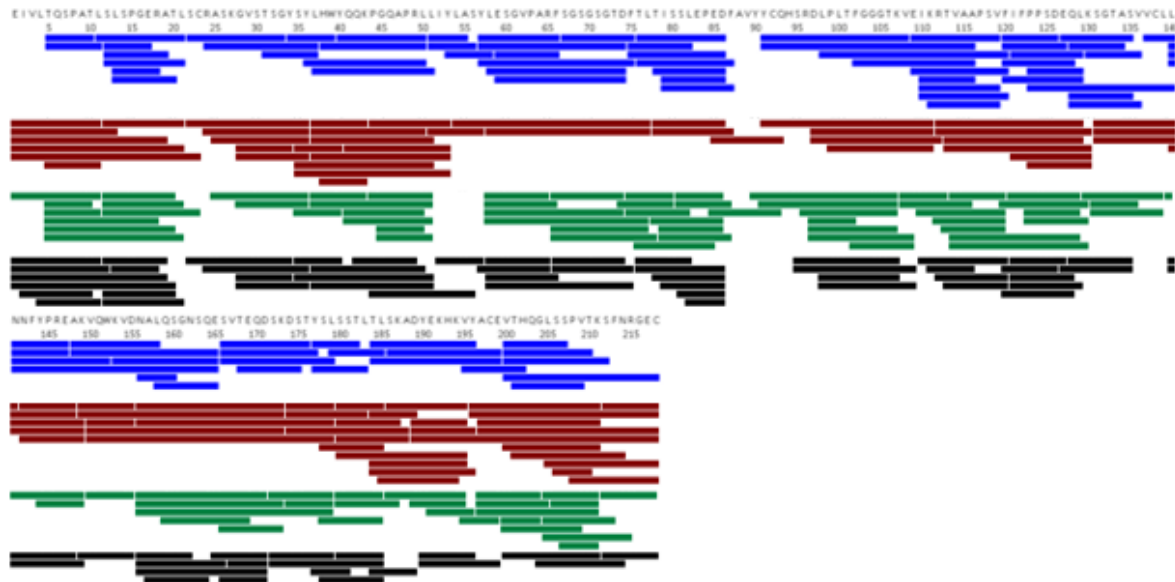
Annexe 4

Carte de couverture de séquence obtenue en HDX-MS pour la phosphorylase B avec les différentes enzymes de digestion (pepsine, bleu ; nepenthesine I, rouge ; nepenthesine II, vert ; rhizopuspepsine, orange) réalisé grâce à ¹.



Annexe 5

Carte de couverture de séquence obtenue en HDX-MS pour la chaîne légère du pembrolizumab avec les différentes enzymes de digestion (pepsine, bleu ; nepenthesine I, rouge ; nepenthesine II, vert ; rhizopuspepsine, noir) réalisé grâce à ¹.



Annexe 6

Conditions expérimentales utilisées en HDX-MS avec et sans l'IM activée.

| | PhosB | PXR | Nivolumab | EPAC | PRTM9 |
|----------------------------|---|---|---|--|---|
| Tampon | 10 mM K ₂ HPO ₄ /KH ₂ PO ₄ pH 7 | 150 mM NaCl, 20 mM Tris-HCl, 1 mM DTT pH 8 | 10 mM K ₂ HPO ₄ /KH ₂ PO ₄ pH 7 | 150 mM NaCl, 50 mM HEPES, 2 mM MgCl ₂ , 2mM β-mercaptoethanol pH 8 | 20mM Tris-HCl, 100mM NaCl, 2mM TCEP pH 8 |
| Tampon Q | 100 mM K ₂ HPO ₄ /KH ₂ PO ₄ pH 2,4 | 2M Guanidine HCl, 100 mM Glycine, 1 mM DTT pH 2,2 | 4M Guanidine HCl, 100 mM Glycine, 500 mM TCEP pH 2,2 | 4M Guanidine HCl, 150 mM Glycine pH 1,5 | 2M Guanidine HCl, 100 mM Glycine pH 2,1 |
| Protéase | Pepsine | | | | |
| Quantité injectée | 40 pmoles | 50 pmoles | 40 pmoles | 19 pmoles | 30 pmoles |
| Débit lors de la digestion | 200 µL/min | 200 µL/min | 200 µL/min | 100 µL/min | 200 µL/min |

Annexe 7

Tableau décrivant la couverture de séquence, la redondance et le nombre de peptides identifiés du Nivolumab, EPAC1, PhosB, PRMT9 et PXR avec ou sans l'IM activée suivant les filtres d'énergies appliqués (LE et HE)

| | | Nivolumab | EPAC1 | PhosB | PRMT9 | PXR | |
|----------------------|------------|-------------------------------|-------|-------|-------|------|------|
| LE : 250 HE : 100 | sans IM | couverture de séquence (%) | 76,5 | 65,6 | 96,4 | 92,1 | 99,4 |
| | | redondance | 2,6 | 1,5 | 3,6 | 3,7 | 5,9 |
| | | nombre de peptides identifiés | 115 | 81 | 267 | 256 | 179 |
| | IM activée | couverture de séquence (%) | 74,7 | 79,5 | 94,1 | 89,1 | 98,2 |
| | | redondance | 3,1 | 2,3 | 3,9 | 3 | 5,2 |
| | | nombre de peptides identifiés | 130 | 143 | 289 | 219 | 161 |
| LE : 120 HE : 20 | sans IM | couverture de séquence (%) | 84,4 | 71,6 | 98,5 | 92,9 | 100 |
| | | redondance | 3,6 | 2 | 4,1 | 4,1 | 6,5 |
| | | nombre de peptides identifiés | 174 | 117 | 307 | 292 | 199 |
| | IM activée | couverture de séquence (%) | 92,9 | 92,7 | 94,9 | 87 | 98,8 |
| | | redondance | 4,7 | 3,7 | 4,2 | 4,5 | 5,8 |
| | | nombre de peptides identifiés | 235 | 250 | 325 | 304 | 174 |

Annexe 8

Tableau décrivant la couverture de séquence, la redondance et le nombre de peptides identifiés du Nivolumab (Lc et Hc), PhosB et PXR avec ou sans l'IM activée en HDX-MS avant et après traitement des données.

| | | Nivolumab | | PhosB | PXR | |
|------------------|------------|-------------------------------|------|-------|------|------|
| | | Lc | Hc | | | |
| avant traitement | sans IM | couverture de séquence (%) | 80,4 | 74,7 | 96,4 | 99,4 |
| | | redondance | 3 | 2,4 | 3,6 | 5,9 |
| | | nombre de peptides identifiés | 46 | 69 | 267 | 179 |
| | IM activée | couverture de séquence (%) | 74,8 | 74,7 | 94,1 | 98,2 |
| | | redondance | 3,5 | 2,9 | 3,9 | 5,2 |
| | | nombre de peptides identifiés | 49 | 81 | 289 | 161 |
| après traitement | sans IM | couverture de séquence (%) | 80,4 | 69,6 | 91,8 | 94,4 |
| | | redondance | 3 | 2,2 | 2,9 | 4,5 |
| | | nombre de peptides identifiés | 41 | 60 | 200 | 135 |
| | IM activée | couverture de séquence (%) | 74,8 | 74,7 | 81,5 | 98,2 |
| | | redondance | 3,5 | 2,8 | 3,5 | 4,5 |
| | | nombre de peptides identifiés | 49 | 79 | 249 | 142 |

Annexe 9

Conditions expérimentales utilisées en HDX-MS pour la comparaison des outils de validation statistique.

| | VDR/W1 vs. VDR/W1/ZK12813 | EPAC1/AMPc vs. EPAC1/AMPc/liposomes | PPARγWT vs. F310S | PRMT2 vs. PRMT2/RSF1 |
|----------------------------|---|---|--|--|
| Tampon | 200 mM NaCl, 20 mM Tris-HCl, 1 mM TCEP pH 8 | 150 mM NaCl, 50 mM HEPES, 2 mM MgCl ₂ , 2mM β -mercaptoethanol pH 8 | 20 mM Tris-HCl, 200 mM NaCl pH 8 | 20mM HEPES, 100mM NaCl, 0,5mM DTT pH 7 |
| Tampon Q | 2M Guanidine HCl, 200 mM Glycine pH 2,2 | 4M Guanidine HCl, 150 mM Glycine pH 2,2 | 2M Guanidine HCl, 100 mM Glycine pH 2,2 | 2M Guanidine HCl, 100 mM Glycine pH 2,1 |
| Protéase | Pepsine | | | |
| Quantité injectée | 20 pmoles (VDR), 100 pmoles (W1), 40 pmoles (Z12813K) | 19 pmoles (EPAC1), 4,5 mM (AMPc), 3,125 mM (liposomes) | 30 pmoles | 20 pmoles |
| Débit lors de la digestion | 200 μ L/min | 100 μ L/min | 200 μ L/min | 100 μ L/min |

Annexe 11

Protocole expérimental du gel 1D SDS-PAGE à 10%.

Gel 1D 10% : 1.5mm d'épaisseur, 10 puits, préparé 1 jour avant l'électrophorèse

| Gel de séparation (10mL) | 10% |
|--------------------------|--------|
| H ₂ O | 4 mL |
| Tris 1.5M pH 8.8 | 2.5 mL |
| Acryl/Bis 30% | 3.4 mL |
| SDS 10% | 100 µL |
| Temed | 10 µL |
| APS 10% | 30 µL |

| Gel de concentration (3mL) | 5% |
|----------------------------|--------|
| H ₂ O | 2 mL |
| Tris 1M pH 6.8 | 400 µL |
| Acryl/Bis 30% | 500 µL |
| SDS 10% | 30 µL |
| Temed | 6 µL |
| APS 10% | 24 µL |

Préparation de l'échantillon :

Un mix de 50/50 (v/v) de l'échantillon et d'un tampon Laemmli 2x (10mM Tris pH 6,8, 1mM EDTA, 5% β-mercaptoéthanol, 5% SDS, 10% glycérol, 0,01% Bleu de bromophénol) est chauffé à 95°C pendant 5 minutes puis déposé.

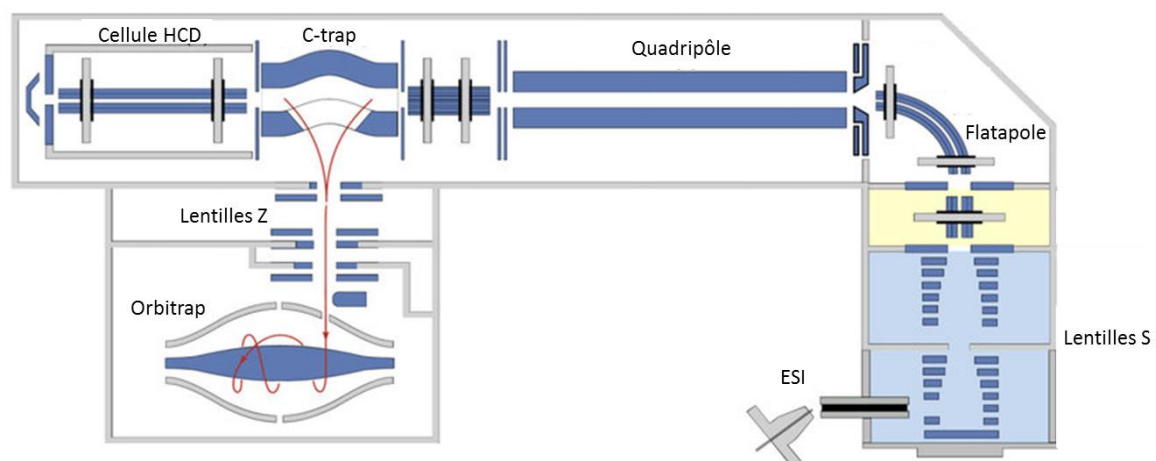
Migration : 50-120 V pendant 1h30

Fixation : Bain composé de 50% EtOH, 3% d'acide phosphorique pendant 15min (3 cycles de lavage à l'eau après la fixation, 5 min)

Coloration : Solution de colorant CBB (« Coomassie Brilliant Blue ») pendant 30min-1h (3 cycles de lavage à l'eau après la fixation, 5 min)

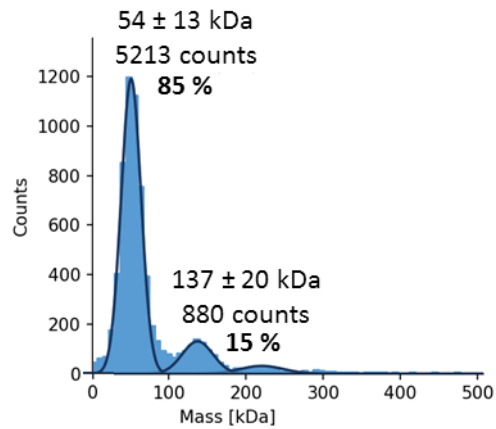
Annexe 12

Représentation de la géométrie du Orbitrap Q-Exactive Plus ou du Orbitrap Q-Exactive HF-X (Thermo Fisher Scientific), d'après ³.



Annexe 13

Histogramme de distribution de contraste obtenue en M-Ph de la BSA (monomère 66 kDa et dimère 132 kDa).



Annexe 14

Tableau relatant des conditions expérimentales des trois systèmes protéiques étudiés (BSA, PRMT2/RSF1 et R2D) pour comparer les logiciels de traitement des données.

| | BSA | PRMT2/RSF1 | R2D |
|-----------------------|--|--|--|
| Tampon | 25mM K ₂ HPO ₄ , 25mM KH ₂ PO ₄ , 0,5mM TCEP pH 7,8 | 20mM HEPES, 100mM NaCl, 0,5mM TCEP pH 7,5 | 20mM HEPES, 100mM NaCl, 0,5mM TCEP pH 7,5 |
| Excès d'agent pontant | 200 | 100 | 100 |
| Système MS | Q-Exactive HF-X Orbitrap | Q-Exactive HF-X Orbitrap | Q-Exactive Plus Orbitrap |

Annexe 15

Tableau descriptif des peptides cross-linkés (protéines, sites, logiciels et distances Ca-Ca) identifiés suite au XL-MS de la BSA ainsi que des complexes PRMT2/RSF1 et R2D.

BSA :

| Intra-crosslink Site 1 | Intra-crosslink Site 2 | MaxQuant | Merox | distance Ca-Ca | Intra-crosslink Site 1 | Intra-crosslink Site 2 | MaxQuant | Merox | distance Ca-Ca |
|------------------------|------------------------|----------|-------|----------------|------------------------|------------------------|----------|-------|----------------|
| K28 | K36 | x | | 12,9 | K235 | K374 | x | x | 13,2 |
| K28 | K88 | x | x | 9,1 | K235 | K401 | x | | 23,5 |
| K28 | S89 | x | | 10,1 | K235 | K437 | | x | 32 |
| K28 | T92 | x | | 14,7 | K235 | K455 | x | x | 28,8 |
| K28 | K155 | x | | 31,3 | K235 | S466 | x | | 26,8 |
| K28 | K235 | x | x | 30,1 | K235 | K489 | x | x | 25,5 |

| Intra-crosslink | | MaxQuant | Merox | distance Ca-Ca | Intra-crosslink | | MaxQuant | Merox | distance Ca-Ca |
|-----------------|--------|----------|-------|-------------------|-----------------|--------|----------|-------|-------------------|
| Site 1 | Site 2 | | | | Site 1 | Site 2 | | | |
| K28 | K245 | x | x | 37,8 | K235 | S493 | x | | 29,8 |
| K28 | K248 | x | x | 39,7 | K235 | K495 | x | x | 27,5 |
| K28 | K263 | x | x | 20,8 | K235 | S503 | | x | 15,2 |
| K28 | K455 | x | x | 49,9 | K235 | K548 | x | x | 41,8 |
| K28 | T458 | x | | 52,1 | K235 | T550 | x | | 38,5 |
| K28 | K548 | | x | 58,4 | S238 | K266 | x | | 12,2 |
| S29 | K36 | x | | 10,8 | K245 | K248 | x | | 8,6 |
| S29 | K155 | x | | 27,9 | K245 | K256 | x | | 15 |
| K36 | S82 | x | | 15,7 | K245 | K263 | x | x | 19,6 |
| K36 | K155 | x | | 20,4 | K245 | K266 | x | x | 21,5 |
| K36 | K156 | x | | 22,2 | K245 | Y286 | x | | 16,4 |
| K36 | Y184 | x | | 25,9 | K245 | S296 | x | | 10,4 |
| K36 | K235 | x | x | 29,8 | K245 | K297 | x | | 9,3 |
| K36 | K245 | x | x | 32,5 | K245 | K299 | x | x | 14,1 |
| K36 | K248 | x | x | 35,5 | K245 | K304 | x | x | 19,6 |
| K36 | K266 | x | | 22,5 | K245 | K309 | x | x | 16,3 |
| K36 | K285 | x | x | 16,8 | K245 | S310 | x | x | 14,7 |
| K36 | K309 | | x | 19,9 | K245 | K401 | x | x | 22 |
| K36 | S310 | x | | 19,5 | K245 | K455 | | x | 26 |
| K36 | K455 | x | x | 43,5 | K245 | T458 | x | | 24,9 |
| K36 | T458 | x | | 44,9 | K245 | K463 | | x | 19,2 |
| K36 | K548 | | x | 52,8 | K245 | S466 | x | | 17,3 |
| K36 | T550 | x | | 51 | K245 | K489 | x | x | 35,4 |
| K44 | K155 | x | | 11,1 | K245 | T490 | x | | 38,1 |
| K44 | K156 | x | | 11,3 | K245 | T497 | x | | 33,4 |
| K44 | K245 | x | | 33,2 | K245 | S503 | | x | 25,2 |
| K44 | T458 | x | | 38,2 | K245 | K548 | x | x | 42,5 |
| T71 | K155 | x | | 15,8 | K245 | T550 | x | | 40,9 |
| T71 | K156 | x | | 15,7 | K248 | K263 | x | | 23,3 |
| S82 | K155 | x | | 31,9 | K248 | S296 | x | | 9,1 |
| S82 | K156 | x | | 32,4 | K248 | K297 | x | x | 11,2 |
| S82 | K235 | x | x | 38,2 | K248 | K299 | | x | 15 |
| T92 | K245 | x | | 38,4 | K248 | S310 | x | | 20,2 |
| K100 | S225 | | x | 32,3 | K248 | K374 | x | x | 23,6 |
| K100 | K235 | | x | 37,1 | K248 | K401 | x | x | 23,5 |
| S103 | K235 | x | | 39,6 | K248 | K455 | x | x | 34,2 |
| S103 | K245 | x | | 47,7 | K248 | K489 | x | x | 42,4 |
| S128 | K140 | x | | 29,8 | K248 | K548 | x | x | 50,8 |
| S128 | K489 | x | | 10,2 | K248 | T550 | x | | 49,2 |
| K130 | K140 | x | | 24,2 | K256 | S296 | x | | 18,3 |
| K138 | K160 | | x | 14,9 | K256 | K297 | | x | 19,2 |
| K138 | K455 | x | x | 19,6 | K256 | K437 | | x | 42,5 |
| K138 | T458 | x | | 25,2 | K256 | K455 | | x | 36,2 |
| K140 | K155 | x | | 21 | K256 | T458 | x | | 36,9 |
| K140 | K156 | x | | 18,9 | K256 | T550 | x | | 47,7 |
| K140 | K160 | x | x | 14,7 | T262 | K266 | x | | 5,9 |

| Intra-crosslink | | MaxQuant | Merox | distance Ca-Ca | Intra-crosslink | | MaxQuant | Merox | distance Ca-Ca |
|-----------------|--------|----------|-------|-------------------|-----------------|-----------|----------|-------|-------------------|
| Site 1 | Site 2 | | | | Site 1 | Site 2 | | | |
| K140 | K235 | x | x | 38,5 | T262 | K309 | x | | 18,8 |
| K140 | K245 | x | | 37,8 | T262 | K336 | x | | 33,2 |
| K140 | K455 | | x | 21,5 | T262 | K401 | x | | 29,2 |
| K140 | T458 | x | | 23,8 | T262 | K489 | x | | 25,5 |
| K140 | K489 | x | | 28,9 | K263 | K455 | | x | 32,4 |
| K140 | T490 | x | | 29,4 | K263 | T458 | x | | 34,5 |
| T145 | K156 | x | | 14,7 | K263 | K489 | x | x | 26,2 |
| K151 | K156 | x | | 10,2 | K266 | K455 | x | | 29,6 |
| K151 | K197 | x | x | 13,6 | K266 | T458 | x | | 32,4 |
| K151 | K235 | x | | 44 | K266 | K489 | | x | 21,4 |
| K151 | K455 | x | | 35,9 | K266 | T490 | x | | 25,1 |
| K155 | K156 | x | | 3,8 | K285 | K309 | x | x | 9,5 |
| K155 | Y161 | x | | 9,8 | K285 | S310 | x | | 9,2 |
| K155 | K183 | x | | 12,9 | S296 | K304 | x | | 16 |
| K155 | Y184 | x | | 11,6 | S296 | K318 | x | | 9,1 |
| K155 | K228 | x | | 37,8 | S296 | K401 | x | | 30,4 |
| K155 | K235 | x | | 36,9 | K297 | K304 | x | | 13,9 |
| K155 | K245 | x | | 35,9 | K299 | S310 | x | | 12,7 |
| K155 | K248 | x | | 42,2 | K304 | S310 | x | | 8,6 |
| K155 | K256 | x | | 37 | K336 | K346 | x | | 16,1 |
| K155 | T262 | x | | 31,6 | K336 | K401 | x | | 21,2 |
| K155 | K266 | x | | 29,3 | K346 | K374 | x | x | 12,4 |
| K155 | S296 | x | | 36,7 | Y357 | K401 | | x | 11,7 |
| K155 | K297 | x | | 34 | S358 | K401 | x | | 13,9 |
| K155 | K304 | x | | 21,7 | K374 | K399 | x | | 11,9 |
| K155 | K336 | x | | 63,6 | K374 | K437 | | x | 26,3 |
| K155 | K374 | x | | 47,7 | K374 | K455 | | x | 28,9 |
| K155 | K401 | x | | 53,9 | K374 | T458 | x | | 30,8 |
| K155 | T490 | x | | 39,1 | K374 | S466 | x | | 24,1 |
| K156 | K235 | x | x | 37,4 | K374 | K498 | | x | 18 |
| K156 | K245 | x | x | 37,5 | T395 | K401 | x | | 10,6 |
| K156 | K248 | x | x | 44,1 | K401 | K437 | | x | 28,8 |
| K160 | S225 | | x | 30,2 | K401 | K455 | x | x | 29 |
| K160 | K235 | | x | 34,1 | K401 | K463 | | x | 23,9 |
| K160 | K245 | x | x | 35,1 | K401 | S466 | x | | 17 |
| K160 | K489 | | x | 29,2 | K401 | K498 | | x | 26,3 |
| Y161 | K489 | x | | 28 | K401 | T563 | x | | 37,8 |
| Y163 | K248 | x | | 40,1 | K412 | K437 | | x | 20,9 |
| Y173 | K245 | x | | 20,7 | K420 | K437 | | x | 23 |
| Y180 | K455 | x | x | 22,9 | K420 | T458 | x | | 12,7 |
| K183 | S238 | x | | 26,6 | K437 | T501/S503 | | x | 17,1 |
| K183 | K245 | x | x | 25,3 | K437 | K561 | | x | 11,4 |
| Y184 | K235 | x | | 32 | K437 | K568 | | x | 20 |
| Y184 | K245 | x | | 28,2 | K455 | K463 | x | | 13,2 |
| Y184 | K304 | x | | 17,7 | K455 | S466 | x | | 18,7 |

| Intra-crosslink | | MaxQuant | Merox | distance Ca-Ca | Intra-crosslink | | MaxQuant | Merox | distance Ca-Ca |
|-----------------|--------|----------|-------|-------------------|-----------------|--------|----------|-------|-------------------|
| Site 1 | Site 2 | | | | Site 1 | Site 2 | | | |
| T207 | K455 | x | | 14,6 | K455 | Y475 | x | x | 9,7 |
| S225 | K245 | x | x | 21,2 | K455 | S477 | x | | 13,8 |
| K228 | K235 | x | | 13,5 | K455 | K489 | x | x | 25,2 |
| K228 | K489 | | x | 13,1 | K455 | K498 | | x | 25,3 |
| K235 | K245 | x | | 15,4 | K455 | K548 | | x | 17,2 |
| K235 | K248 | x | x | 19,1 | T458 | K463 | x | | 8,9 |
| K235 | K256 | | x | 10,9 | T458 | K489 | x | | 30,5 |
| K235 | T262 | x | | 6,3 | T458 | K548 | x | | 20,9 |
| K235 | K263 | x | x | 9,7 | T458 | K561 | x | | 29,8 |
| K235 | K266 | x | x | 10,4 | K489 | K548 | | x | 26,1 |
| K235 | K285 | x | x | 20,6 | K489 | S503 | | x | 17,7 |
| K235 | S296 | x | | 23,6 | T490 | K548 | x | | 24,3 |
| K235 | K297 | x | | 23,1 | K495 | K561 | | x | 18,7 |
| K235 | K304 | x | | 27,4 | K495 | T563 | x | | 23,3 |
| K235 | K309 | | x | 22,8 | T497 | K561 | x | | 23,3 |
| K235 | S310 | x | x | 19,3 | K544 | K548 | x | | 5,8 |
| K235 | K336 | x | | 27,7 | K548 | K568 | x | x | 14,2 |
| K235 | K346 | x | x | 14,1 | T550 | K568 | x | | 16,1 |

PRMT2/RSF1 :

| Intra-crosslink | | | | Merox | Maxlynx | distance Ca-Ca |
|-----------------|-----------|--------|--------|-------|---------|-------------------|
| Protein 1 | Protein 2 | Site 1 | Site 2 | | | modèle PRMT2/RSF1 |
| PRMT2 | PRMT2 | K68 | K136 | x | | 50,3 |
| PRMT2 | PRMT2 | K68 | Y137 | | x | 49,1 |
| PRMT2 | PRMT2 | K68 | S139 | | x | 53,9 |
| PRMT2 | PRMT2 | K68 | S298 | | x | 30,1 |
| PRMT2 | PRMT2 | K136 | S290 | x | x | 28,6 |
| PRMT2 | PRMT2 | K136 | S298 | | x | 34,3 |
| PRMT2 | PRMT2 | Y270 | K294 | x | x | 12,3 |
| PRMT2 | PRMT2 | S272 | K294 | | x | 14,5 |
| PRMT2 | PRMT2 | S290 | K294 | x | x | 6,3 |
| RSF1 | RSF1 | S2 | K19 | x | | 24 |
| RSF1 | RSF1 | T6 | K19 | | x | 16,9 |
| RSF1 | RSF1 | K19 | Y29 | | x | 15,8 |
| RSF1 | RSF1 | K19 | Y122 | | x | - |
| RSF1 | RSF1 | K19 | Y131 | | x | - |
| PRMT2 | RSF1 | K68 | K19 | x | x | 22,3 |
| PRMT2 | RSF1 | K68 | Y122 | x | x | - |
| PRMT2 | RSF1 | K68 | Y131 | | x | - |
| PRMT2 | RSF1 | K136 | S133 | | x | - |
| PRMT2 | RSF1 | K294 | S-1 | x | x | - |

| Intra-crosslink | | | | Merox | Maxlynx | distance |
|-----------------|-----------|--------|--------|-------|---------|-------------------|
| Protein 1 | Protein 2 | Site 1 | Site 2 | | | Ca-Ca |
| | | | | | | modèle PRMT2/RSF1 |
| PRMT2 | RSF1 | K294 | T6 | | x | 28,4 |
| PRMT2 | RSF1 | K294 | K19 | x | | 26 |
| PRMT2 | RSF1 | K294 | Y122 | x | x | - |
| PRMT2 | RSF1 | K294 | S133 | | x | - |
| PRMT2 | RSF1 | S298 | K19 | x | x | 23,5 |
| PRMT2 | RSF1 | S298 | K28 | | x | 37,1 |
| PRMT2 | RSF1 | S298 | K79 | | x | 33,4 |

R2D (RUVBL1-2/DPCD) :

| Intra-crosslink | | | | Merox | Maxlynx | Intra-crosslink | | | | Merox | Maxlynx |
|-----------------|-----------|--------|--------|-------|---------|-----------------|-----------|--------|--------|-------|---------|
| Protein 1 | Protein 2 | Site 1 | Site 2 | | | Protein 1 | Protein 2 | Site 1 | Site 2 | | |
| DPCD | RuvBL1 | N-ter | K15 | x | | DPCD | RuvBL2 | N-ter | K279 | x | |
| DPCD | RuvBL1 | N-ter | K20 | x | | DPCD | RuvBL2 | N-ter | K368 | x | |
| DPCD | RuvBL1 | N-ter | K181 | x | | DPCD | RuvBL2 | N-ter | K417 | x | |
| DPCD | RuvBL1 | N-ter | K195 | x | | DPCD | RuvBL2 | N-ter | K427 | x | |
| DPCD | RuvBL1 | N-ter | K287 | x | | DPCD | RuvBL2 | K18 | K186 | x | |
| DPCD | RuvBL1 | N-ter | K294 | x | | DPCD | RuvBL2 | K18 | K197 | x | |
| DPCD | RuvBL1 | N-ter | K435 | x | | DPCD | RuvBL2 | K18 | T367 | | x |
| DPCD | RuvBL1 | N-ter | K458 | x | | DPCD | RuvBL2 | K18 | K427 | x | x |
| DPCD | RuvBL1 | K18 | K181 | x | x | DPCD | RuvBL2 | T19 | K444 | | x |
| DPCD | RuvBL1 | K28 | S50 | | x | DPCD | RuvBL2 | K37 | K279 | x | x |
| DPCD | RuvBL1 | K28 | K294 | x | | DPCD | RuvBL2 | K58 | K9 | | x |
| DPCD | RuvBL1 | K58 | K294 | | x | DPCD | RuvBL2 | K58 | K279 | | x |
| DPCD | RuvBL1 | K116 | T176 | | x | DPCD | RuvBL2 | K116 | K427 | x | |
| DPCD | RuvBL1 | K116 | K389 | x | | DPCD | RuvBL2 | K126 | K67 | | x |
| DPCD | RuvBL1 | Y119 | K389 | | x | DPCD | RuvBL2 | K126 | K234 | x | |
| DPCD | RuvBL1 | S122 | K35 | | x | DPCD | RuvBL2 | K126 | K236 | x | |
| DPCD | RuvBL1 | K126 | T176 | | x | DPCD | RuvBL2 | K126 | K427 | x | x |
| DPCD | RuvBL1 | K126 | K435 | | x | DPCD | RuvBL2 | K137 | K164 | | x |
| DPCD | RuvBL1 | K137 | K46 | | x | DPCD | RuvBL2 | K137 | T367 | | x |
| DPCD | RuvBL1 | K137 | K181 | x | x | DPCD | RuvBL2 | K137 | K427 | | x |
| DPCD | RuvBL1 | S198 | K184 | | x | DPCD | RuvBL2 | K142 | K236 | x | |
| DPCD | RuvBL1 | S198 | K294 | | x | DPCD | RuvBL2 | K142 | T367 | | x |
| DPCD | RuvBL1 | S198 | K458 | | x | DPCD | RuvBL2 | S198 | K427 | | x |
| DPCD | RuvBL1 | K204 | K287 | x | x | DPCD | RuvBL2 | K204 | K427 | x | |
| DPCD | RuvBL1 | K204 | K294 | | x | | | | | | |

| Intra-crosslink | | | | Merox | Maxlynx | distance Ca-Ca |
|-----------------|-----------|--------|--------|-------|---------|-------------------|
| Protein 1 | Protein 2 | Site 1 | Site 2 | | | |
| DPCD | DPCD | N-ter | K18 | x | | 29,6 |
| DPCD | DPCD | N-ter | K37 | x | | 23,7 |
| DPCD | DPCD | N-ter | K126 | x | | 21 |
| DPCD | DPCD | N-ter | S144 | x | | 18,2 |
| DPCD | DPCD | N-ter | K188 | x | | 25,8 |
| DPCD | DPCD | K18 | K18 | x | | 71,9 |
| DPCD | DPCD | K18 | K28 | x | x | 12,3 |
| DPCD | DPCD | K18 | K126 | | x | 40,1 |
| DPCD | DPCD | K18 | K137 | | x | 39,3 |
| DPCD | DPCD | K18 | K194 | x | | 4,3 |
| DPCD | DPCD | K18 | K204 | x | | 10,5 |
| DPCD | DPCD | K28 | S198 | | x | 15,8 |
| DPCD | DPCD | K46 | K102 | x | | 13,8 |
| DPCD | DPCD | K46 | K194 | x | | 21,4 |
| DPCD | DPCD | K46 | S198 | | x | 27,1 |
| DPCD | DPCD | S48 | K102 | | x | 11 |
| DPCD | DPCD | K102 | K177 | | x | 11,2 |
| DPCD | DPCD | K102 | K188 | x | x | 13,2 |
| DPCD | DPCD | K102 | K191 | | x | 16,8 |
| DPCD | DPCD | S104 | K188 | | x | 18,4 |
| DPCD | DPCD | K126 | K137 | | x | 26,9 |
| DPCD | DPCD | K126 | K142 | x | x | 12 |
| DPCD | DPCD | K126 | S144 | | x | 9,8 |

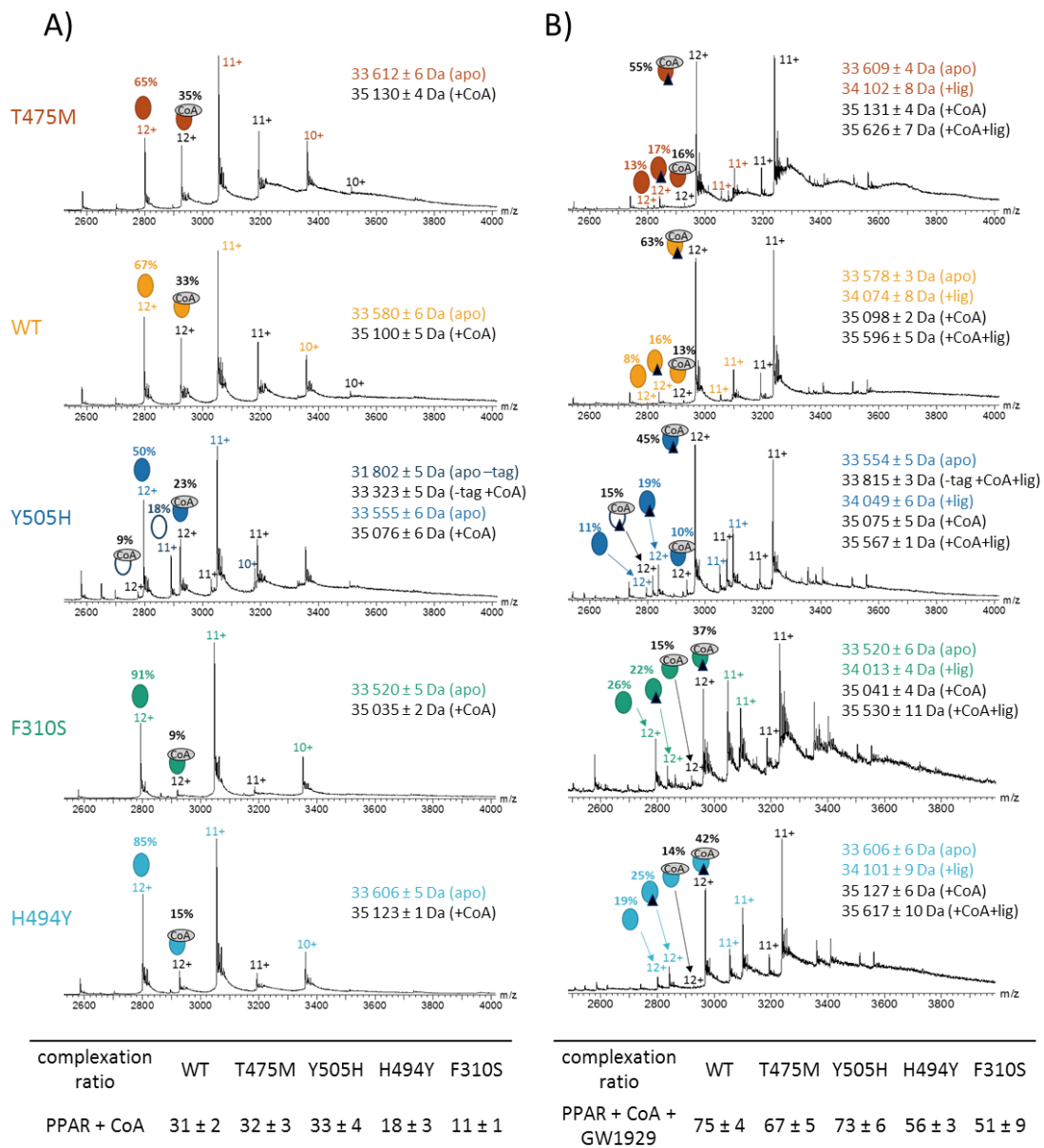
| Intra-crosslink | | | | Merox | Maxlynx | distance Ca-Ca | Intra-crosslink | | | | Merox | Maxlynx | distance Ca-Ca |
|-----------------|-----------|--------|--------|-------|---------|-------------------|-----------------|-----------|--------|--------|-------|---------|-------------------|
| Protein 1 | Protein 2 | Site 1 | Site 2 | | | | Protein 1 | Protein 2 | Site 1 | Site 2 | | | |
| RuvBL1 | RuvBL1 | K15 | K24 | x | | - | RuvBL2 | RuvBL2 | K184 | K201 | x | x | 9,5 |
| RuvBL1 | RuvBL1 | K15 | T25 | | x | - | RuvBL2 | RuvBL2 | K184 | S203 | x | x | 12 |
| RuvBL1 | RuvBL1 | K15 | K141 | | x | - | RuvBL2 | RuvBL2 | K184 | K204 | | x | 14,8 |
| RuvBL1 | RuvBL1 | K15 | K195 | x | x | - | RuvBL2 | RuvBL2 | K186 | K201 | x | x | 10,5 |
| RuvBL1 | RuvBL1 | K15 | K214 | x | x | - | RuvBL2 | RuvBL2 | K186 | K204 | | x | 10,2 |
| RuvBL1 | RuvBL1 | K15 | K281 | x | | - | RuvBL2 | RuvBL2 | K186 | T222 | | x | 18,4 |
| RuvBL1 | RuvBL1 | K15 | K287 | x | x | - | RuvBL2 | RuvBL2 | K186 | K417 | x | | 55,1 |
| RuvBL1 | RuvBL1 | K15 | K294 | x | x | - | RuvBL2 | RuvBL2 | K186 | K427 | x | x | 63,8 |
| RuvBL1 | RuvBL1 | K15 | K431 | | x | - | RuvBL2 | RuvBL2 | T199 | K223 | | x | 26 |
| RuvBL1 | RuvBL1 | K20 | K35 | x | | - | RuvBL2 | RuvBL2 | K201 | K236 | x | x | 16,1 |
| RuvBL1 | RuvBL1 | K20 | S102 | | x | - | RuvBL2 | RuvBL2 | K201 | K279 | | x | 30,3 |
| RuvBL1 | RuvBL1 | K20 | K195 | x | x | - | RuvBL2 | RuvBL2 | K201 | K427 | | x | 63,2 |
| RuvBL1 | RuvBL1 | K20 | K214 | x | | - | RuvBL2 | RuvBL2 | S203 | K236 | | x | 16,4 |
| RuvBL1 | RuvBL1 | K20 | K281 | x | | - | RuvBL2 | RuvBL2 | K204 | K223 | | x | 13,8 |
| RuvBL1 | RuvBL1 | K20 | K389 | x | | - | RuvBL2 | RuvBL2 | K223 | K236 | | x | 24,9 |
| RuvBL1 | RuvBL1 | S21 | K35 | | x | - | RuvBL2 | RuvBL2 | K223 | K427 | | x | 62,9 |
| RuvBL1 | RuvBL1 | T23 | K181 | | x | - | RuvBL2 | RuvBL2 | K234 | K236 | x | | 6,6 |

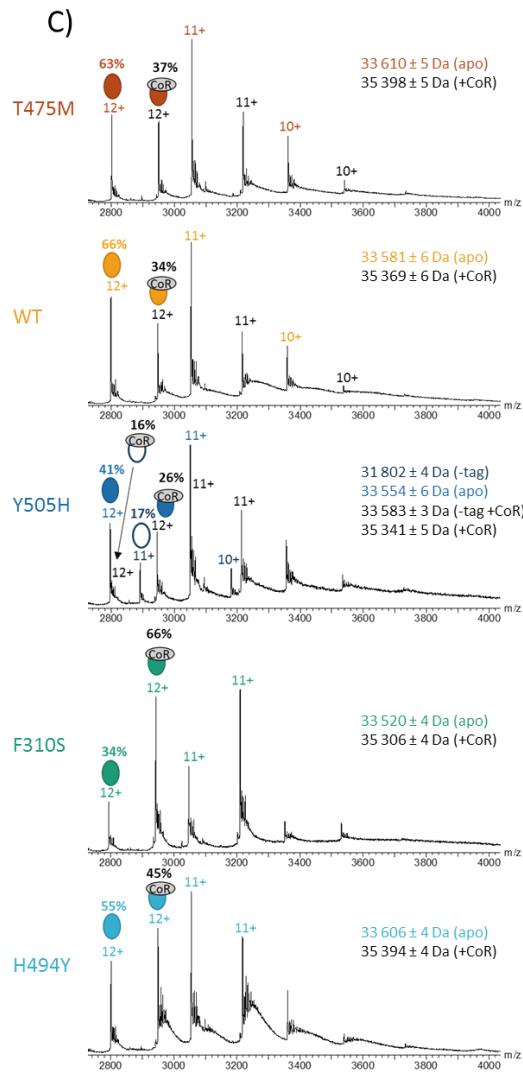
| Intra-crosslink | | | | Merox | Maxlynx | distance Ca-Ca | Intra-crosslink | | | | Merox | Maxlynx | distance Ca-Ca |
|-----------------|-----------|--------|--------|-------|---------|-------------------|-----------------|-----------|-----------|--------|-------|---------|-------------------|
| Protein 1 | Protein 2 | Site 1 | Site 2 | | | SOAF | Protein 1 | Protein 2 | Site 1 | Site 2 | | | SOAF |
| RuvBL1 | RuvBL1 | T23 | K195 | | x | - | RuvBL2 | RuvBL2 | S270 | K279 | | x | 15,4 |
| RuvBL1 | RuvBL1 | T23 | K214 | | x | - | RuvBL2 | RuvBL2 | T367 | K427 | | x | 18 |
| RuvBL1 | RuvBL1 | T23 | K278 | | x | - | RuvBL2 | RuvBL2 | K427 | K427 | x | x | 63,8 |
| RuvBL1 | RuvBL1 | T23 | K389 | | x | - | RuvBL1 | RuvBL2 | K20 | K67 | x | | - |
| RuvBL1 | RuvBL1 | T23 | K431 | | x | - | RuvBL1 | RuvBL2 | T23 | K67 | | x | - |
| RuvBL1 | RuvBL1 | K24 | K35 | x | | 20,3 | RuvBL1 | RuvBL2 | K24 | K67 | x | | 12 |
| RuvBL1 | RuvBL1 | T25 | K35 | | x | 18 | RuvBL1 | RuvBL2 | T25 | K67 | | x | 8,8 |
| RuvBL1 | RuvBL1 | T25 | K431 | | x | 20,6 | RuvBL1 | RuvBL2 | K46 | T367 | | x | 37,7 |
| RuvBL1 | RuvBL1 | K35 | K389 | x | x | 16,2 | RuvBL1 | RuvBL2 | K46 | K417 | x | x | 21,1 |
| RuvBL1 | RuvBL1 | K46 | K181 | | x | 59,2 | RuvBL1 | RuvBL2 | K46 | T419 | | x | 26,4 |
| RuvBL1 | RuvBL1 | K46 | K389 | x | x | 17,4 | RuvBL1 | RuvBL2 | K46 | K427 | x | x | 21,8 |
| RuvBL1 | RuvBL1 | S50 | K287 | | x | 47,3 | RuvBL1 | RuvBL2 | K70/S71 | K417 | x | x | 16,7 |
| RuvBL1 | RuvBL1 | S71 | K138 | | x | 21,1 | RuvBL1 | RuvBL2 | K70 | T419 | | x | 18,6 |
| RuvBL1 | RuvBL1 | K120 | K287 | x | x | 19,1 | RuvBL1 | RuvBL2 | K73 | T7 | | x | - |
| RuvBL1 | RuvBL1 | K121 | K281 | | x | 11,8 | RuvBL1 | RuvBL2 | K73 | K417 | | x | 17,8 |
| RuvBL1 | RuvBL1 | K138 | K214 | | x | 26 | RuvBL1 | RuvBL2 | K138 | K9 | | x | - |
| RuvBL1 | RuvBL1 | T140 | K238 | | x | 27,3 | RuvBL1 | RuvBL2 | T176 | K365 | | x | 78,3 |
| RuvBL1 | RuvBL1 | Y144 | K181 | | x | 9,5 | RuvBL1 | RuvBL2 | K178 | K427 | | x | 67 |
| RuvBL1 | RuvBL1 | S168 | K184 | | x | 10,7 | RuvBL1 | RuvBL2 | T180 | K9 | | x | - |
| RuvBL1 | RuvBL1 | K175 | K181 | x | | 6,6 | RuvBL1 | RuvBL2 | K181 | K9 | | x | - |
| RuvBL1 | RuvBL1 | T176 | K238 | | x | 6,5 | RuvBL1 | RuvBL2 | K181 | K67 | | x | 48,3 |
| RuvBL1 | RuvBL1 | T176 | K243 | | x | 12,4 | RuvBL1 | RuvBL2 | K181 | K223 | | x | 56,2 |
| RuvBL1 | RuvBL1 | K178 | K238 | | x | 7 | RuvBL1 | RuvBL2 | T180/K181 | K427 | x | x | 72,6 |
| RuvBL1 | RuvBL1 | T180 | K238 | | x | 12 | RuvBL1 | RuvBL2 | K184 | K427 | | x | 67,5 |
| RuvBL1 | RuvBL1 | K184 | K287 | | x | 39,6 | RuvBL1 | RuvBL2 | K195 | K67 | | x | 46,4 |
| RuvBL1 | RuvBL1 | K195 | K214 | x | | 10,1 | RuvBL1 | RuvBL2 | K281 | K269 | x | | 13,5 |
| RuvBL1 | RuvBL1 | K278 | K281 | | x | 8,2 | RuvBL1 | RuvBL2 | K281 | K279 | x | | 27,2 |
| RuvBL1 | RuvBL1 | K278 | K287 | | x | 20,4 | RuvBL1 | RuvBL2 | K281 | K365 | x | | 49,8 |
| RuvBL1 | RuvBL1 | K278 | T282 | | x | 11,7 | RuvBL1 | RuvBL2 | K281 | K427 | | x | 42,2 |
| RuvBL1 | RuvBL1 | K278 | T285 | | x | 19,6 | RuvBL1 | RuvBL2 | T282 | K279 | | x | 29,2 |
| RuvBL1 | RuvBL1 | T282 | K294 | | x | 20,5 | RuvBL1 | RuvBL2 | T285 | K279 | | x | 23,9 |
| RuvBL1 | RuvBL1 | K281 | K294 | x | x | 22,8 | RuvBL1 | RuvBL2 | K287 | T8 | | x | - |
| RuvBL1 | RuvBL1 | T285 | K294 | | x | 13,2 | RuvBL1 | RuvBL2 | K287 | K9 | | x | - |
| RuvBL1 | RuvBL1 | K287 | K294 | x | x | 10,5 | RuvBL1 | RuvBL2 | K287 | K269 | x | | 21,8 |
| RuvBL1 | RuvBL1 | K287 | K431 | | x | 37,6 | RuvBL1 | RuvBL2 | K287 | S270 | | x | 25 |
| RuvBL1 | RuvBL1 | K385 | T415 | | x | 19,9 | RuvBL1 | RuvBL2 | K287 | T367 | | x | 35,6 |
| RuvBL1 | RuvBL1 | K413 | S437 | | x | 23,5 | RuvBL1 | RuvBL2 | K287 | K427 | x | x | 43,8 |
| RuvBL1 | RuvBL1 | K413 | S455 | | x | 11,3 | RuvBL1 | RuvBL2 | K294 | T8 | | x | - |
| RuvBL1 | RuvBL1 | T415 | K440 | | x | 18,5 | RuvBL1 | RuvBL2 | K294 | K9 | | x | - |
| RuvBL1 | RuvBL1 | K435 | K389 | x | | 18,6 | RuvBL1 | RuvBL2 | K294 | T183 | | x | 35,1 |
| RuvBL1 | RuvBL1 | K454 | K466 | x | | - | RuvBL1 | RuvBL2 | K294 | T199 | | x | 27,8 |
| RuvBL1 | RuvBL1 | K454 | Y467 | | x | - | RuvBL1 | RuvBL2 | K294 | K201 | | x | 24,2 |

| Intra-crosslink | | | | Merox | Maxlynx | distance Ca-Ca | Intra-crosslink | | | | Merox | Maxlynx | distance Ca-Ca |
|-----------------|-----------|--------|--------|-------|---------|-------------------|-----------------|-----------|--------|--------|-------|---------|-------------------|
| Protein 1 | Protein 2 | Site 1 | Site 2 | | | 5OAF | Protein 1 | Protein 2 | Site 1 | Site 2 | | | 5OAF |
| RuvBL2 | RuvBL2 | T8 | K427 | | x | - | RuvBL1 | RuvBL2 | K294 | K269 | x | | 24 |
| RuvBL2 | RuvBL2 | K9 | K201 | | x | - | RuvBL1 | RuvBL2 | K294 | S270 | | x | 36 |
| RuvBL2 | RuvBL2 | K9 | S270 | | x | - | RuvBL1 | RuvBL2 | K389 | T367 | | x | 45,3 |
| RuvBL2 | RuvBL2 | K9 | T367 | | x | - | RuvBL1 | RuvBL2 | K389 | K427 | x | x | 44,6 |
| RuvBL2 | RuvBL2 | K9 | T419 | | x | - | RuvBL1 | RuvBL2 | K431 | K67 | | x | 13 |
| RuvBL2 | RuvBL2 | K9 | K427 | | x | - | RuvBL1 | RuvBL2 | K435 | K160 | x | | 95,8 |
| RuvBL2 | RuvBL2 | K132 | K201 | | x | 22,2 | RuvBL1 | RuvBL2 | K435 | K186 | | x | 63,5 |
| RuvBL2 | RuvBL2 | S156 | K186 | | x | 21,2 | RuvBL1 | RuvBL2 | K435 | T367 | | x | 37,7 |
| RuvBL2 | RuvBL2 | K157 | T183 | | x | 12,9 | RuvBL1 | RuvBL2 | S437 | K9 | | x | - |
| RuvBL2 | RuvBL2 | K157 | K184 | | x | 16,3 | RuvBL1 | RuvBL2 | K458 | K427 | x | x | 40,2 |
| RuvBL2 | RuvBL2 | K157 | K186 | x | | 18,9 | RuvBL1 | RuvBL2 | K466 | K427 | x | | - |
| RuvBL2 | RuvBL2 | K160 | T199 | | x | 19 | RuvBL1 | RuvBL2 | K466 | Y442 | | x | - |
| RuvBL2 | RuvBL2 | K160 | K427 | | x | 78,2 | RuvBL1 | RuvBL2 | K466 | K444 | x | | - |
| RuvBL2 | RuvBL2 | T183 | K201 | | x | 11,4 | RuvBL1 | RuvBL2 | K469 | K427 | x | x | - |

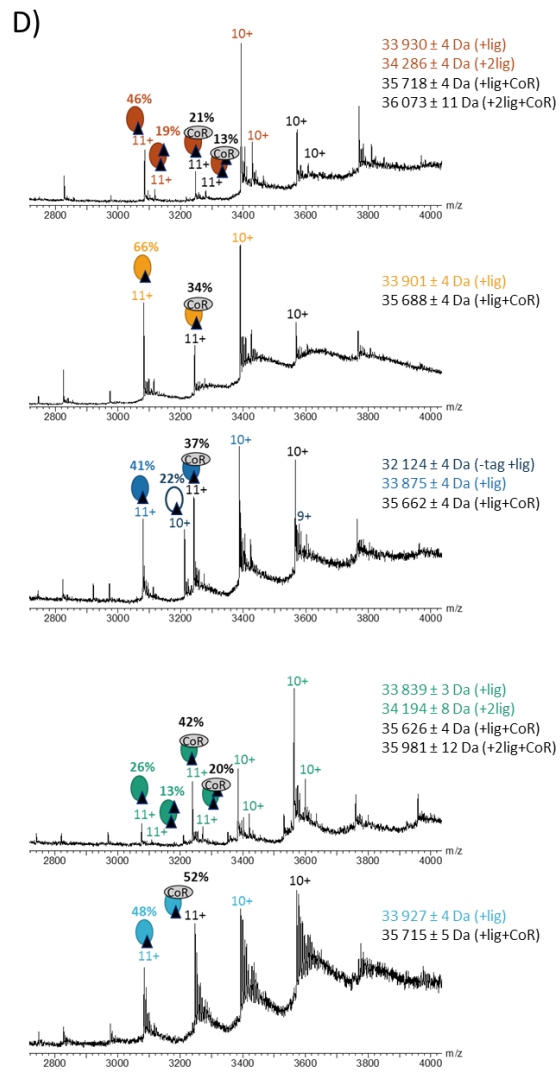
Annexe 16

Spectres de masses obtenus en MS native d'un réplicat des analyses effectuées sur chacune des protéines (T475M, WT, Y505H, H494Y et F310S) en présence de (A) PGC1 α (CoA), (B) PGC1 α et GW1929 (lig), (C) NCoR (CoR), (D) NCoR et SR16832 (lig) et (E) NCoR (CoR) et T0070907 (lig). Les masses, espèces et taux de complexation sont décrits. Les tableaux récapitulatifs des moyennes et écarts types des taux de complexation de PPAR γ ^{LBD} avec les co-régulateurs et ligands.

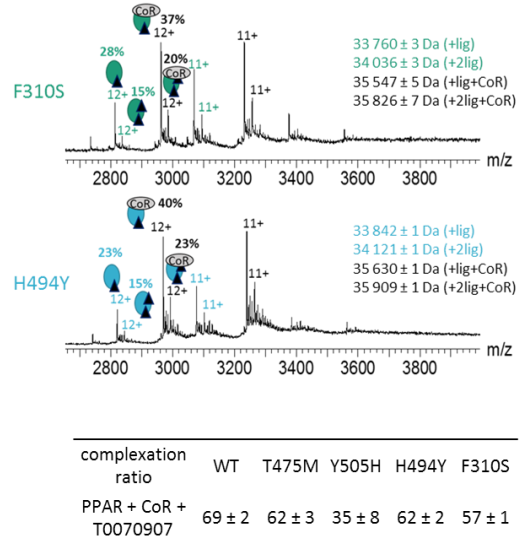
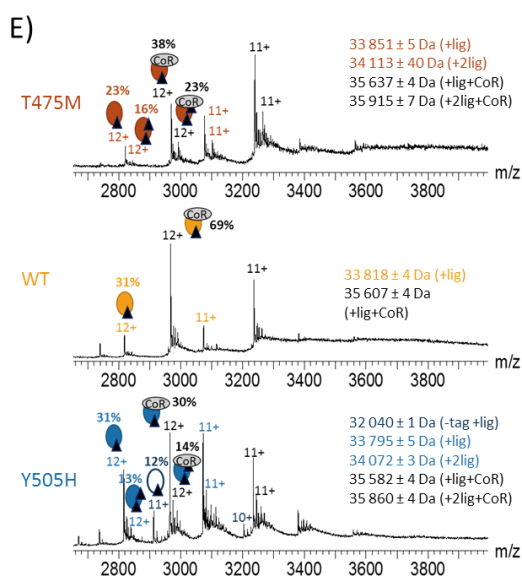




| complexation ratio | WT | T475M | Y505H | H494Y | F310S |
|--------------------|--------|--------|--------|--------|--------|
| PPAR + CoR | 35 ± 1 | 40 ± 4 | 43 ± 1 | 45 ± 2 | 63 ± 3 |



| complexation ratio | WT | T475M | Y505H | H494Y | F310S |
|----------------------|--------|--------|--------|--------|--------|
| PPAR + CoR + SR16832 | 38 ± 4 | 39 ± 4 | 37 ± 3 | 51 ± 2 | 54 ± 6 |

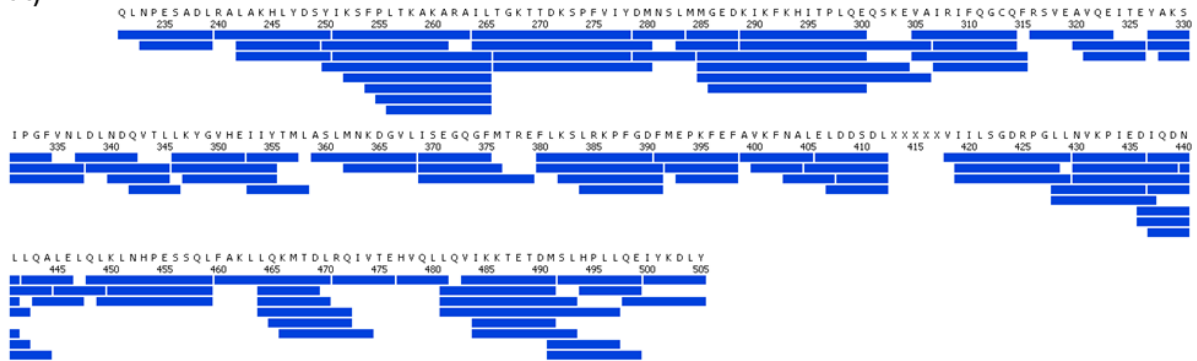


| complexation ratio | WT | T475M | Y505H | H494Y | F310S |
|-----------------------|--------|--------|--------|--------|--------|
| PPAR + CoR + T0070907 | 69 ± 2 | 62 ± 3 | 35 ± 8 | 62 ± 2 | 57 ± 1 |

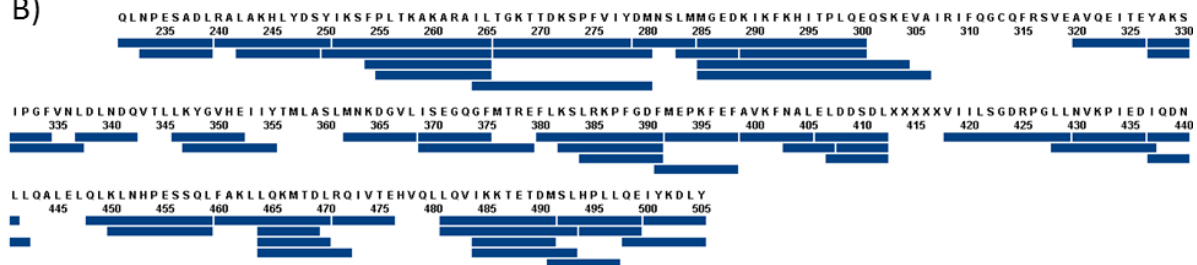
Annexe 17

Cartes de couverture de séquence obtenus en HDX-MS pour PPAR γ LBD (A) WT avant deutération, (B) F310S et (C) WT après deutération.

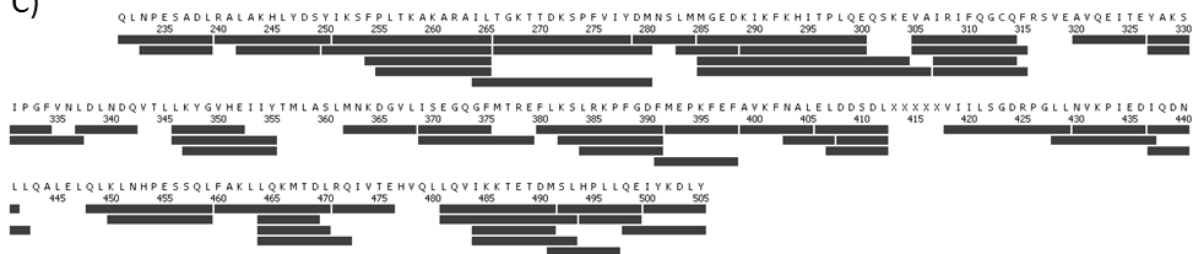
A)



B)

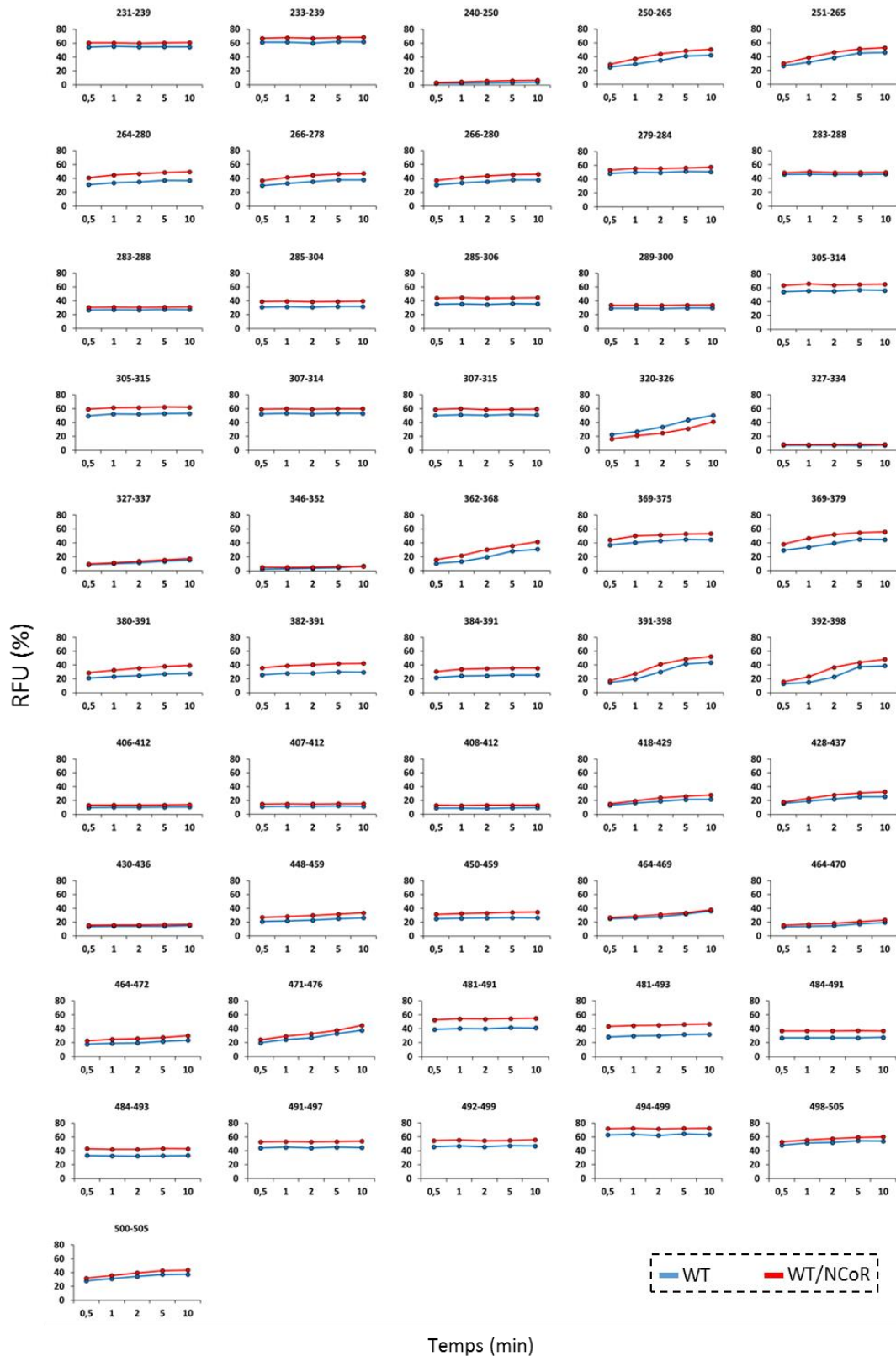


C)



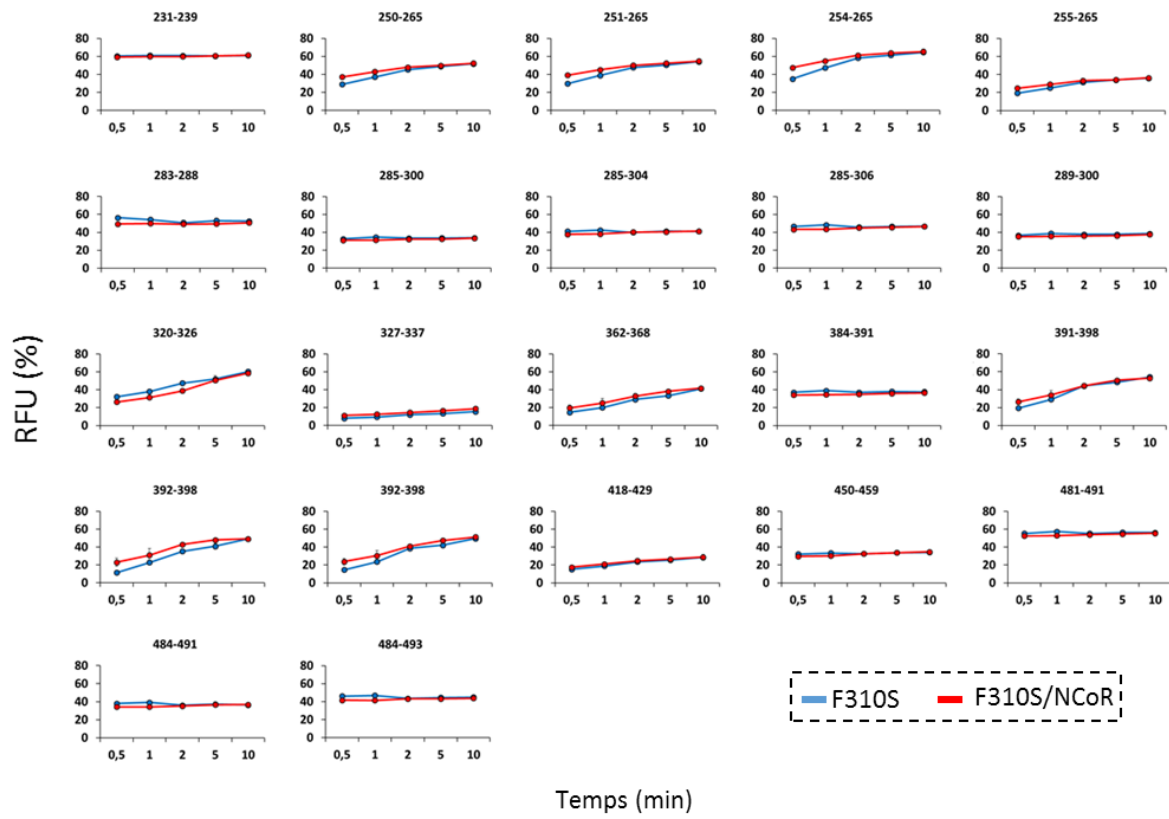
Annexe 18

Incorporation relative en D aux différents temps de deutération de chaque peptide de PPAR γ ^{LBD} statistiquement significatif (MEMHDX avec une p-value à 0,01) obtenus en HDX-MS (WT – WT/NCoR).



Annexe 19

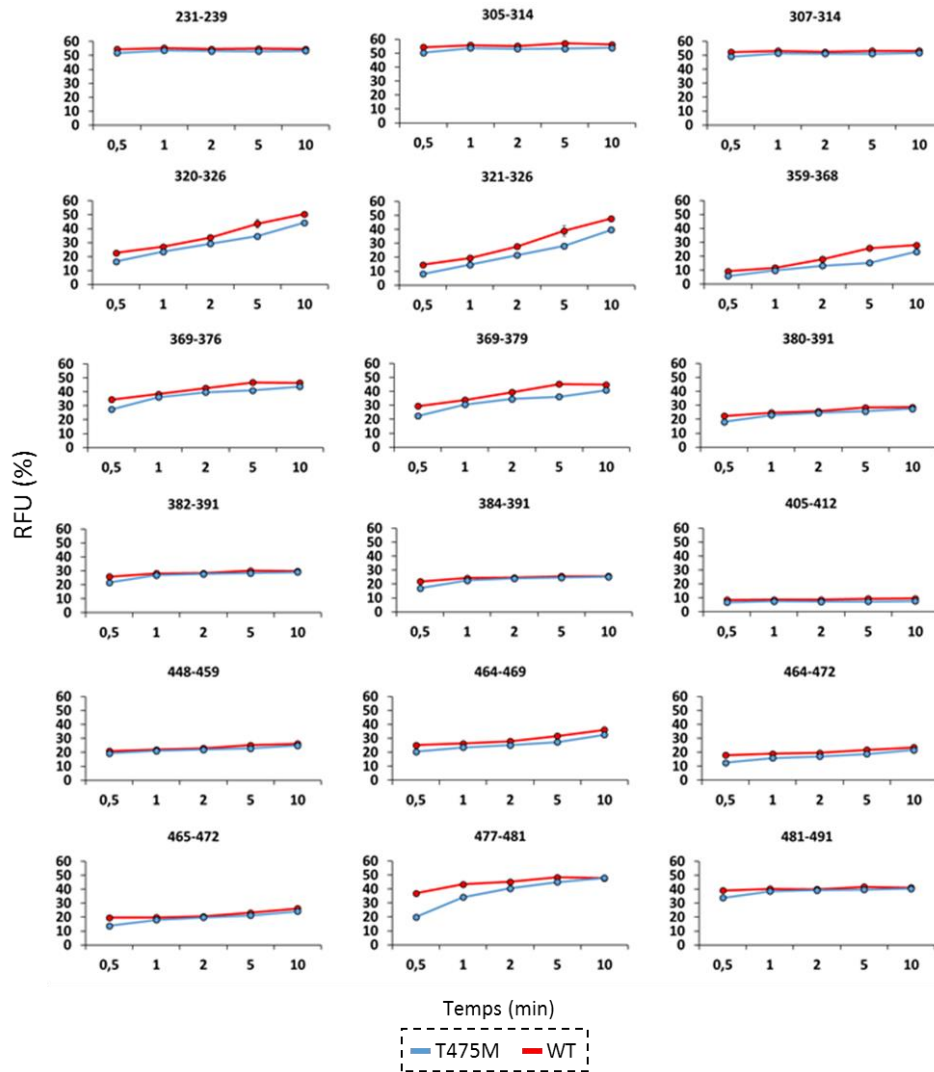
Incorporation relative en D aux différents temps de deutération de chaque peptide de PPAR γ ^{LBD} statistiquement significatif (MEMHDX avec une p-value à 0,01) obtenus en HDX-MS (F310S – F310S/NCoR).



