



UNIVERSITÉ PARIS-SUD

ÉCOLE DOCTORALE 425 INNOVATION THÉRAPEUTIQUE: *DU FONDAMENTAL À L'APPLIQUÉ*

INSTITUT DE CHIMIE DES SUBSTANCES NATURELLES

CNRS, UPR2301, GIF-SUR-YVETTE

Thèse de doctorat de l'Université Paris-Sud

Présentée par :

Safa LASSOUED

Pour obtenir le titre de :

Docteur en Biochimie structurale

Étude par RMN de la structure et des interactions de protéines du
Virus Respiratoire Syncytial humain

Soutenance le 2 novembre 2015, devant le jury composé de :

Pr Herman VAN TILBEURGH :

Président

Dr Olivier LEQUIN :

Rapporteur

Dr Nadia IZADI-PRUNEYRE :

Rapporteur

Dr Bernard DELMAS :

Examineur

Dr Rémy COUDERC :

Examineur

Dr Christina SIZUN :

Directrice de thèse

Numéro national d'enregistrement :

REMERCIEMENTS

A l'issue de la rédaction de cette recherche, je suis convaincue que la thèse est loin d'être un travail solitaire. En effet, je n'aurais jamais pu réaliser ce travail doctoral sans le soutien d'un grand nombre de personnes.

En premier lieu, je tiens à remercier le Dr Eric GUITTET pour m'avoir accueillie au sein de son équipe pendant ces trois années.

Je tiens surtout à remercier ma directrice de thèse, le Dr Christina SIZUN, pour la confiance qu'elle m'a accordée en acceptant d'encadrer ce travail doctoral, pour ses multiples conseils et pour toutes les heures qu'elle a consacrées à diriger cette recherche. J'aimerais également lui dire à quel point j'ai apprécié sa grande disponibilité, j'ai été extrêmement sensible à ses qualités humaines d'écoute, de compréhension et de son soutien tout au long de ce travail.

Mes remerciements les plus sincères à tous les membres de l'équipe 64, ces années ont été extrêmement enrichissantes pour moi et j'ai beaucoup appris à vos côtés. Et votre soutien dans les moments difficiles m'a encouragé et m'a permis de ne jamais perdre espoir.

Je remercie le Dr Olivier LEQUIN et le Dr Nadia IZADI-PRUNEYRE d'avoir accepté de juger mon travail en qualité de rapporteurs. J'adresse aussi mes remerciements au Dr Bernard DELMAS, Dr Rémy COUDERC et au Pr Herman VAN TILBEURGH qui m'ont fait l'honneur d'examiner ce travail et de participer au jury de soutenance.

J'ai eu un grand plaisir à collaborer avec le Dr Jean-François ÉLEOUËT et le Dr Marie GALLOUX. Je les remercie pour cette collaboration et de leur implication dans l'avancement du projet.

Je n'oublie pas les expériences de mesure de la thermostabilité des protéines par fluorescence qui ont été réalisées par les Dr Eric Jacquet et Naima Nhiri à la plateforme de criblage de thermostabilité des protéines par fluorescence sur le campus de Gif-sur-Yvette. Je les remercie tous les deux et plus particulièrement, le Dr Eric Jacquet pour sa bonne humeur et son soutien. Je vous suis profondément reconnaissante pour ce que vous avez fait pour moi. Une chose est sûre : je n'oublierai jamais.

Je voudrais remercier plus particulièrement les personnes que j'ai rencontré au cours de ces années passées à l'ICSN, avec qui j'étais plus proche et qui sont devenues maintenant des amis pour moi : Prishila, Oriane, Célia, Fataneh.

Un grand MERCI au Pr Fathi MOUSSA, qui m'a fait confiance et m'a permis de venir en France pour continuer mes études.

Un énorme MERCI pour des personnes tout simplement géniales ! Qu'aurais-je fais si vous n'étiez pas là? Assurément pas grand chose ! Merci encore mes amis, vous êtes des anges ! Pierre, Selma, Mouna, Diana, Nora, Khalil... Je ne peux pas citer tout le monde car j'ai la chance d'être entourée par plein de gens formidables.

Je ne saurais assez remercier mes parents, Je voulais vous remercier, d'avoir fait de moi ce que je suis aujourd'hui et des valeurs que vous avez si bien su m'inculquer, la gentillesse, le respect et le dévouement. Je suis si fière d'être votre fille. Vous avez toujours été là, à mon écoute, ouverts et compréhensifs, tolérants, tendres et aimants... même pendant mes "dérapages" de jeunesse. Alors merci... pour tout ce que vous m'avez donné et me donnerez encore.

Et enfin, je remercie mon frère Oussama, ma belle soeur Vanessa, ma soeur Jihene. Ils m'ont donné beaucoup d'amour et ils étaient toujours là pour moi. Et mes deux petites fleurs d'amour Yasmine et Camélia qui sont venues au monde pendant ma thèse et m'ont apporté beaucoup de bonheur, surtout une nouvelle responsabilité : c'est celle d'être TaTa.

Abstract

Respiratory syncytial virus (RSV) is a member of the Paramyxoviridae family of non segmented, negative sense single stranded RNA viruses. Human RSV (hRSV) is a major cause of respiratory diseases in children, and is a prioritized vaccine target. Our aim is to get answers about the structure and dynamics of different components of the RNA dependent RNA polymerase complex (RdRp) of RSV, in particular about their interactions, by using Nuclear Magnetic Resonance (NMR), as a prerequisite to rational drug design. My work focused on two RSV proteins: the phosphoprotein, which is the main polymerase co-factor and necessary for both viral transcription and replication, and M2-1, which functions as an essential transcriptional cofactor of the viral RdRp complex by increasing polymerase processivity. My first aim was to characterize the M2-1 protein of hRSV. M2-1 is a modular RNA binding protein that also interacts with the viral phosphoprotein P. These binding properties are related to the core region of M2-1. Knowing the structure of the corresponding domain, I determined the binding surfaces of RNA and P on the monomeric M2-1[58-177] fragment by NMR. I showed that they partially overlap, which accounts for the previously observed competitive behavior of RNA versus P in M2-1 binding. The NMR results were confirmed by functional tests by a collaboration at INRA, Jouy-en-Josas. However the low binding affinity of M2-1[58-177] for RNA, as compared to that reported by others for full-length M2-1, led me to investigate the possibility to work with full-length M2-1 using NMR. I investigated the structure of M2-1 by starting with its N-terminal zinc-binding domain and the tetramerization domain. I also studied the impact of deletions of these domains by NMR. My second aim was to investigate the structure of hRSV P, which is also a modular protein. I designed experiments using a paramagnetic tag to find out long range contacts inside P. Finally by using deletion mutants of P, I could determine the binding region of M2-1 in P. In a last part, I characterized the stability of RSV non structural protein NS1, used by RSV to escape the host interferon response, for further NMR characterization.

Keywords : respiratory syncytial virus, polymerase activity, transcription, phosphoprotein, M2-1, Nuclear Magnetic Resonance.

Résumé

Le virus respiratoire syncytial (VRS) est un membre de la famille des Paramyxoviridae, des virus à ARN simple brin non segmenté de polarité négative. Le VRS humain (VRSh) est la principale cause des maladies respiratoires chez les enfants, et il est la priorité des cibles vaccinales. Notre objectif est d'obtenir des réponses sur la structure et la dynamique des différents composants du complexe ARN polymérase ARN dépendant du VRS (RdRp) et sur leurs interactions, en utilisant la résonance magnétique nucléaire (RMN), en espérant pouvoir proposer des molécules qui inhibent ou perturbent ces interactions afin de développer des médicaments antiviraux spécifiques. Mon travail a porté sur deux protéines du complexe RdRp: la phosphoprotéine P, qui est le co-facteur principal de la polymérase et est nécessaire à la fois pour la transcription virale et pour la réplication, et M2-1, qui est un cofacteur du complexe RdRp essentiel pour la transcription virale en augmentant la processivité de la polymérase. Mon premier objectif était de caractériser la protéine M2-1 du VRSh. M2-1 est une protéine modulaire qui lie l'ARN et interagit également avec la phosphoprotéine virale P. Ces propriétés de liaison sont liées au domaine globulaire de M2-1. Connaissant la structure du domaine globulaire de M2-1, j'ai déterminé les surfaces de liaison de l'ARN et de P sur le fragment monomère de M2-1[58-177] par RMN. J'ai montré qu'elles se chevauchent partiellement, expliquant pourquoi cette protéine se lie à la phosphoprotéine P et à l'ARN de manière compétitive. Les résultats de RMN ont été confirmés par des tests fonctionnels au travers d'une collaboration avec l'INRA, Jouy-en-Josas. Cependant la faible affinité de M2-1[58-177] pour l'ARN, comparée à celle rapportée par d'autres pour M2-1 entière, m'a conduite à étudier la possibilité de travailler avec la protéine M2-1 pleine longueur en RMN. J'ai étudié la structure de M2-1 en commençant par son domaine N-terminal de liaison au zinc et le domaine de tétramérisation. J'ai aussi évalué l'impact de la suppression de ces domaines par RMN. Mon deuxième objectif a porté sur la structure de la protéine P du VRSh, qui est aussi une protéine modulaire. J'ai mis au point des expériences avec une sonde paramagnétique pour détecter des contacts à longue distance à l'intérieur de P. Enfin, en utilisant des mutants de délétion de P, j'ai pu déterminer la région de liaison de M2-1 dans P. Dans une dernière partie j'ai caractérisé la stabilité de la protéine non structurale NS1 du VRS, impliquée dans le contournement de la réponse interféron de l'hôte, pour des études par RMN ultérieures.

Mots clés : Virus respiratoire syncytial, activité polymérase, transcription, phosphoprotéine, M2-1, résonance magnétique nucléaire.

SOMMAIRE

Chapitre 1 : Introduction	23
1 Généralités sur le VRS	23
1.1 Historique	23
1.2 Pathologie du VRS humain	24
1.3 Taxonomie	26
2 Organisation du VRS	27
2.1 Virion du VRS	27
2.2 Génome	28
2.3 Cycle viral.....	29
2.3.1 L'entrée du virus : attachement et fusion	30
2.3.1.1 Attachement.....	30
2.3.1.2 Fusion	31
2.3.2 La transcription virale.....	31
2.3.2.1 Composition du complexe de transcription.....	31
2.3.2.2 Signaux en « cis » pour la transcription.....	32
2.3.2.3 Coiffage des ARNm.....	32
2.3.3 La réplication virale.....	34
2.3.4 Corps d'inclusion de protéines du VRS	35
2.3.5 Assemblage et bourgeonnement	36
2.4 Les protéines virales	37
2.4.1 Les protéines non structurales NS1 et NS2.....	37
2.4.2 Les protéines d'enveloppe.....	38
2.4.2.1 La protéine de fusion F.....	38
2.4.2.2 La glycoprotéine G	41
2.4.2.3 La protéine hydrophobe SH	42
2.4.2.4 La protéine de matrice M	43
2.4.3 Les protéines du complexe polymérase	44
2.4.3.1 La nucléoprotéine N	45
2.4.3.2 La polymérase L, sous-unité catalytique portant les activités de transcription et de réplication virale	47
2.4.3.3 La phosphoprotéine P cofacteur principal de la polymérase du VRS.....	48
2.4.3.4 Les protéines M2-1 et M2-2	49
3 Epidémiologie et traitements de l'infection par le VRSH	50
3.1 Epidémiologie.....	50

3.2	Traitements de l'infection par le VRSH	51
3.2.1	Les médicaments antiviraux	51
3.2.1.1	La ribavirine	51
3.2.1.2	Autres molécules antivirales	52
3.2.2	Immunoprophylaxie avec des anticorps	52
3.2.3	Les vaccins	54
3.2.3.1	Les vaccins sous-unitaires	55
3.2.3.2	Les vaccins obtenus par génétique inverse	56
3.2.3.3	Les vaccins vectorisés	56
3.2.3.4	Vaccins pour la forme bovine du VRS	57
4	L'objectif de ma thèse	58
Chapitre 2 : Etude structurale par RMN de la protéine M2-1 du VRSH 61		
1	Objectifs	61
2	Résultats antérieurs à ma thèse	62
2.1	Organisation en domaines de M2-1	62
2.2	Structure du domaine globulaire de la protéine M2-1	63
3	Analyse du N-terminal de la protéine M2-1 du VRS	65
3.1	Constructions de M2-1 comportant le domaine N-terminal	65
3.2	Expression et purification de fragments N-terminaux de M2-1	66
3.3	Analyse par RMN des fragments N-terminaux de M2-1	69
3.3.1	Spectres préliminaires HSQC ¹ H- ¹⁵ N	69
3.3.2	Influence de la température	70
3.4	Analyse par RMN d'un peptide synthétique contenant le motif "3Cys-1His" de M2-1	71
3.4.1	Analyse de la séquence du domaine N-terminal de M2-1	71
3.4.2	Essais de reconstitution du complexe ZNF29:Zn	72
3.4.2.1	Conditions de tampon pour reconstituer le complexe ZNF29:Zn	72
3.4.2.2	Reconstitution du complexe ZNF29:Zn par titrage du pH	73
3.4.2.3	Complexation du cadmium	74
3.4.3	Analyse par RMN homonucléaire du peptide ZNF29	74
4	Analyse par RMN de la protéine M2-1 dépourvue du domaine d'oligomérisation	77
4.1	Expression et purification de M2-1Δ[35-58]	77
4.2	Optimisation du tampon de M2-1Δ[35-58]	78

4.3 Attribution des fréquences de résonance du squelette peptidique de la construction M2-1Δ[35-58]	80
4.3.1 Principe de l'attribution séquentielle des fréquences de résonance	80
4.3.2 Attribution du squelette peptidique de M2-1Δ[35-58].....	81
5 Analyse par RMN d'une construction de M2-1 sans domaine	
N-terminal	84
5.1 Production du fragment M2-1Δ[1-31]	84
5.2 Analyse par RMN du fragment M2-1Δ[1-31]	85
6 Comparaison avec des résultats cristallographiques	87
Chapitre 3 : Etude des interactions de la protéine M2-1	
du VRSH par RMN	90
1 Introduction	90
2 Caractérisation des interactions de protéines par RMN	90
2.1 Régimes d'échange en RMN	90
2.2 Cartographie des variations de déplacement chimique	91
2.3 Détermination d'une constante de dissociation	92
2.4 Intensités des signaux en RMN	92
2.5 Expériences de saturation croisée transférée.....	93
3 Préparation des échantillons de protéine et d'ARN	94
3.1 Expression et purification du domaine central de M2-1.....	94
3.2 Synthèse et la purification des ARN	95
4 Etude des interactions du domaine central de M2-1 avec des ARN	96
4.1 Etude préliminaire avec un ARNt de levure	96
4.2 Liaison de séquences d'ARN courtes correspondant à des signaux de transcription du VRS.....	98
4.2.1 Choix de courtes séquences d'acides nucléiques pour des expériences d'interaction par RMN	98
4.2.2 Surface d'interaction sur M2-1[58-177]	101
4.2.3 Affinités comparées de différents ARN	103
4.2.3.1 Détermination de la constante de dissociation	103
4.2.3.2 Comparaison des affinités déterminées avec différents acides nucléiques.....	104
4.3 Discussion	106

5 Etude des interactions du domaine central de M2-1 avec la phosphoprotéine du VRS	109
5.1 Résultats antérieurs en titration calorimétrique isotherme	109
5.2 Spectres HSQC en présence de la protéine P du VRS.....	110
5.3 Spectres HSQC en présence du fragment P100-160	111
5.4 Confirmation de la surface d'interaction de P sur M2-1 par saturation croisée transférée	112
6 Validation des résultats RMN par des tests fonctionnels	114
6.1 Mesure de la transcription in vitro par la polymérase du VRS.....	114
6.2 Discrimination des résidus critiques pour l'interaction M2-1:P et M2-1:ARN	115
6.3 Analyse par immunofluorescence de la localisation de mutants de M2-1	115
6.4 Conclusion.....	116
Chapitre 4 : Etude structurale par RMN de la phosphoprotéine du VRS....	120
1 Introduction.....	120
1.1 La protéine P, plaque tournante des interactions dans le complexe polymérase du VRS.....	120
1.2 Structure de P	120
1.3 Eléments de réponse sur la structure de P par RMN.....	122
1.3.1 Utilisation de fragments de P	122
1.3.2 Perte de signal dans les régions structurées	124
1.3.3 Analyse de la structure secondaire des régions désordonnées de P	125
2 Recherche de contacts à longue distance dans P	126
2.1 Préparation d'échantillons avec une sonde paramagnétique	126
2.1.1 Préparation de mutants cystéines.....	126
2.1.2 Ajout d'une étiquette paramagnétique iodoacétamido-PROXYL	127
2.2 Analyse des spectres des protéines avec une sonde paramagnétique.....	128
2.2.1 Vérification par RMN des mutants cystéines	128
2.2.1.1 Oxydation partielle des cystéines.....	129
2.2.1.2 Disparition de signaux dans la partie N-terminale de P	129
2.2.2 Analyse par RMN des protéines avec étiquette paramagnétique.....	131
3 Etude de l'interaction de la phosphoprotéine avec le domaine globulaire	

de M2-1 par RMN.....	134
3.1 Expériences d'interaction par RMN	134
4 Conclusion	137
Chapitre 5 : Les protéines NS1 et NS2 du VRS.....	138
1 Introduction et objectifs.....	138
2 Rôles et fonctions des protéines non structurales du VRS	139
2.1 Les voies interférons	139
2.2 Rôle de la protéine NS1	140
2.3 Rôle de la protéine NS2	141
2.4 Oligomérisation de NS1 et NS2	142
3 Production et purification de la protéine NS1	143
3.1 Plasmides	143
3.2 Expression de MBP-NS1	143
3.3 Purification de MBP-NS1 par affinité sur résine amylose	144
3.4 Etape de polissage « polishing » par filtration sur gel	145
3.5 Difficultés rencontrées pour la production de NS1	147
4 Analyse par RMN de la protéine NS1-¹⁵N	147
5 Mesure de la thermostabilité de NS1 par fluorescence	149
5.1 Principe.....	149
5.2 Application du Thermofluor à NS1	150
4.1.1.1 Influence de la force ionique sur la thermostabilité de NS1.....	153
4.1.1.2 Influence du pH sur la thermostabilité de NS1.....	153
6 Conclusion	156
Chapitre 6 : Conclusion et perspectives.....	158
Bibliographie.....	162

TABLE DES ILLUSTRATIONS

- Figure 1 : Radiographies pulmonaires d'un enfant infecté par le virus respiratoire syncytial démontrant la formation de syncytium et la mise en évidence d'une bronchiolite virale aiguë.** 23
- Figure 2 : Arbre phylogénétique des membres de la famille des Paramyxoviridae. Il y a deux sous-familles : les Pneumovirinae, qui comporte les pneumovirus (VRS humain (hRSV), VRS bovin (bRSV), VRS ovin (bRSV), pneumovirus de la souris (PVM pour Pneumonia Virus of Mice)) et les métapneumovirus (métapneumovirus respiratoire humain (hMPV), pneumovirus aviaires (APV)), et les Paramyxovirinae, qui comprend les respirovirus (virus Parainfluenza humain 1 et 3 (hPIV1 et hPIV3), virus de Sendai (SeV)), les henipah (Hendra virus (Hev), Nipah virus (NiV)), les morbillivirus (canine distemper virus (CDV), morbillivirus de phoque (PDV), morbillivirus du dauphin (DMV), virus de la rougeole (MeV), Pinderpest virus (RPV)), les rubulavirus (virus Parainfluenza humain 2 et 4 (hPIV2 et hPIV4), Mumps virus (MuV), Simian virus 5 (SV5), Simian virus 41 (SV41)), le virus de la maladie de Newcastle et les tupaia paramyxovirus (TPMV).** 26
- Figure 3 : Micrographies électroniques à coloration négative, montrant les caractères morphologiques de VRS sous forme arrondie (Gauche) sous forme filamenteuse (Droite).** 27
- Figure 4 : Carte du génome de sens négatif du VRS. (Souche A2) à l'échelle approximative. Les numéros en dessous du schéma indiquent la longueur en nombre de nucléotides (nt): le 3_ leader (le) et le 5_ trailer (tr) et les régions intergéniques sont soulignées et la longueur de chevauchement des gènes est entre parenthèses. Les numéros en gras au dessus indiquent les longueurs des protéines non modifiées en acide aminé (aa). Les protéines virales sont les suivantes : G, glycoprotéine d'attachement; F, glycoprotéine de fusion; SH, petite glycoprotéine hydrophobe; M, protéine de matrice; N, nucléoprotéine; P, phosphoprotéine; L, la grande unité polymérase; M2-1, produit du premier ORF de l'ARNm M2; M2-2; produit du deuxième ORF de l'ARNm M2; NS1, protéine non structurale 1; NS2, protéine non structurale 2. Les cadres de lecture ouverts de M2-1 et M2-2 se chevauchent sur le gène. (Collins and Melero 2011)** 29
- Figure 5 : Représentation schématique du cycle viral du virus respiratoire syncytial . (A) Le VRS se fixe sur la membrane cellulaire via la protéine G, (B) Cette étape est suivie de la fusion médiée par la protéine F, ce qui conduit à l'incorporation de l'enveloppe virale dans la membrane cellulaire et la libération subséquente de la nucléocapside dans le cytosol, (C) pour permettre à l'ARN viral d'entrer dans la cellule. (D) Transcription de l'ARNm, (E) Réplication de l'ARN, (F) les protéines M, F et G forment des complexes avec la membrane plasmique, alors que les protéines N, P, L et M2-1 forment des inclusions dans le cytoplasme. (G) Grâce à une série d'interaction, la protéine M coordonne l'assemblage des protéines de l'enveloppe avec les protéines de la nucléocapside. (H) Bourgeonnement du nouveau virion mature prêt à infecter une cellule voisine (Bawage et al. 2013).** 30

- Figure 6 : Schéma de la composition du complexe de transcription virale chez le VRS, avec les cofacteurs essentiels de la polymérase virale (orange) (unité catalytique L (orange), et la phosphoprotéine (bleu) et la protéine M2-1 (rose))** 31
- Figure 7 : Schéma du gradient décroissant d'expression des gènes (ARNm) en fonction de leur éloignement par rapport à l'extrémité 3'.** 33
- Figure 8 : Schéma de la composition du complexe de réplication virale chez le VRS , avec les cofacteurs essentiels de la polymérase virale (Orange) (l'unité catalytique L en orange, la phosphoprotéine (bleu) et la protéine M2-2 (rouge)).** 35
- Figure 9 : Observation de corps d'inclusion après une infection des cellules BSR-T7 avec le VRSH. Marquage indirect 24 h post-infection avec un anticorps monoclonal de souris anti-P (rouge) et un anticorps polyclonal de lapin anti-M2-1 et anti-N (vert) (Blondot et al. 2012)** 35
- Figure 10 : Schéma d'une particule virale du VRS. Modifié à partir de http://education.expasy.org/images/Metapneumovirus_virion.jpg.** 37
- Figure 11 : Structure cristallographique de l'ectodomaine de la protéine F du VRSH sous la forme postfusion avec une résolution de 3.2 Å (A) Représentation schématique de la séquence primaire de l'ectodomaine de la protéine F. Le numéro des résidus délimitant chaque segment est indiqué ainsi que l'endroit des sites de clivage aux furines. DI-III, domaines I-III ; p27, peptide excisé ; FP, peptide fusion ; HRA,-B,-C, heptapeptides répétés A, B et C. (B) et (C) Représentation en ruban d'un monomère et d'un trimère de F sous la forme post-fusion. Adapté de (Swanson et al. 2011)** 39
- Figure 12 : Schéma des régions dynamiques de la protéine F qui se replie pour initier la fusion. Les trois hélices α en N-terminal dans le domaine pré-HRA sont reliées par deux peptides non hélicoïdaux sous une forme métastable. Lors du déclenchement, ces deux peptides se replient en hélice α , rallongeant ainsi les hélices α HRA permettant l'insertion du peptide de fusion F1 en N-terminal dans la membrane de la cellule cible. De même, le trimère de HRA (gris foncé) se plie alors en deux, et les hélices α HRB (gris clair) s'insèrent dans les sillons à la surface du trimère HRA formant un faisceau stable de 6 hélices (6HB). Par conséquent, les virions et les membranes cellulaires sont rassemblés et la fusion membranaire commence. La région de «tête» de la protéine F ne se réorganise pas pendant les événements de déclenchement et de repliement et, par conséquent, elle n'est pas représentée sur ce dessin. (Costello et al. 2012)** 40
- Figure 13 : Modèle de la protéine SH du VRS, basé sur la structure de la protéine SH fonctionnelle observée par microscopie électronique et des données fonctionnels existants (Carter et al. 2010). Le modèle moléculaire de la protéine SH hexamérique vue de dessus ou en bas insérée dans une membrane. La flèche indique le canal formé.** 43
- Figure 14 : Structure cristallographique de la protéine M du VRSH formant un dimère avec une résolution de 1,7 Å. Les deux protomères sont en rose et en vert (PDB : 4D4T) et la similarité de structure avec les protéines M des métagneumovirus humain (MPVH) (PDB : 4LP7) et de celle du virus de la maladie de Newcastle (NDV) (PDB : 4GIG) (Forster et al. 2015).** 44

- Figure 15 : Gauche : Structure radiocristallographique d'un anneau contenant 10 protomères de nucléoprotéine du VRSH avec une résolution de 3,29 Å, chaque sous-unité de protéine N étant représentée avec une couleur différente. Droite : le NTD est en jaune, le CTD en rouge et les bras N- et C-terminaux en bleu. L'ARN est en cyan et les bases en contact avec un monomère de protéine N sont en noir. PDB : 2WJ8 (Tawar et al. 2009) 45**
- Figure 16 : Nombre hebdomadaire de recours aux urgences hospitalières pour bronchiolite des enfants de moins de 2 ans dans les 354 établissements hospitaliers participants à la surveillance entre 2012 et 2015, Données de l'institut de veille sanitaire (InVS). 50**
- Figure 17 : Organisation en domaines de la protéine M2-1. Prédiction de la structure secondaire de la protéine M2-1 du VRSH avec PSI-PRED (<http://bioinf.cs.ucl.ac.uk/psipred/>). 62**
- Figure 18: (A) Comparaison du domaine globulaire de M2-1 avec le domaine C-terminal de la protéine VP30 du virus Ebola (Hartlieb et al. 2007). La couleur varie du bleu vers le rouge du N- vers le C-terminal. Pour M2-1[58-177], seul le modèle RMN de plus basse énergie est donné. Les extrémités N- et C-terminales désordonnées ne sont pas représentées. L'alignement structural entre M2-1 et VP30 a été généré sur le serveur Dali (Holm and Rosenstrom 2010). L'hélice $\alpha 7$ du domaine C-terminal de la VP30 n'est pas représentée puisqu'elle n'a pas d'équivalent dans M2-1. Les résidus hydrophobes identiques sur les hélices $\alpha 5$ et $\alpha 6$ sont représentés en bâtons. 64**
- Figure 19 : Alignement des séquences des fragments de la protéine M2-1. Les couleurs correspondent à la polarité des chaînes latérales des acides aminés : acide (rouge), basique (bleu), polaire (vert), apolaire (noir). Les étoiles bleues indiquent le motif 3Cys-1His. M2-1Delta correspond à une forme sans le domaine d'oligomérisation M2-1 Δ [35-58]. 65**
- Figure 20 : Gel d'expression SDS-PAGE des trois constructions de la partie N-terminale de la protéine M2-1 avant et après induction par l'IPTG. 66**
- Figure 21 : Gel SDS-PAGE après coupure à la thrombine des trois constructions de la partie N-terminale de la protéine M2-1 (MP = marqueur de poids moléculaire, R = résine glutathion avec GST après coupure à la thrombine, FT = fraction non retenue après coupure à la thrombine contenant les protéines M2-1[1-32], M2-1[1-59] et M2-1[1-65]). 67**
- Figure 22 : Gel de protéines SDS-PAGE des 3 fragments de la protéine M2-1 (M2-1[1-32], M2-1[1-59] et M2-1[1-65]) éluées avec trois tampons différents donnés dans le Tableau 2 (MP = marqueur de poids moléculaire). 68**
- Figure 23 : Profils d'absorption en spectrométrie UV-visible des trois constructions de la partie N-terminale de la protéine M2-1. Les différentes préparations (1 et 2) correspondent à deux lavages successifs avec le même tampon (Tampon 1, Tableau 2). 68**
- Figure 24 : Spectres HSQC ^1H - ^{15}N acquis à 14,1 T des fragments M2-1[1-32] (tampon Tris pH 6.5, NaCl 200 mM, ZnCl_2 20 μM , température 293 K) et M2-1[1-59] (tampon MES pH 6.8, NaCl 150 mM, ZnCl_2 20 μM , TCEP 0.5 mM, température 293 K). 69**
- Figure 25 : Spectres HSQC ^1H - ^{15}N du fragment M2-1[1-59] acquis à 14,1 T et à température variable comprise entre 288 K et 303 K (tampon MES pH 6.8, NaCl 150 mM, ZnCl_2 20 μM , TCEP 0.5 mM) 70**

- Figure 26 : Alignement des séquences du peptide synthétique ZNF29 correspondant à la séquence N-terminale de M2-1 avec les séquences des doigts de zinc en tandem (TZF) des protéines TIS11d et Nup475. 71
- Figure 27 : Modèle par homologie du domaine N-terminal de M2-1 (vert) superposé à la structure RMN de la protéine TIS11d en complexe avec un ARNm. Les trois cystéines et l'histidine de TIS11d sont représentées en bâtons rouges et magenta, respectivement. 71
- Figure 28 : Diagramme potentiel-pH du zinc pour une concentration d'espèces en solution 10^{-3} mol/l. 72
- Figure 29 : Zoom dans la région des protons amides et aromatiques de spectres RMN ^1H du peptide synthétique ZNF29 à 335 μM à deux pH différents (pH 3,6 en violet et pH 7,8 en bleu) dans l'eau et en présence de 2 mM DTT-d10 acquis à 14,1 T et 298 K. 73
- Figure 30 : Diagramme potentiel-pH du cadmium tracé pour une concentration d'espèces en solution 10^{-2} mol/l 74
- Figure 31 : Détermination de l'enchaînement séquentiel dans un peptide en utilisant des spectres TOCSY et NOESY : les flèches bleues montrent les corrélations intra-résiduelles observées en TOCSY et les flèches rouges les corrélations inter-résiduelles observées en NOESY pour un peptide non structuré. 75
- Figure 32 : Zoom dans la région des corrélations HN-H α dans les spectres TOCSY (vert, spin lock 60 ms) et NOESY (rose, temps de mélange 120 ms) ^1H - ^1H du peptide synthétique ZNF29 (concentration 335 μM , pH 3,6, 2mM DTT), acquis à 14,1 T et 298 K. 75
- Figure 33 : Extraits des spectres TOCSY (vert, spin lock 60 ms) et NOESY (rose, temps de mélange 120 ms) ^1H - ^1H du peptide synthétique ZNF29 (concentration 335 μM , pH 3,6, 2mM DTT), acquis à 14,1 T et 298 K montrant les corrélations des protons amide HN vers les protons aliphatiques et les corrélations entre protons aliphatiques. 76
- Figure 34 : Gel SDS-PAGE après coupure à la thrombine de la construction M2-1 Δ [35-58] (MP = marqueur de poids moléculaire, R = résine glutathion avec GST après coupure à la thrombine, FT1, 2, 3 = fractions non retenues après coupure à la thrombine contenant la protéines M2-1 Δ [35-58]) 77
- Figure 35 : Spectre d'absorption UV-VIS de deux fractions obtenues après coupure à la thrombine pour M2-1 Δ [35-58]. 78
- Figure 36 : Spectre HSQC ^1H - ^{15}N de la protéine M2-1 Δ [35-58] (en rouge) à 14,1 T et 303K à une concentration de 90 μM dans un tampon 20 mM MES pH 6,2, 150 mM NaCl, 0.5 M arginine, 20 μM ZnCl $_2$, 2 mM β -mercaptoéthanol. Le spectre de M2-1[58-177] dans les mêmes conditions (noir, avec attribution des signaux amides). 79
- Figure 37 : Transferts d'aimantation dans les expériences tridimensionnelles triple-noyau utilisées en RMN des protéines. Les flèches représentent les transferts d'aimantation partant du proton amide vers les noyaux édités (en rose) via les couplages scalaires. Les noyaux en bleu sont utilisés comme relais de l'aimantation, mais ils ne sont pas édités. (<http://www.protein-nmr.org.uk/>) 81

- Figure 38 : (A) Variations de déplacements chimiques combinés ^1H et ^{15}N des amides du domaine globulaire de M2-1 entre M2-1 Δ [35-58] et M2-1[58-177]. Le graphe est coupé et la valeur pour Lys177 (0.66 ppm) n'est pas reportée. (B) Les résidus avec les variations les plus importantes ont été cartographiés en rouge sur la structure de M2-1[58-177], la partie N-terminale (E58-N88) en jaune et la petite hélice prédite pour les résidus E70 à A73 en bleu.** 82
- Figure 39 : Spectre HSQC ^1H - ^{15}N de la protéine M2-1 Δ [35-58] à 14,1 T et 303K (90 μM dans 20 mM MES pH 6,2, 150 mM NaCl, 0.5 M arginine, 20 μM ZnCl_2 , 2 mM β -mercaptoéthanol). L'attribution de la partie C-terminale désordonnée est en rouge. Le seuil des contours est élevé et les signaux de la partie globulaire apparaissent très faibles. Seuls quelques signaux du domaine globulaire sont annotés, mais l'attribution de ce domaine est complète.** 83
- Figure 40 : Comparaison des séquences en acides aminés des constructions de M2-1 correspondant à des délétions de domaines : M2-1 Δ [1-31], M2-1 entier et M2-1[58-177].** 84
- Figure 41 : Gel SDS-PAGE du suivi de culture de M2-1 Δ [1-31]. (MP = marqueur de poids moléculaire, NI et I = extraits bactériens avant et après induction à l'IPTG).** 85
- Figure 42 : Spectres HSQC ^1H - ^{15}N superposés des fragments M2-1 Δ [1-31] (en vert ; 100 μM dans un tampon 20 mM MES pH 6,5, 150 mM NaCl, 2 mM β -mercaptoéthanol) M2-1[58-177] (noir) acquis dans les mêmes conditions (14,1 T et 293 K). Les attributions des signaux amides sont données pour M2-1 Δ [1-31] en vert et M2-1[58-177] en noir.** 85
- Figure 43 : (A) Variations de déplacements chimiques combinés ^1H et ^{15}N des amides du domaine globulaire de M2-1 entre M2-1 Δ [1-31] et M2-1[58-177]. Le graphe est coupé et la valeur pour Lys177 (0.68 ppm) n'est pas reportée. (B) Les résidus avec les variations les plus importantes ont été cartographiés en magenta sur la structure de M2-1[58-177], la partie N-terminale (E58- N98) en orange.** 86
- Figure 44 : Structure cristallographique de la protéine M2-1 du VRSH avec une résolution de 2,52 Å (PDB 4C3D) (Tanner et al. 2014). (A) Model ruban du tétramère de la protéine M2-1 du VRSH, avec zoom sur la partie N-terminale. Chaque protomère a une couleur différente et l'atome de zinc est représenté par une sphère grise. (B) Protomère de M2-1 où les régions sont colorées par domaine : domaine de liaison au zinc (orange); domaine de tétramérisation en hélices (magenta); domaine globulaire (bleu); domaine partiellement structuré (gris); résidus appartenant au vecteur de clonage (noir). Les extrémités N- et C-terminales sont marquées avec des sphères bleues et rouges, respectivement.** 87
- Figure 45 : Alignement de la structure cristallographique d'un protomère de la protéine M2-1 (Tanner et al. 2014) avec la structure RMN du domaine globulaire M2-1[58-177] . Les hélices sont numérotées selon les désignations faites pour la protéine M2-1 entière.** 88
- Figure 46: Principe de l'étude des interactions entre deux protéines par saturation croisée transférée (TCS) et marquages isotopiques.** 93
- Figure 47 : Gel d'expression de polyacrylamide obtenu par électrophorèse en milieu dénaturant (SDS-PAGE) pour M2-1[58-177].MP = marqueur de poids moléculaire. NI = extrait bactérien avant induction par l'IPTG. I = extrait bactérien après induction. La bande correspondant à GST-M2-1[58-177] est indiquée par une flèche. (B) Gel de protéine SDS-PAGE après coupure à**

la thrombine (MP = marqueur de poids moléculaire, R = résine GSH-sépharose, NR = fraction non retenue qui contient la protéine sans la GST).	94
Figure 48 : Cycle de synthèse de l'ARN en phase solide.	95
Figure 49 : Gradient de concentration de NaCl (entre 0 et 1M) et chromatogramme obtenu sur colonne anionique au cours de la purification de l'ARN "leader positif".	96
Figure 50 : Superposition de spectres ^1H - ^{15}N -HSQC de M2-1[58-177] en présence et en absence d'un ARN de levure commercial avec un rapport molaire de 1:1 (50 μM , T = 298 K, B_0 = 14,1 T, 600 MHz). Spectre de référence (rouge), avec l'ARN (bleu), et après digestion par la RNase A (orange). Les pics correspondants aux résonances des chaînes latérales de l'asparagine et de la glutamine sont reliés par des lignes entre les deux protons géminaux de l'amide. Les attributions des résidus avec des variations de déplacement chimique significatives sont indiquées.	97
Figure 51 : Variations de déplacement chimique ^1H et ^{15}N , résidu par résidu pour M2-1[58-177] en présence de l'ARN de levure. Le code couleur correspond à l'amplitude des variations : les valeurs les plus basses sont en rouge, les plus élevées en bleu. La moyenne et l'écart type sont indiqués par des traits pleins et en pointillés.	98
Figure 52 : La terminaison de la transcription est sensible aux jonctions des gènes et à la présence de la protéine M2-1 (Hardy et al. 1999). Les différentes jonctions des gènes sont représentées sur l'axe X. L'axe Y donne le pourcentage de l'abondance molaire des transcrits dicistroniques par rapport à des transcrits monocistroniques, obtenus avec un système de miniréplicon, en présence et en absence de la protéine M2-1. Un transcrit dicistronique reflète l'échec de la terminaison de transcription.	99
Figure 53 : Extrait de la région des protons imino des spectre RMN ^1H des ARN « GE-F positif » (rouge), « Leader positif » (noir) et « GE-F double brin » (vert) à 200 μM montrant l'appariements de paires de bases dans ces ARN (B_0 = 14.1 T, T = 293 K).	100
Figure 54 : Superposition de spectres ^1H - ^{15}N -HSQC de M2-1[58-177] en présence de différentes concentrations de l'ARN « polyA » (50 μM , T = 298K, B_0 = 14,1 T). L'ARN est rajouté en % : 0, 10, 20, 40, 60, 80, 100, 150, 250, 400 et 600%.	101
Figure 55 : Variations de déplacement chimique résidu par résidu induit par l'ARN « polyA » dans le spectres ^1H - ^{15}N -HSQC de M2-1[58-177] avec un rapport molaire ARN/protéine de 6/1. Le code couleur correspond à l'amplitude des variations : les valeurs les plus basses sont en rouge, les plus élevées en bleu. Cartographie des variations de déplacement chimique sur la structure du domaine central de M2-1. Les atomes d'azote des liaisons amides des résidus qui présentent une variation significative sont indiqués par des sphères rouges.	102
Figure 56 : Zoom sur les résonances des chaînes latérales des arginines (C δ -H δ) et des lysines (C ϵ -H ϵ) dans les spectres ^1H - ^{13}C -HSQC de M2-1[58-177] en présence de concentrations croissantes de l'ARN « Leader » (50 μM , T = 293K, B_0 = 24,3 T). Spectre de référence (rouge), en présence de 1 équivalent (vert) et 2 équivalents (bleu)	102
Figure 57 : Récapitulatif des expériences d'interaction par RMN de M2-1[58-177] avec de courtes séquences d'ARN par cartographie des variations de déplacement chimique sur la structure	

du domaine central de M2-1. Les atomes d'azote des liaisons amides des résidus qui présentent une variation significative pour toutes les séquences d'ARN génomique testées (données dans le Tableau 3) sont indiqués par des sphères rouges. Les chaînes latérales de R151 et K92, pour lesquels l'ARN a induit des variations de déplacement chimique $^{13}\text{C}\delta\text{-}^1\text{H}\delta$ et $^{13}\text{C}\epsilon\text{-}^1\text{H}\epsilon$ significatives, sont dans la représentation en bâton. La surface formée par l'ensemble de ces résidus est colorée en rouge. P153, pour laquelle aucune information n'est disponible, s'affiche en noir. 103

Figure 58 : (A) Extrait du spectre HSQC du titrage avec l'ARN polyA pour illustrer les variations de déplacement chimique de quelques résidus (G62, N88, N89, N121 et D155) en présence de l'ARN. (B) Courbes expérimentale et calculée des variations de déplacement chimique combinées ^1H et ^{15}N en fonction du rapport [ARN]/[protéine] dans le cas de N88. La constante de dissociation est $\sim 20 \mu\text{M}$. 104

Figure 59 : Récapitulatif des expériences de titrage de la protéine M2-1[58-177] avec les ARN et détermination de constantes de dissociation (K_d) par ajustement des variations de déplacement chimique en ^1H et ^{15}N du squelette peptidique ou les chaînes latérales de la glutamine (notés H, N, H ϵ 21, H ϵ 22 et H ϵ 2 respectivement), en fonction de la concentration relative de ligand par rapport à la concentration de la protéine, avec un modèle de liaison avec une stœchiométrie 1:1. Seuls les résidus avec des perturbations significatives ont été utilisés. Les mesures ont été effectuées à 293 K sur un spectromètre RMN à 14,1 T. Les spectres HSQC $^1\text{H}\text{-}^{15}\text{N}$ ont été acquis avec des ajouts croissants d'ARN à des solutions de M2-1[58-177] à 50-60 μM . Les courbes correspondent aux séquences d'ARN : (a) UGA2 (épingle à cheveux), (b) Leader, (c) Gene Start NS1, (d) Gene End SH, (e) Gene End F, (f) poly U long. Les K_d sont indiquées pour chaque ARN. 105

Figure 60 : Expérience d'ITC pour mesurer l'interaction entre P et M2-1[58-177] (Blondot et al. 2012). (A) Données brutes : chaleur absorbée à chaque ajout de M2-1[58-177] (2 μl par injection, la concentration de protéine étant de 335 μM) dans une cellule contenant P (200 μl de volume initial; 33 pM concentration initiale). (B) Courbe de titrage obtenue après intégration des données brutes montrant la chaleur absorbée intégrée par mole de ligand ajouté. Les carrés pleins représentent les points expérimentaux. La ligne continue correspond à la courbe calculée. La valeur de K_d moyenne est égale à 3 μM correspondant à une stœchiométrie de 1:1. 109

Figure 61 : Superposition de spectres HSQC $^1\text{H}\text{-}^{15}\text{N}$ de M2-1[58-177] à 50 μM (spectre de référence en rouge) en présence de 0,25 (orange) et 0,5 (violet) équivalent de la phosphoprotéine P du VRS (T = 298 K, $B_0 = 14,1 \text{ T}$). 110

Figure 62 : Superposition de spectres HSQC $^1\text{H}\text{-}^{15}\text{N}$ de M2-1[58-177] en présence de deux équivalents de P[100-166] (T = 298 K, $B_0 = 14,1 \text{ T}$). 111

Figure 63 : Variations d'intensité induites par P[100-166] sur le spectre HSQC $^1\text{H}\text{-}^{15}\text{N}$ de M2-1[58-177], résidu par résidu, dans les conditions de la Figure 62. Le code couleur correspond à l'amplitude des variations : les valeurs les plus basses sont en rouge, les plus élevées en bleu. Ces variations sont cartographiées sur la structure de M2-1[58-177]. Les atomes d'azote des

- liaisons amides des résidus qui présentent une variation d'intensité significative sont colorés en bleu si $\Delta I/I_0 > 75\%$ et en cyan si $\Delta I/I_0 > 70\%$. 112
- Figure 64 : Rapport des intensités résidu par résidu dans les spectre HSQC $^1\text{H}-^{15}\text{N}$ de $150\ \mu\text{M}$ M2-1[58-177] $]-^2\text{H}^{15}\text{N}$ dans 91% D_2O obtenues en irradiant la région des méthyles de P (I_{sat}) ou non (I_0), en présence de 10 % (rouge) ou en absence (noir) de la phosphoprotéine du VRS ($T = 298\text{K}$, $B_0 = 14,1\ \text{T}$, 150). Cartographie de la saturation croisée transférée sur la structure de M2-1[58-177]. Les atomes d'azote des liaisons amides des résidus qui présentent une variation significative sont indiqués par des sphères violettes. La surface formée par l'ensemble de ces résidus est colorée en violet. 113
- Figure 65 : Schéma du minigénome du VRS codant pour la luciférase et mesure de l'activité polymérase in vitro du VRS via la mesure de l'activité luciférase pour différents mutants de M2-1. 114
- Figure 66: Expériences de retard sur gel (Gel tris glycine 1% agarose) avec la protéine GST-M2-1[58-177] ($100\ \mu\text{M}$) et des ARNt de levure ($50\ \mu\text{M}$) (Blondot et al. 2012) 115
- Figure 67: Expériences de GST-pulldown par GST-M2-1 de P et de mutants de P. Les étoiles rouges montrent les mutations qui ont induit une baisse d'interaction de plus que 50% par rapport à la protéine M2-1 sauvage. 115
- Figure 68: Immunofluorescence 24 h après co-transfection en cellules BSR-T7 de P, N et M2-1 sauvage ou mutée. Marquage indirect avec un anticorps polyclonal de lapin anti-P (rouge) et anticorps monoclonal de souris anti-M2-1(vert) (Blondot et al. 2012). 116
- Figure 69: Les surfaces de contact avec P (en bleu) et avec l'ARN (en rouge) sur le domaine globulaire de M2-1 font apparaître une superposition partielle qui pourrait correspondre à une surface de compétition (en orange). 117
- Figure 70 : Schéma récapitulatif des différentes informations connues sur la structure de la phosphoprotéine P du VRS. Un motif "soyuz1" en N-terminal (résidus 8-20, rectangle beige) est conservé chez les Mononégavirales (Karlin and Belshaw 2012). Le domaine d'oligomérisation résistant à la digestion à la trypsine est représenté par des rectangles bleus (Llorente et al. 2006). Un modèle par homologie a été construit à partir du domaine d'oligomérisation de la phosphoprotéine du virus Sendaï. La région déterminée par digestion à la chymotrypsine est en bleu foncé (Llorente et al. 2008). La région C-terminale (résidus 232-241, rectangle vert) interagit avec la nucléoprotéine N de la nucléocapside (Tran et al. 2007). 121
- Figure 71 : Constructions utilisées pour l'étude de la structure de la phosphoprotéine par RMN : P pleine longueur, 2 fragments C-terminaux sans et avec le domaine d'oligomérisation, 2 fragments N-terminaux sans et avec le domaine d'oligomérisation et un fragment où le domaine d'oligomérisation a été supprimé P Δ [126-161]. La forme courte du domaine d'oligomérisation est indiquée par un rectangle bleu foncé. 122
- Figure 72 : Superposition des spectres HSQC $^1\text{H}-^{15}\text{N}$ de la protéine P pleine longueur P[1-241] en rouge et de deux fragments correspondant à la partie N-terminale de P avec le domaine d'oligomérisation P[1-163] (bleu) et à la partie C-terminale avec le domaine

d'oligomérisation P[127-241] (vert). Les spectres ont été acquis à un champ de 14.1 T et une température de 288 K, à des concentrations de 50 μ M dans un tampon phosphate 20 mM, NaCl 100 mM, pH 6.5.	123
Figure 73: Spectres HSQC ^1H - ^{15}N attribués de deux fragments de P correspondant à la partie C-terminale sans domaine d'oligomérisation P[1-126] et avec domaine d'oligomérisation : P[1-163] (bleu), P[1-126] (rouge). L'acquisition a été faite à un champ de 14.1 T, à 288 K, à une concentration de 50 μ M en tampon phosphate salin (pH 6.5).	124
Figure 74 : Régions de P avec des structures en hélices partielles déterminées par RMN.	125
Figure 75 : Position des sites mutés en cystéine dans la séquence de P par rapport aux régions structurées ou partiellement structurées.	126
Figure 76 : Alignement de séquence des différentes constructions de P : P sauvage P[1-241], P(S23C), P(S99C), P(S143C), P(S156C), P(E179C), P(E193C) et P(S237C).	127
Figure 77 : Spectres HSQC ^1H - ^{15}N superposés du mutant P(S237C) (en bleu ; 50 μ M dans un tampon 20 mM phosphate pH 6,5, 100 mM NaCl, 0,1% Azoture de sodium) et de P sauvage (rouge) acquis dans les mêmes conditions (14,1 T et 288 K). Les pics correspondant à la sérine 237 et lysine 238 sont indiqués par des flèches rouges.	128
Figure 78 : Agrandissement des spectres HSQC ^1H - ^{15}N de P(237C) marqué ^{15}N (rouge) et marqué $^{15}\text{N}^{13}\text{C}$ (vert) autour du résidu F241, indiquant la présence de deux formes de P(237C) dans l'échantillon marqué ^{15}N .	129
Figure 79 : Agrandissements des spectres HSQC ^1H - ^{15}N superposés du mutant P(S23C) (en brun ; 50 μ M dans 20 mM Phosphate pH 6,5, 100 mM NaCl, 0,1% NaN_3) et P sauvage (en bleu) acquis dans les mêmes conditions (14,1 T et 288 K).	130
Figure 80 : Agrandissement des spectres HSQC ^1H - ^{15}N superposés du mutant P(S99C) (en rouge ; 50 μ M dans 20 mM Phosphate pH 6,5, 100 mM NaCl, 0,1% NaN_3) et P sauvage(en bleu) acquis dans les mêmes conditions (14,1 T et 288 K), montrant des élargissement de signaux pour des résidus près du site de mutation (S94, F98, T105, T108).	131
Figure 81 : Agrandissements des spectres HSQC ^1H - ^{15}N du mutant P(S237C) avant (en rouge) et après traitement avec l'IAP (en bleu) acquis dans les mêmes conditions (50 μ M dans 20 mM Phosphate pH 6,5, 100 mM NaCl, 0,1% NaN_3 ; 14,1 T et 288 K), montrant la disparition du signal pour la forme 1 des résidus D235, E239, D240 et F241 liée à la proximité de la sonde paramagnétique sur C237.	132
Figure 82 : Rapport des intensité des signaux des spectres HSQC ^1H - ^{15}N de mutant P(S237C) avec une étiquette paramagnétique IAP (Ipara) et avant traitement à l'IAP (Idia). La position de la sonde paramagnétique est indiquée par un trait rouge.	133
Figure 83 : Rapport des intensité des signaux des spectres HSQC ^1H - ^{15}N des mutants P(S156C) et P(E193C) avec une étiquette paramagnétique IAP (Ipara) et avant traitement à l'IAP (Idia). Ces mutants étaient intacts dans la région N-terminale. La position de la sonde paramagnétique est indiquée par un trait rouge.	133
Figure 84 : Superposition des spectres HSQC ^1H - ^{15}N de P[1-126] seule 50 μ M (bleu) et en présence de M2-1[58-177] (vert). (T = 288 K, B_0 = 14,1 T)	134

- Figure 85 : Rapport des intensité résidu par résidu des signaux de P[1-126] en présence et en absence du domaine globulaire de M2-1. Les régions avec des variations significatives apparaissent en rouges (T18-I39) et (I91-E113). 135
- Figure 86 : Superposition des spectres HSQC 1H- 15N de P[1-163] seule (bleu) et en présence de M2-1[58-177] (rouge). (T = 288 K, B0 = 14,1 T). 136
- Figure 87 : Rapport des intensité résidu par résidu des signaux la partie N-terminale de la phosphoprotéine P avec le domaine d'oligomérisation P[1-163] en présence et en absence du domaine globulaire de M2-1. La régions avec un baisse significative apparaît en rouge (I91-125).Le domaine d'oligomérisation pour lequel aucun signal n'est observé apparaît sur fond vert 136
- Figure 88 : Récapitulatif des information fonctionnelles et structurales sur la phosphoprotéine du VRS. 137
- Figure 89 : Rôle de la protéine NS1 du VRSH dans la voie d'induction d'interféron de la cellule. NS1 a pour effet de diminuer les concentrations stationnaires de TRAF3 et IKK ϵ (flèches rouges), ce qui diminue le taux de transcription des gènes des IFN de type I. En effet, TRAF3 et IKK ϵ activent les facteurs de régulation d'interféron IRF3 et IRF7 par phosphorylation. Ceux-ci sont ensuite transloqués vers le noyau, où ils activent la transcription des gènes des IFN de type I. L'activation est symbolisée par des flèches noires. 140
- Figure 90 : Rôle de la protéine NS2 du VRSH dans la voie de réponse aux interférons. Les interférons de type I se lient à leurs récepteurs IFNAR2c et IFNAR1, ce qui stimule la phosphorylation des facteurs de transcription STAT1 et STAT2 par les kinases Tyk2 et JAK1 et la formation d'un complexe entre STAT2, STAT1 et le facteur de régulation d'interféron IRF9. Celui-ci est ensuite transloqué dans le noyau pour activer la transcription de gènes codant pour des protéines antivirales. L'effet de NS2 est de réduire le taux de STAT2 (flèche rouge). 141
- Figure 91 : Domaines fonctionnels des protéines NS1 et NS2. L'analyse des propriétés de liaison aux protéines TRAF3, IKK ϵ et STAT2 de mutants de délétion de NS1 et NS2 a permis de définir des domaines fonctionnels. La séquence térapeptidique C-terminale DLNP, commune à NS1 et NS2, n'intervient pas dans la formation d'oligomères. Schéma adapté de (Swedan et al. 2011). 142
- Figure 92: Alignement de séquence pour la protéine NS1 (plasmides pMal et pGEX). 143
- Figure 93 : Gel d'expression de polyacrylamide pour NS1 obtenu par électrophorèse en milieu dénaturant (SDS-PAGE). MP = marqueur de poids moléculaire. NI = extrait bactérien avant induction par l'IPTG. I = extrait bactérien après induction. La bande correspondant à MBP-NS1 est indiquée par une flèche. 144
- Figure 94 : Gel de protéine SDS-PAGE après élution au maltose de la fusion MBP-NS1 (MP : marqueur de poids moléculaire, R : résine amylose, FT1 à FT6 : fractions non retenues qui contiennent la fusion MBP-NS1. 144

- Figure 95 : Gel de protéine SDS-PAGE après coupure à la thrombine de NS1. MP : marqueur de poids moléculaire, AvC : fusion MBP-NS1 avant coupure, ApC: protéines après coupure avec deux bandes correspondant à la MBP et à NS1. 145**
- Figure 96 : Profil d'élution du mélange par suivi de l'absorption à trois longueurs d'onde (282 nm en bleu, 258 nm en rouge et 215 nm en rose) et analyse par SDS-PAGE des fractions collectées lors de la purification par filtration sur gel sur colonne Superdex-75. Le premier pic (fraction B4) semble contenir des agrégats. Le deuxième pic (fraction D3) correspond à la taille de la MBP. Le troisième pic indiqué par une flèche épaisse est celui qui correspond au volume d'élution attendu pour NS1 sous forme monomère. Plusieurs fractions ont été déposées (E5, E6, E7, E8, F2, F3, F4) : elles correspondent à la protéine d'intérêt NS1. 146**
- Figure 97 : A). Superposition des spectres HSQC ^1H - ^{15}N de la protéine NS1 à différentes concentrations (10, 40 et 100 μM) acquis à un champ de 18,8 T (256 scans), à une température de 293 K et dans un même tampon ("tampon final" : 10 mM Tris-HCl pH 7,5, 200 mM NaCl, 1 mM DTT et 0,5 mM EDTA). 148**
- Figure 98 : Principe de la mesure de stabilité d'une protéine par "ThermoFluor" (<http://www.jenabioscience.com>). La protéine est symbolisée par des cylindres oranges, la molécule fluorescente par des étoiles bleues ou rouge, suivant l'intensité de fluorescence. 150**
- Figure 99 : Mesure en fonction de la température de la fluorescence du SYPRO avec des quantités croissantes de la protéine NS1 dans le "tampon final de purification". Les mesures ont été effectuées en duplicata. 151**
- Figure 100 : A) Dérivée de la courbe de fluorescence de SYPRO orange en présence de NS1 dans son tampon de conservation. Elle présente un minimum à $\sim 55^\circ\text{C}$, correspondant à la température de demi-dénaturation (T_m). B) Courbes de dénaturation de NS1 mesurées par dichroïsme circulaire montrant que l'état de dénaturation est irréversible (Pretel et al. 2013). 151**
- Figure 101 : Mesure de la température de dénaturation en utilisant les dérivées des courbes de fluorescence pour trois tampons de type "Tris" : (1) Tampon Tris à pH 8,5, (2) tampon de final de purification de la protéine (10 mM Tris à pH 7.5 avec 200 mM NaCl, 1 mM DTT et 0,5 mM EDTA), (3) tampon Tris à pH 8,5 avec 500 mM NaCl. 152**
- Figure 102 : Conditions de stabilité de NS1 dans différents tampons donnés en abscisse. L'axe des ordonnées correspond à la température de dénaturation mesurée par ThermoFluor. Le tampon final de purification de la protéine est en rouge. 153**
- Figure 103 : Mesure de la fluorescence de SYPRO orange en présence de la protéine NS1 en fonction de la concentration en NaCl (100, 300 et 600 mM) en tampon Tris PH 7,5. Les températures de dénaturation sont respectivement de 53, 58 et 61 $^\circ\text{C}$. 154**
- Figure 104 : A) Effet du pH sur la stabilisation de NS1. Les barres indiquent les températures de dénaturation. B) Courbes de fluorescence du SYPRO orange en présence de NS1 dans un tampon phosphate pH 4,4 avec 500 mM NaCl. 154**
- Figure 105 : Courbes de fluorescence du SYPRO orange en présence de la fusion MBP-NS1, de NS1 seule et de la MBP seule, dans le "tampon final" montrant que la protéine NS1 expose beaucoup de régions hydrophobes. 155**

LISTE DES TABLEAUX

Tableau 1 : Les différents candidats pour des médicaments contre le VRS et leurs protéines cibles (Empey et al. 2010).	54
Tableau 2 : Tampons utilisés pour essayer de solubiliser les trois constructions N-terminales de la protéine M2-1	67
Tableau 3 : Récapitulatif des valeurs des constantes de dissociations K_d déterminées par RMN pour l'association de différents oligonucléotides avec M2-1[58-177].	106

LES ABREVIATIONS

aa / AA	acide aminé
ADN	Acide DésoxyriboNucléique
ADNc	ADN complémentaire
AMPV	Avian Metapneumovirus ou Métapneumovirus aviaire
ARN	Acide RiboNucléique
ARNm	ARN messenger
ARNt	ARN de transfert
ARN T7	ARN du bactériophage T7
APV	Avian Pneumovirus
ATP	Adénosine Triphosphate
BDV	Borna Disease Virus, virus de la maladie de Borna
CCA	Chimpanzee coryza, Coryza du chimpanzé
Complexe RdRp	Complexe RNA dependent RNA polymerase
Complexe RNP	Complexe RiboNucléoProtéique
CTD	C-terminal domain, domaine C-terminal
DTT	Dithiothréitol
<i>E. coli</i>	<i>Echerichia coli</i>
GAGs	GlycosAminoGlycanes
GE	Gene End
GS	Gene Start
GST	Gluthation S-transférase
GDP	Guanosine DiPhosphate
GTP	Guanosine TriPhosphate
HRA,-B,-C	Heptapeptides Répétés A, B et C
HEp-2	Human Epithelial-2 cells
HMPV	Human Metapneumovirus, Métapneumovirus humain
HSP70	Heat Shock Protein 70
HSQC	Heteronuclear Single Quantum Correlation
IFN	Interféron
InVS	Institut de Veille Sanitaire
IGIV	ImmunoGlobuline IntraVeineuse
IgA	Immunoglobuline A
IPTG	IsoPropyl-β-D-Thio-Galactoside
IRF3	Interferon Regulatory Factor 3
Kb	Kilobase
Kd	Constante de dissociation
kDa	kiloDalton
le	leader
tr	trailer
MAVS	Mitochondrial AntiViral Signaling protein
MPV	MétaPneumoVirus

MHz	Méga Hertz
MDA5	Melanoma Differentiation-Associated gene 5
milieu LB	milieu Luria Bertani
NS1-2	Protéines NS 1 et 2
nt	nucléotide
NTD	N-terminal Domain, Domaine N-terminal
ORF	Open Reading Frame, phase ouverte de lecture
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PCR multiplexe	Multiplex Polymerase Chain Reaction
PIV	Virus ParaInfluenza
PRE	Paramagnetic Relaxation Enhancement
CnPnV	Canine PneumoVirus ou Pneumovirus du chien
Protéine F	Protéine de fusion
Protéine G	Glycoprotéine
Protéine L	Polymérase L
Protéine M	Protéine de matrice
Protéine N	Nucléoprotéine
Protéines NS	Protéines non structurales
Protéine P	Phosphoprotéine
Protéine SH	Petite protéine hydrophobe (SH pour Small Hydrophobic)
PVM	Pneumonia Virus of Mice, Virus de la Pneumonie Murine
RE	Reticulum Endoplasmique
RIG-I	Retinoic-acid-Inducible Gene-I
RMN	Résonance Magnétique Nucléaire
RNase	RiboNucléase
RNP	Ribonucléoprotéine
rpm	rotation par minute
RT	Room Temperature, température ambiante
SDS	Sodium Dodécyl Sulfate
SDS-PAGE	Sodium Dodecyl Sulfate PolyAcrylamide Gel Electrophoresis
STAT2	Signal Transducer and Activator of Transcription 2
TNF- α	α Tumor necrosis factor ou facteur de nécrose tumorale α
VIH	Virus d'Immunodéficience Humaine
VP30	Viral Protein (30 kDa)
VRS	Virus Respiratoire Syncytial
VRSB-H-O	VRS bovin-humain-ovin
VSV	Vesicular Stomatitis Virus, Virus de la Stomatite Vésiculaire
WHO	World Health Organization ou OMS
WT	Wild Type, sauvage

Chapitre 1 : Introduction

1 Généralités sur le VRS

1.1 Historique

Le virus respiratoire syncytial (VRS) a été isolé pour la première fois en 1955 à partir d'un chimpanzé qui présentait un coryza, et a été alors dénommé CCA pour agent du coryza du chimpanzé (Blount et al. 1956). Il a rapidement été identifié aux Etats-Unis comme virus humain et agent pathogène respiratoire important chez les enfants (Chanock and Finberg 1957; Chanock 1957). Le VRS est maintenant reconnu comme le plus important agent viral de la maladie des voies respiratoires chez les nourrissons du monde entier (Hall et al. 2009) (Nair et al. 2010). Il est désigné comme tel en raison des effets cytopathogènes observés dans les cultures infectées. En effet, il provoque la fusion des cellules qu'il infecte donnant naissance à une cellule géante plurinucléée, un syncytium (Figure 1). En France, le VRS a été décrit pour la première fois chez des bébés prématurés lors d'une épidémie d'infections respiratoires en mars 1960 (Breton and Dubois 1961).



Figure 1 : Radiographies pulmonaires d'un enfant infecté par le virus respiratoire syncytial démontrant la formation de syncytium et la mise en évidence d'une bronchiolite virale aiguë.

En 1970, à l'institut national de santé animale au Japon, un VRS ayant pour hôte naturel le bovin a été isolé (Inaba et al. 1970). Celui-ci présentait de fortes similitudes avec la souche humaine d'un point de vue génomique, antigénique, épidémiologique et pathogénique. Le VRS est aujourd'hui reconnu comme étant l'agent étiologique principal des bronchiolites chez les nourrissons, les personnes âgées et immunodéprimées.

1.2 Pathologie du VRS humain

Les symptômes de la maladie au cours de l'infection par le VRS varient largement entre les individus. L'infection peut soit se limiter au tractus respiratoire supérieur, comme dans le cas de rhinopharyngites ou impliquer à la fois les voies supérieures et inférieures, comme dans le cas de bronchiolites ou de pneumonies. Des anomalies après l'infection de la fonction respiratoire et une sensibilisation éventuelle à l'asthme peuvent persister pendant l'adolescence. L'incidence annuelle des infections par le VRS est estimée à 64 millions de cas par l'organisation mondiale de la santé (OMS) et à 160 000 décès annuels (The World Health Report 2006, WHO 2006). Chez le nourrisson, entre 60 à 80 % des bronchiolites sont dues au VRS et conduisent à une hospitalisation dans 42 à 45 % des cas en Europe (Simoes and Carbonell-Estrany 2003). Ce virus peut entraîner aussi des pneumonies du nourrisson. Chez certains nourrissons, en particulier les prématurés, elles sont très sévères, entraînant une insuffisance respiratoire aiguë qui nécessite des mesures de réanimation. La mortalité des bronchiolites hospitalisées en réanimation pédiatrique est de 2 %. Pour toutes ces raisons, le VRS constitue une priorité en santé publique.

La mise en évidence du VRSH est basée sur la détection du virus qui est présent au niveau des sécrétions nasales ou respiratoires chez les personnes infectées. Le virus n'est détectable généralement que dans les premiers jours de l'infection. Il existe plusieurs méthodes pour détecter le virus, mais la plus utilisée est la recherche d'antigène de VRS. La recherche rapide d'antigène de VRS est réalisée en France en laboratoire d'analyse médicale, avec des résultats disponibles en moins d'une heure. La technique d'immunofluorescence directe consiste à étaler sur une lame des cellules des sécrétions respiratoires et les mettre en contact avec un anticorps spécifique anti-VRS marqué par un fluorochrome. La lecture au microscope à fluorescence révèle la présence d'une fluorescence cytoplasmique des cellules bronchiques traduisant la présence du VRS.

Les techniques détectant le matériel génétique viral comme la PCR multiplexe (Multiplex Polymerase Chain Reaction) ont l'avantage d'identifier non seulement le VRS mais aussi un grand nombre de virus ou d'agents pathogènes respiratoires. Cependant, les résultats de ces examens sont plus longs à obtenir que le test de détection rapide du VRS et sont également beaucoup plus coûteux. Ils sont donc réservés au diagnostic des infections respiratoires sévères de sujets hospitalisés.

La réaction immunitaire dirigée contre le VRS associe une réaction humorale et une réaction immunitaire cellulaire. La réaction humorale qui comporte une réponse locale avec

synthèse d'immunoglobulines A (IgA) sécrétoires et une réponse systémique avec des anticorps neutralisants. La réaction immunitaire cellulaire joue un rôle important dans la défense contre l'infection étant donné que les patients ayant un déficit de l'immunité cellulaire (sujets greffés ou infectés par le virus d'immunodéficience humaine (VIH)) font des infections à VRS sévères. Les anomalies dues à l'infection au VRS s'expliquent par un effet cytopathique direct du VRS sur les cellules bronchiques, c'est à dire des altérations métaboliques, biochimiques et morphologiques de la cellule infectée par le virus, et par une réaction immuno-pathologique qui a été mise en évidence dans les années 1960 lorsque des essais vaccinaux réalisés avec des préparations vaccinales contenant le virus inactivé ont été un échec complet. En effet, ces vaccins étaient administrés par voie générale et entraînaient une très forte réponse immunitaire humorale et cellulaire systémique, mais aucune réponse immunitaire locale. Lorsque l'enfant vacciné était infecté, le VRS se multipliait dans les cellules de la muqueuse bronchique non protégée par des IgA sécrétoires et cette infection locale entraînait l'afflux de cellules immunitaires systémiques dont des cellules T cytotoxiques (CD4) spécifiques du VRS entraînant une destruction très importante des cellules bronchiques infectées. La conséquence clinique de cette réponse immunitaire déséquilibrée était la survenue de bronchiolites sévères beaucoup plus fréquentes que chez les enfants non vaccinés.

La contamination s'effectue par les voies nasales supérieures en conditions naturelles. A partir de la muqueuse nasale, le virus se dissémine dans le reste de l'appareil respiratoire par aspiration des mucosités et des sécrétions contenant des particules virales, ce qui entraîne la fusion entre les cellules infectées et non infectées voisines entraînant ainsi la formation de syncytia et permettant la colonisation du virus sans passage par le milieu extra cellulaire (Collins et al. 2001; Collins and Murphy 2002). Ce mode de contamination permet au virus d'échapper à la réaction immunitaire humorale de l'hôte. Il ne semble pas exister de phase de virémie, le virus restant localisé dans le tractus respiratoire. Les études *in vitro*, bien que ne reflétant pas toujours les observations faites *in vivo*, ont permis de mieux comprendre les effets cytopathiques du VRS (Tristram et al. 1998) (Wright et al. 2005) (Zhang et al. 2002). Les lignées cellulaires utilisées pour effectuer des infections efficaces sont des cellules de l'épithélium respiratoire telles que les HEp-2 issues d'un carcinome de larynx humain, pour le VRS humain.

1.3 Taxonomie

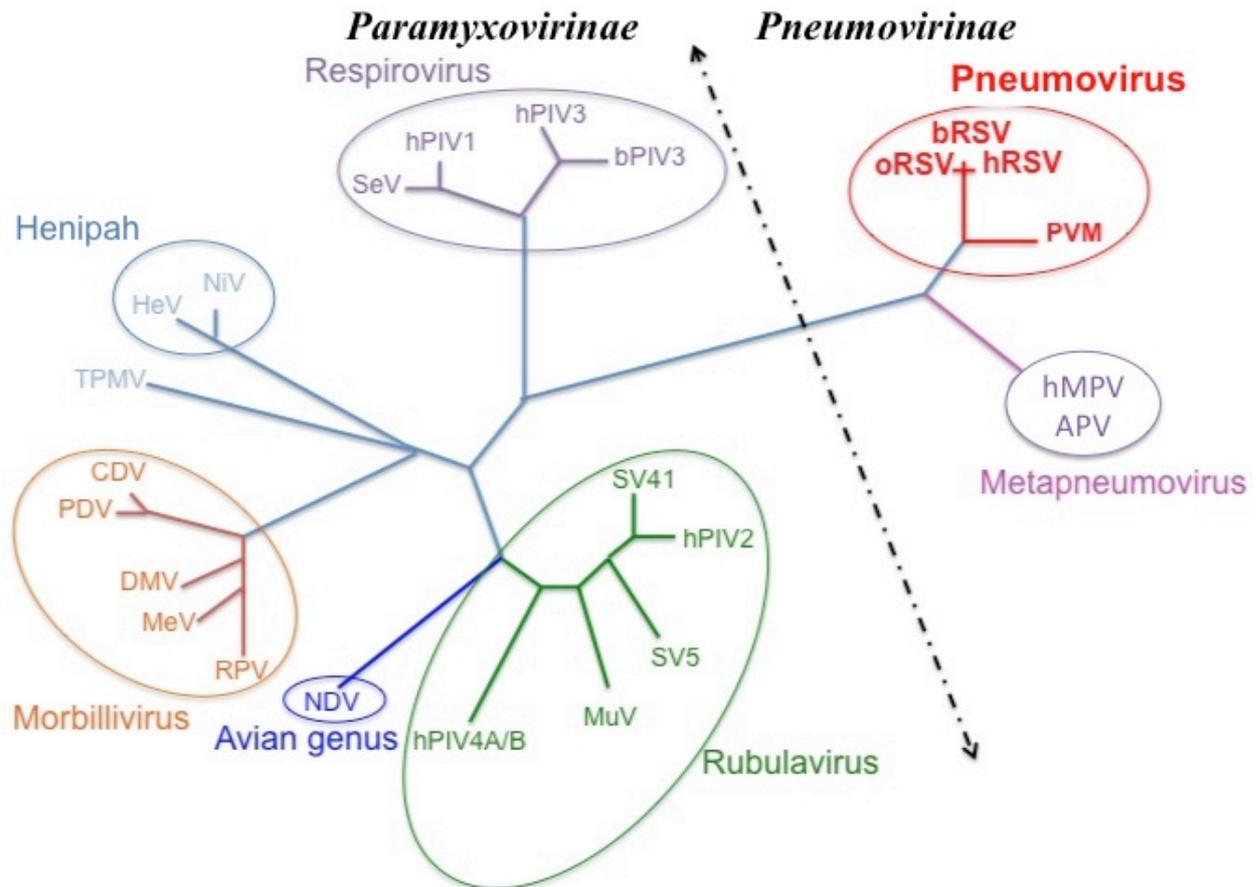


Figure 2 : Arbre phylogénétique des membres de la famille des Paramyxoviridae. Il y a deux sous-familles : les Pneumovirinae, qui comporte les pneumovirus (VRS humain (hRSV), VRS bovin (bRSV), VRS ovin (oRSV), pneumovirus de la souris (PVM pour Pneumonia Virus of Mice)) et les métapneumovirus (métapneumovirus respiratoire humain (hMPV), pneumovirus aviaires (APV)), et les Paramyxovirinae, qui comprend les respirovirus (virus Parainfluenza humain 1 et 3 (hPIV1 et hPIV3), virus de Sendai (SeV)), les henipah (Hendra virus (Hev), Nipah virus (NiV)), les morbillivirus (canine distemper virus (CDV), morbillivirus de phoque (PDV), morbillivirus du dauphin (DMV), virus de la rougeole (MeV), Pinderpest virus (RPV)), les rubulavirus (virus Parainfluenza humain 2 et 4 (hPIV2 et hPIV4), Mumps virus (MuV), Simian virus 5 (SV5), Simian virus 41 (SV41)), le virus de la maladie de Newcastle et les tupaia paramyxovirus (TPMV).

Le VRS est un virus enveloppé à ARN de sens négatif non segmenté de l'ordre des Mononegavirales. Il est classé dans la famille des *Paramyxoviridae*. Cette famille a deux sous-familles : (i) celle des *Pneumovirinae*, qui comporte les Pneumovirus (le VRS humain

(VRSH), le VRS bovin (VRSB), le VRS ovin (VRSO), le virus de la pneumonie murine (PVM pour Pneumonia Virus of Mice) et le Pneumovirus du chien (CnPnV pour Canine Pneumovirus)) et les Méta-pneumovirus, et (ii) celle des *Paramyxovirinae*, qui comprend le virus de Sendai, les virus Parainfluenza (HPIV), virus de la rougeole, et un nombre d'autres agents pathogènes humains et animaux (Figure 2). Les Pneumovirus se différencient des Méta-pneumovirus par l'ordre des gènes, la présence de gènes supplémentaires ainsi que par des séquences génomiques distinctes.

Concernant le VRS, il existe 73 % d'identité en moyenne entre les séquences en acide aminé des souches humaine et bovine.

2 Organisation du VRS

2.1 Virion du VRS

La structure du VRS observée en microscopie électronique est variable, le virus étant fréquemment observé sous forme filamenteuse avec un diamètre de 60-110 nm et une longueur de 2 à 8 μm (Roberts et al. 1995) ou arrondie avec un diamètre de 100-350 nm (Figure 3).



Figure 3 : Micrographies électroniques à coloration négative, montrant les caractères morphologiques de VRS sous forme arrondie (Gauche) sous forme filamenteuse (Droite).

Le VRS est un virus enveloppé, l'enveloppe virale étant une bicouche lipidique formée autour du virion lors de sa fusion avec la membrane cytoplasmique de la cellule, dans laquelle viennent s'insérer trois glycoprotéines virales : la protéine d'attachement G, la protéine de fusion F et la petite protéine hydrophobe SH. Les protéines G et F forment des

spicules de 11 à 20 nm de long à la surface de la particule virale. L'intérieur de l'enveloppe virale est recouvert par la protéine de matrice M ayant un rôle dans la morphogénèse virale.

L'ARN génomique est associé de façon constitutive à la nucléoprotéine virale (N), formant ainsi un complexe ribonucléoprotéique (RNP) de symétrie hélicoïdale, appelé nucléocapside, résistant aux RNases (Collins et al. 2001) (Collins and Murphy 2002). Cette nucléocapside sert de matrice à la fois pour la transcription et la réplication du génome par l'ARN polymérase virale (RNA-dependent RNA polymerase, RdRp) (Grosfeld et al. 1995) (Khattar et al. 2001). Le complexe RdRp est constitué a minima de la polymérase L et de son cofacteur, la phosphoprotéine P. Lorsque la polymérase fonctionne en mode transcription, le RdRp est complété par le cofacteur viral M2-1 (antitermineur de la transcription). Lors de la réplication c'est le co-facteur M2-2 qui est associé au RdRp.

2.2 Génome

Le génome du VRS consiste en un ARN non segmenté de polarité négative de 15 kilobases (kb) et de 10 gènes (Figure 4). L'organisation du génome viral est la suivante :

- En 3', une région « leader » « de tête », extra génique, de 44 nucléotides qui contient vraisemblablement un promoteur viral majeur (Mink et al. 1991).
- En 5', après le gène L, une région « trailer » « de fin » qui est constituée de 155 nucléotides.
- Deux segments nucléotidiques agissant en Cis et localisés entre 5' et 3' qui paraissent indispensables pour la transcription, la réplication et l'assemblage viral. En 3', cette région est constituée de 88 nucléotides, comprenant la région « leader » « de tête », le signal d'initiation de transcription du gène codant pour la protéine non structurale (NS1) et son extrémité non codante (Zamora and Samal 1992). En 5', cette région est constituée de 192 nucléotides, incluant l'extrémité 5' non codante du gène de la polymérase, le signal de terminaison de transcription, ainsi que l'extrémité 5' extra génique du génome (Yunus et al. 1998).
- Entre les régions 5' et 3', 88 % du génome codent pour les dix gènes du VRS humain. Chaque gène débute par un signal d'initiation de transcription de 9 à 10 nucléotide, dont la séquence est conservée sur presque tous les gènes, noté GS « Gene Start » (3'-CCCCGUUUA-5') (Zamora and Samal 1992), et se termine par un signal de terminaison de la transcription qui ordonne la fin de la transcription et la polyadénylation (Kuo et al. 1996). Ce signal de terminaison est semi-conservé entre chaque gène et a une longueur de 12 à 13 nucléotides et est noté GE « Gene End » (3'-UCANNNNNNUUUU-5').

Les 9 premiers gènes sont non chevauchants et séparés par des régions d'une longueur de 1 à 52 nucléotides. Les gènes M2 et L se chevauchent sur 67 nucléotides (Collins et al. 1987). Chez le VRSH, il a été postulé par le groupe de Rachel Fearn, à l'université de Leeds en Angleterre, que ce chevauchement aurait une fonction régulatrice de la transcription de la polymérase (Munday et al. 2015).

