



**HAL**  
open science

# Caractérisation structurale par RMN des interactions entre protéines du complexe polymérase du virus respiratoire syncytial et des protéines partenaires cellulaires

Christophe Cardone

► **To cite this version:**

Christophe Cardone. Caractérisation structurale par RMN des interactions entre protéines du complexe polymérase du virus respiratoire syncytial et des protéines partenaires cellulaires. Biologie structurale [q-bio.BM]. Université Paris-Saclay, 2019. Français. NNT : 2019SACLS586 . tel-02462886

**HAL Id: tel-02462886**

**<https://theses.hal.science/tel-02462886>**

Submitted on 31 Jan 2020

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

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

# Caractérisation structurale par RMN des interactions entre protéines du complexe polymérase du virus respiratoire syncytial et des protéines partenaires cellulaires

Thèse de doctorat de l'Université Paris-Saclay  
préparée à l'Université Paris-Sud

École doctorale n° 569 Innovation thérapeutique du fondamental à  
l'appliqué (ITFA)  
Spécialité de doctorat : Biochimie et biologie structurale

Thèse présentée et soutenue à Gif-sur-Yvette, le 16 décembre 2019, par

**Christophe Cardone**

Composition du Jury :

Herman VAN TILBEURGH, Professeur des universités, I2BC Université Paris Saclay	Président
Malene RINGKJØBING JENSEN Directrice de recherche, CNRS, IBS Grenoble	Rapporteuse
Serge BOUAZIZ Directeur de recherche, CNRS, CiTCoM Université Paris Descartes	Rapporteur
Benjamin BARDIAUX Chargé de recherche, CNRS, Institut Pasteur Paris	Examineur
Christina SIZUN Chargée de recherche, CNRS, ICSN Gif-sur-Yvette	Directrice de thèse

## Résumé

Le virus respiratoire syncytial humain (hRSV) est le principal agent pathogène responsable des bronchiolites. Le complexe ARN polymérase (RdRp), du hRSV, nécessaire à la réplication de son génome, est composé *a minima* de la sous-unité catalytique (L), de son principal cofacteur qu'est la phosphoprotéine (P) et de la nucléoprotéine (N) qui assure l'encapsidation du génome viral. Le cœur de mon projet doctoral a été l'étude dynamique et structurale de domaines des protéines N et P du hRSV ainsi que leurs interactions avec certaines protéines cellulaires principalement par résonance magnétique nucléaire.

Dans un premier temps j'ai étudié une potentielle interaction entre 2 domaines appartenant à la protéine N et à la protéine cellulaire Tax1BP1 impliquée notamment dans la régulation de l'autophagie. Ensuite, j'ai entrepris une étude structurale et dynamique de hRSV-P isolée notamment dans le but de déterminer des contacts transitoires au sein de la protéine et d'obtenir la structure tridimensionnelle du domaine d'oligomérisation de P. Enfin, j'ai participé à la caractérisation de l'interaction entre la protéine hRSV-P et le cofacteur de transcription du hRSV hRSV-M2-1, puis entre hRSV P et la phosphatase cellulaire PP1 $\alpha$ , afin d'en cartographier les régions de contacts.

## Abstract

Human respiratory syncytial virus (hRSV) is the main pathogen responsible for bronchiolitis. The RNA polymerase complex (RdRp) of hRSV, necessary for the replication of its genome, is composed at least of the catalytic subunit (L), its main cofactor phosphoprotein (P) and nucleoprotein (N), which encapsidates the viral genome. At the heart of my doctoral project was the dynamic and structural study of domains of the proteins N and P of the hRSV as well as of their interactions with several cellular proteins, mainly by nuclear magnetic resonance.

Firstly, I studied a potential interaction between 2 domains belonging to the N protein and to the Tax1BP1 cellular protein involved notably in regulation of autophagy. Secondly, I undertook a structural and dynamic study of isolated hRSV-P in order to determine transient contacts within the protein and to obtain the three-dimensional structure of the P oligomerization domain. Last, I participated in the characterization of the interaction between the hRSV-P protein and the hRSV transcription cofactor hRSV-M2-1, and between hRSV P and the cellular phosphatase PP1 $\alpha$  to map the contact regions.

**Mots-clés** : virus respiratoire syncytial, résonance magnétique nucléaire, interactions de protéines, structure de protéine, dynamique de protéines, phosphoprotéine, nucléoprotéine



# Remerciements

Pour commencer, je tiens à remercier Christina Sizun qui m'a accepté au laboratoire d'abord comme encadrante de stages puis comme directrice de thèse. En effet, elle a toujours été un soutien indéfectible pour mes travaux et au-delà. Je suis très reconnaissant pour le temps qu'elle m'a accordé. Christina a, sans l'ombre d'un doute, été pour moi une excellente directrice de thèse tant sur le plan de la transmission des connaissances et du savoir-faire que sur le plan pédagogique et social.

Egalement, je remercie Eric Guittet et surtout Carine van Heijenoort de m'avoir accepté au sein du laboratoire de chimie et biologie structurales de l'ICSN.

Je remercie chaque membre de mon jury de thèse (Malene, Serge, Benjamin et Herman) d'avoir accepté de juger mes travaux.

Je suis très heureux d'avoir pu faire la connaissance de chaque permanent (Carine, Nadine, François Bontems, Ewen, Nelly, Annie, François Giraud, Eric Jaquet, Jean-François Gallard, Jean-Nicolas, Naïma). En pensant à chacun je revois tous les bons moments passés à vos côtés et toutes vos contributions sur le plan scientifique. De même, je suis favorisé d'avoir pu rencontrer tous les étudiants passés par le laboratoire et en particulier Florian, Louise, Claire-Marie et Camille avec qui nous avons cheminé durant ces 3 années sinon plus pour certain ; cela a été très agréable d'être avec eux et leur aide m'a été précieuse. Je remercie aussi tous les professionnels de l'ICSN, notamment Alda, Emilie, Guillaume et David m'ayant accompagné dans divers domaines.

Je suis également privilégié d'avoir ma formidable épouse, Monique, qui m'accompagne et me soutient fidèlement en m'encourageant depuis le lycée jusqu'ici. Je remercie mes enfants, Claire, Chalom et Cleya, qui m'ont apporté beaucoup de joie et d'amour tout au long de ces années. Merci à tous les membres de ma famille et de ma belle-famille, notamment ma mère sur qui je peux compter et Nadia qui me pousse toujours dans le bon sens. Je remercie également la communauté Providence, et en particulier Jellil et Shirley do Rego, pour leurs conseils et leur soutien à tous les niveaux au quotidien.

# Abréviations

ARN : Acide RiboNucléique

BEST TROSY : Band-selective Excitation Short-Transient Transverse relaxation-optimized spectroscopy

C arm : bras C-terminal (de N)

DSS : Sodium trimethylsilylpropanesulfonate

DTT : Dithiothréitol

GST : Glutathion S-Transférase

hMPV : Human MetaPneumoVirus – MetaPneumoVirus Humain

hRSV : Human Respiratory Syncytial Virus – Virus Respiratoire Syncytial Humain

HSQC : Heteronuclear Single Quantum Correlation

IDP : Intrinsically Disordered Protein – Protéine Intrinsèquement Désordonnée

IDR : Intrinsically Disordered Region – Région Intrinsèquement Désordonnée

IPTG : IsoPropyl béta-D-1-ThioGalactopyranoside

Kd : Constante de dissociation

kDa : kiloDalton

M2-1/2 : Protéine de matrice 2-1/2

N : Nucléoprotéine

OD : Oligomerization Domain – Domaine d'Oligomérisation

P : Phosphoprotéine

PP1 $\alpha$  : Protein Phosphatase 1 $\alpha$  – Protéine Phosphatase 1 $\alpha$

PRE : Paramagnetic Relaxation Enhancement – Augmentation de la relaxation par effet paramagnétique

RdRp : RNA-dependent RNA polymerase – ARN polymérase ARN dépendante

SAXS : Small Angle X-ray Scattering – Diffusion des rayons X aux petits angles

SDS-PAGE : Sodium Dodecyl Sulfate PolyAcrylamide Gel Electrophoresis – Électrophorèse sur gel de polyacrylamide en présence de dodécylsulfate de sodium

TAX1BP1 : Tax1-binding protein 1 – Protéine 1 liant Tax1

TCEP : Tris(2-carboxyethyl)phosphine

UBZ1+2 : Ubiquitin-Binding Zinc finger domain 1 and 2– Domaine en doigt de zinc liant l'ubiquitine

# Liste des figures

**Figure 1 : Épidémiologie mensuelle du RSV en Europe.** Le virus se propage généralement du sud au nord et d'ouest en est, montrant une activité sur le continent d'octobre à mai dans certains pays. L'activité maximale est généralement observée en hiver dans la plupart des pays, à l'exception des pays les plus septentrionaux, où une forte activité est observée au début du printemps. Concernant la France, cette étude correspond aux années 2011 à 2017. Cette figure est issue de Obando-Pacheco et al., 2018.

..... 26

**Figure 2 : Caractérisation morphologique de base des trois morphologies du hRSV.** Particules représentatives des trois catégories de morphologie présentes dans l'échantillon de RSV : sphérique (A à C), asymétrique (D à F) et filamenteuse (G à I). Les images sont des coupes centrales de 7,5 nm issues de reconstructions tomographiques. La figure est tirée de Kiss et al., 2014. Barres d'échelle, 100 nm..... 27

**Figure 3 : Arbre phylogénétique des Pneumoviridae en 2017.** Des séquences de l'ARN polymérase L alignées provenant des membres représentatifs des Pneumoviridae ont été utilisées pour produire un arbre phylogénétique par maximum de vraisemblance. Le pourcentage d'arbres dans lesquels les taxons associés sont regroupés est indiqué pour des valeurs supérieures à 70 %. Les barres indiquent les affectations d'espèces de ces isolats de virus. Source : [https://talk.ictvonline.org/cfs-file/\\_key/communityserver-wikis-components-files/00-00-00-00-90/ONSR.Pne.Fig3.v3.png](https://talk.ictvonline.org/cfs-file/_key/communityserver-wikis-components-files/00-00-00-00-90/ONSR.Pne.Fig3.v3.png)..... 28

**Figure 4 : Carte génétique du RSV.** La figure montre une carte du génome à sens négatif (souche A2 du RSV), approximativement à l'échelle. Les cadres de lecture ouvert M2-1 et M2-2 qui se chevauchent sont montrés sur le gène. Les chiffres en dessous de la carte indiquent les longueurs en nucléotides (nt) : celles des régions du leader 3' (le), du trailer 5'(tr) et des régions intergéniques sont soulignées, et les longueurs du chevauchement des gènes sont indiquées entre parenthèses. Les chiffres en gras et en italique sur la carte indiquent les longueurs en acides aminés (aa) des protéines non modifiées. Les protéines virales dans l'ordre sont : G, glycoprotéine d'attachement ; F, glycoprotéine de fusion ; SH, petite glycoprotéine hydrophobe ; M, protéine matricielle ; N, nucléoprotéine ; P, phosphoprotéine ; L, grande polymérase ; M2-1, produit du premier ORF dans l'ARNm M2 ; M2-2, produit du deuxième ORF dans l'ARN M2 ; NS1, protéine non structurale 1 ; NS2, protéine non structurale 2. Figure issue de Collins et Melero, 2011. .... 29

**Figure 5 : Schéma du virion et des protéines du hRSV.** Un focus est mis sur le complexe ARN polymérase ARN dépendante (RdRp) entouré en rouge. Le virion ne contient pas les protéines non structurales NS1 et NS2. .... 30

**Figure 6 : Structure cristallographique de la nucléoprotéine du hRSV sous forme décamerique liée à l'ARN.** Une sous-unité N est colorée selon les différents domaines : jaune et rouge, respectivement pour NTD et CTD. Les bras N- et C-terminaux sont en bleu. L'ARN est affiché en cyan

et les bases azotées en noir. La polarité de l'ARN est indiquée dans l'anneau vu de haut (à droite). La barre d'échelle à gauche indique 100 Å. .... 31

**Figure 7 : Complémentarité de surface N-NTD et un peptide de P.** On voit des vues agrandies de la liaison de P2 (peptide constitué des 2 derniers résidus de P). La surface de van der Waals sur N-NTD est colorée en fonction du potentiel électrostatique : potentiels positif en bleu et négatif en rouge. Le ligand P2 est affiché sous forme de sphères de van der Waals colorées par type d'atome, avec les atomes de carbone en blanc. Figure provenant de Ouizougoun-Oubari et al., 2015. .... 32

**Figure 8 : Modèle structural d'une forme de RSV N monomère (N<sup>0</sup>) obtenue avec le mutant NΔ30 – P40 en solution.** Représentation en ruban du meilleur modèle structural du complexe RSV NΔ30-P40 (40 premiers acides aminés de P). La protéine N (résidus 31 à 382) est représentée en bleu, N-(C arm) (ici : résidus 383 à 391) en violet et le peptide P40 en vert. Les régions de liaison sont représentées par des sphères. .... 32

**Figure 9 : Prédiction d'hélices α de la phosphoprotéine du RSV par Llorente et al., 2008.** Le domaine de liaison de la protéine M2-1 du hRSV se situe en amont du domaine d'oligomérisation de P (Mason et al., 2003 ; Tran et al., 2009 ; Blondot et al., 2012 ; Richard et al., 2018 ; Selvaraj et al., 2018). .... 33

**Figure 10 : Structure cristallographique de hRSV-M2-1.** Un protomère de M2-1 est représenté avec des régions colorées de la manière suivante : domaine de liaison au zinc (orange) avec son zinc (sphère grise) ; hélice de tétramérisation (magenta) ; domaine core (bleu) ; domaine non structuré dans la plupart des molécules (gris) ; résidus du vecteur de clonage (noir). Les extrémités N- et C-terminales sont marquées avec des sphères bleues et rouges, respectivement. Figure issue de Tanner et al., 2014. .... 34

**Figure 11 : Structure cristallographique de PP1α.** La surface de la phosphatase est représentée en gris partiellement transparent afin d'apercevoir les structures secondaires de la protéine avec le code couleur suivant : hélice α en rouge, brin β en jaune, boucle en vert. Les 2 sphères correspondent à 2 ions manganèse appartenant au site catalytique de la protéine. Les sites de fixation de substrats et le site catalytique sont indiqués. La structure est issue du fichier pdb 3E7A et la figure a été générée à l'aide du logiciel Pymol (Schrödinger). .... 36

**Figure 12 : Schéma des domaines de Tax1BP1 et structures connues des domaines.** La répartition des domaines de Tax1BP1 et leurs bornes sont indiquées sur le schéma. En dessous, sont représentées les structures déjà obtenues des domaines SKICH et UBZ1+2 (Ceregido et al., 2014 ; Yang et al., 2015). .... 37

**Figure 13 : Schéma de construction des protéines recombinantes.** Les extrémités N- et C-terminales sont indiquées de part et d'autre. En rouge est indiqué le site de coupure à la thrombine et le trait noir représente l'extension (GS)<sub>n</sub>. .... 40

**Figure 14 : Principe de l'attribution séquentielle des fréquences de résonance du squelette peptidique d'une protéine comportant un enchaînement Ala-Ser-Gly (<http://www.protein-nmr.org.uk>).**..... 46

**Figure 15 : Validation de l'interaction directe entre TAX1BP1 et N par GST pulldown et immunoprécipitation.** (A) Analyse sur gel acrylamide 12% coloré en bleu de Coomassie des protéines GST, GST-TAXBP1, et N recombinantes, et des échantillons correspondant au test de pulldown de N en présence de GST ou GST-TAX1BP1. (B) Des lysats cellulaires Hep2 infectés par le RSV-mCherry pendant 48h ont été incubés en présence de billes GST seule ou GST-TAX1BP1 sur la nuit à 4°C. La présence de N associée aux billes a été analysée par western blot avec un anticorps anti-N. (M : Marqueur de poids moléculaire). A noter que la bande correspondant à la protéine GST est une détection non spécifique par les anticorps. (C) Les cellules ont été co-transfectées avec les plasmides codant pour pFLAG-TAX1BP1 et pcDNA-N. Une immunoprécipitation en présence d'anticorps anti-FLAG a ensuite été réalisée sur les lysats cellulaires. Les protéines TAX1BP1 et N ont été révélées par western blot en présence d'anticorps anti-FLAG-HRP et anti-N. (Rapport de stage de M2 Sarah MADRIERES) ..... 54

**Figure 16 : Etude de l'impact de Tax1BP1 sur l'activité polymérase.** (A) Principe du test fonctionnel de minigénome pour mesurer l'activité du complexe polymérase : l'activité luciférase est proportionnelle à l'activité du complexe polymérase et quantifiée. (B) Test minigénome en présence de concentrations croissantes de pFlag-Tax1BP1 ou pFlag. L'activité luciférase a été quantifiée 24h post-transfection. (C) Analyse de l'expression de Tax1BP1 et N par western blot. (Rapport de stage de M2 Vanny KOUBANA) ..... 55

**Figure 17 : Etude in vivo de la déplétion de TAX1BP1 sur l'infection par le RSV chez la souris.** (A) Des poumons de souris sauvages ou TAX1BP1 <sup>-/-</sup> infectées par le RSV-luciférase ont été lysés et l'activité luciférase a été quantifiée (B, C). Quantification de l'expression des gènes de l'autophagie à partir de cellules (majoritairement des macrophages) issues de lavages broncho-alvéolaires de souris sauvages et TAX1BP1<sup>-/-</sup> infectées. L'expression des gènes codant pour des protéines impliquées dans l'autophagie, LC3 et ATG5 a été quantifiée par qPCR. Les mesures ont été réalisées en duplicats. Les résultats ont été normalisés par rapport au gène de ménage Hprt en calculant le Ct (Ct moyen (gène étudié) – Ct moyen (Hprt)). Les moyennes sont représentées en histogramme. Le nombre d'animaux testés est n = 4 souris sauvages infectées par le RSV et n = 2 souris TAX1BP1<sup>-/-</sup> infectées par le RSV. Ct correspond à l'acronyme cycle threshold (seuil de cycle) pour chaque gène. Cette valeur correspond au nombre de cycles nécessaires de qPCR pour que la quantité d'ADNc présente dans l'échantillon dépasse un seuil fixé manuellement. (Rapport de stage de M2 Sarah MADRIERES) ..... 56

**Figure 18 : Identification du domaine d'interaction de la protéine Tax1BP1 avec la protéine N par GST pulldown.** (A) Les protéines recombinantes GST-fragments de Tax1BP1 purifiés sur billes ont été incubés en présence de N entière ou N $\Delta$ 30 pendant 4 h. Les échantillons ont été analysés sur un

gel d'acrylamide coloré au bleu de Coomassie. (B) Les cellules HEP2 ont été infectées par le RSV (MOI=2 pendant 48 h). Les lysats des cellules non infectées (NI) et infectées (I) ont été incubés avec les billes de GST-fragments de Tax1BP1. La protéine N a été révélée par western blot avec un anticorps anti-protéine N. (Rapport de stage de M2 Vanny KOUBANA) ..... 57

**Figure 19 : Identification du domaine d'interaction de la protéine Tax1BP1 avec la protéine N par immunoprécipitation avec une étiquette Flag.** Les fragments de Tax1BP1 ont été clonés dans un vecteur pFlag permettant une expression en cellules eucaryotes. Immunoprécipitation de la protéine N dans un contexte de surexpression et des fragments de Tax1BP1. Les produits de l'expérience ont été analysés en utilisant un anticorps anti-N et un anticorps anti-Flag HRP. .... 58

**Figure 20 : Description des constructions des domaines UBZ de Tax1BP1 et C-terminal de N.** (A) Séquences des domaines Tax1BP1(UBZ1+2) et N(C arm). L'alignement de séquence des 2 domaines UBZ de Tax1BP1 a été réalisé avec Clustal Omega (Madeira et al., 2019) et préparé avec ESPRIPT 3.0 (Robert et Gouet, 2014). Les extensions restantes après coupure à la thrombine sont indiquées en violet. (B) SDS PAGE des deux domaines purifiés. Les marqueurs de tailles et leurs poids moléculaires sont indiqués sur les pistes de gauche en kDa..... 59

**Figure 21 : Spectres HSQC et BEST-TROSY  $^1\text{H}$ - $^{15}\text{N}$  attribués de  $^{13}\text{C}^{15}\text{N}$ -N(C arm).** A gauche est représenté le spectre HSQC  $^1\text{H}$ - $^{15}\text{N}$  attribué de  $^{13}\text{C}^{15}\text{N}$ -N(C arm) à 1 mM dans un tampon phosphate 20 mM, NaCl 100 mM, 7,5%  $\text{D}_2\text{O}$ , TCEP 1 mM, pH 6,5). A droite est montré le spectre TROSY attribué et replié de  $^{13}\text{C}^{15}\text{N}$ -N(C arm) à 100  $\mu\text{M}$  dans le même tampon à pH 6,9. Tous les deux ont été enregistrés à 600 MHz et à 298 K (25°C). Les corrélations  $^1\text{H}^{15}\text{N}$  du squelette peptidique ont été indiquées par le numéro et le nom du résidu auxquelles elles appartiennent. Les traits noirs sur le spectre de gauche représentent les paires de protons  $^1\text{H}^{\text{N}}$  portées par un même azote sur les chaînes latérales d'asparagines ou de glutamines. Les lignes de couleur magenta sur le spectre de droite représentent les limites de la largeur spectrale réellement enregistrée ..... 60

**Figure 22 : Prédiction de structures secondaires de chaque résidu de N(C arm).** Le calcul des probabilités de structures secondaire a été réalisé à partir du logiciel TALOS-N en ligne avec les déplacements chimiques  $^1\text{H}_{\text{N}}$ ,  $^{15}\text{N}_{\text{H}}$ ,  $^{13}\text{C}_{\alpha}$ ,  $^{13}\text{C}_{\beta}$  et  $^{13}\text{C}'$  et  $^1\text{H}_{\alpha}$  en utilisant  $^{13}\text{C}^{15}\text{N}$ -N(C arm) à 1 mM dans le tampon phosphate 20 mM, NaCl 100 mM, 7,5%  $\text{D}_2\text{O}$ , TCEP 1 mM, pH 6,5. En bas du graphique, est représenté un schéma théorique de la structure secondaire de la protéine. .... 62

**Figure 23: Paramètre d'ordre de chaque résidus de N(C arm).** Le calcul du paramètre d'ordre a été fait à partir du logiciel TALOS-N selon de modèle de Wishart à partir des même données que celles utilisées pour la **Figure 22**. En bas du graphique, est représenté un schéma théorique de la structure secondaire de la protéine. .... 63

**Figure 24 : Alignement des structures du domaine C arm de la nucléoprotéine du RSV cristallisée.** L'alignement a été réalisé avec le logiciel Pymol à partir de la structure de 2 décimères de N en complexe avec de l'ARN (code PDB : 2WJ8) et le résultat est donné avec les 20 structures partielles de N(C arm) présentes dans l'unité asymétrique en représentation cartoon. En bleu marine et en rouge

sont indiquées respectivement les extrémités N et C-terminale. Les premiers et derniers résidus des 2 hélices  $\alpha$  représentées sont indiqués par leur numéro et leur code à une lettre. .... 64

**Figure 25 : Effets de la concentration de N(C arm) sur ses déplacements chimiques  $^{13}\text{C}\alpha$  et  $^{13}\text{C}\beta$ .** Les histogrammes en barres vertes et marrons ont été fait à partir des différences de perturbations selon l'équation  $\Delta\delta^{13}\text{C} = \delta^{13}\text{C}(2 \text{ mM}) - \delta^{13}\text{C}(0,5 \text{ mM})$ . L'histogramme en barres oranges correspond à la soustraction des différences  $\Delta\delta^{13}\text{C}_{\alpha-\beta} = \Delta\delta^{13}\text{C}_{\alpha} - \Delta\delta^{13}\text{C}_{\beta}$ . Les lignes en pointillés noirs représentent la moyenne plus (pour tous les graphiques) et moins (pour  $^{13}\text{C}\beta$ ) l'écart-type. En bas du graphique, est représenté le schéma théorique de la structure secondaire de la protéine. .... 65

**Figure 26 : Différence de probabilité de former une hélice  $\alpha$  en fonction des résidus de N(C arm) à deux concentrations différentes.** Ce calcul a été fait à partir des probabilités de former une structure secondaire prédit par le logiciel Talos-N en ligne avec les déplacements chimiques  $^1\text{H}_\text{N}$ ,  $^{15}\text{N}$ ,  $^{13}\text{C}_\alpha$  et  $^{13}\text{C}_\beta$  attribués en utilisant le domaine  $^{13}\text{C}^{15}\text{N}$ -N(C arm) à 2 mM et à 0,5 mM dans le tampon phosphate 20 mM, NaCl 100 mM, 7,5%  $\text{D}_2\text{O}$ , TCEP 1 mM, pH 6,5. En bas du graphique, est représenté le schéma théorique de la structure secondaire de la protéine. .... 66

**Figure 27 : Représentation « cartoon » des 20 conformères finaux de N(C arm) calculés par CYANA.** En bleu et en rouge sont représentées respectivement les extrémités N- et C-terminales. L'hélice  $\alpha$  des résidus 372 à 380 ainsi que 2 hélices transitoires en C-terminal sont signalées. .... 69

**Figure 28 : Représentation sous forme « bâtonnets » des chaînes latérales de l'hélice  $\alpha$  des résidus 372 à 380.** En vert et en jaunes sont montrées respectivement les extrémités N- (jusqu'au résidu 371) et C-terminales (jusqu'au résidu 381). Les résidus dont le numéro et le code à 1 lettre sont indiqués correspondent à ceux qui ont une orientation semblable ou presque. .... 69

**Figure 29 : Spectres HSQC  $^1\text{H}$ - $^{15}\text{N}$  de  $^{15}\text{N}$  Tax1BP1(UBZ1+2).** Celui du haut a été enregistré à 800 MHz par Ceregido et al. (J. Mol. Biol., 2014) avec  $^{13}\text{C}^{15}\text{N}$  Tax1BP1(UBZ1+2) à 700  $\mu\text{M}$  dans un tampon phosphate salin (PBS) 50 mM, 7,5%  $\text{D}_2\text{O}$ , 1 mM TCEP à pH 6,0. Le spectre du bas, a été enregistré à 700 MHz par moi-même avec  $^{15}\text{N}$  Tax1BP1(UBZ1+2) concentrée à 100  $\mu\text{M}$  dans un tampon phosphate 20 mM, NaCl 100 mM, 7,5%  $\text{D}_2\text{O}$ , TCEP 1 mM à pH 6,5. Les 2 spectres ont été enregistrés à 298 K (25°C). Les corrélations  $^1\text{H}^{15}\text{N}$  du squelette peptidique ont été indiquées par le numéro du résidu auxquelles elles appartiennent. Les traits (bleu ou noir) sur les spectres représentent les paires de protons  $^1\text{H}^\text{N}$  portées par un même azote sur les chaînes latérales d'asparagines ou de glutamines. .... 71

**Figure 30 : Spectres BEST-TROSY  $^1\text{H}$ - $^{15}\text{N}$  superposés de  $^{13}\text{C}^{15}\text{N}$ -N(C arm) avec différentes concentrations de Tax1BP1(UBZ1+2).** Le spectre de couleur magenta est celui de  $^{13}\text{C}^{15}\text{N}$ -N(C arm) seul et concentré à 100  $\mu\text{M}$  ; il correspond au spectre de référence. Les spectres vert et mauve sont les spectres issus du titrage de  $^{13}\text{C}^{15}\text{N}$ -N(C arm) par Tax1BP1(UBZ1+2) correspondant respectivement aux ratios molaires 1:19 et 1:33 (N(C arm):Tax1BP1(UBZ1+2)). Les numéros et codes à une lettre des résidus qui ont subi le plus de perturbation de leur déplacement chimique sont indiqués. Tous les spectres ont été enregistrés à 298 K (25°C) et à 600 MHz et le tampon utilisé était du phosphate 20 mM, NaCl 100 mM, 7,5%  $\text{D}_2\text{O}$ , TCEP 1 mM à pH 6,9. Les résidus qui semblent être présents plusieurs fois sur le

spectre de manière périodique sont le résultat du dépliement de la fenêtre spectrale dans la dimension  $^{15}\text{N}$  ..... 73

**Figure 31 : Variations des déplacements chimiques  $^1\text{H}$ ,  $^{15}\text{N}$ ,  $^1\text{H}^{15}\text{N}$ , et des intensités des signaux HN par résidu de N(C arm) en présence de 33 équivalents de Tax1BP1(UBZ1+2).** Les barres d'erreurs pour les calculs de perturbation de déplacements chimiques ont été produites à partir d'additions de la résolution spectrale des expériences qui ont servi à l'extraction des données. Les barres d'erreur des ratios d'intensités ont été calculées à partir du bruit des spectres HSQC  $^1\text{H}$ - $^{15}\text{N}$  (erreur =  $(I*\text{bruit}_0+I_0*\text{bruit}) / I_0^2$ ). Les lignes en pointillés noirs représentent la moyenne plus (pour  $^1\text{H}$ ,  $^{15}\text{N}$  et  $^1\text{H}^{15}\text{N}$ ) ou moins (pour  $^1\text{H}$ ,  $^{15}\text{N}$  et  $I/I_0$ ) l'écart-type. En bas du graphique, est représenté le schéma théorique de la structure secondaire de la protéine..... 74

**Figure 32 : Perturbations des déplacements chimiques  $^1\text{H}^{\text{N}}$  de N(C arm) en fonction du nombre d'équivalent en peptide Tax1BP1(UBZ1+2).** Seuls les résidus dont le déplacement chimique a été décalé d'au moins la valeur de la moyenne des déplacements chimiques plus ou moins un écart-type (en présence de 33 équivalents de Tax1BP1(UBZ1+2)) sont montrés. Le nom et le numéro des résidus sélectionnés sont écrits en titre de graphique. Les barres d'erreurs pour les calculs de perturbation de déplacements chimiques ont été produites à partir d'additions de la résolution spectrale des expériences qui ont servi à l'extraction des données..... 75

**Figure 33 : Perturbations des déplacements chimiques  $^{15}\text{N}$  de N(C arm) en fonction du nombre d'équivalent en peptide Tax1BP1(UBZ1+2).** Seuls les résidus dont le déplacement chimique a été décalé d'au moins la valeur de la moyenne des déplacements chimiques plus ou moins un écart-type (en présence de 33 équivalents de Tax1BP1(UBZ1+2)) sont montrés. Le nom et le numéro des résidus sélectionnés sont écrits en titre de graphique. Les barres d'erreurs pour les calculs de perturbation de déplacements chimiques ont été produites à partir d'additions de la résolution spectrale des expériences qui ont servi à l'extraction des données..... 76

**Figure 34 : Perturbations des déplacements chimiques combinés  $^1\text{H}^{15}\text{N}$  de N(C arm) en fonction du nombre d'équivalent en Tax1BP1(UBZ1+2).** Seuls les résidus dont le déplacement chimique a été décalé d'au moins la valeur de la moyenne des déplacements chimiques plus un écart-type (en présence de 33 équivalents de Tax1BP1(UBZ1+2)) sont montrés. Le nom et le numéro des résidus sélectionnés sont écrits en titre de graphique. Les barres d'erreurs pour les calculs de perturbation de déplacements chimiques ont été produites à partir de la résolution spectrale des expériences qui ont servi à l'extraction des données..... 77

**Figure 35 : Perturbations des déplacements chimiques  $^{13}\text{C}$   $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$  et  $\epsilon$  par résidu de N(C arm) en présence de 30 équivalents de Tax1BP1(UBZ1+2).** Les lignes en pointillés noirs représentent la moyenne plus ou moins l'écart-type de toutes les données. Les spectres HSQC  $^1\text{H}$ - $^{13}\text{C}$  ayant servi à l'extraction des données ont été enregistrés à 298 K (25°C) et à 600 MHz. Les protéines étaient dans un tampon phosphate 20 mM, NaCl 100 mM, 7,5%  $\text{D}_2\text{O}$ , TCEP 1 mM à pH 6,5..... 78

**Figure 36 : Variations des ratios d'intensités des signaux HN par résidu de N(C arm) en présence de 33 équivalents de Tax1BP1(UBZ1+2) ou en présence de 5 équivalents de BSA.** Les barres d'erreur des ratios d'intensités ont été calculées à partir du bruit des spectres HSQC  $^1\text{H}$ - $^{15}\text{N}$  (erreur =  $(I^*\text{bruit}_0+I_0*\text{bruit}) / I_0^2$ ). Les lignes en pointillés noirs représentent la moyenne moins l'écart-type. 79

**Figure 37 : Représentation des résidus les plus affectés en termes de déplacements chimiques combinés HN sur la structure de N(C arm) par Tax1BP1(UBZ1+2).** Les résidus dont les déplacements chimiques combinés HN sont au-dessus de 1 moyenne + 1 écart-type et ceux seulement au-dessus de 1 moyenne sont montrés en violet et en saumon respectivement..... 80

**Figure 38 : Spectres HSQC  $^1\text{H}$ - $^{15}\text{N}$  superposés de  $^{15}\text{N}$ -Tax1BP1(UBZ1+2) à 100  $\mu\text{M}$  seul et avec 45 équivalents de N(C arm).** Le spectre de couleur noire est celui de  $^{15}\text{N}$ -Tax1BP1(UBZ1+2) seul et concentré à 100  $\mu\text{M}$  ; il correspond au spectre de référence. Le spectre rouge est issu du titrage de  $^{15}\text{N}$ -Tax1BP1(UBZ1+2) par N(C arm) avec, ici, un ratio 1:45 (Tax1BP1(UBZ1+2) : N(C arm)). Les spectres ont été enregistrés à 298 K (25°C) et à 700 MHz et le tampon utilisé était du phosphate 50 mM, NaCl 100 mM, 7,5%  $\text{D}_2\text{O}$ , TCEP 1 mM à pH 6,5. Les numéros et codes à une lettre des résidus qui ont varié le plus par rapport à leurs déplacements chimiques sont indiqués. .... 81

**Figure 39 : Variations des déplacements chimiques  $^1\text{H}$ ,  $^{15}\text{N}$ ,  $^1\text{H}^{15}\text{N}$ , et des intensités des signaux HN par résidu de Tax1BP1(UBZ1+2) en présence de 45 équivalents de N(C arm).** Les barres d'erreurs pour les calculs de perturbation de déplacements chimiques ont été produites à partir d'additions de la résolution spectrale des expériences qui ont servi à l'extraction des données. Les barres d'erreur des ratios d'intensités ont été calculées à partir du bruit des spectres HSQC  $^1\text{H}$ - $^{15}\text{N}$  (erreur =  $(I^*\text{bruit}_0+I_0*\text{bruit}) / I_0^2$ ). Les lignes en pointillés noirs représentent la moyenne plus (pour  $^{15}\text{N}$   $^1\text{H}^{15}\text{N}$ ) ou moins (pour  $^1\text{H}$ ,  $^{15}\text{N}$  et  $I/I_0$ ) l'écart-type. .... 82

**Figure 40 : Perturbations des déplacements chimiques  $^1\text{H}^{\text{N}}$  de Tax1BP1(UBZ1+2) en fonction du nombre d'équivalent en peptide N(C arm).** Seuls les résidus dont le déplacement chimique a été décalé d'au moins la valeur de la moyenne des déplacements chimiques plus un écart-type (en présence de 45 équivalents de N(C arm)) sont montrés. Le nom et le numéro des résidus sélectionnés sont écrits en titre de graphique. Les barres d'erreurs pour les calculs de perturbation de déplacements chimiques ont été produites à partir d'additions de la résolution spectrale des expériences qui ont servis à l'extraction des données. .... 84

**Figure 41 : Perturbations des déplacements chimiques  $^{15}\text{N}$  de Tax1BP1(UBZ1+2) en fonction du nombre d'équivalent en peptide N(C arm).** Seuls les résidus dont le déplacement chimique a été décalé d'au moins la valeur de la moyenne des déplacements chimiques moins un écart-type (en présence de 45 équivalents de N(C arm)) sont montrés. Le nom et le numéro des résidus sélectionnés sont écrits en titre de graphique. Les barres d'erreurs pour les calculs de perturbation de déplacements chimiques ont été produites à partir d'additions de la résolution spectrale des expériences qui ont servis à l'extraction des données. .... 85

**Figure 42 : Perturbations des déplacements chimiques combinés  $^1\text{H}^{15}\text{N}$  de Tax1BP1(UBZ1+2) en fonction du nombre d'équivalent en peptide N(C arm).** Seuls les résidus dont les déplacements chimiques ont été décalés d'au moins la valeur de la moyenne des déplacements chimiques plus un écart-type (en présence de 45 équivalents de N(C arm)) sont montrés. Le nom et le numéro des résidus sélectionnés sont écrits en titre de graphique. Les barres d'erreurs pour les calculs de perturbation de déplacements chimiques ont été produites à partir de la résolution spectrale des expériences qui ont servis à l'extraction des données. .... 86

**Figure 43 : Variations des ratios d'intensités des signaux  $^1\text{H}^{15}\text{N}$  par résidu de Tax1BP1(UBZ1+2) en présence de 45 équivalents de N(C arm) ou en présence de 5 équivalents de BSA.** Les barres d'erreur des ratios d'intensités ont été calculées à partir du bruit des spectres HSQC  $^1\text{H}^{15}\text{N}$  (erreur =  $(I \cdot \text{bruit}_0 + I_0 \cdot \text{bruit}) / I_0^2$ ). Les lignes en pointillés noirs représentent la moyenne moins l'écart-type..... 87

**Figure 44 : Structure de Tax1BP1(UBZ1+2) avec les résidus les plus affectés par la présence de N(C arm).** La structure représentée est un des conformères du domaine UBZ1+2 calculée à partir des données RMN de Ceregido et al (pdb 2M7Q) ; la représentation est montrée sans l'étiquette GPHM en N-terminal du peptide. (A) et (B) : Représentations en « cartoon » du peptide. Les couleurs cyan et orange clair indique respectivement les domaines N- (UBZ1) et C-terminaux (UBZ2). Les sphères jaunes montrent les atomes de zinc chélatés par des systèmes tétraédriques de type  $\text{Cis}_2\text{-His}_2$  dont les chaînes latérales sont représentées en bâtonnets jaunes. Les sphères violettes correspondent aux atomes d'azote de la chaîne peptidique dont la perturbation de déplacement chimique combinée HN est la plus affectée (les numérotations et noms des résidus en question apparaissent de la même couleur). Les bâtonnets violets indiquent les chaînes latérales impliquées dans le cœur hydrophobe et pour lesquels la perturbation de déplacement chimique combinée HN est la plus affectée. (C) et (D) : Représentations de la structure en « surface » avec le même code couleur que (A) et (B). (E) et (F) : Représentations de la structure sous forme de « surface électrostatique » avec en rouge les parties chargées plutôt négativement, en bleu les régions plutôt positivement chargées et en blanc les zones plutôt neutres. Les parties (B), (D) et (F) sont simplement une vue de dos des figures de gauche. .... 88

**Figure 45 : Alignement de séquences des domaines UBZ 1 et 2 de Tax1BP1 montrant les résidus dont les déplacements chimiques combinés  $^1\text{H}^{15}\text{N}$  sont les plus perturbés en présence de 45 équivalents de N(C arm).** Les résidus avec un triangle violet au-dessus ou en-dessous sont ceux dont les déplacements chimiques combinés  $^1\text{H}^{15}\text{N}$  sont les plus perturbés. L'alignement de séquence a été réalisé avec Clustal Omega et préparé avec ESPRIPT 3.0. .... 89

**Figure 46 : Gels des expériences de GST pulldown de GST-Tax1BP1(UBZ1+2)—N(C arm) et GST- N(C arm)—Tax1BP1(UBZ1+2).** Les expériences ont été faites en mélangeant 100  $\mu\text{M}$  de billes sépharose-GSH-GST-\* (pour 50  $\mu\text{L}$  total) avec soit du tampon phosphate pH 6,5 ou plusieurs équivalents molaires de N(C<sub>arm</sub>) ou UBZ1+2 ainsi qu'une quantité suffisante de tampon pour 50  $\mu\text{L}$

totaux. Les pistes avec Tax1BP1(UBZ1+2) ou N(C arm) seules contiennent l'équivalent de 100  $\mu$ M de protéines. .... 91

**Figure 47 : Réplicats des expériences de GST pulldown de GST-Tax1BP1(UBZ1+2)—N(C arm) et GST- N(C arm)—Tax1BP1(UBZ1+2).** Les expériences ont été faites en mélangeant 100  $\mu$ M de billes sépharose-GSH-GST-\* (pour 50  $\mu$ L total) avec soit du tampon phosphate pH 6,5 ou plusieurs équivalents molaires de N(C<sub>arm</sub>) ou UBZ1+2 ainsi qu'une quantité suffisante de tampon pour 50  $\mu$ L totaux. Les pistes avec Tax1BP1(UBZ1+2) ou N(C arm) seules contiennent l'équivalent de 100  $\mu$ M de protéines. Ici, 5 équivalents molaires des protéines proies ont été déposées avec les billes GST\*, il y a eu seulement 2 lavages des billes et 5 fois plus d'échantillons ont été déposés sur gel. .... 92