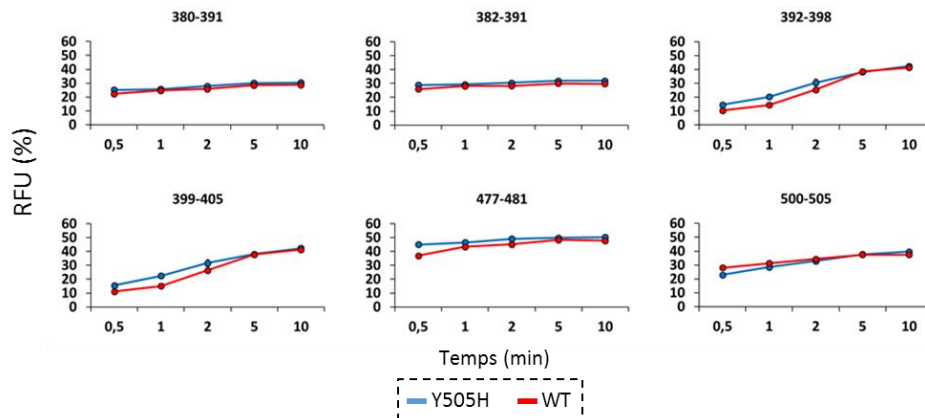
Annexe 20

Incorporation relative en D aux différents temps de deutération de chaque peptide de PPAR γ ^{LBD} statistiquement significatif (MEMHDX avec une p-value à 0,01) obtenus en HDX-MS (WT-mutant).

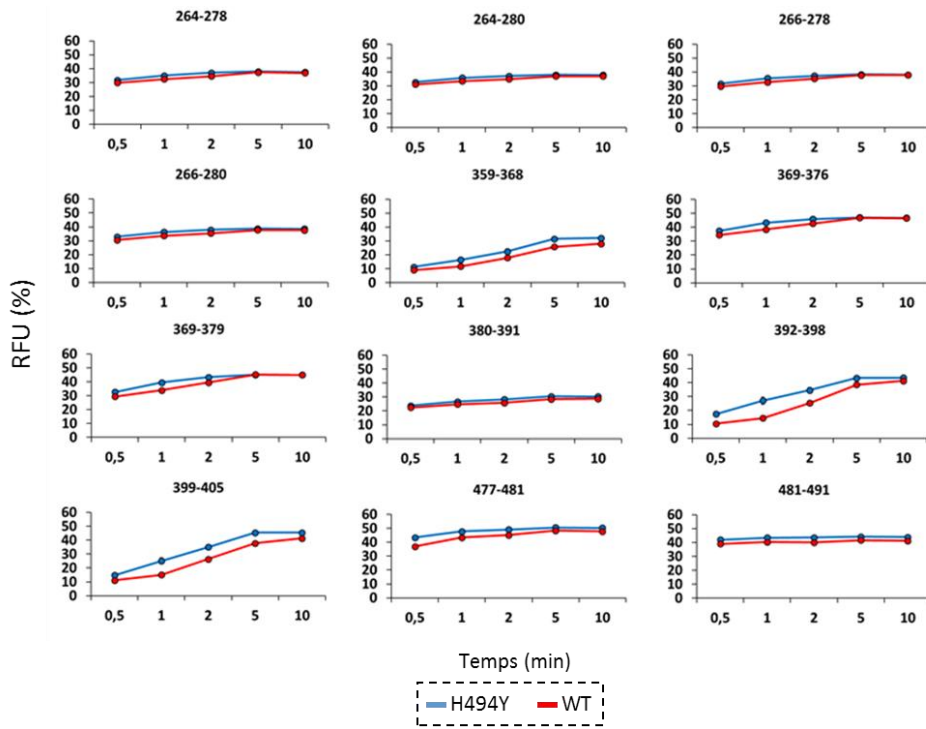
WT-T475M :



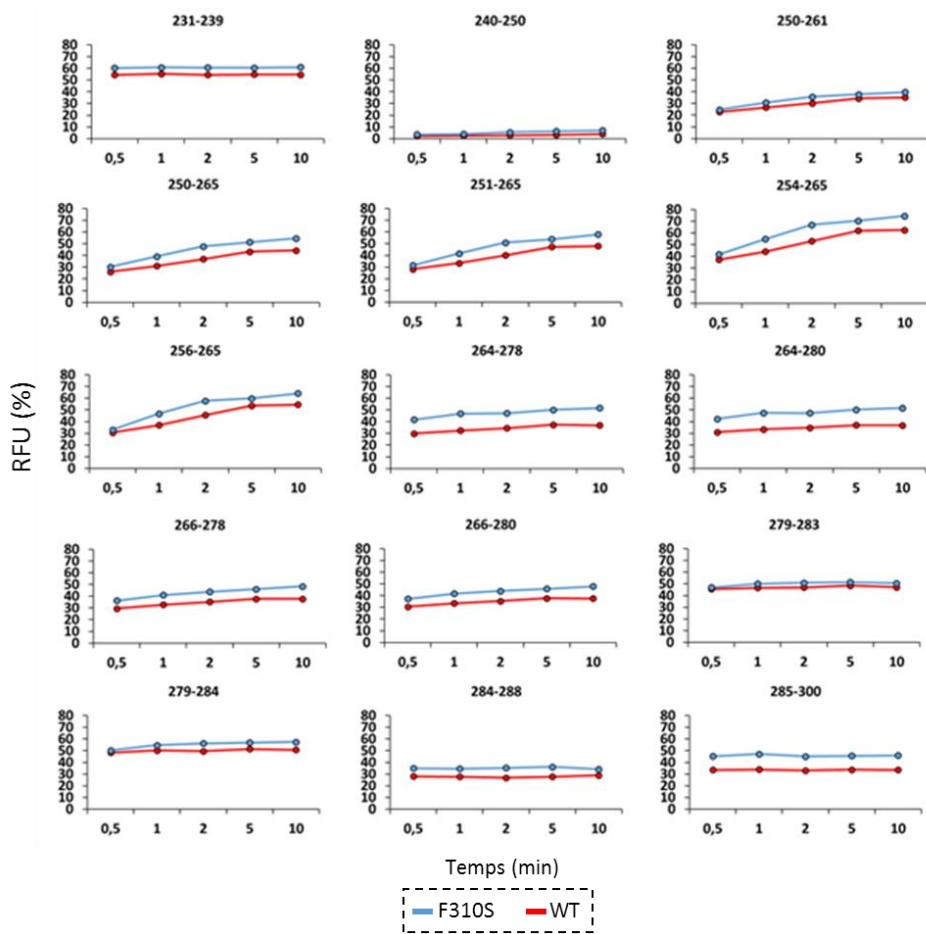
WT-Y505H :

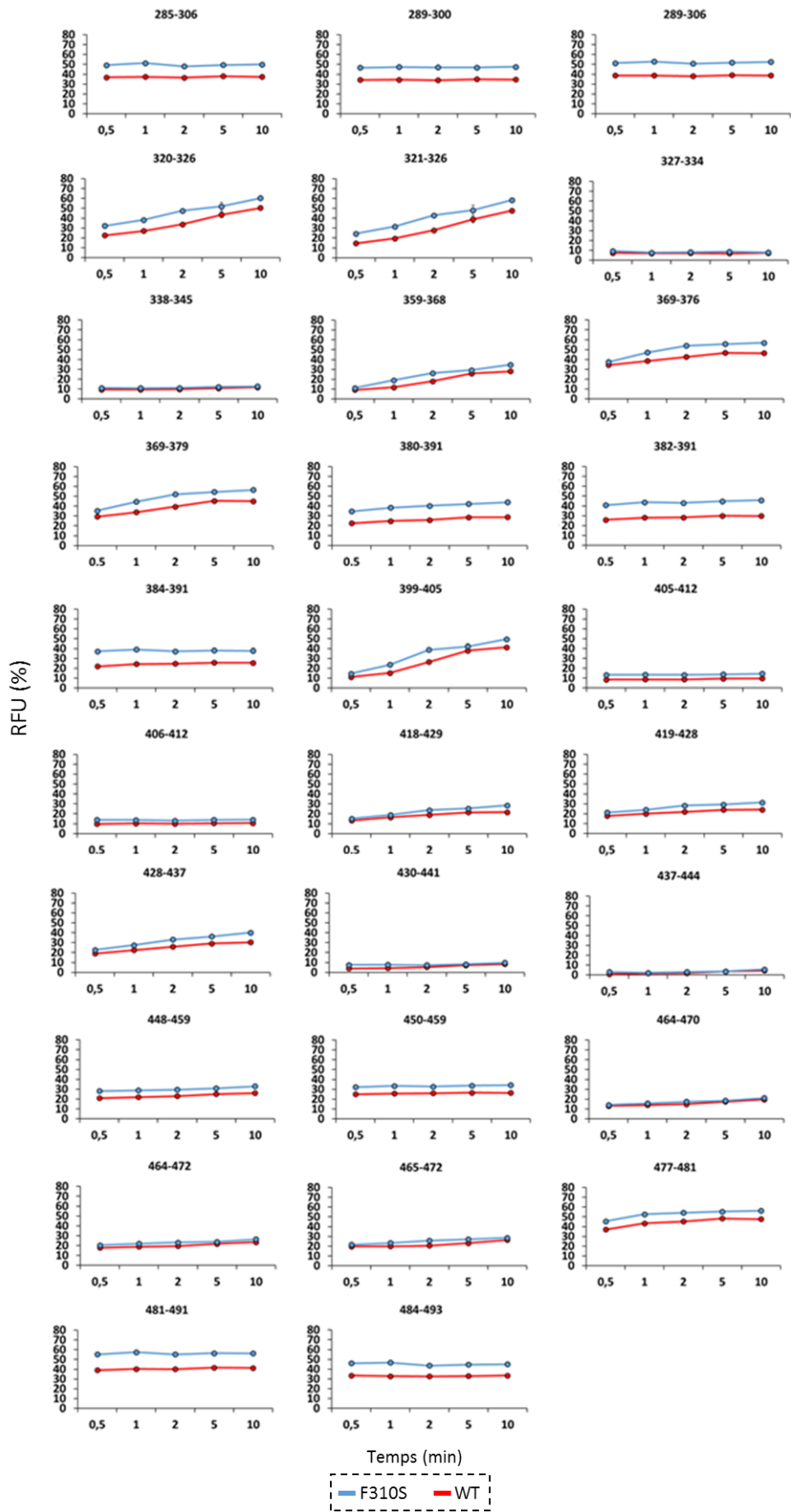


WT-H494Y :



WT-F310S :

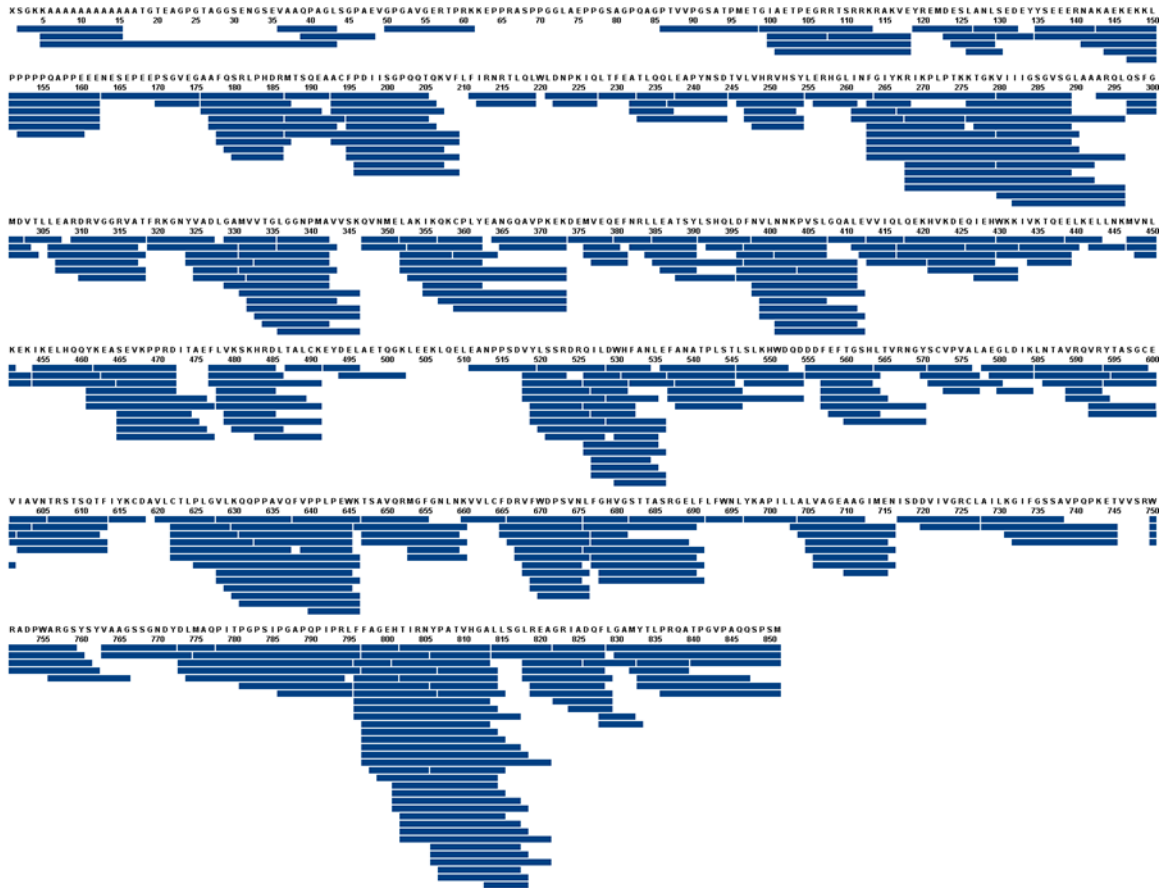




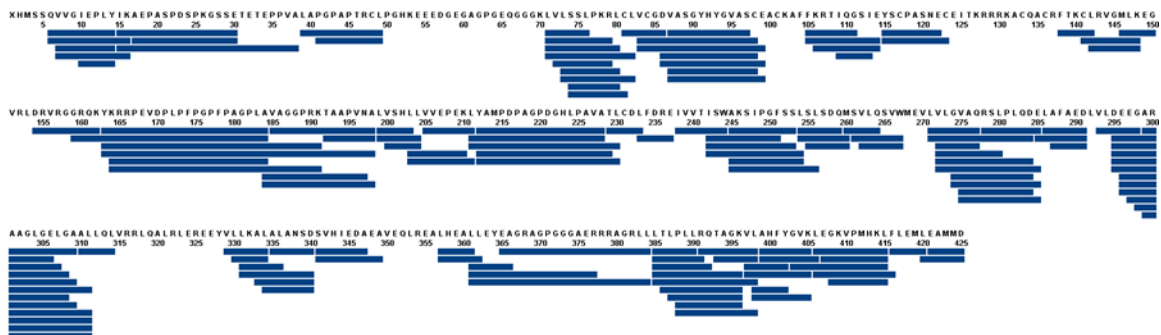
Annexe 21

Cartes de couverture de séquence obtenus en HDX-MS pour ERR α et LSD1 (A) seule, (B) en complexe avant deutération et (C) après deutération.

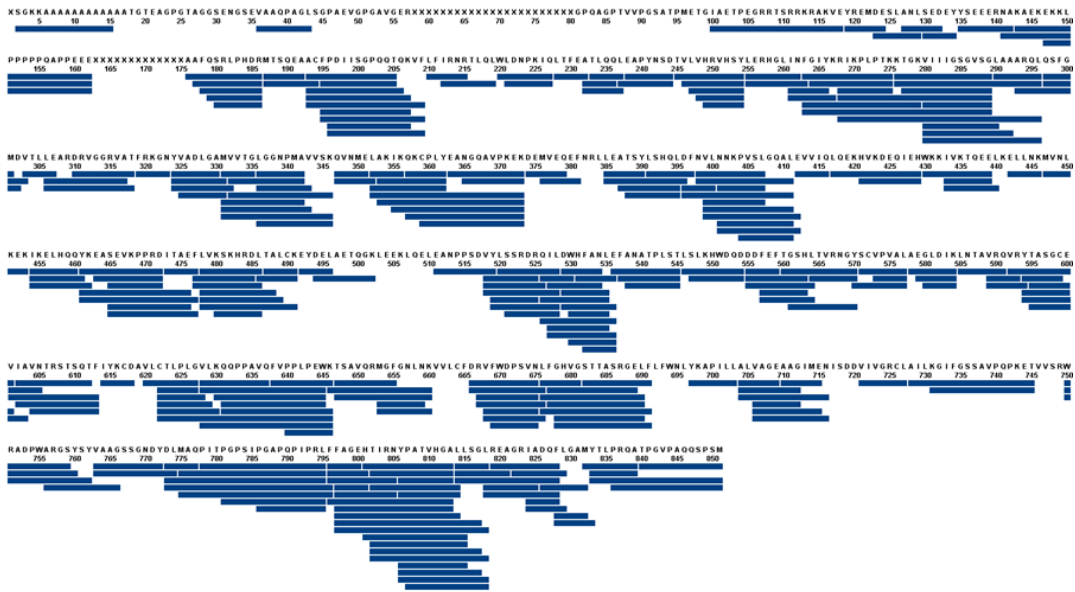
A) LSD1



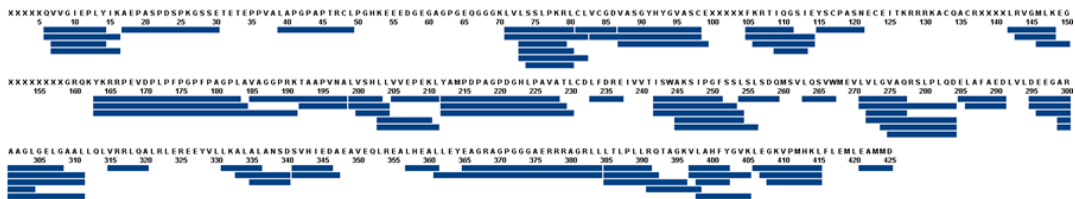
ERR α



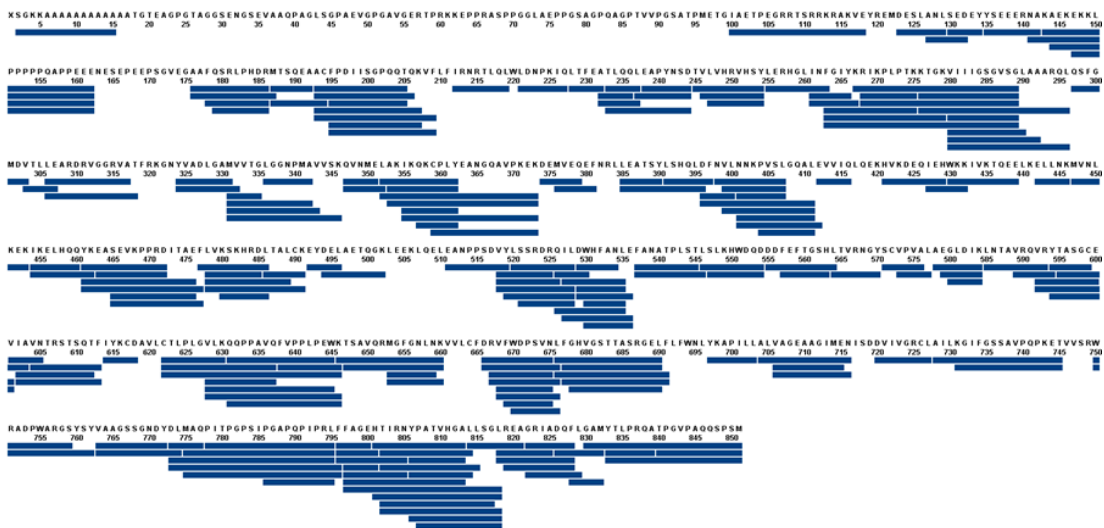
B) LSD1



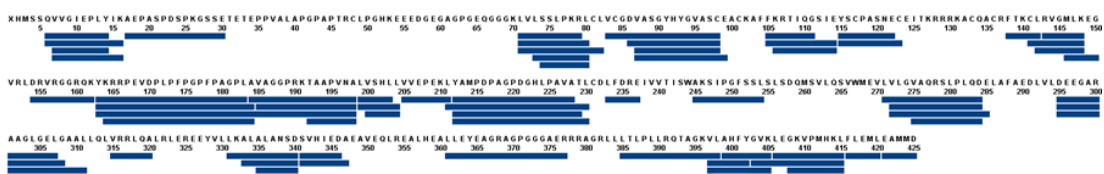
ERR α



C) LSD1

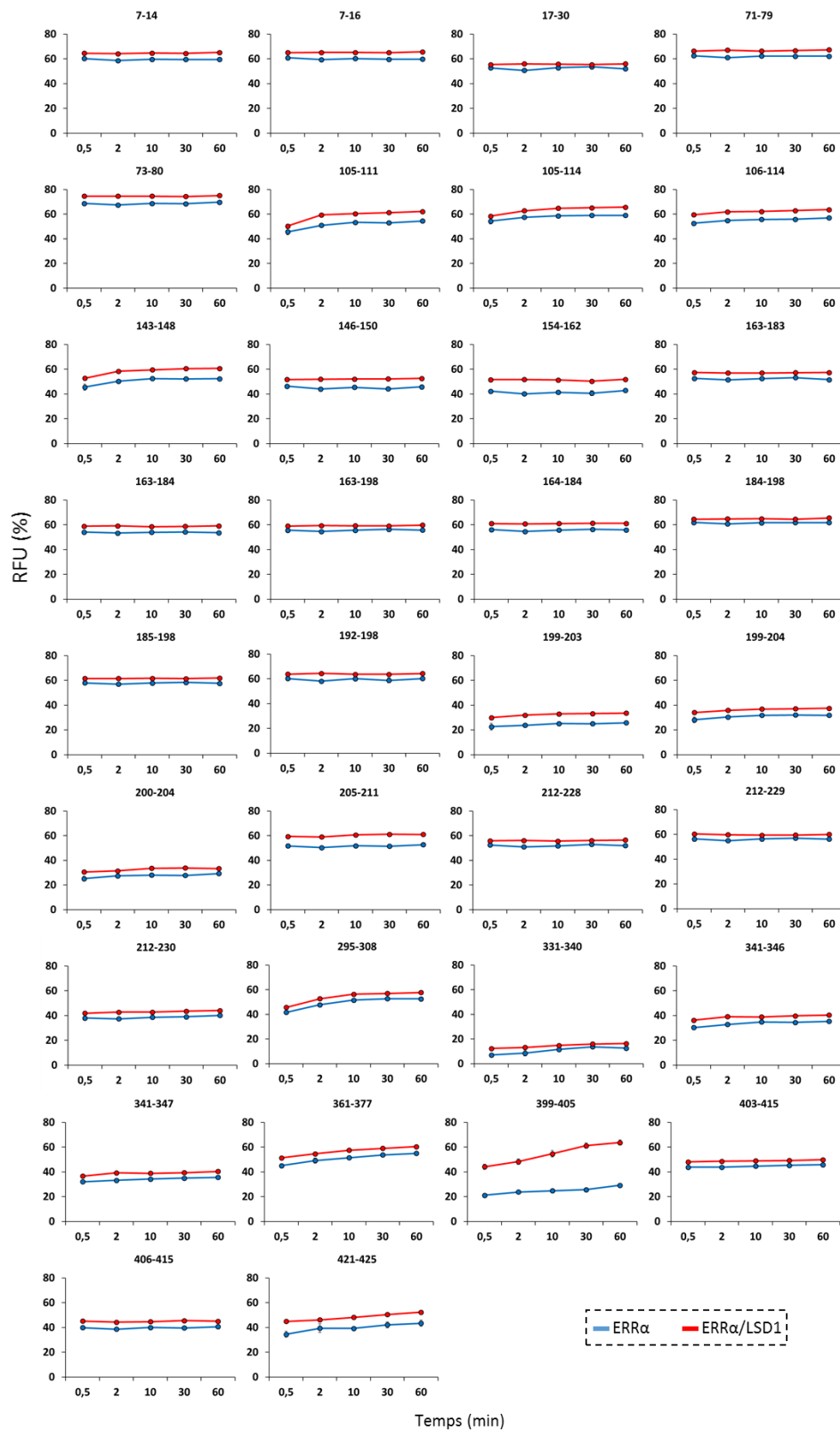


ERR α



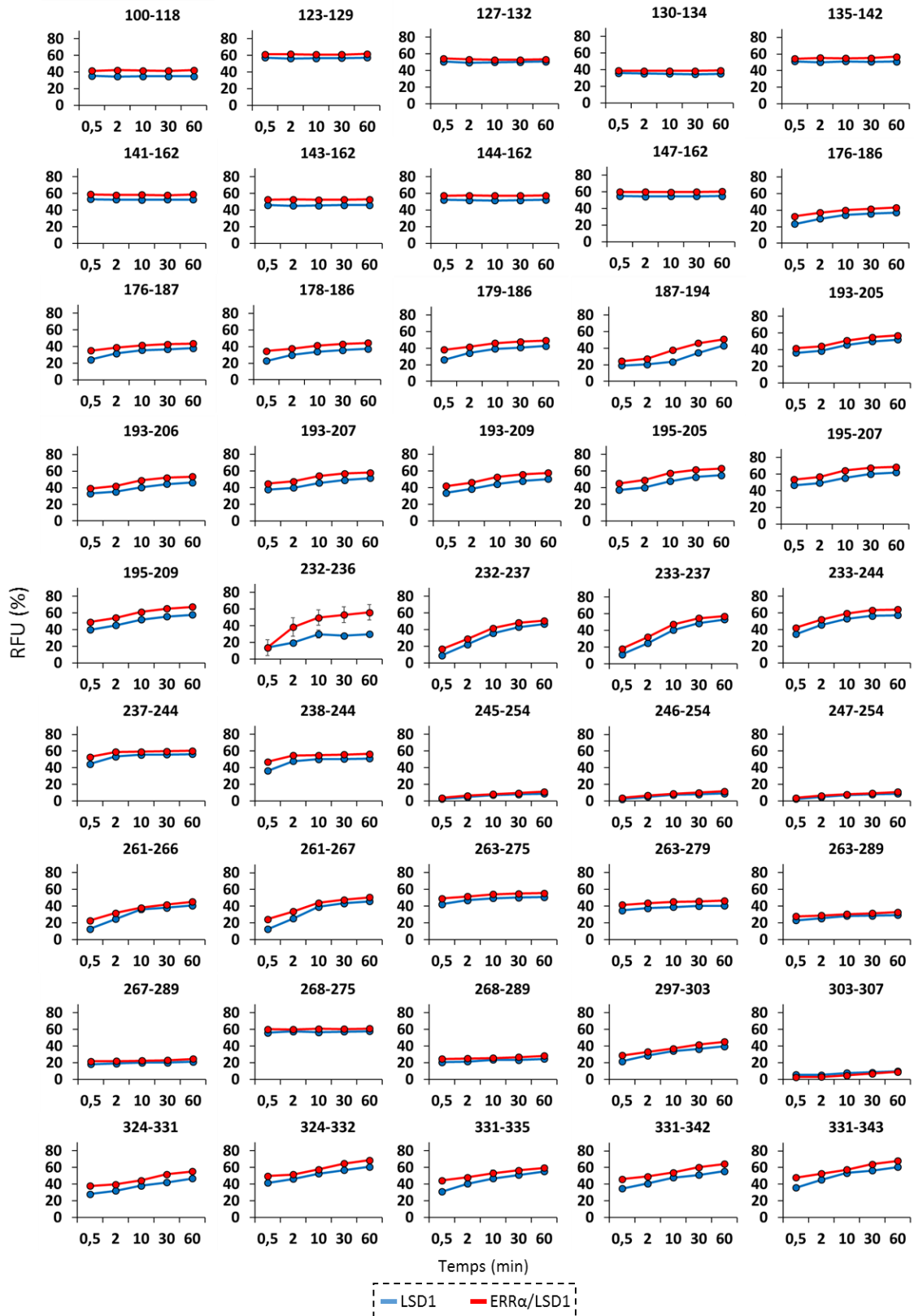
Annexe 22

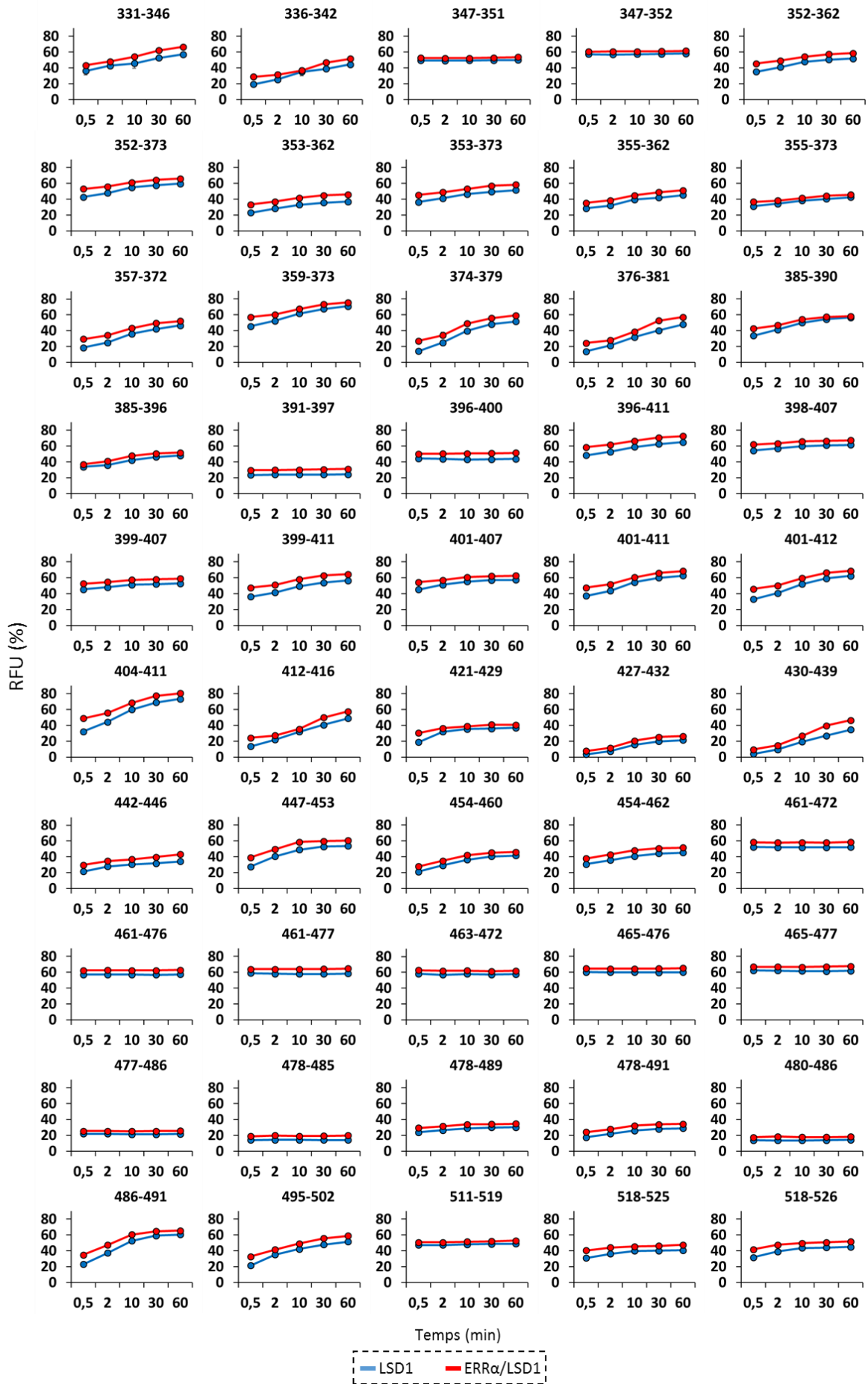
Incorporation relative en D aux différents temps de deutération de chaque peptide de ERR α statistiquement significatif (MEMHDX avec une p-value à 0,01) obtenus en HDX-MS (ERR α – ERR α /LSD1).

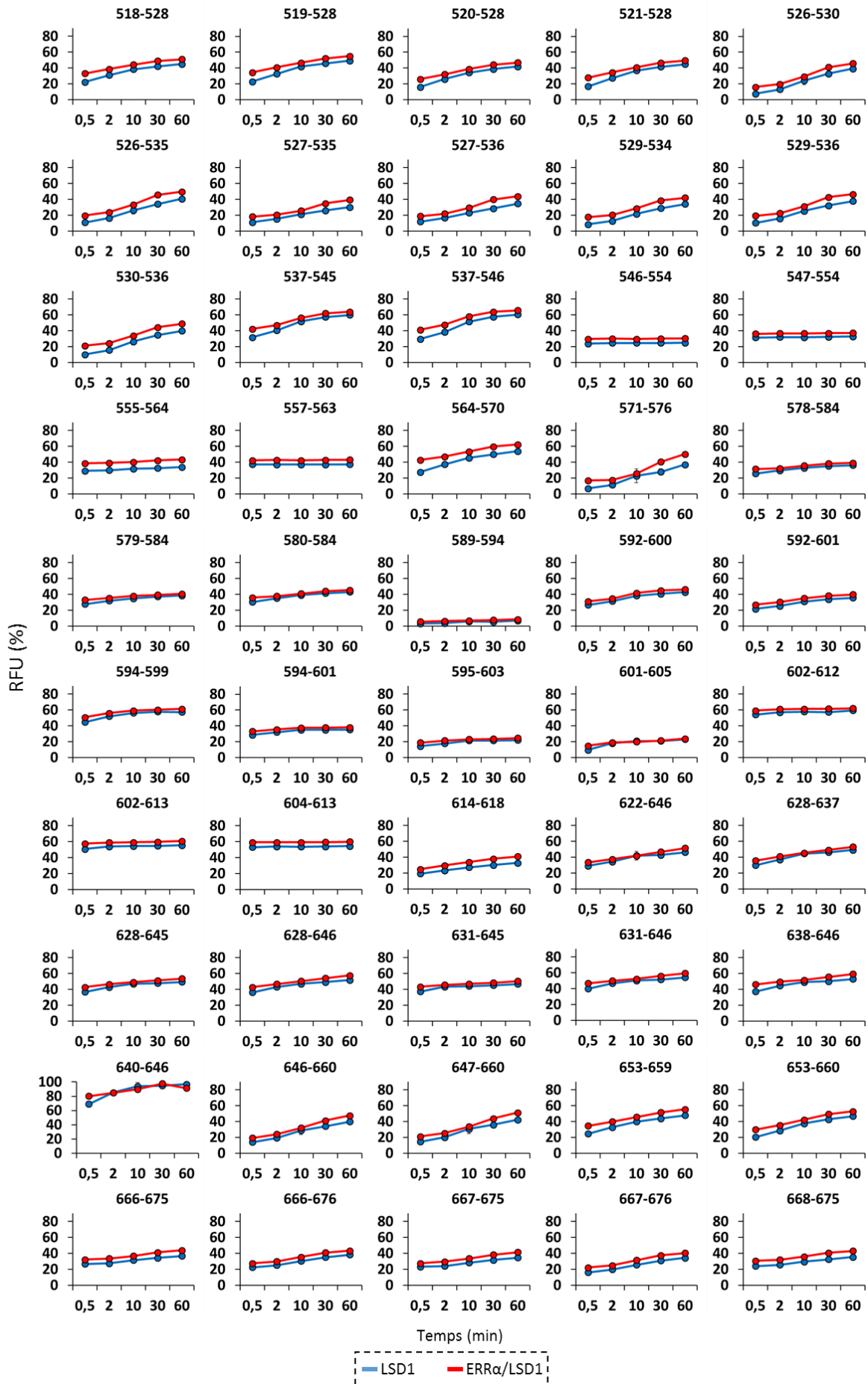


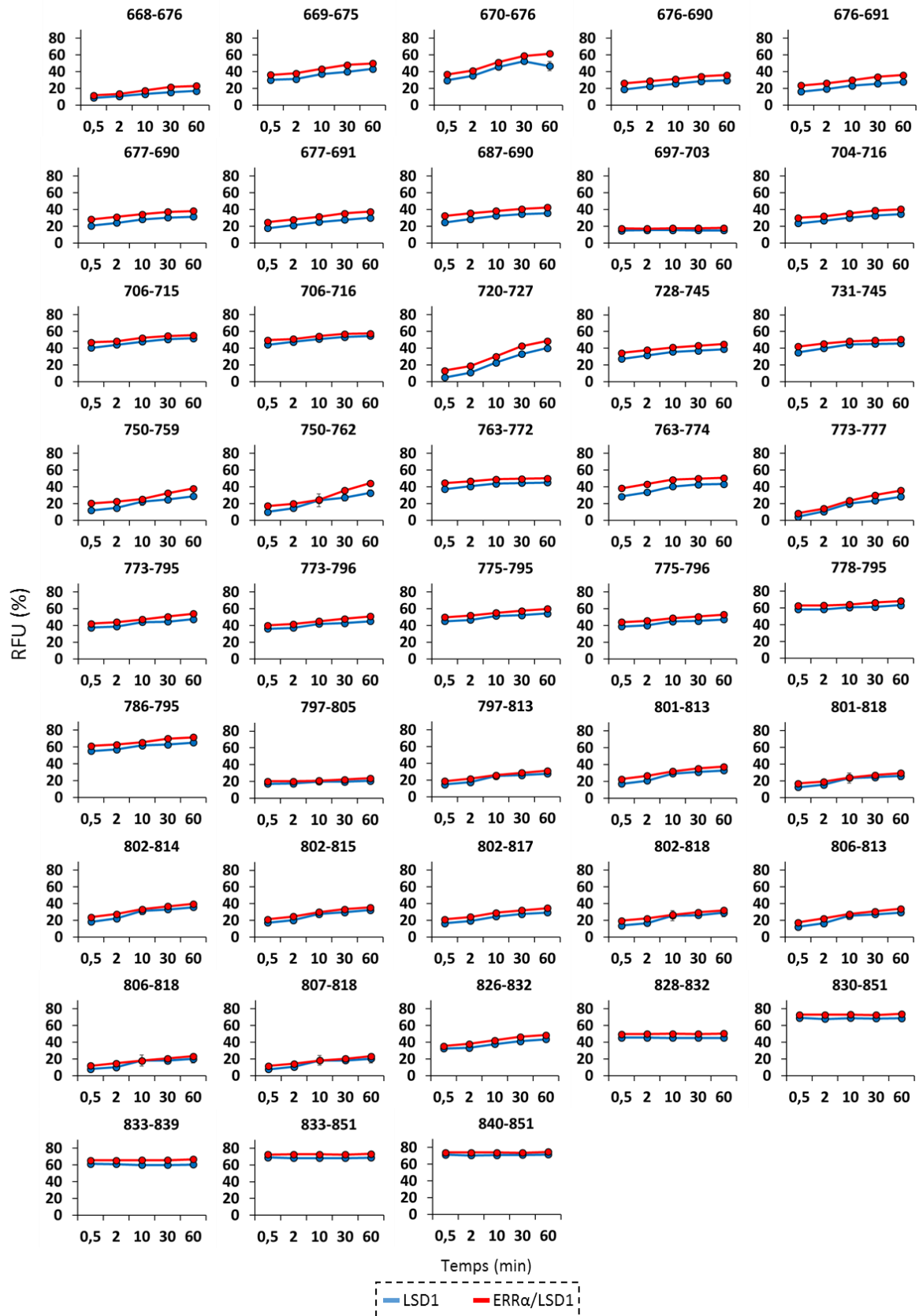
Annexe 23

Incorporation relative en D aux différents temps de deutération de chaque peptide de LSD1 statistiquement significatif (MEMHDX avec une p-value à 0,01) obtenus en HDX-MS (LSD1 – ERR α /LSD1).









Annexe 24

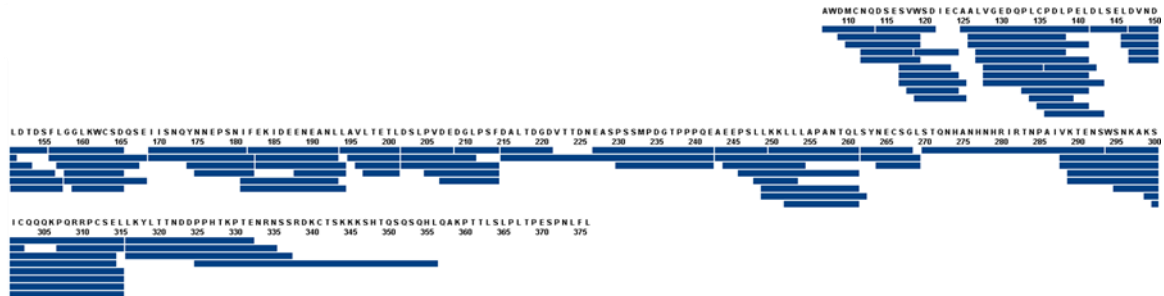
Tableau descriptif des peptides intra-ERR α et intra-LSD1 cross-linkés (protéines, sites) identifiés suite au XL-MS du complexe ERR α /LSD1.

| protéine 1 | protéine 2 | site 1 | site 2 |
|------------|------------|--------|--------|
| LSD1 | LSD1 | K62 | K116 |
| LSD1 | LSD1 | K62 | K267 |
| LSD1 | LSD1 | K116 | K267 |
| LSD1 | LSD1 | K116 | K270 |
| LSD1 | LSD1 | K116 | K354 |
| LSD1 | LSD1 | K116 | K371 |
| LSD1 | LSD1 | K116 | K451 |
| LSD1 | LSD1 | K270 | Y833 |
| LSD1 | LSD1 | K275 | K616 |
| LSD1 | LSD1 | K279 | K584 |
| LSD1 | LSD1 | K279 | K616 |
| LSD1 | LSD1 | K354 | Y362 |
| LSD1 | LSD1 | K373 | K743 |
| LSD1 | LSD1 | K431 | K441 |
| LSD1 | LSD1 | K431 | K506 |
| LSD1 | LSD1 | K435 | K506 |
| LSD1 | LSD1 | K451 | K462 |
| LSD1 | LSD1 | K451 | K468 |
| LSD1 | LSD1 | K451 | K480 |
| LSD1 | LSD1 | K451 | K491 |
| LSD1 | LSD1 | K480 | K491 |
| ERR | ERR | K78 | K140 |
| ERR | ERR | K106 | K131 |
| ERR | ERR | K106 | K148 |

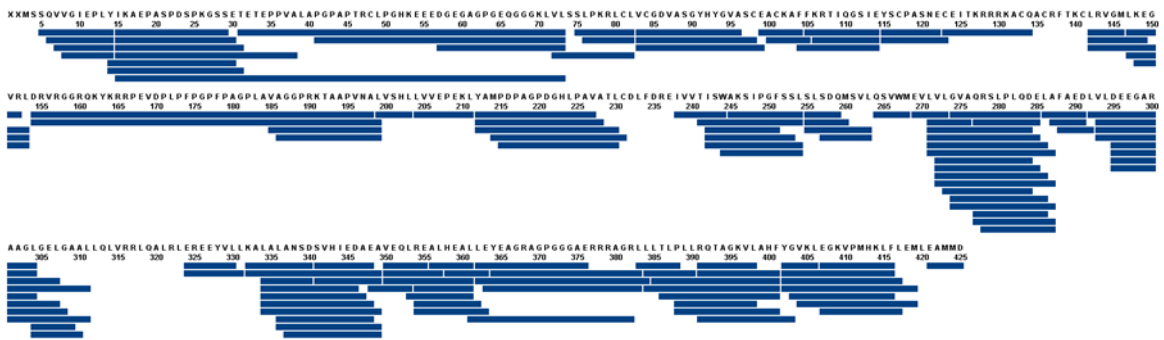
Annexe 25

Cartes de couverture de séquence obtenus en HDX-MS pour ERR α et PGC1 α (A) seule, (B) en complexe avant deutération et (C) après deutération.

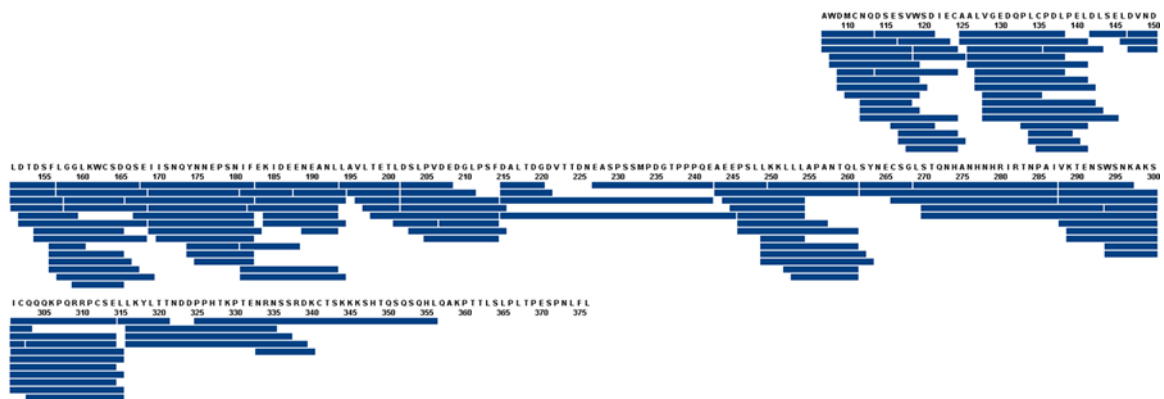
A) PGC1 α



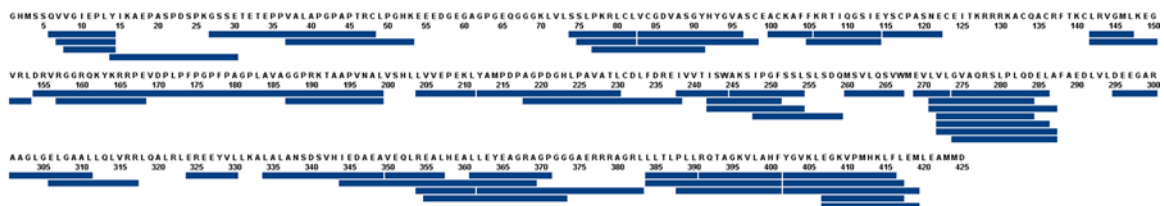
ERR α



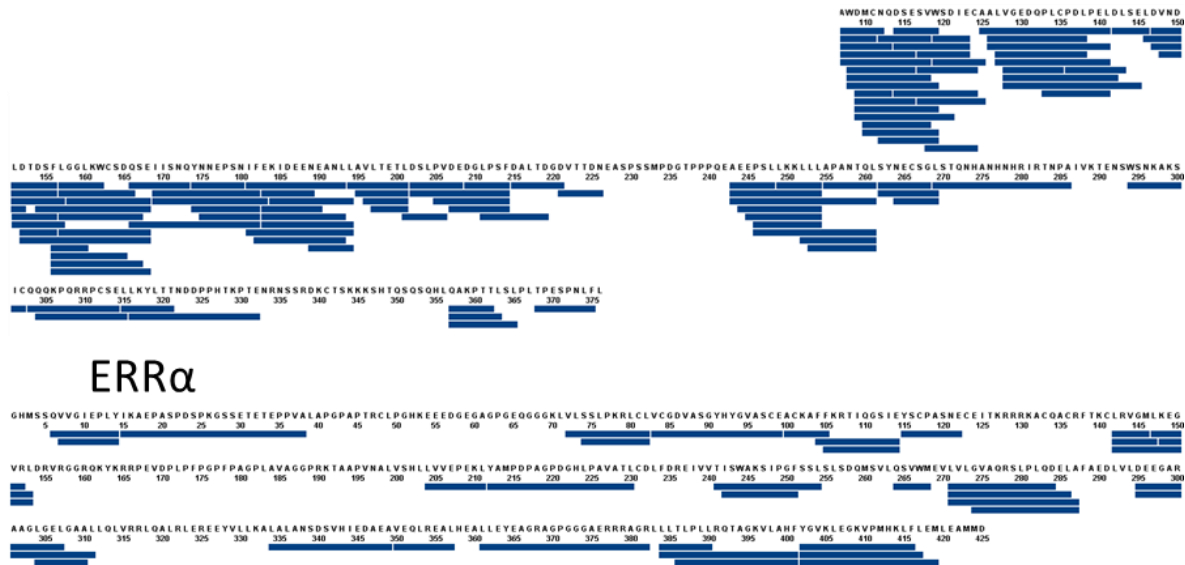
B) PGC1 α



ERR α

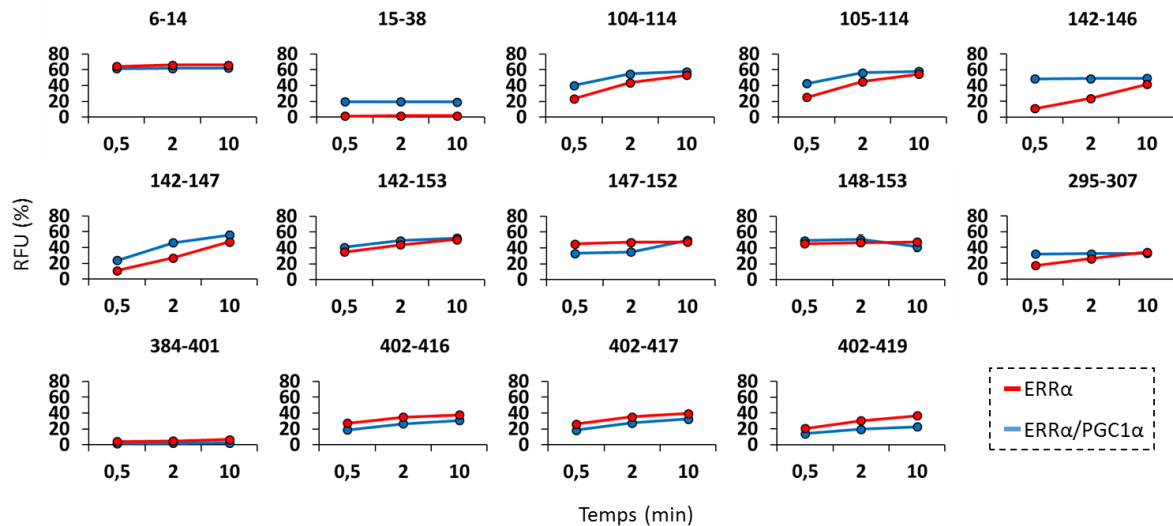


C) PGC1 α



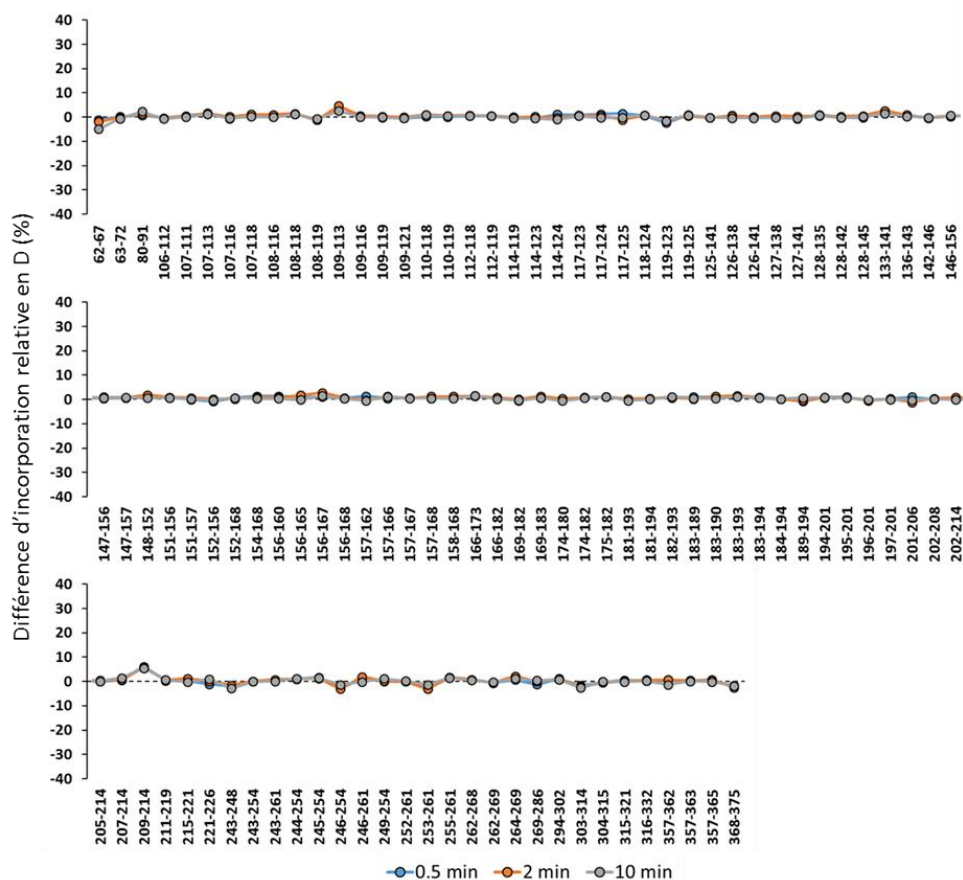
Annexe 26

Incorporation relative en D aux différents temps de deutération de chaque peptide de ERR α statistiquement significatif (MEMHDX avec une p-value à 0,01) obtenus en HDX-MS (ERR α – ERR α /PGC1 α).



Annexe 27

Représentation des différences d'incorporation relative en D entre PGC1 α et en complexe avec ERR α , pour chaque peptide identifié de PGC1 α aux différents temps de deutération.



Annexe 28

Tableau descriptif des peptides intra-ERR α et intra-PGC1 α cross-linkés (protéines, sites, cross-linker, domaines et distances Ca-Ca) identifiés suite au XL-MS du complexe ERR α /PGC1 α .

| Protéine 1 | Protéine 2 | Site 1 | Site 2 | DSAU | DSBU | Domaine site 1 | Domaine site 2 | Distance Ca-Ca (Å) ERR α /PGC1 α |
|------------|------------|--------|--------|------|------|----------------|----------------|---|
| ERR | ERR | K51 | K146 | | x | AF-1 | DBD | - |
| ERR | ERR | K51 | K189 | x | x | AF-1 | Hinge | - |
| ERR | ERR | K76 | K189 | | x | DBD | Hinge | - |
| ERR | ERR | S87 | K189 | | x | DBD | Hinge | - |
| ERR | ERR | K104 | K138 | | x | DBD | DBD | 14,6 |
| ERR | ERR | K104 | K146 | | x | DBD | DBD | 9,1 |
| ERR | ERR | K125 | K189 | | x | DBD | Hinge | - |
| ERR | ERR | K129 | K189 | x | x | DBD | Hinge | - |
| ERR | ERR | K138 | K138 | | x | DBD | DBD | 27,6 |
| ERR | ERR | K146 | K189 | x | x | DBD | Hinge | - |
| ERR | ERR | K189 | K189 | x | x | Hinge | Hinge | - |
| PGC1 | PGC1 | K306 | K329 | x | | - | - | - |

Annexe 29

Description des mutations (rag23 et rant21) sur la séquence du récepteur A_{2A}.

A_{2A} :

MWSHPQFEKHHHHHHHENLYFQGPIMGSSVYITVELAIAVLAILGNVLCWAVWLNSNLQNVNTNYFV
VSLAAADI L VGVLAIPFAITISTGFCAACHGCL A IACFVLVL A QSSIFSLLAIAIDRYIAIRIPLRYN
GLVTGARA A GIIAICWVLSFAIGLTPMLGWNNCGQPKEGKNHSQGC GEGQVACL FEDVVP MNYMVYFN
F F L CVLVPLLLMLGVYL A IFLAARRQ A KQMESQPLPGERARSTLQKEVHAAKSLAII A GLFALCWLPL
HIINCFTFFCPDCSHAPLWLMY A AIVLSHTNSV VNPFIYAYRIREFRQTFRKIIRSHVLRQQEPFKAA
GTSARVLA AHGSDGEQVSLRLNGHPPGVWANGSAPHPERRPNGYALGLVSGGSAQESQNTGLPDVEL
LSHELKGVCEPPGLDDPLAQDGAGVS

A_{2A}rag23 :

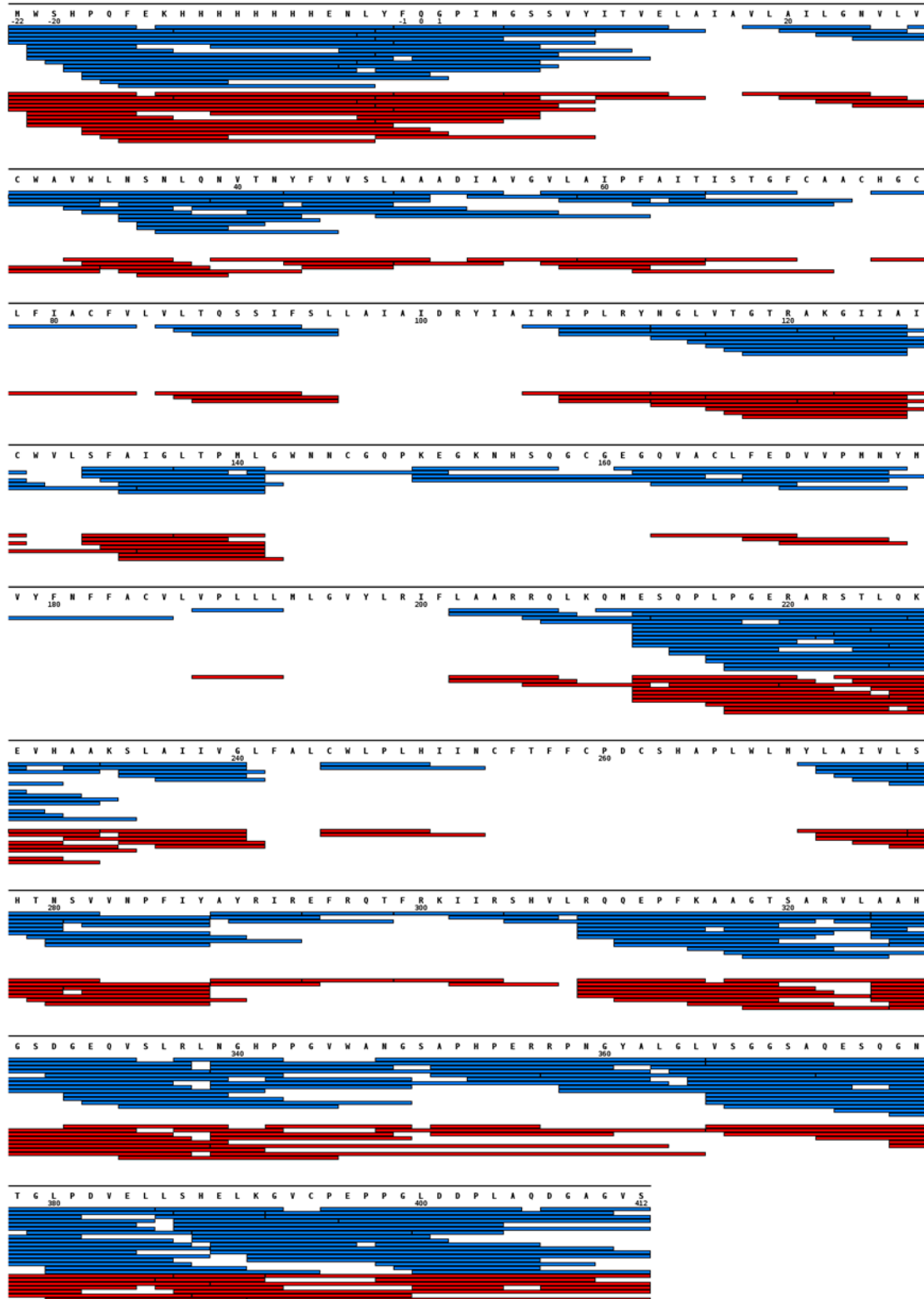
MWSHPQFEKHHHHHHHENLYFQGPIMGSSVYITVELAIAVLAILGNVLCWAVWLNSNLQNVNTNYFV
VSLAAADIL VGVLAIPFAITISTGFCAACHGCL F IACFVLVLA QSSIFSLLAIAIDRYIAIRIPLRYN
GLVTGTRAAGIIAICWVLSFAIGLTPMLGWNNCGQPKEGKNHSQGC GEGQVACL FEDVVP MNYMVYFN
F F A CVLVPLLLMLGVYL R IFLAARRQ L KQMESQPLPGERARSTLQKEVHAAKSLAII A GLFALCWLPL
HIINCFTFFCPDCSHAPLWLMY L AIVLSHTNSV VNPFIYAYRIREFRQTFRKIIRSHVLRQQEPFKAA
GTSARVLA AHGSDGEQVSLRLNGHPPGVWANGSAPHPERRPNGYALGLVSGGSAQESQNTGLPDVEL
LSHELKGVCEPPGLDDPLAQDGAGVS

A_{2A}rant21 :

MWSHPQFEKHHHHHHHENLYFQGPIMGSSVYITVELAIAVLAILGNVLCWAVWLNSNLQNVNTNYFV
VSLAAADI A VGVLAIPFAITISTGFCAACHGCL F IACFVLVL T QSSIFSLLAIAIDRYIAIRIPLRYN
GLVTGTRA K GIIAICWVLSFAIGLTPMLGWNNCGQPKEGKNHSQGC GEGQVACL FEDVVP MNYMVYFN
F F A CVLVPLLLMLGVYL R IFLAARRQ L KQMESQPLPGERARSTLQKEVHAAKSLAII V GLFALCWLPL
HIINCFTFFCPDCSHAPLWLMY L AIVLSHTNSV VNPFIYAYRIREFRQTFRKIIRSHVLRQQEPFKAA
GTSARVLA AHGSDGEQVSLRLNGHPPGVWANGSAPHPERRPNGYALGLVSGGSAQESQNTGLPDVEL
LSHELKGVCEPPGLDDPLAQDGAGVS

Annexe 30

Cartes de couverture de séquence obtenus en HDX-MS pour A_{2A} rant21 avec un minimum de « produit par acide aminé » de 0,2 (bleu) ou de 0,3 (rouge) réalisé grâce à ¹.

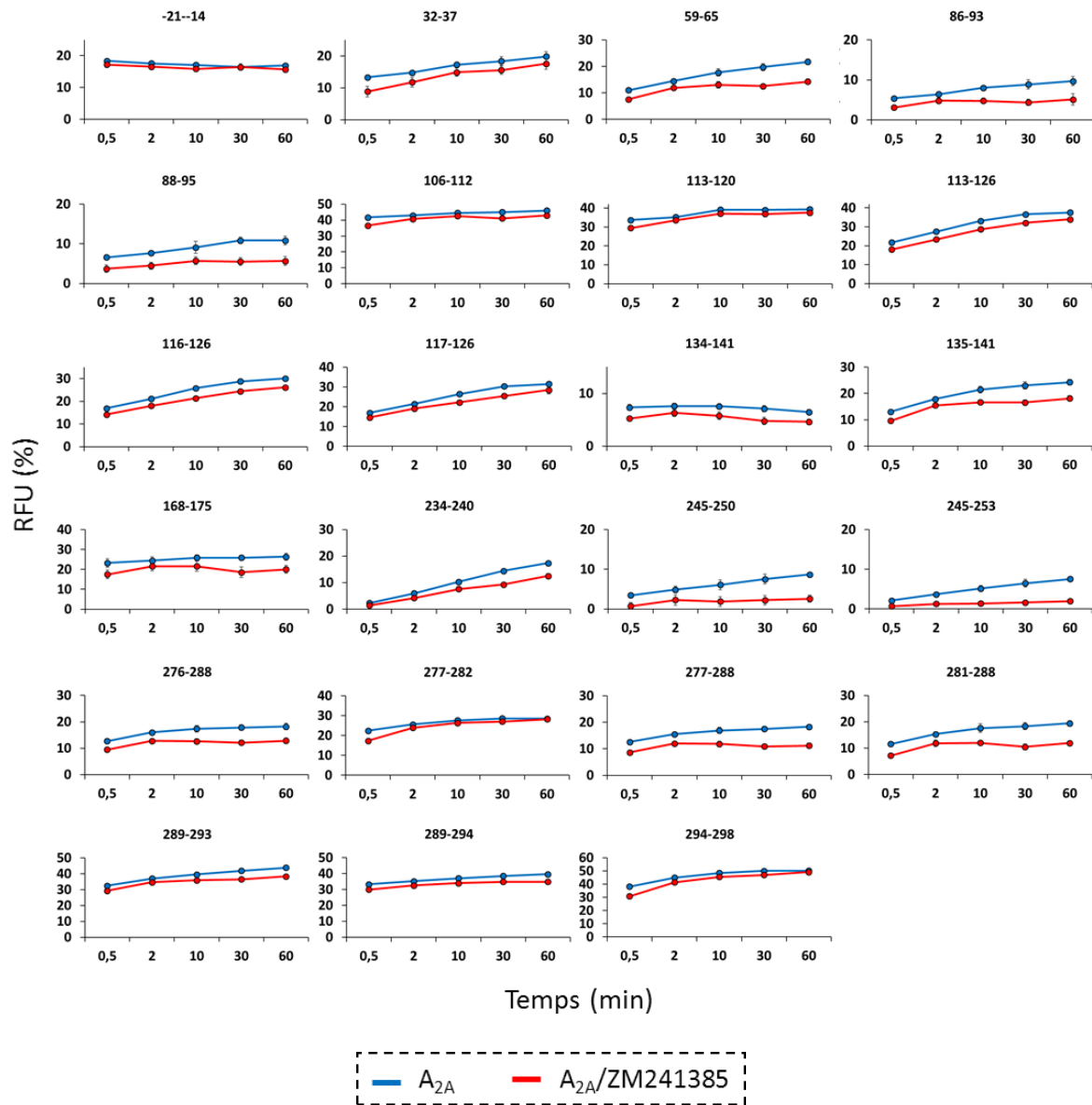


Produit par acide aminé : ■ 0,2 ■ 0,3

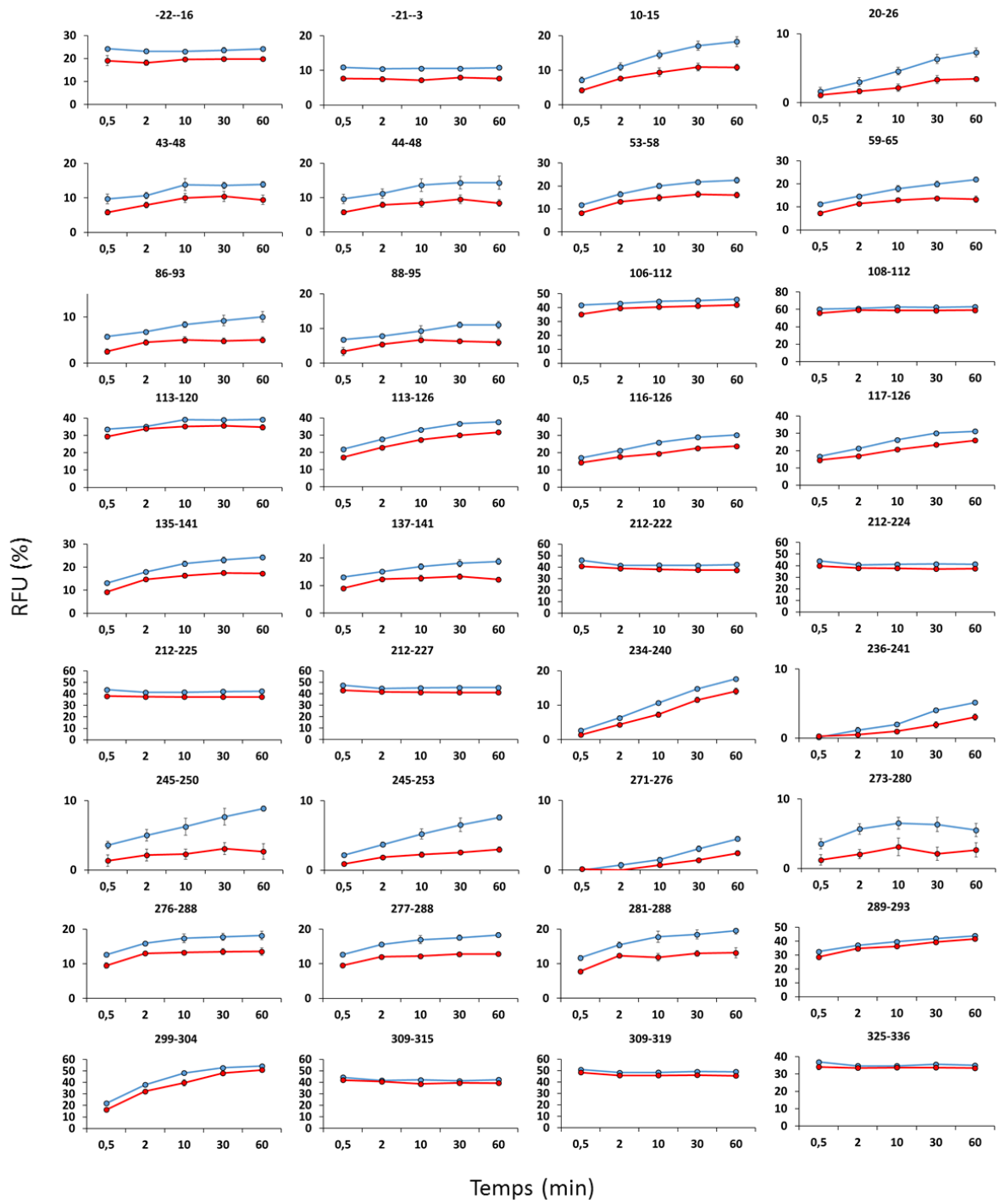
Annexe 31

Incorporation relative en D aux différents temps de deutération de chaque peptide de A_{2A} rant21 statistiquement significatif (MEMHDX avec une p-value à 0,01) obtenus en HDX-MS avec (A) A_{2A} rant21 - A_{2A} rant21/ZM241385 et (B) A_{2A} rant21 - A_{2A} rant21/CGS21680.