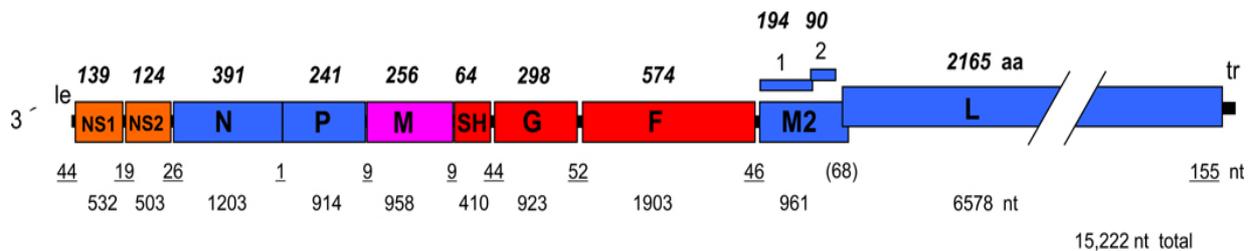


Figure 4 : Carte du génome de sens négatif du VRS. (Souche A2) à l'échelle approximative. Les numéros en dessous du schéma indiquent la longueur en nombre de nucléotides (nt): le 3' leader (le) et le 5' trailer (tr) et les régions intergéniques sont soulignées et la longueur de chevauchement des gènes est entre parenthèses. Les numéros en gras au dessus indiquent les longueurs des protéines non modifiées en acide aminé (aa). Les protéines virales sont les suivantes : G, glycoprotéine d'attachement; F, glycoprotéine de fusion; SH, petite glycoprotéine hydrophobe; M, protéine de matrice; N, nucléoprotéine; P, phosphoprotéine; L, la grande unité polymérase; M2-1, produit du premier ORF de l'ARNm M2; M2-2; produit du deuxième ORF de l'ARNm M2; NS1, protéine non structurale 1; NS2, protéine non structurale 2. Les cadres de lecture ouverts de M2-1 et M2-2 se chevauchent sur le gène. (Collins and Melero 2011)

2.3 Cycle viral

Une fois que le virus a pénétré à l'intérieur de la cellule hôte, il s'ensuit de multiples événements dans le cytoplasme afin de former de nouvelles particules virales et de propager l'infection (Figure 5). Ces étapes ne nécessitent pas un passage dans le noyau car les *Pneumovirus* possèdent leur propre polymérase pour assurer la transcription et la réplication.

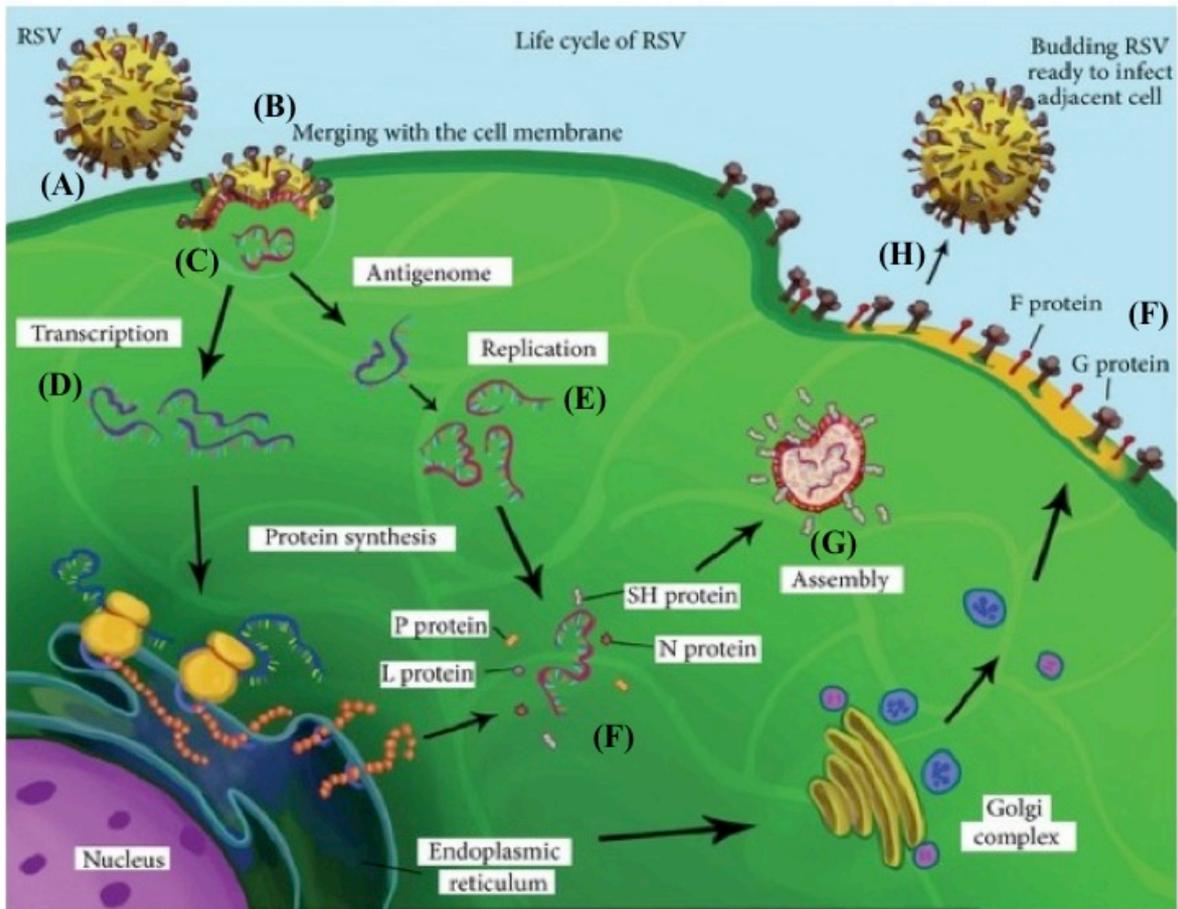


Figure 5 : Représentation schématique du cycle viral du virus respiratoire syncytial . (A) Le VRS se fixe sur la membrane cellulaire via la protéine G, (B) Cette étape est suivie de la fusion médiée par la protéine F, ce qui conduit à l'incorporation de l'enveloppe virale dans la membrane cellulaire et la libération subséquente de la nucléocapside dans le cytosol, (C) pour permettre à l'ARN viral d'entrer dans la cellule. (D) Transcription de l'ARNm, (E) Réplication de l'ARN, (F) les protéines M, F et G forment des complexes avec la membrane plasmique, alors que les protéines N, P, L et M2-1 forment des inclusions dans le cytoplasme. (G) Grâce à une série d'interactions, la protéine M coordonne l'assemblage des protéines de l'enveloppe avec les protéines de la nucléocapside. (H) Bourgeonnement du nouveau virion mature prêt à infecter une cellule voisine (Bawage et al. 2013).

2.3.1 L'entrée du virus : attachement et fusion

2.3.1.1 Attachement

La première étape du cycle viral consiste en l'attachement des particules à la surface de la cellule hôte. Il a été démontré que l'interaction entre la protéine G et les glycosaminoglycanes (GAGs) est essentielle à l'attachement de la particule virale à la

membrane cellulaire (Hallak et al. 2007). Les GAGs sont composés de sous-unités de disaccharides répétées présentes dans la matrice extracellulaire, et la protéine G semblerait s'attacher préférentiellement aux GAGs contenant des disaccharides héparane sulfate. Il est possible que l'interaction entre les GAGs et la protéine G ne soit que le premier événement d'attachement et que celui-ci soit suivi par d'autres étapes impliquant une ou plusieurs des trois protéines de surface et cette hypothèse était posée quand des virus où les gènes G et SH étaient enlevés se répliquaient très bien *in vitro*. Ceci suppose que parmi les trois glycoprotéines d'enveloppe, seule la protéine de fusion F est essentielle pour l'entrée du virus dans la cellule hôte.

2.3.1.2 Fusion

La fixation du virion à la surface de la cellule hôte induit ensuite un changement conformationnel de la protéine F conduisant à la fusion de l'enveloppe virale avec la membrane cellulaire. Le processus d'entrée est indépendant du pH et conduit à la libération de la nucléocapside dans le cytoplasme de la cellule infectée (Tayyari et al. 2011).

2.3.2 La transcription virale

2.3.2.1 Composition du complexe de transcription

Contrairement aux autres virus de l'ordre des Mononegavirales, la transcription et la réplication virale des Pneumovirus font intervenir des mécanismes de régulation supplémentaires liés respectivement aux protéines M2-1 et M2-2. Lorsque le virus pénètre à l'intérieur de la cellule, celui-ci libère sa nucléocapside dans le cytoplasme. L'ARN polymérase ARN dépendante peut alors initier la transcription du génome en des ARNm coiffés en 5' et polyadénylés en 3' (Figure 6).

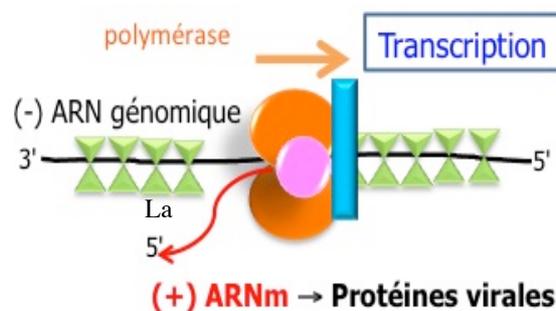


Figure 6 : Schéma de la composition du complexe de transcription virale chez le VRS, avec les cofacteurs essentiels de la polymérase virale (orange) (unité catalytique L (orange), la phosphoprotéine (bleu) et la protéine M2-1 (rose))

2.3.2.2 Signaux en « cis » pour la transcription

Les 10 derniers résidus de la séquence leader, région riche en U, sont nécessaires pour une synthèse optimale des ARNm (Fearn et al. 2000). Donc tout d'abord, avant de synthétiser les ARNm correspondant aux différents gènes, la polymérase virale doit obligatoirement transcrire la séquence «leader» en 3' du génome (Cowton and Fearn 2005). De plus, par la suite, celle-ci peut continuer la transcription de l'extrémité 3' à l'extrémité 5' sans se détacher. La polymérase virale transcrit les gènes de manière séquentielle selon des signaux d'initiation de transcription "Gene Start", comprenant la région de la coiffe et d'initiation de la synthèse d'ARNm (des séquences conservées (3'-CCCCGUUUA-5')) et de terminaison "Gene End", où se trouve une série de 4 à 7 uraciles responsables de la polyadénylation, d'une courte région intergénique non transcrite (des séquences semi-conservées,). Celles-ci sont présentes dans chacune des régions intergéniques. Lorsque la polymérase en cours de transcription rencontre la série d'uraciles dans la séquence terminatrice du gène, elle libère l'ARNm nouvellement synthétisé et progresse jusqu'à la rencontre d'une séquence de commencement d'un gène. Les GE sont parfois inefficaces et deux GS se succèdent sur un même ARN. Ceci conduit à la synthèse de deux protéines sur le même ARN de fusion qui est dit ARNm bicistronique. Ces ARN messenger bicistroniques représentent environ 10 % des ARNm viraux totaux.

2.3.2.3 Coiffage des ARNm

Durant la transcription virale, l'ARNm naissant est guanylylé puis méthylé par le complexe polymérase afin de produire la coiffe qui permettra leur traduction par les ribosomes de la cellule hôte (Barik 1993). La structure de la coiffe des ARNm des Mononegavirales est similaire à celle des ARNm cellulaires, de type 5' m7GpppNm (Abraham et al. 1975). Chez les Mononegavirales, la démonstration de l'activité de coiffe des ARNm par la polymérase L a été mise en évidence la première fois avec le virus Sendaï grâce à la mise au point d'un système *in vitro* par l'équipe de Kenji Mizumoto à Tokyo (Ogino et al. 2005). L'ajout de la coiffe des ARNm viraux a également été caractérisé pour le VSV par la même équipe (Ogino and Banerjee 2007).

Les auteurs ont proposé un mécanisme non conventionnel, expliquant le coiffage cotranscriptionnel des ARNm viraux par la protéine L différent de celui utilisé par la machinerie cellulaire et de la plupart des virus. Dans un premier temps, la polymérase L hydrolyserait un GTP en GDP (activité GTPase), puis, durant la transcription virale, le pré-

ARNm 5' tri-phosphate naissant (pppApApCpApG) représenterait le substrat de la protéine L aboutissant à la formation d'un complexe L/pré-ARNm intermédiaire 5' monophosphate (L/pApApCpApG) et à la libération d'un pyrophosphate. Enfin, l'ARN serait transféré au GDP via l'activité guanylyltransférase de la polymérase L (activité ARN:GDP polyribonucleotidyltransférase) pour former la structure de la coiffe (GpppApApCpApG). Par analogie avec le VSV, il est supposé que les pré-ARNm viraux néosynthétisés sont coiffés cotranscriptionnellement en 5' et polyadénylés en 3' par la protéine L du VRS (Barik 1993; Liuzzi et al. 2005). D'autre part, la région semi-conservée des GE contient une séquence riche en U marquant la position à laquelle la polyadénylation est ajoutée à l'ARNm par analogie avec le VSV (Barr et al. 1997).

2.3.2.1. Gradient de transcription

Parfois, la polymérase virale se détache des régions intergéniques. Mais elle est alors incapable d'initier la synthèse de l'ARNm du gène suivant. Cela occasionne donc un arrêt de la transcription des gènes se trouvant sur le reste du génome jusqu'à l'extrémité 5'. Par conséquent, la quantité d'ARNm de chacun des gènes décroît selon leur position dans le génome par rapport à l'extrémité 3'. Les gènes proximaux sont favorisés et plus fréquemment transcrits que les gènes se rapprochant de l'extrémité 5'. La synthèse complète d'un ARNm est dépendante de la protéine M2-1 jouant le rôle d'antitermateur de la transcription (Fearn and Collins 1999) (Figure 7).

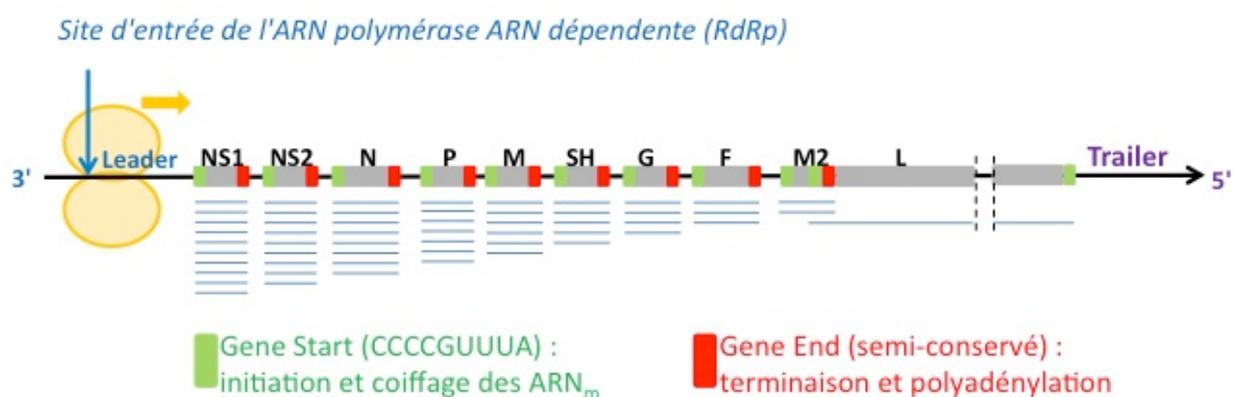


Figure 7 : Schéma du gradient décroissant d'expression des gènes (ARNm) en fonction de leur éloignement par rapport à l'extrémité 3'.

De plus, il a été démontré que la polymérase ne peut s'attacher à l'intérieur du génome. Cependant, celle-ci pourrait avancer et reculer sur le génome, ce qui expliquerait la transcription des gènes M2-1 et M2-2 contenant des cadres de lecture ouverts chevauchants, et du gène L chevauchant le gène M2.

Lorsque les protéines virales atteignent une concentration suffisante dans le cytoplasme de la cellule infectée, le complexe polymérase décroît progressivement son activité de transcription et bascule vers la réplication. La protéine M2-2 assurerait ce basculement dans le programme de synthèse d'ARN en faveur de l'ARN génomique pour permettre l'assemblage de nouveaux virions (Bermingham and Collins 1999).

2.3.3 La réplication virale

La réplication du génome viral débute suite à la traduction des transcrits et à l'accumulation de protéines virales. Tout d'abord, le génome ARN négatif du VRS est répliqué par l'intermédiaire d'un ARN antigénomique (+), qui consiste en une copie d'ARN parfaitement complémentaire et de même longueur que le génome (-). Cet antigénome ne contient aucun cadre de lecture ouvert et aucune protéine ne semble être transcrite à partir de ce dernier. Son unique fonction serait de servir d'intermédiaire afin de refaire des copies du génome (-) qui seront au centre des particules virales nouvellement formées. Lors de la réplication virale, la synthèse d'ARN est assurée par le complexe polymérase constitué de la protéine L ainsi que de la protéine P. Il semblerait qu'au cours de la réplication, la polymérase virale ignorerait simplement les jonctions entre les gènes et synthétiserait une copie complémentaire et exacte du génome initial (-). L'ARN néosynthétisé est encapsidé par la nucléoprotéine N de façon concomitante.

Le processus de décapsidation de l'ARN matrice permettant l'accessibilité au complexe polymérase reste à ce jour inconnu, de même que celui de l'encapsidation du brin néosynthétisé. L'encapsidation de l'ARN génomique par la protéine N n'est pas propre aux Mononegavirales puisqu'on la retrouve chez d'autres virus à ARN négatif tels que les Orthomyxovirus (Ye et al. 2006), les Bunyavirus (Ferron et al. 2011) et les Arenavirus (Hastie et al. 2011).

L'équilibre entre la transcription (Figure 6) et la réplication (Figure 8) serait contrôlé par la concentration intracellulaire de la protéine M2-2. En début d'infection, la quantité de M2-2 est faible et on observe un haut taux de transcription, alors que plus la concentration de M2-2 augmente, plus la transcription est inhibée et la réplication favorisée. (Bermingham and Collins 1999) (Figure 8)

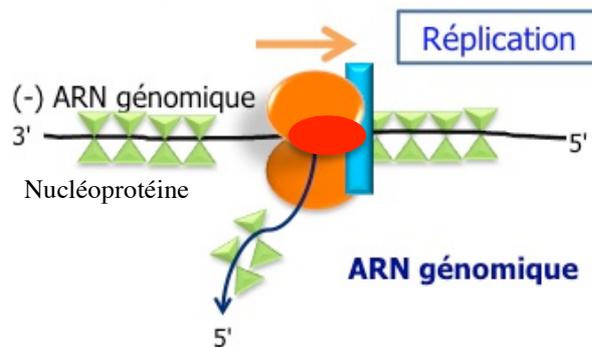


Figure 8 : Schéma de la composition du complexe de réplication virale chez le VRS, avec les cofacteurs essentiels de la polymérase virale (Orange) (l'unité catalytique L en orange, la phosphoprotéine (bleu) et la protéine M2-2 (rouge).

2.3.4 Corps d'inclusion de protéines du VRS

L'infection de cellules *in vitro* par le VRS provoque la formation de corps d'inclusion dans le cytoplasme des cellules infectées qui grossissent en fonction du temps (Figure 9).

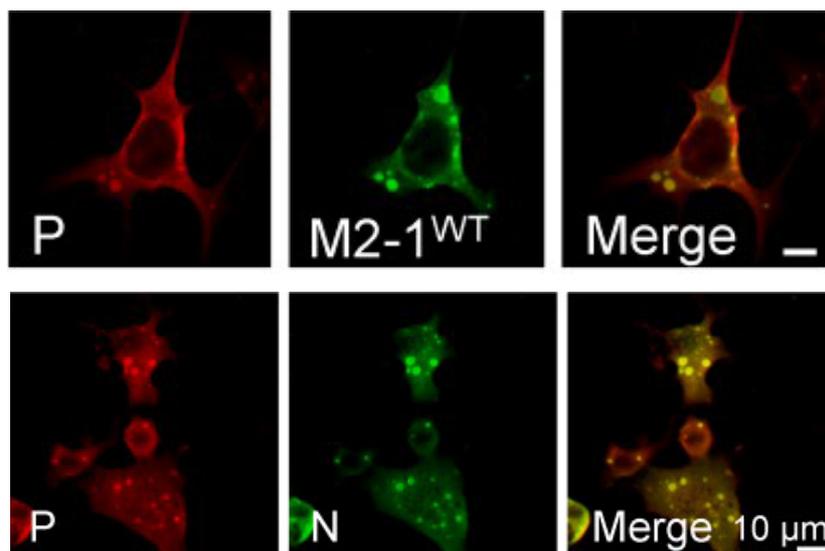


Figure 9 : Observation de corps d'inclusion après une infection des cellules BSR-T7 avec le VRS. Marquage indirect 24 h post-infection avec un anticorps monoclonal de souris anti-P (rouge) et un anticorps polyclonal de lapin anti-M2-1 et anti-N (vert) (Blondot et al. 2012)

Ce phénomène a été décrit pour la première fois par l'équipe de Norrby en 1970 pour le VRS (Norrby et al. 1970), mais avait déjà été observé pour d'autres virus tel que le virus de la pneumonie de souris (pneumonia virus of mice (PVM)) (Compans et al. 1967). La composition de ces corps d'inclusion a été caractérisée par marquages avec des anticorps dirigés contre les principales protéines structurales, montrant qu'ils contiennent toutes les protéines du complexe polymérase : la phosphoprotéine P, la nucléoprotéine N, la protéine

M2-1 et la polymérase L (Garcia et al. 1993; Carromeu et al. 2007). Ceci est illustré dans la Figure 9 pour les protéines P, M2-1 et N qui sont co-localisées au niveau de corps d'inclusion (Blondot et al. 2012).

L'équipe du Dr Reena Ghildyal, à l'université de Canberra en Australie, a mis en évidence la présence de la protéine de matrice M dans les corps d'inclusion lors des phases tardives par des cinétiques d'infection (Ghildyal et al. 2002).

Des expériences de co-transfection ont montré que l'expression des protéines N et P était suffisante pour induire la formation de corps d'inclusion cytoplasmiques similaires à ceux observés durant une infection par le VRSH (Garcia et al. 1993). L'ARN génomique viral est associé à ces corps d'inclusion lorsqu'ils sont encore petits (Santangelo et al. 2006; Lifland et al. 2012). Des protéines cellulaires sont également retrouvées associées à ces corps d'inclusion telles que la protéine chaperon Hsp70 ainsi que l'actine qui colocalisent avec les protéines du complexe polymérase (Brown G. et al. 2005). Des expériences d'immunoprécipitation montrent que la protéine Hsp70 interagit avec le complexe ribonucléoprotéique (RNP), mais son rôle n'est pas élucidé. D'autre part, l'équipe de Philip Santangelo à l'institut de technologie de Georgia a mis en évidence la présence des protéines MDA5, MAVS et RIG-I dans les corps d'inclusion cytoplasmiques lors d'une infection par le VRSH ou de la co-transfection des protéines N et P dans des cellules Hep-2 eucaryotes (Lifland et al. 2012). Ces trois protéines ayant un rôle majeur lors de la réponse antivirale, le virus les séquestrerait dans les corps d'inclusion afin d'inhiber la production d'IFN. Par analogie avec les corps de Negri des *Rhabdoviridae* (Lahaye et al. 2009), ces corps d'inclusion pourraient constituer les sites actifs de réplication et de transcription du VRS au début de l'infection (Heinrich et al. 2010; Lifland et al. 2012).

2.3.5 Assemblage et bourgeonnement

L'assemblage des nouvelles particules virales survient suite à la transcription et à la réplication. Ce phénomène s'effectue en deux étapes distinctes. Premièrement, les protéines de surface F, G et SH sont synthétisées dans le réticulum endoplasmique, subissent leurs dernières modifications dans le Golgi et sont transportées à la membrane plasmique. Le mécanisme par lequel les protéines virales s'assemblent à la surface de la cellule n'est cependant pas encore élucidé. Deuxièmement, les sous-unités de la protéine N s'associent à l'ARN génomique, puis au complexe polymérase (P-L) dans le cytoplasme de la cellule. Par la suite, le transport de la nucléocapside nouvellement formée à l'endroit approprié où se trouvent les protéines virales dans la membrane plasmique serait effectué par la protéine de

matrice M. Finalement, le virus acquiert son enveloppe en bourgeonnant à la surface de la cellule et les nouvelles particules virales sont libérées (Figure 10). Dans les cellules polarisées, les virus bourgeonnent seulement à la partie apicale, réduisant ainsi la dissémination du virus à partir de l'épithélium des voies respiratoires.

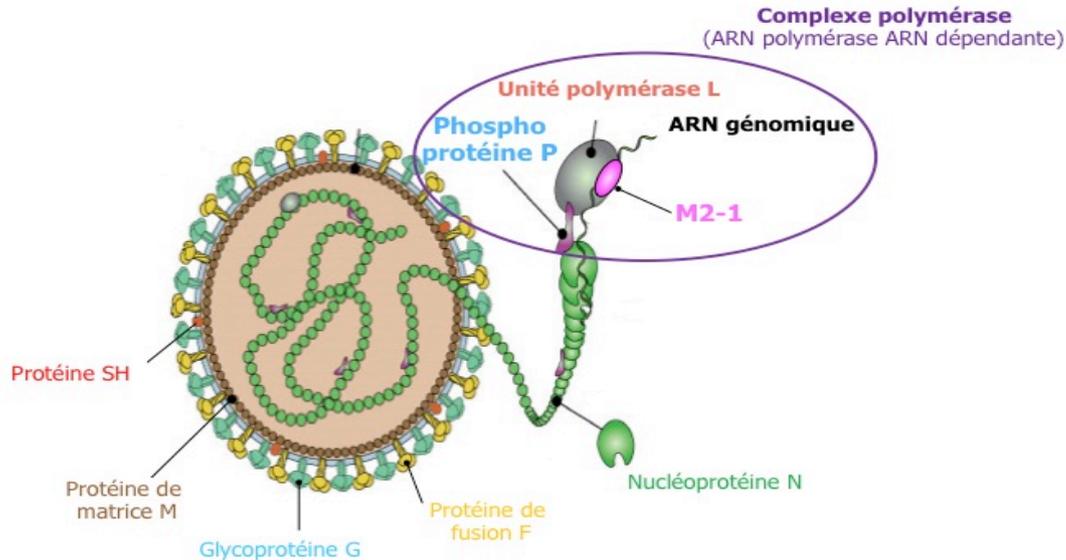


Figure 10 : Schéma d'une particule virale du VRS. Modifié à partir de http://education.expasy.org/images/Metapneumovirus_virion.jpg.

2.4 Les protéines virales

Comme indiqué précédemment, les 10 gènes du VRS permettent la transcription de 11 ARN messagers qui, eux-mêmes sont traduits en 11 protéines (le gène M2 codant pour 2 protéines M2-1 et M2-2). Deux d'entre elles sont dites non structurales NS1 et NS2, car elles ne sont pas retrouvées dans les particules virales. Parmi les neuf protéines restantes, quatre (F, G, SH et M) vont constituer l'enveloppe virale et cinq autres (N, P, L, M2-1 et M2-2) participent à l'activité du complexe polymérase.

2.4.1 Les protéines non structurales NS1 et NS2

Les gènes les plus abondamment transcrits du VRS codent pour les protéines non structurales NS1 et NS2. Ces protéines sont spécifiques des *Pneumovirus*, avec aucune homologie apparente pour une autre protéine. Il s'agit de deux petites protéines de 139 et 124 acides aminés pour NS1 et NS2 respectivement. Elles ont des effets sur plusieurs protéines appartenant à la voie de signalisation de l'interféron, mais on ne sait pas actuellement quelles molécules dans ces voies interagissent directement avec ces protéines (Bossert et al. 2003; Spann et al. 2005; Swedan et al. 2009). Elles jouent un rôle dans le mécanisme d'évasion du

système immunitaire et la spécificité d'hôte. En outre leur expression semble jouer un rôle dans l'inhibition de l'apoptose précoce au cours de l'infection par le VRS.

Pour l'instant rien n'a été publié sur les structures tridimensionnelles de ces protéines ou les mécanismes moléculaires par lesquels elles exercent leurs fonctions. Toutefois, des études de co-immunoprécipitation dans des cellules eucaryotes ont montré que les protéines NS sont capables d'interagir pour former des homodimères et hétérodimères (Swedan et al. 2009; Swedan et al. 2011). Des études ont été faites portant sur l'analyse de la stabilité de ces protéines, les changements conformationnels menant à la formation des oligomères sphériques solubles de type amyloïde ou avec des structures répétitives en feuillet β et aussi l'équilibre conformationnel entre l'état monomère et l'état oligomère (Pretel et al. 2013). Les protéines NS1 et NS2 ne semblent pas essentielles à la réplication virale *in vitro* puisque la délétion simple ou double des gènes codant les protéines NS1 et NS2 n'est pas létale pour le VRS. Un virus comportant une de ces délétions pousse plus lentement en culture cellulaire (Teng and Collins 1999) (Schlender et al. 2000) et donne un phénotype atténué par rapport à la souche sauvage lors d'une infection chez le chimpanzé (Whitehead et al. 1999) (Teng et al. 2000).

2.4.2 Les protéines d'enveloppe

Les particules virales des Pneumovirus comportent une enveloppe qui est composée de trois protéines transmembranaires : la protéine de fusion (F), la glycoprotéine d'attachement (G) et la petite protéine hydrophobe (SH). Ces trois protéines s'associent en homo-oligomère et les protéines F et G forment des projections en spicule à la surface de l'enveloppe virale.

2.4.2.1 La protéine de fusion F

Le gène F code pour une protéine transmembranaire de 574 acides aminés. Le rôle de la protéine F chez les paramyxovirus est de favoriser la fusion de l'enveloppe virale avec la membrane de la cellule hôte afin que le virus puisse pénétrer à l'intérieur et se répliquer. Récemment, il a été mis en évidence une interaction spécifique entre F et la nucléoline permettant l'entrée du VRS dans les cellules hôtes *in vitro*. La nucléoline jouerait ainsi le rôle de récepteur cellulaire du VRS (Tayyari et al. 2011). La protéine de fusion est d'abord synthétisée en un précurseur F₀. Il est clivé dans le trans-Golgi par des protéases cellulaires afin de former un hétérodimère F2-F1 lié par des ponts disulfures (Figure 11).

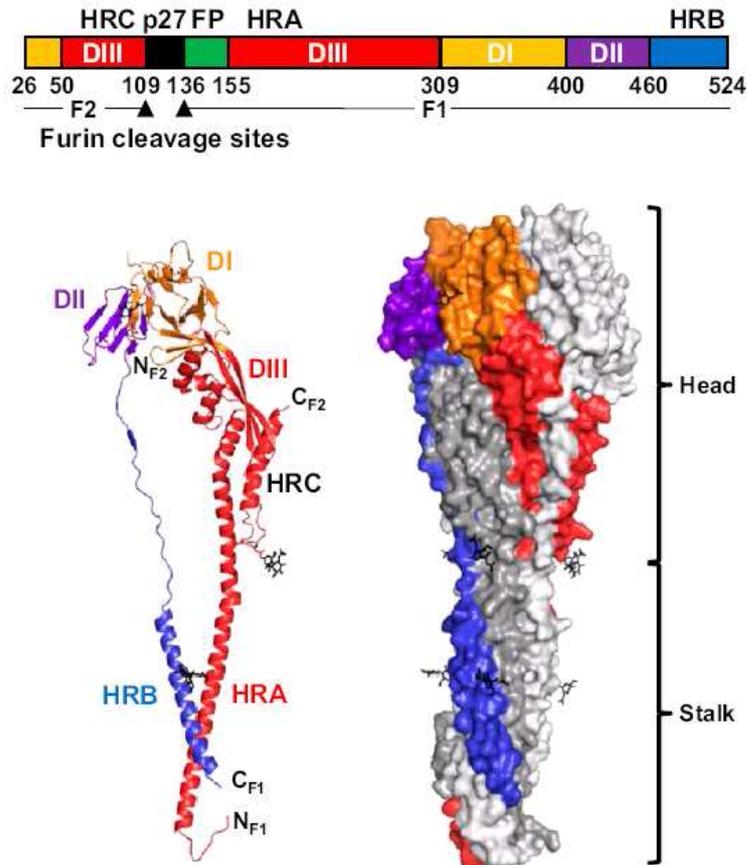


Figure 11 : Structure cristallographique de l'ectodomaine de la protéine F du VRSH sous la forme postfusion avec une résolution de 3.2 Å (A) Représentation schématique de la séquence primaire de l'ectodomaine de la protéine F. Le numéro des résidus délimitant chaque segment est indiqué ainsi que l'endroit des sites de clivage aux furines. DI-III, domaines I-III ; p27, peptide excisé ; FP, peptide fusion ; HRA,-B,-C, heptapeptides répétés A, B et C. (B) et (C) Représentation en ruban d'un monomère et d'un trimère de F sous la forme post-fusion. Adapté de (Swanson et al. 2011)

Chez le VRS, les deux sites de clivage sont généralement composés des acides aminés R-X-R/K-R, motif reconnu par les furines cellulaires. La protéine F transite ensuite vers la membrane plasmique de la cellule infectée où elle s'ancre sous la forme dite de pré-fusion. La forme pré-fusion subit des changements conformationnels importants pour se retrouver sous une forme post-fusion.

Deux groupes différents ont résolu récemment la structure atomique de l'ectodomaine de la protéine F du VRSH dans sa conformation post-fusion, cette étude est basée sur l'analyse d'une version de la protéine conçue pour éliminer le peptide de fusion et le domaine

transmembranaire ainsi que la queue cytoplasmique (McLellan et al. 2011) ; (Swanson et al. 2011). Celle-ci a une forme de cône similaire à celle de l'ectodomaine des protéines F du virus parainfluenza humain 3 (HPIV3) (Yin et al. 2005) et du virus de la maladie de Newcastle (Swanson et al. 2010).

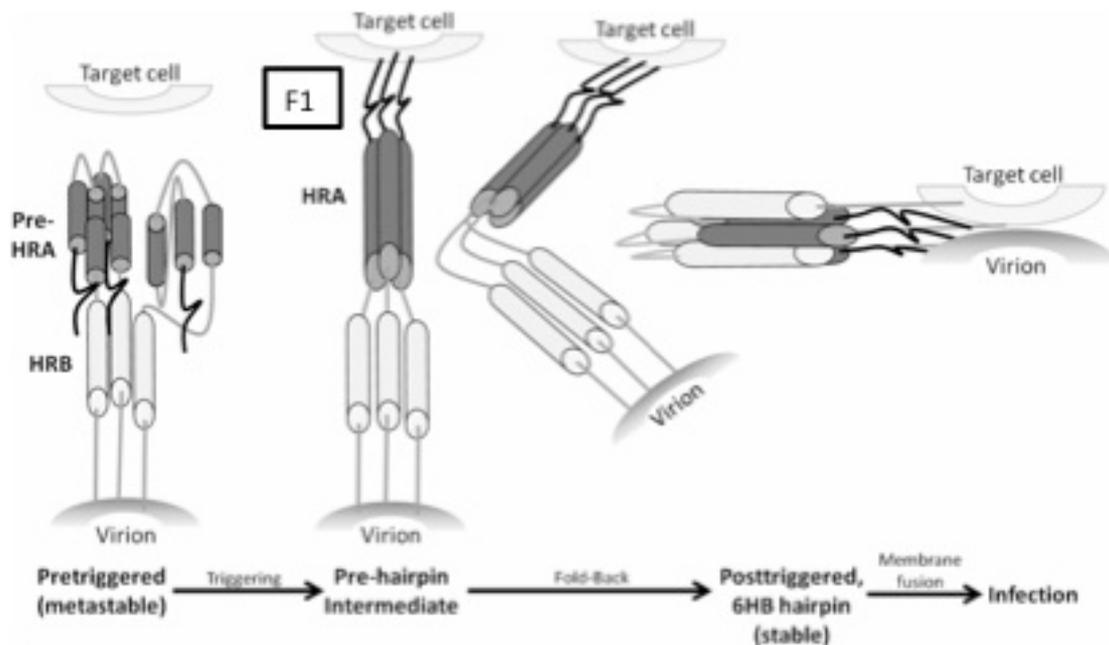


Figure 12 : Schéma des régions dynamiques de la protéine F qui se replie pour initier la fusion. Les trois hélices α en N-terminal dans le domaine pré-HRA sont reliées par deux peptides non hélicoïdaux sous une forme métastable. Lors du déclenchement, ces deux peptides se replient en hélice α , rallongeant ainsi les hélices α HRA permettant l'insertion du peptide de fusion F1 en N-terminal dans la membrane de la cellule cible. De même, le trimère de HRA (gris foncé) se plie alors en deux, et les hélices α HRB (gris clair) s'insèrent dans les sillons à la surface du trimère HRA formant un faisceau stable de 6 hélices (6HB). Par conséquent, les virions et les membranes cellulaires sont rassemblés et la fusion membranaire commence. La région de «tête» de la protéine F ne se réorganise pas pendant les événements de déclenchement et de repliement et, par conséquent, elle n'est pas représentée sur ce dessin. (Costello et al. 2012)

Le clivage du précurseur F0 est nécessaire afin que la protéine effectue sa fonction de fusion. Suite à la formation des fragments F2 et F1, le peptide de fusion adjacent au site de clivage devient alors accessible et constitue alors l'extrémité N-terminale du segment F1 (Figure 12). Ce peptide est suffisamment hydrophobe pour agir comme un domaine d'ancrage transmembranaire et serait directement impliqué dans l'insertion de la protéine F à la

membrane cellulaire. Deux motifs particuliers sur le fragment F1 sont nécessaires à la fusion virale et le rapprochement des membranes cellulaires, les heptapeptides répétés A (HRA) et les heptapeptides répétés B (HRB) (Figure 12) permettant ainsi la formation des syncytia. De cette manière, la fusion d'une cellule infectée avec une cellule non infectée permet la dissémination des virions. Dans cette même étude, les auteurs ont montré que l'immunisation de rats cotonniers avec la forme trimérique post-fusion de la protéine F était capable d'induire des taux d'anticorps neutralisants très élevés et protecteurs après une infection par le VRSH (Swanson et al. 2011). De plus, la protéine F étant essentielle au virus et hautement conservée entre souches, elle représente une bonne cible vaccinale. Elle est même considérée comme un antigène majeur pour la vaccination.

2.4.2.2 La glycoprotéine G

Le gène G code pour une glycoprotéine de 282 à 319 acides aminés selon la souche de VRS avec une masse de 32 kDa sous sa forme non glycosylée. Celle-ci a été identifiée comme étant la protéine majeure d'attachement à la cellule hôte (Levine et al. 1987). Cette découverte a été remise en cause quand il a été démontré que les virus recombinants où le gène G avait été enlevé se répliquaient aussi efficacement que le virus sauvage en culture cellulaire (Karron et al. 1997; Karger et al. 2001) et provoquaient une infection atténuée chez les souris BALB/c (Teng et al. 2001). Ces observations suggèrent que le VRS pourrait utiliser d'autres mécanismes pour s'attacher aux cellules hôtes. Toutefois, chez le VRS, il a été décrit que l'interaction entre la protéine G et les glycosaminoglycanes (GAGs) est essentielle à l'attachement de la particule virale à la membrane cellulaire. De plus, l'interaction de cette protéine avec le récepteur CX3CR1 (mastocytes, leucocytes) faciliterait l'infection et aurait un rôle crucial à jouer dans la sévérité de la pathogénèse associée à l'infection virale. De récentes études ont aussi proposé que la protéine G pourrait interagir avec la L-sélectine et l'annexine II. L'utilisation de ces composés ou d'antagonistes inhibe l'infection par le VRS *in vitro* et *in vivo* et ces résultats suggèrent que la L-sélectine et l'annexine II pourraient être des récepteurs du VRS sur les leucocytes et les cellules épithéliales.

L'analyse du profil d'hydrophobicité de la protéine G a révélé qu'il s'agirait d'une glycoprotéine transmembranaire de type II. L'extrémité N-terminale hydrophile et cytoplasmique est suivie d'un petit domaine transmembranaire hautement hydrophobe, puis d'un grand fragment hydrophile extracellulaire, constituant l'extrémité C-terminale. La

glycoprotéine G est la plus variable entre les souches puisqu'elle est souvent soumise à de fortes pressions immunologiques, certaines protéines de surface des pathogènes étant appelées à muter afin d'éviter la réponse du système immunitaire. De plus, l'analyse des séquences du VRS démontre une utilisation variable du codon stop, résultant en la traduction de protéines de différentes longueurs. On retrouve néanmoins treize résidus conservés ainsi qu'un segment contenant quatre cystéines formant un motif CX₃C. La principale caractéristique de cette protéine G est qu'elle subit de nombreuses N- et O-glycosylations au niveau de son domaine extracellulaire ajoutant au total 55 kDa supplémentaires à la protéine. De plus, elle contient un très grand nombre de résidus sérines et thréonines et la partie extracellulaire contient un nombre élevé de résidus prolines (8 %) qui sont des déterminants majeurs de la structure tridimensionnelle des protéines (Johnson et al. 1987); (Wertz et al. 1985). Cette composition particulière en acides aminés ainsi que la présence de nombreux sites potentiels de glycosylation sont des caractéristiques propres aux glycoprotéines d'origine muqueuse et suggèrent que la protéine G aurait une structure de type mucine hautement glycosylée. Actuellement, aucune analogie structurale avec les autres protéines d'attachement d'autres *Paramyxoviridae* n'a été rapportée.

2.4.2.3 La protéine hydrophobe SH

Le gène SH du VRS code pour une petite protéine de 64 acides aminés. L'analyse du profil d'hydrophobicité de la protéine SH révèle qu'il s'agirait d'une protéine transmembranaire de type II presque entièrement hydrophobe. Son rôle exact au sein de la particule virale n'est pas connu pour le VRS, mais elle est très conservée entre les souches de VRS (Collins et al. 1990; Chen et al. 2000). La protéine SH n'est ni essentielle à l'attachement, ni à l'infectivité ou à l'assemblage des particules virales en culture cellulaire.

De plus, il a été démontré que les virus mutants ne possédant pas le gène SH se répliquaient efficacement dans des modèles d'infection chez des animaux tels que la souris et le chimpanzé avec des titres inférieurs au virus sauvage de 10 à 40 fois respectivement (Bukreyev et al. 1997; Whitehead et al. 1999). Des études de modèles moléculaires suggèrent que la protéine SH se structurerait en canal ionique pouvant ainsi jouer le rôle de viroporine (Gonzalez and Carrasco 2003). Les protéines SH s'insèrent dans la membrane des cellules infectées et induisent la perméabilité aux ions et aux petites molécules. En accord avec cette hypothèse, la protéine SH augmenterait la perméabilité de la membrane lorsqu'elle est exprimée dans des bactéries (Perez et al. 1997), et lorsqu'elle est incorporée dans des membranes artificielles elle formerait des structures poreuses en forme pentamères et

hexamères ayant une activité de canal de cations (Gan et al. 2008; Carter et al. 2010) (Figure 13).

La protéine SH a été signalée pour réduire l'apoptose et moduler la réponse immune en inhibant la voie de signalisation du TNF- α , mais les effets semblent faibles (Fuentes et al. 2007).

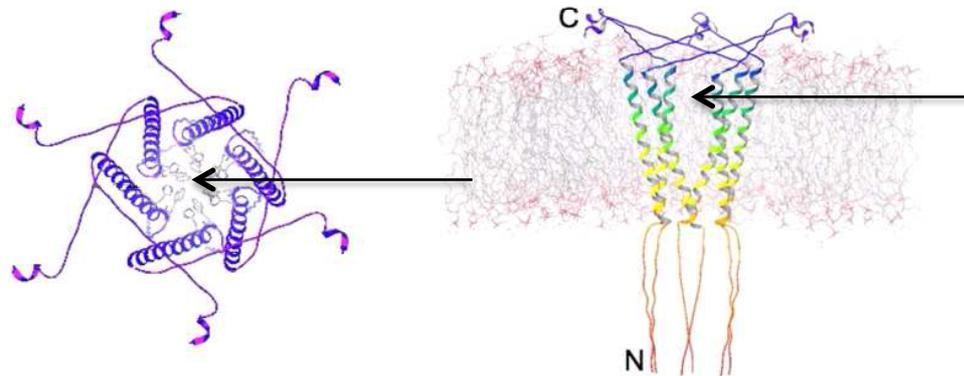


Figure 13 : Modèle de la protéine SH du VRS, basé sur la structure de la protéine SH fonctionnelle observée par microscopie électronique et des données fonctionnelles existantes (Carter et al. 2010). Le modèle moléculaire de la protéine SH hexamérique vue de dessus ou en bas insérée dans une membrane. La flèche indique le canal formé.

2.4.2.4 La protéine de matrice M

Le gène M du VRSH code pour une protéine non glycosylée de 256 acides aminés (28,7 kDa). La formation de filaments viraux et la libération de particules sont des étapes essentielles dans le cycle de vie du VRSH qui nécessitent l'assemblage d'une couche régulière de protéines M. La localisation de la protéine M varie en fonction du cycle viral. Au début de l'infection, la protéine M est retrouvée principalement dans le noyau des cellules infectées, ce qui permettrait d'inhiber la transcription de gènes cellulaires. Au contraire, elle est localisée en majorité dans le cytoplasme et particulièrement dans les corps d'inclusion durant les phases tardives de l'infection, permettant l'inhibition de la transcription du VRS et favorisant la morphogénèse et le bourgeonnement du virus (Ghildyal et al. 2003). De nombreuses interactions avec la protéine M ont été décrites dans des cellules infectées par le VRSH : avec le complexe RNP (Ghildyal et al. 2002), avec la protéine M2-1 impliquant les 110 premiers résidus en N-terminal de la protéine M (Li et al. 2008) et permettant sa localisation dans les corps d'inclusion, avec la protéine G (Ghildyal et al. 2005) ou avec de l'ARN (Rodriguez et al. 2004).

Une étude récente montre que la protéine M forme des dimères en solution et lorsqu'elle est cristallisée. La dimérisation et l'oligomérisation sont essentielles pour le bourgeonnement viral et la libération des nouveaux virions. La première structure cristalline (Money et al. 2009) de la protéine M du VRS donne un monomère avec deux domaines composés en majorité de feuillets β , mais cette structure a été remise en cause récemment avec une autre structure mettant en évidence la véritable interface de dimérisation (Forster et al. 2015). Celle-ci rappelle la structure des protéines M d'autres paramyxovirus illustrés dans la Figure 14 (Forster et al. 2015). En outre, le bourgeonnement et la libération de particules pseudo-virales sont empêchés dans des mutants de l'interface de dimérisation. Cette étude a surtout montré que la protéine M était aussi sous forme dimère en solution et qu'elle était biologiquement active sous cette forme (Forster et al. 2015).

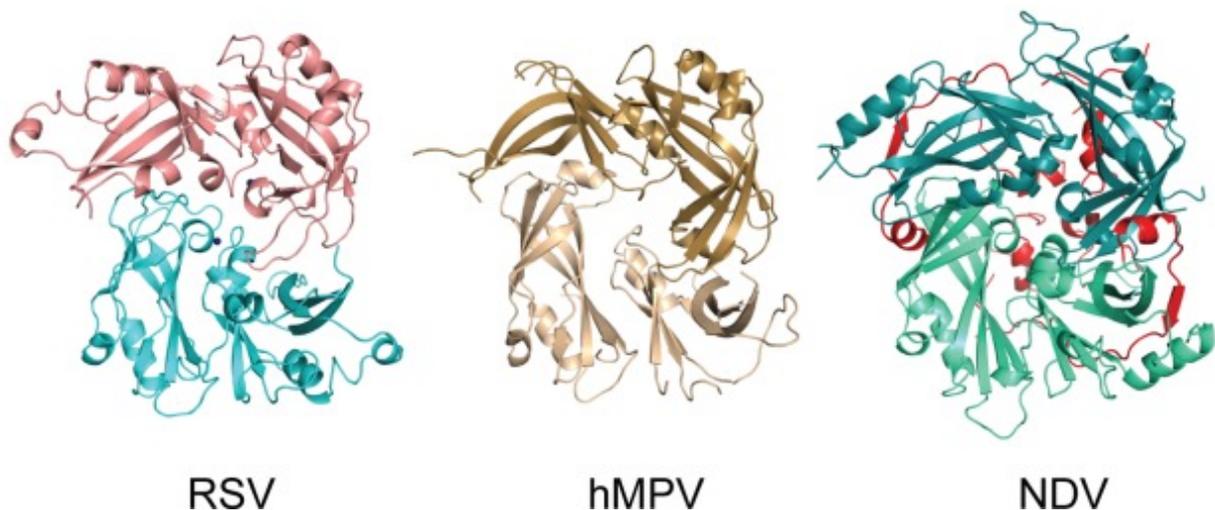


Figure 14 : Structure cristallographique de la protéine M du VRS formant un dimère avec une résolution de 1,7 Å. Les deux protomères sont en rose et en vert (PDB : 4D4T) et la similarité de structure avec la protéines M des métapneumovirus humain (MPVH) (PDB : 4LP7) et de celle du virus de la maladie de Newcastle (NDV) (PDB : 4GIG) (Forster et al. 2015).

2.4.3 Les protéines du complexe polymérase

Le complexe ARN polymérase ARN dépendant (RdRp) du VRS est responsable de la synthèse de l'ARN, messager ou génomique, à partir de la matrice d'ARN génomique, au cours des processus de transcription et répllication. Ce complexe fait intervenir un petit nombre de protéines encodées par le virus : la sous-unité catalytique L, la phosphoprotéine P

qui est un facteur essentiel de la polymérase, la nucléoprotéine N qui encapside l'ARN génomique et deux facteurs de régulation de la transcription et de la réplication, les protéines M2-1 et M2-2.

2.4.3.1 La nucléoprotéine N

La nucléoprotéine est le principal constituant protéique de la nucléocapside. Elle est présente en grande quantité dans les virions et les cellules infectées. En s'associant à l'ARN génomique viral, elle forme le complexe RNP et ainsi elle protégerait l'ARN génomique des dégradations dues aux RNases. D'autre part, la fixation de la protéine N à l'ARN empêche la formation de structures secondaires des ARN viraux néosynthétisés. Au sein des *Paramyxoviridae*, le VRSH possède la plus petite nucléoprotéine N avec 391 résidus contre 489 à 553 pour les autres membres de la famille.

Pour étudier le fonctionnement de ce complexe protéine:ARN, l'équipe de Felix Rey à l'Institut Pasteur l'a cristallisé. La Figure 15 montre comment les nucléoprotéines s'associent en oligomère de 10 ou 11 protomères en formant des anneaux, à l'aide de « bras » permettant d'associer les protomères entre eux tout le long de l'ARN (Tawar et al. 2009).

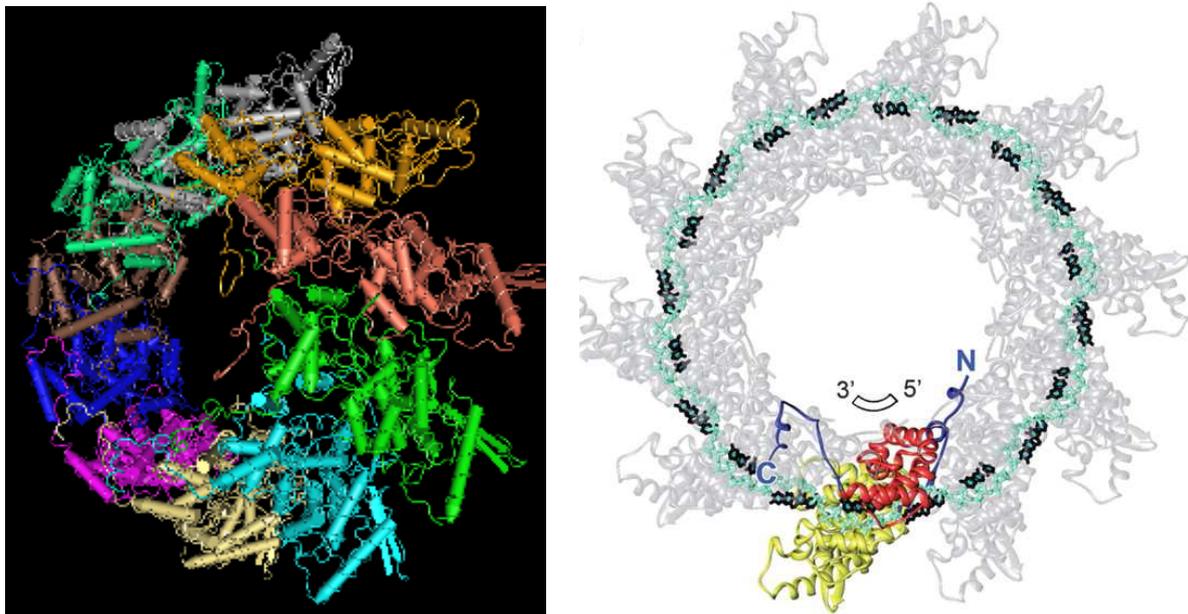


Figure 15 : Gauche : Structure radiocristallographique d'un anneau contenant 10 protomères de nucléoprotéine du VRSH avec une résolution de 3,29 Å, chaque sous-unité de protéine N étant représentée avec une couleur différente. Droite : le NTD est en jaune, le CTD en rouge et les bras N- et C-terminaux en bleu. L'ARN est en cyan et les bases en contact avec un monomère de protéine N sont en noir. PDB : 2WJ8 (Tawar et al. 2009)

Chaque protéine N du VRS est organisée en quatre domaines : un bras N-terminal (1-30), un bras C-terminal (361-391) et deux régions globulaires en N et C-terminal appelées NTD (32-252) et CTD (253-360). Le domaine globulaire C-terminal est formé de quatre hélices α . Le domaine globulaire N-terminal est constitué de 10 hélices α et d'un feuillet β . La présence d'un feuillet β dans le NTD est une spécificité structurale de la nucléoprotéine N du VRS, par rapport aux nucléoprotéines d'autres Mononégavirus comme le virus de la stomatite vésiculeuse (VSV) ou le virus de la maladie de Borna (BDV). (Tawar et al. 2009)

Les deux domaines C et N terminaux se referment autour de l'ARN comme une pince. Ces deux domaines sont séparés par une charnière flexible. Une hypothèse a été faite que lors de la multiplication virale, la pince s'ouvre et laisse uniquement passer l'enzyme, permettant une lecture de l'information génétique contenue dans la séquence d'ARN. La présence d'ARN stabilise l'interaction entre les monomères de protéine N et l'ARN viral serait ainsi toujours protégé au sein de ce complexe (Figure 15) (Tawar et al. 2009).

Comme pour l'ensemble des Mononegavirales, la protéine N complexée à l'ARN interagit avec la phosphoprotéine P. Cette interaction P:N-ARN est critique pour la reconnaissance du complexe RNP par la polymérase L. L'équipe de Jean-François Elouët à l'INRA a montré que le domaine NTD interagit directement avec l'extrémité C-terminale de la phosphoprotéine P (Tran et al. 2007; Galloux et al. 2012). La région du NTD impliquée dans cette interaction a été caractérisée et correspond à une poche constituée de résidus hydrophobes et basiques. Ce domaine d'interaction avec la protéine P a une localisation différente des protéines N des virus du genre Morbillivirus (rougeole), Respirivirus (Sendai), et Rhabdovirus (rage et VSV) pour lesquelles l'extrémité C-terminale de la nucléoprotéine interagit avec la protéine P (Krishnamurthy and Samal 1998; Mason et al. 2003) (Lahaye et al. 2009; Tran et al. 2009) (Ruigrok and Crepin 2010), mais similaire à celui décrit pour le genre Rubulavirus (oreillons) qui correspond à l'extrémité N-terminale de la protéine N (Kingston et al. 2004).

D'autre part, des études menées sur la nucléoprotéine du virus Sendai, de la rougeole et plus récemment du VSV, ont montré l'existence d'une forme néo-synthétisée de nucléoprotéine notée N_0 (Curran et al. 1995) (Spehner et al. 1997) (Leyrat et al. 2011). La protéine N_0 dépourvue d'ARN semble être maintenue dans cet état grâce à l'interaction avec la phosphoprotéine P. Cette dernière jouerait le rôle d'une protéine chaperon en empêchant la polymérisation de la nucléoprotéine N ainsi que l'encapsidation non spécifique de l'ARN de la cellule hôte.

L'équipe de Jean-François Eléouët à l'INRA a montré récemment qu'une forme mutante de N monomère, et incapable de se lier à l'ARN, interagit toujours avec P et pourrait ainsi imiter le monomère N₀ (Galloux et al. 2015). Ils ont utilisé cette N mutée, désignée N_{mono}, comme un substitut à N₀, afin de caractériser les régions de P impliquées dans la formation du complexe N₀-P. En utilisant une série de fragments de P, ils ont déterminé par GST pulldown que les extrémités N et C de P sont capables d'interagir avec N_{mono}. Ils ont analysé le rôle fonctionnel des résidus en N-terminal de P par mutagenèse ponctuelle, en mesurant l'activité de la polymérase du VRS dans un miniréplicon de VRS humain, et ils ont déterminé plusieurs résidus critiques pour la synthèse de l'ARN viral. Ces résidus critiques sont impliqués dans l'interaction entre le domaine P [1-40] et N_{mono} *in vitro*. Et enfin, ils ont montré que la surexpression du peptide P [1-29] peut inhiber l'activité de la polymérase à l'aide d'un miniréplicon de VRS. (Galloux et al. 2015)

2.4.3.2 La polymérase L, sous-unité catalytique portant les activités de transcription et de réplication virale

Deux des gènes de VRS (M2 et L) se chevauchent sur 68 nucléotides (Collins et al. 1987). Des études avec des mini-réplicons ont montré que, suite à la transcription du gène M2, la polymérase virale scanne à la fois en amont et en aval pour localiser le site de départ du gène L (Fearn and Collins 1999). Comme son nom l'indique, la protéine L (pour Large protein) est très grande puisqu'elle est constituée de 2165 résidus pour le VRSH. Par analogie avec d'autres virus, la protéine L des Pneumovirus est considérée comme le principal composant du complexe ARN-polymérase ARN-dépendante du VRS. Elle est responsable de la synthèse des ARNm (transcription), des ARN génomiques et antigénomiques (réplication) et elle assure les activités de coiffage et de polyadénylation des ARNm.

Par des alignement de séquences de différentes protéines L, Poch et collaborateurs ont décrit six domaines conservés pour les protéines de la polymérase du virus à ARN brin négatif conservés notés de I à VI dont trois domaines fonctionnels (II, III et VI), et les régions entre les domaines distincts permettant l'assemblage des domaines fonctionnels dans une conformation appropriée (Poch, Sauvaget et al. 1989; Simonin, Menissier-de Murcia et al. 1990). Le domaine I semble être impliqué dans l'interaction avec la phosphoprotéine P (Cevik et al. 2004). Le domaine II composé de résidus chargés positivement, serait responsable de la reconnaissance de la matrice ARN. Des mutations dans ce domaine conservé abolissent la synthèse d'ARN (Smallwood et al. 1999). Le domaine III contient un motif GDNQ retrouvé parmi toutes les RdRp de Mononegavirales et correspondant au site catalytique de la

polymérase. Le domaine IV contient un site putatif de liaison de l'ATP et Cartee et collaborateurs ont démontré qu'un seul changement d'acide aminé (N1049D) sur ce domaine conduisait à la terminaison de la transcription sans altérer le taux de réplication du virus (Cartee et al. 2003). Le domaine V serait impliqué dans les activités nécessaires à la coiffe des ARNm viraux (Li et al. 2008).

L'équipe de R. Fearnls a publié en 2012 la purification d'un complexe P-L fonctionnel pour le VRSH à partir de cellules d'insectes infectées par un baculovirus recombinant (Noton et al. 2012). Ils ont montré que la polymérase initiait la transcription majoritairement au nucléotide +3 et non +1 sur l'ARN matrice *in vitro*. De plus, si les deux dernières bases du génome sont mutées, la RdRp du VRS corrige et incorpore le dinucléotide AC sur le brin complémentaire en 5' (Noton et al. 2010). Ceci suggère que le RdRp pourrait être chargé en dinucléotide AC voire trinucleotide ACG avant d'entrer en 3' du génome ou de l'antigénome pour initier la synthèse d'ARN "Leader" ou "Trailer".

2.4.3.3 La phosphoprotéine P cofacteur principal de la polymérase du VRS

La phosphoprotéine du VRSH est composée de 241 AA. Elle est beaucoup plus courte que les autres phosphoprotéines de Paramyxovirus. Cette protéine est phosphorylée principalement au niveau de la sérine 232 (Barik et al. 1995) (Sanchez-Seco et al. 1995). D'autres sites de phosphorylation mineurs ont été identifiés (Navarro et al. 1991) (Asenjo et al. 2005).

Bien que le mécanisme d'action exact de la phosphoprotéine P reste incertain, il s'agit du principal co-facteur de la polymérase du VRS et elle sert de plaque tournante au cours des processus de transcription ou réplication, puisqu'elle est capable de se lier aux protéines L, N et M2-1 du VRS ainsi qu'à diverses protéines cellulaires. Le domaine d'interaction avec la protéine L sur la phosphoprotéine serait localisé sur le domaine d'oligomérisation entre les résidus 121 à 160 (Khattar et al. 2001). Le domaine C-terminal de la phosphoprotéine interagit avec le complexe N-ARN (Garcia-Barreno et al. 1996). Il a été montré par l'équipe de J-F Elouët que les 9 derniers résidus de la protéine P suffisaient pour cette interaction et que les résidus acides et hydrophobes de ce domaine étaient essentiels pour l'interaction avec la protéine N (Tran et al. 2007). Elle se lie aussi à la forme libre nouvellement synthétisée de N pour empêcher sa liaison aux ARN non viraux et pour pouvoir la libérer lors de la réplication du génome (Curran et al. 1995) (Castagne et al. 2004)

La protéine P forme des homotétramères et son domaine d'oligomérisation a été localisé entre les résidus 104 à 163 (Castagne et al. 2004) (Llorente et al. 2006). A part ce

domaine structuré en hélice α , les régions N-terminale (1-103) et C-terminale (200 à 241) semblent intrinsèquement désordonnées (Llorente et al. 2008; Simabuco et al. 2011). Ces prédictions structurales ont été confirmées par des études biochimiques et biophysiques (Esperante et al. 2012). Cette flexibilité de la phosphoprotéine est nécessaire pour lui permettre d'interagir avec ses nombreux partenaires.

2.4.3.4 Les protéines M2-1 et M2-2

Le gène M2 n'est présent que pour les virus de la sous-famille des *Pneumovirinae*. Ce gène contient deux ORF chevauchants codant pour les deux protéines M2-1 et M2-2. La protéine M2-1 est traduite en premier suivie par M2-2, la machinerie ribosomale cellulaire permettant la traduction de l'ARNm, issu de la transcription du gène M2, à partir d'un codon initiateur situé en amont du codon stop de M2-1. Le phénomène d'ORF chevauchants n'est pas unique au VRS, il s'agirait d'un mécanisme évolutif en réponse à la pression de sélection, il existe aussi pour le virus de la pneumonie de souris (PVM) et le pneumovirus aviaire (APV) (Gould and Easton 2007). Plusieurs virus ont adopté ce mécanisme et ce qui leurs permet une régulation de l'expression des gènes viraux.

- **La protéine M2-2**

C'est une protéine de 11 kDa composée de 83 à 90 AA selon le site d'initiation de la traduction puisque la deuxième ORF contient trois codons AUG en position 1, 3 et 7 (Ahmadian et al. 2000). Cette protéine virale est l'espèce la moins abondante lors de l'infection virale des cellules. L'absence de l'expression du gène M2-2, entraîne un retard dans la vitesse de croissance virale, une diminution de la réplication, mais par contre la transcription augmente. La protéine M2-2 assure l'équilibre entre la réplication et la transcription virale mais favorise aussi l'assemblage des nouveaux virions (Bermingham and Collins 1999) ce qui peut expliquer que la concentration la plus élevée de protéine M2-2 est observée tard dans l'infection (Collins et al. 1996).

- **La protéine M2-1**

C'est une protéine de 22,15 kDa composée de 194 AA et codée par le premier ORF en 5' du gène M2. Cette protéine a un rôle majeur dans le cycle viral et elle est essentielle à la formation de nouveaux virions infectieux (Collins et al. 1995). C'est une phosphoprotéine mais au cours de l'infection la forme phosphorylée est minoritaire (Lambert et al. 1988), alors que si la protéine est exprimée seule, cette forme phosphorylée est majoritaire. Ceci peut s'expliquer par le fait que lorsque la protéine M2-1 est associée au complexe polymérase, ces sites potentiels de phosphorylation ne sont plus accessibles. (Cuesta et al. 2000). Cette

protéine est un homotétramère qui se lie à la phosphoprotéine P et à l'ARN de manière compétitive (Tran et al. 2009). C'est un cofacteur du complexe RdRp essentiel pour la transcription virale mais pas pour la réplication (Collins et al. 1996). Elle possède deux activités distinctes : elle permet d'une part, l'élongation de la transcription, et d'autre part empêche la terminaison de la transcription au niveau des séquences GE.

3 Epidémiologie et traitements de l'infection par le VRSH

3.1 Epidémiologie

Le VRSH est un virus très contagieux qui peut infecter l'être humain tout au long de sa vie. L'homme est le seul hôte naturel du VRS, bien que le virus puisse facilement infecter et parfois causer une maladie grave chez les primates (Blount et al. 1956; Kondgen et al. 2008). Les infections au VRS se produisent surtout lors d'épidémies annuelles pendant les mois d'hiver dans les pays tempérés ou pendant la saison des pluies dans les pays tropicaux. La plupart des enfants sont infectés par le VRS au cours de la première année de vie et presque tous sont infectés avant l'âge de deux ans (Glezen et al. 1986). La

Figure 16 montre que les profils dynamiques des épidémies se répètent.

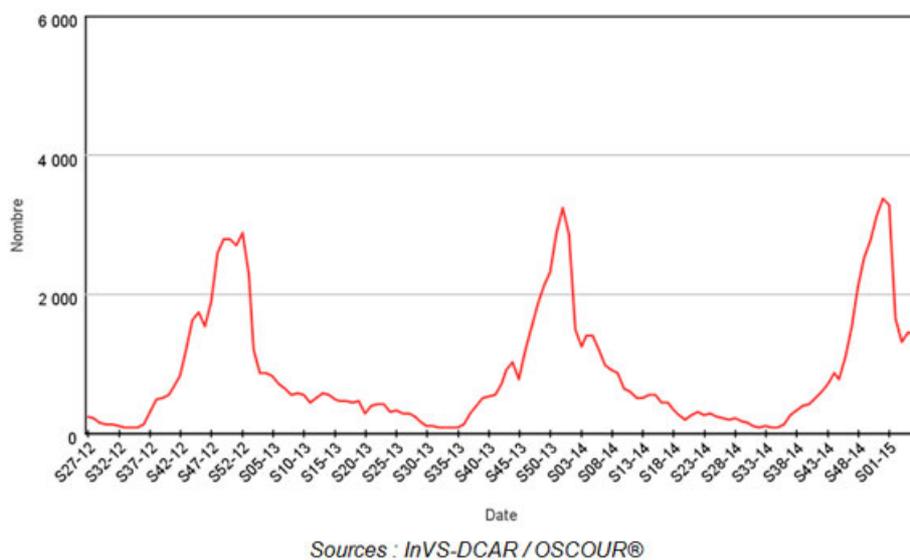


Figure 16 : Nombre hebdomadaire de recours aux urgences hospitalières pour bronchiolite des enfants de moins de 2 ans dans les 354 établissements hospitaliers participant à la surveillance entre 2012 et 2015, Données de l'institut de veille sanitaire (InVS).

La réinfection est fréquente au cours des premières années de la vie. Dans une étude prospective, des enfants qui avaient été infectés au cours de la première année de vie ont été réinfectés à 47% et 45% au cours de leur deuxième et troisième années, respectivement (Glezen et al. 1986). Les isolats de VRSH sont classés en deux groupes antigéniques (A et B) (Mufson et al. 1985) qui représentent des lignées génétiques distinctes dont la divergence, qui a été récemment calculée, semble avoir eu lieu il y a environ 350 années (Zlateva et al. 2005). La majorité de ces divergences réside dans la séquence codant pour le gène de la glycoprotéine G et en particulier au niveau de deux régions hypervariables situées dans l'ectodomaine de la protéine (Johnson et al. 1987). La protéine G est l'un des produits de gènes viraux les plus variables présents dans n'importe quel virus humain. Il est possible que cette variabilité reflète simplement l'extrême plasticité de G pour recevoir non seulement des substitutions d'acides aminés, mais aussi, d'autres changements de séquence plus drastiques, tels que des insertions, des délétions et des codons d'arrêt prématurés (Melero et al. 1997). Des génotypes multiples dans les sous-groupes A et B ont été identifiés et peuvent co-circuler dans la même saison et au sein d'une même communauté, avec un ou deux génotypes dominants (Cane 2001). De plus, une étude menée aux Etats-Unis sur des jeunes enfants hospitalisés suite à une infection par le VRSH suggère que la souche A soit responsable de symptômes plus sévères que celle du groupe B (McConnochie et al. 1990).

3.2 Traitements de l'infection par le VRSH

3.2.1 Les médicaments antiviraux

3.2.1.1 *La ribavirine*

La ribavirine est un analogue de nucléoside qui interfère avec la réplication d'un certain nombre de virus à ARN et d'ADN. Elle inhibe l'activité de l'ARN polymérase et elle a été le premier médicament homologué pour le traitement de l'hépatite C et de l'infection par le VRS chez les humains. Bien que cette molécule présente une activité puissante contre le VRS dans les cultures tissulaires et sur les animaux de laboratoire (Hruska et al. 1982), l'utilisation de la ribavirine en clinique est à présent très limitée en raison du manque d'efficacité prouvée (Rodriguez et al. 1994), de la difficulté d'administration (c'est-à-dire, le plus souvent sous forme d'aérosol), et les risques de toxicité. Actuellement, la ribavirine n'est pas recommandée pour le traitement de routine de l'infection par le VRS (2006). Elle est utilisée parfois dans le traitement de l'infection par le VRS en combinaison avec

l'immunoglobuline intraveineuse et les corticostéroïdes neutralisants chez les individus à haut risque, ceux dont le système immunitaire est affaibli par malnutrition ou sous-alimentation ou bien ceux qui ont des maladies préexistantes « telles qu'une infection par le VIH ou par la rougeole », mais leur efficacité reste controversée.

3.2.1.2 Autres molécules antivirales

Un certain nombre d'autres petites molécules anti-VRS ont été décrites au cours des 20 dernières années. La plupart d'entre elles ont été découvertes par criblage de chimiothèques de produits naturels. De façon remarquable, la plupart des inhibiteurs de VRS ciblent la protéine F (responsable de la fusion de l'enveloppe virale avec la membrane cellulaire), peut-être parce que la fusion des membranes virales et cellulaires se produit hors de la cellule et que les inhibiteurs n'ont donc pas besoin de traverser la membrane. Un exemple de ces inhibiteurs est le RS637, une petite molécule inhalable qui bloque la fusion de la protéine F à la surface de la cellule cible. D'autres molécules antivirales ont été trouvées ciblant d'autres protéines, comme la MBX-300 qui bloque la protéine G (Kimura et al. 2000) (Douglas 2004), YM-53403 ciblant la grande sous-unité polymérase L (Sudo et al. 2005) et RSV-604 bloquant la nucléoprotéine N (Chapman et al. 2007) (Challa et al. 2015). Une série de dérivés de benzimidazole interagit avec la protéine F. Malgré leur différence de structure, ils semblent tous se lier à une poche hydrophobe de la chaîne de F1 (Zhao et al. 2000).

Des nouvelles petites molécules chimiques ayant pour cible des protéines cellulaires nécessaires à la réplication du virus ont été testées comme antiviraux. De cette manière, le spectre d'action de ces molécules ne se restreint pas à un seul virus mais s'attaque à des voies métaboliques communes, utilisées par plusieurs virus (Bonavia et al. 2011).

Une autre piste pour le traitement du VRS a été étudiée, en utilisant des petits ARN interférents (siRNA), des courtes séquences d'ARN double brin pouvant se lier spécifiquement à une séquence d'ARN messagers et ainsi empêcher l'expression d'un gène en clivant cet ARN. L'ALN-RSV01 cible les deux gènes des protéines P et N ainsi que le gène de la protéine non structurale NS1. Ce traitement a été associé à une réduction de l'infection du VRS chez des adultes volontaires sains (DeVincenzo et al. 2008) et une amélioration de la fonction pulmonaire (Zamora et al. 2011).

3.2.2 Immunoprophylaxie avec des anticorps

L'immunoprophylaxie systémique avec des anticorps neutralisant le VRS fournit une protection substantielle contre l'infection à VRS grave chez les nourrissons et les enfants à

haut risque ((2009); (Mejias and Ramilo 2008);(Wu et al. 2007)). Le premier produit à usage humain est une immunoglobuline intraveineuse (IGIV-VRS, RespigamTM, produite par MedImmune), composée d'anticorps sériques purifiés à partir de donneurs dépistés à haute activité neutralisante du VRS. Il a été autorisé en 1996. Une administration mensuelle pendant la saison du VRS réduit la fréquence d'hospitalisation pour le VRS de 55%, ou plus ainsi que les jours passés en soins intensifs de 97% ((1997);(Groothuis et al. 1993); (Mejias and Ramilo 2008)). Cependant, l'inconvénient de ce produit était le mode d'administration (perfusion intraveineuse) et aussi le risque possible d'interférence avec les autres vaccins pour les enfants en raison de son caractère polyvalent (Wu et al. 2007).

Ce produit a été remplacé par le palivizumab qui est un neutralisant du VRS, développé par humanisation d'un anticorps monoclonal murin (Beeler and van Wyke Coelingh 1989);(Johnson et al. 1997);(Wu et al. 2007)). Licencié en 1998, cet anticorps monoclonal est de 50 à 100 fois plus efficace sur une base de poids que l'IGIV-VRS, permettant l'administration d'un volume beaucoup plus faible par injection intramusculaire mensuelle. Son efficacité clinique est similaire à celle de l'IGIV-VRS (1998). L'IGIV-VRS n'était pas recommandée pour une utilisation chez les nourrissons ayant une maladie cardiaque, en raison du volume important administré (Groothuis et al. 1993), alors que le palivizumab est efficace dans cette population (Feldes et al. 2003). Le palivizumab est largement utilisé pour la prophylaxie des nourrissons et des jeunes enfants à haut risque d'infection sévère par le VRS, dû à la prématurité et/ou une maladie sous-jacente, et peut également être utilisé chez les receveurs de cellules souches hématopoïétiques (Boeckh et al. 2001). Cet anticorps monoclonal représente ainsi le premier et unique moyen disponible pour réduire spécifiquement l'incidence de l'infection sévère par le VRS. Le palivizumab a été modifié par maturation d'affinité *in vitro*, qui permet d'augmenter l'affinité respective des lymphocytes B pour leur antigène et de sélectionner les meilleures immunoglobulines, pour créer un dérivé plus puissant, appelé motavizumab (MEDI-524 ou Numax, MedImmune) (Wu et al. 2007).

Dans une étude clinique portant sur 6635 nouveau-nés prématurés, le groupe ayant reçu du motavizumab présentait 26% de réduction des hospitalisations dues au VRSH par rapport au palivizumab (Carbonell-Estrany et al. 2010); (Gill and Welliver 2009). En 2010, le comité consultatif américain des médicaments antiviraux a refusé de soutenir cette demande de médicament, et son développement dans la prophylaxie de VRS n'est pas poursuivi. Ces produits n'empêchent pas une infection par le VRSH, mais réduisent suffisamment la réplication virale dans les voies respiratoires inférieures pour atténuer les symptômes.

Toutefois, du fait de son coût (environ 5000 euros/patient), ce traitement est restreint aux enfants à risque présentant des maladies respiratoires chroniques ou aux bébés prématurés (Harkensee et al. 2006). Le Tableau 1 récapitule les différents candidats de médicaments jusqu'à 2010 contre le VRS ainsi que leurs protéines cibles (Empey et al. 2010).

Protéine cible	Candidat	mécanisme
Protéine F	Motavizumab	Anticorps-monoclonal
	MEDI-524	Inhibiteur de fusion
	TMC-353121	Inhibiteur de fusion
	BMS-433771	Inhibiteur de fusion
	RFI-641	Inhibiteur de fusion
	VP-14637	Inhibiteur de fusion
	BTA9881	Inhibiteur de fusion
Protéine G	MBX-300	Inhibiteur d'attachement
	mAb 131-2G	Anticorps monoclonal
Nucléoprotéine N	ALN-RSV01 RSV-604	siRNA (petit ARN interféron) inhibiteur de protéine
Protéine polymérase L	YM-53403	Dérivé de la benzazépine

Tableau 1 : Les différents candidats pour des médicaments contre le VRS et leurs protéines cibles (Empey et al. 2010).

3.2.3 Les vaccins

Le VRS affectant principalement les nourrissons (et les jeunes veaux pour la forme bovine), l'idéal est donc de vacciner les nouveau-nés dès leurs premières semaines de vie. Malheureusement, l'immunisation à un âge précoce se heurte à l'immaturation du système immunitaire des individus ainsi qu'à la présence d'anticorps maternels qui neutralisent les épitopes présents dans le vaccin (Crowe and Williams 2003). La difficulté dans la vaccination contre le VRS réside donc dans la capacité du vaccin à stimuler une réponse immune efficace, équilibrée et durable durant les premiers mois de vie. Ceci nécessite plusieurs doses de vaccin données dans le cadre du programme de vaccination de routine à 2, 4 et 6 mois d'âge.

Le développement d'un vaccin pédiatrique contre le VRS a été compliqué par l'expérience avec un vaccin inactivé contre le VRS à la formaline (FI-VRS) évalué dans les années 1960 chez les nourrissons et les jeunes enfants. Ce vaccin, un virus inactivé au formol concentré avec un adjuvant à l'hydroxyde d'aluminium, a été administré par voie intramusculaire à des enfants âgés de six mois à neuf ans. À la suite d'une exposition naturelle au VRS, 80 % des enfants vaccinés ont dû être hospitalisés contre 5 % dans le groupe

contrôle, et deux sont morts (Kim et al. 1969). Une analyse ultérieure a indiqué que la réponse immunitaire des sujets vaccinés au FI-VRS était nettement différente de celle induite par une infection au VRS naturel, entraînant une mauvaise induction d'anticorps sériques neutralisant le virus (Murphy et al. 1988) et une réponse exagérée de lymphocytes T CD4 + (Kim et al. 1976). Ces anticorps sont incapables de limiter efficacement la réplication virale, mais ils se lient aux antigènes et créent des complexes anticorps-antigène qui activent le complément et génèrent la maladie (Melendi et al. 2007). De plus, la fixation au formol a créé des groupes carbonyle qui améliorent la réponse Th2 (Moghaddam et al. 2006).

En résumé, les réponses aberrantes au vaccin FI-RSV sont dues à l'induction d'anticorps qui se lient aux antigènes viraux mais ne neutralisent pas l'infection, ce qui génère une mauvaise stimulation des cellules immunitaires et des voies de signalisation intracellulaire importantes dans la régulation immunitaire, et aussi l'augmentation de la stimulation des cellules Th2. Trois types de vaccins sont à l'étude : les vaccins sous-unitaires, les vaccins atténués construits par la technique dite de génétique inverse et les vaccins vectorisés.