**Figure 48 : Réplicats des expériences de GST pulldown de GST-Tax1BP1(UBZ1+2)—N(C arm) et GST- N(C arm)—Tax1BP1(UBZ1+2).** Les expériences ont été faites en mélangeant 100  $\mu$ M de billes sépharose-GSH-GST-\* (pour 50  $\mu$ L total) avec soit du tampon 20 mM tris, 150 mM NaCl, pH 8,5 ou plusieurs équivalents molaires de N(C<sub>arm</sub>) ou UBZ1+2 ainsi qu'une quantité suffisante de tampon pour 50  $\mu$ L totaux. Les pistes avec Tax1BP1(UBZ1+2) ou N(C arm) seules contiennent l'équivalent de 100  $\mu$ M de protéines. Ici, 5 équivalents molaires des protéines proies ont été déposées avec les billes GST\*, il y a eu seulement 2 lavages des billes et 5 fois plus d'échantillons ont été déposés sur gel. .... 93

**Figure 49 : Bornes pour les fragments de P et leurs séquences en acides aminés.** A) Schéma des différents fragments. Les régions absentes sont en pointillés. A droite sont indiqués les poids moléculaires des différents fragments et leurs points isoélectriques (pI). B, Séquence en acides aminés de P entière et indication des emplacements des fragments de P représentés par des lignes colorées au-dessus de la séquence primaire. .... 96

**Figure 50 : Analyse par SDS PAGE de la protéine P entière et de différentes constructions de P purifiées.** A gauche des pistes de chaque gel sont indiquées les tailles des marqueurs de poids moléculaires. .... 97

**Figure 51 : Spectres HSQC <sup>1</sup>H-<sup>15</sup>N de la phosphoprotéine du RSV marquée <sup>15</sup>N.** (A) Représentation schématique des limites des fragments de protéine P utilisés en RMN. Les régions supprimées sont symbolisées par des cadres en pointillés. Les zones hachurées indiquent les régions de la protéine avec des attributions amides manquantes. (B) Les spectres HSQC <sup>1</sup>H-<sup>15</sup>N de P<sub>FL</sub>, P<sub>ND+OD</sub>, P<sub>OD+CD</sub>, P<sub>ND</sub>, P<sub>CD</sub>, et P <sub>$\Delta$ OD</sub> ont été acquis dans des conditions expérimentales identiques (50–100  $\mu$ M, 288 K, 14,1 T). Une couleur différente a été utilisée avec chaque construction pour représenter les contours des signaux et pour montrer les attributions. Pour tous les mutants de délétions, le spectre est superposé avec celui de P<sub>FL</sub> pour comparer. Les attributions de résonance amide sont indiquées pour P<sub>ND+OD</sub>, P<sub>OD+CD</sub>, P<sub>ND</sub> et P<sub>CD</sub> (numéro de résidu et code à 1 lettre). Les signaux des NH<sub>2</sub> des chaînes latérales des asparagines et glutamines ne sont pas individuellement attribués. Pour P<sub>ND</sub> le cadre isolé en haut à gauche montre les attributions de la région centrale du spectre contenant un amas de signaux. .... 99

**Figure 52 : Caractérisation de l'hétérogénéité structurale de P par analyse des déplacements chimiques du squelette peptidique et expériences de relaxation  $^{15}\text{N}$ .** (A) Diagrammes empilés qui représentent les propensions de structures secondaires (hélices  $\alpha$ , brins  $\beta$ , pelote statistique) déterminées à partir des déplacements chimiques  $^1\text{H}_\text{N}$ ,  $^{15}\text{N}$ ,  $^1\text{H}_\alpha$ ,  $^{13}\text{C}'$ ,  $^{13}\text{C}_\alpha$  et  $^{13}\text{C}_\beta$  avec le logiciel Talos+ sur différentes constructions de P. Des régions notables sont signalées par des couleurs de fonds. Les régions hachurées correspondent aux régions de la protéine où les signaux amides sont absents. (B) Organisation structurale de hRSV-P en lien avec les résultats montrés dans la partie A. Les hélices  $\alpha$  et le brin  $\beta$  transitoires sont représentés par des rectangles ayant le même code couleur que dans la partie A. La propension maximale des structures secondaires est indiquée à l'intérieur des rectangles. (C) Les temps de relaxation  $^{15}\text{N-R}_1$ ,  $^{15}\text{N-R}_2$  et NOE hétéronucléaires  $^1\text{H-}^{15}\text{N}$  ont été mesurés pour 4 fragments de P à 14.1 T et 288 K. Les barres d'erreur pour chaque point correspondent aux valeurs d'écart types calculées à partir des différences entre les modèles d'ajustement mono-exponentiel et les jeux de données expérimentaux correspondant. Les lignes horizontales en pointillés correspondent aux valeurs moyennes en tenant uniquement en compte des résidus désordonnés. La localisation des régions en hélices  $\alpha$  et du brin  $\beta$  instables ainsi que du domaine d'oligomérisation est indiquée avec le même code couleur que dans les parties A et B. La région hachurée sur les graphiques correspond aux résidus dont les signaux  $^1\text{H-}^{15}\text{N}$  sont absents sur les spectres RMN correspondant; tandis que le rectangle grisé entouré de pointillés indique la séquence supprimée dans le fragment P $\Delta$ OD. .... 101

**Figure 53 : Caractérisation de l'hétérogénéité structurale de hRSV-P par analyse d'expériences de relaxation  $^{15}\text{N}$  du squelette peptidique.** Les temps de relaxation  $^{15}\text{N-R}_1$  et  $^{15}\text{N-R}_2$  et les NOE hétéronucléaires  $^1\text{H-}^{15}\text{N}$  ont été mesurés pour la protéine P entière à un champ magnétique de 14,1 T et à 288 K. Les barres d'erreur pour chaque point correspondent aux valeurs d'écart types calculées à partir des différences entre les modèles d'ajustement mono-exponentiel et les jeux de données expérimentaux correspondant. Les lignes horizontales en pointillés correspondent aux valeurs moyennes en tenant uniquement en compte des résidus désordonnés. La localisation des régions en hélices  $\alpha$  et du brin  $\beta$  instables ainsi que du domaine d'oligomérisation est indiquée au-dessus des graphiques. La région hachurée sur les graphiques correspond aux résidus dont les signaux  $^1\text{H-}^{15}\text{N}$  sont absents sur les spectres RMN correspondant. .... 102

**Figure 54 : Mise en évidence de contacts à longue distance dans la protéine P entière par des expériences d'augmentation de la relaxation par effet paramagnétique (PRE).** Les PREs ont été mesurés avec un marquage radicalaire de type nitroxyde aux positions 23, 99, 143, 156, 179, 193 et 237 dans la séquence de P. Les zones hachurées indiquent les régions sans signal amide détectable dans l'état diamagnétique. Les barres d'erreur représentent les déviations quadratiques moyennes calculées avec deux jeux de données. Les positions de marquage de spin sont indiquées par des flèches pour chaque diagramme en barres. Les éléments de structure secondaire structure transitoire sont indiqués par des fond colorés. La figure est tirée de l'article de Pereira et al., 2016..... 105

**Figure 55 : Spectre HSQC  $^1\text{H}$ - $^{15}\text{N}$  de P[127-205].** Le spectre a été enregistré sur un spectromètre RMN à 14.1 T à une température de 313 K (40°C). Les corrélations  $^1\text{H}$ - $^{15}\text{N}$  du squelette peptidique ont été indiquées par le numéro et la lettre du résidu auxquelles elles appartiennent. La ligne en trait pointillé sur le spectre représente la région des paires de protons  $^1\text{H}_\text{N}$  portées par les mêmes azotes sur les 2 chaînes latérales d'asparagines..... 107

**Figure 56 : Spectre HSQC  $^1\text{H}$ - $^{15}\text{N}$  de P<sub>OD</sub>.** Le spectre a été enregistré sur un spectromètre RMN à 14.1 T à 298 K (25°C). Les corrélations  $^1\text{H}$ - $^{15}\text{N}$  du squelette peptidique ont été indiquées par le numéro et la lettre du résidu auxquelles elles appartiennent. La ligne en trait pointillé sur le spectre représente la région des paires de protons  $^1\text{H}_\text{N}$  portées par les mêmes azotes sur les chaînes latérales d'asparagines. .... 109

**Figure 57 : Prédiction de structures secondaires pour P[127-205] et P<sub>OD</sub>.** Le calcul des probabilités de structures secondaires a été réalisé en utilisant le serveur TALOS-N (<https://spin.niddk.nih.gov/bax/software/TALOS-N/>) avec les déplacements chimiques  $^1\text{H}_\text{N}$ ,  $^{15}\text{N}_\text{H}$ ,  $^{13}\text{C}_\alpha$ ,  $^{13}\text{C}_\beta$ ,  $^{13}\text{C}'$  et  $^1\text{H}_\alpha$  des 2 fragments. (A) Diagramme empilé représentant les propensions de structures secondaires de P[127-205] à 313 K. Au-dessus du graphique est représenté le schéma des structures secondaires de la protéine selon Pereira et al., 2016. (B) Diagramme empilé représentant les propensions de structures secondaires de P<sub>OD</sub> à 298 K..... 111

**Figure 58 : Résultat du calcul de structure de CYANA d'un protomère de P<sub>OD</sub>.** L'image représente l'ensemble des 20 conformères alignés en représentation cartoon. A gauche de l'image se situe l'extrémité N-terminale et à droite l'extrémité C-terminale. .... 116

**Figure 59 : Superposition de spectres HSQC  $^1\text{H}$ - $^{13}\text{C}$  de tétramères entièrement marqués  $^{13}\text{C}^{15}\text{N}$  et de tétramères mixtes obtenus avec P<sub>OD</sub>\* et P<sub>OD</sub>.** Les spectres ont été acquis dans un tampon 20 mM phosphate, 100 mM NaCl dans D<sub>2</sub>O avec P<sub>OD</sub>/P<sub>OD</sub>\* concentré à 600  $\mu\text{M}$  à un champ magnétique de 22,3 T et à une température de 298 K..... 117

**Figure 60 : Structure du domaine d'oligomérisation de hMPV-P.** A, Alignement des séquences primaires de hRSV-P avec celle de hMPV-P à partir du logiciel Clustal Omega et préparé avec ESPRIPT 3.0. La position de l'hélice  $\alpha$  prédite par TALOS-N pour hRSV-P<sub>OD</sub> et la position de l'hélice  $\alpha$  pour le domaine de tétramérisation de hMPV-P sont indiquées en haut et en bas respectivement. B, La structure du domaine d'oligomérisation de hMPV-P obtenue par diffraction des rayons X est représentée avec vue du côté C-terminal en gris et avec vue latérale des faisceaux d'hélices en couleur arc-en-ciel, avec les extrémités C- et N-terminales en rouge et bleu respectivement (figure issue des travaux de Leyrat et al., 2013). .... 118

**Figure 61 : Structure obtenue par ARIA avec un jeu de données partiel et utilisé comme structure initiale pour un deuxième recuit simulé du tétramère de P<sub>OD</sub>.** Les extrémités N- et C-terminales sont indiquées de part et d'autre de la représentation cartoon du tétramère. .... 119

**Figure 62 : Résultat du calcul de structure de ARIA du tétramère de P<sub>OD</sub>.** L'image représente l'ensemble des 20 conformères alignés en représentation « cartoon ». A gauche de l'image se situe

l'extrémité N-terminale et à droite l'extrémité C-terminale. Les flèches noires indiquent une contorsion locale des hélices autour du résidu E144..... 121

**Figure 63 : Superpositions des chaînes latérales du tétramère de P<sub>OD</sub>.** Seules les chaînes latérales des parties structurées en hélice (P[129-156]) sont indiquées. .... 122

**Figure 64 : Représentation du cœur hydrophobe du tétramère de P<sub>OD</sub>.** A, Mise en valeur des résidus hydrophobes en gris. B, Représentation sous forme de surface électrostatique d'un conformère du tétramère avec en rouge les régions chargées négativement, en blanc les zones hydrophobes et en bleu les régions chargées positivement. .... 122

**Figure 65 : Résultat du second calcul de structure de ARIA du tétramère de P<sub>OD</sub>.** Les 20 conformères alignés avec les chaînes latérales hydrophobes représentées en gris sont représentés sur les deux figures du haut. La seconde représentation de structure de P<sub>OD</sub> est montrée sans le protomère frontal qui était présent dans la structure du haut afin de visualiser le cœur hydrophobe. La structure au bas de la figure montre un tétramère sous forme de surface électrostatique. .... 125

**Figure 66 : Superposition de la structure du tétramère de P<sub>OD</sub> du hRSV avec celle du hMPV.** Les structures en vert correspondent à celles des conformères de hRSV-P<sub>OD</sub> calculés avec ARIA (2<sup>ème</sup> calcul) et en gris sont montrés les tétramères issus des structures de hMPV-P<sub>core</sub> (codes pdb : 5OIX et 5OIIY). .... 126

**Figure 67 : Spectres HSQC <sup>1</sup>H-<sup>15</sup>N des expériences de PRE avec les mutants cystéines de P[127-205].** Les PRE ont été mesurés après fixation de sondes radicalaires IAP aux positions 143 ou 156 sur les mutants cystéines de P[127-205]. Les expériences ont été réalisées sur les mutants P[127-205](S143C) et P[127-205](S156C) concentrés à 200 μM dans un tampon 20 mM phosphate pH 6,5, 100 mM NaCl, 7,5% D<sub>2</sub>O à 313 K et enregistrées sur un spectromètre RMN à 14,1 T avec cryosonde. Les positions de marquage sont indiquées par des flèches rouges sur les schémas de structures secondaires au-dessus de chaque spectre ..... 127

**Figure 68 : Histogrammes des intensités normalisées pour les signaux des mutants de P[127-205] comportant une sonde paramagnétique.** Les fonds colorés correspondent aux domaines et régions de la protéine signalés au-dessus des histogrammes. La flèche blanche indique la position de marquage par la sonde paramagnétique. La ligne pointillée correspond à la moyenne de toutes les intensités. Les barres d'erreur pour chaque intensité correspondent à l'écart type des ratios d'intensité en utilisant des réplicats des expériences. .... 128

**Figure 69 : Ajustement de données PRE de P[127-205] à partir de modèles générés par Flexible-meccano.** Les diagrammes représentent les mêmes données que celles de la **Figure 67** mais en ajoutant les courbes théoriques qui ont chacune été calculées à partir de 1000 conformères générés par le logiciel. En vert clair est indiquée une simulation sans aucun contact à longue distance. En vert foncé, en bleu et en violet sont montrées les courbes des simulations avec des contacts à 15, 10 et 7 Å de distance entre la région 143-156 et 160-199..... 129

<b>Figure 70 : Echantillon de 20 conformères de protomère de P[127-205] calculés par Flexible-meccano avec un contact de 7 Å entre P<sub>OD</sub> et <math>\alpha_{C1/2}</math>.</b> Les structures ont été sélectionnées aléatoirement parmi un ensemble de 1000 conformères calculés par le logiciel. Les conformères ont été alignés sur le domaine P <sub>OD</sub> (ici au niveau des résidus 130-156). En bleu est indiqué le N-terminal et en rouge, le C-terminal.....	130
<b>Figure 71 : SDS PAGE de <sup>15</sup>N-P et M2-1<sub>core</sub> purifiés.</b> Les marqueurs de tailles et leurs masses en kDa sont indiqués dans la piste la plus à gauche. ....	132
<b>Figure 72 : Spectre MALDI-TOF de la protéine P marquée <sup>15</sup>N.</b> La masse mono isotopique attendue de la protéine est indiquée en haut à droite. ....	133
<b>Figure 73 : Spectre MALDI-TOF du domaine core de la protéine M2-1<sub>core</sub>.</b> La masse mono isotopique attendue de la protéine est indiquée en haut à gauche. Les intervalles de numéros en gras au-dessus de certains pics indiquent les fragments de M2-1 auxquels appartiennent les massifs. La séquence primaire de la protéine est également indiquée. Sur cette dernière est indiqué en rouge le résidu restant d'extension de l'étiquette GST après coupure à la thrombine et en gras la région d'interaction avec la protéine P (Blondot et al., 2012). Le cadre en pointillé autour de la séquence indique la partie partiellement ou complètement absente du spectre de masse.....	134
<b>Figure 74 : Spectres HSQC <sup>1</sup>H-<sup>15</sup>N superposés de <sup>15</sup>N-P<sub>FL</sub> seule et en présence de M2-1<sub>core</sub> non marquée.</b> Les numéros indiqués correspondent aux résidus dont l'intensité de leur signal est la plus affectée. ....	135
<b>Figure 75 : Diagramme des intensités normalisées du spectre HSQC <sup>1</sup>H-<sup>15</sup>N de <sup>15</sup>N-P<sub>FL</sub> en présence de M2-1<sub>core</sub> non marquée.</b> Le diagramme représenté a été tracé à partir des données d'intensité des signaux des spectres de la Figure 73. Les lignes droites et pointillées représentent les valeurs moyennes et moyennes $\pm$ RMSD (moyenne quadratique) sur tous les signaux. La position du domaine d'oligomérisation et des régions présentant une structure secondaire transitoire en hélice $\alpha$ ou en brin $\beta$ le long de la séquence de P est indiquée dans le schéma au-dessus du graphique. Le fond rouge indique la région autour de l'hélice transitoire $\alpha_{N2}$ où M2-1 <sub>core</sub> induit une diminution de l'intensité >60%. La zone hachurée correspond à la région de P englobant le domaine OD et $\alpha_{C1/2}$ pour lesquelles les pics sont trop élargis et donc invisibles sur les spectres.....	135
<b>Figure 76 : Diagrammes des intensités normalisées des spectres HSQC <sup>1</sup>H-<sup>15</sup>N de <sup>15</sup>N-P<sub>ND</sub> et <sup>15</sup>N-P<sub>ND+OD</sub> en présence de M2-1<sub>core</sub> non marquée.</b> Les diagrammes ont été fait avec des spectres HSQC <sup>1</sup> H- <sup>15</sup> N de P <sub>ND</sub> à 75 $\mu$ M et P <sub>ND+OD</sub> 35 $\mu$ M marqués <sup>15</sup> N, seuls et en présence d'un équivalent molaire de M2-1 <sub>core</sub> . Les régions autour de $\alpha_{N2}$ avec des intensités normalisées inférieures à la moyenne quadratique sont mises en évidence avec des arrière-plans bleu et cyan pour P <sub>ND</sub> et P <sub>ND+OD</sub> respectivement.....	136
<b>Figure 77 : Roue hélicoïdale montrant les résidus de P critiques pour l'activité de transcription du RSV.</b> Représentation en roue hélicoïdale de l'hélice $\alpha$ putative située entre les résidus S94 et N111 de P (serveur HeliQuest). Les résidus critiques pour la liaison M2-1 sont indiqués par une	

étoile et le résidu phosphorylable T108 est indiqué par deux étoiles. Les résidus chargés positivement sont en bleu, les résidus chargés négativement en rouge, les résidus phosphorylés putatifs en violet et les résidus hydrophobes en jaune. .... 138

**Figure 78 : Modèle d'un complexe P—M2-1.** Le modèle a été obtenu en rapprochant l'hélice D95-F109 de P (en vert) sur un protomère de M2-1<sub>core</sub> (en gris, avec indication des numéros d'hélice). Les résidus de P et M2-1 critiques pour la fixation des deux sont indiqués (en orange pour M2-1). Les résidus impliqués dans le domaine de liaison à l'ARN au sein de M2-1 sont en rouge. .... 139

**Figure 79 : Spectres BEST-TROSY <sup>1</sup>H-<sup>15</sup>N de <sup>15</sup>N-P<sub>ND</sub>/<sup>15</sup>N-P avec GST-PP1 $\alpha$ /GST.** (Panneau supérieur) Spectres BEST-TROSY <sup>1</sup>H-<sup>15</sup>N superposés de <sup>15</sup>N-P<sub>ND</sub> ou de <sup>15</sup>N-P à 25  $\mu$ M seuls (contours magenta) et en présence 2 équivalents molaires de GST-PP1 $\alpha$  (contours noirs) ; les attributions sont également indiquées. (Panneau inférieur) Spectres HSQC <sup>1</sup>H-<sup>15</sup>N superposés de <sup>15</sup>N-P<sub>ND</sub> à 10  $\mu$ M seul (contours verts) et en présence de 10 équivalents molaires de GST (contours noirs). Le tampon de l'échantillon <sup>15</sup>N-P<sub>ND</sub> + GST contenait 25 mM de glutathion qui produisait des signaux d'abondance naturelle <sup>15</sup>N indiqués par des étoiles. .... 141

**Figure 80 : Diagrammes des intensités normalisées des spectres BEST-TROSY <sup>1</sup>H-<sup>15</sup>N de <sup>15</sup>N-P<sub>ND</sub> et <sup>15</sup>N-P<sub>FL</sub> en présence de PP1 $\alpha$  non marquée.** Les rapports d'intensité (I/I<sub>0</sub>) ont été mesurés avec les spectres de <sup>15</sup>N-P<sub>ND</sub> (A) ou <sup>15</sup>N-P (B) entière à 25  $\mu$ M seules et en présence de 2 équivalents molaires de GST-PP1 $\alpha$ . Les lignes droites et pointillées indiquent les valeurs moyennes et moyennes  $\pm$  moyenne quadratique. Les régions présentant une structure secondaire transitoire en hélice  $\alpha$  ou en feuille  $\beta$  sont mises en surbrillance par un arrière-plan coloré avec le même code de couleur que celui de la **Figure 75**. .... 142

**Figure 81 : Changement conformationnel du bras C-terminal de la nucléoprotéine du hMPV.** Les domaines N et C-terminaux de N sont représentés en « cartoon » bleu et cyan respectivement. Le domaine N(C arm) est représenté en « cartoon » rouge : il est orienté vers le bas dans le complexe N<sup>0</sup>-P et vers le haut dans la structure d'anneaux N-ARN, comme indiqué sur la figure (Renner et al., 2016). .... 147

**Figure 82 : Etude du rôle de la phosphorylation de TAX1BP1 dans l'interaction avec N.** (A) Schéma des domaines de l'isoforme 2 de Tax1BP1. Les délimitations des domaines et les positions de phosphorylation sont indiquées par les numéros des résidus correspondants. (B) Immunoprécipitation à partir de cellules BSRT7 co-transfectées avec les plasmides codant pour pFLAG-TAX1BP1 (sauvage ou mutant) et pcDNA-N. Une immunoprécipitation en présence d'anticorps anti-FLAG a ensuite été réalisée sur les lysats cellulaires. Les protéines TAX1BP1 et N ont été révélées par western blot en présence d'anticorps anti-FLAG-HRP et anti-N (Rapport de stage de M2 Sarah MADRIERES). .... 148

**Figure 83 : Structure du complexe P—L du hRSV.** Le L apparaît en surface blanche et les monomères P en « cartoon ». Les 4 protomères sont colorés en arc-en-ciel (bleu à rouge correspond à N- vers C-terminus respectivement). Les images du haut montrent 2 orientations du complexe et en bas les 4 protomères sont représentés séparément. Figure tirée de (Gilman et al., 2019). .... 150

**Figure 84 : Détails de l'interactions M2-1—P révélés par la structure cristallographique de M2-1—P[90–110] (pdb : 6G0Y).** Panneau de gauche : Les structures de M2-1 et du peptide de P sont colorées en bleu et orange respectivement sur la structure en « cartoon ». Les résidus importants dans la surface d'interaction sont mis en évidence. Quelques potentielles interactions électrostatiques et ponts salins sont représentés par des tirets. Panneau de droite : La figure montre l'enfouissement du résidu hydrophobe F109 de P entre 2 hélices du domaine M2-1<sub>core</sub>, à l'intérieur d'une poche hydrophobe représentée en surface électrostatique. La figure de gauche est tirée de (Selvaraj et al., 2018)..... 151

## Table des matières

Chapitre 1. Introduction générale sur le RSV .....	25
1.1. Pathologies et traitements.....	25
1.2. Le RSV au sein des <i>Mononegavirales</i> .....	27
1.3. Génome du RSV.....	29
1.4. Les protéines du complexe ARN polymérase du RSV.....	30
1.4.1. Les fonctions du complexe RdRp.....	30
1.4.2. La nucléoprotéine .....	31
1.4.3. La phosphoprotéine .....	33
1.4.4. Le cofacteur de transcription M2-1 .....	34
1.5. Partenaires cellulaires d'interaction des protéines N et P du RSV .....	35
1.5.1. Interactants de la protéine N.....	35
1.5.2. Interactants de la protéine P.....	35
1.5.3. La protéine PP1 $\alpha$ .....	35
1.5.4. La protéine Tax1BP1.....	36
1.6. Description du sujet de thèse .....	38
Chapitre 2. Matériels et Méthodes .....	39
2.1. Biologie moléculaire .....	39
2.1.1. Mutagenèse dirigée et amplification de plasmide .....	39
2.1.2. Transformation de bactéries par électroporation .....	40
2.2. Production et purification de protéines recombinantes marquées par des isotopes stables ou non .....	40
2.2.1. Culture de bactéries et production de protéines recombinantes .....	40
2.2.2. Purification de protéines recombinantes.....	41
2.2.3. Contrôle de la qualité des purifications .....	42
2.3. Expériences de GST pulldown .....	43
2.4. Utilisation de la RMN pour l'études de protéines .....	44
2.4.1. Acquisition de spectres RMN d'échantillons de protéines.....	44
2.4.2. Attribution séquentielle des fréquences de résonance .....	45
2.4.3. Titrage par perturbation des déplacements chimiques.....	47

2.4.4.	Mesure des vitesses de relaxation de spin nucléaire.....	47
2.4.5.	Augmentation de la relaxation par effet paramagnétique.....	48
2.5.	Logiciels d'analyse de structure et de dynamique de protéines utilisés .....	49
2.5.1.	CYANA.....	49
2.5.2.	ARIA .....	50
2.5.3.	Flexible-meccano .....	50
Chapitre 3. Etude de l'interaction entre la nucléoprotéine du hRSV et la protéine cellulaire		
Tax1BP1	53	
3.1.	Découverte de l'interaction entre la protéine N du HRSV et la protéine Tax1BP1 ..	53
3.2.	Définition de fragments de nucléoprotéine et de Tax1BP1 comme modèles d'étude	57
3.3.	Production et purification des domaines C-terminaux de la nucléoprotéine et de la protéine Tax1BP1.....	59
3.4.	Etude de la structure du domaine N(C arm) isolé .....	60
3.4.1.	Attribution des déplacements chimiques de N(C arm).....	60
3.4.2.	Etude des effets de la concentration de N(C arm) sur les déplacements chimiques.	64
3.4.3.	Obtention de la structure du domaine N(C arm) .....	67
3.5.	Etude du domaine Tax1BP1(UBZ1+2) isolé par RMN .....	70
3.5.1.	Attribution des déplacements chimiques de Tax1BP1(UBZ1+2) .....	70
3.6.	Etude de l'interaction entre les peptide $^{13}\text{C}^{15}\text{N}$ -N(C arm) et Tax1BP1(UBZ1+2) ....	71
3.6.1.	Titration de $^{13}\text{C}^{15}\text{N}$ -N(C arm) par Tax1BP1(UBZ1+2) .....	71
3.6.2.	Contrôle négatif de l'effet de l'albumine sur N(C arm) .....	78
3.6.3.	Cartographie des perturbations structurales de N(C arm) par Tax1BP1(UBZ1+2)	79
3.7.	Etude de l'interaction entre les peptides $^{15}\text{N}$ -Tax1BP1(UBZ1+2) et N(C arm).....	80
3.7.1.	Titration de $^{15}\text{N}$ -Tax1BP1(UBZ1+2) par N(C arm).....	80
3.7.1.	Contrôle négatif de l'effet de l'albumine sur Tax1BP1(UBZ1+2).....	87
3.7.2.	Cartographie des perturbations structurales de Tax1BP1(UBZ1+2) induites par N(C arm)	88
3.7.3.	Alignement de séquence en acide aminé de Tax1BP1(UBZ1+2) .....	89

3.7.4.	Expériences complémentaires de GST pulldown entre Tax1BP1(UBZ1+2) et N(C arm)	90
3.8.	Conclusions .....	93
Chapitre 4.	Caractérisation structurale de la phosphoprotéine (P) du hRSV .....	95
4.1.	Intérêt de la caractérisation structurale de la protéine P du hRSV et stratégie .....	95
4.2.	Production et purification de P entière ou en domaines séparés .....	97
4.3.	Caractérisation structurale de P entière .....	97
4.3.1.	Principales données préliminaires .....	97
4.3.2.	Expériences de relaxation de spin nucléaire sur la protéine P entière .....	102
4.3.3.	Recherche de contacts à longue distance au sein de la protéine P entière .....	103
4.4.	Caractérisation structurale de la partie centrale et des hélices C-terminales de P ...	106
4.4.1.	Attribution des fréquences de résonance du squelette peptidique de P[127-205] et de P <sub>OD</sub>	106
4.4.2.	Prédiction de structures secondaires .....	110
4.4.3.	Détermination de la structure de P <sub>OD</sub> .....	112
4.4.4.	Recherche de contacts à longue distance au sein du fragment P[127-205] .....	126
4.1.	Conclusions .....	130
Chapitre 5.	Etude des interactions entre protéines P—M2-1 et P—PP1 $\alpha$ .....	131
5.1.	Fonctions, interactions et phosphorylation de M2-1 .....	131
5.2.	Etude de l'interaction M2-1—P .....	132
5.2.1.	Production et purification des protéines M2-1 <sub>core</sub> , P marquée <sup>15</sup> N et ses fragments marqués <sup>15</sup> N	132
5.2.1.	Analyse par spectrométrie de masse des protéines M2-1 <sub>core</sub> et P marquée <sup>15</sup> N ...	132
5.2.2.	Etude de l'interaction entre la protéine P marquée <sup>15</sup> N et le domaine core de M2-1 non marquée par RMN .....	134
5.2.3.	Expériences complémentaires pour l'étude de l'interaction P—M2-1 .....	137
5.1.	Déphosphorylation de M2-1 par PP1 $\alpha$ avec l'aide de P .....	139
5.1.1.	Démonstration d'une interaction directe entre PP1 $\alpha$ et P .....	140
5.1.2.	Expériences complémentaires avec P, PP1 $\alpha$ et M2-1 .....	142
5.2.	Conclusions .....	143

Chapitre 6. Synthèse, discussion et perspectives .....	145
6.1. Interaction entre la nucléoprotéine du hRSV et la protéine cellulaire Tax1BP1 .....	145
6.2. Caractérisation structurale de la phosphoprotéine du hRSV .....	149
6.3. Interactions entre hRSV P et les protéines M2-1 et PP1 $\alpha$ .....	150
6.4. Perspectives .....	153
Chapitre 7. Bibliographie .....	155
Chapitre 8. Annexes .....	167
8.1. Déplacements chimiques attribués de N(C arm) .....	167
8.2. Titrage de $^{13}\text{C}^{15}\text{N}$ -N(C arm) par Tax1BP1(UBZ1+2) non marquée.....	173
8.3. Titrage de $^{15}\text{N}$ -Tax1BP1(UBZ1+2) par N(C arm) non marquée.....	174
8.4. Données des expériences de $^{15}\text{N}$ -hetNOE sur P entière .....	176
8.5. Données des expériences de relaxation $^{15}\text{N}$ R <sub>1</sub> sur P entière .....	179
8.6. Données des expériences de relaxation $^{15}\text{N}$ R <sub>2</sub> sur P entière .....	184
8.7. Déplacements chimiques attribués de P <sub>OD</sub> .....	189
8.8. Données des expériences de PRE sur P[127-205](S143C) .....	191
8.9. Données des expériences de PRE sur P[127-205](S156C) .....	193
8.10. Articles scientifiques publiés.....	195



# Chapitre 1. Introduction générale sur le RSV

## 1.1. Pathologies et traitements

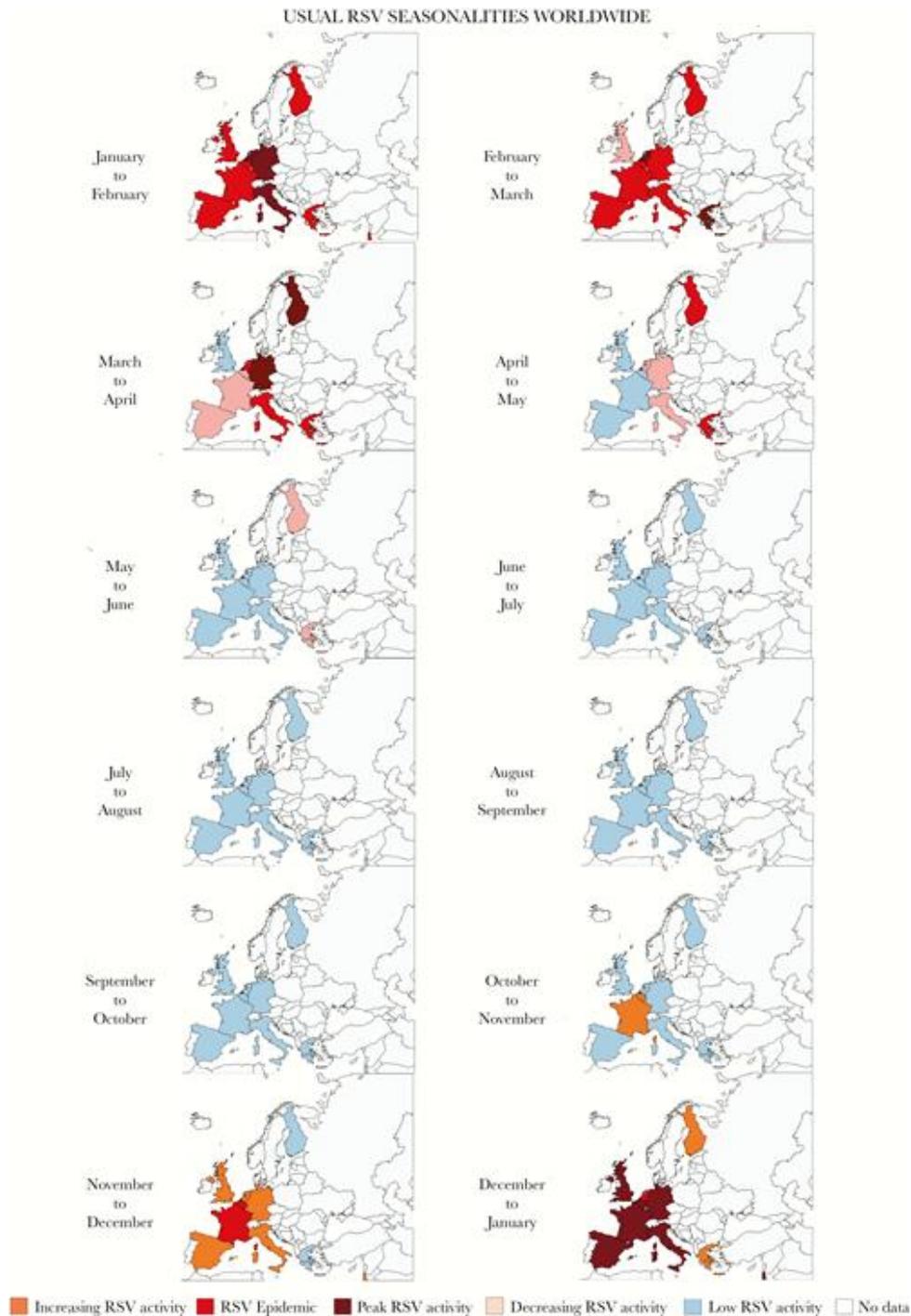
Le virus respiratoire syncytial (RSV) a été isolé pour la première fois en 1955 chez un chimpanzé en captivité atteint de maladie des voies respiratoires supérieures (Blount et al., 1956). Il a ensuite été retrouvé comme pathogène responsable de maladie respiratoire touchant les enfants (Chanock et al., 1957 ; (Chanock et Finberg, 1957)).

Aujourd'hui, le virus est reconnu comme le principal responsable des maladies respiratoires affectant les voies basses et comme cause de la majorité des bronchiolites mais aussi de pneumonies à travers le monde. Le virus touche particulièrement les enfants de moins de 2 ans avec une prévalence de 100 % (Glezen et al., 1986), mais il affecte aussi les personnes âgées et les immunodéprimés (Collins et Melero, 2011). En 2015, chez des enfants de moins de 5 ans, plus de 33 millions de cas d'infection aiguë des voies respiratoires inférieures dues au RSV humain (hRSV) ont été reportés dans le monde, dont environ 10% ont donné lieu à des hospitalisations parmi lesquelles environ 60 000 décès ont été recensés (~120 000 morts en comptant ceux n'ayant pas été hospitalisés ; Shi et al., 2017). L'affection des jeunes enfants a une tendance à causer de l'asthme et des problèmes de respiration sifflante à plus long terme (Rostad, 2019). Dans les pays au climat tempéré, le hRSV atteint de manière prépondérante les populations en automne/hiver avec un pic généralement situé autour du mois de décembre (**Figure 1**).

Il n'existe pas de vaccin ciblant le hRSV sur le marché actuellement. En préventif, il y a uniquement un anticorps monoclonal qui cible la protéine de fusion (F) du hRSV, le Palivizumab. Mais il est onéreux et efficace à environ 50%. Par conséquent il est utilisé seulement de manière préventive pour les cas les plus risqués. Plus de cinquante vaccins sont en cours de développement contre le virus et le plus avancé d'entre eux est un vaccin maternel pour stimuler le transfert d'anticorps transplacentaires (Kombe et al., 2019).

Il n'y a pas non plus d'antiviral spécifique du hRSV sur le marché. Il existe seulement la Ribavirine, qui est un analogue de nucléoside et qui n'est pas spécifique du virus, mais qui est très controversé notamment à cause de ses faibles effets thérapeutiques et de ses effets secondaires. Cependant, de nombreuses molécules inhibitrices sont en phase préclinique ou en phases cliniques dont un grand nombre ciblent la protéine hRSV-F (Heylen et al., 2017). Il existe aussi une approche par ARN interférents (siRNA) dont l'un ciblant la nucléoprotéine (N) du hRSV (Broadbent et al., 2015).

Quoi qu'il en soit, le problème de santé mondial que génère le hRSV et l'absence de traitement optimal contre ce dernier engendrent la nécessité absolue d'une compréhension plus avancée de ses mécanismes d'actions et du développement de nouveau traitement contre le virus.

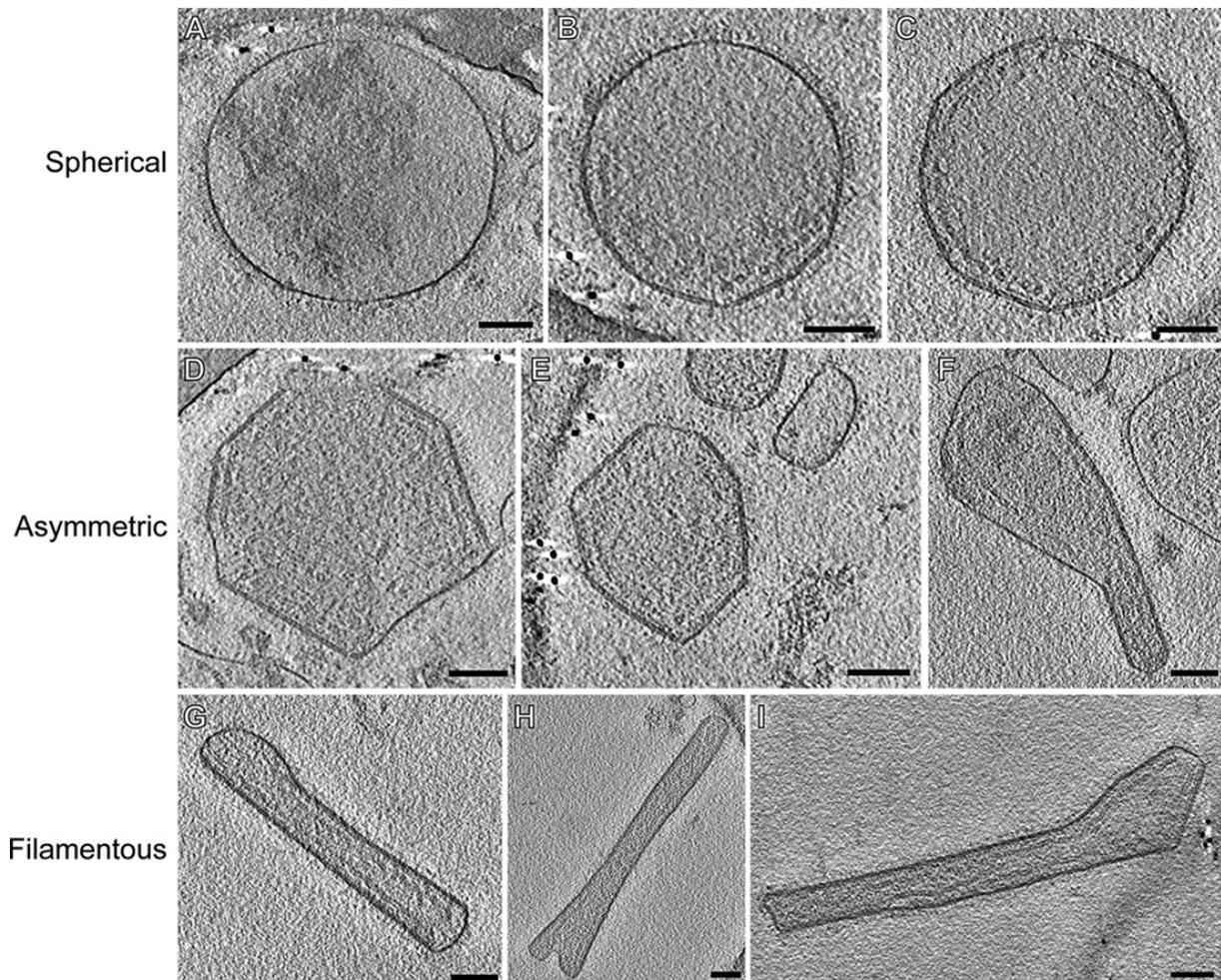


**Figure 1 : Épidémiologie mensuelle du RSV en Europe.** Le virus se propage généralement du sud au nord et d'ouest en est, montrant une activité sur le continent d'octobre à mai dans certains pays. L'activité maximale est généralement observée en hiver dans la plupart des pays, à l'exception des pays les plus septentrionaux, où une forte activité est observée au début du printemps. Concernant la France, cette étude correspond aux années 2011 à 2017. Cette figure est issue de Obando-Pacheco et al., 2018.