A) A_{2A} rant21 - A_{2A} rant21/ZM241385



B) A_{2A} rant21 - A_{2A} rant21/CGS21680



Annexe 32










Banères, J. L.; Botzanowski, T.; Boutin, J. A.; Calamini, B.; Castel, J.; Catoire, L. J.; Cianférani, S.; Demesmay, C.; Ferguson, G.; Ferry, G.; Kniazeff, J.; Krimm, I.; Langer, T.; Lebon, G.; Ley, M.; Nyerges, M.; Schwob, M.; Venien-Bryan, C.; Wagner, R.; Zeder-Lutz, G.; Zilian-Stohrer, C.,

Biophysical Dissection of Isolated GPCRs: The Adenosine A_{2A} Receptor under the

Bistouries. *Receptors* 2023, 2, 47-92.

Review

Biophysical Dissection of Isolated GPCRs: The Adenosine A_{2A} Receptor under the Bistouries

Jean-Louis Banères ¹, Thomas Botzanowski ^{2,†}, Jean A. Boutin ^{3,*} , Barbara Calamini ^{4,‡} , Jérôme Castel ^{2,5}, Laurent J. Catoire ⁶ , Sarah Cianférani ² , Claire Demesmay ⁷, Gavin Ferguson ⁸ , Gilles Ferry ⁹, Julie Kniazeff ¹, Isabelle Krimm ¹⁰ , Thierry Langer ¹¹ , Guillaume Lebon ⁸, Marie Ley ² , Miklos Nyerges ¹² , Magali Schwob ^{5,13}, Catherine Venien-Bryan ¹⁴, Renaud Wagner ¹³, Gabrielle Zeder-Lutz ¹³ and Claudia Zilian-Stohrer ⁶

- ¹ Institut des Biomolécules Max Mousseron (IBMM), UMR 5247, CNRS-Université de Montpellier-ENSCM, 34093 Montpellier, France
 - ² Laboratoire de Spectrométrie de Masse BioOrganique (LSMBO), Institut Pluridisciplinaire Hubert Curien, IPHC, CNRS UMR 7178, Infrastructure Nationale de Protéomique ProFI, FR 2048 CNRS CEA, Université de Strasbourg, 67084 Strasbourg, France
 - ³ Laboratory of Neuroendocrine Endocrine and Germinal Differentiation and Communication (NorDiC), Université de Rouen Normandie, Inserm, UMR 1239, 76000 Rouen, France
 - ⁴ Center for Drug Discovery, Department of Neurobiology, Duke University Medical Center, Durham, NC 27710, USA
 - ⁵ Department of Structural Biology, NovAliX, 67080 Strasbourg, France
 - ⁶ Laboratoire de Biologie Physico-Chimique des Protéines Membranaires, UMR 7099, IBPC, 75005 Paris, France
 - ⁷ Equipe Techsep, Institut des Sciences Analytiques, 69100 Villeurbanne, France
 - ⁸ Institut de Génomique Fonctionnelle, Université de Montpellier, CNRS, INSERM, 34094 Montpellier, France
 - ⁹ Institut de Recherches Servier, 78290 Croissy-sur-Seine, France
 - ¹⁰ Small Molecules for Biological Targets Team, Université Claude Bernard Lyon 1, INSERM 1052, CNRS 5286, Centre Léon Bérard, Centre de Recherche en Cancérologie de Lyon, 69373 Lyon, France
 - ¹¹ Department of Pharmaceutical Sciences, University of Vienna, 1090 Vienna, Austria
 - ¹² Servier Research Institute of Medicinal Chemistry, 1031 Budapest, Hungary
 - ¹³ Plateforme IMPReSs, UMR 7242, Biotechnologie et Signalisation Cellulaire, Ecole Supérieure de Biotechnologie de Strasbourg, 67412 Illkirch, France
 - ¹⁴ Institut de Minéralogie, de Physique des Matériaux et de Cosmochimie, UMR 7590 CNRS, UPMC, IRD, MNHN, 75005 Paris, France
- * Correspondence: ja.boutin.pro@gmail.com
† Present Address: Department of Structural Biology, NovAliX, 67080 Strasbourg, France.
‡ Present Address: Eli Lilly and Company, Boston, MA 02142, USA.



Citation: Banères, J.-L.; Botzanowski, T.; Boutin, J.A.; Calamini, B.; Castel, J.; Catoire, L.J.; Cianférani, S.; Demesmay, C.; Ferguson, G.; Ferry, G.; et al. Biophysical Dissection of Isolated GPCRs: The Adenosine A_{2A} Receptor under the Bistouries.

Receptors **2023**, *2*, 47–92. <https://doi.org/10.3390/receptors2010004>

Academic Editor: Stephen H. Safe

Received: 16 December 2022

Revised: 6 January 2023

Accepted: 11 January 2023

Published: 4 February 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: In an effort to provide an overview of the biophysical approaches used to study G-protein-coupled receptors, we chose to consider the adenosine A_{2A} receptor as a model, as it is widely reported in the literature to explore the way GPCRs are studied nowadays. After a brief introduction of the receptor, we gathered descriptions of the various tools used to investigate the pharmacology and structure of the A_{2A} receptor. We began by describing the key developments which have led to successful studies of GPCRs including the cloning, expression and purification of A_{2A}, and the subsequent characterizations including quality control, binding and functional studies that have been necessary for the further understanding of the receptor. Then, we reviewed the reconstitution of A_{2A} into nanodiscs as well as the use of this biological material in structural mass spectrometry, NMR, calorimetry and various other approaches to gain not only information about the structure and function of A_{2A}, but also the dynamics of the receptor and the tools necessary to pursue such investigations. The body of techniques presented herein are applicable to all GPCRs amenable to purification.

Keywords: adenosine receptor A_{2A}; biophysical approach; NMR; mass spectrometry; molecular pharmacology; expression; purification; reconstitution; ligands; techniques; review

1. Foreword: Biophysical Approaches to Answer Which Questions?

During the last two to three decades, as researchers in the GPCR area, we saw a profound evolution in the study of these membrane-bound proteins. It started with the description of whole genomes and the systematic comparison of sequences that revealed the remarkable conserved topology of this exceptionally large family of proteins. This in turn shed light on the necessity to explore the subtle molecular mechanisms and determinants that were responsible for recognizing such a variety of ligands, to understand how these receptors were selectively activated and how they are involved in so many cellular processes. Initially, the aim was to better understand the pharmacology of these proteins, and thus it became key to be able to study these entities in solution, or in reconstituted and controlled environments. The endpoint is to identify and develop directly interacting molecules that could become novel or “better” drugs. This includes molecules which provide safer pharmacological profiles, improved efficacy, in addition to greater precision, efficiency, and potency. Our scientific community therefore invested intensive efforts to produce and isolate these GPCRs and, hence, to assay and develop integrated analytical techniques to enable the characterization of GPCR structures, functions and dynamics, and further understand the way these receptors are modulated. To make such a characterization, a large diversity of biophysical principles have been exploited, including light, fluorescence, mass and NMR spectroscopies, surface plasmon resonance (SPR), X-ray diffraction and electron microscopy, among others. All of these techniques have been successfully applied to the biophysical dissection of the molecular functioning of GPCRs.

Like many other protein families in similar studies, a small number of prototypical receptors served the development of these GPCR-tailored methodologies, among which is the adenosine A_{2A} receptor (A_{2A}) which we chose to focus on in the present review. This receptor from the Class A GPCR family is indeed of continuing interest as a major therapeutic target in a panel of inflammatory, neurodegenerative and cancer disease types. More importantly, unlike a large proportion of GPCRs that are still recalcitrant to molecular investigations, this receptor rapidly appeared favorably throughout the literature to be produced and isolated in various experimental formats that prove compatible with a remarkable diversity.

The present review summarizes the main biophysical techniques that brought important information on GPCRs through the data obtained on the adenosine A_{2A} receptor. In the first section, we will remind the reader of the main facts known about this receptor. The next sections extensively review the preparative conditions leading to the isolation of the receptor and the many biophysical approaches that contributed to our understanding of the structure and function of GPCRs in general.

2. Background and Introduction

Adenosine is a natural compound formed by an adenine attached to a ribose via a β -N9-glycosidic bond. It occurs widely in nature in the form of many diverse derivatives from multiple sources. All those molecules in which adenosine is a central chemical play important roles in living organisms. It is a part of one of the four building blocks of which the nucleic acids are made and its phosphorylated derivatives are key elements of the energetic system in all living systems. Finally, it exists also under a cyclic form, cAMP, which is a clear element of signaling pathways via the regulation of its key synthesis enzyme, adenylate cyclase. Adenosine is recognized essentially by four GPCRs: A_1 , A_{2A} , A_{2B} and A_3 . A full survey of these receptor characteristics can be found in Fredholm et al. [1]. The human A_{2A} receptor is a 412-amino-acid-long protein encoded by the gene located at chromosome position 22q11.23. The A_{2A} receptor was initially isolated and purified as early as 1974 [2]. Although membrane protein purification under an active form is not trivial, this receptor was widely used because its cloning and expression from human sources were reported early [3]. It was possible to express it in *E. coli* [4] and its characterization began from there.

A complete review of A_{2A} pharmacology exists and is a reference for the matters concerning this receptor [5]. In brief, the A_{2A} expression pattern concerns mainly the brain [6], peripheral T cells [7] and platelets as well as various organs such as the colon, heart and kidneys [5].

The main coupling routes of the receptors are in human platelets, cAMP [8], while coupling to ERK_{1/2} has also been reported on the human cloned receptor [9]. Early works clearly indicated that A_{2A} is not linked to intracellular calcium modifications, at least upon reference to its agonist effects [10]. In 1995, interplays were suspected with other GPCRs, leading to the A_{2A}-dependent reduction in dopamine receptor activation [10]. A very complex picture of the co-regulations of ion channels was drawn later (see, for example, Ikeuchi et al. [11] and Wang and Zhou [12]).

The key step for receptor studies is the binding assay on membranes, whether derived from native tissues or from hosts in which the receptor has been cloned. This is also dependent on the availability of specific radioligands, such as the agonists [³H]-CGS21680 [13] or [³H]-NECA [14], generally preferred over tritiated adenosine [15], and the antagonist [³H]-XAC (xanthine amine congener) [16], with pK_d in the range of 8. New ligands were added, with pK_d in the range of 9, such as [³H]-ZM241385 [17].

A complete survey of the clinical status of A_{2A} receptor ligands as therapeutic drugs can be found in Ijzermann et al. [5]. More than 50 different drug trials were still active in 2022 for targeting A_{2A} as treatments for cancer (Imaradenant), pain (Spongosine), Parkinson's disease (BIIB014, Taminadenant) and sickle cell disease (Regadenoson), for which the infusion of the A_{2A} agonist reduces the production of IFN-γ and enhances the production of IL-13 and CD39 [18].

3. Mass Production of A_{2A} Receptors

3.1. Expression Systems

Owing to their finely regulated functions, GPCRs are naturally scarce in native biological sources. Accordingly, robust recombinant expression systems are needed to produce the milligram amounts required for their biophysical analyses. These systems should not only fit yield criteria but should also offer the most adapted environment for producing properly folded and functional GPCRs. Several systems have been developed with the aim of producing large amounts of the receptor properly folded with various degrees of sophistication and success rates. These include cell-free approaches [19,20], bacterial [21] or yeast [22–24] microorganisms, insect and animal cells using viral-infection or DNA-transfection strategies [25], photoreceptor cells of transgenic animals [26,27] and even whole organisms such as silkworm (*Bombyx mori*) larvae [28].

Among the many GPCRs investigated in these studies, A_{2A} is remarkably one of the very few that has been successfully expressed at relatively high levels in nearly all recombinant systems. It was indeed produced in milligram amounts in *E. coli*, either unfolded in inclusion bodies [29], or as ligand-binding active receptors in the bacterial membranes [30–32]. A_{2A} was also efficiently produced in the membranes of yeast cells, mainly *Saccharomyces cerevisiae* [33–35] and *Pichia pastoris* [22,36–39]. These hosts combine the advantages of handling the simplicity of microorganisms with the complex cellular machineries of eukaryotes. *S. cerevisiae* notably offers a large panel of molecular tools and engineering possibilities that were exploited for functional studies of many GPCRs [40] as well as for extensive screenings of stabilizing mutations and preparative conditions, including for A_{2A} [41–46]. Concerning *P. pastoris*, its peculiar methylotrophic metabolism, the tightly regulatable expression conditions and the exceptional cell densities it can afford, make this yeast well suited for expressing membrane proteins for structural studies. Accordingly, it has not only allowed two of the high-resolution structures of A_{2A} to be solved by X-ray crystallography [38,47], but it has also delivered various isotopically labelled samples for most of the NMR-based studies conducted on A_{2A} [47–50].

Though technically more demanding and more time-consuming than microbes, animal cell systems offer a closer native-like environment and have therefore been extensively

used as recombinant hosts for A_{2A} production. Most notably, recombinant baculovirus infecting either *Trichoplusia ni* (Tni) or *Spodoptera frugiperda* (Sf9) insect cells became the reference systems for generating crystallography-grade GPCR samples over the years. For A_{2A} in particular, the majority of high resolution structures referenced to date were obtained from various forms of the receptor produced with this insect cell expression system [51–55]. Concerning mammalian cells, low expression levels and cost issues initially precluded the direct use of the high-quality receptors they produce for further biophysical analyses. Until recently, CHO or HEK 293 cell lines were thus mainly used as orthogonal recombinant systems for evaluating and validating the pharmacological properties of engineered or mutated A_{2A} receptors that were designed for structural studies with other systems [39,51,56–59]. With the development of HEK 293-derived cell lines adapted to suspension and high cell density culturing in optimized media, combined with their transient transfection with tetracycline-inducible plasmids, GPCRs produced in these quasi-homologous host cells, including A_{2A}, have now become accessible to thorough biophysical and structural studies [60,61].

3.2. Molecular Constructs

The design of A_{2A} sequences that were expressed in these different systems are representative of the strategies classically employed for the large-scale production and purification of membrane proteins. The A_{2A} coding sequence, sometimes optimized to fit the codon preference of the host cell [38], has been N- or C-terminally fused to various tags (most commonly His6, His8, His10 and FLAG) to enable detection and purification schemes, or to reporter proteins (GFP, RFP) [33,41,62,63] to monitor expression and purification optimization. These supplementary sequences were often separated from the A_{2A} coding sequence by a cleavage sequence (TEV or 3C protease sequence) [22,62,64,65]. Finally, signal peptide sequences adapted to the expression host (from influenza hemagglutinin HA in animal cells, or from α -Factor in yeast systems) were almost always added upstream of the fusion constructs to improve protein translocation in the ER and target the plasma membrane [22,33,51,66,67].

More specific modifications were further implemented in the A_{2A} sequence when homogeneity and stability issues had to be addressed, notably for crystallography or cryo-electron microscopy (cryo-EM) structural studies. As further described in Section 8 below, besides a systematic truncation of the last C-terminal amino acids 317 to 412 [38,51,52], two complementary or combined sequence engineering strategies were employed to stabilize A_{2A}. The first one consists of replacing flexible domains of the receptor with more structured and stable soluble proteins, the most common being the phage T4 lysozyme (T4L) [51,58,68] and the apo-cytochrome b562RIL (BRIL) [54,56,61,65,69,70]; however, other less-used proteins such as flavodoxin, xylanase and rubredoxin [71], as well as custom-made de-novo-designed α -helical proteins have been successfully employed [43]. The second approach relied on the introduction of thermostabilizing point mutations, identified via directed-evolution [31,45] or structure-guided [55] approaches, allowing crystallographic studies [52,72] and other biochemical analyses [42,59,62]. An additional point mutation aiming at suppressing a glycosylation site at position N154 to improve receptor homogeneity was often introduced [38,47,52,72]. However, similar constructs bearing no such mutation also led to homogenous and crystallography grade samples [51,56,68]. Finally, for some specific NMR studies, several amino acids exposed at the intracellular surface were replaced by cysteine residues for targeted labeling with ¹⁹F chemical probes [48,73–75].

4. Extraction and Purification

When the expression conditions are met, the next challenge before biophysical investigations of isolated membrane proteins consists of extracting and purifying the expected amounts of pure, homogenous, and correctly folded receptors. As for a large majority of GPCRs and other membrane proteins studied in solution [25,76], A_{2A} was mainly solubilized using the mild detergent n-dodecyl- β -D-maltopyranoside (DDM), independently

of the expression system used [30,38,62,63,67,77], or with the more recent lauryl maltose neopentyl glycol (LMNG) [47,48,50,61]. The physicochemical properties of these two surfactants are particularly suited for destabilizing biological membranes, while providing the receptors a reasonably stable hydrophobic environment to maintain their folding and function throughout their preparation and the planned analyses [78].

In the case of thermostabilized A_{2A} sequences, the higher stability conferred by the selected point mutations allowed the solubilization of the receptors with n-decyl-β-D-maltopyranoside (DM), a harsher maltoside-based detergent. Compared to DDM, DM forms smaller micelles and has a higher critical micelle concentration, which was better adapted to further detergent exchanges and crystallographic studies using vapor diffusion crystallization conditions [52,69,72,79,80].

For the non-thermostabilized A_{2A} constructs, the stability issues during solubilization and purification were addressed via the addition of specific ligands, locking the receptors into a limited number of conformations. Theophylline was the most frequently used [30,38,47–51,55,56,58,71,79–82], even if other ligands such as caffeine [69], DPCPX [77] and ZM241385 [61,63] appeared to also be useful. The stabilizing effect of these ligands [56,83–85] was very often combined with the supplementation of cholesteryl hemisuccinate (CHS), a cholesterol derivative that modulates the shape of detergent micelles in a pseudo-bicelle architecture. The addition of CHS into the purification buffers has been known to be beneficial for improving the stability of numerous solubilized GPCRs [86], including A_{2A} [30,56,63,87]. Finally, other agents commonly applied to minimize the aggregation of membrane proteins in solution were also often added, such as glycerol which likely acts as an amphiphilic interface between hydrophobic domains and the polar solvent [88], and NaCl which contributes to maintaining proteins soluble in aqueous solutions through a “salting in” effect. Incidentally, sodium ions are also considered an allosteric modulator of A_{2A} [89]. Iodoacetamide can also be used to prevent protein aggregation by intermolecular disulfide-bond formation.

The strategies used for purifying detergent-solubilized A_{2A} were essentially based on affinity chromatography (AC) approaches. Since nearly all recombinant receptors were fused to a poly-histidine tag, an immobilized-metal affinity chromatography was most often performed as a first step, either on Co²⁺ [50,51,63,67,90,91] or Ni²⁺-grafted resins [30,34,52,77]. Some procedures also involved agarose resins covalently bound to XAC, a non-specific, high A_{2A} affinity antagonist at adenosine receptors [92], which were implemented as ligand-affinity columns in a first [30] or second purification step [49,50,93,94]. Recently, AC procedures using anti-tag antibodies coupled to agarose resins, such as M1 anti-FLAG [61], or anti-Rho-1D4 [59] antibodies, were also reported.

Purification of A_{2A} has often been finalized by a polishing step on size exclusion chromatography (SEC) Superdex 200 columns, allowing the recovery of homogenous populations of monomeric receptors, and the elimination of additional impurities and specific compounds eluted from the previous AC step(s).

5. Isolation in Lipid Nanoparticles

As detailed in the previous section, a variety of conditions has allowed the isolation of A_{2A} in detergent, in a sufficiently stable and convenient environment for a panel of biochemical and biophysical studies. However, the hydrophobic environment brought by detergents is chemically and physically distinct from the complexity of biological membranes, which poses certain limitations and issues for a number of investigations. Examples of such issues include the interference of detergents with certain analytical methods, insufficient stability for prolonged periods of analyses, absence of critical lipid and/or protein interactants for a receptor’s structure and function [95]. To overcome these difficulties, several methods have been developed over the years to reinsert detergent-isolated GPCRs in a more stable and more native lipid environment, including lipid-detergent bicelles, lipid bilayer nanoparticles or unilamellar lipid vesicles (ULVs) [95–97].

Bicelles are discoidal particles made of a planar phospholipid bilayer surrounded by a belt of detergent. While the reconstitution of several purified GPCRs in bicelles proved useful for their structural and functional investigations [98–100], none of these approaches have been reported so far for the characterization of A_{2A}.

Concerning the implementation of ULVs, or liposomes, the situation is just slightly better documented. In an NMR study on A_{2A} purified in detergent, liposome reconstitution was used as a quality control material to evaluate the functionality of the highly deuterated receptor for its capacity to activate its heterotrimeric G protein [67]. In another study, an original liposome synthesis method called INSYRT (in situ lipid synthesis for protein reconstitution technology) was recently developed and validated with A_{2A} [101]. The receptor was solubilized and purified with DDM thioester analogs that served as precursors for the synthesis of phospholipids via a native chemical ligation (NCL) reaction. The resulting proteoliposomes were further characterized to contain A_{2A}, showing ligand-binding pharmacological properties similar to those described in cellular membranes, and to the receptor reconstituted in the so-called nanodiscs (NDs).

Nanodiscs are nanometric discoidal lipid bilayers stabilized by two copies of a membrane scaffold proteins (MSP) deriving from the Apo-A1 lipoprotein [102]. Since it has been demonstrated that likely every detergent-solubilized membrane protein can self-assemble in such particles and maintain its structure and activity in a stable environment [103], NDs have been the prevailing source of material for the biophysical investigation of numerous GPCRs reconstituted in a lipid bilayer in recent times [104]. This is particularly the case for A_{2A} for which the use of ND particles generated led to the discovery of a number of findings on receptor–ligand interactions and the structural dynamics of the receptor. The different approaches used included surface plasmon resonance (SPR) [62,105], mass spectrometry (MS) [106], miniaturized weak affinity chromatography (nanoWAC) [77], single molecule fluorescence microscopy [75] and ligand- or receptor-oriented NMR techniques [49,50,74,75,107]. In these studies, detergent-purified A_{2A} was successfully assembled with different versions of the engineered MSP, either MSP1ΔH5 [50,74], MSP1D1 [62,80,106,107], MSP1E3D1 [49,77] or the less widespread Apolipoprotein A1 from zebrafish, Zap1 [105], which essentially differ in their molecular size and in the disk size of the generated NDs (from approximately 9 to 12 nm in diameter [103]). Some of these MSP were genetically [62] or chemically [77] modified to allow the grafting of the resulting NDs via their MSP on different supports for SPR and WAC analyses. The lipid composition essentially consisted of POPC:POPG mixtures incubated with MSP and A_{2A} at various molecular ratios, in some cases replaced by or supplemented with other components used to evaluate their effect on the ND assembly efficiency [80] and/or on the receptor function [49,50,74].

Amphiphilic polymers, mainly amphipols, have also been largely used for isolating GPCRs in a membrane-mimicking environment [108,109]. While these particles have proven to be highly stable and allowed many investigations in detergent-free solutions, they are essentially generated after the reconstitution of proteins already purified in detergents. Thus, they may lack some lipids important for the conformational dynamic and function of the receptors. To date, no publication has reported on A_{2A} reconstitution in amphipols.

Alternative amphiphilic co-polymers made of styrene and maleic acid (SMAs) are being increasingly used for the characterization of isolated membrane proteins, notably GPCRs, in lipid bilayers [97,110]. Unlike the standard amphipols, which are acrylate-based polymers grafted with octylamine, SMA compounds hold the ability, under defined conditions, to destabilize biological membranes and spontaneously form SMA lipid disc particles (SMALPs). Those containing the protein of interest can then be directly purified. A_{2A} has been successfully isolated in SMALPs generated from the membranes of *Pichia pastoris* and HEK recombinant cells, and was characterized for its ligand-binding capacity and its increased stability compared to the receptor isolated in detergent [39]. Further studies reported by the same group investigated A_{2A} SMALPs from *Pichia pastoris* membranes to

evaluate ligand-binding events via fluorescence correlation spectroscopy [111], as well as some ligand-induced conformational changes in A_{2A} via fluorescence spectroscopy [112].

6. Quality Controls on Isolated Receptors

Once the protein material is isolated and purified, a panel of biochemical and/or biophysical tests are generally performed to assess the integrity of the receptor including the purity, homogeneity, stability and the activity of the isolated receptors. These are all important or even critical criteria to be met before initiating the functional and structural characterizations that are described in sections below.

Receptor integrity and purity are almost always assessed via classical SDS-PAGE analyses, sometimes complemented with A_{2A}-specific immunodetection. The evaluation of sample homogeneity is often necessary to identify the presence of potential residual contaminant, but also to appreciate the oligomerization states of the receptor. This is essentially assessed through standard size-exclusion chromatography (SEC) analyses [35,51,71,94,106], or by discriminating the particle size either in the final sample via dynamic light scattering (DLS) [80] or during its purification via SEC-MALS [35,43,91,94]. Negative-staining EM imaging may also provide a useful qualitative complementary approach [61,77,91].

Native mass spectrometry (native MS) is often used as a first-line technique after the protein purification of soluble proteins to address oligomerization issues. This technique aims at transferring and detecting intact non-covalent interactions in the gas phase of a mass spectrometer. While native MS has been established for the analysis of soluble proteins and protein complexes, it is however challenging for membrane protein analyses [113,114], but has been successfully applied to A_{2A}. Figure 1, panel B, represents a native mass spectrum of A_{2A} solubilized in DDM, highlighting several species corresponding to different populations of receptors, including full-length, degraded or modified forms.

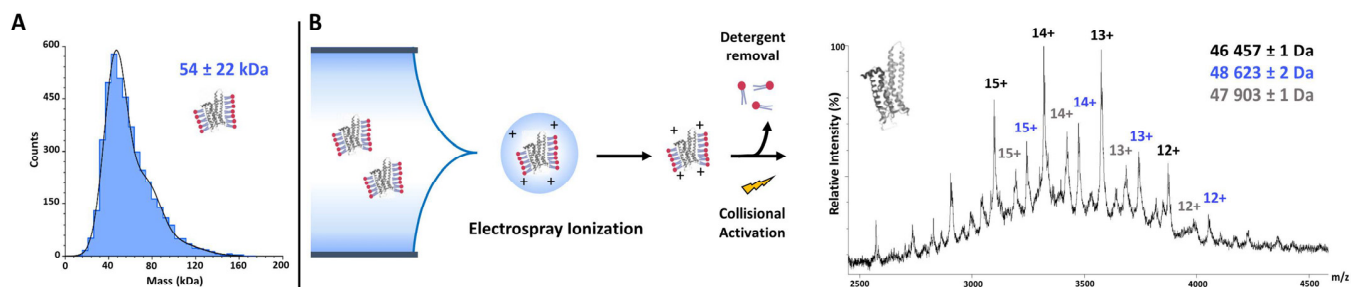


Figure 1. A_{2A} mass measurement using mass photometry and native MS. Mass Photometry (A) is an emerging technique for fast mass and oligomeric state determination of proteins. Analysis of A_{2A} showed a good homogeneity and enabled the determination of a mass of 54 ± 22 kDa (molecular mass distribution histogram of A_{2A}—the solid blue line represents major species fit with a Gaussian function). The quite large mass distribution obtained can be explained by the accuracy of the technique itself but also by the presence of potentially remaining micelles around the protein which may average out the overall mass obtained and thus increase the mass deviation. Native MS (B) is a powerful analytical tool for the study of membrane proteins embedded in detergent. The technique, preserving non-covalent interactions in the gas phase for mass analysis, enabled the visualization of the A_{2A} protein with nano electrospray. Several populations (48.6 kDa, the expected mass of A_{2A}; 46.4 and 47.9 kDa, truncated forms of the protein) have been identified after optimization of the instrumental parameters. These finely tuned and harsh conditions are essential for efficient detergent removal.

The implementation of mass photometry, a very recent label-free technique that quantifies molecular weights of biomolecules at the single-molecule level by interferometric detection of scattered light [115], provides an additional way for the characterization of membrane protein sample homogeneity [116]. In this context, we evaluated the technique for the analyses of A_{2A} samples (Figure 1, panel A). The obtained results highlight the

complementarity of mass photometry and native MS for the characterization of membrane proteins. Requiring low amounts of starting material (approximately 100 ng protein) and no additional buffer exchange steps during the preparation, mass photometry can be used to obtain information on the sample studied with less optimization than native MS, providing heterogeneity and oligomerization information very quickly but with considerably less accuracy.

As mentioned earlier, the stability of the isolated protein is often critical for a number of analyses, in particular for crystallography studies or when the analytical methods require extended periods of measurement. For A_{2A} , this issue has been mostly addressed via methods assessing the thermal stability of the isolated protein. Those were mainly fluorescence-based techniques, either involving the thiol-reactive chemical probe N-[4-(7-diethylamino-4-methyl-3-coumarinyl)phenyl]maleimide that covalently binds to internal cysteines and fluoresces upon thermal denaturation [55,56,71,83], consisting of a fluorescence size-exclusion chromatography (FSEC) evaluation that requires the presence of a fluorescent protein (i.e., GFP) fused to the receptor [62], or label-free tryptophan fluorescence monitoring upon unfolding [117]. A_{2A} thermal stability was also assessed via the analyses of migration profiles of the protein on acrylamide gel after denaturation, either revealed via Western blotting [84] or via clear-native PAGE [63]. As an alternative to thermal denaturation, an isothermal chemical denaturation method using guanidium chloride has been successfully applied to assess the stability of A_{2A} when isolated in the presence of various ligands [90].

Perhaps the most important indication of the quality of a protein sample relies on its capacity to fulfill its function. A variety of techniques have been applied to assess the activity of A_{2A} via its ability either to bind specific ligands or to activate its associated heterotrimeric G-protein partner upon agonist stimulation. These methods include radioligand binding assays in filtration [36,77,91] or scintillation proximity assay (SPA) formats [62,105], GTP γ S activity assays [67], fluorescence analyses [111], and a panel of SPR-, MS- or NMR-based approaches, which are described in more detail in the following sections.

7. Functionality of the Isolated A_{2A}

7.1. Ligand Binding

7.1.1. Radioactive Ligand Binding

As far as specific radiolabeled compounds are available, for decades, this direct evaluation of receptor–ligand interactions has been the preferred technique to study and define the pharmacology of GPCRs. It is a highly sensitive approach based on the counting of the radiation emitted by the radioactive compounds bound to their receptors, from which are deduced precise affinity parameters either in saturation or competition modes. In the most classical methods, receptor-bound and free radioligands are separated via filter-based techniques that are particularly well suited to evaluating GPCRs in membrane fraction samples. When GPCRs are solubilized and isolated, standard filter-based assays are generally not compatible with the small size of the particles to be analyzed, so alternative separation methods were needed. For A_{2A} , these mainly relied on the use of mini-spin gel filtration columns [30–32,36,39,49,118,119], on the filter-based separation of the receptor–ligand complexes either precipitated with polyethylene glycol (PEG) [77,91] or immobilized on agarose beads [34,83], or by immobilization on functionalized scintillating bead supports designed for SPA [47,62,77,79,105]. These studies were essentially conducted to characterize and validate the pharmacology of A_{2A} isolated either in detergent micelles [30,31,36,39,49,91], in NDs [49,62,77,105], in SMALPs [39] or in liposomes [101]. Interestingly, one study conducted on different A_{2A} samples generated from *P. pastoris* or HEK cells, either in membrane fractions or isolated in detergent or in SMA particles [39], determined comparable affinity values for several agonist and antagonist ligands. In addition, this work demonstrated the increased stability overtime of SMALP or ND samples over detergent micelles on the ligand-binding activity of the isolated receptor [39].

Radioactive ligand binding has also been the key method for evaluating the thermal stability of solubilized A_{2A}, either for screening thermostabilizing mutations [31,52] or for evaluating the stabilizing effect of various detergents and additives [119].

7.1.2. Surface Plasmon Resonance

Several alternative biophysical methods, such as SPR, have emerged to determine binding affinities and kinetics, as they are essential for drug discovery and development [120]. Indeed, the kinetics of drug–receptor complex formation, and in particular the residence time of the ligand [121], is a crucial element that affects the functional properties of a ligand and its pharmacological profile [122–124].

SPR is a label-free technique that allows the quantitative analysis of molecular binding. One binding partner is immobilized on a biosensor surface and the other is introduced into a microfluidic system and flows in solution over the surface. The binding causes a small change in the refractive index at the biosensor surface due to the accumulated adsorbed mass. Continuous monitoring of this signal allows the determination of the kinetic parameters such as the association rate constant (k_{on}), the dissociation rate constant (k_{off}) and the affinity (K_d) of the interaction [125].

SPR has become the method of choice for studying ligand binding to various GPCRs [126–129]. This technique requires small amounts of membrane proteins in a native or similar environment. The challenges are to maintain high levels of binding activity of the immobilized receptor and to detect the binding of low-molecular-weight ligands, as this technology is based on mass changes [130–132]. This has been successfully applied to assess the ligand-binding affinities of a number of GPCRs. A variety of A_{2A}/ligand interactions have been studied using SPR over the past decades. Findings from these studies illustrates how SPR technology supports structure-based ligand discovery and fragment screening and contributes to the better understanding of the molecular aspects of signaling. Detergent-solubilized A_{2A}, either non modified or thermostabilized (A_{2A}-StaR), has been widely investigated and generally extracted from insect cells membranes. Several A_{2A}-StaR (purified or as crude solubilized extracts) were captured through their C-terminal 10His tag to a Ni-loaded NTA sensor surface (carboxymethylated dextran functionalized with nitrilotriacetic acid) [90,131,133–135]. Small antagonists (molecular weight ranging from 285 to 345 Da) were shown to interact with a K_d ranging from 0.3 to 160 nM in accordance with the classical competition radioligand binding assay [133]. Importantly, from the screening point of view, SPR is able to detect even weaker GPCR/compound complexes ($K_d > 1 \mu\text{M}$) [131]. The groundwork was laid for the cross-screening of a panel of antagonists with mutated A_{2A}-StaR. The resulting matrix of binding kinetics data was then combined with molecular modelling and docking data from a refined map of the A_{2A} binding site, thereby facilitating the structure-based drug design [135,136]. Affinity and kinetic rate constants measured at different temperatures allow the calculation of enthalpic and entropic energy components as well as the analysis of kinetics of the transition state, constituting the thermodynamic signature of a target-ligand complex [137–139]. Calculating the kinetic transition state components for six ligand–A_{2A}–StaR complexes and combining them with supervised molecular dynamics (MD) and metadynamics approaches clarified the crucial role of receptor and antagonist solvation/desolvation in the transition states of entropic nature. Thermodynamics combined with kinetic analyses allows a better understanding of the details of the molecular interactions. They could also help rationalize the drug design since kinetics better correlates with EC₅₀ than thermodynamic equilibrium affinity constant values for some GPCR–ligand complexes [138].

Solubilized GPCR micelles are unstable and tend to denature and/or aggregate over time. As detailed previously in Section 5, several strategies have been developed to reconstitute GPCRs in a lipid bilayer, such as proteoliposomes or high-density lipoproteins known as NDs, to provide more stable conditions for the receptor compared to detergents. The kinetic characterization, using SPR, of the binding of nine adenosine antagonists to a

A_{2A}-StaR reconstituted in NDs correlates well with those of the same receptor in a membrane environment obtained from the conventional radioligand binding assay. NDs likely correspond to a more native environment than detergent and may avoid the problems of ligand partitioning in micelles [105]. Bocquet et al. conducted different binding assays (SPA and SPR) with the A_{2A}-StaR in membranes, in NDs and in detergent micelles [62]. Different set-ups were evaluated in the SPR assay, with NDs being immobilized via different tags fused either on the receptor or on the MSP of the NDs, without impacting the binding kinetics of small antagonists (337 to 428 Da). In addition, the receptor was shown to be more stable in NDs than in detergent micelles.

The SPR approach has also been validated for fragment screening investigations on A_{2A} (see Section 9). The success of such experiments requires a high density of immobilized GPCRs with high binding capacity, as well as careful experimental design, to reduce false positive results since high concentrations of fragments are tested and tend to non-specifically bind to many targets (reviewed in Shepherd et al. [132] and Coyle and Walser [140]). This approach was first developed for a solubilized A_{2A}-StaR with a panel of 70 fragments derived from xanthine, with molecular weights ranging from 136 to 194 Da. [134]. SPR was also used to validate putative ligands of a A_{2A}-StaR issued from a mass spectrometry fragment screening [141]. In another study, the native, non-stabilized A_{2A} receptor, as well as three other receptor subtypes from the same family (A₁, A_{2B} and A₃), were investigated via SPR. [142].

As GPCRs interact with different intracellular partners of the downstream signaling cascade, SPR was also used to study some of these molecular interactions. It was notably shown that the binding of solubilized A_{2A} to immobilized G α subunit proteins was affected by GDP analogues and that the C-terminus of A_{2A} was necessary for its high-affinity association to this G-protein subunit [143]. In another study on A_{2A} in NDs bound to a full or a partial agonist, the binding characteristics for the heterotrimeric G $\alpha\beta\gamma$ proteins were determined using SPR [50]. No differences in affinity and kinetics were observed, but complementary structural analyses using NMR suggested a model based on conformational changes in the receptor to explain their different functional output. Other aspects of A_{2A} functioning have been explored via SPR, such as the impact of mutations in a cholesterol consensus motif on G-protein binding [59].

7.1.3. Calorimetry

Isothermal titration calorimetry (ITC) is a label-free biophysical technique that can measure kinetics and the thermodynamic parameters of binding, such as affinity, stoichiometry, enthalpy and entropy between macromolecules (e.g., between two proteins) or between a macromolecule and its small molecule ligands [144,145]. Applied to GPCRs, ITC allows for a deeper understanding of the binding interactions. Indeed, it reflects the subtle structural and dynamic changes between bound and unbound forms, identifying the most important regions of the binding interface [146,147] as well as receptor–ligand specificity, cell surface expression, endocytosis and recycling [148]. A_{2A} is of particular interest because it possesses a long and flexible C-terminal region which favors interactions with several proteins, including G-protein-coupled receptor kinases, β -arrestins, α -actinins, calmodulin (CaM), ubiquitin-specific protease 4 (USP4), and neuronal calcium sensor protein 1 (NCS-1) [149–156]. The affinity and stoichiometry of A_{2A} binding to two of its partners, CaM and α -actinin 1, have been recently studied together with their regulation by calcium using ITC [157,158]. In this study, human CaM and two A_{2A} C-terminal regions of different lengths (A_{2A}-ctL comprising 293–412 amino acids and A_{2A}-ctS comprising 321–412 amino acids). CaM was titrated into the protein in the presence of calcium with or without the addition of EDTA. In the absence of EDTA, the binding between A_{2A} and CaM displayed strong affinity ($K_d = 97.9$ nM) with a 1:1 stoichiometry. In the presence of calcium, the binding of CaM to A_{2A}-ctL led to an exothermic reaction with a favorable enthalpy, but unfavorable entropy, possibly indicating conformational changes. No binding occurs in the presence of the calcium chelator, EDTA. The binding site of CaM on the A_{2A}-ctL

region was further identified using the A_{2A}-ctS construct lacking the amino acids 293 to 320. CaM binding in the presence of calcium led to small exothermic peaks signalling a loss of affinity. These results, confirmed by NMR spectroscopy, identified the residues 293–310 as the binding epitope of CaM onto A_{2A} [158].

Similarly, mutating arginine residues in this region abolished CaM binding. These data are in agreement with previously published results [151]. α -Actinin binding to A_{2A} was suspected to be implicated in the internalization of the receptor after agonist activation. To characterize the domain responsible for the interaction between the C-terminal domains of A_{2A} and α -actinin 1, α -actinin1 wild-type constructs limited to the calcium binding domain (CABD), the rod domain or the CABD-rod domains were used. In the absence of calcium, the titration of the rod-CABD domains led to exothermic peaks and to a saturable binding curve. The affinity of rod-CABD domains to α -actinin was found to be in the micromolar range ($K_d = 39 \mu\text{M}$), which was weaker than for A_{2A}-ctL with CaM. Together, the results of the ITC experiments suggest that the A_{2A} C-terminus binds to the distal part of α -actinin 1 C-terminal domain, either in the CABD or at the subsequent C-terminus, a region shown to interact with CaM, and in fact it has been demonstrated that these proteins compete when binding to A_{2A}. It is worth highlighting that ITC is a classical approach to identify GPCR partners and should be systematically used to validate and characterize receptor partners at the molecular level.

7.1.4. Receptor Fluorescent Labeling

Many different techniques have been applied to illuminate the molecular bases of trans-membrane signaling through GPCRs in general, and through A_{2A}, in particular. Among them, fluorescence-based studies appear to be particularly adapted to investigate the functioning of these receptors, as they can be used in many different environments ranging from native and recombinant cell lines to isolated receptors in membrane-mimicking systems [159]. These techniques nevertheless require GPCRs to be specifically labelled with adapted fluorophores. For the last two decades, several approaches have been developed to label receptors for fluorescence-based applications. Besides the fluorescent ligands described later (Section 9), these include protein and peptide tags compatible with measurements in live cells (e.g., GFP, YFP, SnapTag and FAsH-tetracysteine tag) [160]. Several of these labelling strategies have been applied to the A_{2A}. For instance, the concomitant insertion of a Cyan Fluorescent Protein and the FAsH-tetracysteine tag into human A_{2A} was combined with fluorescence resonance energy transfer (FRET) measurements to provide a description of the activation process of this receptor in cellular model systems [161]. More recently, a covalent ligand was used to couple a fluorophore to A_{2A} [162]. Although this latter approach was used in a cell-based system, it could be extended to isolated receptors as well. Besides cell-based studies, strategies to label purified receptors in detergent micelles or nanodiscs have also been developed. In many cases, the insertion of the fluorescent probe primarily involves the labelling of unique reactive cysteines with a fluorescent probe through a maleimide-based reaction. This strategy was applied to recombinant A_{2A} to successfully introduce probes for NMR- and fluorescence-based applications [75]. For the latter, cyanine 3 was attached to the purified receptor expressed in *P. pastoris* and assembled into nanodiscs [163]. This opened the way to single-molecule fluorescence experiments aimed at describing the conformational transitions of A_{2A}. In the same way, a purified mutant receptor containing single reactive cysteines was labeled with an environment-sensitive dye, BODIPY-FL, for fluorescence quenching experiments aimed at delineating the dynamics of the TM6 domain of A_{2A} [164]. In addition, orthogonal labeling strategies were also described in this same study [164]. Specifically, a double-cysteine mutant with reactive Cys residues on TM4 and TM6 was used for double-labeling with a fluorescence donor (Alexa Fluor 488) and acceptor (Alexa Fluor 647), paving the way for single molecule-FRET analyses of receptor conformational fluctuations. In addition to these classical protein engineering approaches, alternative labeling strategies have been described for A_{2A}. In particular, a modified aminoacyl-tRNA-synthetase and a suppressor tRNA were used in

a cell-free system to incorporate an unnatural amino acid that was further coupled to a fluorescent dye [165]. Together, these studies provide evidence that the isolated adenosine A_{2A} receptor in different membrane-mimicking environments can be efficiently labeled with fluorophores, providing a convenient system to describe its energy ensemble and the modulations by the environment using fluorescence-based techniques.

7.1.5. Ligand-Based NMR

Ligand-observed NMR experiments are widely used to measure and analyze small molecule binding to soluble and membrane proteins such as GPCRs [166]. No protein labeling is required, and experiments can be performed in micelles, membranes and nanodiscs. Ligand-observed NMR experiments are suitable for moderate- to low-affinity ligands (10 nM to mM affinities). This is of particular interest as far as screening approaches are concerned, as the net results of most of the screening campaigns are poor affinity binders. Nevertheless, the binding of higher-affinity ligands can be monitored through competition experiments, using the NMR binding signal of a moderate-affinity compound as a reference signal. Saturation transfer difference (STD), nuclear Overhauser effect spectroscopy (NOESY) experiments or ¹⁹F NMR on labeled ligands have been reported on A_{2A} [74,91,107]). Experiments are typically compared to a control experiment where the receptor is either missing or blocked by a high-affinity ligand. The NMR experiments allow the identification of novel ligands and of their binding pockets when a reference compound at a known binding site is available. Furthermore, they may reveal other binding pockets on A_{2A}. For example, the binding of adenosine in an allosteric pocket was observed through STD experiments recorded with A_{2A} bound to ZM241358 [91].

Three-dimensional structures of A_{2A}-ligand complexes can be calculated from NOESY experiments called INPHARMA (Interligand Noes for PHarmacophore Mapping) [107,167]. For example, the structure of the complex between A_{2A} and 3-pyrrolidin-1-ylquinoxalin-2-amine (PQA) was obtained from the interligand NOESY observed between PQA and the antagonist compound ZM241358 [107].

Finally, ligand-observed NMR can be used to investigate allosteric mechanisms, as positive and negative allosteric modulators (PAMs and NAMs), respectively, increasing or decreasing the binding of a ligand and thus its NMR binding signal. Recently, Huang et al. [74] investigated the binding of cholesterol on A_{2A} using ¹⁹F-labeled cholesterol analogs. They observed the broadening of the NMR resonances as well as a 0.5 ppm up-field shift for the ¹⁹F-labeled cholesterol bound to A_{2A} in nanodiscs compared to empty nanodiscs. However, the NMR spectrum was not modified upon the addition of an inverse agonist, full agonist, or a G-protein mimetic, suggesting that the cholesterol interactions with A_{2A} are transient and unspecific, and that the effect of cholesterol is due to changes in the membrane bilayer properties [74]. The investigation of the allosteric modulation of A_{2A} by small molecules using ligand-observed NMR represents a rapid and robust approach that should be further used to explore the biology of A_{2A}.

7.1.6. Native Mass Spectrometry

As described above (see Section 6), native MS is often used to check the quality of a purified protein and to address oligomerization issues. In addition, native MS is also a valuable method to evaluate molecular interactions between the protein and various partners. Yen et al. [168] used native MS to reveal that A_{2A} in its apo state still binds to endogenous phosphatidylserines (PS) and phosphatidylinositol (PI). Native MS was next applied for the affinity ranking of such compounds, demonstrating that PI(4,5)P₂ is a better binder compared to other phospholipids during binding assays [168]. Even more challenging is the native MS detection of a heterotrimeric G-protein complex involving the monomeric A_{2A}, the G α subunit and a nanobody to stabilize the entire assembly [168]. The lipids bound to this receptor complex were measured after quadrupole selection and collisional activation of the whole complex. To obtain such high-quality, high-resolution native MS data of GPCRs, all steps of the native MS workflow must be carefully optimized,

from sample injection using nano-emitters, to MS methods. Furthermore, an optimization of collisional conditions is required to allow efficient desolvation and membrane-mimic removal while maintaining an intact membrane for protein analysis. Comparisons between ligand-binding native MS results obtained from classical NH_4OAc buffer and NaCl -based buffers on $\text{A}_{2\text{A}}$ and glucagon receptors in the presence of different ligands demonstrated that ligand-binding interactions are better conserved on GPCRs using NaCl -based buffers than NH_4OAc buffers [169]. For $\text{A}_{2\text{A}}$, the authors could differentiate agonist from antagonist compounds based on the capacity of the receptor to retain sodium adducts from the storage buffer, thus paving the way for the use of native MS screenings on GPCRs for the detection and characterization of non-covalent interactions under more relevant physiological conditions.

7.2. Activity

7.2.1. G-Protein Coupling and Arrestin Recruitment

Advanced Resonance Energy Transfer (BRET/FRET) strategies have been developed during the last decade to monitor receptor–effector coupling and selectivity in signaling. These approaches have been extensively reviewed [170,171] and have been mostly used with recombinant cell systems expressing receptor and effectors labeled with fluorophores and/or fused to fluorescent/bioluminescent partners or fragments. In the case of isolated receptors, apart from the purification of stable signaling complexes for subsequent structural studies, *in vitro* assays monitoring the coupling to signaling partners are also currently used to assess the functional properties of the isolated GPCRs, in parallel to ligand-binding assays. Initially, these assays focused on the allosteric stabilization of the active state of the receptor and its impact on an agonist's affinity and/or receptor stability. Such assays were used, for instance, to reveal the coupling of the isolated adenosine receptor to engineered mini-G proteins [53]. Specifically, mini-Gs were shown to increase the affinity of the purified $\text{A}_{2\text{A}}$ in detergents for NECA, and significantly increased the stability of the receptor in the presence of the agonist. Besides these assays, experimental designs have been implemented to monitor receptor-catalyzed G-protein activation through GTP turnover (Figure 2). This process can be visualized by monitoring the binding of a radiolabeled [^{35}S]GTP γS [172] or the fluorescent analog of GTP (Bodipy-GTP γS) [173] to the $\text{G}\alpha$ -subunit of the purified $\text{G}\alpha\beta\gamma$ heterotrimer. [^{35}S]GTP γS binding was used to decipher the G-protein coupling to $\text{A}_{2\text{A}}$ in Sf9 membranes [174]. Besides membrane fractions, labeled GTP γS binding assays can also be applied to GPCRs isolated in detergent micelles [175,176] or nanodiscs [177,178]. More recently, a GTP turnover assay has been devised that monitors the amount of free GTP after receptor-catalyzed G-protein activation. Initially described for other GPCRs such as the $\alpha 2$ -adrenergic [179] or ghrelin [180] receptors, this assay has been recently used to reveal the allosteric modulation of the adenosine $\text{A}_{2\text{A}}$ receptor in nanodiscs by cholesterol [74].

Regarding arrestin coupling, although some alternative assays have been reported that are based, for instance, on the use of radiolabeled arrestin [181], a popular *in vitro* assay with isolated GPCRs relies on the use of recombinant arrestin labeled on its finger loop with a specific fluorophore, monobromobimane (mBBR) (Figure 2). This fluorophore is attached to recombinant arrestin through a unique reactive cysteine [182]. The emission properties (intensity, maximum wavelength) of mBBR vary upon coupling to the receptor, and as such, allow for visualization of the GPCR–arrestin interaction. Initially developed with rhodopsin and visual arrestins [183], this assay has since been extended to other GPCRs such as the vasopressin, $\beta 2$ -adrenergic or ghrelin receptors in detergents, amphipols and nanodiscs [178,184–186]. To the best of our knowledge, this assay has not been used with purified $\text{A}_{2\text{A}}$ so far but could be directly applied to this receptor isolated in membrane-mimicking systems, allowing, for instance, screening for particular ligands that would selectively trigger arrestin recruitment.

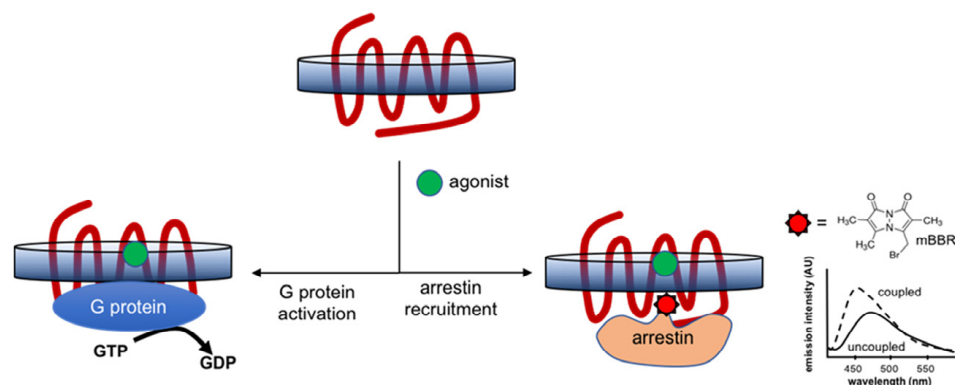


Figure 2. Schematic representation of the *in vitro* assays for G-protein activation and arrestin recruitment. These assays monitor receptor-catalyzed G-protein activation through GDP-to-GTP exchange by measuring either the binding of a labeled GTP analog or the depletion of GTP in the solution. To study arrestin recruitment, the most popular assay relies on monitoring the changes in the emission profile of a fluorophore, monobromobimane (mBBR), attached to the finger loop of recombinant arrestin.

It is clear from the last years in GPCR biology that our understanding of receptor signaling pathways has been completely changed by two concepts: (i) the description of the signaling bias of receptors [187–189] and (ii) the numerous proteins that have been found associated with receptors, such as the MT₁ receptor [190]. For the first point, isolated systems are necessary to better understand how the activation of receptors are controlled. For the second point, an understanding of the activation routes of a receptor will be clearly defined from isolated systems for which potential coupling between purified receptors and their protein binding partners will be studied in the defined environment of nanodiscs.

7.2.2. Functionality of Isolated Receptors: Activity of Receptor Dimers and Oligomers

Even though a monomeric A_{2A} receptor is sufficient to promote G-protein activation or β -arrestin recruitment, this receptor, like many other GPCRs, has been shown to form dimers and larger oligomers [191–194]. GPCR oligomers offer additional inter-receptor communications to potentially regulate their own functional properties. A_{2A} receptor was shown to associate with itself [195], forming A_{2A} homodimers, but also with other receptors to form heterodimers or larger oligomers. In particular, A_{2A} could form heterodimers with the other adenosine receptors A₁, A_{2B} and A₃ [196,197], as well as with the cannabinoid CB₁ receptor [198], the dopamine D₂ and D₃ receptors [199–202] or the glutamate mGlu5 receptor [203]. Most of these studies were performed in heterologous systems, mainly using BRET or FRET approaches. Among the heterodimers, A_{2A}/D₂ has gained a lot of attention. First, its existence is well documented *in vivo* both in animal models [204] and in the post-mortem human brain [205]. Second, A_{2A}/D₂ heterodimers are located in the GABAergic striatopallidal neurons, a region critical in locomotor activity and central in the physiopathology of Parkinson's disease [206]. Interestingly, A_{2A}/D₂ heteromers were shown to operate a reciprocal negative functional crosstalk and the interaction between the two receptors is reduced along the development of the pathology [206,207]. Therefore, A_{2A}/D₂ constitutes a major therapeutic target for the treatment of Parkinson's disease.

While cellular and *in vivo* studies provide a general picture of the functional crosstalk within GPCR oligomers [208] (whether homodimers, heterodimers or heteromers) including A_{2A}/D₂, a precise characterization of the pharmacological, conformational or signaling specificity of each oligomer relative to the monomers or other oligomers is very difficult to reach in these complex systems where several oligomeric species are likely to co-exist. Therefore, isolated systems where a single receptor combination is present, either monomer or oligomer, bring invaluable information to elucidate the molecular, pharmacological and signaling signatures. Additionally, in isolated systems, specific mutations (e.g., non-binding

or non-coupling mutations) can easily be introduced to a given protomer. They can analyze the overall effect of the absence of the oligomer function. To date, a single publication reported A_{2A} oligomers in isolated systems [94]. However, other GPCR oligomers have been analyzed in such systems illustrating the power of the method but also highlighting the technical challenges to overcome [197,209].

As indicated above, purified receptors can be obtained either in detergent micelles, or reconstituted in proteoliposomes or in lipid nanodiscs. Depending on the protein concentration, isolation procedures and purification steps, samples containing the desired receptor composition are obtained. For example, the neurotensin NTS1 receptor is mostly monomeric in detergent at low protein concentrations and becomes dimeric at higher concentrations [210]. Leukotriene BLT2 receptor has been shown to be present in a mixture of monomers and dimers with each species successfully isolated after HPLC fractionation [176]. For nanodiscs, the incorporation of monomeric or dimeric GPCRs can be achieved by acting on the receptor-to-scaffold protein ratio as reported for rhodopsin [211]. Alternatively, the reconstitution of the purified glutamate mGlu2 receptor into nanodiscs has been shown to lead to a mixture of discs containing monomers or dimers that can be isolated by HPLC fractionation [212]. Upon reconstitution, the dimer can assemble in parallel or anti-parallel orientations, with only the former having a physiological relevance. A general assessment of the relative orientation can be conducted using fluorescently labeled protomers either at the N- or C-terminus. Indeed, FRET would occur only at a short distance, i.e., when both fluorophores are located at the same side of the receptor relative to the lipids. Thus, if FRET is detected between probes inserted at the N- and C-terminus, anti-parallel dimers are present [176]. To overcome this phenomenon and to purify only the parallel dimers from a mixture, an original strategy has been developed using the stable CaM binding simultaneously to two similar sequences derived from the CaM binding domain from *Petunia* glutamate decarboxylase (PGD). Accordingly, after the fusion of the 26-residue PGD-tag to the N-terminus of the receptor of interest, only parallel dimers can be purified using a CaM binding step followed by SEC. This method was efficiently used on lipid nanodiscs containing LTB1 receptor dimers or ghrelin GHSR receptor dimers [213]. When considering heterodimers, their formation in detergents or in nanodiscs gives a mixture of several species such as monomers, homodimers and heterodimers. In order to select the heterodimers from this mixture, double tag purification strategies have been developed. Indeed, the fusion of a different tag to each receptor of interest, followed by two successive purification steps, each using one of the tags, lead to the purification of only dimers presenting one of each tag [214]. This method has been used to prepare BLT1 receptor dimers in detergent where one of the subunits only is mutated in its binding site for coupling with a G protein [214]. It is worth noting that some receptors assemble in complexes larger than dimers. Therefore, nanodisc size is not compatible with the insertions of these large complexes. To overcome this limitation, proteoliposomes are an alternative that have been used to reconstitute a heterotetramer composed of two GHSR and two D2 receptors [215]. The heterotetramers were further isolated from the proteoliposomes using SMALPs.

Following the isolation of receptor complexes, functional and conformational experiments can be performed similarly to those carried out on isolated monomeric receptors, highlighting the specific properties of the different complexes. For example, in a rhodopsin dimer, only one of the protomers is in a high-affinity state upon recruitment of G protein while the dimers have a lower affinity for β -arrestin 1 than the monomer [211,216]. The analysis of the GHSR/D2 heteromer highlights that heteromerization directly modulated D2-mediated G_i protein activation. More specifically, such heterocomplexes act directly on the conformation of the $G_{\alpha i}$ subunit [215]. Applying similar strategies to the analysis of the A_{2A} homodimer or A_{2A}/D_2 heteromer analysis would bring important information on the precise mechanism underlying the function of the complexes and would lead to new hypotheses of how to more precisely regulate these pathophysiologically relevant heteromers.

8. Structure of A_{2A}

The early evidence for a necessary transfer of information from structure to function has been reviewed by Piirainen et al. [217] and clearly showed an early path for the rationalization of GPCR structure determination to gain an understanding of the function of the A_{2A} receptor and GPCRs in general. The central role of biophysical techniques in the determination of the structure and function of GPCRs has been nicely reviewed by Langelaan et al. [218]. Furthermore, artificial intelligence approaches, such as with AlphaFold, are powerful tools to predict three-dimensional protein structures and protein–protein interactions. This technology can be accessed for the prediction of the human A_{2A} three-dimensional structure, in addition to almost all other GPCRs [219,220].

8.1. Structural Mass Spectrometry Approaches

Among the large panel of biophysical methods available, structural MS techniques have become a valuable tool for either ligand screening or for more detailed structural characterization of purified GPCRs, including the A_{2A} receptor. The structural MS toolbox encompasses a series of techniques, such as native MS for stoichiometry and oligomerization determination, ion mobility MS for the assessment of global conformation, and labeling methods such as hydrogen/deuterium exchange MS (HDX-MS) and chemical cross-linking to investigate interacting regions and/or conformational changes [221,222].

While native MS and mass photometry have great capacity for the characterization of oligomeric state or binding events (see Sections 6 and 7 above), both techniques are not adapted to provide either amino acid resolution information on binding regions, or the dynamics of assembly or disassembly [113,114]. In recent years, HDX-MS has become a very powerful labeling method for membrane protein analysis, adapted to tackle conformational issues related to ligand binding and even allostery [223]. A typical HDX-MS experiment includes the incubation of a protein of interest in two or more states (apo- or holo-states for example) in deuterated buffer at different time points to allow hydrogens from backbone amides to exchange with deuterium. The exchange reaction is then quenched by switching to acidic pH and low temperature followed by online digestion with acid-functional proteases. The generated peptides are then separated on a reversed phase-HPLC system and detected using MS. Dynamics and binding-site mapping are then extracted by comparing apo- and holo-states. While HDX-MS is becoming a well-established technique for soluble proteins with standardized workflows and protocols [224], the study of membrane proteins remains technically challenging. Thus, extensive work on the optimization of the experimental conditions (composition of buffers, choice of digestion protease, etc.) must be carried out systematically. For instance, nepenthesin protease seems to be particularly suited for membrane protein studies compared to the standard pepsin one. Similarly, the addition of urea as denaturant in quench buffers seems to be beneficial when working with membrane proteins [225]. Other dedicated workflows and methodological developments have been set up to prevent the early deterioration of the instruments caused by the large amounts of detergents and phospholipids in the samples, including an additional off-line cleaning of the LC columns, or the use of zirconium beads for lipid depletion, etc. [226,227].

Figure 3 presents a typical HDX-MS workflow used to monitor the binding of the antagonist ZM241385 on A_{2A} and the induced conformational changes. Online digestion of the protein allowed to cover 70.8% of the sequence with 79 identified peptides and a redundancy of 2.7, due to the use of nepenthesin II protease, was reported as more efficient on GPCRs. Experiments carried out on apo A_{2A} showed a low deuterium incorporation on transmembrane domains, reflecting their known low accessibility to the solvent. Among these domains, some peptides (20-31; 58-62; 86-93; 134-141; 234-241; 245-253; 271-280) exhibited a dramatically low deuterium incorporation typically less than 10%, which is consistent with the known structure of the protein. Upon ZM241385 binding, a significant part of the protein was protected meaning that A_{2A} incorporates less deuterium in the presence of the ligand. Several peptides exhibited statistically significant differences in deuterium incorporation (*p*-value of 0.01). Peptides 168-175, 245-250 and 245-253 in the known ligand

binding region [38] (F168, N253 and I274 residues) presented a H/D protection higher than 5% confirming the ZM241385 binding site. The significant protection of additional peptides 59–65, 276–288, 277–282 and 277–288 supports the known ligand fixation, as involved in the following residues (A63, H278), forming a cleft to accommodate the phenol group of ZM241385, as well as for the end of the transmembrane domain 7 (TM7) represented by residues 281 to 298 [72]. Other regions showing differences in incorporation (TM3, TM4 and the loop in between) may reflect the induced-structural rearrangement of the receptor upon ligand binding. This example highlights the ability of HDX-MS experiments to measure membrane protein conformational dynamics in their native state.

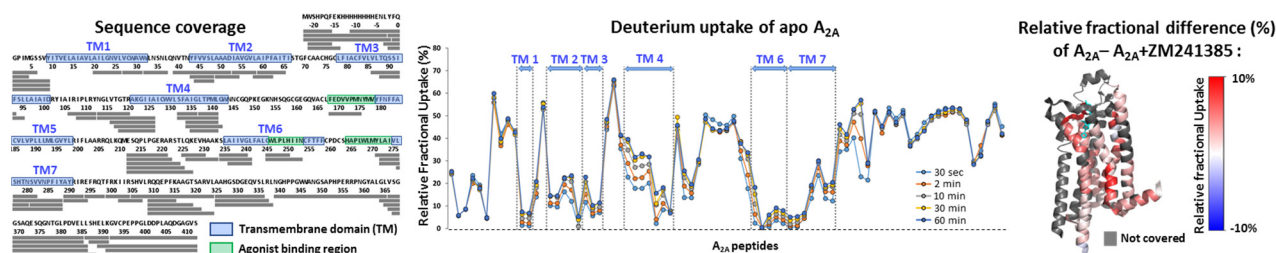


Figure 3. Structural characterization of A_{2A} upon ligand ZM241385 binding with HDX-MS. HDX-MS identified peptides shown on the A_{2A} sequence; transmembrane domains are highlighted in blue and the agonist ligand-binding region in green (left). The relative deuterium uptake of A_{2A} of each identified peptide at each deuteration times are plotted, with peptides in the transmembrane regions framed in blue (middle). Relative fractional uptake differences between A_{2A} and A_{2A} with the ligand ZM241385 are represented on the protein structure (PDB code: 4E1Y [56]) at a 30 min deuteration time; regions not covered are colored in grey (right).

In addition to continuous labeling studies (deuteration time points from a few seconds to hours), HDX-MS can also be used in a time-resolved manner to monitor sequential conformational changes in the protein of interest over time, using very short “pulsed” deuteration time [223] either in the apo state to report on folding and aggregation [228] or in an holo state to report on protein–protein interaction. To our knowledge, this pulsed HDX-MS approach has been reported only once on A_{2A} by Du et al. [229], in order to better understand why GPCR/G-protein complexes are stabilized in a GDP free-state, how this nucleotide is released and to lay the structural basis of coupling specificity between G-protein subtypes and GPCRs. Interestingly, the authors used hydroxyl radical protein footprinting with mass spectrometry [230,231], which aims at looking at protein conformational changes after the irreversible labeling of amino acid side chains by hydroxyl radicals (generated by radiolysis or photolysis reactions), as an orthogonal technique to detect very early events (from a few milliseconds to seconds).

Among the emerging structural MS methods, chemical cross-linking followed by MS analysis is still rarely used for in vitro membrane protein analysis. The first step consists of the covalent binding of a cross-linker agent to the side chain of amino acids, followed by the quenching of the reaction and the reduction, alkylation and digestion step. Then, peptides including cross-linked dipeptides are separated and analyzed using LC-MS/MS and finally identified via different software due to the mass increment due to the cross-linking reagent [232,233]. The A_{2A} receptor was used as a model to assess which chemical cross-linkers are suitable for the characterization of GPCRs by cross-linking mass spectrometry (XL-MS) [234]. Similar to HDX-MS, all steps of the XL-MS workflow have to be adapted to membrane protein analyses, especially the choice of the cross-linker reagent and the cross-linking conditions, enzymatic digestion and sample clean up, along with adapted LC-MS methods and data treatment. By comparing the MS-identified cross-links to the structurally compatible cross-links from available GPCR structures, Jones et al. concluded arginine-arginine cross-linker and lysine-arginine are best adapted for XL-MS analyses of GPCRs and are highly complementary to the existing classical cross-linkers

present in the toolbox of specialists [235]. Altogether, structural MS methods have great potential to complement more classical high-resolution structural biology techniques, and are becoming more and more popular.

8.2. X-ray Crystallography of GPCRs: Uncovering Conserved Activation Mechanisms Using A_{2A} as a Model

GPCRs are highly dynamic and become particularly unstable during the purification process once extracted from the cell membrane. As a consequence, crystallization of GPCRs remains a challenging task. To overcome such difficulties, protein engineering strategies have been developed to reduce the flexibility and to increase the stability of these membrane receptors. A_{2A} was one of the first receptors for which the high-resolution structure was solved using X-ray crystallography. The T4 Lysozyme (T4L) fusion partner, initially established for resolving the high-resolution crystal structure of the beta-2 adrenergic receptor (β_{2A} R) [236], was later applied to solve the first structure of A_{2A} [51]. The T4L increases the hydrophilic surface of detergent-solubilized GPCRs to promote crystal contacts, and reduces the flexibility associated with the TM6 helix [236]. Subsequent developments led to the application of the BRIL fusion partner into intracellular loop 3 to solve the A_{2A} crystal structure at 1.8 Å [56]. Another key engineering strategy to overcome the instability of GPCRs in detergents is by introducing point mutations in the 7TM domain to select thermostabilized ligand-bound conformations. The development of conformational thermostabilization of the A_{2A} [31,237] has been successfully applied for a range of co-crystallized agonist- and antagonist-bound structures [52,72].

Ciancetta pointed out in 2019 that [238] “A_{2A} represents to date (2019) one of the rare Class A GPCR that has been solved in the inactive (R) [79], active-intermediate (R*) [68], and fully active state coupled to an engineered Gs protein α subunit (R * G) [53,64]”. Most of these structures have been obtained using X-ray crystallography. During the last 3 years, more structures were solved in similar configurations for other Class A GPCRs.

The structural determination of multiple A_{2A} conformational states has contributed greatly to the understanding of Class A GPCR activation mechanisms [38,51–53,68,72,239]. The antagonist-binding pocket has been thoroughly characterized by the co-crystallization of A_{2A} bound to ZM241385 (Figure 4A) and to various xanthine-containing compounds including caffeine and XAC, using either thermostabilization or fusion protein strategies [72]. Later developments combining the antagonist-thermostabilized point mutations with the BRIL fusion partner provided the highly utilized construct known as A_{2A}-StaR2-bRIL [69,79,240]. Multiple structures of antagonist-bound states also revealed the sodium allosteric modulation of A_{2A} [56,58]. The allosteric sodium ion-binding site is nested in a set of highly conserved residues, located below the orthosteric binding pocket, and bound to the highly conserved Asp^{2.50} [56,58].

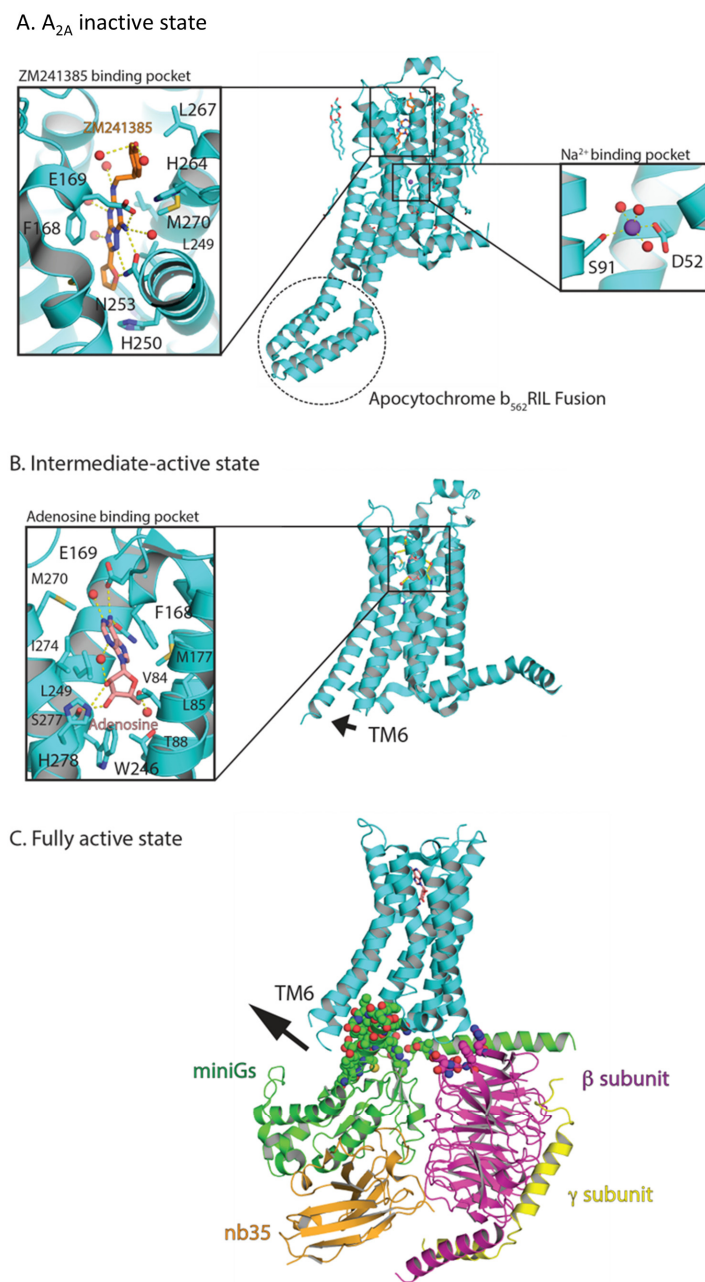


Figure 4. Three-dimensional structures of the human adenosine A_{2A} receptor. The A_{2A} receptor is bound to the antagonist ZM241385 (A; PDB: 4EII [56]), agonist adenosine (B; PDB: 2YDO [52]), NECA and mini-Gs protein (C; PDB: 6GDG [241]). These figures represent the inactive, intermediate-active, and fully active states, respectively. Polar contacts are represented as yellow dashed lines, with the sodium ion depicted as a purple sphere and water molecules as red spheres.