3.2.3.1 Les vaccins sous-unitaires

Les vaccins sous-unitaires sont en cours de développement en santé humaine. Deux approches sont envisagées. La première utilise la protéine F, épitope majeur pour la production d'anticorps neutralisants. Une combinaison des voies d'administration parentérale et intranasale semble nécessaire pour obtenir une réponse protectrice satisfaisante (Meyer et al. 2008) La seconde approche est basée sur l'élaboration d'une protéine recombinante comprenant l'épitope majeur de la région centrale de la protéine G. L'efficacité des candidats vaccins sous-unitaires contre le VRSH a été testée avec plus ou moins de succès soit avec un modèle de rongeur (de Waal et al. 2004), soit sur adultes et enfants âgés de plus de 12 mois. Mais n'était pas très immunogène (Power et al. 2001) et a été associé à des réactions d'hypersensibilité dans certains cas.

Plus récemment, comme déjà indiqué, une forme post-fusion de la protéine F a été produite avec la suppression des grandes régions hydrophobes (McLellan et al. 2011) (Swanson et al. 2011). Au vu des résultats prometteurs sur les rongeurs, l'amélioration de ce vaccin pourrait donner un vaccin sous-unitaire efficace pour le VRSH.

3.2.3.2 *Les vaccins obtenus par génétique inverse*

L'utilisation de la génétique inverse a permis d'identifier certains facteurs de pathogénicité virale du VRSH et les cibles de nouveaux vaccins. Les souches vaccinales les plus prometteuses pour l'obtention d'un vaccin contre le VRSH ont en effet été obtenues par introduction dans le génome viral de mutations spécifiques d'atténuation ou de délétion de gènes pathogènes non essentiels (Whitehead et al. 1999). Plusieurs de ces nouvelles souches de VRS atténué ont été évaluées dans des essais cliniques (Karron et al. 2005) (Wright et al. 2006). Le candidat le plus prometteur, appelé rA2cp248 / 404 / 1030_SH ou MEDI-559, contient un certain nombre de mutations, comprenant deux substitutions d'acides aminés dans la L-polymérase, une substitution dans le gène du signal de démarrage M2, et une délétion du gène SH. Mais il est immunogénique, et provoque des congestions nasales qui génèrent un problème d'alimentation chez les jeunes enfants (Wright et al. 2000) . Un autre vaccin candidat potentiel contient une délétion du gène NS1 (Teng et al. 2000).

Le candidat vaccin le plus prometteur à ce jour est le virus dérivé de l'ADNc (rA2cp248/404/1030 Δ SH) (Karron et al. 2005) qui combine cinq éléments atténuants indépendants. Actuellement une deuxième version, presque identique de ce virus, appelé MEDI-559, est évaluée dans une étude clinique de phase 1/2 (identificateur ClinicalTrials.gov NCT00767416). En outre, un candidat vaccin dans lequel le cadre de lecture ouvert M2-2 a été supprimé est également en cours d'évaluation dans une étude de phase 1 (identificateur ClinicalTrials.gov NCT01459198). (Lin et al. 2006). L'atténuation peut impliquer divers mécanismes. Des mutations ponctuelles dans la polymérase L ou dans le signal de démarrage (GS) sont susceptibles de réduire la synthèse de l'ARN viral. La délétion de M2-2 favorise la transcription au détriment de la répllication de l'ARN. Enfin la suppression de NS1 réduit la capacité du virus à inhiber les interférons (IFN) du système de l'hôte. Donc, il serait intéressant de caractériser les propriétés de ces vaccins chez les humains, mais le vrai défi reste de trouver le juste équilibre entre l'atténuation, la sécurité, la stabilité et l'immunogénicité.

3.2.3.3 *Les vaccins vectorisés*

C'est une autre stratégie de vaccin atténué qui utilise par exemple un virus de parainfluenza (PIV) atténué comme vecteur et qui permet d'exprimer en plus les protéine F et / ou G du VRS en ajoutant un ou deux gènes. Ceci permet d'obtenir un vaccin à la fois contre le vecteur PIV et le VRS (Schmidt et al. 2002) (Tang et al. 2008). D'autres approches de

vaccins vectorisés, à base d'adénovirus (Shao et al. 2009) ou d'alphavirus (Elliott et al. 2007), sont également en cours d'élaboration pour une éventuelle utilisation pédiatrique. Ces vaccins ne doivent pas utiliser des vecteurs qui seraient des agents pathogènes pour l'être humain (Gonzalez et al. 2000). De nombreuses approches pour obtenir un vaccin contre le VRSH sont en cours d'évaluation dans des études pré-cliniques, mais à l'heure actuelle, aucun vaccin anti-VRSH n'est commercialisé.

3.2.3.4 Vaccins pour la forme bovine du VRS

A l'inverse du VRSH, de nombreux vaccins contre le VRSB sont disponibles. Depuis 2000, un vaccin inactivé et de nouveaux vaccins atténués ont été mis sur le marché en Europe et sont distribués à raison de plusieurs millions de doses sans que les systèmes de pharmacovigilance ou les données du terrain n'aient enregistré de cas graves d'accidents liés à un défaut d'innocuité de ces vaccins. Cela suggère que ces vaccins inactivés et atténués sont sûrs quand ils sont administrés correctement dans le cadre de l'autorisation de mise sur le marché (AMM). Ces vaccins protègent cliniquement les animaux et permettent de limiter la circulation du virus. Les deux types de vaccins nécessitent deux injections intramusculaires à trois semaines d'intervalle afin d'optimiser la réponse immunitaire, mais l'immunité est de courte durée, environ 4-5 mois. Cinq formulations commerciales sont proposées en France dont quatre utilisent la voie parentérale, et un vaccin utilise la voie intranasale. Les vaccins atténués utilisent des souches isolées en 1969 et le vaccin inactivé est une souche virale isolée en 1990. Les vaccins inactivés utilisés chez les bovins aujourd'hui, emploient des adjuvants (saponine Quillaja) qui orientent la réponse immunitaire vers un type Th1 (Meyer et al. 2008). Pour certains vaccins inactivés, la réponse humorale paraît satisfaisante selon le rapport anticorps neutralisants sur anticorps totaux effectué grâce à une technique immunoenzymatique, même s'il reste inférieur à celui obtenu avec les vaccins vivants (Ellis 1992; Ellis 1995). La réponse immunitaire à médiation cellulaire n'a été évaluée que par des tests de prolifération lymphoblastique. La comparaison de vaccins inactivés commercialisés ou non et de vaccins vivants modifiés mono ou polyvalents commercialisés montre qu'une réponse cellulaire est induite dans tous les cas (Ellis 1992; Ellis 1995).

4 L'objectif de ma thèse

Comme présenté dans les paragraphes précédents, le virus respiratoire syncytial (VRS) est le principal agent responsable de maladies respiratoires graves chez les bovins ainsi que chez l'homme, causant à lui seul plus de 70% des bronchiolites chez les deux espèces. Il n'existe toujours pas de traitement préventif ou curatif efficace pour l'homme. Certains laboratoires essaient de développer des vaccins, mais il existe des limitations à cette approche, liées principalement au fait que ce sont des enfants en très bas âge, chez qui le système immunitaire n'est pas encore mature, qui sont principalement concernés. L'alternative consiste à privilégier la voie du traitement curatif, une voie pour empêcher la multiplication du virus. Le laboratoire où j'ai préparé ma thèse s'intéresse à la caractérisation fonctionnelle et structurale de certaines protéines du VRSH, impliquées dans cette activité et la façon dont elles s'imbriquent les unes dans les autres. La connaissance fine de ces interactions permettrait de proposer des molécules qui empêcheraient cet échafaudage et donc bloquer les processus moléculaires associés.

Le complexe ARN polymérase ARN dépendant (RdRp) du VRS est une cible de choix puisqu'il est spécifique et sans équivalent chez l'hôte. Il a aussi été établi que le bon fonctionnement du complexe RdRp faisait intervenir un petit nombre de protéines encodées par le virus : la sous-unité catalytique L, la phosphoprotéine ou protéine P qui est un facteur essentiel de la polymérase, la nucléoprotéine N qui encapside l'ARN génomique et deux facteurs de régulation de la transcription/réplication, les protéines M2-1 et M2-2. Mon travail a consisté à approfondir les connaissances structurales de deux cibles potentielles pour des molécules antivirales au sein du complexe ARN polymérase ARN dépendant (RdRp).

La première cible sur laquelle j'ai travaillé est la protéine M2-1 afin de mieux comprendre le fonctionnement de ce facteur d'antiterminaison de la transcription au sein du complexe polymérase et du cycle viral. La résolution de la structure du domaine globulaire isolé de la protéine M2-1 avait déjà été résolue par RMN au laboratoire (Blondot et al. 2012). Dans un premier temps, mon objectif consistait à déterminer finement les interactions entre la protéine M2-1 et l'ARN d'une part, et la phosphoprotéine P d'autre part. Dans un second temps, l'objectif a été d'acquérir une vision globale de la structure de M2-1. Au vu de la taille de la protéine, qui est limitante pour une technique telle que la RMN, nous avons choisi une approche où des fragments correspondant aux différents domaines sont isolés et étudiés séparément, un domaine N-terminal de liaison au zinc et un domaine de tétramérisation, pour déterminer l'agencement des domaines entre eux dans la protéine entière.

La deuxième cible était la phosphoprotéine, qui est un facteur essentiel pour la polymérase, il intervient au cours de la réplication ainsi que la transcription. Notre approche consistait à établir les bases structurales des interactions de la phosphoprotéine à partir d'informations structurales à l'échelle atomique, obtenues par résonance magnétique nucléaire en raison de son caractère intrinsèquement désordonné.

Et enfin, je me suis intéressée aux deux protéines non structurales 1 et 2 (NS1 et NS2), qui jouent un rôle dans le mécanisme d'évasion de système immunitaire et la spécificité d'hôte. Les protéines NS1 et NS2 sont en effet capables de cibler différents acteurs de la voie de signalisation de l'IFN (IRF3, NF-KB, STAT2...). Cependant les mécanismes moléculaires restent encore mal caractérisés (Bossert et al. 2003; Spann et al. 2005) et aucune structure tridimensionnelle des protéines NS du VRS n'a été résolue. Pour cette troisième cible, nous avons travaillé en collaboration avec le Dr Gonzalo Prat-Gay, à l'institut Leloir à Buenos Aires en Argentine, qui nous a fourni le plasmide de NS1-H. Des travaux préliminaires ont été effectués par RMN ainsi que des tests de criblage de la thermostabilité par fluorescence. Ce travail effectué pourra apporter des informations sur la protéine NS1 et sa stabilité afin de trouver les bonnes conditions pour une étude ultérieure par RMN.

Chapitre 2 : Etude structurale par RMN de la protéine M2-1 du VRS

1 Objectifs

La protéine M2-1, spécifique des *Pneumovirinae*, est un cofacteur du complexe RdRp essentiel pour la transcription virale du VRS (Collins et al. 1996). Sa fonction principale est d'empêcher la terminaison de la transcription de l'ARN génomique par la polymérase virale.

L'objectif premier de ma thèse était d'acquérir une vision globale de la structure de M2-1 afin de mieux comprendre certains mécanismes mis en jeu lors de la transcription/réplication du VRS et d'explorer les possibilités qu'offrirait M2-1 en tant que cible thérapeutique pour des inhibiteurs.

L'organisation en domaines de cette protéine va de pair avec une grande multiplicité de partenaires viraux et cellulaires. En septembre 2012, lorsque j'ai débuté ma thèse, cette protéine était caractérisée d'un point de vue biochimique à l'INRA dans l'équipe du Dr Jean-François Eléouët, avec qui nous avons collaboré (Tran et al. 2009). La structure du domaine central de M2-1 venait d'être résolue par RMN au laboratoire. En revanche, aucune structure tridimensionnelle n'était décrite dans la littérature pour la protéine entière, ni pour les autres domaines. Des cristaux avaient été obtenus par nos collaborateurs pour la protéine entière, mais les données de diffraction des rayons X n'étaient pas de qualité suffisante. Or ils avaient aussi montré que si le domaine central contenait la région d'interaction avec la protéine P du VRS et des ARN, le domaine N-terminal et le domaine d'oligomérisation étaient tous les deux essentiels à l'activité de la protéine entière (Tran et al. 2009).

Pour essayer de proposer un modèle structural complet de la protéine M2-1 entière sous forme de tétramère, nous avons privilégié une approche où des fragments correspondant aux différents domaines sont isolés et étudiés séparément, au vu de la taille de la protéine, qui pourrait être limitante pour une technique telle que la RMN. Nous avons envisagé de compléter l'étude structurale des domaines séparés menée par RMN par des données de plus basse résolution, par exemple de diffusion des rayons X aux petits angles, pour assembler les domaines et déterminer l'agencement entre eux dans la protéine entière.

2 Résultats antérieurs à ma thèse

2.1 Organisation en domaines de M2-1

La protéine M2-1 est constituée de 194 acides aminés. Elle se décompose en quatre domaines distincts prédits par bio-informatique (Figure 17) et validés par des approches biochimiques utilisant le pontage avec le glutaraldéhyde et des expériences de dissociation et de dénaturation chimique ou des approches biophysiques utilisant la diffusion de la lumière, statique ou dynamique, le dichroïsme circulaire ou la fluorescence (Tran et al. 2009; Esperante et al. 2011). De plus il a aussi été montré que M2-1 forme des tétramères et est capable de lier le zinc (Esperante et al. 2011).

L'organisation en domaines de M2-1 s'articule de la manière suivante :

- 1) une région présentant un motif "3Cys-1His" (résidus 1-32) (Hardy and Wertz 2000; Zhou et al. 2003) qui pourrait être un doigt de zinc, par analogie avec le facteur de transcription VP30 du virus Ebola, un homologue structural et fonctionnel de M2-1 (Modrof et al. 2003; Blondot et al. 2012)
- 2) une région d'oligomérisation (résidus 33-58) responsable de la formation de tétramères de M2-1 *in vitro* (Tran et al. 2009),
- 3) un domaine globulaire central (résidus 75-171), hélicoïdal, responsable des interactions avec l'ARN et la phosphoprotéine P (Tran et al. 2009; Blondot et al. 2012), appelé "core domain",
- 4) une extrémité C terminale désordonnée.

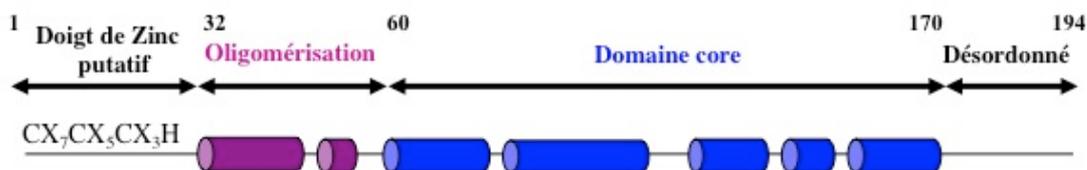


Figure 17 : Organisation en domaines de la protéine M2-1. Prédiction de la structure secondaire de la protéine M2-1 du VRSH avec PSI-PRED (<http://bioinf.cs.ucl.ac.uk/psipred/>).

2.2 Structure du domaine globulaire de la protéine M2-1

La structure du domaine central de M2-1 a été résolue par RMN au laboratoire en utilisant un fragment comportant les résidus 58 à 177, noté M2-1[58-177] (Blondot et al. 2012). Cette construction comprend une zone d'intérêt essentielle, à savoir le domaine d'interaction avec la phosphoprotéine P et avec l'ARN (Tran et al. 2009). Le domaine d'oligomérisation en est absent, de sorte que M2-1[58-177] est un monomère de 14 kDa comparé à des objets de plus de 80 kDa pour la protéine M2-1 entière tétramère. En outre, la queue C-terminale étant désordonnée, elle ne présentait pas d'intérêt pour la structure de la protéine M2-1 et pouvait même représenter un obstacle dans l'obtention d'un spectre RMN. En effet, la présence d'une telle région génère un signal intense qui se superpose à celui du domaine globulaire. De plus, la protéine M2-1 dépourvue de sa queue C-terminale est plus stable dans le temps, car elle est moins sensible à la dégradation par les protéases.

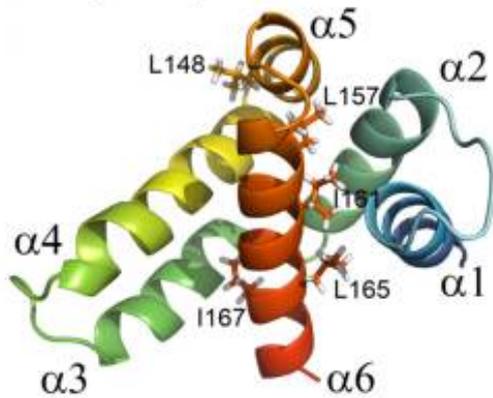
La structure en solution du domaine globulaire de la protéine M2-1 est montrée dans la Figure 18A. Ce domaine est constitué d'un faisceau d'hélices α principal ($\alpha 1$ « G75-G85 », $\alpha 2$ « K92-E105 », $\alpha 5$ « K143-R151 », $\alpha 6$ « D155-I171 ») et d'une épingle à cheveux formée par deux autres hélices ($\alpha 3$ « S108-D117 » et $\alpha 4$ « K124-K140 »). L'extrémité N-terminale de ce domaine globulaire (S58-L74) est désordonnée. Sa surface se divise en deux faces de charges opposées (Figure 18B). La face chargée positivement comporte des clusters de résidus basiques compatibles avec sa propriété de liaison aux acides nucléiques.

La résolution de la structure du fragment M2-1[58-177] a permis de mettre en évidence une homologie structurale avec le domaine C-terminal de la protéine VP30 du virus Ebola (Hartlieb et al. 2007) qui ne pouvait pas être prédite par l'alignement de séquence qui donne un score d'identité très faible de 9 %. Les deux protéines présentent une structure globulaire composée d'hélices α dont quatre sur six se superposent parfaitement. De plus, le même cluster de résidus hydrophobes est retrouvé sur les hélices $\alpha 5$ et $\alpha 6$ des deux protéines (Figure 18B). Tout comme la protéine M2-1, la protéine VP30 du virus Ebola est une phosphoprotéine constituée d'un domaine N-terminal contenant un motif doigt de zinc 3Cys-1His (CX8CX4CX3H) suivi d'un domaine d'oligomérisation. Cette protéine permet une augmentation de la transcription virale (Muhlberger et al. 1999) et elle est essentielle pour l'obtention de virions par génétique inverse (Volchkov et al. 2001). Mais le mode d'action exercée par la protéine VP30 sur la transcription est probablement différent de celui de la protéine M2-1, puisque la protéine VP30 ne semble pas agir comme un facteur

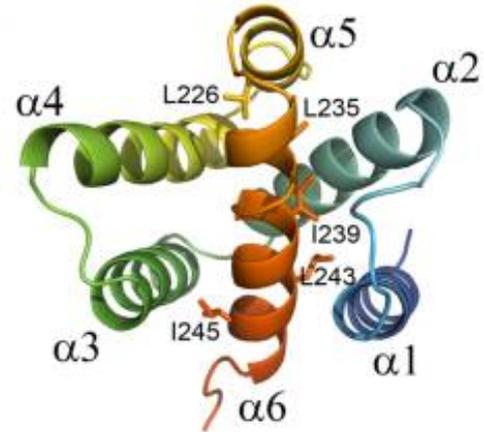
d'antiterminaison de la transcription (Weik et al. 2002). En réalité, la protéine VP30 serait impliquée dans les étapes précoces en permettant l'amorçage de la transcription virale.

A

RSV M2-1(58-177)



EBOV VP30-CTD



B

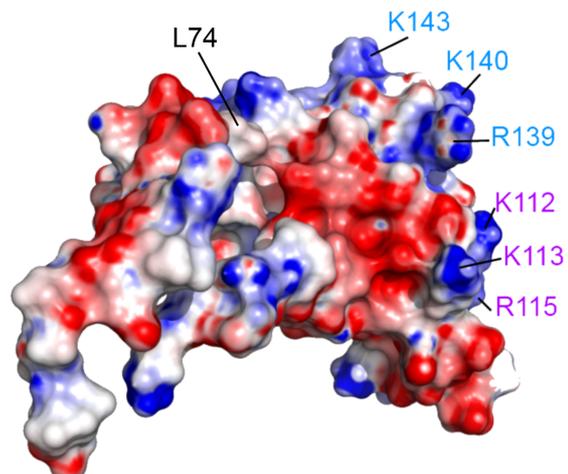
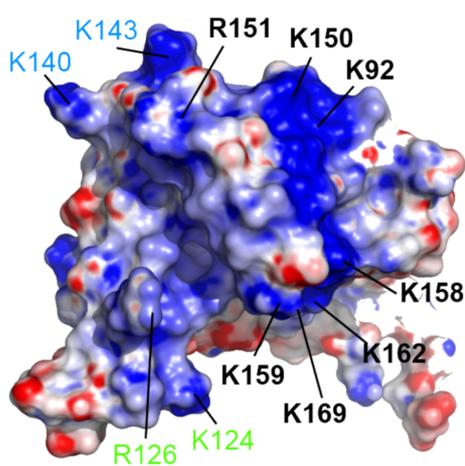


Figure 18: (A) Comparaison du domaine globulaire de M2-1 avec le domaine C-terminal de la protéine VP30 du virus Ebola (Hartlieb et al. 2007). La couleur varie du bleu vers le rouge du N- vers le C-terminal. Pour M2-1[58-177], seul le modèle RMN de plus basse énergie est donné. Les extrémités N- et C-terminales désordonnées ne sont pas représentées.

L'alignement structural entre M2-1 et VP30 a été généré sur le serveur Dali (Holm and Rosenstrom 2010). L'hélice $\alpha 7$ du domaine C-terminal de la VP30 n'est pas représentée puisqu'elle n'a pas d'équivalent dans M2-1. Les résidus hydrophobes identiques sur les hélices $\alpha 5$ et $\alpha 6$ sont représentés en bâtons.

(B) Potentiel de surface électrostatique de M2-1[58-177] calculé avec DELPHI (Rocchia et al. 2001). Deux faces opposées sont observées. Les couleurs pour les charges varient du bleu vers le rouge pour $-6k_B T$ $+6k_B T$ respectivement. Les résidus basiques sont marqués avec des couleurs différentes pour les quatre groupes basiques (Blondot et al. 2012).

3 Analyse du N-terminal de la protéine M2-1 du VRS

3.1 Constructions de M2-1 comportant le domaine N-terminal

Des expériences préliminaires avec la protéine M2-1 pleine longueur n'ont pas donné de résultat. La protéine avait été produite avec un marquage à l'azote 15, mais son spectre HSQC $^1\text{H}-^{15}\text{N}$ indiquait qu'elle était agrégée. Pour essayer de proposer un modèle structural complet de la protéine M2-1 du VRS, nous avons donc voulu commencer par résoudre la structure du domaine N-terminal de M2-1 en utilisant des fragments de M2-1 plus courts. Nous avons travaillé avec trois constructions différentes contenant toutes les trente premiers résidus correspondant au doigt de zinc putatif (Figure 19).

- M2-1[1-32] ne contient que le domaine N-terminal.
- M2-1[1-59] et M2-1[1-65] comprennent aussi le domaine d'oligomérisation.

La région qui pourrait lier le zinc grâce au motif 3Cys-1His, conservé dans tous les *Pneumovirus*, est contenue dans les 30 premiers résidus. Il est important pour la réplication et nécessaire pour le bon fonctionnement de la protéine M2-1 (Hardy and Wertz 2000; Zhou et al. 2003; Tran et al. 2009), mais son rôle exact est encore inconnu.

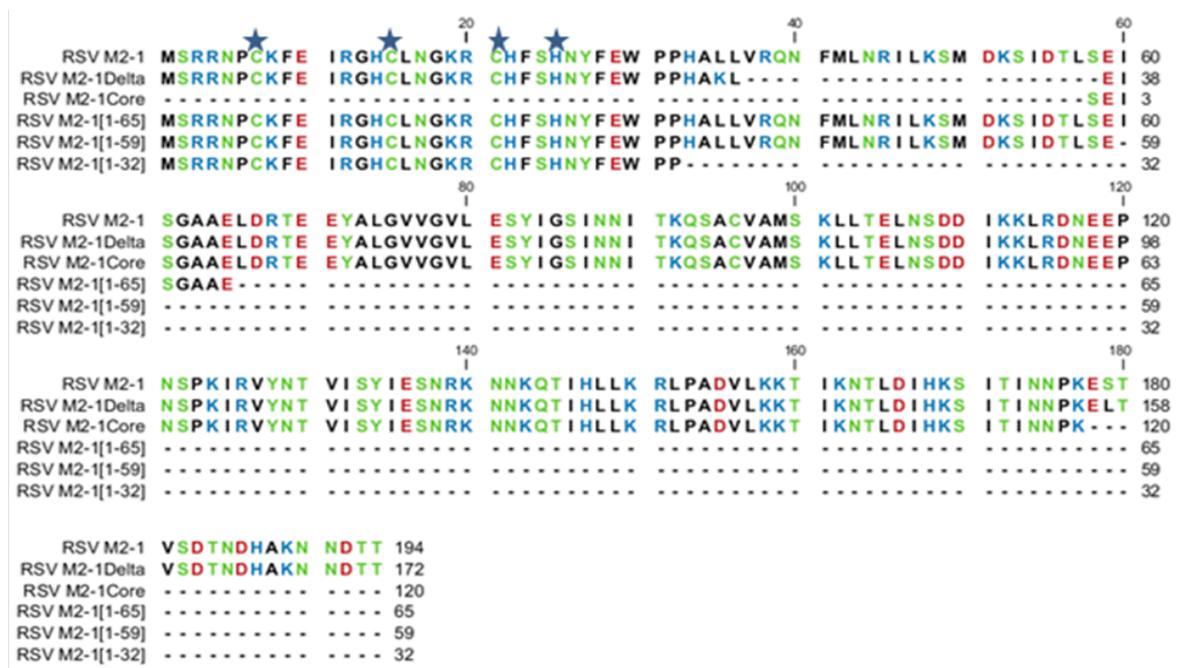


Figure 19 : Alignement des séquences des fragments de la protéine M2-1. Les couleurs correspondent à la polarité des chaînes latérales des acides aminés : acide (rouge), basique (bleu), polaire (vert), apolaire (noir). Les étoiles bleues indiquent le motif 3Cys-1His. M2-1Delta correspond à une forme sans le domaine d'oligomérisation M2-1Δ[35-58].

3.2 Expression et purification de fragments N-terminaux de M2-1

Les plasmides pGEX-M2-1 des trois constructions M2-1[1-32], M2-1[1-59] et M2-1[1-65] ont été fournis par l'équipe de Jean-François Eléouët à l'INRA. Les protéines correspondantes ont été exprimées dans la souche *E. coli* BL21 (DE3) sous la forme de protéines de fusion à la glutathion S-transférase (GST) avec un site de clivage pour la thrombine. Elles sont marquées avec des isotopes stables pour des études par RMN dans un milieu de culture minimum M9 contenant du $^{15}\text{NH}_4\text{Cl}$ comme source d'azote et du glucose marqué ^{13}C ou non comme source de carbone. Le milieu contient aussi du zinc (ZnCl_2), puisqu'on pense que la région comprenant les résidus 1-32 pourrait être un doigt de zinc avec un motif 3Cys-1His.

La surexpression des protéines a été vérifiée en prélevant des échantillons avant induction et en fin des cultures (Figure 20). Des bandes sont observées à des poids moléculaires correspondant à la somme de celui de la GST (~26 kDa) et des fragments M2-1[1-32] (4,1 kDa), M2-1[1-59] (7,2 kDa) et M2-1[1-65] (7,8 kDa).

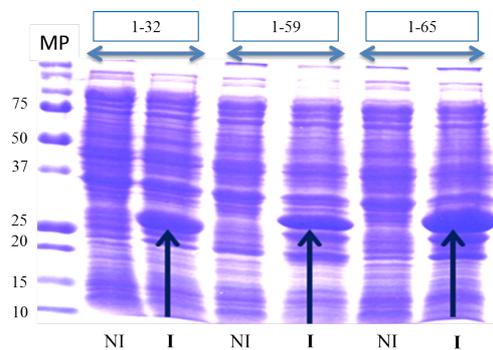


Figure 20 : Gel d'expression SDS-PAGE des trois constructions de la partie N-terminale de la protéine M2-1 avant et après induction par l'IPTG.

La purification est ensuite faite sur une résine glutathion sépharose qui fixe l'étiquette GST. Une fois la résine nettoyée, on l'incube avec de la thrombine biotinylée, une protéase qui coupe une séquence d'acides aminés entre la protéine d'intérêt et la GST. Les protéines de fusion ont bien été retenues par la résine glutathion-sépharose et ont ensuite été traitées par la thrombine. Nous avons vérifié par gel SDS-PAGE que la coupure de l'étiquette GST thrombine avait bien lieu (Figure 21).

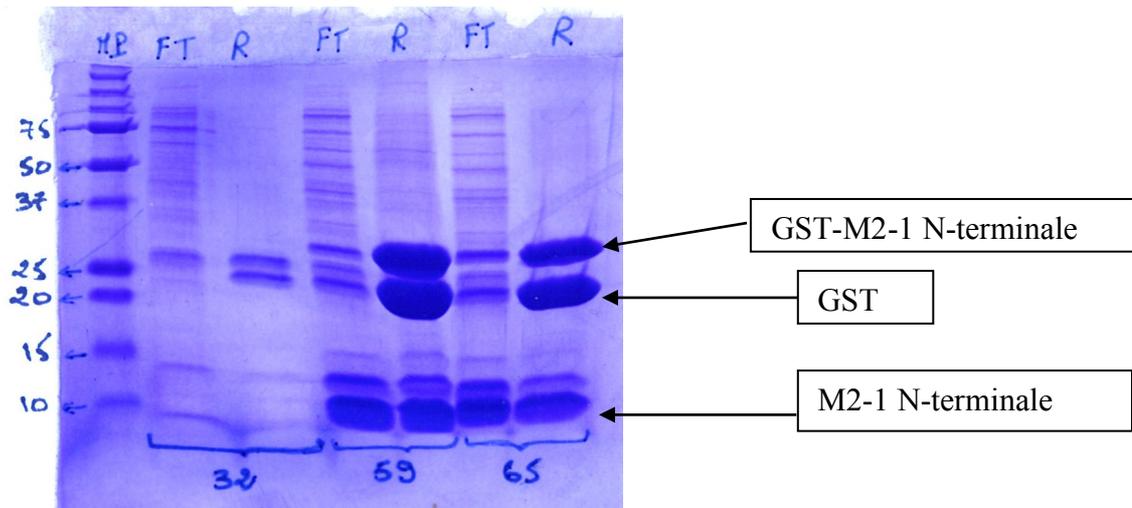


Figure 21 : Gel SDS-PAGE après coupure à la thrombine des trois constructions de la partie N-terminale de la protéine M2-1 (MP = marqueur de poids moléculaire, R = résine glutathion avec GST après coupure à la thrombine, FT = fraction non retenue après coupure à la thrombine contenant les protéines M2-1[1-32], M2-1[1-59] et M2-1[1-65]).

Mais comme nous le voyons sur la Figure 21, une partie des protéines clivées sont restées accrochées à la résine. Pour essayer de résoudre ce problème, des lavages des résines ont été faits avec différents tampons en présence de sel et d'un dénaturant (urée) après traitement à la thrombine. Les différentes conditions sont données dans le Tableau 2.

Tampon 1	Tampon 2	Tampon 3
50 mM Tris (pH=8)	50 mM Tris (pH=8)	50 mM Tris (pH=8)
1 M NaCl	200 mM NaCl	200 mM NaCl
2 mM DTT	2 mM DTT	2 mM DTT
20 μ M ZnCl ₂	20 μ M ZnCl ₂	20 μ M ZnCl ₂
	2 M urée	8 M urée

Tableau 2 : Tampons utilisés pour essayer de solubiliser les trois constructions N-terminales de la protéine M2-1

Nous avons analysé les fractions obtenues après chaque lavage par un gel de protéines SDS-PAGE (Figure 22). Après le dernier lavage à 8 M d'urée, même une partie de la GST était dénaturée, puisqu'on la trouve dans la solution de lavage. La solution contenant le fragment M2-1[1-32] était transparente alors que pour les deux autres, M2-1[1-59] et M2-1[1-65] les solutions étaient opalescentes. Ceci indique que dans les deux derniers cas la protéine était sûrement agrégée.

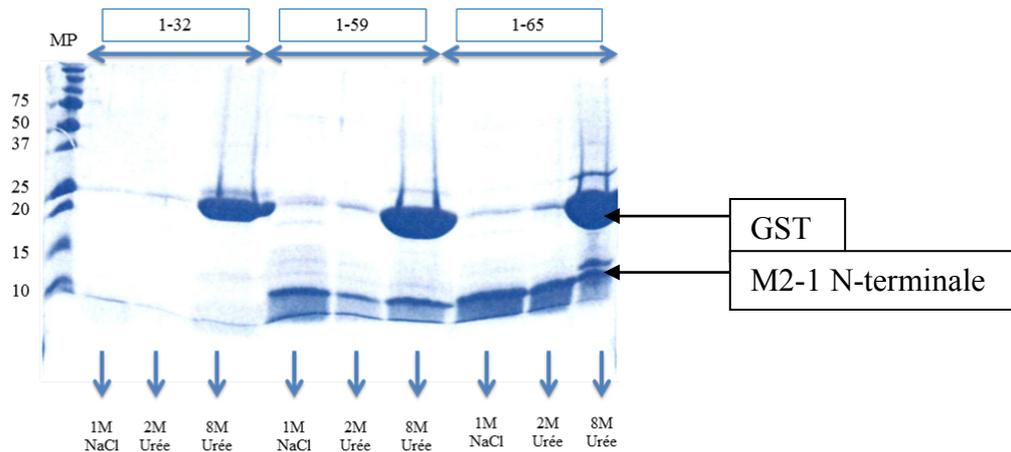


Figure 22 : Gel de protéines SDS-PAGE des 3 fragments de la protéine M2-1 (M2-1[1-32], M2-1[1-59] et M2-1[1-65]) éluées avec trois tampons différents donnés dans le Tableau 2 (MP = marqueur de poids moléculaire).

Les profils d'absorption de la lumière en UV-visible ont été mesurés avec un spectrophotomètre Nanodrop dans des conditions non dénaturantes (Figure 23). D'après ces profils, on constate qu'il y a probablement des acides nucléiques collés aux protéines, car le maximum d'absorption n'est pas centré à 280 nm (résidus aromatiques dans les protéines), mais à 260 nm (acides nucléiques). Or ceci ne correspond pas à ce qui était indiqué dans la littérature, puisque c'est le domaine central de M2-1 qui a été décrit pour fixer les ARN (Tran et al. 2009). Dans le cas des solutions opalescentes de M2-1[1-59] et M2-1[1-65], l'absorbance à 310 nm n'était pas nulle, ce qui est révélateur de la présence d'agrégats.

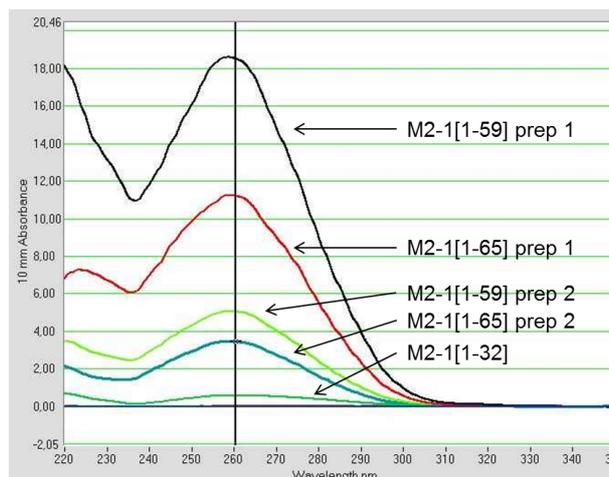


Figure 23 : Profils d'absorption en spectrométrie UV-visible des trois constructions de la partie N-terminale de la protéine M2-1. Les différentes préparations (1 et 2) correspondent à deux lavages successifs avec le même tampon (Tampon 1, Tableau 2).

3.3 Analyse par RMN des fragments N-terminaux de M2-1

3.3.1 Spectres préliminaires HSQC ^1H - ^{15}N

Au vu de ces résultats, nous avons essayé d'éliminer les contaminants sous forme d'acides nucléiques en utilisant des tampons contenant 0,5-1 M de sel lors du cassage des bactéries. Mais nous avons continué à observer une absorbance à 260 nm caractéristique de la présence résiduelle d'acides nucléiques. Des échantillons des trois constructions ont été analysés par RMN, même si les conditions n'étaient pas optimales, puisque les solutions étaient opalescentes. Certains résultats sont présentés dans la Figure 24 sous la forme de spectres HSQC (Heteronuclear Single Quantum Correlation) permettant de visualiser une connexion entre deux noyaux liés par une liaison covalente. Le spectre HSQC ^1H - ^{15}N permet d'observer en particulier les connectivités proton-azote correspondant aux liaisons amides.

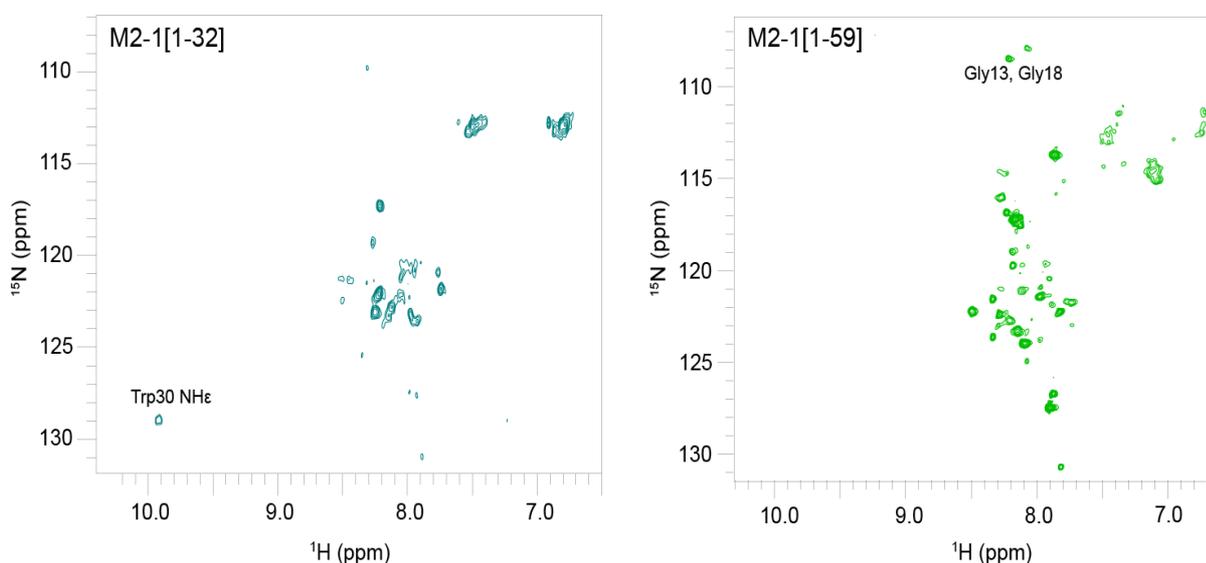


Figure 24 : Spectres HSQC ^1H - ^{15}N acquis à 14,1 T des fragments M2-1[1-32] (tampon Tris pH 6.5, NaCl 200 mM, ZnCl_2 20 μM , température 293 K) et M2-1[1-59] (tampon MES pH 6.8, NaCl 150 mM, ZnCl_2 20 μM , TCEP 0.5 mM, température 293 K).

Pour le fragment M2-1[1-32], des spectres HSQC ^1H - ^{15}N ont été acquis en tampon Tris après avoir baissé le pH en-dessous de 7, mais la faible quantité de signal indique que la protéine n'était pas monomère (Figure 24). Quelques pics intenses sont observés : ils correspondent à des résidus non structurés. On observe aussi un signal correspondant à la chaîne latérale du tryptophane Trp30. Mais cet échantillon n'est pas de qualité suffisante pour envisager de continuer.

Les fragments M2-1[1-59] et M2-1[1-65] ont tendance à précipiter lorsque le pH est abaissé de 8 à 6.5. Le fragment M2-1[1-59] semble être le plus stable des deux. Le spectre

HSQC ^1H - ^{15}N dans la Figure 24 montre qu'on obtient plus de signaux que pour le fragment M2-1[1-32], mais que l'intensité des raies est inhomogène. Les raies fines et intenses avec des déplacements chimiques ^1H entre 7,5 et 8,5 ppm indiquent la présence de régions désordonnées, ce qui n'était pas attendu. A 293 K on observe deux signaux de glycines qui correspondent à Gly13 et Gly18, mais pas la chaîne latérale de Trp30.

3.3.2 Influence de la température

Le fragment M2-1[1-59] a été analysé plus en détail. J'ai étudié plus particulièrement l'influence de la température sur les spectres RMN. La Figure 25 montre que l'intensité des signaux augmente progressivement avec la température de 288 à 303 K. A 303 K on peut observer le signal NH ϵ du tryptophane Trp30. Par contre les signaux restent inhomogènes. Des pics très intenses se superposent à des pics larges et peu intenses et on ne trouve pas la soixantaines de signaux attendus.

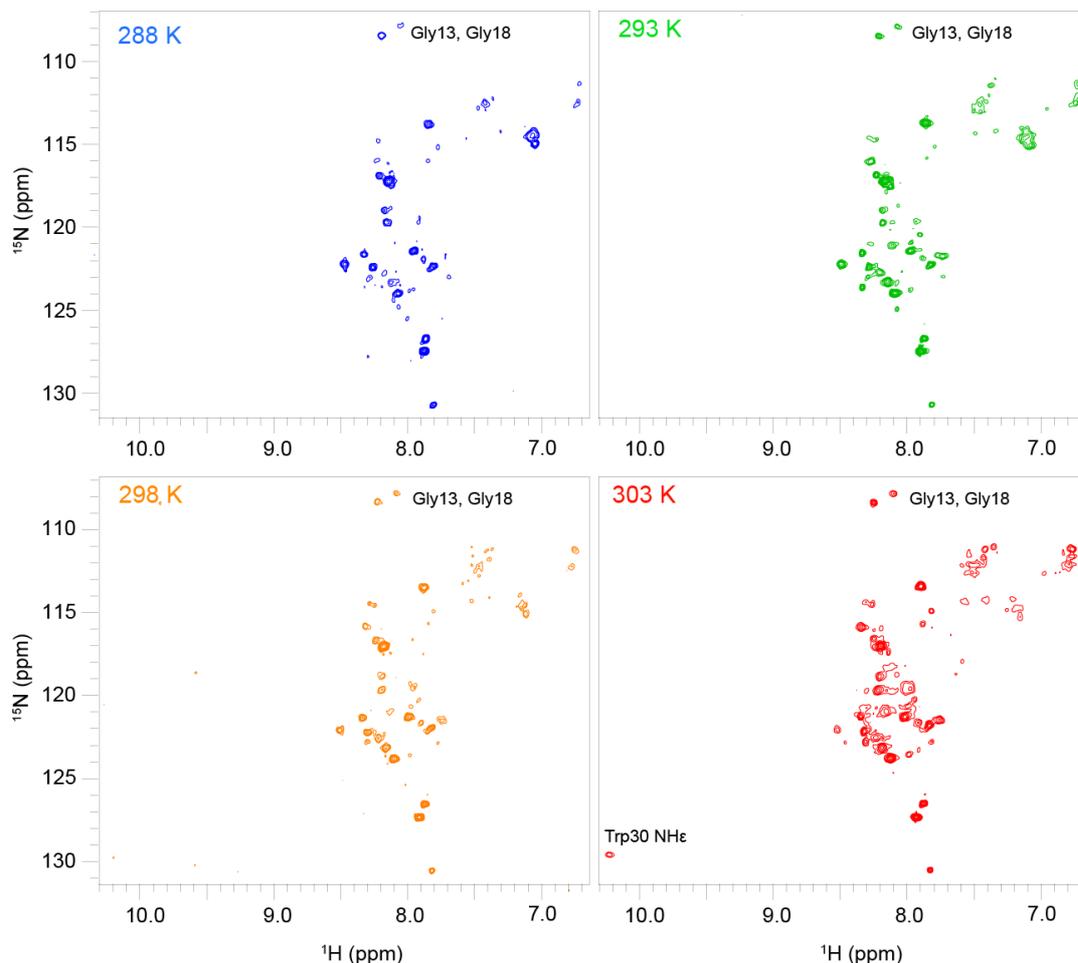


Figure 25 : Spectres HSQC ^1H - ^{15}N du fragment M2-1[1-59] acquis à 14,1 T et à température variable comprise entre 288 K et 303 K (tampon MES pH 6.8, NaCl 150 mM, ZnCl_2 20 μM , TCEP 0.5 mM)

3.4 Analyse par RMN d'un peptide synthétique contenant le motif "3Cys-1His" de M2-1

3.4.1 Analyse de la séquence du domaine N-terminal de M2-1

En raison des problèmes de solubilité et de stabilité des fragments de M2-1 produits sous forme de protéines recombinantes, j'ai choisi de travailler avec un peptide synthétique correspondant au domaine de liaison du zinc, afin de pouvoir mieux contrôler les concentrations relatives en zinc et en peptide. La séquence de ce peptide synthétique a été choisie après alignement avec des doigts de zinc de type 3Cys-1His, plus précisément CX8CX5CX3H, présents sous forme de tandem dans les protéines Nup475 (Amann et al. 2003) et TIS11d (Hudson et al. 2004). Un peptide de 29 acides aminés, ZNF29, a été choisi. Les alignements sont indiqués dans la Figure 26.

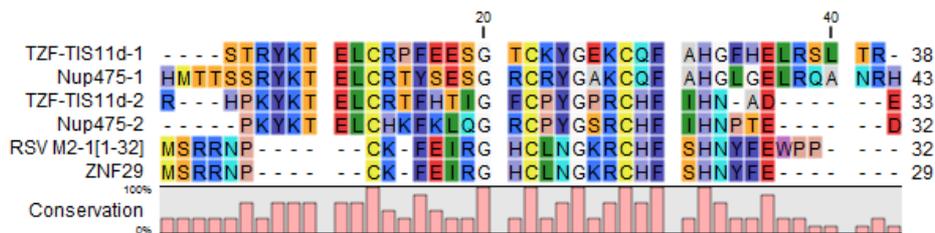


Figure 26 : Alignement des séquences du peptide synthétique ZNF29 correspondant à la séquence N-terminale de M2-1 avec les séquences des doigts de zinc en tandem (TZF) des protéines TIS11d et Nup475.

Un modèle par homologie (Figure 27) a été construit sur la base des quatre structures des doigts de zinc dans Nup475 (PDB 1MO9) et TIS11d (PDB 1RGO) avec le logiciel Modeller (9v8) (Sali and Blundell 1993). Il a servi à définir les bornes du peptide ZNF29.

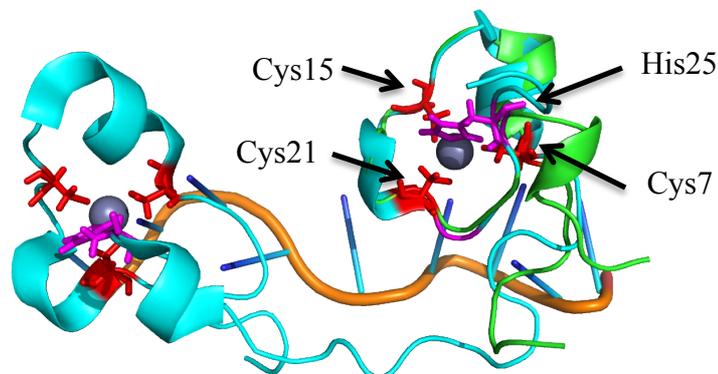


Figure 27 : Modèle par homologie du domaine N-terminal de M2-1 (vert) superposé à la structure RMN de la protéine TIS11d en complexe avec un ARNm. Les trois cystéines et l'histidine de TIS11d sont représentées en bâtons rouges et magenta, respectivement.

3.4.2 Essais de reconstitution du complexe ZNF29:Zn

3.4.2.1 Conditions de tampon pour reconstituer le complexe ZNF29:Zn

Le zinc a un rôle exclusivement structural dans les protéines à doigts de zinc. Le cation Zn^{2+} est lié aux chaînes latérales des cystéines ou histidines pour former un complexe organométallique. Ce complexe présente une coordination tétraédrique, avec les atomes de soufre (SG) des cystéines ou d'azote (ND1 ou NE2) des histidines autour du zinc. Les chaînes latérales des cystéines ou histidine sont alors déprotonées.

Le but recherché était d'identifier des conditions dans lesquelles le peptide ZNF29 se structure en présence de zinc. Les tampons utilisés devaient répondre aux critères suivants :

- Le tampon doit contenir un agent réducteur pour éviter l'oxydation des cystéines. Nous avons testé le DTT ou le TCEP, mais le DTT qui comporte deux fonctions thiols pourrait chélater le cation Zn^{2+} .
- Le phosphate formant un complexe insoluble avec Zn^{2+} , il a été exclu.
- Le peptide possède un point isoélectrique de 9,5. Cela permet d'explorer en théorie une gamme de pH de 1 à 9.
- Le tampon doit contenir le cation Zn^{2+} sous forme soluble, donc à un pH inférieur à 7,5 d'après le diagramme potentiel-pH du zinc (Figure 28).

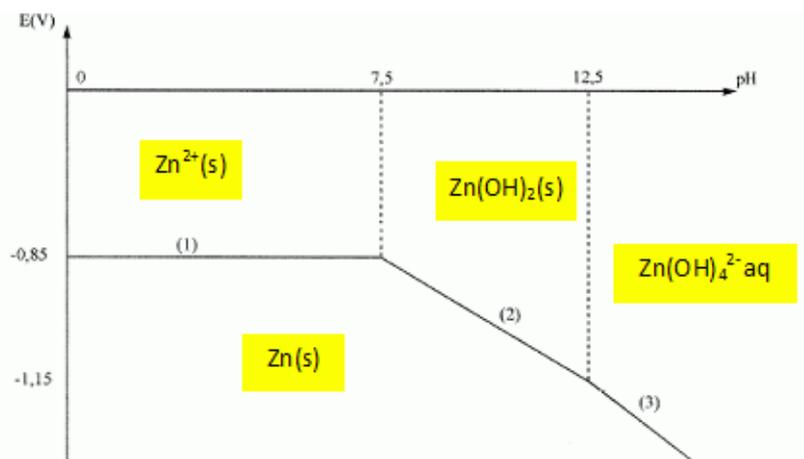


Figure 28 : Diagramme potentiel-pH du zinc pour une concentration d'espèces en solution 10^{-3} mol/l.

Le peptide est initialement sous forme lyophilisée. Lorsqu'il est repris dans de l'eau, son pH est d'environ 4. Afin de trouver les meilleures conditions pour obtenir une protéine stable et bien structurée, nous avons testé différentes concentrations en peptide, nous avons effectué des titrages en pH en présence de zinc, puis essayé de rajouter du zinc à un pH plus élevé et enfin de remplacer le zinc par du cadmium.

3.4.2.2 Reconstitution du complexe ZNF29:Zn par titrage du pH

Les changements de déplacement chimique liés un changement d'état de protonation d'une chaîne latérale peuvent être mesurés à l'aide de spectres RMN ^1H en fonction du pH. Si le peptide se structure en présence de zinc, l'addition d'ions Zn^{2+} donne lieu à une importante variation de déplacements chimiques. Le pKa d'une cystéine libre est de 8,5, mais cette valeur peut baisser jusqu'à 4 lorsque le thiolate est stabilisé par coordination avec un cation métallique, comme c'est le cas dans un doigt de zinc.

Une première série a été faite en partant d'une solution de peptide dissous dans de l'eau avec deux équivalents de ZnCl_2 et en présence de DTT deutérié. Des spectres RMN 1D ^1H ont été enregistrés à 600 MHz pour suivre l'état de repliement de ce peptide en augmentant progressivement le pH par ajout d'une solution de NaOH concentrée à 0,5 M. Ces spectres reflètent ce que l'on voit à l'oeil nu : la solution devient progressivement opalescente et le peptide précipite.

Une deuxième série a été faite en partant d'une solution de peptide dissous en présence de DTT deutérié dans de l'eau, sans ZnCl_2 , et en augmentant progressivement le pH. A pH acide (3,6), la solution est limpide. La faible dispersion des pics de résonance dans la région des protons amides montre que le peptide est non replié, comme illustré dans la Figure 29. Au fur et à mesure que le pH augmente, on observe le même phénomène qu'en présence de ZnCl_2 . Les raies du spectre s'élargissent, comme on peut le voir à pH 7,8 dans la Figure 29. Du zinc a été ajouté jusqu'à un équivalent à pH 6, mais cela n'a pas eu d'effet.

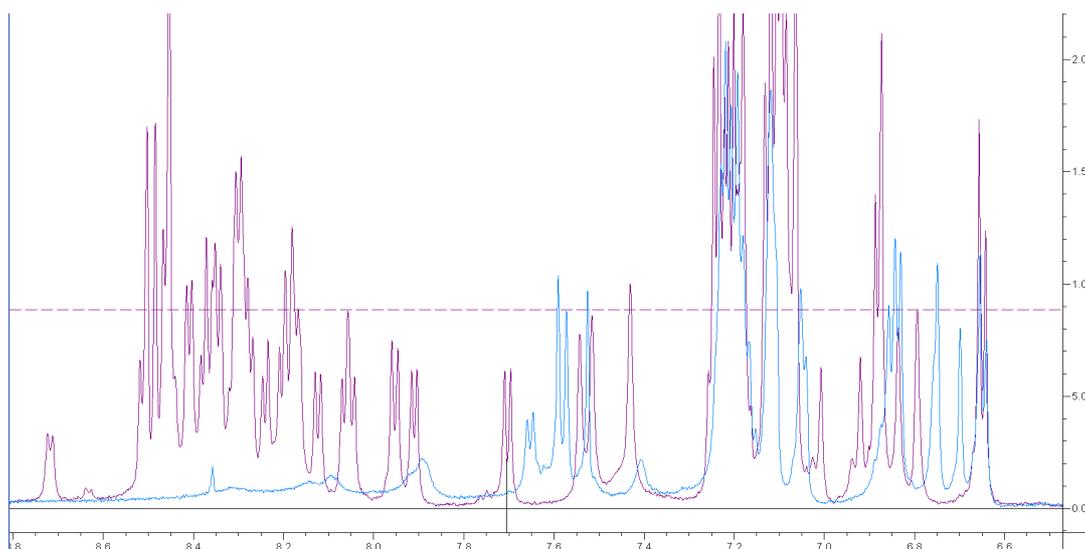


Figure 29 : Zoom dans la région des protons amides et aromatiques de spectres RMN ^1H du peptide synthétique ZNF29 à 335 μM à deux pH différents (pH 3,6 en violet et pH 7,8 en bleu) dans l'eau et en présence de 2 mM DTT-d10 acquis à 14,1 T et 298 K.

3.4.2.3 Complexation du cadmium

N'ayant pas réussi à obtenir de complexe stable avec du zinc, j'ai fait une dernière tentative, en utilisant le cadmium à la place du zinc. Le cadmium figure dans la même colonne du tableau périodique que le zinc et possède donc des propriétés comparables. On peut remplacer le zinc par le cadmium dans des complexes de coordination. Les propriétés électrochimiques sont proches. Si on compare les diagrammes potentiel-pH des deux métaux, le cation Cd^{2+} (Figure 30) paraît plus soluble dans l'eau que Zn^{2+} (Figure 28). Les essais effectués n'ont pas été concluants.

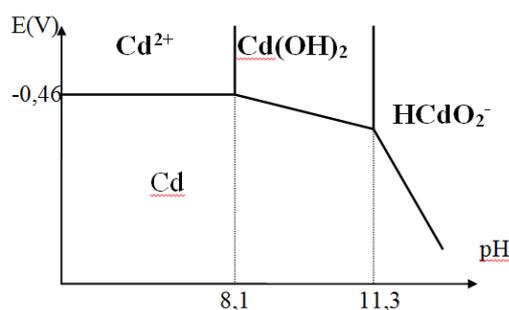


Figure 30 : Diagramme potentiel-pH du cadmium tracé pour une concentration d'espèces en solution 10^{-2} mol/l

3.4.3 Analyse par RMN homonucléaire du peptide ZNF29

Malgré les échecs pour reconstituer un doigt de zinc avec le peptide ZNF29, nous avons voulu obtenir des informations structurales sur ce peptide en solution. Des expériences TOCSY et NOESY ont été acquises durant les étapes de titrage en pH.

Dans une expérience bidimensionnelle NOESY "Nuclear Overhauser Enhancement Spectroscopy" ^1H - ^1H on mesure les effets dus à l'interaction de protons voisins dans l'espace. Elle permet d'obtenir une information sur les distances entre protons distants de moins de 6\AA , grâce à l'intensité des pics de corrélation dus à l'effet Overhauser nucléaire (NOE). Une expérience bidimensionnelle TOCSY "Total Correlation Spectroscopy" ^1H - ^1H est basée sur les couplages scalaires entre protons d'un même système de spin au travers des liaisons covalentes.

L'expérience TOCSY (Figure 32) a permis de déterminer les systèmes de spin correspondant aux 29 acides aminés de ZNF29. L'expérience NOESY a permis de trouver les enchaînements séquentiels grâce aux signaux NOE intenses entre le proton amide d'un résidu i et le proton $\text{H}\alpha$ du résidu en $i-1$ dans des structures de type pelote statistique, comme indiqué dans la Figure 31. Ceci démontre que le peptide synthétisé a bien la bonne séquence et que

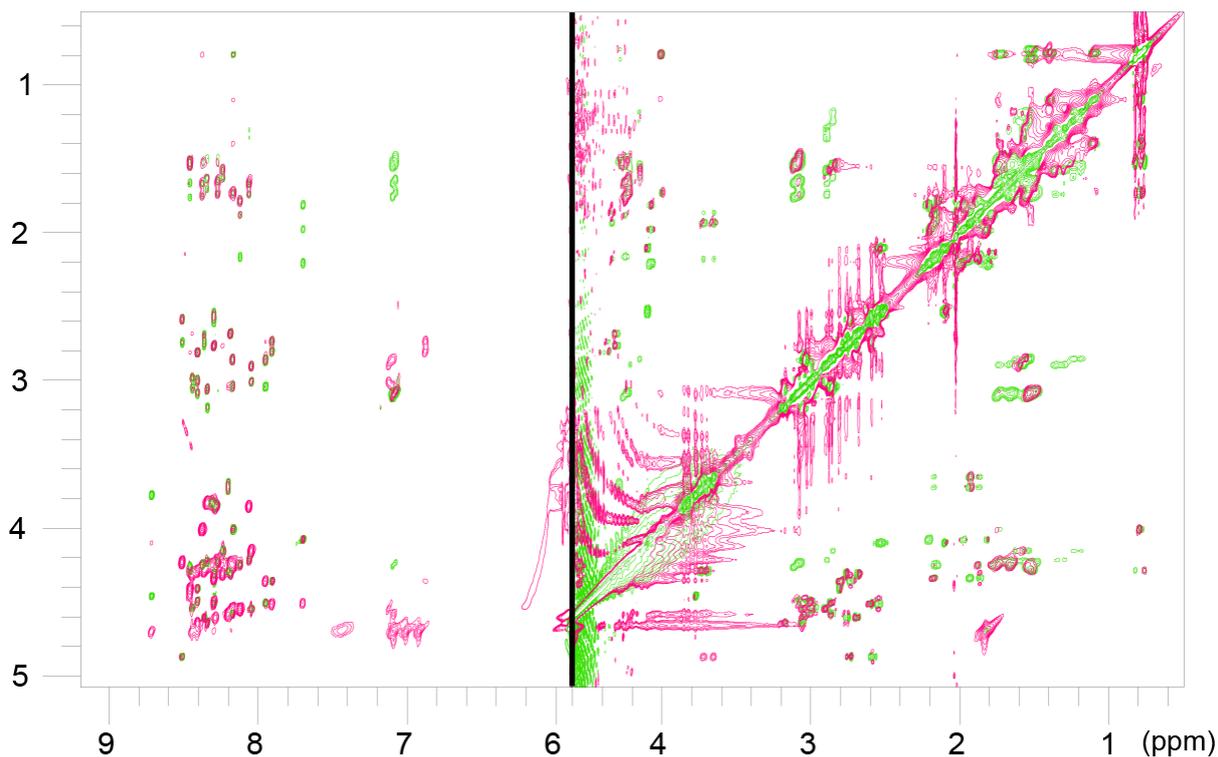


Figure 33 : Extraits des spectres TOCSY (vert, spin lock 60 ms) et NOESY (rose, temps de mélange 120 ms) ^1H - ^1H du peptide synthétique ZNF29 (concentration 335 μM , pH 3,6, 2 mM DTT), acquis à 14,1 T et 298 K montrant les corrélations des protons amide HN vers les protons aliphatiques et les corrélations entre protons aliphatiques.

Dans le cas du peptide ZNF29 à pH 3,6, le spectre NOESY ne contient pas de signaux autres que ceux correspondant à des contacts intra-résiduels (Figure 32, Figure 33). Ceci confirme ce que nous avons vu sur le spectre 1D : le peptide n'est pas structuré dans ces conditions.

Pour conclure, les résultats obtenus avec le peptide ZNF29 semblent indiquer que le domaine N-terminal de M2-1 n'est pas stable à pH neutre. Ceci pourrait être expliqué par le fait que ce domaine a besoin des autres domaines de la protéine M2-1 pour se structurer correctement.

4 Analyse par RMN de la protéine M2-1 dépourvue du domaine d'oligomérisation

4.1 Expression et purification de M2-1 Δ [35-58]

Afin d'éliminer les problèmes de stabilité qui pourraient être liés à la présence du domaine d'oligomérisation et surtout de diminuer la taille globale de la protéine, nous avons choisi de travailler avec une construction qui correspond à la protéine M2-1 sans son domaine d'oligomérisation : M2-1 Δ [35-58] dont la séquence est donnée dans la Figure 19. Cela permet aussi d'avoir une région témoin que nous connaissons déjà, qui est le domaine globulaire de M2-1. Cette construction est constituée de 174 acides aminés et son poids moléculaire est de 19,7 kDa.

Cette protéine a été produite sous forme de protéine de fusion avec la GST, GST-M2-1 Δ [35-58], en utilisant le même protocole de production et de purification que celui utilisé pour les fragments N-terminaux de M2-1, détaillé en Annexe, comportant une étape de capture de la GST sur résine glutathion-sépharose, suivi d'une coupure à la thrombine (Figure 34).

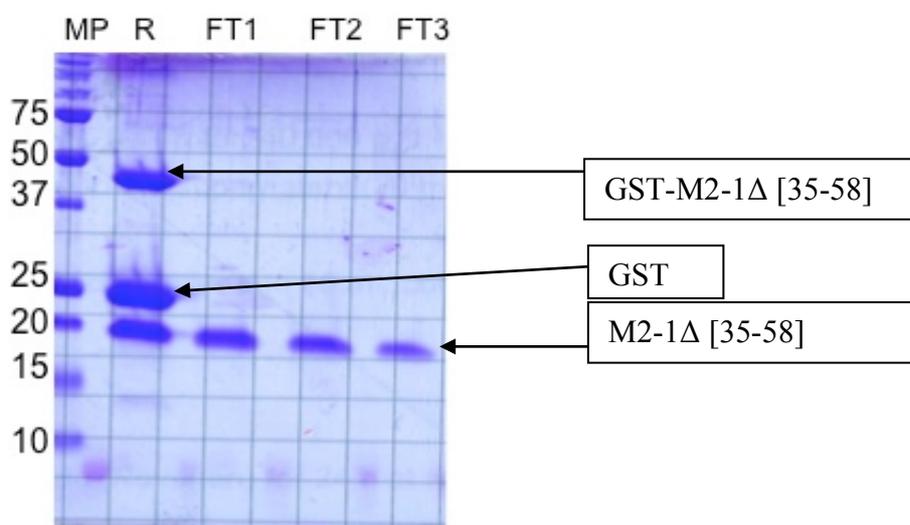


Figure 34 : Gel SDS-PAGE après coupure à la thrombine de la construction M2-1 Δ [35-58] (MP = marqueur de poids moléculaire, R = résine glutathion avec GST après coupure à la thrombine, FT1, 2, 3 = fractions non retenues après coupure à la thrombine contenant la protéine M2-1 Δ [35-58])

En particulier les tampons utilisés contenaient de faibles quantités de Zn^{2+} pour compenser la perte éventuelle de zinc au niveau du doigt de zinc. Le tampon de cassage contenait de grandes quantités de sel et de la benzonase pour éliminer les acides nucléiques

pouvant se lier par exemple au domaine central. Malgré tout, au vu du spectre d'absorption UV-VIS (Figure 35), on peut se demander s'il ne reste pas des acides nucléiques puisque le maximum d'absorption n'est pas complètement à 280 nm.

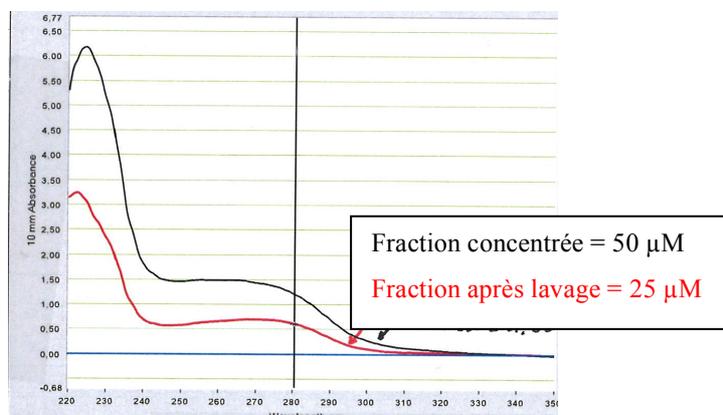


Figure 35 : Spectre d'absorption UV-VIS de deux fractions obtenues après coupure à la thrombine pour M2-1Δ[35-58].

4.2 Optimisation du tampon de M2-1Δ[35-58]

La solution de protéine a été dialysée contre un tampon 20 mM MES pH 6,2, 150 mM NaCl, 1 mM β-mercaptoéthanol, 20 μM ZnCl₂ pour essayer d'être dans les meilleures conditions d'étude pour la RMN. Après dialyse, au moment de concentrer la solution, la protéine a précipité. La protéine ne semble pas supporter des concentrations supérieures à 50 μM dans ces conditions.

Les spectres HSQC ¹H-¹⁵N mesurés à différentes températures, après centrifugation et élimination de la protéine précipitée, étaient de faible intensité et semblaient montrer que les solutions contenaient essentiellement des agrégats. Pour essayer de stabiliser la protéine, j'ai ajouté de l'arginine et/ou du glutamate (L-Arg, L-Glu) à des concentrations de 0,5 M dans la solution de M2-1Δ[35-58]. Ces acides aminés sont efficaces pour prévenir l'agrégation et la précipitation des protéines, car ils réduisent les interactions intermoléculaires. Cela permet d'améliorer la solubilité des protéines et d'augmenter considérablement la stabilité à long terme de l'échantillon. En outre, ils protègent les échantillons de protéines de la dégradation protéolytique.

L'ajout de 0.5 M d'arginine a permis d'obtenir un spectre pour une solution plus concentrée avec des raies plus fines, présenté dans la Figure 36. L'arginine a un effet stabilisateur puisqu'elle possède un groupe guanidinium avec cinq hydrogènes disponibles pour faire des liaisons hydrogènes avec la protéine et peut même participer à des ponts salins doubles avec l'asparagine et la glutamine.

4.3 Attribution des fréquences de résonance du squelette peptidique de la construction M2-1Δ[35-58]

4.3.1 Principe de l'attribution séquentielle des fréquences de résonance

Pour pouvoir interpréter les signaux supplémentaires et les variations dans le spectre HSQC ^1H - ^{15}N de M2-1Δ[35-58] par rapport à celui de M2-1[58-177], il faut attribuer les signaux amides de M2-1Δ[35-58].

Le marquage isotopique des protéines a permis le développement de stratégies originales d'attribution pour les protéines (Kay et al. 1990). De nombreuses expériences tridimensionnelles hétéronucléaires permettent l'attribution des fréquences de résonance des noyaux du squelette. Elles sont notées HNCA, HN(CO)CA, HNCACB, HN(CO)CACB, HNCO, HN(CA)CO et sont basées sur les transferts d'aimantation entre les noyaux actifs en RMN par l'intermédiaire des couplages scalaires homonucléaires et hétéronucléaires J (entre ^{13}C , ^{15}N et leurs protons attachés), comme illustré dans la Figure 37.

La plupart des acides aminés ont un groupe HN amide unique à l'exception de la proline qui n'en a aucun et l'asparagine et la glutamine qui ont une chaîne latérale amide (CONH_2). Dans des expériences 3D classiques pour l'attribution des protéines, des groupes HN sont corrélés avec un autre type de noyau de manière séquentielle ou bidirectionnelle. Ces expériences sont nommées par le noyau impliqué dans le transfert. Les déplacements chimiques des noyaux non édités sont mis entre parenthèses.

Par exemple, l'expérience HNCO donne la corrélation entre les fréquences de résonances $^1\text{H}^{\text{N}}$, ^{15}N du résidu i et la fréquence de résonance $^{13}\text{C}'$ du résidu précédent $i-1$. L'expérience HN(CA)CO donne la corrélation entre les résonances $^1\text{H}^{\text{N}}$, ^{15}N du résidu i et la résonance $^{13}\text{C}'$ du résidu précédent $i-1$, mais aussi du même résidu i via le carbone aliphatique non édité $^{13}\text{C}^{\alpha}$.

La Figure 37 montre les transferts d'aimantation créés pour différentes expériences utilisées. En combinant toutes ces expériences, l'attribution séquentielle des fréquences de résonance du squelette peptidique peut être effectuée. Des bandes ^1H ^{13}C sont extraites des spectres 3D à une fréquence donnée ^{15}N et sont alignées pour permettre l'attribution séquentielle. Certains types d'acides aminés tels que alanine, glycine, serine et thréonine peuvent être identifiés facilement grâce à leurs déplacements chimiques $^{13}\text{C}_{\alpha}$ ou $^{13}\text{C}_{\beta}$ remarquables et ils sont utilisés comme point de départ pour l'attribution séquentielle.

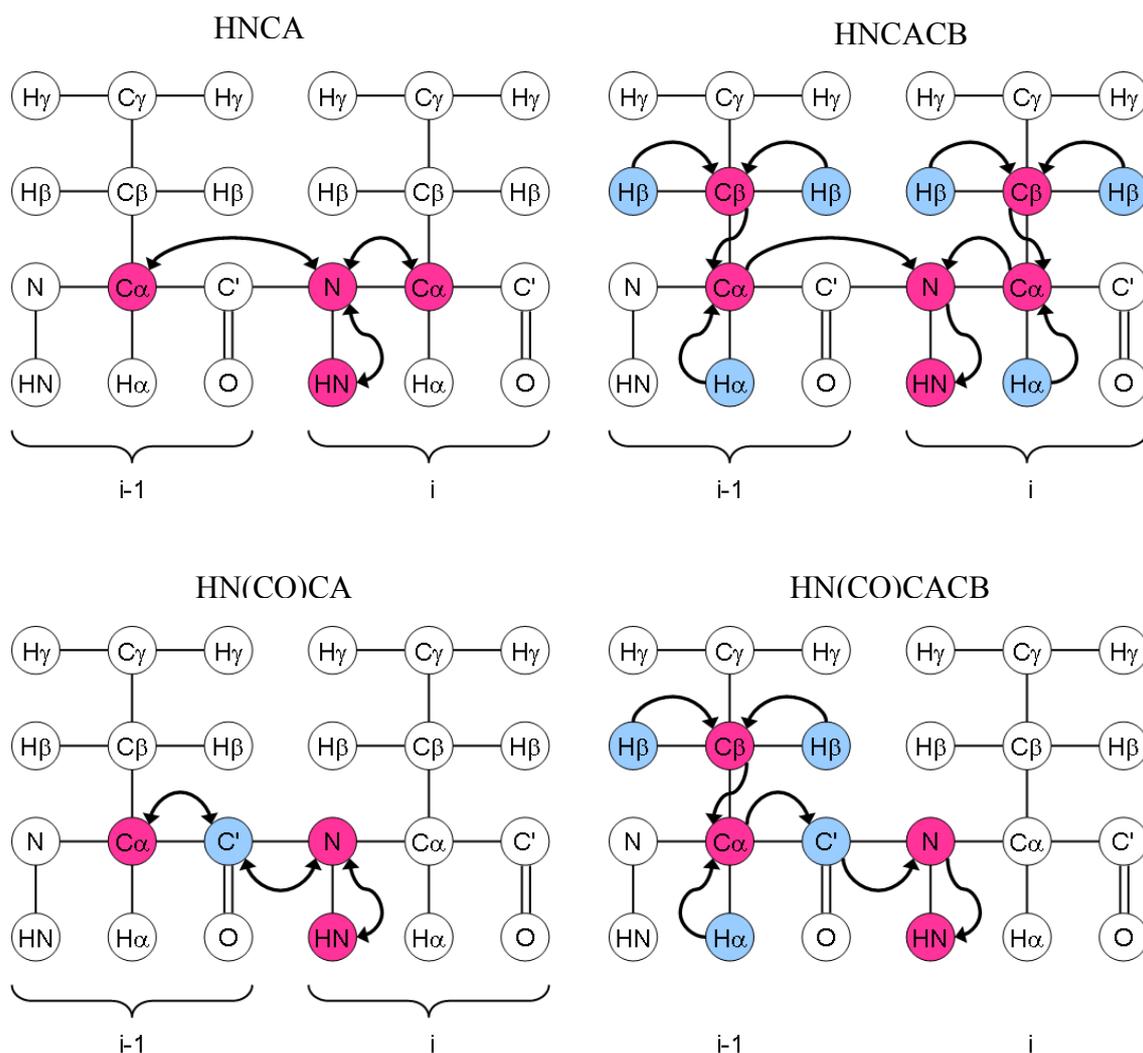


Figure 37 : Transferts d'aimantation dans les expériences tridimensionnelles triple-noyau utilisées en RMN des protéines. Les flèches représentent les transferts d'aimantation partant du proton amide vers les noyaux édités (en rose) via les couplages scalaires. Les noyaux en bleu sont utilisés comme relais de l'aimantation, mais ils ne sont pas édités.

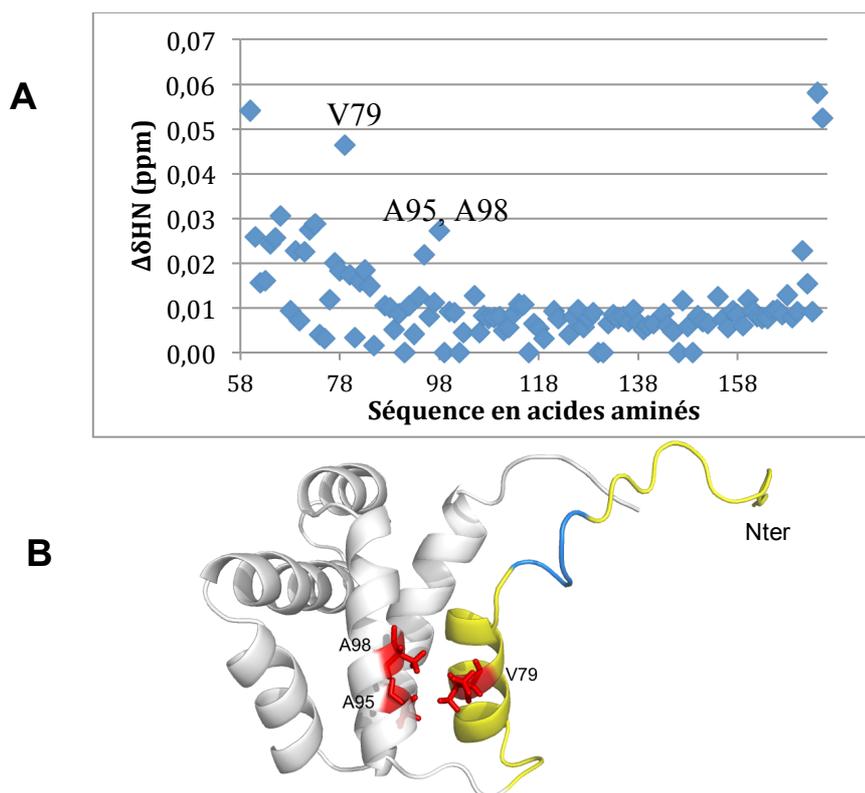
(<http://www.protein-nmr.org.uk/>)

4.3.2 Attribution du squelette peptidique de M2-1 Δ [35-58]

J'ai produit la protéine doublement marquée ^{13}C - ^{15}N . Puis j'ai acquis des spectres HNCA, HN(CO)CA, CBCA(CO)NH sur un spectromètre RMN Avance II avec un champ de 14,1 T équipé d'une cryo-sonde, pour procéder à l'attribution séquentielle des fréquences de résonance de son squelette peptidique, ^1H , ^{15}N , $^{13}\text{C}\alpha$ et dans la mesure du possible $^{13}\text{C}\beta$. Les mesures ont été faites à 303 K pour avoir des raies plus fines et dans un tampon contenant 0,5 M d'arginine. Le jeu de spectres n'est pas complet. La HNCACB n'a pas été lancée en raison de la faible sensibilité intrinsèque de cette expérience et de la faible concentration de

l'échantillon. Mais comme le domaine globulaire de la protéine M2-1 avait été déjà attribué, cela a facilité la tâche. La HNCO n'a pas été acquise non plus, bien qu'elle aurait permis de compléter les données des fréquences de résonance pour prédire les structures secondaires. Je me suis attaquée à l'attribution des fréquences de résonance du squelette peptidique de cette construction afin d'identifier les pics correspondants aux parties N-terminale et C-terminale. L'analyse des spectres a été faite dans CCPNMR (Vranken et al. 2005).

L'attribution des fréquences de résonance du domaine globulaire a pu être retrouvée dans le fragment M2-1 Δ [35-58]. On peut comparer plus finement le spectre HSQC ^1H - ^{15}N de M2-1 Δ [35-58] et de M2-1[58-177] en analysant les variations de déplacements chimiques des différents résidus du domaine globulaire de M2-1 entre M2-1 Δ [35-58] et M2-1[58-177] (Figure 38A). On voit que les variations sont plus petites que la largeur de raie (0,02 ppm), sauf aux extrémités. On constate que la région N-terminale est plus perturbée et que cette région correspond à la première hélice du domaine globulaire de M2-1 (Figure 38B). Les résidus les plus affectés sont situés en vis-à-vis entre l'hélice 1 et l'hélice 2.



Ces résultats semblent indiquer que la présence de l'extension N-terminale perturbe légèrement la structure du domaine globulaire et ainsi que celle du linker entre ces deux régions qui apparaît comme non structuré dans M2-1[58-177]. On peut penser que le linker se structure dans le contexte de la construction M2-1Δ[35-58]. La prédiction de structure secondaire sur la base des déplacements chimiques faite par le logiciel DANGLE (Cheung et al. 2010), implémenté dans CCPNMR, donne une hélice pour les résidus E70 à A73.