## 1.2. Le RSV au sein des *Mononegavirales*

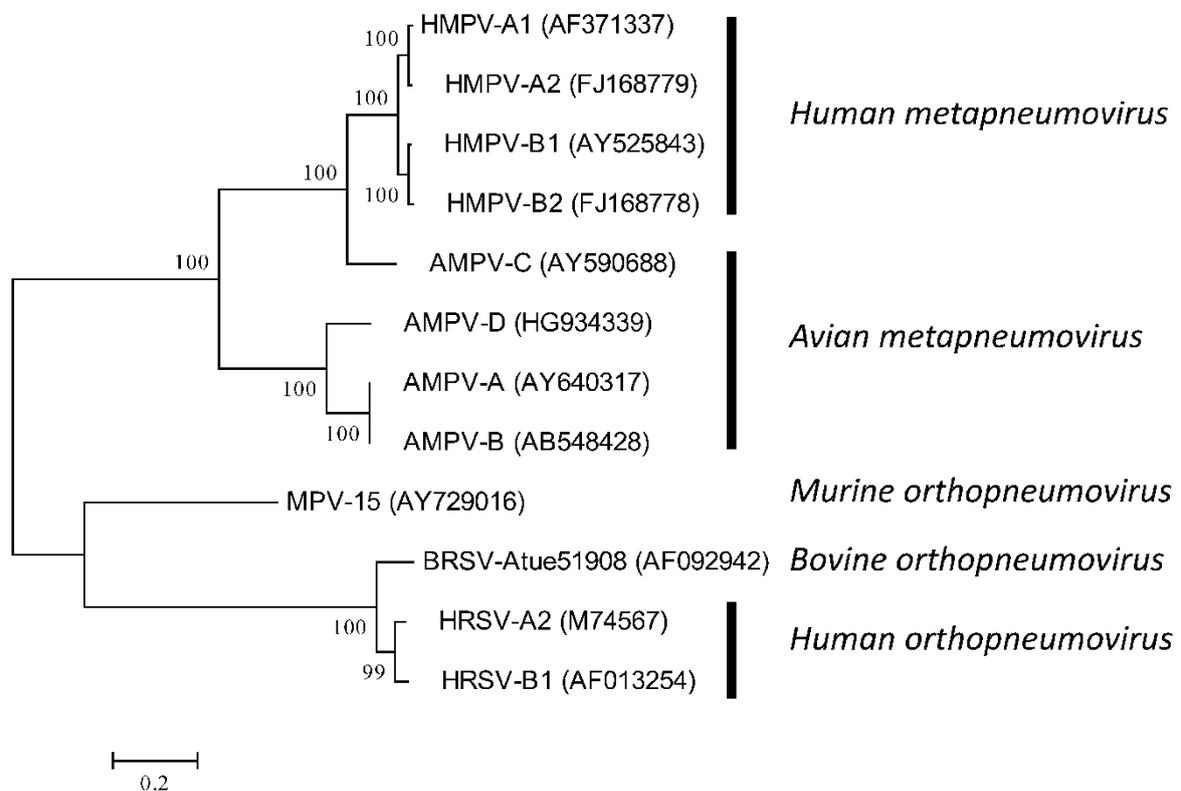
Le virus respiratoire syncytial humaine appartient à l'ordre des *Mononegavirales*. Ces virus ont la particularité d'être des virus enveloppés avec un génome à ARN négatif simple brin. Le hRSV est classé dans le genre *Orthopneumovirus* appartenant à la famille des *Pneumoviridae* (Amarasinghe et al., 2019).

L'architecture du virion du hRSV, avec une taille de 100 nm à 1 µm, est apparu comme étant sphérique, filamenteux ou comme une combinaison des deux, c'est-à-dire asymétrique (Liljeroos et al., 2013 ; **Figure 2**). Parmi les plus proches parents des *Pneumoviridae*, les *Filoviridae* (dont le virus d'Ebola et de Marburg font partie) forment aussi des virion filamenteux avec des tailles comparables (Bharat et al., 2012). Ces deux familles ont également en commun un cofacteur de transcription qui leur est propre (VP30 pour les virus d'Ebola/Marburg et M2-1 pour le hRSV).



**Figure 2 : Caractérisation morphologique de base des trois morphologies du hRSV.** Particules représentatives des trois catégories de morphologie présentes dans l'échantillon de RSV : sphérique (A à C), asymétrique (D à F) et filamenteuse (G à I). Les images sont des coupes centrales de 7,5 nm issues de reconstructions tomographiques. La figure est tirée de Kiss et al., 2014. Barres d'échelle, 100 nm.

La famille des *Pneumoviridae* regroupe les métapneumovirus humain (hMPV) et aviaire, le virus de la pneumonie murine (mPVM) et les virus respiratoires syncytiaux humain (A2 et B1) et bovin (Figure 3). Le hMPV, responsable du même genre de symptômes que le RSV mais dans une moindre mesure, est le plus proche parent du hRSV parmi les virus affectant l'humain. Le hMPV, tout comme le RSV, possède le gène M2 qui code pour les protéines M2-1 et M2-2 impliquées dans la régulation de la transcription et la réplication du génome viral. Une des différences majeures du hMPV par rapport au hRSV est l'absence des protéines NS1 et NS2 (perturbant la réponse des interférons antiviraux d'origine cellulaire) ainsi que l'ordre de certains gènes sur son génome (Rima et al., 2017). Cependant la famille *Paramyxoviridae* (dont le virus de la grippe fait partie), plus éloignée, possède également des protéines NS1 et NS2 dont la première protéine a également un rôle dans la perturbation de la réponse aux interférons post infection.

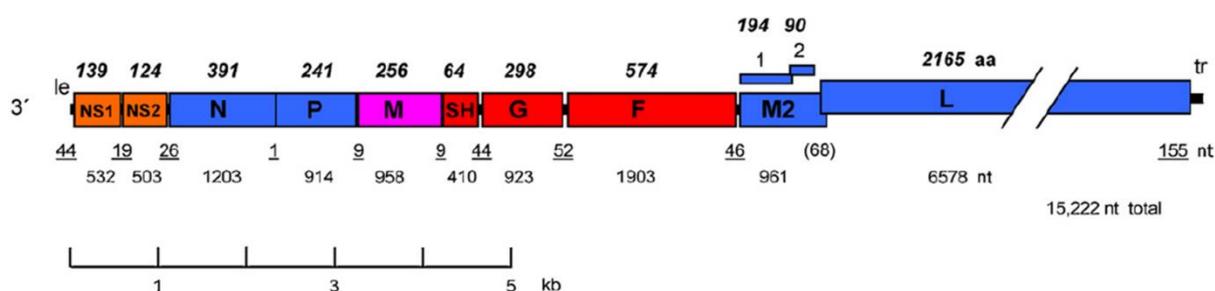


**Figure 3 : Arbre phylogénétique des *Pneumoviridae* en 2017.** Des séquences de l'ARN polymérase L alignées provenant des membres représentatifs des *Pneumoviridae* ont été utilisées pour produire un arbre phylogénétique par maximum de vraisemblance. Le pourcentage d'arbres dans lesquels les taxons associés sont regroupés est indiqué pour des valeurs supérieures à 70 %. Les barres indiquent les affectations d'espèces de ces isolats de virus. Source : [https://talk.ictvonline.org/cfs-file/\\_key/communityserver-wikis-components-files/00-00-00-00-90/ONSR.Pne.Fig3.v3.png](https://talk.ictvonline.org/cfs-file/_key/communityserver-wikis-components-files/00-00-00-00-90/ONSR.Pne.Fig3.v3.png).

### 1.3. Génome du RSV

Après être entré dans la cellule grâce à ses protéines de surface que sont la protéine d'adhésion (G) et la protéine de fusion (F), le hRSV doit transcrire et répliquer son génome au sein de la cellule hôte (Tayyari et al., 2011).

Le génome du hRSV, orienté de 3' vers 5' (virus ARN négatif), a une taille de 15,2 kb et possède 10 gènes distincts dont le gène M2 qui code pour les protéines M2-1 et M2-2 avec des séquences partiellement chevauchantes (**Figure 4**).



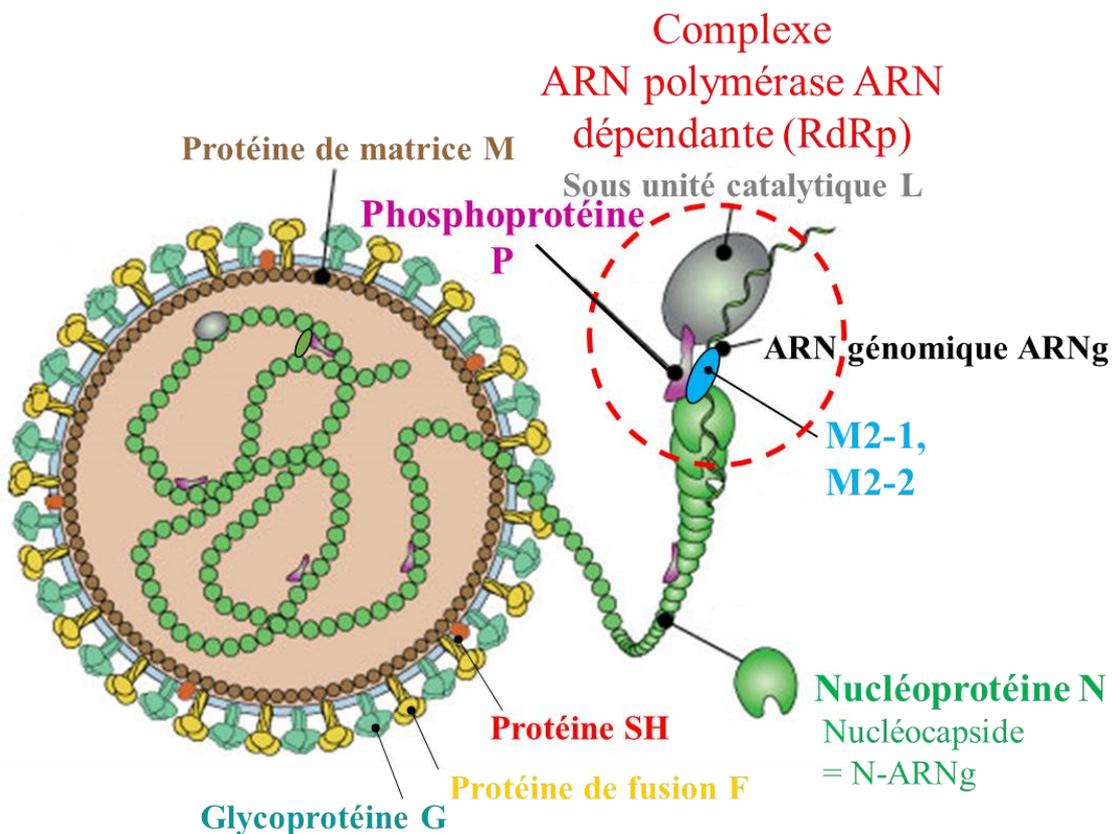
**Figure 4 : Carte génétique du RSV.** La figure montre une carte du génome à sens négatif (souche A2 du RSV), approximativement à l'échelle. Les cadres de lecture ouvert M2-1 et M2-2 qui se chevauchent sont montrés sur le gène. Les chiffres en dessous de la carte indiquent les longueurs en nucléotides (nt) : celles des régions du leader 3' (le), du trailer 5' (tr) et des régions intergéniques sont soulignées, et les longueurs du chevauchement des gènes sont indiquées entre parenthèses. Les chiffres en gras et en italique sur la carte indiquent les longueurs en acides aminés (aa) des protéines non modifiées. Les protéines virales dans l'ordre sont : G, glycoprotéine d'attachement ; F, glycoprotéine de fusion ; SH, petite glycoprotéine hydrophobe ; M, protéine matricielle ; N, nucléoprotéine ; P, phosphoprotéine ; L, grande polymérase ; M2-1, produit du premier ORF dans l'ARNm M2 ; M2-2, produit du deuxième ORF dans l'ARN M2 ; NS1, protéine non structurale 1 ; NS2, protéine non structurale 2. Figure issue de Collins et Melero, 2011.

En 3' du génome est localisée une région dite « leader » (Le) de 55 nucléotides et en 5' une région « trailer » (Tr) de 155 nucléotides. La séquence Le est nécessaire en tant que promoteur viral pour la réplication du génome viral et contient également la première séquence de transcription du gène NS1 tandis que la région Tr, dans le sens opposé, améliore l'efficacité du promoteur de l'anti génome pour reproduire le génome viral (Mink et al., 1991 ; Cowton et Fearn, 2005 ; Hanley et al., 2010). Le hRSV possède sa propre ARN polymérase (L) qui transcrit les gènes du virus de manière séquentielle à partir de signaux d'initiation et de transcription (gene start) qui comprennent la région de coiffe et d'initiation pour la synthèse d'ARNm. De plus, des régions de terminaison (gene end) responsables de la polyadénylation et d'une courte région intergénique non transcrite. Les gene end permettent également à la polymérase L de libérer les ARNm néo synthétisés et de progresser jusqu'au prochain gene start. Les gene start/end sont situées à chaque extrémité des gènes.

## 1.4. Les protéines du complexe ARN polymérase du RSV

### 1.4.1. Les fonctions du complexe RdRp

La machinerie de transcription et de réplication du hRSV, l'ARN polymérase ARN dépendante du virus (RdRp ; **Figure 5**), est composée *a minima* de la sous unité catalytique L et de son principal cofacteur, la phosphoprotéine (P). Le complexe RdRp se fixe sur la nucléocapside qui forme un complexe hélicoïdal (Bakker et al., 2013) comprenant la nucléoprotéine (N) du virus qui encapside l'ARN génomique viral (ARNg). Au moment de la transcription, la protéine M2-1 joue également le rôle de cofacteur de transcription au sein du complexe RdRp. Ce dernier est aussi en charge de la maturation du brin néoformé, de son coiffage en 5' et de la polyadénylation en 3'. Toutes ces fonctions essentielles au hRSV font du complexe RdRp une cible très intéressante et spécifique contre ce virus de par les spécificités de chacune de ses protéines.



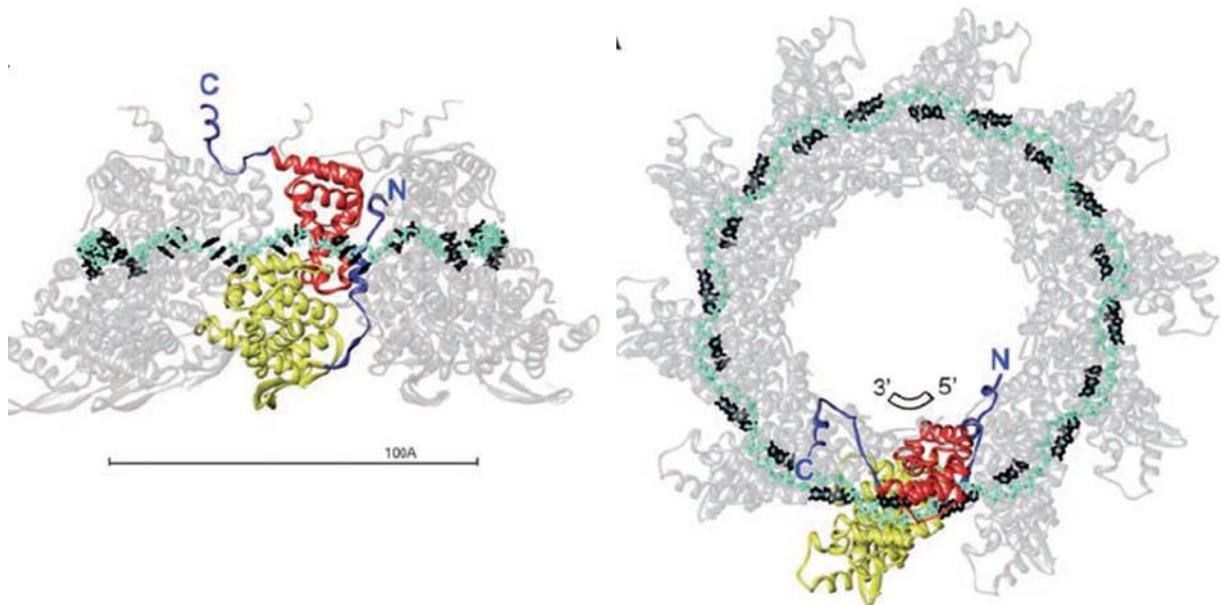
**Figure 5 :** Schéma du virion et des protéines du hRSV. Un focus est mis sur le complexe ARN polymérase ARN dépendante (RdRp) entouré en rouge. Le virion ne contient pas les protéines non structurales NS1 et NS2.

La nucléoprotéine, la phosphoprotéine et le cofacteur de transcription M2-1 du hRSV seront étudiées plus en détail par la suite, car c'est autour de ces 3 protéines que s'articulera mon projet de thèse.

### 1.4.2. La nucléoprotéine

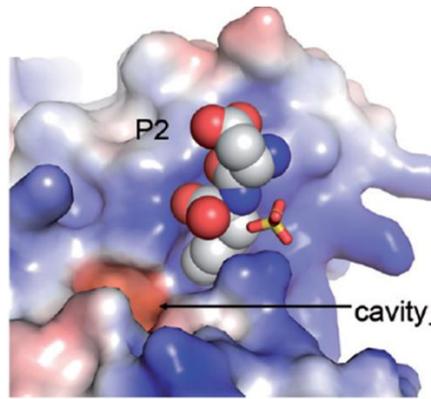
La nucléoprotéine du hRSV joue un rôle de protection de l'ARN génomique viral contre les mécanismes de défense de la cellule hôte en l'encapsidant (Jensen et al., 2011). L'interaction de N avec l'ARN n'est pas spécifique au génome du hRSV (Tawar et al., 2009).

La protéine N fait 43 kDa et est composée de 391 résidus. Elle possède deux domaines globulaires qui sont le domaine N-terminal (N-NTD correspondant aux résidus 36-253) et le domaine C-terminal (N-CTD correspondant aux résidus 254-360) ainsi que 2 extensions côté N- (N arm correspondant aux résidus 1-35) et C-terminal (C arm correspondant aux résidus 361-391) qui semblent structurées dans le contexte de la protéine entière (**Figure 6** ; Tawar et al., 2009). La structure de la protéine a été déterminée sous forme d'anneaux décameriques avec un ARN cyclique qui stabilise l'ensemble. Une stabilisation supplémentaire du décimère est obtenue grâce au N arm qui s'insère entre deux protomères et au C arm qui repose sur un protomère adjacent.



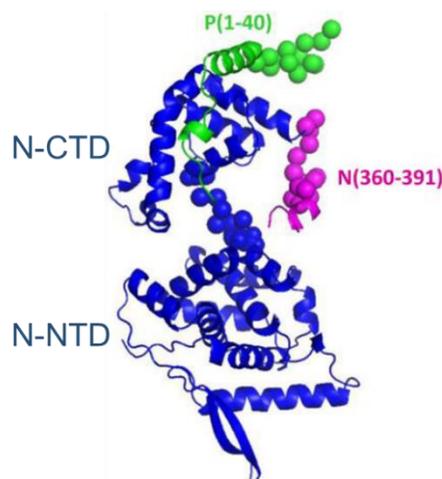
**Figure 6 : Structure cristallographique de la nucléoprotéine du hRSV sous forme décamerique liée à l'ARN.** Une sous-unité N est colorée selon les différents domaines : jaune et rouge, respectivement pour NTD et CTD. Les bras N- et C-terminaux sont en bleus. L'ARN est affiché en cyan et les bases azotées en noir. La polarité de l'ARN est indiquée dans l'anneau vu de haut (à droite). La barre d'échelle à gauche indique 100 Å.

Concernant l'interaction essentielle de la nucléocapside du RSV avec P, une poche hydrophobe au sein du domaine N-NTD a été caractérisée structuralement comme étant le site de fixation de hRSV-P (**Figure 7** ; Ouizougun-Oubari et al., 2015).



**Figure 7 : Complémentarité de surface N-NTD et un peptide de P.** On voit des vues agrandies de la liaison de P2 (peptide constitué des 2 derniers résidus de P). La surface de van der Waals sur N-NTD est colorée en fonction du potentiel électrostatique : potentiels positif en bleu et négatif en rouge. Le ligand P2 est affiché sous forme de sphères de van der Waals colorées par type d'atome, avec les atomes de carbone en blanc. Figure provenant de Ouizougun-Oubari et al., 2015.

Dernièrement, des modèles structuraux (à partir de données de SAXS, d'ultracentrifugation analytique et de biochimie) d'une forme libre et monomérique de N ( $N^0$ ) en interaction avec la région N-terminale de P ont été générés en partant du mutant RSV  $N\Delta 30$  qui reste monomère et ne lie pas l'ARN (**Figure 8** ; Esneau et al., 2019). Ces modèles sont basés sur des structures cristallographiques obtenue pour les formes N-ARN du RSV (Tawar et al., 2009) et  $N^0$ -P du hMPV (Renner et al., 2016).  $N^0$ -P correspond à une forme de N non liée à l'ARN qui permet de maintenir N dans un état monomère et disponible pour encapsider de l'ARN génomique (Collins et Melero, 2011, Galloux et al., 2015). On voit sur ces modèles le domaine N-CTD et l'extension N(C arm) en contact avec un peptide correspondant à la séquence N-terminale de P.

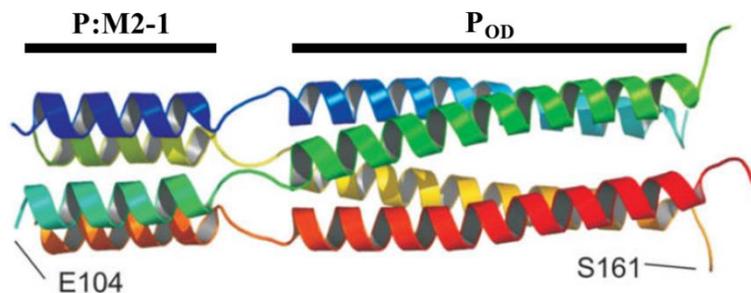


**Figure 8 : Modèle structural d'une forme de RSV N monomère ( $N^0$ ) obtenue avec le mutant  $N\Delta 30$  – P40 en solution.** Représentation en ruban du meilleur modèle structural du complexe RSV  $N\Delta 30$ -P40 (40 premiers acides aminés de P). La protéine N (résidus 31 à 382) est représentée en bleu, N-(C arm) (ici : résidus 383 à 391) en violet et le peptide P40 en vert. Les régions de liaison sont représentées par des sphères.

### 1.4.3. La phosphoprotéine

La phosphoprotéine (P) du RSV est une protéine de 27 kDa comportant 241 résidus. La protéine P peut être phosphorylée sur des thréonines ou des sérines avec un cyclage plus ou moins rapide au niveau des résidus 30, 39, 45, 46, 54, 108, 116, 117, 119, 143, 160, 210, 232 et 237 (Navarro et al., 1991 ; Sánchez-Seco et al., 1995 ; Asenjo et al., 2005 ; Asenjo et al., 2006 ; Asenjo et al., 2008) et pouvant éventuellement affecter l'interaction entre les protéines P et hRSV-M2-1 sans pour autant affecter le niveau de réplication du virus (Villanueva et al., 2000 ; Asenjo et al., 2006 ; Asenjo and Villanueva, 2016).

La phosphoprotéine forme des tétramères constitutifs indépendamment de la phosphorylation de la protéine (Asenjo et Villanueva, 2000 ; Castagné et al., 2004 ; Llorente et al., 2006). Le domaine d'oligomérisation de P, prédit en structure de super hélice, a été localisé entre les résidus 104 à 160 (**Figure 9**), et il est flanqué de part et d'autre par de longues régions désordonnées contenant de potentielles structures secondaires transitoires (Llorente et al., 2006 ; Llorente et al., 2008 ; Simabuco et al., 2011 ; Noval et al., 2016). Le modèle par homologie présenté dans la **Figure 9** comporte deux domaines, un domaine court et un domaine long allant des résidus 130 à 156. Le premier domaine a été présenté pendant longtemps comme une prolongation du domaine d'oligomérisation (Llorente et al., 2008, Noval et al., 2016).



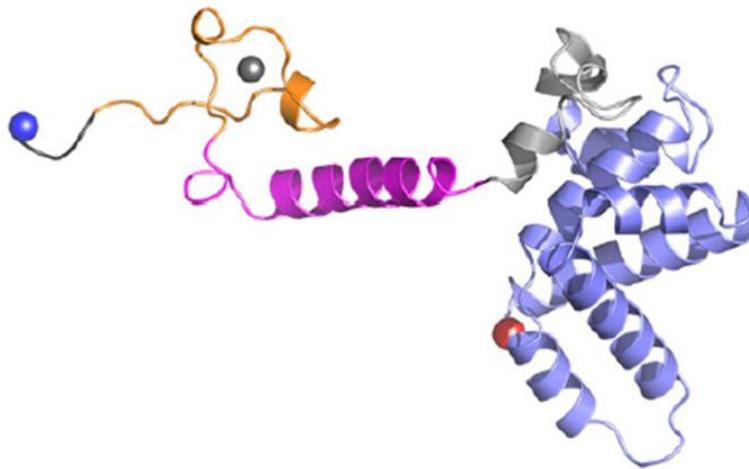
**Figure 9 : Prédiction d'hélices  $\alpha$  de la phosphoprotéine du RSV par Llorente et al., 2008.** Le domaine de liaison de la protéine M2-1 du hRSV se situe en amont du domaine d'oligomérisation de P (Mason et al., 2003 ; Tran et al., 2009 ; Blondot et al., 2012 ; Richard et al., 2018 ; Selvaraj et al., 2018).

La protéine P joue un rôle essentiel dans le rapprochement de la protéine L du hRSV et de la nucléocapside grâce à une interaction directe entre N et P au niveau des 12 derniers résidus de P (García-Barreno et al., 1996 ; Slack and Easton, 1998 ; Tran et al., 2007 ; Galloux et al., 2012 ; Ouizougoun-Oubari et al., 2015) ainsi qu'un rôle de chaperonne pour la protéine N néo-synthétisée, monomérique et sans ARN grâce à une interaction avec les 40 premiers résidus de P (Galloux et al., 2015 ; Pereira et al., 2016 ; Esneau et al., 2019). D'autre part, P interagit avec le cofacteur de transcription M2-1 au niveau de la région impliquant les résidus 90-110 (Mason et al., 2003 ; Tran et al., 2009 ; Blondot et al., 2012 ; Richard et al., 2018 ; Selvaraj et al., 2018 ; **Figure 9**).

#### 1.4.4. Le cofacteur de transcription M2-1

M2-1 est une protéine de 22,15 kDa incluant 194 acides aminés. Cette protéine joue un rôle essentiel pour la transcription du génome du hRSV en ARN messagers (Collins et al., 1995). En augmentant la processivité du complexe RdRp en tant que transcriptase. Celui-ci ayant une tendance à se décrocher du brin matrice après un gene end, la protéine M2-1 agit comme un anti-terminateur de la transcription du RSV en empêchant la terminaison prématurée de la transcription (Fearn and Collins, 1999).

M2-1 forme des tétramères stables en solution. Chaque protomère comprend un domaine de doigt de zinc en N-terminal (résidus 7-25), un motif de tétramérisation  $\alpha$ -hélicoïdal (résidus 32-49), un domaine globulaire hélicoïdal (résidus 69-172) nommé « core » ainsi qu'une extrémité C-terminale désordonnée (résidus 173-194) ; les résidus non cités dans un domaine sont considérés comme flexibles (Tran et al., 2009 ; Blondot et al., 2012 ; Tanner et al., 2014 ; **Figure 10**).



**Figure 10 : Structure cristallographique de hRSV-M2-1.** Un protomère de M2-1 est représenté avec des régions colorées de la manière suivante : domaine de liaison au zinc (orange) avec son zinc (sphère grise) ; hélice de tétramérisation (magenta) ; domaine core (bleu) ; domaine non structuré dans la plupart des molécules (gris) ; résidus du vecteur de clonage (noir). Les extrémités N- et C-terminales sont marquées avec des sphères bleues et rouges, respectivement. Figure issue de Tanner et al., 2014.

M2-1 est une protéine de liaison à l'ARN qui se lie préférentiellement aux ARNm du RSV et aux séquences riches en adénines (Cartee and Wertz, 2001). M2-1 peut également interagir avec P au niveau du domaine core (Blondot et al., 2012). De plus, P et les ARN se lient de manière compétitive à M2-1 avec des surfaces d'interaction qui se recoupent partiellement (Tran et al., 2009 ; Blondot et al., 2012).

## 1.5. Partenaires cellulaires d'interaction des protéines N et P du RSV

### 1.5.1. Interactants de la protéine N

Au-delà de l'implication de la protéine N du RSV dans ses interactions avec l'ARN génomique viral et avec la protéine P, il a été montré par des expériences de co-localisation et Proximity Ligation Assay (de test de lien entre protéines) qu'il existe une interaction entre N et les protéines MDA5 (*Mélanoma* differentiation-associated protein 5) et RIG-I (Retinoic acid-inducible gene I) qui sont des protéines cellulaires impliquées dans le contrôle de la réponse antivirale (Lifland et al., 2012).

Plus tard, des expériences de co-immunoprécipitation suivies de détection par spectrométrie de masse ou par détection avec des anticorps ont permis d'identifier des interactions entre l'arginine méthyltransférase, la protéine méthylosome et la protéine HSP70 (Oliveira et al., 2013). L'arginine méthyltransférase a une fonction histone méthyltransférase sur l'histone H. La protéine méthylosome est une protéine non catalytique appartenant au complexe méthylosome qui est également responsable d'une fonction arginine méthyltransférase qui s'exerce sur des protéines de la particule d'épissage (spliceosome) (Friesen et al., 2002). Enfin, la protéine HSP70 est une protéine chaperonne ubiquitaire.

### 1.5.2. Interactants de la protéine P

Comme pour la protéine N, la protéine P, qui est une véritable plateforme d'interaction pour le RSV (notamment au travers de son interaction avec la nucléoprotéine libre, la nucléocapside, la polymérase L et le cofacteur de transcription M2-1), possède certainement d'autres partenaires étant d'origine cellulaires.

La publication de Oliveira et al., 2013, avec les mêmes méthodes que pour les partenaires cellulaires identifiés pour la protéine N, montre que la protéine P du RSV interagit avec la chaperonne Hsp70 ainsi qu'avec la tropomyosine qui peut fixer l'actine et jouer un rôle central dans la contraction des muscles striés (Janco et al., 2012).

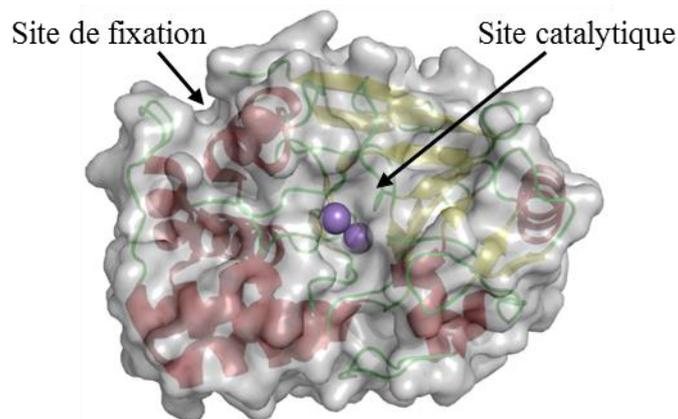
Cependant, il ne semble pas exister d'avantage d'informations sur l'interactome de P avec des protéines cellulaires.

### 1.5.3. La protéine PP1 $\alpha$

La phosphatase cellulaire PP1 $\alpha$  (Protein Phosphatase 1 $\alpha$ ) d'origine humaine a été identifiée comme étant impliquée dans la déphosphorylation de la protéine RSV-M2-1 grâce à une interaction avec la protéine P du RSV par un de nos collaborateurs de l'INRA à Jouy-en-Josas, le Dr Charles-Adrien Richard.

La protéine PP1 $\alpha$  est une phosphatase à sérine / thréonine humaine de 37,5 kDa comprenant 330 acides aminés. PP1 $\alpha$  ne possède qu'un seul domaine. Elle est composée de 10 hélices  $\alpha$  et de 14 brins  $\beta$  formant 3 feuillets dont 2 coplanaires. PP1 $\alpha$  comporte un site de fixation de régulation/substrat sous forme de sillon hydrophobe et un site catalytique distincts (Goldberg et al., 1995 ; **Figure 11**).

PP1 $\alpha$  est très abondante dans les cellules et est capable de déphosphoryler une grande variété de protéines. La protéine est impliquée dans divers processus biologiques dont la synthèse de protéines, la contraction musculaire, le métabolisme des glucides, la signalisation neuronale et dans la progression dans le cycle cellulaire. Cependant, PP1 n'a pas une grande affinité pour ses substrats et se trouve régulièrement complétée par des protéines régulatrices lui permettant d'interagir avec plus de 200 partenaires au total. De plus, la plupart des régulateurs de PP1 $\alpha$  et certains de ses substrats se lient à elle via un motif RVxF qui se lie sur le site de fixation (voir **Figure 11**) de PP1 à 20 Å du site catalytique de la phosphatase (Choy et al., 2014).



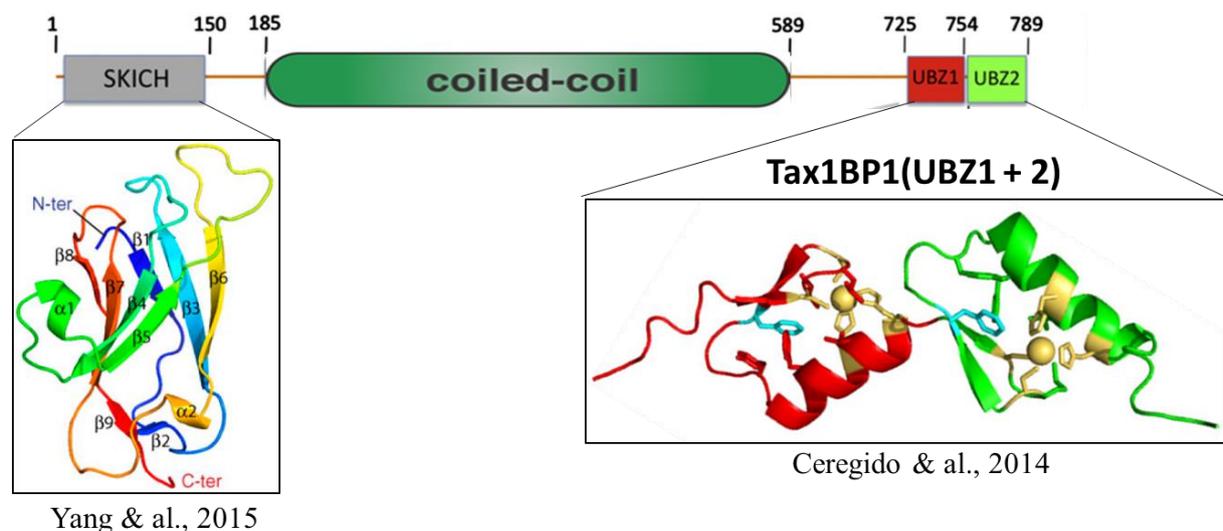
**Figure 11 : Structure cristallographique de PP1 $\alpha$ .** La surface de la phosphatase est représentée en gris partiellement transparent afin d'apercevoir les structures secondaires de la protéine avec le code couleur suivant : hélice  $\alpha$  en rouge, brin  $\beta$  en jaune, boucle en vert. Les 2 sphères correspondent à 2 ions manganèse appartenant au site catalytique de la protéine. Les sites de fixation de substrats et le site catalytique sont indiqués. La structure est issue du fichier pdb 3E7A et la figure a été générée à l'aide du logiciel Pymol (Schrödinger).

#### 1.5.4. La protéine Tax1BP1

La protéine TAX1BP1 (*TAX1 Binding Protein 1*), aussi nommée T6BP, TXBP151, ou encore CALCOCO3 a été identifiée comme un interactant d'origine cellulaire de la protéine N par nos collaborateurs de l'INRA à Jouy-en-Josas, en particulier le Dr Marie Galloux.

Tax1BP1 a d'abord été identifiée comme protéine partenaire de l'oncoprotéine virale Tax du virus HTLV1 (*Human T cell Leukemia/Lymphoma Virus type 1*). TAX1BP1 est un homodimère d'environ 90 kDa avec 789 résidus, exprimé de façon ubiquitaire. La protéine est composée de trois domaines principaux qui sont un domaine SKICH (*SKIP carbonyl homology*) en N-terminal dont

la fonction reste inconnue (Yang et al., 2015), un domaine d'oligomérisation au centre de la protéine (Ling et Goeddel, 2000), et enfin un domaine formé séparé en deux sous domaines UBZ (*Ubiquitin binding zinc finger*) en tandem (Ceregido et al., 2014) (**Figure 12**). Ce dernier possède une structure formée d'une répétition de deux doigts de zinc appelés UBZ1 et UBZ2. Deux résidus sérines (S593 et S624) ont aussi été signalés (sur l'isotype 2 de la protéine dont les résidus 604-645 sont absent par rapport à l'isotype 1 étudié dans ce travail), localisés entre le domaine central sur-enroulé et les domaines UBZ, dont la phosphorylation semble réguler l'interaction de TAX1BP1 avec ses partenaires cellulaires (Shembade et al., 2011).



**Figure 12 : Schéma des domaines de Tax1BP1 et structures connues des domaines.** La répartition des domaines de Tax1BP1 et leurs bornes sont indiquées sur le schéma. En dessous, sont représentées les structures déjà obtenues des domaines SKICH et UBZ1+2 (Ceregido et al., 2014 ; Yang et al., 2015).

De nombreuses fonctions ont été attribuées à la protéine TAX1BP1. Trois rôles majeurs sont à retenir pour cette protéine. Elle intervient dans la régulation négative de l'immunité innée en agissant sur la signalisation des voies NFκB (*Nuclear factor-kappa B*) et IRF3 (*Interferon regulatory factor 3*) (Honda et al., 2006 ; Verstreppe et al., 2011) qui permet le recrutement du complexe multiprotéique A20, possédant une activité déubiquitylase qui inhibe la transduction du signal. En réponse à l'infection virale, TAX1BP1 semble jouer un rôle anti-apoptotique en recrutant la ligase Itch au niveau de l'adaptateur mitochondrial MAVS (*Mitochondrial antiviral-signaling protein*), résultant en son ubiquitination et sa dégradation (Choi et al., 2016). Pour finir, TAX1BP1 joue un rôle dans le recrutement de la machinerie d'autophagie en interagissant à la fois avec la protéine LC3 des autophagosomes, les protéines ubiquitylées à dégrader et la myosine VI impliquée dans le recrutement des lysosomes. Il a notamment été rapporté que la protéine

TAX1BP1 était requise pour la clairance de la bactérie *Salmonella Typhimurium* au travers de l'autophagie (Tumbarello et al., 2015).

## 1.6. Description du sujet de thèse

Etant donné que plusieurs caractérisation d'interaction entre les protéines du hRSV ont déjà été menées ou sont bien avancées, mon projet de thèse s'est axé principalement autour de la caractérisation de nouvelles interactions impliquant cette fois des protéines du RSV et des protéines cellulaires en tant que potentielles futures cibles thérapeutiques.

Les critères retenus pour une étude éventuelle d'interaction entre protéine du RSV et protéine cellulaire étaient que la protéine cellulaire soit productible sous forme recombinante, pas trop grande (pour une étude par résonance magnétique nucléaire) et stable en solution, qu'on ait déjà des informations sur cette protéine et, si l'on pouvait le savoir à l'avance, que le ciblage de cette interaction n'affectait pas trop négativement la cellule.

Je n'ai pas travaillé sur les protéines cellulaires interagissant avec la protéine N ou P qui ont été décrites dans l'introduction car elles sont soit de grande taille ou trop peu renseignées. Aussi, des protéines cellulaires très centrales pour les mécanismes cellulaires, avec de larges interactomes, sont probablement de mauvais choix en tant que cibles thérapeutiques si on ne veut pas déréguler complètement des cellules saines en perturbant de manière croisée des interactions essentielles au bon fonctionnement cellulaire.

Ainsi, j'ai choisi premièrement de travailler sur l'étude de l'interaction entre la nucléoprotéine du RSV et la protéine Tax1BP1 qui peuvent être toutes deux produites par domaines et ont déjà été en partie étudiées par RMN.

Deuxièmement, j'ai travaillé sur l'interaction entre la phosphoprotéine du RSV et la phosphatase PP1 $\alpha$ . La protéine P avait déjà été beaucoup étudiée au laboratoire et la phosphatase est une protéine bien connue de manière générale dont la structure tridimensionnelle est connue.

Enfin, j'ai continué de travailler sur la caractérisation structurale de la protéine P du RSV sur laquelle j'avais déjà commencé quelques travaux durant mes stages précédant la thèse à l'ICSN.