Agonist-binding modes of UK-432097 [68], NECA and adenosine (Figure 4B) to the thermostabilized A_{2A} highlight the molecular interaction of the ribose moiety of all agonists with Ser277 and His278 and contraction of the receptor binding site [52]. Agonist binding also initiates a slight outward movement of TM6 that defines an intermediate-active conformation [52,57,58,68,242]. The crystal structure of the A_{2A} bound to an engineered Gs protein (known as mini-Gs) illustrated the large outward movement of TM6 that is required for the mini-Gs to bind, in addition to slight movements of TM5 and TM7 and rotamer shifts of several amino acids near the G-protein binding site (Figure 4C) [53].

This was further validated by solving the cryo-EM structure of A_{2A} bound to the heterotrimeric mini-Gs (Figure 4C) [241]. However, comparisons between the NECA

binding pocket in the intermediate [52] and fully activated conformation [53] show identical binding modes. Together, these studies provided an overview of the conserved activation mechanisms of Class A GPCRs [243].

High-resolution crystal structures of A_{2A} have also led to the advancement of structural-based drug design (SBDD) [70,135]. Recent reports have outlined chemical modifications of an adenosine scaffold, commonly attributed to A_{2A} agonist activity, which led to the alteration of the compound from an agonist to an antagonist [244]. Although the ribose moiety is accepted to be important for A_{2A} agonist activity, recent structural and functional data have demonstrated the binding mode of a non-riboside partial agonist [242]. Although allosteric regulation of the human A_{2A} remains relatively unexplored, differences between antagonist-bound conformations have suggested the possibility of an allosteric binding site within the 7TM binding pocket [245]. Taken together, these studies highlight future avenues for the design of selective orthosteric A_{2A} modulators and the possibility of development of allosteric ligands.

Technical advances have been made in macromolecular crystallography for speeding up the workflow for screening, data collection and processing, proving to be essential for the progress of GPCR structural biology. The infrastructure developments of synchrotron microfocus beamlines have also been key for solving A_{2A} structures and GPCRs in general. Moreover, the use X-ray free-electron lasers (XFELs) and lipid cubic phase (LCP) injectors for sample delivery has allowed the structural determination of A_{2A} at room temperature, opening new possibilities for investigating the highly dynamic nature of GPCRs [78,81,246–251]. Currently, X-ray crystallography has become one of the more accessible approaches for the structural determination of GPCRs. Understanding the dynamic activation mechanism of A_{2A} remains of strong interest. Accordingly, the use of integrated approaches, such as cryo-EM, will uncover more details of A_{2A} in complex with its signaling partners in the future.

8.3. Cryo-Electron Microscopy

Cryo-EM is a powerful technique for solving structures of challenging and dynamic proteins that are difficult to obtain using X-ray crystallography. While there is a debate on the respective virtues of cryo-EM versus X-ray crystallography for the determination of protein structures, particularly those of membrane-bound and integral membrane proteins [252–254], it is clear to us that the complementarity between those approaches opens avenues for a better description of membrane-bound proteins in the future, as reviewed by others [255,256].

A key advantage of cryo-EM is that it does not require the preparation of protein crystals, rather it allows for the imaging of single isolated particles in isomorphous ice prepared directly from purified samples. Conversely, the application of the cryo-EM technology for proteins or complexes smaller than 60 kDa remains challenging due to the inherent low signal-to-noise ratio (or low contrast) that limits the accurate alignment of projections during image analyses. Accordingly, the relatively small size of a ligand-bound Class A GPCR such as A_{2A} (approximately 35–45 kDa) is typically not adapted to cryo-EM studies but is suitable for X-ray crystallography which remains the preferred technique for the determination of GPCR structure in the inactive state [257].

This size limitation has been recently circumvented by increasing the molecular weight of GPCRs isolated in protein complexes. Most remarkably, advances of the near-atomic resolution of activated GPCRs in complex with G-protein binding partners have been achievable with the use of cryo-EM [258]. Cryo-EM has since been successfully employed to solve the structure of an engineered A_{2A} thioredoxin fusion protein (TrxA- A_{2A}) in complex with the mini-Gs protein [241]. A recent study made use of an anti-BRIL Fab in complex with the A_{2A} -BRIL to increase the molecular weight of A_{2A} and to facilitate cryo-EM data processing [61].

One method recently developed, known as microcrystal electron diffraction (MicroED), combines crystallization in LCP with the acquisition of cryo-EM for structural determina-

tion. Here, a beam of electrons is transmitted through protein microcrystals grown in LCP, rather than X-rays by conventional crystallography. A_{2A} is currently the only GPCR for which this method has been successfully employed [259]. Future work ought to apply the use of cryo-EM to investigate the structures of the purified A_{2A} in complex with signaling partners such as the full-length Gs heterotrimer and other intracellular signaling partners.

8.4. NMR Approaches

At a molecular level, GPCR signaling can be described through the concept of free energy landscape [260] where orthosteric ligands, intracellular effectors and allosteric modulators influence both the population and lifetime of various inactive, intermediate and active states. High-resolution NMR spectroscopy is a powerful tool to investigate these complex energy landscapes. While NMR can be used to obtain de novo structures of proteins, in the case of large complexes such as GPCRs associated with various surfactants or lipids, most of the NMR studies that are conducted in solution concern conformational landscapes or interactions, with considerations sometimes relating to chemical exchanges. They rely advantageously on the plethora of atomic scale structures obtained from X-ray crystallography and cryo-EM to highlight the coexistence of several inactive and active sub-states, including lowly populated states, but also, to a much lesser extent, the detection of some chemical exchanges. Along with the $\beta_{2A}R$, A_{2A} is one of the GPCRs most studied with NMR. These studies are based on different isotopic labeling strategies and/or expression systems.

An appreciable amount of NMR studies of A_{2A} is based on fluorine (¹⁹F) NMR [48,50,73,93]. This nucleus can be easily incorporated in proteins via biosynthetic pathways or chemical reactions and has the advantage of displaying a large chemical shift dispersion compared to other nuclei that are traditionally detected to study biomolecules [261,262]. This partly counterbalances the fast ¹⁹F transversal relaxation due to the large chemical shift anisotropy, which becomes the major source of linewidth in high magnetic fields. As GPCRs are usually expressed in eukaryotic expression systems, ¹⁹F labeling via chemical reactions is generally preferred. However, in this case, only solvent-exposed residues can be labeled due to the accessibility for the chemical reaction to occur and, consequently, all these studies are based on single extra-membrane ¹⁹F probes. Following such a strategy, the conformational landscape of the receptor was observed using one-dimensional (1D) ¹⁹F NMR at the three following positions that are all located at the extracellular tip of TM helices: V229C^{6.31} (TM helix VI, TM6) [48,93], L225C^{6.27} (TM6) and A289C^{7.54} (TM7) [73]. 1D ¹⁹F spectra in these locations display a pattern characteristic of a complex conformational ensemble, composed of several inactive and active sub-states in equilibrium. Their population strongly depends on the presence or absence of ligands and their associated pharmacological properties, as well as the presence or absence of effectors. This illustrates that signal transduction by A_{2A} relies on a dynamic equilibrium of several conformations that can be detected through a fluorine probe. Interestingly, several coexisting active states, whose populations are related to specific ligands were observed, fitting with the concept of partial agonism, i.e., when a GPCR is activated to a sub-maximal level.

By investigating these three positions of the ¹⁹F probe, different distributions of populations in the conformational ensemble could be identified in the apo or holo (e.g., in the presence of the full agonist NECA) states. Between V229C^{6.31} [48,93], or L225C^{6.27} [73], and A289C^{7.54} [73], this may reflect a different response to drug efficacy at the intracellular ends of TM6 and TM7 [73]. In contrast, it is most puzzling that the distribution of populations is not the same between V229C^{6.31} and L225C^{6.27}, as these two residues are localized very closely with each other at the extracellular tip of TM6 [48,73]. Moreover, these two studies were performed in the same detergent solution. This discrepancy could be associated with different possible effects of aromatic ring current fields in the vicinity of these two ¹⁹F NMR probes that could partially respond differently to conformational changes [263].

Interestingly, looking at the conformational ensemble of V229C^{6.31} in the apo state or in the presence of NECA of A_{2A} embedded in a lipid nanodiscs, the 1D ¹⁹F NMR spectra

are very different from what has been previously observed in detergent [50] (Figure 5), highlighting the importance of studying the energy landscape of membrane proteins in a native-like lipid bilayer environment, even if using very small nanodiscs where the receptor is surrounded by just one layer of lipid (lipoprotein Δ H5 [264]), such as those used in Huang et al. [50]. Furthermore, this study is also remarkable as they could additionally depict the conformational ensemble of V229C^{6.31} in the presence of a heterotrimeric G protein, revealing a G-protein-bound intermediate and distinct nucleotide-free state in the presence of either a partial or a full agonist. This study further suggests a scaffolding role for the G $\beta\gamma$ -associated subunits, facilitating the allosteric transmission through the transmembrane part of the receptor from the orthosteric ligand site to the intra-cellular side in a ligand-dependent manner [50].

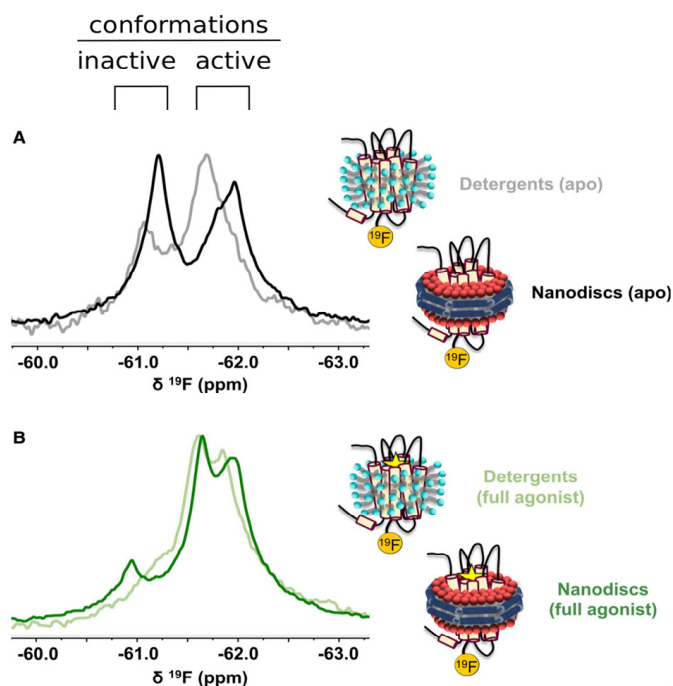


Figure 5. A_{2A} exhibits a different distribution of states and slower exchange dynamics in a lipid bilayer environment than in detergent micelles (A,B). Comparison of the ^{19}F NMR spectra of apo (A) or agonist-bound (B) A_{2A} reconstituted in either LMNG micelles or phospholipid nanodiscs (legend and figure adapted from Huang et al. [50] with automatic permission from Elsevier).

A common feature identified in all of these studies is that most of the sub-states sampled by the ^{19}F probe suggest a slow chemical exchange at the ^{19}F chemical shift timescale (millisecond timescale), under the assumption that some chemical exchanges are present between the conformations revealed by the fluorine probe. A strong indication that such chemical exchanges are indeed present is given by the observed variations in either chemical shifts and/or signal linewidths in association with measurements of the transversal relaxation time (T_2) upon the addition of various ligands and/or a heterotrimeric G protein or a G α mimetic peptide. 1D ^{19}F NMR can also be expanded with a second ^{19}F frequency axe to directly visualized conformational exchanges by 2D [^{19}F - ^{19}F] exchange spectroscopy (EXSY) spectra [73] (Figure 6).

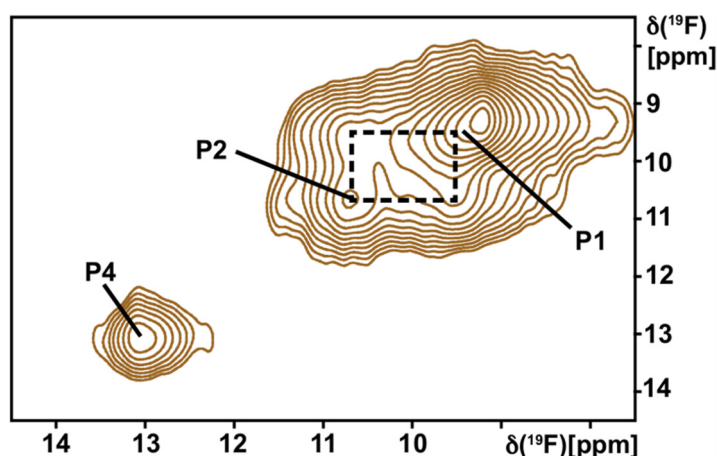


Figure 6. Conformational exchange in the A_{2A} complex with the full agonist NECA observed using 2D exchange spectroscopy. A contour plot is shown of a 2D $[^{19}\text{F}, ^{19}\text{F}]$ -EXSY spectrum collected at 280 °K with a mixing time of 100 ms. The diagonal peak positions of the active-like sub-states P1, P2, and P4 are labeled, and a dashed box indicates cross-peaks observed between conformations P1 and P2. This spectrum indicates the existence of a low-energy barrier for interconversion between sub-states P1 and P2 (legend and figure adapted from Susac et al. [73] with automatic permission from PNAS).

Among the various allosteric modulators that have been identified in the activation of GPCRs, lipids play an important role by either binding to specific sites, e.g., cholesterol or phosphatidylinositol molecules acting as specific cofactors [180], or by modulating the physical properties of the lipid bilayer [265], thereby impacting the activation of these receptors. Docosahexaenoic acid (DHA) phospholipid chains are well represented in the membrane of cells where A_{2A} is highly expressed. With the help of the introduction of a ^{13}C -labeled and protonated methyl group in methionine residues, Mizumura et al. investigated the effect of these acyl chains on the activation of A_{2A} [49]. Thanks to deuteration, this study could be performed on quite large nanodiscs stabilized by the lipoprotein MSP1E3 [103]. Based on 2D $[^1\text{H}, ^{13}\text{C}]$ NMR spectra, they deduced that in the presence of a partial agonist, A_{2A} is in an equilibrium between inactive and active states, while a full agonist shifts the equilibrium towards the active state only. Interestingly, the addition of DHA enhances the G-protein activation by A_{2A} , a positive allosteric effect that has also been observed, for instance, with phosphatidylinositol (4,5)-bisphosphate and an analog of cholesterol in the activation of ghrelin [180] and leukotriene receptors [266], respectively. Based on these observations, they make the assumption that DHA acyl chains modify the conformational equilibrium centered at the NPxxY motif that would induce a large clockwise rotation of TM6, which is based on the lipid-dependent chemical shift variations of methionine residues located at the cytoplasmic ends of TM3 and TM6 (M106 and M232, respectively).

Another important class of allosteric modulators of GPCRs are cations. Deciphering A_{2A} conformation landscape in various situations, Ye et al. confirmed the negative allosteric modulation of A_{2A} activity by sodium ions, as already observed in pharmacological and structural studies and confirmed as being common to many Class A GPCRs [56]. Interestingly, they observed an opposite trend upon the addition of divalent cations, e.g., Ca^{2+} and Mg^{2+} , which had positive allosteric effects that were boosted by the presence of an agonist and a G-protein-derived peptide [93]. With the aid of molecular dynamic (MD) simulations, they propose that divalent cations would establish bridges between extra-cellular acidic residues belonging to TM5 and TM6, which would allosterically contribute to the opening of the G-protein binding cavity on the intracellular side of the receptor.

^{19}F NMR can be useful to study GPCR conformational ensembles thanks to a large chemical shift dispersion, a high sensitivity (a gyromagnetic ratio close to the proton) and the fact that there are no background signals. However, if using just a single incorporated

probe, in particular, localized in an extra-membrane region, it is very difficult to obtain a global and coherent view of the molecular mechanism governing activation. To overcome these limitations, Eddy et al. used uniform-labeling of the receptor with ^{15}N and ^2H isotopes to investigate the allosteric coupling between the drug binding event and intracellular signaling [47]. The authors took advantage of the ^1H , ^{15}N correlation peak chemical shifts of exchangeable Gly amide (^1HN) and Trp indole protons, as their $^1\text{H}/^{15}\text{N}$ cross-peaks resonate in well-isolated parts in typical 2D [^1H , ^{15}N] correlation spectra. Thanks to an additional partial deuteration (~70%), this strategy enabled them to identify well resolved NMR cross-peaks that could be assigned to six individual Trp indole protons and eight Gly amide protons by single mutagenesis. As discussed in this study, multiplying the number of sampling points in the receptor is only interesting if most of the assigned residues are located in the vicinity of “ring current fields” to increase the sensitivity of these nuclei to subtle changes in the conformational polymorphism of the receptor. Hence, significant ^1H and/or ^{15}N chemical shift variations for 11 of the 14 assigned residues were observed between $\text{A}_{2\text{A}}$ bound to an agonist and to an antagonist. Incidentally, no data were reported on the apo state in this work. The synthesis of all their observations leads to a model where there is a strong interplay between the very-well-conserved residue of A52^{2,50} and W256^{6,48}; the latter, which is located just below the ligand orthosteric pocket, has previously been hypothesized to change its rotameric state upon activation (“toggle switch”) [267]. The authors also extended the use of indole protons of tryptophan to introduce extrinsic tryptophan residues in $\text{A}_{2\text{A}}$ to sample the energy landscape of the receptor at additional locations, confirming that the receptor responds differently to ligands with variable efficacy and that the intracellular end of TM6 samples wide conformational variations upon activation [268]. More recently, they applied this same strategy to compare the conformational ensembles of $\text{A}_{2\text{A}}$ associated with two different partial agonists (LUF5834 and Regadenoson) and the full agonist NECA [269]. Their study suggests a different signal propagation pathway based on variations in populations of simultaneously coexisting conformations at equilibrium [269].

Solution-state NMR undoubtedly represents a key ally to classical structural approaches using crystallography or electron microscopy to further explore the energy landscape of very dynamic proteins, such as GPCRs. Numerous studies performed at or nearly at physiological temperatures with no or just a few chemical modifications depict a complex conformational landscape often associated with high kinetic barriers. $\text{A}_{2\text{A}}$ is no exception, and beyond representing an important target for several human diseases, it also portrays an interesting model to study the fundamental properties of GPCRs, such as biased or partial agonism and the impact of allosteric modulators in the activation process. From the studies performed so far, it appears to be better to sample several locations at the same time within the receptor, whenever possible. In order to connect different parts of the receptor in the molecular activation mechanism(s), it will also be important to sample novel positions beyond the water-exposed regions. This should also limit the discrepancies between different approaches relying on nuclei with different physical properties and/or sample conditions. Moreover, as a membrane protein in solution is highly sensitive to its swimming belt, it seems to be essential to try to study the receptor in a membrane-like environment. Furthermore, in consideration of the importance of the constitutive activity in GPCR signaling, it will be crucial to include the *apo* state in all subsequent studies, as this has not been conducted systematically in the studies mentioned here. One major objective for NMR remains the determination of the kinetic barriers separating the different sub-states observed, which, in addition to the determination of populations, will finally provide fundamental kinetic and thermodynamic data to better understand the signaling process through these integral membrane receptors such as $\text{A}_{2\text{A}}$. One possible path to help to increase the amount of information gained from an NMR experiment would be to improve the deuteration level of the receptor in conjunction with the use of appropriate isotope-labeling schemes, high magnetic fields and other suitable methodologies [270]. This would allow not only base NMR-derived activation models on chemical shift variations or

intensities of NMR signals, but would also be of particular advantage to investigate the dynamics in GPCRs using the relaxation phenomena of nuclear spins [271].

8.5. Future Trends for GPCR Structure Studies

8.5.1. Cryo-Electron Tomography

Cryo-electron tomography (cryo-ET) data can be collected with the same transmission electron microscopes as single-particle analysis data. Cryo-ET is label-free and provides three-dimensional snapshots of organelles and proteins at nanometer resolutions within their functional cellular environments, allowing users to visualize and understand how they, and other molecules, work together to carry out major processes in a cell.

The majority of Class A GPCRs are relatively small and thus very difficult to identify *in vivo* in a cellular environment. However, one can cite the very interesting example on visual rhodopsin, a member of the GPCR family [272]. In this work, the authors identified the highly ordered arrangement of rhodopsin molecules in the disk membrane of intact photoreceptors with molecular resolution. Due to the small size of rhodopsin (40 kDa), the resulting molecular map was not amenable to fitting with the X-ray structure of rhodopsin. However, the resolution was sufficient to identify that rhodopsin was forming dimers as building blocks of a hierarchical supramolecular architecture.

This structure is characterized by aligned rows of dimers, and pairs of rows. The authors proposed that GPCR dimerization and oligomerization offer possibilities for the allosteric regulation of GPCR activity. These results support previously published work, which show for the first time that rhodopsin molecules assemble in dimers and form paracrystals in native rod outer segment membranes using atomic force microscopy (AFM) images [273].

8.5.2. Structural Flexibility of GPCR Investigated Using Single Molecule Fluorescence and Atomic Force Microscopy

Evidence has emerged that, rather than being activated by ligands in an on/off manner switching from an inactive to an active state, GPCRs exhibit high structural flexibility in the absence and even in the presence of ligands. The physiological as well as pharmacological impact of this structural flexibility remains puzzling. Single molecule fluorescence offers the advantage of monitoring the structural dynamics of biomolecules in real-time, with minimal structural invasiveness and in the context of complex biological environments; in particular, single-molecule Förster resonance energy transfer is a key method to measure actual distance changes between two probes, and highlight conformational changes occurring at timescales relevant for protein conformational movements [159].

AFM-based high-resolution imaging and force spectroscopy have been also successfully used to better understand the molecular mechanism details of how GPCR conformations are modulated by ligand-binding or GPCR-transducer complex formation [274].

9. Pharmacology of the Isolated Protein A_{2A}

9.1. Background

While finding inhibitors or ligands for soluble proteins is relatively easy, and basically concerns the number of molecules one can screen, for membrane-associated proteins, such as GPCRs, it is far more complicated. A key to understand GPCR function still resides in the availability of molecules with chemically diverse structures—in other words, structurally diverse ligands (agonists as well as antagonists) are needed. Although this is more important for some GPCRs (melatonin, dopamine) than for others (histamine, serotonin), it is crucial that efforts are made in screening large libraries of molecules with the hope of finding new ligands. The availability of new approaches, particularly *in silico*, will assist in the discovery of new molecules, although many obstacles remain on these avenues.

9.2. Background on Ligand Chemistry

9.2.1. A_{2A} Receptor Agonists

Not surprisingly, most adenosine receptor agonists are derivatives of adenosine, while only a few examples of non-nucleoside agonists have been found, as shown in Figure 7 [5,275]. Due to the well-known fact that adenine is the “address portion” of the adenosine scaffold, while ribose encodes the “message portion”, most modifications of the ribose part usually resulted in a decrease in efficacy. The only well-tolerated modification identified so far was the exchange of the primary hydroxyl with an ethyl amide. Besides that, a thioether analogue, GS9667, and a molecule containing a rigid pseudo-ribose moiety, MRS-5151, also proved to be successful agonists [276]. In contrast, adenines lacking the entire ribose entity most often act as antagonists. Modifications at the C-2 and C-6 positions on the adenine ring were well explored and tolerated, and led to the discovery of the first A_{2A} selective agonist, CGS 21,680 [277]. There are two main families of non-nucleoside agonists, based on other heterocyclic scaffolds such as the dicyanopyridines [278].

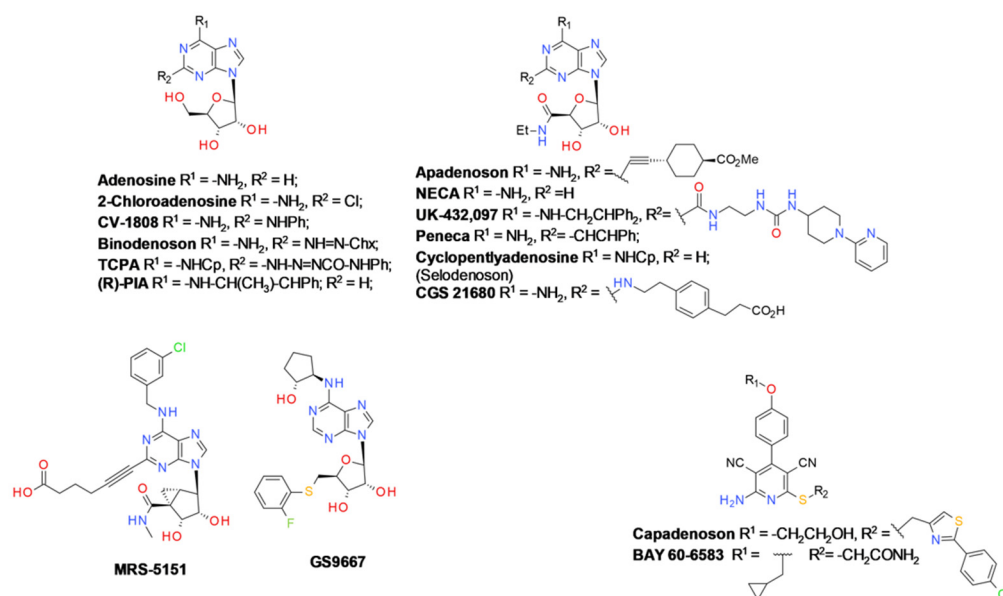


Figure 7. Structures of the main A_{2A} receptors non-nucleoside based agonists.

Most of these compounds have pK_i ranging from 9.4 (Apadenoson) to 5 (MRS-3558) and display some specificity over the other adenosine receptor subtypes. However, the specificity profile of these compounds remains poorly explored outside the adenosine receptor family. Activities of agonists at A_{2A} and other receptors have generally been poorly reported, and most of them came from industrial origins (e.g., Pfizer, AstraZeneca, Ciba-Geigy). Comprehensive molecular details of available A_{2A} receptor ligands, including functional motif, binding sites, EC₅₀/IC₅₀, pK_i, pA₂, and related references can be found either in the IUPHAR data base [1] or in IJzermann et al. [5].

9.2.2. A_{2A} Receptor Antagonists

For the discovery of antagonists, theophylline and adenine initially served as a nature-derived starting points (shown in Figures 8 and 9), from which compounds with diverse modifications were produced. Nevertheless, alternative scaffolds including flavonoids or antimalarial quinolines were also used to design antagonists of adenosine receptors (Figure 10). Alkylxanthines, such as CSC or Rolophylline, are the prototypical adenosine receptor antagonists and hundreds of analogues have been reported, leading to selective compounds for each adenosine receptor subtype [279]. The best A_{2A} receptor antagonists were developed with the incorporation of 8-styryl groups on the xanthine scaffold, resulting in the first approved A_{2A} antagonist drug, Istradefylline (Figure 8). A large

number of adenine derivatives without the ribose moiety have been reported as adenosine receptor antagonists, starting with the A₁-selective N-0840 first reported in 1987 [280]. It was followed by many alternative but closely related heterocycles, including fused five- and six-membered rings. One of the tricyclic series has progressed to Preladenant [281], which reached the clinical phase (Figure 9). It is worth noting that the most powerful antagonist is SCH442416 [282,283], although the actual antagonistic property seemed not to be experimentally determined, but was based on the fact that this compound displaced the potent antagonist, SCH58261 [284]. This compound and its reported analogues were also structurally related to adenine. The development of this molecule led to a positron emission tomography (PET) imaging compound applied to visualize the A_{2A} expression in vivo [283,285].

Molecular pharmacology shows that the A_{2A} agonist and antagonist families are well-populated. They form a large ensemble of molecules with a large panel of properties, although their chemical diversity is somewhat restricted, as is often the case for small molecule GPCRs, such as melatonin [286] or serotonin [287] receptors.

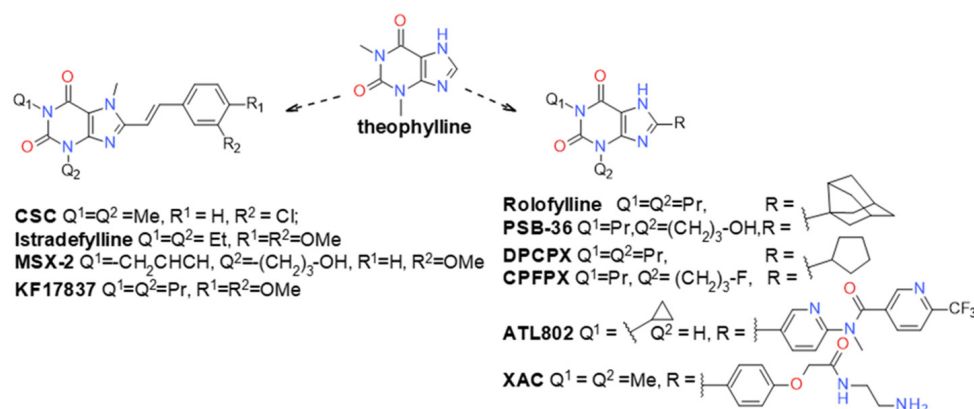


Figure 8. Structures of the main A_{2A} receptors agonists derived from theophylline.

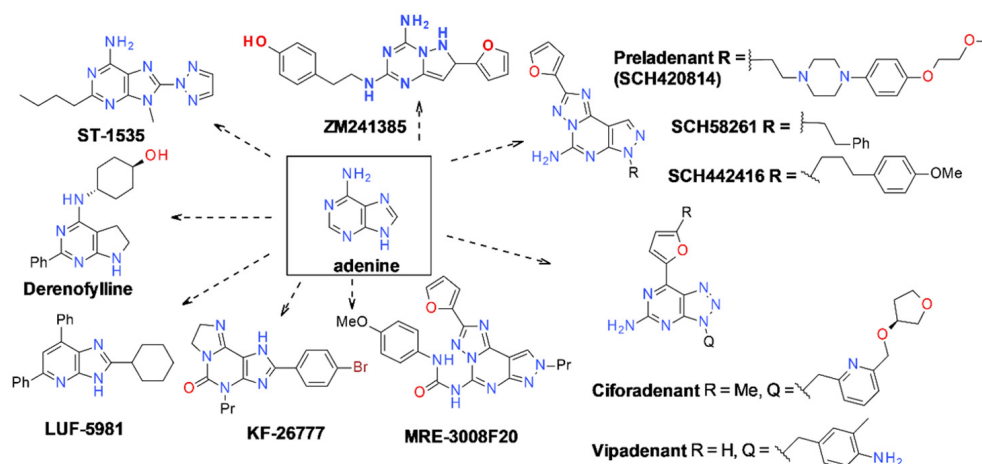


Figure 9. Structures of the main A_{2A} receptors agonists derived from adenine.

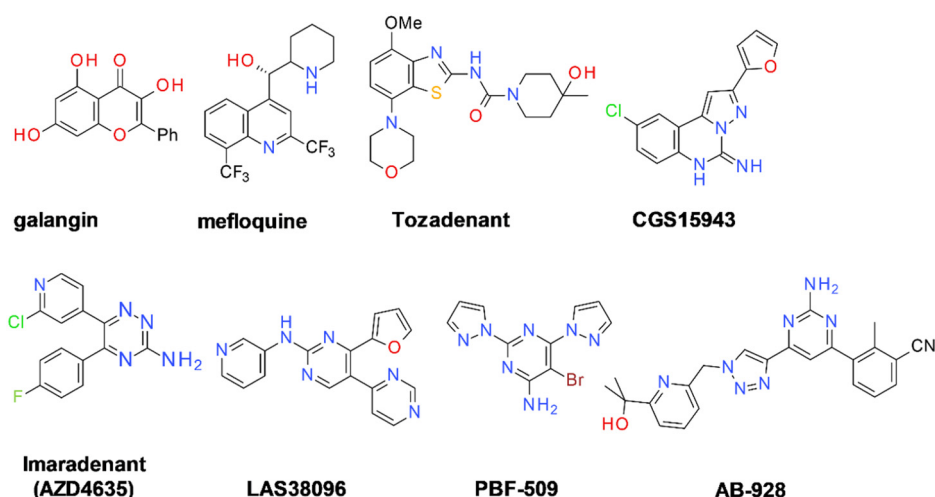


Figure 10. Structures of the main A_{2A} receptors antagonists.

9.3. Allosteric Modulators

Among the new descriptions of GPCR ligands, besides the classical antagonists, agonists and inverse agonists, are also the allosteric modulators of GPCR signaling which have previously been reviewed in depth [288–290]. In brief, such compounds bind to the receptor at a site different to the orthosteric site, leading to subtle modifications of the area, and hence to changes in either the affinity for the agonist or of the activity of orthosteric ligands. Two major classes of allosteric ligands have been described, i.e., negative allosteric modulators (NAMs) or positive allosteric modulators (PAMs). NAMs inhibit the activation of the receptor and the downstream signal transduction. On the other hand, PAMs increase the activation of the receptor. Furthermore, ions are also considered as allosteric modulators for multiple GPCRs, such as Na⁺ for the adenosine A_{2A} receptor [89,291] or Zn²⁺ for the β₂-adrenergic receptor [292] for which they have been shown to bind at allosteric sites which differ from the orthosteric site as reviewed by van der Westhuizen et al. [290]. While they offer new ways to modulate and fine-tune GPCR activity, they are quite difficult to screen for, as they might be dependent on the nature of the agonist they work in conjunction with. This poses a great challenge for the screening format, especially when functional assays are performed on living cells. Furthermore, to characterize this novel, more peculiar type of binding sites of NAMs and PAMs, a large panel of biophysical techniques is required, including molecular dynamics (MD) simulations, ligand structure–activity relationship studies, and in-solution NMR analyses, to name a few [293].

The main example of an A_{2A} PAM is perhaps DFIABA (3,4-difluoro-2-((2-fluoro-4-iodophenyl)amino)benzoic acid), a compound that enhances adenosine signaling at the A_{2A} receptor, and thereby induces slow wave sleep without affecting body temperature in mice [294]. This remarkable feature exemplifies the role of these types of molecules, promising precise modulations of signaling processes with reduced side effects. Other molecules with similar functions include, AEA061 [295], or inosine. Inosine is an endogenous molecule, that could open avenues in the concept of natural PAMs or NAMs, adding to the arsenal of molecules modulating GPCRs functions [296]. The interplay between several receptors, in pathological models illustrates the complexity of various pharmacological ligands, as soon as they interfere with several profoundly different GPCRs, e.g., M4 and A_{2A} [297], mGlu₄ and A_{2A} [298] or (TLR)4 and A_{2A} [299]. Adenosine receptors have also been screened for NAMs, as adenosine receptor subtype-specific negative allosteric modulators would offer access to new potential treatments for analgesia, possibly with minimum negative off-target effects, an area where immense treatment progress remains to be done [300].

9.4. Fluorescent Ligands

Adenosine receptors are of therapeutic interest and, as such, selective ligands attract particular attention. Indeed, a plethora of ligands with well-defined pharmacological profiles have been developed during the last years for A_{2A} . Due to the accumulation of the three-dimensional structures of A_{2A} in different activation states, the design of ligands is now largely structure-based. Different compounds with particular pharmacological profiles have been identified and reviewed exhaustively [5], as well as their modified counterparts such as radiolabeled, covalent or bivalent ligands [301]. We will focus here on a particular class of labeled compounds, fluorescent ligands. Indeed, such compounds are of major interest when developing ligand-binding assays with isolated receptors [302]. A full series of fluorescent A_{2A} -selective ligands have been devised to be used in native membranes or cell-based systems [303]. These include agonists such as the initial compound FITC-APEC that was derived from the conjugation of the A_{2A} agonist APEC to fluorescein [304]. An alternative version of FITC-APEC, with Alexa Fluor 488 as the fluorophore instead of fluorescein, has been more recently used to study the kinetics of A_{2A} redistribution in living cells [305]. Besides agonists, fluorescent antagonists have also been designed, for instance by coupling the selective A_{2A} antagonist preladenant to Alexa Fluor 647 [306], or using either of the two pharmacophores, 1,3-dipropyl-8-phenylxanthine and triazolo [1,5-c]quinazolin-5-yl)amine [307]. Regarding isolated receptors, the number of studies with fluorescent ligands is much more limited, and these compounds have been mostly used in the context of ligand-binding experiments. A fluorescent antagonist, CA600245, was recently used to assess the ligand-binding ability of A_{2A} encapsulated into SMALPs generated by solubilizing membranes of *P. pastoris* expressing A_{2A} with the SMA polymer. In this case, the fluorescent ligand comprised the antagonist xanthine amine moiety linked to a red BODIPY630/650 fluorophore (Figure 11A). The resulting study demonstrated that A_{2A} in SMALPS is a powerful model system to characterize ligand-receptor complexes when used in combination with fluorescence spectroscopy [111]. More recently, the fluorescent compound FITC-APEC described above (Figure 11B) was used to provide an experimental and theoretical framework for monitoring ligand-binding kinetics to the A_{2A} in detergent micelles under ligand depletion conditions using fluorescence anisotropy measurements [308]. Taken together, the availability of different fluorescent probes from distinct pharmacological classes now provides a toolbox of labeled ligands to develop radioactive-free ligand-binding assays with purified receptors, as well as to probe the pharmacological behavior of A_{2A} isolated in detergent micelles and nanodiscs.

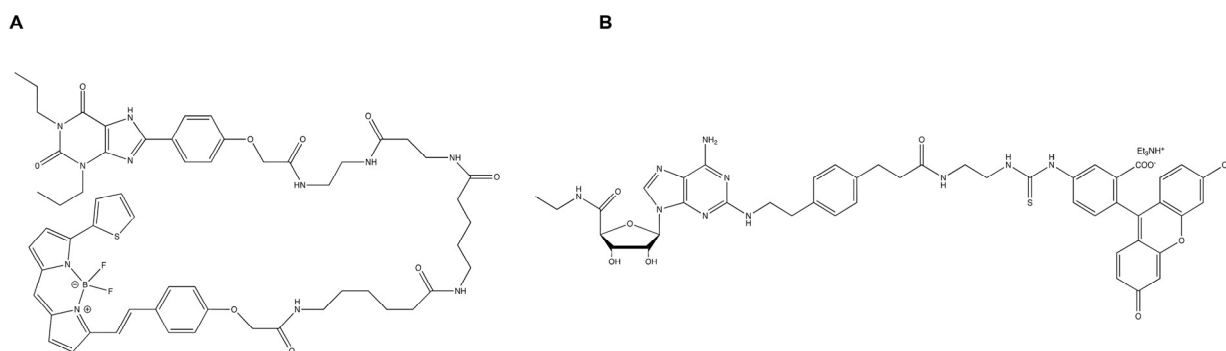


Figure 11. Chemical structures of the fluorescent ligands used for in vitro assays with purified A_{2A} . (A) CA600245 was used to assess the functional properties of the isolated receptor in SMALPS [111]. (B) FITC-APEC was used to monitor the ligand-binding kinetics to the A_{2A} in detergent micelles [308].

9.5. Structure-Based Ligand Discovery for A_{2A}

Structure-based methods in drug discovery comprise a variety of different computer-assisted strategies, including either ligand- or structure-based virtual screening approaches, as well as pharmacophore-based, fragment-based and comparative molecular field anal-

ysis [309–312]. These methods are widely used to identify novel ligands exerting either an agonistic or an antagonistic effect on a specific protein target and they also have had considerable impact on the drug-discovery-related pharmacology of A_{2A}. There are several examples in the literature that illustrate the practical application of structural information for the discovery of A_{2A} antagonists which have been reviewed by Jazayeri et al. [313]. Furthermore, molecular docking has also been employed for the design of either A_{2A} agonists or PDE10A inhibitors, and has been a recent focus in the search of novel antiproliferative agents [314]. In addition, MD simulations have been used extensively to investigate the conformational dynamics at the A_{2A} or PDE10A [315–321]. In a recent paper, Kalash et al. studied a novel structure-based method for identifying ligands that activate A_{2A}, while simultaneously inhibiting the PDE10A [322]. The focus of this approach was on the key interacting residues that are reported in the literature to discriminate between agonist and antagonist activity of A_{2A} ligands [52,57,64,68]. It has been postulated that the motion of the residue Val84 in the TM3, upon A_{2A} ligand binding, might discriminate between agonist and antagonist activity, which had not been studied previously by any of the MD simulations of A_{2A} [64,323]. Kalash et al. show that the motion of this residue can be used as a conformational descriptor for the characterization of receptor activation by A_{2A} ligands [322]. Furthermore, it is suggested that a selective interaction of ligands with His²⁵⁰ accounts for A_{2A} sub-type selectivity. This interaction has only been detected in co-crystal structures with selective A_{2A} agonists, such as CGS21680 [57] and UK432097 [58]. Moreover, MD simulation analysis demonstrated that this residue undergoes conformational changes only when selective A_{2A} agonists are bound and not when non-selective agonists bind to A_{2A} [322]. Additionally, the analysis of the MD trajectories highlighted the motion of Val84 in TM3 as an essential requirement for A_{2A} activation, which is a strong indicator for the potential application of structure-based methods in the discovery of novel A_{2A} ligands.

9.6. Fragment-Based Ligand Discovery for A_{2A}

Fragment-based ligand design is a well-established approach for the generation of novel protein inhibitors or modulators [324]. In a first step, fragment hits are identified from the screening of a fragment library against the protein target. In a second step, selected fragments are iteratively optimized using three-dimensional structures of the protein–ligand complex until a high-affinity compound is obtained. Fragment libraries typically regroup thousands of highly soluble small molecules with low molecular weight (<300 g/mol). Fragment screening using biophysical methods has been achieved for A_{2A} using target immobilized NMR screening [325], SPR [134], STD-NMR [91] and Affinity Mass Spectrometry [293] with A_{2A} samples prepared in detergents, or weak-affinity chromatography with A_{2A} prepared in nanodiscs [77]. Whereas main studies are conducted with the wild-type receptor, a thermostabilized receptor was used in the study reported by Congreve et al. [134]. All techniques, except for STD-NMR, require the immobilization of A_{2A} onto a solid support. To avoid the identification of false positive hits, the binding signal observed in the presence of A_{2A} is compared to the one observed in the absence of A_{2A} or in the presence of the receptor blocked with a reference ligand. In addition to experimental biophysical screening, fragment-based computer-aided design was reported by several groups [293,311,326–329]. To probe the conformational changes in the receptor upon fragment-based ligand binding, serial crystallography could play a key role in the future as a method for obtaining high-resolution conformational dynamics of protein–inhibitor complexes.

9.7. Mass Spectrometry-Based Ligand Screening

Among the available modern screening methods described in this Section, affinity selection-mass spectrometry (AS-MS) is emerging as a technique of choice for the identification and confirmation of binders from large compound libraries, due to its high throughput capacity. AS-MS has been applied by the group of Shui to find new A_{2A} receptor binders [106]. As a proof of concept, they first developed AS-MS using a small model

library consisting of a fifteen-compound mixture including eight known binders (such as adenosine, NECA, caffeine or ZM241385) and A_{2A} reconstituted in nanodiscs. All the known compounds were identified as positive binders from the AS-MS screen, validating the set-up for hit discovery on GPCRs. They next highlighted the potential of AS-MS to screen large compound libraries (approximately 20,000 compounds) on A_{2A} in one pool, leading to the discovery of three new antagonist compounds [141]. Interestingly, affinity MS experiments were successfully performed on a purified A_{2A} immobilized on nickel agarose beads as well as on the receptor embedded in cell-membranes. More recently [293], the same authors adapted their methodology to screen a library of 1100 fragments with the purpose of finding new NAMs of A_{2A} . Among the 28 initial hits, 17 have been confirmed to be specific binders at A_{2A} , and 9 as strong binders by SPR. These nine binders were then analyzed using affinity mass spectrometry in the presence of A_{2A} , free or bound to ZM241385. Thanks to an extensive optimization of immobilization and incubation conditions (see scheme in Figure 12), their method was initially validated as it allowed the detection of three known weak antagonists with binding affinity above 20 μ M (theophylline, NG-52 and LH-846). The six fragments that retained 50% binding in the presence of ZM241385 were further tested in competition with HMA (5-(N,N-hexamethylene)amiloride), one of the most potent NAMs of A_{2A} . Two new HMA-competitive binders were discovered using this approach and confirmed as NAMs using a competitive radioligand binding assay, molecular docking and 19 F-labeled protein NMR experiments [293].

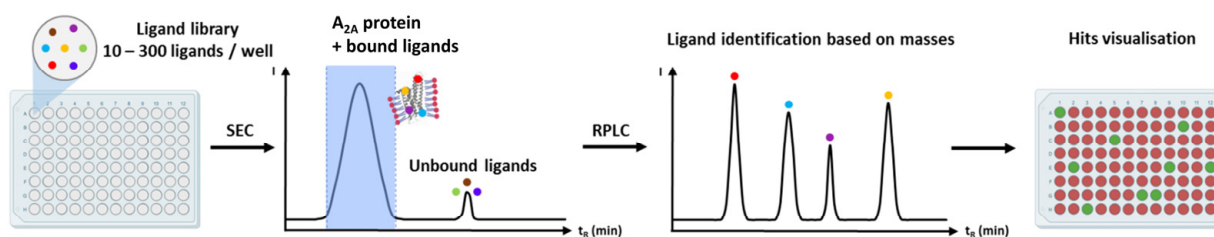


Figure 12. Schematic workflow of modern affinity selection mass spectrometry (AS-MS) used for A_{2A} hit identification. Additionally, called automated ligand identification system (ALIS), this technique is performed to identify small molecules that engage a specific target, such as A_{2A} . The protein of interest is incubated with a mixture of compounds, each with unique molar mass. The protein–ligand complexes are separated from unbound compounds using SEC. The protein peak (depicted in blue on the figure) is then directed to a reverse-phase liquid chromatography (RPLC) column for dissociation, desalting, and elution of any ligands into an ESI-MS system for identification. Alternatively to this 2D-LC set-up, an ultracentrifugation or an offline gel filtration step can be performed prior to LC-MS. This high-throughput screening technique enables hit identification for large collections of ligands.

10. Concluding Remarks and Perspectives

As the isolation of A_{2A} and its reconstitution in an active form were achieved relatively early compared to most other GPCRs, its handling became more streamlined compared to other GPCRs. Accordingly, A_{2A} has become a highly amenable model for the application and development of novel innovative approaches for the biophysical characterization of GPCRs. Moreover, A_{2A} is also a very important drug target in several human disease areas, and it goes without saying that the accumulation of knowledge of the structure and function of A_{2A} was also driven by the desire to discover novel and more specific pharmacological compounds. Indeed, the progress made in many of the cellular biology and cellular chemistry domains, together with the *in silico* approach of modelling the receptor structures and the three-dimensional structure of ligands, lead to the discovery of a series of novel ligands including compounds containing novel chemical scaffolds. Nonetheless, it remains to be seen if this knowledge will be transferable to the studies of other, more challenging to obtain GPCRs, especially from other families.

Many other areas could be further explored for A_{2A} that might be more advanced in the context of other receptors. For example, the structural effects of the variety of agonists existing for this receptor should be further investigated by cryo-EM, possibly also in complex with different G proteins, to investigate the possibility of biased agonism at this receptor and how it can be explained at the structural level. Similarly, the structural effects on signal transduction of allosteric modulators are still largely unknown for the A_{2A} receptor. To further understand these mechanisms, as well as the effects of subtle variations of the agonists on the structure of the G-protein binding site, dynamics kinetics need to further be studied. To this end, single molecule studies, molecular dynamics and also ensemble NMR studies will be invaluable.

While studies of the isolated receptor are crucial for understanding the mechanisms of signal transduction, it is of the utmost importance to also study receptors in their cellular context. Indeed, the progress recently made in the chemical alkylation of proteins in situ (for instance by click chemistry) might eventually allow high resolution studies of the receptor in a “native” milieu, surrounded by its “natural” neighbors in A_{2A}-richly expressing membranes isolated from human tissues. However, the prerequisite of such experimental endeavors will be to introduce a handle at the receptor that can be recognized specifically by the chemical reaction. Combining the use of stem cells with the expression of such handles in a neutral genomic area of the transgene could provide a closer model of the physiological situation.

Finally, A_{2A} has been proven to be an excellent model in many cases of biophysical, structural and even pharmacological characterizations of a GPCR. This is probably due to the relative easiness with which the receptor has been cloned, expressed and isolated, together with a natural stability that helped keep it active once purified.

Funding: The following funding sources are acknowledged. LJC: ANR UnderPressure (ANR-22-CE29-0020-03); ANR DYNALAND; Laboratoire d’Excellence (LabEx) DYNAMO (ANR-11-LABX-0011); Equipements d’Excellence (EQUIPEX) CACSICE (ANR-11-EQPX-0008); SC: CNRS, the University of Strasbourg, the Agence Nationale de la Recherche, and the French Proteomic Infrastructure (ProFI; ANR-10-INBS-08-03); GL: Switch-On (ANR-20-CE11-0019); IK, CD & RW: ANR-21-CE29-0012-02; CVB: AFM-Téléthon #23207.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

A_{2A}, adenosine A_{2A} receptor; AC, affinity chromatography; AFM, atomic force microscopy; AS-MS, affinity selection-mass spectrometry; BRIL, apo-cytochrome b562RIL; CABD, calcium binding domain; CaM, calmodulin; cryo-EM, electron cryo-microscopy; cryo-ET, cryo-electron tomography; DDM, n-dodecyl-β-D-maltopyranoside; DHA, docosahexaenoic acid; EXSY, exchange spectroscopy; FRET, fluorescence resonance energy transfer; GFP, green fluorescent protein; HDX, H/D exchange mass spectrometry; ITC, isothermal titration calorimetry; LCP, lipid cubic phase; LMNG, lauryl maltose neopentyl glycol; mBBR, monobromobimane; MS, mass spectrometry; MSP, membrane scaffold proteins; NAM, negative allosteric modulators; NDs, nanodiscs; NOESY, nuclear Overhauser effect spectroscopy; PAM, positive allosteric modulators; PGD, *petunia* glutamate decarboxylase; PI, Phosphatidylinositol; PQA, 3-pyrrolidin-1-ylquinoxalin-2-amine; PS, Phosphatidylserines; SEC, size-exclusion chromatography; Sf9, *Spodoptera frugiperda*; SMALPs, SMA lipid disc particles; SPA, Scintillation Proximity Assay; SPR, surface plasmon resonance; StaR, thermostabilized; STD, saturation transfer difference; T4L, T4 Lysozyme; TM, GPCR transmembrane domain; ULVs, unilamellar lipid vesicles; XAC, xanthine amine congener; XFEL, X-ray free-electron lasers; XL-MS, Cross-linking mass spectrometry.

References

1. Fredholm, B.B.; Frenguelli, B.G.; Hills, R.; IJzerman, A.P.; Jacobson, K.A.; Klotz, K.-N.; Linden, J.; Müller, C.E.; Schwabe, U.; Stiles, G.L. Adenosine receptors in GtoPdb v.2021.2. *IUPHAR/BPS Guide Pharmacol. CITE* **2021**, 2021. [[CrossRef](#)]
2. Blackmore, R.V.; Linn, S. Partial purification and characterization of four endodeoxyribonuclease activities from *Escherichia coli* K-12. *Nucleic Acids Res.* **1974**, *1*, 1–17. [[CrossRef](#)]
3. Pierce, K.D.; Furlong, T.J.; Selbie, L.A.; Shine, J. Molecular cloning and expression of an adenosine A_{2b} receptor from human brain. *Biochem. Biophys. Res. Commun.* **1992**, *187*, 86–93. [[CrossRef](#)] [[PubMed](#)]
4. Michalke, K.; Huyghe, C.; Lichière, J.; Gravière, M.-E.; Siponen, M.; Sciara, G.; Lepaul, I.; Wagner, R.; Magg, C.; Rudolph, R.; et al. Mammalian G protein-coupled receptor expression in *Escherichia coli*: II. Refolding and biophysical characterization of mouse cannabinoid receptor 1 and human parathyroid hormone receptor 1. *Anal. Biochem.* **2010**, *401*, 74–80. [[CrossRef](#)] [[PubMed](#)]
5. IJzerman, A.P.; Jacobson, K.A.; Müller, C.E.; Cronstein, B.N.; Cunha, R.A. International Union of Basic and Clinical Pharmacology. CXII: Adenosine Receptors: A Further Update. *Pharmacol. Rev.* **2022**, *74*, 340–372. [[CrossRef](#)]
6. Yu, L.; Frith, M.C.; Suzuki, Y.; Peterfreund, R.A.; Gearan, T.; Sugano, S.; Schwarzschild, M.A.; Weng, Z.; Fink, J.S.; Chen, J.-F. Characterization of genomic organization of the adenosine A_{2A} receptor gene by molecular and bioinformatics analyses. *Brain Res.* **2004**, *1000*, 156–173. [[CrossRef](#)] [[PubMed](#)]
7. Koshiba, M.; Rosin, D.L.; Hayashi, N.; Linden, J.; Sitkovsky, M.V. Patterns of A_{2A} extracellular adenosine receptor expression in different functional subsets of human peripheral T cells. Flow cytometry studies with anti-A_{2A} receptor monoclonal antibodies. *Mol. Pharmacol.* **1999**, *55*, 614–624.
8. Cooper, J.A.; Hill, S.J.; Alexander, S.P.; Rubin, P.C.; Horn, E.H. Adenosine receptor-induced cyclic AMP generation and inhibition of 5-hydroxytryptamine release in human platelets. *Br. J. Clin. Pharmacol.* **1995**, *40*, 43–50. [[CrossRef](#)]
9. Schulte, G.; Fredholm, B.B. Human adenosine A₁, A_{2A}, A_{2B}, and A₃ receptors expressed in Chinese hamster ovary cells all mediate the phosphorylation of extracellular-regulated kinase $\frac{1}{2}$. *Mol. Pharmacol.* **2000**, *58*, 477–482. [[CrossRef](#)]
10. Yang, S.N.; Dasgupta, S.; Lledo, P.M.; Vincent, J.D.; Fuxe, K. Reduction of dopamine D₂ receptor transduction by activation of adenosine A_{2a} receptors in stably A_{2a}/D₂ (long-form) receptor co-transfected mouse fibroblast cell lines: Studies on intracellular calcium levels. *Neuroscience* **1995**, *68*, 729–736. [[CrossRef](#)]
11. Ikeuchi, Y.; Nishizaki, T.; Mori, M.; Okada, Y. Adenosine activates the K⁺ channel and enhances cytosolic Ca²⁺ release via a P_{2Y} purinoceptor in hippocampal neurons. *Eur. J. Pharmacol.* **1996**, *304*, 191–199. [[CrossRef](#)] [[PubMed](#)]
12. Wang, Q.; Zhou, F.M. cAMP-producing chemogenetic and adenosine A_{2a} receptor activation inhibits the inwardly rectifying potassium current in striatal projection neurons. *Neuropharmacology* **2019**, *148*, 229–243. [[CrossRef](#)]
13. Varani, K.; Gessi, S.; Dalpiaz, A.; Borea, P.A. Pharmacological and biochemical characterization of purified A_{2a} adenosine receptors in human platelet membranes by [³H]-CGS 21680 binding. *Br. J. Pharmacol.* **1996**, *117*, 1693–1701. [[CrossRef](#)] [[PubMed](#)]
14. Buckley, N.; Burnstock, G. Autoradiographic demonstration of peripheral adenosine binding sites using [³H]NECA. *Brain Res.* **1983**, *269*, 374–377. [[CrossRef](#)] [[PubMed](#)]
15. Schütz, W.; Brugger, G. Characterization of [³H]-adenosine binding to media membranes of hog carotid arteries. *Pharmacology* **1982**, *24*, 26–34. [[CrossRef](#)]
16. Ji, X.D.; Stiles, G.L.; van Galen, P.J.; Jacobson, K.A. Characterization of human striatal A₂-adenosine receptors using radioligand binding and photoaffinity labeling. *J. Recept. Res.* **1992**, *12*, 149–169. [[CrossRef](#)]
17. Alexander, S.P.; Millns, P.J. [³H]ZM241385—An antagonist radioligand for adenosine A_{2A} receptors in rat brain. *Eur. J. Pharmacol.* **2001**, *411*, 205–210. [[CrossRef](#)]
18. Yu, J.C.; Lin, G.; Field, J.J.; Linden, J. Induction of antiinflammatory purinergic signaling in activated human iNKT cells. *JCI Insight* **2018**, *3*, e91954. [[CrossRef](#)]
19. Klammt, C.; Schwarz, D.; Eifler, N.; Engel, A.; Piehler, J.; Haase, W.; Hahn, S.; Dötsch, V.; Bernhard, F. Cell-free production of G protein-coupled receptors for functional and structural studies. *J. Struct. Biol.* **2007**, *158*, 482–493. [[CrossRef](#)]
20. Kögler, L.M.; Stichel, J.; Beck-Sickinger, A.G. Structural investigations of cell-free expressed G protein-coupled receptors. *Biol. Chem.* **2019**, *401*, 97–116. [[CrossRef](#)]
21. Banères, J.-L.; Popot, J.-L.; Mouillac, B. New advances in production and functional folding of G-protein-coupled receptors. *Trends Biotechnol.* **2011**, *29*, 314–322. [[CrossRef](#)]
22. André, N.; Cherouati, N.; Prual, C.; Steffan, T.; Zeder-Lutz, G.; Magnin, T.; Pattus, F.; Michel, H.; Wagner, R.; Reinhart, C. Enhancing functional production of G protein-coupled receptors in *Pichia pastoris* to levels required for structural studies via a single expression screen. *Protein Sci.* **2006**, *15*, 1115–1126. [[CrossRef](#)]
23. Krettler, C.; Reinhart, C.; Bevans, C.G. Expression of GPCRs in *Pichia pastoris* for structural studies. *Methods Enzymol.* **2013**, *520*, 1–29. [[CrossRef](#)]
24. Blocker, K.M.; Britton, Z.T.; Naranjo, A.N.; McNeely, P.M.; Young, C.L.; Robinson, A.S. Recombinant G protein-coupled receptor expression in *Saccharomyces cerevisiae* for protein characterization. *Methods Enzymol.* **2015**, *556*, 165–183. [[CrossRef](#)]
25. Errey, J.C.; Fiez-Vandal, C. Production of membrane proteins in industry: The example of GPCRs. *Protein Expr. Purif.* **2020**, *169*, 105569. [[CrossRef](#)] [[PubMed](#)]
26. Panneels, V.; Kock, I.; Krijnse-Locker, J.; Rezaoui, M.; Sinning, I. Drosophila photoreceptor cells exploited for the production of eukaryotic membrane proteins: Receptors, transporters and channels. *PLoS ONE* **2011**, *6*, e18478. [[CrossRef](#)]

27. Zhang, L.; Salom, D.; He, J.; Okun, A.; Ballesteros, J.; Palczewski, K.; Li, N. Expression of functional G protein-coupled receptors in photoreceptors of transgenic *Xenopus laevis*. *Biochemistry* **2005**, *44*, 14509–14518. [[CrossRef](#)] [[PubMed](#)]
28. Kajikawa, M.; Sasaki, K.; Wakimoto, Y.; Toyooka, M.; Motohashi, T.; Shimojima, T.; Takeda, S.; Park, E.Y.; Maenaka, K. Efficient silkworm expression of human GPCR (nociceptin receptor) by a *Bombyx mori* bacmid DNA system. *Biochem. Biophys. Res. Commun.* **2009**, *385*, 375–379. [[CrossRef](#)]
29. Michalke, K.; Gravière, M.-E.; Huyghe, C.; Vincentelli, R.; Wagner, R.; Pattus, F.; Schroeder, K.; Oschmann, J.; Rudolph, R.; Cambillau, C.; et al. Mammalian G-protein-coupled receptor expression in *Escherichia coli*: I. High-throughput large-scale production as inclusion bodies. *Anal. Biochem.* **2009**, *386*, 147–155. [[CrossRef](#)] [[PubMed](#)]
30. Weiss, H.M.; Grishammer, R. Purification and characterization of the human adenosine A_{2a} receptor functionally expressed in *Escherichia coli*. *Eur. J. Biochem.* **2002**, *269*, 82–92. [[CrossRef](#)] [[PubMed](#)]
31. Magnani, F.; Shibata, Y.; Serrano-Vega, M.J.; Tate, C.G. Co-evolving stability and conformational homogeneity of the human adenosine A_{2a} receptor. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 10744–10749. [[CrossRef](#)]
32. Lebon, G.; Bennett, K.; Jazayeri, A.; Tate, C.G. Thermostabilisation of an Agonist-Bound Conformation of the Human Adenosine A_{2A} Receptor. *J. Mol. Biol.* **2011**, *409*, 298–310. [[CrossRef](#)]
33. Niebauer, R.T.; Robinson, A.S. Exceptional total and functional yields of the human adenosine (A_{2a}) receptor expressed in the yeast *Saccharomyces cerevisiae*. *Protein Expr. Purif.* **2006**, *46*, 204–211. [[CrossRef](#)]
34. O'Malley, M.A.; Lazarova, T.; Britton, Z.T.; Robinson, A.S. High-level expression in *Saccharomyces cerevisiae* enables isolation and spectroscopic characterization of functional human adenosine A_{2a} receptor. *J. Struct. Biol.* **2007**, *159*, 166–178. [[CrossRef](#)] [[PubMed](#)]
35. Schonenbach, N.S.; Rieth, M.D.; Han, S.; O'Malley, M.A. Adenosine A_{2a} receptors form distinct oligomers in protein detergent complexes. *FEBS Lett.* **2016**, *590*, 3295–3306. [[CrossRef](#)] [[PubMed](#)]
36. Fraser, N.J. Expression and functional purification of a glycosylation deficient version of the human adenosine 2a receptor for structural studies. *Protein Expr. Purif.* **2006**, *49*, 129–137. [[CrossRef](#)] [[PubMed](#)]
37. Singh, S.; Gras, A.; Fiez-Vandal, C.; Ruprecht, J.; Rana, R.; Martinez, M.; Strange, P.G.; Wagner, R.; Byrne, B. Large-scale functional expression of WT and truncated human adenosine A_{2A} receptor in *Pichia pastoris* bioreactor cultures. *Microb. Cell Fact.* **2008**, *7*, 28. [[CrossRef](#)]
38. Hino, T.; Arakawa, T.; Iwanari, H.; Yurugi-Kobayashi, T.; Ikeda-Suno, C.; Nakada-Nakura, Y.; Kusano-Arai, O.; Weyand, S.; Shimamura, T.; Nomura, N.; et al. G-protein-coupled receptor inactivation by an allosteric inverse-agonist antibody. *Nature* **2012**, *482*, 237–240. [[CrossRef](#)]
39. Jamshad, M.; Charlton, J.; Lin, Y.-P.; Routledge, S.J.; Bawa, Z.; Knowles, T.J.; Overduin, M.; Dekker, N.; Dafforn, T.R.; Bill, R.M.; et al. G-protein coupled receptor solubilization and purification for biophysical analysis and functional studies, in the total absence of detergent. *Biosci. Rep.* **2015**, *35*, e00188. [[CrossRef](#)]
40. Lengger, B.; Jensen, M.K. Engineering G protein-coupled receptor signalling in yeast for biotechnological and medical purposes. *FEMS Yeast Res.* **2020**, *20*, foz087. [[CrossRef](#)]
41. Shiroishi, M.; Tsujimoto, H.; Makyio, H.; Asada, H.; Yurugi-Kobayashi, T.; Shimamura, T.; Murata, T.; Nomura, N.; Haga, T.; Iwata, S.; et al. Platform for the rapid construction and evaluation of GPCRs for crystallography in *Saccharomyces cerevisiae*. *Microb. Cell Fact.* **2012**, *11*, 78. [[CrossRef](#)]
42. Berthelme, N.; Strege, A.; Bunting, S.E.; Dowell, S.J.; Byrne, B. Arginine 199 and leucine 208 have key roles in the control of adenosine A_{2A} receptor signalling function. *PLoS ONE* **2014**, *9*, e89613. [[CrossRef](#)]
43. Mitsumoto, M.; Sugaya, K.; Kazama, K.; Nakano, R.; Kosugi, T.; Murata, T.; Koga, N. State-Targeting Stabilization of Adenosine A_{2A} Receptor by Fusing a Custom-Made de Novo Designed α -Helical Protein. *Int. J. Mol. Sci.* **2021**, *22*, 12906. [[CrossRef](#)]
44. Wang, X.; van Westen, G.J.P.; Heitman, L.H.; IJzerman, A.P. G protein-coupled receptors expressed and studied in yeast. The adenosine receptor as a prime example. *Biochem. Pharmacol.* **2021**, *187*, 114370. [[CrossRef](#)]
45. Meltzer, M.; Zvagelsky, T.; Hadad, U.; Papo, N.; Engel, S. Yeast-based directed-evolution for high-throughput structural stabilization of G protein-coupled receptors (GPCRs). *Sci. Rep.* **2022**, *12*, 8657. [[CrossRef](#)]
46. Yamamoto, T.; Yasuda, S.; Kasai, R.S.; Nakano, R.; Hikiri, S.; Sugaya, K.; Hayashi, T.; Ogasawara, S.; Shiroishi, M.; Fujiwara, T.K.; et al. A methodology for creating mutants of G-protein coupled receptors stabilized in active state by combining statistical thermodynamics and evolutionary molecular engineering. *Protein Sci.* **2022**, *31*, e4425. [[CrossRef](#)] [[PubMed](#)]
47. Eddy, M.T.; Lee, M.-Y.; Gao, Z.-G.; White, K.L.; Didenko, T.; Horst, R.; Audet, M.; Stanczak, P.; McClary, K.M.; Han, G.W.; et al. Allosteric Coupling of Drug Binding and Intracellular Signaling in the A_{2A} Adenosine Receptor. *Cell* **2018**, *172*, 68–80.e12. [[CrossRef](#)]
48. Ye, L.; van Eps, N.; Zimmer, M.; Ernst, O.P.; Prosser, R.S. Activation of the A_{2A} adenosine G-protein-coupled receptor by conformational selection. *Nature* **2016**, *533*, 265–268. [[CrossRef](#)] [[PubMed](#)]
49. Mizumura, T.; Kondo, K.; Kurita, M.; Kofuku, Y.; Natsume, M.; Imai, S.; Shiraiishi, Y.; Ueda, T.; Shimada, I. Activation of adenosine A_{2A} receptor by lipids from docosahexaenoic acid revealed by NMR. *Sci. Adv.* **2020**, *6*, eaay8544. [[CrossRef](#)] [[PubMed](#)]
50. Huang, S.K.; Pandey, A.; Tran, D.P.; Villanueva, N.L.; Kitao, A.; Sunahara, R.K.; Sljoka, A.; Prosser, R.S. Delineating the conformational landscape of the adenosine A_{2A} receptor during G protein coupling. *Cell* **2021**, *184*, 1884–1894.e14. [[CrossRef](#)] [[PubMed](#)]