L'attribution de la queue C-terminale a aussi été obtenue, comme indiqué sur le spectre HSQC ^1H - ^{15}N de la Figure 39. Un tableau avec tous les déplacements chimiques est donné en Annexe. La prédiction de structure secondaire par DANGLE confirme qu'il s'agit d'une région désordonnée.

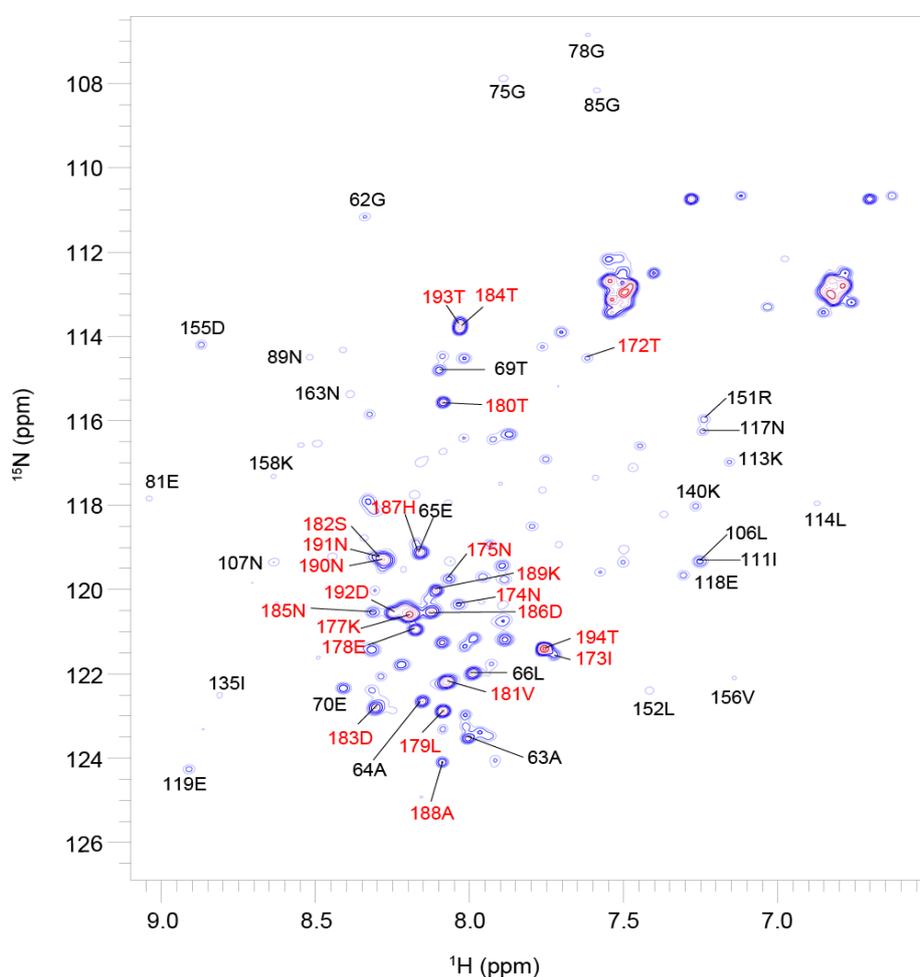


Figure 39 : Spectre HSQC ^1H - ^{15}N de la protéine M2-1Δ[35-58] à 14,1 T et 303K (90 μM dans 20 mM MES pH 6,2, 150 mM NaCl, 0.5 M arginine, 20 μM ZnCl_2 , 2 mM β -mercaptoéthanol). L'attribution de la partie C-terminale désordonnée est en rouge. Le seuil des contours est élevé et les signaux de la partie globulaire apparaissent très faibles. Seuls quelques signaux du domaine globulaire sont annotés, mais l'attribution de ce domaine est complète.

Il reste des signaux non attribués qui doivent correspondre à la partie N-terminale, mais leurs intensités sont trop faibles pour pouvoir faire une attribution séquentielle : dans les spectres 3D, on ne voit pas de corrélation avec le carbone 13. Ceci indique que ce domaine N-terminal n'est probablement pas structuré correctement et adopte peut-être des conformations multiples en échange.

5 Analyse par RMN d'une construction de M2-1 sans domaine N-terminal

5.1 Production du fragment M2-1 Δ [1-31]

La construction M2-1 Δ [35-58] montre que la partie doigt de zinc pose problème, donc je suis revenue sur une construction de M2-1 dépourvu de ce domaine, mais comprenant le domaine d'oligomérisation, noté M2-1 Δ [1-31].

Cette protéine a été produite sous forme de protéine de fusion avec la GST, GST-M2-1 Δ [1-31]. La séquence est donnée dans la Figure 40.

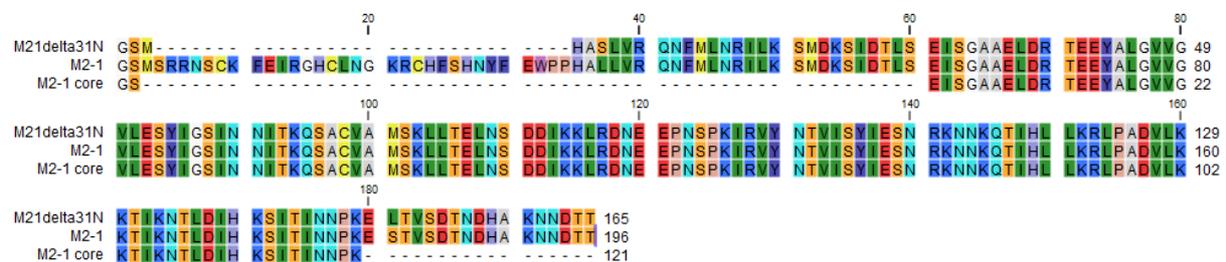


Figure 40 : Comparaison des séquences en acides aminés des constructions des M2-1 correspondant à des délétions de domaines : M2-1 Δ [1-31], M2-1 entier et M2-1[58-177].

Le même protocole de production et de purification a été utilisé que pour GST-M2-1 Δ [35-58], mais sans ajouter de ZnCl₂ dans les tampons. L'analyse par SDS-PAGE des différentes étapes est illustrée dans la Figure 41. Ce fragment de protéine est soluble et paraît stable dans des conditions de tampons usuelles de RMN (ici tampon MES pH 6,5 NaCl 150 mM).

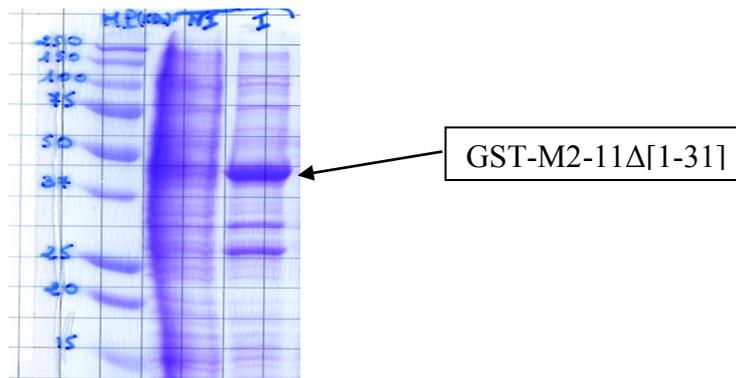


Figure 41 : Gel SDS-PAGE du suivi de culture de M2-1Δ[1-31]. (MP = marqueur de poids moléculaire, NI et I = extraits bactériens avant et après induction à l'IPTG).

5.2 Analyse par RMN du fragment M2-1Δ[1-31]

La protéine M2-1Δ[1-31] a été produite avec un marquage ^{15}N et son spectre HSQC ^1H - ^{15}N comparé à celui du fragment M2-1[58-177] dans les mêmes conditions expérimentales (Figure 42).

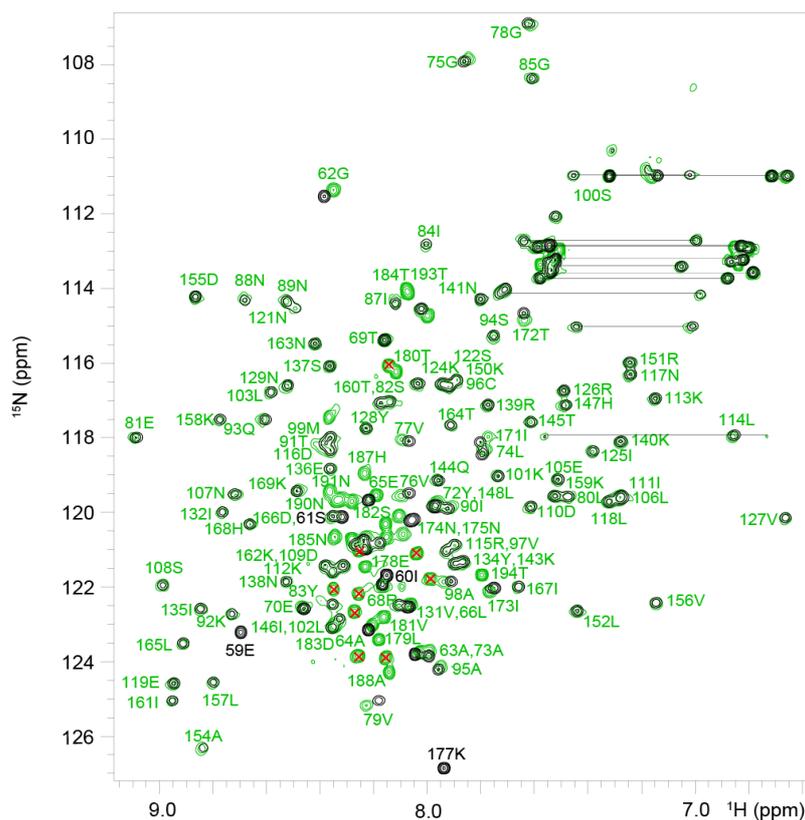


Figure 42 : Spectres HSQC ^1H - ^{15}N superposés des fragments M2-1Δ[1-31] (en vert ; 100 μM dans un tampon 20 mM MES pH 6,5, 150 mM NaCl, 2 mM β -mercaptoéthanol) et M2-1[58-177] (noir) acquis dans les mêmes conditions (14,1 T et 293 K). Les attributions des signaux amides sont données pour M2-1Δ[1-31] en vert et M2-1[58-177] en noir.

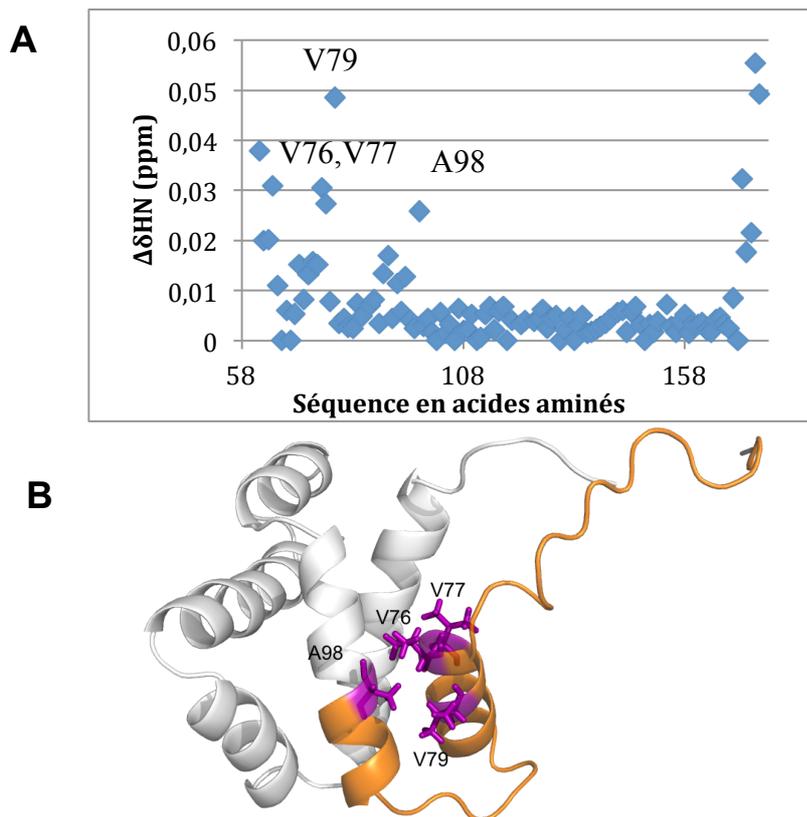


Figure 43 : (A) Variations de déplacements chimiques combinés ^1H et ^{15}N des amides du domaine globulaire de M2-1 entre M2-1 Δ [1-31] et M2-1[58-177]. Le graphe est coupé et la valeur pour Lys177 (0.68 ppm) n'est pas reportée. (B) Les résidus avec les variations les plus importantes ont été cartographiés en magenta sur la structure de M2-1[58-177], la partie N-terminale (E58- N98) en orange.

L'attribution séquentielle n'a pas été faite pour cette construction, mais les attributions du domaine globulaire et de la queue C-terminale déterminées avec les autres fragments peuvent être reportées facilement. Une dizaine de signaux intenses n'a pas pu être attribuée. Ces signaux sont marqués d'une croix rouge sur la Figure 42. Ils correspondent à des signaux de la partie N-terminale. Il manque aussi une vingtaine de signaux par rapport à la séquence en acides aminés. En particulier il n'a pas été possible de trouver les signaux des résidus Glu59-Ser61, soit parce qu'ils se sont déplacés à cause du changement d'environnement électronique lié à la présence d'acides aminés N-terminaux, soit parce que leurs signaux sont élargis. On suppose qu'ils appartiennent au domaine d'oligomérisation. L'origine de l'élargissement de raie est probablement différent de celui observé dans le fragment M2-1 Δ [35-58]. En effet, si l'on imagine que le domaine de tétramérisation est formé d'hélices, celles-ci peuvent glisser un peu les unes par rapport aux autres, ce qui induirait des élargissements par de l'échange en régime intermédiaire. Enfin, la comparaison des

déplacements chimiques du domaine globulaire dans les fragments M2-1 Δ [1-31] et M2-1[58-177] (Figure 43A) montre que la structure de ce domaine est globalement conservée, avec des réarrangements possibles entre les résidus S58 et A98, comme cela était le cas dans M2-1 Δ [35-58].

6 Comparaison avec des résultats cristallographiques

Si l'on compare les résultats obtenus avec les différentes constructions, il apparaît que c'est le doigt de zinc putatif qui était à l'origine des problèmes de solubilité et/ou de stabilité des fragments étudiés. Or en décembre 2013 l'équipe du Dr John Barr à l'université de Leeds a publié la structure cristallographique de la protéine M2-1 entière, présentée dans la Figure 44 (Tanner et al. 2014).

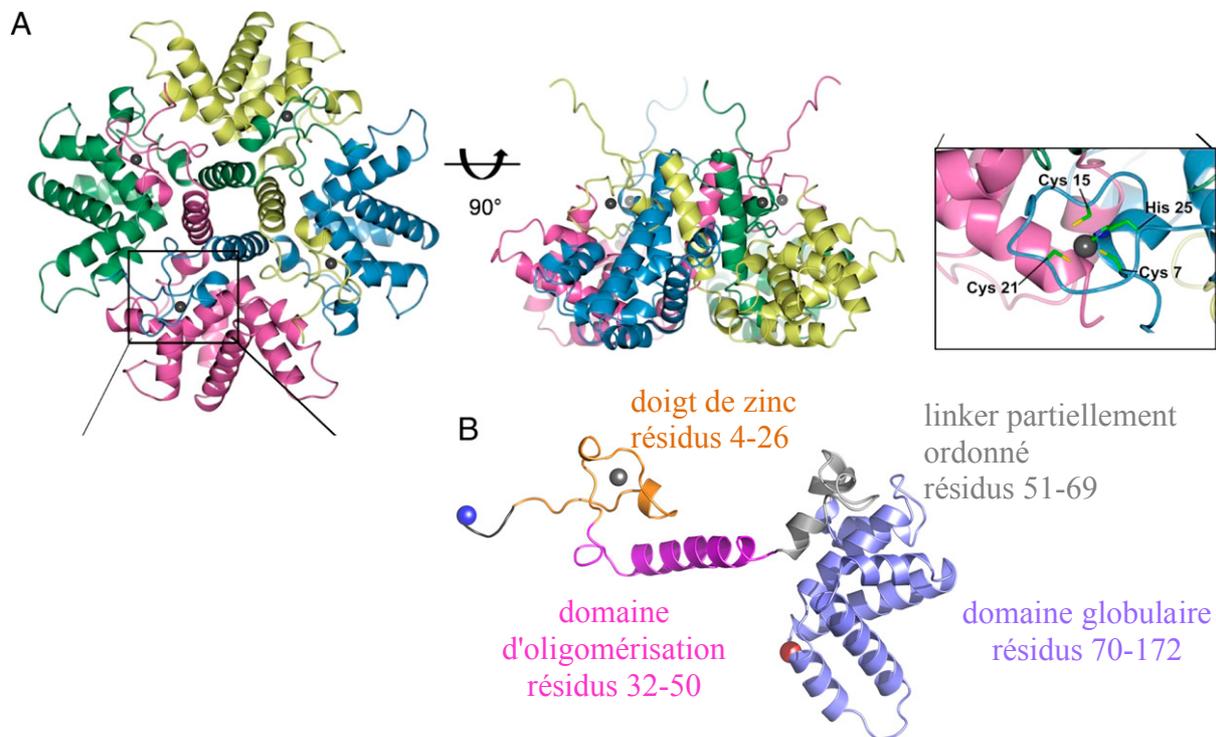


Figure 44 : Structure cristallographique de la protéine M2-1 du VRSH avec une résolution de 2,52 Å (PDB 4C3D) (Tanner et al. 2014). (A) Model ruban du tétramère de la protéine M2-1 du VRSH, avec zoom sur la partie N-terminale. Chaque protomère a une couleur différente et l'atome de zinc est représenté par une sphère grise. (B) Protomère de M2-1 où les régions sont colorées par domaine : domaine de liaison au zinc (orange); domaine de tétramérisation en hélices (magenta); domaine globulaire (bleu); domaine partiellement structuré (gris); résidus appartenant au vecteur de clonage (noir). Les extrémités N- et C-terminales sont marquées avec des sphères bleues et rouges, respectivement.

L'alignement de la structure cristallographique d'un protomère de M2-1 (Tanner et al. 2014) avec la structure du domaine globulaire résolue par RMN (Blondot et al. 2012) donné dans la Figure 45 montre que la structure en solution est bien retrouvée dans le tétramère. L'hélice notée α_2 , qui correspond à la première hélice du domaine globulaire, se prolonge de 2-3 résidus en N-terminal dans la protéine M2-1 entière. Cela peut être relié au morceau d'hélice supplémentaire qui est prédit à partir des déplacements chimiques pour M2-1 Δ [35-58] (Figure 38). La structure du tétramère confirme que le domaine N-terminal de M2-1 est bien un doigt de zinc et montre que le domaine d'oligomérisation est constitué de quatre hélices.

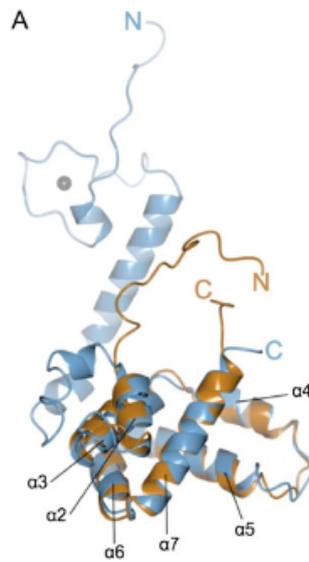


Figure 45 : Aligment de la structure cristallographique d'un protomère de la protéine M2-1 (Tanner et al. 2014) avec la structure RMN du domaine globulaire M2-1[58-177] . Les hélices sont numérotées selon les désignations faites pour la protéine M2-1 entière.

La structure du tétramère est très compacte. Les différents domaines sont proches les uns des autres et les protomères sont empilés de manière enchevêtrée : le domaine doigt de zinc du protomère i est en contact avec le domaine globulaire du protomère $i+1$, voire $i+2$ (Figure 44). Ceci pourrait expliquer pourquoi nous n'avons pas réussi à travailler sur des domaines isolés de la protéine. Ces interactions dans le tétramère pourraient stabiliser le domaine doigt de zinc.

D'autre part les domaines globulaires apparaissent comme indépendants du reste de la protéine dans les fragments M2-1 Δ [35-58] et M2-1 Δ [1-31] où manquent soit le domaine d'oligomérisation, soit le domaine doigt de zinc. Les résultats obtenus en RMN montrent donc que les trois domaines sont nécessaires pour que M2-1 adopte une structure compacte.

Chapitre 3 : Etude des interactions de la protéine M2-1 du VRSH par RMN

1 Introduction

Un deuxième volet de la caractérisation de M2-1 du VRSH a été l'étude de ses interactions au niveau atomique avec des partenaires identifiés précédemment : les ARN et la phosphoprotéine de VRSH (Tran et al. 2009). Pour cela, j'ai travaillé avec un fragment de M2-1 ne contenant que le domaine central globulaire, la région d'interaction avec ces deux partenaires ayant aussi déjà été réduite à ce domaine (Tran et al. 2009). Dans la mesure du possible, j'ai travaillé avec une protéine marquée en azote 15 et j'ai essayé de suivre les interactions avec des spectres HSQC ^1H - ^{15}N . Dans le cas des ARN, cela a suffi. Pour les interactions avec la phosphoprotéine, des expériences complémentaires ont dû être mises en œuvre. Pour pouvoir exploiter les résultats des expériences d'interaction par RMN, il est indispensable de disposer des attributions du domaine central de M2-1. Celles-ci avaient déjà été obtenues au laboratoire avant mon arrivée (Dubosclard et al. 2011).

2 Caractérisation des interactions de protéines par RMN

2.1 Régimes d'échange en RMN

La RMN est une méthode très efficace pour étudier les interactions entre des macromolécules à l'échelle atomique. En effet, les déplacements chimiques mesurés par RMN sont très sensibles aux changements structuraux et environnementaux. Ceci nous permettra d'identifier les acides aminés dont l'environnement local a été modifié en contact avec un partenaire.

Nous nous intéressons à un équilibre du type : $\text{P.L} \rightleftharpoons \text{P} + \text{L}$, où P désigne une protéine et L un ligand. Il est caractérisé par une constante thermodynamique de dissociation, K_d , et des constantes cinétiques d'association k_{on} et de dissociation k_{off} . On peut aussi définir une constante d'échange $k_{\text{ex}} = k_{\text{on}} \cdot L + k_{\text{off}}$. Si on considère qu'un noyau résonant à la fréquence ν_{libre} dans la protéine libre et à la fréquence $\nu_{\text{lié}}$ dans le complexe, l'évolution du signal de la protéine en présence d'une concentration

croissante de ligand dépend du rapport entre k_{ex} et $\Delta\nu$, où $\Delta\nu$ correspond à la différence des fréquences du noyau $\Delta\nu = |\nu_{libre} - \nu_{lié}|$.

Si $k_{ex} < \Delta\nu$, ceci correspond à un **régime d'échange lent**. On observe deux jeux de signaux, l'un correspondant au signal du noyau libre, l'autre à celui de la forme liée. Au cours du titrage l'intensité du signal de la forme libre diminue pendant que celui de la forme liée augmente.

Inversement, si $k_{ex} > \Delta\nu$, ceci correspond à un régime **d'échange rapide**, et le signal se déplace au cours du titrage avec une fréquence apparente donnée par

$$\nu = p_{libre} \cdot \nu_{libre} + p_{lié} \cdot \nu_{lié}$$

où p_{libre} et $p_{lié}$ sont les proportions de formes libres et liées.

Si $k_{ex} \approx \Delta\nu$, le **régime d'échange** est **intermédiaire**. Les signaux s'élargissent fortement et se déplacent en même temps avec une fréquence apparente intermédiaire entre ν_{libre} et $\nu_{lié}$.

Un échange lent correspond généralement à une forte affinité, un échange rapide à une affinité faible.

2.2 Cartographie des variations de déplacement chimique

S'il y a formation d'un complexe, l'environnement électronique des couples amides ^1H - ^{15}N avec lesquels le ligand interagit change. Lorsque le régime d'échange est rapide, on observe des variations au niveau des déplacements chimiques des pics correspondants, comme indiqué plus haut. On peut aussi définir des différences globales ou combinées de déplacements chimiques, calculées selon l'équation :

$$\Delta\delta^{1\text{H}^{15}\text{N}} (\text{ppm}) = \sqrt{(\delta_{1\text{H}} - \delta_{1\text{H}}^{\text{free}})^2 + ((\delta_{15\text{N}} - \delta_{15\text{N}}^{\text{free}})/10)^2}$$

où $\delta_{1\text{H}}$ et $\delta_{15\text{N}}$ désignent les déplacements chimiques ^1H et ^{15}N à un point donné du titrage en ppm, $\delta_{1\text{H}}^{\text{free}}$ et $\delta_{15\text{N}}^{\text{free}}$ les déplacements chimiques de la protéine libre. La différence de facteur de 10 entre les échelles du proton et de l'azote 15 provient de la différence des rapports gyromagnétiques entre ces deux noyaux.

En régime d'échange rapide la valeur de l'écart entre les déplacements chimiques des formes libre et liée permet alors de déterminer les résidus impliqués dans l'interaction entre les deux partenaires. Si l'on dispose de la structure tridimensionnelle de la protéine, ces résidus peuvent ensuite être cartographiés sur cette structure pour déterminer une région d'interaction.

En régime d'échange intermédiaire, ce sera l'intensité des signaux des résidus impliqués dans l'interaction entre les deux partenaires qui variera. On peut cartographier les variations d'intensité pour déterminer la région d'interaction.

2.3 Détermination d'une constante de dissociation

Dans le cas d'un régime rapide d'échange, la constante de dissociation K_d peut être déterminée en réalisant un titrage de la protéine marquée par un ligand non marqué, dans les mêmes conditions de pureté d'échantillon, de concentration de la protéine marquée, de tampon et de température. Le suivi est fait en mesurant des spectres HSQC. Les variations de déplacement chimique ^1H et ^{15}N sont ensuite ajustées en fonction de la quantité relative de ligand. Dans un modèle de liaison avec une stœchiométrie 1:1, on utilise une fonction définie comme suit :

$$\delta - \delta_{\text{free}} = \alpha * (\delta_{\text{bound}} - \delta_{\text{free}}) = B * (A + r - \sqrt{(A + r)^2 - 4r})$$

où $B = \frac{1}{2}(\delta_{\text{bound}} - \delta_{\text{free}})$, A est défini par $K_d = (A - 1) * [N - \text{NTD}]_{\text{tot}}$, et δ_{free} et δ_{bound} désignent les déplacements chimiques de la protéine libre et saturée en ligand.

2.4 Intensités des signaux en RMN

Un second paramètre à prendre en compte est l'intensité des raies. Celle-ci est globalement inversement proportionnelle à la vitesse de relaxation R_2 . En première approximation, si on suppose que la protéine est rigide, tous les protons de la protéine libre relaxent à la même vitesse $R_{2 \text{ libre}}$, tandis que tous ceux de la protéine liée relaxent à la même vitesse $R_{2 \text{ lié}}$. Plus l'objet est gros, plus la vitesse de relaxation R_2 est rapide, et donc plus l'intensité des raies est large et plus leur intensité est faible. L'évolution des largeurs de raies, en particulier celle des protons ne se situant pas à l'interface et pour lesquels la variation des déplacements chimiques est nulle dépendra des valeurs relatives de la différence des vitesses de relaxation transverse $\Delta R_2 = R_{2 \text{ lié}} - R_{2 \text{ libre}}$.

En **échange lent** ($k_{\text{ex}} < \Delta R_2$), les signaux de la protéine libre vont avoir une intensité dépendante de $R_{2 \text{ libre}}$ et ceux de la protéine liée dépendante de $R_{2 \text{ lié}}$. Pour un petit ligand, si la différence de masse n'est pas significative entre P et P.L, toutes les raies vont être semblables. Si le ligand est gros, alors les intensités des signaux de la forme liée sont beaucoup plus larges que ceux de la forme libre.

En **échange rapide** ($k_{\text{ex}} > \Delta R_2$), la vitesse de relaxation d'un signal est donnée par $R_2 = p_{2 \text{ libre}} \cdot R_{2 \text{ libre}} + p_{2 \text{ lié}} \cdot R_{2 \text{ lié}} + R_{\text{ex}}$.

Pour un petit ligand, $R_{2 \text{ libre}} = R_{2 \text{ lié}}$: les raies des noyaux non affectés vont avoir une intensité constante au cours du titrage, tandis que les raies des noyaux impliqués dans l'interaction pourront subir une variation d'intensité due au terme R_{ex} .

Pour un ligand de grande taille, tous les signaux vont être affectés et peuvent, à la limite, disparaître dès que $p_2 \text{ lié} \cdot R_{2 \text{ lié}}$ devient suffisamment grand. Dans ce cas il est difficile de distinguer les signaux dans et hors interface. En échange rapide, pour un ligand de grande taille, l'effet observé sur le spectre HSQC ^1H - ^{15}N de la protéine marquée ne sera donc pas forcément une variation de déplacement chimique, mais une variation d'intensité.

2.5 Expériences de saturation croisée transférée

En RMN, les expériences de perturbation classiques (de déplacement chimique et d'intensité) sont non seulement sensibles aux interactions avec un partenaire, mais aussi à des changements conformationnels dans des régions plus distales, entraînés par la fixation d'un partenaire. Pour mettre en évidence les régions d'interaction directe, des expériences plus spécifiques doivent être mises en œuvre, par exemple la saturation croisée transférée dans le cas de systèmes de haut poids moléculaire (Nakanishi et al. 2002). Le principe est illustré dans la Figure 46. La saturation des protons d'une grosse protéine partenaire par diffusion de spin, entraîne le transfert de saturation sur la zone de contact avec la protéine observée, ce qui se traduit par une baisse de l'intensité des signaux. Pour cela il faut que l'échange entre formes libres et liée de la protéine observée (I) soit rapide et que celle-ci soit en excès par rapport à son partenaire (II). Pour éviter que cet effet ne se propage pas sur l'ensemble de la protéine observée, celle-ci est marquée au deutérium qui empêche la diffusion de spin.

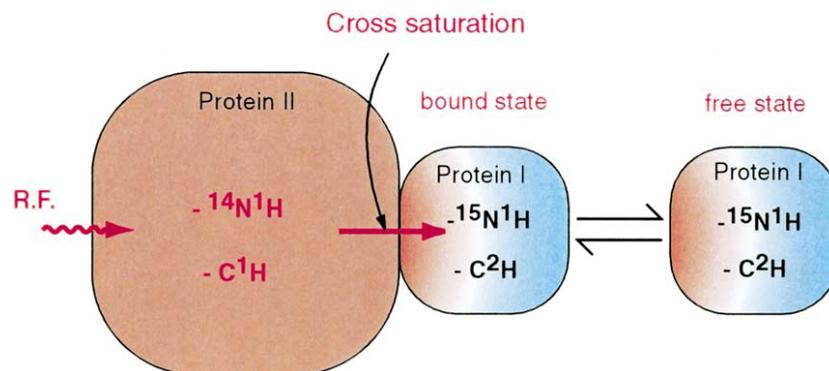


Figure 46: Principe de l'étude des interactions entre deux protéines par saturation croisée transférée (TCS) et marquages isotopiques.

3 Préparation des échantillons de protéine et d'ARN

3.1 Expression et purification du domaine central de M2-1

Dans un premier temps le fragment de protéine correspondant au domaine central de M2-1 a été produit. Le plasmide pGEX-M2-1[58-177] a été fourni par l'équipe de l'INRA. Cette protéine est exprimée dans la souche *E. coli* BL21(DE3) sous la forme d'une protéine de fusion avec la glutathion-S-transférase (GST) contenant un site de clivage de la thrombine. Elle doit être marquée par des isotopes pour l'étude par RMN. Elle a donc été produite dans un milieu de culture minimum M9 contenant du $^{15}\text{NH}_4\text{Cl}$ comme source d'azote 15 et du glucose- ^{13}C comme source de carbone 13. Le protocole détaillé de production est donné dans les annexes à la fin du manuscrit. Le gel d'expression présenté dans la Figure 47A montre une bande intense, entre 30 et 45 kDa après induction, qui correspond à la protéine M2-1[58-177] (13.5 kDa) en fusion avec la GST (~26 kDa).

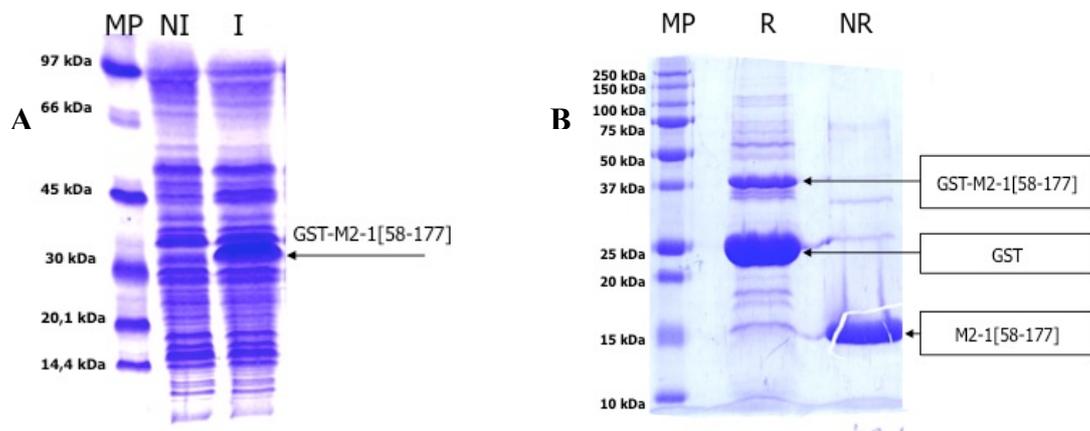


Figure 47 : Gel d'expression de polyacrylamide obtenu par électrophorèse en milieu dénaturant (SDS-PAGE) pour M2-1[58-177]. MP = marqueur de poids moléculaire. NI = extrait bactérien avant induction par l'IPTG. I = extrait bactérien après induction. La bande correspondant à GST-M2-1[58-177] est indiquée par une flèche. (B) Gel de protéine SDS-PAGE après coupure à la thrombine (MP = marqueur de poids moléculaire, R = résine GSH-sépharose, NR = fraction non retenue qui contient la protéine sans la GST).

La purification est faite sur une résine glutathion sépharose qui fixe l'étiquette GST. Une fois la résine lavée, elle est incubée avec de la thrombine biotinylée, une protéase qui reconnaît la séquence Leu-Val-Pro-Arg-Gly-Ser et coupe la liaison peptidique entre Arg et Gly située entre la protéine d'intérêt et la GST. L'efficacité de la coupure et la pureté de la préparation de protéine sont vérifiées par SDS-PAGE en comparant la résine et la fraction non

retenue où l'on s'attend à trouver la protéine M2-1[58-177] (Figure 47B). La fraction non retenue contient la protéine pure avec la thrombine, qui est ensuite traitée avec une résine streptavidine qui fixe la thrombine biotinylée.

La concentration de protéine est mesurée avec un test colorimétrique (Bradford) et par spectrophotométrie en mesurant l'absorption de la protéine à 280 nm. Un échantillon pour une étude d'interaction par RMN a un volume d'environ 200 μL à une concentration de 50 μM .

3.2 Synthèse et la purification des ARN

Différentes séquences d'ARN (10-15 nucléotides) ont été synthétisées en phase solide pour réaliser des expériences d'interaction entre M2-1[58-177] et des ARN. La synthèse est basée sur quatre étapes (déprotection, couplage, "capping" et oxydation), illustrées sur la Figure 48.

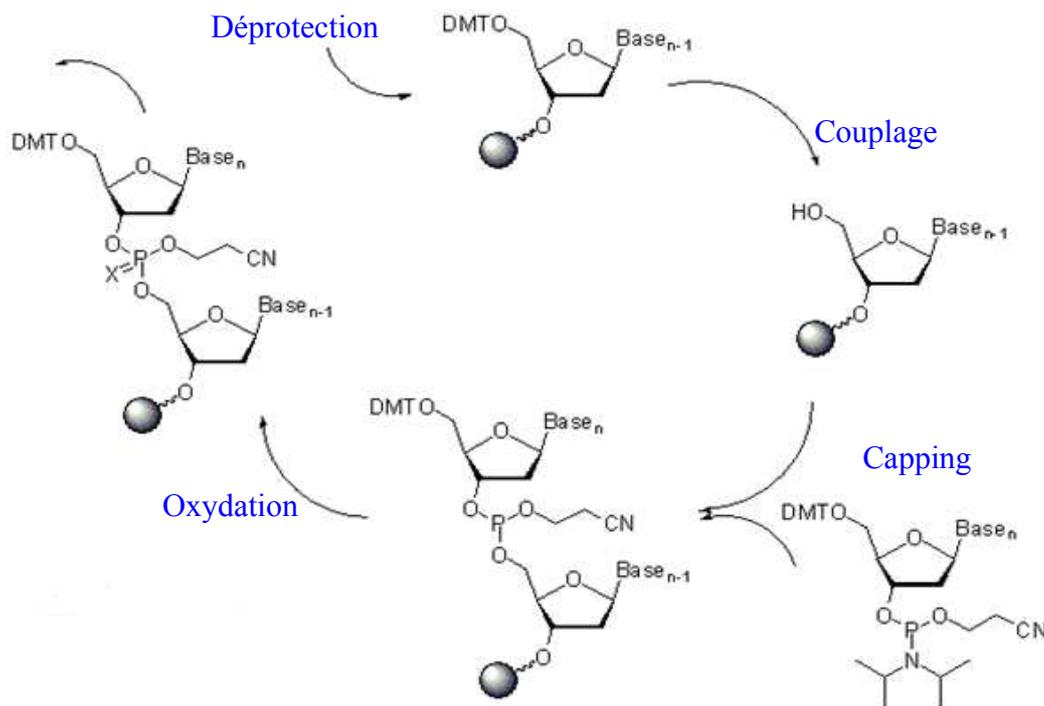


Figure 48 : Cycle de synthèse de l'ARN en phase solide.

Une fois synthétisés, les produits ont été purifiés par chromatographie anionique sur une colonne DEAE-Sepharose. L'élution a été réalisée avec un gradient linéaire de 0 à 1 M de NaCl puisque l'ARN est chargé négativement de par ses groupements phosphates. Cette étape est importante pour éliminer des fragments d'ARN plus courts issus de l'arrêt prématuré de la synthèse. Un exemple de purification d'un ARN est présenté dans la Figure 49.

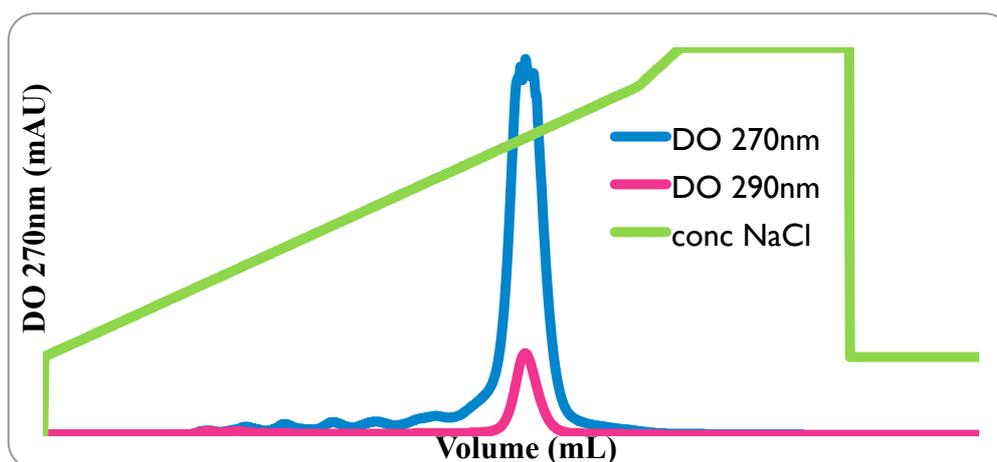


Figure 49 : Gradient de concentration de NaCl (entre 0 et 1M) et chromatogramme obtenu sur colonne anionique au cours de la purification de l'ARN "leader positif".

Comme mentionné, la synthèse peut être interrompue et donc des fragments d'ARN plus petits sont produits. On les voit pour des volumes entre 20 et 45 mL dans la Figure 49, le segment d'ARN recherché correspondant au pic majoritaire, avec le volume de rétention le plus grand. Les acides nucléiques sont observés à deux longueurs d'onde, 270 et 290 nm. Celles-ci sont décalées par rapport à la longueur d'onde d'absorbance maximale des acides nucléiques à 260 nm pour éviter de saturer le détecteur. La quantité d'ARN pur obtenue à la fin est de l'ordre de 30 DO à 260 nm. L'ARN est stocké et utilisé sous forme lyophilisée.

4 Etude des interactions du domaine central de M2-1 avec des ARN

4.1 Etude préliminaire avec un ARNt de levure

Pour avoir une première idée de l'influence d'un ARN sur le spectre HSQC ^1H - ^{15}N de M2-1[58-177], un ARNt de levure commercial (Roche) a été ajouté avec un rapport molaire de 1:1 environ.

L'interaction avec l'ARNt de levure montre une perturbation sur quelques pics de la protéine M2-1[58-177] (Figure 50). Il y a des variations de déplacement chimiques en plus d'une perte d'intensité. Certains pics sont élargis jusqu'à perte du signal. On observe bien la fixation de l'ARNt sur M2-1[58-177]. Les variations de déplacement chimique, résidu par résidu sont reportées dans la Figure 51, et permettent de mettre en évidence plusieurs régions

qui pourraient être des zones d'interaction : entre G58 et R68, entre S94 et A98, entre L148 et V156.

Après ajout de la RNase A, le spectre revient quasiment à l'état initial en l'absence d'ARNt, mais pas complètement. Les pics qui présentaient une variation de déplacement chimique se trouvent à un état intermédiaire entre les deux états (sans l'ARN et avec l'ARN). Le fait que nous ne retrouvons pas le spectre de référence après digestion à la RNase A s'explique peut-être par le fait que certaines régions de l'ARN de levure soient résistantes à la RNase A et que ces régions peuvent se fixer sur M2-1[58-177].

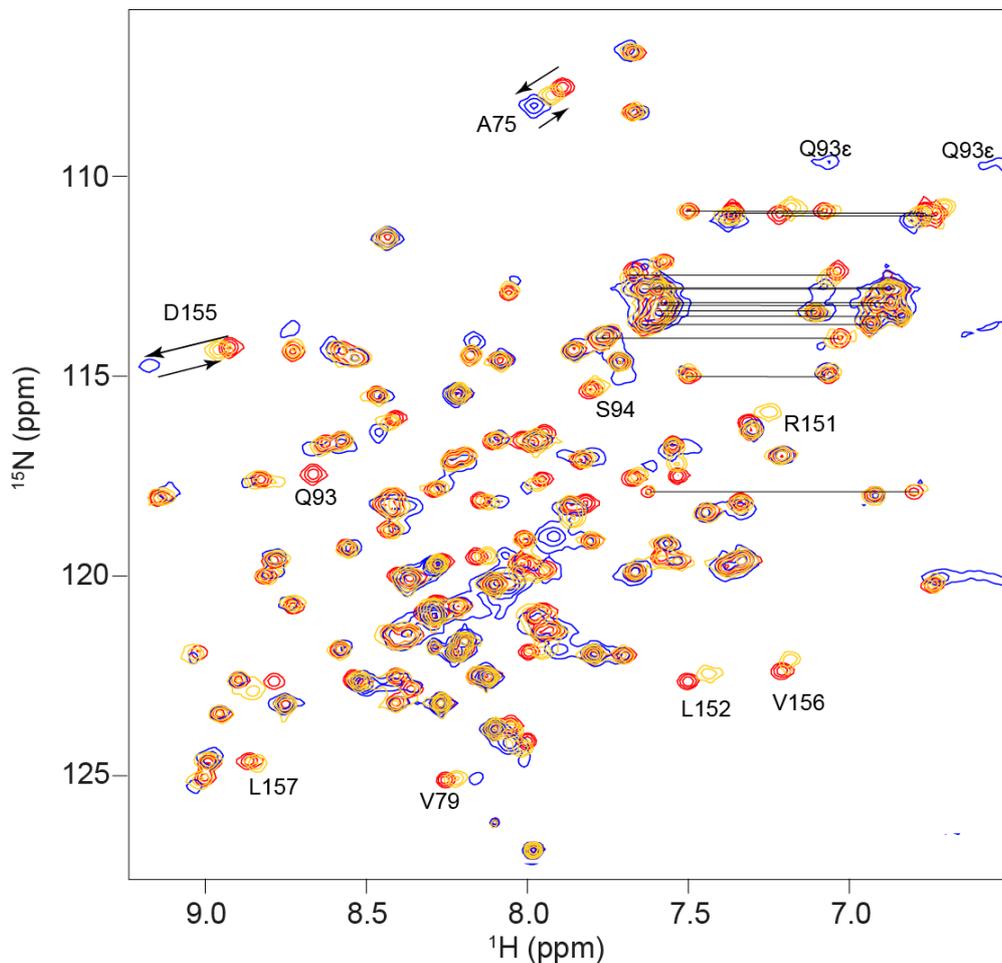


Figure 50 : Superposition de spectres ^1H - ^{15}N -HSQC de M2-1[58-177] en présence et en absence d'un ARN de levure commercial avec un rapport molaire de 1:1 ($50\ \mu\text{M}$, $T = 298\ \text{K}$, $B_0 = 14,1\ \text{T}$, $600\ \text{MHz}$). Spectre de référence (rouge), avec l'ARN (bleu), et après digestion par la RNase A (orange). Les pics correspondant aux résonances des chaînes latérales de l'asparagine et de la glutamine sont reliés par des lignes entre les deux protons géminaux de l'amide. Les attributions des résidus avec des variations de déplacement chimique significatives sont indiquées.

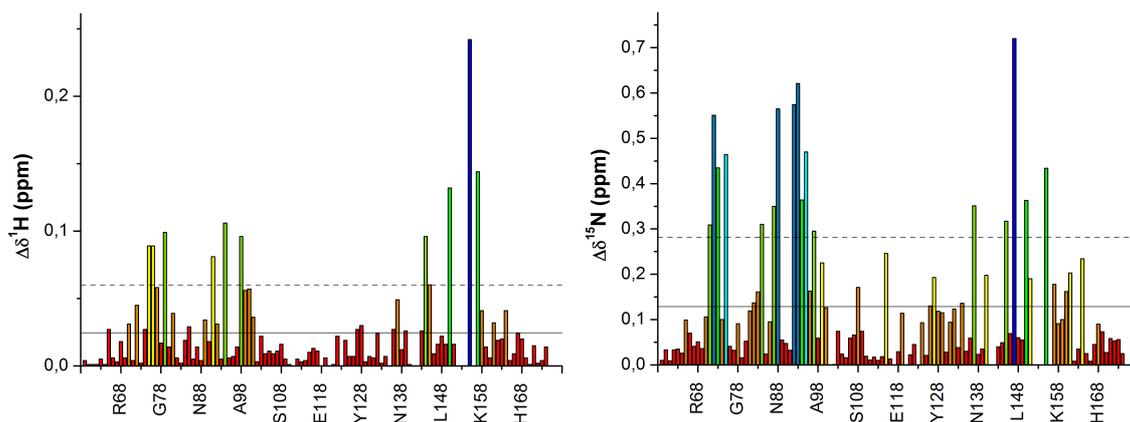


Figure 51 : Variations de déplacement chimique ^1H et ^{15}N , résidu par résidu pour M2-1[58-177] en présence de l'ARN de levure. Le code couleur correspond à l'amplitude des variations : les valeurs les plus basses sont en rouge, les plus élevées en bleu. La moyenne et l'écart type sont indiqués par des traits pleins et en pointillés.

4.2 Liaison de séquences d'ARN courtes correspondant à des signaux de transcription du VRS

4.2.1 Choix de courtes séquences d'acides nucléiques pour des expériences d'interaction par RMN

Cette première série d'expériences a fourni une preuve de concept pour travailler avec des ARN ayant une signification biologique pour la transcription du VRS.

La transcription s'effectue selon un mécanisme séquentiel d'"arrêt-redémarrage" au cours duquel la polymérase reconnaît des séquences spécifiques situées dans les régions intergéniques. Il s'agit de séquences d'arrêt (Gene End, GE) et de redémarrage (Gene Start, GS) de la transcription (Sutherland et al. 2001). Les séquences de terminaison de la transcription « GE » sont des séquences semi-conservées de 13 nucléotides (3'-UC(A/C)(A/U)UXXXUUUUX-5'), responsables de la polyadénylation et la libération de l'ARNm. La séquence de démarrage de la transcription « GS » est conservée et constituée de 9 nucléotides (3'-CCCCGUUUA-5'). Elle permet le démarrage de la transcription et le coiffage des ARN messagers. Les jonctions intergéniques sont indiquées sur la Figure 52.

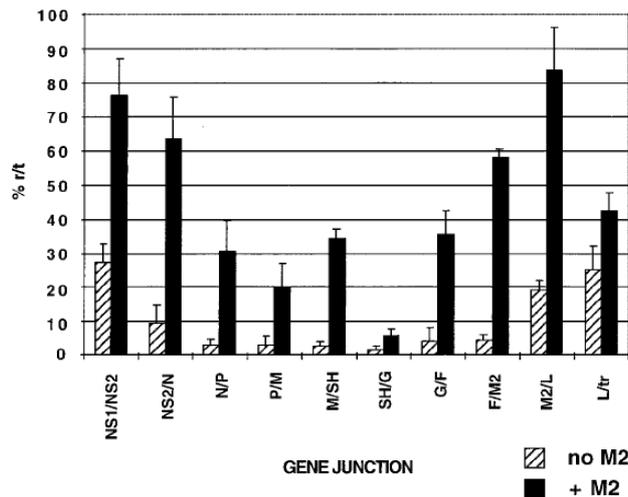


Figure 52 : La terminaison de la transcription est sensible aux jonctions des gènes et à la présence de la protéine M2-1 (Hardy et al. 1999). Les différentes jonctions des gènes sont représentées sur l'axe X. L'axe Y donne le pourcentage de l'abondance molaire des transcrits dicistroniques par rapport à des transcrits monocistroniques, obtenus avec un système de miniréplicon, en présence et en absence de la protéine M2-1. Un transcrit dicistronique reflète l'échec de la terminaison de transcription.

La polymérase a tendance à se dissocier de l'ARN génomique au niveau des séquences GE, mais une fois dissociée elle ne peut pas ré-initier la transcription au niveau de la séquence GS du gène qui suit, ce qui conduit à un gradient de transcription qui diminue de l'extrémité 3' à l'extrémité 5' du génome (Fearn and Collins 1999). Pour tous les pneumovirus, l'activité du complexe RdRp en mode transcription dépend de la protéine M2-1, spécifique de cette sous-famille. M2-1 est un facteur d'anti-terminaison de la transcription permettant à la fois à la polymérase de finir la synthèse d'un ARN messager et de traverser les régions intergéniques en produisant des ARN messagers polycistroniques. Il est donc très important pour une synthèse efficace de tous les ARNm (Collins et al. 1996).

Il a été montré que M2-1 agissait au niveau des jonctions des gènes, mais que l'efficacité de réduire la terminaison de la transcription en présence de M2-1 variait entre ces différentes jonctions, comme illustré dans la Figure 52 (Fearn and Collins 1999; Hardy et al. 1999). Par exemple M2-1 réduit la terminaison de la transcription au niveau de la jonction entre les gènes de F et M2. Cependant, les mécanismes par lesquels M2-1 agit restent à éclaircir. En particulier il n'est pas clair si M2-1 reconnaît des séquences spécifiques. Des résultats divergents ont été publiés. D'un côté il a été montré dans le groupe du Dr Gail Wertz qu'au cours d'une infection par le VRS, M2-1 liait essentiellement des ARN messagers viraux, mais que ces ARN ne contenaient pas la séquence Leader (Cartee and Wertz 2001). Un autre

groupe a montré que M2-1 liait de façon préférentielle de courtes séquences virales contenant la séquence Leader positive *in vitro*, mais qu'elle liait aussi des séquences d'ARN plus longues sans spécificité de séquence (Cuesta et al. 2000).

Concernant les tests d'interaction entre M2-1 et l'ARN, nous avons choisi de tester un large panel d'ARN en raison des résultats contradictoires sur la spécificité et l'affinité, publiés jusqu'alors (Hardy et al. 1999; Cuesta et al. 2000; Cartee and Wertz 2001). Toutes les séquences d'ARN ont été synthétisées au laboratoire.

Une première série de mesures a été faite avec des séquences d'ARN génomique du VRS, de polarité négative : la séquence d'entrée de la polymérase sur l'ARN génomique (Leader), une séquence riche en uridine en 5' du Leader (polyU), la séquence GS et plusieurs séquences GE. Ces séquences avaient été testées dans l'hypothèse où ce serait directement l'ARN génomique (une fois décapsidé) qui pourrait être reconnu par M2-1 et modulerait son activité d'anti-terminaison de la transcription.

Une autre hypothèse est que ce serait l'ARN messager naissant qui pourrait en être à l'origine. Nous avons donc testé les ARN complémentaires avec des signaux de transcription de sens positif : le site d'entrée de la polymérase (Leader positif), une courte séquence riche en adénosine (polyA) ainsi que les GE-SH et GE-F positifs. Les oligonucléotides ont été conçus pour minimiser l'auto-association ou la formation de structures secondaires. Les calculs ont été faits sur le serveur Web *μ*fold (<http://mfold.rna.albany.edu/?q=mfold>) et l'absence d'auto-association a été évaluée par RMN ¹H en vérifiant l'absence de signaux de protons imino de guanine ou d'uridine entre 12 et 14 ppm. Dans le cas du leader positif et de GE-F positif, nous avons néanmoins observé deux signaux de protons imino, ce qui indiquerait la formation possible de duplex ou d'auto-association de ces ARN (Figure 53). Mais ces signaux restent faibles comparés à celui obtenu avec un duplex d'ARN constitué de brins GE-F négatifs et positifs.

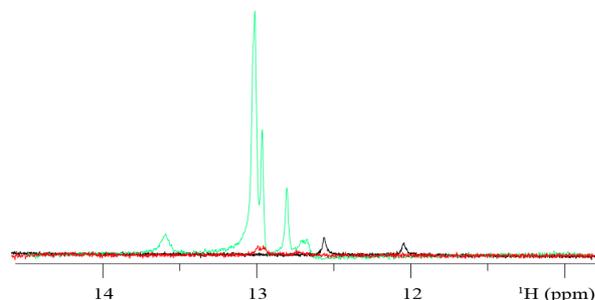


Figure 53 : Extrait de la région des protons imino des spectre RMN ¹H des ARN « GE-F positif » (rouge), « Leader positif » (noir) et « GE-F double brin » (vert) à 200 μ M montrant l'appariement de paires de bases dans ces ARN ($B_0 = 14.1$ T, $T = 293$ K).

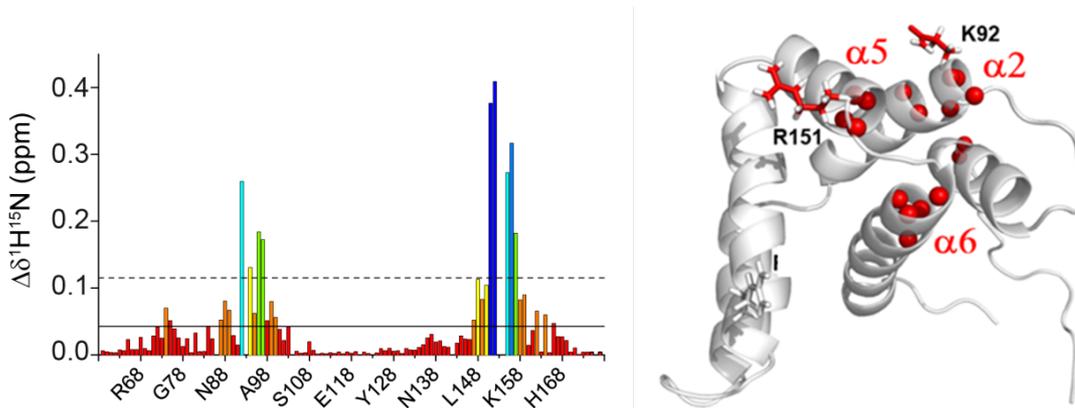


Figure 55 : Variations de déplacement chimique résidu par résidu induit par l'ARN « polyA » dans le spectres ^1H - ^{15}N -HSQC de M2-1[58-177] avec un rapport molaire ARN/protéine de 6/1. Le code couleur correspond à l'amplitude des variations : les valeurs les plus basses sont en rouge, les plus élevées en bleu. Cartographie des variations de déplacement chimique sur la structure du domaine central de M2-1. Les atomes d'azote des liaisons amides des résidus qui présentent une variation significative sont indiqués par des sphères rouges.

Des variations de déplacement chimique dans les expériences ^1H - ^{13}C HSQC ont été observées pour des résidus dans la même région (Figure 56).

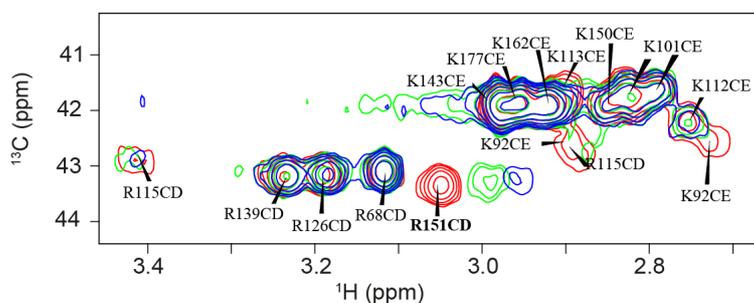


Figure 56 : Zoom sur les résonances des chaînes latérales des arginines ($\text{C}\delta\text{-H}\delta$) et des lysines ($\text{C}\epsilon\text{-H}\epsilon$) dans les spectres ^1H - ^{13}C -HSQC de M2-1[58-177] en présence de concentrations croissantes de l'ARN « Leader » ($50\ \mu\text{M}$, $T = 293\text{K}$, $B_0 = 24,3\ \text{T}$). Spectre de référence (rouge), en présence de 1 équivalent (vert) et 2 équivalents (bleu)

Les profils des variations de déplacement chimique étaient semblables pour tous les ARN testés, les résidus délimités par les hélices $\alpha 2$ (K92), $\alpha 5$ (K150, R151) et $\alpha 6$ (K158, K159, K162, K169) forment une surface continue chargée positivement, ce qui est cohérent avec une surface de liaison à l'ARN, comme indiqué dans la Figure 57.

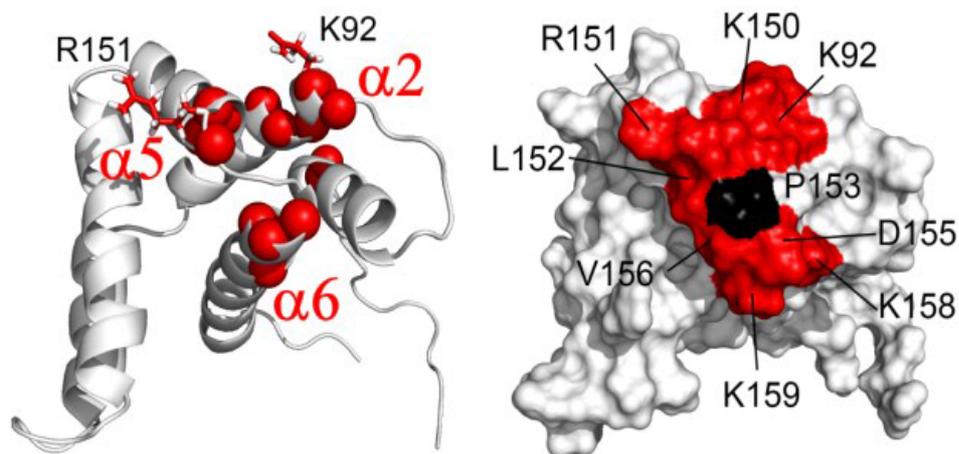


Figure 57 : Récapitulatif des expériences d'interaction par RMN de M2-1[58-177] avec de courtes séquences d'ARN par cartographie des variations de déplacement chimique sur la structure du domaine central de M2-1. Les atomes d'azote des liaisons amides des résidus qui présentent une variation significative pour toutes les séquences d'ARN génomique testées (données dans le Tableau 3) sont indiqués par des sphères rouges. Les chaînes latérales de R151 et K92, pour lesquels l'ARN a induit des variations de déplacement chimique $^{13}\text{C}\delta\text{-}^1\text{H}\delta$ et $^{13}\text{C}\epsilon\text{-}^1\text{H}\epsilon$ significatives, sont représentées en bâton. La surface formée par l'ensemble de ces résidus est colorée en rouge. P153, pour laquelle aucune information n'est disponible, est affichée en noir.

4.2.3 Affinités comparées de différents ARN

4.2.3.1 Détermination de la constante de dissociation

Dans le cas d'un régime d'échange rapide, la constante de dissociation K_d peut être déterminée en ajustant les variations de déplacement chimique en ^1H et ^{15}N en fonction de la quantité relative de ligand. Un exemple de courbe de variation de déplacement chimique est donné dans Figure 58 pour le résidu N88 dans le cas de l'ARN « polyA ».

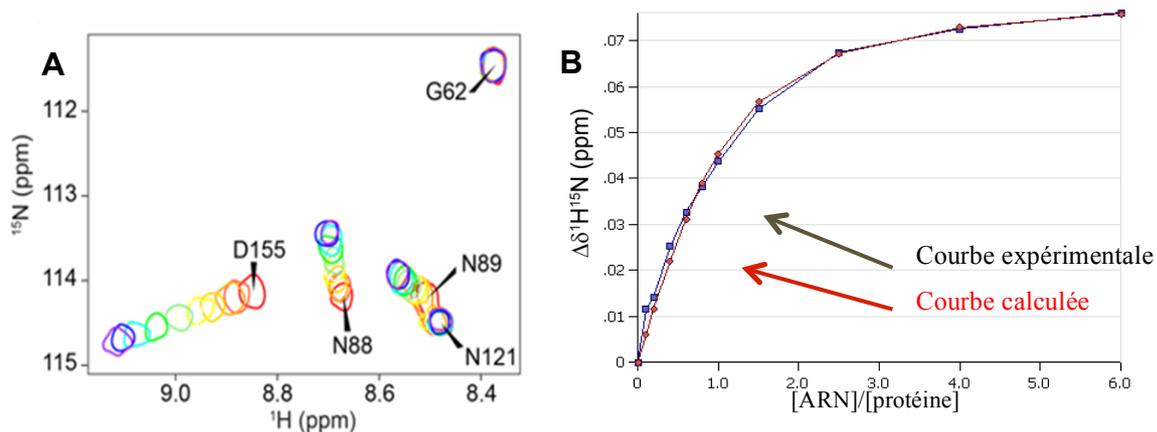


Figure 58 : (A) Extrait du spectre HSQC du titrage avec l'ARN polyA pour illustrer les variations de déplacement chimique de quelques résidus (G62, N88, N89, N121 et D155) en présence de l'ARN. (B) Courbes expérimentale et calculée des variations de déplacement chimique combinées ^1H et ^{15}N en fonction du rapport $[\text{ARN}]/[\text{protéine}]$ dans le cas de N88. La constante de dissociation est $\sim 20 \mu\text{M}$.

4.2.3.2 Comparaison des affinités déterminées avec différents acides nucléiques

Les courbes de titrage sont données dans la Figure 59. Le récapitulatif des valeurs des constantes de dissociation pour les différents oligonucléotides est donné dans le Tableau 3. Les valeurs de K_d s'étendent de $2,5 \mu\text{M}$ à plus de $600 \mu\text{M}$. Pour les séquences riches en uridines de sens négatif les valeurs varient entre 75 et $125 \mu\text{M}$, alors que les séquences riches en adénosines de sens positif (Leader positif, GE-SH positif, GE-F positif et polyA) présentent des affinités globalement meilleures pour M2-1[58-177] : de l'ordre de $2,5$ - $22 \mu\text{M}$.

Le contrôle que nous avons effectué avec la séquence courte d'ADN simple brin, équivalente à la séquence d'ARN GE-F négatif, montre que l'ADN se lie à M2-1[58-177] avec une affinité ($250 \mu\text{M}$) légèrement inférieure à la séquence d'ARN ($85 \mu\text{M}$). Enfin, le double brin (GE-F double-brin) présente une meilleure affinité que le simple brin.

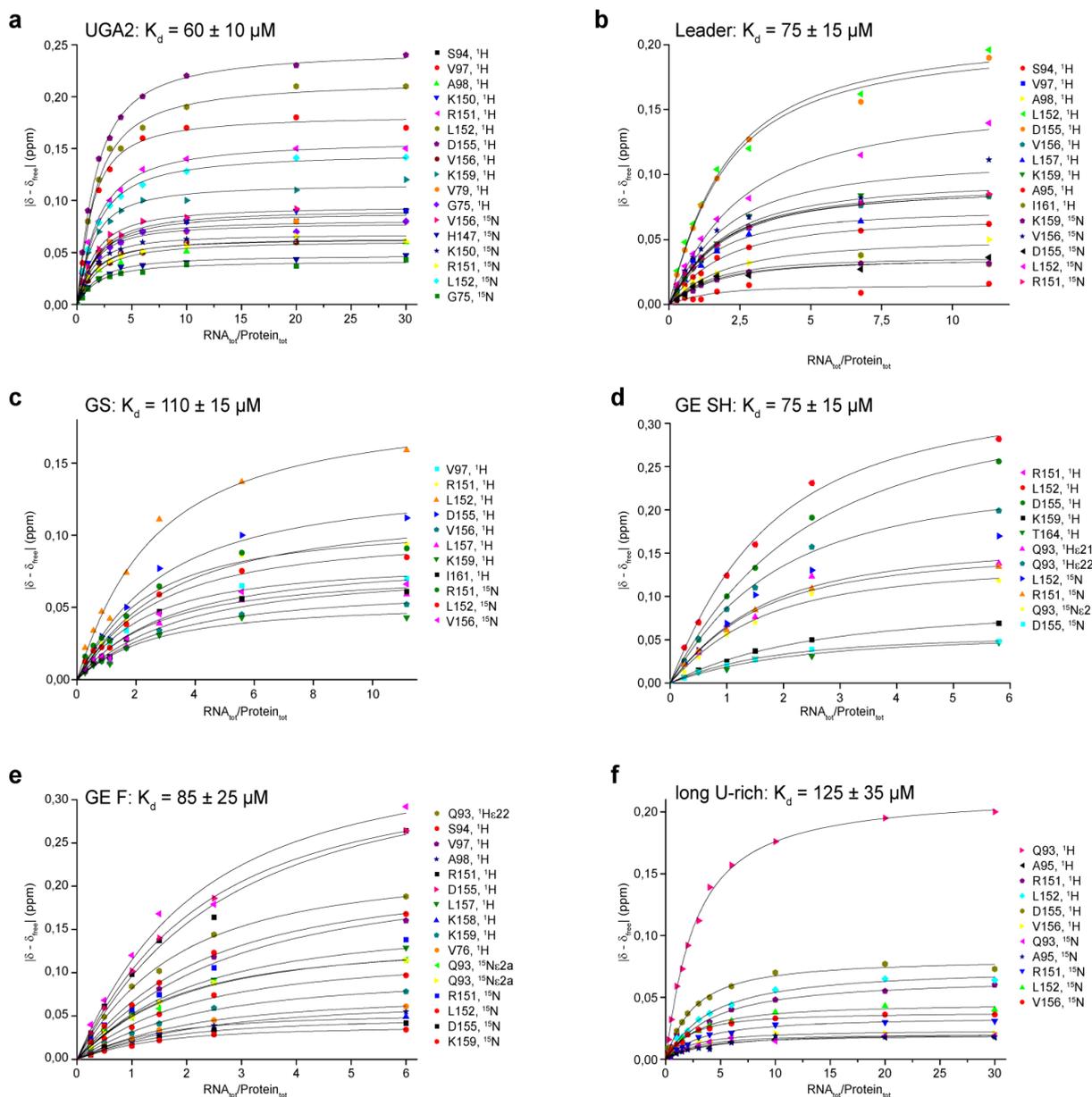


Figure 59 : Récapitulatif des expériences de titrage de la protéine M2-I[58-177] avec les ARN et détermination de constantes de dissociation (K_d) par ajustement des variations de déplacement chimique en ^1H et ^{15}N du squelette peptidique ou les chaînes latérales de la glutamine (notés H, N, H ϵ 21, H ϵ 22 et H ϵ 2 respectivement), en fonction de la concentration relative de ligand par rapport à la concentration de la protéine, avec un modèle de liaison avec une stœchiométrie 1:1. Seuls les résidus avec des perturbations significatives ont été utilisés. Les mesures ont été effectuées à 293 K sur un spectromètre RMN à 14,1 T. Les spectres HSQC ^1H - ^{15}N ont été acquis avec des ajouts croissants d'ARN à des solutions de M2-I[58-177] à 50-60 μM . Les courbes correspondent aux séquences d'ARN : (a) UGA2 (épingle à cheveux), (b) Leader, (c) Gene Start NS1, (d) Gene End SH, (e) Gene End F, (f) poly U long. Les K_d sont indiquées pour chaque ARN.

Désignation	Séquence d'oligonucléotides	K _d (μM)
Leader négatif	3'-UGC GC UUUUUU AC GC-5'	75 ± 15
polyU long	3'-UUG GUUUUUUU ACCC-5'	125 ± 35
polyU court	3'- GG UUUUUU UA-5'	> 600
GS négatif	3'-CCCCG UUU AU-5'	110 ± 15
GE_SH négatif	3'-UCA AUU AA UUU UU-5'	75 ± 15
Leader positif	5'-ACG CG AAAAA AU-3'	2.5 ± 1.5
polyA	5'-CC AAAAA AU-3'	22 ± 6
GE_SH positif	5'-AGU UA AUU AAAA-3'	13 ± 5
GE_F positif	5'-AGU UA AUU AAAA-3'	11 ± 5
GE_F négatif	3'-UCA AU AUU UU-5'	85 ± 25
GE_F double-brin	double-brin	4.5 ± 2.5
GE_F ADN négatif	3'-TCA AT ATAT TTT-5'	250 ± 70

Tableau 3 : Récapitulatif des valeurs des constantes de dissociations K_d déterminées par RMN pour l'association de différents oligonucléotides avec M2-1[58-177].

4.3 Discussion

L'étude de l'interaction avec l'ARN montre trois clusters constituant une surface continue chargée positivement sur la protéine M2-1 : K92 à V97 (α2), L149 à L152 (α5) et D155 à K159 (α6) (Figure 57). La spécificité de l'interaction avec l'ARN peut être remise en question si la seule présence de résidus chargés positivement suffit à lier une molécule chargée négativement telle qu'un ARN. Or l'affinité observée avec une séquence d'ADN est plus faible que pour la même séquence en ARN (250 μM vs 85 μM). Cette interaction n'est pas pertinente dans un contexte physiologique, puisque M2-1 est retrouvée exclusivement dans le cytoplasme des cellules infectées, mais indique que l'interaction avec le squelette phosphaté d'un oligonucléotide n'y contribue pas uniquement. L'interaction entre la protéine M2-1 et l'ARN a un rôle essentiel dans le cycle viral, ce qui pousse à penser que la nature de cette interaction est quand-même spécifique.

Les données de la littérature décrivent une grande variabilité d'ARN pouvant interagir avec la protéine M2-1. C'est pour cela que nous avons déterminé leur affinité par RMN avec plusieurs séquences d'ARN. Nous avons choisi de nous focaliser sur des séquences

spécifiques du VRS, en particulier le leader ainsi que les GE du gène SH et du gène F, testées à la fois en sens positif et négatif. Nos résultats montrent une affinité plus forte de la protéine M2-1 pour les séquences d'ARN de sens positif (K_d 2,5-22 μ M) que pour les séquences de sens négatif (K_d 50-100 μ M). Cette observation serait en faveur d'une interaction entre la protéine M2-1 avec les ARNm viraux plutôt qu'avec le brin d'ARN matrice. De plus, cela corrobore les données publiées par le groupe du Dr Gail Wertz (Cartee and Wertz 2001) dans lesquelles des expériences de co-immunoprécipitation ont permis de démontrer une interaction préférentielle entre la protéine M2-1 et des ARNm viraux, réalisées sur des lysats de cellules infectées par le VRSH.

Cuesta et collaborateurs ont mis en évidence une interaction forte avec un K_d de 90 nM entre la protéine M2-1 et des séquences d'ARN courtes de sens positif contenant la totalité du leader (44 nt) (Cuesta et al. 2000). Nos résultats sont en accord avec ces résultats puisqu'avec les douze premières bases de cette séquence, nous avons obtenu le K_d le plus petit (2.5 μ M). Cependant un ordre de grandeur sépare ces deux valeurs. Ceci peut résider dans la taille de la séquence utilisée, mais aussi dans la méthode de quantification employée. En effet, leur technique consiste à calculer le K_d sur gel retard en suivant la formation du complexe entre la protéine M2-1 et les ARN d'intérêt radiomarqués.

Par ailleurs nous avons observé que l'affinité de la protéine M2-1 pour un ARN double brin correspondant au GE du gène F était plus élevée que celle avec l'ARN simple brin de sens positif (K_d de 4,5 et 11,5 μ M). Ce résultat peut être interprété de deux manières. Soit la protéine M2-1 est capable d'interagir avec l'ARNm tout en restant accrochée à la matrice, soit la protéine M2-1 possède une activité supplémentaire antivirale en séquestrant les ARN viraux doubles brins afin qu'ils ne soient pas détectés par les récepteurs cellulaires de l'immunité innée (PRR = Pattern Recognition Receptor), capable d'identifier des motifs moléculaires caractéristiques des pathogènes, comme la protéine VP35 du virus Ebola (Cardenas et al. 2006; Leung et al. 2010).

Par contre, nos résultats ne permettent pas d'expliquer la sélectivité de M2-1 au niveau des séquences intergéniques GE pour empêcher la terminaison de la transcription (Hardy et al. 1999). En utilisant un système de miniréplicon, Hardy et collaborateurs ont montré que les GE SH/G et F/M2 sont tous deux des terminateurs forts, mais SH/G est insensible à M2-1, contrairement à F/M2 M2-1 (Figure 52). Mais nous n'avons trouvé aucune différence

d'affinité entre les GE du gène SH et du gène F. Les différences observées pourraient s'expliquer du fait des techniques différentes employées, mais la différence essentielle entre notre étude et celle décrite dans (Hardy et al. 1999) réside dans les constructions utilisées. Nous avons utilisé une protéine M2-1 tronquée (résidus 58-177) pour notre étude par RMN, tandis que Hardy et collaborateurs ont utilisé la protéine M2-1 entière pour des expériences en cellules eucaryotes. En admettant que le système du miniréplicon soit plus proche de ce qui se passe lors d'une infection par le VRS *in vivo*, on peut imaginer plusieurs hypothèses concernant la spécificité d'interaction entre la protéine M2-1 selon la nature de la séquence GE.

La spécificité peut être liée à la partie N-terminale de la protéine M2-1 (résidus 1-58), absente dans nos expériences de RMN. Dans ce cas, on peut supposer qu'une interaction peut avoir lieu au niveau de cette région ou bien que les domaines N-terminaux de la protéine M2-1, même s'ils ne sont pas directement impliqués dans l'interaction avec l'ARN, sont nécessaires à sa fonction de reconnaissance. Cette hypothèse est appuyée par le fait que l'intégrité du motif "3Cys-1His" dans le domaine N-terminal est essentielle à la fonction d'antitermineur de la protéine M2-1 (Hardy and Wertz 2000). Dans le chapitre précédent il a aussi été montré que ce domaine possédait potentiellement des propriétés de liaison aux acides nucléiques.

La tétramérisation de la protéine M2-1 pourrait donner lieu à un effet coopératif en présentant une surface de liaison plus importante.

5 Etude des interactions du domaine central de M2-1 avec la phosphoprotéine du VRS

5.1 Résultats antérieurs en titration calorimétrique isotherme

L'interaction du domaine globulaire de M2-1 avec la phosphoprotéine avait été étudiée précédemment grâce à des expériences de titration calorimétrique isotherme (ITC, Isothermal Titration Calorimetry), illustrées dans la (Figure 60) (Blondot et al. 2012). Une valeur de 3 μM avait été déterminée pour la constante de dissociation avec une stœchiométrie de 1:1.

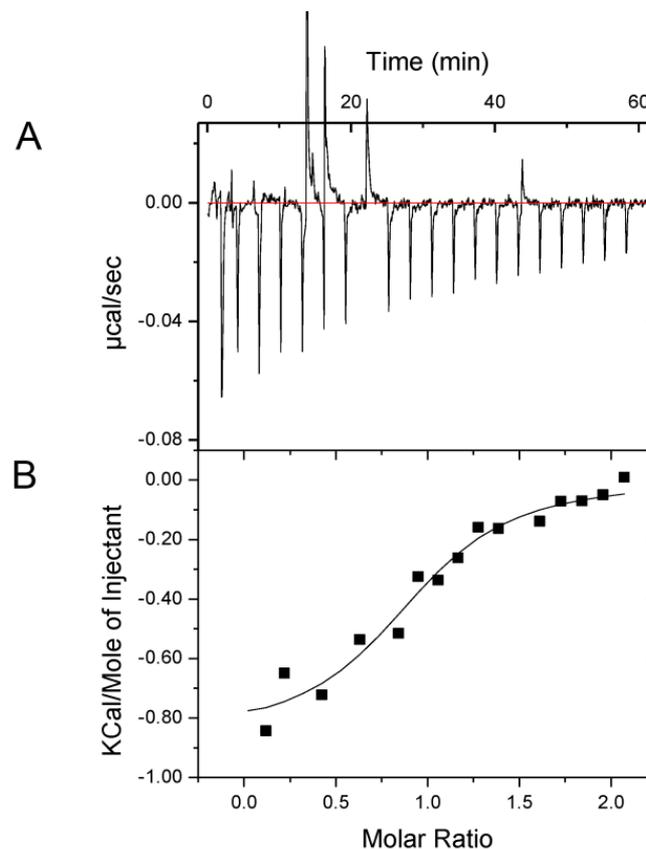


Figure 60 : Expérience d'ITC pour mesurer l'interaction entre P et M2-1[58-177] (Blondot et al. 2012). (A) Données brutes : chaleur absorbée à chaque ajout de M2-1[58-177] (2 μl par injection, la concentration de protéine étant de 335 μM) dans une cellule contenant P (200 μl de volume initial; 33 μM concentration initiale). (B) Courbe de titrage obtenue après intégration des données brutes montrant la chaleur absorbée intégrée par mole de ligand ajouté. Les carrés pleins représentent les points expérimentaux. La ligne continue correspond à la courbe calculée. La valeur de K_d moyenne est égale à 3 μM correspondant à une stœchiométrie de 1:1.