# Chapitre 2. Matériels et Méthodes

## 2.1. Biologie moléculaire

### 2.1.1. Mutagenèse dirigée et amplification de plasmide

Les expériences de mutagenèse dirigée ont été réalisées avec des kits QuickChange II (Stratagene). Les constructions qui ont été faites sont indiquées dans le **Erreur ! Source du renvoi introuvable.** en indiquant les séquences en acides aminés des protéines de départ et des protéines modifiées ainsi que les oligonucléotides utilisés. **Figure 1**

Séquence d'origine	Oligonucléotides	Séquence finale
P[127-184] : GSGSGSGSTNDNITARLDRI DEKLSEILGMLHTLVVASA GPTSARDGIRDAMVGLREE MIEKIRTE	<b>P127-163stop_S :</b> 5'-aggacctacatctgctaggta aggtataagagatgccatgg-3' <b>P127-163stop_AS :</b> 5'-ccatggcatctcttatacctta cctagcagatgtaggtcct-3'	<b>P[127-163]/P<sub>OD</sub> :</b> insertion <b>d'un codon stop</b> GSGSGSGSTNDNITARLDRI DEKLSEILGMLHTLVVASA GPTSAR
P[161-241] : GSSARDGIRDAMVGLREEM IEKIRTEALMTNDRLEAMA RLRNEESEKMAKDTSDSEVS LNPTSEKLNNLLEGNDSDN DLSLEDF	<b>P161-205stop_S :</b> 5'-cacttcacatctgatgtgtcttt Tgcttacttttcacttttcctcatt cctgag-3' <b>P161-205stop_AS :</b> 5'-ctcaggaatgaggaaagtga aagtaagcaaaagacacatcagat gaagtg-3'	<b>P[161-205] :</b> insertion d'un <b>codon stop</b> GSSARDGIRDAMVGLREEM IEKIRTEALMTNDRLEAMA RLRNEESEK
P[127-205] : GSGSGSGSTNDNITARLDRI DEKLSEILGMLHTLVVASA GPTSARDGIRDAMVGLREE MIEKIRTEALMTNDRLEAM ARLRN	<b>P-S143C_Sens</b> 5'-CTGCAAGATTAGATAGGATTGA TGAAAAATTATGTGAAATACTAGGA ATG-3' <b>P-S143C_Antisens :</b> 5'-CATTCCCTAGTATTTACATAAT TTTTCATCAATCCTATCTAATCTTG CAG-3'	<b>P[127-205](S143C) :</b> mutation <b>ponctuelle</b> GSGSGSGSTNDNITARLDRI DEKL <b>C</b> EILGMLHTLVVASA GPTSARDGIRDAMVGLREE MIEKIRTEALMTNDRLEAM ARLRN
P[127-205] : GSGSGSGSTNDNITARLDRI DEKLSEILGMLHTLVVASA GPTSARDGIRDAMVGLREE MIEKIRTEALMTNDRLEAM ARLRN	<b>P-S156C_Sens</b> 5'-CACACATTAGTAGTAGCATGTG CAGGACCTACATCTG-3' <b>P-S156C_Antisens</b> 5'-CAGATGTAGGTCCTGCACATGC TACTACTAATGTGTG-3'	<b>P[127-205](S156C) :</b> mutation <b>ponctuelle</b> GSGSGSGSTNDNITARLDRI DEKLSEILGMLHTLVV <b>A</b> CA GPTSARDGIRDAMVGLREE MIEKIRTEALMTNDRLEAM ARLRN

**Tableau 1 : Constructions plasmidiques modifiées par mutagenèse dirigée.** Les constructions montrées correspondent aux séquences des protéines traduites après clivage à la thrombine d'une étiquette GST (Glutathion S-transférase). En rouge, sont indiqués les résidus du linker restant après

*coupure. Les mutations ponctuelles de cystéines sont indiquées en grande lettre et en gras. Les oligonucléotides utilisés pour la mutagénèse dirigée sont indiqués.*

Chaque plasmide muté a été envoyé au séquençage pour vérifier les mutations.

Concernant les productions et purification de plasmides stocks, j'ai utilisé des kits de mini prep Fermentas (désormais : Thermo Fisher Scientific) ainsi que des kits Macherey-Nagel pour les midi et maxi prep.

### 2.1.2. Transformation de bactéries par électroporation

J'ai utilisé des aliquots de cellules électro-compétentes X11blue (pour les productions de stock de plasmides) de 40-60  $\mu\text{L}$  (1 volume de cellules / 1 volume de glycérol 10%) que j'ai mélangés avec 1-10  $\mu\text{L}$  de plasmide à concentration  $\geq 0,08 \mu\text{g}/\mu\text{L}$  dans une cuve d'électroporation de 2 mm de large.

J'ai ensuite utilisé un réalisé la transformation par électroporation à 2500 V durant  $\sim 5$  ms.

Les cellules sont resuspendues dans 1 mL de milieu riche (LB ou 2YT ou SOC) sans antibiotique et incubées sous agitation pendant 1 h à  $37^\circ\text{C}$ .

Enfin, les cellules transformées sont sélectionnées en étalant 10-100  $\mu\text{L}$  de suspension de cellules sur boîtes de Pétri avec de l'agar-agar + milieu riche (le même que pour la resuspension en général) + antibiotique adapté au plasmide d'intérêt (kanamycine ou ampicilline) à  $37^\circ\text{C}$  durant une nuit ou plus.

## 2.2. Production et purification de protéines recombinantes marquées par des isotopes stables ou non

### 2.2.1. Culture de bactéries et production de protéines recombinantes

Pour chaque protéine recombinante que j'ai produite, j'ai utilisé des plasmides pGex exprimant une protéine GST suivie d'un site de coupure à la thrombine puis d'une courte extension  $(\text{GS})_n$  (où n varie entre 1 et 4), sauf pour la protéine P entière qui contenait une étiquette GSI en N-terminal, et enfin de la protéine d'intérêt (**Figure 13**).



**Figure 13 : Schéma de construction des protéines recombinantes.** Les extrémités N- et C-terminales sont indiquées de part et d'autre. En rouge est indiqué le site de coupure à la thrombine et le trait noir représente l'extension  $(\text{GS})_n$ .

D'abord, j'ai transformé par électroporation des bactéries *E. coli* BL21(DE3) selon le même protocole que celui indiqué au début du Chapitre 2 Matériels et Méthodes.

A partir d'une colonie sélectionnée après transformation, j'ai ensuite inoculé une préculture avec 10-50 mL de milieu LB + antibiotique adapté au plasmide que j'ai incubé à 37°C sur la nuit.

J'ai inoculé 10-20 mL de préculture dans 500 mL milieu LB (ou milieu M9 contenant du  $^{15}\text{NH}_4\text{Cl}$  à 1 g/L et du glucose uniformément marqué  $^{13}\text{C}$  à 3 g/L si les protéines étaient marquées  $^{13}\text{C}$ - $^{15}\text{N}$  ou non marqué pour des protéines uniquement marquées  $^{15}\text{N}$ ) + antibiotique et j'ai incubé la culture à 37°C sur agitateur durant ~7h en suivant la densité optique (DO) à 600 nm toutes les heures pour produire la biomasse. Après 7h, avec une  $\text{DO}_{600}\sim 3$  en LB ou une  $\text{DO}_{600}\sim 1,6$  en milieu M9, j'ai induit la production de protéines recombinantes avec un ajout de 500 mL du même milieu + 80  $\text{mg}\cdot\text{L}^{-1}$  d'IPTG (isopropyl  $\beta$ -D-1-thiogalactopyranoside) (+ 50  $\mu\text{M}$  de  $\text{ZnCl}_2$  pour les protéines contenant une structure en doigt de zinc) et laissé à incuber sur la nuit à 28°C. Enfin, j'ai centrifugé mes cultures pour récupérer mes culots de cellules et les congeler à 80°C avant la lyse et la purification des protéines.

### 2.2.2. Purification de protéines recombinantes

Les étapes suivantes ont été faites autant que possible sur de la glace ou à 4°C.

Le premier jour, j'ai resuspendu mes culots de cellules congelées dans du tampon de lyse (Tris (tris(hydroxyméthyl)aminométhane) 50 mM, NaCl 60 mM, Triton X-100 0,2 %, pH 7,8 ; + 50  $\mu\text{M}$  de  $\text{ZnCl}_2$  et 200  $\mu\text{M}$  de TCEP pour les protéines contenant une structure en doigts de zinc) supplémenté d'un cachet de cocktail d'inhibiteur de protéase (Roche complete EDTA-free) pour obtenir des cellules diluées entre 5 et 10 fois. Ensuite, j'ai cassé les cellules avec un cell disruptor (Constant Systems) à 1,5 kbar de pression. Puis, j'ai centrifugé à 40.000 g, 1 h, 4°C les lysats pour les clarifier et retirer les débris cellulaires. J'ai ensuite filtré les surnageants sur une membrane avec un cut-off de 0,2  $\mu\text{m}$ . Enfin, j'ai ajouté des billes de sépharose-GSH pour capturer les protéines d'intérêt fusionnées à une étiquette GST (2 mL / litre de culture) au surnageant que j'ai laissé incuber une nuit à 4°C avec une très faible agitation.

Le second jour, j'ai récupéré les billes sépharose-GSH (GE Healthcare) dans une colonne vide avec un fritté. Après élution de la solution de lysat, j'ai rincé les billes avec ~10 mL de tampon de lyse puis en faisant 3-4 lavages avec ~10 mL de tampon de coupure à la thrombine (Tris 20 mM, NaCl 150 mM,  $\text{CaCl}_2$  2,5 mM, pH 8,4). Ensuite, j'ai resuspendu les billes dans le même volume de tampon thrombine et j'ai incubé les suspensions de billes avec de la thrombine biotinylée (10-20 unité / 2 mL de billes ; Merck Millipore) pendant environ une nuit à 4°C sur roue.

Le troisième jour, j'ai élué les protéines clivées par la thrombine (considérant que l'étiquette GST reste accrochée sur la résine de billes sépharose-GSH). Pour les protéines avec doigts de zinc, j'ai ajouté 1 mM de TCEP en solution à cette étape. J'ai ajouté des billes agarose-streptavidine (200  $\mu\text{L}$  de billes

(dilution au  $\frac{1}{2}^e$ ) / 10 U de thrombine biotinylée ; Merck Millipore) dans l'échantillon de protéine clivées contenant la thrombine biotinylée afin de fixer cette dernière avec la streptavidine en laissant tourner l'ensemble sur roue pendant environ ~1 h. Enfin, j'ai élué mes protéines d'intérêt tout en filtrant les billes ayant capturé la thrombine puis j'ai dialysé mes échantillons avec mes protéines d'intérêt afin de les passer dans la solution tampon adaptée aux expériences RMN que j'ai menées (en général tampon phosphate 20 mM, NaCl 100 mM, pH 6-7 ; + 1 mM de TCEP pour les protéines contenant une structure en doigts de zinc à la fin des dialyses).

## 2.2.3. Contrôle de la qualité des purifications

### 2.2.3.1. SDS PAGE

Pour contrôler la qualité et l'avancement de mes purifications de protéines recombinantes ou pour d'autres raisons j'ai été amené à réaliser un grand nombre de gels SDS PAGE (électrophorèse sur gel de polyacrylamide en présence de dodécylsulfate de sodium).

Pour ce faire, j'ai utilisé des gels de séparation avec une concentration en acrylamide/bis acrylamide entre 10 et 20% (v/v) selon les intervalles des masses de protéines que souhaitais observer. En général, j'ai utilisé un mélange de marqueurs de tailles de 250, 150, 100, 75, 50, 37, 25, 20, 15 et 10 kDa (Biorad Precision Plus). J'ai coloré mes gels avec au bleu de Coomassie et décoloré avec de l'acide acétique 7% + éthanol 12%. Les échantillons de protéines ont été dénaturés dans du tampon de Laemmli (Sigma-Aldrich, ~1X final). La migration des gels s'est faite à 20-50 mA par gel et à 200 V.

### 2.2.3.2. Détermination de la concentration de protéines par spectrophotométrie

Dans le cas où j'ai eu des protéines recombinantes contenant au moins un tryptophane ou une tyrosine, j'ai utilisé leur absorbance maximale à 280 nm et le rapport  $DO_{260}/DO_{280}$  (où DO signifie densité optique et le rapport étant un indicateur de pureté de protéines notamment s'il y a une présence d'acides nucléiques, sachant qu'ils ont un maximum d'absorbance à 260 nm) pour faire les mesures de concentrations en protéines grâce à un spectrophotomètre à petit volume (Nanodrop, Thermo Fisher). Cet appareil permet de faire la mesure des absorbances à différentes longueurs d'onde à partir d'échantillon de ~2  $\mu$ L. Le « blanc » a été réalisé avec les mêmes solutions tampons que celles dans lesquelles les protéines étudiées étaient. J'ai évalué le coefficient d'extinction molaire,  $\epsilon$ , de mes protéines à partir des séquences en acides aminés sur le serveur ProtParam (<http://web.expasy.org/protparam/>) afin de remonter à la concentration de mes protéines grâce à la relation de Beer-Lambert ( $A = \epsilon \times [\text{protéine}] \times \text{longueur de trajet optique}$ ). En pratique le logiciel du nanodrop donne l'absorbance à 280 nm rapportée à un trajet optique de 1 cm (= DO).

Pour le suivi des croissances de cultures de bactéries, j'ai utilisé des cuves de 1 mL et mesuré sur un spectrophotomètre classique la DO à 600 nm en faisant le « blanc » avec de l'eau sachant que les milieux de culture utilisés ont une absorption à 600 nm très proche de celle de l'eau.

#### 2.2.3.1. Détermination de la concentration de protéines par test colorimétrique (Test de Bradford)

Dans le cas où les protéines recombinantes étudiées n'avaient pas de tryptophane ou de tyrosine qui absorbe à 280 nm, j'ai utilisé le test colorimétrique de Bradford basée sur le changement de couleur du bleu de Coomassie en présence de protéines qui induisent un maximum d'absorption à 595 nm (couleur bleue) au lieu de ~470 nm (couleur brune).

L'expérience a consisté à ajouter 200 µL de réactif de Bradford + 800 µL eau (sans ajout supplémentaire pour la solution de référence) + 1 à 10 µL de solution de protéine. J'ai ensuite mesuré la DO à 595 nm. J'ai ensuite calculé la concentration massique ( $[\text{protéines}](\text{g/L}) = DO_{595}/0,066$  ; la constante 0,066 avait été calculée à partir d'une gamme étalon) puis la concentration molaire ( $[\text{protéine}](\text{mM}) = [\text{protéine}](\text{g/L}) * 1000/\text{MM}$  ; où MM représente la masse molaire de la protéine).

#### 2.2.3.2. Spectrométrie de masse

Des expériences de contrôles par spectrométrie de masse pour vérifier l'état de certaines de mes protéines ont été réalisées par Jean-Pierre Le Caer (Plateforme de spectrométrie de masse, ICSN Gif-sur-Yvette). Parmi les méthodes d'ionisation utilisées, l'électro nébulisation (ESI) et la désorption-ionisation laser assistée par matrice (MALDI) ont été utilisées. La masse des protéines analysées a été mesurée avec des analyseurs de type orbitrap et temps de vol des ions (TOF), respectivement.

Pour ces expériences, les échantillons ont été dessalés sur colonnes en phase inverse et les échantillons analysés par MALDI ont été cristallisés avec de l'acide sinapinique.

## 2.3. Expériences de GST pulldown

Dans le cadre du projet d'étude de l'interaction entre la protéine hRSV-N et la protéine Tax1BP1, j'ai réalisé des expériences de GST pulldown.

Pour cela, j'ai procédé comme suit :

- Production et purification des domaines d'intérêt de hRSV-N et de Tax1BP1 dont l'un ou l'autre a été purifié en gardant l'étiquette GST fusionnée (pour les contrôles négatifs nous avons aussi utilisé de la GST seule) en récupérant directement les billes sépharose-GSH.
- Les protéines purifiées ont été dialysées dans la solution tampon des expériences de GST pulldown (tampon 20 mM phosphate, 100 mM NaCl, 1 mM TCEP, pH 6,5 ou tampon 20 mM tris, 150 mM NaCl, pH 8,5).

- Quantification des protéines sans GST fusionnée par nanodrop.
- Dénaturation d'échantillons de billes sépharose-GSH + protéines de fusion avec la GST avec de la solution tampon + 4% SDS, agitation au vortex à vitesse maximale et incubation à 95°C sous agitation plusieurs fois (pour décrocher les GST des billes sépharose-GSH), centrifugation à 14k g 2' puis quantification des protéines de fusion par nanodrop.
- Mélange de 100 µM de GST-\* (GST-\* fait référence à la GST avec ou sans partenaire de fusion) sur billes sépharose-GSH avec ou sans la protéine partenaire à différente concentration ainsi qu'une quantité suffisante de solution tampon pour un volume total de v50 µL.
- Agitation au vortex à vitesse moyenne.
- Incubation à 25°C pendant 1 h sous légère agitation à 400 rpm (re-passage au vortex à mi-temps et à la fin).
- Centrifugation 5' à 500 g à température ambiante.
- Elimination du surnageant.
- Lavage avec 500 µL de solution tampon, resuspension des billes au vortex à vitesse moyenne, centrifugation 5' à 500 g à température ambiante, élimination du surnageant (2 à 4 fois).
- Incubation des billes lavées avec 30 µL de tampon Laemmli 2x.
- Incubation à 95°C pendant 5'.
- Dépôt sur gel SDS-PAGE 15% et migration à 20 mA par gel puis révélation au bleu de Coomassie.

## 2.4. Utilisation de la RMN pour l'études de protéines

### 2.4.1. Acquisition de spectres RMN d'échantillons de protéines

Pour préparer les échantillons à analyser par RMN, j'ai prélevé 185 µL de solution d'intérêt que j'ai mélangée avec 15 µL de D<sub>2</sub>O (pour le verrouillage du champ magnétique) + 0,5 µL de DSS(acide 4,4-diméthyl-4-silapentane-1-sulfonique) à différentes concentration (50-200 mM en fonction de la concentration en protéines en solution) pour la calibration des déplacements chimiques dans la dimension proton. J'ai ensuite éliminé d'éventuels agrégats et/ou bulles par centrifugation à ~14k g uniquement ou en utilisant des tubes avec des filtres ayant des pores de 0,2 µm. Puis, j'ai inséré l'échantillon dans des tubes RMN de 3 mm de diamètre.

Sur le spectromètre RMN, j'ai modifié la température de l'appareil, inséré mon tube pré positionné dans le spinner, accordé la sonde, ajusté l'homogénéité du champ magnétique, calibré l'impulsion 90° <sup>1</sup>H pour les expériences à enregistrer. Enfin, j'ai affiné certains paramètres d'acquisition si nécessaires (impulsions autres que <sup>1</sup>H, gain, nombre de scans, nombre d'incrémentations de temps,

fenêtres d'enregistrement par dimension, impulsion de découplage  $^{13}\text{C}/^{15}\text{N}$  ou autre) avant de lancer les enregistrements d'expériences.

## 2.4.2. Attribution séquentielle des fréquences de résonance

L'attribution des déplacements chimiques d'une protéine permet l'analyse et l'interprétation d'expériences RMN qui permettent d'accéder à des informations telles que la structure, la dynamique et les interactions d'une protéine à l'échelle du résidu. Elle est donc essentielle dans ce projet.

### 2.4.2.1. Attribution des fréquences de résonance du squelette peptidique

Le principe de l'attribution du squelette peptidique repose sur les couplages scalaires entre les noyaux  $^1\text{H}$ ,  $^{13}\text{C}$  et  $^{15}\text{N}$  présents dans une protéine uniformément marquée  $^{13}\text{C}/^{15}\text{N}$ . L'objectif premier est d'attribuer à chaque pic du spectre de corrélation 2D HSQC  $^1\text{H}-^{15}\text{N}$  (Heteronuclear Single Quantum Coherence) un acide aminé qui en est la source (Kay et al., 1990). Le spectre HSQC  $^1\text{H}-^{15}\text{N}$  corréle les protons et azotes liés par une liaison amide et constitue une empreinte digitale pour une protéine.

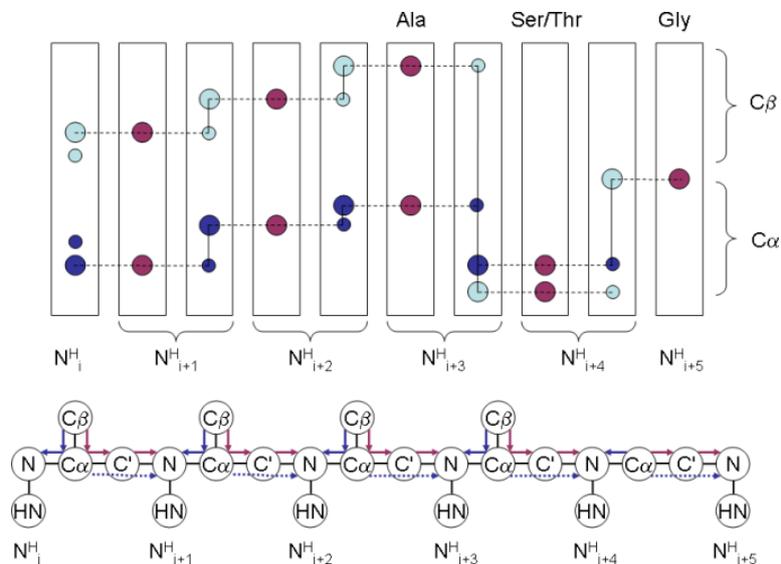
Pour ce faire, on doit acquérir des spectres triple résonances ( $^1\text{H}$ ,  $^{13}\text{C}$  et  $^{15}\text{N}$ ) en 3 dimensions (3D). Ainsi, le spectre HSQC  $^1\text{H}-^{15}\text{N}$  se trouve prolongé dans une troisième dimension et pour chaque coordonnée ( $^1\text{H}$ ,  $^{15}\text{N}$ ) d'un pic on peut trouver jusqu'à cinq fréquences de résonance (ou déplacements chimiques) correspondant aux  $^{13}\text{C}_\alpha$  et  $^{13}\text{C}_\beta$  du même acide aminé, ainsi qu'au  $^{13}\text{C}$  du groupe carbonyle et aux  $^{13}\text{C}_\alpha$  et  $^{13}\text{C}_\beta$  de l'acide aminé précédent.

Pour cela, on acquiert une série de spectres 3D ayant des dimensions  $^1\text{H}$  et  $^{15}\text{N}$  identiques à la HSQC, mais des dimensions carbones différentes :

- HNCO : cette expérience permet de relier le groupe amide ( $^1\text{H}$  et  $^{15}\text{N}$ ) d'un acide aminé  $i$  avec le groupe carbonyle de l'acide aminé précédent en  $i-1$ .
- HNCA : elle relie le groupe amide d'un résidu  $i$  avec les  $\text{C}_\alpha$  d'un même résidu  $i$  et celui en  $i-1$ .
- HN(CO)CA : elle relie le groupe amide de l'acide aminé  $i$  avec le  $\text{C}_\alpha$  de l'acide aminé en  $i-1$ .
- HNCACB : elle relie le groupe amide d'un résidu  $i$  avec les  $\text{C}_\alpha$  et  $\text{C}_\beta$  du résidu  $i$  et ceux en  $i-1$ .
- HN(CO)CACB ou CBCA(CO)NH : elles relient le groupe amide de l'acide aminé  $i$  avec les  $\text{C}_\alpha$  et  $\text{C}_\beta$  de l'acide aminé en  $i-1$ .

A partir de là, on cherche à relier les systèmes de spins entre eux. Pour cela, on utilise le fait que les fréquences de résonance  $^1\text{H}$  et  $^{15}\text{N}$  du groupe amide d'un acide aminé  $i$  peuvent être reliées à ses  $\text{C}_\alpha$  et  $\text{C}_\beta$ , mais aussi aux  $\text{C}_\alpha$  et  $\text{C}_\beta$  de l'acide aminé en  $i-1$ . En partant du système de spins correspondant à un résidu  $i$   $\{^1\text{H}(i), ^{15}\text{N}(i), ^{13}\text{C}_\alpha(i), ^{13}\text{C}_\beta(i)\}$  et des valeurs  $\{^{13}\text{C}_\alpha(i-1), ^{13}\text{C}_\beta(i-1)\}$  on cherche tous les systèmes de spins ayant des fréquences de résonance  $\{^{13}\text{C}_\alpha(i), ^{13}\text{C}_\beta(i)\}$  correspondant aux fréquences de résonance ( $i-1$ ). On trouve parfois plusieurs acides aminés possibles (la largeur de raie des signaux ne permettant pas de les différencier). Mais, il existe d'autres informations pour nous permettre de relier

les différentes pièces de ce puzzle. Il s'agit de la séquence d'acides aminés qui est connue, ainsi que des déplacements chimiques remarquables de certains acides aminés tels que l'alanine, la glycine, la sérine ou la thréonine. Le principe de l'attribution séquentielle des fréquences de résonance du squelette peptidique d'une protéine est illustré dans la **Figure 14**.



**Figure 14 : Principe de l'attribution séquentielle des fréquences de résonance du squelette peptidique d'une protéine comportant un enchaînement Ala-Ser-Gly (<http://www.protein-nmr.org.uk>).**

Le travail d'attribution des signaux ainsi que de la plupart des analyses des spectres a été fait avec le logiciel d'analyse de données de RMN des protéines CcpNmr Analysis (Vranken et al., 2005).

#### 2.4.2.2. Attribution des fréquences de résonance des chaînes latérales

Ici, le but est d'attribuer à chaque pic du spectre HSQC  $^1\text{H}$ - $^{13}\text{C}$  les paires ( $^1\text{H}$ ,  $^{13}\text{C}$ ) d'atomes aliphatiques qui y correspondent. Le principe repose cette fois sur les couplages scalaires entre les noyaux  $^1\text{H}$  et  $^{13}\text{C}$  présents dans une protéine uniformément marquée  $^{13}\text{C}^{15}\text{N}$ . Il existe plusieurs méthodes d'attribution des déplacements chimiques des chaînes latérales, mais je présente ici uniquement celle que j'ai majoritairement utilisées.

Dans un premier temps, on doit compléter l'attribution du squelette peptidique par les déplacements chimiques  $^1\text{H}_\alpha$  et  $^1\text{H}_\beta$ . On se sert du spectre 3D HBHA(CO)NH qui permet d'observer les corrélations entre le groupe amide d'un résidu  $i$ , observé sur le spectre HSQC  $^1\text{H}$ - $^{15}\text{N}$ , et les  $^1\text{H}_\alpha$  et  $^1\text{H}_\beta$  du résidu  $i-1$  via les couplages scalaires. Les attributions ont été vérifiées sur un spectre 3D  $^{15}\text{N}$ -NOESY-HSQC (Nuclear Overhauser Effect Spectroscopy) qui permet d'observer les corrélations entre les groupes amides observés sur le spectre HSQC  $^1\text{H}$ - $^{15}\text{N}$  et les protons  $^1\text{H}$  dans une troisième dimension, grâce à l'effet nOe qui repose sur les couplages dipolaires, à travers l'espace, entre protons. Ainsi, pour chaque coordonnée ( $^1\text{H}$ ,  $^{15}\text{N}$ ) d'un pic HSQC  $^1\text{H}$ - $^{15}\text{N}$  on peut trouver des résonances correspondant aux  $^1\text{H}$  amides ou aliphatiques ( $^1\text{H}_\alpha$ ,  $^1\text{H}_\beta$ ,  $^1\text{H}_\gamma$ , etc) pourvu qu'ils soient proches dans l'espace du proton amide

observé (entre 0 et 6 Å). Pour la vérification de l'attribution des protons  $^1\text{H}_\alpha$  ce sont les corrélations entre les  $^1\text{H}_\text{N}$  d'un résidu  $i$  et les  $^1\text{H}_\alpha$  du même résidu et du résidu  $i-1$  qui ont été utilisées.

Ensuite, on utilise la séquence 3D HCCH-TOCSY (TOCSY : Total Correlation Spectroscopy). Ainsi, le spectre HSQC  $^1\text{H}$ - $^{13}\text{C}$  se trouve prolongé dans une troisième dimension et pour chaque coordonnée ( $^1\text{H}$ ,  $^{13}\text{C}$ ) d'un pic on peut trouver toutes les fréquences de résonance correspondant aux  $^{13}\text{C}$  aliphatiques contigus du même acide aminé, grâce aux couplages scalaires longues distances et un temps de mélange pendant lequel on applique un verrouillage de champ magnétique (20-40 ms). Avec ce spectre, grâce aux données initiales  $\{^1\text{H}_\alpha(i), ^1\text{H}_\beta(i), ^{13}\text{C}_\alpha(i), ^{13}\text{C}_\beta(i)\}$  pour un spin  $i$  et les tables de référence de déplacements chimiques des  $^1\text{H}$  et  $^{13}\text{C}$  aliphatiques pour chaque type de résidu, on peut alors attribuer les autres corrélations sur le spectre HSQC  $^1\text{H}$ - $^{13}\text{C}$ .

Il reste toutefois le méthyle en position  $\epsilon$  de la méthionine qui est séparé des autres groupements  $\text{CH}_n$  par un hétéroatome et qui ne peut donc pas être attribué de cette manière. Cependant, les fréquences de résonances  $^1\text{H}$  et  $^{13}\text{C}$  sont remarquables et le signal peut être attribué par la suite grâce aux nOe.

### 2.4.3. Titrage par perturbation des déplacements chimiques

Dans le but d'étudier des interactions entre 2 protéines, j'ai notamment utilisé la RMN en suivant l'évolution des spectres 2D HSQC pour différents ratios ligand/protéine.

Pour ce faire, à partir d'une protéine au moins marquée  $^{15}\text{N}$  j'ai utilisé une concentration constante de cette dernière que j'ai mélangée avec différentes concentrations de protéine partenaire non marquée isotopiquement. J'ai ajouté la protéine partenaire soit lyophilisée dans de l'eau sans tampon ni sel soit en mélangeant des volumes de solutions de protéine non marquée et de protéine marquée dans les mêmes solutions tampons en faisant en sorte de rester à concentration constante en protéine marquée. Après homogénéisation des solutions et insertion dans des tubes RMN de 3 mm, j'ai enregistré des expériences HSQC  $^1\text{H}^{15}\text{N}$  (ou  $^1\text{H}^{13}\text{C}$  dans certains cas) avec les mêmes paramètres d'acquisition ou de traitement tout au long du titrage. Après avoir traité les spectres RMN, j'ai attribué chaque signal et analysé les déplacements chimiques (protons, hétéro atomes et combinés) et les intensités des signaux en fonction du rapport protéines partenaires non marquées / protéines marquées.

### 2.4.4. Mesure des vitesses de relaxation de spin nucléaire

Pour étudier la dynamique interne de protéines, j'ai mesuré des vitesses de relaxation de spin nucléaire  $^{15}\text{N}$  longitudinale ( $1/T_1$ ) et transversale ( $1/T_2$ ) ainsi que des NOE hétéronucléaires  $^1\text{H}$ - $^{15}\text{N}$  avec des séquences d'impulsion standard. Des acquisitions avec différents temps délais de relaxation ont été enregistrées pour les vitesses de relaxation  $^{15}\text{N}$ . Les données ont été acquises comme des spectres « pseudo » 3D où la 3<sup>ème</sup> dimension correspond aux différents délais. Les signaux pour chaque délai ont

été attribués, puis les intensités ont été ajustées avec des courbes exponentielles en fonction du délai de relaxation ( $t$ ) selon l'équation :  $I(t) = I_0 \times e^{-\frac{t}{T_{1,2}}}$ , en faisant varier  $I_0$  et les temps de relation  $T_1$  et  $T_2$ .

Pour les expériences NOE hétéronucléaires, deux spectres avec saturation on et off résonance ont été extraits d'une seule expérience enregistrée. Après attributions des signaux des deux spectres, j'ai calculé le ratio des intensités des signaux des spectres on résonance / off résonance.

L'analyse des signaux et des intensités ainsi que les ajustements d'exponentielles des intensités des spectres ont été fait avec le logiciel CcpNmr Analysis.

#### 2.4.5. Augmentation de la relaxation par effet paramagnétique

Dans la continuité des études de relaxation, j'ai réalisé des expériences d'augmentation de la relaxation par effet paramagnétique (PRE) par RMN sur des mutants de la protéine P du RSV. Cette expérience consiste à positionner une sonde paramagnétique sur un résidu de la protéine à étudier qui diminuera le temps de relaxation  $T_2$  et donc l'intensité des signaux des atomes proches ( $\sim 15$  Å) de la sonde. Les PRE permettent de détecter notamment des contacts transitoires à longue distance dans les protéines de type IDP (Lietzow et al., 2002).

J'ai utilisé comme sonde paramagnétique le radical libre 3-(2-iodoacetamido)-PROXYL (IAP) qui peut se fixer sur des cystéines libres (Gillespie et Shortle, 1997). La séquence de P sauvage ne contenant pas de cystéines. Nous avons donc muté certaines sérines ou glutamates en cystéines.

De manière générale, comme pour les autres protéines recombinantes, j'ai produit et purifiée mes protéines mutantes mais j'ai ajouté dès le début de la lyse des bactéries jusqu'en fin de purification 5 à 10 mM de DTT (dithiothréitol) dans mes solutions tampons afin d'éviter l'oxydation des cystéines.

Après avoir purifié et dialysé les protéines, j'ai concentré mes échantillons dans des volumes de 200  $\mu$ L. J'ai passé mes échantillons au travers de micro-colonnes d'exclusion de taille (Bio-Rad) conditionnées dans un tampon 50 mM Tris/ 200 mM NaCl/ pH 8 afin d'éliminer le DTT et d'équilibrer la protéine dans un tampon adéquat pour la réaction avec l'IAP sur les cystéines libres. Rapidement, j'ai incubé les échantillons à l'abri de la lumière à 15 °C sous agitation à  $\sim 300$  rpm avec  $\sim 10$  équivalents molaires d'IAP durant une nuit. Le lendemain, j'ai passé trois fois les échantillons dans les colonnes conditionnées dans du tampon phosphate 20mM/ NaCl 100 mM/ pH 6,5 sans DTT afin d'éliminer l'IAP en excédent et de conditionner les protéines dans le tampon pour les acquisitions d'expériences RMN. Après l'acquisition de spectres HSQC  $^1\text{H}$ - $^{15}\text{N}$ , j'ai réduit les étiquettes IAP avec 10 équivalents molaires de vitamine C à température à  $\sim 25$  ° C sur une nuit. Enfin, j'ai enregistré les spectres HSQC  $^1\text{H}$ - $^{15}\text{N}$  des protéines avec IAP réduite (+ spectres 3D pour la réattribution de certains signaux si des signaux étaient trop décalés par rapport à la protéine sauvage).

J'ai ensuite attribué les signaux amides sur les spectres HSQC  $^1\text{H}$ - $^{15}\text{N}$  des formes paramagnétiques de mutants P-IAP et diamagnétique (IAP réduit), puis j'ai calculé les intensités normalisées de chaque signal amide ( $I_{\text{IAP paramagnétique}}/I_{\text{IAP diamagnétique}}$ ).

## 2.5. Logiciels d'analyse de structure et de dynamique de protéines utilisés

En dehors du cadre des informations d'interactions et de dynamique des protéines explicitées dans les parties précédentes, je me suis servi des déplacements chimiques attribués sur le squelette peptidique ( $^1\text{H}_\text{N}$ ,  $^{15}\text{N}_\text{H}$ ,  $^{13}\text{C}_\alpha$ ,  $^{13}\text{C}_\beta$ ,  $^{13}\text{C}'$  et  $^1\text{H}_\alpha$ ) de protéines étudiées pour prédire leurs angles dièdres  $\phi$  et  $\psi$  associés à l'aide du serveur TALOS-N (Shen et Bax, 2013). J'ai également acquis plusieurs expériences de type NOESY HSQC  $^1\text{H}$ - $^{15}\text{N}$  ou  $^1\text{H}$ - $^{13}\text{C}$  dont les intensités, grâce à la relation  $I_{\text{NOE}} = \frac{k}{r_{ij}^6}$  (où  $k$  est une constante inhérente à chaque expérience NOESY et  $r_{ij}$  est la distance séparant un atome  $i$  d'un atome  $j$ ), m'ont donné indirectement accès à des informations de courtes distances inter atomes ( $\leq 5 \text{ \AA}$ ). Cependant, à cause de superpositions de signaux il n'est pas toujours aisé d'attribuer les protons impliqués dans les NOE.

Pour attribuer les NOE de manière automatique et calculer des structures tridimensionnelles de protéines (à l'aide également des contraintes d'angles dièdres) j'ai utilisé les logiciels CYANA (Güntert et al., 1997 ; Herrmann et al., 2002) et ARIA (Rieping et al., 2007).

### 2.5.1. CYANA

Le logiciel CYANA (algorithme combiné d'attribution et de dynamique pour les applications de RMN) utilise un algorithme de dynamique dans l'espace des angles de torsion pour faire du recuit simulé de protéines et fonctionne par défaut avec 7 cycles de calcul afin de faire converger de manière itérative la structure à partir des données fournies en essayant d'obtenir le meilleur ensemble d'attributions possibles pour les pics NOE.

Après la première itération, les structures de départ du cycle en cours correspondent aux meilleurs résultats (notamment en termes de satisfaction des contraintes de distance) du cycle précédent. Le logiciel utilise les structures de l'étape précédente, les déplacements chimiques avec une tolérance indiquée dans les paramètres initiaux et le nombre de liaisons covalentes séparant les atomes pour attribuer les contraintes NOE en satisfaisant ces dernières au mieux sans prendre de « risque » ainsi qu'en diminuant progressivement le nombre de possibilités d'attribution par contrainte NOE. Lors de la dernière itération, un filtre conserve uniquement les NOE avec une attribution unique. Une fonction cible tendant à être minimisée rend compte à chaque cycle d'éventuelles violations de contraintes de distance ou simplement de son éloignement par rapport aux distances moyennes des structures mesurées durant le cycle.

Je me suis servi de CYANA dans le cas de calcul de monomères ou de protomères.

### 2.5.2. ARIA

Le logiciel ARIA (contraintes ambiguës pour attribution itérative) permet aussi l'attribution automatisée de NOE et un calcul de structure d'ensemble de conformères par recuit simulé dans l'espace des angles de torsions ou l'espace cartésien avec 8 itérations ainsi qu'une étape d'affinement de dynamique moléculaire dans un solvant explicite.

Après conversion des intensités ou volumes des pics NOE en contraintes de distance à chaque cycle (après calibration de la constante reliant les intensités NOE et les distances au 6<sup>ème</sup>), attribue les contraintes de distances à un ou plusieurs protons. En plus des contraintes non ambiguës (une attribution pour une contrainte), ARIA utilise des attributions ambiguës (avec plusieurs possibilités) où toutes les possibilités d'affectation contribuent à la contrainte de distance calculée mais avec une pondération qui évolue au cours des itérations. Cependant, la plupart des attributions sont incohérentes et ne peuvent donc pas être remplies simultanément par une même structure. Durant le protocole itératif, le logiciel commence chaque cycle par une correction de la liste de contraintes en filtrant les attributions et les pics de bruit improbables avec les meilleures structures du cycle précédent puis les contraintes de distances redondantes restantes sont convergées avant de lancer le calcul d'un nouvel ensemble de structure.

Parmi ses nombreuses fonctions, ARIA permet de calculer des multimères sans ou avec symétries de type  $C_n$  ou  $D_2$ . C'est d'ailleurs une des fonctionnalités du logiciel que j'ai utilisées.

### 2.5.3. Flexible-meccano

Flexible-meccano est un logiciel qui se tourne plutôt vers l'étude des protéines intrinsèquement désordonnées.

Le logiciel permet de calculer des ensembles conformationnels sans champs de force mais à partir d'une base de données d'angles  $\phi$  et  $\psi$  recensés dans des structures cristallographiques à haute résolutions sur des résidus non impliqués dans des structures secondaires. Il est possible de fournir au logiciel des contraintes d'angles  $\phi / \psi$  avec une erreur donnée et une propension donnée. De même, des structures secondaires avec une propension donnée ainsi que des contacts entre au moins deux  $C_\alpha$  dans deux régions distinctes avec une propension et une distance données peuvent être définis dans le logiciel. Ensuite, flexible-meccano se charge de générer un ensemble de  $N$  conformères pour lesquels il peut ensuite calculer les RDC (couplages dipolaires résiduels), les couplages scalaires entre  $H_N$  et  $H_\alpha$  ( $^3J_{NH\alpha}$ ), les PRE (augmentation de la relaxation par effet paramagnétique), et le diagramme de Ramachandran correspondant. Le logiciel peut également, à l'aide d'autres logiciels, calculer les déplacements chimiques  $^{13}C_\alpha$ ,  $^{13}C_\beta$ ,  $^{13}C'$  et  $H_\alpha$  (notamment avec le logiciel SPARTA) pour chaque conformère ainsi les courbes de SAXS (avec le logiciel CRY SOL) moyennées sur l'ensemble des conformères.

En superposant les données calculées sur l'ensemble de conformères générés par le logiciel avec les données expérimentales il est donc possible d'avoir une meilleure compréhension de la dynamique et de la structuration des protéines étudiées.