51. Jaakola, V.-P.; Griffith, M.T.; Hanson, M.A.; Cherezov, V.; Chien, E.Y.T.; Lane, J.R.; IJzerman, A.P.; Stevens, R.C. The 2.6 angstrom crystal structure of a human A_{2A} adenosine receptor bound to an antagonist. *Science* **2008**, *322*, 1211–1217. [[CrossRef](#)] [[PubMed](#)]
52. Lebon, G.; Warne, T.; Edwards, P.C.; Bennett, K.; Langmead, C.J.; Leslie, A.G.W.; Tate, C.G. Agonist-bound adenosine A_{2A} receptor structures reveal common features of GPCR activation. *Nature* **2011**, *474*, 521–525. [[CrossRef](#)]
53. Carpenter, B.; Nehmé, R.; Warne, T.; Leslie, A.G.W.; Tate, C.G. Structure of the adenosine A_{2A} receptor bound to an engineered G protein. *Nature* **2016**, *536*, 104–107. [[CrossRef](#)]
54. Miyagi, H.; Asada, H.; Suzuki, M.; Takahashi, Y.; Yasunaga, M.; Suno, C.; Iwata, S.; Saito, J.-I. The discovery of a new antibody for BRIL-fused GPCR structure determination. *Sci. Rep.* **2020**, *10*, 11669. [[CrossRef](#)]
55. Claff, T.; Klapschinski, T.A.; Tiruttani Subhramanyam, U.K.; Vaaßen, V.J.; Schlegel, J.G.; Vielmuth, C.; Voß, J.H.; Labahn, J.; Müller, C.E. Single Stabilizing Point Mutation Enables High-Resolution Co-Crystal Structures of the Adenosine A_{2A} Receptor with Preladenant Conjugates. *Angew. Chem. Int. Ed Engl.* **2022**, *61*, e202115545. [[CrossRef](#)]
56. Liu, W.; Chun, E.; Thompson, A.A.; Chubukov, P.; Xu, F.; Katritch, V.; Han, G.W.; Roth, C.B.; Heitman, L.H.; IJzerman, A.P.; et al. Structural basis for allosteric regulation of GPCRs by sodium ions. *Science* **2012**, *337*, 232–236. [[CrossRef](#)] [[PubMed](#)]
57. Lebon, G.; Edwards, P.C.; Leslie, A.G.W.; Tate, C.G. Molecular Determinants of CGS21680 Binding to the Human Adenosine A_{2A} Receptor. *Mol. Pharmacol.* **2015**, *87*, 907–915. [[CrossRef](#)] [[PubMed](#)]
58. White, K.L.; Eddy, M.T.; Gao, Z.-G.; Han, G.W.; Lian, T.; Deary, A.; Patel, N.; Jacobson, K.A.; Katritch, V.; Stevens, R.C. Structural Connection between Activation Microswitch and Allosteric Sodium Site in GPCR Signaling. *Structure* **2018**, *26*, 259–269.e5. [[CrossRef](#)]
59. McGraw, C.; Koretz, K.S.; Oseid, D.; Lyman, E.; Robinson, A.S. Cholesterol Dependent Activity of the Adenosine A_{2A} Receptor Is Modulated via the Cholesterol Consensus Motif. *Molecules* **2022**, *27*, 3529. [[CrossRef](#)]
60. Yeliseev, A.; van den Berg, A.; Zoubak, L.; Hines, K.; Stepnowski, S.; Williston, K.; Yan, W.; Gawrisch, K.; Zmuda, J. Thermostability of a recombinant G protein-coupled receptor expressed at high level in mammalian cell culture. *Sci. Rep.* **2020**, *10*, 16805. [[CrossRef](#)]
61. Zhang, K.; Wu, H.; Hoppe, N.; Manglik, A.; Cheng, Y. Fusion protein strategies for cryo-EM study of G protein-coupled receptors. *Nat. Commun.* **2022**, *13*, 4366. [[CrossRef](#)]
62. Bocquet, N.; Kohler, J.; Hug, M.N.; Kusznir, E.A.; Rufer, A.C.; Dawson, R.J.; Hennig, M.; Ruf, A.; Huber, W.; Huber, S. Real-time monitoring of binding events on a thermostabilized human A_{2A} receptor embedded in a lipid bilayer by surface plasmon resonance. *Biochim. Biophys. Acta* **2015**, *1848*, 1224–1233. [[CrossRef](#)] [[PubMed](#)]
63. Suzuki, N.; Takamuku, Y.; Asakawa, T.; Inai, M.; Hino, T.; Iwata, S.; Kan, T.; Murata, T. An efficient screening method for purifying and crystallizing membrane proteins using modified clear-native PAGE. *Anal. Biochem.* **2018**, *548*, 7–14. [[CrossRef](#)] [[PubMed](#)]
64. Carpenter, B.; Lebon, G. Human Adenosine A_{2A} Receptor: Molecular Mechanism of Ligand Binding and Activation. *Front. Pharmacol.* **2017**, *8*, 898. [[CrossRef](#)]
65. Cui, M.; Zhou, Q.; Xu, Y.; Weng, Y.; Yao, D.; Zhao, S.; Song, G. Crystal structure of a constitutive active mutant of adenosine A_{2A} receptor. *IUCr* **2022**, *9*, 333–341. [[CrossRef](#)] [[PubMed](#)]
66. Lundstrom, K.; Wagner, R.; Reinhart, C.; Desmyter, A.; Cherouati, N.; Magnin, T.; Zeder-Lutz, G.; Courtot, M.; Prual, C.; André, N.; et al. Structural genomics on membrane proteins: Comparison of more than 100 GPCRs in 3 expression systems. *J. Struct. Funct. Genom.* **2006**, *7*, 77–91. [[CrossRef](#)] [[PubMed](#)]
67. Clark, L.D.; Dikiy, I.; Chapman, K.; Rödröm, K.E.; Aramini, J.; LeVine, M.V.; Khelashvili, G.; Rasmussen, S.G.; Gardner, K.H.; Rosenbaum, D.M. Ligand modulation of sidechain dynamics in a wild-type human GPCR. *eLife* **2017**, *6*, e28505. [[CrossRef](#)] [[PubMed](#)]
68. Xu, F.; Wu, H.; Katritch, V.; Han, G.W.; Jacobson, K.A.; Gao, Z.-G.; Cherezov, V.; Stevens, R.C. Structure of an agonist-bound human A_{2A} adenosine receptor. *Science* **2011**, *332*, 322–327. [[CrossRef](#)] [[PubMed](#)]
69. Cheng, R.K.Y.; Segala, E.; Robertson, N.; Deflorian, F.; Doré, A.S.; Errey, J.C.; Fiez-Vandal, C.; Marshall, F.H.; Cooke, R.M. Structures of Human A₁ and A_{2A} Adenosine Receptors with Xanthines Reveal Determinants of Selectivity. *Structure* **2017**, *25*, 1275–1285.e4. [[CrossRef](#)]
70. Jespers, W.; Verdon, G.; Azuaje, J.; Majellaro, M.; Keränen, H.; García-Mera, X.; Congreve, M.; Deflorian, F.; de Graaf, C.; Zhukov, A.; et al. X-ray Crystallography and Free Energy Calculations Reveal the Binding Mechanism of A_{2A} Adenosine Receptor Antagonists. *Angew. Chem. Int. Ed Engl.* **2020**, *59*, 16536–16543. [[CrossRef](#)] [[PubMed](#)]
71. Chun, E.; Thompson, A.A.; Liu, W.; Roth, C.B.; Griffith, M.T.; Katritch, V.; Kunken, J.; Xu, F.; Cherezov, V.; Hanson, M.A.; et al. Fusion partner toolchest for the stabilization and crystallization of G protein-coupled receptors. *Structure* **2012**, *20*, 967–976. [[CrossRef](#)] [[PubMed](#)]
72. Doré, A.S.; Robertson, N.; Errey, J.C.; Ng, I.; Hollenstein, K.; Tehan, B.; Hurrell, E.; Bennett, K.; Congreve, M.; Magnani, F.; et al. Structure of the adenosine A_{2A} receptor in complex with ZM241385 and the xanthines XAC and caffeine. *Structure* **2011**, *19*, 1283–1293. [[CrossRef](#)] [[PubMed](#)]
73. Sušac, L.; Eddy, M.T.; Didenko, T.; Stevens, R.C.; Wüthrich, K. A_{2A} adenosine receptor functional states characterized by 19F-NMR. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 12733–12738. [[CrossRef](#)]
74. Huang, S.K.; Almurad, O.; Pejana, R.J.; Morrison, Z.A.; Pandey, A.; Picard, L.-P.; Nitz, M.; Sljoka, A.; Prosser, R.S. Allosteric modulation of the adenosine A_{2A} receptor by cholesterol. *eLife* **2022**, *11*, e73901. [[CrossRef](#)]

75. Thakur, N.; Wei, S.; Ray, A.P.; Lamichhane, R.; Eddy, M.T. Production of human A_{2A}AR in lipid nanodiscs for 19F-NMR and single-molecule fluorescence spectroscopy. *STAR Protoc.* **2022**, *3*, 101535. [[CrossRef](#)] [[PubMed](#)]
76. Choy, B.C.; Cater, R.J.; Mancía, F.; Pryor, E.E. A 10-year meta-analysis of membrane protein structural biology: Detergents, membrane mimetics, and structure determination techniques. *Biochim. Biophys. Acta Biomembr.* **2021**, *1863*, 183533. [[CrossRef](#)]
77. Lecas, L.; Hartmann, L.; Caro, L.; Mohamed-Bouteben, S.; Raingeval, C.; Krimm, I.; Wagner, R.; Dugas, V.; Demesmay, C. Miniaturized weak affinity chromatography for ligand identification of nanodiscs-embedded G-protein coupled receptors. *Anal. Chim. Acta* **2020**, *1113*, 26–35. [[CrossRef](#)]
78. Lee, M.-Y.; Geiger, J.; Ishchenko, A.; Han, G.W.; Barty, A.; White, T.A.; Gati, C.; Batyuk, A.; Hunter, M.S.; Aquila, A.; et al. Harnessing the power of an X-ray laser for serial crystallography of membrane proteins crystallized in lipidic cubic phase. *IUCr* **2020**, *7*, 976–984. [[CrossRef](#)]
79. Segala, E.; Guo, D.; Cheng, R.K.Y.; Bortolato, A.; Deflorian, F.; Doré, A.S.; Errey, J.C.; Heitman, L.H.; IJzerman, A.P.; Marshall, F.H.; et al. Controlling the Dissociation of Ligands from the Adenosine A_{2A} Receptor through Modulation of Salt Bridge Strength. *J. Med. Chem.* **2016**, *59*, 6470–6479. [[CrossRef](#)]
80. Reis, R.I.; Moraes, I. Probing Membrane Protein Assembly into Nanodiscs by In Situ Dynamic Light Scattering: A_{2A} Receptor as a Case Study. *Biology* **2020**, *9*, 400. [[CrossRef](#)]
81. Batyuk, A.; Galli, L.; Ishchenko, A.; Han, G.W.; Gati, C.; Popov, P.A.; Lee, M.-Y.; Stauch, B.; White, T.A.; Barty, A.; et al. Native phasing of x-ray free-electron laser data for a G protein-coupled receptor. *Sci. Adv.* **2016**, *2*, e1600292. [[CrossRef](#)] [[PubMed](#)]
82. Dikiy, I.; Clark, L.D.; Gardner, K.H.; Rosenbaum, D.M. Isotopic Labeling of Eukaryotic Membrane Proteins for NMR Studies of Interactions and Dynamics. *Methods Enzymol.* **2019**, *614*, 37–65. [[CrossRef](#)] [[PubMed](#)]
83. O'Malley, M.A.; Naranjo, A.N.; Lazarova, T.; Robinson, A.S. Analysis of Adenosine A_{2A} Receptor Stability: Effects of Ligands and Disulfide Bonds. *Biochemistry* **2010**, *49*, 9181–9189. [[CrossRef](#)] [[PubMed](#)]
84. Ashok, Y.; Nanekar, R.; Jaakola, V.-P. Defining thermostability of membrane proteins by western blotting. *Protein Eng. Des. Sel.* **2015**, *28*, 539–542. [[CrossRef](#)]
85. Zhang, X.; Stevens, R.C.; Xu, F. The importance of ligands for G protein-coupled receptor stability. *Trends Biochem. Sci.* **2015**, *40*, 79–87. [[CrossRef](#)]
86. Thompson, A.A.; Liu, J.J.; Chun, E.; Wacker, D.; Wu, H.; Cherezov, V.; Stevens, R.C. GPCR stabilization using the bicelle-like architecture of mixed sterol-detergent micelles. *Methods* **2011**, *55*, 310–317. [[CrossRef](#)]
87. O'Malley, M.A.; Helgeson, M.E.; Wagner, N.J.; Robinson, A.S. The morphology and composition of cholesterol-rich micellar nanostructures determine transmembrane protein (GPCR) activity. *Biophys. J.* **2011**, *100*, L11–L13. [[CrossRef](#)]
88. Vagenende, V.; Yap, M.G.S.; Trout, B.L. Mechanisms of protein stabilization and prevention of protein aggregation by glycerol. *Biochemistry* **2009**, *48*, 11084–11096. [[CrossRef](#)]
89. Wittmann, H.-J.; Seifert, R.; Strasser, A. Mathematical analysis of the sodium sensitivity of the human histamine H₃ receptor. *Silico Pharmacol.* **2014**, *2*, 1. [[CrossRef](#)]
90. Ross, P.; Weihofen, W.; Siu, F.; Xie, A.; Katakia, H.; Wright, S.K.; Hunt, I.; Brown, R.K.; Freire, E. Isothermal chemical denaturation to determine binding affinity of small molecules to G-protein coupled receptors. *Anal. Biochem.* **2015**, *473*, 41–45. [[CrossRef](#)]
91. Igonet, S.; Raingeval, C.; Cecon, E.; Pučić-Baković, M.; Lauc, G.; Cala, O.; Baranowski, M.; Perez, J.; Jockers, R.; Krimm, I.; et al. Enabling STD-NMR fragment screening using stabilized native GPCR: A case study of adenosine receptor. *Sci. Rep.* **2018**, *8*, 8142. [[CrossRef](#)]
92. Ukena, D.; Jacobson, K.A.; Kirk, K.L.; Daly, J.W. A [³H]amine congener of 1,3-dipropyl-8-phenylxanthine. *FEBS Lett.* **1986**, *199*, 269–274. [[CrossRef](#)]
93. Ye, L.; Neale, C.; Sljoka, A.; Lyda, B.; Pichugin, D.; Tsuchimura, N.; Larda, S.T.; Pomès, R.; García, A.E.; Ernst, O.P.; et al. Mechanistic insights into allosteric regulation of the A_{2A} adenosine G protein-coupled receptor by physiological cations. *Nat. Commun.* **2018**, *9*, 1372. [[CrossRef](#)] [[PubMed](#)]
94. Nguyen, K.D.Q.; Vigers, M.; Sefah, E.; Seppälä, S.; Hoover, J.P.; Schonenbach, N.S.; Mertz, B.; O'Malley, M.A.; Han, S. Homooligomerization of the human adenosine A_{2A} receptor is driven by the intrinsically disordered C-terminus. *eLife* **2021**, *10*, e66662. [[CrossRef](#)]
95. Jones, A.J.Y.; Gabriel, F.; Tandale, A.; Nietlispach, D. Structure and Dynamics of GPCRs in Lipid Membranes: Physical Principles and Experimental Approaches. *Molecules* **2020**, *25*, 4729. [[CrossRef](#)]
96. Serebryany, E.; Zhu, G.A.; Yan, E.C.Y. Artificial membrane-like environments for in vitro studies of purified G-protein coupled receptors. *Biochim. Biophys. Acta* **2012**, *1818*, 225–233. [[CrossRef](#)] [[PubMed](#)]
97. Overduin, M.; Trieber, C.; Prosser, R.S.; Picard, L.-P.; Sheff, J.G. Structures and Dynamics of Native-State Transmembrane Protein Targets and Bound Lipids. *Membranes* **2021**, *11*, 451. [[CrossRef](#)]
98. Park, S.H.; Das, B.B.; Casagrande, F.; Tian, Y.; Nothnagel, H.J.; Chu, M.; Kiefer, H.; Maier, K.; de Angelis, A.A.; Marassi, F.M.; et al. Structure of the chemokine receptor CXCR1 in phospholipid bilayers. *Nature* **2012**, *491*, 779–783. [[CrossRef](#)] [[PubMed](#)]
99. Duc, N.M.; Du, Y.; Zhang, C.; Lee, S.Y.; Thorsen, T.S.; Kobilka, B.K.; Chung, K.Y. Effective application of bicelles for conformational analysis of G protein-coupled receptors by hydrogen/deuterium exchange mass spectrometry. *J. Am. Soc. Mass Spectrom.* **2015**, *26*, 808–817. [[CrossRef](#)]
100. Schmidt, P.; Bender, B.J.; Kaiser, A.; Gulati, K.; Scheidt, H.A.; Hamm, H.E.; Meiler, J.; Beck-Sickinger, A.G.; Huster, D. Improved in Vitro Folding of the Y2 G Protein-Coupled Receptor into Bicelles. *Front. Mol. Biosci.* **2017**, *4*, 100. [[CrossRef](#)]

101. Brea, R.J.; Cole, C.M.; Lyda, B.R.; Ye, L.; Prosser, R.S.; Sunahara, R.K.; Devaraj, N.K. In Situ Reconstitution of the Adenosine A_{2A} Receptor in Spontaneously Formed Synthetic Liposomes. *J. Am. Chem. Soc.* **2017**, *139*, 3607–3610. [[CrossRef](#)]
102. Bayburt, T.H.; Grinkova, Y.V.; Sligar, S.G. Self-Assembly of Discoidal Phospholipid Bilayer Nanoparticles with Membrane Scaffold Proteins. *Nano Lett.* **2002**, *2*, 853–856. [[CrossRef](#)]
103. Ritchie, T.K.; Grinkova, Y.V.; Bayburt, T.H.; Denisov, I.G.; Zolnerciks, J.K.; Atkins, W.M.; Sligar, S.G. Chapter 11—Reconstitution of membrane proteins in phospholipid bilayer nanodiscs. *Methods Enzymol.* **2009**, *464*, 211–231. [[CrossRef](#)]
104. Lavington, S.; Watts, A. Lipid nanoparticle technologies for the study of G protein-coupled receptors in lipid environments. *Biophys. Rev.* **2020**, *12*, 1287–1302. [[CrossRef](#)]
105. Segala, E.; Errey, J.C.; Fiez-Vandal, C.; Zhukov, A.; Cooke, R.M. Biosensor-based affinities and binding kinetics of small molecule antagonists to the adenosine A_{2A} receptor reconstituted in HDL like particles. *FEBS Lett.* **2015**, *589*, 1399–1405. [[CrossRef](#)] [[PubMed](#)]
106. Ma, J.; Lu, Y.; Wu, D.; Peng, Y.; Loa-Kum-Cheung, W.; Peng, C.; Quinn, R.J.; Shui, W.; Liu, Z.-J. Ligand identification of the adenosine A_{2A} receptor in self-assembled nanodiscs by affinity mass spectrometry. *Anal. Methods* **2017**, *9*, 5851–5858. [[CrossRef](#)]
107. Fredriksson, K.; Lottmann, P.; Hinz, S.; Onila, I.; Shymanets, A.; Harteneck, C.; Müller, C.E.; Griesinger, C.; Exner, T.E. Nanodiscs for INPHARMA NMR Characterization of GPCRs: Ligand Binding to the Human A_{2A} Adenosine Receptor. *Angew. Chem. Int. Ed Engl.* **2017**, *56*, 5750–5754. [[CrossRef](#)]
108. Mary, S.; Damian, M.; Rahmeh, R.; Mouillac, B.; Marie, J.; Granier, S.; Banères, J.-L. Amphipols in G protein-coupled receptor pharmacology: What are they good for? *J. Membr. Biol.* **2014**, *247*, 853–860. [[CrossRef](#)]
109. Zoonens, M.; Popot, J.-L. Amphipols for each season. *J. Membr. Biol.* **2014**, *247*, 759–796. [[CrossRef](#)] [[PubMed](#)]
110. Overduin, M.; Klumperman, B. Advancing membrane biology with poly(styrene-co-maleic acid)-based native nanodiscs. *Eur. Polym. J.* **2019**, *110*, 63–68. [[CrossRef](#)]
111. Grime, R.L.; Goulding, J.; Uddin, R.; Stoddart, L.A.; Hill, S.J.; Poyner, D.R.; Briddon, S.J.; Wheatley, M. Single molecule binding of a ligand to a G-protein-coupled receptor in real time using fluorescence correlation spectroscopy, rendered possible by nano-encapsulation in styrene maleic acid lipid particles. *Nanoscale* **2020**, *12*, 11518–11525. [[CrossRef](#)]
112. Routledge, S.J.; Jamshad, M.; Little, H.A.; Lin, Y.-P.; Simms, J.; Thakker, A.; Spickett, C.M.; Bill, R.M.; Dafforn, T.R.; Poyner, D.R.; et al. Ligand-induced conformational changes in a SMALP-encapsulated GPCR. *Biochim. Biophys. Acta Biomembr.* **2020**, *1862*, 183235. [[CrossRef](#)] [[PubMed](#)]
113. Marty, M.T.; Hoi, K.K.; Robinson, C.V. Interfacing Membrane Mimetics with Mass Spectrometry. *Acc. Chem. Res.* **2016**, *49*, 2459–2467. [[CrossRef](#)] [[PubMed](#)]
114. Keener, J.E.; Zhang, G.; Marty, M.T. Native Mass Spectrometry of Membrane Proteins. *Anal. Chem.* **2021**, *93*, 583–597. [[CrossRef](#)]
115. Young, G.; Hundt, N.; Cole, D.; Fineberg, A.; Andrecka, J.; Tyler, A.; Olerinyova, A.; Ansari, A.; Marklund, E.G.; Collier, M.P.; et al. Quantitative mass imaging of single biological macromolecules. *Science* **2018**, *360*, 423–427. [[CrossRef](#)]
116. Olerinyova, A.; Sonn-Segev, A.; Gault, J.; Eichmann, C.; Schimpf, J.; Kopf, A.H.; Rudden, L.S.P.; Ashkinadze, D.; Bomba, R.; Frey, L.; et al. Mass Photometry of Membrane Proteins. *Chem* **2021**, *7*, 224–236. [[CrossRef](#)]
117. Schlick, K.H.; Lange, C.K.; Gillispie, G.D.; Cloninger, M.J. Characterization of protein aggregation via intrinsic fluorescence lifetime. *J. Am. Chem. Soc.* **2009**, *131*, 16608–16609. [[CrossRef](#)] [[PubMed](#)]
118. Singh, S.; Zhang, M.; Bertheleme, N.; Kara, E.; Strange, P.G.; Byrne, B. Radioligand binding analysis as a tool for quality control of GPCR production for structural characterization: Adenosine A_{2A}R as a template for study. *Curr. Protoc. Protein Sci.* **2012**, *29*, 29.3. [[CrossRef](#)] [[PubMed](#)]
119. Lee, S.; Ghosh, S.; Jana, S.; Robertson, N.; Tate, C.G.; Vaidehi, N. How Do Branched Detergents Stabilize GPCRs in Micelles? *Biochemistry* **2020**, *59*, 2125–2134. [[CrossRef](#)]
120. Holdgate, G.; Embrey, K.; Milbradt, A.; Davies, G. Biophysical methods in early drug discovery. *ADMET DMPK* **2019**, *7*, 222–241. [[CrossRef](#)] [[PubMed](#)]
121. van der Velden, W.J.C.; Heitman, L.H.; Rosenkilde, M.M. Perspective: Implications of Ligand-Receptor Binding Kinetics for Therapeutic Targeting of G Protein-Coupled Receptors. *ACS Pharmacol. Transl. Sci.* **2020**, *3*, 179–189. [[CrossRef](#)]
122. Copeland, R.A. The drug-target residence time model: A 10-year retrospective. *Nat. Rev. Drug Discov.* **2016**, *15*, 87–95. [[CrossRef](#)]
123. Guo, D.; Heitman, L.H.; IJzerman, A.P. Kinetic Aspects of the Interaction between Ligand and G Protein-Coupled Receptor: The Case of the Adenosine Receptors. *Chem. Rev.* **2017**, *117*, 38–66. [[CrossRef](#)] [[PubMed](#)]
124. IJzerman, A.P.; Guo, D. Drug-Target Association Kinetics in Drug Discovery. *Trends Biochem. Sci.* **2019**, *44*, 861–871. [[CrossRef](#)]
125. Rich, R.L.; Myszka, D.G. Advances in surface plasmon resonance biosensor analysis. *Curr. Opin. Biotechnol.* **2000**, *11*, 54–61. [[CrossRef](#)]
126. Locatelli-Hoops, S.; Yeliseev, A.A.; Gawrisch, K.; Gorshkova, I. Surface plasmon resonance applied to G protein-coupled receptors. *Biomed. Spectrosc. Imaging* **2013**, *2*, 155–181. [[CrossRef](#)] [[PubMed](#)]
127. Patching, S.G. Surface plasmon resonance spectroscopy for characterisation of membrane protein-ligand interactions and its potential for drug discovery. *Biochim. Biophys. Acta* **2014**, *1838*, 43–55. [[CrossRef](#)]
128. Aristotelous, T.; Hopkins, A.L.; Navratilova, I. Surface plasmon resonance analysis of seven-transmembrane receptors. *Methods Enzymol.* **2015**, *556*, 499–525. [[CrossRef](#)]
129. Olaru, A.; Bala, C.; Jaffrezic-Renault, N.; Aboul-Enein, H.Y. Surface plasmon resonance (SPR) biosensors in pharmaceutical analysis. *Crit. Rev. Anal. Chem.* **2015**, *45*, 97–105. [[CrossRef](#)]

130. Giannetti, A.M. From experimental design to validated hits a comprehensive walk-through of fragment lead identification using surface plasmon resonance. *Methods Enzymol.* **2011**, *493*, 169–218. [[CrossRef](#)] [[PubMed](#)]
131. Rich, R.L.; Errey, J.; Marshall, F.; Myszkka, D.G. Biacore analysis with stabilized G-protein-coupled receptors. *Anal. Biochem.* **2011**, *409*, 267–272. [[CrossRef](#)] [[PubMed](#)]
132. Shepherd, C.A.; Hopkins, A.L.; Navratilova, I. Fragment screening by SPR and advanced application to GPCRs. *Prog. Biophys. Mol. Biol.* **2014**, *116*, 113–123. [[CrossRef](#)] [[PubMed](#)]
133. Robertson, N.; Jazayeri, A.; Errey, J.; Baig, A.; Hurrell, E.; Zhukov, A.; Langmead, C.J.; Weir, M.; Marshall, F.H. The properties of thermostabilised G protein-coupled receptors (StaRs) and their use in drug discovery. *Neuropharmacology* **2011**, *60*, 36–44. [[CrossRef](#)] [[PubMed](#)]
134. Congreve, M.; Rich, R.L.; Myszkka, D.G.; Figaroa, F.; Siegal, G.; Marshall, F.H. Fragment screening of stabilized G-protein-coupled receptors using biophysical methods. *Methods Enzymol.* **2011**, *493*, 115–136. [[CrossRef](#)]
135. Congreve, M.; Andrews, S.P.; Doré, A.S.; Hollenstein, K.; Hurrell, E.; Langmead, C.J.; Mason, J.S.; Ng, I.W.; Tehan, B.; Zhukov, A.; et al. Discovery of 1,2,4-triazine derivatives as adenosine A_{2A} antagonists using structure based drug design. *J. Med. Chem.* **2012**, *55*, 1898–1903. [[CrossRef](#)]
136. Zhukov, A.; Andrews, S.P.; Errey, J.C.; Robertson, N.; Tehan, B.; Mason, J.S.; Marshall, F.H.; Weir, M.; Congreve, M. Biophysical mapping of the adenosine A_{2A} receptor. *J. Med. Chem.* **2011**, *54*, 4312–4323. [[CrossRef](#)]
137. Roos, H.; Karlsson, R.; Nilshans, H.; Persson, A. Thermodynamic analysis of protein interactions with biosensor technology. *J. Mol. Recognit.* **1998**, *11*, 204–210. [[CrossRef](#)]
138. Babazada, H.; Alekberli, T.; Hajieva, P.; Farajov, E. Biosensor-based kinetic and thermodynamic characterization of opioids interaction with human μ -opioid receptor. *Eur. J. Pharm. Sci.* **2019**, *138*, 105017. [[CrossRef](#)]
139. Deganutti, G.; Zhukov, A.; Deflorian, F.; Federico, S.; Spalluto, G.; Cooke, R.M.; Moro, S.; Mason, J.S.; Bortolato, A. Impact of protein-ligand solvation and desolvation on transition state thermodynamic properties of adenosine A_{2A} ligand binding kinetics. *Silico Pharmacol.* **2017**, *5*, 16. [[CrossRef](#)]
140. Coyle, J.; Walser, R. Applied Biophysical Methods in Fragment-Based Drug Discovery. *SLAS Discov.* **2020**, *25*, 471–490. [[CrossRef](#)]
141. Lu, Y.; Qin, S.; Zhang, B.; Dai, A.; Cai, X.; Ma, M.; Gao, Z.-G.; Yang, D.; Stevens, R.C.; Jacobson, K.A.; et al. Accelerating the Throughput of Affinity Mass Spectrometry-Based Ligand Screening toward a G Protein-Coupled Receptor. *Anal. Chem.* **2019**, *91*, 8162–8169. [[CrossRef](#)] [[PubMed](#)]
142. Shepherd, C.; Robinson, S.; Berizzi, A.; Thompson, L.E.J.; Bird, L.; Culurgioni, S.; Varzandeh, S.; Rawlins, P.B.; Olsen, R.H.J.; Navratilova, I.H. Surface Plasmon Resonance Screening to Identify Active and Selective Adenosine Receptor Binding Fragments. *ACS Med. Chem. Lett.* **2022**, *13*, 1172–1181. [[CrossRef](#)]
143. Koretz, K.S.; McGraw, C.E.; Stradley, S.; Elbaradei, A.; Malmstadt, N.; Robinson, A.S. Characterization of binding kinetics of A_{2A}R to G α s protein by surface plasmon resonance. *Biophys. J.* **2021**, *120*, 1641–1649. [[CrossRef](#)] [[PubMed](#)]
144. Damian, L. Isothermal titration calorimetry for studying protein-ligand interactions. *Methods Mol. Biol.* **2013**, *1008*, 103–118. [[CrossRef](#)]
145. Pierce, M.M.; Raman, C.S.; Nall, B.T. Isothermal titration calorimetry of protein-protein interactions. *Methods* **1999**, *19*, 213–221. [[CrossRef](#)]
146. Drackowski, P.; Matusiuk, D.; Jozwiak, K. Isothermal titration calorimetry in membrane protein research. *J. Pharm. Biomed. Anal.* **2014**, *87*, 313–325. [[CrossRef](#)]
147. Rajarathnam, K.; Rösger, J. Isothermal titration calorimetry of membrane proteins-progress and challenges. *Biochim. Biophys. Acta* **2014**, *1838*, 69–77. [[CrossRef](#)]
148. Magalhaes, A.C.; Dunn, H.; Ferguson, S.S. Regulation of GPCR activity, trafficking and localization by GPCR-interacting proteins. *Br. J. Pharmacol.* **2012**, *165*, 1717–1736. [[CrossRef](#)] [[PubMed](#)]
149. Klaasse, E.C.; IJzerman, A.P.; de Grip, W.J.; Beukers, M.W. Internalization and desensitization of adenosine receptors. *Purinergic Signal.* **2008**, *4*, 21–37. [[CrossRef](#)]
150. Burgueño, J.; Blake, D.J.; Benson, M.A.; Tinsley, C.L.; Esapa, C.T.; Canela, E.I.; Penela, P.; Mallol, J.; Mayor, F.; Lluís, C.; et al. The adenosine A_{2A} receptor interacts with the actin-binding protein alpha-actinin. *J. Biol. Chem.* **2003**, *278*, 37545–37552. [[CrossRef](#)]
151. Woods, A.S.; Marcellino, D.; Jackson, S.N.; Franco, R.; Ferré, S.; Agnati, L.F.; Fuxe, K. How calmodulin interacts with the adenosine A_{2A} and the dopamine D₂ receptors. *J. Proteome Res.* **2008**, *7*, 3428–3434. [[CrossRef](#)]
152. Milojevic, T.; Reiterer, V.; Stefan, E.; Korkhov, V.M.; Dorostkar, M.M.; Ducza, E.; Ogris, E.; Boehm, S.; Freissmuth, M.; Nanoff, C. The ubiquitin-specific protease Usp4 regulates the cell surface level of the A_{2A} receptor. *Mol. Pharmacol.* **2006**, *69*, 1083–1094. [[CrossRef](#)] [[PubMed](#)]
153. Canela, L.; Luján, R.; Lluís, C.; Burgueño, J.; Mallol, J.; Canela, E.I.; Franco, R.; Ciruela, F. The neuronal Ca²⁺-binding protein 2 (NECAB2) interacts with the adenosine A_{2A} receptor and modulates the cell surface expression and function of the receptor. *Mol. Cell. Neurosci.* **2007**, *36*, 1–12. [[CrossRef](#)] [[PubMed](#)]
154. Sun, C.-N.; Cheng, H.-C.; Chou, J.-L.; Lee, S.-Y.; Lin, Y.-W.; Lai, H.-L.; Chen, H.-M.; Chern, Y. Rescue of p53 blockage by the A_{2A} adenosine receptor via a novel interacting protein, translin-associated protein X. *Mol. Pharmacol.* **2006**, *70*, 454–466. [[CrossRef](#)]
155. Gsandtner, I.; Charalambous, C.; Stefan, E.; Ogris, E.; Freissmuth, M.; Zezula, J. Heterotrimeric G protein-independent signaling of a G protein-coupled receptor. Direct binding of ARNO/cytohesin-2 to the carboxyl terminus of the A_{2A} adenosine receptor is necessary for sustained activation of the ERK/MAP kinase pathway. *J. Biol. Chem.* **2005**, *280*, 31898–31905. [[CrossRef](#)]

156. Navarro, G.; Hradsky, J.; Lluís, C.; Casadó, V.; McCormick, P.J.; Kreutz, M.R.; Mikhaylova, M. NCS-1 associates with adenosine A_{2A} receptors and modulates receptor function. *Front. Mol. Neurosci.* **2012**, *5*, 53. [[CrossRef](#)] [[PubMed](#)]
157. Piirainen, H.; Hellman, M.; Tossavainen, H.; Permi, P.; Kursula, P.; Jaakola, V.-P. Human adenosine A_{2A} receptor binds calmodulin with high affinity in a calcium-dependent manner. *Biophys. J.* **2015**, *108*, 903–917. [[CrossRef](#)] [[PubMed](#)]
158. Piirainen, H.; Taura, J.; Kursula, P.; Ciruela, F.; Jaakola, V.-P. Calcium modulates calmodulin/ α -actinin 1 interaction with and agonist-dependent internalization of the adenosine A_{2A} receptor. *Biochim. Biophys. Acta Mol. Cell Res.* **2017**, *1864*, 674–686. [[CrossRef](#)]
159. Quast, R.B.; Margeat, E. Studying GPCR conformational dynamics by single molecule fluorescence. *Mol. Cell. Endocrinol.* **2019**, *493*, 110469. [[CrossRef](#)] [[PubMed](#)]
160. Tian, H.; Fürstenberg, A.; Huber, T. Labeling and Single-Molecule Methods To Monitor G Protein-Coupled Receptor Dynamics. *Chem. Rev.* **2017**, *117*, 186–245. [[CrossRef](#)]
161. Hoffmann, C.; Gaietta, G.; Bünemann, M.; Adams, S.R.; Oberdorff-Maass, S.; Behr, B.; Vilardaga, J.-P.; Tsien, R.Y.; Ellisman, M.H.; Lohse, M.J. A FIAsh-based FRET approach to determine G protein-coupled receptor activation in living cells. *Nat. Methods* **2005**, *2*, 171–176. [[CrossRef](#)] [[PubMed](#)]
162. Stoddart, L.A.; Kindon, N.D.; Otun, O.; Harwood, C.R.; Patera, F.; Veprintsev, D.B.; Woolard, J.; Briddon, S.J.; Franks, H.A.; Hill, S.J.; et al. Ligand-directed covalent labelling of a GPCR with a fluorescent tag in live cells. *Commun. Biol.* **2020**, *3*, 722. [[CrossRef](#)] [[PubMed](#)]
163. Wei, S.; Thakur, N.; Ray, A.P.; Jin, B.; Obeng, S.; McCurdy, C.R.; McMahon, L.R.; Gutiérrez-de-Terán, H.; Eddy, M.T.; Lamichhane, R. Slow conformational dynamics of the human A_{2A} adenosine receptor are temporally ordered. *Structure* **2022**, *30*, 329–337.e5. [[CrossRef](#)]
164. Fernandes, D.D.; Neale, C.; Gomes, G.-N.W.; Li, Y.; Malik, A.; Pandey, A.; Oraziotti, A.P.; Wang, X.; Ye, L.; Scott Prosser, R.; et al. Ligand modulation of the conformational dynamics of the A_{2A} adenosine receptor revealed by single-molecule fluorescence. *Sci. Rep.* **2021**, *11*, 5910. [[CrossRef](#)] [[PubMed](#)]
165. Zemella, A.; Richter, T.; Thoring, L.; Kubick, S. A Combined Cell-Free Protein Synthesis and Fluorescence-Based Approach to Investigate GPCR Binding Properties. *Methods Mol. Biol.* **2019**, *1947*, 57–77. [[CrossRef](#)]
166. Raingeval, C.; Krimm, I. NMR investigation of protein-ligand interactions for G-protein coupled receptors. *Future Med. Chem.* **2019**, *11*, 1811–1825. [[CrossRef](#)]
167. Stauch, B.; Orts, J.; Carlomagno, T. The description of protein internal motions aids selection of ligand binding poses by the INPHARMA method. *J. Biomol. NMR* **2012**, *54*, 245–256. [[CrossRef](#)]
168. Yen, H.-Y.; Hoi, K.K.; Liko, I.; Hedger, G.; Horrell, M.R.; Song, W.; Wu, D.; Heine, P.; Warne, T.; Lee, Y.; et al. PtdIns(4,5)P₂ stabilizes active states of GPCRs and enhances selectivity of G-protein coupling. *Nature* **2018**, *559*, 423–427. [[CrossRef](#)]
169. Agasid, M.T.; Sørensen, L.; Urner, L.H.; Yan, J.; Robinson, C.V. The Effects of Sodium Ions on Ligand Binding and Conformational States of G Protein-Coupled Receptors-Insights from Mass Spectrometry. *J. Am. Chem. Soc.* **2021**, *143*, 4085–4089. [[CrossRef](#)] [[PubMed](#)]
170. Olsen, R.H.J.; English, J.G. Advancements in G protein-coupled receptor biosensors to study GPCR-G protein coupling. *Br. J. Pharmacol.* **2022**. [[CrossRef](#)]
171. Zhou, Y.; Meng, J.; Xu, C.; Liu, J. Multiple GPCR Functional Assays Based on Resonance Energy Transfer Sensors. *Front. Cell Dev. Biol.* **2021**, *9*, 611443. [[CrossRef](#)] [[PubMed](#)]
172. Milligan, G. Principles: Extending the utility of [35S]GTP γ S binding assays. *Trends Pharmacol. Sci.* **2003**, *24*, 87–90. [[CrossRef](#)]
173. McEwen, D.P.; Gee, K.R.; Kang, H.C.; Neubig, R.R. Fluorescent BODIPY-GTP analogs: Real-time measurement of nucleotide binding to G proteins. *Anal. Biochem.* **2001**, *291*, 109–117. [[CrossRef](#)]
174. Ravyn, V.; Bostwick, J.R. Functional coupling of the Galpha(olf) variant XLGalpha(olf) with the human adenosine A_{2A} receptor. *J. Recept. Signal Transduct. Res.* **2006**, *26*, 241–258. [[CrossRef](#)] [[PubMed](#)]
175. Beckner, R.L.; Zoubak, L.; Hines, K.G.; Gawrisch, K.; Yeliseev, A.A. Probing thermostability of detergent-solubilized CB2 receptor by parallel G protein-activation and ligand-binding assays. *J. Biol. Chem.* **2020**, *295*, 181–190. [[CrossRef](#)]
176. Arcemibéhère, L.; Sen, T.; Boudier, L.; Balestre, M.-N.; Gaibelet, G.; Detouillon, E.; Orcel, H.; Mendre, C.; Rahmeh, R.; Granier, S.; et al. Leukotriene BLT2 receptor monomers activate the G(i2) GTP-binding protein more efficiently than dimers. *J. Biol. Chem.* **2010**, *285*, 6337–6347. [[CrossRef](#)] [[PubMed](#)]
177. Inagaki, S.; Ghirlando, R.; White, J.F.; Gvozdenovic-Jeremic, J.; Northup, J.K.; Grisshammer, R. Modulation of the interaction between neurotensin receptor NTS1 and Gq protein by lipid. *J. Mol. Biol.* **2012**, *417*, 95–111. [[CrossRef](#)] [[PubMed](#)]
178. Damian, M.; Marie, J.; Leyris, J.-P.; Fehrentz, J.-A.; Verdié, P.; Martinez, J.; Banères, J.-L.; Mary, S. High constitutive activity is an intrinsic feature of ghrelin receptor protein: A study with a functional monomeric GHS-R1a receptor reconstituted in lipid discs. *J. Biol. Chem.* **2012**, *287*, 3630–3641. [[CrossRef](#)]
179. Gregorio, G.G.; Masurel, M.; Hilger, D.; Terry, D.S.; Juette, M.; Zhao, H.; Zhou, Z.; Perez-Aguilar, J.M.; Hauge, M.; Mathiasen, S.; et al. Single-molecule analysis of ligand efficacy in β 2AR-G-protein activation. *Nature* **2017**, *547*, 68–73. [[CrossRef](#)]
180. Damian, M.; Louet, M.; Gomes, A.A.S.; M’Kadmi, C.; Denoyelle, S.; Cantel, S.; Mary, S.; Bisch, P.M.; Fehrentz, J.-A.; Catoire, L.J.; et al. Allosteric modulation of ghrelin receptor signaling by lipids. *Nat. Commun.* **2021**, *12*, 3938. [[CrossRef](#)] [[PubMed](#)]

181. Bayburt, T.H.; Vishnivetskiy, S.A.; McLean, M.A.; Morizumi, T.; Huang, C.-C.; Tesmer, J.J.G.; Ernst, O.P.; Sligar, S.G.; Gurevich, V.V. Monomeric rhodopsin is sufficient for normal rhodopsin kinase (GRK1) phosphorylation and arrestin-1 binding. *J. Biol. Chem.* **2011**, *286*, 1420–1428. [[CrossRef](#)] [[PubMed](#)]
182. Srivastava, A.; Baidya, M.; Dwivedi-Agnihotri, H.; Shukla, A.K. Site-directed labeling of β -arrestin with monobromobimane for measuring their interaction with G protein-coupled receptors. *Methods Enzymol.* **2020**, *633*, 271–280. [[CrossRef](#)] [[PubMed](#)]
183. Sommer, M.E.; Farrens, D.L. Arrestin can act as a regulator of rhodopsin photochemistry. *Vis. Res.* **2006**, *46*, 4532–4546. [[CrossRef](#)] [[PubMed](#)]
184. Rahmeh, R.; Damian, M.; Cottet, M.; Orcel, H.; Mendre, C.; Durroux, T.; Sharma, K.S.; Durand, G.; Pucci, B.; Trinquet, E.; et al. Structural insights into biased G protein-coupled receptor signaling revealed by fluorescence spectroscopy. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 6733–6738. [[CrossRef](#)]
185. Kumari, P.; Srivastava, A.; Banerjee, R.; Ghosh, E.; Gupta, P.; Ranjan, R.; Chen, X.; Gupta, B.; Gupta, C.; Jaiman, D.; et al. Functional competence of a partially engaged GPCR- β -arrestin complex. *Nat. Commun.* **2016**, *7*, 13416. [[CrossRef](#)]
186. Staus, D.P.; Wingler, L.M.; Choi, M.; Pani, B.; Manglik, A.; Kruse, A.C.; Lefkowitz, R.J. Sortase ligation enables homogeneous GPCR phosphorylation to reveal diversity in β -arrestin coupling. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 3834–3839. [[CrossRef](#)]
187. Prosser, R.S.; Ye, L.; Pandey, A.; Oraziotti, A. Activation processes in ligand-activated G protein-coupled receptors: A case study of the adenosine A_{2A} receptor. *Bioessays* **2017**, *39*, 1700072. [[CrossRef](#)] [[PubMed](#)]
188. Welihinda, A.A.; Kaur, M.; Greene, K.; Zhai, Y.; Amento, E.P. The adenosine metabolite inosine is a functional agonist of the adenosine A_{2A} receptor with a unique signaling bias. *Cell. Signal.* **2016**, *28*, 552–560. [[CrossRef](#)]
189. Kolb, P.; Kenakin, T.; Alexander, S.P.H.; Bermudez, M.; Bohn, L.M.; Breinholt, C.S.; Bouvier, M.; Hill, S.J.; Kostenis, E.; Martemyanov, K.A.; et al. Community guidelines for GPCR ligand bias: IUPHAR review 32. *Br. J. Pharmacol.* **2022**, *179*, 3651–3674. [[CrossRef](#)]
190. Benleulmi-Chaachoua, A.; Chen, L.; Sokolina, K.; Wong, V.; Jurisica, I.; Emerit, M.B.; Darmon, M.; Espin, A.; Stagljar, I.; Tafelmeyer, P.; et al. Protein interactome mining defines melatonin MT1 receptors as integral component of presynaptic protein complexes of neurons. *J. Pineal Res.* **2016**, *60*, 95–108. [[CrossRef](#)]
191. Overton, M.C.; Blumer, K.J. The extracellular N-terminal domain and transmembrane domains 1 and 2 mediate oligomerization of a yeast G protein-coupled receptor. *J. Biol. Chem.* **2002**, *277*, 41463–41472. [[CrossRef](#)] [[PubMed](#)]
192. Mahesh, G.; Jaiswal, P.; Dey, S.; Sengupta, J.; Mukherjee, S. Cloning, Expression, Purification and Characterization of Oligomeric States of the Native 5HT_{2A} G-Protein-Coupled Receptor. *Protein Pept. Lett.* **2018**, *25*, 390–397. [[CrossRef](#)]
193. Fernández-Dueñas, V.; Bonaventura, J.; Aso, E.; Luján, R.; Ferré, S.; Ciruela, F. Overcoming the Challenges of Detecting GPCR Oligomerization in the Brain. *Curr. Neuropharmacol.* **2022**, *20*, 1035–1045. [[CrossRef](#)]
194. Mansoor, S.; Kayık, G.; Durdagi, S.; Sensoy, O. Mechanistic insight into the impact of a bivalent ligand on the structure and dynamics of a GPCR oligomer. *Comput. Struct. Biotechnol. J.* **2022**, *20*, 925–936. [[CrossRef](#)] [[PubMed](#)]
195. Canals, M.; Burgueño, J.; Marcellino, D.; Cabello, N.; Canela, E.I.; Mallol, J.; Agnati, L.; Ferré, S.; Bouvier, M.; Fuxe, K.; et al. Homodimerization of adenosine A_{2A} receptors: Qualitative and quantitative assessment by fluorescence and bioluminescence energy transfer. *J. Neurochem.* **2004**, *88*, 726–734. [[CrossRef](#)] [[PubMed](#)]
196. Ciruela, F.; Casadó, V.; Rodrigues, R.J.; Luján, R.; Burgueño, J.; Canals, M.; Borycz, J.; Rebola, N.; Goldberg, S.R.; Mallol, J.; et al. Presynaptic control of striatal glutamatergic neurotransmission by adenosine A₁-A_{2A} receptor heteromers. *J. Neurosci.* **2006**, *26*, 2080–2087. [[CrossRef](#)]
197. Franco, R.; Cordero, A.; Del Llinas Torrent, C.; Lillo, A.; Serrano-Marín, J.; Navarro, G.; Pardo, L. Structure and function of adenosine receptor heteromers. *Cell. Mol. Life Sci.* **2021**, *78*, 3957–3968. [[CrossRef](#)] [[PubMed](#)]
198. Carriba, P.; Ortiz, O.; Patkar, K.; Justinova, Z.; Stroik, J.; Themann, A.; Müller, C.; Woods, A.S.; Hope, B.T.; Ciruela, F.; et al. Striatal adenosine A_{2A} and cannabinoid CB1 receptors form functional heteromeric complexes that mediate the motor effects of cannabinoids. *Neuropsychopharmacology* **2007**, *32*, 2249–2259. [[CrossRef](#)] [[PubMed](#)]
199. Hillion, J.; Canals, M.; Torvinen, M.; Casado, V.; Scott, R.; Terasmaa, A.; Hansson, A.; Watson, S.; Olah, M.E.; Mallol, J.; et al. Coaggregation, cointernalization, and codesensitization of adenosine A_{2A} receptors and dopamine D₂ receptors. *J. Biol. Chem.* **2002**, *277*, 18091–18097. [[CrossRef](#)] [[PubMed](#)]
200. Canals, M.; Marcellino, D.; Fanelli, F.; Ciruela, F.; de Benedetti, P.; Goldberg, S.R.; Neve, K.; Fuxe, K.; Agnati, L.F.; Woods, A.S.; et al. Adenosine A_{2A}-dopamine D₂ receptor-receptor heteromerization: Qualitative and quantitative assessment by fluorescence and bioluminescence energy transfer. *J. Biol. Chem.* **2003**, *278*, 46741–46749. [[CrossRef](#)]
201. Kamiya, T.; Saitoh, O.; Yoshioka, K.; Nakata, H. Oligomerization of adenosine A_{2A} and dopamine D₂ receptors in living cells. *Biochem. Biophys. Res. Commun.* **2003**, *306*, 544–549. [[CrossRef](#)]
202. Torvinen, M.; Marcellino, D.; Canals, M.; Agnati, L.F.; Lluís, C.; Franco, R.; Fuxe, K. Adenosine A_{2A} receptor and dopamine D₃ receptor interactions: Evidence of functional A_{2A}/D₃ heteromeric complexes. *Mol. Pharmacol.* **2005**, *67*, 400–407. [[CrossRef](#)]
203. Cabello, N.; Gandía, J.; Bertarelli, D.C.G.; Watanabe, M.; Lluís, C.; Franco, R.; Ferré, S.; Luján, R.; Ciruela, F. Metabotropic glutamate type 5, dopamine D₂ and adenosine A_{2A} receptors form higher-order oligomers in living cells. *J. Neurochem.* **2009**, *109*, 1497–1507. [[CrossRef](#)] [[PubMed](#)]
204. Trifilieff, P.; Rives, M.-L.; Urizar, E.; Piskorowski, R.A.; Vishwasrao, H.D.; Castrillon, J.; Schmauss, C.; Slättman, M.; Gullberg, M.; Javitch, J.A. Detection of antigen interactions ex vivo by proximity ligation assay: Endogenous dopamine D₂-adenosine A_{2A} receptor complexes in the striatum. *Biotechniques* **2011**, *51*, 111–118. [[CrossRef](#)]

205. Fernández-Dueñas, V.; Gómez-Soler, M.; Valle-León, M.; Watanabe, M.; Ferrer, I.; Ciruela, F. Revealing Adenosine A_{2A}-Dopamine D₂ Receptor Heteromers in Parkinson's Disease Post-Mortem Brain through a New AlphaScreen-Based Assay. *Int. J. Mol. Sci.* **2019**, *20*, 3600. [[CrossRef](#)] [[PubMed](#)]
206. Beggiato, S.; Antonelli, T.; Tomasini, M.C.; Borelli, A.C.; Agnati, L.F.; Tanganelli, S.; Fuxe, K.; Ferraro, L. Adenosine A_{2A}-D₂ receptor-receptor interactions in putative heteromers in the regulation of the striato-pallidal gaba pathway: Possible relevance for parkinson's disease and its treatment. *Curr. Protein Pept. Sci.* **2014**, *15*, 673–680. [[CrossRef](#)]
207. Ferré, S.; Bonaventura, J.; Tomasi, D.; Navarro, G.; Moreno, E.; Cortés, A.; Lluís, C.; Casadó, V.; Volkow, N.D. Allosteric mechanisms within the adenosine A_{2A}-dopamine D₂ receptor heterotetramer. *Neuropharmacology* **2016**, *104*, 154–160. [[CrossRef](#)]
208. Romero-Fernandez, W.; Taura, J.J.; Crans, R.A.J.; Lopez-Cano, M.; Fores-Pons, R.; Narváez, M.; Carlsson, J.; Ciruela, F.; Fuxe, K.; Borroto-Escuela, D.O. The mGlu5 Receptor Protomer-Mediated Dopamine D₂ Receptor Trans-Inhibition Is Dependent on the Adenosine A_{2A} Receptor Protomer: Implications for Parkinson's Disease. *Mol. Neurobiol.* **2022**, *59*, 5955–5969. [[CrossRef](#)] [[PubMed](#)]
209. Schuler, M.A.; Denisov, I.G.; Sligar, S.G. Nanodiscs as a new tool to examine lipid-protein interactions. *Methods Mol. Biol.* **2013**, *974*, 415–433. [[CrossRef](#)] [[PubMed](#)]
210. White, J.F.; Grodnitzky, J.; Louis, J.M.; Trinh, L.B.; Shiloach, J.; Gutierrez, J.; Northup, J.K.; Grishammer, R. Dimerization of the class A G protein-coupled neurotensin receptor NTS1 alters G protein interaction. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 12199–12204. [[CrossRef](#)] [[PubMed](#)]
211. Bayburt, T.H.; Leitz, A.J.; Xie, G.; Oprian, D.D.; Sligar, S.G. Transducin activation by nanoscale lipid bilayers containing one and two rhodopsins. *J. Biol. Chem.* **2007**, *282*, 14875–14881. [[CrossRef](#)]
212. El Moustaine, D.; Granier, S.; Doumazane, E.; Scholler, P.; Rahmeh, R.; Bron, P.; Mouillac, B.; Banères, J.-L.; Rondard, P.; Pin, J.-P. Distinct roles of metabotropic glutamate receptor dimerization in agonist activation and G-protein coupling. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 16342–16347. [[CrossRef](#)]
213. Mary, S.; Fehrentz, J.-A.; Damian, M.; Gaibelet, G.; Orcel, H.; Verdié, P.; Mouillac, B.; Martinez, J.; Marie, J.; Banères, J.-L. Heterodimerization with Its splice variant blocks the ghrelin receptor 1a in a non-signaling conformation: A study with a purified heterodimer assembled into lipid discs. *J. Biol. Chem.* **2013**, *288*, 24656–24665. [[CrossRef](#)]
214. Mesnier, D.; Banères, J.-L. Cooperative conformational changes in a G-protein-coupled receptor dimer, the leukotriene B(4) receptor BLT1. *J. Biol. Chem.* **2004**, *279*, 49664–49670. [[CrossRef](#)]
215. Damian, M.; Pons, V.; Renault, P.; M'Kadmi, C.; Delort, B.; Hartmann, L.; Kaya, A.I.; Louet, M.; Gagne, D.; Ben Haj Salah, K.; et al. GHSR-D2R heteromerization modulates dopamine signaling through an effect on G protein conformation. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 4501–4506. [[CrossRef](#)]
216. Tsukamoto, H.; Sinha, A.; DeWitt, M.; Farrens, D.L. Monomeric rhodopsin is the minimal functional unit required for arrestin binding. *J. Mol. Biol.* **2010**, *399*, 501–511. [[CrossRef](#)] [[PubMed](#)]
217. Piirainen, H.; Ashok, Y.; Nanekar, R.T.; Jaakola, V.-P. Structural features of adenosine receptors: From crystal to function. *Biochim. Biophys. Acta* **2011**, *1808*, 1233–1244. [[CrossRef](#)]
218. Langelaan, D.N.; Ngweniform, P.; Rainey, J.K. Biophysical characterization of G-protein coupled receptor-peptide ligand binding. *Biochem. Cell Biol.* **2011**, *89*, 98–105. [[CrossRef](#)] [[PubMed](#)]
219. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596*, 583–589. [[CrossRef](#)] [[PubMed](#)]
220. Varadi, M.; Anyango, S.; Deshpande, M.; Nair, S.; Natassia, C.; Yordanova, G.; Yuan, D.; Stroe, O.; Wood, G.; Laydon, A.; et al. AlphaFold Protein Structure Database: Massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* **2022**, *50*, D439–D444. [[CrossRef](#)] [[PubMed](#)]
221. Marcoux, J.; Cianférani, S. Towards integrative structural mass spectrometry: Benefits from hybrid approaches. *Methods* **2015**, *89*, 4–12. [[CrossRef](#)]
222. Dafun, A.S.; Marcoux, J. Structural mass spectrometry of membrane proteins. *Biochim. Biophys. Acta Proteins Proteom.* **2022**, *1870*, 140813. [[CrossRef](#)] [[PubMed](#)]
223. Khanal, A.; Pan, Y.; Brown, L.S.; Konermann, L. Pulsed hydrogen/deuterium exchange mass spectrometry for time-resolved membrane protein folding studies. *J. Mass Spectrom.* **2012**, *47*, 1620–1626. [[CrossRef](#)]
224. Masson, G.R.; Burke, J.E.; Ahn, N.G.; Anand, G.S.; Borchers, C.; Brier, S.; Bou-Assaf, G.M.; Engen, J.R.; Englander, S.W.; Faber, J.; et al. Recommendations for performing, interpreting and reporting hydrogen deuterium exchange mass spectrometry (HDX-MS) experiments. *Nat. Methods* **2019**, *16*, 595–602. [[CrossRef](#)]
225. Möller, I.R.; Slivacka, M.; Hausner, J.; Nielsen, A.K.; Pospíšilová, E.; Merkle, P.S.; Lišková, R.; Polák, M.; Loland, C.J.; Kádek, A.; et al. Improving the Sequence Coverage of Integral Membrane Proteins during Hydrogen/Deuterium Exchange Mass Spectrometry Experiments. *Anal. Chem.* **2019**, *91*, 10970–10978. [[CrossRef](#)]
226. Trabjerg, E.; Nazari, Z.E.; Rand, K.D. Conformational analysis of complex protein states by hydrogen/deuterium exchange mass spectrometry (HDX-MS): Challenges and emerging solutions. *TrAC Trends Anal. Chem.* **2018**, *106*, 125–138. [[CrossRef](#)]
227. Martens, C.; Politis, A. A glimpse into the molecular mechanism of integral membrane proteins through hydrogen-deuterium exchange mass spectrometry. *Protein Sci.* **2020**, *29*, 1285–1301. [[CrossRef](#)] [[PubMed](#)]

228. Zhang, Y.; Rempel, D.L.; Zhang, J.; Sharma, A.K.; Mirica, L.M.; Gross, M.L. Pulsed hydrogen-deuterium exchange mass spectrometry probes conformational changes in amyloid beta (A β) peptide aggregation. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 14604–14609. [[CrossRef](#)]
229. Du, Y.; Duc, N.M.; Rasmussen, S.G.F.; Hilger, D.; Kubiak, X.; Wang, L.; Bohon, J.; Kim, H.R.; Wegrecki, M.; Asuru, A.; et al. Assembly of a GPCR-G Protein Complex. *Cell* **2019**, *177*, 1232–1242.e11. [[CrossRef](#)]
230. Konermann, L.; Tong, X.; Pan, Y. Protein structure and dynamics studied by mass spectrometry: H/D exchange, hydroxyl radical labeling, and related approaches. *J. Mass Spectrom.* **2008**, *43*, 1021–1036. [[CrossRef](#)]
231. Wang, L.; Chance, M.R. Protein Footprinting Comes of Age: Mass Spectrometry for Biophysical Structure Assessment. *Mol. Cell. Proteom.* **2017**, *16*, 706–716. [[CrossRef](#)] [[PubMed](#)]
232. Iacobucci, C.; Piotrowski, C.; Aebersold, R.; Amaral, B.C.; Andrews, P.; Bernfur, K.; Borchers, C.; Brodie, N.I.; Bruce, J.E.; Cao, Y.; et al. First Community-Wide, Comparative Cross-Linking Mass Spectrometry Study. *Anal. Chem.* **2019**, *91*, 6953–6961. [[CrossRef](#)]
233. Steigenberger, B.; Albanese, P.; Heck, A.J.R.; Scheltema, R.A. To Cleave or Not To Cleave in XL-MS? *J. Am. Soc. Mass Spectrom.* **2020**, *31*, 196–206. [[CrossRef](#)] [[PubMed](#)]
234. Xia, L.; Ma, Z.; Tong, J.; Tang, Y.; Li, S.; Qin, S.; Lou, R.; Zhao, S.; Lei, X.; Shui, W. Evaluation of chemical cross-linkers for in-depth structural analysis of G protein-coupled receptors through cross-linking mass spectrometry. *Anal. Chim. Acta* **2020**, *1102*, 53–62. [[CrossRef](#)]
235. Jones, A.X.; Cao, Y.; Tang, Y.-L.; Wang, J.-H.; Ding, Y.-H.; Tan, H.; Chen, Z.-L.; Fang, R.-Q.; Yin, J.; Chen, R.-C.; et al. Improving mass spectrometry analysis of protein structures with arginine-selective chemical cross-linkers. *Nat. Commun.* **2019**, *10*, 3911. [[CrossRef](#)]
236. Rosenbaum, D.M.; Cherezov, V.; Hanson, M.A.; Rasmussen, S.G.F.; Thian, F.S.; Kobilka, T.S.; Choi, H.-J.; Yao, X.-J.; Weis, W.I.; Stevens, R.C.; et al. GPCR engineering yields high-resolution structural insights into beta2-adrenergic receptor function. *Science* **2007**, *318*, 1266–1273. [[CrossRef](#)]
237. Magnani, F.; Serrano-Vega, M.J.; Shibata, Y.; Abdul-Hussein, S.; Lebon, G.; Miller-Gallacher, J.; Singhal, A.; Strege, A.; Thomas, J.A.; Tate, C.G. A mutagenesis and screening strategy to generate optimally thermostabilized membrane proteins for structural studies. *Nat. Protoc.* **2016**, *11*, 1554–1571. [[CrossRef](#)] [[PubMed](#)]
238. Ciancetta, A.; Rubio, P.; Lieberman, D.I.; Jacobson, K.A. A₃ adenosine receptor activation mechanisms: Molecular dynamics analysis of inactive, active, and fully active states. *J. Comput. Aided Mol. Des.* **2019**, *33*, 983–996. [[CrossRef](#)] [[PubMed](#)]
239. Venkatakrisnan, A.J.; Deupi, X.; Lebon, G.; Tate, C.G.; Schertler, G.F.; Babu, M.M. Molecular signatures of G-protein-coupled receptors. *Nature* **2013**, *494*, 185–194. [[CrossRef](#)]
240. Rucktooa, P.; Cheng, R.K.Y.; Segala, E.; Geng, T.; Errey, J.C.; Brown, G.A.; Cooke, R.M.; Marshall, F.H.; Doré, A.S. Towards high throughput GPCR crystallography: In Meso soaking of Adenosine A_{2A} Receptor crystals. *Sci. Rep.* **2018**, *8*, 41. [[CrossRef](#)]
241. García-Nafria, J.; Lee, Y.; Bai, X.; Carpenter, B.; Tate, C.G. Cryo-EM structure of the adenosine A_{2A} receptor coupled to an engineered heterotrimeric G protein. *eLife* **2018**, *7*, e35946. [[CrossRef](#)]
242. Amelia, T.; van Veldhoven, J.P.D.; Falsini, M.; Liu, R.; Heitman, L.H.; van Westen, G.J.P.; Segala, E.; Verdon, G.; Cheng, R.K.Y.; Cooke, R.M.; et al. Crystal Structure and Subsequent Ligand Design of a Nonriboside Partial Agonist Bound to the Adenosine A_{2A} Receptor. *J. Med. Chem.* **2021**, *64*, 3827–3842. [[CrossRef](#)] [[PubMed](#)]
243. Venkatakrisnan, A.J.; Deupi, X.; Lebon, G.; Heydenreich, F.M.; Flock, T.; Miljus, T.; Balaji, S.; Bouvier, M.; Veprintsev, D.B.; Tate, C.G.; et al. Diverse activation pathways in class A GPCRs converge near the G-protein-coupling region. *Nature* **2016**, *536*, 484–487. [[CrossRef](#)]
244. Shiriaeva, A.; Park, D.; Kim, G.; Lee, Y.; Hou, X.; Jarhad, D.B.; Kim, G.; Yu, J.; Hyun, Y.E.; Kim, W.; et al. GPCR Agonist-to-Antagonist Conversion: Enabling the Design of Nucleoside Functional Switches for the A_{2A} Adenosine Receptor. *J. Med. Chem.* **2022**, *65*, 11648–11657. [[CrossRef](#)] [[PubMed](#)]
245. Sun, B.; Bachhawat, P.; Chu, M.L.-H.; Wood, M.; Ceska, T.; Sands, Z.A.; Mercier, J.; Lebon, F.; Kobilka, T.S.; Kobilka, B.K. Crystal structure of the adenosine A_{2A} receptor bound to an antagonist reveals a potential allosteric pocket. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 2066–2071. [[CrossRef](#)]
246. Martin-Garcia, J.M.; Conrad, C.E.; Nelson, G.; Stander, N.; Zatsepin, N.A.; Zook, J.; Zhu, L.; Geiger, J.; Chun, E.; Kissick, D.; et al. Serial millisecond crystallography of membrane and soluble protein microcrystals using synchrotron radiation. *IUCr* **2017**, *4*, 439–454. [[CrossRef](#)] [[PubMed](#)]
247. Weinert, T.; Olieric, N.; Cheng, R.; Brünle, S.; James, D.; Ozerov, D.; Gashi, D.; Vera, L.; Marsh, M.; Jaeger, K.; et al. Serial millisecond crystallography for routine room-temperature structure determination at synchrotrons. *Nat. Commun.* **2017**, *8*, 542. [[CrossRef](#)]
248. Martin-Garcia, J.M.; Zhu, L.; Mendez, D.; Lee, M.-Y.; Chun, E.; Li, C.; Hu, H.; Subramanian, G.; Kissick, D.; Ogata, C.; et al. High-viscosity injector-based pink-beam serial crystallography of microcrystals at a synchrotron radiation source. *IUCr* **2019**, *6*, 412–425. [[CrossRef](#)]
249. Ishchenko, A.; Stauch, B.; Han, G.W.; Batyuk, A.; Shiriaeva, A.; Li, C.; Zatsepin, N.; Weierstall, U.; Liu, W.; Nango, E.; et al. Toward G protein-coupled receptor structure-based drug design using X-ray lasers. *IUCr* **2019**, *6*, 1106–1119. [[CrossRef](#)] [[PubMed](#)]

250. Shimazu, Y.; Tono, K.; Tanaka, T.; Yamanaka, Y.; Nakane, T.; Mori, C.; Terakado Kimura, K.; Fujiwara, T.; Sugahara, M.; Tanaka, R.; et al. High-viscosity sample-injection device for serial femtosecond crystallography at atmospheric pressure. *J. Appl. Crystallogr.* **2019**, *52*, 1280–1288. [[CrossRef](#)]
251. Ihara, K.; Hato, M.; Nakane, T.; Yamashita, K.; Kimura-Someya, T.; Hosaka, T.; Ishizuka-Katsura, Y.; Tanaka, R.; Tanaka, T.; Sugahara, M.; et al. Isoprenoid-chained lipid EROCO17+4: A new matrix for membrane protein crystallization and a crystal delivery medium in serial femtosecond crystallography. *Sci. Rep.* **2020**, *10*, 19305. [[CrossRef](#)]
252. García-Nafria, J.; Tate, C.G. Structure determination of GPCRs: Cryo-EM compared with X-ray crystallography. *Biochem. Soc. Trans.* **2021**, *49*, 2345–2355. [[CrossRef](#)] [[PubMed](#)]
253. Vénien-Bryan, C.; Li, Z.; Vuillard, L.; Boutin, J.A. Cryo-electron microscopy and X-ray crystallography: Complementary approaches to structural biology and drug discovery. *Acta Crystallogr. F Struct. Biol. Commun.* **2017**, *73*, 174–183. [[CrossRef](#)] [[PubMed](#)]
254. Boutin, J.A.; Li, Z.; Vuillard, L.; Vénien-Bryan, C. La cryo-microscopie, une alternative à la cristallographie aux rayons X? *Med. Sci.* **2016**, *32*, 758–767. [[CrossRef](#)] [[PubMed](#)]
255. Rossmann, M.G.; Battisti, A.J.; Plevka, P. Future prospects. *Adv. Protein Chem. Struct. Biol.* **2011**, *82*, 101–121. [[CrossRef](#)]
256. Blundell, T.L.; Chaplin, A.K. The resolution revolution in X-ray diffraction, Cryo-EM and other Technologies. *Prog. Biophys. Mol. Biol.* **2021**, *160*, 2–4. [[CrossRef](#)]
257. Glukhova, A.; Draper-Joyce, C.J.; Sunahara, R.K.; Christopoulos, A.; Wootten, D.; Sexton, P.M. Rules of Engagement: GPCRs and G Proteins. *ACS Pharmacol. Transl. Sci.* **2018**, *1*, 73–83. [[CrossRef](#)]
258. Liang, Y.-L.; Khoshouei, M.; Radjainia, M.; Zhang, Y.; Glukhova, A.; Tarrasch, J.; Thal, D.M.; Furness, S.G.B.; Christopoulos, G.; Coudrat, T.; et al. Phase-plate cryo-EM structure of a class B GPCR-G-protein complex. *Nature* **2017**, *546*, 118–123. [[CrossRef](#)]
259. Martynowycz, M.W.; Shiriaeva, A.; Ge, X.; Hattne, J.; Nannenga, B.L.; Cherezov, V.; Gonen, T. MicroED structure of the human adenosine receptor determined from a single nanocrystal in LCP. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2106041118. [[CrossRef](#)]
260. Frauenfelder, H.; Sligar, S.G.; Wolynes, P.G. The energy landscapes and motions of proteins. *Science* **1991**, *254*, 1598–1603. [[CrossRef](#)]
261. Kitevski-LeBlanc, J.L.; Prosser, R.S. Current applications of 19F NMR to studies of protein structure and dynamics. *Prog. Nucl. Magn. Reson. Spectrosc.* **2012**, *62*, 1–33. [[CrossRef](#)]
262. Picard, L.-P.; Prosser, R.S. Advances in the study of GPCRs by 19F NMR. *Curr. Opin. Struct. Biol.* **2021**, *69*, 169–176. [[CrossRef](#)]
263. Liu, D.; Wüthrich, K. Ring current shifts in (19)F-NMR of membrane proteins. *J. Biomol. NMR* **2016**, *65*, 1–5. [[CrossRef](#)] [[PubMed](#)]
264. Hagn, F.; Etzkorn, M.; Raschle, T.; Wagner, G. Optimized phospholipid bilayer nanodiscs facilitate high-resolution structure determination of membrane proteins. *J. Am. Chem. Soc.* **2013**, *135*, 1919–1925. [[CrossRef](#)]
265. Pozza, A.; Giraud, F.; Cece, Q.; Casiraghi, M.; Point, E.; Damian, M.; Le Bon, C.; Moncoq, K.; Banères, J.-L.; Lescop, E.; et al. Exploration of the dynamic interplay between lipids and membrane proteins by hydrostatic pressure. *Nat. Commun.* **2022**, *13*, 1780. [[CrossRef](#)] [[PubMed](#)]
266. Casiraghi, M.; Damian, M.; Lescop, E.; Point, E.; Moncoq, K.; Morellet, N.; Levy, D.; Marie, J.; Guittet, E.; Banères, J.-L.; et al. Functional Modulation of a G Protein-Coupled Receptor Conformational Landscape in a Lipid Bilayer. *J. Am. Chem. Soc.* **2016**, *138*, 11170–11175. [[CrossRef](#)]
267. Nygaard, R.; Frimurer, T.M.; Holst, B.; Rosenkilde, M.M.; Schwartz, T.W. Ligand binding and micro-switches in 7TM receptor structures. *Trends Pharmacol. Sci.* **2009**, *30*, 249–259. [[CrossRef](#)]
268. Eddy, M.T.; Gao, Z.-G.; Mannes, P.; Patel, N.; Jacobson, K.A.; Katritch, V.; Stevens, R.C.; Wüthrich, K. Extrinsic Tryptophans as NMR Probes of Allosteric Coupling in Membrane Proteins: Application to the A_{2A} Adenosine Receptor. *J. Am. Chem. Soc.* **2018**, *140*, 8228–8235. [[CrossRef](#)]
269. Eddy, M.T.; Martin, B.T.; Wüthrich, K. A_{2A} Adenosine Receptor Partial Agonism Related to Structural Rearrangements in an Activation Microswitch. *Structure* **2021**, *29*, 170–176.e3. [[CrossRef](#)] [[PubMed](#)]
270. Casiraghi, M.; Damian, M.; Lescop, E.; Banères, J.-L.; Catoire, L.J. Illuminating the Energy Landscape of GPCRs: The Key Contribution of Solution-State NMR Associated with *Escherichia coli* as an Expression Host. *Biochemistry* **2018**, *57*, 2297–2307. [[CrossRef](#)]
271. Sekhar, A.; Kay, L.E. NMR paves the way for atomic level descriptions of sparsely populated, transiently formed biomolecular conformers. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 12867–12874. [[CrossRef](#)]
272. Gunkel, M.; Schöneberg, J.; Alkhalidi, W.; Irsen, S.; Noé, F.; Kaupp, U.B.; Al-Amoudi, A. Higher-order architecture of rhodopsin in intact photoreceptors and its implication for phototransduction kinetics. *Structure* **2015**, *23*, 628–638. [[CrossRef](#)]
273. Fotiadis, D.; Liang, Y.; Filipek, S.; Saperstein, D.A.; Engel, A.; Palczewski, K. Atomic-force microscopy: Rhodopsin dimers in native disc membranes. *Nature* **2003**, *421*, 127–128. [[CrossRef](#)]
274. Sapra, K.T.; Spoerri, P.M.; Engel, A.; Alsteens, D.; Müller, D.J. Seeing and sensing single G protein-coupled receptors by atomic force microscopy. *Curr. Opin. Cell Biol.* **2019**, *57*, 25–32. [[CrossRef](#)] [[PubMed](#)]
275. Jacobson, K.A.; Tosh, D.K.; Jain, S.; Gao, Z.-G. Historical and Current Adenosine Receptor Agonists in Preclinical and Clinical Development. *Front. Cell. Neurosci.* **2019**, *13*, 124. [[CrossRef](#)]
276. Melman, A.; Gao, Z.-G.; Kumar, D.; Wan, T.C.; Gizewski, E.; Auchampach, J.A.; Jacobson, K.A. Design of (N)-methanocarba adenosine 5'-uronamides as species-independent A₃ receptor-selective agonists. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 2813–2819. [[CrossRef](#)]