5.2 Spectres HSQC en présence de la protéine P du VRS

Pour avoir une idée plus précise de l'interaction entre M2-1[58-177] et P, nous avons regardé l'influence de P sur les spectres HSQC ^1H - ^{15}N de M2-1[58-177] marquée ^{15}N . Les mesures ont été effectuées à 298 K sur un spectromètre RMN à 14,1 T (fréquence proton de 600 MHz). La Figure 61 montre que l'intensité des signaux décroît drastiquement en fonction de la quantité de P ajoutée. A 0,5 équivalent molaire de P, beaucoup de signaux ne sont plus observables. L'élargissement des raies est un élargissement global, lié à la fixation d'un partenaire de haut poids moléculaire. En effet P est un tétramère de 4*27,3 kDa. Il est difficile de déterminer dans ce cas une zone de fixation de P sur M2-1[58-177].

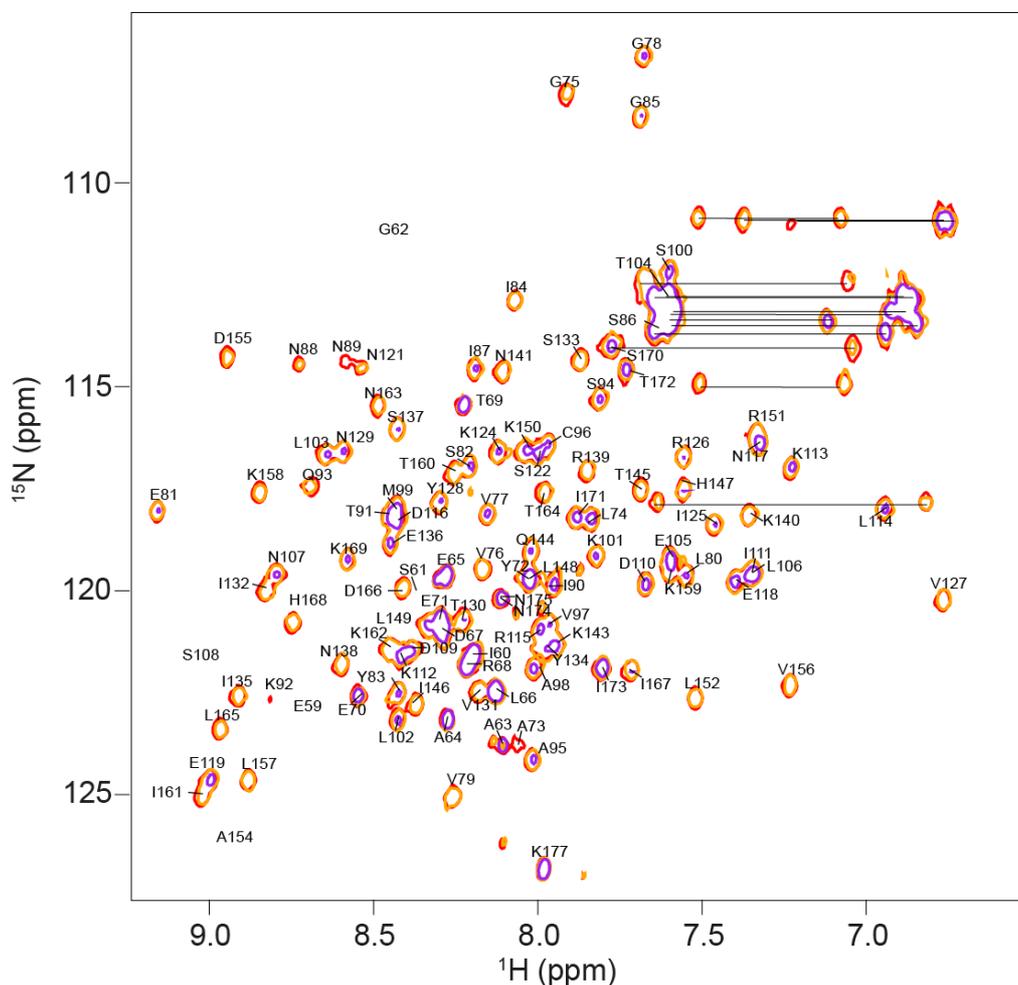


Figure 61 : Superposition de spectres HSQC ^1H - ^{15}N de M2-1[58-177] à 50 μM (spectre de référence en rouge) en présence de 0,25 (orange) et 0,5 (violet) équivalent de la phosphoprotéine P du VRS ($T = 298\text{ K}$, $B_0 = 14,1\text{ T}$).

5.3 Spectres HSQC en présence du fragment P100-160

Or il a été proposé par Mason et collaborateurs que la région de la phosphoprotéine impliquée dans l'interaction avec la protéine M2-1 se situait entre les résidus 100 à 120 de P (Mason et al. 2003). Ils ont procédé par chromatographie d'affinité, en immobilisant la protéine M2-1 recombinante purifiée et en ajoutant le lysat cellulaire contenant la protéine P, et inversement. Avec un système miniréplicon ils ont pu en outre mettre en évidence l'importance de trois résidus de la phosphoprotéine P, L101, Y102 et F109, pour l'interaction avec la protéine M2-1. En effet, la mutation de ces résidus en alanine provoque une diminution drastique de l'activité polymérase du VRS *in vitro*.

C'est pour cela que le fragment de P, P[100-160], contenant cette région a été utilisé pour refaire des expériences d'interaction avec M2-1[58-177] par RMN, toujours en utilisant des spectres HSQC ^1H - ^{15}N (Figure 62). Le poids moléculaire de P[100-160], qui contient le domaine de tétramérisation de P, est de 4*7 kDa, comparé à 4*27 kDa pour P entière.

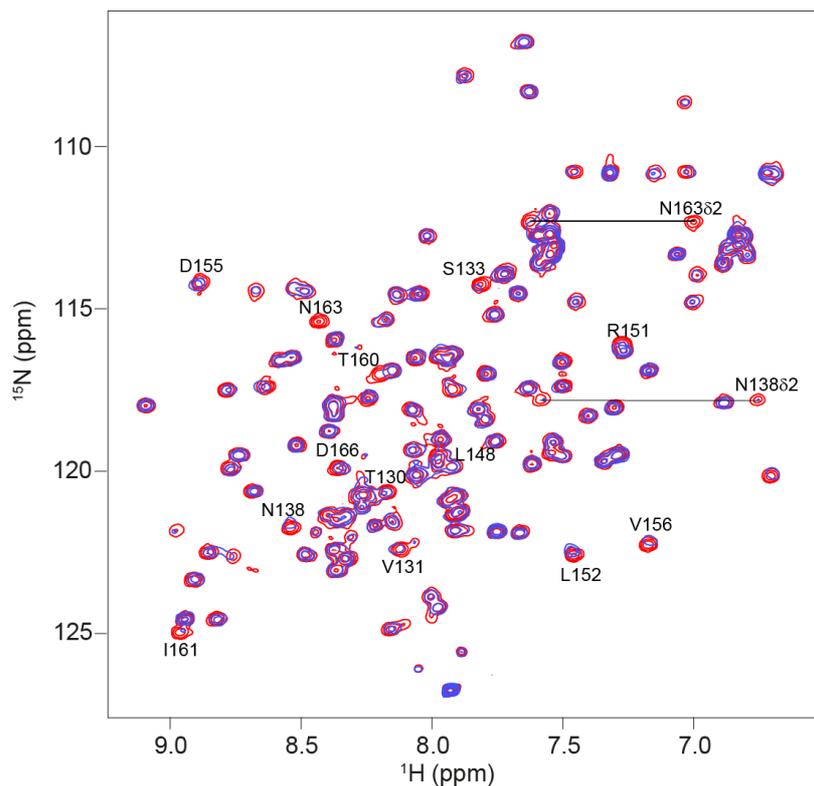


Figure 62 : Superposition de spectres HSQC ^1H - ^{15}N de M2-1[58-177] en présence de deux équivalents de P[100-166] ($T = 298\text{ K}$, $B_0 = 14,1\text{ T}$).

P[100-160] n'induit pas un élargissement aussi drastique des signaux du spectre HSQC ^1H - ^{15}N de M2-1[58-177] que P entière. Mais on observe aussi un phénomène de variation d'intensité des signaux. Ces variations sont montrées dans la Figure 62. Si l'on cartographie

les résidus les plus affectés par l'ajout de P[100-160], on constate qu'ils se trouvent tous dans une même région qui correspond à une crevasse entre les hélices $\alpha 4$ et $\alpha 6$ de M2-1[58-177] (Figure 63).

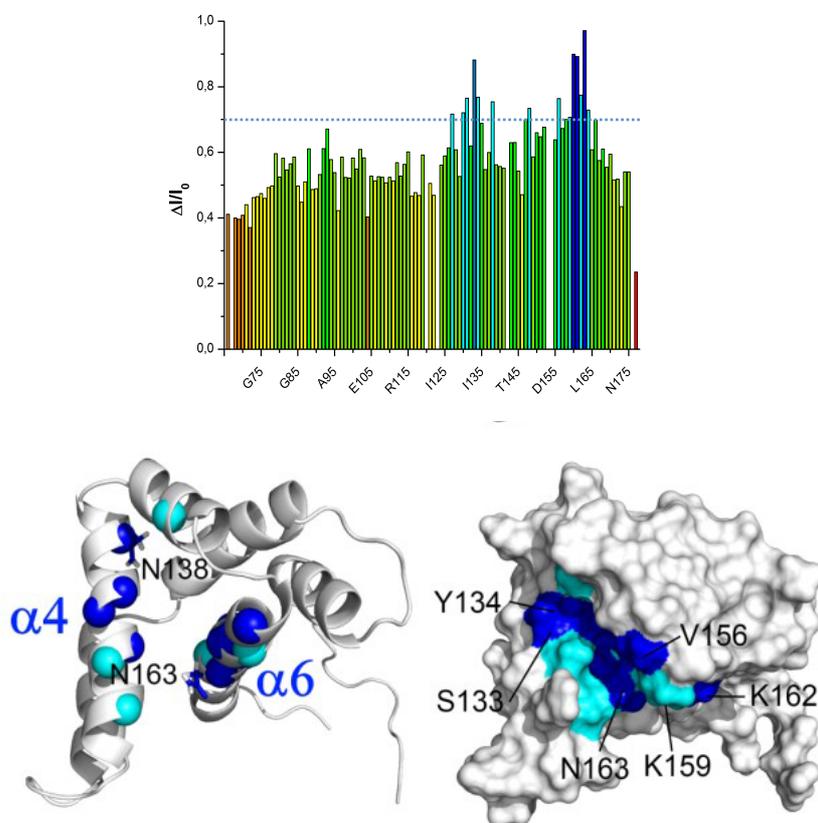


Figure 63 : Variations d'intensité induites par P[100-166] sur le spectre HSQC ^1H - ^{15}N de M2-1[58-177], résidu par résidu, dans les conditions de la Figure 62. Le code couleur correspond à l'amplitude des variations : les valeurs les plus basses sont en rouge, les plus élevées en bleu. Ces variations sont cartographiées sur la structure de M2-1[58-177]. Les atomes d'azote des liaisons amides des résidus qui présentent une variation d'intensité significative sont colorés en bleu si $\Delta I/I_0 > 75\%$ et en cyan si $\Delta I/I_0 > 70\%$.

5.4 Confirmation de la surface d'interaction de P sur M2-1 par saturation croisée transférée

Pour confirmer ces résultats, des expériences de saturation croisée transférée (TCS) ont été réalisées. Le marquage de la protéine M2-1 au deutérium a été nécessaire. Cette fois-ci la protéine P entière a été choisie pour disposer d'une protéine partenaire de haut poids moléculaire. La saturation des protons des méthyles de P donne lieu à une baisse de l'intensité de certains pics sur le spectre de la protéine M2-1[58-177] (Figure 64). Les expériences de

saturation croisée transférée ont été réalisées avec 150 μM de M2-1[58-177] marqué $^2\text{H}^{15}\text{N}$ dans 91% D_2O et en présence de 15 μM (10%) de la protéine P entière (résidus 1-241) non marquée. Des spectres ont été enregistrés sans et avec saturation et des expériences de contrôle ont été effectuées sans P afin de prendre en considération la diffusion de spin et l'effet des protons aliphatiques résiduels qui ne sont pas échangeables dans l'eau. Une surface contiguë est trouvée (Figure 64) qui est compatible avec la surface déterminée avec P[100-160] (Figure 63), mais qui se recouvre aussi partiellement avec la région de fixation des ARN (Figure 57).

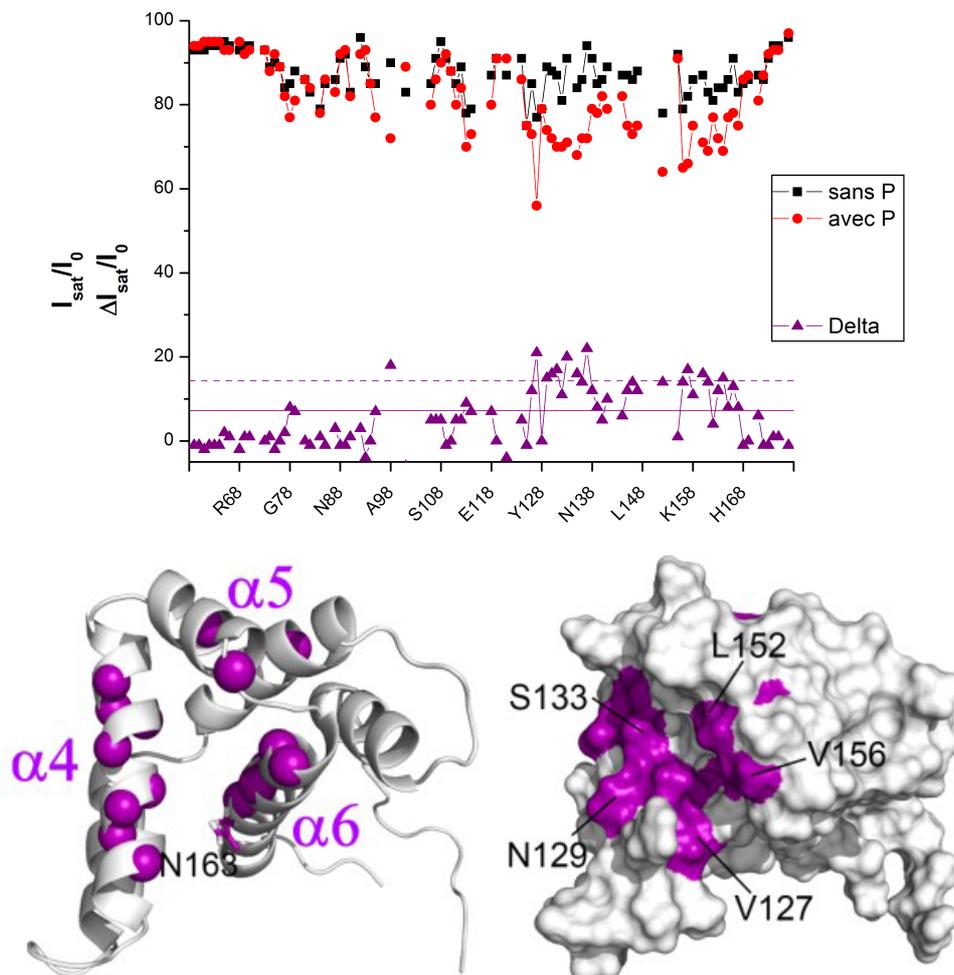


Figure 64 : Rapport des intensités résidu par résidu dans les spectre HSQC $^1\text{H}-^{15}\text{N}$ de 150 μM M2-1[58-177] $^{-2}\text{H}^{15}\text{N}$ dans 91% D_2O obtenues en irradiant la région des méthyles de P (I_{sat}) ou non (I_0), en présence de 10 % (rouge) ou en absence (noir) de la phosphoprotéine du VRS ($T = 298\text{K}$, $B_0 = 14,1\text{ T}$). Cartographie de la saturation croisée transférée sur la structure de M2-1[58-177]. Les atomes d'azote des liaisons amides des résidus qui présentent une variation significative sont indiqués par des sphères violettes. La surface formée par l'ensemble de ces résidus est colorée en violet.

6 Validation des résultats RMN par des tests fonctionnels

Les surfaces d'interaction de la protéine P et l'ARN sur la protéine M2-1 ainsi que leur recouvrement partiel ont ensuite été validés par l'équipe de J-F Eléouët à l'INRA. À partir des données obtenues par RMN, ils ont cherché à identifier et valider les résidus critiques de la protéine M2-1 pour l'interaction avec la protéine P et l'ARN. Pour cela, ils ont généré des mutations dans le domaine globulaire de la protéine M2-1 et ont mesuré l'impact de ces mutations sur l'activité de la protéine M2-1 avec un système de miniréplicon.

6.1 Mesure de la transcription *in vitro* par la polymérase du VRS

L'efficacité de la transcription par la polymérase du VRS en présence de protéines M2-1 mutées a été mesurée dans un système de miniréplicon. Ce système consiste à co-transfecter des cellules avec des plasmides codant pour les protéines formant le complexe polymérase du VRS (N, P, M2-1, L), transcrits par l'ARN polymérase T7, et un plasmide permettant de produire un ARN contenant des signaux essentiels pour la polymérase du VRS, à savoir la séquence Leader et deux séquences GS et GE (Figure 65). La transcription de cet ARN permet l'expression d'un gène rapporteur, la luciférase, et dépend de l'état fonctionnel des quatre protéines impliquées (Figure 65) (Blondot et al. 2012).

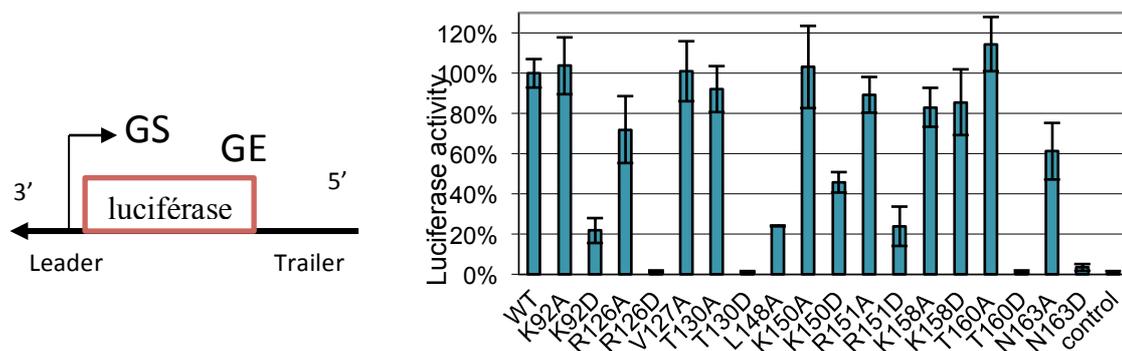


Figure 65 : Schéma du minigénome du VRS codant pour la luciférase et mesure de l'activité polymérase *in vitro* du VRS via la mesure de l'activité luciférase pour différents mutants de M2-1.

Huit mutations réduisent de plus de 50 % l'activité polymérase *in vitro*. Elles se situent toutes dans les zones d'interaction avec les ARN ou avec P déterminées par RMN. Des expériences d'interaction *in vitro* ont donc été menées pour essayer de discriminer entre résidus importants pour l'interaction avec P et avec l'ARN.

6.2 Discrimination des résidus critiques pour l'interaction M2-1:P et M2-1:ARN

Des expériences *in vitro* de retard sur gel ont permis de montrer que trois des huit mutants identifiés avec le miniréplicon (K92D, K150D et R151D) présentaient un défaut d'interaction avec l'ARN (Figure 66) (Blondot et al. 2012).

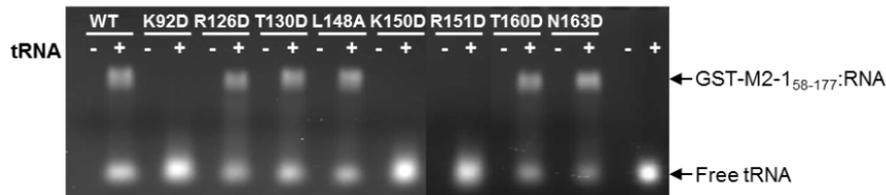


Figure 66: Expériences de retard sur gel (Gel tris glycine 1% agarose) avec la protéine GST-M2-1[58-177] (100 μ M) et des ARNt de levure (50 μ M) (Blondot et al. 2012)

Des expériences de GST pull-down ont permis de montrer que cinq des huit mutants identifiés avec le miniréplicon (R126D, T130D, L148A, T160D et N163D) présentaient un défaut d'interaction avec l'ARN (Figure 67) (Blondot et al. 2012).

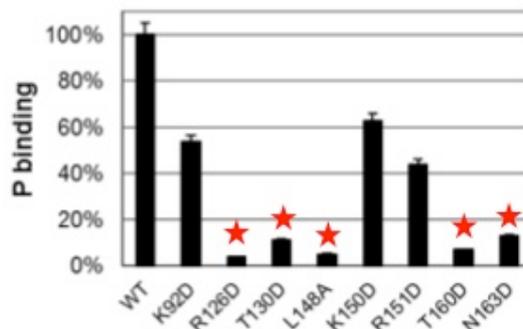


Figure 67: Expériences de GST-pulldown par GST-M2-1 de P et de mutants de P. Les étoiles rouges montrent les mutations qui ont induit une baisse d'interaction de plus de 50% par rapport à la protéine M2-1 sauvage.

6.3 Analyse par immunofluorescence de la localisation de mutants de M2-1

Enfin l'impact de ces mutations sur la localisation intracellulaire de la protéine M2-1 par immunofluorescence, comme illustré dans la Figure 68. Deux phénotypes ont été observés. Le premier correspond à la co-localisation de P et de M2-1 dans des corps d'inclusion cytoplasmiques lorsqu'elle est co-exprimée avec les protéines P et N du VRS

(Garcia et al. 1993). Il est observé avec la protéine M2-1 sauvage et les mutants K92D, K150D et R151D. Ces mutants correspondent à la fixation d'ARN, comme indiqué plus haut. Dans le deuxième phénotype, M2-1 est diffus dans le cytoplasme et exclus des inclusions cytoplasmiques qui contiennent P. Il est observé avec les mutants R126D, T130D, L148A, T160D et N163D. Ces mutants correspondent à la perte d'interaction avec P, comme déterminé plus haut. On peut donc corréliser la perte d'activité du complexe polymérase avec un défaut de recrutement de la protéine M2-1 dans les corps d'inclusion, avec la phosphoprotéine P.

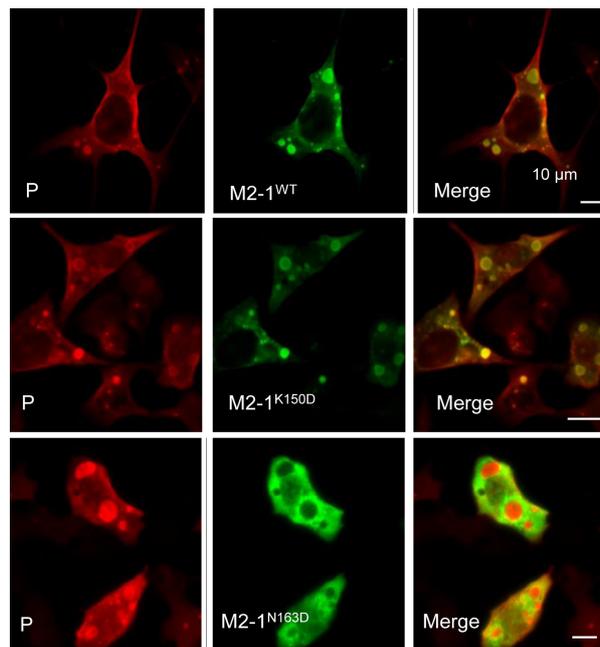


Figure 68: Immunofluorescence 24 h après co-transfection en cellules BSR-T7 de P, N et M2-1 sauvage ou mutée. Marquage indirect avec un anticorps polyclonal de lapin anti-P (rouge) et anticorps monoclonal de souris anti-M2-1 (vert) (Blondot et al. 2012).

6.4 Conclusion

Pour résumer, les expériences d'interactions par RMN ont permis de cartographier les zones de fixation de l'ARN et de la protéine P. Pour chacune des deux interactions, trois hélices formant une surface uniforme sont impliquées : α_2 , α_5 et α_6 pour l'interaction avec l'ARN et α_4 , α_5 et α_6 pour la phosphoprotéine P.

Ces surfaces présentent un recouvrement partiel de ces zones sur la protéine M2-1, illustré dans la Figure 69. La proximité de ces surfaces d'interaction peut ainsi expliquer la compétition observée *in vitro* entre l'ARN et la protéine P pour la protéine M2-1 (Tran et al. 2009). L'affinité déterminée par ITC pour la phosphoprotéine P et est du même ordre de

grandeur que celle des séquences d'ARN en sens positif pour M2-1[58-177]. Ces résultats sont cohérents avec les observations de compétitivité entre la protéine M2-1 et ces deux partenaires.

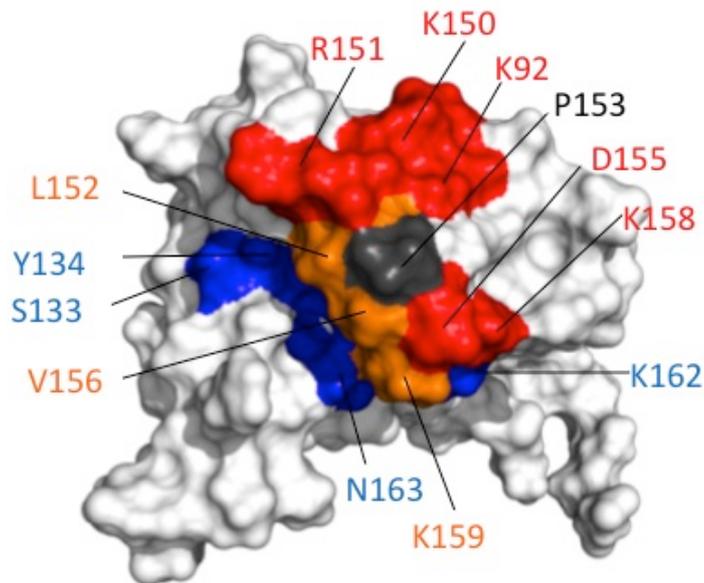


Figure 69: Les surfaces de contact avec P (en bleu) et avec l'ARN (en rouge) sur le domaine globulaire de M2-1 font apparaître une superposition partielle qui pourrait correspondre à une surface de compétition (en orange).

Ces résultats ont posé les bases pour des tests fonctionnels qui ont permis l'identification de huit résidus de la protéine M2-1 critiques pour l'activité polymérase du VRS, dont cinq essentiels pour l'interaction avec la phosphoprotéine et trois pour l'interaction avec l'ARN. Ces résultats soulignent l'importance des interactions M2-1:P et M2-1:ARN pour la fonction de la protéine M2-1 au cours de la transcription virale. Les résultats concernant la co-localisation des protéines virales par immunofluorescence permettent également de formuler des hypothèses quant à l'enchaînement des événements de fixation de P ou des ARN sur M2-1. Dans un premier temps, la protéine M2-1 pourrait interagir avec la phosphoprotéine P qui permettrait son recrutement dans les corps d'inclusion cytoplasmiques qui contiennent aussi la protéine N du VRS. Ces sites sont supposés être les sites de synthèse d'ARN viral. L'interaction M2-1:P serait alors être rompue en faveur d'une interaction M2-1:ARN viral, via un processus où la protéine P elle-même se lierait à un autre partenaire.

Dans l'équipe du Dr. Gonzalo de Prat-Gay l'interaction entre la protéine P et la protéine M2-1 a été déterminée par spectrofluorimétrie (Esperante et al. 2012). Ils ont

déterminé un K_d de l'ordre de 10 nM, soit une affinité 1000 fois plus élevée que ce que nous avons déterminé par RMN. L'hypothèse principale mise en avant pour expliquer une telle différence repose sur l'effet de tétramérisation de la protéine M2-1, car ils ont utilisé la protéine M2-1 entière pour leur expérience. On peut alors supposer que le même phénomène puisse se produire lors de l'interaction entre la protéine M2-1 et l'ARN. Cette observation favorise l'idée d'un effet coopératif global lié à l'oligomérisation de la protéine M2-1 et permet d'expliquer les différences de K_d également observé avec l'ARN.

Comme la structure radiocristallographique de M2-1 est connue maintenant, on peut envisager de caractériser ces mêmes interaction par RMN, mais en utilisant cette fois-ci la protéine entière, pour savoir comment la tétramérisation de M2-1 influe sur les propriétés de fixation du domaine globulaire. Des marquages isotopiques permettant de travailler avec des protéines de taille > 25 kDa seront nécessaires, entre autre le marquage de M2-1 au deutérium.

Chapitre 4 : : Etude structurale par RMN de la phosphoprotéine du VRS

1 Introduction

1.1 La protéine P, plaque tournante des interactions dans le complexe polymérase du VRS

La phosphoprotéine P est le principal co-facteur de la polymérase du VRS. Elle sert de plaque tournante au cours des processus de transcription et de réplication, puisqu'elle est capable de se lier aux protéines L, N et M2-1 du VRS ainsi qu'à diverses protéines cellulaires. Ces propriétés de liaison sont étroitement liées à l'organisation en domaines de la protéine P, qui comporte un domaine central d'oligomérisation et des domaines N- et C-terminaux dont les structures sont prédites comme principalement désordonnées (Llorente et al. 2006; Simabuco et al. 2011).

La reconnaissance de N, à l'intérieur de la nucléocapside, par P est une étape-clé de la synthèse d'ARN viral du VRS. Elle permet la décapsidation de l'ARN génomique et son accès à la sous-unité L. La partie C-terminale de P est impliquée dans ce processus, dont le mécanisme exact n'est pas connu.

La protéine P se lie également au facteur de transcription M2-1, et la surface d'interaction de P sur M2-1 a été identifiée, comme je l'ai montré plus tôt dans ce manuscrit. Le but dans cette partie était d'en savoir plus sur le complexe M2-1:P en observant P. Pour cela il est important d'acquérir une vision plus précise de la structure de P dans un premier temps.

1.2 Structure de P

Les informations disponibles sur la structure de la phosphoprotéine P sont résumées dans la Figure 70. Un motif conservé chez les Mononégavirales a été identifié en N-terminal par Karlin et Belshaw (Karlin and Belshaw 2012). Il pourrait être impliqué dans la formation du complexe N0:P. Les bornes du domaine d'oligomérisation de P ont été étudiées par digestion à la trypsine dans le groupe du Dr José Méléro (Llorente et al. 2006). Ils ont postulé que le domaine d'oligomérisation correspondait aux résidus 104 à 163. Une forme plus courte a été déterminée plus tard par digestion à la chymotrypsine, comprenant les résidus 119 à 160

(Llorente et al. 2008). Enfin le groupe du Dr Jean-François Eléouët a identifié la région C-terminale comme zone d'interaction avec la nucléocapside (Tran et al. 2007).

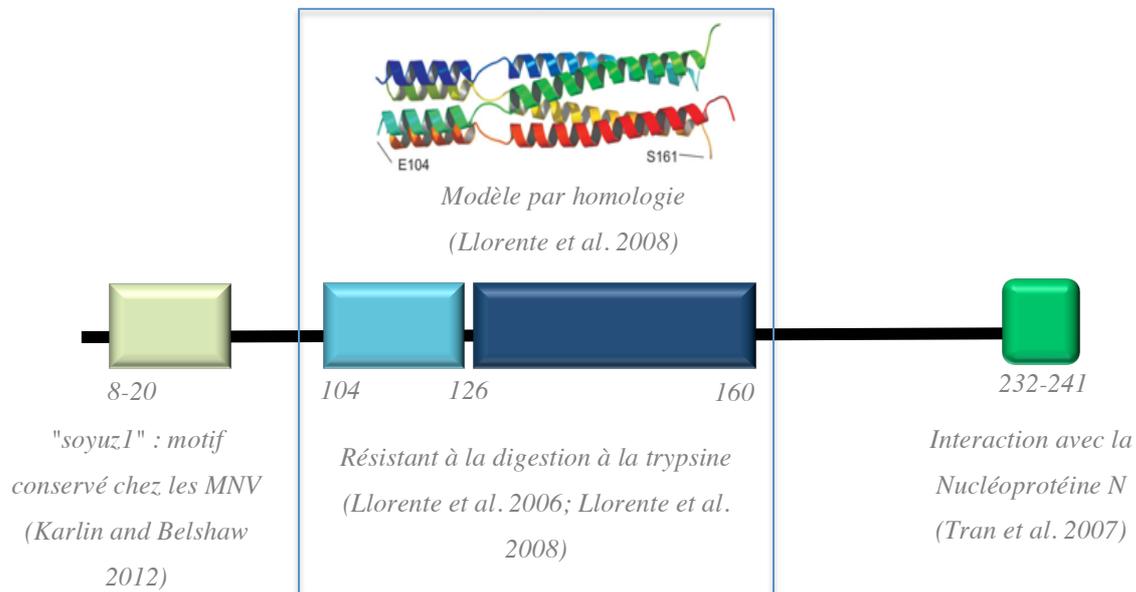


Figure 70 : Schéma récapitulatif des différentes informations connues sur la structure de la phosphoprotéine P du VRS. Un motif "soyuz1" en N-terminal (résidus 8-20, rectangle beige) est conservé chez les Mononégavirales (Karlin and Belshaw 2012). Le domaine d'oligomérisation résistant à la digestion à la trypsine est représenté par des rectangles bleus (Llorente et al. 2006). Un modèle par homologie a été construit à partir du domaine d'oligomérisation de la phosphoprotéine du virus Sendai. La région déterminée par digestion à la chymotrypsine est en bleu foncé (Llorente et al. 2008). La région C-terminale (résidus 232-241, rectangle vert) interagit avec la nucléoprotéine N de la nucléocapside (Tran et al. 2007).

Esperante et collaborateurs ont réussi à visualiser un changement de conformation de la protéine P en présence de la protéine M2-1 (Esperante et al. 2012). En effet, la phosphoprotéine seule est constituée d'une région N et C-terminale désordonnée et d'un domaine d'oligomérisation structuré en hélice α . Or, lors de la formation du complexe entre les protéines P et M2-1, ils ont observé une globularisation ainsi qu'une augmentation en hélice α de la protéine P. La partie N-terminale de la protéine P, dans laquelle se situe le domaine d'interaction avec la protéine M2-1, serait ainsi stabilisée par cette dernière, passant d'un état désordonné à un état au moins partiellement structuré. Cela suggère que la structure de la protéine P serait modulée en fonction des interactions avec ses partenaires. On peut ainsi expliquer la compétitivité observée avec l'ARN en imaginant que la protéine P changerait de

conformation en présence d'un nouveau partenaire permettant le relargage de la protéine M2-1, libre d'interagir avec l'ARN ou un autre partenaire. La présence de régions intrinsèquement désordonnées au sein d'une protéine lui permettrait de supporter des changements conformationnels perpétuels afin de médier des interactions hautement spécifiques avec de nombreux partenaires (Babu et al. 2012).

1.3 Eléments de réponse sur la structure de P par RMN

1.3.1 Utilisation de fragments de P

Bien que la méthode consistant à produire des fragments correspondant à des domaines n'ait pas fonctionné pour M2-1, cette stratégie a néanmoins été appliquée à la phosphoprotéine du VRS. Nous avons travaillé avec plusieurs fragments, en nous basant sur la structure en domaines de P donnée dans la Figure 70. Différents groupes travaillant sur P ont estimé que le domaine d'oligomérisation s'étend des résidus 104 à 163 (Llorente et al. 2008; Esperante et al. 2012). Cependant les résultats présentés dans le paragraphe suivant montrent que ce domaine est probablement réduit aux résidus 126 à 163. Cette définition plus stricte est retenue dans la suite. La Figure 71 récapitule les constructions définies à partir des bornes du domaine d'oligomérisation.

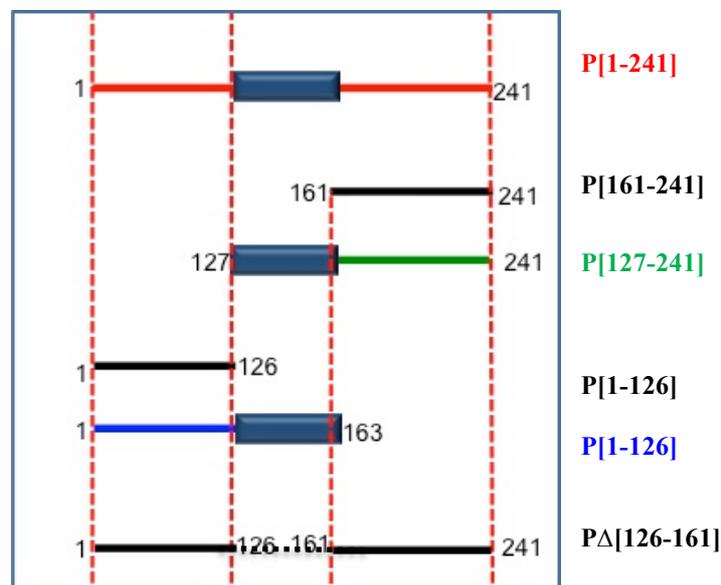


Figure 71 : Constructions utilisées pour l'étude de la structure de la phosphoprotéine par RMN : P pleine longueur, 2 fragments C-terminaux sans et avec le domaine d'oligomérisation, 2 fragments N-terminaux sans et avec le domaine d'oligomérisation et un fragment où le domaine d'oligomérisation a été supprimé PΔ[126-161]. La forme courte du domaine d'oligomérisation est indiquée par un rectangle bleu foncé.

Dans un premier temps nous avons donc travaillé avec la protéine entière P[1-241], deux fragments N-terminaux, contenant ou non le domaine d'oligomérisation, P[1-163] et P[1-126], et deux fragments C-terminaux, contenant ou non le domaine d'oligomérisation, P[127-241] et P[161-241].

La Figure 72 permet de comparer les spectres HSQC ^1H - ^{15}N de fragments de phosphoprotéine N- et C-terminaux contenant le domaine d'oligomérisation, P[1-163] et P[127-241] respectivement, avec le spectre de la phosphoprotéine entière. On constate que les signaux se superposent : à chaque signal de P se superpose un signal de P[1-163] ou de P[127-241]. Cela indique que les fragments représentent bien la protéine P et que les domaines N- et C-terminaux sont indépendants l'un de l'autre.

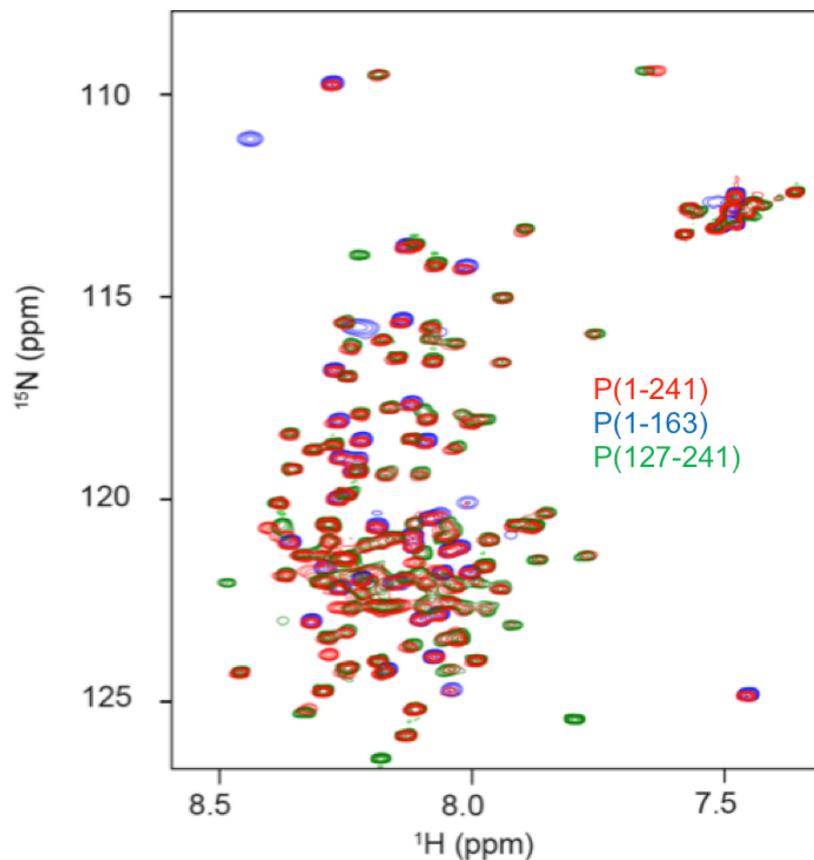


Figure 72 : Superposition des spectres HSQC ^1H - ^{15}N de la protéine P pleine longueur P[1-241] en rouge et de deux fragments correspondant à la partie N-terminale de P avec le domaine d'oligomérisation P[1-163] (bleu) et à la partie C-terminale avec le domaine d'oligomérisation P[127-241] (vert). Les spectres ont été acquis à un champ de 14,1 T et une température de 288 K, à des concentrations de 50 μM dans un tampon phosphate 20 mM, NaCl 100 mM, pH 6,5.

1.3.2 Perte de signal dans les régions structurées

Si l'on compte le nombre de signaux amides dans les spectres montrés dans la Figure 72, on constate qu'il en manque quasiment 80 pour la protéine P, 40 pour P[1-163] et 80 pour P[127-241]. La faible dispersion des signaux dans la dimension ^1H , entre 8 et 8,5 ppm, et leur forte intensité indiquent que les signaux visibles correspondent à des régions désordonnées. Ces résultats indiquent aussi que les signaux absents correspondent vraisemblablement au domaine d'oligomérisation, qui est le point commun entre les fragments P[1-163] et P[127-241]. Ils confirment donc que la phosphoprotéine comporte de larges extensions N- et C-terminales intrinsèquement désordonnées.

Pour déterminer plus exactement les limites du domaine d'oligomérisation, nous avons comparé les spectres de deux constructions N-terminales sans et avec le domaine d'oligomérisation, P[1-126] et P[1-163] (Figure 73).

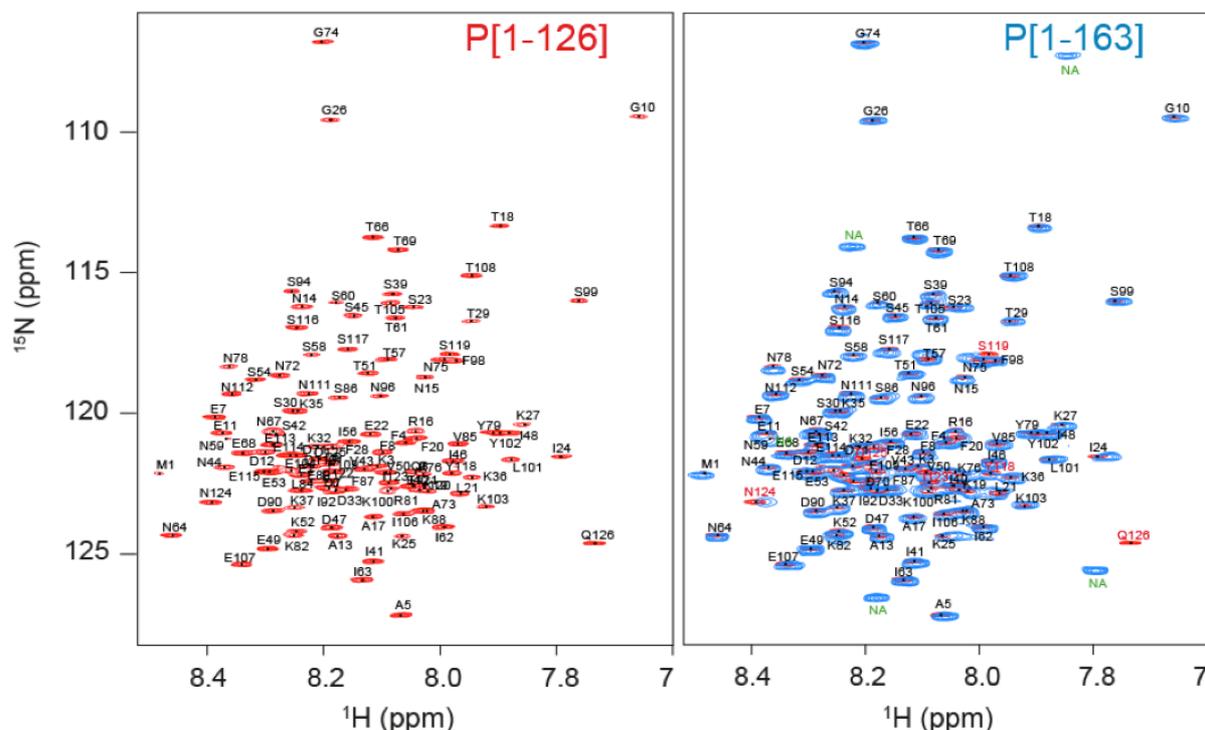


Figure 73: Spectres HSQC ^1H - ^{15}N attribués de deux fragments de P correspondant à la partie C-terminale sans domaine d'oligomérisation P[1-126] et avec domaine d'oligomérisation : P[1-163] (bleu), P[1-126] (rouge). L'acquisition a été faite à un champ de 14,1 T, à 288 K, à une concentration de 50 μM en tampon phosphate salin (pH 6,5).

L'attribution des fréquences de résonance du squelette peptidique avait été faite pour les deux protéines. Pour P[1-126] tous les signaux amides sont visibles et les déplacements chimiques ^1H entre 8 et 8,5 ppm, indiquent que cette protéine est complètement désordonnée.

Les spectres HSQC ^1H - ^{15}N des deux fragments N-terminaux de la phosphoprotéine se superposent bien. Cela montre que les résidus 1-120 dans P[1-163] sont également désordonnés. Les signaux pour les résidus 121-160, correspondant au domaine d'oligomérisation, ne sont pas visibles. Cela indique aussi que le domaine d'oligomérisation commence à partir du résidu 126 et non pas à partir du résidu 104, contrairement à ce qui a été prédit par (Llorente et al. 2008).

1.3.3 Analyse de la structure secondaire des régions désordonnées de P

L'étude du déplacement chimique secondaire et de la relaxation nucléaire de l'azote 15 sur les autres constructions indiquées dans la Figure 71 ont permis de montrer que des hélices α transitoires sont formées dans les extrémités N- et C-terminales de la phosphoprotéine. A l'extrémité C-terminale, les hélices sont presque complètement formées et semblent prolonger le domaine d'oligomérisation. A l'extrémité N-terminale deux régions sont identifiées, avec de très faibles propensions en hélice α (Figure 74). La première pourrait correspondre à la région "mir" ou "soyuz" (Figure 70) et donc à un second site potentiel d'interaction avec la protéine N pour former le complexe N_0 : P, identifié sur la base de sa ressemblance avec une séquence retrouvée dans d'autre virus. La deuxième pourrait correspondre au site d'interaction avec le cofacteur de transcription M2-1, comme je vais le montrer par des expériences d'interaction par RMN dans le chapitre suivant.



Figure 74 : Régions de P avec des structures en hélices partielles déterminées par RMN.

La représentation des structures secondaires transitoires alignées dans la Figure 74 traduit le fait que peu de renseignements sont disponibles sur la structure tridimensionnelle de P. Esperante et collaborateurs ont observé une globularisation de P en présence de M2-1 (Esperante et al. 2012). Nous avons donc voulu explorer le potentiel de P à former un ensemble plus compact dans sa forme tétramère, et en particulier à savoir si des contacts pouvaient se former entre les domaines N- et C-terminaux.

2 Recherche de contacts à longue distance dans P

Pour chercher des informations de contacts à longue distance entre les parties N- et C-terminales de la phosphoprotéine, j'ai réalisé des expériences d'augmentation de la vitesse de relaxation nucléaire en présence d'une sonde paramagnétique (PRE, Paramagnetic Relaxation Enhancement) (Cloue et al. 2007; Iwahara et al. 2007). L'augmentation des vitesses de relaxation R_2 des noyaux proches d'une sonde paramagnétique se traduit par un élargissement de raie et donc par une baisse d'intensité du signal. Ces expériences fournissent des informations sur des distances de l'ordre de 15-24 Å.

2.1 Préparation d'échantillons avec une sonde paramagnétique

2.1.1 Préparation de mutants cystéines

Pour réaliser ces expériences, j'ai muté des résidus de la phosphoprotéine en cystéines pour fixer une étiquette paramagnétique iodoacétamido-PROXYL (IAP). Sept mutants ponctuels de substitution par une cystéine ont été produits. Les sites de mutation ont été choisis de manière à distribuer des sondes tout au long de la séquence. S23C et S99C sont situés dans la partie N-terminale, E179C, E193C et S237C dans la partie C-terminale, S143C et S156C à l'intérieur du domaine d'oligomérisation, comme indiqué dans la Figure 75. Dans la mesure du possible des sérines ont été choisies pour perturber le moins possible la séquence, puisque la mutation correspond à une substitution $OH\gamma$ par $SH\gamma$. Les deux résidus du domaine d'oligomérisation ont été choisis en prenant des résidus de surface sur le modèle structural par homologie de P.



Figure 75 : Position des sites mutés en cystéine dans la séquence de P par rapport aux régions structurées ou partiellement structurées.

Les mutants ponctuels ont été obtenus par mutagenèse dirigée en utilisant le kit Quickchange (Stratagene®). Le mode opératoire et les séquences des amorces utilisées pour chaque mutant sont donnés en Annexe. Une fois ces mutations ponctuelles effectuées, tous les mutants ont été séquencés. La Figure 76 montre l'alignement de séquence des différents mutants de P avec la séquence de la forme sauvage de P.

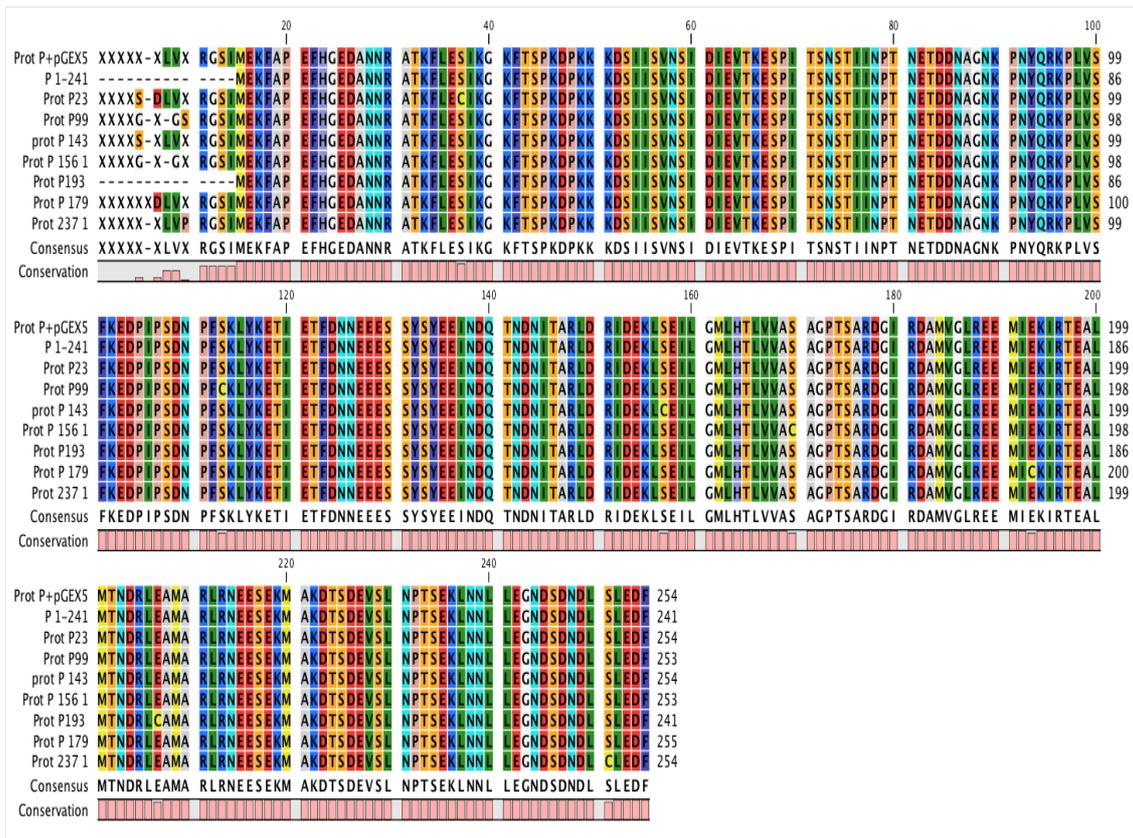


Figure 76 : Alignement de séquence des différentes constructions de P : P sauvage P[1-241], P(S23C), P(S99C), P(S143C), P(S156C), P(E179C), P(E193C) et P(S237C).

2.1.2 Ajout d'une étiquette paramagnétique iodoacétamido-PROXYL

Le même protocole de production et de purification a été appliqué pour toutes les constructions et mutants de la phosphoprotéine. Les détails figurent dans l'Annexe "matériels et méthodes".

Pour la fixation de l'étiquette IAP, les échantillons protéiques ont d'abord été réduits par l'addition de 2 mM de dithiothréitol (DTT) et dialysés dans 50 mM de Tris pH 8,0 et 200 mM de NaCl. Ils ont ensuite été incubés avec trois équivalents molaires de 3-(2-iodoacétamido)-2,2,5,5 tétraméthyl-1-pyrrolidinyloxy radical (iodoacétamido-PROXYL; Sigma-Aldrich) dissous dans du méthanol. Après une nuit de réaction dans l'obscurité à 4°C ou 30-60 minutes à température ambiante, la protéine modifiée est passée sur colonne de dessalage pour enlever l'excès du marqueur qui n'a pas réagi et changer le tampon. Les échantillons n'ont pas été vérifiés par spectrométrie de masse.

Pour chaque mutant, quatre spectres HSQC ^1H - ^{15}N ont été acquis : 1) un spectre de la

protéine avant modification, 2) un spectre avec l'étiquette paramagnétique, 3) un spectre après traitement à la vitamine C pour réduire le radical libre (par ajout de 6 équivalents molaires d'acide ascorbique) servant de référence diamagnétique, et 4) un spectre de contrôle de la protéine non modifiée en présence de vitamine C pour vérifier que celle-ci n'induit pas de modifications spectrales.

2.2 Analyse des spectres des protéines avec une sonde paramagnétique

2.2.1 Vérification par RMN des mutants cystéines

La position des mutations sur la protéine peut être vérifiée en mesurant les spectres HSQC ^1H - ^{15}N pour les mutants correspondant à des sites qui sont observables, c'est-à-dire situés dans les régions 1-120 et 205-241. Cela exclut les mutants S143C, S156C, E179C et E193C.

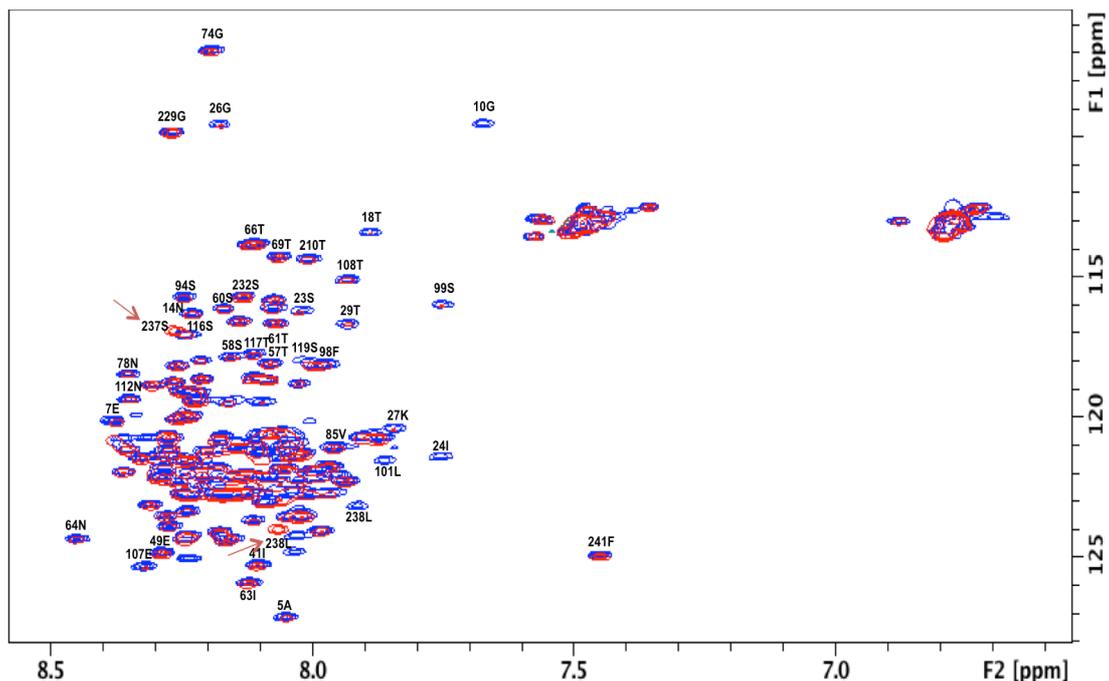


Figure 77 : Spectres HSQC ^1H - ^{15}N superposés du mutant P(S237C) (en bleu ; 50 μM dans un tampon 20 mM phosphate pH 6,5, 100 mM NaCl, 0,1% Azoture de sodium) et de P sauvage (rouge) acquis dans les mêmes conditions (14,1 T et 288 K). Les pics correspondant à la sérine 237 et lysine 238 sont indiqués par des flèches rouges.

Par exemple, en superposant les spectres HSQC ^1H - ^{15}N du mutant P(S237C) et de la protéine P sauvage (Figure 77), on constate que la majorité des pics se superposent à ceux de P sauvage. Certains pics ont bougé. Ils correspondent aux résidus à proximité du site de

mutation, comme Lys 238. Dans le cas de P(S237C) l'attribution séquentielle complète des fréquences de résonance du squelette peptidique a été refaite à l'aide d'expériences triple-noyau (HNCA, HN(CO)CA, HN(CO)CACB, HNCACB, HNCO) pour vérifier les attributions des signaux sur le spectre HSQC ^1H - ^{15}N .

2.2.1.1 Oxydation partielle des cystéines

Deux échantillons de P(237C) ont été produits avec deux types de marquage différent, l'un ^{15}N , le deuxième $^{15}\text{N}^{13}\text{C}$. L'échantillon marqué ^{15}N présente un dédoublement des pics autour du site de la mutation, indiquant que deux formes co-existent en solution. Ceci est illustré dans la Figure 78 pour le résidu F241. Les résultats présentés plus loin montrent que seule la forme 1 réagit avec l'IAP. Cela laisse penser que la forme 2 pourrait correspondre à une cystéine oxydée ou un adduit formé par exemple avec le β -mercaptoéthanol utilisé lors de la purification, alors que la forme 1 correspond à une cystéine réduite et libre.

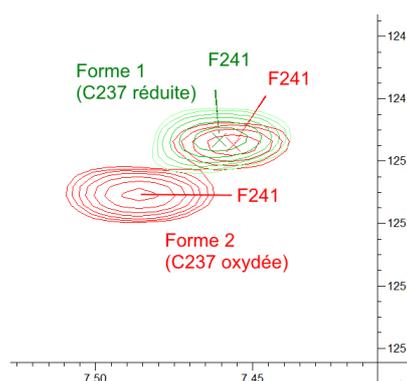


Figure 78 : Agrandissement des spectres HSQC ^1H - ^{15}N de P(237C) marqué ^{15}N (rouge) et marqué $^{15}\text{N}^{13}\text{C}$ (vert) autour du résidu F241, indiquant la présence de deux formes de P(237C) dans l'échantillon marqué ^{15}N .

2.2.1.2 Disparition de signaux dans la partie N-terminale de P

Un autre problème est apparu avec le mutant P(S23C). En effet, une grande partie des signaux semble avoir disparu, en particulier la totalité des signaux de la partie N-terminale (résidus 1-29) (Figure 79). Cela a aussi été observé pour le mutant P(E179C), mais uniquement pour les résidus 1-10.

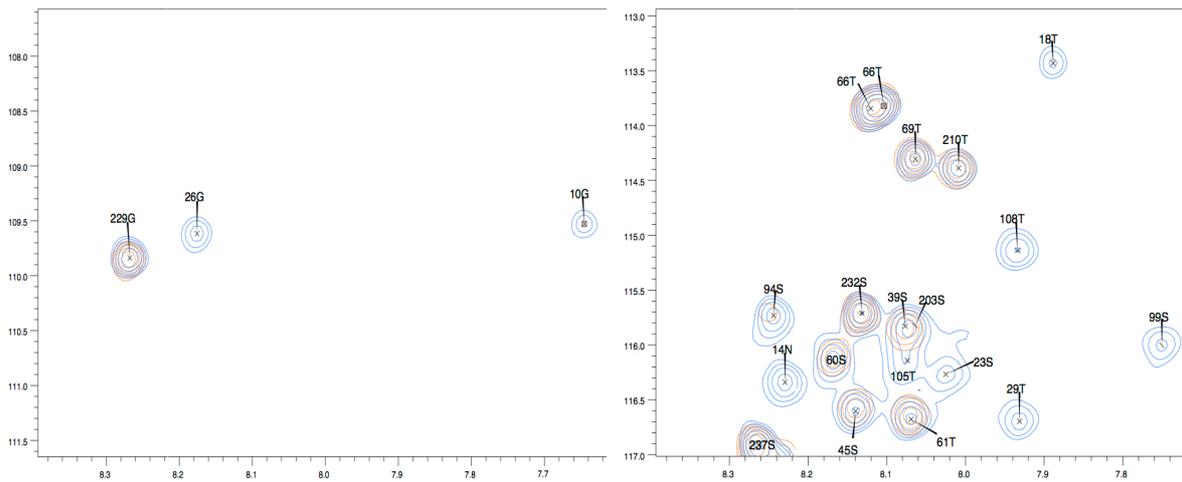


Figure 79 : Agrandissements des spectres HSQC ^1H - ^{15}N superposés du mutant P(S23C) (en brun ; 50 μM dans 20 mM Phosphate pH 6,5, 100 mM NaCl, 0,1% NaN_3) et P sauvage (en bleu) acquis dans les mêmes conditions (14,1 T et 288 K).

Ces résultats indiquent que les mutants de P sont sensibles à la protéolyse dans la partie N-terminale, une région désordonnée, comme démontré plus tôt. Cela est probablement dû au cocktail d'antiprotéases utilisées au cours de la purification. Il avait été changé contre un cocktail moins complet.

Néanmoins la Figure 79 montre aussi une disparition du signal dans la région N90-D110 pour le mutant P(S23C). Le même phénomène est observé pour le mutant P(S99C) (Figure 80) et il ne peut pas être expliqué directement par un effet de protéolyse. D'ailleurs aucun des deux spectres RMN ne comporte de trace évidente d'un deuxième C-terminal généré par une coupure. Tout se passe comme si les signaux étaient élargis par un changement de dynamique du domaine N-terminal. On peut envisager une structuration, par exemple en hélice, comme c'est le cas pour les résidus du domaine d'oligomérisation et ceux des hélices C-terminales en aval du domaine d'oligomérisation. L'hypothèse la plus probable serait que les cystéines dans la partie N-terminale forment des ponts disulfures qui associent les régions N-terminales et ainsi la rigidifient, ce qui se traduirait par un élargissement de raie. Une analyse par spectrométrie de masse aurait peut-être permis de mieux identifier ce problème et de trancher sur la question qui est de savoir si les effets observés sont dus à une dégradation par protéolyse du domaine N-terminal ou à changements structuraux, ou aux deux. Il aurait aussi fallu doser les cystéines libres par exemple avec le réactif d'Ellman.

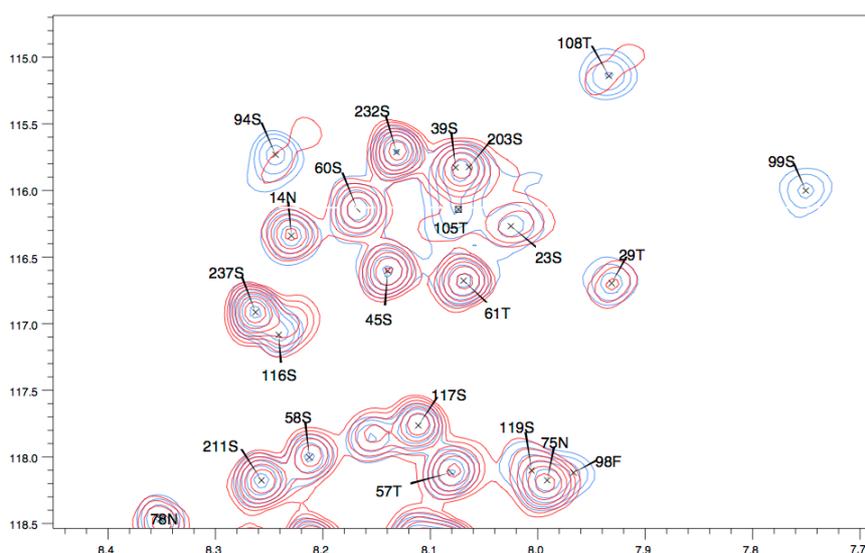


Figure 80 : Agrandissement des spectres HSQC ^1H - ^{15}N superposés du mutant P(S99C) (en rouge ; 50 μM dans 20 mM Phosphate pH 6,5, 100 mM NaCl, 0,1% NaN_3) et P sauvage(en bleu) acquis dans les mêmes conditions (14,1 T et 288 K), montrant des élargissement de signaux pour des résidus près du site de mutation (S94, F98, T105, T108).

2.2.2 Analyse par RMN des protéines avec étiquette paramagnétique

L'exemple de P(S237C) montre que la présence d'une sonde paramagnétique induit bien un élargissement des raies dans le spectre HSQC ^1H - ^{15}N pour les résidus à proximité de C237 (Figure 81). L'exemple du résidu F241 permet de confirmer la présence de deux formes, dont une seule a réagi avec le réactif IAP.

Pour les autres mutants de P, les signaux de la cystéine et des résidus proches sont déjà élargis en l'absence de sonde paramagnétique. On ne peut donc pas vérifier la fixation de la sonde paramagnétique par RMN. Cela est aussi vrai pour les mutants du domaine d'oligomérisation pour lesquels aucun signal n'est observé, comme décrit plus tôt.

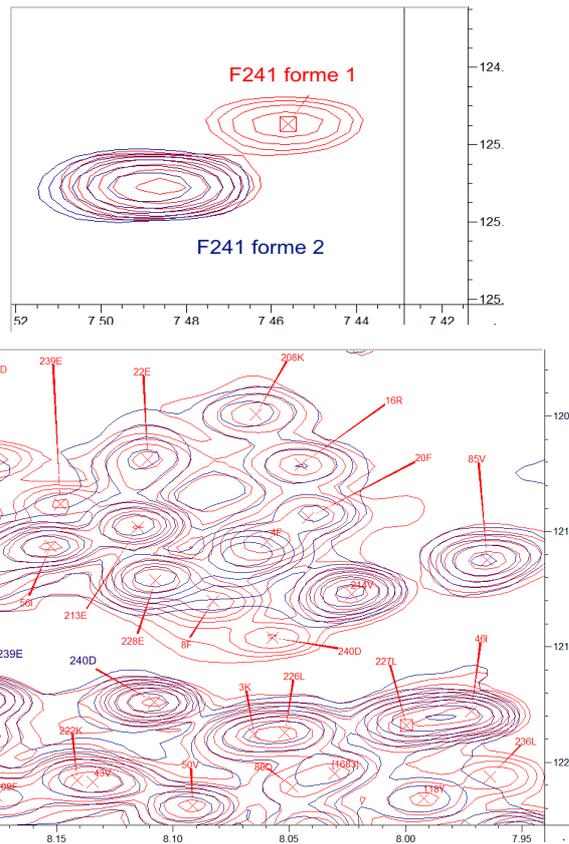


Figure 81 : Agrandissements des spectres HSQC $^1\text{H}-^{15}\text{N}$ du mutant P(S237C) avant (en rouge) et après traitement avec l'IAP (en bleu) acquis dans les mêmes conditions (50 μM dans 20 mM Phosphate pH 6,5, 100 mM NaCl, 0,1% NaN_3 ; 14,1 T et 288 K), montrant la disparition du signal pour la forme 1 des résidus D235, E239, D240 et F241 liée à la proximité de la sonde paramagnétique sur C237.

En principe l'augmentation de relaxation paramagnétique (PRE) due à l'étiquette de spin est déterminée en comparant les intensités des signaux dans l'état paramagnétique et l'état diamagnétique ($I_{\text{para}} / I_{\text{dia}}$) à partir des spectres HSQC $^1\text{H}-^{15}\text{N}$ de la protéine modifiée avant et après traitement par l'acide ascorbique (vitamine C) qui permet de neutraliser le radical libre du PROXYL. Or après incubation avec la vitamine C, aucun signal n'est réapparu, ce qui laisse penser que le traitement a été inefficace. Les intensités de référence ont donc été prises dans le spectre des mutants avant fixation de la sonde paramagnétique. Les résultats sont donnés dans la Figure 82 pour P(S237C). On constate que les élargissements dus à l'effet PRE restent localisés autour du site de la mutation. Cela semble dire que la sonde n'est pas détectée par des résidus éloignés, même si on observe également un élargissement pour les résidus S45, S58 et S60. Pour confirmer ce résultat il faudrait recommencer l'expérience en prenant comme intensités de référence celles mesurées après neutralisation du radical libre PROXYL par la vitamine C.

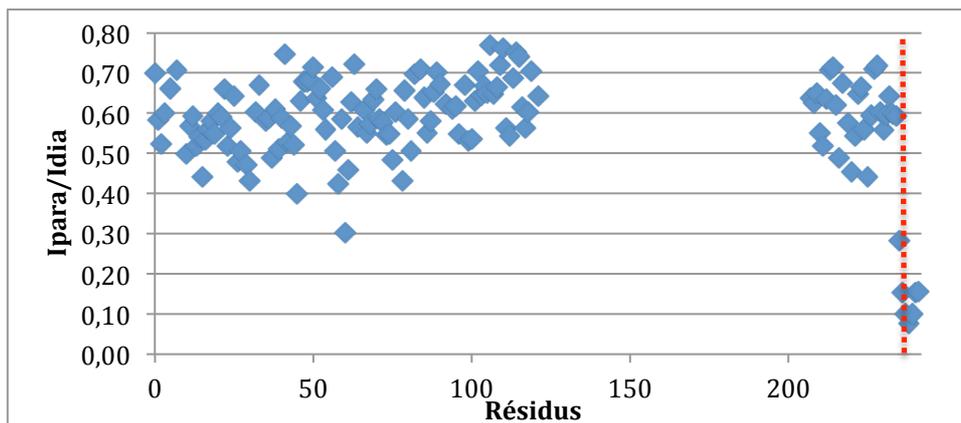


Figure 82 : Rapport des intensité des signaux des spectres HSQC ^1H - ^{15}N du mutant P(S237C) avec une étiquette paramagnétique IAP (Ipara) et avant traitement à l'IAP (Idia). La position de la sonde paramagnétique est indiquée par un trait rouge.

La Figure 83 montre que les mutations dans la région des hélices C-terminales ne permettent pas non plus de mettre en évidence des contacts à longue distance. Cependant les incertitudes sur l'état des échantillons ne permettent pas de conclure. Ces expériences devront être refaites en contrôlant précisément l'état des protéines et les rendements des différentes étapes de modifications chimiques. D'autre part les mutations portant sur des contacts possibles entre le domaine d'oligomérisation et les hélices C-terminales devront être faites sur un fragment de P plus court, P[127-205], pour lequel tous les signaux sont observables, contrairement à P entière.

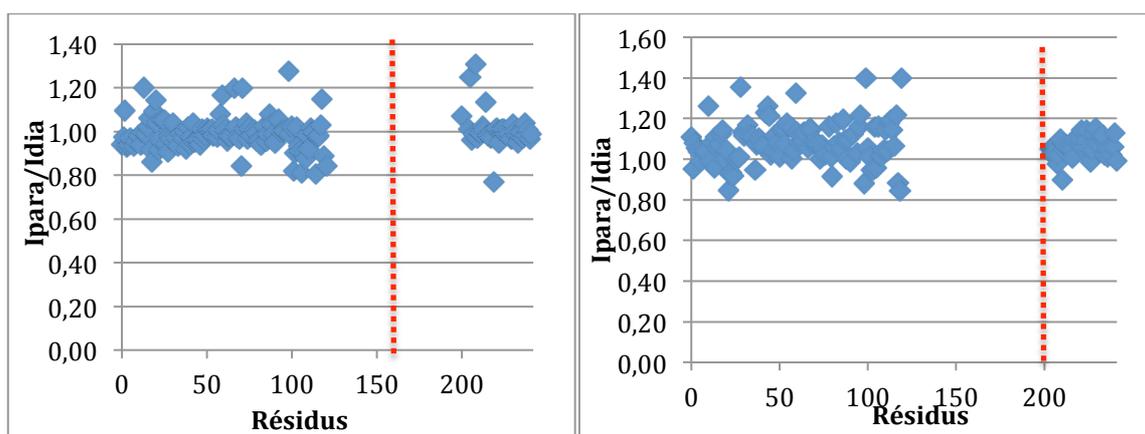


Figure 83 : Rapport des intensité des signaux des spectres HSQC ^1H - ^{15}N des mutants P(S156C) et P(E193C) avec une étiquette paramagnétique IAP (Ipara) et avant traitement à l'IAP (Idia). Ces mutants étaient intacts dans la région N-terminale. La position de la sonde paramagnétique est indiquée par un trait rouge.

3 Etude de l'interaction de la phosphoprotéine avec le domaine globulaire de M2-1 par RMN

3.1 Expériences d'interaction par RMN

Dans le chapitre précédent j'ai exposé les résultats obtenus en observant le domaine globulaire de la protéine M2-1 marquée en isotopes stables par RMN en présence du fragment P[100-160]. J'ai montré qu'ils étaient en accord avec les résultats de Mason et collaborateurs (Mason et al. 2003). Pour compléter ces expériences, nous avons voulu suivre la liaison du domaine globulaire de M2-1 à P en observant P afin de déterminer la zone d'interaction de M2-1 sur P. Les fragments de la phosphoprotéine correspondant à P[1-126] et P[1-163] ont été produits sous forme de protéines de fusion avec la glutathion-S-transférase (GST) contenant un site de clivage de la thrombine. Des expériences HSQC ^1H - ^{15}N des deux fragments N-terminaux de la phosphoprotéine P[1-126] et P[1-163] ont été mesurées en présence du domaine globulaire de M2-1 non marqué. L'expérience d'interaction avec P[1-126] est montrée dans la Figure 84.

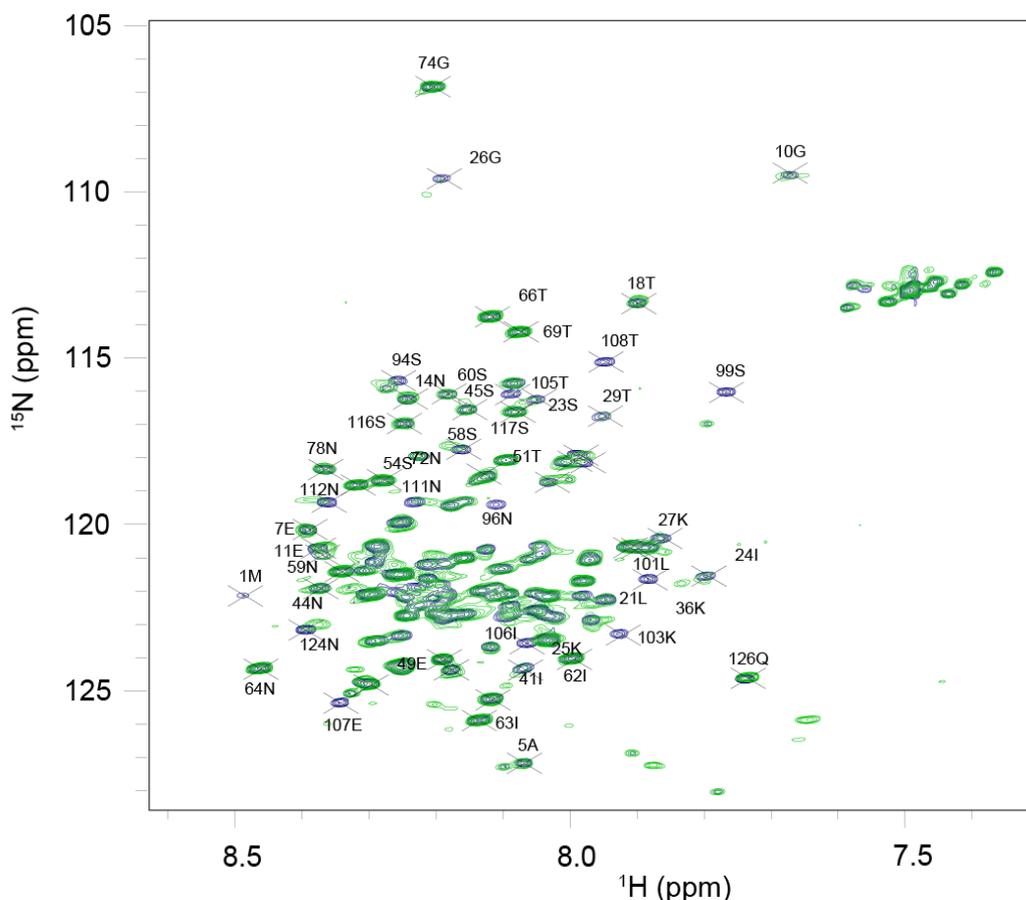


Figure 84 : Superposition des spectres HSQC ^1H ^{15}N de P[1-126] seule 50 μM (bleu) et en présence de M2-1[58-177] (vert). ($T = 288\text{ K}$, $B_0 = 14,1\text{ T}$)

La Figure 84 montre que dans des conditions équimolaires, une partie des signaux de P[1-126] est fortement élargie en présence du domaine globulaire de M2-1. Si l'on représente les variations d'intensité en fonction de la séquence en acides aminés, on constate que deux régions éloignées de P sont affectées (Figure 85) : d'une part la région des résidus I91-E113 pour laquelle le signal disparaît quasiment complètement, et d'autre part la région T18-I39, où les signaux sont encore visibles, mais fortement atténués. La première région contient les résidus identifiés par Mason et collaborateurs (Mason et al. 2003) comme étant critiques pour l'interaction avec M2-1. C'est aussi une région avec une petite propension de structuration en hélice α . On peut faire l'hypothèse que cette région dans la protéine P se structure complètement en hélice si elle s'associe avec M2-1. La deuxième région correspond aussi à une région avec une faible propension en hélice α . Elle pourrait former des contacts transitoires avec M2-1. On peut aussi imaginer qu'elle serait en contact avec la première région et que la fixation de P perturberait ces contacts.