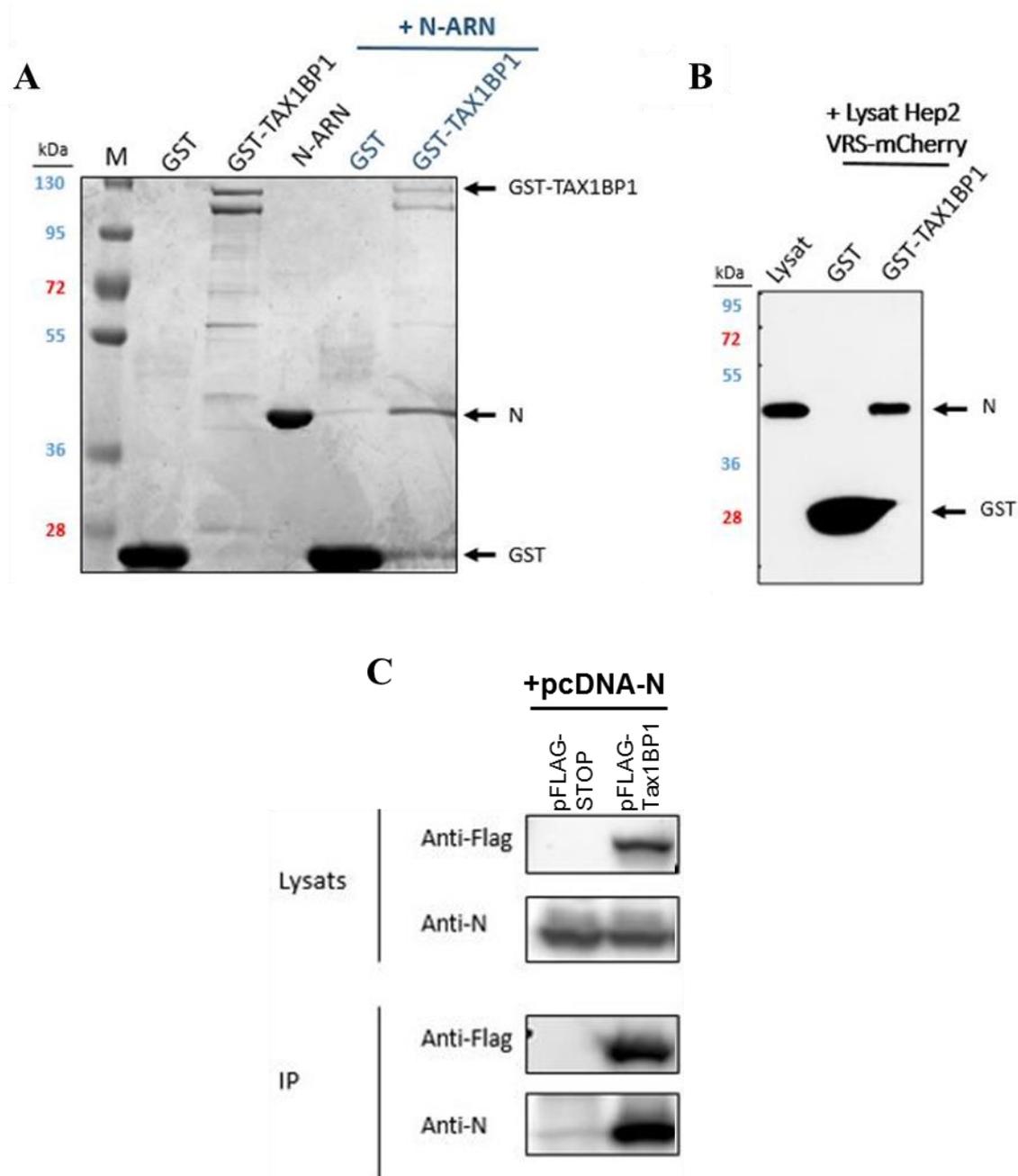
# Chapitre 3. Etude de l'interaction entre la nucléoprotéine du hRSV et la protéine cellulaire Tax1BP1

## 3.1. Découverte de l'interaction entre la protéine N du HRSV et la protéine Tax1BP1

Dans le cadre d'une étude approfondie de l'interactome de la protéine N du RSV avec le protéome humain, un crible double-hybride en levure a été réalisé par le Dr Pierre-Olivier VIDALAIN (Institut Pasteur) en collaboration avec le Dr Marie GALLOUX (VIM, INRA Jouy-en-Josas). Le crible a été entrepris à partir de la forme mutante monomérique de la protéine N (N<sup>0</sup>) comme appât sur deux banques d'ADNc : ORFeome Humain (<http://horfdb.dfci.harvard.edu/hv7/>) et une librairie d'ADNc du foie humain (Life Technologies). Parmi les différents partenaires cellulaires identifiés, la protéine Tax1BP1 a été isolée à partir de 40 clones différents. Ces résultats suggèrent donc une interaction directe entre N et Tax1BP1.

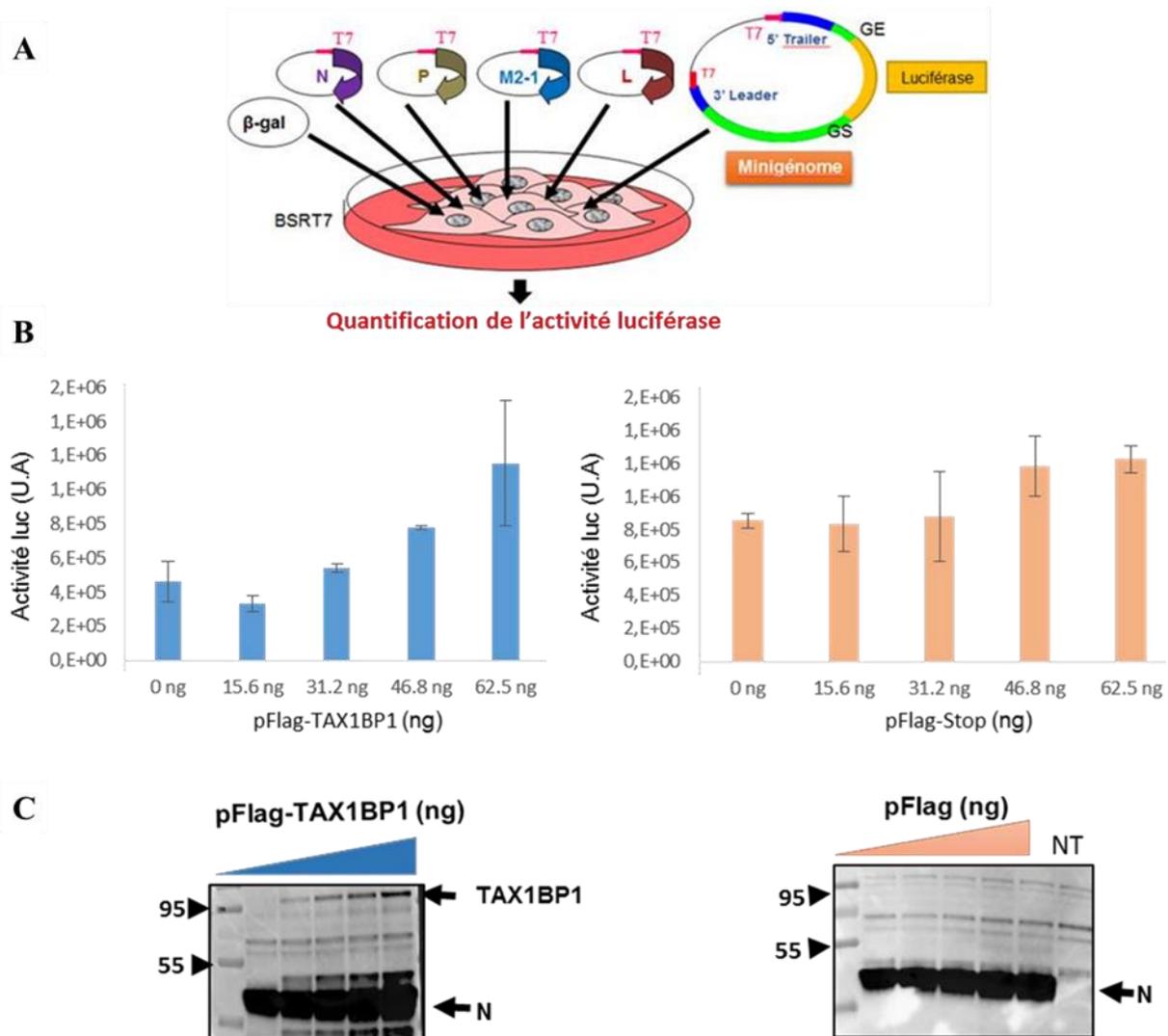
L'interaction directe entre N, à la fois pour la forme monomérique N $\Delta$ 30 sans ARN et pour la forme N entière avec ARN et (**Figure 18**), et Tax1BP1 a ensuite été confirmée par des expériences de GST pulldown et des expériences d'immunoprécipitation effectuées par le Dr Marie GALLOUX à l'INRA et plusieurs de ses stagiaires : Vanny KOUBANA (M2, 2015-2016), Venice GRAF (AgroParisTech 2<sup>e</sup> année, 2015-2016) et Sarah MADRIERES (M2, 2016-2017) (**Figure 15**).

Tous les résultats présentés dans le paragraphe 3.1 ont été obtenus dans l'équipe du Dr Marie GALLOUX et n'ont pas encore été publiés.



**Figure 15 : Validation de l'interaction directe entre TAX1BP1 et N par GST pulldown et immunoprécipitation.** (A) Analyse sur gel acrylamide 12% coloré en bleu de Coomassie des protéines GST, GST-TAXBP1, et N recombinantes, et des échantillons correspondant au test de pulldown de N en présence de GST ou GST-TAX1BP1. (B) Des lysats cellulaires Hep2 infectés par le RSV-mCherry pendant 48h ont été incubés en présence de billes GST seule ou GST-TAX1BP1 sur la nuit à 4°C. La présence de N associée aux billes a été analysée par western blot avec un anticorps anti-N. (M : Marqueur de poids moléculaire). A noter que la bande correspondant à la protéine GST est une détection non spécifique par les anticorps. (C) Les cellules ont été co-transfectées avec les plasmides codant pour pFLAG-TAX1BP1 et pcDNA-N. Une immunoprécipitation en présence d'anticorps anti-FLAG a ensuite été réalisée sur les lysats cellulaires. Les protéines TAX1BP1 et N ont été révélées par western blot en présence d'anticorps anti-FLAG-HRP et anti-N. (Rapport de stage de M2 Sarah MADRIERES)

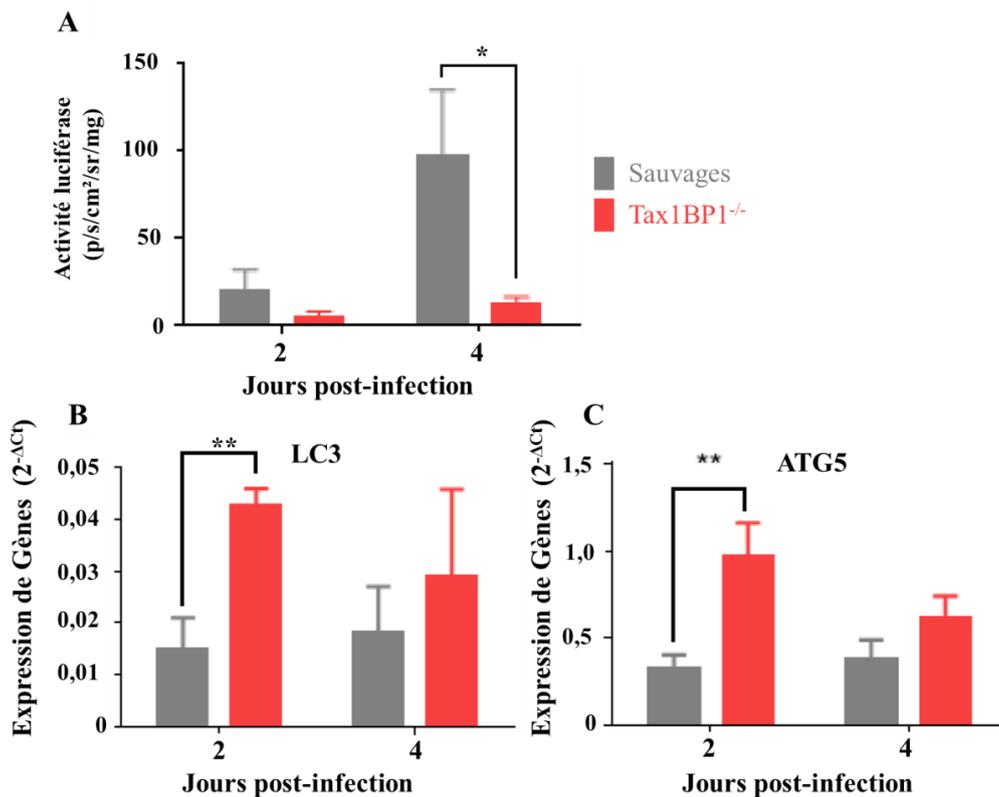
Etant donné que la protéine Tax1BP1 a une implication notable dans la régulation de l'immunité innée et de l'autophagie, des études fonctionnelles de cette protéine et de son impact sur le RSV ont été menées afin de vérifier l'intérêt de cibler l'interaction entre N et Tax1BP1 comme cible thérapeutique potentielle. Dans ce cadre, des expériences de quantification de l'activité polymérase du RSV à l'aide d'un minigénome (Tran et al., 2009) en présence d'une quantité croissante de protéine Tax1BP1 ont prouvé que cette dernière a un effet positif sur l'activité ARN polymérase du RSV (fait par Vanny Koubana sous la tutelle de Marie Galloux à l'INRA ; **Figure 16**).



**Figure 16 : Etude de l'impact de Tax1BP1 sur l'activité polymérase.** (A) Principe du test fonctionnel de minigénome pour mesurer l'activité du complexe polymérase : l'activité luciférase est proportionnelle à l'activité du complexe polymérase et quantifiée. (B) Test minigénome en présence de concentrations croissantes de pFlag-Tax1BP1 ou pFlag. L'activité luciférase a été quantifiée 24h post-transfection. (C) Analyse de l'expression de Tax1BP1 et N par western blot. (Rapport de stage de M2 Vanny KOUBANA)

A contrario, la déplétion de TAX1BP1 conduirait quant à elle à une diminution de la réplication virale et une augmentation de la quantité de protéine N (données non montrées). D'autres essais ont été

faits *in vivo*, au cours du stage de M2 de Sarah MADRIERES, avec des souris sauvages ou KO pour le gène Tax1BP1 (TAX1BP1<sup>-/-</sup>). Ils ont permis de mettre en évidence un défaut de réplication du RSV dans les poumons des souris TAX1BP1<sup>-/-</sup> par rapport aux animaux sauvages (**Figure 17A**). D'autre part l'infection par le RSV de macrophages alvéolaires issus de souris TAX1BP1<sup>-/-</sup> induirait l'expression de molécules majeures de la machinerie d'autophagie que sont Lc3 et Atg5 (**Figure 17B et C**). Pour information, Lc3 (light chain 3), est une protéine soluble qui au cours de l'autophagie est capable de se conjuguer à une phosphatidyléthanolamine pour ensuite être recrutée dans la membrane d'un autophagosome (Tanida et al., 2008). Atg5 (Autophagy related) est une protéine constitutivement liée à la protéine Atg12 qui, ensemble, permettent la modification lipidique de LC3 (Otomo et al., 2013).

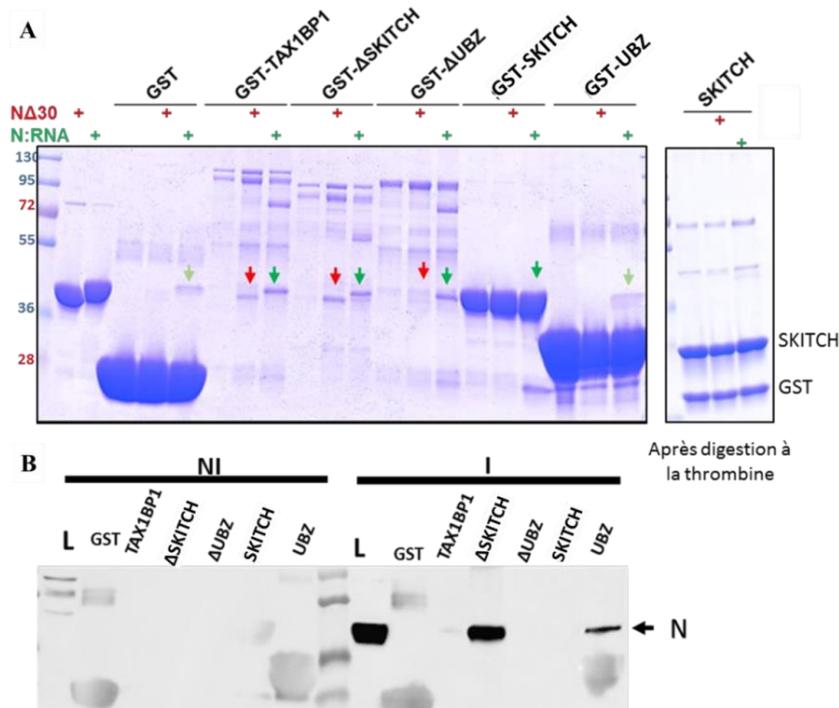


**Figure 17 : Etude *in vivo* de la déplétion de TAX1BP1 sur l'infection par le RSV chez la souris.** (A) Des poumons de souris sauvages ou TAX1BP1<sup>-/-</sup> infectées par le RSV-luciférase ont été lysés et l'activité luciférase a été quantifiée (B, C). Quantification de l'expression des gènes de l'autophagie à partir de cellules (majoritairement des macrophages) issues de lavages broncho-alvéolaires de souris sauvages et TAX1BP1<sup>-/-</sup> infectées. L'expression des gènes codant pour des protéines impliquées dans l'autophagie, LC3 et ATG5 a été quantifiée par qPCR. Les mesures ont été réalisées en duplicats. Les résultats ont été normalisés par rapport au gène de ménage Hprt en calculant le Ct (Ct moyen (gène étudié) – Ct moyen (Hprt)). Les moyennes sont représentées en histogramme. Le nombre d'animaux testés est n = 4 souris sauvages infectées par le RSV et n = 2 souris TAX1BP1<sup>-/-</sup> infectées par le RSV. Ct correspond à l'acronyme cycle threshold (seuil de cycle) pour chaque gène. Cette valeur correspond au nombre de cycles nécessaires de qPCR pour que la quantité d'ADNc présente dans l'échantillon dépasse un seuil fixé manuellement. (Rapport de stage de M2 Sarah MADRIERES)

Ces résultats préliminaires suggèrent que suite à une infection par le RSV, la protéine TAX1BP1 exercerait une fonction régulatrice de la biogénèse des autophagosomes des macrophages alvéolaires, comme cela a pu être observé dans des cellules épithéliales intestinales infectées par *Salmonelle Typhimurium* (Tumbarello et al., 2015). La délétion de TAX1BP1 entraînant de façon significative l'augmentation des transcrits de gènes de l'autophagie suite à une infection par le RSV, l'interaction de la protéine virale N avec TAX1BP1 apparaît comme un moyen pour le RSV de limiter l'induction de l'autophagie afin de limiter la réponse antivirale de l'hôte (stage de M2 de Sarah MADRIERES, INRA, 2016-2017).

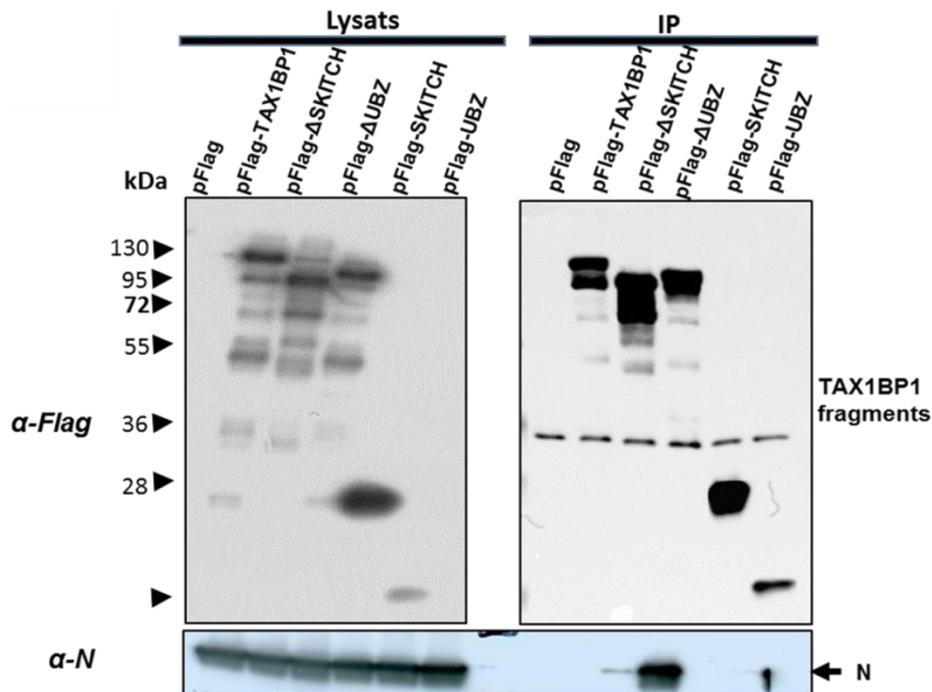
### 3.2. Définition de fragments de nucléoprotéine et de Tax1BP1 comme modèles d'étude

Afin d'obtenir plus d'indications sur le ou les domaines de Tax1BP1 impliqués dans l'interaction avec la protéine N, Vanny KOUBANA (stagiaire M2 du Dr Marie GALLOUX, INRA, 2015-2016) a entrepris des expériences de GST pulldown (Figure 18) et d'immunoprécipitation (Figure 19) avec différents fragments de Tax1BP1 et la protéine N entière ou N $\Delta$ 30 (seulement par GST pulldown).



**Figure 18 : Identification du domaine d'interaction de la protéine Tax1BP1 avec la protéine N par GST pulldown.** (A) Les protéines recombinantes GST-fragments de Tax1BP1 purifiés sur billes ont été incubées en présence de N entière ou N $\Delta$ 30 pendant 4 h. Les échantillons ont été analysés sur un gel d'acrylamide coloré au bleu de Coomassie. (B) Les cellules HEP2 ont été infectées par le RSV (MOI=2 pendant 48 h). Les lysats des cellules non infectées (NI) et infectées (I) ont été incubés avec les billes de GST-fragments de Tax1BP1. La protéine N a été révélée par western blot avec un anticorps anti-protéine N. (Rapport de stage de M2 Vanny KOUBANA)

Les résultats des expériences de GST pulldown (**Figure 18**) montrent une interaction avec N—ARN (N entière) et NΔ30 pour la protéine Tax1BP1 entière et le fragment Tax1BP1ΔSKITCH. Dans la **Figure 18A**, le fragment ΔUBZ semble présenter une interaction seulement avec N entière alors que pour le fragment UBZ il y a une quantité tellement faible de N entière que cela pourrait être non spécifique (voir contrôle négatif) quand on met en présence seulement les partenaires deux à deux. Cependant, dans le contexte d’une expérience de pulldown en lysat de cellule infectées par le RSV c’est l’inverse : le fragment UBZ semble se fixer à la protéine N entière et le fragment ΔUBZ non. Avec le domaine SKITCH, il ne semble pas y avoir la moindre trace d’une éventuelle présence de N entière ou NΔ30.



**Figure 19 : Identification du domaine d’interaction de la protéine Tax1BP1 avec la protéine N par immunoprécipitation avec une étiquette Flag.** Les fragments de Tax1BP1 ont été clonés dans un vecteur pFlag permettant une expression en cellules eucaryotes. Immunoprécipitation de la protéine N dans un contexte de surexpression et des fragments de Tax1BP1. Les produits de l’expérience ont été analysés en utilisant un anticorps anti-N et un anticorps anti-Flag HRP.

Sur la **Figure 19**, on observe que Tax1BP1 entière et le fragment ΔSKITCH présentent une interaction avec la protéine N. Concernant la piste avec le domaine UBZ de Tax1BP1, il semble y avoir une trace de la protéine N entière et par conséquent une potentielle interaction avec cette dernière, tout comme avec l’expérience de GST pulldown.

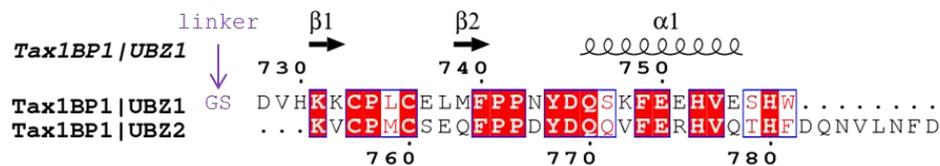
Concernant la recherche d’un domaine d’interaction de la protéine N avec Tax1BP1, par GST pulldown avec un lysat de cellules et des billes GST-N(C arm), la protéine Tax1BP1 a été détectée par spectrométrie de masse uniquement (données non montrées). Par conséquent, la suite du travail consistera à étudier l’interaction entre le domaine N(C arm) et le domaine Tax1BP1(UBZ1+2).

### 3.3. Production et purification des domaines C-terminaux de la nucléoprotéine et de la protéine Tax1BP1

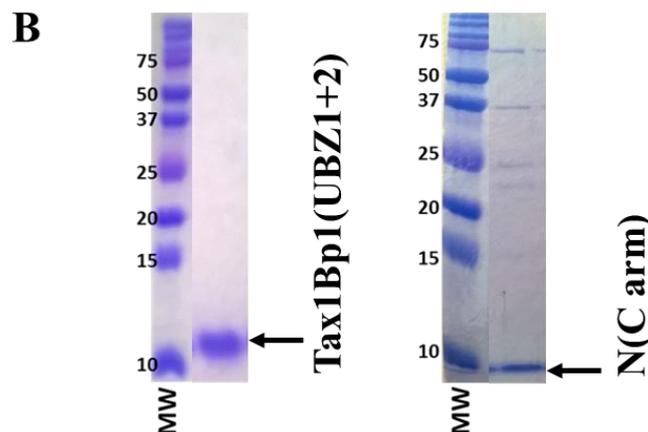
Les domaines N(C arm) et Tax1BP1(UBZ1+2) ont été produits dans *E. coli* sous forme de protéines recombinantes avec des étiquettes GST, puis purifiés et les étiquettes GST ont été clivées à la thrombine comme indiqué dans le Chapitre 2 (Matériels et Méthodes). Il est à noter que pour le domaine Tax1BP1(UBZ1+2), qui contient deux doigts de zinc, j'ai ajouté du zinc lors de l'induction, puis lors de la lyse, ainsi que du TCEP tout au long de la purification de la protéine et durant les expériences avec le domaine indiqué.

Les 2 fragments sont de très petite taille (**Figure 20**) et donc pratiques à étudier par RMN.

#### A Tax1BP1(UBZ1+2) : Ubiquitin binding zinc finger (8,5 kDa ; pI = 5,05)



#### N(C-arm) (3,6 kDa ; pI = 4,19)



**Figure 20 : Description des constructions des domaines UBZ de Tax1BP1 et C-terminal de N. (A) Séquences des domaines Tax1BP1(UBZ1+2) et N(C arm). L'alignement de séquence des 2 domaines UBZ de Tax1BP1 a été réalisé avec Clustal Omega (Madeira et al., 2019) et préparé avec ESPRIPT 3.0 (Robert et Gouet, 2014). Les extensions restantes après coupure à la thrombine sont indiquées en violet. (B) SDS PAGE des deux domaines purifiés. Les marqueurs de tailles et leurs poids moléculaires sont indiqués sur les pistes de gauche en kDa.**

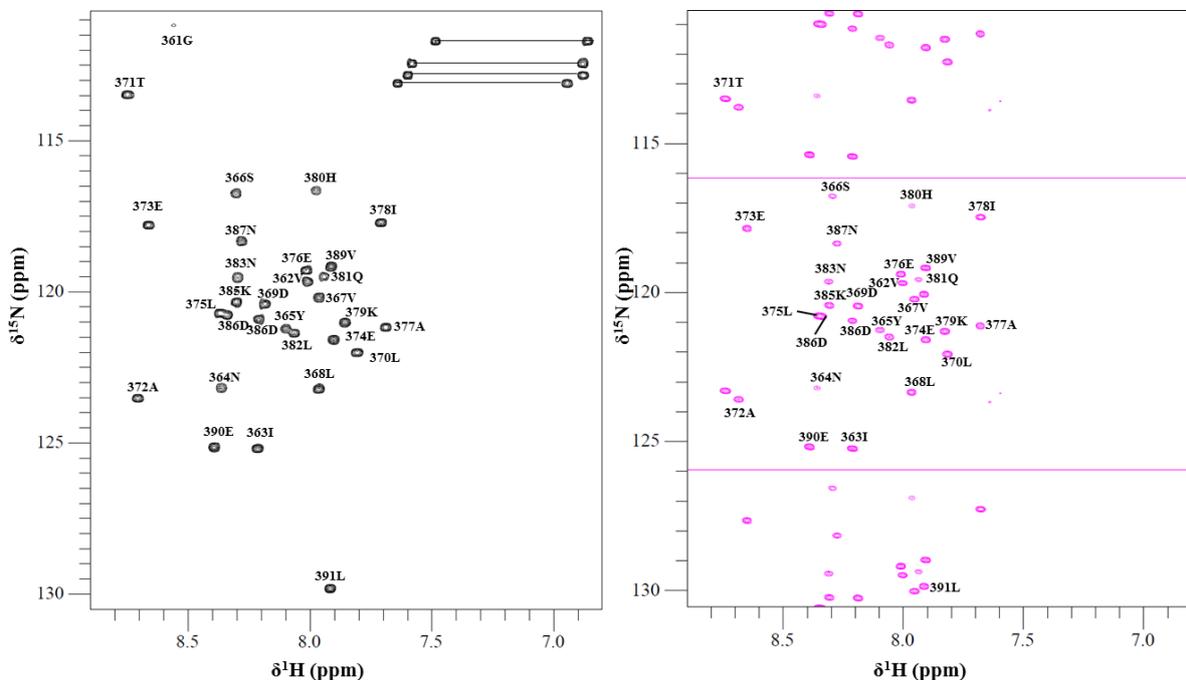
## 3.4. Etude de la structure du domaine N(C arm) isolé

### 3.4.1. Attribution des déplacements chimiques de N(C arm)

Les spectres HSQC et BEST-TROSY  $^1\text{H}$ - $^{15}\text{N}$  de N(C arm) (**Figure 21**) montrent des signaux qui sont plutôt regroupés au centre du spectre ce qui semble indiquer un peptide non replié ou en hélice.

En vue d'une attribution séquentielle du squelette peptidique et de l'attribution des chaînes latérales du domaine C-terminal de N, des expériences de type HSQC  $^1\text{H}$ - $^{13}\text{C}$ , HNC(O), HNCA, HN(CO)CA, HNCACB, HN(CO)CACB, HBHANH, HBHA(CO)NH, HCC(CO)NH et NOESY-HSQC  $^1\text{H}$ - $^{15}\text{N}$  sur spectromètre RMN à 600 MHz avec cryosonde. L'échantillon utilisé est N(C arm) marqué  $^{13}\text{C}$  et  $^{15}\text{N}$  concentré à 1 mM dans un tampon phosphate 20 mM, NaCl 100 mM, 7,5%  $\text{D}_2\text{O}$ , TCEP 1 mM, pH 6,5. Le référencement des déplacements chimiques dans la dimension proton a été fait avec du DSS (acide 4,4-diméthyl-4-silapentane-1-sulfonique) pour ces expériences et toutes les suivantes. La protéine étant concentrée à 1 mM, je n'ai pas eu besoin d'enregistrer l'expérience HSQC  $^1\text{H}$ - $^{13}\text{C}$  dans de l'eau lourde pour éliminer le signal de l'eau sur le spectre.

L'intégralité des résidus en dehors de ceux de l'étiquette N-terminale (G359-S360) est bien visible sur les spectres HSQC  $^1\text{H}$ - $^{15}\text{N}$  et BEST-TROSY. Les statistiques d'attribution de déplacements chimiques de N(C arm) sont données dans le **Tableau 2**.



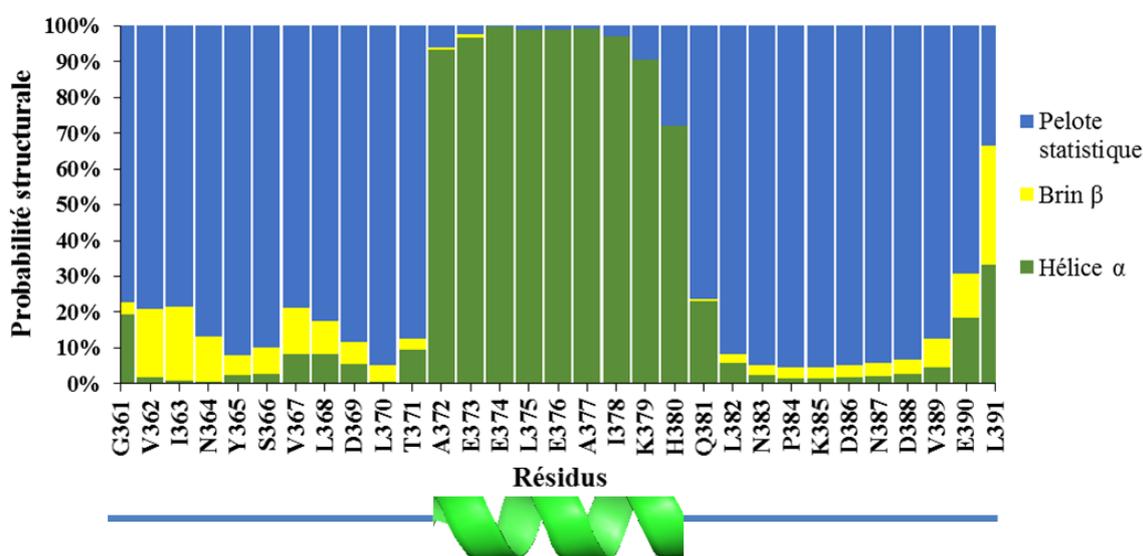
**Figure 21 :** Spectres HSQC et BEST-TROSY  $^1\text{H}$ - $^{15}\text{N}$  attribués de  $^{13}\text{C}^{15}\text{N}$ -N(C arm). A gauche est représenté le spectre HSQC  $^1\text{H}$ - $^{15}\text{N}$  attribué de  $^{13}\text{C}^{15}\text{N}$ -N(C arm) à 1 mM dans un tampon phosphate 20 mM, NaCl 100 mM, 7,5%  $\text{D}_2\text{O}$ , TCEP 1 mM, pH 6,5). A droite est montré le spectre TROSY attribué et replié de  $^{13}\text{C}^{15}\text{N}$ -N(C arm) à 100  $\mu\text{M}$  dans le même tampon à pH 6,9. Tous les deux ont été enregistrés à 600 MHz et à 298 K (25°C). Les corrélations  $^1\text{H}^{15}\text{N}$  du squelette peptidique ont été indiquées par le

numéro et le nom du résidu auxquelles elles appartiennent. Les traits noirs sur le spectre de gauche représentent les paires de protons  $^1\text{H}^N$  portées par un même azote sur les chaînes latérales d'asparagines ou de glutamines. Les lignes de couleur magenta sur le spectre de droite représentent les limites de la largeur spectrale réellement enregistrée

Catégorie	Disponibles	Attribués	Complétude (%)	$\delta$ minimum	$\delta$ moyen	$\delta$ maximum
Elément C	156	92	58,97	18,16	88,17	180,20
Elément H	188	172	91,49	0,63	3,58	8,75
Elément N	41	30	73,17	111,16	120,36	129,83
Amide	64	60	93,75	-	-	-
Squelette	99	91	91,92	-	-	-
H chaînes latérales	122	109	89,34	0,63	2,09	7,09
Type N	33	30	90,91	111,16	120,36	129,83
Type N $\delta$	4	0	0,00	-	-	-
Type N $\epsilon$	2	0	0,00	-	-	-
Type N $\zeta$	2	0	0,00	-	-	-
Type H	31	30	96,77	7,69	8,16	8,75
Type H $\alpha$	35	33	94,29	3,93	4,31	5,00
Type H $\beta$	54	54	100,00	1,47	2,34	4,63
Type H $\gamma$	34	33	97,06	0,79	1,61	2,39
Type H $\delta$	26	17	65,38	0,63	1,75	7,09
Type H $\epsilon$	8	5	62,50	2,90	3,73	6,82
Type C	33	29	87,88	173,77	176,48	180,20
Type C $\alpha$	33	32	96,97	45,19	56,76	63,54
Type C $\beta$	31	31	100,00	18,16	37,97	70,97
Type C $\gamma$	32	0	0,00	-	-	-
Type C $\delta$	22	0	0,00	-	-	-
Type C $\epsilon$	4	0	0,00	-	-	-
Type C $\zeta$	1	0	0,00	-	-	-
Résidu Ala	2	2	100,00	-	-	-
Résidu Asn	3	3	100,00	-	-	-
Résidu Asp	3	3	100,00	-	-	-
Résidu Gln	1	1	100,00	-	-	-
Résidu Glu	4	4	100,00	-	-	-
Résidu Gly	2	1	50,00	-	-	-
Résidu His	1	1	100,00	-	-	-
Résidu Ile	2	2	100,00	-	-	-
Résidu Leu	5	5	100,00	-	-	-
Résidu Lys	2	2	100,00	-	-	-
Résidu Pro	1	1	100,00	-	-	-
Résidu Ser	2	2	100,00	-	-	-
Résidu Thr	1	1	100,00	-	-	-
Résidu Tyr	1	1	100,00	-	-	-
Résidu Val	3	3	100,00	-	-	-
Tous les résidus	33	32	96,97	-	-	-

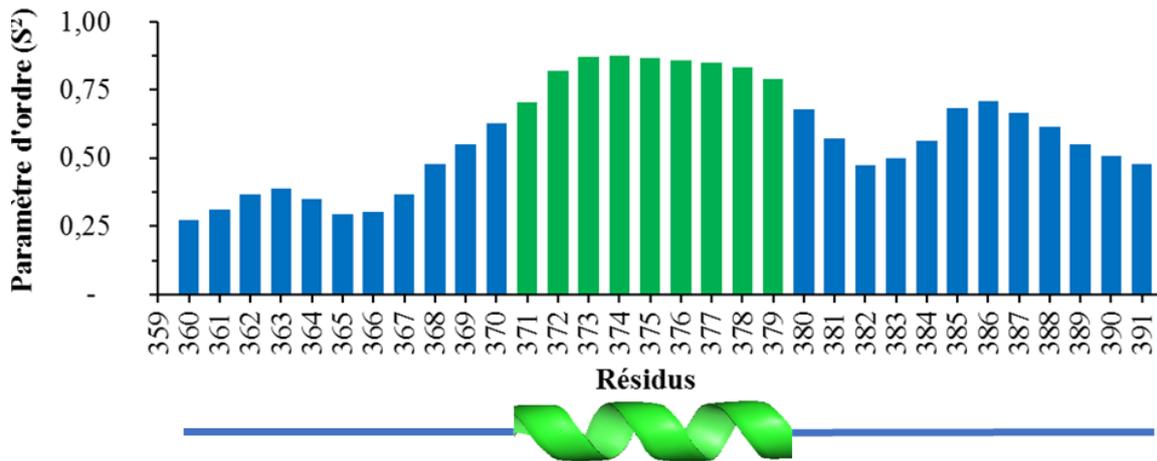
**Tableau 2 : Statistiques de l'attribution des déplacements chimiques de N(C arm)**

A partir de l'attribution des déplacements chimiques  $^1\text{H}_\text{N}$ ,  $^{15}\text{N}_\text{H}$ ,  $^{13}\text{C}_\alpha$ ,  $^{13}\text{C}_\beta$ ,  $^{13}\text{C}'$  et  $^1\text{H}_\alpha$ , une prédiction de structure secondaire par résidu a été faite avec le logiciel Talos-N (Shen et Bax, 2013 ; **Figure 22**). Les résultats de prédiction de structure secondaire montrent une très forte propension d'hélice  $\alpha$  (probabilité > 90 % en moyenne) du résidu 372 au résidu 380 et majoritairement une structure en pelote statistique pour le reste du peptide. D'autre part, à partir du résidu 379, la propension d'hélice  $\alpha$  décroît fortement mais pas totalement jusqu'au résidu 383. De même, une probabilité de former une hélice  $\alpha$  de l'ordre de 10% existe pour le résidu 371 en amont de l'hélice ainsi qu'une plus faible probabilité d'hélice  $\alpha$  pour les résidus 365 à 369 (de l'ordre de moins de 10 %). La fin de l'hélice en C-ter est certainement due à la présence de la proline 384.



**Figure 22 : Prédiction de structures secondaires de chaque résidu de N(C arm).** Le calcul des probabilités de structures secondaires a été réalisé à partir du logiciel TALOS-N en ligne avec les déplacements chimiques  $^1\text{H}_\text{N}$ ,  $^{15}\text{N}_\text{H}$ ,  $^{13}\text{C}_\alpha$ ,  $^{13}\text{C}_\beta$  et  $^{13}\text{C}'$  et  $^1\text{H}_\alpha$  en utilisant  $^{13}\text{C}^{15}\text{N}$ -N(C arm) à 1 mM dans le tampon phosphate 20 mM, NaCl 100 mM, 7,5%  $\text{D}_2\text{O}$ , TCEP 1 mM, pH 6,5. En bas du graphique, est représenté un schéma théorique de la structure secondaire de la protéine.