277. Ivanov, A.A.; Jacobson, K.A. Molecular modeling of a PAMAM-CGS21680 dendrimer bound to an A_{2A} adenosine receptor homodimer. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 4312–4315. [[CrossRef](#)] [[PubMed](#)]
278. Dal Ben, D.; Lambertucci, C.; Buccioni, M.; Martí Navia, A.; Marucci, G.; Spinaci, A.; Volpini, R. Non-Nucleoside Agonists of the Adenosine Receptors: An Overview. *Pharmaceuticals* **2019**, *12*, 150. [[CrossRef](#)] [[PubMed](#)]
279. Müller, C.E.; Jacobson, K.A. Xanthines as adenosine receptor antagonists. *Handb. Exp. Pharmacol.* **2011**, *200*, 151–199. [[CrossRef](#)]
280. Ukena, D.; Padgett, W.L.; Hong, O.; Daly, J.W.; Daly, D.T.; Olsson, R.A. N6-substituted 9-methyladenines: A new class of adenosine receptor antagonists. *FEBS Lett.* **1987**, *215*, 203–208. [[CrossRef](#)]
281. Zhou, X.; Khanapur, S.; Huizing, A.P.; Zijlma, R.; Schepers, M.; Dierckx, R.A.J.O.; van Waarde, A.; de Vries, E.F.J.; Elsinga, P.H. Synthesis and preclinical evaluation of 2-(2-furanyl)-7-2-4-4-(2-11Cmethoxyethoxy)phenyl-1-piperazineethyl-7H-pyrazolo[4,3-e]1,2,4-triazolo[1,5-c]pyrimidine-5-amine (11CPreladenant) as a PET tracer for the imaging of cerebral adenosine A_{2A} receptors. *J. Med. Chem.* **2014**, *57*, 9204–9210. [[CrossRef](#)]
282. Shinkre, B.A.; Kumar, T.S.; Gao, Z.-G.; Deflorian, F.; Jacobson, K.A.; Trenkle, W.C. Synthesis and evaluation of 1,2,4-triazolo[1,5-c]pyrimidine derivatives as A_{2A} receptor-selective antagonists. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 5690–5694. [[CrossRef](#)] [[PubMed](#)]
283. Todde, S.; Moresco, R.M.; Simonelli, P.; Baraldi, P.G.; Cacciari, B.; Spalluto, G.; Varani, K.; Monopoli, A.; Matarrese, M.; Carpinelli, A.; et al. Design, radiosynthesis, and biodistribution of a new potent and selective ligand for in vivo imaging of the adenosine A_{2A} receptor system using positron emission tomography. *J. Med. Chem.* **2000**, *43*, 4359–4362. [[CrossRef](#)] [[PubMed](#)]
284. Zocchi, C.; Ongini, E.; Conti, A.; Monopoli, A.; Negretti, A.; Baraldi, P.G.; Dionisotti, S. The non-xanthine heterocyclic compound SCH 58261 is a new potent and selective A_{2A} adenosine receptor antagonist. *J. Pharmacol. Exp. Ther.* **1996**, *276*, 398–404. [[PubMed](#)]
285. Marques, T.R.; Natesan, S.; Rabiner, E.A.; Searle, G.E.; Gunn, R.; Howes, O.D.; Kapur, S. Adenosine A_{2A} receptor in schizophrenia: An in vivo brain PET imaging study. *Psychopharmacology* **2021**, *239*, 3439–3445. [[CrossRef](#)]
286. Boutin, J.A.; Witt-Enderby, P.A.; Sotriffer, C.; Zlotos, D.P. Melatonin receptor ligands: A pharmaco-chemical perspective. *J. Pineal Res.* **2020**, *69*, e12672. [[CrossRef](#)]
287. Bojarski, A.J. Pharmacophore models for metabotropic 5-HT receptor ligands. *Curr. Top. Med. Chem.* **2006**, *6*, 2005–2026. [[CrossRef](#)]
288. Gao, Z.-G.; Jacobson, K.A. Allosteric modulation and functional selectivity of G protein-coupled receptors. *Drug Discov. Today Technol.* **2013**, *10*, e237–e243. [[CrossRef](#)]
289. Lütjens, R.; Rocher, J.-P. Recent advances in drug discovery of GPCR allosteric modulators for neurodegenerative disorders. *Curr. Opin. Pharmacol.* **2017**, *32*, 91–95. [[CrossRef](#)]
290. van der Westhuizen, E.T.; Valant, C.; Sexton, P.M.; Christopoulos, A. Endogenous allosteric modulators of G protein-coupled receptors. *J. Pharmacol. Exp. Ther.* **2015**, *353*, 246–260. [[CrossRef](#)]
291. Madsen, J.J.; Ye, L.; Frimurer, T.M.; Olsen, O.H. Mechanistic basis of GPCR activation explored by ensemble refinement of crystallographic structures. *Protein Sci.* **2022**, *31*, e4456. [[CrossRef](#)]
292. Swaminath, G.; Lee, T.W.; Kobilka, B. Identification of an allosteric binding site for Zn²⁺ on the beta2 adrenergic receptor. *J. Biol. Chem.* **2003**, *278*, 352–356. [[CrossRef](#)]
293. Lu, Y.; Liu, H.; Yang, D.; Zhong, L.; Xin, Y.; Zhao, S.; Wang, M.-W.; Zhou, Q.; Shui, W. Affinity Mass Spectrometry-Based Fragment Screening Identified a New Negative Allosteric Modulator of the Adenosine A_{2A} Receptor Targeting the Sodium Ion Pocket. *ACS Chem. Biol.* **2021**, *16*, 991–1002. [[CrossRef](#)] [[PubMed](#)]
294. Korkutata, M.; Saitoh, T.; Cherasse, Y.; Ioka, S.; Duo, F.; Qin, R.; Murakoshi, N.; Fujii, S.; Zhou, X.; Sugiyama, F.; et al. Enhancing endogenous adenosine A_{2A} receptor signaling induces slow-wave sleep without affecting body temperature and cardiovascular function. *Neuropharmacology* **2019**, *144*, 122–132. [[CrossRef](#)] [[PubMed](#)]
295. Welihinda, A.A.; Amento, E.P. Positive allosteric modulation of the adenosine A_{2A} receptor attenuates inflammation. *J. Inflamm.* **2014**, *11*, 37. [[CrossRef](#)] [[PubMed](#)]
296. Welihinda, A.A.; Kaur, M.; Raveendran, K.S.; Amento, E.P. Enhancement of inosine-mediated A_{2A}R signaling through positive allosteric modulation. *Cell. Signal.* **2018**, *42*, 227–235. [[CrossRef](#)]
297. Salovich, J.M.; Sheffler, D.J.; Vinson, P.N.; Lamsal, A.; Utley, T.J.; Blobaum, A.L.; Bridges, T.M.; Le, U.; Jones, C.K.; Wood, M.R.; et al. *Probe Reports from the NIH Molecular Libraries Program: Discovery of a Novel Structural Class of M4 Positive Allosteric Modulators: Characterization of ML293, N-(4-Methoxy-7-Methylbenzodiazol-2-yl)isonicotinamide, with CNS Exposure in Rats*; National Center for Biotechnology Information (US): Bethesda, MD, USA, 2010.
298. Jones, C.K.; Bubser, M.; Thompson, A.D.; Dickerson, J.W.; Turle-Lorenzo, N.; Amalric, M.; Blobaum, A.L.; Bridges, T.M.; Morrison, R.D.; Jadhav, S.; et al. The metabotropic glutamate receptor 4-positive allosteric modulator VU0364770 produces efficacy alone and in combination with L-DOPA or an adenosine 2A antagonist in preclinical rodent models of Parkinson's disease. *J. Pharmacol. Exp. Ther.* **2012**, *340*, 404–421. [[CrossRef](#)] [[PubMed](#)]
299. Pinhal-Enfield, G.; Ramanathan, M.; Hasko, G.; Vogel, S.N.; Salzman, A.L.; Boons, G.-J.; Leibovich, S.J. An Angiogenic Switch in Macrophages Involving Synergy between Toll-Like Receptors 2, 4, 7, and 9 and Adenosine A_{2A} Receptors. *Am. J. Pathol.* **2003**, *163*, 711–721. [[CrossRef](#)]
300. Jung, S.-M.; Peyton, L.; Essa, H.; Choi, D.-S. Adenosine receptors: Emerging non-opioids targets for pain medications. *Neurobiol. Pain* **2022**, *11*, 100087. [[CrossRef](#)]
301. Federico, S.; Lassiani, L.; Spalluto, G. Chemical Probes for the Adenosine Receptors. *Pharmaceuticals* **2019**, *12*, 168. [[CrossRef](#)]

302. Soave, M.; Briddon, S.J.; Hill, S.J.; Stoddart, L.A. Fluorescent ligands: Bringing light to emerging GPCR paradigms. *Br. J. Pharmacol.* **2020**, *177*, 978–991. [[CrossRef](#)]
303. Kozma, E.; Jayasekara, P.S.; Squarcialupi, L.; Paoletta, S.; Moro, S.; Federico, S.; Spalluto, G.; Jacobson, K.A. Fluorescent ligands for adenosine receptors. *Bioorg. Med. Chem. Lett.* **2013**, *23*, 26–36. [[CrossRef](#)]
304. McCabe, R.T.; Skolnick, P.; Jacobson, K.A. 2-2-4-2-2-1,3-Dihydro- 1,1-bis (4-hydroxyphenyl)-3-oxo-5-isobenzofuranthioureydi lethylaminocarbonylethylphenyl ethylamino-5'-N-ethylcarboxamidoadenosine (FITC-APEC): A Fluorescent Ligand for A_{2A}-Adenosine Receptors. *J. Fluoresc.* **1992**, *2*, 217–223. [[CrossRef](#)]
305. Brand, F.; Klutz, A.M.; Jacobson, K.A.; Fredholm, B.B.; Schulte, G. Adenosine A_{2A} receptor dynamics studied with the novel fluorescent agonist Alexa488-APEC. *Eur. J. Pharmacol.* **2008**, *590*, 36–42. [[CrossRef](#)]
306. Comeo, E.; Kondon, N.D.; Soave, M.; Stoddart, L.A.; Kilpatrick, L.E.; Scammells, P.J.; Hill, S.J.; Kellam, B. Subtype-Selective Fluorescent Ligands as Pharmacological Research Tools for the Human Adenosine A_{2A} Receptor. *J. Med. Chem.* **2020**, *63*, 2656–2672. [[CrossRef](#)]
307. Toti, K.S.; Campbell, R.G.; Lee, H.; Salmaso, V.; Suresh, R.R.; Gao, Z.-G.; Jacobson, K.A. Fluorescent A_{2A} and A₃ adenosine receptor antagonists as flow cytometry probes. *Purinergic Signal.* **2022**. [[CrossRef](#)] [[PubMed](#)]
308. McNeely, P.M.; Naranjo, A.N.; Forsten-Williams, K.; Robinson, A.S. A_{2A}R Binding Kinetics in the Ligand Depletion Regime. *SLAS Discov.* **2017**, *22*, 166–175. [[CrossRef](#)] [[PubMed](#)]
309. Rieger, J.M.; Brown, M.L.; Sullivan, G.W.; Linden, J.; Macdonald, T.L. Design, synthesis, and evaluation of novel A_{2A} adenosine receptor agonists. *J. Med. Chem.* **2001**, *44*, 531–539. [[CrossRef](#)]
310. Chen, J.-B.; Liu, E.M.; Chern, T.-R.; Yang, C.-W.; Lin, C.-I.; Huang, N.-K.; Lin, Y.-L.; Chern, Y.; Lin, J.-H.; Fang, J.-M. Design and synthesis of novel dual-action compounds targeting the adenosine A_{2A} receptor and adenosine transporter for neuroprotection. *ChemMedChem* **2011**, *6*, 1390–1400. [[CrossRef](#)] [[PubMed](#)]
311. Yuan, G.; Gedeon, N.G.; Jankins, T.C.; Jones, G.B. Novel approaches for targeting the adenosine A_{2A} receptor. *Expert Opin. Drug Discov.* **2015**, *10*, 63–80. [[CrossRef](#)] [[PubMed](#)]
312. Williams, R.; Hosila, F.J. The effect of different extraction sites upon incisor retraction. *Am. J. Orthod.* **1976**, *69*, 388–410. [[CrossRef](#)]
313. Jazayeri, A.; Andrews, S.P.; Marshall, F.H. Structurally Enabled Discovery of Adenosine A_{2A} Receptor Antagonists. *Chem. Rev.* **2017**, *117*, 21–37. [[CrossRef](#)]
314. Hu, E.; Kunz, R.K.; Rumpf, S.; Andrews, K.L.; Li, C.; Hitchcock, S.A.; Lindstrom, M.; Treanor, J. Use of structure based design to increase selectivity of pyridyl-cinnoline phosphodiesterase 10A (PDE10A) inhibitors against phosphodiesterase 3 (PDE3). *Bioorg. Med. Chem. Lett.* **2012**, *22*, 6938–6942. [[CrossRef](#)]
315. Deganutti, G.; Moro, S. Supporting the Identification of Novel Fragment-Based Positive Allosteric Modulators Using a Supervised Molecular Dynamics Approach: A Retrospective Analysis Considering the Human A_{2A} Adenosine Receptor as a Key Example. *Molecules* **2017**, *22*, 818. [[CrossRef](#)] [[PubMed](#)]
316. McGraw, C.; Yang, L.; Levental, I.; Lyman, E.; Robinson, A.S. Membrane cholesterol depletion reduces downstream signaling activity of the adenosine A_{2A} receptor. *Biochim. Biophys. Acta Biomembr.* **2019**, *1861*, 760–767. [[CrossRef](#)] [[PubMed](#)]
317. Yuan, S.; Hu, Z.; Filipek, S.; Vogel, H. W246(6.48) opens a gate for a continuous intrinsic water pathway during activation of the adenosine A_{2A} receptor. *Angew. Chem. Int. Ed Engl.* **2015**, *54*, 556–559. [[CrossRef](#)] [[PubMed](#)]
318. Guo, D.; Pan, A.C.; Dror, R.O.; Mocking, T.; Liu, R.; Heitman, L.H.; Shaw, D.E.; IJzerman, A.P. Molecular Basis of Ligand Dissociation from the Adenosine A_{2A} Receptor. *Mol. Pharmacol.* **2016**, *89*, 485–491. [[CrossRef](#)] [[PubMed](#)]
319. Mondal, C.; Halder, A.K.; Adhikari, N.; Jha, T. Structural findings of cinnolines as anti-schizophrenic PDE10A inhibitors through comparative chemometric modeling. *Mol. Divers.* **2014**, *18*, 655–671. [[CrossRef](#)]
320. Rodríguez, D.; Gao, Z.-G.; Moss, S.M.; Jacobson, K.A.; Carlsson, J. Molecular docking screening using agonist-bound GPCR structures: Probing the A_{2A} adenosine receptor. *J. Chem. Inf. Model.* **2015**, *55*, 550–563. [[CrossRef](#)]
321. Novikov, G.V.; Sivozhelezov, V.S.; Shaïtan, K.V. Investigation of the conformational dynamics of the adenosine A_{2A} receptor by means of molecular dynamics simulation. *Biofizika* **2013**, *58*, 618–634.
322. Kalash, L.; Winfield, I.; Safitri, D.; Bermudez, M.; Carvalho, S.; Glen, R.; Ladds, G.; Bender, A. Structure-based identification of dual ligands at the A_{2A}R and PDE10A with anti-proliferative effects in lung cancer cell-lines. *J. Cheminform.* **2021**, *13*, 17. [[CrossRef](#)] [[PubMed](#)]
323. Ng, H.W.; Laughton, C.A.; Doughty, S.W. Molecular dynamics simulations of the adenosine A_{2A} receptor: Structural stability, sampling, and convergence. *J. Chem. Inf. Model.* **2013**, *53*, 1168–1178. [[CrossRef](#)] [[PubMed](#)]
324. Hajduk, P.J.; Greer, J. A decade of fragment-based drug design: Strategic advances and lessons learned. *Nat. Rev. Drug Discov.* **2007**, *6*, 211–219. [[CrossRef](#)] [[PubMed](#)]
325. Chen, D.; Ranganathan, A.; IJzerman, A.P.; Siegal, G.; Carlsson, J. Complementarity between in silico and biophysical screening approaches in fragment-based lead discovery against the A_{2A} adenosine receptor. *J. Chem. Inf. Model.* **2013**, *53*, 2701–2714. [[CrossRef](#)]
326. Chen, D.; Errey, J.C.; Heitman, L.H.; Marshall, F.H.; IJzerman, A.P.; Siegal, G. Fragment screening of GPCRs using biophysical methods: Identification of ligands of the adenosine A_{2A} receptor with novel biological activity. *ACS Chem. Biol.* **2012**, *7*, 2064–2073. [[CrossRef](#)]

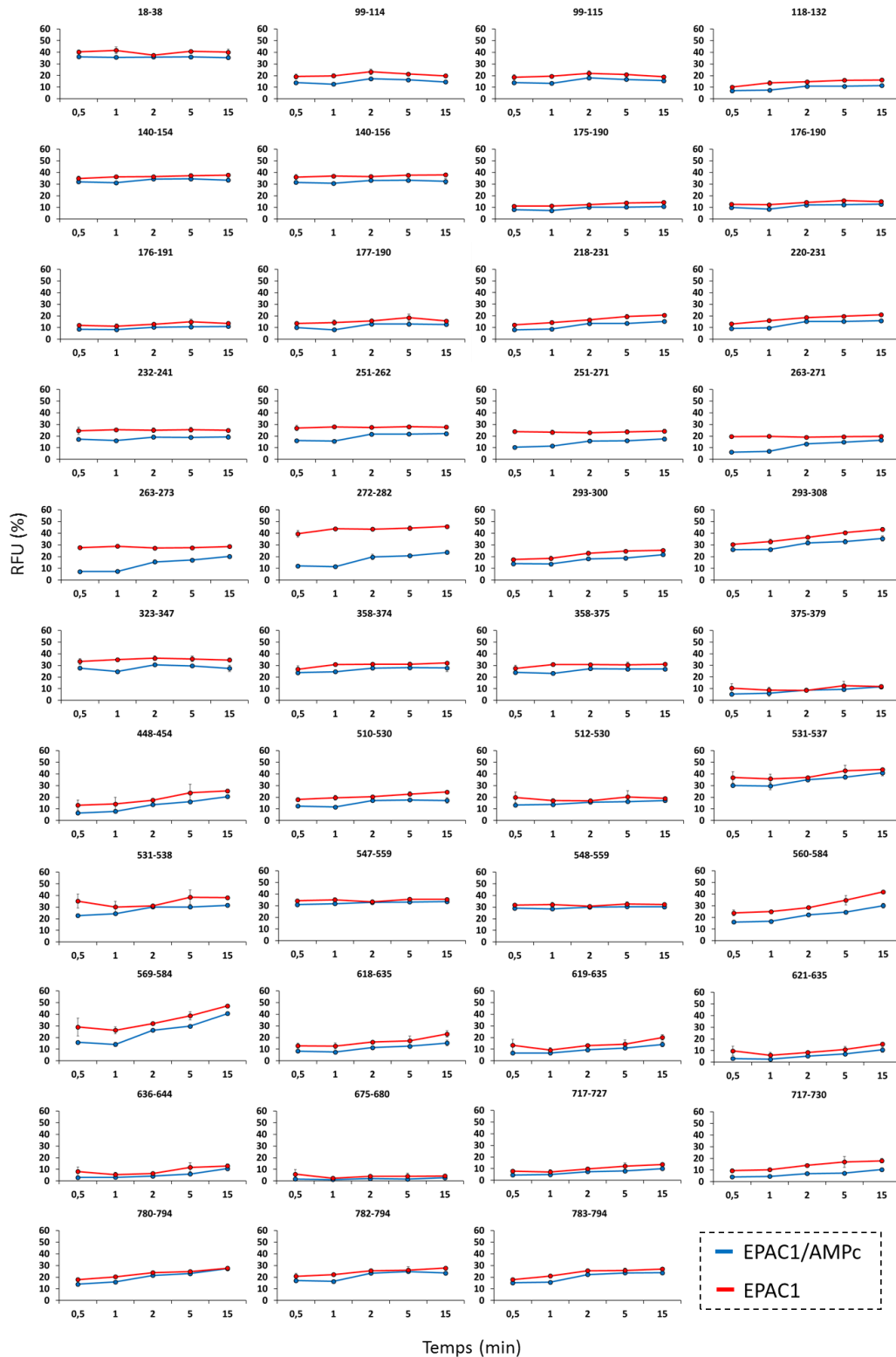
327. Matricon, P.; Ranganathan, A.; Warnick, E.; Gao, Z.-G.; Rudling, A.; Lambertucci, C.; Marucci, G.; Ezzati, A.; Jaiteh, M.; Dal Ben, D.; et al. Fragment optimization for GPCRs by molecular dynamics free energy calculations: Probing druggable sub-pockets of the A_{2A} adenosine receptor binding site. *Sci. Rep.* **2017**, *7*, 6398. [[CrossRef](#)]
328. Tosh, D.K.; Phan, K.; Gao, Z.-G.; Gakh, A.A.; Xu, F.; Deflorian, F.; Abagyan, R.; Stevens, R.C.; Jacobson, K.A.; Katritch, V. Optimization of adenosine 5'-carboxamide derivatives as adenosine receptor agonists using structure-based ligand design and fragment screening. *J. Med. Chem.* **2012**, *55*, 4297–4308. [[CrossRef](#)]
329. Zhang, M.; Fan, S.; Zhou, X.; Xie, F.; Li, S.; Zhong, W. Design, synthesis and biological evaluation of 2-hydrazinyladenosine derivatives as A_{2A} adenosine receptor ligands. *Eur. J. Med. Chem.* **2019**, *179*, 310–324. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

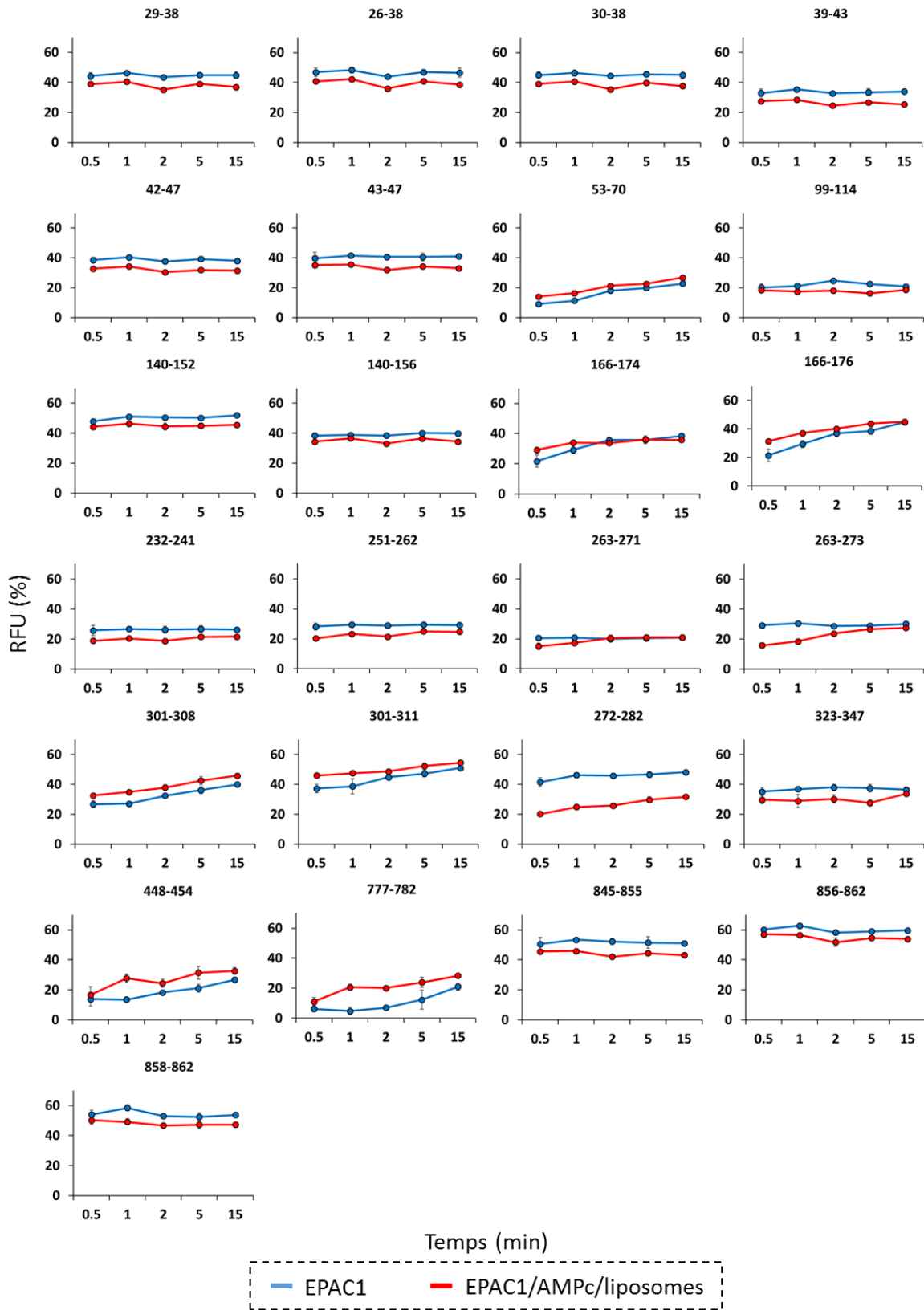
Annexe 33

Incorporation relative en D aux différents temps de deutération de chaque peptide de EPAC1 statistiquement significatif (MEMHDX avec une p-value à 0,01) obtenus en HDX-MS avec (A) EPAC1 – EPAC1/AMPc et (B) EPAC1 – EPAC1/AMPc/liposomes.

A) EPAC1 – EPAC1/AMPc

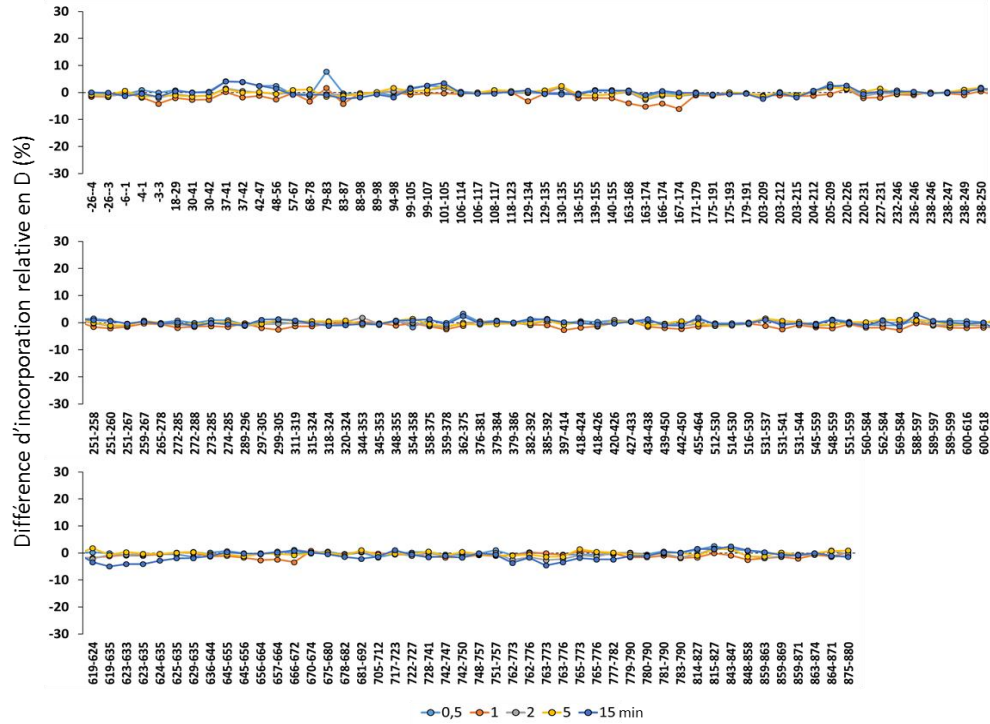


B) EPAC1 – EPAC1/AMPC/liposomes



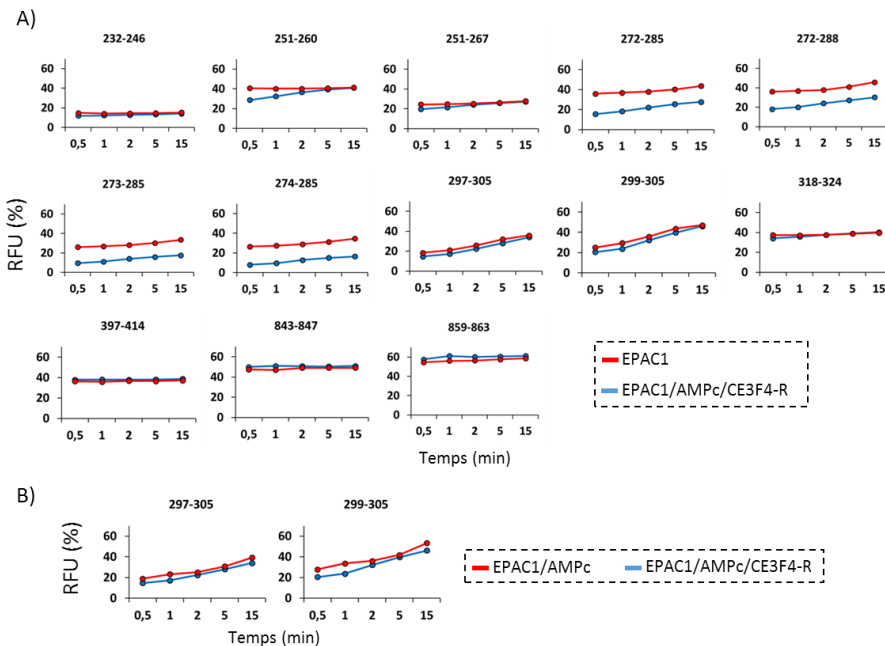
Annexe 34

Représentation des différences d'incorporation relative en D entre EPAC1 seule et en complexe avec AMPc/CE3F4-R, pour chaque peptide identifié de EPAC1.



Annexe 35

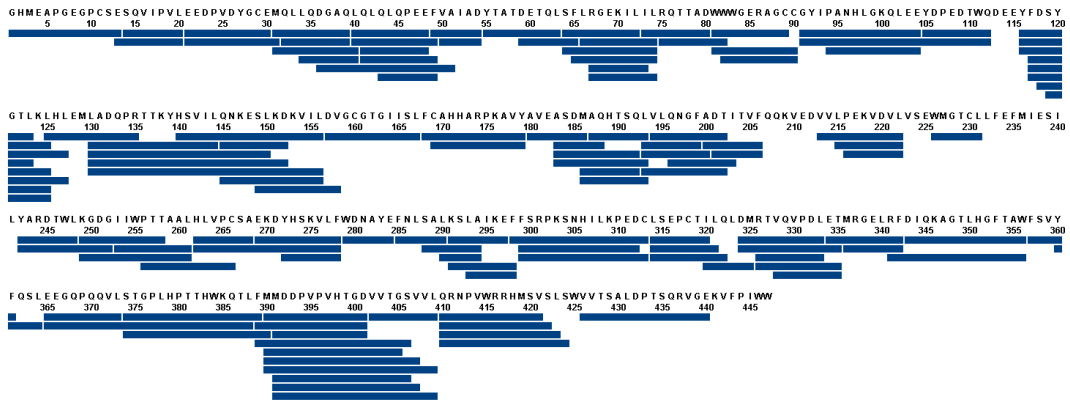
Incorporation relative en D aux différents temps de deutération de chaque peptide de EPAC1 statistiquement significatif (MEMHDX avec une p-value à 0,01) obtenus en HDX-MS avec (A) EPAC1 – EPAC1/AMPc/CE3F4-R et (B) EPAC1/AMPc – EPAC1/AMPc/CE3F4-R.



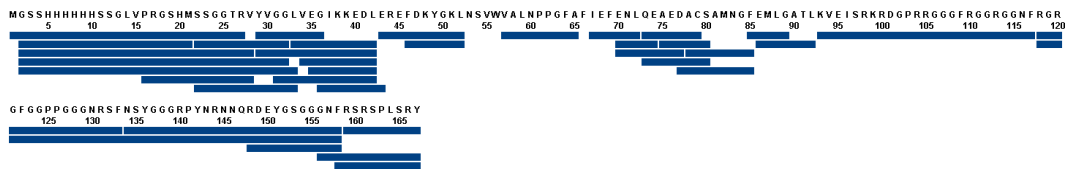
Annexe 36

Cartes de couverture de séquence obtenus en HDX-MS pour PRMT2, RSF1 et le complexe PRMT2/RSF1.

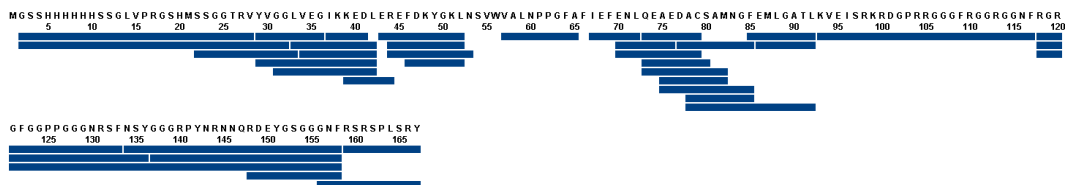
PRMT2 :



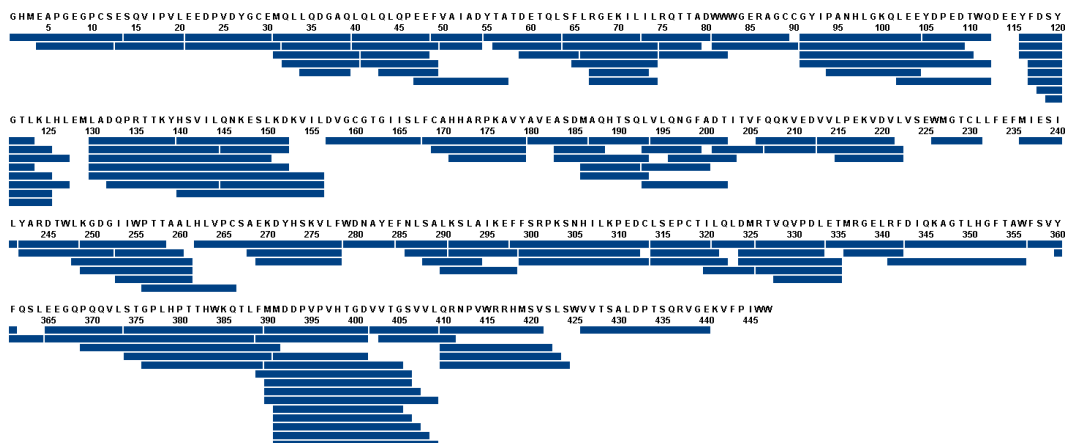
RSF1 :



Complexe PRMT2/RSF1 : RSF1

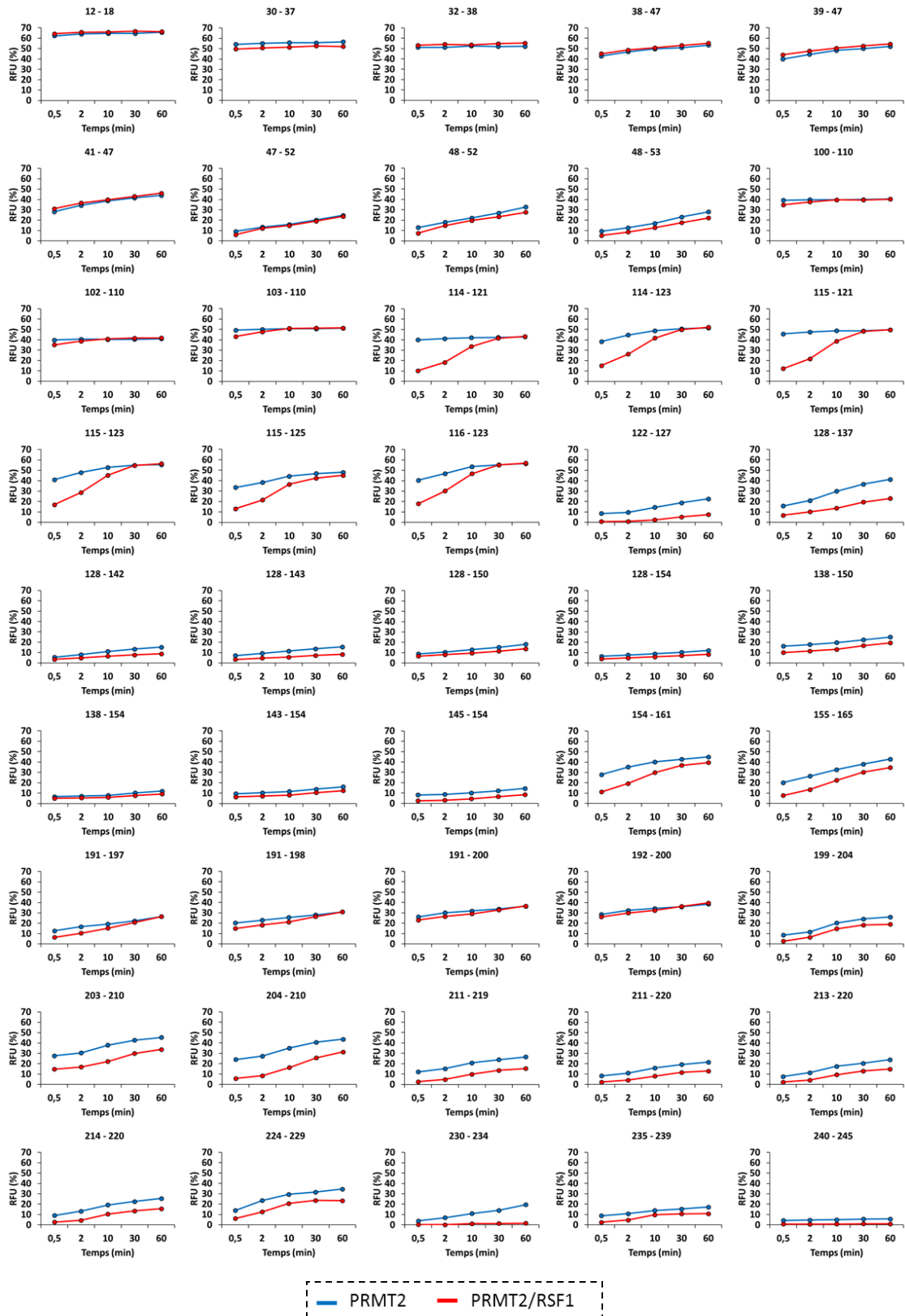


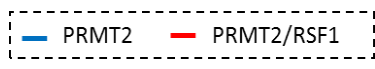
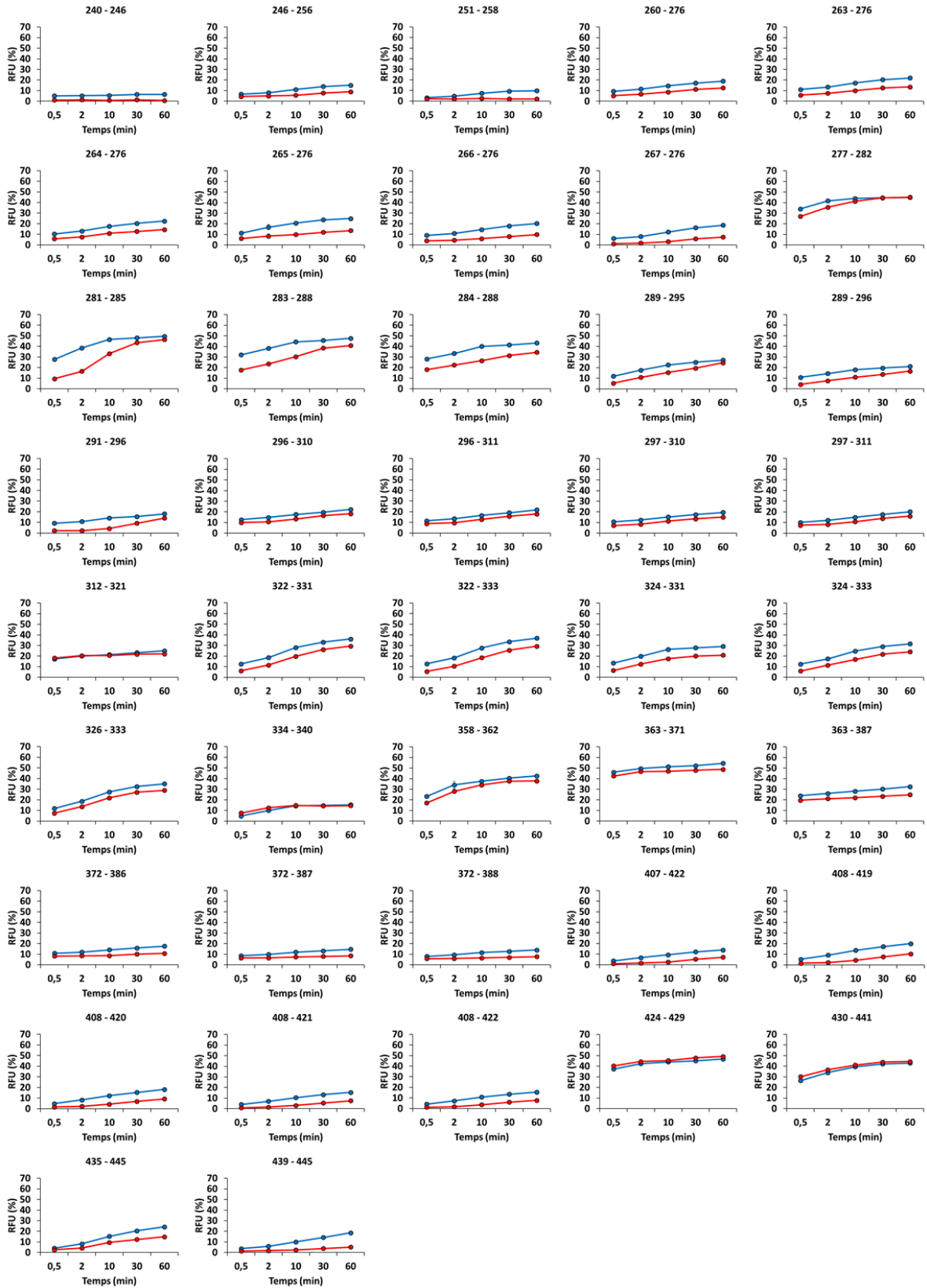
Complexe PRMT2/RSF1 : PRMT2



Annexe 37

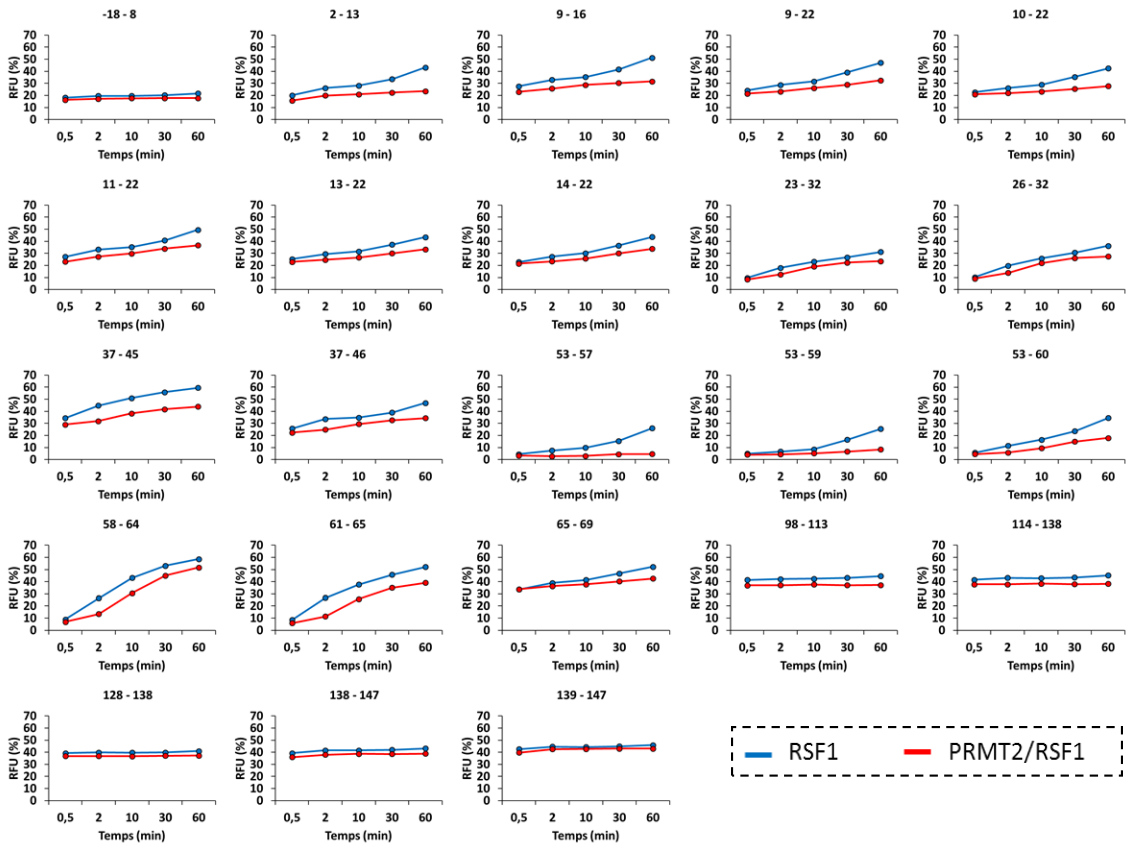
Incorporation relative en D aux différents temps de deutération de chaque peptide de PRMT2 statistiquement significatif (MEMHDX avec une p-value à 0,01) obtenus en HDX-MS (PRMT2-PRMT2/RSF1).





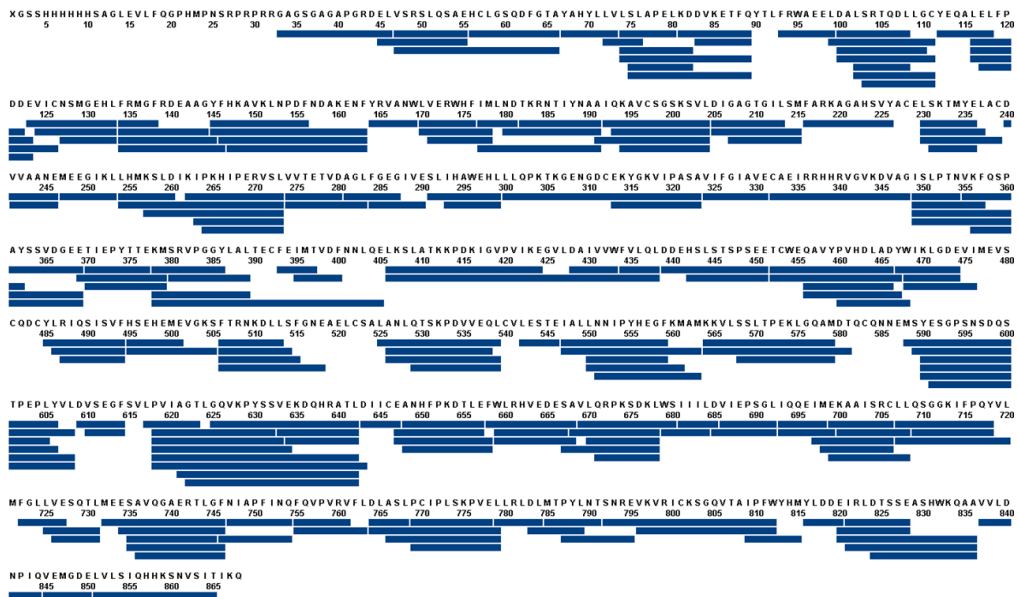
Annexe 38

Incorporation relative en D aux différents temps de deutération de chaque peptide de RSF1 statistiquement significatif (MEMHDX avec une p-value à 0,01) obtenus en HDX-MS (PRMT2-PRMT2/RSF1).



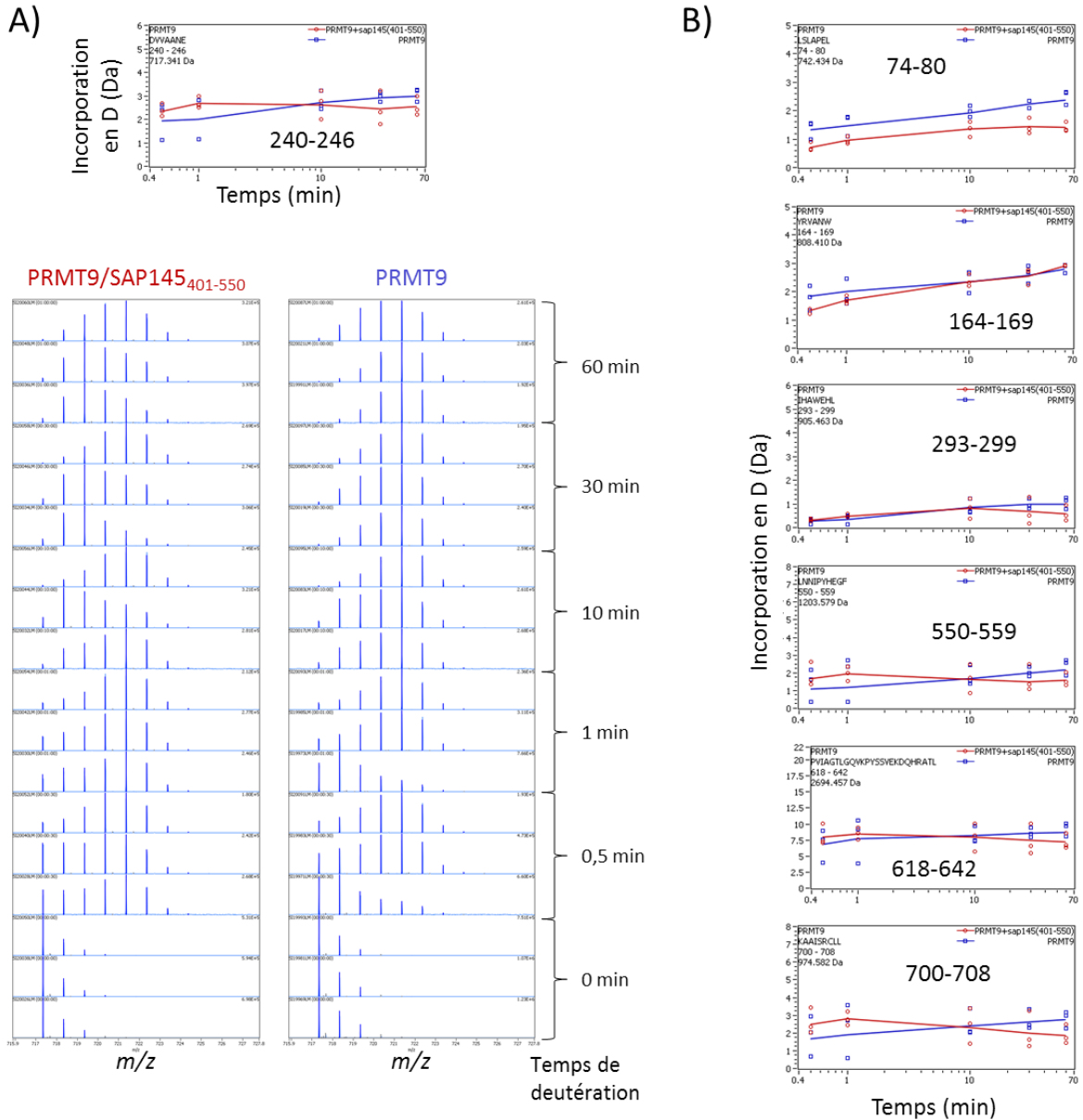
Annexe 39

Carte de couverture de séquence obtenus en HDX-MS pour PRMT9.



Annexe 40

(A) Incorporation en D (Da) du peptide 240-246 de PRMT9 aux différents temps de deutération et les profils isotopiques de chaque réplicats correspondant obtenus en HDX-MS (PRMT9-PRMT9/SAP145401-550). (B) Incorporation en D (Da) de certains peptides (74-80 ; 164-169 ; 293-299 ; 550-559 ; 618-642 ; 700-708) de PRMT9 aux différents temps de deutération obtenus en HDX-MS (PRMT9-PRMT9/SAP145401-550).



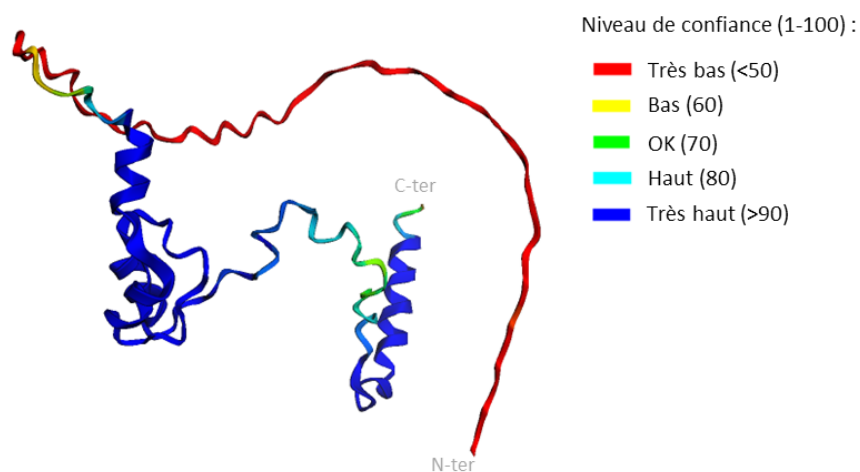
Annexe 41

Tableau descriptif des peptides intra-SAP145 cross-linkés (protéines, sites, cross-linker, domaines et distances Cα-Cα) identifiés suite au XL-MS du complexe PRMT9/SAP145₄₀₁₋₅₅₀.

| Protéine 1 | Protéine 2 | Site 1 | Site 2 | DSBU | DSAU | Domaine site 1 | Domaine site 2 | distance Cα-Cα modèle Alphafold |
|------------|------------|--------|--------|------|------|----------------|----------------|---------------------------------|
| SAP145 | SAP145 | N-ter | N-ter | | x | - | - | - |
| SAP145 | SAP145 | N-ter | K9 | x | | - | - | 27,2 |
| SAP145 | SAP145 | N-ter | K13 | x | | - | - | 41 |
| SAP145 | SAP145 | N-ter | K16 | x | | - | - | 50,4 |
| SAP145 | SAP145 | N-ter | K96 | x | x | - | - | 64,6 |
| SAP145 | SAP145 | N-ter | K118 | x | | - | - | 59,7 |
| SAP145 | SAP145 | K7 | K24 | x | | - | - | 52,8 |
| SAP145 | SAP145 | K9 | K24 | x | | - | - | 46,1 |
| SAP145 | SAP145 | K16 | K70 | | x | - | - | 48,9 |
| SAP145 | SAP145 | K16 | K96 | x | | - | - | 49,4 |
| SAP145 | SAP145 | K24 | K70 | x | | - | - | 45,5 |
| SAP145 | SAP145 | K70 | K133 | x | | - | - | 41,2 |
| SAP145 | SAP145 | K70 | K147 | x | x | - | - | 52,7 |
| SAP145 | SAP145 | K96 | K118 | x | x | - | - | 27,9 |
| SAP145 | SAP145 | K96 | K123 | x | | - | - | 27,8 |
| SAP145 | SAP145 | K96 | K133 | x | | - | - | 31 |
| SAP145 | SAP145 | K96 | K147 | x | | - | - | 40,6 |
| SAP145 | SAP145 | K118 | K123 | x | | - | - | 9,1 |
| SAP145 | SAP145 | K118 | K147 | x | | - | - | 31,4 |
| SAP145 | SAP145 | K123 | K147 | x | | - | - | 25,3 |
| SAP145 | SAP145 | K123 | K151 | x | | - | - | 28,2 |
| SAP145 | SAP145 | K133 | K147 | x | x | - | - | 20,5 |
| SAP145 | SAP145 | K133 | K154 | x | | - | - | 31,2 |

Annexe 42

Modèle AlphaFold de SAP145₄₀₁₋₅₅₀ avec la représentation des niveaux de confiance donnés.



Annexe 43

Tableau descriptif des peptides intra-PRMT9 cross-linkés (protéines, sites, cross-linker, domaines et distances C α -C α) identifiés suite au XL-MS du complexe PRMT9/SAP145₄₀₁₋₅₅₀.

| Protéine 1 | Protéine 2 | Site 1 | Site 2 | DSBU | DSAU | Domaine site 1 | Domaine site 2 | distance C α -C α 6PDM |
|------------|------------|--------|--------|------|------|-------------------|-------------------|--------------------------------------|
| PRMT9 | PRMT9 | N-ter | K132 | x | x | - | TRP | - |
| PRMT9 | PRMT9 | N-ter | K246 | x | x | - | RF1 | - |
| PRMT9 | PRMT9 | N-ter | S691 | x | | - | RF2 | - |
| PRMT9 | PRMT9 | K132 | S691 | x | | TRP | RF2 | - |
| PRMT9 | PRMT9 | K141 | K848 | | x | - | tonneau β_2 | - |
| PRMT9 | PRMT9 | K164 | K200 | x | | - | RF1 | 18,2 |
| PRMT9 | PRMT9 | S179 | K200 | x | | - | RF1 | 14,8 |
| PRMT9 | PRMT9 | S181 | K200 | x | | - | RF1 | 13,3 |
| PRMT9 | PRMT9 | K182 | K295 | x | x | - | RF1 | 10,3 |
| PRMT9 | PRMT9 | K287 | K393 | x | | RF1 | tonneau β_1 | - |
| PRMT9 | PRMT9 | K325 | K485 | x | x | tonneau β_1 | - | - |
| PRMT9 | PRMT9 | K397 | K485 | x | x | tonneau β_1 | - | - |
| PRMT9 | PRMT9 | K756 | K783 | x | | tonneau β_2 | tonneau β_2 | 16 |

Annexe 44

Tableau descriptif des peptides inter-PRMT9/SAP145₄₀₁₋₅₅₀ cross-linkés (protéines, sites, cross-linker et domaines) identifiés suite au XL-MS du complexe PRMT9/SAP145₄₀₁₋₅₅₀.

| Protéine 1 | Protéine 2 | Site 1 | Site 2 | DSBU | DSAU | Domaine site 1 | Domaine site 2 |
|------------|------------|--------|--------|------|------|----------------|----------------|
| PRMT9 | SAP145 | N-ter | N-ter | x | x | - | - |
| PRMT9 | SAP145 | N-ter | K7 | x | | - | - |
| PRMT9 | SAP145 | N-ter | K16 | x | | - | - |
| PRMT9 | SAP145 | N-ter | K133 | x | | - | - |
| PRMT9 | SAP145 | S17 | K16 | x | | - | - |
| PRMT9 | SAP145 | S17 | K24 | x | | - | - |
| PRMT9 | SAP145 | S17 | K70 | x | | - | - |
| PRMT9 | SAP145 | S17 | K133 | x | | - | - |
| PRMT9 | SAP145 | S29 | N-ter | x | x | TRP | - |
| PRMT9 | SAP145 | S29 | K16 | | x | TRP | - |
| PRMT9 | SAP145 | S29 | K24 | | x | TRP | - |
| PRMT9 | SAP145 | K129 | K133 | x | | TRP | - |
| PRMT9 | SAP145 | K129 | K147 | x | | TRP | - |
| PRMT9 | SAP145 | K132 | K133 | x | x | TRP | - |
| PRMT9 | SAP145 | K132 | K147 | x | x | TRP | - |
| PRMT9 | SAP145 | K132 | K151 | x | | TRP | - |
| PRMT9 | SAP145 | K141 | K147 | x | | - | - |
| PRMT9 | SAP145 | K200 | K147 | x | | RF1 | - |
| PRMT9 | SAP145 | K200 | K151 | x | | RF1 | - |
| PRMT9 | SAP145 | K238 | K118 | x | x | RF1 | - |
| PRMT9 | SAP145 | K238 | K123 | x | x | RF1 | - |
| PRMT9 | SAP145 | K243 | K118 | x | | RF1 | - |
| PRMT9 | SAP145 | K243 | K123 | x | | RF1 | - |
| PRMT9 | SAP145 | K246 | N-ter | x | | RF1 | - |
| PRMT9 | SAP145 | K246 | K118 | x | x | RF1 | - |
| PRMT9 | SAP145 | K246 | K123 | x | x | RF1 | - |
| PRMT9 | SAP145 | K246 | K133 | x | | RF1 | - |
| PRMT9 | SAP145 | K246 | K147 | x | | RF1 | - |
| PRMT9 | SAP145 | K655 | N-ter | x | | RF2 | - |
| PRMT9 | SAP145 | K655 | K16 | x | | RF2 | - |
| PRMT9 | SAP145 | K655 | K70 | x | | RF2 | - |
| PRMT9 | SAP145 | K655 | K90 | x | x | RF2 | - |
| PRMT9 | SAP145 | K655 | K96 | x | x | RF2 | - |
| PRMT9 | SAP145 | S691 | N-ter | x | | RF2 | - |
| PRMT9 | SAP145 | S691 | K70 | x | | RF2 | - |
| PRMT9 | SAP145 | S691 | K90 | x | | RF2 | - |
| PRMT9 | SAP145 | S691 | K96 | x | | RF2 | - |
| PRMT9 | SAP145 | S691 | K133 | x | | RF2 | - |
| PRMT9 | SAP145 | K778 | N-ter | | x | tonneau β2 | - |

Annexe 45

Tableau descriptif des peptides inter-R1R2, intra-R1 et intra-R2 cross-linkés (protéines, sites, domaines, distances Cα-Cα mesurées sur les structures 7OLE et 5OAF) identifiés suite au XL-MS du complexe R1R2.

| Protéine 1 | Protéine 2 | Site 1 | Site 2 | Domaine site 1 | Domaine site 2 | Distance Cα-Cα | |
|------------|------------|--------|--------|----------------|----------------|----------------|------|
| | | | | | | 7OLE | 5OAF |
| RuvBL2 | RuvBL2 | K157 | K197 | DII | DII | 12,4 | 21,5 |
| RuvBL2 | RuvBL2 | K160 | K160 | DII | DII | 123,3 | 58,7 |
| RuvBL2 | RuvBL2 | K160 | K197 | DII | DII | 11,1 | 19,4 |
| RuvBL2 | RuvBL2 | K184 | K201 | DII | DII | 26,1 | 9,5 |
| RuvBL2 | RuvBL2 | K184 | K204 | DII | DII | 18,3 | 14,8 |
| RuvBL2 | RuvBL2 | K186 | K204 | DII | DII | 21,4 | 10,2 |
| RuvBL2 | RuvBL2 | K197 | K197 | DII | DII | 104,6 | 57,3 |
| RuvBL2 | RuvBL2 | K201 | K223 | DII | DII | 35,8 | 21,7 |
| RuvBL2 | RuvBL2 | K201 | K236 | DII | DII | 18,4 | 16,7 |
| RuvBL2 | RuvBL2 | K427 | Y442 | DIII | DIII | 18 | 20,1 |
| RuvBL1 | RuvBL2 | K15 | K67 | - | DI | - | - |
| RuvBL1 | RuvBL2 | K15 | K201 | - | DII | - | - |
| RuvBL1 | RuvBL2 | K15 | K279 | - | DII | - | - |
| RuvBL1 | RuvBL2 | K20 | K67 | DI | DI | 22,2 | - |
| RuvBL1 | RuvBL2 | K20 | K197 | DI | DII | 69,2 | - |
| RuvBL1 | RuvBL2 | K24 | K67 | DI | DI | 15,2 | 12 |
| RuvBL1 | RuvBL2 | K35 | K67 | DI | DI | 26 | 20,5 |
| RuvBL1 | RuvBL2 | K46 | K365 | DI | DIII | 35,7 | 41,3 |
| RuvBL1 | RuvBL2 | K46 | K427 | DI | DIII | 17,6 | 21,1 |
| RuvBL1 | RuvBL2 | K70 | K417 | DI | DIII | 23,5 | 16,4 |
| RuvBL1 | RuvBL2 | K73 | K184 | DI | DII | 57,4 | 52,7 |
| RuvBL1 | RuvBL2 | K73 | K186 | DI | DII | 52 | 47,4 |
| RuvBL1 | RuvBL2 | K73 | K204 | DI | DII | 67 | 38 |
| RuvBL1 | RuvBL2 | K73 | K417 | DI | DIII | 29,5 | 17,8 |
| RuvBL1 | RuvBL2 | K138 | K160 | DI | DII | 46,8 | 42,2 |
| RuvBL1 | RuvBL2 | K175 | K67 | DII | DI | 44,1 | 46,3 |
| RuvBL1 | RuvBL2 | K175 | K160 | DII | DII | 65,5 | 33,4 |
| RuvBL1 | RuvBL2 | K178 | K67 | DII | DI | 46,6 | 43,6 |
| RuvBL1 | RuvBL2 | K181 | K427 | DII | DIII | 62,5 | 68,3 |
| RuvBL1 | RuvBL2 | K195 | K67 | DII | DI | 24,8 | 46,4 |
| RuvBL1 | RuvBL2 | K435 | K67 | DIII | DI | 19,6 | 24,2 |
| RuvBL1 | RuvBL2 | K435 | K160 | DIII | DII | 67,8 | 74,6 |
| RuvBL1 | RuvBL2 | K435 | K201 | DIII | DII | 58,8 | 58,6 |
| RuvBL1 | RuvBL2 | K458 | K427 | DIII | DIII | 41,3 | 40,2 |
| RuvBL1 | RuvBL2 | K469 | K427 | DIII | DIII | - | - |

| Protéine 1 | Protéine 2 | Site 1 | Site 2 | Domaine site 1 | Domaine site 2 | Distance Cα-Cα | |
|------------|------------|--------|-----------|----------------|----------------|----------------|------|
| | | | | | | 7OLE | 5OAF |
| RuvBL1 | RuvBL1 | K15 | K15 | - | - | - | - |
| RuvBL1 | RuvBL1 | K15 | K20 | - | DI | - | - |
| RuvBL1 | RuvBL1 | K15 | K35 | - | DI | - | - |
| RuvBL1 | RuvBL1 | K15 | K46 | - | DI | - | - |
| RuvBL1 | RuvBL1 | K15 | K138 | - | DI | - | - |
| RuvBL1 | RuvBL1 | K15 | S168 | - | DII | - | - |
| RuvBL1 | RuvBL1 | K15 | K175 | - | DII | - | - |
| RuvBL1 | RuvBL1 | K15 | K178 | - | DII | - | - |
| RuvBL1 | RuvBL1 | K15 | K184 | - | DII | - | - |
| RuvBL1 | RuvBL1 | K15 | K244 | - | DII | - | - |
| RuvBL1 | RuvBL1 | K15 | K281 | - | DII | - | - |
| RuvBL1 | RuvBL1 | K15 | K287 | - | DII | - | - |
| RuvBL1 | RuvBL1 | K15 | K389 | - | DIII | - | - |
| RuvBL1 | RuvBL1 | K15 | K435 | - | DIII | - | - |
| RuvBL1 | RuvBL1 | K15 | K458 | - | DIII | - | - |
| RuvBL1 | RuvBL1 | K15 | K466/K469 | - | DIII | - | - |
| RuvBL1 | RuvBL1 | K20 | K35 | DI | DI | 25,2 | - |
| RuvBL1 | RuvBL1 | K20 | K73 | DI | DI | 33,6 | - |
| RuvBL1 | RuvBL1 | K20 | K175 | DI | DII | 32,4 | - |
| RuvBL1 | RuvBL1 | K20 | K181 | DI | DII | 32,3 | - |
| RuvBL1 | RuvBL1 | K20 | K184 | DI | DII | 29,5 | - |
| RuvBL1 | RuvBL1 | K20 | K195 | DI | DII | 17,8 | - |
| RuvBL1 | RuvBL1 | K20 | K214 | DI | DII | 17,4 | - |
| RuvBL1 | RuvBL1 | K20 | K281 | DI | DII | - | - |
| RuvBL1 | RuvBL1 | K20 | K287 | DI | DII | 21,8 | - |
| RuvBL1 | RuvBL1 | K20 | K435 | DI | DIII | 39 | - |
| RuvBL1 | RuvBL1 | K20 | K454 | DI | DIII | 39,3 | - |
| RuvBL1 | RuvBL1 | K24 | K73 | DI | DI | 30,5 | 40,7 |
| RuvBL1 | RuvBL1 | K24 | K195 | DI | DII | 23,6 | 37,7 |
| RuvBL1 | RuvBL1 | K24 | K281 | DI | DII | - | 40 |
| RuvBL1 | RuvBL1 | K35 | K178 | DI | DII | 38 | 42,9 |
| RuvBL1 | RuvBL1 | K35 | K389 | DI | DIII | 16,6 | 16,2 |
| RuvBL1 | RuvBL1 | K46 | K389 | DI | DIII | 26 | 17,4 |
| RuvBL1 | RuvBL1 | K46 | K435 | DI | DIII | 43,5 | 35,6 |
| RuvBL1 | RuvBL1 | K175 | K435 | DII | DIII | 49,8 | 60,9 |
| RuvBL1 | RuvBL1 | K181 | K181 | DII | DII | 115,4 | 60,4 |
| RuvBL1 | RuvBL1 | K181 | K435 | DII | DIII | 54,3 | 64,1 |
| RuvBL1 | RuvBL1 | K181 | K458 | DII | DIII | 73,5 | 78,7 |
| RuvBL1 | RuvBL1 | K287 | K435 | DII | DIII | 54,7 | 46,5 |
| RuvBL1 | RuvBL1 | K389 | K435 | DIII | DIII | 17,8 | 17,5 |
| RuvBL1 | RuvBL1 | K413 | K458 | DIII | DIII | 15,1 | 14 |
| RuvBL1 | RuvBL1 | K458 | K466/K469 | DIII | DIII | - | - |