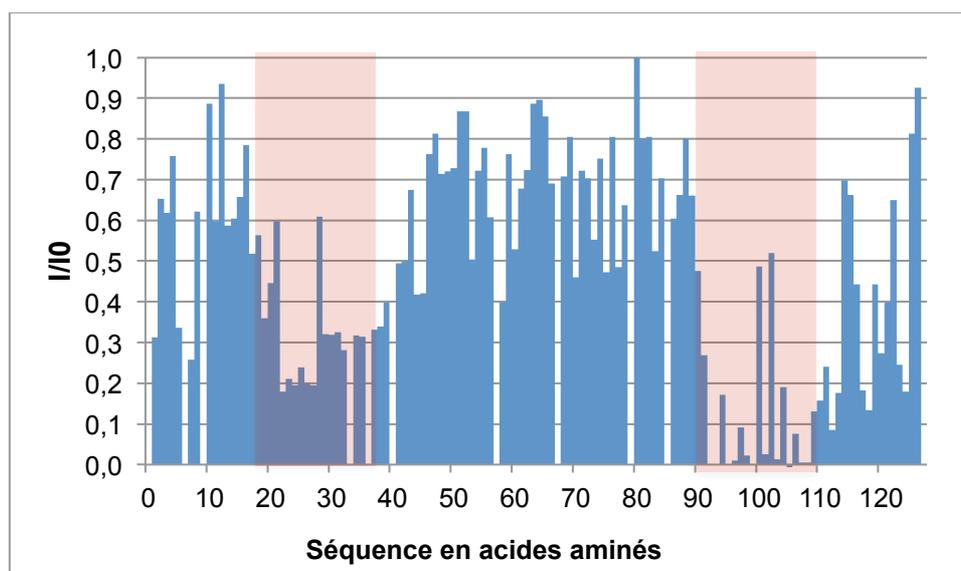


Figure 85 : Rapport des intensités résidu par résidu des signaux de P[1-126] en présence et en absence du domaine globulaire de M2-1. Les régions avec des variations significatives apparaissent en rouge (T18-I39) et (I91-E113).

L'expérience d'interaction avec P[1-163] est montrée dans la Figure 86. Les variations d'intensité en fonction de la séquence en acides aminés sont représentées dans la Figure 87. On retrouve la quasi disparition du signal de la région 90-110 déjà observée avec P[1-127]. Mais cette fois-ci toute la région 90-155 est absente du spectre. Par contre la baisse de signal dans la région 20-40 est moins prononcée.

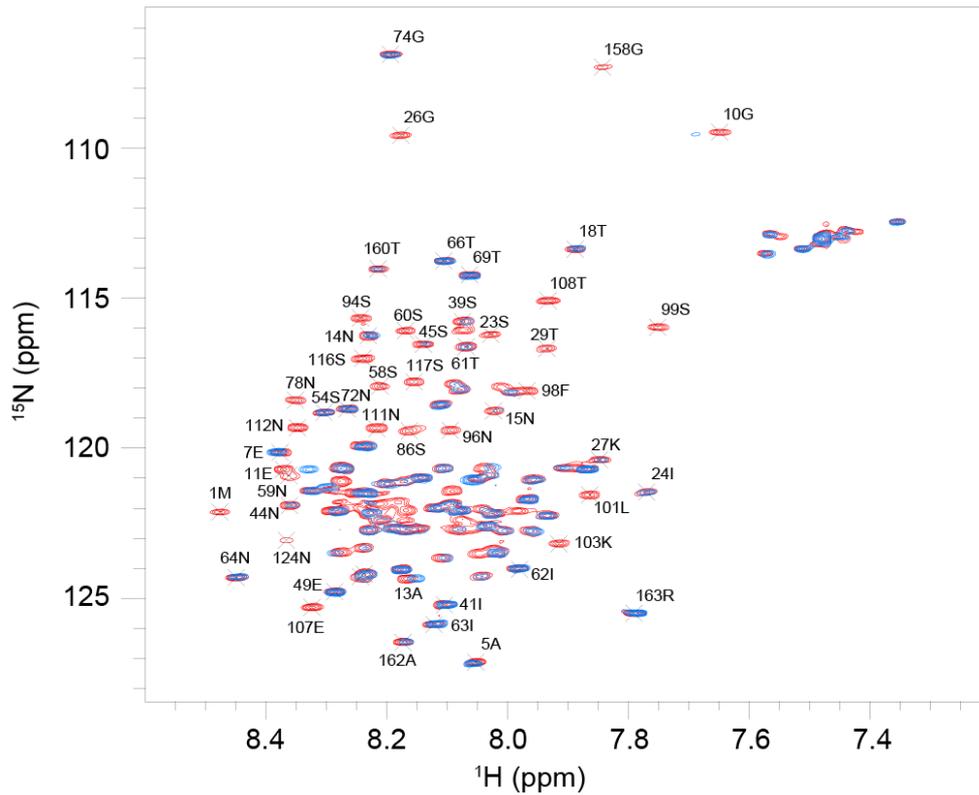


Figure 86 : Superposition des spectres HSQC ^1H - ^{15}N de P[1-163] seule (bleu) et en présence de M2-1[58-177] (rouge). ($T = 288\text{ K}$, $B_0 = 14,1\text{ T}$).

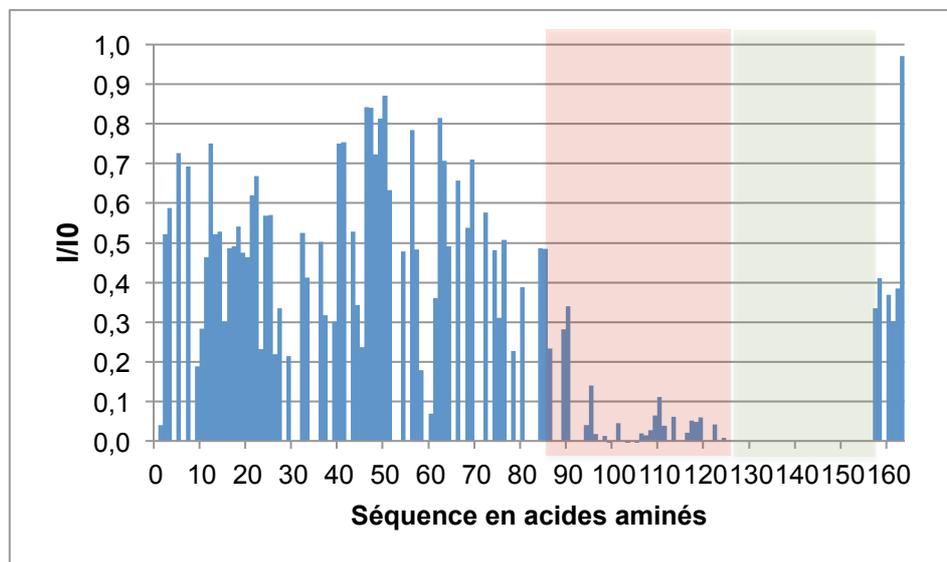


Figure 87 : Rapport des intensité résidu par résidu des signaux de la partie N-terminale de la phosphoprotéine P avec le domaine d'oligomérisation P[1-163] en présence et en absence du domaine globulaire de M2-1. La région avec un baisse significative apparaît en rouge (191-125). Le domaine d'oligomérisation pour lequel aucun signal n'est observé apparaît sur fond vert

4 Conclusion

Pour résumer, les expériences d'interaction entre les deux fragments N-terminaux de P avec le domaine globulaire de la protéine M2-1 mettent en évidence un domaine d'interaction entre ces deux protéines dans la partie correspondant aux résidus 95-104. Ceci est en accord avec les résultats publiés dans (Mason et al. 2003) qui montrent que la région de la phosphoprotéine P impliquée dans l'interaction avec la protéine M2-1 se situe entre les résidus 100 et 120. Cette région forme une hélice transitoire et pourrait se structurer dans le complexe avec M2-1.

Une autre région de P semble être sensible à la présence de M2-1, comprenant les résidus 13-39. Elle correspond à une autre partie de P avec une propension en hélice α dans la région 13-25. Cette région est prédite pour interagir avec la nucléoprotéine N et former le complexe N_0 : P. Le double effet de M2-1 sur les spectres de P est surprenant et pourrait signifier que ces deux régions ne sont pas indépendantes, mais pourraient interagir entre elles.

En revanche nous n'avons pas pu mettre en évidence des contacts à longue distance entre les domaines N et C-terminaux de P, ni à l'intérieur du domaine N-terminal de P, par des expériences de PRE. Cependant ces expériences seraient à refaire, en contrôlant l'état des échantillons et en utilisant des fragments plus courts, contenant toutefois le domaine d'oligomérisation de P.

La Figure 88 résume les informations obtenues sur la structure de la phosphoprotéine. Elle montre que sauf pour la partie C-terminale de P, les régions d'interaction avec des partenaires de P correspondent à des zones de structuration partielle en hélice α . La région des hélices C-terminales (172-205) par contre ne correspond pas à une zone de fixation d'un partenaire identifié.

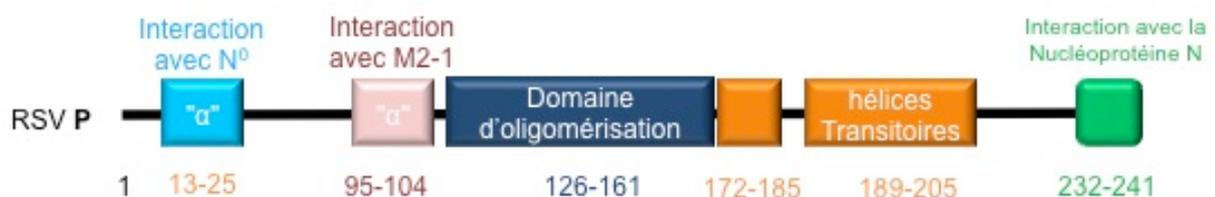


Figure 88 : Récapitulatif des informations fonctionnelles et structurales sur la phosphoprotéine du VRS.

Chapitre 5 : : Les protéines NS1 et NS2 du VRS

1 Introduction et objectifs

Les gènes les plus abondamment transcrits du VRS encodent les protéines non structurales NS1 et NS2. Il s'agit de deux petites protéines de 139 et 124 acides aminés, pour NS1 et NS2 du VRSH respectivement, qui sont spécifiques des *Pneumovirus*. Elles jouent un rôle dans le mécanisme d'évasion du système immunitaire de la cellule par le VRS et dans la spécificité de l'hôte.

Il a été montré que la souche A2 du VRS était résistante aux effets antiviraux des interférons (IFN) (Atreya and Kulkarni 1999). Or les protéines NS1 et NS2 du VRS sont capables de cibler différents acteurs de la voie de signalisation de l'IFN tels que IRF3, NF-KB ou STAT2 (Bossert et al. 2003; Spann et al. 2005; Swedan et al. 2009). Cependant les mécanismes moléculaires restent mal caractérisés et on ne sait pas actuellement quelles molécules dans ces voies interagissent directement avec ces protéines

D'autre part, la délétion simple ou double des gènes encodant les protéines NS1 et NS2 n'est pas létale pour le VRSH, mais le virus pousse plus lentement en culture cellulaire (Teng and Collins 1999; Schlender et al. 2000) et donne un phénotype atténué par rapport à la souche sauvage lors d'une infection chez le chimpanzé (Whitehead et al. 1999; Teng et al. 2000).

Aucune structure tridimensionnelle des protéines non structurales du VRS n'a été résolue pour l'instant. Toutefois, des études de co-immunoprécipitation dans des cellules eucaryotes ont montré que les protéines NS du VRS sont capables d'interagir pour former des homodimères, mais aussi des hétérodimères (Swedan et al. 2009; Swedan et al. 2011). NS1 se présente sous forme monomère dans certaines conditions, mais dans d'autres conditions comme à faible pH ou à des températures physiologiques, NS1 peut former des oligomères qui présentent des propriétés amyloïdes (Pretel et al. 2013). Dans ce contexte j'ai cherché à caractériser la structure de ces protéines. La taille semblait appropriée à une étude par RMN. L'enjeu était donc de produire ces protéines sous leur forme monomère dans un premier temps.

2 Rôles et fonctions des protéines non structurales du VRS

2.1 Les voies interférons

Le succès de la réponse antivirale de l'hôte repose sur plusieurs familles de récepteurs qui reconnaissent des structures conservées d'agents pathogènes, ce qui aboutit à la production de cytokines et de chimiokines pour lutter contre l'infection virale et induire une réponse antivirale acquise. Trois grandes familles de récepteurs d'agents pathogènes ont été décrites : les récepteurs transmembranaires Toll-like (TLR), les récepteurs cytosoliques Nod-like (NLR) et les récepteurs cytosoliques RIG-I-like (RLR). Au niveau cellulaire, la production d'interféron de type I est régulée à plusieurs niveaux. L'action synergique des facteurs de transcription activés tels que NFκB et les facteurs de régulation d'interféron (Interferon Regulatory Factor, IRF) sur les promoteurs des gènes de l'IFN de type I déclenche une réponse immédiate qui est amplifiée par la voie de signalisation JAK-STAT. Les voies interférons de la cellule peuvent être divisées en deux groupes fonctionnels : d'une part la voie d'induction d'interférons, dans laquelle les cellules produisent les interférons, et d'autre part la voie de réponse aux interférons, dans laquelle les cellules répondent à des interférons exogènes.

L'une des premières étapes dans la première voie est l'induction de l'activation des hélicases à ARN cytoplasmiques de la famille RIG-I (Yoneyama et al. 2004; Yoneyama et al. 2005; Takahasi et al. 2008; Myong et al. 2009). Ceci déclenche une cascade d'événements de signalisation où RIG-I interagit avec et active la protéine mitochondriale MAVS via des séquences CARD ou CARD-like (Kawai et al. 2005; Chau et al. 2008). MAVS activée active ensuite TRAF3 (Saha and Cheng 2006; Chau et al. 2008), qui active à son tour deux kinases, IKKε et TBK1. Ces dernières sont des sérine/thréonine kinases qui phosphorylent le domaine C-terminal des facteurs de régulation d'interféron 3 et 7 (IRF3 et IRF7), conduisant à leur activation et leur translocation du cytoplasme vers le noyau (Chau et al. 2008), où ils activent les gènes de la transcription d'interféron de type I (Marie et al. 1998; Sato et al. 2000).

Ces interférons synthétisés sont sécrétés hors de la cellule, où ils se lient à des récepteurs d'interférons de type I, ce qui déclenche la voie de réponse aux interférons. Celle-ci consiste en une autre cascade de signalisation qui aboutit à la phosphorylation et l'activation des facteurs de transcription STAT1 et STAT2 et conduit ainsi à l'activation de la transcription d'une grande famille de gènes sensibles aux interférons. De nombreux gènes

dans cette famille codent pour des protéines antivirales, ce qui conduit au final à un état de résistance de la cellule au virus (Basler and Garcia-Sastre 2002).

2.2 Rôle de la protéine NS1

Swedan et collaborateurs ont montré que la protéine NS1 du VRS humain contribuait au contournement de la réponse immunitaire innée de l'hôte par le VRS en intervenant dans la voie d'induction des interférons en diminuant les concentrations stationnaires de TRAF3 et de IKK ϵ (Swedan et al. 2009). Ces résultats sont résumés dans la Figure 89.

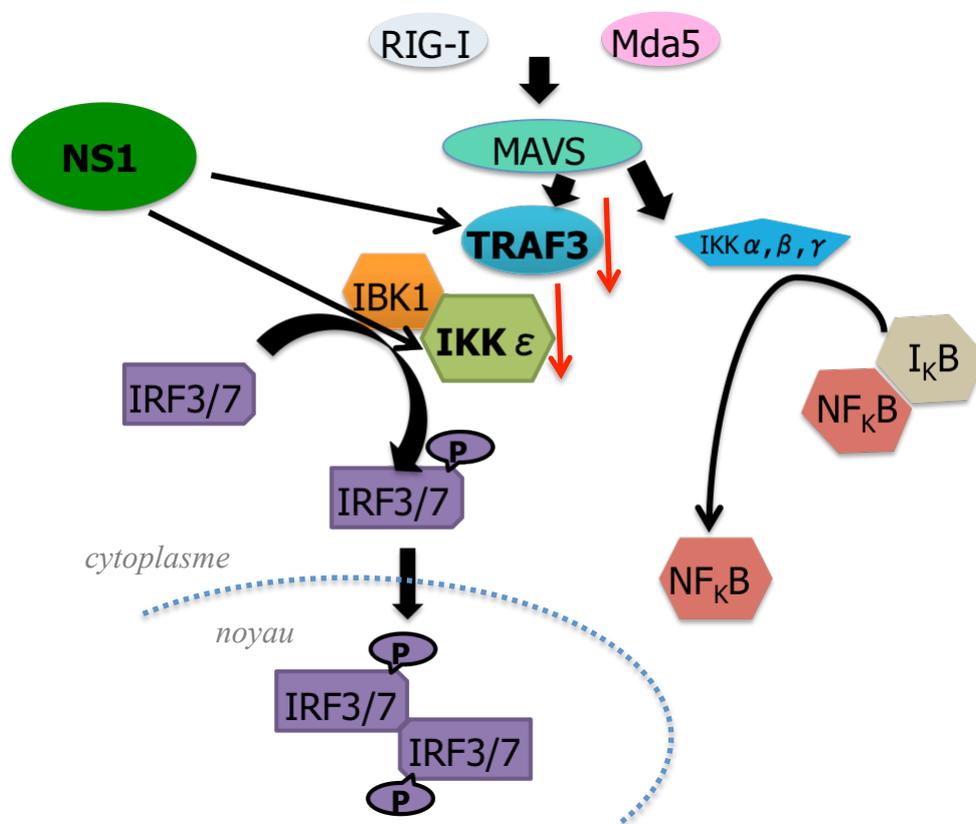


Figure 89 : Rôle de la protéine NS1 du VRS dans la voie d'induction d'interféron de la cellule. NS1 a pour effet de diminuer les concentrations stationnaires de TRAF3 et IKK ϵ (flèches rouges), ce qui diminue le taux de transcription des gènes des IFN de type I. En effet, TRAF3 et IKK ϵ activent les facteurs de régulation d'interféron IRF3 et IRF7 par phosphorylation. Ceux-ci sont ensuite transloqués vers le noyau, où ils activent la transcription des gènes des IFN de type I. L'activation est symbolisée par des flèches noires.

2.3 Rôle de la protéine NS2

L'effet prédominant de la protéine NS2 est de diminuer le niveau d'équilibre du facteur de transcription STAT2 (Signal Transducers and Activators of Transcription 2) (Swedan et al. 2011). La cascade de signalisation de la voie de réponse aux interférons qui passe par la phosphorylation des facteurs de transcription STAT1 et STAT2 est illustrée dans la Figure 90.

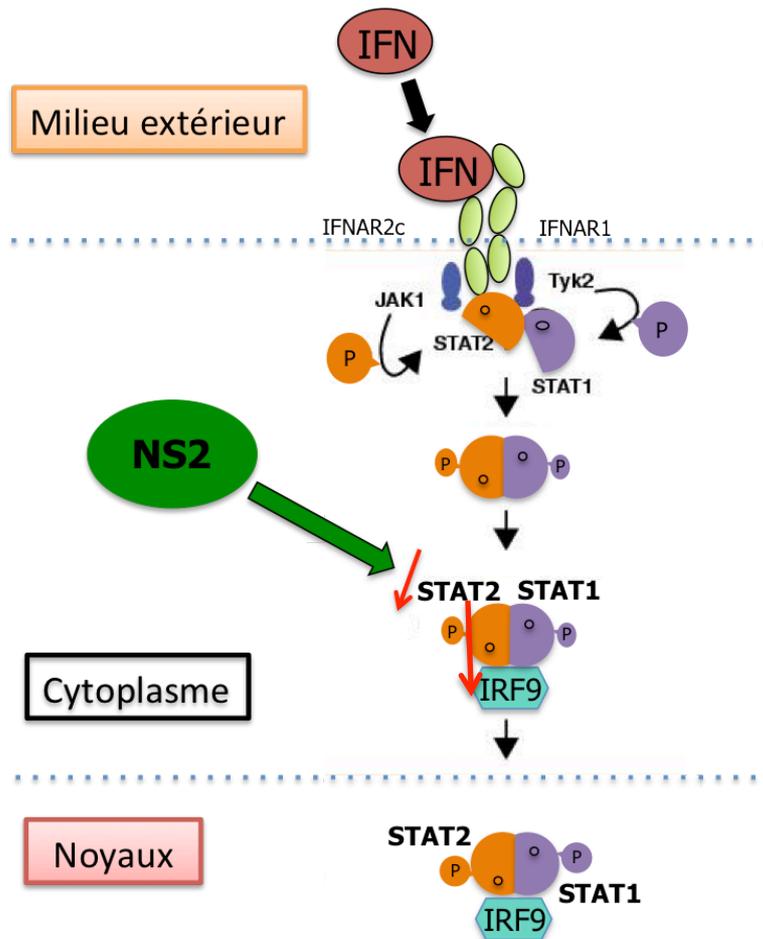


Figure 90 : Rôle de la protéine NS2 du VRSV dans la voie de réponse aux interférons. Les interférons de type I se lient à leurs récepteurs IFNAR2c et IFNAR1, ce qui stimule la phosphorylation des facteurs de transcription STAT1 et STAT2 par les kinases Tyk2 et JAK1 et la formation d'un complexe entre STAT2, STAT1 et le facteur de régulation d'interféron IRF9. Celui-ci est ensuite transloqué dans le noyau pour activer la transcription de gènes codant pour des protéines antivirales. L'effet de NS2 est de réduire le taux de STAT2 (flèche rouge).

2.4 Oligomérisation de NS1 et NS2

Bien que les rôles individuels de NS1 et NS2 soient décrits dans les différentes voies des interférons, il semblerait que ces protéines fonctionnent en synergie. En effet dans des mutants de virus recombinants bovins, les protéines NS1 et NS2 sont nécessaires de manière coopérative pour conférer la résistance à la réponse des interférons (Schlender et al. 2000).

Il a été montré plus récemment que NS1 et NS2 pouvaient se présenter sous forme monomère ou homo-oligomère (par exemple, NS1-NS1) (Swedan et al. 2009). Les protéines NS1 et NS2 peuvent aussi se lier les unes aux autres pour former des oligomères hétérogènes (Swedan et al. 2011). Le nombre exact de sous-unités dans un multimère (homo- ou hétéro-oligomères) n'est pas connu. Ces résultats ont fourni une explication pour le double effet suppresseur d'interférons des protéines non structurales du VRS, à la fois sur l'induction d'interférons et sur les voies de réponse aux interférons. La formation d'hétéro-oligomères donne lieu à la possibilité d'autres interactions plus complexes avec leurs cibles et avec les protéines de l'hôte (Swedan et al. 2011).

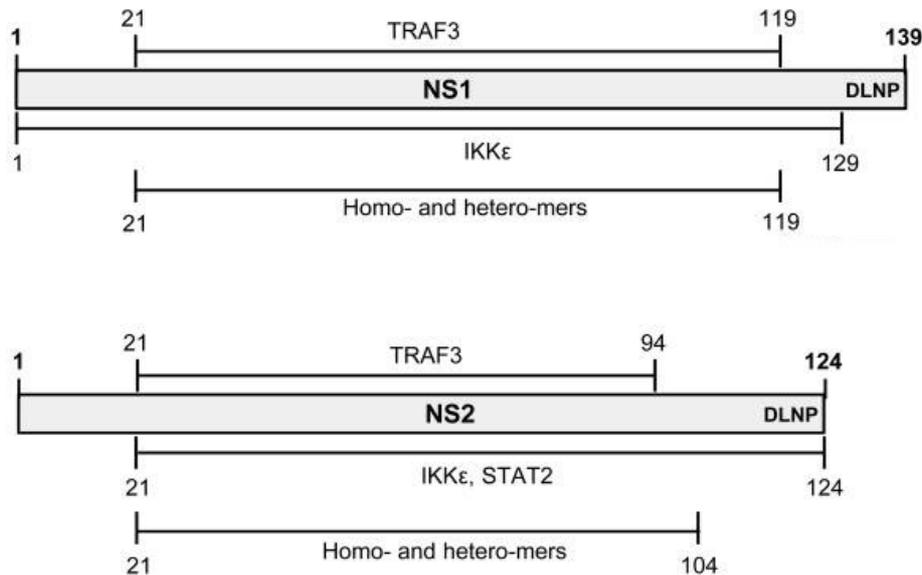


Figure 91 : Domaines fonctionnels des protéines NS1 et NS2. L'analyse des propriétés de liaison aux protéines TRAF3, IKKε et STAT2 de mutants de délétion de NS1 et NS2 a permis de définir des domaines fonctionnels. La séquence térapeptidique C-terminale DLNP, commune à NS1 et NS2, n'intervient pas dans la formation d'oligomères. Schéma adapté de (Swedan et al. 2011).

3 Production et purification de la protéine NS1

3.1 Plasmides

Au cours de ma thèse je me suis intéressée à une seule de ces protéines, à savoir NS1. Un premier plasmide avait été fourni par le Dr Jean-François Eléouët, correspondant à la protéine NS1 de la souche *Long* en fusion avec la GST. Un autre plasmide permettant de produire NS1 de la souche A en fusion avec la protéine de liaison au maltose (MPB) a été fournie par le Dr Gonzalo Prat-Gay (Institut Leloir, Buenos Aires, Argentine). Le résultat du séquençage des deux plasmides est donné dans la Figure 92.

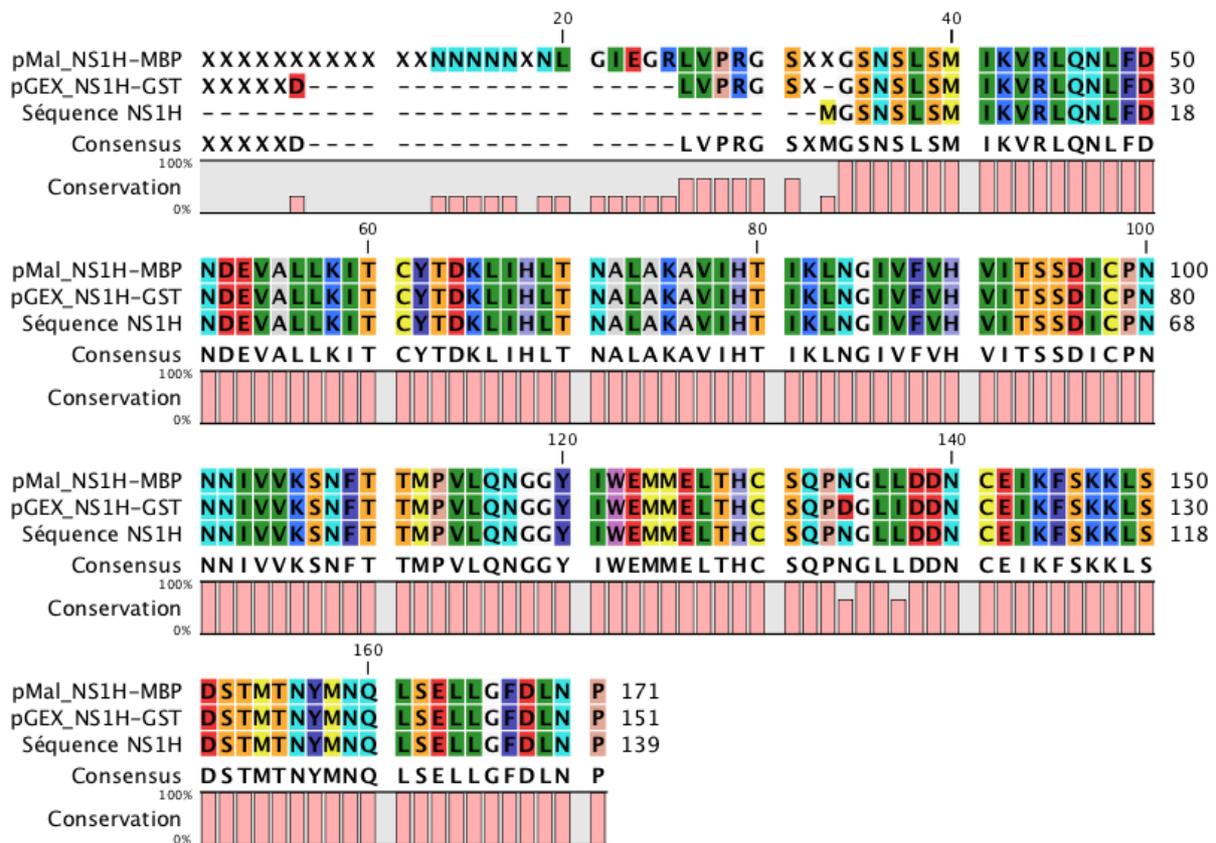


Figure 92: Aligement de séquence pour la protéine NS1 (plasmides pMal et pGEX).

3.2 Expression de MBP-NS1

J'ai choisi de produire la protéine NS1 sous forme de protéine de fusion avec la MBP (Maltose Binding Protein), le plasmide pMal-C. NS1 a été exprimée dans la souche bactérienne *E. coli* BL21 (DE3) avec un site de clivage pour la thrombine. Elle doit être marquée avec des isotopes stables pour l'étude par RMN. Elle a donc été produite dans un milieu de culture minimum M9 contenant du $^{15}\text{NH}_4\text{Cl}$ comme source d'azote ^{15}N . Le protocole détaillé de production est décrit en Annexe. Le gel d'expression présenté dans la Figure 93

montre une bande intense après induction, un peu au-dessus de 50 kDa, qui correspond à la protéine NS1 (15,6 kDa) en fusion avec la MBP (~40 kDa).

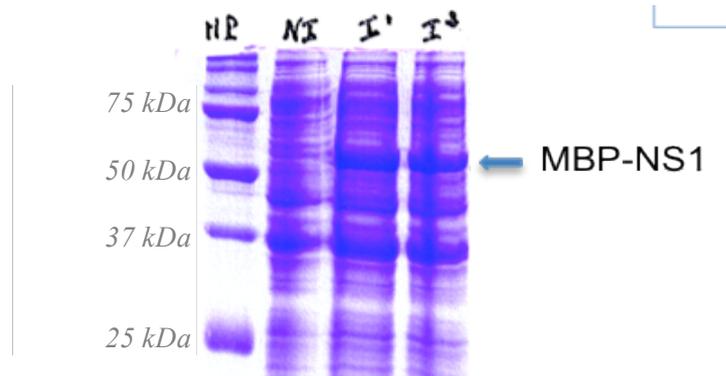


Figure 93 : Gel d'expression de polyacrylamide pour NS1 obtenu par électrophorèse en milieu dénaturant (SDS-PAGE). MP = marqueur de poids moléculaire. NI = extrait bactérien avant induction par l'IPTG. I = extrait bactérien après induction. La bande correspondant à MBP-NS1 est indiquée par une flèche.

3.3 Purification de MBP-NS1 par affinité sur résine amylose

La première étape consiste en une précipitation au sulfate d'ammonium (50 %), suivie d'une dialyse pour éliminer le sulfate d'ammonium, selon des conditions mises au point dans le laboratoire du Dr Gonzalo Prat-Gay.

La deuxième étape comporte la fixation de la protéine sur résine amylose grâce à l'étiquette MBP. La protéine est éluée avec un tampon qui contient 20 mM de maltose. L'efficacité de cette étape de purification a été vérifiée par SDS-PAGE (Figure 94).

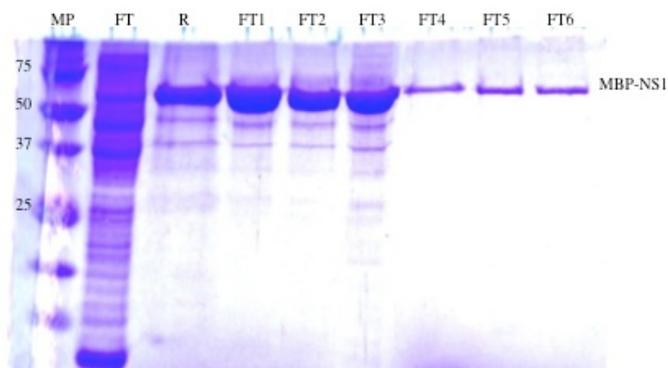


Figure 94 : Gel de protéine SDS-PAGE après élution au maltose de la fusion MBP-NS1 (MP : marqueur de poids moléculaire, R : résine amylose, FT1 à FT6 : fractions non retenues qui contiennent la fusion MBP-NS1).

Ensuite, on rassemble et concentre les fractions contenant la fusion MBP-NS1 sur membrane jusqu'à l'obtention d'une concentration finale de 10 mg/mL dans 2 mL.

L'étape suivante consiste à incuber l'échantillon avec de la thrombine, une protéase qui coupe la séquence d'acides aminés LVPRGS introduite entre la protéine d'intérêt et la MBP. L'efficacité de la coupure et la pureté de la préparation de protéine sont vérifiées par SDS-PAGE en comparant l'échantillon avant et après incubation. On trouve bien deux bandes, l'une vers 15 kDa correspondant à la protéine d'intérêt NS1, et l'autre vers 40 kDa correspondant à l'étiquette MBP (Figure 95).

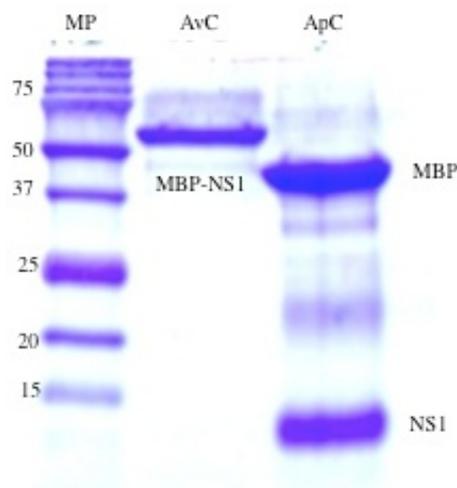


Figure 95 : Gel de protéine SDS-PAGE après coupure à la thrombine de NS1. MP : marqueur de poids moléculaire, AvC : fusion MBP-NS1 avant coupure, ApC: protéines après coupure avec deux bandes correspondant à la MBP et à NS1.

3.4 Etape de polissage « polishing » par filtration sur gel

La solution à ce stade contient un mélange de la protéine NS1, la MBP et la thrombine, ce qui nécessite une étape supplémentaire de filtration sur gel. Le principe d'éluion d'une colonne de filtration sur gel consiste à séparer les constituants d'un mélange selon la taille des particules. Le tampon choisi est celui envisagé pour les études par RMN. Le résultat est montré dans la

Figure 96.

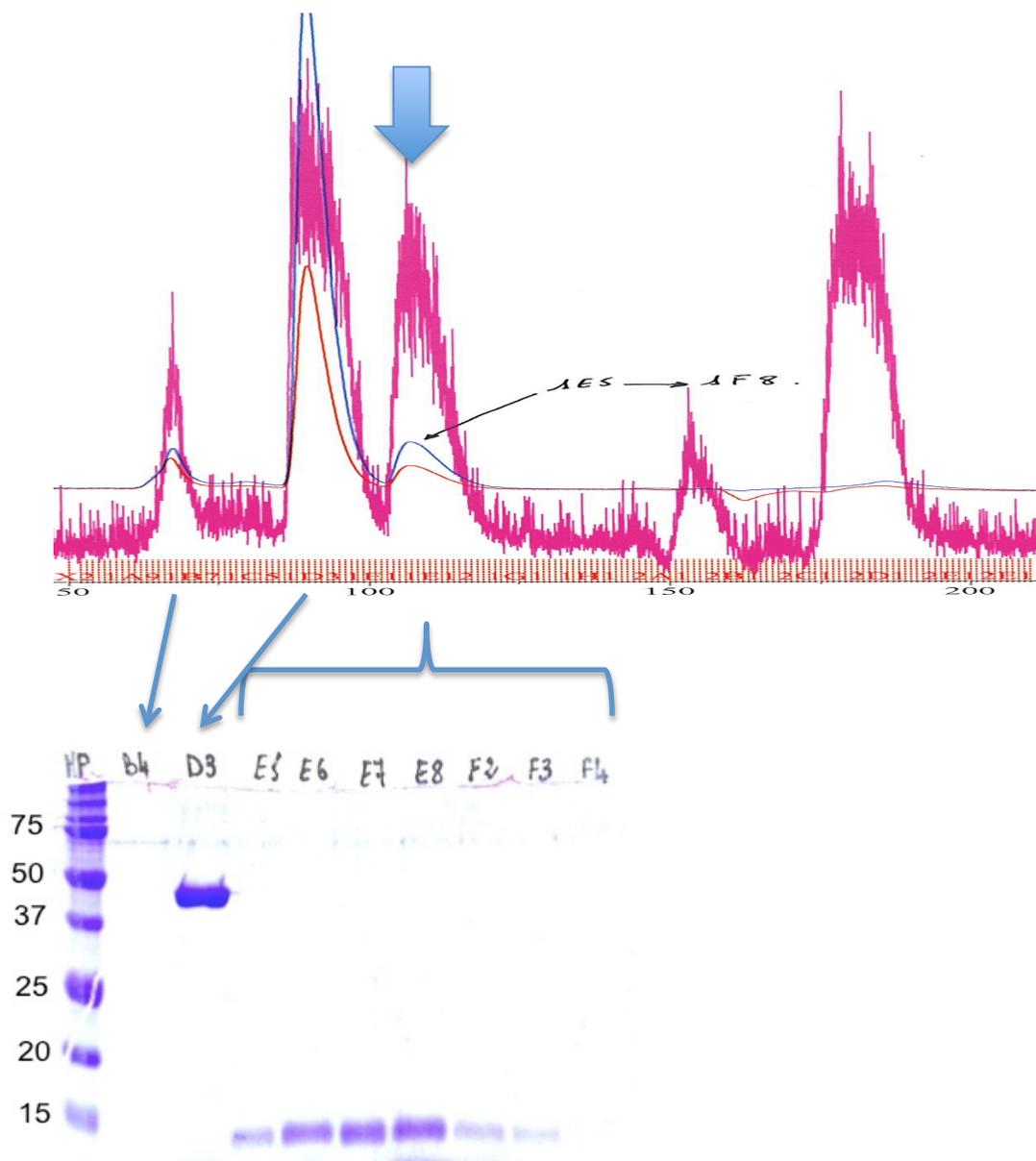


Figure 96 : Profil d'éluion du mélange par suivi de l'absorption à trois longueurs d'onde (282 nm en bleu, 258 nm en rouge et 215 nm en rose) et analyse par SDS-PAGE des fractions collectées lors de la purification par filtration sur gel sur colonne Superdex-75. Le premier pic (fraction B4) semble contenir des agrégats. Le deuxième pic (fraction D3) correspond à la taille de la MBP. Le troisième pic indiqué par une flèche épaisse est celui qui correspond au volume d'éluion attendu pour NSI sous forme monomère. Plusieurs fractions ont été déposées (E5, E6, E7, E8, F2, F3, F4) : elles correspondent à la protéine d'intérêt NSI.

3.5 Difficultés rencontrées pour la production de NS1

Cette protéine commence directement à former des agrégats après coupure à la thrombine. J'ai donc testé différentes conditions à chaque étape de la purification afin d'assurer la stabilité et la solubilité de la protéine. Toutes les étapes de purification ont été réalisées à 4°C. J'ai testé plusieurs tampons de purification en gel filtration. Des profils différents ont été trouvés, mais je n'ai réussi à récupérer la protéine NS1 monomère que dans le tampon 10 mM Tris-HCl pH 7,5, 200 mM NaCl, 1 mM DTT et 0,5 mM EDTA, noté "tampon final d'élution" dans la suite (

Figure 96). Une grande quantité de la protéine était déjà agrégée sur le profil d'élution (Fraction B4 sur la

Figure 96).

A la sortie de la colonne de gel filtration la protéine était trop diluée (10 µM) pour pouvoir être utilisée directement pour la RMN. Au départ j'ai essayé de la concentrer jusqu'à 100 µM par centrifugation sur membrane (Amicon Ultra Centrifugal Filters de 10 kDa à 4°C), mais le problème majeur était de concentrer cette protéine sans former des agrégats. J'ai remarqué qu'elle ne supportait pas des concentrations au-delà de 40 µM dans le tampon final.

Bien que les conditions n'étaient pas optimales pour analyser la protéine par RMN, j'ai acquis des spectres RMN pour voir s'il était possible d'obtenir du signal malgré l'agrégation de la protéine.

4 Analyse par RMN de la protéine NS1-¹⁵N

Des spectres RMN HSQC ¹H-¹⁵N ont été acquis à différentes concentrations de NS1 marquée en azote 15 et dans différents tampons (Figure 97). Toutes les mesures ont été effectuées à 293 K sur un spectromètre RMN à 14,1 T ou 18,8 T (600 ou 800 MHz en ¹H). Les meilleurs résultats ont été obtenus avec un échantillon à 40 µM dans le "tampon final d'élution" correspondant à du Tris salin à pH 7,5. Au-delà de cette concentration, il n'y avait que très peu de signal indiquant que la protéine n'était probablement plus monomérique et que des oligomères s'étaient formés.

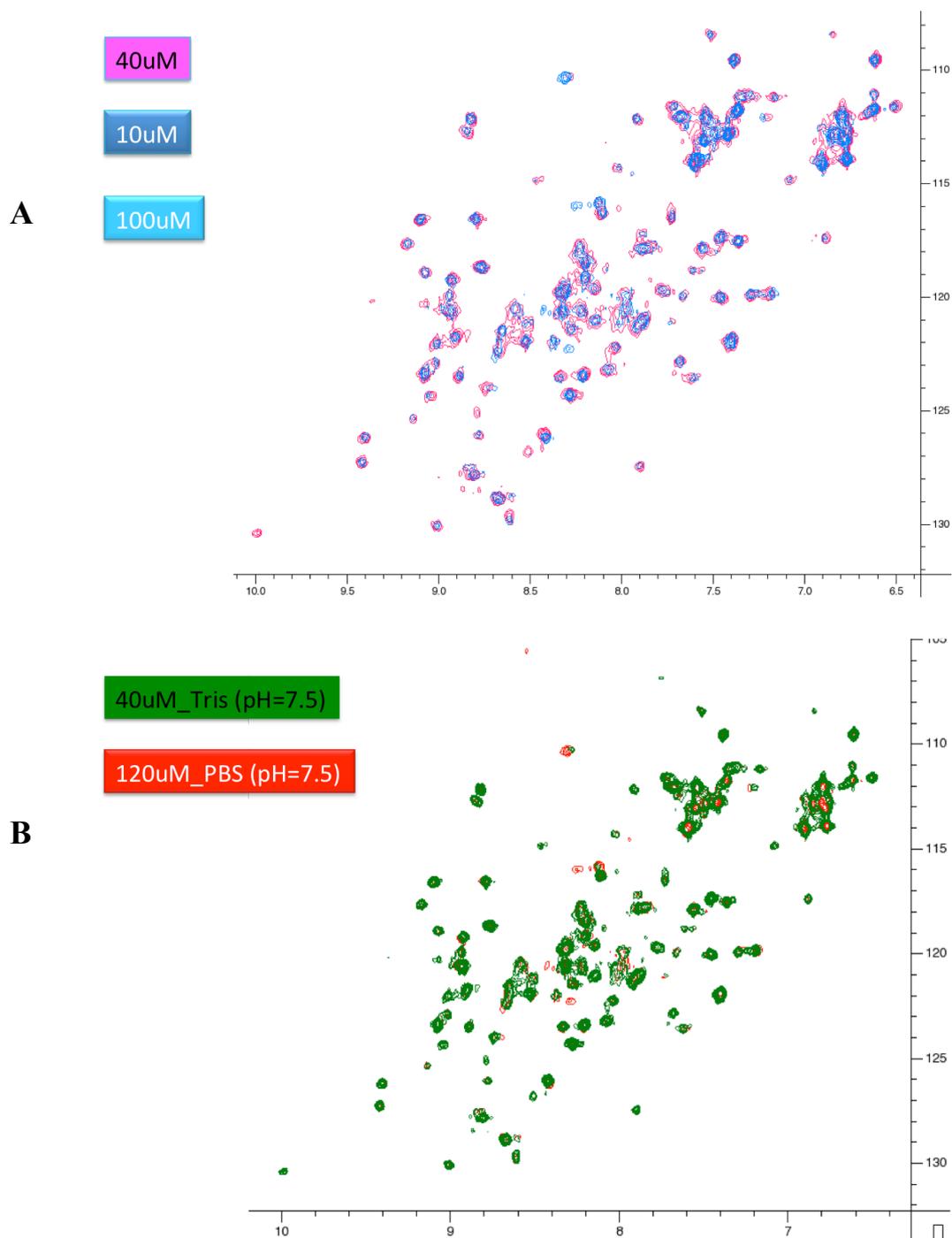


Figure 97 : A). Superposition des spectres HSQC $^1\text{H}-^{15}\text{N}$ de la protéine NSI à différentes concentrations (10, 40 et 100 μM) acquis à un champ de 18,8 T (256 scans), à une température de 293 K et dans un même tampon ("tampon final" : 10 mM Tris-HCl pH 7,5, 200 mM NaCl, 1 mM DTT et 0,5 mM EDTA).

B) Superposition des spectres HSQC $^1\text{H}-^{15}\text{N}$ de la protéine NSI à différentes concentrations et dans deux tampons différents, "tampon final" en vert et PBS en rouge.

Les spectres HSQC ^1H - ^{15}N montrent que NS1 est une protéine structurée, puisque les signaux sont bien éclatés dans la dimension ^1H . Même si à ce stade on n'a pas accès à sa structure secondaire, on peut supposer qu'elle contient du feuillet β , puisque beaucoup de déplacements chimiques sont supérieurs à 9 ppm. Cela serait cohérent avec des études précédentes en dichroïsme circulaire faites par le groupe du Dr Gonzalo Prat-Gay montrant qu'elle a une forte tendance à former des feuilletts β (Pretel et al. 2013).

Pour pouvoir continuer les études par RMN, nous avons besoin d'une protéine monomère et stable, à des concentrations de l'ordre de 100 μM . Le nombre de scans nécessaires pour obtenir un spectre HSQC ^1H - ^{15}N de bonne qualité ne permet pas pour le moment d'envisager des expériences tridimensionnelles pour l'attribution du squelette peptidique. J'ai donc effectué des mesures de thermostabilité par fluorescence dans le but de trouver des meilleures conditions de stabilité.

5 Mesure de la thermostabilité de NS1 par fluorescence

5.1 Principe

Pour la réalisation de cette étude, j'ai travaillé avec les Dr Eric Jacquet et Naima Nhiri à la plateforme de criblage de thermostabilité des protéines par fluorescence sur le campus de Gif-sur-Yvette. Cette plateforme utilise un appareil de PCR en temps réel (ABI 7900 HT) qui permet de mesurer la dénaturation des protéines induite par un gradient de température, en suivant la fluorescence d'une molécule fluorescente ajoutée à la solution de protéine. Lorsque le fluorophore (SYPRO orange) est ajouté à une protéine native en solution, il est exposé à un environnement aqueux et l'émission de fluorescence est faible. Lorsque la température augmente, la protéine subit une dénaturation et expose ses régions hydrophobes. Le fluorophore se lie à ces régions hydrophobes et devient fortement fluorescent. Cela se traduit par une augmentation du signal de fluorescence traduisant une dénaturation de la protéine. Cette méthode, dite "Thermofluor", a été mise au point pour tester la stabilité relative d'une protéine en présence de ligands et est aussi connue sous le terme de "Thermal Shift Assay", puisque la stabilisation/déstabilisation par des ligands entraîne un décalage de la température de dénaturation de la protéine.

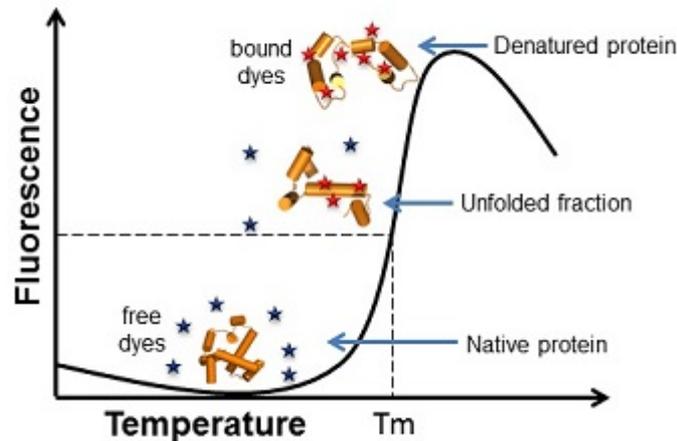


Figure 98 : Principe de la mesure de stabilité d'une protéine par "Thermofluor" (<http://www.jenabioscience.com>). La protéine est symbolisée par des cylindres oranges, la molécule fluorescente par des étoiles bleues ou rouge, suivant l'intensité de fluorescence.

Cette méthode permet de cribler rapidement de nombreuses conditions de tampon (pH, conditions salines) avec seulement quelques microgrammes de protéine, afin de déterminer celles qui seraient les plus favorables à la stabilité de la protéine, ceci dans le but d'optimiser son rendement de purification et sa conservation sous forme native. Dans mon cas la protéine doit être suffisamment stable pour permettre de réaliser des expériences de RMN qui nécessitent plusieurs heures, voire plusieurs jours de mesure.

5.2 Application du Thermofluor à NS1

5.2.1 Détermination de la température de dénaturation de NS1

La Figure 99 montre qu'avec des concentrations croissantes de la protéine NS1 dans le "tampon final", le signal global de fluorescence augmente. On peut distinguer une transition vers 56°C. Mais le bruit de fond à 25°C augmente aussi beaucoup avec la concentration et donc il est probable que la protéine expose déjà beaucoup de régions hydrophobes à cette température.

La Figure 100A montre la dérivée de la courbe de fluorescence qui permet de mesurer la température à demi-dénaturation. Elle est d'environ 55°C dans le tampon final de conservation de NS1. Cette valeur est compatible avec la température de dénaturation mesurée par dichroïsme circulaire par l'équipe du Dr Gonzalo Prat-Gay (Pretel et al. 2013) qui a aussi montré que l'état de dénaturation était irréversible (Figure 100B).

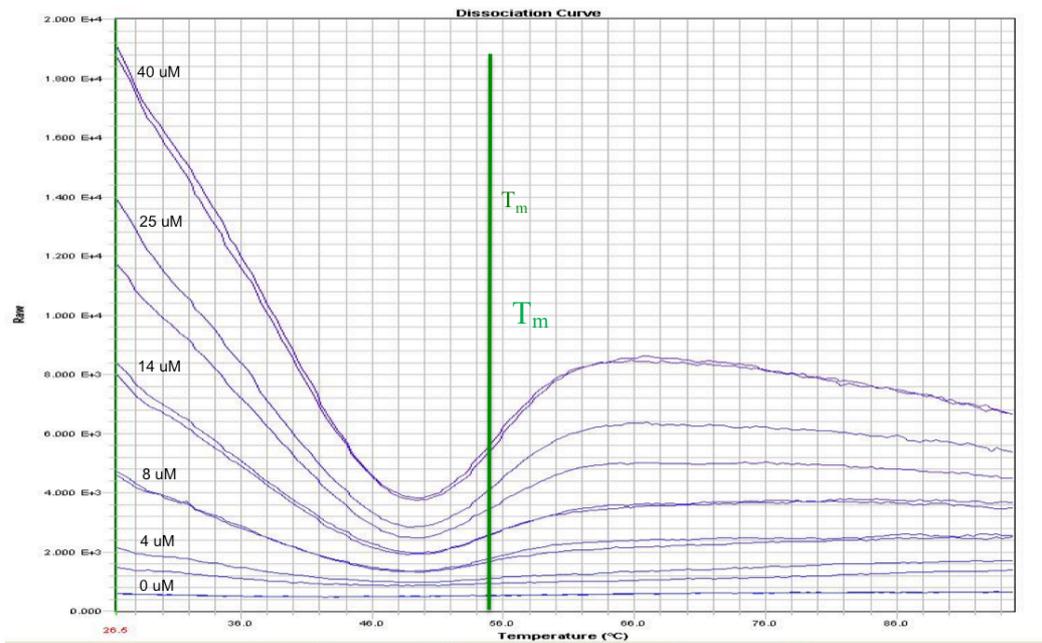


Figure 99 : Mesure en fonction de la température de la fluorescence du SYPRO avec des quantités croissantes de la protéine NSI dans le "tampon final de purification". Les mesures ont été effectuées en duplicata.

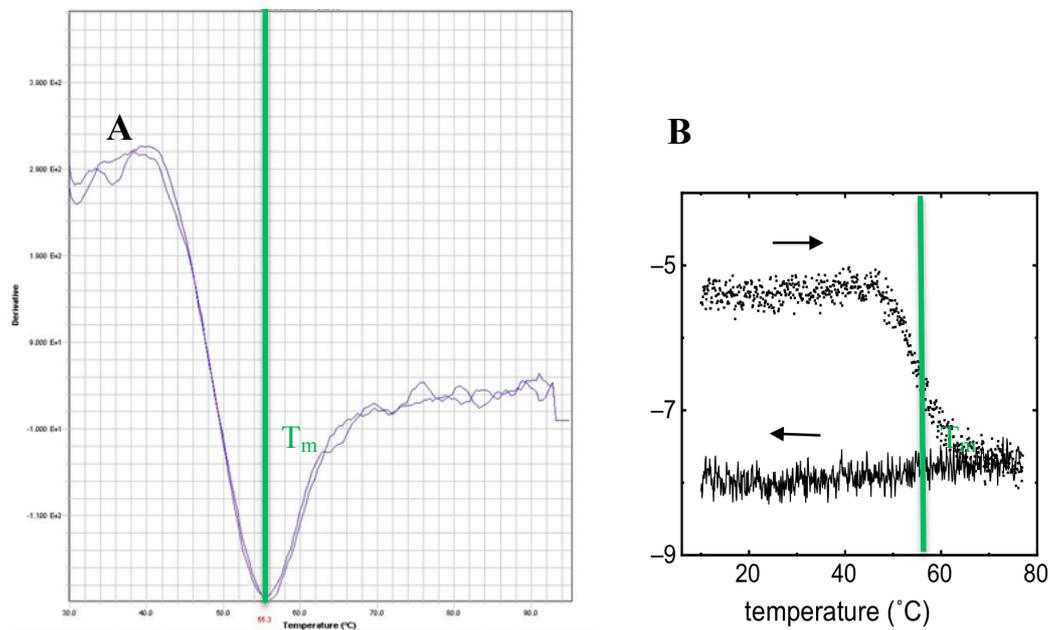


Figure 100 : A) Dérivée de la courbe de fluorescence de SYPRO orange en présence de NSI dans son tampon de conservation. Elle présente un minimum à $\sim 55^{\circ}\text{C}$, correspondant à la température de demi-dénaturation (T_m). B) Courbes de dénaturation de NSI mesurées par dichroïsme circulaire montrant que l'état de dénaturation est irréversible (Pretel et al. 2013).

5.2.2 Criblage de différentes conditions de tampon

Afin de trouver des meilleures conditions de stabilité de NS1, nous avons commencé par tester différents tampons, correspondant à des tampons couramment utilisés pour la purification des protéines, en utilisant le critère de température de dénaturation thermique mesurée par Thermofluor.

Dans un premier temps nous avons voulu comparer le tampon de purification final avec les tampons Tris de la gamme de tampons disponibles. La Figure 101 montre les courbes dérivées de fluorescence pour trois tampons : (1) Tris à pH 8,5, (2) tampon final (Tris pH 7,5 et 200 mM NaCl) et (3) Tris à pH 8,5 avec 500 mM de NaCl. Les températures de dénaturation correspondantes sont respectivement de 45, 50 et 55°C, ce qui montre que la présence de sel semble stabiliser NS1.

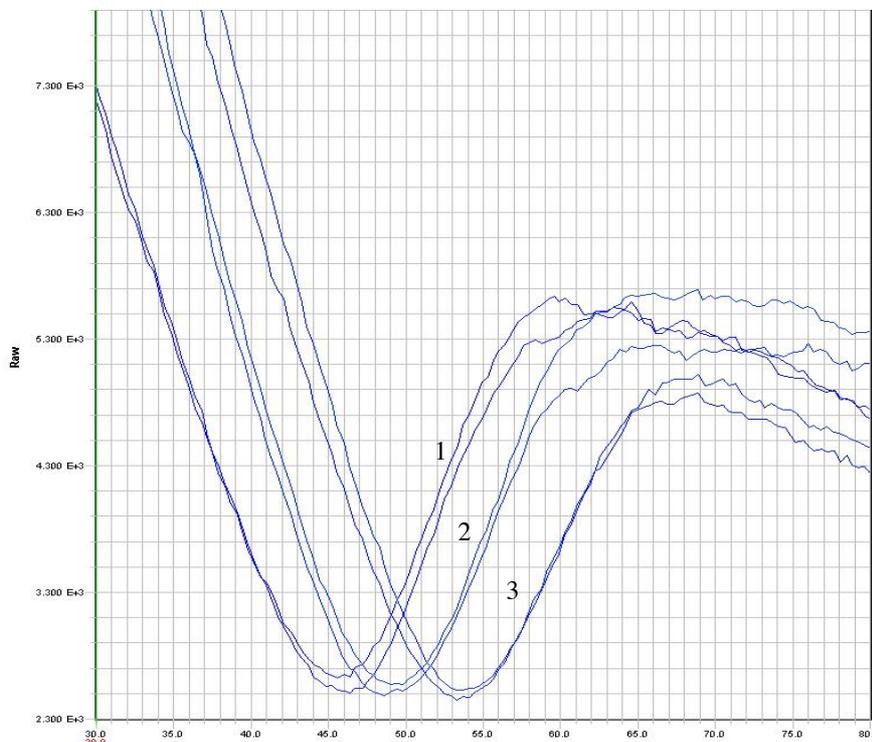


Figure 101 : Mesure de la température de dénaturation en utilisant les dérivées des courbes de fluorescence pour trois tampons de type "Tris" : (1) Tampon Tris à pH 8,5, (2) tampon final de purification de la protéine (10 mM Tris à pH 7,5 avec 200 mM NaCl, 1 mM DTT et 0,5 mM EDTA), (3) tampon Tris à pH 8,5 avec 500 mM NaCl.

Dans un deuxième temps nous avons testé tous les autres tampons disponibles à la plateforme. La Figure 102 donne le bilan des différents tampons testés. Le meilleur tampon est un tampon phosphate à pH 7,5 avec 500 mM NaCl et 10% de Glycérol.

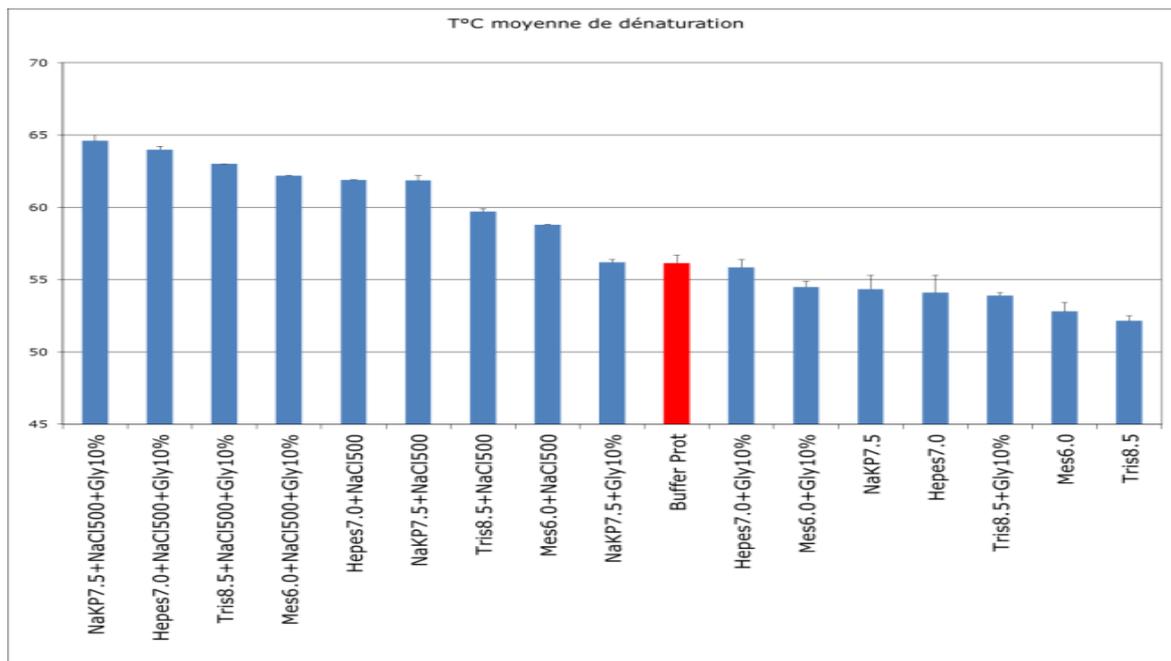


Figure 102 : Conditions de stabilité de NS1 dans différents tampons donnés en abscisse. L'axe des ordonnées correspond à la température de dénaturation mesurée par Thermofluor. Le tampon final de purification de la protéine est en rouge.

5.2.3 Affinement des conditions de stabilité de NS1

4.1.1.1 Influence de la force ionique sur la thermostabilité de NS1

Les tests avec les différents tampons montrent que les meilleures conditions de stabilité correspondent à la présence de 500 mM de sel (Figure 102). Nous avons donc testé différentes concentrations en sel. La Figure 103 montre que plus la concentration en NaCl augmente, plus la température de dénaturation augmente, signifiant que la protéine devient plus stable.

4.1.1.2 Influence du pH sur la thermostabilité de NS1

Le deuxième paramètre testé plus en détail était le pH. Comme le tampon dans lequel la protéine est la plus stable est le phosphate, nous avons fait des tests avec ce tampon à différents pH. La Figure 104 montre les courbes de fluorescence à trois pH différents (4,2; 5,5; et 8). La température de dénaturation ne varie pas beaucoup, mais le bruit de fond à basse température diminue énormément à pH acide. En rajoutant 500 mM NaCl au tampon phosphate à pH 4,4 le bruit de fond à 25°C disparaît quasiment, montrant que la combinaison de ces deux paramètres a un effet stabilisateur sur la protéine (Figure 104).

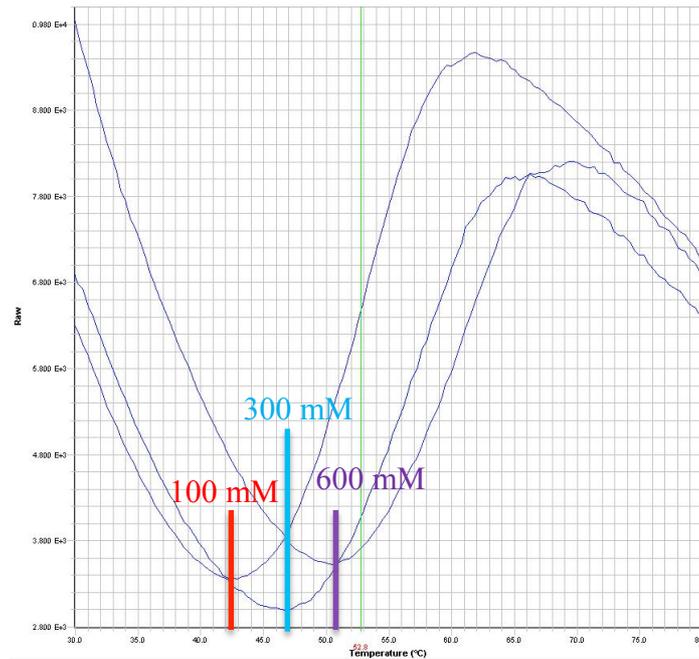


Figure 103 : Mesure de la fluorescence de SYPRO orange en présence de la protéine NSI en fonction de la concentration en NaCl (100, 300 et 600 mM) en tampon Tris pH 7,5. Les températures de dénaturation sont respectivement de 53, 58 et 61°C.

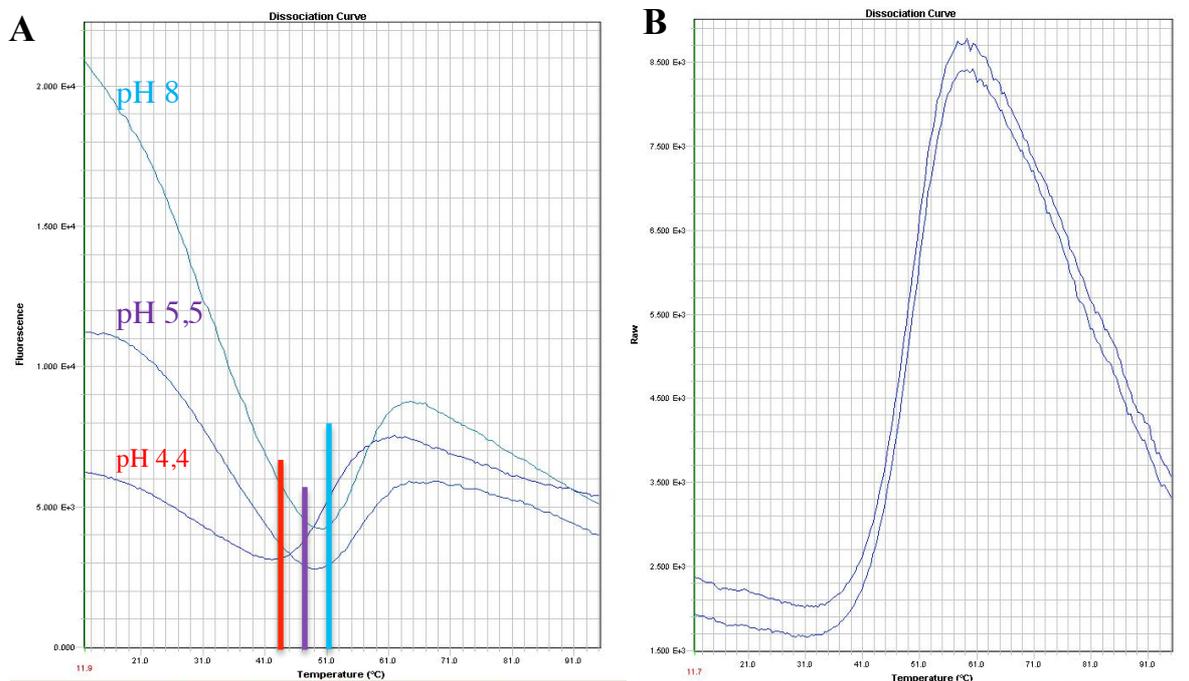


Figure 104 : A) Effet du pH sur la stabilisation de NSI. Les barres indiquent les températures de dénaturation. B) Courbes de fluorescence du SYPRO orange en présence de NSI dans un tampon phosphate pH 4,4 avec 500 mM NaCl.

5.2.4 Influence de la protéine de fusion

Nous avons trouvé le meilleur rapport signal/bruit pour les courbes de fluorescence à pH acide (Figure 104). Or il a été rapporté qu'à bas pH la protéine formait des oligomères (Swedan et al. 2011). Donc le phénomène observé semble correspondre à la formation d'oligomères qui exposent beaucoup moins de régions hydrophobes. D'après Swedan et collaborateurs ce serait la région correspondant aux résidus 21-119 qui serait impliquée (Figure 91).

J'ai remarqué que NS1 commençait à former des agrégats juste après coupure à la thrombine. Cela semble indiquer que la protéine de fusion, la MBP, participe à la stabilisation de la forme monomère de NS1. La taille de la protéine de fusion pourrait empêcher la formation d'oligomères. J'ai donc fait une dernière série de mesures de ThermoFluor pour comparer NS1 en fusion avec la MBP aux deux protéines isolées. La Figure 105 montre que le signal du bruit de fond à 25°C de la fusion MBP-NS1 est négligeable par rapport au signal de bruit de fond à 25°C de la protéine NS1 seule. Le bruit de fond est aussi absent pour la MBP seule. Ces résultats semblent indiquer que la MBP cacherait bien les régions hydrophobes exposées de la protéine NS1.

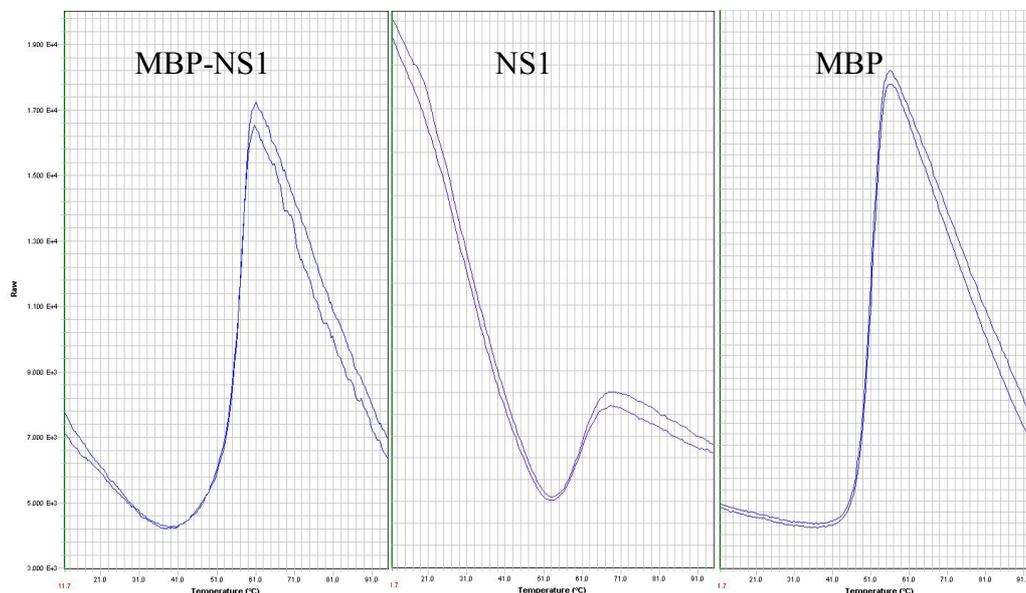


Figure 105 : Courbes de fluorescence du SYPRO orange en présence de la fusion MBP-NS1, de NS1 seule et de la MBP seule, dans le "tampon final" montrant que la protéine NS1 expose beaucoup de régions hydrophobes.

6 Conclusion

Le but de cette étude était de produire la protéine NS1 sous sa forme monomère et dans un état stable pour une étude RMN, mais j'ai constaté la formation d'agrégats juste après l'étape de coupure à la thrombine et ceci était confirmé à la fin de la purification car à la sortie de la colonne gel filtration, la concentration de la protéine était trop faible et le profil d'élution montre que nous avons perdu une partie de la protéine sous forme d'agrégats.

Lorsque j'ai essayé de concentrer la protéine, j'ai remarqué que la protéine ne supportait pas des concentrations au-delà de 40 μM . Or pour la caractérisation structurale par RMN, nous aurions idéalement besoin d'une protéine à une concentration de l'ordre de 100 μM . Nous avons alors effectué des mesures de thermostabilité par fluorescence pour trouver les meilleures conditions de stabilité de la protéine NS1. Pour résumer ces mesures, la température de dénaturation pour des concentrations croissantes de la protéine a montré que le bruit de fond à 25°C augmente énormément, ce qui peut être expliqué par le fait que la protéine NS1 expose beaucoup de régions hydrophobes. Mais il y a une forte réduction de ce bruit de fond à bas pH, ce qui peut s'expliquer par le fait qu'à pH acide, la protéine forme des oligomères et donc elle expose beaucoup moins de régions hydrophobes. Mais pour les études RMN, nous avons besoin de la protéine sous sa forme monomère et le pH utilisé doit être autour du pH physiologique pour minimiser l'échange des protons amides avec le solvant.

Le bilan de différents tampons testés montre que les meilleurs tampons, dans lesquels la protéine est la plus stable, sont le tampon phosphate à pH 7,5 et le tampon Hepes à pH 7, tous les deux avec 500 mM NaCl et 10% de glycérol. Puisque la protéine avait l'air plus stable dans ces tampons en présence de 500 mM de sel, nous avons testé l'effet de ce paramètre sur la stabilité de NS1 et nous avons constaté que le NaCl a un effet stabilisateur sur notre protéine. Mais en RMN, l'augmentation de la concentration en sel entraîne une diminution du rapport signal/bruit. Même si les conditions n'étaient pas optimales pour une étude RMN, des spectres HSQC ^1H - ^{15}N ont été acquis, montrant que la protéine est structurée.

Pour conclure, les études de thermostabilité nous ont aussi permis d'expliquer la formation d'agrégats juste après l'étape de coupure à la thrombine, puisque la MBP a l'air de participer à la stabilité de la protéine NS1 en masquant les régions hydrophobes et empêchant la formation d'oligomères. Pour des études ultérieures, je pense qu'il faut essayer de purifier la protéine dans les meilleurs tampons testés avec des concentrations élevées en sel ou bien envisager de travailler avec la protéine NS1 en fusion avec une étiquette polyhistidine qui permettra de purifier directement notre protéine sans éliminer le tag. En parallèle, on pourrait

aussi essayer peut être de faire un criblage des résidus par mutagenèse dirigée pour identifier les résidus critiques à la formation d'oligomère, et trouver un mutant stable sous la forme monomère pour une étude par RMN.

Chapitre 6 : Conclusion et perspectives

Le virus respiratoire syncytial (VRS) est le principal agent viral responsable de la bronchiolite chez les jeunes enfants. Il est aussi la cause d'infections des voies respiratoires inférieures chez des personnes immunodéprimées ou âgées. Actuellement il n'existe pas de traitement efficace, ni préventif ni curatif. L'approche par vaccination est limitée à la fois pour des raisons techniques dues à l'immaturité du système immunitaire chez des enfants en très bas âge, mais aussi des raisons sociétales liées aux réticences à la vaccination. Il existe actuellement deux médicaments antiviraux pour des patients à haut risque, la ribavirine utilisée en application curative et le palivizumab utilisé en application préventive. Ces deux médicaments présentent de nombreux inconvénients, notamment liés à leur faible efficacité. De plus la toxicité de la ribavirine est démontrée et le coût du traitement par le palivizumab très élevé. Or la recherche rationnelle d'inhibiteurs serait une solution possible pour découvrir de nouveaux médicaments plus spécifiques. Nous nous plaçons en amont de cette démarche, puisque cette approche requiert la caractérisation structurale à haute résolution de cibles virales.

Le VRS est un pneumovirus de la famille des Mononegavirales. Son complexe ARN polymérase ARN dépendant (RdRp) est responsable de la synthèse de l'ARN, messenger ou génomique, à partir de la matrice d'ARN génomique, au cours des processus de transcription et répllication, respectivement. Il représente une cible de choix, puisqu'il est spécifique du VRS et sans équivalent chez l'hôte. Il a été établi que le bon fonctionnement du complexe RdRp faisait intervenir un petit nombre de protéines encodées par le virus : la sous-unité catalytique L, la phosphoprotéine P qui est un facteur essentiel de la polymérase, la nucléoprotéine N qui encapside l'ARN génomique et deux facteurs de régulation de la transcription et de la répllication, les protéines M2-1 et M2-2. Les surfaces d'interaction entre ces protéines peuvent être ciblées par des petites molécules inhibant la formation d'un complexe fonctionnel nécessaire à l'activité polymérase.

Nous nous sommes intéressé en particulier aux deux co-facteurs de la polymérase du VRS, la protéine M2-1 et la phosphoprotéine P. Le principal outil du laboratoire est la

spectroscopie RMN qui permet d'étudier à l'échelle atomique la structure, la dynamique et les interactions de protéines. Cette méthode implique un certain nombre de contraintes, en particulier le marquage des échantillons avec des isotopes stables. Les marquages en azote 15 et carbone 13 sont effectués en routine, le marquage au deutérium dans certains cas particuliers. Les protéines analysées au laboratoire sont obtenues sous forme recombinante dans *E. coli*. Une deuxième contrainte est la taille des objets étudiés par RMN. Avec des marquages isotopiques de routine, les objets doivent être de petite taille (<25 kDa), même si l'on peut observer des complexes avec des partenaires de taille plus importante, mais non marqués.

Puisque pour la première fois, une partie de la structure de la protéine M2-1 d'un Pneumovirus a pu être déterminée par Résonance Magnétique Nucléaire. La résolution du domaine globulaire de la protéine M2-1, constitué des résidus 58 à 177 par mon équipe en collaboration avec J.-F. Eleouët (INRA, Jouy-en-Josas) a représenté la base du début de thèse (Dubosclard et al. 2011).

L'objectif premier de la thèse était de caractériser les deux autres domaines de M2-1, un domaine N-terminal de liaison au zinc et un domaine de tétramérisation, puis de déterminer l'agencement des domaines entre eux dans la protéine entière. Pour la caractérisation structurale par RMN différentes constructions de ces domaines de M2-1 ont été produites sous forme de protéines recombinantes avec des marquages isotopiques en azote 15 et carbones 13. Nous avons constaté que le domaine doigt de zinc n'était stable ni seul ni en fusion avec le domaine d'oligomérisation. Nous avons donc dû remettre en question notre approche de détermination structurale par de petits fragments de protéine. D'un autre côté nous avons pu observer par RMN que les domaines globulaires n'étaient pas associés entre eux en présence du domaine d'oligomérisation.

En parallèle de ça nous avons aussi effectué des études d'interaction de ce domaine avec la phosphoprotéine du VRS et différents ARN. Ces expériences par RMN ont permis de mettre en évidence un recouvrement partiel des surfaces d'interaction de la protéine P et de l'ARN sur la protéine M2-1 impliquant trois hélices α . Ce résultat permet d'expliquer et de confirmer la compétition observée *in vitro* entre la protéine P et l'ARN. Huit résidus de la protéine M2-1, identifiés par mutagenèse dirigée, s'avèrent cruciaux pour l'interaction soit avec la protéine P soit avec l'ARN. Un défaut d'interaction avec l'un ou l'autre des partenaires provoque une diminution de l'activité du complexe polymérase, testée avec le miniréplicon. De plus, l'affinité plus importante de la protéine M2-1 pour les séquences d'ARN de sens positif laisse supposer une interaction de la protéine M2-1 en faveur des

ARNm viraux. L'ensemble de ces résultats nous a permis de proposer un modèle quant au rôle clé de la protéine M2-1 lors du cycle viral. (Blondot et al, PLoS Pathogens 2012).

En décembre 2013 nous avons appris que la structure tétramérique de M2-1 venait d'être résolue par une équipe concurrente à Leeds (Tanner et al, PNAS 2014) : cette structure est très compacte et montre que les différents domaines de M2-1 sont en contact les uns avec les autres, en particulier le domaine N-terminal et le domaine globulaire. Nous avons alors préféré mettre l'accent sur l'analyse de la structure de la phosphoprotéine du VRS.

Si la méthode par fragments n'a pas fonctionné pour M2-1, elle a en revanche permis de mieux comprendre la structure de la phosphoprotéine du VRS. Cette structure centrée sur un court domaine de tétramérisation hélical est très dynamique par ailleurs. Ceci a été révélé en analysant méthodiquement par RMN plusieurs constructions de P, correspondant à des mutants de délétion des domaines centraux, N- et C-terminaux. Bien que mobiles, ces régions présentent des structures plus ou moins transitoires. Du côté C-terminal les hélices du domaine d'oligomérisation sont prolongées par des hélices de la même longueur qui ne sont pas arrangées sous forme de faisceau. Cette région n'a pour l'instant pas été formellement identifiée comme une région d'interaction, mais des résultats préliminaires semblent indiquer qu'il s'agit de la région d'interaction avec L. On peut imaginer que les hélices C-terminales permettent de rigidifier partiellement la structure de P. En même temps P conserve une certaine plasticité et les domaines C-terminaux correspondant à des protomères différents restent indépendants de manière à interagir avec différentes nucléoprotéines de la nucléocapside, permettant ainsi la progression de la polymérase.