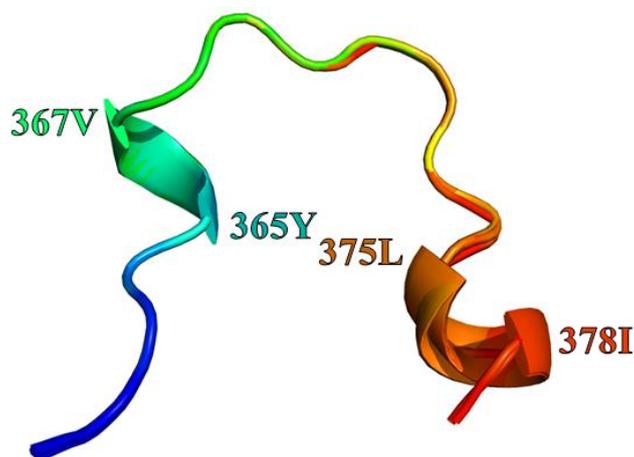
Le logiciel Talos-N, avec les mêmes données que précédemment, a également permis de calculer un paramètre d'ordre ( $S^2$ ) par résidu de N(C arm) à partir du modèle de Wishart (Berjanskii et Wishart, 2005 ; **Figure 23**). Ces résultats confirment la présence d'une rigidification de la région en hélice  $\alpha$  aux résidus 372-380. Mais cela suggère également une possible rigidification locale en amont et en aval de cette hélice au niveau des résidus 362-364 et plus fortement au niveau des résidus 385-387 (juste après la proline 384). De manière plus générale, l'extrémité C-terminale semble plus rigide que l'extrémité N-terminale.



**Figure 23: Paramètre d'ordre de chaque résidu de N(C arm).** Le calcul du paramètre d'ordre a été fait à partir du logiciel TALOS-N selon de modèle de Wishart à partir des mêmes données que celles utilisées pour la **Figure 22**. En bas du graphique, est représenté un schéma théorique de la structure secondaire de la protéine.

La structure intégrale de la nucléoprotéine du virus respiratoire syncytial résolue par diffraction des rayons X (Tawar et al., 2009) m'a permis de récupérer au moins une partie de la structure atomique du domaine N(C arm) dans ce contexte. En effet, l'unité asymétrique du cristal contient 2 anneaux de décimères de N avec ARN. Pour chaque protomère de N, une partie plus ou moins longue du domaine C-terminal est visible. Une conformation en hélice  $\alpha$  pour les résidus 365 à 367 est toujours observée. Aussi, quelques-uns des protomères de N présentent une structure en hélice  $\alpha$  plus ou moins longues pour les résidus 375 à 378. Au-delà du résidu 379 dans les 20 protomères de N de l'unité asymétrique, aucun des résidus C-terminaux n'est visible à cause de leur flexibilité. Ces observations confirment une grande partie de la propension de structure secondaire que j'ai calculée avec Talos-N. Néanmoins, l'hélice  $\alpha$  systématiquement présente pour les résidus 365 à 367 dans le cristal ne représente qu'une infime propension en solution en plus d'être plus longue en solution. Vice-versa, la région avec une propension très élevée à former une hélice  $\alpha$  en solution pour les résidus 372 à 380 n'est pas formée dans le cristal en plus d'être écourtée aux résidus 375 à 378. Ces différences de conformations locales peuvent s'expliquer par la contrainte de repliement due à l'empilement cristallin par rapport à une forme relâchée en solution. Il est à noter aussi que la structure du domaine N(C arm) obtenue par cristallographie a été déterminée dans le cadre de la nucléoprotéine entière complexée avec de l'ARN alors que par RMN j'ai travaillé sur le domaine N(C arm) isolé.

J'ai aligné les structures des domaines N(C arm) présents dans les 20 protomères de N de l'unité asymétrique du cristal à l'aide du logiciel Pymol (Schrödinger) (**Figure 24**) afin de mieux visualiser la structure du domaine. L'alignement montre une excellente superposition du résidu 361 à 379 en dépit des différentes longueurs de protomères de N pour le domaine C-terminal.



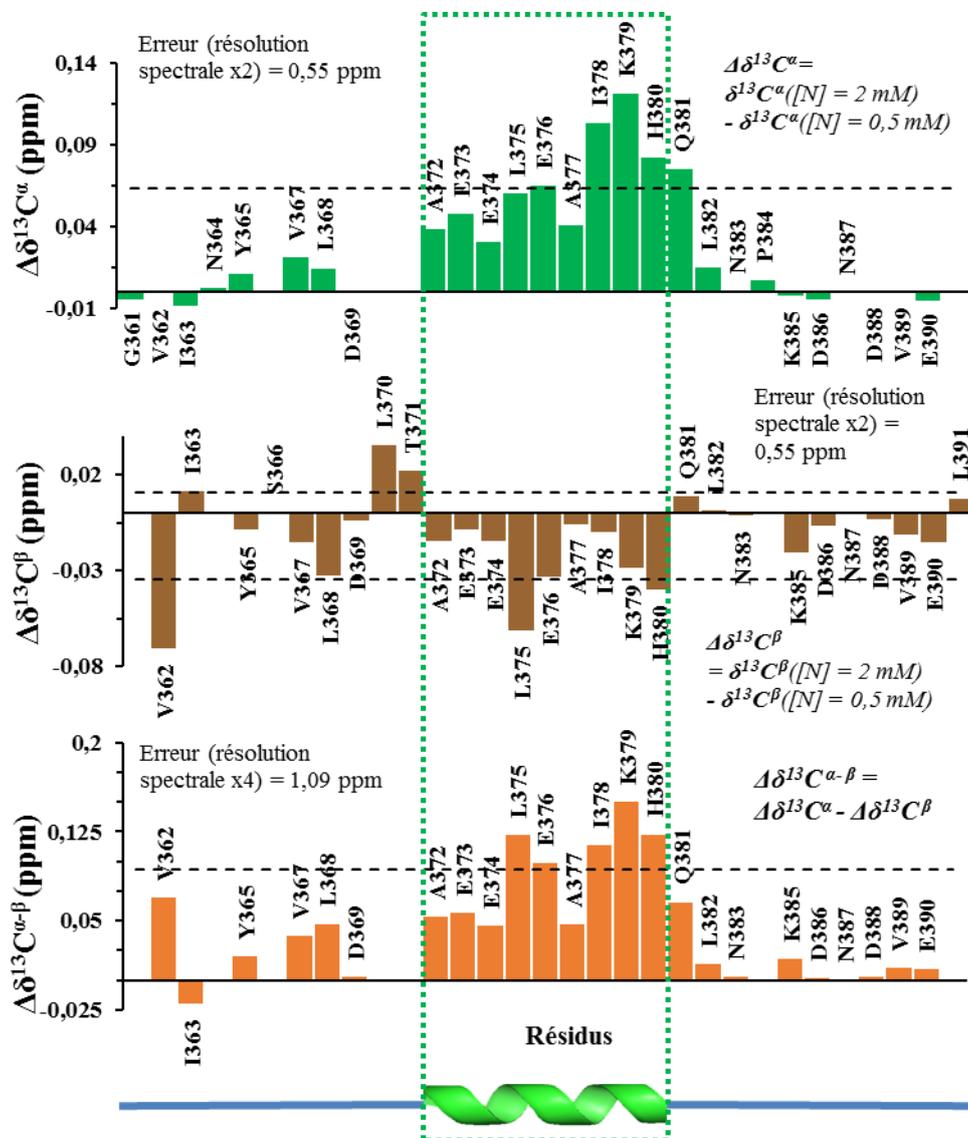
**Figure 24 :** *Alignement des structures du domaine C arm de la nucléoprotéine du RSV cristallisée. L'alignement a été réalisé avec le logiciel Pymol à partir de la structure de 2 décimères de N en complexe avec de l'ARN (code PDB : 2WJ8) et le résultat est donné avec les 20 structures partielles de N(C arm) présentes dans l'unité asymétrique en représentation cartoon. En bleu marine et en rouge sont indiquées respectivement les extrémités N et C-terminale. Les premiers et derniers résidus des 2 hélices  $\alpha$  représentées sont indiqués par leur numéro et leur code à une lettre.*

### 3.4.2. Etude des effets de la concentration de N(C arm) sur les déplacements chimiques.

Afin d'étudier la stabilité et la structure de N(C arm), des enregistrements de spectres HNCACB du peptide ont été réalisés à deux différentes concentrations (500  $\mu$ M et 2 mM), mais dans les mêmes conditions de tampons. Après réattribution des déplacements chimiques  $^1\text{H}^{\text{N}}$ ,  $^{15}\text{N}$ ,  $^{13}\text{C}_{\alpha}$ , et  $^{13}\text{C}_{\beta}$  à partir des déplacements chimiques du peptide à 1 mM, les différences de déplacements chimiques  $^{13}\text{C}_{\alpha}$ , et  $^{13}\text{C}_{\beta}$  à ces deux concentrations ont été analysées pour chaque résidu (**Figure 25**).

Ces résultats (**Figure 25**) montrent clairement qu'à plus forte concentration de N(C arm) on observe une légère augmentation de la valeur des déplacements chimiques  $^{13}\text{C}_{\alpha}$  et une diminution de la valeur des déplacements chimiques  $^{13}\text{C}_{\beta}$  dans la région de l'hélice  $\alpha$ , du résidu 372 au résidu 380. Une variation positive de la valeur des déplacements chimiques  $^{13}\text{C}_{\alpha}$  est aussi observée pour le résidu 381 et un moins importante pour le résidu 382. Mais il n'y a qu'une infime perturbation de la valeur des déplacements chimiques  $^{13}\text{C}_{\beta}$  pour les résidus 382-383 contrairement au résidu 381 qui montre une variation positive de déplacements chimique  $^{13}\text{C}_{\beta}$  à l'inverse de la région en hélice  $\alpha$  juste avant. Quoiqu'il en soit, en calculant la différence des différences  $\Delta\delta^{13}\text{C}_{\alpha} - \Delta\delta^{13}\text{C}_{\beta}$ , le signe du résultat de cette opération est positif pour toute la région de l'hélice  $\alpha$  et pour les résidus C-terminaux suivants (381-383). De même, il est à noter que la valeur de  $\Delta\delta^{13}\text{C}_{\alpha} - \Delta\delta^{13}\text{C}_{\beta}$  pour les résidus 365 à 369 est positive. Or, il a été démontré que l'augmentation de la valeur des déplacements chimiques  $^{13}\text{C}_{\alpha}$  et  $^{13}\text{C}'$  ainsi que la diminution de la valeur des déplacements chimiques  $^{13}\text{C}_{\beta}$  vont généralement dans le sens d'une formation d'hélice  $\alpha$  (Saitô, 1986; Spera and Bax, 1991; Wishart et al., 1991). Cette information nous permet de supposer, d'après les résultats de la **Figure 25**, qu'il y a un renforcement de la probabilité de

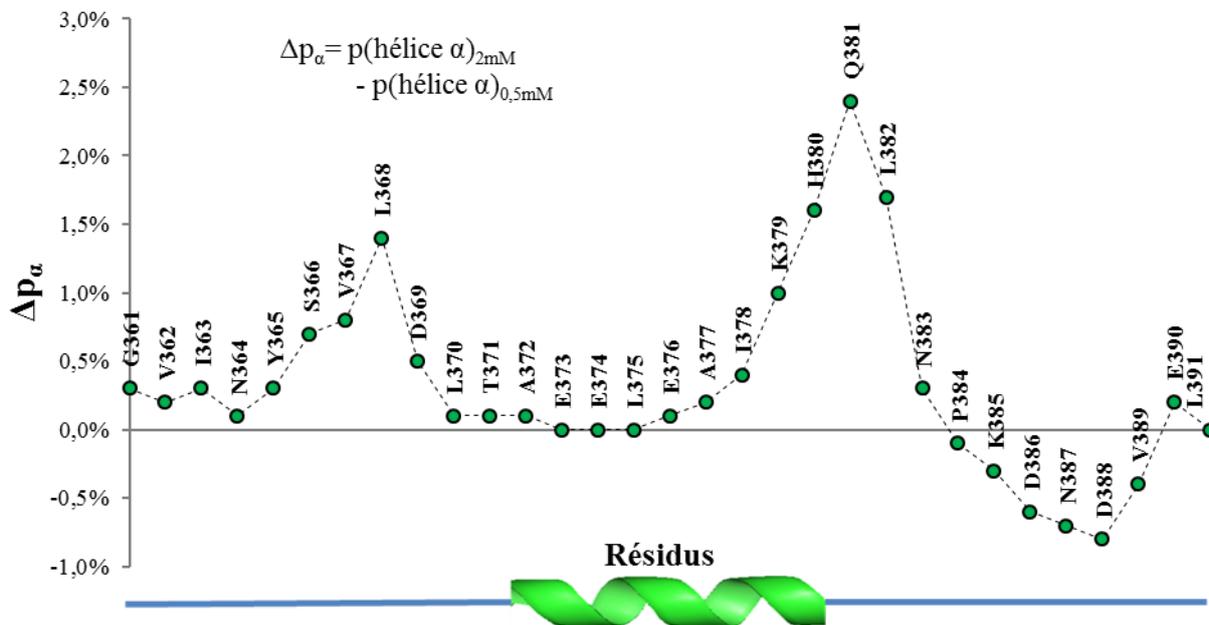
former une hélice  $\alpha$  dans la région englobant les résidus 372 à 383, particulièrement pour les 6 derniers résidus, en plus d'une éventuelle formation transitoire d'hélice  $\alpha$  pour les résidus 365 à 369. Cependant, tous les résultats décrits ici sont sujets à une erreur expérimentale et théorique bien trop importante pour rationnellement s'appuyer sur ces conclusions.



**Figure 25 :** Effets de la concentration de N(C arm) sur ses déplacements chimiques  $^{13}C_{\alpha}$  et  $^{13}C_{\beta}$ . Les histogrammes en barres vertes et marrons ont été fait à partir des différences de perturbations selon l'équation  $\Delta\delta^{13}C = \delta^{13}C(2 \text{ mM}) - \delta^{13}C(0,5 \text{ mM})$ . L'histogramme en barres oranges correspond à la soustraction des différences  $\Delta\delta^{13}C_{\alpha-\beta} = \Delta\delta^{13}C_{\alpha} - \Delta\delta^{13}C_{\beta}$ . Les lignes en pointillés noirs représentent la moyenne plus (pour tous les graphiques) et moins (pour  $^{13}C_{\beta}$ ) l'écart-type. En bas du graphique, est représenté le schéma théorique de la structure secondaire de la protéine.

En utilisant les déplacements chimiques  $^1H^N$ ,  $^{15}N$ ,  $^{13}C_{\alpha}$  et  $^{13}C_{\beta}$  pour faire 2 autres prédictions de structure secondaire (SSP) du peptide avec le logiciel Talos-N (aux 2 concentrations différentes utilisées), j'ai calculé la différence de probabilité d'adopter une conformation en hélice  $\alpha$  aux 2 différentes concentrations,  $p(\text{hélice } \alpha)_{2\text{mM}} - p(\text{hélice } \alpha)_{0,5\text{mM}}$  (**Figure 26**). La courbe de  $p(\text{hélice } \alpha)_{2\text{mM}} -$

$p(\text{hélice } \alpha)_{0,5\text{mM}}$  en fonction des résidus de N(C arm) montre clairement une légère augmentation de chance de former une hélice  $\alpha$  pour les résidus 378 à 383 (en fin de l'hélice  $\alpha$  précédemment identifiée), de même, de manière moins élevée, pour la région en N-terminal des résidus 366 à 369. De manière surprenante, on observe également une diminution de la probabilité de former une hélice  $\alpha$  après la proline 384 qui vient tronquer l'hélice 372-383. Cette diminution de propension en hélice  $\alpha$  en C-terminal du peptide laisse à penser qu'il y aurait peut-être des hélices  $\alpha$  transitoirement présentes dans cette région. Toutefois, les résultats de la **Figure 26** sont à prendre avec précaution car je n'ai pas utilisé de déplacements chimiques  $^1\text{H}_\alpha$  et  $^{13}\text{C}'$  dans les calculs de propensions alors qu'ils ont beaucoup de poids pour la prédiction de structure secondaire.



**Figure 26 :** Différence de probabilité de former une hélice  $\alpha$  en fonction des résidus de N(C arm) à deux concentrations différentes. Ce calcul a été fait à partir des probabilités de former une structure secondaire prédit par le logiciel Talos-N en ligne avec les déplacements chimiques  $^1\text{H}_\text{N}$ ,  $^{15}\text{N}$ ,  $^{13}\text{C}_\alpha$  et  $^{13}\text{C}_\beta$  attribués en utilisant le domaine  $^{13}\text{C}^{15}\text{N}$ -N(C arm) à 2 mM et à 0,5 mM dans le tampon phosphate 20 mM, NaCl 100 mM, 7,5%  $\text{D}_2\text{O}$ , TCEP 1 mM, pH 6,5. En bas du graphique, est représenté le schéma théorique de la structure secondaire de la protéine.

En définitive, les données de la **Figure 25** et de la **Figure 26** vont dans le sens d'une légère élongation de l'hélice  $\alpha$  en amont et en aval des bornes trouvées à 1 mM pour le domaine N(C arm) lorsque la concentration augmente. Si on regarde de plus près, la région 366-369 formant probablement transitoirement une hélice  $\alpha$  (selon les prédictions de propensions d'hélice  $\alpha$  donnés par TALOS-N sur N(C arm) à 1 mM et la structure cristallographiques de N-ARN) n'apparaît pas comme étant continue avec l'hélice  $\alpha$  372-380 contrairement à l'extension de l'hélice  $\alpha$  aux résidus 381 à 383. Cependant, ce renforcement ou cette apparition d'une probable hélice  $\alpha$  en augmentant la concentration du peptide en solution pourrait être dû à un phénomène de début d'agrégation.

### 3.4.3. Obtention de la structure du domaine N(C arm)

#### 3.4.3.1. Contraintes utilisées pour le calcul de structure

Ayant connaissance des angles dièdres prédits par le logiciel Talos-N et ayant à disposition un spectre NOESY-HSQC  $^1\text{H}$ - $^{15}\text{N}$  manuellement attribué du domaine N(C arm), j'ai utilisé ces données comme contraintes d'angles dièdres et sources de contraintes de distances interatomiques afin de calculer la structure tridimensionnelle du peptide avec le logiciel CYANA.

Dans le **Tableau 3** sont indiquées les valeurs de contraintes d'angles dièdres pour chaque résidu.

Numéro de résidu	Nom du résidu	Angle $\phi$ minimum	Angle $\phi$ maximum	Angle $\psi$ minimum	Angle $\psi$ maximum	Confiance de prédiction
370	LEU	-94,3	-54,3	125,5	168,5	Forte
371	THR	-95,3	-55,3	144,3	184,3	Forte
372	ALA	-78,4	-38,4	-60	-20	Forte
373	GLU	-87,6	-47,6	-60,4	-20,4	Forte
374	GLU	-87,8	-47,8	-60,1	-20,1	Forte
375	LEU	-85,5	-45,5	-59,5	-19,5	Forte
376	GLU	-86,8	-46,8	-60	-20	Forte
377	ALA	-85,1	-45,1	-57,3	-17,3	Forte
378	ILE	-90,1	-50,1	-57,8	-17,8	Forte
379	LYS	-86,8	-46,8	-50,2	-10,2	Forte
380	HIS	-99,1	-59,1	-34,9	5,1	Forte
385	LYS	-158,5	-61,3	72,7	196,1	Généreuse
386	ASP	-103,6	-38,2	99,1	173	Généreuse

**Tableau 3 :** Contraintes d'angles dièdres sélectionnées pour le calcul de structure de N(C arm). Les lignes surlignées en vert clair correspondent à la région prédite par TALOS-N en hélice  $\alpha$ .

Concernant le spectre NOESY-HSQC  $^1\text{H}$ - $^{15}\text{N}$ , 160 pics attribués correspondant à des nOe ont été introduits dans CYANA et réattribués par le logiciel.

#### 3.4.3.2. Calcul de la structure de N(C arm)

En plus, des contraintes d'angles dièdres et des pics de NOESY-HSQC  $^1\text{H}$ - $^{15}\text{N}$ , j'ai ajouté au logiciel CYANA :

- La séquence des résidus de N(C arm),
- 1 liste de déplacements chimiques  $^1\text{H}$ ,  $^{15}\text{N}$  et  $^{13}\text{C}$ .

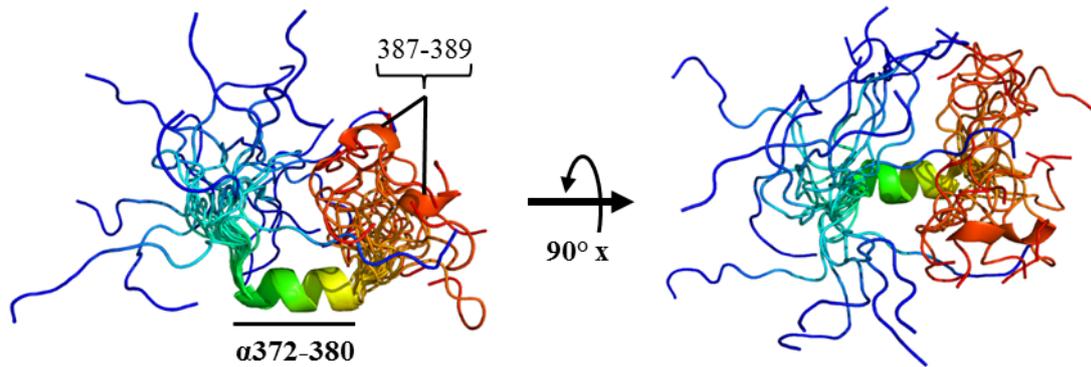
Dans le **Tableau 4** sont données les statistiques du calcul de structure de N(C arm) par CYANA.

Cycles	1	2	3	4	5	6	7	final
<b>Pics</b>								
Sélectionnés	160	160	160	160	160	160	160	160
Attribués	126	135	135	135	135	134	132	
Non attribués	34	25	25	25	25	26	28	
Avec attribution dans la diagonale	0	0	0	0	0	0	0	
<b>Pics croisés</b>								
Attribution hors diagonale	126	135	135	135	135	134	132	
Avec attribution unique	104	112	116	117	124	124	121	
Avec attribution courte distance $ i-j  \leq 1$	122	124	124	124	124	124	124	
Avec attribution distance moyenne $1 <  i-j  < 5$	4	11	11	11	11	10	8	
Avec attribution longue distance $ i-j  \geq 5$	0	0	0	0	0	0	0	
<b>Comparaison avec les attributions de référence</b>								
Pics croisés avec attribution de référence	158	158	158	158	158	158	158	
Avec attribution identique à la référence	90	95	96	96	99	99	97	
Avec attribution compatible à la référence	105	109	108	108	107	106	105	
Avec attribution incompatible à la référence	21	26	27	27	28	28	27	
Avec attribution en plus de la référence	32	23	23	23	23	24	26	
Avec attribution aditionnelle	0	0	0	0	0	0	0	
<b>Limites supérieures de distance</b>								
Total	77	82	81	79	77	77	79	84
Courte distance $ i-j  \leq 1$	74	74	73	71	69	70	73	78
Distance moyenne $1 <  i-j  < 5$	3	8	8	8	8	7	6	6
Longue distance $ i-j  \geq 5$	0	0	0	0	0	0	0	0
Nombre d'attribution moyenne/contraite	1,26	1,29	1,15	1,15	1,1	1,1	1	1
Valeur moyenne de la fonction cible	0,02	0,22	0,39	0,11	0,07	0,02	0,01	0
<b>RMSD (résidus 372-380)</b>								
RMSD moyen du squelette peptidique (Å)	0,43	0,45	0,25	0,32	0,19	0,3	0,25	0,23
RMSD moyen des atomes lourds (Å)	1,27	1,26	1,09	1,16	1,05	1,13	1,1	1

**Tableau 4 : Statistiques du calcul de structure de N(C arm) par CYANA.**

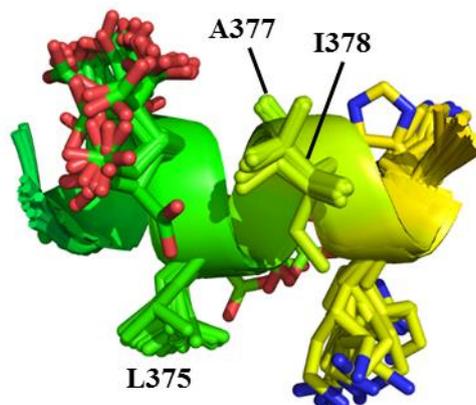
Sur ce tableau, on remarque qu'environ deux tiers des attributions de pics faites par CYANA sont identiques ou compatibles avec les attributions de références.

Après avoir réaligné, avec le logiciel PYMOL, les différents conformères de N(C arm) sur l'hélice  $\alpha$  des résidus 372 à 380, j'ai représenté sous forme « cartoon » les 20 conformères correspondant à la dernière itération du calcul de structure avec CYANA (**Figure 27**). Cette figure montre une très bonne convergence du squelette peptidique de l'hélice  $\alpha$ 372-380, les deux extrémités étant désordonnées. Aussi, l'extrémité C-terminale semble plus compacte en termes d'exploration de l'espace par rapport à l'extrémité N-terminale ; cette observation semble aller dans le même sens que les résultats de paramètres d'ordres (**Figure 23**). De plus, une hélice calculée sur 2 conformères au niveau des résidus 387 à 389 apparaît, mais ne semble pas avoir de lien avec les contraintes d'angles dièdres qui s'arrêtent au résidu 386.



**Figure 27 :** Représentation « cartoon » des 20 conformères finaux de N(C arm) calculés par CYANA. En bleu et en rouge sont représentées respectivement les extrémités N- et C-terminales. L'hélice  $\alpha$  des résidus 372 à 380 ainsi que 2 hélices transitoires en C-terminal sont signalées.

En ce qui concerne les chaînes latérales de N(C arm) de la région structurée en hélice  $\alpha$  (**Figure 28**), seule l'alanine 377 a une unique conformation. 19 des conformères de N(C arm) présentent l'isoleucine 378 dans la même orientation. La leucine 375 adopte aussi des conformations orientées de manière assez semblable mais avec une convergence relativement faible. Les autres chaînes latérales de l'hélice, qui sont au contraire plutôt hydrophiles, ont des orientations beaucoup plus confuses. Le manque de convergence concernant les chaînes latérales est certainement dû au faible nombre de contraintes de distances initialement utilisées pour ce calcul de structure.



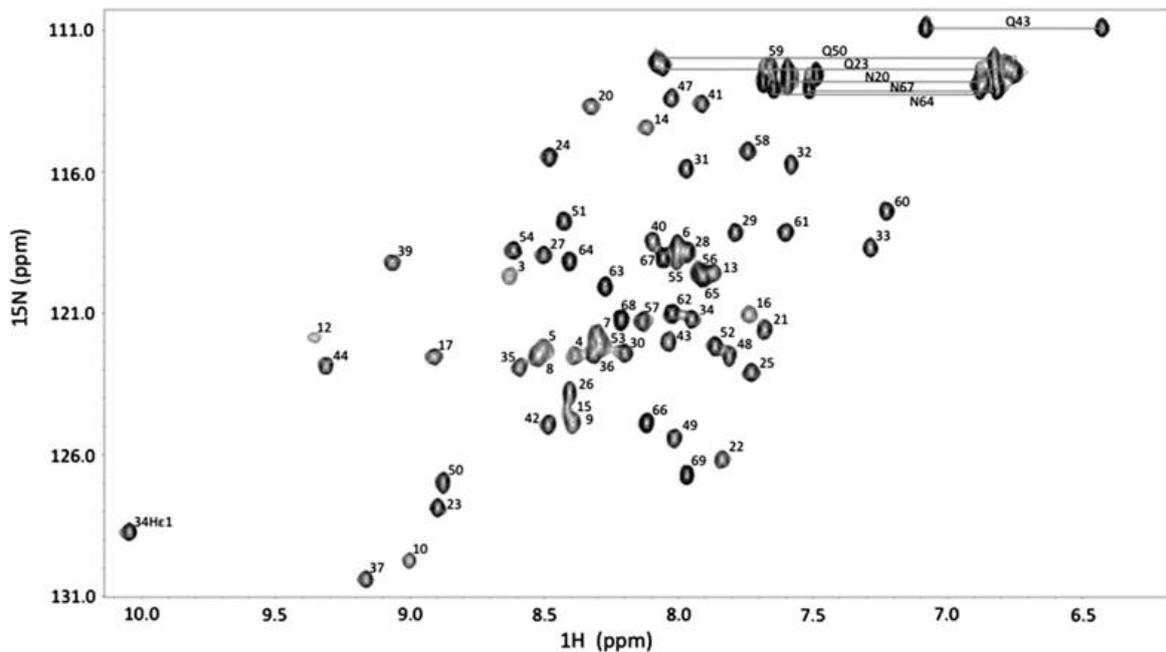
**Figure 28 :** Représentation sous forme « bâtonnets » des chaînes latérales de l'hélice  $\alpha$  des résidus 372 à 380. En vert et en jaunes sont montrées respectivement les extrémités N- (jusqu'au résidu 371) et C-terminales (jusqu'au résidu 381). Les résidus dont le numéro et le code à 1 lettre sont indiqués correspondent à ceux qui ont une orientation semblable ou presque.

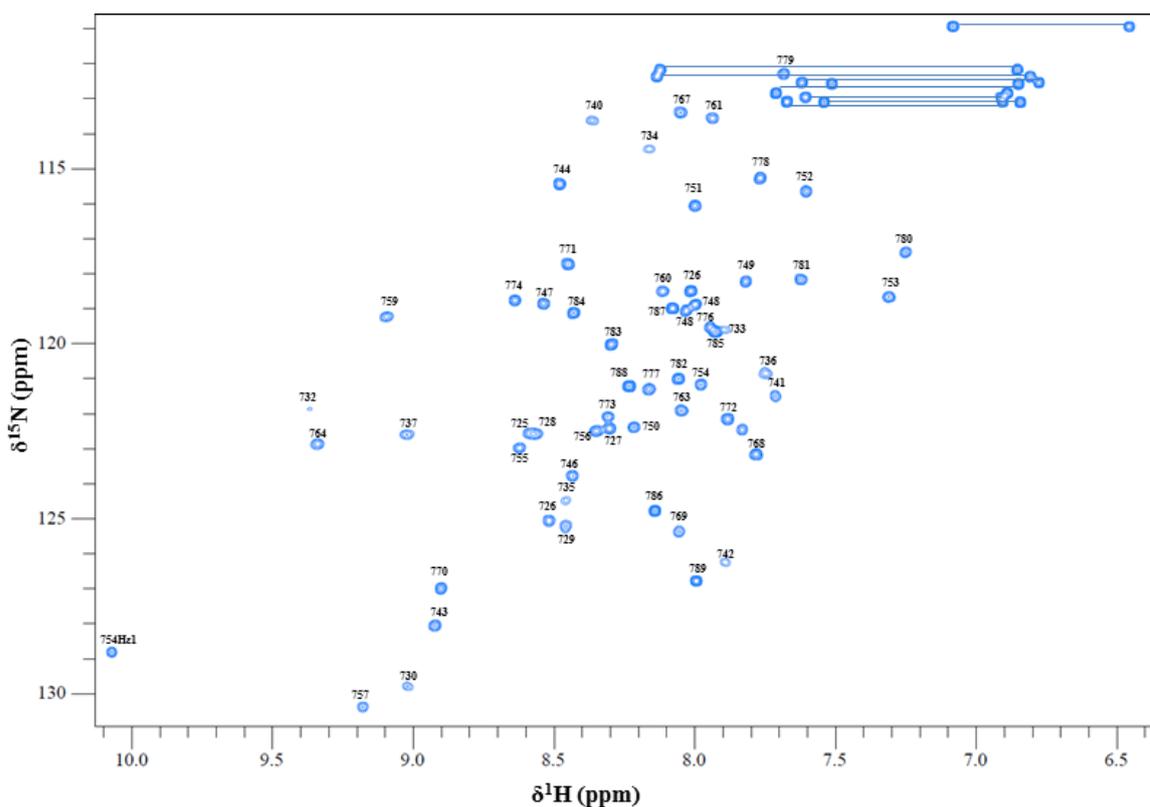
### 3.5. Etude du domaine Tax1BP1(UBZ1+2) isolé par RMN

#### 3.5.1. Attribution des déplacements chimiques de Tax1BP1(UBZ1+2)

Après la production et la purification du domaine  $^{15}\text{N}$ -UBZ1+2 de Tax1BP1, j'ai acquis des spectres HSQC  $^1\text{H}$ - $^{15}\text{N}$  du peptide à différentes concentrations avec un spectromètre 700 MHz dans un tampon phosphate 20 mM, NaCl 100 mM, 7,5%  $\text{D}_2\text{O}$ , TCEP 1 mM à pH 6,5 à 298 K (25°C).

A partir du spectre HSQC  $^1\text{H}$ - $^{15}\text{N}$  de  $^{15}\text{N}$ -UBZ1+2 et des attributions de déplacements chimiques déposés dans la BMRB sous le code 19201 par (Ceregido et al., 2014), j'ai pu réattribuer assez simplement le spectre du domaine UBZ1+2 et les déplacements chimiques amides. Je l'ai fait aisément en dépit des légers changements de tampon en termes de pH et de concentration du tampon ainsi que de la présence en N-ter de résidus résultant de la coupure de l'étiquette GST, GS au lieu de GPHM (**Figure 29**).





**Figure 29 :** Spectres HSQC  $^1\text{H}$ - $^{15}\text{N}$  de  $^{15}\text{N}$  Tax1BP1(UBZ1+2). Celui du haut a été enregistré à 800 MHz par Ceregido et al. (J. Mol. Biol., 2014) avec  $^{13}\text{C}^{15}\text{N}$  Tax1BP1(UBZ1+2) à 700  $\mu\text{M}$  dans un tampon phosphate salin (PBS) 50 mM, 7,5%  $\text{D}_2\text{O}$ , 1 mM TCEP à pH 6,0. Le spectre du bas, a été enregistré à 700 MHz par moi-même avec  $^{15}\text{N}$  Tax1BP1(UBZ1+2) concentrée à 100  $\mu\text{M}$  dans un tampon phosphate 20 mM, NaCl 100 mM, 7,5%  $\text{D}_2\text{O}$ , TCEP 1 mM à pH 6,5. Les 2 spectres ont été enregistrés à 298 K (25°C). Les corrélations  $^1\text{H}^{15}\text{N}$  du squelette peptidique ont été indiquées par le numéro du résidu auxquelles elles appartiennent. Les traits (bleu ou noir) sur les spectres représentent les paires de protons  $^1\text{H}^{\text{N}}$  portées par un même azote sur les chaînes latérales d'asparagines ou de glutamines.

### 3.6. Etude de l'interaction entre les peptide $^{13}\text{C}^{15}\text{N}$ -N(C arm) et Tax1BP1(UBZ1+2)

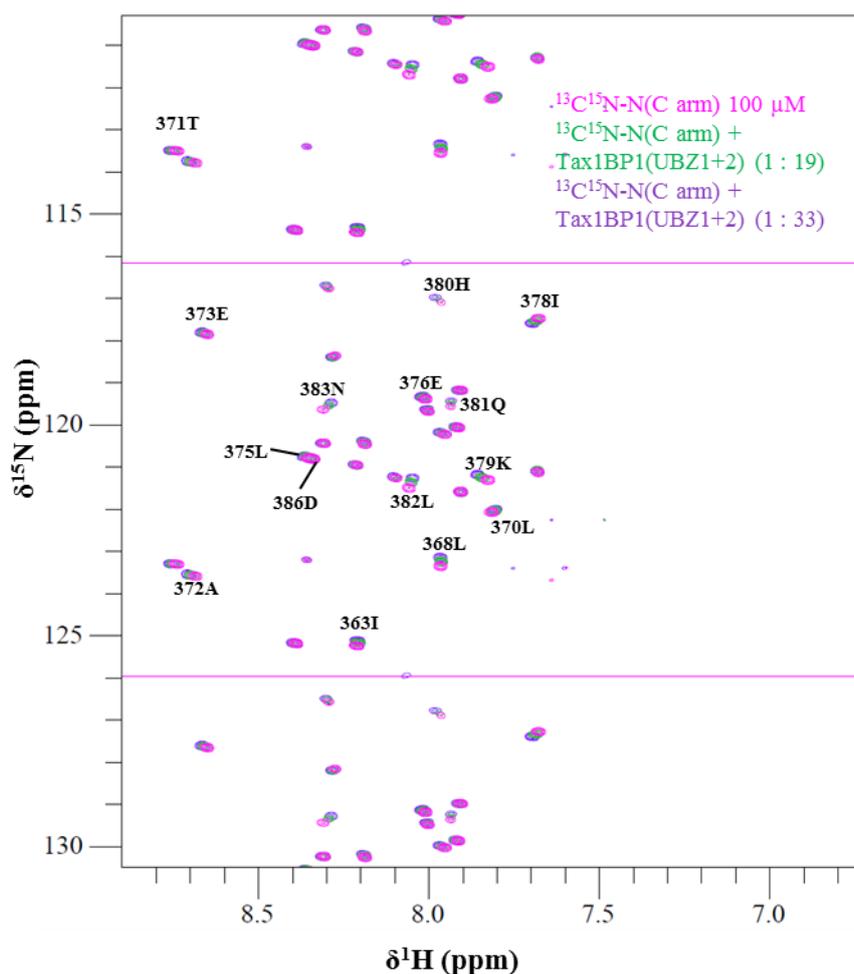
#### 3.6.1. Titration de $^{13}\text{C}^{15}\text{N}$ -N(C arm) par Tax1BP1(UBZ1+2)

J'ai réalisé la production et la purification du domaine C arm de N marqué  $^{13}\text{C}^{15}\text{N}$  et du domaine UBZ1+2 de Tax1BP1 qui se sont déroulées sans difficultés. Les 2 protéines ont été dialysées ensemble dans un tampon phosphate 20 mM, NaCl 100 mM, 7,5%  $\text{D}_2\text{O}$ , TCEP 1 mM à pH 6,9 à 298 K (25°C). J'ai ensuite fait un titrage de  $^{13}\text{C}^{15}\text{N}$ -N(C arm) par Tax1BP1(UBZ1+2) par mélange de différentes solutions concentrées. A chaque point de titrage, pour différentes concentrations de Tax1BP1(UBZ1+2) et à concentration constante de 100  $\mu\text{M}$  de  $^{13}\text{C}^{15}\text{N}$ -N(C arm), j'ai acquis des spectres 1D  $^1\text{H}$  et BEST-TROSY  $^1\text{H}$ - $^{15}\text{N}$  sur un spectromètre 600 MHz avec cryosonde.

J'ai ensuite réattribué les déplacements chimiques des signaux  $^1\text{H}^{\text{N}}$  et  $^{15}\text{N}$  de N(C arm) sur les spectres BEST-TROSY  $^1\text{H}$ - $^{15}\text{N}$  du titrage (**Figure 30**). Dans un premier temps, on observe que ce sont

principalement des signaux au niveau des résidus en hélices  $\alpha$  qui semblent être déplacés d'une largeur de raie ou un peu plus en termes de déplacements chimiques quand on ajoute plusieurs dizaines d'équivalents Tax1BP1(UBZ1+2). Mais ces perturbations de déplacements chimiques sont de faible amplitude, à des concentrations en Tax1BP1(UBZ1+2) de l'ordre de plusieurs millimolaire et le décalage des signaux semble se faire de manière linéaire par rapport à la concentration.

J'ai ensuite calculé les perturbations de déplacements chimiques  $\Delta^1H^N$ ,  $\Delta^{15}N$  et combinées  $\Delta^1H^{15}N$  ( $= \sqrt{(\delta^1H - \delta^1H_0)^2 + 0,14 \times (\delta^{15}N - \delta^{15}N_0)}$ ) et des intensités relatives,  $I/I_0$ , des signaux de N(C arm) sur les BEST-TROSY  $^1H$ - $^{15}N$  en présence de 33 équivalents de Tax1BP1(UBZ1+2). Les résultats sont présentés sous forme d'histogrammes (**Figure 31**). L'amplitude des perturbations des déplacements chimiques amides  $^1H^{15}N$  de N(C arm) en présence d'une forte concentration de Tax1BP1(UBZ1+2) est faible. En dépit de cette observation, on voit que les valeurs  $\Delta^{15}N$  et  $\Delta^1H^{15}N$  dans la région C-terminale de l'hélice  $\alpha$  372-380 et pour le résidu 368 sont plus perturbées que les autres. On peut remarquer que ces perturbations plus accentuées se situent dans des régions supposées comme étant en hélices  $\alpha$  transitoires. Concernant les intensités relatives, la concentration élevée de Tax1BP1(UBZ1+2) par rapport à N(C arm) ne semble pas affecter significativement les ratios  $I/I_0$  des signaux HN de N(C arm).



**Figure 30 : Spectres BEST-TROSY  $^1\text{H}$ - $^{15}\text{N}$  superposés de  $^{13}\text{C}^{15}\text{N}$ -N(C arm) avec différentes concentrations de Tax1BP1(UBZ1+2).** Le spectre de couleur magenta est celui de  $^{13}\text{C}^{15}\text{N}$ -N(C arm) seul et concentré à 100  $\mu\text{M}$  ; il correspond au spectre de référence. Les spectres vert et mauve sont les spectres issus du titrage de  $^{13}\text{C}^{15}\text{N}$ -N(C arm) par Tax1BP1(UBZ1+2) correspondant respectivement aux ratios molaires 1:19 et 1:33 (N(C arm):Tax1BP1(UBZ1+2)). Les numéros et codes à une lettre des résidus qui ont subi le plus de perturbation de leur déplacement chimique sont indiqués. Tous les spectres ont été enregistrés à 298 K (25°C) et à 600 MHz et le tampon utilisé était du phosphate 20 mM, NaCl 100 mM, 7,5%  $\text{D}_2\text{O}$ , TCEP 1 mM à pH 6,9. Les résidus qui semblent être présents plusieurs fois sur le spectre de manière périodique sont le résultat du dépliement de la fenêtre spectrale dans la dimension  $^{15}\text{N}$ .