Annexe 46

Tableau descriptif des peptides intra-DPCD cross-linkés (protéines, sites, domaines, distances Ca-Ca) identifiés suite au XL-MS de DPCD.

| Protéine 1 | Protéine 2 | Site 1 | Site 2 | Domaine site 1 | Domaine site 2 | Distance Ca-Ca (DPCD modèle) |
|------------|------------|--------|--------|----------------|----------------|------------------------------|
| DPCD | DPCD | N-ter | K18 | - | - | 28 |
| DPCD | DPCD | N-ter | K37 | - | - | 25,2 |
| DPCD | DPCD | N-ter | K126 | - | CS putative | 20,5 |
| DPCD | DPCD | N-ter | K141 | - | CS putative | 28,4 |
| DPCD | DPCD | N-ter | K142 | - | CS putative | 25,1 |
| DPCD | DPCD | N-ter | K188 | - | - | 24,5 |
| DPCD | DPCD | N-ter | K194 | - | - | 26,6 |
| DPCD | DPCD | K18 | K116 | - | CS putative | 45,7 |
| DPCD | DPCD | K18 | K194 | - | - | 3,7 |
| DPCD | DPCD | K126 | K142 | CS putative | CS putative | 12,7 |

Annexe 47

Tableau descriptif des peptides cross-linkés (protéines, sites, domaines, distances Ca-Ca) identifiés suite au XL-MS de R1R2/DPCD.

| Protéine 1 | Protéine 2 | Site 1 | Site 2 | Domaine site 1 | Domaine site 2 | Distance Ca-Ca (R1R2/DPCD) |
|------------|------------|--------|--------|----------------|----------------|----------------------------|
| DPCD | DPCD | N-ter | K18 | - | - | 22,9 |
| DPCD | DPCD | N-ter | K37 | - | - | 18,2 |
| DPCD | DPCD | N-ter | K126 | - | CS putative | 25,1 |
| DPCD | DPCD | N-ter | S144 | - | CS putative | 24,4 |
| DPCD | DPCD | N-ter | K188 | - | - | 27,4 |
| DPCD | DPCD | K18 | K18 | - | - | 68,9 |
| DPCD | DPCD | K18 | K28 | - | - | 10,4 |
| DPCD | DPCD | K18 | K194 | - | - | 9,7 |
| DPCD | DPCD | K18 | K204 | - | - | 24,5 |
| DPCD | DPCD | K46 | K102 | - | - | 11,7 |
| DPCD | DPCD | K46 | K194 | - | - | 17,7 |
| DPCD | DPCD | K102 | K188 | CS putative | - | 15,3 |
| DPCD | DPCD | K126 | K142 | CS putative | CS putative | 11,6 |
| RuvBL1 | RuvBL1 | K15 | K24 | - | DI | - |
| RuvBL1 | RuvBL1 | K15 | K195 | - | DII | - |
| RuvBL1 | RuvBL1 | K15 | K214 | - | DII | - |
| RuvBL1 | RuvBL1 | K15 | K281 | - | DII | - |
| RuvBL1 | RuvBL1 | K15 | K287 | - | DII | - |
| RuvBL1 | RuvBL1 | K15 | K294 | - | DII | - |
| RuvBL1 | RuvBL1 | K20 | K35 | DI | DI | 18,4 |
| RuvBL1 | RuvBL1 | K20 | K195 | DI | DII | 16,6 |
| RuvBL1 | RuvBL1 | K20 | K214 | DI | DII | 13,8 |

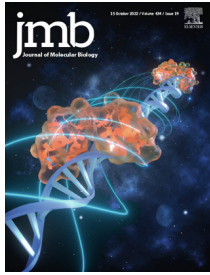
| Protéine 1 | Protéine 2 | Site 1 | Site 2 | Domaine site 1 | Domaine site 2 | Distance Ca-Ca (R1R2/DPCD) |
|------------|------------|--------|--------|----------------|----------------|----------------------------|
| RuvBL1 | RuvBL1 | K20 | K281 | DI | DII | - |
| RuvBL1 | RuvBL1 | K20 | K389 | DI | DIII | 30 |
| RuvBL1 | RuvBL1 | K24 | K35 | DI | DI | 16,6 |
| RuvBL1 | RuvBL1 | K35 | K389 | DI | DIII | 15,9 |
| RuvBL1 | RuvBL1 | K46 | K389 | DI | DIII | 18,5 |
| RuvBL1 | RuvBL1 | K120 | K287 | DI | DII | 12,2 |
| RuvBL1 | RuvBL1 | K175 | K181 | DII | DII | 6,1 |
| RuvBL1 | RuvBL1 | K195 | K214 | DII | DII | 9,4 |
| RuvBL1 | RuvBL1 | K281 | K294 | DII | DII | - |
| RuvBL1 | RuvBL1 | K287 | K294 | DII | DII | 13,2 |
| RuvBL1 | RuvBL1 | K435 | K389 | DIII | DIII | 18,8 |
| RuvBL1 | RuvBL1 | K454 | K466 | DIII | DIII | - |
| RuvBL2 | RuvBL2 | K157 | K186 | DII | DII | - |
| RuvBL2 | RuvBL2 | K184 | K201 | DII | DII | 10,3 |
| RuvBL2 | RuvBL2 | K184 | S203 | DII | DII | 16,8 |
| RuvBL2 | RuvBL2 | K186 | K201 | DII | DII | 5,7 |
| RuvBL2 | RuvBL2 | K186 | K417 | DII | DIII | 39,2 |
| RuvBL2 | RuvBL2 | K186 | K427 | DII | DIII | 50,2 |
| RuvBL2 | RuvBL2 | K201 | K236 | DII | DII | 16,6 |
| RuvBL2 | RuvBL2 | K234 | K236 | DII | DII | 6,8 |
| RuvBL2 | RuvBL2 | K427 | K427 | DIII | DIII | 79,3 |
| DPCD | RuvBL1 | N-ter | K15 | - | DI | - |
| DPCD | RuvBL1 | N-ter | K20 | - | DI | 51,5 |
| DPCD | RuvBL1 | N-ter | K181 | - | DII | 41,6 |
| DPCD | RuvBL1 | N-ter | K195 | - | DII | 56,3 |
| DPCD | RuvBL1 | N-ter | K287 | - | DII | - |
| DPCD | RuvBL1 | N-ter | K294 | - | DII | 42,7 |
| DPCD | RuvBL1 | N-ter | K435 | - | DIII | 83,6 |
| DPCD | RuvBL1 | N-ter | K458 | - | DIII | 86,9 |
| DPCD | RuvBL1 | K18 | K181 | - | DII | 27,5 |
| DPCD | RuvBL1 | K28 | K294 | - | DII | 20,2 |
| DPCD | RuvBL1 | K116 | K389 | CS putative | DIII | 68,6 |
| DPCD | RuvBL1 | K137 | K181 | CS putative | DII | 37 |
| DPCD | RuvBL1 | K204 | K287 | - | DII | 25,4 |
| DPCD | RuvBL2 | N-ter | K279 | - | DII | 41,7 |
| DPCD | RuvBL2 | N-ter | K368 | - | DIII | 81,9 |
| DPCD | RuvBL2 | N-ter | K417 | - | DIII | 76,4 |
| DPCD | RuvBL2 | N-ter | K427 | - | DIII | 85,3 |
| DPCD | RuvBL2 | K18 | K186 | - | DII | 46,2 |
| DPCD | RuvBL2 | K18 | K197 | - | DII | 32,4 |
| DPCD | RuvBL2 | K18 | K427 | - | DIII | 68,9 |
| DPCD | RuvBL2 | K37 | K279 | - | DII | 31,9 |
| DPCD | RuvBL2 | K116 | K427 | CS putative | DIII | 85,6 |
| DPCD | RuvBL2 | K126 | K234 | CS putative | DII | 49,6 |
| DPCD | RuvBL2 | K126 | K236 | CS putative | DII | 51,6 |

| Protéine 1 | Protéine 2 | Site 1 | Site 2 | Domaine site 1 | Domaine site 2 | Distance Cα-Cα (R1R2/DPCD) |
|------------|------------|-----------|--------|----------------|----------------|----------------------------|
| DPCD | RuvBL2 | K126 | K427 | CS putative | DIII | 93,9 |
| DPCD | RuvBL2 | K142 | K236 | CS putative | DII | 51,1 |
| DPCD | RuvBL2 | K204 | K427 | - | DIII | 64,2 |
| RuvBL1 | RuvBL2 | K20 | K67 | DI | DI | 24,3 |
| RuvBL1 | RuvBL2 | K24 | K67 | DI | DI | 18,3 |
| RuvBL1 | RuvBL2 | K46 | K417 | DI | DIII | 19,4 |
| RuvBL1 | RuvBL2 | K46 | K427 | DI | DIII | 20,6 |
| RuvBL1 | RuvBL2 | K70/S71 | K417 | DI | DIII | 28,4 |
| RuvBL1 | RuvBL2 | T180/K181 | K427 | DII | DIII | 65,8 |
| RuvBL1 | RuvBL2 | K281 | K269 | DII | DII | - |
| RuvBL1 | RuvBL2 | K281 | K279 | DII | DII | - |
| RuvBL1 | RuvBL2 | K281 | K365 | DII | DI | - |
| RuvBL1 | RuvBL2 | K287 | K269 | DII | DII | 16,4 |
| RuvBL1 | RuvBL2 | K287 | K427 | DII | DIII | 43,2 |
| RuvBL1 | RuvBL2 | K294 | K269 | DII | DII | 23,1 |
| RuvBL1 | RuvBL2 | K389 | K427 | DIII | DIII | 35,1 |
| RuvBL1 | RuvBL2 | K435 | K160 | DIII | DII | 64,5 |
| RuvBL1 | RuvBL2 | K458 | K427 | DIII | DIII | 36,8 |
| RuvBL1 | RuvBL2 | K466 | K427 | DIII | DIII | - |
| RuvBL1 | RuvBL2 | K466 | K444 | DIII | DIII | - |
| RuvBL1 | RuvBL2 | K469 | K427 | DIII | DIII | - |

Annexe 48

Dos Santos Morais*, R.; Santo, P. E.*; Ley, M.*; et al., **Deciphering cellular and molecular determinants of human DPCD protein in complex with RUVBL1/RUVBL2 AAA-ATPases.**

Journal of Molecular Biology **2022**, 434, 167760.



Deciphering cellular and molecular determinants of human DPCD protein in complex with RUVBL1/RUVBL2 AAA-ATPases

Raphael Dos Santos Morais^{1†*}, Paulo E. Santo^{2,3†}, Marie Ley^{4,5†},
 Cédric Schelcher¹, Yoann Abel^{6,7}, Laura Plassart⁸, Evolène Deslignière^{4,5},
 Marie-Eve Chagot¹, Marc Quinternet⁹, Ana C. F. Paiva^{2,3}, Steve Hessmann^{4,5},
 Nelly Morellet¹⁰, Pedro M. F. Sousa^{2,3}, Franck Vandermoere¹¹,
 Edouard Bertrand^{6,7}, Bruno Charpentier¹, Tiago M. Bandejas^{2,3}, Célia Plisson-
 Chastang⁸, Céline Verheggen^{6,7}, Sarah Cianférani^{4,5*} and Xavier Manival^{1*}

1 - *IMoPA, CNRS, Université de Lorraine, Nancy F-54000, France*

2 - *iBET, Instituto de Biologia Experimental e Tecnológica, Apartado 12, 2781-901 Oeiras, Portugal*

3 - *Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Av. da República, 2780-157 Oeiras, Portugal*

4 - *Laboratoire de Spectrométrie de Masse BioOrganique, Université de Strasbourg, CNRS, IPHC UMR 7178, Strasbourg 67000, France*

5 - *Infrastructure Nationale de Protéomique ProFI – FR2048 CNRS, 67087 Strasbourg, France*

6 - *IGH, CNRS, Univ Montpellier, Montpellier, France*

7 - *Equipe labélisée Ligue Nationale Contre le Cancer, 34293 Montpellier, France*

8 - *MCD, Centre de Biologie Intégrative (CBI), University of Toulouse, CNRS, UPS, 31062 Toulouse, France*

9 - *Université de Lorraine, CNRS, INSERM, IBSLor, Biophysics and Structural Biology Core Facility, F-54000, France*

10 - *Très Grandes Infrastructures de Recherche, Institut de Chimie des Substances Naturelles – CNRS, 1 avenue de la Terrasse, 91190 Gif-sur-Yvette, France*

11 - *IGF, CNRS, INSERM, Univ Montpellier, Montpellier, France*

Correspondence to Raphael Dos Santos Morais, Sarah Cianférani and Xavier Manival:

<https://doi.org/10.1016/j.jmb.2022.167760>

Edited by J. Buchner

Abstract

DPCD is a protein that may play a role in cilia formation and whose absence leads to primary ciliary dyskinesia (PCD), a rare disease caused by impairment of ciliated cells. Except for high-throughput studies that identified DPCD as a possible RUVBL1 (R1) and RUVBL2 (R2) partner, no in-depth cellular, biochemical, and structural investigation involving DPCD have been reported so far. R1 and R2 proteins are ubiquitous highly conserved AAA + family ATPases that assemble and mature a plethora of macromolecular complexes and are pivotal in numerous cellular processes, especially by guaranteeing a co-chaperoning function within R2TP or R2TP-like machineries. In the present study, we identified DPCD as a new R1R2 partner *in vivo*. We show that DPCD interacts directly with R1 and R2 *in vitro* and in cells. We characterized the physico-chemical properties of DPCD in solution and built a 3D model of DPCD. In addition, we used a variety of orthogonal biophysical techniques including small-angle X-ray scattering, structural mass spectrometry and electron microscopy to assess the molecular determinants of DPCD interaction with R1R2. Interestingly, DPCD disrupts the dodecameric state of R1R2 complex upon binding and this interaction occurs mainly via the DII domains of R1R2.

© 2022 The Authors. Published by Elsevier Ltd.

Introduction

Primary Ciliary Dyskinesia (PCD) is a genetically heterogeneous autosomal recessive disorder caused by mutations and characterized by immotile, dysmotile, or absent cilia. Cilia are highly conserved microtubule-based organelles that can be divided into two major types: motile cilia and non-motile cilia. Non-motile cilia, also called primary cilia, serve as sensory organelles, while motile cilia serve as flagella. The dysfunction of motile cilia causes ciliopathies with many phenotypes, including neonatal respiratory distress, chronic pulmonary disease, female and male infertility, organ laterality defects, hydrocephalus, and congenital heart defects.¹ Motile cilia are complex ultrastructures composed of a hundred proteins including axonemal dyneins that form the outer- and inner-arms (ODA and IDA) along doublet microtubules.² These multicomponent microtubule motors provide the force for ciliary beating resulting from ATPase activity and are built around heavy chain (HC) motor units that associate with WD-repeat intermediate chain (ICs) and various light chains (LCs).³ Since ribosomes are absent from cilia, all axonemal dynein must be synthesized in the cytoplasm of the cell body and then transported and delivered precisely to the ciliary end.⁴ Mutations in PCD genes usually cause the absence of ODA and/or IDA associated with microtubular disorganization and/or central pair organization defects.⁵ In mice, among these genes, Zariwala and collaborators identified *DPCD*, which stands for *deleted in a mouse model of primary ciliary dyskinesia*.^{6–7} Northern blots on about 20 human tissues showed an abundance of *DPCD* mRNA but only in testis⁷ while *DPCD* protein was identified in human sperm.⁸ In culture of human bronchial epithelial cells that differentiate to form a heavily ciliated epithelium, *DPCD* expression increased over time like other cilia-specific genes.⁷ Examination of a group of PCD patients with mutations in *DPCD* showed defects in the IDA and ODA but none of the sequence variants was confirmed as the direct cause of the disease.

DPCD is poorly described at the protein level. It comprises 203 amino acids (23 kDa), has no functional homologs, and its function and 3D structure are currently unknown. *DPCD* was found to associate with the RUVBL1 (R1, also called Pontin) and RUVBL2 (R2, also called Reptin) in proteomic studies in humans^{9–12} and in mouse cells,¹³ which is consistent with R1 being found essential in the pre-assembly of the ciliary protein complex.¹³ In mouse models, *DPCD* was found to be associated with DNAAF1 (Dynein axonemal assembly factor 1) in a co-immunoprecipitation assay.¹³ This strengthens the potential role of *DPCD* in dynein arm assembly and more generally in cilia formation.

Given the complexity of axonemal dynein motors, many factors are required for their stable assembly in cytoplasm and their transport to the axonemal super-structure.¹⁴ Among these factors, R2TP, a multiprotein complex acting in concert with the HSP70 and HSP90 chaperones, enables the stability, maturation, and quaternary assembly of functional macromolecular complexes. R2TP was originally identified in yeast¹⁵ and is involved in many different cellular pathways. R2TP includes R1 (Rvb1 in yeast) and R2 (Rvb2 in yeast), two highly conserved ATPases Associated with diverse cellular Activities (AAA+) that form a functional alternating heterohexamer, and RNA polymerase II-associated protein 3 (RPAP3, Tah1p in yeast) and PIH1 domain-containing protein 1 (PIH1D1, Pih1p in yeast) proteins constituting a heterodimer. RPAP3 is the adapter for HSP70/90, and PIH1D1 mediates client protein recruitment. The sequence of R1R2 encompasses three domains: the $\alpha\beta\alpha$ domain I (DI) and the α -helical domain III (DIII) that form the alternating AAA ring (ATPase or AAA-face), while a central and protruding domain II (DII-face, at the opposite of the AAA-face) represents a specific insertion into the core AAA+ domains. The DII domain is subdivided into DII_{int} including an α -helical bundle and a long flexible antiparallel β stalk, and DII_{ext} structurally equivalent to an oligonucleotide/oligosaccharide-binding (OB)-fold, which is an ssDNA binding domain. The DII-face represents a scaffolding platform within the R1R2 hexamer to recruit the majority of their clients/cofactors.

In vivo, ATP binding or hydrolysis by R1 or R2 is essential for all their postponed activities. For example, R1R2 ATPase activity is required for the maturation and/or disassembly of the PAQosome complex.¹⁶ To date, all analysis of the 3D structures of free R1R2 revealed that nucleotide, most frequently ADP, is present in the binding pocket of each AAA-subunit even if it was not added during purification.^{17–19} The binding pocket requires for nucleotide hydrolysis the arginine finger motif provided by the adjacent subunit in the hexamer ring. The hexamerization of R1-ADP blocks the nucleotide binding pocket, thus rendering an exchange from ADP to ATP impossible.¹⁹ This point partially explains the weak ATPase activity of recombinant purified proteins,¹⁷ unless DII domains are truncated, most likely mimicking cofactor regulation. Recent crystal structures suggest communication between the DII and the outer faces of ATPase to regulate this exchange. The ordered N-terminus of the R2 subunit, comprising the His-(Ser)-His motif, interacts with both the nucleotide and DII, thereby preventing the nucleotide to be released from its pocket.²⁰ The cryo-EM structure of the human R2TP revealed its structural organization as well as the pivotal role of PIH1D1.^{21–22} These authors proposed that the binding of one PIH1D1 molecule

to the DII of one R2 subunit disorders the N-terminal segment of R2 and releases the ADP nucleotide likely to promote ATP turnover. Similarly, the binding of DHX34 to R1R2 displaces the N-terminal region in each R2 subunit and leads to the loss of nucleotides, allowing DHX34 to down-regulate R2 ATPase activity.²³ Very recently, ZNHIT2, a specific adaptor of the U5 snRNP biogenesis, was shown to induce changes in N-terminal regions of RUVBL2 and nucleotides were lost in all R2 protomers.²⁴ However, this type of mechanism remains hypothetical, as other authors were unable to observe significant displacement of the nucleotide from its active site when R2TP formed.²⁵

Recently, it was suggested that other multiprotein complexes, homologous to R2TP, could form in cells in a tissue-specific manner to interact with particular adaptors and protein clients^{11,26}: (i) R2SP where RPAP3 and PIH1D1 are replaced by SPAG1 (Sperm-associated antigen 1) and PIH1D2 (PIH1 domain-containing protein 2) respectively forming the SP heterodimer, (ii) R2SD where RPAP3 and PIH1D1 are replaced by SPAG1 and DNAAF2 (Dynein axonemal assembly factor 2, also Kintoun), (iii) D2 with DYX1C1 (Dyslexia susceptibility 1 candidate gene 1 protein, also DNAAF4)-a (a isoform) and DNAAF2, and (iv) DP with DYX1C1-c (c isoform) and PIH1D3. In mammals, R2TP associates with a Prefoldin-like (PFDL) module to form the PAQosome (Particle for Arrangement of Quaternary structure), a 12-subunit chaperone complex.^{27,28} The PFDL module forms a hexameric structure that contains two α subunits (UXT and URI1) and four β subunits (PFDN2, PFDN6, PDRG1, and ASDURF for ASNSD1 upstream open reading frame protein). The PAQosome also includes two additional components, the RNA Pol I/II/III shared subunit POLR2E, and the Tryptophan/Aspartate Repeat 92 (WDR92, also known as Monad).²⁷ WDR92 has been reported to associate with DPCD⁹. It is only expressed in organisms that encode axonemal dynein.²⁹ FlyAtlas 2 adult expression data show that *Drosophila Wdr92* is highly and specifically expressed in testis in developing spermatocytes and is confined to the cytoplasm.³⁰ Recent biochemical data on *Chlamydomonas* revealed that Wdr92 is required for the cytoplasmic stability of ODA HCs.²⁹ The R2TP subunit RPAP3/PIH1D1 associates with Wdr92 which, in turn, is associated with dynein chains and DNAAFs.³¹ Moreover, the association of Wdr92 protein with Spag1 is facilitated by the presence of R1 and R2.³¹ The loss of Wdr92 leads to the malfunction of dynein arm assembly into sperm flagella and sensory neuronal cilia. To summarize, the R2TP/R2SP/R2SD/PAQosome complexes are thought to regulate the correct architectural pre-assembly of ciliary dynein motor, and malfunction of any of these key steps leads to PCD. In *Drosophila*, Wdr92 is hypothesized to be a new dedicated DNAAF.

In the present study, we identified *in vivo* DPCD partners, unveiled the first cellular, biochemical, and structural informations concerning human DPCD using fluorescence microscopy, SILAC- and LUMIER-IPs, and a series of structural biology techniques among which SAXS, NMR and structural mass spectrometry (structural MS). We first built a 3D model of the free DPCD structure based on experimental structural data. We next identified a direct sub-micromolar affinity interaction between up to 3 DPCD and one hexameric R1R2 (R1R2D complex) thanks to native MS and ITC. Additional structural data on the R1R2D complex obtained by SEC-SAXS, negative staining electron microscopy and cross-linking MS (XL-MS) confirmed that the interaction of DPCD occurs on R1R2 hexamers and dissociates the R1R2 dodecameric assembly. Finally, we discuss the putative role and biological relevance of DPCD in R2TP, R2TP-like or R1R2-related complexes.

Material and methods

Cell culture and cell line construction

HeLa Flp-In cells were a gift from S. Emiliani (Institut Cochin, Paris). HEK-293 T cells were obtained from the ATCC collection. All cells were grown at 37 °C, 5% CO₂ and cDNA transfections were done with JetPrime (Ozyme). HeLa Flp-In and HEK-293 T cells were grown in Dulbecco's Modified Eagle Medium (DMEM) containing 10% fetal bovine serum (FBS), glutamine (2.9 mg/mL), and penicillin/streptomycin (10 U/mL). For SILAC, after plasmid recombination at the FRT (Flp Recombination Target) site of HeLa Flp-In by co-transfection of FlpO expression plasmid and pcDNA₅-GFP-DPCD, clones were selected in Hygromycin B (150 μ g/mL), picked individually, and characterized by Western blot and fluorescence microscopy.

SILAC-IP and proteomic analysis

For the SILAC experiment, a HeLa Flp-In clone expressing GFP-DPCD was grown for 15 days in isotopically semi-heavy label media containing L-Lysine-²HCl (²H₄, 96–98%) and L-Arginine-HCl (¹³C₆, 99%) (percentages represent the isotopic purity of the labeled amino acids, CIL/Eurisotop). The parental cell line that does not express GFP-DPCD was grown in light non-labeled media (R0K0, L). Eight 15-cm diameter plates were used per condition. Cells were rinsed with PBS, trypsinized and cryoground in liquid nitrogen, and the resulting powder was resuspended in HNT lysis buffer (20 mM HEPES, pH 7.4, 150 mM NaCl, 0.5% triton X-100, protease inhibitor cocktail (cComplete, Roche). Extracts were incubated for 20 min at 4 °C and clarified by centrifugation for

10 min at $20,000 \times g$. Before immunoprecipitation, extracts were pre-cleared by incubation with Protein G Sepharose beads (GE healthcare) for 1 h at 4 °C. Each extract was then incubated with 50 μ L of GFP-Trap beads (Chromotek) for 90 min at 4 °C, washed 5 times with HNT buffer, and the beads from the different isotopic conditions were pooled. Bound proteins were eluted by adding 1% SDS to the beads and boiling for 10 min. Proteomic analysis was performed as previously described.¹¹

Lumier-IP

HEK-293 T cells were grown in 24-well plates and co-transfected with 450 ng of the RL fusion and 50 ng of the 3x FLAG-FL fusion using JetPrime transfection reagent (Ozyme). After 48 h, cells were extracted in 500 μ L of ice-cold HNTG buffer (20 mM HEPES, pH 7.9, 150 mM NaCl, 1% Triton X-100, 10% glycerol, 1 mM MgCl₂, 1 mM EGTA) containing protease inhibitor cocktail (Roche), incubated for 15 min at 4 °C and spun down for 15 min at 4 °C at $20,000 \times g$. Aliquots (100 μ L) of the extract were placed in two wells in a 96-well plate, one well-being coated with anti-FLAG antibody (10 μ g/mL in 1X PBS, F1804 Sigma-Aldrich), and one control well with no antibodies. The plates were incubated for 3 h at 4 °C, and then washed 5 times with 300 μ L of ice-cold HNTG for 10 min at 4 °C for each wash. After the last wash, 10 μ L of 1X PLB (Promega) was added in each well. To measure the input, 2 μ L of extract and 8 μ L of 1X PLB were mixed in new wells. The plates were then incubated for 5 min at room temperature, and FL and RL luciferase activities were measured in IP and input wells using the dual luciferase kit (Promega). Co-IP efficiency was defined as the RL/FL ratio in the pellet, divided by the RL/FL ratio in the input. Unless otherwise stated, statistical significance was evaluated using a Z-test to test if co-IP efficiency in the anti-FLAG IP was more than 11 times higher than the mean values obtained in the control IP, i.e. with no antibodies.¹¹ In assays where the CB6644 drug was used,³² the drug was added at a final concentration of 0.5 mM in the culture medium 4 h after transfection and left to stand for 48 h. The control, with no drug, was with DiMethyl SulfOxide (DMSO).

Protein expression and purification

Recombinant His-DPCD protein (Uniprot: Q9BVM2, 203 residues, optimized synthetic gene encoding the L156S natural variant, Life technologies) was overexpressed in the *E. coli* BL21(DE3) pRARE2 strain using a pNEA-3CH vector.³³

Bacteria were transformed with the corresponding plasmid and a clone was selected on LB agar medium with ampicillin. Bacteria were grown in LB medium at 37 °C and when the OD of 600 nm reached ~ 0.7 – 1 , protein production was induced with 0.2 mM IsoPropyl β -D-1-ThioGalactopyranoside (IPTG). Bacterial cultures were incubated overnight at 20 °C in a New Brunswick (Innova[®]) 44R Shaker at 180 rpm, and then harvested by centrifugation at $4,500 \times g$. Bacterial cells were broken down by sonication in lysis buffer (25 mM HEPES pH 7.5, 300 mM NaCl, 5% glycerol, 1 mM TCEP, 10 mM imidazole). The suspension was supplemented with 5 U/mL of DENARASE[®] (c-LEcta) and centrifuged to pellet cellular debris (30 min, 4 °C, $45,000 \times g$). The supernatant was filtered at 0.22 μ m and injected onto a HiTrap TALON (Cytiva). The 5 mL prepacked column was washed extensively with lysis buffer (without TCEP), and then proteins were eluted in the same buffer but containing 300 mM imidazole. The His-tag was removed by Human RhinoVirus 3C (HRV-3C, Cytiva) cleavage overnight, unless otherwise stated. DPCD was finally polished in a final size exclusion chromatography step using Superdex 75 or 200 columns (16/60, Cytiva) equilibrated in HNGT buffer (20 mM HEPES, 150 mM NaCl, 1% glycerol, 1 mM TCEP). Protein purity was assessed by SDS-PAGE stained with Coomassie blue and the concentration was determined spectrophotometrically using a molar extinction coefficient of 32 430 M⁻¹.cm⁻¹. For the SEC-SLS of DPCD alone and circular dichroism (CD) experiments, the 6xHis tag was exceptionally preserved.

Recombinant R1R2 proteins (Uniprot: Q9Y265 and Q9Y230) were purified as described in¹¹ with the following modifications. R2 has a C-terminal FLAG_FH8 Tag preceded by HRV-3C protease cleavage site. Peak fractions collected from the HisTrap were incubated with 5 mM CaCl₂ for 1 h and loaded onto an HiPrep Octyl FF 16/10 column (Cytiva) equilibrated in Buffer B (20 mM HEPES pH 8.0, 200 mM NaCl, 5 mM CaCl₂, 0.5 mM TCEP). Bound proteins were eluted using Buffer C (Buffer B without NaCl, CaCl₂, and supplemented with 5 mM EDTA). The FLAG-FH8 tag in the collected samples was removed by incubation for 18 h at 4 °C with 1% (w/w) HRV-3C protease (Thermo Fisher Scientific). A Superose 6 column equilibrated with Buffer D (20 mM HEPES pH 8.0, 150 mM NaCl, 0.5 mM TCEP) was used to allow the isolation of the main oligomeric species. The peak fractions were pooled and concentrated to 17.3 mg/mL using a 30 kDa Cut-off Amicon Ultra centrifugal filter (Millipore). R1 Δ DIIR2 Δ DII and R1_D302NR2_D299N purifications were performed as described above except that all buffers were supplemented with ATP in the latter purifica-

tion. All purification steps were carried out at room temperature and were monitored by NuPage Bis-Tris gels (Invitrogen, NP0302).

Far-UV circular dichroism analysis

The Far-UV circular dichroism spectrum of His-DPCD was recorded on a Chirascan spectropolarimeter (Applied Photophysics) from 190 to 260 nm in 0.01 cm pathlength Hellma cells. Blanks corresponding to the buffer were subtracted from the respective spectra. His-DPCD was analyzed at 50 μM and units are expressed as Molar Residue Ellipticity (MRE, in degrees. cm^2 . dmol^{-1} . residue^{-1}). The measurement temperature was 20 °C. The secondary structure content was estimated using the BeStSel server (<http://bestsel.elte.hu/index.php>).^{34–35}

Size exclusion chromatography coupled with static light scattering (SEC-SLS)

SEC-SLS experiments were performed at 30 °C with an AKTA Explorer HPLC system (Cytiva) coupled to light scattering and refractive index detectors (Viscotek TDA302, Malvern Panalytical). The SEC columns (Shodex KW803 or KW804, 7 μm , 8 mm ID \times 300 mm, void volume \sim 6 mL, total volume \sim 12.5 mL or Agilent BioSEC5-500 Å, 4.6 mm ID \times 300 mm, void volume \sim 2 mL, total volume \sim 5 mL) were equilibrated either with PBS (pH 7.4, Gibco, Thermo Fisher Scientific) or HNGT buffer (5 mM TCEP). The flow rate ranged from 0.2 to 0.5 mL/min. Data were processed with OmniSEC software (v5.12, Malvern Panalytical) using RALS/LALS detectors. The refractometer was used to determine the protein concentration and the refractive index increment value (dn/dc) used to determine the molecular weight (MW) was 0.185 mL/g. Hydrodynamic radii were estimated with SEC protein standards (BEH200, Waters).

Nuclear magnetic resonance (NMR)

$^{13}\text{C}/^{15}\text{N}$ and $^2\text{H}/^{13}\text{C}/^{15}\text{N}$ samples of DPCD were prepared as described above, except those bacteria were grown in aqueous or deuterated minimal M9 media supplemented with $^{15}\text{NH}_4\text{Cl}$ and ^{13}C -d6-Glucose as sole nutrient source. The recombinant proteins were purified as described above, the final buffer being 10 mM NaPi, pH 6.5, 300 mM NaCl, 5 mM DTT deuterated. ^1H - ^{15}N TROSY-HSQC, ^1H - ^{15}N NOESY-HSQC and triple resonance HNCA/trHNCA, HNCO, HNCACB/trHNCACB and CBCACONH/trCBCACONH spectra were recorded at 298 K on 600 MHz and 950 MHz spectrometers equipped with cryoprobes. Data were processed with TopSpin-3.0 (Bruker) and analyzed with CARRA.³⁶ Assigned HN, N, CO, C α and C β chemical shifts were derived into secondary structures using TALOS-N.³⁷

Isothermal titration calorimetry (ITC)

Protein/protein interactions were characterized at 30 °C using a microcalorimeter (MicroCal iTC200, Malvern Panalytical) in HNGT buffer. DPCD was placed into the syringe at a concentration of 203 μM and 6 μM of R1/R2 hexamers were injected into the cell. Nineteen successive 2 μL injections were performed, except for the first injection which was only 0.5 μL to prime the syringe. The rotation of the blade was set to 300 rpm. Data were processed with NITPIC³⁸ (v1.2.7) and NanoAnalyze (v3.10.0, TA instruments) software to determine the dissociation constant (K_d), the stoichiometry (n), and the thermodynamic parameters: enthalpy (ΔH), entropy (ΔS) and free energy (ΔG).

Chromatographic co-elution assay

For the formation of R1R2D complex, 1.93 nmol DPCD (23.2 kDa) were combined with different ratios of R1R2 and incubated for 18 h at 4 °C. Buffer D was added to reach a final volume of 200 μL immediately before injection in Superose 6 10/30 GL Increase (Cytiva). A 100- μL loop was used with a flow rate of 0.75 mL/min using an AKTA purifier system (Cytiva). The eluting protein was detected by UV absorbance at 280 nm. This elution profile was compared to injections of 1.93 nmol R1/R2 and DPCD alone. All chromatographic steps were carried out at room temperature and were monitored by NuPage Bis-Tris gels (Invitrogen, NP0302). Protein molecular masses were estimated by comparison with six molecular mass standards (Sigma, St. Louis, MI, USA): Thyroglobulin (669 kDa), Apoferritin (443 kDa), β -Amylase (200 kDa), Alcohol Dehydrogenase (150 kDa), Bovine Serum Albumin (66 kDa), and Cytochrome C (12.4 kDa).

Native mass spectrometry (MS) and ion mobility (IM-MS) measurements

For native MS analysis, samples were buffer exchanged against ammonium acetate (150 mM, pH 7.5, Sigma, St. Louis, MO, USA), using Zeba Spin desalting columns (Thermo Fisher Scientific, Rockford, IL, USA). Sample concentrations were determined by UV absorbance using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, France). Native MS analyses were carried out on an Orbitrap Exactive Plus EMR (Thermo Fisher Scientific, Bremen, Germany) coupled to an automated chip-based nanoelectrospray source (TriVersa NanoMate, Advion Biosciences, Ithaca, U.S.A.) operating in positive ion mode. The in-source collision-induced dissociation (CID) was set to 50 eV. The higher-energy collisional dissociation (HCD) cell was set to 200 eV for each analysis to improve the desolvation. The trapping gas pressure was set to

7 a.u. (corresponding to an ultra-high vacuum of 1.10^{-9} mbar). To improve the transmission of the high MW species, the voltages on the injection-, inter-, and bent-flatapoles were set to 2 V. Acquisitions were performed in the m/z range 1,000–20,000 with a 3 s scan time and a resolution of 8,750 at 200 m/z with an automatic gain control (AGC) target set to 1.10^6 and a maximum injection time set to 100 ms. External calibration was performed using singly charged ions produced by a 2 g/L solution of cesium iodide in 2-propanol/water (50/50 v/v). Orbitrap MS data were interpreted using BioPharmaFinder v3.2 (Thermo Fisher Scientific, Bremen, Germany).

Native IM-MS experiments were performed on a hybrid quadrupole/travelling wave ion mobility spectrometry (TWIMS)/time-of-flight mass spectrometer (Synapt G2 HDMS, Waters, Wilmslow, UK). The Synapt was operated in positive ionization mode. Parameters – sample cone 5 V, trap gas 5 mL/min (argon) and backing pressure 6 mbar – were optimized to preserve noncovalent interactions and to minimize ion activation. For DPCD, ions were focused in the helium cell (150 mL/min), prior to separation in the IM cell with a nitrogen flow rate of 25 mL/min. The wave velocity (WV) and height (WH) were set to 1,250 m/s and 40 V, respectively. For R1R2 and R1R2D complexes, gas flow rates of 120 mL/min (He) and 60 mL/min (N_2) were used, and WH and WV parameters were fixed to 550 m/s and 40 V, respectively. IM-MS experiments were recorded in triplicate under identical instrumental conditions. Arrival time distributions (ATD) were extracted with MassLynx v4.1 (Waters, Manchester, UK). ATDs were converted into collision cross sections ($^{TW}CCS_{N_2}$) using alcohol dehydrogenase, pyruvate kinase and glutamate dehydrogenase as external calibrants, as reported elsewhere.^{39–40} Experimental $^{TW}CCS_{N_2}$ values were then compared with theoretical CCS using PDB files with the IMoS v1.10 software.⁴¹ Nitrogen was considered as the drift gas, with a corresponding Van der Waals radius of 1.5 Å and gas polarizability of 1.7 \AA^3 . Calculations were performed using the projection approximation (PA) and the exact hard sphere scattering (EHSS) methods. Averaged theoretical CCS values were obtained from six independent calculations. For straightforward comparison of IM-MS results with solution-based techniques, IM radii (R_{IM-MS}) were derived from their corresponding $^{TW}CCS_{N_2}$ assuming a spherical shape of the protein, with $R_{IM-MS} = (^{TW}CCS_{N_2}/\pi)^{1/2}$.⁴²

Differential scanning fluorimetry (DSF)

DSF was performed in MicroAmp EnduraPlate Optical 96-Well Clear Reaction Plates with Barcode (Applied Biosystems, Life Technologies, California, USA) using a QuantStudio 7 Flex Real-Time PCR System (Applied Biosystems). Protein samples were briefly centrifuged before

preparation for the assay. The final reaction mixture (total volume 20 μL) contained either DPCD, R1R2, or R1R2D, and 4-fold Protein Thermal Shift Dye (Applied Biosystems) diluted in Buffer D. Melting curve data were recorded from 25 °C to 90 °C at an increment rate of 0.016 C. s^{-1} . Excitation and emissions filters were applied for Protein Thermal Shift Dye (470 nm and 520 nm, respectively) and for ROX reference dye (580 nm and 623 nm, respectively). The melting temperatures were obtained by calculating the midpoint of each transition, using Protein Thermal Shift Software™ v1.3. All the samples were tested in duplicate.

Dynamic light scattering (DLS) Analysis

A SpectroLight 610 (XtalConcepts GmbH, Hamburg, Germany) was used to perform serial DLS measurements. All the samples were centrifuged (15–30 min, 4 °C, $17,200 \times g$) in a benchtop centrifuge before measurement and were pipetted (each sample in duplicate, 1 μL per well) onto a 96 well Vapor Batch Plate (Jena Bioscience). Prior to use, the plates were filled with paraffin oil (Cat N. 18512, Merck) to protect the sample solutions from drying out. The laser wavelength was 660 nm at a power of 100 mW. The scattering angle for placement of the detector was set to 142°. All tested samples were kept in HEPES buffer so the refractive index (1.34) and viscosity (1.006 cP) of water were used for calculations. All the samples were measured at a constant temperature of 20 °C, one scan per drop with 20 measurements of 20 seconds each. R1/R2 and R1R2D were tested at a concentration of 1 mg/mL.

Small-angle X-ray scattering (SAXS) experiments and de novo protein prediction

SAXS data were obtained using SEC-SAXS on the SWING beam line at the French SOLEIL synchrotron (Saint-Aubin, France). Data were collected using an Eiger 4 M detector at a sample-detector distance of 2 m and at a wavelength of $\lambda = 1.033 \text{ \AA}$ ($I(Q)$ vs Q , where $Q = 4\pi\sin\theta/\lambda$, and 2θ is the scattering angle), covering a Q -range of $0.005\text{--}0.5 \text{ \AA}^{-1}$. SEC-SAXS was performed at 15 °C with a Bio SEC5-500 Å HPLC column (4.6 mm ID \times 300 mm, void volume \sim 2 mL, total volume \sim 5 mL, Agilent) equilibrated in HNGT buffer (5 mM TCEP) at 0.3 mL/min. Fifty microliters of samples at 5–12 mg/mL were injected into the column. The data were collected through the SEC peak of the protein as a series of 1-s exposures. Data reduction to absolute unit, frame averaging, and subtraction were performed with FOXTROT. Curves displaying a constant R_g in a R_g versus frame number plot, were averaged and used for further characterization. Data were analyzed with the ATSAS suite⁴³ following the man-

ufacturer's guidelines unless indicated otherwise. The MW was determined using the volume of correlation method.⁴⁴ Ten *ab initio* models were generated with GASBOR software, the models were aligned and averaged with DAMAVER software. The models shown are those with the lowest χ^2 value. For DPCD, Tr-Rosetta⁴⁵ was used to predict *de novo* models. RoseTTAFold⁴⁶ and AlphaFold2⁴⁷ were also employed. The generated models were compared to SAXS data with CRY SOL. The model with the lowest χ^2 value was selected for further processing and was subjected to normal mode analysis using SREFLEX⁴⁸ through the ATSAS online interface (www.embl-hamburg.de/biosaxs/atsas-online). Data have been deposited in the SASBDB⁴⁹: DPCD (SASDF5), and R1R2D complex (SASDFH5).

Negative staining electron microscopy

R1R2 or DPCD-containing complexes were formed and purified as described above. Aliquots (3.5 μ L) of the complexes (concentration: 20 μ g/mL) were deposited on glow-discharged, 300 Mesh Copper carbon-coated grids (Quantifoil, Germany). After 1 min of incubation at room temperature, excess liquid was blotted, and the grids were stained with two successive drops of 1% uranyl acetate (3.5 μ L was used for each blot). The first drop was immediately blotted, the second was left for \sim 20 s on the grids before blotting. Finally, the grids were air-dried before observation on the electron microscopes of the METI platform (*Centre de Biologie Intégrative*, Toulouse, France). Initial images were recorded on a Jeol JEM-1400-HC operating at 80 kV and equipped with a Gatan Orius SC1000B CCD camera and DigitalMicrograph (Gatan, Inc.), at a magnification corresponding to a calibrated pixel size of 2.39 Å. For 2D classification, images were recorded on a JEOL JEM-2100 electron microscope (LaB6 cathode, operating voltage: 200 kV) with a Gatan US1000 CCD camera and DigitalMicrograph (Gatan, Inc.). Defoci ranged between -0.5 and -1.6 μ m, and the calibrated pixel size corresponded to 2.35 Å. Single particle analysis (CTF estimation, particle picking and extraction, 2D classification) was performed using Relion 3.1.⁵⁰ For R1R2 complexes, 250 micrographs were recorded corresponding to 42,860 extracted particles that were further classified into 150 2D classes. For R1R2D complexes, 300 micrographs were recorded that yielded 59,629 extracted particles, subsequently classified into 250 2D classes. For R1R2 Δ DII, 300 micrographs were acquired from which 33,028 particles were extracted and classified into 150 2D classes.

Cross-linking MS (XL-MS)

Cross-linking reactions were conducted with 1.75 mg/mL protein solution. Freshly prepared

10 mM stock solution of disuccinimidyl dibutyric urea (DSBU, which reacts specifically with Lys residues) was added 100-fold molar excess. The reaction lasted 45 min at room temperature. The reaction was quenched for 20 min using NH_4HCO_3 to reach a final concentration of 20 mM. Disulfide was reduced by incubating the cross-linked complex solution with 5 mM DTT for 30 min at 37 °C. Alkylation was performed with 15 mM iodoacetamide (IAA) for 60 min in the dark. Trypsin/Lys-C mix (Promega, Madison, USA) was then added to reach a 1:50 enzyme: substrate ratio. Samples were incubated overnight at 37 °C. Digestion was quenched with 1% TFA. Peptides were cleaned up using SPE cartridges and samples were concentrated in a SpeedVac concentrator before LC/MS/MS analysis. NanoLC-MS/MS analysis was performed using a nanoAcquity Ultra-Performance-LC (Waters, Milford, USA) coupled to the Q-Exactive Plus Orbitrap mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) with a nanoSpray source. Samples were trapped on a nanoACQUITY UPLC precolumn (C18, 180 μ m \times 20 mm, 5 μ m particle size), and the peptides were separated on a nanoACQUITY UPLC column (C18, 75 μ m \times 250 mm with 1.7 μ m particle size, Waters, Milford, USA) maintained at 60 °C. Mobile phase A was 0.1% (v/v) formic acid in water, and mobile phase B was 0.1% (v/v) formic acid in acetonitrile. The sample were first injected with a 105 min gradient (1–3% B for 2 min, 3–35% B for 77 min, 35–90% B for 1 min, 90% B for 5 min, 90–1% B for 2 min and maintained 1% B for 2 min) at a flow rate of 400 nL/min. The Q Exactive Plus Orbitrap source temperature was set to 250 °C and the spray voltage to 1.8 kV. Full scan MS spectra (300–1,800 m/z) were acquired in positive mode at a resolution of 140,000, a maximum injection time of 50 ms and an AGC target value of 3.106 charges, with the lock-mass option enabled (polysiloxane ion from ambient air at 445.12 m/z). The 10 most intense multiple charged peptides per full scan (charge states > 2) were isolated using a 2 m/z window and fragmented using HCD (30 normalized collision energy, $\pm 3\%$). MS/MS spectra were acquired with a resolution of 35,000, a maximum injection time of 100 ms, an AGC target value of 1.105, and dynamic exclusion set to 60 s. The system was fully controlled by XCalibur software v3.0.63, 2013 (Thermo Scientific) and a NanoACQUITY UPLC console v1.51.3347 (Waters). Raw data collected were processed and converted into.mgf format. The MS/MS data were analyzed using MeroX software v2.0.1.4 (Götze et al., 2015). Mass tolerances of 5 ppm for precursor ions and of 10 ppm for product ions were applied. A 5% FDR cut-off and a signal-to-noise ≥ 2 were applied. Lys and Arg residues were considered as protease sites with a

maximum of three missed cleavages. Carbamidomethylation of cysteine was set as static modification and oxidation of methionine as variable modification (max. mod. 2). Primary amino groups (Lys side chains and N-termini) as well as primary hydroxyl groups (Ser, Thr and Tyr side chains) were considered as cross-linking sites. The cRap database was used in combination with the reporter ion scan event (RISEUP) mode with a maximum missing ion of 1. Cross-links composed of consecutive amino acid sequences were ignored. Each cross-linked product automatically annotated with MeroX was manually validated. Three different experiments were performed, and cross-linked peptides were validated when seen in 2 out of 3 experiments. The XL-MS data set has been deposited to the ProteomeXchange Consortium via the PRIDE⁵¹ partner repository with the dataset identifier PXD031339.

Results

In vivo characterization of several complexes with DPCD

DPCD was first identified in association with R2TP or R2TP/PFDL in several high throughput studies.^{9–10,13} To identify new DPCD key partners in a cellular environment, we first generated a 3xFLAG-GFP-DPCD fusion protein for cellular localization and quantitative proteomics.

Using fluorescence microscopy in HeLa cells, DPCD was found to be distributed throughout the cell in both the cytosol and the nucleoplasm (Figure 1(A)). Next, we created a HeLa Flip-In clone that stably expressed the fusion protein and performed quantitative proteomics using SILAC followed by immunoprecipitation (SILAC-IP). Proteomic analysis of anti-GFP IP obtained from this clone revealed two levels of protein enrichment (Figure 1(B)). At the first level, we found R1 and R2, which showed major enrichment in DPCD (SILAC ratio) and high intensity MS detection (comparable to the bait, Table S1),⁶² suggesting a strong association between DPCD and R1R2. Interestingly, this is the first evidence to the existence of a unique R1R2D complex *in vivo*. The second level corresponds to proteins with milder but still significant (Bsignificance < 0.05) enrichment in DPCD and milder MS intensity (Table S1), suggesting fewer expressed interacting proteins or more labile protein–protein interactions. This second class includes R2TP core proteins such as RPAP3 and PIH1D1, R2TP clients like the U5 snRNP protein EFTUD2 and the U4, U6, U5 tri-snRNP protein AAR2, co-factors like ZNHIT2 and C2orf44/WDCP/MMAP, and core proteins of PAQosome like WDR92/Monad and POLR2E.^{9,11,27} In this group, we also found HSP90/70 chaperones and their co-chaperones such as STIP1. We therefore conclude

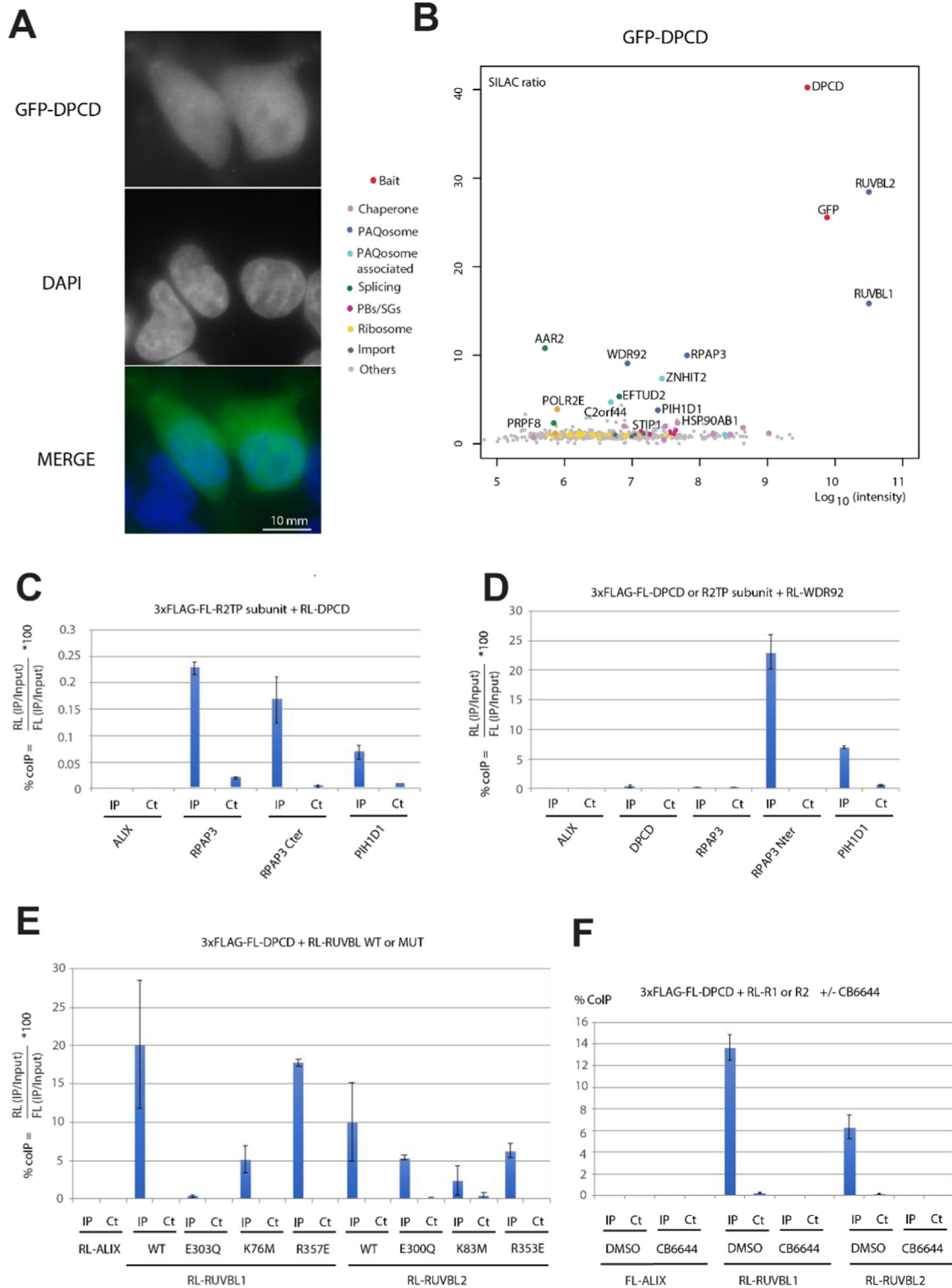
that DPCD associates with R2TP-HSP70/90-PAQosome complexes.

To confirm the interactions between DPCD and R2TP and some co-factors, we performed LUMIER-IP experiments, i.e. a quantitative immunoprecipitation assay that uses pairs of over-expressed proteins.⁵² In these experiments, DPCD had a low percentage of CoIP with RPAP3, RPAP3-Cter and PIH1D1 (Figure 1(C)) and did not interact with WDR92 (Figure 1(D)). Conversely, RPAP3 N-terminal domain and PIH1D1 show in our assay with WDR92, high and intermediate signals, respectively, suggesting a strong interaction (Figure 1(D)). Surprisingly, full-size RPAP3 did not interact at all with WDR92. As signals for DPCD interactions with R1 WT and R2 WT were respectively, high and intermediate (Figure 1(E)), indirect association could explain the presence of RPAP3, PIH1D1 and WDR92 in the SILAC-IP of GFP-DPCD (Figure 1(B)).

We also investigated the interactions between DPCD and point functional mutants of R1 and R2 (Figure 1(E)). Concerning R1 mutants, DPCD still interacted with R357E (deletion of the trans-arginine finger) whereas *in vivo* binding with E303Q was lost (no ATP hydrolysis) and was reduced with K76M (no ATP binding). Concerning R2 mutants, the interaction was also reduced with E300Q (no ATP hydrolysis), K83M (no ATP binding), and R353E (deletion of the trans-arginine finger) compared to R2 WT. Moreover, we tested the effect of a selective (non-ATP-competitive) R1R2 ATPase inhibitor, the CB6644 drug, known to stabilize the ATP-bound state, on the interaction between DPCD and each individual RUVBL, in transfected cells.³² This treatment eliminated the interaction between DPCD and R1 and R2 (Figure 1(F)). These data suggest that the interaction between DPCD and R1 and R2 relies on ATPase activity-dependent events that might be linked to subtle DII conformational changes during the ATPase cycle.

Biochemical and 3D structural characterization of DPCD in solution

At the protein level, DPCD is rarely described in the literature and almost no structural information is available, prompting us to characterize this protein in solution. The L156S natural variant (Uniprot: Q9BVM2) of DPCD (further referred as DPCD or D) was obtained with a high degree of purity (> 95%), assessed by SDS-PAGE with a band at ~ 25 kDa in line with the theoretical amino acid sequence and MS analysis (23,503.9 ± 1.0 Da) (Figure S1(A)). The calculated hydrodynamic radius was 24 Å (Table S2) and the oligomeric state of His-DPCD was determined using SEC-SLS with a molecular weight (MW) of ~ 30 kDa at the top of the peak, attesting that His-DPCD is monomeric in solution with HNGT buffer (Figure S1(B)). This was further confirmed



by native MS analysis, which also attested to the monomeric state of DPCD (Figure S1(C)). It is worth noting that DPCD tended to dimerize if the buffer was not supplemented with a reducing agent, meaning that oligomerization occurs through an inter-molecular disulfide bond.

Supplementing the buffer with TCEP prevented the formation of multimeric forms (Figure S1(C)).

To assess the structural determinants of DPCD, the secondary structure was predicted using PSIPRED⁵³ (Figure S2(A)). Analysis estimated ~ 35% β-sheet and 15% α-helix, while the

remaining 50% amino acids are likely organized in a coil. Circular dichroism was used to confirm these *in silico* results experimentally and revealed that His-DPCD has the same secondary structure (Figure S2(B)). Although folded, half DPCD is composed of random coil structures.

Indeed, apart from the prediction of the CHORD-containing proteins and SGT1 (CS) domain in the C-terminal, no experimental 3D structural data are available concerning full-length DPCD either free or bound to potential protein partners. Despite several attempts using X-ray crystallography (XRC) at high resolution, we failed to obtain structural information concerning the free form. We then moved to solution-state NMR spectroscopy with $^{13}\text{C}/^{15}\text{N}$ and $^2\text{H}/^{13}\text{C}/^{15}\text{N}$ labeled samples of DPCD and a set of 2D and 3D spectra recorded at 600 and 950 MHz. First, the $^1\text{H}-^{15}\text{N}$ TROSY-HSQC spectrum displayed the features of a folded monomeric protein, in agreement with SEC-SLS and native MS analyses (Figure S2(C)). Unfortunately, in-depth analysis of the 3D spectra revealed missing backbone resonances at various spots in the protein, preventing complete and trustworthy assignment of DPCD resonances. Changes in temperature or in the concentration of the sample did not modify its dynamic behavior. This phenomenon was likely due to unfavorable exchange regimes between several fragments of the protein. Nevertheless, around 75% of the H_N , N , C_O , C_α and C_β nuclei in the protein were successfully assigned. Even if it was not possible to transform these partial NMR data into an accurate 3D structure, the values of the chemical shifts obtained for some of the residues were derived into secondary structures using the TALOS-N software (Figure S2(A)).³⁷ The predicted (PSIPRED, upper) and experimental (NMR, lower) patterns of secondary structures are in good agreement, confirming the global folded structure of the protein. Respectively, 13%, 37% and 50% of the assigned residues were folded in a helix, sheet, and coil, in agreement with the CD data. Interestingly, the NMR data support the presence of a CS domain ranging from residue K99 to residue K172

(Figure S2(C)). Note that the long unstructured regions in DPCD may impair protein crystallization.

To circumvent these difficulties, we changed to low-resolution SEC-SAXS experiments (Figure S3) to obtain structural insights and molecular clues about DPCD conformation in solution (Figure 2 and Tables S2 and S3). First, using the Guinier approximation, a radius of gyration (R_g) value of ~ 22 Å was obtained. This value is in good agreement with the $R_{\text{IM-MS}}$ value of ~ 27 Å (monomeric DPCD, charge state 9+) derived from $^{\text{TW}}\text{CCS}_{\text{N}_2}$ measurements obtained by IM-MS (Table S2). $P(r)$ analysis (Figure 2(A), inset) showed a peak around 25 Å and a decrease to a D_{max} value of ~ 78 Å. Dimensionless Kratky plots and Porod-Debye plots are typical of a protein with both folded and unfolded, or asymmetrical domains, but with preserved compactness (Figures S3(B) and S3(C)). To shed light on DPCD 3D organization, *ab initio* low-resolution models were generated (Figure S3(D)). The *ab initio* models consisted of a disk-like shaped domain with an extended tail, in line with the $P(r)$ analysis. To go further in the 3D characterization of DPCD, *de novo* models were generated with the trRosetta algorithm⁴⁵ (Figures 2(B) and S3(E)), leading to four models with good scores (0.78). All the secondary structures predicted by the models are in line with PSIPRED prediction and TALOS-N determination (Figure S2(A)). In these models, a CS domain (K99-K172) is preceded by an additional domain of 5 antiparallel strands (Q14-Q66) (Figures 2(B) and S3(E)). Finally, the two α helices found in the N-term and C-term parts in the DPCD sequence are predicted to be close in space. To validate the accuracy and the reliability of these *de novo* models,⁴⁵ they were compared with SAXS experimental data (Figure 2(A)). The theoretical SAXS curves of 4 of the models fit the experimental data remarkably well with equivalent χ^2 (5.5–6.5) whereas the fifth, which differs in the position of an unstructured loop, did not fit the data properly ($\chi^2 \sim 12.5$). Derived R_g values of ~ 20 Å are in good agreement with the value of 22 Å determined experimentally. Of note, newly



Figure 1. *In vivo* characterization of the interaction between DPCD and RUVBL1/2. (A) Location of GFP-DPCD. HeLa cells were transfected with plasmid coding for GFP-DPCD fusion. DAPI was used for DNA staining. (B) Proteomic analysis of GFP-DPCD partners by SILAC-IP. Graphs show SILAC ratios (y axis, specific versus control IP) as a function of signal abundance (x axis, $\log_{10}(\text{intensity})$). Each dot represents a protein and is color coded according to the classification shown between panels A and B. The labeled dots highlight proteins that are relevant for this study. The full hit list with Significance B values is given in Table S1. (C-F) LUMIER-IP analysis of interactions. Graph plotting the % of co-IP efficiency for the interaction between co-expressed 3xFLAG-FL-R2TP subunits and RL-DPCD (C), 3xFLAG-FL-R2TP subunits or DPCD and RL-WDR92 (D), 3xFLAG-FL-DPCD and RL-R1 or R2 WT or mutant (E), 3xFLAG-FL-DPCD and RL-R1 or R2 with or without CB6644 treatment (F). The percentage of co-IP was calculated from the co-precipitated RL fusion protein (IP/Input ratios of RL activity), normalized with the IP/Input ratio of FL activity obtained in the anti-FLAG IP. The values are means of two or three experiments, each performed with duplicated values. ALIX was used as negative control. For each bait, Ct IP was done without antibody. Error bars represent standard deviations.

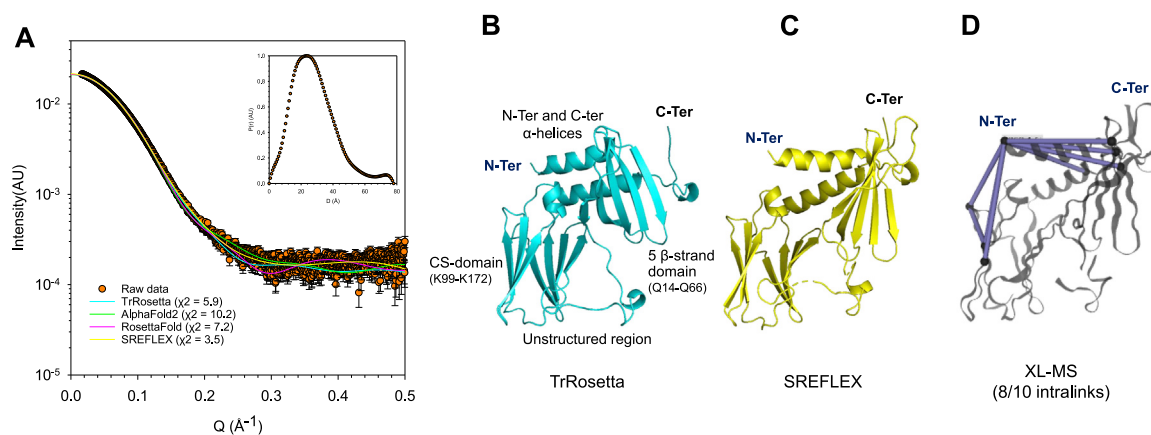


Figure 2. SEC-SAXS of DPCD. (A) Scattering intensity $I(q)$ vs Q plot (orange dots). Lines correspond to Crystol fits of the models described in B (blue), C (yellow). The comparison with AlphaFold2 (green) and RoseTTAFold (pink) models are also displayed. Pair distribution function analysis is shown in the inset. (B) The trRosetta model of DPCD with the lowest χ^2 value. (C) DPCD trRosetta model after SAXS/NMA refinement. (D) Intralinks identified by XL-MS analysis and in line with the DPCD model.

developed methods, RoseTTAFold⁴⁶ and AlphaFold2⁴⁷ produced similar DPCD models, but with lower agreement to SAXS data (Figure 2(A)). Despite a good initial fit of SAXS data, the trRosetta model with the lowest χ^2 value was subjected to SAXS/NMA structural refinement to improve agreement with experimental data (Figure 2(C)). As expected, few domain shifts were observed with increased data fitting ($\chi^2 \sim 3.5$ – 3.8). We next used XL-MS with DSBU as chemical reagent to further validate the *in silico* DPCD 3D model. Interestingly, most of the intra-DPCD cross-links (8/10) were compatible considering $C\alpha$ - $C\alpha$ distances ranging from 6 to 30 Å (Figure 2(D), Table S4). Seven of these cross-links include the N-term part of DPCD, which is composed of residual residues from 3C protease cleavage (GPH) and turned out to be useful to validate the DPCD model, since $C\alpha$ - $C\alpha$ distances are fully compatible with the DPCD model. A reliable model based on *de novo* prediction and in line with experimental SAXS and XL-MS data is therefore proposed as a basis for further structural investigations.

Three DPCD molecules interact with R1R2 heterohexamer with sub-micromolar affinity

The formation of the R1R2-DPCD (R1R2D) complex was first analyzed using SEC-UV followed by SDS-PAGE analysis (Figure 3(A), top), which confirmed co-elution of DPCD, R1 and R2 with subsequent formation of a ternary complex. A shift of the SEC-UV elution peak towards lower hydrodynamic radii for the R1R2D complex compared to R1R2 alone was observed, suggesting a change in R1R2 oligomeric state from dodecamer to hexamer upon DPCD binding. To assess the specificity of the R1R2D interaction, similar SEC-UV analyses were conducted with

R1R2 DII-truncated constructs (Figure 3(A), bottom). No co-elution was observed with R1R2 lacking DII domains, suggesting the interaction between R1R2 and DPCD mainly occurred through DII domains. The hydrodynamic radius of the R1R2D complex determined by SEC-UV using globular protein standards was smaller (73 Å) than that of R1R2 alone (94 Å). The same trend was confirmed by DLS in which the difference in the calculated hydrodynamic radius decreased from ~ 70 to 63 Å for R1R2 and R1R2D, respectively, (Figure S4) in line with the proposed shift from dodecamer to hexamer in the R1R2 ring scaffold.

To evaluate the stoichiometry of the R1R2D complex through MW measurements, SEC-SLS experiments were performed and allowed a MW measurement of ~ 370 kDa at the top of the peak, in agreement with one R1R2 hexamer interacting with three DPCD molecules (Figure 3(B)). To determine the R1R2D stoichiometry precisely, native MS experiments were performed and the binding of up to three DPCD molecules per R1R2 hexamer was validated ($383,611 \pm 66$ Da, Figure 3(C)) again confirming that DPCD binding can change the R1R2 oligomeric state from dodecamer to hexamer (that coexist in absence of DPCD, Figure S5).

ITC (Figure 3(D)) enabled measurement of a dissociation constant (K_d) of ~ 350 nM, attesting the high affinity between the protein partners. The calculated reaction stoichiometry (N-value) is ~ 2.7 DPCD per 1 R1R2 hexamer, in agreement with native SEC-SLS and native MS data. The thermodynamic signature is characteristic of hydrogen bonding with unfavorable conformational modifications with negative values for ΔG and ΔH and a positive value for $-T\Delta S$ (Figure 3(D), inset).

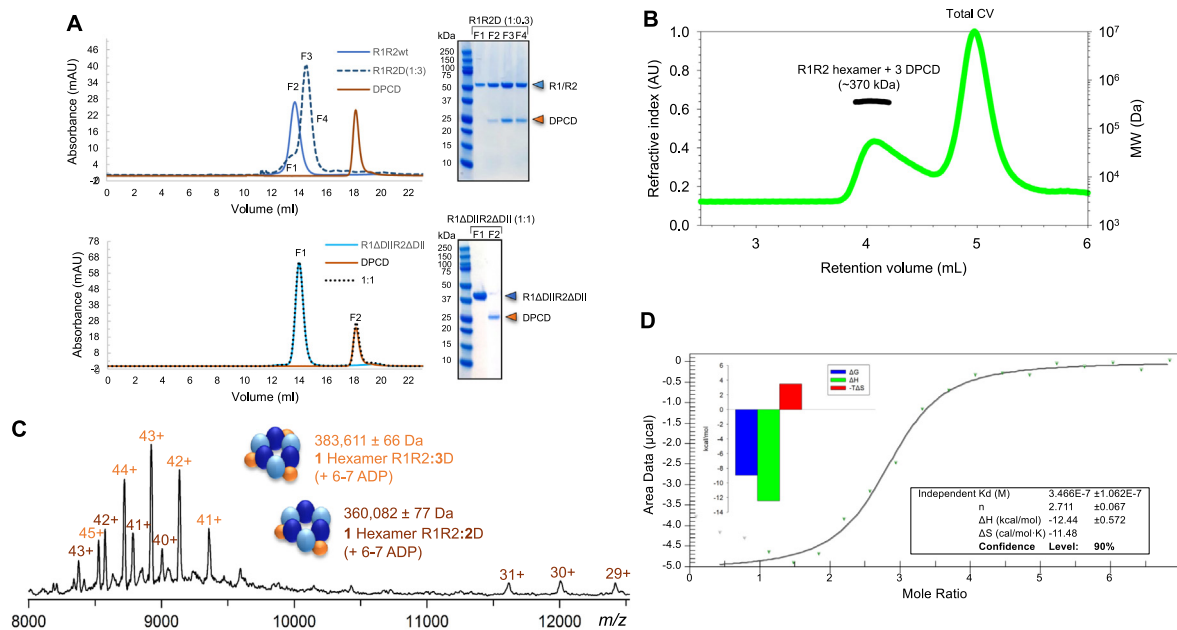


Figure 3. DPCD interacts directly with R1/R2. (A) Chromatographic co-elution assay followed by SDS-PAGE analysis. (B) SEC-SLS analysis highlighting a 3:3:3 stoichiometry of the R1R2D complex. CV and MW are the column volume and the molecular weight, respectively. (C) High resolution native MS analysis confirming the stoichiometry, i. e., 3 DPCD per 1 R1/R2 hexamer. (D) ITC analysis of the interaction. Dissociation constant (K_d), stoichiometry (n), variation in enthalpy (ΔH), in entropy (ΔS), and in free energy (ΔG) are reported along with the thermodynamic signature.