Du côté N-terminal nous avons identifié deux régions avec des structures hélicales très transitoires. Ces régions correspondent 1) à la région putative de liaison de la nucléoprotéine sous sa forme N_0 non liée à l'ARN et 2) à la région de liaison putative avec M2-1. Les structures transitoires pourraient donc intervenir dans la reconnaissance de ces deux partenaires et être stabilisées à l'intérieur des complexes. Pour l'instant aucune structure à haute résolution de ces complexes n'a été obtenue. En revanche nous avons pu identifier une partie de la zone d'interaction de P sur N_0 en utilisant uniquement le domaine N-terminal de N, dont la structure est déjà connue. Pour le complexe M2-1:P, nous avons établi précédemment la zone d'interaction de P sur M2-1 mais aussi la zone d'interaction de M2-1 sur P. Dans les deux cas, la caractérisation du mécanisme de reconnaissance moléculaire permettra d'assister un criblage *in silico* de banques de données de petites molécules, pour la recherche d'inhibiteurs potentiels.

Enfin, en collaboration avec les équipes des Dr. F. Rey et J.-F. Eléouët, nous avons pu obtenir la structure d'un troisième complexe correspondant à la fixation de P à la nucléocapside. Dans ce cas nous avons pu mettre à profit les données structurales du complexe pour déterminer des inhibiteurs. Ces molécules ont permis de fournir une preuve de concept pour la conception de molécules antivirales spécifiques du VRS (Ouizougun-Oubari et al, 2015).

Pour la protéine M2-1, l'utilisation de fragments de protéines s'est révélée inadéquate. Mais une structure cristallographique est maintenant disponible sur laquelle il est possible de s'appuyer pour caractériser des zones d'interaction par RMN. Les études peuvent reprendre mais en utilisant la protéine entière avec des marquages isotopiques permettant de travailler avec des protéines de taille > 25 kDa, entre autre en deutérant la protéine. Des expériences d'interaction peuvent aussi être effectuées avec la phosphoprotéine et avec des acides nucléiques pour savoir comment la tétramérisation de M2-1 influe sur les propriétés de fixation du domaine globulaire. Et un criblage in silico peut être effectué sur la surface d'interaction du domaine globulaire de M2-1 avec la phosphoprotéine du VRS afin de trouver des inhibiteurs. Plusieurs études structurales de la phosphoprotéine restent à faire pour mieux comprendre certains mécanismes mis en jeu lors de la transcription/réplication du VRS et explorer les possibilités qu'offre la phosphoprotéine P en tant que cible thérapeutique pour des inhibiteurs.

Finalement, les études préliminaires effectuées sur la protéine NS1 du VRS ouvrent des pistes à découvrir concernant cette protéine surtout qu'aucune structure tridimensionnelle des protéines NS du VRS n'a été résolue.

Bibliographie

- (1997). "Reduction of respiratory syncytial virus hospitalization among premature infants and infants with bronchopulmonary dysplasia using respiratory syncytial virus immune globulin prophylaxis. The PREVENT Study Group." *Pediatrics* **99**(1): 93-99.
- (1998). "Palivizumab, a humanized respiratory syncytial virus monoclonal antibody, reduces hospitalization from respiratory syncytial virus infection in high-risk infants. The IMpact-RSV Study Group." *Pediatrics* **102**(3 Pt 1): 531-537.
- (2006). "Diagnosis and management of bronchiolitis." *Pediatrics* **118**(4): 1774-1793.
- (2009). "From the American Academy of Pediatrics: Policy statements--Modified recommendations for use of palivizumab for prevention of respiratory syncytial virus infections." *Pediatrics* **124**(6): 1694-1701.
- Abraham, G., D. P. Rhodes and A. K. Banerjee (1975). "The 5' terminal structure of the methylated mRNA synthesized in vitro by vesicular stomatitis virus." *Cell* **5**(1): 51-58.
- Ahmadian, G., J. S. Randhawa and A. J. Easton (2000). "Expression of the ORF-2 protein of the human respiratory syncytial virus M2 gene is initiated by a ribosomal termination-dependent reinitiation mechanism." *The EMBO journal* **19**(11): 2681-2689.
- Amann, B. T., M. T. Worthington and J. M. Berg (2003). "A Cys3His zinc-binding domain from Nup475/tristetraprolin: a novel fold with a disklike structure." *Biochemistry* **42**(1): 217-221.
- Asenjo, A., L. Rodriguez and N. Villanueva (2005). "Determination of phosphorylated residues from human respiratory syncytial virus P protein that are dynamically dephosphorylated by cellular phosphatases: a possible role for serine 54." *The Journal of general virology* **86**(Pt 4): 1109-1120.
- Atreya, P. L. and S. Kulkarni (1999). "Respiratory syncytial virus strain A2 is resistant to the antiviral effects of type I interferons and human MxA." *Virology* **261**(2): 227-241.
- Babu, M. M., R. W. Kriwacki and R. V. Pappu (2012). "Structural biology. Versatility from protein disorder." *Science* **337**(6101): 1460-1461.
- Barik, S. (1993). "The structure of the 5' terminal cap of the respiratory syncytial virus mRNA." *The Journal of general virology* **74** (Pt 3): 485-490.
- Barik, S., T. McLean and L. C. Dupuy (1995). "Phosphorylation of Ser232 directly regulates the transcriptional activity of the P protein of human respiratory syncytial virus: phosphorylation of Ser237 may play an accessory role." *Virology* **213**(2): 405-412.
- Barr, J. N., S. P. Whelan and G. W. Wertz (1997). "cis-Acting signals involved in termination of vesicular stomatitis virus mRNA synthesis include the conserved AUAC and the U7 signal for polyadenylation." *Journal of virology* **71**(11): 8718-8725.
- Basler, C. F. and A. Garcia-Sastre (2002). "Viruses and the type I interferon antiviral system: induction and evasion." *International reviews of immunology* **21**(4-5): 305-337.

- Bawage, S. S., P. M. Tiwari, S. Pillai, V. Dennis and S. R. Singh (2013). "Recent advances in diagnosis, prevention, and treatment of human respiratory syncytial virus." *Advances in virology* **2013**: 595768.
- Beeler, J. A. and K. van Wyke Coelingh (1989). "Neutralization epitopes of the F glycoprotein of respiratory syncytial virus: effect of mutation upon fusion function." *Journal of virology* **63**(7): 2941-2950.
- Bermingham, A. and P. L. Collins (1999). "The M2-2 protein of human respiratory syncytial virus is a regulatory factor involved in the balance between RNA replication and transcription." *Proceedings of the National Academy of Sciences of the United States of America* **96**(20): 11259-11264.
- Blondot, M. L., V. Dubosclard, J. Fix, S. Lassoued, M. Aumont-Nicaise, F. Bontems, J. F. Eleouet and C. Sizun (2012). "Structure and functional analysis of the RNA- and viral phosphoprotein-binding domain of respiratory syncytial virus M2-1 protein." *PLoS pathogens* **8**(5): e1002734.
- Blount, R. E., Jr., J. A. Morris and R. E. Savage (1956). "Recovery of cytopathogenic agent from chimpanzees with coryza." *Proceedings of the Society for Experimental Biology and Medicine. Society for Experimental Biology and Medicine* **92**(3): 544-549.
- Boeckh, M., M. M. Berrey, R. A. Bowden, S. W. Crawford, J. Balsley and L. Corey (2001). "Phase 1 evaluation of the respiratory syncytial virus-specific monoclonal antibody palivizumab in recipients of hematopoietic stem cell transplants." *The Journal of infectious diseases* **184**(3): 350-354.
- Bonavia, A., M. Franti, E. Pusateri Keaney, K. Kuhlen, M. Seepersaud, B. Radetich, J. Shao, A. Honda, J. Dewhurst, K. Balabanis, J. Monroe, K. Wolff, C. Osborne, L. Lanieri, K. Hoffmaster, J. Amin, J. Markovits, M. Broome, E. Skuba, I. Cornella-Taracido, G. Joberty, T. Bouwmeester, L. Hamann, J. A. Tallarico, R. Tommasi, T. Compton and S. M. Bushell (2011). "Identification of broad-spectrum antiviral compounds and assessment of the druggability of their target for efficacy against respiratory syncytial virus (RSV)." *Proceedings of the National Academy of Sciences of the United States of America* **108**(17): 6739-6744.
- Bossert, B., S. Marozin and K. K. Conzelmann (2003). "Nonstructural proteins NS1 and NS2 of bovine respiratory syncytial virus block activation of interferon regulatory factor 3." *Journal of virology* **77**(16): 8661-8668.
- Breton, A. and O. Dubois (1961). "[Remote results of treatments by specific hyposensitization in bronchial asthma in children]." *La semaine des hopitaux : organe fonde par l'Association d'enseignement medical des hopitaux de Paris* **37**: 3676-3678.
- Bukreyev, A., S. S. Whitehead, B. R. Murphy and P. L. Collins (1997). "Recombinant respiratory syncytial virus from which the entire SH gene has been deleted grows efficiently in cell culture and exhibits site-specific attenuation in the respiratory tract of the mouse." *Journal of virology* **71**(12): 8973-8982.
- Cane, P. A. (2001). "Molecular epidemiology of respiratory syncytial virus." *Reviews in medical virology* **11**(2): 103-116.
- Carbonell-Estrany, X., E. A. Simoes, R. Dagan, C. B. Hall, B. Harris, M. Hultquist, E. M. Connor and G. A. Losonsky (2010). "Motavizumab for prophylaxis of respiratory syncytial virus in high-risk children: a noninferiority trial." *Pediatrics* **125**(1): e35-51.
- Cardenas, W. B., Y. M. Loo, M. Gale, Jr., A. L. Hartman, C. R. Kimberlin, L. Martinez-Sobrido, E. O. Saphire and C. F. Basler (2006). "Ebola virus VP35 protein binds

- double-stranded RNA and inhibits alpha/beta interferon production induced by RIG-I signaling." *Journal of virology* **80**(11): 5168-5178.
- Carroneu, C., F. M. Simabuco, R. E. Tamura, L. E. Farinha Arcieri and A. M. Ventura (2007). "Intracellular localization of human respiratory syncytial virus L protein." *Archives of virology* **152**(12): 2259-2263.
- Cartee, T. L., A. G. Megaw, A. G. Oomens and G. W. Wertz (2003). "Identification of a single amino acid change in the human respiratory syncytial virus L protein that affects transcriptional termination." *Journal of virology* **77**(13): 7352-7360.
- Cartee, T. L. and G. W. Wertz (2001). "Respiratory syncytial virus M2-1 protein requires phosphorylation for efficient function and binds viral RNA during infection." *Journal of virology* **75**(24): 12188-12197.
- Carter, S. D., K. C. Dent, E. Atkins, T. L. Foster, M. Verow, P. Gorny, M. Harris, J. A. Hiscox, N. A. Ranson, S. Griffin and J. N. Barr (2010). "Direct visualization of the small hydrophobic protein of human respiratory syncytial virus reveals the structural basis for membrane permeability." *FEBS letters* **584**(13): 2786-2790.
- Castagne, N., A. Barbier, J. Bernard, H. Rezaei, J. C. Huet, C. Henry, B. Da Costa and J. F. Eleouet (2004). "Biochemical characterization of the respiratory syncytial virus P-P and P-N protein complexes and localization of the P protein oligomerization domain." *The Journal of general virology* **85**(Pt 6): 1643-1653.
- Cevik, B., D. E. Holmes, E. Vrotsos, J. A. Feller, S. Smallwood and S. A. Moyer (2004). "The phosphoprotein (P) and L binding sites reside in the N-terminus of the L subunit of the measles virus RNA polymerase." *Virology* **327**(2): 297-306.
- Challa, S., A. D. Scott, O. Yuzhakov, Y. Zhou, C. L. Tiong-Yip, N. Gao, J. Thresher and Q. Yu (2015). "Mechanism of action for respiratory syncytial virus inhibitor RSV604." *Antimicrobial agents and chemotherapy* **59**(2): 1080-1087.
- Chanock, R. and L. Finberg (1957). "Recovery from infants with respiratory illness of a virus related to chimpanzee coryza agent (CCA). II. Epidemiologic aspects of infection in infants and young children." *American journal of hygiene* **66**(3): 291-300.
- Chanock, R. M. (1957). "Recovery of a new type of myxovirus from infants with croup." *Annals of the New York Academy of Sciences* **67**(8): 287-295.
- Chapman, J., E. Abbott, D. G. Alber, R. C. Baxter, S. K. Bithell, E. A. Henderson, M. C. Carter, P. Chambers, A. Chubb, G. S. Cockerill, P. L. Collins, V. C. Dowdell, S. J. Keegan, R. D. Kelsey, M. J. Lockyer, C. Luongo, P. Najarro, R. J. Pickles, M. Simmonds, D. Taylor, S. Tyms, L. J. Wilson and K. L. Powell (2007). "RSV604, a novel inhibitor of respiratory syncytial virus replication." *Antimicrobial agents and chemotherapy* **51**(9): 3346-3353.
- Chau, T. L., R. Gioia, J. S. Gatot, F. Patrascu, I. Carpentier, J. P. Chappelle, L. O'Neill, R. Beyaert, J. Piette and A. Chariot (2008). "Are the IKKs and IKK-related kinases TBK1 and IKK-epsilon similarly activated?" *Trends in biochemical sciences* **33**(4): 171-180.
- Chen, M. D., M. Vazquez, L. Buonocore and J. S. Kahn (2000). "Conservation of the respiratory syncytial virus SH gene." *The Journal of infectious diseases* **182**(4): 1228-1233.
- Cheung, M. S., M. L. Maguire, T. J. Stevens and R. W. Broadhurst (2010). "DANGLE: A Bayesian inferential method for predicting protein backbone dihedral angles and secondary structure." *J Magn Reson B* **202**(2): 223-233.

- Clore, G. M., C. Tang and J. Iwahara (2007). "Elucidating transient macromolecular interactions using paramagnetic relaxation enhancement." *Curr Opin Struct Biol* **17**(5): 603-616.
- Collins, P. L., R. M. Chanock and B. R. Murphy (2001). *Respiratory Syncytial Virus*. Philadelphia, Lippincott Williams.
- Collins, P. L., M. G. Hill, E. Camargo, H. Grosfeld, R. M. Chanock and B. R. Murphy (1995). "Production of infectious human respiratory syncytial virus from cloned cDNA confirms an essential role for the transcription elongation factor from the 5' proximal open reading frame of the M2 mRNA in gene expression and provides a capability for vaccine development." *Proceedings of the National Academy of Sciences of the United States of America* **92**(25): 11563-11567.
- Collins, P. L., M. G. Hill, J. Cristina and H. Grosfeld (1996). "Transcription elongation factor of respiratory syncytial virus, a nonsegmented negative-strand RNA virus." *Proceedings of the National Academy of Sciences of the United States of America* **93**(1): 81-85.
- Collins, P. L. and J. A. Melero (2011). "Progress in understanding and controlling respiratory syncytial virus: Still crazy after all these years." *Virus research* **162**(1-2): 80-99.
- Collins, P. L. and B. R. Murphy (2002). "Respiratory Syncytial Virus: Reverse Genetics and Vaccine Strategies." *Virology* **296**(2): 204-211.
- Collins, P. L., R. A. Olmsted and P. R. Johnson (1990). "The small hydrophobic protein of human respiratory syncytial virus: comparison between antigenic subgroups A and B." *The Journal of general virology* **71** (Pt 7): 1571-1576.
- Collins, P. L., R. A. Olmsted, M. K. Spriggs, P. R. Johnson and A. J. Buckler-White (1987). "Gene overlap and site-specific attenuation of transcription of the viral polymerase L gene of human respiratory syncytial virus." *Proceedings of the National Academy of Sciences of the United States of America* **84**(15): 5134-5138.
- Compans, R. W., D. H. Harter and P. W. Choppin (1967). "Studies on pneumonia virus of mice (PVM) in cell culture. II. Structure and morphogenesis of the virus particle." *The Journal of experimental medicine* **126**(2): 267-276.
- Costello, H. M., W. C. Ray, S. Chaiwatpongsakorn and M. E. Peeples (2012). "Targeting RSV with vaccines and small molecule drugs." *Infectious disorders drug targets* **12**(2): 110-128.
- Cowton, V. M. and R. Fearn (2005). "Evidence that the respiratory syncytial virus polymerase is recruited to nucleotides 1 to 11 at the 3' end of the nucleocapsid and can scan to access internal signals." *Journal of virology* **79**(17): 11311-11322.
- Crowe, J. E., Jr. and J. V. Williams (2003). "Immunology of viral respiratory tract infection in infancy." *Paediatric respiratory reviews* **4**(2): 112-119.
- Cuesta, I., X. Geng, A. Asenjo and N. Villanueva (2000). "Structural phosphoprotein M2-1 of the human respiratory syncytial virus is an RNA binding protein." *Journal of virology* **74**(21): 9858-9867.
- Curran, J., J. B. Marq and D. Kolakofsky (1995). "An N-terminal domain of the Sendai paramyxovirus P protein acts as a chaperone for the NP protein during the nascent chain assembly step of genome replication." *Journal of virology* **69**(2): 849-855.
- de Waal, L., U. F. Power, S. Yuksel, G. van Amerongen, T. N. Nguyen, H. G. Niesters, R. L. de Swart and A. D. Osterhaus (2004). "Evaluation of BBG2Na in infant macaques:

- specific immune responses after vaccination and RSV challenge." *Vaccine* **22**(8): 915-922.
- DeVincenzo, J., J. E. Cehelsky, R. Alvarez, S. Elbashir, J. Harborth, I. Toudjarska, L. Nechev, V. Murugaiah, A. Van Vliet, A. K. Vaishnav and R. Meyers (2008). "Evaluation of the safety, tolerability and pharmacokinetics of ALN-RSV01, a novel RNAi antiviral therapeutic directed against respiratory syncytial virus (RSV)." *Antiviral research* **77**(3): 225-231.
- Douglas, J. L. (2004). "In search of a small-molecule inhibitor for respiratory syncytial virus." *Expert review of anti-infective therapy* **2**(4): 625-639.
- Dubosclard, V., M. L. Blondot, J. F. Eleouet, F. Bontems and C. Sizun (2011). "1H, 13C, and 15N resonance assignment of the central domain of hRSV transcription antitermination factor M2-1." *Biomolecular NMR assignments* **5**(2): 237-239.
- Elliott, M. B., T. Chen, N. B. Terio, S. Y. Chong, R. Abdullah, A. Luckay, M. A. Egan, L. A. Boutilier, K. Melville, R. A. Lerch, D. Long, J. H. Eldridge, C. L. Parks, S. A. Udem and G. E. Hancock (2007). "Alphavirus replicon particles encoding the fusion or attachment glycoproteins of respiratory syncytial virus elicit protective immune responses in BALB/c mice and functional serum antibodies in rhesus macaques." *Vaccine* **25**(41): 7132-7144.
- Ellis, R. W. (1992). "Vaccine development: progression from target antigen to product." *Advances in experimental medicine and biology* **327**: 263-271.
- Ellis, R. W. (1995). "Varicella vaccine." *Vaccine* **13**(15): 1387-1388.
- Empey, K. M., R. S. Peebles, Jr. and J. K. Kolls (2010). "Pharmacologic advances in the treatment and prevention of respiratory syncytial virus." *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **50**(9): 1258-1267.
- Esperante, S. A., L. B. Chemes, I. E. Sanchez and G. de Prat-Gay (2011). "The respiratory syncytial virus transcription antiterminator M(2-1) is a highly stable, zinc binding tetramer with strong pH-dependent dissociation and a monomeric unfolding intermediate." *Biochemistry* **50**(40): 8529-8539.
- Esperante, S. A., L. B. Chemes, I. E. Sanchez and G. D. Prat-Gay (2011). "The respiratory syncytial virus transcription antiterminator M2-1 is highly stable, zinc binding tetramer with a strong pH dependent dissociation and a monomeric unfolding intermediate." *Biochemistry* **50**(40): 8529-8539.
- Esperante, S. A., G. Paris and G. de Prat-Gay (2012). "Modular unfolding and dissociation of the human respiratory syncytial virus phosphoprotein p and its interaction with the m(2-1) antiterminator: a singular tetramer-tetramer interface arrangement." *Biochemistry* **51**(41): 8100-8110.
- Fearn, R. and P. L. Collins (1999). "Model for polymerase access to the overlapped L gene of respiratory syncytial virus." *Journal of virology* **73**(1): 388-397.
- Fearn, R. and P. L. Collins (1999). "Role of the M2-1 transcription antitermination protein of respiratory syncytial virus in sequential transcription." *Journal of virology* **73**(7): 5852-5864.
- Fearn, R., P. L. Collins and M. E. Peeples (2000). "Functional analysis of the genomic and antigenomic promoters of human respiratory syncytial virus." *Journal of virology* **74**(13): 6006-6014.
- Feltes, T. F., A. K. Cabalka, H. C. Meissner, F. M. Piazza, D. A. Carlin, F. H. Top, Jr., E. M. Connor and H. M. Sondheimer (2003). "Palivizumab prophylaxis reduces hospitalization due to respiratory syncytial virus in young children with

- hemodynamically significant congenital heart disease." *The Journal of pediatrics* **143**(4): 532-540.
- Ferron, F., Z. Li, E. I. Danek, D. Luo, Y. Wong, B. Coutard, V. Lantez, R. Charrel, B. Canard, T. Walz and J. Lescar (2011). "The hexamer structure of Rift Valley fever virus nucleoprotein suggests a mechanism for its assembly into ribonucleoprotein complexes." *PLoS pathogens* **7**(5): e1002030.
- Forster, A., G. N. Maertens, P. J. Farrell and M. Bajorek (2015). "Dimerization of matrix protein is required for budding of respiratory syncytial virus." *Journal of virology* **89**(8): 4624-4635.
- Fuentes, S., K. C. Tran, P. Luthra, M. N. Teng and B. He (2007). "Function of the respiratory syncytial virus small hydrophobic protein." *Journal of virology* **81**(15): 8361-8366.
- Galloux, M., G. Gabiane, J. Sourimant, C. A. Richard, P. England, M. Moudjou, M. Aumont-Nicaise, J. Fix, M. A. Rameix-Welti and J. F. Eleouet (2015). "Identification and characterization of the binding site of the respiratory syncytial virus phosphoprotein to RNA-free nucleoprotein." *Journal of virology* **89**(7): 3484-3496.
- Galloux, M., B. Tarus, I. Blazevic, J. Fix, S. Duquerroy and J. F. Eleouet (2012). "Characterization of a viral phosphoprotein binding site on the surface of the respiratory syncytial nucleoprotein." *Journal of virology* **86**(16): 8375-8387.
- Gan, S. W., L. Ng, X. Lin, X. Gong and J. Torres (2008). "Structure and ion channel activity of the human respiratory syncytial virus (hRSV) small hydrophobic protein transmembrane domain." *Protein science : a publication of the Protein Society* **17**(5): 813-820.
- Garcia-Barreno, B., T. Delgado and J. A. Melero (1996). "Identification of protein regions involved in the interaction of human respiratory syncytial virus phosphoprotein and nucleoprotein: significance for nucleocapsid assembly and formation of cytoplasmic inclusions." *Journal of virology* **70**(2): 801-808.
- Garcia, J., B. Garcia-Barreno, A. Vivo and J. A. Melero (1993). "Cytoplasmic inclusions of respiratory syncytial virus-infected cells: formation of inclusion bodies in transfected cells that coexpress the nucleoprotein, the phosphoprotein, and the 22K protein." *Virology* **195**(1): 243-247.
- Ghildyal, R., C. Baulch-Brown, J. Mills and J. Meanger (2003). "The matrix protein of Human respiratory syncytial virus localises to the nucleus of infected cells and inhibits transcription." *Archives of virology* **148**(7): 1419-1429.
- Ghildyal, R., A. Ho, K. M. Wagstaff, M. M. Dias, C. L. Barton, P. Jans, P. Bardin and D. A. Jans (2005). "Nuclear import of the respiratory syncytial virus matrix protein is mediated by importin beta1 independent of importin alpha." *Biochemistry* **44**(38): 12887-12895.
- Ghildyal, R., J. Mills, M. Murray, N. Vardaxis and J. Meanger (2002). "Respiratory syncytial virus matrix protein associates with nucleocapsids in infected cells." *The Journal of general virology* **83**(Pt 4): 753-757.
- Gill, M. A. and R. C. Welliver (2009). "Motavizumab for the prevention of respiratory syncytial virus infection in infants." *Expert opinion on biological therapy* **9**(10): 1335-1345.
- Glezen, W. P., L. H. Taber, A. L. Frank and J. A. Kasel (1986). "Risk of primary infection and reinfection with respiratory syncytial virus." *American journal of diseases of children* **140**(6): 543-546.

- Gonzalez, I. M., R. A. Karron, M. Eichelberger, E. E. Walsh, V. W. Delagarza, R. Bennett, R. M. Chanock, B. R. Murphy, M. L. Clements-Mann and A. R. Falsey (2000). "Evaluation of the live attenuated cpts 248/404 RSV vaccine in combination with a subunit RSV vaccine (PFP-2) in healthy young and older adults." *Vaccine* **18**(17): 1763-1772.
- Gonzalez, M. E. and L. Carrasco (2003). "Viroporins." *FEBS letters* **552**(1): 28-34.
- Gould, P. S. and A. J. Easton (2007). "Coupled translation of the second open reading frame of M2 mRNA is sequence dependent and differs significantly within the subfamily Pneumovirinae." *Journal of virology* **81**(16): 8488-8496.
- Groothuis, J. R., E. A. Simoes, M. J. Levin, C. B. Hall, C. E. Long, W. J. Rodriguez, J. Arrobio, H. C. Meissner, D. R. Fulton, R. C. Welliver and et al. (1993). "Prophylactic administration of respiratory syncytial virus immune globulin to high-risk infants and young children. The Respiratory Syncytial Virus Immune Globulin Study Group." *The New England journal of medicine* **329**(21): 1524-1530.
- Grosfeld, H., M. G. Hill and P. L. Collins (1995). "RNA replication by respiratory syncytial virus (RSV) is directed by the N, P, and L proteins; transcription also occurs under these conditions but requires RSV superinfection for efficient synthesis of full-length mRNA." *Journal of virology* **69**(9): 5677-5686.
- Hall, C. B., G. A. Weinberg, M. K. Iwane, A. K. Blumkin, K. M. Edwards, M. A. Staat, P. Auinger, M. R. Griffin, K. A. Poehling, D. Erdman, C. G. Grijalva, Y. Zhu and P. Szilagyi (2009). "The burden of respiratory syncytial virus infection in young children." *The New England journal of medicine* **360**(6): 588-598.
- Hallak, L. K., S. A. Kwilas and M. E. Peeples (2007). "Interaction between respiratory syncytial virus and glycosaminoglycans, including heparan sulfate." *Methods in molecular biology* **379**: 15-34.
- Hardy, R. W., S. B. Harmon and G. W. Wertz (1999). "Diverse gene junctions of respiratory syncytial virus modulate the efficiency of transcription termination and respond differently to M2-mediated antitermination." *Journal of virology* **73**(1): 170-176.
- Hardy, R. W. and G. W. Wertz (2000). "The Cys(3)-His(1) motif of the respiratory syncytial virus M2-1 protein is essential for protein function." *Journal of virology* **74**(13): 5880-5885.
- Harkensee, C., M. Brodli, N. D. Embleton and M. McKean (2006). "Passive immunisation of preterm infants with palivizumab against RSV infection." *The Journal of infection* **52**(1): 2-8.
- Hartlieb, B., T. Muziol, W. Weissenhorn and S. Becker (2007). "Crystal structure of the C-terminal domain of Ebola virus VP30 reveals a role in transcription and nucleocapsid association." *Proceedings of the National Academy of Sciences of the United States of America* **104**(2): 624-629.
- Hastie, K. M., T. Liu, S. Li, L. B. King, N. Ngo, M. A. Zandonatti, V. L. Woods, Jr., J. C. de la Torre and E. O. Saphire (2011). "Crystal structure of the Lassa virus nucleoprotein-RNA complex reveals a gating mechanism for RNA binding." *Proceedings of the National Academy of Sciences of the United States of America* **108**(48): 19365-19370.
- Heinrich, B. S., D. K. Cureton, A. A. Rahmeh and S. P. Whelan (2010). "Protein expression redirects vesicular stomatitis virus RNA synthesis to cytoplasmic inclusions." *PLoS pathogens* **6**(6): e1000958.
- Holm, L. and P. Rosenstrom (2010). "Dali server: conservation mapping in 3D." *Nucleic acids research* **38**(Web Server issue): W545-549.

- Hruska, J. F., P. E. Morrow, S. C. Suffin and R. G. Douglas, Jr. (1982). "In vivo inhibition of respiratory syncytial virus by ribavirin." *Antimicrobial agents and chemotherapy* **21**(1): 125-130.
- Hudson, B. P., M. A. Martinez-Yamout, H. J. Dyson and P. E. Wright (2004). "Recognition of the mRNA AU-rich element by the zinc finger domain of TIS11d." *Nature structural & molecular biology* **11**(3): 257-264.
- Inaba, Y., Y. Tanaka, T. Omori and M. Matumoto (1970). "Isolation of bovine respiratory syncytial virus." *The Japanese journal of experimental medicine* **40**(6): 473-474.
- Iwahara, J., C. Tang and G. Marius Clore (2007). "Practical aspects of (1)H transverse paramagnetic relaxation enhancement measurements on macromolecules." *J Magn Reson* **184**(2): 185-195.
- Johnson, P. R., M. K. Spriggs, R. A. Olmsted and P. L. Collins (1987). "The G glycoprotein of human respiratory syncytial viruses of subgroups A and B: extensive sequence divergence between antigenically related proteins." *Proceedings of the National Academy of Sciences of the United States of America* **84**(16): 5625-5629.
- Johnson, S., C. Oliver, G. A. Prince, V. G. Hemming, D. S. Pfarr, S. C. Wang, M. Dormitzer, J. O'Grady, S. Koenig, J. K. Tamura, R. Woods, G. Bansal, D. Couchenour, E. Tsao, W. C. Hall and J. F. Young (1997). "Development of a humanized monoclonal antibody (MEDI-493) with potent in vitro and in vivo activity against respiratory syncytial virus." *The Journal of infectious diseases* **176**(5): 1215-1224.
- Karger, A., U. Schmidt and U. J. Buchholz (2001). "Recombinant bovine respiratory syncytial virus with deletions of the G or SH genes: G and F proteins bind heparin." *The Journal of general virology* **82**(Pt 3): 631-640.
- Karlin, D. and R. Belshaw (2012). "Detecting Remote Sequence Homology in Disordered Proteins: Discovery of Conserved Motifs in the N-Termini of Mononegavirales phosphoproteins." *PLoS One* **7**(3): e31719.
- Karron, R. A., D. A. Buonagurio, A. F. Georgiu, S. S. Whitehead, J. E. Adamus, M. L. Clements-Mann, D. O. Harris, V. B. Randolph, S. A. Udem, B. R. Murphy and M. S. Sidhu (1997). "Respiratory syncytial virus (RSV) SH and G proteins are not essential for viral replication in vitro: clinical evaluation and molecular characterization of a cold-passaged, attenuated RSV subgroup B mutant." *Proceedings of the National Academy of Sciences of the United States of America* **94**(25): 13961-13966.
- Karron, R. A., P. F. Wright, R. B. Belshe, B. Thumar, R. Casey, F. Newman, F. P. Polack, V. B. Randolph, A. Deatly, J. Hackell, W. Gruber, B. R. Murphy and P. L. Collins (2005). "Identification of a recombinant live attenuated respiratory syncytial virus vaccine candidate that is highly attenuated in infants." *The Journal of infectious diseases* **191**(7): 1093-1104.
- Kawai, T., K. Takahashi, S. Sato, C. Coban, H. Kumar, H. Kato, K. J. Ishii, O. Takeuchi and S. Akira (2005). "IPS-1, an adaptor triggering RIG-I- and Mda5-mediated type I interferon induction." *Nature immunology* **6**(10): 981-988.
- Kay, L. E., M. Ikura, R. Tschudin and A. Bax (1990). "Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins." *J Magn Reson* **89**: 496-514.
- Khattar, S. K., A. S. Yunus, P. L. Collins and S. K. Samal (2001). "Deletion and substitution analysis defines regions and residues within the phosphoprotein of bovine respiratory syncytial virus that affect transcription, RNA replication, and interaction with the nucleoprotein." *Virology* **285**(2): 253-269.
- Kim, H. W., J. G. Canchola, C. D. Brandt, G. Pyles, R. M. Chanock, K. Jensen and R. H. Parrott (1969). "Respiratory syncytial virus disease in infants despite prior

- administration of antigenic inactivated vaccine." *American journal of epidemiology* **89**(4): 422-434.
- Kim, H. W., S. L. Leikin, J. Arrobio, C. D. Brandt, R. M. Chanock and R. H. Parrott (1976). "Cell-mediated immunity to respiratory syncytial virus induced by inactivated vaccine or by infection." *Pediatric research* **10**(1): 75-78.
- Kimura, K., S. Mori, K. Tomita, K. Ohno, K. Takahashi, S. Shigeta and M. Terada (2000). "Antiviral activity of NMSO3 against respiratory syncytial virus infection in vitro and in vivo." *Antiviral research* **47**(1): 41-51.
- Kingston, R. L., W. A. Baase and L. S. Gay (2004). "Characterization of nucleocapsid binding by the measles virus and mumps virus phosphoproteins." *Journal of virology* **78**(16): 8630-8640.
- Kondgen, S., H. Kuhl, P. K. N'Goran, P. D. Walsh, S. Schenk, N. Ernst, R. Biek, P. Formenty, K. Matz-Rensing, B. Schweiger, S. Junglen, H. Ellerbrok, A. Nitsche, T. Briese, W. I. Lipkin, G. Pauli, C. Boesch and F. H. Leendertz (2008). "Pandemic human viruses cause decline of endangered great apes." *Current biology : CB* **18**(4): 260-264.
- Krishnamurthy, S. and S. K. Samal (1998). "Identification of regions of bovine respiratory syncytial virus N protein required for binding to P protein and self-assembly." *The Journal of general virology* **79 (Pt 6)**: 1399-1403.
- Kuo, L., H. Grosfeld, J. Cristina, M. G. Hill and P. L. Collins (1996). "Effects of mutations in the gene-start and gene-end sequence motifs on transcription of monocistronic and dicistronic minigenomes of respiratory syncytial virus." *Journal of virology* **70**(10): 6892-6901.
- Lahaye, X., A. Vidy, C. Pomier, L. Obiang, F. Harper, Y. Gaudin and D. Blondel (2009). "Functional characterization of Negri bodies (NBs) in rabies virus-infected cells: Evidence that NBs are sites of viral transcription and replication." *Journal of virology* **83**(16): 7948-7958.
- Lambert, D. M., J. Hambor, M. Diebold and B. Galinski (1988). "Kinetics of synthesis and phosphorylation of respiratory syncytial virus polypeptides." *The Journal of general virology* **69 (Pt 2)**: 313-323.
- Leung, D. W., R. S. Shabman, M. Farahbakhsh, K. C. Prins, D. M. Borek, T. Wang, E. Muhlberger, C. F. Basler and G. K. Amarasinghe (2010). "Structural and functional characterization of Reston Ebola virus VP35 interferon inhibitory domain." *Journal of molecular biology* **399**(3): 347-357.
- Levine, S., R. Klaiber-Franco and P. R. Paradiso (1987). "Demonstration that glycoprotein G is the attachment protein of respiratory syncytial virus." *The Journal of general virology* **68 (Pt 9)**: 2521-2524.
- Leyrat, C., F. Yabukarski, N. Tarbouriech, E. A. Ribeiro, Jr., M. R. Jensen, M. Blackledge, R. W. Ruigrok and M. Jamin (2011). "Structure of the vesicular stomatitis virus N(0)-P complex." *PLoS pathogens* **7**(9): e1002248.
- Li, D., D. A. Jans, P. G. Bardin, J. Meanger, J. Mills and R. Ghildyal (2008). "Association of respiratory syncytial virus M protein with viral nucleocapsids is mediated by the M2-1 protein." *Journal of virology* **82**(17): 8863-8870.
- Li, J., A. Rahmeh, M. Morelli and S. P. Whelan (2008). "A conserved motif in region v of the large polymerase proteins of nonsegmented negative-sense RNA viruses that is essential for mRNA capping." *Journal of virology* **82**(2): 775-784.
- Lifland, A. W., J. Jung, E. Alonas, C. Zurla, J. E. Crowe, Jr. and P. J. Santangelo (2012). "Human respiratory syncytial virus nucleoprotein and inclusion bodies antagonize the innate immune response mediated by MDA5 and MAVS." *Journal of virology* **86**(15): 8245-8258.

- Lin, Y. H., A. M. Deatly, W. Chen, L. Z. Miller, R. Lerch, M. S. Sidhu, S. A. Udem and V. B. Randolph (2006). "Genetic stability determinants of temperature sensitive, live attenuated respiratory syncytial virus vaccine candidates." *Virus research* **115**(1): 9-15.
- Liuzzi, M., S. W. Mason, M. Cartier, C. Lawetz, R. S. McCollum, N. Dansereau, G. Bolger, N. Lapeyre, Y. Gaudette, L. Lagace, M. J. Massariol, F. Do, P. Whitehead, L. Lamarre, E. Scouten, J. Bordeleau, S. Landry, J. Rancourt, G. Fazal and B. Simoneau (2005). "Inhibitors of respiratory syncytial virus replication target cotranscriptional mRNA guanylation by viral RNA-dependent RNA polymerase." *Journal of virology* **79**(20): 13105-13115.
- Llorente, M. T., B. Garcia-Barreno, M. Calero, E. Camafeita, J. A. Lopez, S. Longhi, F. Ferron, P. F. Varela and J. A. Melero (2006). "Structural analysis of the human respiratory syncytial virus phosphoprotein: characterization of an alpha-helical domain involved in oligomerization." *J Gen Virol* **87**(Pt 1): 159-169.
- Llorente, M. T., B. Garcia-Barreno, M. Calero, E. Camafeita, J. A. Lopez, S. Longhi, F. Ferron, P. F. Varela and J. A. Melero (2006). "Structural analysis of the human respiratory syncytial virus phosphoprotein: characterization of an alpha-helical domain involved in oligomerization." *The Journal of general virology* **87**(Pt 1): 159-169.
- Llorente, M. T., I. A. Taylor, E. Lopez-Vinas, P. Gomez-Puertas, L. J. Calder, B. Garcia-Barreno and J. A. Melero (2008). "Structural properties of the human respiratory syncytial virus P protein: evidence for an elongated homotetrameric molecule that is the smallest orthologue within the family of paramyxovirus polymerase cofactors." *Proteins* **72**(3): 946-958.
- Marie, I., J. E. Durbin and D. E. Levy (1998). "Differential viral induction of distinct interferon-alpha genes by positive feedback through interferon regulatory factor-7." *The EMBO journal* **17**(22): 6660-6669.
- Mason, S. W., E. Aberg, C. Lawetz, R. DeLong, P. Whitehead and M. Liuzzi (2003). "Interaction between human respiratory syncytial virus (RSV) M2-1 and P proteins is required for reconstitution of M2-1-dependent RSV minigenome activity." *Journal of virology* **77**(19): 10670-10676.
- McConnochie, K. M., C. B. Hall, E. E. Walsh and K. J. Roghmann (1990). "Variation in severity of respiratory syncytial virus infections with subtype." *The Journal of pediatrics* **117**(1 Pt 1): 52-62.
- McLellan, J. S., Y. Yang, B. S. Graham and P. D. Kwong (2011). "Structure of respiratory syncytial virus fusion glycoprotein in the postfusion conformation reveals preservation of neutralizing epitopes." *Journal of virology* **85**(15): 7788-7796.
- Mejias, A. and O. Ramilo (2008). "Review of palivizumab in the prophylaxis of respiratory syncytial virus (RSV) in high-risk infants." *Biologics : targets & therapy* **2**(3): 433-439.
- Melendi, G. A., S. J. Hoffman, R. A. Karron, P. M. Irusta, F. R. Laham, A. Humbles, B. Schofield, C. H. Pan, R. Rabold, B. Thumar, A. Thumar, N. P. Gerard, W. Mitzner, S. R. Barnum, C. Gerard, S. R. Kleeberger and F. P. Polack (2007). "C5 modulates airway hyperreactivity and pulmonary eosinophilia during enhanced respiratory syncytial virus disease by decreasing C3a receptor expression." *Journal of virology* **81**(2): 991-999.
- Melero, J. A., B. Garcia-Barreno, I. Martinez, C. R. Pringle and P. A. Cane (1997). "Antigenic structure, evolution and immunobiology of human respiratory

- syncytial virus attachment (G) protein." *The Journal of general virology* **78 (Pt 10)**: 2411-2418.
- Meyer, G., M. Deplanche and F. Schelcher (2008). "Human and bovine respiratory syncytial virus vaccine research and development." *Comparative immunology, microbiology and infectious diseases* **31(2-3)**: 191-225.
- Mink, M. A., D. S. Stec and P. L. Collins (1991). "Nucleotide sequences of the 3' leader and 5' trailer regions of human respiratory syncytial virus genomic RNA." *Virology* **185(2)**: 615-624.
- Modrof, J., S. Becker and E. Muhlberger (2003). "Ebola virus transcription activator VP30 is a zinc-binding protein." *Journal of virology* **77(5)**: 3334-3338.
- Moghaddam, A., W. Olszewska, B. Wang, J. S. Tregoning, R. Helson, Q. J. Sattentau and P. J. Openshaw (2006). "A potential molecular mechanism for hypersensitivity caused by formalin-inactivated vaccines." *Nature medicine* **12(8)**: 905-907.
- Money, V. A., H. K. McPhee, J. A. Mosely, J. M. Sanderson and R. P. Yeo (2009). "Surface features of a Mononegavirales matrix protein indicate sites of membrane interaction." *Proceedings of the National Academy of Sciences of the United States of America* **106(11)**: 4441-4446.
- Mufson, M. A., C. Orvell, B. Rafnar and E. Norrby (1985). "Two distinct subtypes of human respiratory syncytial virus." *The Journal of general virology* **66 (Pt 10)**: 2111-2124.
- Muhlberger, E., M. Weik, V. E. Volchkov, H. D. Klenk and S. Becker (1999). "Comparison of the transcription and replication strategies of marburg virus and Ebola virus by using artificial replication systems." *Journal of virology* **73(3)**: 2333-2342.
- Munday, D. C., W. Wu, N. Smith, J. Fix, S. L. Noton, M. Galloux, O. Touzelet, S. D. Armstrong, J. M. Dawson, W. Aljabr, A. J. Easton, M. A. Rameix-Welti, A. P. de Oliveira, F. M. Simabuco, A. M. Ventura, D. J. Hughes, J. N. Barr, R. Fearn, P. Digard, J. F. Eleouet and J. A. Hiscox (2015). "Interactome analysis of the human respiratory syncytial virus RNA polymerase complex identifies protein chaperones as important cofactors that promote L-protein stability and RNA synthesis." *Journal of virology* **89(2)**: 917-930.
- Murphy, B. R., R. A. Olmsted, P. L. Collins, R. M. Chanock and G. A. Prince (1988). "Passive transfer of respiratory syncytial virus (RSV) antiserum suppresses the immune response to the RSV fusion (F) and large (G) glycoproteins expressed by recombinant vaccinia viruses." *Journal of virology* **62(10)**: 3907-3910.
- Myong, S., S. Cui, P. V. Cornish, A. Kirchhofer, M. U. Gack, J. U. Jung, K. P. Hopfner and T. Ha (2009). "Cytosolic viral sensor RIG-I is a 5'-triphosphate-dependent translocase on double-stranded RNA." *Science* **323(5917)**: 1070-1074.
- Nair, H., D. J. Nokes, B. D. Gessner, M. Dherani, S. A. Madhi, R. J. Singleton, K. L. O'Brien, A. Roca, P. F. Wright, N. Bruce, A. Chandran, E. Theodoratou, A. Sutanto, E. R. Sedyaningsih, M. Ngama, P. K. Munywoki, C. Kartasasmita, E. A. Simoes, I. Rudan, M. W. Weber and H. Campbell (2010). "Global burden of acute lower respiratory infections due to respiratory syncytial virus in young children: a systematic review and meta-analysis." *Lancet* **375(9725)**: 1545-1555.
- Nakanishi, T., M. Miyazawa, M. Sakakura, H. Terasawa, H. Takahashi and I. Shimada (2002). "Determination of the interface of a large protein complex by transferred cross-saturation measurements." *J Mol Biol* **318(2)**: 245-249.
- Navarro, J., C. Lopez-Otin and N. Villanueva (1991). "Location of phosphorylated residues in human respiratory syncytial virus phosphoprotein." *The Journal of general virology* **72 (Pt 6)**: 1455-1459.

- Norrby, E., H. Marusyk and C. Orvell (1970). "Morphogenesis of respiratory syncytial virus in a green monkey kidney cell line (Vero)." *Journal of virology* **6**(2): 237-242.
- Noton, S. L., V. M. Cowton, C. R. Zack, D. R. McGivern and R. Fearnls (2010). "Evidence that the polymerase of respiratory syncytial virus initiates RNA replication in a nontemplated fashion." *Proceedings of the National Academy of Sciences of the United States of America* **107**(22): 10226-10231.
- Noton, S. L., L. R. Deflube, C. Z. Tremaglio and R. Fearnls (2012). "The respiratory syncytial virus polymerase has multiple RNA synthesis activities at the promoter." *PLoS pathogens* **8**(10): e1002980.
- Ogino, T. and A. K. Banerjee (2007). "Unconventional mechanism of mRNA capping by the RNA-dependent RNA polymerase of vesicular stomatitis virus." *Molecular cell* **25**(1): 85-97.
- Ogino, T., M. Kobayashi, M. Iwama and K. Mizumoto (2005). "Sendai virus RNA-dependent RNA polymerase L protein catalyzes cap methylation of virus-specific mRNA." *The Journal of biological chemistry* **280**(6): 4429-4435.
- Perez, M., B. Garcia-Barreno, J. A. Melero, L. Carrasco and R. Guinea (1997). "Membrane permeability changes induced in *Escherichia coli* by the SH protein of human respiratory syncytial virus." *Virology* **235**(2): 342-351.
- Power, U. F., T. N. Nguyen, E. Rietveld, R. L. de Swart, J. Groen, A. D. Osterhaus, R. de Groot, N. Corvaia, A. Beck, N. Bouveret-Le-Cam and J. Y. Bonnefoy (2001). "Safety and immunogenicity of a novel recombinant subunit respiratory syncytial virus vaccine (BBG2Na) in healthy young adults." *The Journal of infectious diseases* **184**(11): 1456-1460.
- Pretel, E., G. Camporeale and G. de Prat-Gay (2013). "The Non-Structural NS1 Protein Unique to Respiratory Syncytial Virus: A Two-State Folding Monomer in Quasi-Equilibrium with a Stable Spherical Oligomer." *PLoS One* **8**(9): e74338.
- Roberts, S. R., R. W. Compans and G. W. Wertz (1995). "Respiratory syncytial virus matures at the apical surfaces of polarized epithelial cells." *Journal of virology* **69**(4): 2667-2673.
- Rocchia, W., E. Alexov and B. Honig (2001). "Extending the applicability of the nonlinear Poisson-Boltzmann equation: Multiple dielectric constants and multivalent ions." *J Phys Chem B* **105**: 6507-6514.
- Rodriguez, L., I. Cuesta, A. Asenjo and N. Villanueva (2004). "Human respiratory syncytial virus matrix protein is an RNA-binding protein: binding properties, location and identity of the RNA contact residues." *The Journal of general virology* **85**(Pt 3): 709-719.
- Rodriguez, W. J., C. B. Hall, R. Welliver, E. A. Simoes, M. E. Ryan, H. Stutman, G. Johnson, R. Van Dyke, J. R. Groothuis, J. Arrobio and et al. (1994). "Efficacy and safety of aerosolized ribavirin in young children hospitalized with influenza: a double-blind, multicenter, placebo-controlled trial." *The Journal of pediatrics* **125**(1): 129-135.
- Ruigrok, R. W. and T. Crepin (2010). "Nucleoproteins of negative strand RNA viruses; RNA binding, oligomerisation and binding to polymerase co-factor." *Viruses* **2**(1): 27-32.
- Saha, S. K. and G. Cheng (2006). "TRAF3: a new regulator of type I interferons." *Cell cycle* **5**(8): 804-807.
- Sali, A. and T. L. Blundell (1993). "Comparative protein modelling by satisfaction of spatial restraints." *J Mol Biol* **234**(3): 779-815.

- Sanchez-Seco, M. P., J. Navarro, R. Martinez and N. Villanueva (1995). "C-terminal phosphorylation of human respiratory syncytial virus P protein occurs mainly at serine residue 232." *The Journal of general virology* **76 (Pt 2)**: 425-430.
- Santangelo, P., N. Nitin, L. LaConte, A. Woolums and G. Bao (2006). "Live-cell characterization and analysis of a clinical isolate of bovine respiratory syncytial virus, using molecular beacons." *Journal of virology* **80(2)**: 682-688.
- Sato, M., H. Suemori, N. Hata, M. Asagiri, K. Ogasawara, K. Nakao, T. Nakaya, M. Katsuki, S. Noguchi, N. Tanaka and T. Taniguchi (2000). "Distinct and essential roles of transcription factors IRF-3 and IRF-7 in response to viruses for IFN-alpha/beta gene induction." *Immunity* **13(4)**: 539-548.
- Schlender, J., B. Bossert, U. Buchholz and K. K. Conzelmann (2000). "Bovine respiratory syncytial virus nonstructural proteins NS1 and NS2 cooperatively antagonize alpha/beta interferon-induced antiviral response." *Journal of virology* **74(18)**: 8234-8242.
- Schmidt, A. C., D. R. Wenzke, J. M. McAuliffe, M. St Claire, W. R. Elkins, B. R. Murphy and P. L. Collins (2002). "Mucosal immunization of rhesus monkeys against respiratory syncytial virus subgroups A and B and human parainfluenza virus type 3 by using a live cDNA-derived vaccine based on a host range-attenuated bovine parainfluenza virus type 3 vector backbone." *Journal of virology* **76(3)**: 1089-1099.
- Shao, H. Y., S. L. Yu, C. Sia, Y. Chen, E. Chitra, I. H. Chen, N. Venkatesan, C. H. Leng, P. Chong and Y. H. Chow (2009). "Immunogenic properties of RSV-B1 fusion (F) protein gene-encoding recombinant adenoviruses." *Vaccine* **27(40)**: 5460-5471.
- Simabuco, F. M., J. M. Asara, M. C. Guerrero, T. A. Libermann, L. F. Zerbini and A. M. Ventura (2011). "Structural analysis of human respiratory syncytial virus p protein: identification of intrinsically disordered domains." *Brazilian journal of microbiology : [publication of the Brazilian Society for Microbiology]* **42(1)**: 340-345.
- Simabuco, F. M., J. M. Asara, M. C. Guerrero, T. A. Libermann, L. F. Zerbini and A. M. Ventura (2011). "Structural analysis of human respiratory syncytial virus P protein: identification of intrinsically disordered domains." *Brazilian Journal of Microbiology* **42**: 340-345.
- Simoes, E. A. and X. Carbonell-Estrany (2003). "Impact of severe disease caused by respiratory syncytial virus in children living in developed countries." *The Pediatric infectious disease journal* **22(2 Suppl)**: S13-18; discussion S18-20.
- Smallwood, S., C. D. Eason, J. A. Feller, S. M. Horikami and S. A. Moyer (1999). "Mutations in conserved domain II of the large (L) subunit of the Sendai virus RNA polymerase abolish RNA synthesis." *Virology* **262(2)**: 375-383.
- Spann, K. M., K. C. Tran and P. L. Collins (2005). "Effects of nonstructural proteins NS1 and NS2 of human respiratory syncytial virus on interferon regulatory factor 3, NF-kappaB, and proinflammatory cytokines." *Journal of virology* **79(9)**: 5353-5362.
- Spehner, D., R. Drillien and P. M. Howley (1997). "The assembly of the measles virus nucleoprotein into nucleocapsid-like particles is modulated by the phosphoprotein." *Virology* **232(2)**: 260-268.
- Sudo, K., Y. Miyazaki, N. Kojima, M. Kobayashi, H. Suzuki, M. Shintani and Y. Shimizu (2005). "YM-53403, a unique anti-respiratory syncytial virus agent with a novel mechanism of action." *Antiviral research* **65(2)**: 125-131.

- Sutherland, K. A., P. L. Collins and M. E. Peeples (2001). "Synergistic effects of gene-end signal mutations and the M2-1 protein on transcription termination by respiratory syncytial virus." *Virology* **288**(2): 295-307.
- Swanson, K., X. Wen, G. P. Leser, R. G. Paterson, R. A. Lamb and T. S. Jardetzky (2010). "Structure of the Newcastle disease virus F protein in the post-fusion conformation." *Virology* **402**(2): 372-379.
- Swanson, K. A., E. C. Settembre, C. A. Shaw, A. K. Dey, R. Rappuoli, C. W. Mandl, P. R. Dormitzer and A. Carfi (2011). "Structural basis for immunization with postfusion respiratory syncytial virus fusion F glycoprotein (RSV F) to elicit high neutralizing antibody titers." *Proceedings of the National Academy of Sciences of the United States of America* **108**(23): 9619-9624.
- Swedan, S., J. Andrews, T. Majumdar, A. Musiyenko and S. Barik (2011). "Multiple functional domains and complexes of the two nonstructural proteins of human respiratory syncytial virus contribute to interferon suppression and cellular location." *Journal of virology* **85**(19): 10090-10100.
- Swedan, S., J. Andrews, T. Majumdar, A. Musiyenko and S. Barik (2011). "Multiple functional domains and complexes of the two nonstructural proteins of human respiratory syncytial virus contribute to interferon suppression and cellular location." *J Virol* **85**(19): 10090-10100.
- Swedan, S., A. Musiyenko and S. Barik (2009). "Respiratory syncytial virus nonstructural proteins decrease levels of multiple members of the cellular interferon pathways." *Journal of virology* **83**(19): 9682-9693.
- Takahasi, K., M. Yoneyama, T. Nishihori, R. Hirai, H. Kumeta, R. Narita, M. Gale, Jr., F. Inagaki and T. Fujita (2008). "Nonself RNA-sensing mechanism of RIG-I helicase and activation of antiviral immune responses." *Molecular cell* **29**(4): 428-440.
- Tang, R. S., R. R. Spaete, M. W. Thompson, M. MacPhail, J. M. Guzzetta, P. C. Ryan, K. Reisinger, P. Chandler, M. Hilty, R. E. Walker, M. M. Gomez and G. A. Lososky (2008). "Development of a PIV-vectored RSV vaccine: preclinical evaluation of safety, toxicity, and enhanced disease and initial clinical testing in healthy adults." *Vaccine* **26**(50): 6373-6382.
- Tanner, S. J., A. Ariza, C. A. Richard, H. F. Kyle, R. L. Dods, M. L. Blondot, W. Wu, J. Trincao, C. H. Trinh, J. A. Hiscox, M. W. Carroll, N. J. Silman, J. F. Eleouet, T. A. Edwards and J. N. Barr (2014). "Crystal structure of the essential transcription antiterminator M2-1 protein of human respiratory syncytial virus and implications of its phosphorylation." *Proceedings of the National Academy of Sciences of the United States of America* **111**(4): 1580-1585.
- Tawar, R. G., S. Duquerroy, C. Vonrhein, P. F. Varela, L. Damier-Piolle, N. Castagne, K. MacLellan, H. Bedouelle, G. Bricogne, D. Bhella, J. F. Eleouet and F. A. Rey (2009). "Crystal structure of a nucleocapsid-like nucleoprotein-RNA complex of respiratory syncytial virus." *Science* **326**(5957): 1279-1283.
- Tayyari, F., D. Marchant, T. J. Moraes, W. Duan, P. Mastrangelo and R. G. Hegele (2011). "Identification of nucleolin as a cellular receptor for human respiratory syncytial virus." *Nature medicine* **17**(9): 1132-1135.
- Teng, M. N. and P. L. Collins (1999). "Altered growth characteristics of recombinant respiratory syncytial viruses which do not produce NS2 protein." *Journal of virology* **73**(1): 466-473.
- Teng, M. N., S. S. Whitehead, A. Bermingham, M. St Claire, W. R. Elkins, B. R. Murphy and P. L. Collins (2000). "Recombinant respiratory syncytial virus that does not

- express the NS1 or M2-2 protein is highly attenuated and immunogenic in chimpanzees." *Journal of virology* **74**(19): 9317-9321.
- Teng, M. N., S. S. Whitehead and P. L. Collins (2001). "Contribution of the respiratory syncytial virus G glycoprotein and its secreted and membrane-bound forms to virus replication in vitro and in vivo." *Virology* **289**(2): 283-296.
- Tran, T. L., N. Castagne, D. Bhella, P. F. Varela, J. Bernard, S. Chilmonczyk, S. Berkenkamp, V. Benhamo, K. Grznarova, J. Grosclaude, C. Nespoulos, F. A. Rey and J. F. Eleouet (2007). "The nine C-terminal amino acids of the respiratory syncytial virus protein P are necessary and sufficient for binding to ribonucleoprotein complexes in which six ribonucleotides are contacted per N protein protomer." *The Journal of general virology* **88**(Pt 1): 196-206.
- Tran, T. L., N. Castagne, V. Dubosclard, S. Noinville, E. Koch, M. Moudjou, C. Henry, J. Bernard, R. P. Yeo and J. F. Eleouet (2009). "The respiratory syncytial virus M2-1 protein forms tetramers and interacts with RNA and P in a competitive manner." *J Virol* **83**(13): 6363-6374.
- Tran, T. L., N. Castagne, V. Dubosclard, S. Noinville, E. Koch, M. Moudjou, C. Henry, J. Bernard, R. P. Yeo and J. F. Eleouet (2009). "The respiratory syncytial virus M2-1 protein forms tetramers and interacts with RNA and P in a competitive manner." *Journal of virology* **83**(13): 6363-6374.
- Tristram, D. A., W. Hicks, Jr. and R. Hard (1998). "Respiratory syncytial virus and human bronchial epithelium." *Archives of otolaryngology--head & neck surgery* **124**(7): 777-783.
- Volchkov, V. E., V. A. Volchkova, E. Muhlberger, L. V. Kolesnikova, M. Weik, O. Dolnik and H. D. Klenk (2001). "Recovery of infectious Ebola virus from complementary DNA: RNA editing of the GP gene and viral cytotoxicity." *Science* **291**(5510): 1965-1969.
- Vranken, W. F., W. Boucher, T. J. Stevens, R. H. Fogh, A. Pajon, M. Llinas, E. L. Ulrich, J. L. Markley, J. Ionides and E. D. Laue (2005). "The CCPN data model for NMR spectroscopy: development of a software pipeline." *Proteins* **59**(4): 687-696.
- Weik, M., J. Modrof, H. D. Klenk, S. Becker and E. Muhlberger (2002). "Ebola virus VP30-mediated transcription is regulated by RNA secondary structure formation." *Journal of virology* **76**(17): 8532-8539.
- Wertz, G. W., P. L. Collins, Y. Huang, C. Gruber, S. Levine and L. A. Ball (1985). "Nucleotide sequence of the G protein gene of human respiratory syncytial virus reveals an unusual type of viral membrane protein." *Proceedings of the National Academy of Sciences of the United States of America* **82**(12): 4075-4079.
- Whitehead, S. S., A. Bukreyev, M. N. Teng, C. Y. Firestone, M. St Claire, W. R. Elkins, P. L. Collins and B. R. Murphy (1999). "Recombinant respiratory syncytial virus bearing a deletion of either the NS2 or SH gene is attenuated in chimpanzees." *Journal of virology* **73**(4): 3438-3442.
- Wright, P. F., M. R. Ikizler, R. A. Gonzales, K. N. Carroll, J. E. Johnson and J. A. Werkhaven (2005). "Growth of respiratory syncytial virus in primary epithelial cells from the human respiratory tract." *Journal of virology* **79**(13): 8651-8654.
- Wright, P. F., R. A. Karron, R. B. Belshe, J. Thompson, J. E. Crowe, Jr., T. G. Boyce, L. L. Halburnt, G. W. Reed, S. S. Whitehead, E. L. Anderson, A. E. Wittek, R. Casey, M. Eichelberger, B. Thumar, V. B. Randolph, S. A. Udem, R. M. Chanock and B. R. Murphy (2000). "Evaluation of a live, cold-passaged, temperature-sensitive, respiratory syncytial virus vaccine candidate in infancy." *The Journal of infectious diseases* **182**(5): 1331-1342.

- Wright, P. F., R. A. Karron, S. A. Madhi, J. J. Treanor, J. C. King, A. O'Shea, M. R. Ikizler, Y. Zhu, P. L. Collins, C. Cutland, V. B. Randolph, A. M. Deatly, J. G. Hackell, W. C. Gruber and B. R. Murphy (2006). "The interferon antagonist NS2 protein of respiratory syncytial virus is an important virulence determinant for humans." *The Journal of infectious diseases* **193**(4): 573-581.
- Wu, H., D. S. Pfarr, S. Johnson, Y. A. Brewah, R. M. Woods, N. K. Patel, W. I. White, J. F. Young and P. A. Kiener (2007). "Development of motavizumab, an ultra-potent antibody for the prevention of respiratory syncytial virus infection in the upper and lower respiratory tract." *Journal of molecular biology* **368**(3): 652-665.
- Ye, Q., R. M. Krug and Y. J. Tao (2006). "The mechanism by which influenza A virus nucleoprotein forms oligomers and binds RNA." *Nature* **444**(7122): 1078-1082.
- Yin, H. S., R. G. Paterson, X. Wen, R. A. Lamb and T. S. Jardetzky (2005). "Structure of the uncleaved ectodomain of the paramyxovirus (hPIV3) fusion protein." *Proceedings of the National Academy of Sciences of the United States of America* **102**(26): 9288-9293.
- Yoneyama, M., M. Kikuchi, K. Matsumoto, T. Imaizumi, M. Miyagishi, K. Taira, E. Foy, Y. M. Loo, M. Gale, Jr., S. Akira, S. Yonehara, A. Kato and T. Fujita (2005). "Shared and unique functions of the DExD/H-box helicases RIG-I, MDA5, and LGP2 in antiviral innate immunity." *Journal of immunology* **175**(5): 2851-2858.
- Yoneyama, M., M. Kikuchi, T. Natsukawa, N. Shinobu, T. Imaizumi, M. Miyagishi, K. Taira, S. Akira and T. Fujita (2004). "The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses." *Nature immunology* **5**(7): 730-737.
- Yunus, A. S., P. L. Collins and S. K. Samal (1998). "Sequence analysis of a functional polymerase (L) gene of bovine respiratory syncytial virus: determination of minimal trans-acting requirements for RNA replication." *The Journal of general virology* **79** (Pt 9): 2231-2238.
- Zamora, M. and S. K. Samal (1992). "Gene junction sequences of bovine respiratory syncytial virus." *Virus research* **24**(1): 115-121.
- Zamora, M. R., M. Budev, M. Rolfe, J. Gottlieb, A. Humar, J. Devincenzo, A. Vaishnav, J. Cehelsky, G. Albert, S. Nochur, J. A. Gollob and A. R. Glanville (2011). "RNA interference therapy in lung transplant patients infected with respiratory syncytial virus." *American journal of respiratory and critical care medicine* **183**(4): 531-538.
- Zhang, L., M. E. Peeples, R. C. Boucher, P. L. Collins and R. J. Pickles (2002). "Respiratory syncytial virus infection of human airway epithelial cells is polarized, specific to ciliated cells, and without obvious cytopathology." *Journal of virology* **76**(11): 5654-5666.
- Zhao, X., M. Singh, V. N. Malashkevich and P. S. Kim (2000). "Structural characterization of the human respiratory syncytial virus fusion protein core." *Proceedings of the National Academy of Sciences of the United States of America* **97**(26): 14172-14177.
- Zhou, H., X. Cheng and H. Jin (2003). "Identification of amino acids that are critical to the processivity function of respiratory syncytial virus M2-1 protein." *Journal of virology* **77**(9): 5046-5053.
- Zlateva, K. T., P. Lemey, E. Moes, A. M. Vandamme and M. Van Ranst (2005). "Genetic variability and molecular evolution of the human respiratory syncytial virus subgroup B attachment G protein." *Journal of virology* **79**(14): 9157-9167.

ANNEXE 1 : Matériels et méthodes

Amplification de plasmide, transformation de cellules électro-compétentes et vérification de l'insert amplifié

Amplification et extraction d'ADN

- Dans la glace : cuve d'électroporation, plasmide, cellules électro-compétentes (exemple : XL1-Blue)
- Préparer 2 uL de plasmide (entre 40 et 80 ng/uL) + 50 uL de cellules électro-compétentes et mettre le tout dans la cuve
- Préparer un eppendorf contenant 1 mL de milieu riche (LB, 2×YT)
- Mettre la cuve d'électroporation dans l'appareil (electroporator 2510 d'eppendorf) voltage 2500V et pulser (le temps doit être d'environ 5ms)
- Sortir la cuve et reprendre les 52 uL avec 500 uL de milieu riche puis transvaser le tout dans le reste du milieu riche et incuber le tube pendant 1h à 37°C
- Après 1h, étaler 20 uL sur boîte de sélection à 37°C et laisser pousser la nuit
- Sélectionner un clone de la boîte le lendemain, le mettre en culture dans 5 mL de milieu riche avec antibiotique sur la nuit pour avoir une préculture prête pour extraire le plasmide
- Pour l'extraction, suivre le protocole d'extraction Miniprep du kit QIAprep Spin
- La concentration du plasmide est mesurée sur un nanodrop ND 1000 Spectrophotometer de Labtech à 260 nm
- Pour le séquençage, envoyer 30 uL de plasmide entre 30 et 100 ng/uL

Milieus de production et tampons de purification des protéines « M2-1 » et « P »

Milieu minimum simplement 15N ou doublement marqué 15N-13C pour la culture:

3 g KH₂PO₄ (22 mM final)

6,4 g Na₂HPO₄ ou 8 g Na₂HPO₄.2H₂O ou 15,75 g Na₂HPO₄.12H₂O (45 mM final)

0,5 g NaCl

1 g 15NH₄Cl

→ Ajuster le pH à 7,4 et autoclaver le tout ou filtrer à 0,22µm

20 mL de glucose 20% ou 2-2,5 g glucose ^{13}C (pour le marquage en carbone 13)
2 mL de MgSO_4 @1M
100 μL de CaCl_2 @0,1M
1 mL d'Ampicilline @ 0,5 M
500 μl ZnCl_2 @ 0,1M (uniquement pour la partie N-terminale) de la protéine "M2-1".
→ Toutes ces solutions sont stériles et le milieu est complété avec de l'eau milliQ

Tampon de lyse :

50 mM Tris
350 mM NaCl jusqu'à 1M
1 mM β -mercaptoéthanol
0,5 % Triton X-100
10 mM MgSO_4
1 mM CaCl_2
20 μM ZnCl_2 (uniquement pour la partie N-terminale) de la protéine "M2-1".
→ Compléter avec l'eau milliQ et ajuster le pH à 7,4

Tampon de clivage thrombine :

20 mM Tris
200 mM NaCl
1 mM β -mercaptoéthanol ou 0,5 mM TCEP
100 μM ZnCl_2 (uniquement pour la partie N-terminale) de la protéine "M2-1".
→ Compléter avec l'eau milliQ et ajuster le pH à 8,4.

Tampon de dialyse pour la RMN :

20 mM Mes
200 mM NaCl
1 mM β -mercaptoéthanol
20 μM ZnCl_2 (uniquement pour la partie N-terminale) de la protéine "M2-1".
→ Compléter avec l'eau milliQ et ajuster le pH à 6,2.

Production et purification de fragments de la protéine M2-1 et de la phosphoprotéine du VRS

Jour 1 :

- Transformer des bactéries *E. coli* BL21(DE3) (Novagen) (50 µl) avec le plasmide pGEX “M2-1” (1 µl) par électroporation et incuber dans 1ml de milieu LB pendant 1 heure à 37°C.
- Sélectionner les bactéries transformées par dépôt sur des boîtes de milieu LB-agar contenant de l'antibiotique (ampicilline) et incuber les boîtes toute la journée à 37°C.
- Pré-culture : gratter une colonie sur une boîte et ensemercer 20 ml de milieu LB medium (+ 20 µl d'ampicilline) toute la nuit à 37°C avec agitation.

Jour 2 :

Culture : dans 1 litre de milieu M9 minimum déjà prêt (3 g KH₂PO₄ (22 mM final), 45 mM de Na₂HPO₄ ou Na₂HPO₄.2H₂O, 0,5 g NaCl, 1 g ¹⁵NH₄Cl et pH 7,3).

Ajouter : 2 µM MgSO₄, 10 µM CaCl₂, 1ml ampicilline, 2-2,5 g glucose ¹³C (pour le marquage en carbone 13) et 100 µM ZnCl₂ (uniquement pour la partie N-terminale).

- Inoculer 500 ml du milieu minimum M9 déjà préparé avec 15 ml de pré-culture et incuber sur agitateur à 37°C.
- Surveiller la DO à 600 nm jusqu'à saturation, DO 600 nm > 1,4.
- A saturation, diluer avec 500 ml du milieu M9 minimum et induire avec 80 mg/L d'IPTG (isopropyl β-D-1-thiogalactopyranoside), puis laisser incuber sur la nuit à 28°C.

Jour 3 :

- Collecter les cellules par centrifugation (rotor Beckman JLA9.1000, pot de 1 litre, centrifuger à 7000 rpm pendant 20 min à 4°C) et récupérer le culot.
- Cassage des cellules : reprendre le culot dans 50 ml de tampon de lyse, 1 ml d'anti-protéase et 100 mg/ml de lysozyme et laisser incuber 1 heure sur la glace.

[\(Pour M2-1 : Laisser incuber 45 min sur la glace et 15 min à température ambiante en rajoutant \(3,4 g NaCl\) pour avoir une concentration finale de 1,5 M\)](#)

Utilisation du casseur de cellules

- Rinçage du casseur avec 200-300 mL d'eau milliQ (0,8-1 bar).
- Equilibrer avec 200 mL de tampon de lyse (0,8-1 bar).

- Casser les cellules en suspension 2 fois (1,5 bar).
- Rincer avec 50 mL du tampon de lyse (1,5 bar).
- Dernier rinçage avec 30 mL du tampon de lyse (1,5 bar).
- Laver le casseur avec 500 mL d'eau milliQ (0,8-1 bar).
- Laver le casseur avec 200 mL NaOH à 0,1M (0,8-1 bar).
- Laisser incuber avec un petit volume de NaOH à 0,1M pendant 1 heure.
- Laver le casseur avec l'eau milliQ jusqu'à obtenir un pH ~7 (0,8-1 bar).
- Enfin remplir avec un peu d'éthanol et laisser jusqu'à la prochaine utilisation.

A la fin, obtenir un volume de ~150 ml maximum de solution de lyse.

- Partager le volume de la solution de lyse dans des tubes d'ultracentrifugation (rotor Beckman 70 Ti) et centrifuger à 40 000 g pendant 1 heure à 4°C.
- Purification : récupérer le surnageant, ajouter 2 ~3 ml de résine GSH-sépharose (GE Healthcare) et laisser incuber avec une petite agitation pendant 15 heures ou plus à 4°C.

Jour 4 :

- Filtrer la résine dans des tubes de filtration et laver 5 fois avec 10 ml de tampon spécial thrombine.
- Clivage avec la thrombine : ajouter à la résine 5 µl de thrombine biotinyllée (Novagen) et 3 ml de tampon de clivage de thrombine ensuite laisser le tube tourner (sur le soleil) pendant 24 heures.
-

Jour 5 :

- Filtrer la résine et récupérer la fraction non retenue (~ 3 ml), vérifier sur gel si la thrombine a bien coupé entre la protéine et la GST.
- Capturer la thrombine de la solution protéique avec la résine streptavidine.
- Mesurer les concentrations avec le test colorimétrique (Bradford) et un test spectrophotométrique (Nanodrop).
- Dialyser contre 1 litre de tampon pour l'étude par RMN (20 mM Mes pH 6,2, 200 mM NaCl, 1 mM β-mercaptoethanol et 20 µM ZnCl₂ pour la partie N-terminale).

Milieux de production et tampons de purification de la protéine « NS1 »

Milieu minimum simplement marqué 15N pour la culture:

3 g KH₂PO₄ (22 mM final)

6,4 g Na₂HPO₄ ou 8 g Na₂HPO₄.2H₂O ou 15,75 g Na₂HPO₄.12H₂O (45 mM final)

0,5 g NaCl

1 g 15NH₄Cl

→ Ajuster le pH à 7,4 et autoclaver le tout ou filtrer à 0,22µm

20 mL de glucose 20%

2 mL de MgSO₄ @1M

100 µL de CaCl₂ @0,1M

1 mL d'Ampicilline @ 0,5 M

→ Toutes ces solutions sont stériles et le milieu est complété avec de l'eau milliQ

Tampon de lyse :

20 mM Tris-HCl

200 mM NaCl

5 mM β-mercaptoéthanol

1 mM EDTA

→ Compléter avec l'eau milliQ

et ajuster le pH à 8

Tampon de dialyse :

10 mM Tris-HCl

200 mM NaCl

2 mM β-mercaptoéthanol

→ Compléter avec l'eau milliQ et

ajuster le pH à 7,5

Tampon d'élution de la fusion MBP-NS1 :

10 mM Tris-HCl

200 mM NaCl

2 mM β-mercaptoéthanol

20 mM maltose

→ Compléter avec l'eau milliQ et ajuster le pH à 7,5

Tampon final d'élution (Superdex-75 gel filtration) :

10 mM Tris-HCl

200 mM NaCl

1 mM DTT

0,5 mM EDTA

→ Compléter avec l'eau milliQ et ajuster le pH à 7,5

Production et purification de la protéine « NS1 »

Jour 1 :

- Transformer des bactéries *E. coli* BL21(DE3) (Novagen) (50 µl) avec le plasmide pGEX “M2-1” (1 µl) par électroporation et incuber dans 1ml de milieu LB pendant 1 heure à 37°C.
- Sélectionner les bactéries transformées par dépôt sur des boîtes de milieu LB-agar contenant de l'antibiotique (ampicilline) et incuber les boîtes toute la journée à 37°C.
- Pré-culture : gratter une colonie sur une boîte et ensemercer 20 ml de milieu LB medium (+ 20 µl d'ampicilline) toute la nuit à 37°C avec agitation.

Jour 2 :

Culture : dans 1 litre de milieu M9 minimum déjà prêt (3 g KH_2PO_4 (22 mM final), 45 mM de Na_2HPO_4 ou $\text{Na}_2\text{HPO}_4 \cdot 2\text{H}_2\text{O}$, 0,5 g NaCl , 1 g $^{15}\text{NH}_4\text{Cl}$ et pH 7,3).

Ajouter : 2 µM MgSO_4 , 10 µM CaCl_2 , 1ml ampicilline, 2-2,5 g glucose ^{13}C (pour le marquage en carbone 13) et 100 µM ZnCl_2 (uniquement pour la partie N-terminale).

- Inoculer 1 litre du milieu minimum M9 déjà préparé avec 15 ml de pré-culture et incuber sur agitateur à 37°C.
- Surveiller la DO à 600 nm jusqu'à DO 600 nm \approx 0,6.
- Incuber 30 min à 20°C avant induction
- Induire avec 80 mg/L d'IPTG (isopropyl β -D-1-thiogalactopyranoside), puis laisser incuber sur la nuit à 20°C.

Jour 3 :

- Collecter les cellules par centrifugation (rotor Beckman JLA9.1000, pot de 1 litre, centrifuger à 7000 rpm pendant 20 min à 4°C) et récupérer le culot.
- Cassage des cellules : reprendre le culot dans 50 ml de tampon de lyse, 1 ml d'anti-protéase. (le protocole de cassage est décrit dans le paragraphe de production et purification de fragments de protéine M2-1 et de la phosphoprotéine)
- Centrifuger la solution de lyse à 15000 g pendant 30 min à 4°C.
- Purification : La précipitation différentielle est une des premières étapes de purification = Précipiter le surnageant par l'ajout de sulfate d'ammonium* et avoir une solution saturée à 50%.

- Collecter les protéines précipitées par centrifugation douce, récupérer le culot dans le tampon de dialyse
- Dialyser le volume récupéré contre le même tampon de dialyse pour éliminer le sulfate d'ammonium.

(*:Ce sel est très soluble en solution aqueuse et permet d'atteindre des forces ioniques très élevées. Il est très hydrophile et compétitionne efficacement avec les protéines pour l'eau causant leur déshydratation. Les ions sulfates et ammonium sont relativement petits et peuvent facilement s'approcher des résidus chargés des protéines pour les neutraliser. Ce sel a aussi l'avantage de peu dénaturer les protéines et permet de maximiser l'obtention de protéines biologiquement actives)

Jour 4 :

- Purification : après dialyse, ajouter 2 ~3 ml de résine amylose (New England Biolabs, Hitchin, UK.) et laisser incuber avec une petite agitation pendant 15 heures ou plus à 4°C.

Jour 5 :

- Filtrer la résine dans des tubes de filtration et laver 5 fois avec 10 ml de tampon d'élution sans le maltose.
- Eluer la fusion MBP-NS1 accrochée sur la résine avec le tampon d'élution + 20 mM maltose et vérifier sur gel si on a bien récupéré la fusion MBP-NS1.
- Concentrer le volume récupéré sur des membranes de concentration Amicon ultra Centrifugal Filters de 10 KD jusqu'à 10 mg/mL dans 2 mL.
- Clivage avec la thrombine : ajouter au 2 mL, 100 µl de thrombine (Sigma-Aldrich, St Louis, MO, USA) et laisser le tube eppendorf tourner (sur le soleil) pendant 16 heures à 4°C.
- Vérifier sur gel si la thrombine a bien coupé entre la protéine NS1 et la MBP.
- Injecter directement le mélange sur colonne Superdex 75 gel filtration, pour séparer les constituants du mélange selon leurs poids moléculaires et éluer avec le tampon final d'élution.
- Mesurer les concentrations avec le test colorimétrique (Bradford) et un test spectrophotométrique (Nanodrop).

Synthèse et purification des acides nucléiques

Synthèse en phase solide

Les séquences d'ARN sont synthétisées en phase solide avec un synthétiseur « Pharmacia Gene Assembler Plus » en utilisant une résine avec une amorce universelle sans première base fixée en 3' et avec un programme adapté à la synthèse des ARN. La synthèse est basée sur 4 étapes :

- **Détritilation** (déprotection) : la DMT (diméthyltryptamine) est éliminée à l'aide d'un acide (tel que trichloroacétique), il en résulte donc un groupement 5' hydroxyle libre sur la première base.