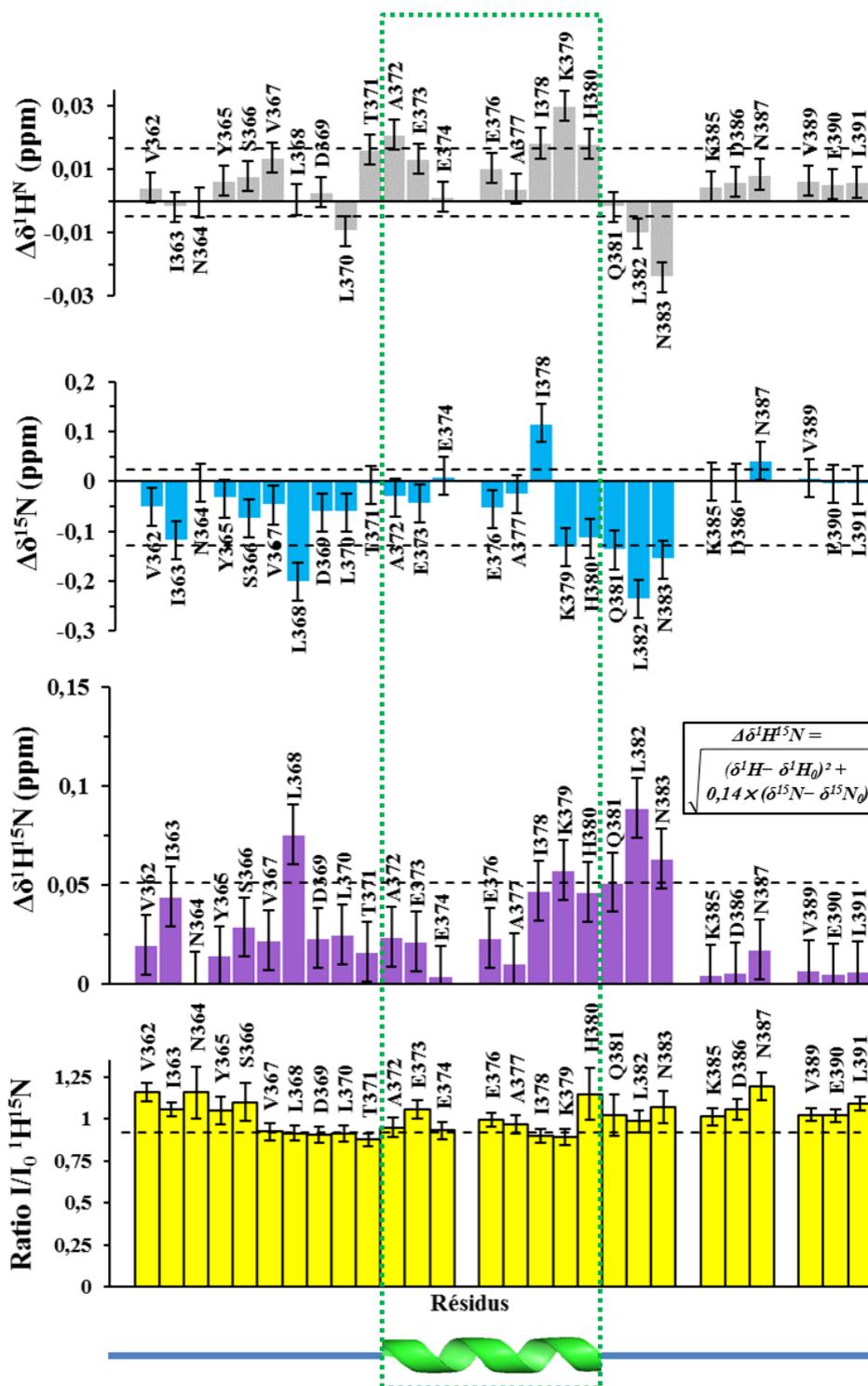
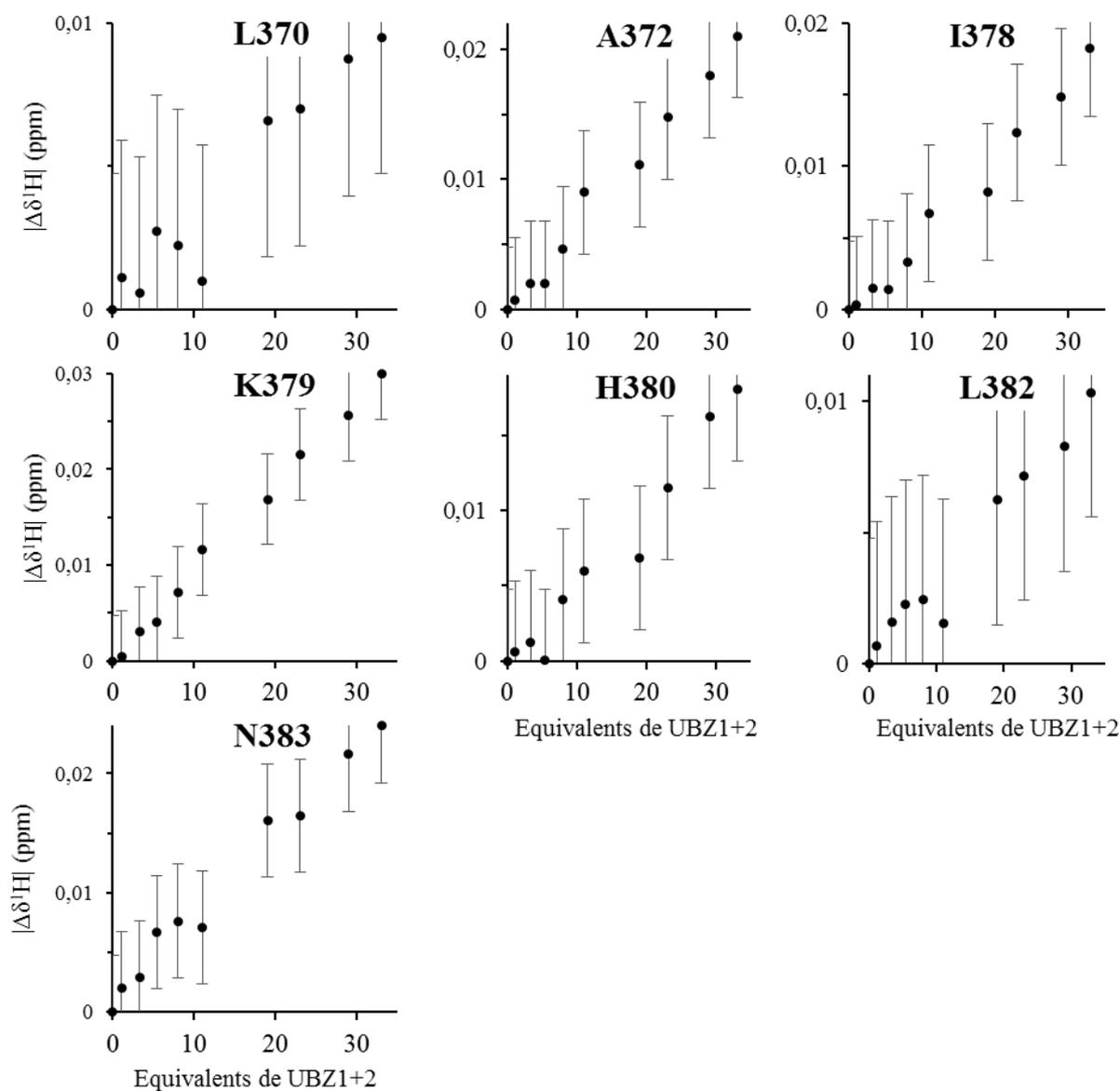
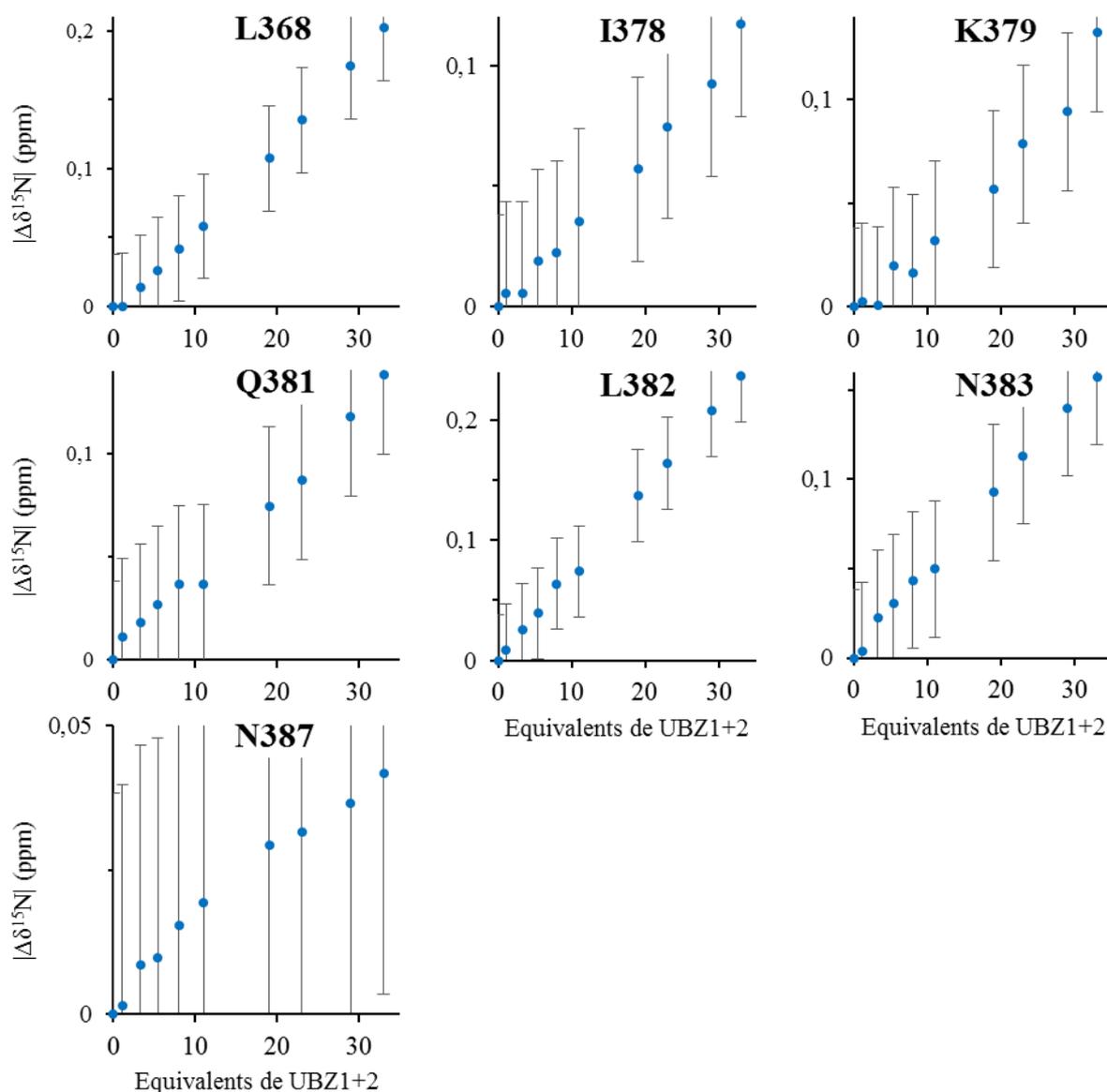


Figure 31 : Variations des déplacements chimiques  $^1\text{H}$ ,  $^{15}\text{N}$ ,  $^1\text{H}^{15}\text{N}$ , et des intensités des signaux HN par résidu de N(C arm) en présence de 33 équivalents de Tax1BP1(UBZ1+2). Les barres d'erreurs pour les calculs de perturbation de déplacements chimiques ont été produites à partir d'additions de la résolution spectrale des expériences qui ont servi à l'extraction des données. Les barres d'erreur des ratios d'intensités ont été calculées à partir du bruit des spectres HSQC  $^1\text{H}$ - $^{15}\text{N}$  (erreur =  $(I \cdot \text{bruit}_0 + I_0 \cdot \text{bruit}) / I_0^2$ ). Les lignes en pointillés noirs représentent la moyenne plus (pour  $^1\text{H}$ ,  $^{15}\text{N}$  et  $^1\text{H}^{15}\text{N}$ ) ou moins (pour  $^1\text{H}$ ,  $^{15}\text{N}$  et  $I/I_0$ ) l'écart-type. En bas du graphique, est représenté le schéma théorique de la structure secondaire de la protéine.

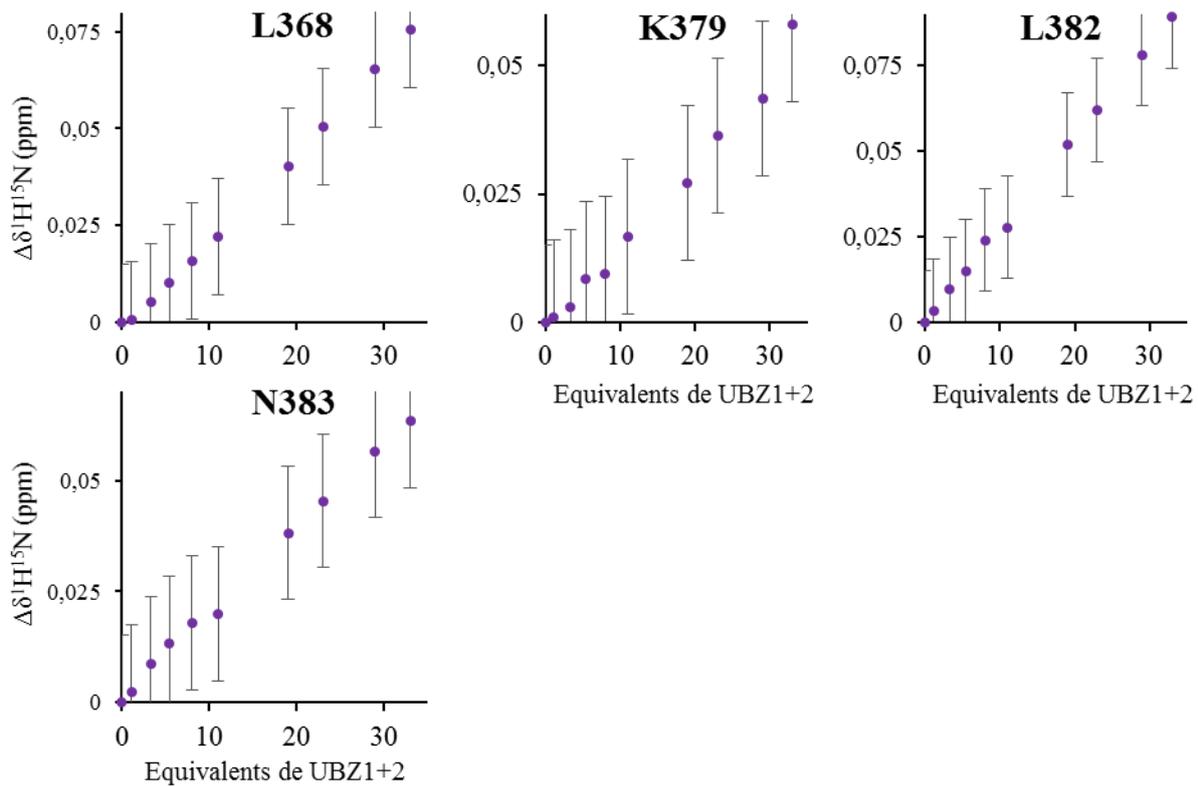
La perturbation des déplacements chimiques  $^1\text{H}^{\text{N}}$  de N(C arm) en fonction du nombre d'équivalent en peptide Tax1BP1(UBZ1+2) (**Figure 32**) ne montre pas de plateau de saturation pour les résidus dont les déplacements chimiques sont les plus affectés. A la place d'un plateau, on observe une évolution plutôt linéaire des valeurs de déplacements chimiques  $^1\text{H}^{\text{N}}$ , ce qui ressemble à un profil d'interaction non spécifique. Le résultat est similaire pour les perturbations des déplacements chimiques  $^{15}\text{N}$  (**Figure 33**) et combinés,  $^1\text{H}^{15}\text{N}$  (**Figure 34**).



**Figure 32 :** Perturbations des déplacements chimiques  $^1\text{H}^{\text{N}}$  de N(C arm) en fonction du nombre d'équivalent en peptide Tax1BP1(UBZ1+2). Seuls les résidus dont le déplacement chimique a été décalé d'au moins la valeur de la moyenne des déplacements chimiques plus ou moins un écart-type (en présence de 33 équivalents de Tax1BP1(UBZ1+2)) sont montrés. Le nom et le numéro des résidus sélectionnés sont écrits en titre de graphique. Les barres d'erreurs pour les calculs de perturbation de déplacements chimiques ont été produites à partir d'additions de la résolution spectrale des expériences qui ont servi à l'extraction des données.

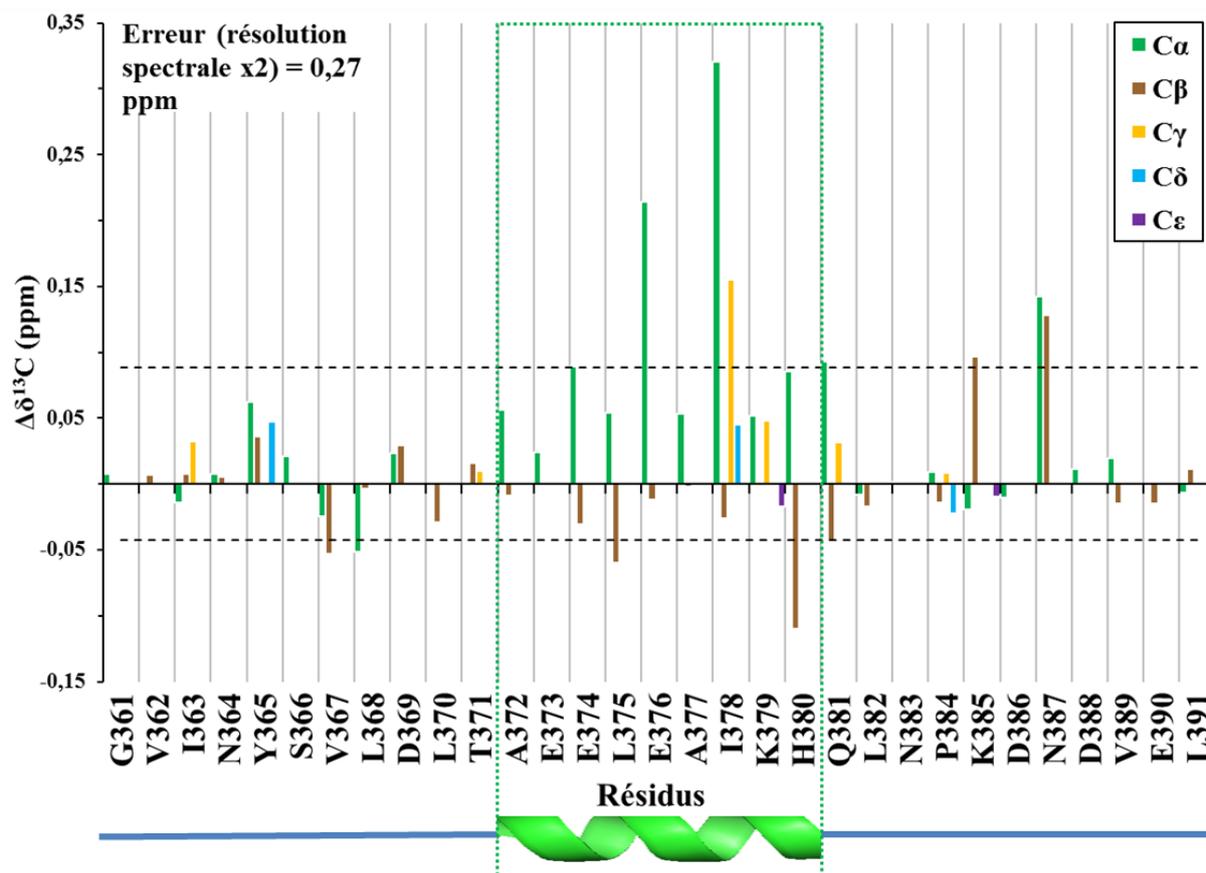


**Figure 33 : Perturbations des déplacements chimiques  $^{15}\text{N}$  de N(C arm) en fonction du nombre d'équivalent en peptide Tax1BP1(UBZ1+2).** Seuls les résidus dont le déplacement chimique a été décalé d'au moins la valeur de la moyenne des déplacements chimiques plus ou moins un écart-type (en présence de 33 équivalents de Tax1BP1(UBZ1+2)) sont montrés. Le nom et le numéro des résidus sélectionnés sont écrits en titre de graphique. Les barres d'erreurs pour les calculs de perturbation de déplacements chimiques ont été produites à partir d'additions de la résolution spectrale des expériences qui ont servi à l'extraction des données.



**Figure 34 : Perturbations des déplacements chimiques combinés  $^1\text{H}^{15}\text{N}$  de N(C arm) en fonction du nombre d'équivalent en Tax1BP1(UBZ1+2).** Seuls les résidus dont le déplacement chimique a été décalé d'au moins la valeur de la moyenne des déplacements chimiques plus un écart-type (en présence de 33 équivalents de Tax1BP1(UBZ1+2)) sont montrés. Le nom et le numéro des résidus sélectionnés sont écrits en titre de graphique. Les barres d'erreurs pour les calculs de perturbation de déplacements chimiques ont été produites à partir de la résolution spectrale des expériences qui ont servi à l'extraction des données.

Bien que le profil de l'interaction semble non spécifique, j'ai analysé les perturbations de déplacement chimique  $^{13}\text{C}\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$  et  $\epsilon$  par résidu de N(C arm) en présence de 30 équivalents de Tax1BP1(UBZ1+2). J'ai obtenu ces déplacements chimiques avec des expériences HSQC  $^1\text{H}$ - $^{13}\text{C}$  sur le même spectromètre que précédemment et j'ai utilisé les mêmes solutions tampons avec la seule différence d'avoir un pH de 6,5 au lieu de pH 6,9. Concernant l'attribution des déplacements chimiques j'ai utilisé comme référence ceux de  $^{13}\text{C}^{15}\text{N}$ -N(C arm) à 1 mM dans le tampon phosphate 20 mM, NaCl 100 mM, 7,5%  $\text{D}_2\text{O}$ , TCEP 1 mM, pH 6,5. Les résultats, sur la **Figure 35**, montrent notamment une augmentation de la valeur de déplacements chimiques  $^{13}\text{C}\alpha$  et une diminution de la valeur de déplacements chimiques  $^{13}\text{C}\beta$  pour les résidus de l'hélice  $\alpha$  372-380 ainsi que pour le résidu 381. Ces résultats semblent indiquer une rigidification de l'hélice  $\alpha$  de N(C arm) en présence de 30 équivalents de Tax1BP1(UBZ1+2) sans pour autant démontrer une quelconque région d'interaction avec Tax1BP1(UBZ1+2).

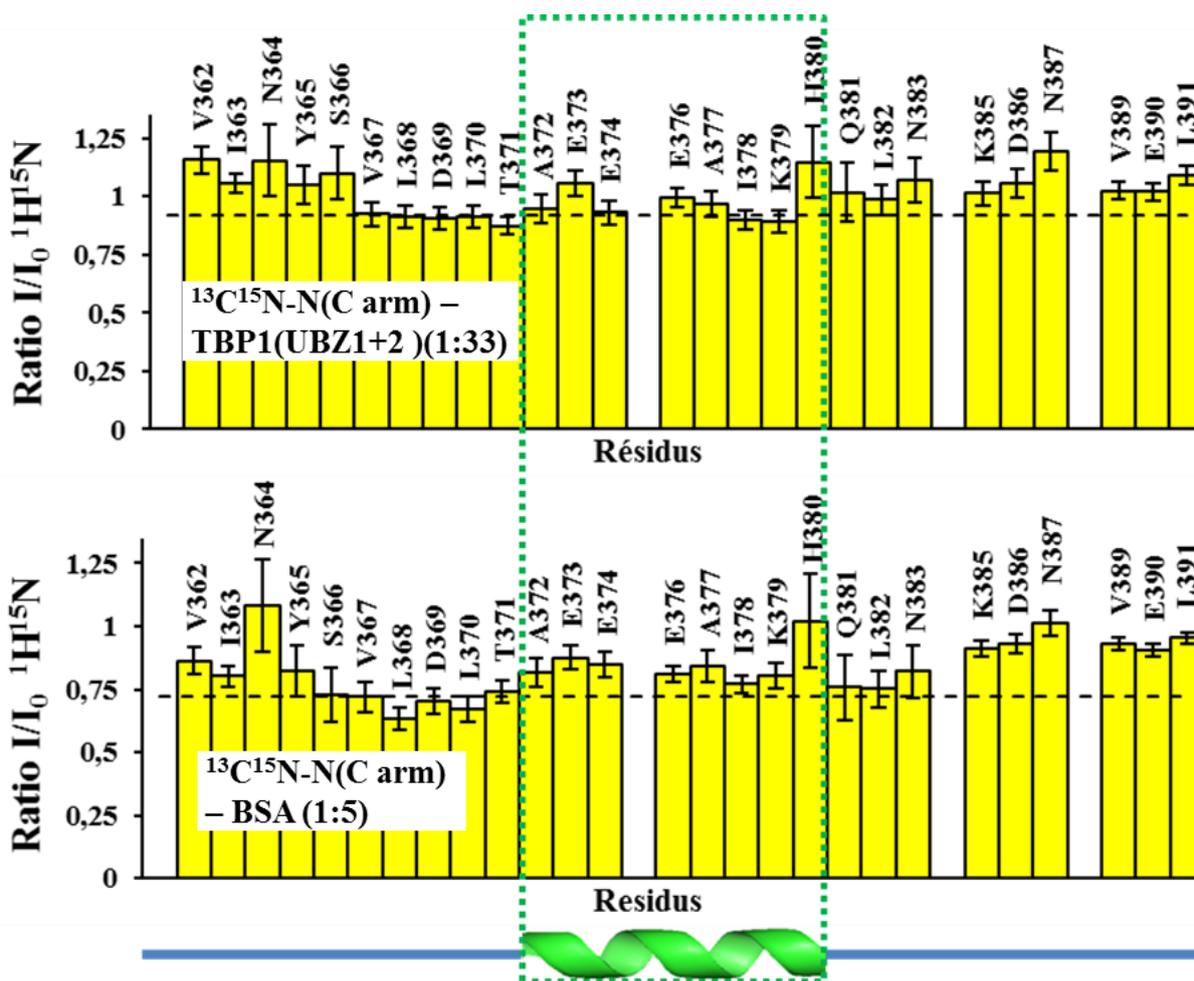


**Figure 35 : Perturbations des déplacements chimiques  $^{13}\text{C}$   $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$  et  $\epsilon$  par résidu de N(C arm) en présence de 30 équivalents de Tax1BP1(UBZ1+2).** Les lignes en pointillés noirs représentent la moyenne plus ou moins l'écart-type de toutes les données. Les spectres HSQC  $^1\text{H}$ - $^{13}\text{C}$  ayant servi à l'extraction des données ont été enregistrés à 298 K (25°C) et à 600 MHz. Les protéines étaient dans un tampon phosphate 20 mM, NaCl 100 mM, 7,5%  $\text{D}_2\text{O}$ , TCEP 1 mM à pH 6,5.

### 3.6.2. Contrôle négatif de l'effet de l'albumine sur N(C arm)

Dans l'objectif de vérifier si la baisse d'intensité des signaux HN dans la région centrale de N(C arm) est un effet spécifique de la présence de Tax1BP1(UBZ1+2) ou non, j'ai utilisé comme contrôle négatif l'albumine sérique bovine (BSA) qui, *a priori*, n'interagit pas avec N(C arm). Cette dernière a tendance à être visqueuse en solution dès que l'on approche des concentrations de l'ordre du millimolaire. Plutôt que de travailler avec des concentrations molaires équivalentes à celles de Tax1BP1(UBZ1+2), j'ai choisi de travailler avec 5 équivalents molaires de BSA, soit 500  $\mu\text{M}$ , ce qui correspond à 32,5 g/mol en concentration massique (masse molaire  $\approx$  65 g/mol) contre 26,4 g/mol pour la concentration de Tax1BP1(UBZ1+2) à 3,3 mM. Cela permet d'avoir une viscosité raisonnable, proche de celle avec 3,3 mM de Tax1BP1(UBZ1+2) et de faciliter la préparation d'échantillon. L'échantillon de N(C arm)—BSA (1:5) a été préparé dans les mêmes conditions que pour le titrage de N(C arm) par Tax1BP1(UBZ1+2), c'est-à-dire avec la même solution tampon et le même pH. La BSA a été resolubilisée dans le tampon final à partir d'un stock de BSA lyophilisée (Fraction V, Sigma) puis dialysée en compagnie de la solution mère de N(C arm) dans le tampon final.

Les résultats du contrôle sont présentés dans la **Figure 36**. Le profil des ratios d'intensité de N(C arm) en présence de la BSA montre une diminution locale au niveau N-terminal entre les résidus 365 et 372 et au niveau de l'hélice  $\alpha$  entre les résidus 374 et 379. Concernant les perturbations de déplacements chimiques, ils sont très faibles en présence BSA (données non montrées). De manière, plus générale le profil des ratios d'intensité par résidu ressemble très fortement à celui de N(C arm)—Tax1BP1(UBZ1+2) avec un profil de diminution sous forme de cloche. Cette forme en cloche fait très probablement écho à un effet de viscosité en solution dû à une de fortes concentrations massiques des protéines et non à la présence d'une protéine en particulier.

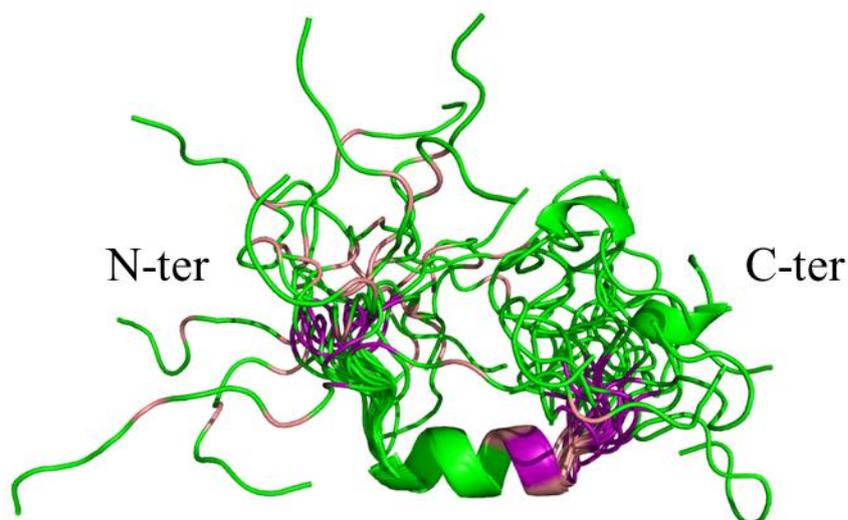


**Figure 36 :** Variations des ratios d'intensités des signaux HN par résidu de N(C arm) en présence de 33 équivalents de Tax1BP1(UBZ1+2) ou en présence de 5 équivalents de BSA. Les barres d'erreur des ratios d'intensités ont été calculées à partir du bruit des spectres HSQC  $^1\text{H}$ - $^{15}\text{N}$  (erreur =  $(I \cdot \text{bruit}_0 + I_0 \cdot \text{bruit}) / I_0^2$ ). Les lignes en pointillés noirs représentent la moyenne moins l'écart-type.

### 3.6.3. Cartographie des perturbations structurales de N(C arm) par Tax1BP1(UBZ1+2)

A partir de la structure calculée de N(C arm), j'ai tout de même représenté les résidus les plus affectés par l'ajout de 33 équivalents de Tax1BP1(UBZ1+2) au niveau des déplacements chimiques

combinés  $^1\text{H}^{15}\text{N}$  (**Figure 37**). Ici, j'ai également représenté les résidus ayant seulement une valeur de perturbation de déplacements chimiques combinés  $^1\text{H}^{15}\text{N}$  au-dessus de la valeur moyenne (résidus 363, 366, 378 et 380-381). On constate que l'ensemble de l'extrémité de l'hélice  $\alpha$  est perturbée et que les résidus concernés sont orientés sur un tour d'hélice entier. Du côté N-terminal, les résidus semblent avoir été affectés de manière plus confuse sans former une potentielle surface d'interaction. Toutefois, il est à noter que le résidu 380 en fin d'hélice est une histidine qui, par une légère modification de pH, pourrait induire des perturbations des résidus autour de ce dernier.



*Figure 37 : Représentation des résidus les plus affectés en termes de déplacements chimiques combinés HN sur la structure de N(C arm) par Tax1BP1(UBZ1+2). Les résidus dont les déplacements chimiques combinés HN sont au-dessus de 1 moyenne + 1 écart-type et ceux seulement au-dessus de 1 moyenne sont montrés en violet et en saumon respectivement.*

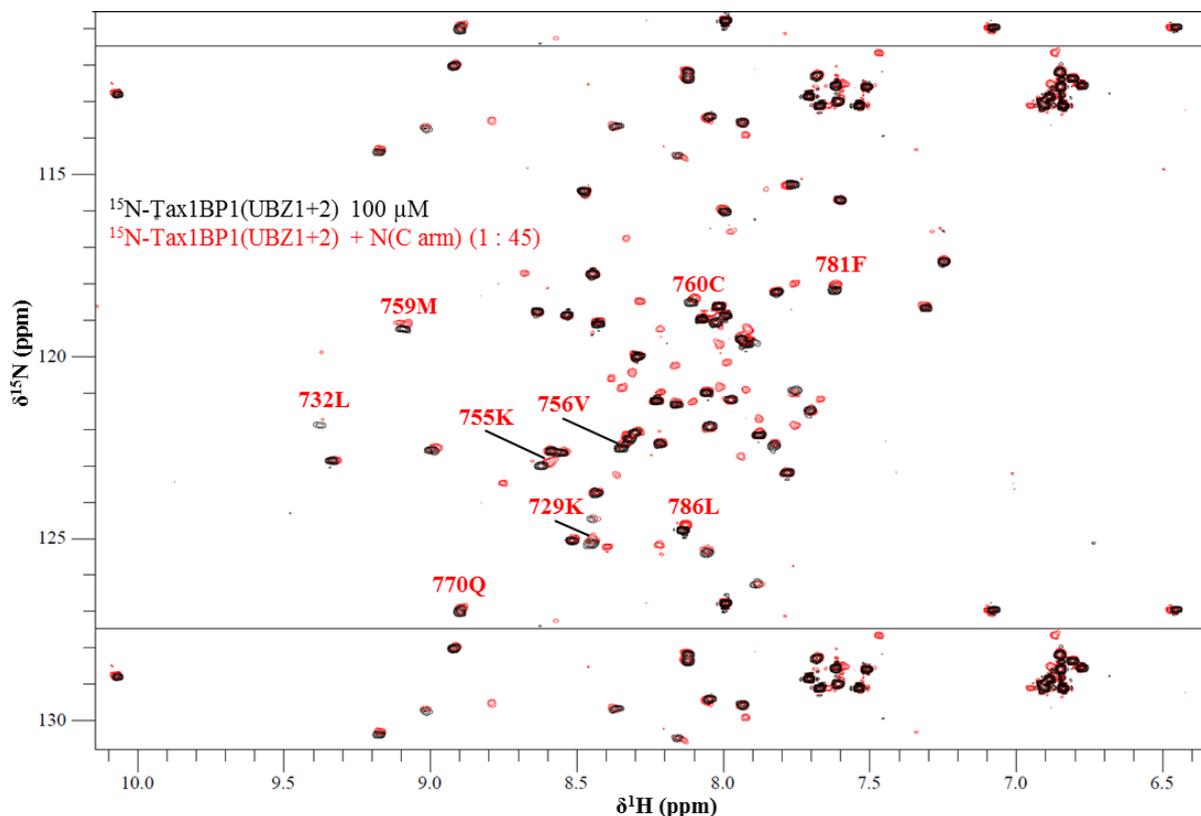
### 3.7. Etude de l'interaction entre les peptides $^{15}\text{N}$ -Tax1BP1(UBZ1+2) et N(C arm)

#### 3.7.1. Titration de $^{15}\text{N}$ -Tax1BP1(UBZ1+2) par N(C arm)

Après production et purification de Tax1BP1(UBZ1+2) marqué  $^{15}\text{N}$  et de N(C arm) j'ai fait le titrage de Tax1BP1(UBZ1+2) par N(C arm) par mélange de solutions concentrées ?. Comme pour le précédent titrage, j'ai acquis des spectres 1D  $^1\text{H}$  et HSQC  $^1\text{H}$ - $^{15}\text{N}$  sur un spectromètre 700 MHz pour chaque point du titrage ayant différentes concentrations de N(C arm) et une concentration constante de 100  $\mu\text{M}$  de  $^{15}\text{N}$ -Tax1BP1(UBZ1+2). Les 2 protéines ont été dialysées dans les mêmes bains avec un tampon phosphate 50 mM, NaCl 100 mM, 7,5%  $\text{D}_2\text{O}$ , TCEP 1 mM à pH 6,5 à 298 K (25°C).

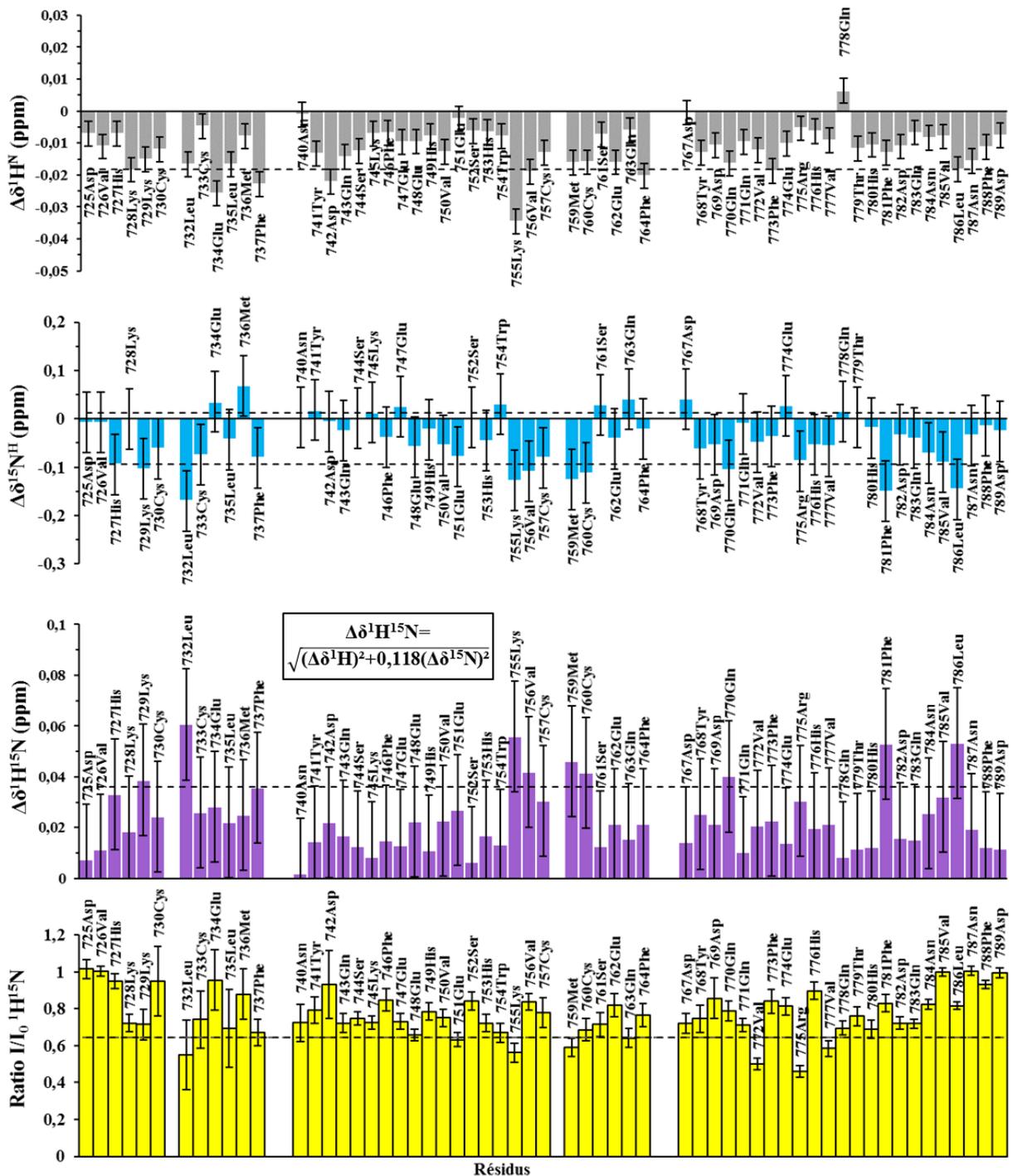
Après avoir reporté les déplacements chimiques des signaux amides  $^1\text{H}$  et  $^{15}\text{N}$  de Tax1BP1(UBZ1+2) des spectres HSQC  $^1\text{H}$ - $^{15}\text{N}$  du titrage (**Figure 38**), on peut observer qu'il n'y a pas de perturbation très importante des spectres en présence de 45 équivalents de N(C arm), mais seulement quelques signaux qui ont été décalés d'environ 1 ou 2 largeurs de raie au plus. Ces résidus un peu plus

perturbés au niveau des déplacements chimiques  $^1\text{H}^{\text{N}}$  et  $^{15}\text{N}$ , sont surtout localisés dans le domaine UBZ2 (à partir du résidu 755 jusqu'à la fin), à l'exception des résidus 729 et 732.