The assembly of the R1R2D complex was also assessed by DSF (Figure S6) to evaluate how complex association influences overall thermostability and how DPCD modulates the R1R2 ring oligomeric state *in-vitro*. R1R2 presents two thermal melting transitions (T_m), 48 °C and 54 °C (Figures S6(A and C)) that were tentatively assigned to the DII domains and the core ring region, respectively. This assignment is supported by the single melting transition observed for the R1R2_ΔDII complex, with a T_m of 53 °C (Figures S6(B and C)). DPCD alone presented a single melting transition at 55 °C, while the R1R2D complex showed a T_m value at 50 °C, suggesting complex formation via DII, thus stabilizing the lower observed T_m component assigned to DII. Conversely, no similar effect was observed for R1R2_ΔDII (DII_truncated) when incubated with DPCD (Figures S6(B and C)), as no interaction was expected. The emerging R1R2D hexameric complex may also justify overall lower thermostability compared to the R1R2 dodecamer.

Altogether, these complementary and orthogonal biophysical techniques allowed us to unambiguously demonstrate that DPCD binds to DII domains of R1 and R2 with a K_d in the sub-micromolar range and displaces the R1R2 dodecameric/hexameric equilibrium towards a stable R1R2D complex consisting of three DPCD per one R1R2 hexamer.

Topological investigation of the interaction between R1R2 and DPCD

First, SEC-SAXS experiments were performed on the R1R2D complex (Figures 4(A–D) and Tables S2 and S3). Remarkably, the R1R2D complex harbored a constant R_g in a R_g versus a frame number plot, attesting to the homogeneity and the stability of the complex even after several dilution/concentration steps (Figure 4(A)). The scattering curve of the R1R2D complex is shown in Figure 4(B). The R_g value determined for R1R2D was 54 Å (Table S3). In comparison, a R_g value of ~ 52 Å was determined for a pure R2 hexamer.²⁰ The same theoretical R_g value of ~ 52 Å was obtained for R1R2 hexamers both having DII up conformation (DII_{ext} stretched upward in the direction of the solvent, PDB: 7OLE). On the other hand, a theoretical R_g value of ~ 45 Å was determined for hexameric R1R2 with DII down conformation (DII_{ext} stretched downward), in the context of INO80 (PDB: 5OAF). The R_g value of R1R2D (54 Å) could be due to DII conformation and/or DPCD contribution which might be buried in DII domains or exposed to the solvent. The P(r) profiles reached maximum around 60 Å and a D_{max} of 159 Å (Figure 4(C)) in line with the dimension of DPCD with hexameric R1R2 and not with dodecameric R1R2. On dimensionless Kratky-plots (Figure 4(D))⁵⁴ the maximum observed at ~ 1.1 for a

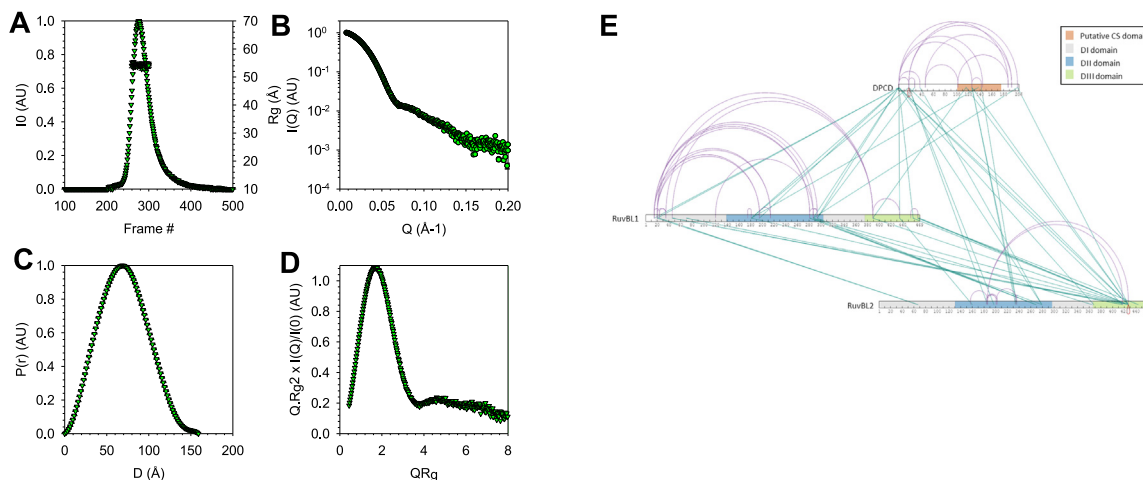


Figure 4. Comparative analysis by SAXS and XL-MS. (A) SEC-SAXS profile. $I(0)$ and R_g of different fractions measured across the single SEC peak. (B) Scattering intensity $I(Q)$ vs Q plot. (C) Pair distribution function analysis. (D) Dimensionless Kratky plots. R1R2 are in red and R1R2D in green. (E) XL-MS map (inter-links in green, intra-links in magenta).

QR_g value of 1.7 ($\sqrt{3}$) is typical of globular proteins, while higher QR_g values suggest the presence of flexible domains.

To use an orthogonal method to SEC-SAXS, we performed IM-MS measurements on R1R2 hexamers (Figure S5) and R1R2D under native conditions, which yielded collision cross section ($^{TW}CCS_{N_2}$) values of $13,081 \pm 10 \text{ \AA}^2$ and $15,225 \pm 34 \text{ \AA}^2$ for hexameric R1R2 (36 + charge state) and R1R2D (41 + charge state), respectively (Table S2). This $\sim 16\%$ increase in $^{TW}CCS_{N_2}$ upon DPCD binding to hexameric R1R2 agrees with both mass-based CCS estimations ($\sim 15\%$) and in-solution SAXS data ($\sim 20\%$). In addition, when $^{TW}CCS_{N_2}$ was converted into an experimental R_{IM-MS} radius for globular proteins, a value of $70 \pm 3 \text{ \AA}$ was obtained for the R1R2D complex, in agreement with $R_h = 73 \text{ \AA}$ as determined with SEC (Table S2), showing that the native global conformation of R1R2D is preserved in the gas phase. Taken together, the SEC-SAXS and IM-MS experiments support the hypothesis that the global shape of the hexameric R1R2 is maintained upon DPCD binding.

Knowing that the interaction occurs mainly via the DII domain, we next mapped which region of DPCD interacts with R1R2 using XL-MS experiments with DSBU, which reacts specifically with Lys residues, as chemical reagent (Figure 4(E)). To evaluate the reliability of our XL-MS experiments, we first compared R1R2 inter- and intra-cross-links on available R1R2 cryo-EM structures either in DII up (PDB: 7OLE) or DII down conformation (PDB: 5OAF). Considering a constraint distance cut-off of 6 to 30 \AA , more cross-linked peptides (both intra and inter) were validated for the DII down stretched structure (23/32 XL peptides, 72% for 5OAF compared to 20/31, 65% for 7OLE),

suggesting that DPCD rather binds to R1R2 in its DII down conformation. Concerning the R1R2D complex, 13 inter-cross-links (8 involving DPCD N-ter) were identified between DPCD and R1, while 14 cross-links (4 involving DPCD N-ter) were identified with R2. Interestingly, the inter-R1R2-DPCD cross-links identified included three different DPCD regions (Figure 4(E)): one at the N-terminus, another at the C-terminus parts corresponding to predicted α -helices located close together, the last region in the putative predicted CS domain. On R1 or R2 sides, DPCD cross-links were identified mainly in the DII domain (8/13 and 7/14 for R1 and R2, respectively) as expected (Figure 4(E)). Interestingly, comparison of intra cross-linked peptides for DPCD alone or in complex with R1R2 revealed mostly similar residues involved. However, two additional regions involving K46 and K102 were detected in R1R2D complex, suggesting conformational rearrangements within DPCD upon R1R2D complex formation (Figure S7).

Finally, negative staining transmission electron microscopy assays, followed by 2D classifications of the different observed complexes were conducted. Comparative observations of R1R2 (Figures 5(A–D) and S8, respectively with and without their DII domains) and *in vitro* reconstituted R1R2D complexes (Figures 5(C–D)) revealed clear morphological differences between R1R2 and R1R2D complexes. Most R1R2 particles (80% of the analyzed population) appeared to be rectangular, and their shape and dimensions ($\sim 13 \times 11 \text{ nm}$) corresponded to side views of dodecameric assemblies, made of two rings of hexamers (Figure 5(B)). The remaining 20% of the analyzed particles was grouped into 2D classes with round/hexameric shapes, of

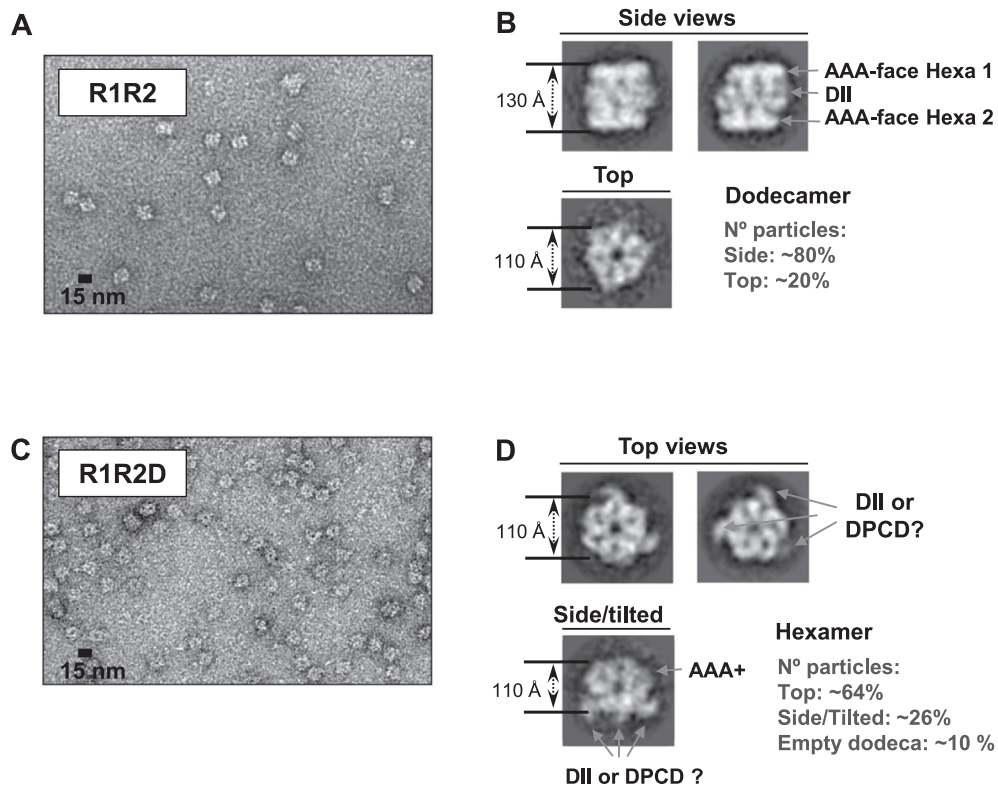


Figure 5. Overall structure of the R1R2D complex. Negative stain electron microscopy images of R1R2 (A) and R1R2D (C). Typical 2D class sums of R1R2 (B) and R1R2D (D) obtained after single particle analysis. The distribution of the analyzed particles according to their orientation (top, side, or tilted views) is indicated in the lower right panel.

11 nm in diameter, which were attributed to top views of this assembly (i.e., views of the hexameric rings of the dodecamers, along their 6-fold symmetry axis) (Figure 5(B) and S8). Similar observations were made for R1 Δ DII/R2 Δ DII particles, clearly harboring double-ringed side views typical of dodecameric assemblies (Figure S8). However, R1 Δ DII/R2 Δ DII particles seem to adopt a more rectangular shape than their full-length counterpart, which resembles more a parallelogram, suggesting a skewed piling of the two hexameric rings for the full-length R1R2 particles. In contrast, most R1R2D complexes (64%) appeared to be round/hexameric in shape, with a diameter of \sim 11 nm, which would be in good agreement with hexamers seen from the top. The presence of three prominences outside the round could correspond to the DIIs or DPCDs. Other views, representing 26% of the observed R1R2D complexes, were attributed to side or tilted views of single-ringed hexamers, while only 10% seem to correspond to side views of double-ringed hexamers, thus dodecameric assemblies. These observations strongly suggest that DPCD association to R1/R2 disrupts its dodecameric assembly.

Discussion

To date, although DPCD has appeared in several proteomic analyses, very little is known about its function. In this study, we have characterized the DPCD protein at the cellular, biochemical, molecular and topological level. DPCD associates with R1R2 *in vivo* to form a specific and strong R1R2D complex. It also associates with other R2TP/R2TP-like/PAQosome components or cofactors like RPAP3, PIH1D1, WDR92, POLR2E, ZNHIT2, and EFTUD2. According to the predicted 3D structure, DPCD is a monomeric protein likely comprising one β -sheet at the N-terminus, a flexible internal linker region, followed by a well-folded CS domain at the C-terminus, and finally two additional N- and C-ter α -helices. The direct interaction between DPCD and the R1R2 complex was evidenced and thoroughly investigated using a combination of biophysical and structural techniques. *In vitro*, the R1R2 complex includes up to 3 DPCD molecules in direct interaction with R1R2 in a hexameric ring (Figure 3(C)). The specific and high-affinity (sub-micromolar range) interaction with the R1R2 ring establishes it as a new AAA + ATPase cofactor. The interaction

occurs mainly via the DII domains of two consecutive R1R2 within hexameric state and through three distant regions on DPCD monomer, including the N-terminal helix, the putative CS domain but also the region at the C-terminal extremity. Finally, the negative stain electron microscopy images of R1R2D (Figure 5(D)) indicate the presence of three protuberances on the lower outer side of the DII-face of hexameric R1R2 which may correspond either to the DII_{ext} of 3R1 or 3R2 in up conformation or to the three molecules of DPCD.

The oligomeric state of R1(Rvb1)R2(Rvb2) in bound form is hexameric with the DII-face as a scaffold platform with the exception of INO80 Chromatin Remodeling Complexes for which are still being debated.^{55–57} It has been proposed that the dodecameric form of R1R2 functions as a storage form.^{18,58} In solution, the R1R2 complex alone with the RUVBL1 protomers N-terminally tagged, is predominantly dodecameric (Figures 3(A), 5(A–B), and S5). We used this oligomeric state as a tool to clearly demonstrate the formation of a complex between the R1R2 complex and DPCD. This observation showed that DPCD alters the dodecameric state of R1R2 *in vitro*, by interacting predominantly with the DIIs. This interaction also raises the question of whether DPCD can regulate the RUVBL's ATPase activity cycle. Using LUMIER-IP, we confirmed the strong interaction of DPCD with R1 WT and R2 WT but we detected no interaction with the R1^{E303Q} mutant form (weak for R2^{E300Q}). The ATP-hydrolysis deficient R1^{E303Q}R2^{E300Q} exhibited similar oligomerization to the wild-type complex,²³ which excludes the change in oligomerization of R1R2 as a driver of the absence of interaction in cells. Freezing of the ATP bound molecule in the nucleotide pocket by the mutants (inactive ATP state) seems to negatively affect the affinity between DPCD and the R1R2 complex. On the other hand, ADP had no effect on the interaction between R1R2 and DPCD *in vitro* and probably not *in vivo*. Despite the absence of ADP molecules in our RUVBL purification and analysis buffers, our native MS analysis showed trapping of 6 ADP molecules per hexamer. These nucleotides probably originate from *E. coli*. The interaction between R1R2 and DPCD therefore takes place in the presence of ADP *in vitro* (Figure 3(C)). As the number of ADP molecules in R1R2 did not change during the interaction, DPCD may drive the switch from dodecameric to hexameric forms of R1R2. From a mechanistic point of view, any factor that influences this dodecamer/hexamer transition is essential because this transition is the basis of the core activity of RUVBLs/Rvbs, associated with their chaperone function.⁵⁹

The influence of ATP on the R1R2 active site has previously been reported to affect partner binding. R1 and R2 associate with partners of the PAQosome assembly pathways as well as some

PAQosome substrates more efficiently in presence of one selective ATPase inhibitor, the compound B.¹⁶ This molecule, like its daughter, CB-6644, binds and stabilizes only the ATP-bound state of hetero-assembled R1R2.³² ATP hydrolysis could serve to mature the client or/and disengage hetero-hexameric R1R2 from its partners and clients. In cells, we showed that DPCD does not associate with the R1R2 in the presence of the CB-6644 inhibitor. We hypothesize that the *in vivo* ATP-bound R1R2 WT may become activated in a conformation where different partner proteins preferentially associate with R1R2, thus preventing or limiting DPCD binding.

After R1R2, our GFP:DPCD SILAC-IP results showed two enriched partners (Figure 1(B)), the components of human PAQosome, WDR92 and RPAP3. LUMIER-IP also revealed a strong association, suggesting direct *in cellulo* interaction between WDR92 and the N-terminal region of RPAP3 (Figure 1(D)). The association loss observed for full-size RPAP3 could be explained by the inaccessibility of its N-terminal part via homo-dimerization. Our finding confirms and refines, at the domain level, recent reports in *Drosophila* and *Chlamydomonas*.^{58,31} For years, it has been reported that human WDR92 associates with the TPR or C-ter domains of RPAP3.⁶¹ These domains allow the recruitment of chaperones HSP70/90 and RUVBLs, respectively. As an example, SILAC-IP experiments in HEK293T with GFP:RPAP3-Cter as bait did not show WDR92 to be the enriched partner but rather first, R1R2 and second, DPCD.¹¹ Human WDR92 and Wdr92 in *Drosophila* associate respectively with SPAG1, a paralog of RPAP3 in human and Spag1, a paralog of Spaghetti (Spag) in *Drosophila*.^{8,31} Like RPAP3 protein, human SPAG1 includes several TPRs and a RPAP3-C domains but *Drosophila* Spag1 only comprises a N-terminal region with a single TPR domain. Our demonstration of the RPAP3 N-terminal region as a WDR92 recognition site instead of its TPR or C-terminal domains explains why *Drosophila* Spag1 conserves the dynein assembly function of human SPAG1 proteins in Zur Lage's results.

The direct interaction between DPCD and R1R2 has just been assessed, but the DPCD behavior towards other R2TP or R2TP-like components remained to be elucidated. Notably, DPCD is expressed in ciliated cells, particularly in testis,^{7–8} like R2SP, which is also enriched in testis.¹¹ AP-MS analysis of *Drosophila* testis revealed association between Wdr92 and the prefoldin subunits, both dynein HCs and ICs, Spag, Spag1 and Dnaaf1.³¹ Wdr92 interacts directly with both dynein HCs and ICs. It is tempting to speculate that DPCD, SPAG1 and WDR92 are involved in regulating dynein assembly in the testis. It raises the question of whether DPCD could be a specific adapter of R2SP, a R2TP/PAQosome client, or an exclusive

partner of R1R2. Perhaps new DPCD partners still need to be identified to better understand its biological function. To this end, working with other cellular models is probably necessary. Use ciliated or testicular cells appear to be essential to fully decipher DPCD cellular function.

This study unveiled important functional and molecular determinants suggesting DPCD as a novel direct interactor of the R1R2 that regulate the oligomerization state of R1R2 complex and raises new questions about the regulatory mechanism of the ATPase activity cycle of the RUVBLs that is essential for cell metabolism. A high-resolution structural study by cryo-EM of the R1R2D complex would pave the way for more in-depth of interacting regions. The role of the R1R2D complex in primary ciliary dyskinesia should be elucidated in future research.

Data Availability

The NMR chemical shift have been deposited in the BMRB - Biological Magnetic Resonance Bank with entry 51533.

DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to thank S. Balor and V. Soldan (METI platform, Toulouse), for technical support in electron microscopy. The negative stain EM image analysis was performed using High Performance Computing resources from CALMIP (grant 2020-P1406). We thank the Biophysics and Structural Biology (B2S) platform of IBSLor (UMS 2008/US 40) for access to the NMR, ITC, CD, and SLS core facilities (<https://umsibslor.univ-lorraine.fr/en/facility/biophysics-structural-biology-b2s>). We also thank the Functional Proteomic Platform of Montpellier (FFP). Financial support from the IR-RMN-THC FR3050 CNRS for conducting the research is gratefully acknowledged. We also thank Javiez Perez and Aurélien Thureau for assistance with the SWING beamline. SOLEIL Synchrotron is acknowledged for providing beamtime and synchrotron radiation facilities. This work was supported by the French *Centre National de la Recherche Scientifique* (CNRS), Université de Lorraine (UL), Université de Strasbourg (Unistra), the *Agence Nationale de la Recherche* [ANR-16-CE11-0032-04], the *Ligue Nationale Contre le Cancer* (*équipe labellisée*), *Institut national du Cancer* [INCa PLBio 2016-161], and the French Proteomics Infrastructure (ProFI; ANR-10-INBS-08-03). The

authors would like to thank GIS IBISA and *Région Alsace* for financial support in purchasing a Synapt G2 HDMS and Orbitrap Exactive Plus EMR instrument, respectively. E.D. acknowledges the French Ministry for Education and Research for funding of her PhD fellowship. M.L. was granted by the ANR and S.H. by UCB. R.D.S.M. received young researcher funding from the *région Grand-Est*. Y.A. was supported by INCa and ANR. iNOVA4Health – UIDB/04462/2020 and UIDP/04462/2020, a program financially supported by *Fundação para a Ciência e Tecnologia / Ministério da Ciência, Tecnologia e Ensino Superior*, Portugal, through national funds is acknowledged.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jmb.2022.167760>.

Received 4 February 2022;

Accepted 21 July 2022;

Available online 25 July 2022

Keywords:

DPCD;
RUVBL1;
RUVBL2;
SILAC-IP;
LUMIER-IP;
SAXS;
structural MS;
EM

† Authors contributed equally.

References

1. Kobayashi, Y., Watanabe, M., Okada, Y., Sawa, H., Takai, H., Nakanishi, M., Kawase, Y., Suzuki, H., et al., (2002). Hydrocephalus, Situs Inversus, Chronic Sinusitis, and Male Infertility in DNA Polymerase λ -Deficient Mice: Possible Implication for the Pathogenesis of Immotile Cilia Syndrome. *Mol. Cell. Biol.* **22**, 2769–2776. <https://doi.org/10.1128/MCB.22.8.2769-2776.2002>.
2. King, S.M., Sale, W.S., (2018). Fifty years of microtubule sliding in cilia. *Mol. Biol. Cell* **29**, 698–701. <https://doi.org/10.1091/mbc.E17-07-0483>.
3. King, S.M., (2018). *Dyneins: Structure, Biology and Disease*. Elsevier: Academic Press, Oxford, UK.
4. Oda, T., Yanagisawa, H., Kamiya, R., Kikkawa, M., (2014). A molecular ruler determines the repeat length in eukaryotic cilia and flagella. *Science* **346**, 857–860. <https://doi.org/10.1126/science.1260214>.
5. Knowles, M.R., Daniels, L.A., Davis, S.D., Zariwala, M.A., Leigh, M.W., (2013). Primary Ciliary Dyskinesia. Recent Advances in Diagnostics, Genetics, and Characterization of Clinical Disease. *Am. J. Respir. Crit. Care Med.* **188**, 913–922. <https://doi.org/10.1164/rccm.201301-0059CI>.
6. Vogel, P., Read, R., Hansen, G.M., Freay, L.C., Zambrowicz, B.P., Sands, A.T., (2010). Situs inversus in

- Dpccd/Poll^{-/-}, Nme7^{-/-}, and Pkd11l1^{-/-} mice. *Vet. Pathol.* **47**, 120–131. <https://doi.org/10.1177/0300985809353553>.
7. Zariwala, M., O'Neal, W.K., Noone, P.G., Leigh, M.W., Knowles, M.R., Ostrowski, L.E., (2004). Investigation of the possible role of a novel gene, DPCD, in primary ciliary dyskinesia. *Am. J. Respir. Cell Mol. Biol.* **30**, 428–434. <https://doi.org/10.1165/rcmb.2003-0338RC>.
 8. Li, S., Ao, L., Yan, Y., Jiang, J., Chen, B., Duan, Y., Shen, F., Chen, J., et al., (2019). Differential motility parameters and identification of proteomic profiles of human sperm cryopreserved with cryostraw and cryovial. *Clin. Proteomics* **16**, 24. <https://doi.org/10.1186/s12014-019-9244-2>.
 9. Cloutier, P., Poitras, C., Durand, M., Hekmat, O., Fiola-Masson, É., Bouchard, A., Faubert, D., Chabot, B., et al., (2017). R2TP/Prefoldin-like component RUVBL1/RUVBL2 directly interacts with ZNHIT2 to regulate assembly of U5 small nuclear ribonucleoprotein. *Nature Commun* **8**, 15615. <https://doi.org/10.1038/ncomms15615>.
 10. Huttlin, E.L., Bruckner, R.J., Paulo, J.A., Cannon, J.R., Ting, L., Baltier, K., Colby, G., Gebreab, F., et al., (2017). Architecture of the human interactome defines protein communities and disease networks. *Nature* **545**, 505–509. <https://doi.org/10.1038/nature22366>.
 11. Maurizy, C., Quinternet, M., Abel, Y., Verheggen, C., Santo, P.E., Bourguet, M., Paiva, C.F., A., Bragantini, B., Chagot, M.-E., Robert, M.-C., (2018). The RPAP3-Cterminal domain identifies R2TP-like quaternary chaperones. *Nature Commun.* **9** <https://doi.org/10.1038/s41467-018-04431-1>.
 12. Rolland, T., Taşan, M., Charlotheaux, B., Pevzner, S.J., Zhong, Q., Sahni, N., Yi, S., Lemmens, I., et al., (2014). A proteome-scale map of the human interactome network. *Cell* **159**, 1212–1226. <https://doi.org/10.1016/j.cell.2014.10.050>.
 13. Dafinger, C., Rinschen, M.M., Borgal, L., Ehrenberg, C., Basten, S.G., Franke, M., Höhne, M., Rauh, M., et al., (2018). Targeted deletion of the AAA-ATPase Ruvbl1 in mice disrupts ciliary integrity and causes renal disease and hydrocephalus. *Exp. Mol. Med.* **50**, 75. <https://doi.org/10.1038/s12276-018-0108-z>.
 14. Mitchell, D.R., (2018). *Cytoplasmic preassembly and trafficking of axonemal dyneins. In: Dyneins: Structure, Biology and Disease. Elsevier: Academic Press, (Oxford, UK, pp. 141–161.*
 15. Zhao, R., Davey, M., Hsu, Y.-C., Kaplanek, P., Tong, A., Parsons, A.B., Krogan, N., Cagney, G., et al., (2005). Navigating the Chaperone Network: An Integrative Map of Physical and Genetic Interactions Mediated by the Hsp90 Chaperone. *Cell* **120**, 715–727. <https://doi.org/10.1016/j.cell.2004.12.024>.
 16. Yenerall, P., Das, A.K., Wang, S., Kollipara, R.K., Li, L.S., Villalobos, P., Flaming, J., Lin, Y.-F., et al., (2020). RUVBL1/RUVBL2 ATPase Activity Drives PAQosome Maturation, DNA Replication and Radioresistance in Lung Cancer. *Cell Chemical Biology* **27**, 105–121.e14. <https://doi.org/10.1016/j.chembiol.2019.12.005>.
 17. Gorynia, S., Bandejas, T.M., Pinho, F.G., McVey, C.E., Vonrhein, C., Round, A., Svergun, D.I., Donner, P., et al., (2011). Structural and functional insights into a dodecameric molecular machine – The RuvBL1/RuvBL2 complex. *J. Struct. Biol.* **176**, 279–291. <https://doi.org/10.1016/j.jsb.2011.09.001>.
 18. Lakomek, K., Stoeber, G., Tosi, A., Schmailzl, M., Hopfner, K.P., (2015). Structural basis for dodecameric assembly states and conformational plasticity of the full-length AAA+ ATPases Rvb1. Rvb2. *Structure* **23**, 483–495. <https://doi.org/10.1016/j.str.2014.12.015>.
 19. Matias, P.M., Gorynia, S., Donner, P., Carrondo, M.A., (2006). Crystal structure of the human AAA+ protein RuvBL1. *J. Biol. Chem.* **281**, 38918–38929. <https://doi.org/10.1074/jbc.M605625200>.
 20. Silva, S.T.N., Brito, J.A., Arranz, R., Sorzano, C.Ó.S., Ebel, C., Douth, J., Tully, M.D., Carazo, J.-M., et al., (2018). X-ray structure of full-length human RuvB-Like 2 – mechanistic insights into coupling between ATP binding and mechanical action. *Sci. Rep.* **8**, 13726. <https://doi.org/10.1038/s41598-018-31997-z>.
 21. Martino, F., Pal, M., Muñoz-Hernández, H., Rodríguez, C. F., Núñez-Ramírez, R., Gil-Carton, D., Degliesposti, G., Skehel, J.M., et al., (2018). RPAP3 provides a flexible scaffold for coupling HSP90 to the human R2TP co-chaperone complex. *Nature Commun.* **9**, 1501. <https://doi.org/10.1038/s41467-018-03942-1>.
 22. Muñoz-Hernández, H., Pal, M., Rodríguez, C.F., Fernandez-Leiro, R., Prodromou, C., Pearl, L.H., Llorca, O., (2019). Structural mechanism for regulation of the AAA-ATPases RUVBL1-RUVBL2 in the R2TP co-chaperone revealed by cryo-EM. *Science. Advances* **5**, eaaw1616. <https://doi.org/10.1126/sciadv.aaw1616>.
 23. López-Perrote, A., Hug, N., González-Corpas, A., Rodríguez, C.F., Serna, M., García-Martín, C., Boskovic, J., Fernandez-Leiro, R., et al., (2020). Regulation of RUVBL1-RUVBL2 AAA-ATPases by the nonsense-mediated mRNA decay factor DHX34, as evidenced by Cryo-EM. *ELife* **9**, e63042.
 24. Serna, M., González-Corpas, A., Cabezudo, S., López-Perrote, A., Degliesposti, G., Zarzuela, E., Skehel, J.M., Muñoz, J., Llorca, O., (2021). CryoEM of RUVBL1–RUVBL2–ZNHIT2, a complex that interacts with pre-mRNA-processing-splicing factor 8. *Acids Res. Nucl.* <https://doi.org/10.1093/nar/gkab1267>.
 25. Seraphim, T.V., Nano, N., Cheung, Y.W.S., Aluksanasuwan, S., Colletti, C., Mao, Y.-Q., Bhandari, V., Young, G., et al., (2021). Assembly principles of the human R2TP chaperone complex reveal the presence of R2T and R2P complexes. *Structure.* <https://doi.org/10.1016/j.str.2021.08.002>.
 26. Coulombe, B., Cloutier, P., Gauthier, M.-S., (2018). How do our cells build their protein interactome? *Nature Commun.* **9**, 2955. <https://doi.org/10.1038/s41467-018-05448-2>.
 27. Houry, W.A., Bertrand, E., Coulombe, B., (2018). The PAQosome, an R2TP-Based Chaperone for Quaternary Structure Formation. *Trends Biochem. Sci.* **43**, 4–9. <https://doi.org/10.1016/j.tibs.2017.11.001>.
 28. Cloutier, P., Poitras, C., Faubert, D., Bouchard, A., Blanchette, M., Gauthier, M.-S., Coulombe, B., (2020). Upstream ORF-Encoded ASDURF Is a Novel Prefoldin-like Subunit of the PAQosome. *J. Proteome Res.* **19**, 18–27. <https://doi.org/10.1021/acs.jproteome.9b00599>.
 29. Patel-King, R.S., Sakato-Antoku, M., Yankova, M., King, S. M., (2019). WDR92 is required for axonemal dynein heavy chain stability in cytoplasm. *MBoC* **30**, 1834–1845. <https://doi.org/10.1091/mbc.E19-03-0139>.
 30. Leader, D.P., Krause, S.A., Pandit, A., Davies, S.A., Dow, J.A.T., (2018). FlyAtlas 2: a new version of the Drosophila melanogaster expression atlas with RNA-Seq, miRNA-Seq

- and sex-specific data. *Nucleic Acids Res.* **46**, D809–D815. <https://doi.org/10.1093/nar/gkx976>.
31. Zur Lage, P., Stefanopoulou, P., Styczynska-Soczka, K., Quinn, N., Mali, G., von Kriegsheim, A., Mill, P., Jarman, A. P., (2018). Ciliary dynein motor preassembly is regulated by Wdr92 in association with HSP90 co-chaperone, R2TP. *The Journal of Cell Biology* **217**, 2583–2598. <https://doi.org/10.1083/jcb.201709026>.
 32. Assimon, V.A., Tang, Y., Vargas, J.D., Lee, G.J., Wu, Z.Y., Lou, K., Yao, B., Menon, M.-K., et al., (2019). CB-6644 Is a Selective Inhibitor of the RUVBL1/2 Complex with Anticancer Activity. *ACS Chem. Biol.* **14**, 236–244. <https://doi.org/10.1021/acscchembio.8b00904>.
 33. Diebold, M.L., Fribourg, S., Koch, M., Metzger, T., Romier, C., (2011). Deciphering correct strategies for multiprotein complex assembly by co-expression: Application to complexes as large as the histone octamer. *J. Struct. Biol.* <https://doi.org/10.1016/j.jsb.2011.02.001>.
 34. Micsonai, A., Wien, F., Kernya, L., Lee, Y.-H., Goto, Y., Réfrégiers, M., Kardos, J., (2015). Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy. *PNAS* **112**, E3095–E3103. <https://doi.org/10.1073/pnas.1500851112>.
 35. Micsonai, A., Wien, F., Bulyáki, É., Kun, J., Moussong, É., Lee, Y.-H., Goto, Y., Réfrégiers, M., et al., (2018). BeStSel: a web server for accurate protein secondary structure prediction and fold recognition from the circular dichroism spectra. *Nucleic Acids Res.* **46**, W315–W322. <https://doi.org/10.1093/nar/gky497>.
 36. Keller, R.L.J., (2004). *Computer-Aided Resonance Assignment Tutorial*. CANTINA Verlag, Zürich.
 37. Shen, Y., Bax, A., (2013). Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J. Biomol. NMR* **56**, 227–241. <https://doi.org/10.1007/s10858-013-9741-y>.
 38. Keller, S., Vargas, C., Zhao, H., Piszczek, G., Brautigam, C.A., Schuck, P., (2012). High-Precision Isothermal Titration Calorimetry with Automated Peak-Shape Analysis. *Anal. Chem.* **84**, 5066–5073. <https://doi.org/10.1021/ac3007522>.
 39. Bush, M.F., Hall, Z., Giles, K., Hoyes, J., Robinson, C.V., Ruotolo, B.T., (2010). Collision Cross Sections of Proteins and Their Complexes: A Calibration Framework and Database for Gas-Phase Structural Biology. *Anal. Chem.* **82**, 9557–9565. <https://doi.org/10.1021/ac1022953>.
 40. Ruotolo, B.T., Benesch, J.L.P., Sandercock, A.M., Hyung, S.-J., Robinson, C.V., (2008). Ion mobility–mass spectrometry analysis of large protein complexes. *Nature Protoc.* **3**, 1139–1152. <https://doi.org/10.1038/nprot.2008.78>.
 41. Shrivastav, V., Nahin, M., Hogan, C.J., Larriba-Andaluz, C., (2017). Benchmark Comparison for a Multi-Processing Ion Mobility Calculator in the Free Molecular Regime. *J. Am. Soc. Mass Spectrom.* **28**, 1540–1551. <https://doi.org/10.1021/jasms.8b05613>.
 42. Atmanene, C., Chaix, D., Bessin, Y., Declerck, N., Van Dorsselaer, A., Sanglier-Cianferani, S., (2010). Combination of Noncovalent Mass Spectrometry and Traveling Wave Ion Mobility Spectrometry Reveals Sugar-Induced Conformational Changes of Central Glycolytic Genes Repressor/DNA Complex. *Anal. Chem.* **82**, 3597–3605. <https://doi.org/10.1021/ac902784n>.
 43. Franke, D., Petoukhov, M.V., Konarev, P.V., Panjkovich, A., Tuukkanen, A., Mertens, H.D.T., Kikhney, A.G., Hajizadeh, et al., (2017). ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J Appl Cryst* **50**, 1212–1225. <https://doi.org/10.1107/S1600576717007786>.
 44. Rambo, R.P., Tainer, J.A., (2013). Accurate assessment of mass, models and resolution by small-angle scattering. *Nature* **496**, 477–481. <https://doi.org/10.1038/nature12070>.
 45. Yang, J., Anishchenko, I., Park, H., Peng, Z., Ovchinnikov, S., Baker, D., (2020). Improved protein structure prediction using predicted interresidue orientations. *PNAS* **117**, 1496–1503. <https://doi.org/10.1073/pnas.1914677117>.
 46. Baek, M., DiMaio, F., Anishchenko, I., Dauparas, J., Ovchinnikov, S., Lee, G.R., Wang, J., Cong, Q., et al., (2021). Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **373**, 871–876. <https://doi.org/10.1126/science.abj8754>.
 47. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., et al., (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589. <https://doi.org/10.1038/s41586-021-03819-2>.
 48. Panjkovich, A., Svergun, D.I., (2016). Deciphering conformational transitions of proteins by small angle X-ray scattering and normal mode analysis. *PCCP* **18**, 5707–5719. <https://doi.org/10.1039/C5CP04540A>.
 49. Valentini, E., Kikhney, A.G., Previtali, G., Jeffries, C.M., Svergun, D.I., (2015). SASBDB, a repository for biological small-angle scattering data. *Nucleic Acids Res.* **43**, D357–D363. <https://doi.org/10.1093/nar/gku1047>.
 50. Zivanov, J., Nakane, T., Forsberg, B.O., Kimanius, D., Hagen, W.J., Lindahl, E., Scheres, S.H., (2018). New tools for automated high-resolution cryo-EM structure determination in RELION-3. *ELife* **7**, e42166.
 51. Perez-Riverol, Y., Bai, J., Bandla, C., Garcia-Seisdedos, D., Hewapathirana, S., Kamatchinathan, S., Kundu, D.J., Prakash, A., et al., (2022). The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res.* **50**, D543–D552. <https://doi.org/10.1093/nar/gkab1038>.
 52. Barrios-Rodiles, M., Brown, K.R., Ozdamar, B., Bose, R., Liu, Z., Donovan, R.S., Shinjo, F., Liu, Y., et al., (2005). High-throughput mapping of a dynamic signaling network in mammalian cells. *Science* **307**, 1621–1625. <https://doi.org/10.1126/science.1105776>.
 53. Buchan, D.W.A., Jones, D.T., (2019). The PSIPRED Protein Analysis Workbench: 20 years on. *Nucleic Acids Res.* **47**, W402–W407. <https://doi.org/10.1093/nar/gkz297>.
 54. Receveur-Brechot, V., Durand, D., (2012). How random are intrinsically disordered proteins? A small angle scattering perspective. *Curr. Protein Pept. Sci.* **13**, 55–75. <https://doi.org/10.2174/138920312799277901>.
 55. Ayala, R., Willhoft, O., Aramayo, R.J., Wilkinson, M., McCormack, E.A., Ocloo, L., Wigley, D.B., Zhang, X., (2018). Structure and regulation of the human INO80–nucleosome complex. *Nature* **556**, 391–395. <https://doi.org/10.1038/s41586-018-0021-6>.
 56. Eustermann, S., Schall, K., Kostrewa, D., Lakomek, K., Strauss, M., Moldt, M., Hopfner, K.-P., (2018). Structural basis for ATP-dependent chromatin remodelling by the INO80 complex. *Nature* **556**, 386–390. <https://doi.org/10.1038/s41586-018-0029-y>.
 57. Tosi, A., Haas, C., Herzog, F., Gilmozzi, A., Berninghausen, O., Ungewickell, C., Gerhold, C.B., Lakomek, K., et al., (2013). Structure and subunit

- topology of the INO80 chromatin remodeler and its nucleosome complex. *Cell* **154**, 1207–1219. <https://doi.org/10.1016/j.cell.2013.08.016>.
58. Silva-Martin, N., Dauden, M.I., Glatt, S., Hoffmann, N.A., Kastritis, P., Bork, P., Beck, M., Muller, C.W., (2016). The Combination of X-Ray Crystallography and Cryo-Electron Microscopy Provides Insight into the Overall Architecture of the Dodecameric Rvb1/Rvb2 Complex. *PLoS ONE* **11**, e0146457.
59. Zhou, C.Y., Stoddard, C.I., Johnston, J.B., Trnka, M.J., Echeverria, I., Palovcak, E., Sali, A., Burlingame, A.L., et al., (2017). Regulation of Rvb1/Rvb2 by a Domain within the INO80 Chromatin Remodeling Complex Implicates the Yeast Rvbs as Protein Assembly Chaperones. *Cell Reports* **19**, 2033–2044. <https://doi.org/10.1016/j.celrep.2017.05.029>.
60. Liu, G., Wang, L., Pan, J., (2018). Chlamydomonas WDR92 in association with R2TP-like complex and multiple DNAAFs to regulate ciliary dynein preassembly. *J. Mol. Cell. Biol.* <https://doi.org/10.1093/jmcb/mjy067>.
61. Itsuki, Y., Saeki, M., Nakahara, H., Egusa, H., Irie, Y., Terao, Y., Kawabata, S., Yatani, H., et al., (2008). Molecular cloning of novel Monad binding protein containing tetratricopeptide repeat domains. *FEBS Letter* **582**, 2365–2370. <https://doi.org/10.1016/j.febslet.2008.05.041>.
62. Cox, J., Hein, M.Y., Luber, C.A., Paron, I., Nagaraj, N., Mann, M., (2014). Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Molecular & Cellular Proteomics : MCP* **13**, 2513–2526. <https://doi.org/10.1074/mcp.M113.031591>.

Annexe 49

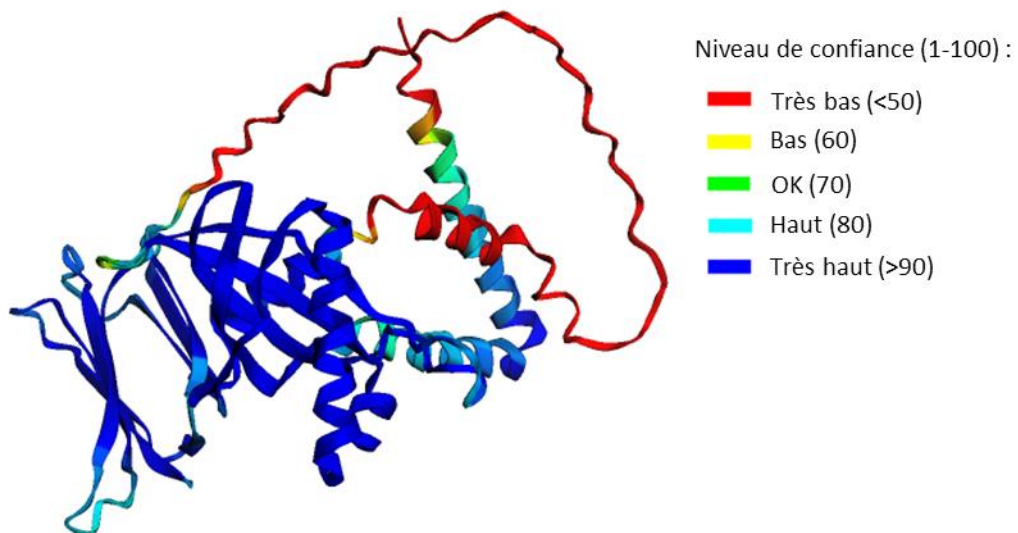
Tableau descriptif des peptides cross-linkés (protéines, sites, domaines, distances Ca-Ca) identifiés suite au XL-MS de SPAG1₆₂₂₋₉₂₆/PIH1D2.

| Protéine 1 | Protéine 2 | Site 1 | Site 2 | Domaine site 1 | Domaine site 2 | Distance Ca-Ca | |
|------------|------------|--------|--------|----------------|----------------|----------------|------|
| | | | | | | R1R2SP | 6I57 |
| SPAG1 | SPAG1 | K627 | K655 | TPR3 | TPR3 | - | 13,1 |
| SPAG1 | SPAG1 | K627 | K692 | TPR3 | TPR3 | - | 16,4 |
| SPAG1 | SPAG1 | K637 | K669 | TPR3 | TPR3 | - | 8,6 |
| SPAG1 | SPAG1 | K637 | K692 | TPR3 | TPR3 | - | 18,5 |
| SPAG1 | SPAG1 | K640 | K669 | TPR3 | TPR3 | - | 8,5 |
| SPAG1 | SPAG1 | K645 | K669 | TPR3 | TPR3 | - | 15,6 |
| SPAG1 | SPAG1 | K655 | K692 | TPR3 | TPR3 | - | 15,4 |
| SPAG1 | SPAG1 | Y667 | K692 | TPR3 | TPR3 | - | 14 |
| SPAG1 | SPAG1 | K669 | K702 | TPR3 | TPR3 | - | 13,3 |
| SPAG1 | SPAG1 | K702 | K741 | TPR3 | - | - | 13,8 |
| SPAG1 | SPAG1 | K705 | K741 | - | - | - | 10,4 |
| SPAG1 | SPAG1 | K705 | K743 | - | - | - | - |
| SPAG1 | SPAG1 | K705 | K749 | - | - | - | - |
| SPAG1 | SPAG1 | K705 | K755 | - | - | - | - |
| SPAG1 | SPAG1 | K709 | K741 | - | - | - | 5,6 |
| SPAG1 | SPAG1 | K709 | K743 | - | - | - | - |
| SPAG1 | SPAG1 | K709 | K749 | - | - | - | - |
| SPAG1 | SPAG1 | K709 | K755 | - | - | - | - |
| SPAG1 | SPAG1 | K741 | K755 | - | - | - | - |
| SPAG1 | SPAG1 | K741 | K787 | - | RPAP3 | - | - |
| SPAG1 | SPAG1 | K743 | K755 | - | - | - | - |
| SPAG1 | SPAG1 | K749 | K755 | - | - | - | - |
| SPAG1 | SPAG1 | K755 | K787 | - | RPAP3 | - | - |
| SPAG1 | SPAG1 | K755 | K822 | - | RPAP3 | - | - |
| SPAG1 | SPAG1 | K787 | K878 | RPAP3 | RPAP3 | - | - |
| SPAG1 | SPAG1 | K797 | K822 | RPAP3 | RPAP3 | 9,9 | - |
| SPAG1 | SPAG1 | K822 | K891 | RPAP3 | RPAP3 | 21,7 | - |
| PIH1D2 | PIH1D2 | N-ter | N-ter | - | - | - | - |
| PIH1D2 | PIH1D2 | N-ter | K42 | - | PIH | - | - |
| PIH1D2 | PIH1D2 | N-ter | K63 | - | PIH | - | - |
| PIH1D2 | PIH1D2 | N-ter | K65 | - | PIH | - | - |
| PIH1D2 | PIH1D2 | N-ter | K128 | - | PIH | - | - |
| PIH1D2 | PIH1D2 | N-ter | K159 | - | PIH | - | - |
| PIH1D2 | PIH1D2 | N-ter | K166 | - | PIH | - | - |
| PIH1D2 | PIH1D2 | K42 | K159 | PIH | PIH | - | - |
| PIH1D2 | PIH1D2 | K42 | K166 | PIH | PIH | - | - |
| PIH1D2 | PIH1D2 | K45 | K166 | PIH | PIH | - | - |
| PIH1D2 | SPAG1 | N-ter | K669 | - | TPR3 | - | - |
| PIH1D2 | SPAG1 | N-ter | K741 | - | - | - | - |
| PIH1D2 | SPAG1 | N-ter | K743 | - | - | - | - |

| Protéine 1 | Protéine 2 | Site 1 | Site 2 | Domaine site 1 | Domaine site 2 | Distance α - α | |
|------------|------------|--------|--------|----------------|----------------|------------------------------|------|
| | | | | | | R1R2SP | 6I57 |
| PIH1D2 | SPAG1 | N-ter | K822 | - | RPAP3 | - | - |
| PIH1D2 | SPAG1 | K42 | K627 | PIH | TPR3 | - | - |
| PIH1D2 | SPAG1 | K42 | K640 | PIH | TPR3 | - | - |
| PIH1D2 | SPAG1 | K42 | K669 | PIH | TPR3 | - | - |
| PIH1D2 | SPAG1 | K42 | K741 | PIH | - | - | - |
| PIH1D2 | SPAG1 | K42 | K743 | PIH | - | - | - |
| PIH1D2 | SPAG1 | K63 | K640 | PIH | TPR3 | - | - |
| PIH1D2 | SPAG1 | K63 | K705 | PIH | - | - | - |
| PIH1D2 | SPAG1 | K63 | K787 | PIH | RPAP3 | - | - |
| PIH1D2 | SPAG1 | K63 | K822 | PIH | RPAP3 | - | - |
| PIH1D2 | SPAG1 | K65 | K640 | PIH | TPR3 | - | - |
| PIH1D2 | SPAG1 | K65 | K787 | PIH | RPAP3 | - | - |
| PIH1D2 | SPAG1 | K159 | K640 | - | TPR3 | - | - |
| PIH1D2 | SPAG1 | K159 | K741 | - | - | - | - |
| PIH1D2 | SPAG1 | K159 | K743 | - | - | - | - |
| PIH1D2 | SPAG1 | K159 | K787 | - | RPAP3 | - | - |
| PIH1D2 | SPAG1 | K159 | K822 | - | RPAP3 | - | - |
| PIH1D2 | SPAG1 | K293 | K741 | CS | - | - | - |

Annexe 50

Modèle AlphaFold de PIH1D2 avec la représentation des niveaux de confiance donnés.



Annexe 51

Tableaux descriptifs des peptides cross-linkés (protéines, sites, domaines et des distances Cα-Cα) identifiés suite au XL-MS de R2SP.

| Protéine 1 | Protéine 2 | Site 1 | Site 2 | Domaine site 1 | Domaine site 2 | Distance Cα-Cα | | |
|------------|------------|-----------|--------|----------------|----------------|----------------|-------|------|
| | | | | | | R1R2SP | 7OLE | 5OAF |
| RuvBL1 | RuvBL1 | K15 | K287 | - | DII | - | - | - |
| RuvBL1 | RuvBL1 | K15 | K294 | - | DII | - | - | - |
| RuvBL1 | RuvBL1 | K15 | K389 | - | DIII | - | - | - |
| RuvBL1 | RuvBL1 | K15 | K435 | - | DIII | - | - | - |
| RuvBL1 | RuvBL1 | K20 | K35 | DI | DI | 19,8 | 23,3 | - |
| RuvBL1 | RuvBL1 | K20 | K195 | DI | DII | - | 18,9 | - |
| RuvBL1 | RuvBL1 | K20 | K214 | DI | DII | - | 15,4 | - |
| RuvBL1 | RuvBL1 | K20 | K281 | DI | DII | - | - | - |
| RuvBL1 | RuvBL1 | K20 | K287 | DI | DII | 23,9 | 21,4 | - |
| RuvBL1 | RuvBL1 | K24 | K281 | DI | DII | - | - | 40 |
| RuvBL1 | RuvBL1 | K35 | K214 | DI | DII | - | 26,6 | 42,2 |
| RuvBL1 | RuvBL1 | K35 | K389 | DI | DIII | 15,9 | 16,4 | 16,2 |
| RuvBL1 | RuvBL1 | K46 | K389 | DI | DIII | 17,4 | 18,3 | 17,4 |
| RuvBL1 | RuvBL1 | K120 | K281 | DI | DII | - | - | - |
| RuvBL1 | RuvBL1 | K120 | K287 | DI | DII | 12,5 | 16,6 | 19,1 |
| RuvBL1 | RuvBL1 | K175 | K238 | DII | DII | - | 12,6 | 10,2 |
| RuvBL1 | RuvBL1 | K178 | K181 | DII | DII | - | 10,3 | 10,1 |
| RuvBL1 | RuvBL1 | K385 | K413 | DIII | DIII | 13,6 | 13,6 | 12,5 |
| RuvBL1 | RuvBL1 | K389 | K435 | DIII | DIII | 18,5 | 18,4 | 21,3 |
| RuvBL1 | RuvBL1 | K413 | K440 | DIII | DIII | 17,3 | 16,4 | 16,1 |
| RuvBL1 | RuvBL1 | K413 | K454 | DIII | DIII | 10,2 | 11,5 | 10,3 |
| RuvBL1 | RuvBL1 | K413 | K458 | DIII | DIII | 13,9 | 15,3 | 14,7 |
| RuvBL1 | RuvBL1 | K458 | K466 | DIII | DIII | - | - | - |
| RuvBL1 | RuvBL1 | K458 | K469 | DIII | DIII | - | - | - |
| RuvBL1 | RuvBL1 | K469 | K469 | DIII | DIII | - | - | - |
| RuvBL2 | RuvBL2 | K67 | K184 | DI | DII | - | 51,5 | 49,1 |
| RuvBL2 | RuvBL2 | K132 | K236 | DI | DII | - | 8,6 | 11,9 |
| RuvBL2 | RuvBL2 | K157 | K184 | DII | DII | - | 19,3 | 16,3 |
| RuvBL2 | RuvBL2 | K157 | K186 | DII | DII | - | 22,2 | 18,9 |
| RuvBL2 | RuvBL2 | K157 | K197 | DII | DII | - | 12,4 | 21,5 |
| RuvBL2 | RuvBL2 | K177 | K197 | DII | DII | - | 14 | 12,2 |
| RuvBL2 | RuvBL2 | S181/T183 | K197 | DII | DII | - | 14,3 | 13 |
| RuvBL2 | RuvBL2 | K184 | K201 | DII | DII | - | 26,1 | 9,5 |
| RuvBL2 | RuvBL2 | K184 | K223 | DII | DII | - | 30,4 | 24,1 |
| RuvBL2 | RuvBL2 | K186 | K204 | DII | DII | - | 21,4 | 10,2 |
| RuvBL2 | RuvBL2 | K186 | K223 | DII | DII | - | 29,4 | 18,3 |
| RuvBL2 | RuvBL2 | K197 | K197 | DII | DII | - | 104,6 | 58,4 |
| RuvBL2 | RuvBL2 | K197 | K269 | DII | DII | - | 44,6 | 36,7 |
| RuvBL2 | RuvBL2 | K201 | K223 | DII | DII | - | 32,6 | 21,7 |
| RuvBL2 | RuvBL2 | K201 | K234 | DII | DII | - | 14 | 18,2 |

| Protéine 1 | Protéine 2 | Site 1 | Site 2 | Domaine site 1 | Domaine site 2 | Distance Cα-Cα | | |
|------------|------------|--------|--------|----------------|----------------|----------------|------|------|
| | | | | | | R1R2SP | 7OLE | 5OAF |
| RuvBL2 | RuvBL2 | K201 | K236 | DII | DII | - | 18,4 | 16,1 |
| RuvBL2 | RuvBL2 | K201 | K279 | DII | DII | - | 41,5 | 29,7 |
| RuvBL2 | RuvBL2 | K223 | K236 | DII | DII | - | 30,3 | 24,9 |
| RuvBL2 | RuvBL2 | K427 | Y442 | DIII | DIII | 18,9 | 18 | 19,4 |
| RuvBL1 | RuvBL2 | K15 | K67 | - | DI | - | - | - |
| RuvBL1 | RuvBL2 | K20 | K67 | DI | DI | 23,6 | 22,2 | - |
| RuvBL1 | RuvBL2 | K24 | K67 | DI | DI | 16,4 | 15,2 | 9,9 |
| RuvBL1 | RuvBL2 | K46 | K417 | DI | DIII | 20,6 | 14,5 | 21,1 |
| RuvBL1 | RuvBL2 | K46 | K427 | DI | DIII | 20,5 | 17,6 | 21,8 |
| RuvBL1 | RuvBL2 | K70 | K417 | DI | DIII | 17,4 | 23,5 | 16,7 |
| RuvBL1 | RuvBL2 | K73 | K184 | DI | DII | - | 45,6 | 44,5 |
| RuvBL1 | RuvBL2 | K73 | K186 | DI | DII | - | 45,6 | 41,2 |
| RuvBL1 | RuvBL2 | K73 | K417 | DI | DIII | 18,9 | 29,5 | 17,9 |
| RuvBL1 | RuvBL2 | K120 | K269 | DI | DII | 20,4 | 21 | 23,4 |
| RuvBL1 | RuvBL2 | K138 | K184 | DI | DII | - | 38,3 | 26,3 |
| RuvBL1 | RuvBL2 | K195 | K67 | DII | DI | - | 24,8 | 48,2 |
| RuvBL1 | RuvBL2 | K281 | K269 | DII | DII | - | - | 14,5 |
| RuvBL1 | RuvBL2 | K281 | K279 | DII | DII | - | - | 28,5 |
| RuvBL1 | RuvBL2 | K287 | K269 | DII | DII | 14,7 | 31,1 | 26,9 |
| RuvBL1 | RuvBL2 | K294 | K197 | DII | DII | - | 37,2 | 27,5 |
| RuvBL1 | RuvBL2 | K294 | K201 | DII | DII | - | 43,1 | 21,5 |
| RuvBL1 | RuvBL2 | K294 | S203 | DII | DII | - | 39,3 | 19,8 |
| RuvBL1 | RuvBL2 | K294 | K269 | DII | DII | 22,9 | 21,3 | 26 |
| RuvBL1 | RuvBL2 | K454 | K444 | DIII | DIII | 15,9 | 23,4 | 21,9 |
| RuvBL1 | RuvBL2 | K466 | K427 | DIII | DIII | - | - | - |
| RuvBL1 | RuvBL2 | K466 | K444 | DIII | DIII | - | - | - |
| RuvBL1 | RuvBL2 | K469 | K368 | DIII | DIII | - | - | - |
| RuvBL1 | RuvBL2 | K469 | K427 | DIII | DIII | - | - | - |
| RuvBL1 | RuvBL2 | K469 | K444 | DIII | DIII | - | - | - |

| Protéine 1 | Protéine 2 | Site 1 | Site 2 | Domaine site 1 | Domaine site 2 | Distance Cα-Cα | |
|------------|------------|--------|--------|----------------|----------------|----------------|------|
| | | | | | | R1R2SP | 6I57 |
| RuvBL1 | SPAG1 | K35 | K787 | DI | - | - | - |
| RuvBL1 | SPAG1 | K46 | K765 | DI | - | - | - |
| RuvBL1 | SPAG1 | K46 | K784 | DI | - | - | - |
| RuvBL1 | SPAG1 | K46 | K822 | DI | RPAP3 | 20,9 | - |
| RuvBL1 | SPAG1 | S50 | K784 | DI | - | - | - |
| RuvBL1 | SPAG1 | K138 | K755 | DI | - | - | - |
| RuvBL1 | SPAG1 | K175 | K743 | DII | - | - | - |
| RuvBL1 | SPAG1 | K178 | K741 | DII | - | - | - |
| RuvBL1 | SPAG1 | K178 | K743 | DII | - | - | - |
| RuvBL1 | SPAG1 | K178 | K749 | DII | - | - | - |
| RuvBL1 | SPAG1 | K178 | K755 | DII | - | - | - |
| RuvBL1 | SPAG1 | K181 | K741 | DII | - | - | - |
| RuvBL1 | SPAG1 | K385 | K797 | DIII | - | 14,3 | - |

| Protéine 1 | Protéine 2 | Site 1 | Site 2 | Domaine site 1 | Domaine site 2 | Distance Cα-Cα | |
|------------|------------|--------|-----------|----------------|----------------|----------------|------|
| | | | | | | R1R2SP | 6157 |
| RuvBL1 | SPAG1 | K389 | K784 | DIII | - | - | - |
| RuvBL1 | SPAG1 | K389 | K787 | DIII | - | - | - |
| RuvBL1 | SPAG1 | K435 | K627 | DIII | TPR3 | - | - |
| RuvBL1 | SPAG1 | K435 | K651 | DIII | TPR3 | - | - |
| RuvBL1 | SPAG1 | K435 | K784 | DIII | - | - | - |
| RuvBL1 | SPAG1 | K435 | K787 | DIII | - | - | - |
| RuvBL2 | SPAG1 | K186 | K749 | DII | - | - | - |
| RuvBL2 | SPAG1 | K368 | K891 | DIII | RPAP3 | 21,5 | - |
| RuvBL2 | SPAG1 | K417 | K765 | DIII | - | - | - |
| RuvBL2 | SPAG1 | K417 | K784 | DIII | - | - | - |
| RuvBL2 | SPAG1 | K417 | K787 | DIII | - | - | - |
| RuvBL2 | SPAG1 | K427 | K787 | DIII | - | - | - |
| RuvBL2 | SPAG1 | K427 | K820 | DIII | RPAP3 | 17,4 | - |
| RuvBL2 | SPAG1 | K427 | K822 | DIII | RPAP3 | 21,4 | - |
| SPAG1 | SPAG1 | K627 | K655 | TPR3 | TPR3 | - | 13,1 |
| SPAG1 | SPAG1 | K627 | Y660/T661 | TPR3 | TPR3 | - | 10,9 |
| SPAG1 | SPAG1 | K627 | K692 | TPR3 | TPR3 | - | 16,4 |
| SPAG1 | SPAG1 | K637 | K669 | TPR3 | TPR3 | - | 8,6 |
| SPAG1 | SPAG1 | K637 | K692 | TPR3 | TPR3 | - | 18,5 |
| SPAG1 | SPAG1 | K640 | K669 | TPR3 | TPR3 | - | 8,5 |
| SPAG1 | SPAG1 | K645 | K669 | TPR3 | TPR3 | - | 15,6 |
| SPAG1 | SPAG1 | K669 | K702 | TPR3 | TPR3 | - | 13,3 |
| SPAG1 | SPAG1 | K705 | K741 | TPR3 | - | - | 10,4 |
| SPAG1 | SPAG1 | K705 | K743 | TPR3 | - | - | - |
| SPAG1 | SPAG1 | K705 | K749 | TPR3 | - | - | - |
| SPAG1 | SPAG1 | K705 | K755 | TPR3 | - | - | - |
| SPAG1 | SPAG1 | K709 | K741 | TPR3 | - | - | 15,6 |
| SPAG1 | SPAG1 | K709 | K743 | TPR3 | - | - | - |
| SPAG1 | SPAG1 | K709 | K749 | TPR3 | - | - | - |
| SPAG1 | SPAG1 | K743 | K755 | - | - | - | - |
| SPAG1 | SPAG1 | S776 | K822 | - | RPAP3 | - | - |
| SPAG1 | SPAG1 | S782 | K822 | - | RPAP3 | - | - |
| SPAG1 | SPAG1 | K784 | K822 | - | RPAP3 | - | - |
| SPAG1 | SPAG1 | K797 | K822 | - | RPAP3 | 9,9 | - |
| SPAG1 | SPAG1 | K797 | K835 | - | RPAP3 | 26,4 | - |
| SPAG1 | SPAG1 | K797 | K864 | - | RPAP3 | 20,5 | - |
| SPAG1 | SPAG1 | K802 | K835 | RPAP3 | RPAP3 | 13,9 | - |
| SPAG1 | SPAG1 | K822 | K864 | RPAP3 | RPAP3 | 14,1 | - |
| SPAG1 | SPAG1 | K822 | K891 | RPAP3 | RPAP3 | 21,7 | - |
| SPAG1 | SPAG1 | K835 | K878 | RPAP3 | RPAP3 | 11,2 | - |

| Protéine 1 | Protéine 2 | Site 1 | Site 2 | Domaine site 1 | Domaine site 2 |
|------------|------------|--------|-----------|----------------|----------------|
| PIH1D2 | PIH1D2 | N-ter | N-ter | - | - |
| PIH1D2 | PIH1D2 | N-ter | K42 | - | PIH |
| PIH1D2 | PIH1D2 | N-ter | K45 | - | PIH |
| PIH1D2 | PIH1D2 | N-ter | K63 | - | PIH |
| PIH1D2 | PIH1D2 | N-ter | K65 | - | PIH |
| PIH1D2 | PIH1D2 | N-ter | K127 | - | PIH |
| PIH1D2 | PIH1D2 | N-ter | K128 | - | PIH |
| PIH1D2 | PIH1D2 | N-ter | K159 | - | - |
| PIH1D2 | PIH1D2 | N-ter | K166 | - | - |
| PIH1D2 | PIH1D2 | N-ter | K210 | - | - |
| PIH1D2 | PIH1D2 | K42 | K42 | PIH | PIH |
| PIH1D2 | PIH1D2 | K42 | K63 | PIH | PIH |
| PIH1D2 | PIH1D2 | K42 | K127 | PIH | PIH |
| PIH1D2 | PIH1D2 | K42 | K166 | PIH | - |
| PIH1D2 | PIH1D2 | K45 | K166 | PIH | - |
| PIH1D2 | PIH1D2 | K63 | K159 | PIH | - |
| PIH1D2 | PIH1D2 | K65 | K65 | PIH | PIH |
| PIH1D2 | PIH1D2 | K65 | K159 | PIH | - |
| PIH1D2 | PIH1D2 | K128 | K155 | PIH | - |
| PIH1D2 | PIH1D2 | K128 | K159 | PIH | - |
| PIH1D2 | PIH1D2 | K159 | K159 | - | - |
| PIH1D2 | PIH1D2 | K159 | S194/S195 | - | - |
| PIH1D2 | PIH1D2 | K166 | K210 | - | - |
| PIH1D2 | PIH1D2 | K252 | K303 | CS | CS |
| PIH1D2 | RuvBL1 | N-ter | K15 | - | - |
| PIH1D2 | RuvBL1 | N-ter | K20 | - | DI |
| PIH1D2 | RuvBL1 | N-ter | K73 | - | DI |
| PIH1D2 | RuvBL1 | N-ter | K138 | - | DI |
| PIH1D2 | RuvBL1 | N-ter | K175 | - | DII |
| PIH1D2 | RuvBL1 | N-ter | K181 | - | DII |
| PIH1D2 | RuvBL1 | N-ter | K281 | - | DII |
| PIH1D2 | RuvBL1 | N-ter | K287 | - | DII |
| PIH1D2 | RuvBL1 | N-ter | K294 | - | DII |
| PIH1D2 | RuvBL1 | N-ter | K389 | - | DIII |
| PIH1D2 | RuvBL1 | N-ter | K435 | - | DIII |
| PIH1D2 | RuvBL1 | K42 | K15 | PIH | - |
| PIH1D2 | RuvBL1 | K42 | K20 | PIH | DI |
| PIH1D2 | RuvBL1 | K42 | K287 | PIH | DII |
| PIH1D2 | RuvBL1 | K42 | K435 | PIH | DIII |
| PIH1D2 | RuvBL1 | K63 | K178 | PIH | DII |
| PIH1D2 | RuvBL1 | K63 | K287 | PIH | DII |
| PIH1D2 | RuvBL1 | K63 | K435 | PIH | DIII |
| PIH1D2 | RuvBL1 | K65 | K20 | PIH | DI |
| PIH1D2 | RuvBL1 | K65 | K178 | PIH | DII |
| PIH1D2 | RuvBL1 | K65 | K281 | PIH | DII |

Références bibliographiques

1. Kavan, D.; Man, P., MStools — Web based application for visualization and presentation of HXMS data. *Int. J. Mass Spectrom.* **2011**, *302*, 53-58.
2. Piersimoni, L.; Kastritis, P. L.; Arlt, C.; Sinz, A., Cross-Linking Mass Spectrometry for Investigating Protein Conformations and Protein-Protein Interactions horizontal line A Method for All Seasons. *Chem Rev* **2022**, *122* (8), 7500-7531.
3. Ben-Nissan, G.; Belov, M. E.; Morgenstern, D.; Levin, Y.; Dym, O.; Arkind, G.; Lipson, C.; Makarov, A. A.; Sharon, M., Triple-Stage Mass Spectrometry Unravels the Heterogeneity of an Endogenous Protein Complex. *Anal Chem* **2017**, *89* (8), 4708-4715.

Développements méthodologiques et combinaisons d'approches de spectrométrie de masse structurale pour la caractérisation de complexes multi-protéiques

Résumé

Ce travail de thèse porte sur des développements et optimisations de méthodes de spectrométrie de masse (MS) structurale, le pontage chimique (XL) et l'échange H/D (HDX) suivi par MS, pour la caractérisation de différents systèmes protéiques. L'apport de la mobilité ionique en HDX-MS a notamment pu être déterminé et un protocole de XL-MS avec des agents pontants MS clivables a été mis en place au laboratoire. Les stratégies bio-informatiques du traitement des données des deux techniques ont été évaluées. Ces développements ont ensuite été appliqués à la caractérisation structurale de différents complexes protéiques impliquant des récepteurs nucléaires, des protéines membranaires, des protéines arginines méthyltransférase et enfin des assemblages impliquant des protéines AAA+. Des approches de MS native et de photométrie de masse ont également été employées dans le but de compléter les informations obtenues en HDX-MS et XL-MS.

Mots-clés : spectrométrie de masse structurale, échange hydrogène/deuterium, pontage chimique, spectrométrie de masse native, photométrie de masse, systèmes protéiques complexes, stratégies bio-informatiques

Résumé en anglais

This PhD thesis focuses on the development and optimisation of structural mass spectrometry (MS) methods for the characterisation of different protein systems. Two structural MS approaches were developed with chemical cross-linking (XL) and H/D exchange (HDX) followed by MS. In particular, the contribution of ion mobility in HDX-MS could be determined and an XL-MS protocol with cleavable MS cross-linker was set up in the laboratory. The bioinformatics strategies for data processing of both techniques were evaluated. These developments were then applied to the structural characterisation of different protein complexes with nuclear receptors, membrane proteins, arginine methyltransferase proteins and finally assemblies involving AAA+ proteins. Native MS and mass photometry approaches were also used to complement the information obtained in HDX-MS and XL-MS.

Keywords: structural mass spectrometry, hydrogen/deuterium exchange, cross-linking, native mass spectrometry, mass photometry, complex protein systems, bioinformatic strategies