- **Couplage** : Le phosphoramidite protégé est introduit et son groupement phosphate réagit avec l'hydroxyle en 5' qui vient d'être déprotégé. Le couplage s'effectue en présence de tétrazole (un acide faible servant d'activateur) pour permettre le couplage par une liaison phosphate et enfin les bases non fixées sont éliminées par lavage.

- **Capping** : environ 1% des groupements 5'OH ne réagissent pas avec une nouvelle base, il est donc nécessaire de les bloquer afin d'éviter les réactions aspécifiques et la formation d'oligonucléotides comportant des délétions et ceci se fait par l'ajout d'un groupement protecteur et il en résulte une acétylation.

- **Oxydation** : la liaison phosphite entre la première et la seconde base doit être stabilisée en rendant le groupement pentavalent (phosphate):

A la fin de la synthèse, des plusieurs étapes sont indispensables :

- **Coupage de la liaison oligonucléotide-résine** : L'oligonucléotide doit être détaché de son support solide par un traitement avec une solution 3 : 1 (Ammoniac «NH₃» / Ethanol) et incubation 60 minutes à 65°C. Mettre le tube dans un gros tube « falcon » contenant quelques ml d'ammoniac pour égaliser la pression et parafilmer le tube pour éviter les fuites de NH₃. Attention : laisser refroidir le tube avant de l'ouvrir car NH₃ est expulsé sous pression.

- **Déprotection des groupes NH₂ protégés** : après filtration de la résine sur laine de verre dans une pipette pasteur, on lave la résine avec 10-15 ml d'un mélange ternaire 9 : 3 : 3 (Ethanol/H₂O/ Acétonitrile) car l'ARN reste accrocher à la résine. Lyophiliser à sec (speed-vac)

- **Retrait du groupe Silyle** : après lyophilisation à sec, on redissout toutes les fractions réunies dans un seul tube NUNC, ce sont des tubes particuliers qui peuvent résister à ces réactifs corrosifs, dans environ 1,5 ml d'un mélange ternaire 1,5 : 0,75 : 1 (N-méthyl-pyrrolidinone / triéthylamine / solution TEA ; 3HF) et on incube 90 minutes à 65°C au bain mari sec,. Ensuite on dilue la solution environ 2 fois dans l'eau milli Q et on dialyse contre de l'eau toute la nuit. Congeler ce qui a été récupéré dans l'azote liquide et lyophiliser à sec (speed-vac).

Purification

Une fois synthétisés, les produits ont été purifiés par chromatographie anionique (Akta, GE Healthcare) sur une colonne DEAE-Sepharose (Pharmacia) dans 5 mM de tampon phosphate pH = 6,2. L'élution est réalisée par un gradient linéaire de 0 à 1 M de NaCl en 80 minutes.

Dissoudre l'ARN dans 0,5-1 ml de tampon phosphate 10 mM et chauffer 15 s à 80°C. Il est très recommandé de faire la purification à haute température (50-65°C) pour une meilleure résolution car à température ambiante la chromatographie peut être résolutive sur colonne analytique et mal résolue en préparatif (risque d'agrégation à haute concentration). On ne rencontre pas ce problème avec l'ADN car il est purifié dans un tampon dénaturant qui contient du NaOH.

La concentration est déterminée avec la valeur de l'absorbance à 260 nm. Les quantités obtenues sont de l'ordre de 30 DO 260 ce qui équivaut à 0,2 µmoles. Les aliquots obtenus sont lyophilisés et stockés à -20°C.

Mutagenèse dirigée

L'ADN codant pour la protéine d'intérêt (la phosphoprotéine P) est amplifié par PCR avec de la polymérase haute fidélité Pfu Turbo (5U, Stratagene®) à partir de 10 ng d'ADN matrice et 100 ng d'amorces. L'ADN est ensuite digéré par les enzymes de restrictions adéquats ainsi que le vecteur. La réaction de ligation est effectuée avec l'enzyme ADN T4 ligase (MBI Fermentas®) selon les recommandations du fabricant.

Fixation d'une étiquette paramagnétique

ANNEXE 2 : Conditions de tampons utilisées pour solubiliser M2-1Δ[35-58]

	Conditions de tampon	Concentration	pH	Autres paramètres
M2-1 Δ 1 ^{ère} production	20mM MES, 150mM NaCl, 20μM ZnCl ₂ , 0,5mM TCEP	100μM	6,2	298K
	20mM MES, 150mM NaCl, 20μM ZnCl ₂ , 0,5mM TCEP		6,2	298K
	1 mM Mg ²⁺ , 20 mM TRIS, 2mM Beta OH, 20mM NaCl, 20μM ZnCl ₂		8,5	Ajouter NaCl
	20mM MES, 150mM NaCl, 2mM Beta OH, 20μM ZnCl ₂	40 μM	6,2	
	20mM MES, 150mM NaCl, 2mM Beta OH, 20μM ZnCl ₂	40μM	6,2	303K, 298K, 283K respectivement
	20mM MES, 150mM NaCl, 2mM Beta OH, 20μM ZnCl ₂	40 μM	6,2	298K compléter à 0,5M NaCl
	20 mM MES, 0,5M NaCl, 2mM Beta OH, 20μM ZnCl ₂	40 μM	6,2	POLY A @ 150- 200 μM
	20mM MES, 150mM NaCl, 1mM Beta OH, 20μM ZnCl ₂	40μM	6,2	Diminuer le sel

2 ^{ème} Production	20mM MES , 200mM NaCl , 1mM Beta OH , 20µM ZnCl₂	90 µM	6,2	298K
	20mM MES , 200mM NaCl , 1mM Beta OH , 20µM ZnCl₂	90 µM	6,2	303K
	20mM MES , 200mM NaCl , 1mM Beta OH , 20µM ZnCl₂	135 µM	6,2	303K
	20mM MES , 200mM NaCl , 1mM Beta OH , 20µM ZnCl₂	90 µM	6,2	303K + 0,5M Arg HCl
3 ^{ème} Production	20mM BisTris , 200mM NaCl , 1mM Beta OH , 10µM ZnCl₂	100 µM	6,2	293K + 0,5M Arg et 0,5M Glu
	20mM BisTris , 200mM NaCl , 1mM Beta OH , 10µM ZnCl₂	100 µM	6,2	303K + 0,5M Arg et 0,5M Glu
	20mM BisTris , 200mM NaCl , 2mM TCEP , 10µM ZnCl₂	400 µM	6,2	298K + 0,5M Arg et 0,5M Glu
	20mM BisTris , 200mM NaCl , 2mM TCEP , 10µM ZnCl₂	400 µM	6,2	303K + 0,5M Arg et 0,5M Glu
	20mM BisTris , 200mM NaCl , 2mM TCEP , 10µM ZnCl₂	400 µM	6,2	303K + 0,5M Arg et 0,5M Glu

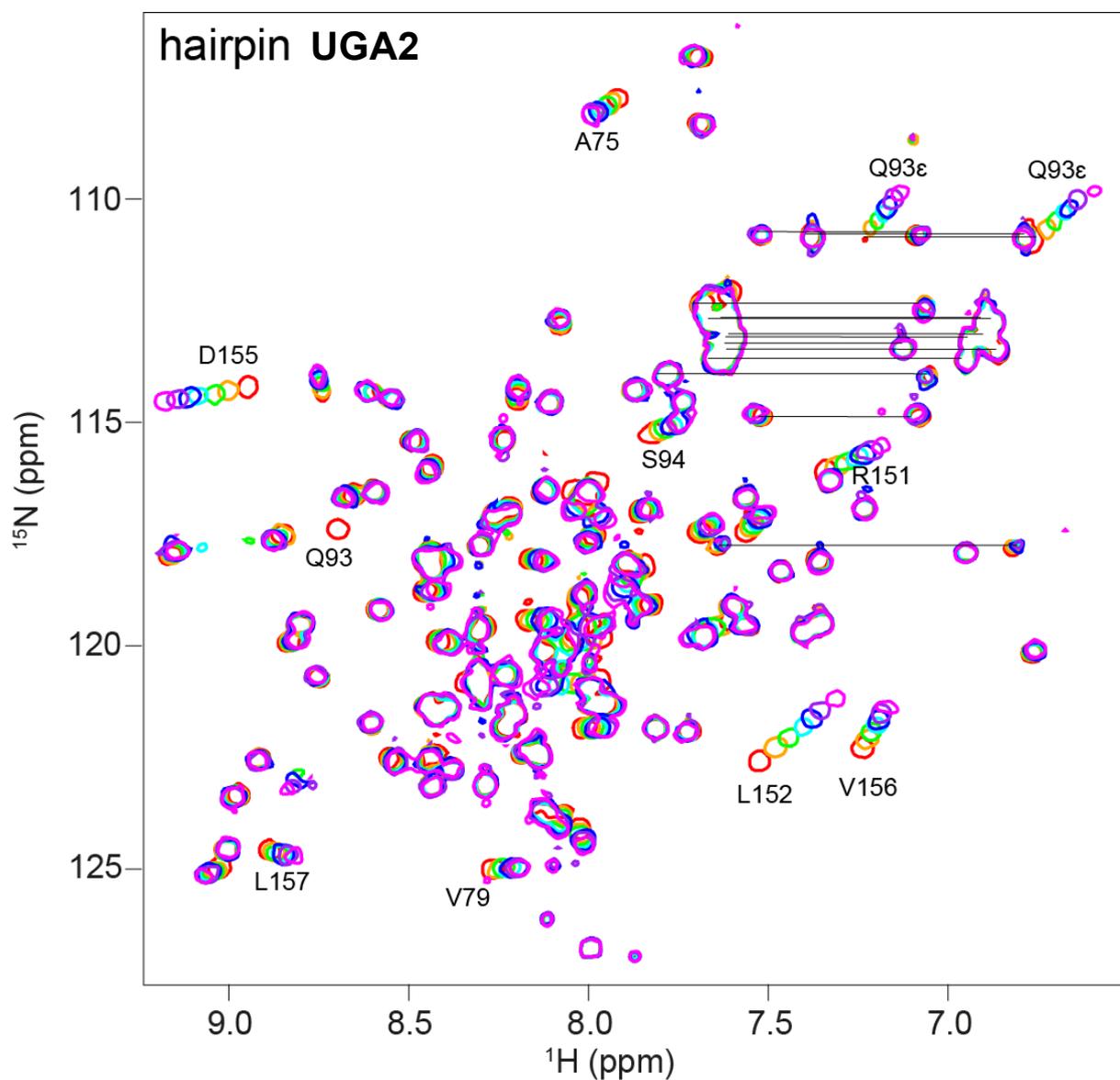
ANNEXE 3 : Conditions de tampons utilisées pour solubiliser le peptide synthétique ZNF29

Conditions de tampon	concentration	pH	autres paramètres
1) H ₂ O, D ₂ O, ZnCl ₂ (2*[P])	~ 350µM	1--->9 dans les deux sens	
2) H ₂ O, D ₂ O, 2mM DTT	~ 200 µM	3,6 à 6,4 sans ZnCl ₂ 5,5 à 6,5 Titration ZnCl ₂	TFE, glycérol
3) H ₂ O, D ₂ O, 2mM DTT, ZnCl ₂ ??	~ 350 µM	4,22 → 7,8	
4) H ₂ O, D ₂ O, 2mM DTT, 300mM NaCl ZnCl ₂ ??	~ 200 µM	3,90 → 8	
5) H ₂ O, D ₂ O, 2mM DTT, 300mM NaCl	~ 200 µM	4,4	cadmium
6) Tampon Hepes/ NaCl, 2,5 mM TCEP	25 µM → 140 µM	Entre 6 et 6,5	1 mM cadmium

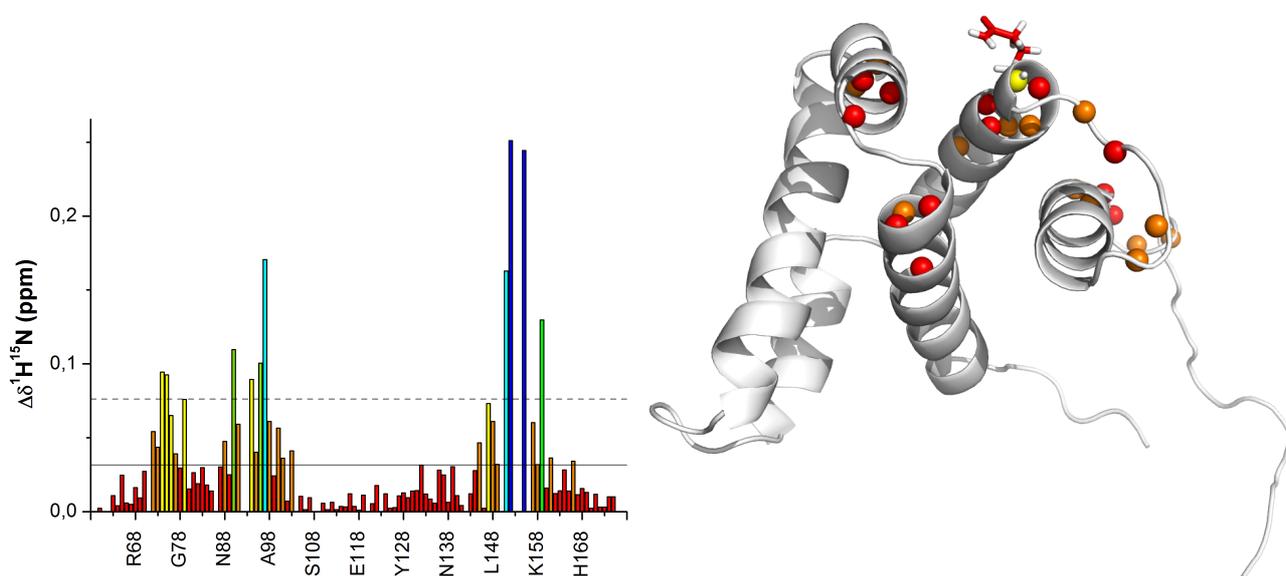
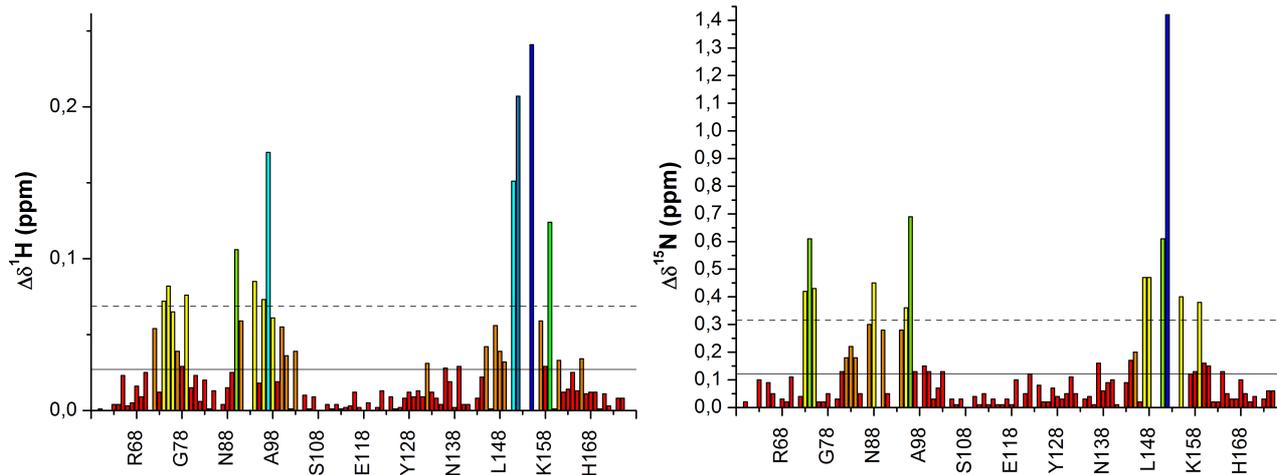
ANNEXE 4 : Oligonucléotides utilisés pour les mutations en cystéine de la phosphoprotéine du VRS pour la fixation de sondes paramagnétiques

Mutation	Séquences des amorces (5' --> 3')
S23C : TCA → TGC C68G-A69C sens C68G-A69C anti-sens	5' -CAGGGCTACTAAATTCCTAGAATGCATAAAGGGCAAATTCACATCACC-3' 5' -GGTGATGTGAATTTGCCCTTTATGCATTCTAGGAATTTAGTAGCCCTG-3'
S99C : TCA → TGC C296G-A297C sens C296G-A297C anti-sens	5' -CCTATACCAAGTGATAATCCCTTTTGCAAACCTATACAAAGAAACCATAGAG-3' 5' -CTCTATGGTTTCTTTGTATAGTTTGCAAAGGGATTATCACTTGGTATAGG-3'
S143C : AGT → TGT A427T sens A427T anti-sens	5' -CTGCAAGATTAGATAGGATTGATGAAAAATTATGTGAAATACTAGGAATG-3' 5' -CATTCCCTAGTATTTTACATAATTTTTTCATCAATCCTATCTAATCTTGCAG-3'
S156C : AGT → TGT A466T sens A466T anti-sens	5' -CACACATTAGTAGTAGCATGTGCAGGACCTACATCTG-3' 5' -CAGATGTAGGTCCTGCACATGCTACTACTAATGTGTG-3'
E179C : GAA → TGC G535T-A536G-A537C sens G535T-A536G-A537C anti-sens	5' -GGTTGGTTTAAAGAGAAGAAATGATATGCAAATCAGAAGCA-3' 5' - TGCTTCAGTTCTGATTTTGCATATCATTCTTCTCTTAAACCAACC -3'
E193C : GAA → TGC G577T-A578G-A579C sens G577T-A578G-A579C anti-sens	5' -GAAGCATTAAATGACCAATGACAGATTATGCGCTATGGCAAGACTCAGGAATGA-3' 5' -TCATTCCCTGAGTCTTGCCATAGCGCATAATCTGTCATTGGTCATTAATGCTTC-3'
S237C : TCA → TGC C710G-A711C sens C710G-A711C anti-sens	5' -AAGGGAATGATAGTGACAATGATCTATGCCTTGAAGATTTCTGATTAGTTAC-3' 5' -GTAACATAATCAGAAATCTTCAAGGCATAGATCATTGTCACTATCATTCCCTT-3'

ANNEXE 5 : Titrages de M2-1[58-177] par de courtes séquences d'ARN

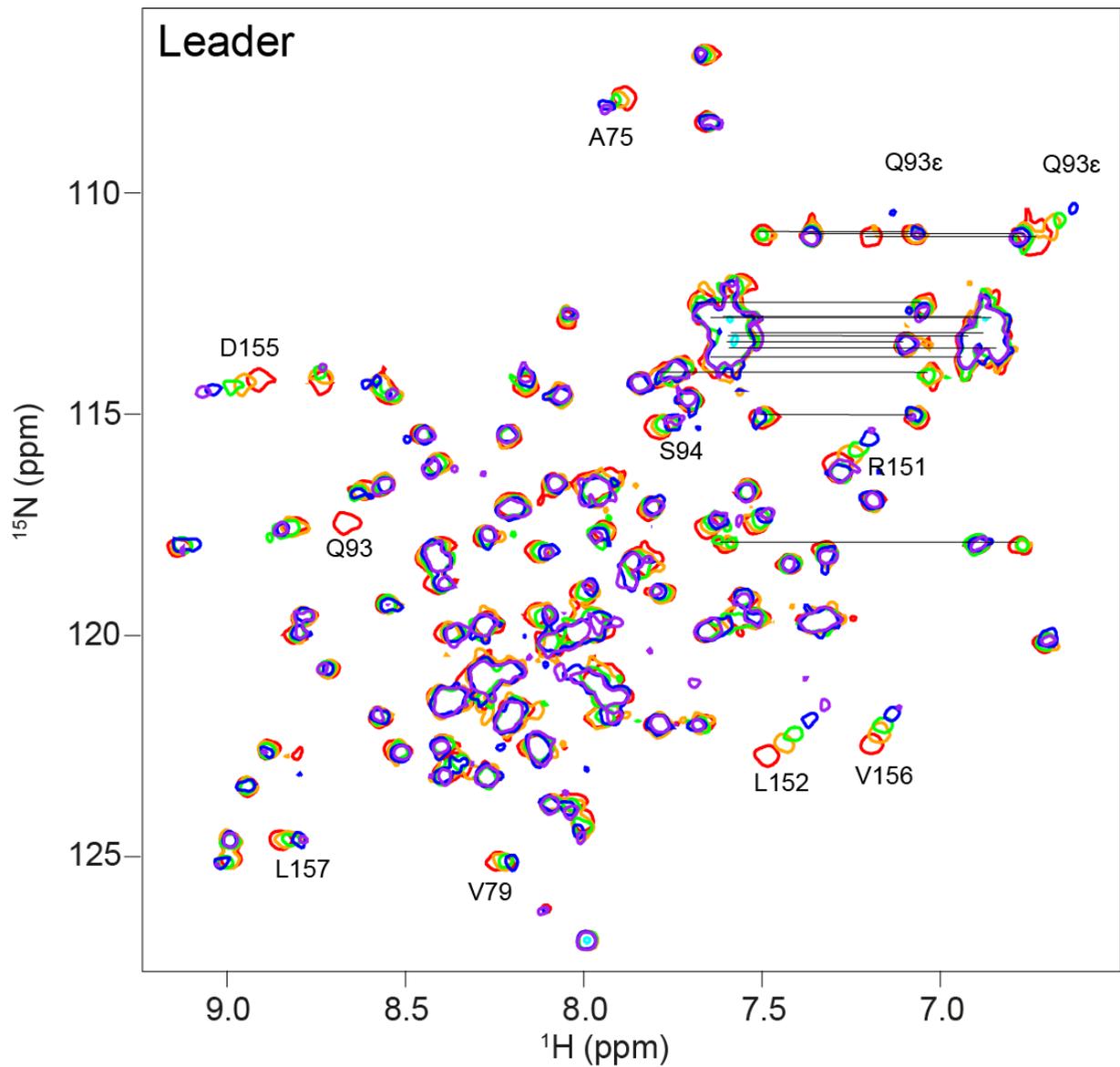


Superposition de spectres ^1H - ^{15}N -HSQC de M2-1[58-177] en présence de différentes concentrations de l'ARN épingle à cheveux « UGA2» (50 μM , $T = 298\text{K}$, $B_0 = 14,1\text{ T}$, 600 MHz). Spectre de référence (rouge), en présence de 0,5 équivalent (orange), 1 équivalent (vert), 2 équivalents (bleu clair), 3 équivalents (bleu), 6 équivalents (Violet pourpre) et 20 équivalents (magenta).

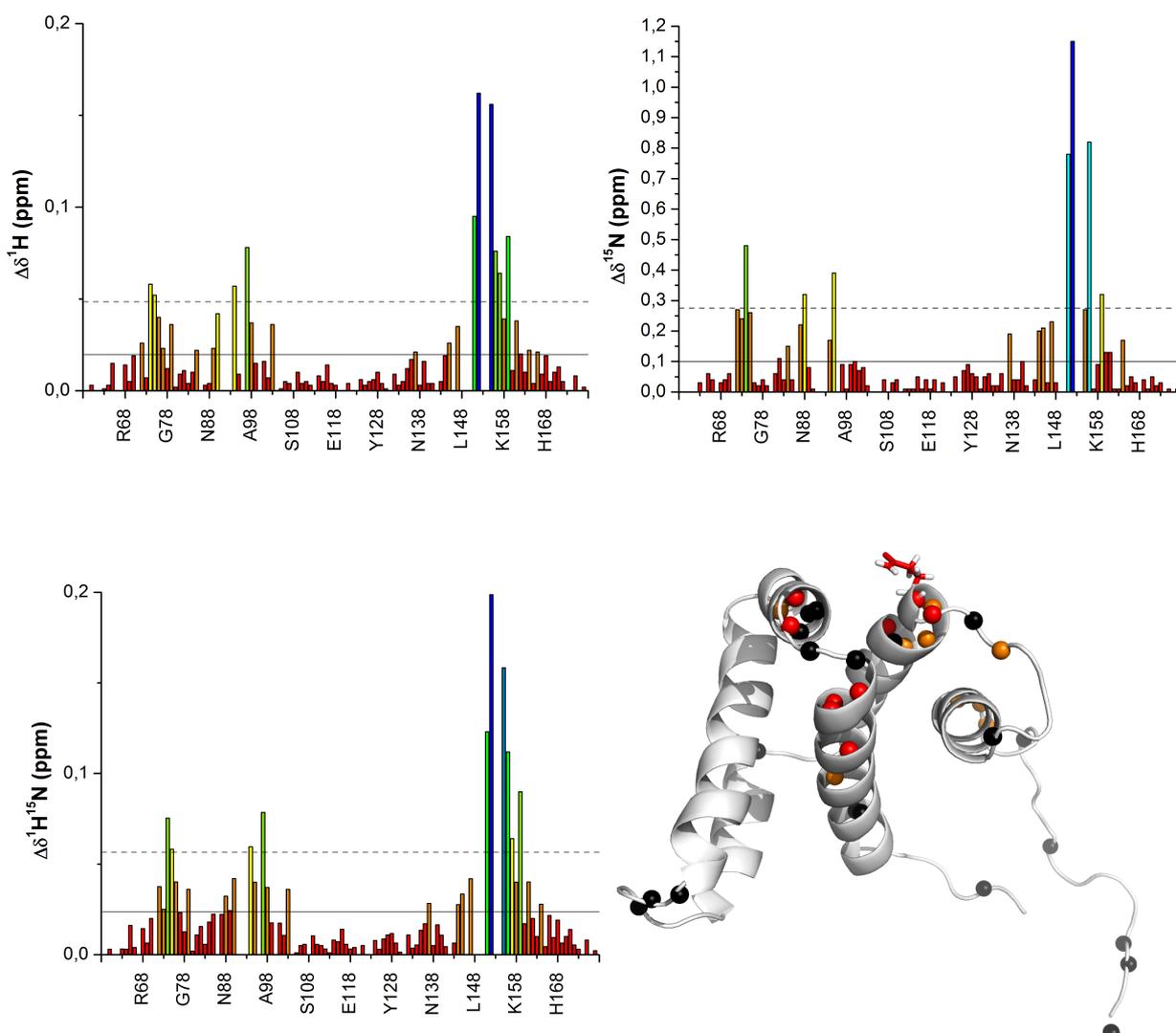


Variations de déplacement chimique en ^1H , en ^{15}N et en $^1\text{H}^{15}\text{N}$ combinées, résidu par résidu pour M2-1[58-177] en présence de l'ARN « UGA2 » avec un rapport molaire ARN/protéine de 20/1. Le code couleur correspond à l'amplitude des variations : les valeurs les plus basses sont en rouge, les plus élevées en bleu. La moyenne et l'écart type moyen sont indiqués par des traits pleins et pointillés.

Cartographie des variations de déplacement chimique sur la structure du domaine central de M2-1. Les atomes d'azote des liaisons amides des résidus qui présentent une variation significative sont indiqués par des sphères et la glutamine 93 en bâtonnets.

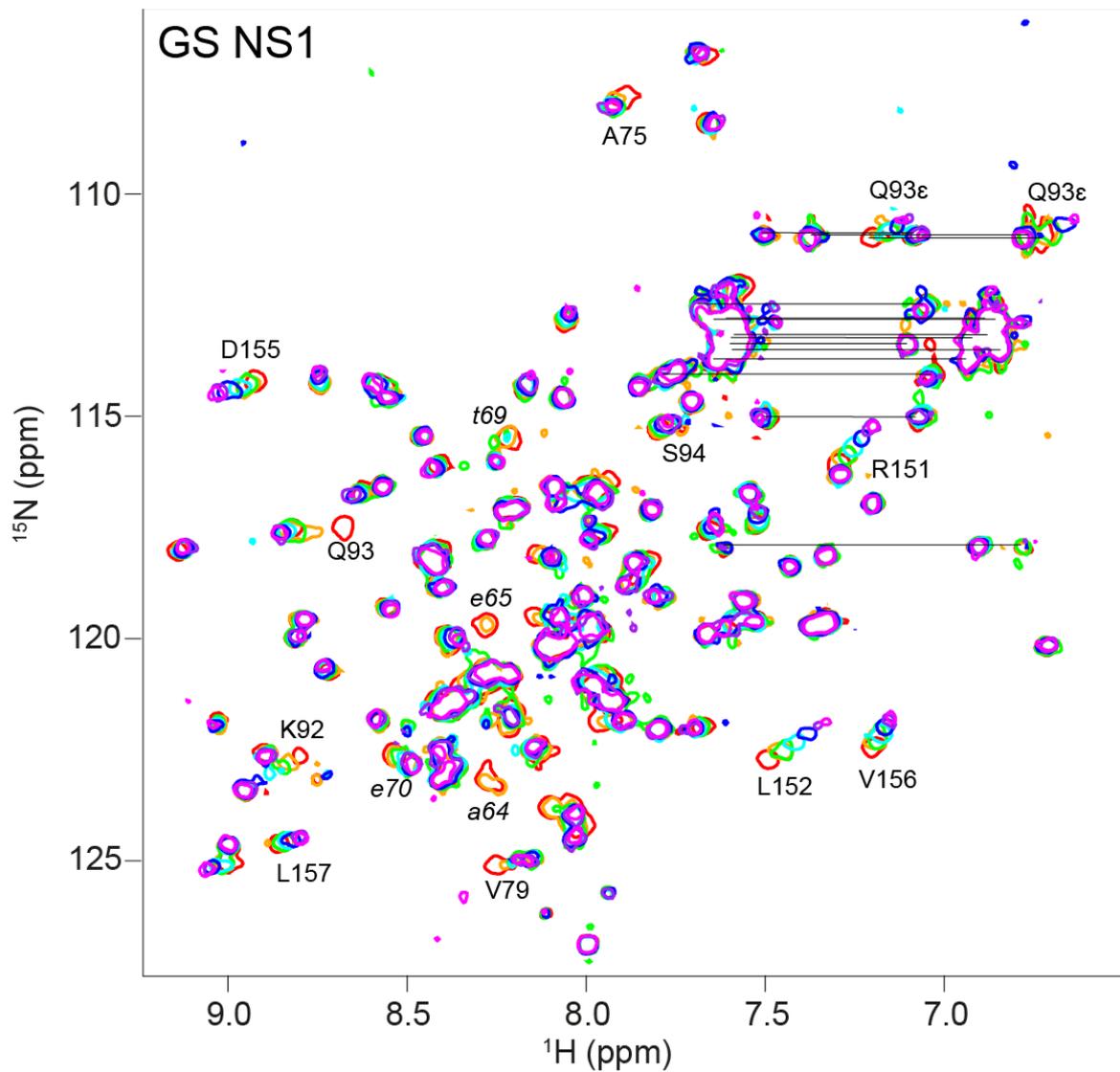


Superposition de spectres ^1H - ^{15}N -HSQC de M2-1[58-177] en présence de différentes concentrations de l'ARN « Leader» (50 μM , $T = 293\text{K}$, $B_0 = 14,1\text{ T}$, 600 MHz). Spectre de référence (rouge), en présence de 0,5 équivalent (orange), 1 équivalent (vert), 1,5 équivalents (bleu clair), 2,5 équivalents (bleu) et 6 équivalents (Violet pourpre).

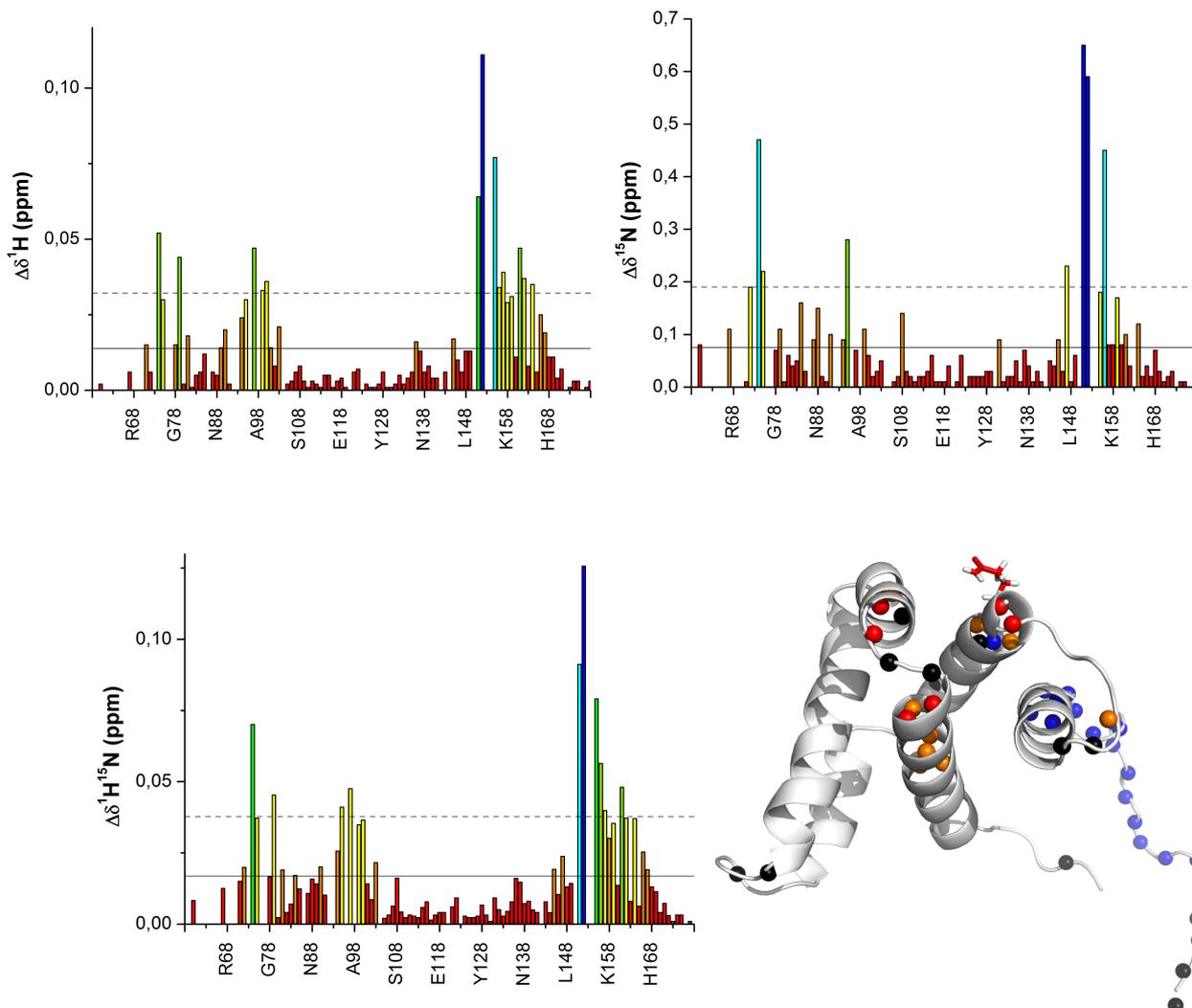


Variations de déplacement chimique en ^1H , en ^{15}N et en $^1\text{H}^{15}\text{N}$ combinées, résidu par résidu pour M2-1[58-177] en présence de l'ARN « Leader» avec un rapport molaire ARN/protéine de 6/1. Le code couleur correspond à l'amplitude des variations : les valeurs les plus basses sont en rouge, les plus élevées en bleu. La moyenne et l'écart type moyen sont indiqués par des traits pleins et pointillés.

Cartographie des variations de déplacement chimique sur la structure du domaine central de M2-1. Les atomes d'azote des liaisons amides des résidus qui présentent une variation significative sont indiqués par des sphères.

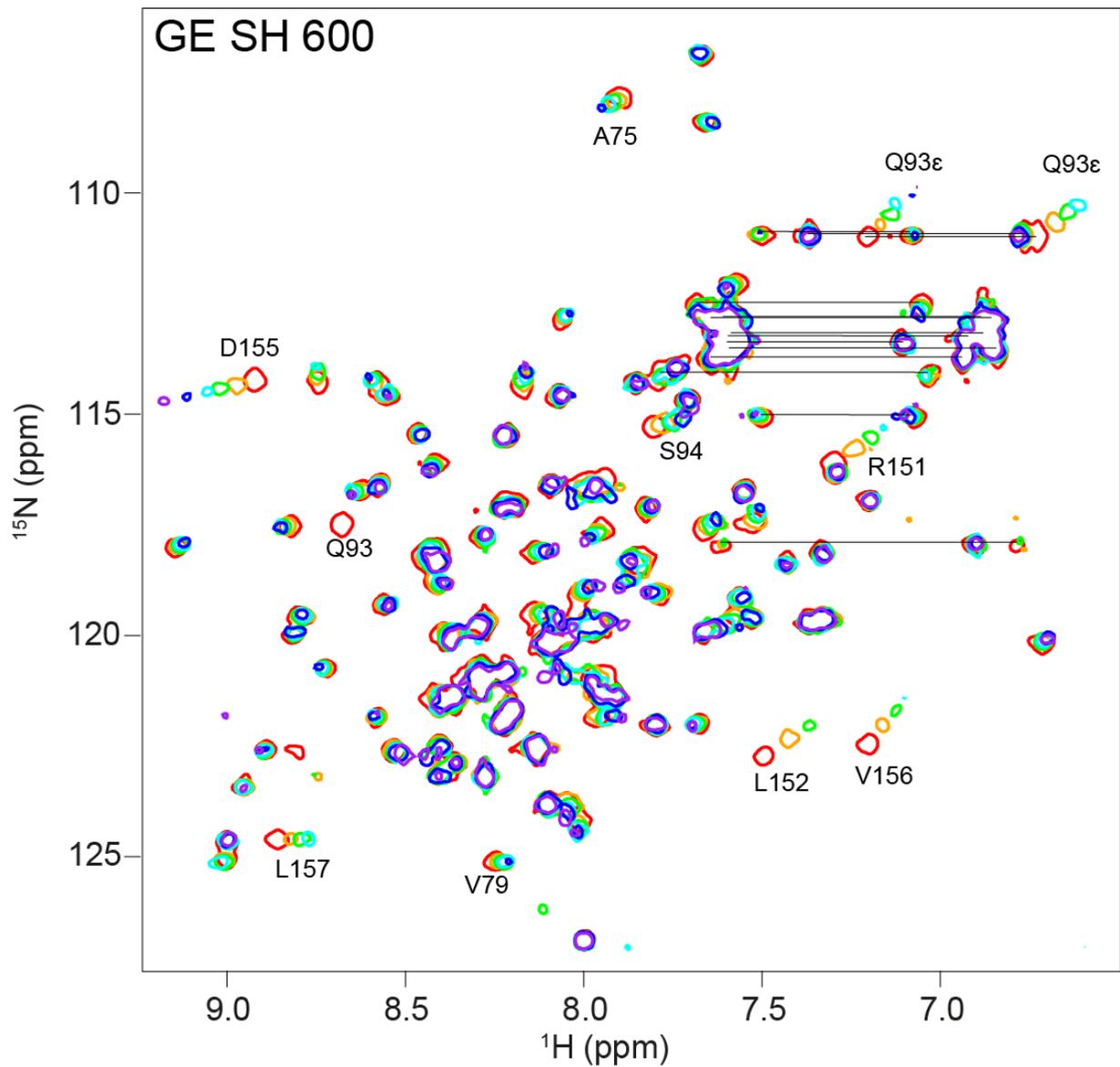


Superposition de spectres ^1H - ^{15}N -HSQC de M2-1[58-177] en présence de différentes concentrations de l'ARN « Gene Start NS1» (50 μM , T = 293K, $B_0 = 14,1$ T, 600 MHz). Spectre de référence (rouge), en présence de 0,5 équivalent (orange), 1 équivalent (vert), 1,5 équivalents (bleu clair), 2,5 équivalents (bleu) 5 équivalents (Violet pourpre) et 10 équivalents (magenta).

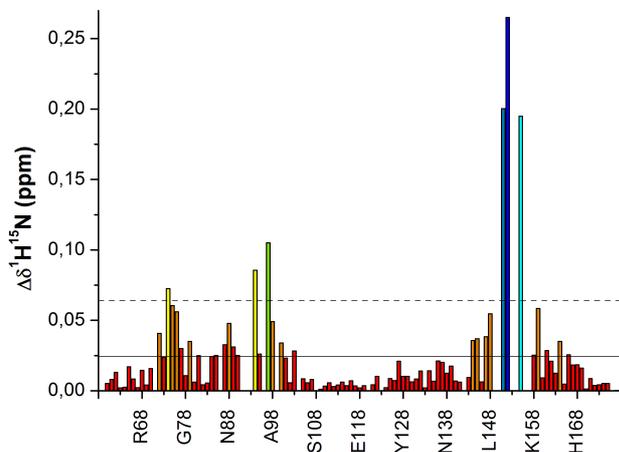
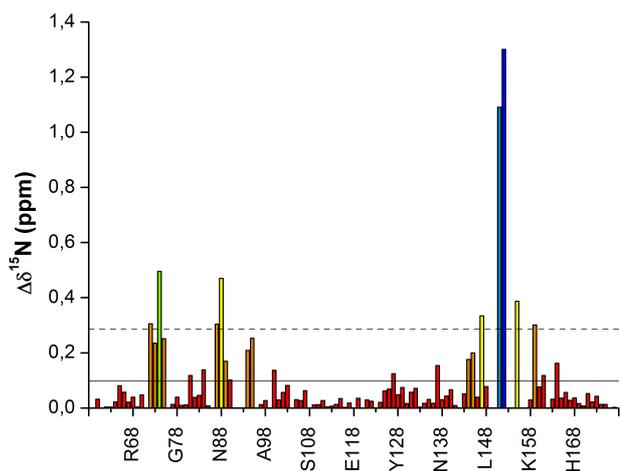
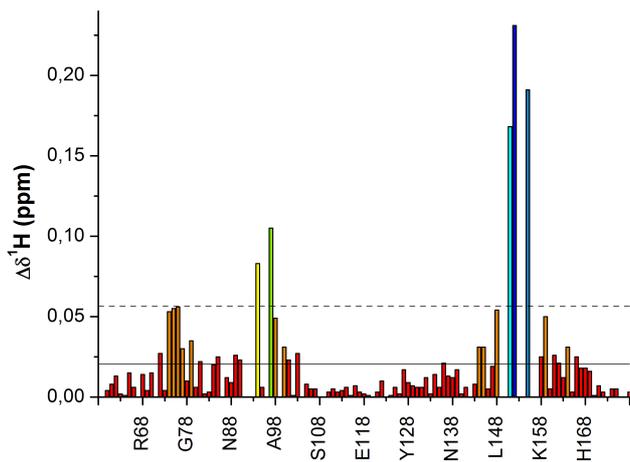


Variations de déplacement chimique en ^1H , en ^{15}N et en $^1\text{H}^{15}\text{N}$ combinées, résidu par résidu pour M2-1[58-177] en présence de l'ARN « Gene Start NS1» avec un rapport molaire ARN/protéine de 2,5/1. Le code couleur correspond à l'amplitude des variations : les valeurs les plus basses sont en rouge, les plus élevées en bleu. La moyenne et l'écart type moyen sont indiqués par des traits pleins et pointillés.

Cartographie des variations de déplacement chimique sur la structure du domaine central de M2-1. Les atomes d'azote des liaisons amides des résidus qui présentent une variation significative sont indiqués par des sphères.

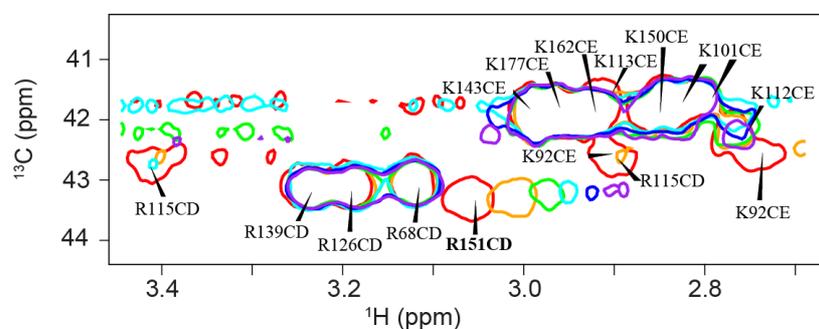
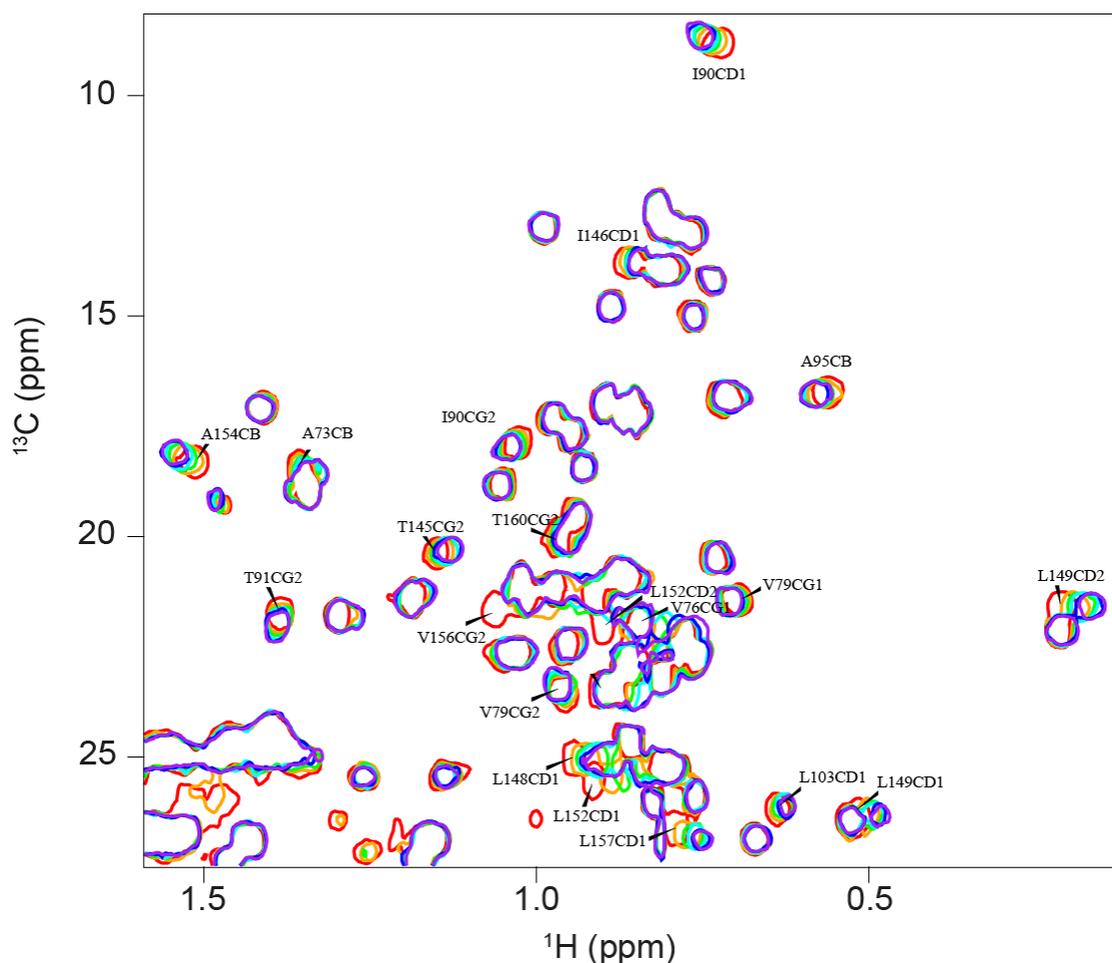


Superposition de spectres ^1H - ^{15}N -HSQC de M2-1[58-177] en présence de différentes concentrations de l'ARN « Gene End SH» (50 μM , $T = 293\text{K}$, $B_0 = 14,1\text{ T}$, 600 MHz). Spectre de référence (rouge), en présence de 0,5 équivalent (orange), 1 équivalent (vert), 1,5 équivalents (bleu clair), 2,5 équivalents (bleu) et 6 équivalents (Violet pourpre).

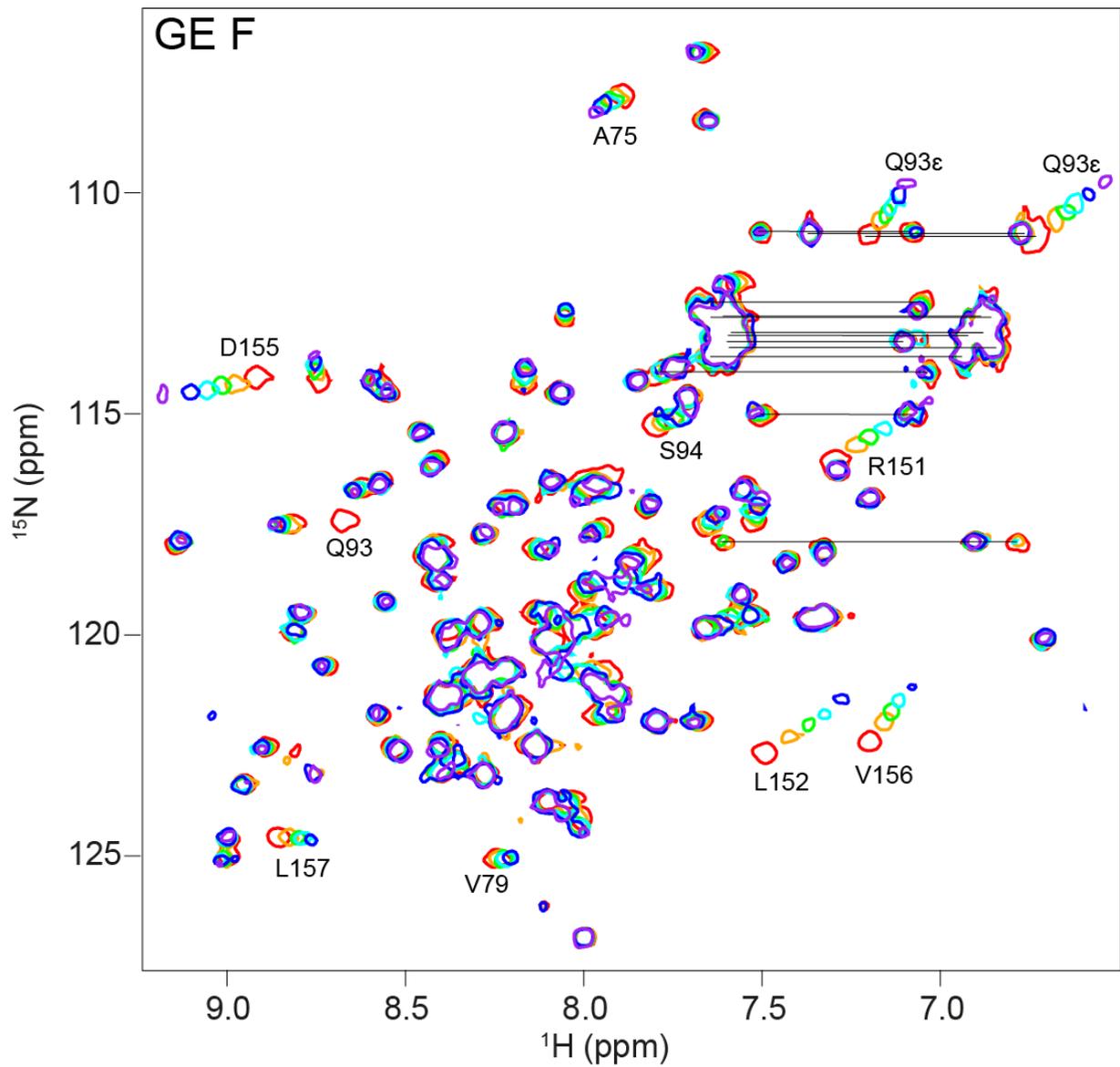


Variations de déplacement chimique en ^1H , en ^{15}N et en $^1\text{H}^{15}\text{N}$ combinées, résidu par résidu pour M2-1[58-177] en présence de l'ARN « Gene End SH» avec un rapport molaire ARN/protéine de 1,5/1. Le code couleur correspond à l'amplitude des variations : les valeurs les plus basses sont en rouge, les plus élevées en bleu. La moyenne et l'écart type moyen sont indiqués par des traits pleins et pointillés.

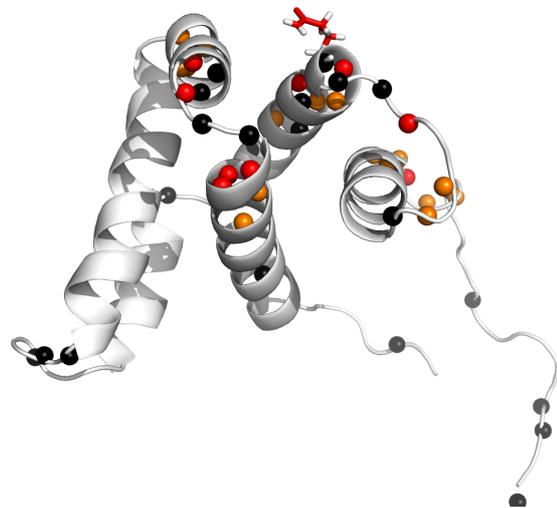
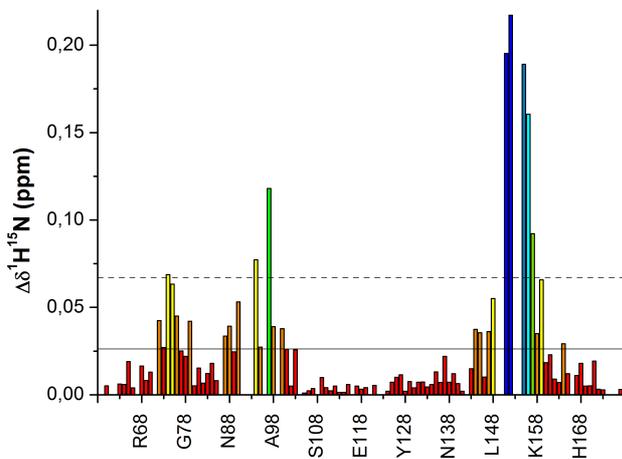
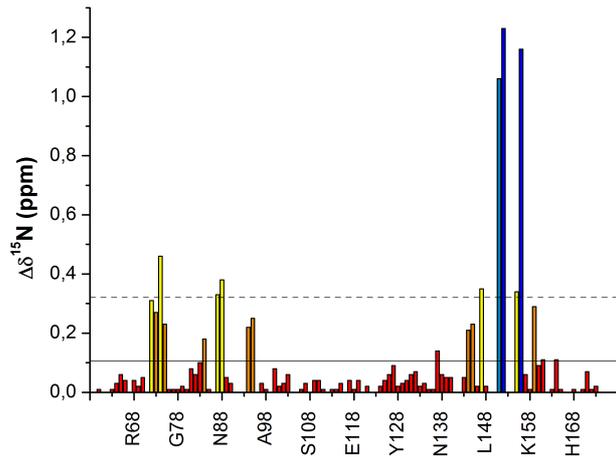
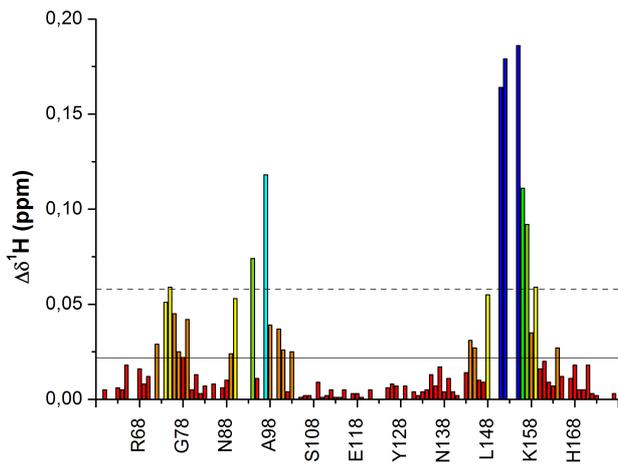
Cartographie des variations de déplacement chimique sur la structure du domaine central de M2-1. Les atomes d'azote des liaisons amides des résidus qui présentent une variation significative sont indiqués par des sphères.



Superposition de spectres ^1H - ^{13}C -HSQC de M2-1[58-177] en présence de différentes concentrations de l'ARN « Gene End SH» (50 μM , $T = 293\text{K}$, $B_0 = 14,1\text{ T}$, 950 MHz). Spectre de référence (rouge), en présence de 0,5 équivalent (orange), 1 équivalent (vert), 1,5 équivalents (bleu clair), 2,5 équivalents (bleu) 5 équivalents (Violet pourpre), avec un zoom sur la région des méthyles et un zoom sur les résonances des chaînes latérales des arginines ($\text{C}\delta\text{-H}\delta$) et des lysines ($\text{C}\epsilon\text{-H}\epsilon$)

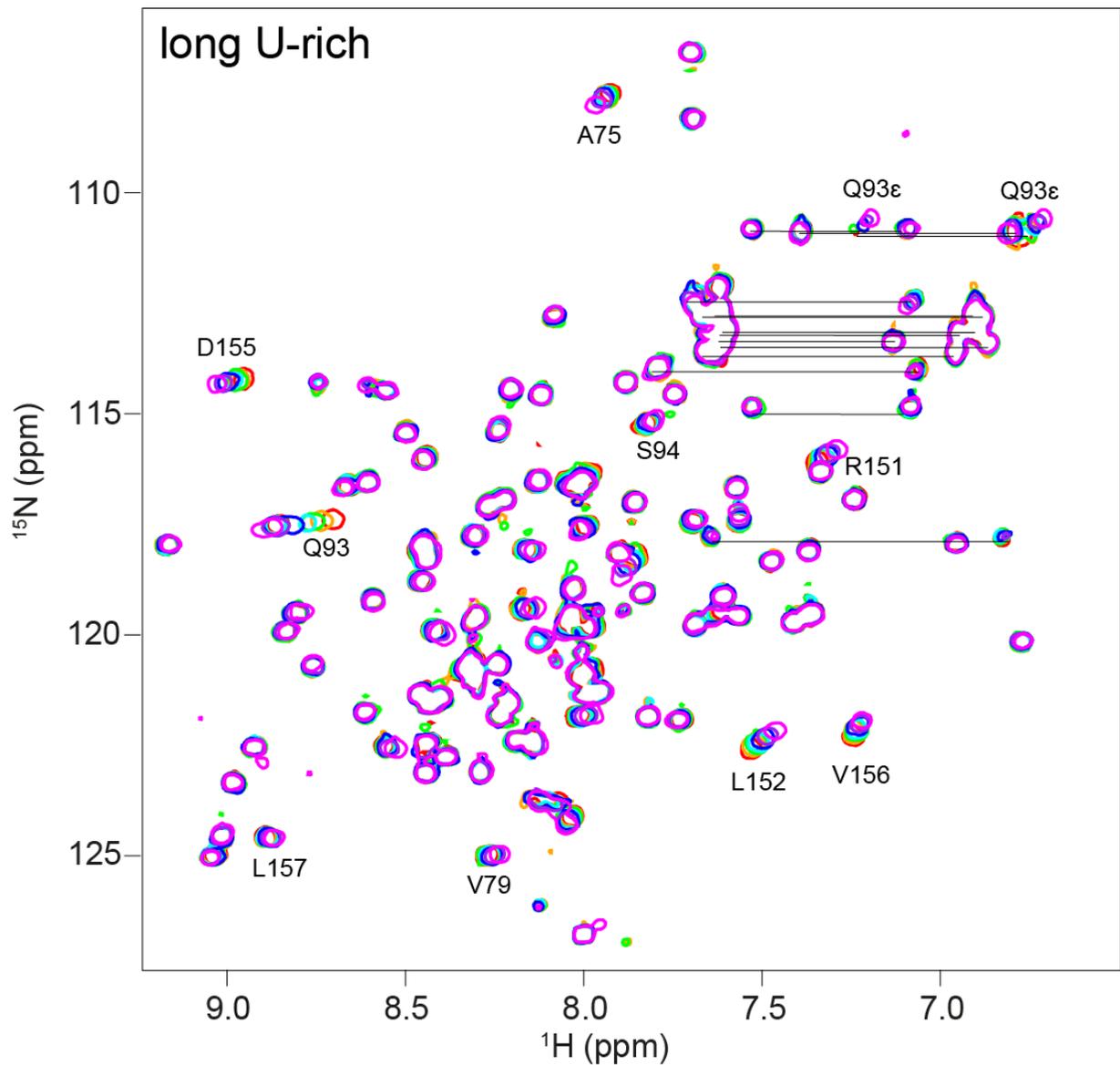


Superposition de spectres ^1H - ^{15}N -HSQC de M2-1[58-177] en présence de différentes concentrations de l'ARN « Gene end F» (50 μM , $T = 293\text{K}$, $B_0 = 14,1\text{ T}$, 600 MHz). Spectre de référence (rouge), en présence de 0,5 équivalent (orange), 1 équivalent (vert), 1,5 équivalents (bleu clair), 2,5 équivalents (bleu) et 6 équivalents (Violet pourpre).

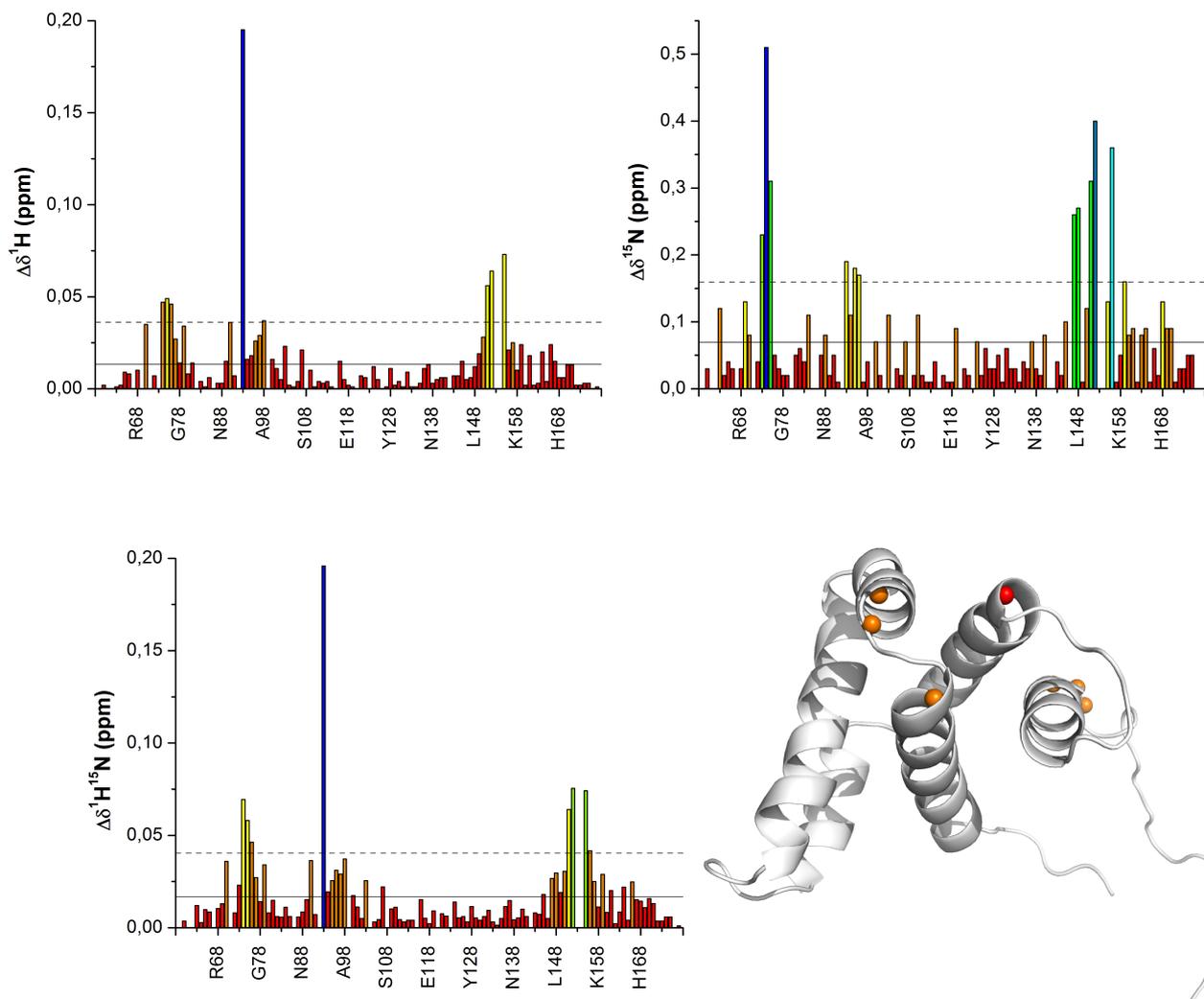


Variations de déplacement chimique en ^1H , en ^{15}N et en $^1\text{H}^{15}\text{N}$ combinées, résidu par résidu pour M2-1[58-177] en présence de l'ARN « Gene end F » avec un rapport molaire ARN/protéine de 2,5/1. Le code couleur correspond à l'amplitude des variations : les valeurs les plus basses sont en rouge, les plus élevées en bleu. La moyenne et l'écart type moyen sont indiqués par des traits pleins et pointillés.

Cartographie des variations de déplacement chimique sur la structure du domaine central de M2-1. Les atomes d'azote des liaisons amides des résidus qui présentent une variation significative sont indiqués par des sphères.



Superposition de spectres ^1H - ^{15}N -HSQC de M2-1[58-177] en présence de différentes concentrations de l'ARN « poly U long » (50 μM , T = 298K, $B_0 = 14,1$ T, 600 MHz). Spectre de référence (rouge), en présence de 0,5 équivalent (orange), 1 équivalent (vert), 1,5 équivalents (bleu clair), 3 équivalents (bleu) 6 équivalents (Violet pourpre) et 20 équivalents (magenta).



Variations de déplacement chimique en ^1H , en ^{15}N et en $^1\text{H}^{15}\text{N}$ combinées, résidu par résidu pour M2-1[58-177] en présence de l'ARN « poly U long » avec un rapport molaire ARN/protéine de 30/1. Le code couleur correspond à l'amplitude des variations : les valeurs les plus basses sont en rouge, les plus élevées en bleu. La moyenne et l'écart type moyen sont indiqués par des traits pleins et pointillés.

Cartographie des variations de déplacement chimique sur la structure du domaine central de M2-1. Les atomes d'azote des liaisons amides des résidus qui présentent une variation significative sont indiqués par des sphères.

ANNEXE 6 : Déplacements chimiques ^1H pour ZNF29

NUM	AA	H	HA	HB	HG	HD	HE
1	Met	8.12	4.10	-	2.53,2.66	-	2.11
2	Ser	8.71	4.45	3.78	-	-	-
3	Arg	8.46	4.27	1.76,1.66	1.52,1.56	3.12	-
4	Arg	8.35	4.23	1.64,1.70	1.53,1.49	-	-
5	Asn	8.51	4.87	2.74,2.59	-	7.54,6.87	-
6	Pro	-	4.34	2.18,1.87	1.93	3.72,3.65	-
7	Cys	8.30	4.31	2.77,2.77	-	-	-
8	Lys	8.24	4.15	1.63,1.14	1.19,1.24	1.57	2.85
9	Phe	8.05	4.55	2.90,3.01	-	7.13	-
10	Glu	8.12	4.25	1.78,1.88	2.17	-	-
11	Ile	8.17	4.01	1.73	1.10,1.40	0.79	-
12	Arg	8.38	4.25	1.67,1.75	1.52,1.55	3.09	7.10
13	Gly	8.31	3.83	-	-	-	-
14	His	8.34	4.64	3.18,3.06	-	7.18	8.50
15	Cys	8.41	4.41	2.81,2.81	-	-	-
16	Leu	8.55	4.28	1.52,1.52	1.52	0.81,0.77	-
17	Asn	8.36	4.61	2.70,2.75	-	7.51,6.84	-
18	Gly	8.29	3.85	-	-	-	-
19	Lys	8.06	4.22	1.67,1.74	1.35,1.32	1.57	2.89
20	Arg	8.27	4.24	1.73,1.65	1.53,1.49	3.08	7.07
21	Cys	8.19	4.31	2.68,2.68	-	-	-
22	His	8.44	4.54	2.98,3.05	-	7.06	-
23	Phe	8.17	4.58	3.04,2.86	-	-	-
24	Ser	8.20	4.29	3.74,3.70	-	-	-
25	His	8.41	4.50	3.01,3.08	-	7.06	-
26	Asn	8.30	4.51	2.59,2.55	-	7.43,6.79	-
27	Tyr	7.91	4.36	2.74,2.80	-	6.88	6.65
28	Phe	7.95	4.51	3.04,2.86	-	7.10	7.20
29	Glu	7.70	4.08	1.81,1.97	2.21,2.21	-	-

ANNEXE 7 : Déplacements chimiques du squelette peptidique pour M2-1Δ[35-58]

NUM	AA	H	N	CA	CB	ND	NE	HD	HE	SS
59	Glu	-	-	55.95	30.64	-	-	-	-	C
60	Ile	8.19	121.71	60.91	38.77	-	-	-	-	C
61	Ser	8.28	119.50	58.43	63.79	-	-	-	-	C
62	Gly	8.34	111.18	45.34	-	-	-	-	-	None
63	Ala	8.00	123.53	52.59	19.21	-	-	-	-	H
64	Ala	8.15	122.65	52.67	19.09	-	-	-	-	C
65	Glu	8.15	119.15	56.60	30.01	-	-	-	-	C
66	Leu	7.99	121.98	55.29	42.24	-	-	-	-	H
67	Asp	8.19	120.60	54.38	41.03	-	-	-	-	None
68	Arg	8.09	121.29	56.17	30.40	-	-	-	-	C
69	Thr	8.10	114.81	63.03	69.43	-	-	-	-	C
70	Glu	8.40	122.35	57.44	29.77	-	-	-	-	H
71	Glu	8.22	120.51	57.78	29.78	-	-	-	-	H
72	Tyr	7.89	119.49	58.57	38.22	-	-	-	-	H
73	Ala	7.94	123.54	53.11	18.85	-	-	-	-	H
74	Leu	7.80	118.60	55.52	43.48	-	-	-	-	C
75	Gly	7.89	107.92	44.53	-	-	-	-	-	C
76	Val	8.05	119.34	66.85	-	-	-	-	-	H
77	Val	8.06	117.94	67.08	-	-	-	-	-	H
78	Gly	7.61	106.86	47.51	-	-	-	-	-	H
79	Val	8.18	124.97	-	-	-	-	-	-	H
80	Leu	7.47	119.43	58.33	-	-	-	-	-	H
81	Glu	9.04	117.82	59.64	-	-	-	-	-	H
82	Ser	8.08	116.74	62.54	-	-	-	-	-	H
83	Tyr	8.32	122.40	-	-	-	-	-	-	H
84	Ile	7.96	112.64	64.83	-	-	-	-	-	H
85	Gly	7.59	108.16	44.78	-	-	-	-	-	C
86	Ser	7.56	113.38	-	-	-	-	-	-	C
87	Ile	8.09	114.40	60.19	-	-	-	-	-	C
88	Asn	8.61	114.35	54.76	37.66	114.58	-	6.97,7.40	-	C
89	Asn	8.51	114.50	53.82	37.55	113.25	-	7.50,6.78	-	C
90	Ile	7.89	119.77	58.58	-	-	-	-	-	E
91	Thr	8.33	117.89	60.75	-	-	-	-	-	C
92	Lys	8.79	122.75	-	-	-	-	-	-	H
93	Gln	8.63	117.35	59.93	-	-	110.75	-	7.12,6.63	H
94	Ser	7.71	115.16	-	-	-	-	-	-	H
95	Ala	7.92	124.06	54.96	16.94	-	-	-	-	H
96	Cys	7.87	116.27	65.08	-	-	-	-	-	H
97	Val	7.89	120.74	66.85	-	-	-	-	-	H
98	Ala	7.92	121.72	55.29	17.24	-	-	-	-	H
99	Met	8.32	117.82	57.76	-	-	-	-	-	H
100	Ser	7.51	112.17	61.60	-	-	-	-	-	H
101	Lys	7.71	118.90	59.68	-	-	-	-	-	H

102	Leu	8.36	122.94	-	-	-	-	-	-	H
103	Leu	8.55	116.59	57.96	-	-	-	-	-	H
104	Thr	7.51	112.55	65.50	69.34	-	-	-	-	H
105	Glu	7.49	119.03	56.80	30.16	-	-	-	-	H
106	Leu	7.25	119.35	54.13	-	-	-	-	-	H
107	Asn	8.63	119.33	51.80	39.59	113.32	-	7.04	-	C
108	Ser	9.00	121.93	62.35	-	-	-	-	-	H
109	Asp	8.31	121.41	57.97	-	-	-	-	-	H
110	Asp	7.59	119.58	-	-	-	-	-	-	H
111	Ile	7.26	119.35	-	-	-	-	-	-	H
112	Lys	8.33	121.39	-	-	-	-	-	-	H
113	Lys	7.16	116.98	-	-	-	-	-	-	H
114	Leu	6.87	117.95	-	-	-	-	-	-	H
115	Arg	7.93	120.81	59.52	-	-	-	-	-	H
116	Asp	8.31	118.12	56.09	39.63	-	-	-	-	H
117	Asn	7.24	116.26	53.10	39.74	112.74	-	6.85,7.55	-	H
118	Glu	7.30	119.65	55.29	30.81	-	-	-	-	None
119	Glu	8.91	124.24	55.18	-	-	-	-	-	C
120	Pro	-	-	64.41	31.74	-	-	-	-	C
121	Asn	8.41	114.33	53.14	37.66	113.45	-	7.54,6.86	-	None
122	Ser	7.87	116.34	55.73	-	-	-	-	-	C
123	Pro	-	-	64.42	31.84	-	-	-	-	H
124	Lys	8.01	116.37	57.87	32.49	-	-	-	-	H
125	Ile	7.36	118.20	66.27	-	-	-	-	-	H
126	Arg	7.44	116.63	59.07	30.04	-	-	-	-	H
127	Val	6.67	120.06	66.15	-	-	-	-	-	H
128	Tyr	8.17	117.73	60.17	-	-	-	-	-	H
129	Asn	8.49	116.54	55.94	-	110.71	-	7.43,6.96	-	H
130	Thr	8.19	120.60	-	-	-	-	-	-	H
131	Val	8.12	122.24	-	-	-	-	-	-	H
132	Ile	8.71	119.83	66.70	-	-	-	-	-	H
133	Ser	7.76	114.25	62.08	-	-	-	-	-	H
134	Tyr	7.89	121.20	61.39	-	-	-	-	-	H
135	Ile	8.81	122.48	66.44	-	-	-	-	-	H
136	Glu	8.34	118.76	59.34	-	-	-	-	-	H
137	Ser	8.32	115.86	61.86	62.96	-	-	-	-	H
138	Asn	8.49	121.59	56.67	-	117.53	-	6.75,7.52	-	H
139	Arg	7.75	116.91	58.12	29.97	-	-	-	-	H
140	Lys	7.26	118.03	58.52	32.80	-	-	-	-	H
141	Asn	8.01	114.48	53.83	-	113.98	-	7.69,6.97	-	H
142	Asn	-	-	56.95	37.08	113.18	-	6.84,7.54	-	H
143	Lys	7.87	121.17	59.84	-	-	-	-	-	H
144	Gln	7.93	118.90	57.93	28.17	-	110.83	-	7.29,6.70	H
145	Thr	7.58	117.34	67.73	-	-	-	-	-	H
146	Ile	8.29	122.57	66.94	-	-	-	-	-	H
147	His	7.47	117.08	59.73	29.92	-	-	-	-	H
148	Leu	7.95	119.73	58.27	-	-	-	-	-	H
149	Leu	8.25	120.70	57.76	41.75	-	-	-	-	H

150	Lys	7.92	116.45	58.75	32.40	-	-	-	-	H
151	Arg	7.23	115.95	56.01	-	-	-	-	-	C
152	Leu	7.41	122.40	53.42	-	-	-	-	-	C
153	Pro									C
154	Ala	8.89	126.35	56.35	18.44	-	-	-	-	H
155	Asp	8.86	114.18	57.28	39.36	-	-	-	-	H
156	Val	7.13	122.13	65.48	-	-	-	-	-	H
157	Leu	8.78	124.52	58.63	-	-	-	-	-	H
158	Lys	8.76	117.48	61.50	-	-	-	-	-	H
159	Lys	7.51	119.38	-	32.26	-	-	-	-	H
160	Thr	8.16	117.00	67.26	-	-	-	-	-	H
161	Ile	8.93	124.95	66.35	-	-	-	-	-	H
162	Lys	8.37	121.29	60.26	31.99	-	-	-	-	H
163	Asn	8.38	115.38	55.93	-	112.61	-	6.99,7.59	-	H
164	Thr	7.91	117.55	-	-	-	-	-	-	H
165	Leu	8.87	123.30	58.07	41.76	-	-	-	-	H
166	Asp	8.30	120.01	57.63	39.86	-	-	-	-	H
167	Ile	7.66	121.85	65.31	-	-	-	-	-	H
168	His	8.62	120.27	60.53	-	-	-	-	-	H
169	Lys	8.44	119.23	59.35	32.13	-	-	-	-	H
170	Ser	7.70	113.90	61.36	63.42	-	-	-	-	H
171	Ile	7.76	117.64	62.81	38.33	-	-	-	-	H
172	Thr	7.61	114.53	63.52	69.25	-	-	-	-	H
173	Ile	7.72	121.57	61.98	38.34	-	-	-	-	H
174	Asn	8.03	120.36	53.25	39.07	113.03	-	6.85,7.51	-	H
175	Asn	8.06	119.75	51.36	-	112.86	-	6.81,7.48	-	C
176	Pro	-	-	63.46	31.98	-	-	-	-	C
177	Lys	8.19	120.64	56.56	32.72	-	-	-	-	None
178	Glu	8.17	120.96	56.40	30.10	-	-	-	-	C
179	Leu	8.08	122.91	55.09	42.31	-	-	-	-	C
180	Thr	8.08	115.56	61.69	69.82	-	-	-	-	None
181	Val	8.07	122.20	62.09	32.71	-	-	-	-	C
182	Ser	8.27	119.25	58.20	63.87	-	-	-	-	C
183	Asp	8.30	122.80	54.40	41.10	-	-	-	-	C
184	Thr	8.03	113.74	61.94	69.43	-	-	-	-	C
185	Asn	8.31	120.54	53.33	38.82	-	-	-	-	C
186	Asp	8.12	120.54	54.47	40.90	-	-	-	-	C
187	His	8.17	118.91	55.89	29.38	-	-	-	-	H
188	Ala	8.08	124.07	52.68	19.07	-	-	-	-	C
189	Lys	8.10	120.03	56.26	32.78	-	-	-	-	C
190	Asn	8.27	119.29	53.20	38.87	-	-	-	-	C
191	Asn	8.30	119.20	53.27	38.90	-	-	-	-	C
192	Asp	8.24	120.55	54.44	41.17	-	-	-	-	C
193	Thr	8.03	113.76	61.70	69.71	-	-	-	-	
194	Thr	7.75	121.39	63.60	-	-	-	-	-	