**Figure 38 :** Spectres HSQC  $^1\text{H}$ - $^{15}\text{N}$  superposés de  $^{15}\text{N}$ -Tax1BP1(UBZ1+2) à 100  $\mu\text{M}$  seul et avec 45 équivalents de N(C arm). Le spectre de couleur noire est celui de  $^{15}\text{N}$ -Tax1BP1(UBZ1+2) seul et concentré à 100  $\mu\text{M}$  ; il correspond au spectre de référence. Le spectre rouge est issu du titrage de  $^{15}\text{N}$ -Tax1BP1(UBZ1+2) par N(C arm) avec, ici, un ratio 1:45 (Tax1BP1(UBZ1+2) : N(C arm)). Les spectres ont été enregistrés à 298 K (25°C) et à 700 MHz et le tampon utilisé était du phosphate 50 mM, NaCl 100 mM, 7,5%  $\text{D}_2\text{O}$ , TCEP 1 mM à pH 6,5. Les numéros et codes à une lettre des résidus qui ont varié le plus par rapport à leurs déplacements chimiques sont indiqués.

Comme pour l'autre titrage, j'ai fait les calculs des perturbations de déplacements chimiques  $\Delta^1\text{H}^{\text{N}}$ ,  $\Delta^{15}\text{N}$ , combinés ( $\Delta^1\text{H}^{15}\text{N}$ ) et des intensités relatives,  $I/I_0$ , des signaux de Tax1BP1(UBZ1+2) sur la BEST-TROSY en présence de 45 équivalents de N(C arm) (**Figure 39**). Les résultats confirment que l'amplitude des perturbations de déplacement chimique HN de Tax1BP1(UBZ1+2) est relativement faible. Cependant, pour les valeurs de déplacements combinés,  $\Delta^1\text{H}^{15}\text{N}$ , les résidus 729 et 732 pour UBZ1 ainsi que les résidus 755, 756, 759, 760, 770, 781 et 786 pour UBZ2 sont les plus affectés comme l'ont montré les spectres HSQC  $^1\text{H}$ - $^{15}\text{N}$ . Concernant les différences de déplacements chimiques  $\Delta^{15}\text{N}$  et  $\Delta^1\text{H}$ , les mêmes résidus et des résidus autour de ceux qui viennent d'être relevés sont les plus affectés en présence de 45 équivalents de N(C arm). On peut également remarquer que tous les résidus dont les déplacements chimiques combinés  $\Delta^1\text{H}^{15}\text{N}$  sont les plus affectés ont des différences de déplacements chimiques  $\Delta^{15}\text{N}$  négatifs.



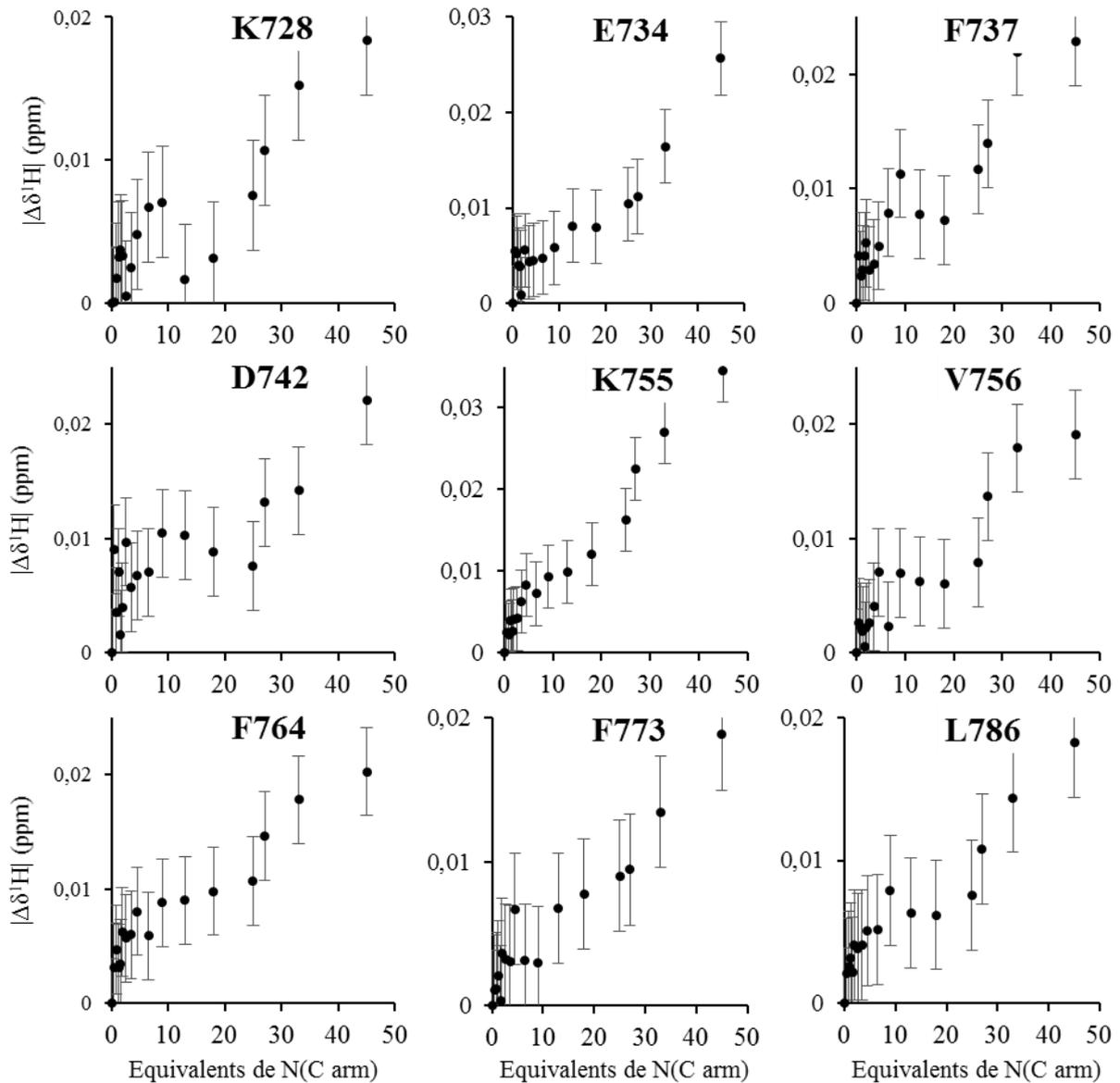
**Figure 39 :** Variations des déplacements chimiques  $^1\text{H}$ ,  $^{15}\text{N}$ ,  $^1\text{H}^{15}\text{N}$ , et des intensités des signaux HN par résidu de Tax1BP1(UBZ1+2) en présence de 45 équivalents de N(C arm). Les barres d'erreurs pour les calculs de perturbation de déplacements chimiques ont été produites à partir d'additions de la résolution spectrale des expériences qui ont servi à l'extraction des données. Les barres d'erreur des ratios d'intensités ont été calculées à partir du bruit des spectres HSQC  $^1\text{H}$ - $^{15}\text{N}$  (erreur =  $(I \cdot \text{bruit}_0 + I_0 \cdot \text{bruit}) / I_0^2$ ). Les lignes en pointillés noirs représentent la moyenne plus (pour  $^{15}\text{N}$   $^1\text{H}^{15}\text{N}$ ) ou moins (pour  $^1\text{H}$ ,  $^{15}\text{N}$  et  $I/I_0$ ) l'écart-type.

Avec une concentration élevée de N(C arm) par rapport à Tax1BP1(UBZ1+2), la courbe des intensités relatives par résidu prend une légère allure de courbe en cloche. Cette diminution des ratios

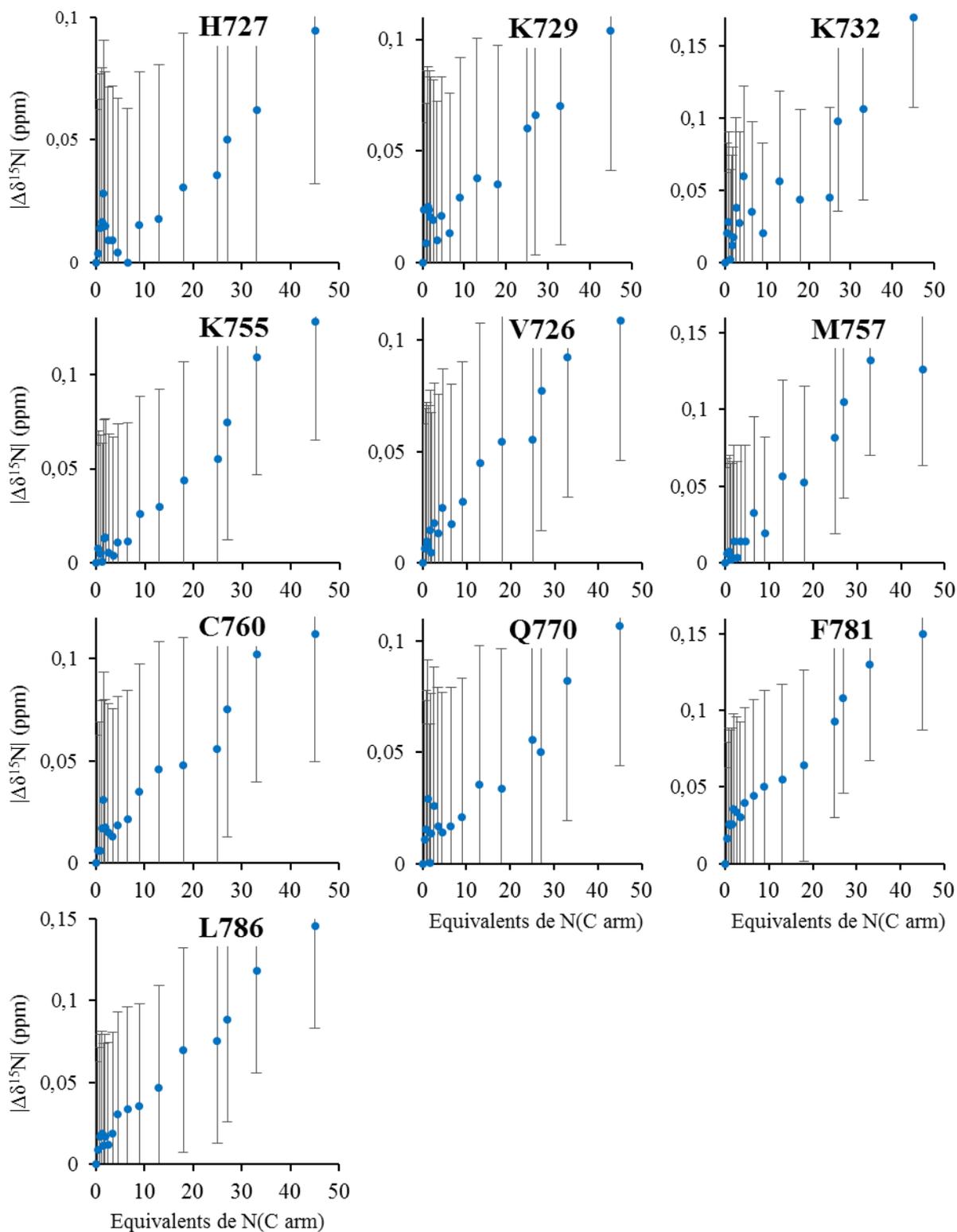
$I/I_0$  sur les 2 domaines UBZ préférentiellement aux extrémités du peptide peut être éventuellement le signe d'une augmentation de la viscosité du tampon avec la forte concentration de N(C arm) à 4,5 mM. Cette dernière hypothèse peut être vérifiée en remplaçant N(C arm) par une protéine non spécifique de Tax1BP1(UBZ1+2) (comme la BSA par exemple) à une concentration massique égale à celle de N(C arm) à 4,5 mM et en enregistrant les mêmes expériences tout en analysant de la même manière les données.

La perturbation des déplacements chimiques  $^1\text{H}^{\text{N}}$  de Tax1BP1(UBZ1+2) en fonction du nombre d'équivalents en peptide N(C arm) pour certains résidus plus affectés que d'autres (**Figure 40**) semble montrer un pseudo plateau de saturation à partir d'environ 10 équivalents de N(C arm) puis une ré-augmentation de la différence  $\Delta\delta^1\text{H}$  sans l'atteinte d'un autre plateau. On observe à peu près le même profil de courbes pour les perturbations de déplacements chimiques  $\Delta\delta^{15}\text{N}$  (**Figure 41**) notamment pour le résidu 781. Les perturbations des déplacements chimiques combinés  $^1\text{H}^{15}\text{N}$  montrent des courbes assez similaires aux perturbations déplacements chimiques individuelles (**Figure 42**). Le pseudo plateau semble correspondre à un mode de fixation spécifique de N(C arm) sur le domaine UBZ1+2 de Tax1BP1 avec une constante d'affinité  $K_d$  de l'ordre de 100-200  $\mu\text{M}$  environ. Pour ce qui est de l'augmentation des perturbations après le pseudo plateau, elle semble linéaire en générale. Cela peut donc correspondre à une fixation non spécifique de N(C arm) sur Tax1BP1(UBZ1+2) ou d'une perturbation autre liée au titrage.

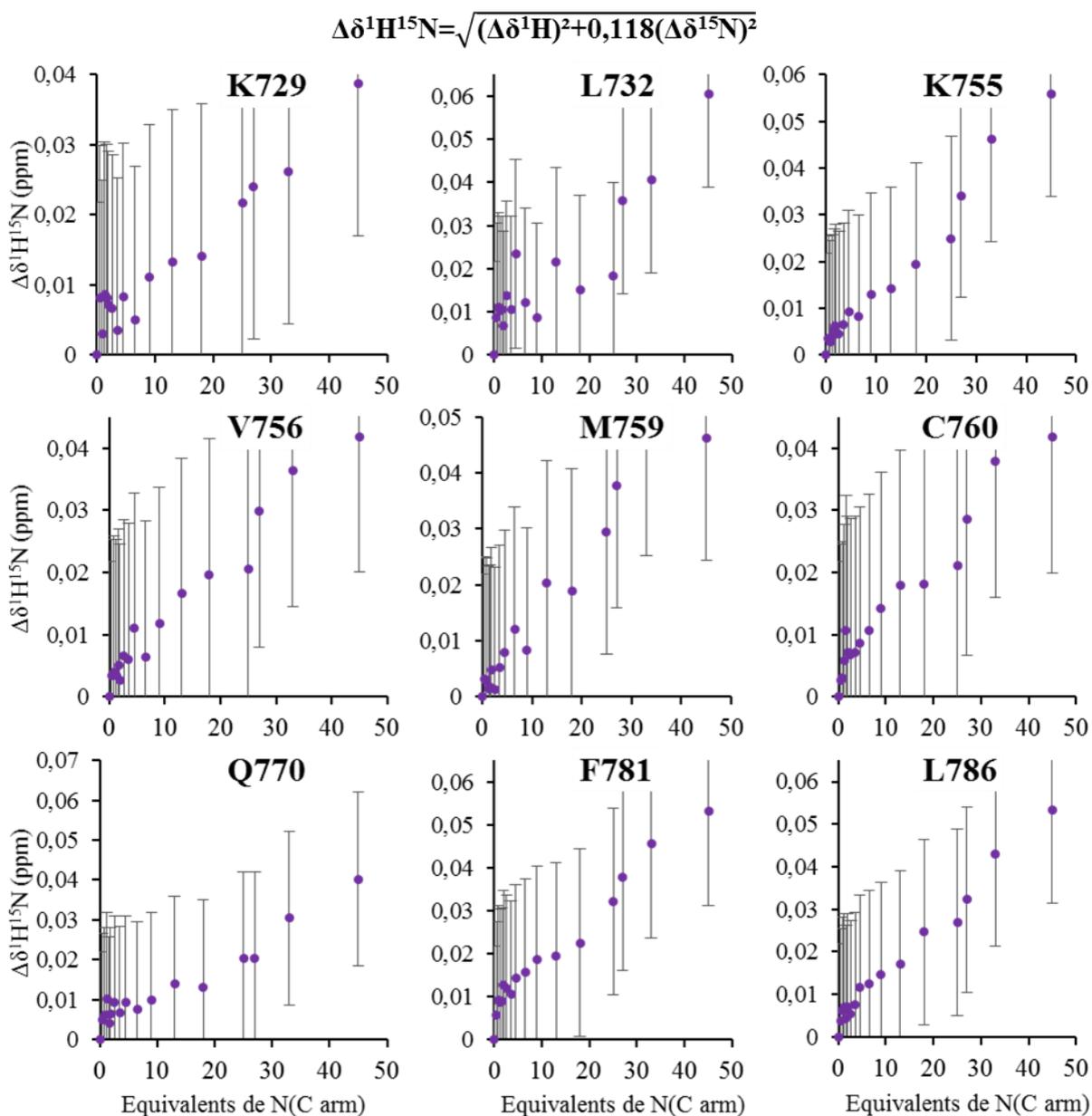
Malgré cette convergence des résultats de perturbation des déplacements chimiques de Tax1BP1(UBZ1+2) en fonction du nombre d'équivalent en peptide N(C arm) pour les résidus les plus affectés, l'amplitude de perturbation reste faible par rapport aux erreurs de mesure. Il est donc nécessaire de faire des expériences complémentaires pour confirmer ces résultats. Pour ce faire, on pourrait utiliser par exemple des acquisitions de spectres RMN donnant accès à des déplacements chimiques carbones (HNCACB, HNCOC, HSQC  $^1\text{H}$ - $^{13}\text{C}$ , HNCA) avec environ 5 points bien choisis d'un titrage de  $^{13}\text{C}^{15}\text{N}$ -Tax1BP1(UBZ1+2) par N(C arm) afin de voir si les mêmes résidus sont touchés au niveaux des déplacements chimiques carbones (en plus de découvrir d'éventuelles informations supplémentaires sur la structure secondaire ou autre). On peut aussi vérifier les résultats avec des expériences de dispersions de relaxation (RD) afin d'obtenir des informations supplémentaires sur les régimes d'échanges entre différents états de Tax1BP1(UBZ1+2) pour chacun de ces résidus en absence et en présence de N(C arm) concentrée.



**Figure 40 : Perturbations des déplacements chimiques  $^1H^N$  de Tax1BP1(UBZ1+2) en fonction du nombre d'équivalent en peptide N(C arm). Seuls les résidus dont le déplacement chimique a été décalé d'au moins la valeur de la moyenne des déplacements chimiques plus un écart-type (en présence de 45 équivalents de N(C arm)) sont montrés. Le nom et le numéro des résidus sélectionnés sont écrits en titre de graphique. Les barres d'erreurs pour les calculs de perturbation de déplacements chimiques ont été produites à partir d'additions de la résolution spectrale des expériences qui ont servis à l'extraction des données.**



**Figure 41 : Perturbations des déplacements chimiques  $^{15}\text{N}$  de Tax1BP1(UBZ1+2) en fonction du nombre d'équivalent en peptide N(C arm).** Seuls les résidus dont le déplacement chimique a été décalé d'au moins la valeur de la moyenne des déplacements chimiques moins un écart-type (en présence de 45 équivalents de N(C arm)) sont montrés. Le nom et le numéro des résidus sélectionnés sont écrits en titre de graphique. Les barres d'erreurs pour les calculs de perturbation de déplacements chimiques ont été produites à partir d'additions de la résolution spectrale des expériences qui ont servis à l'extraction des données.

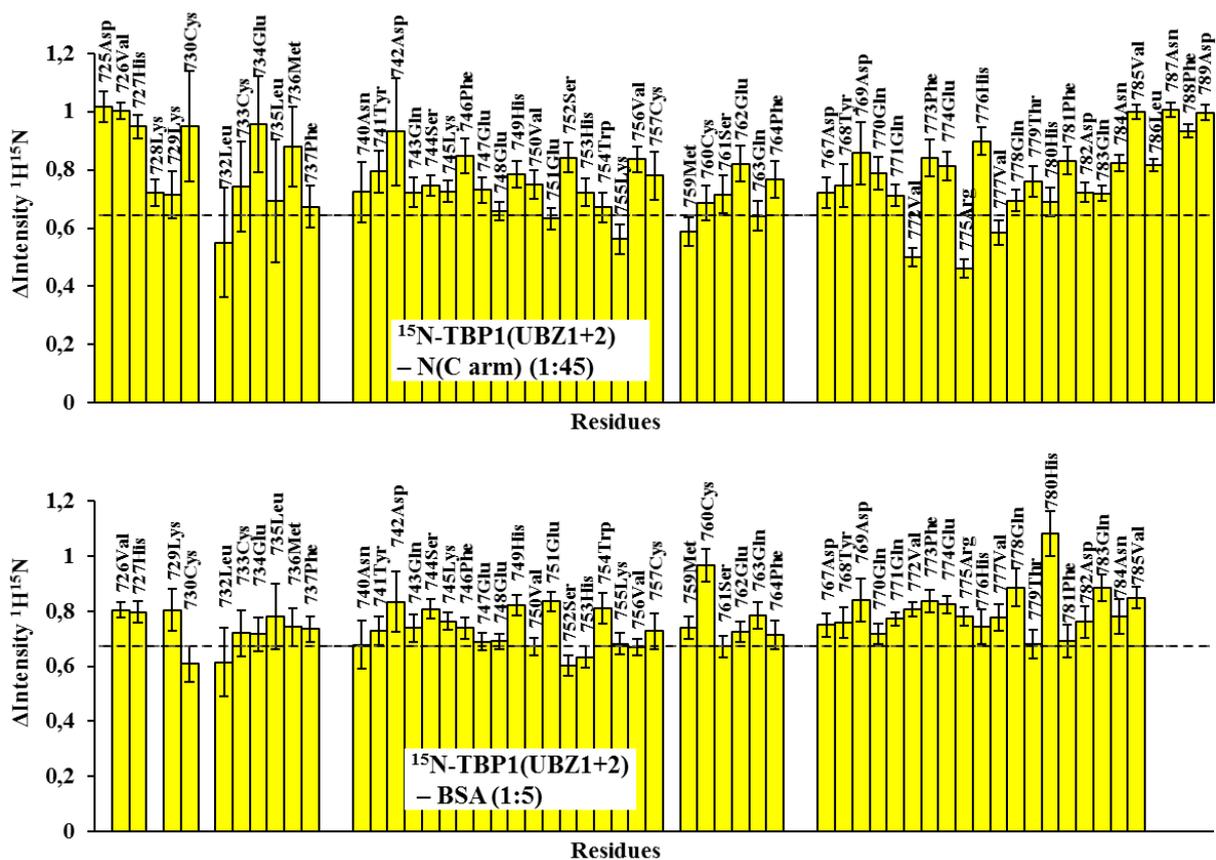


**Figure 42 :** Perturbations des déplacements chimiques combinés  $^1\text{H}^{15}\text{N}$  de *Tax1BP1(UBZ1+2)* en fonction du nombre d'équivalent en peptide N(C arm). Seuls les résidus dont les déplacements chimiques ont été décalés d'au moins la valeur de la moyenne des déplacements chimiques plus un écart-type (en présence de 45 équivalents de N(C arm)) sont montrés. Le nom et le numéro des résidus sélectionnés sont écrits en titre de graphique. Les barres d'erreurs pour les calculs de perturbation de déplacements chimiques ont été produites à partir de la résolution spectrale des expériences qui ont servis à l'extraction des données.

### 3.7.1. Contrôle négatif de l'effet de l'albumine sur Tax1BP1(UBZ1+2)

Comme pour N(C arm), nous avons voulu vérifier si la baisse d'intensité des signaux  $^1\text{H}^{15}\text{N}$  dans la région centrale de Tax1BP1(UBZ1+2) est un effet spécifique de la présence de N(C arm) ou non, en utilisant la BSA. J'ai choisi une fois de plus de travailler avec 32,5 g/mol contre 16,2 g/mol pour la concentration de N(C arm) à 4,5 mM. Les conditions de l'expérience en termes de solutions tampons sont aussi les mêmes qu'avec Tax1BP1(UBZ1+2)—N(C arm) (1:45).

Sur le **Figure 43** le profil de ratios d'intensité de Tax1BP1(UBZ1+2) en présence de BSA montre une diminution générale sans épargner les extrémités du domaine. Là aussi, les perturbations de déplacements chimiques sont presque inexistantes en présence BSA (données non montrées). Même s'il n'y a pas de diminution sous forme de cloche au niveau des diminutions d'intensité, ces résultats en présence de BSA suffisent à expliquer la baisse d'intensité des signaux  $^1\text{H}^{15}\text{N}$  en présence de N(C arm). Il semble encore y avoir encore un effet de viscosité dû à de fortes concentrations en protéines et non dû à la présence d'une protéine en particulier.

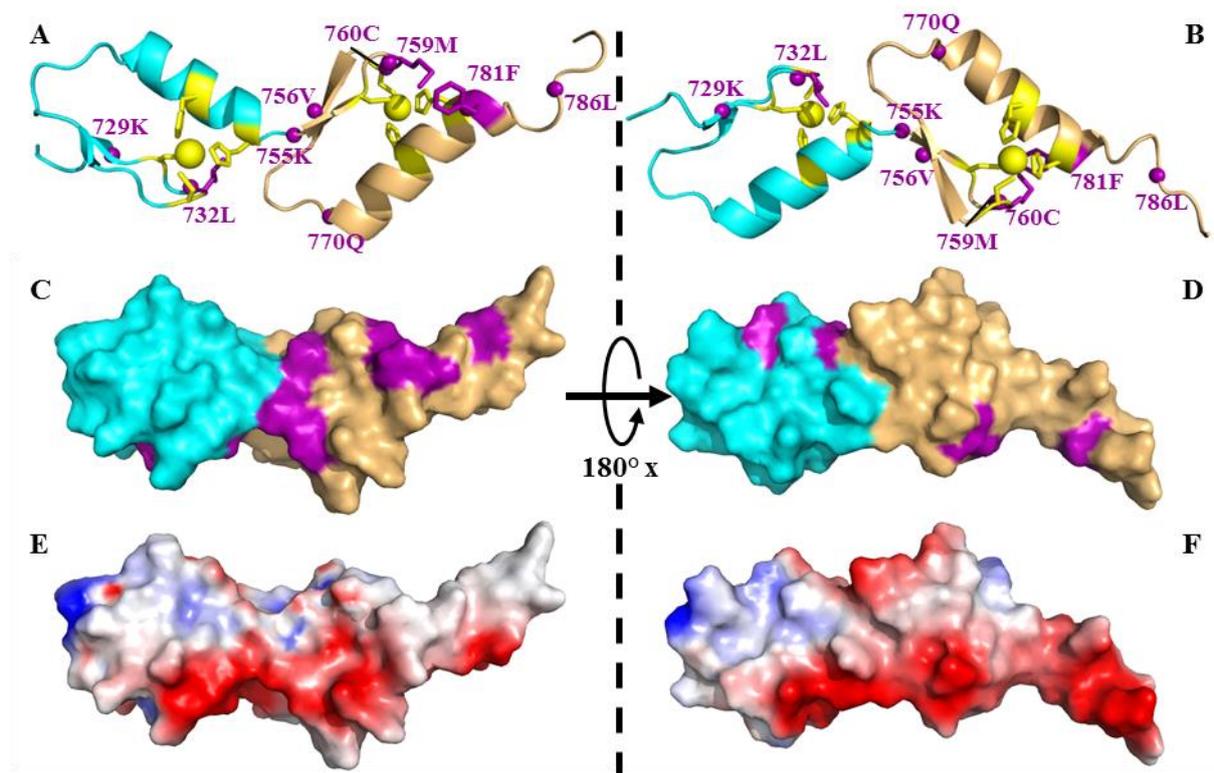


**Figure 43 :** Variations des ratios d'intensités des signaux  $^1\text{H}^{15}\text{N}$  par résidu de Tax1BP1(UBZ1+2) en présence de 45 équivalents de N(C arm) ou en présence de 5 équivalents de BSA. Les barres d'erreur des ratios d'intensités ont été calculées à partir du bruit des spectres HSQC  $^1\text{H}^{15}\text{N}$  (erreur =  $(I \cdot \text{bruit}_0 + I_0 \cdot \text{bruit}) / I_0^2$ ). Les lignes en pointillés noirs représentent la moyenne moins l'écart-type.

### 3.7.2. Cartographie des perturbations structurales de Tax1BP1(UBZ1+2) induites par N(C arm)

Ensuite, j'ai représenté sur une des structures RMN de Tax1BP1(UBZ1+2) (Ceregido et al., 2014) les atomes N amides les plus perturbés par N(C arm) en termes de déplacements chimiques  $\Delta^1\text{H}^{15}\text{N}$  (**Figure 44**). Le résultat montre bien que la plupart des résidus affectés par N(C arm) sont regroupés sur le domaine UBZ2. De plus, sur la **Figure 44C** on remarque sur la représentation en surface de la structure de Tax1BP1(UBZ1+2) que les résidus affectés se regroupent sur une face du domaine UBZ2 plutôt qu'ailleurs. De même, pour les résidus affectés sur le domaine UBZ1 ils semblent se situer côte à côte spatialement. Les **Figure 44C** et **D** montrent que les résidus affectés sur les 2 domaines UBZ sont sur des faces opposés. Cette observation peut laisser penser à 2 éventuels modes de fixation de N(C arm) qui sont :

- Une stœchiométrie de 1 : 2 (Tax1BP1(UBZ1+2) : N(C arm) ) où deux peptides N(C arm) se fixent de part et d'autres des 2 domaines UBZ de manière antiparallèle ou parallèle.
- Une stœchiométrie de 1 : 1 (Tax1BP1(UBZ1+2) : N(C arm) ) où un peptide N(C arm) se fixe autour des 2 domaines UBZ de manière hélicoïdale.

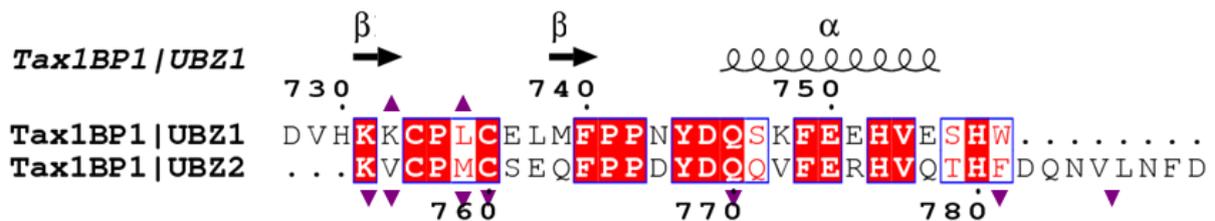


**Figure 44 : Structure de Tax1BP1(UBZ1+2) avec les résidus les plus affectés par la présence de N(C arm).** La structure représentée est un des conformères du domaine UBZ1+2 calculée à partir des données RMN de Ceregido et al (pdb 2M7Q) ; la représentation est montrée sans l'étiquette GPHM en N-terminal du peptide. (A) et (B) : Représentations en « cartoon » du peptide. Les couleurs cyan et orange clair indique respectivement les domaines N- (UBZ1) et C-terminaux (UBZ2). Les sphères

jaunes montrent les atomes de zinc chélatés par des systèmes tétraédriques de type  $Cis_2-His_2$  dont les chaînes latérales sont représentées en bâtonnets jaunes. Les sphères violettes correspondent aux atomes d'azote de la chaîne peptidique dont la perturbation de déplacement chimique combinée HN est la plus affectée (les numérotations et noms des résidus en question apparaissent de la même couleur). Les bâtonnets violets indiquent les chaînes latérales impliquées dans le cœur hydrophobe et pour lesquels la perturbation de déplacement chimique combinée HN est la plus affectée. (C) et (D) : Représentations de la structure en « surface » avec le même code couleur que (A) et (B). (E) et (F) : Représentations de la structure sous forme de « surface électrostatique » avec en rouge les parties chargées plutôt négativement, en bleu les régions plutôt positivement chargées et en blanc les zones plutôt neutres. Les parties (B), (D) et (F) sont simplement une vue de dos des figures de gauche.

### 3.7.3. Alignement de séquence en acide aminé de Tax1BP1(UBZ1+2)

Avec les séquences en acides aminés de Tax1BP1(UBZ1+2), j'ai aligné et représenté les résidus affectés par la présence de N(C arm) en solution (**Figure 45**). Le résultat montre que les résidus affectés au début de chaque domaine UBZ sont affectés au même niveau de la région où sont situées les cystéines impliquées dans la chélation du zinc pour les 2 domaines UBZ. Concernant les résidus en aval de ces régions il ne semble pas y avoir de résidus affectés au niveau des perturbations des déplacements chimiques combinés  $^1H^{15}N$  qui soient homologues. En dépit du fait que tous les résidus affectés sur les deux domaines UBZ en tandem ne soient pas alignés séquentiellement, il y a quand même une corrélation entre les amplitudes de perturbations de déplacements chimiques individuelles ( $\Delta^1H$  et  $\Delta^{15}N$ ) (**Figure 39**). Cette dernière observation est relativement vraie de manière générale sauf pour la phénylalanine 781 et le tryptophane 754 au niveau de la perturbation de déplacements chimiques  $\Delta^{15}N$  qui n'ont pas la même amplitude. Donc, même si N(C arm) était bien spécifique de Tax1BP1(UBZ1+2) il reste encore des interrogations sur le mode de fixation des deux domaines.



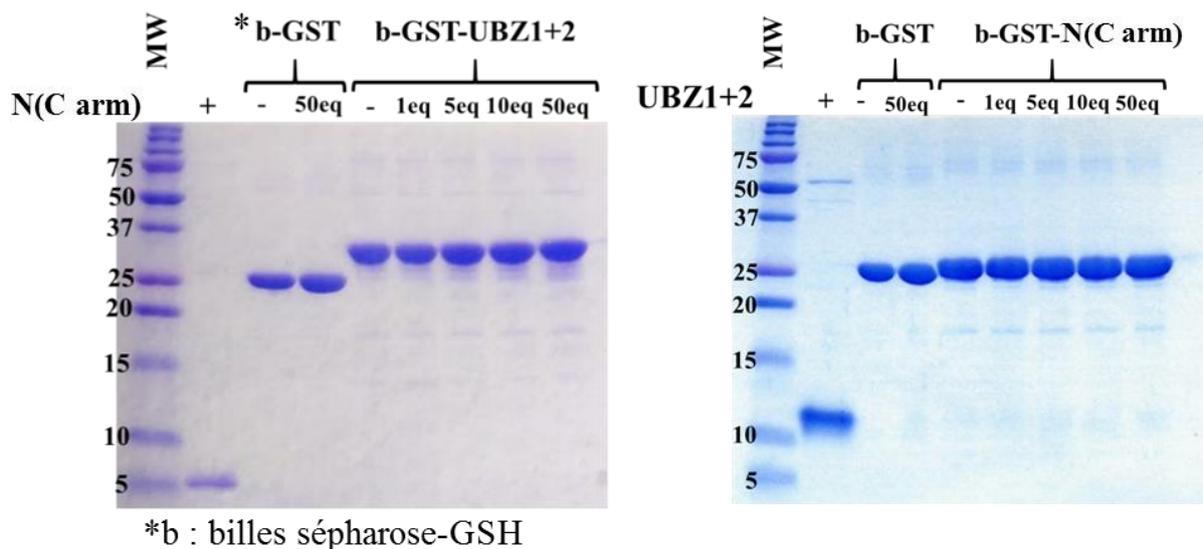
**Figure 45 : Alignement de séquences des domaines UBZ 1 et 2 de Tax1BP1 montrant les résidus dont les déplacements chimiques combinés  $^1H^{15}N$  sont les plus perturbés en présence de 45 équivalents de N(C arm). Les résidus avec un triangle violet au-dessus ou en-dessous sont ceux dont les déplacements chimiques combinés  $^1H^{15}N$  sont les plus perturbés. L'alignement de séquence a été réalisé avec Clustal Omega et préparé avec ESPRIPT 3.0.**

### 3.7.4. Expériences complémentaires de GST pulldown entre Tax1BP1(UBZ1+2) et N(C arm)

Les expériences d'interaction par RMN entre Tax1BP1(UBZ1+2) et N(C arm) n'ayant pas été concluant, nous avons refait des expériences de GST pulldown entre les 2 partenaires en solution pour essayer de comprendre ces résultats.

Pour cela, j'ai produit et purifié sans marquage isotopique GST-Tax1BP1(UBZ1+2) et GST-N(C arm) ainsi que la protéine GST seule comme contrôle négatif. Une partie des protéines chimériques a été clivée par la thrombine pour supprimer l'étiquette GST afin d'isoler Tax1BP1(UBZ1+2) et N(C arm) comme proies du GST pulldown. L'autre partie des protéines GST-Tax1BP1(UBZ1+2) et GST-N(C arm) non clivées ont été mises de côté pour être utilisés comme appâts de l'expérience, en plus de la GST. Enfin toutes les protéines ont été préparées dans la même solution tampon 20 mM phosphate, 100 mM NaCl, 1 mM TCEP, pH 6,5 (tampon identique aux précédentes expériences de titrage par RMN) et quantifiées. Après optimisation des conditions de l'expérience de GST pulldown (temps d'interaction des partenaires, nombre de lavages, pourcentage du gel en acrylamide et voltage pour la migration des protéines), j'ai opté pour une incubation des protéines à 25°C pendant 1 h sous légère agitation, 3 ou 4 lavages des billes GST-\* (GST avec ou sans protéine de fusion) puis migration sur gel SDS-PAGE 15% à 20 mA (à temps variable) par gel puis révélation au bleu de Coomassie (détaillé dans le Chapitre 2 (Matériel et méthodes)). De plus, les billes GST-\* sont à une concentration de l'ordre 100 µM.

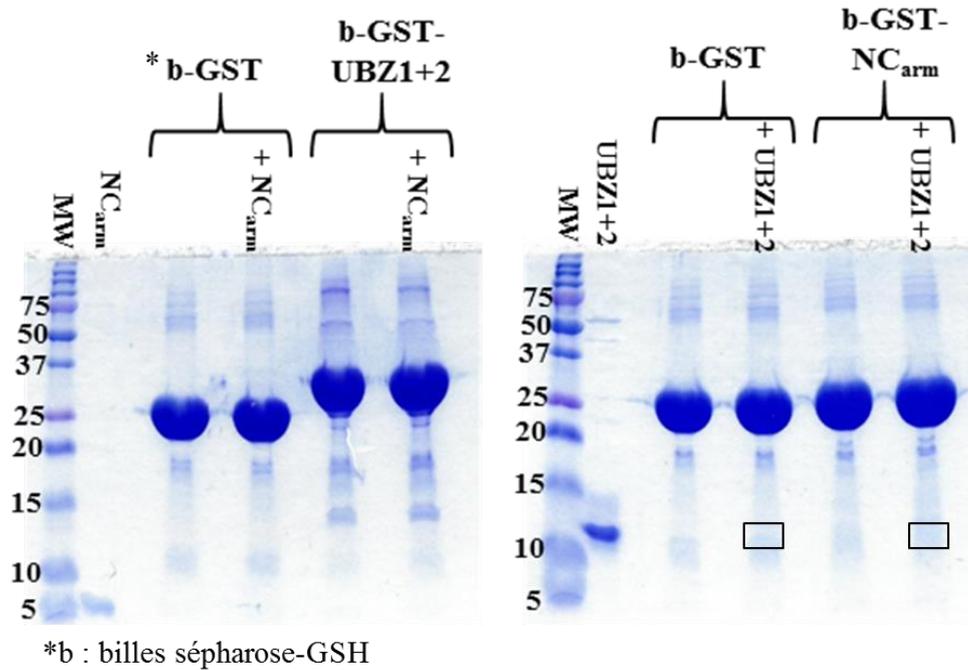
Les gels de GST pulldown de GST-Tax1BP1(UBZ1+2)—N(C arm) et GST- N(C arm)—Tax1BP1(UBZ1+2) sont montrés dans la **Figure 46**. Pour l'expérience GST-Tax1BP1(UBZ1+2)—N(C arm), on ne voit pas de bande correspondant à la taille de N(C arm) avec GST seule ou avec GST-Tax1BP1(UBZ1+2). Pour le pulldown de Tax1BP1(UBZ1+2) avec GST-N(C arm), il ne semble pas y avoir non plus de bande correspondant à la taille de Tax1BP1(UBZ1+2) que ce soit avec GST seule ou avec GST-N(C arm). Il est important de noter que je suis monté jusqu'à 50 équivalents molaires, soit 5 mM de la protéine proie pour chacune des expériences de pulldown.



**Figure 46 :** Gels des expériences de GST pulldown de GST-Tax1BP1(UBZ1+2)—N(C arm) et GST-N(C arm)—Tax1BP1(UBZ1+2). Les expériences ont été faites en mélangeant 100  $\mu\text{M}$  de billes sépharose-GSH-GST-\* (pour 50  $\mu\text{L}$  total) avec soit du tampon phosphate pH 6,5 ou plusieurs équivalents molaires de N(C<sub>arm</sub>) ou UBZ1+2 ainsi qu'une quantité suffisante de tampon pour 50  $\mu\text{L}$  totaux. Les pistes avec Tax1BP1(UBZ1+2) ou N(C arm) seules contiennent l'équivalent de 100  $\mu\text{M}$  de protéines.

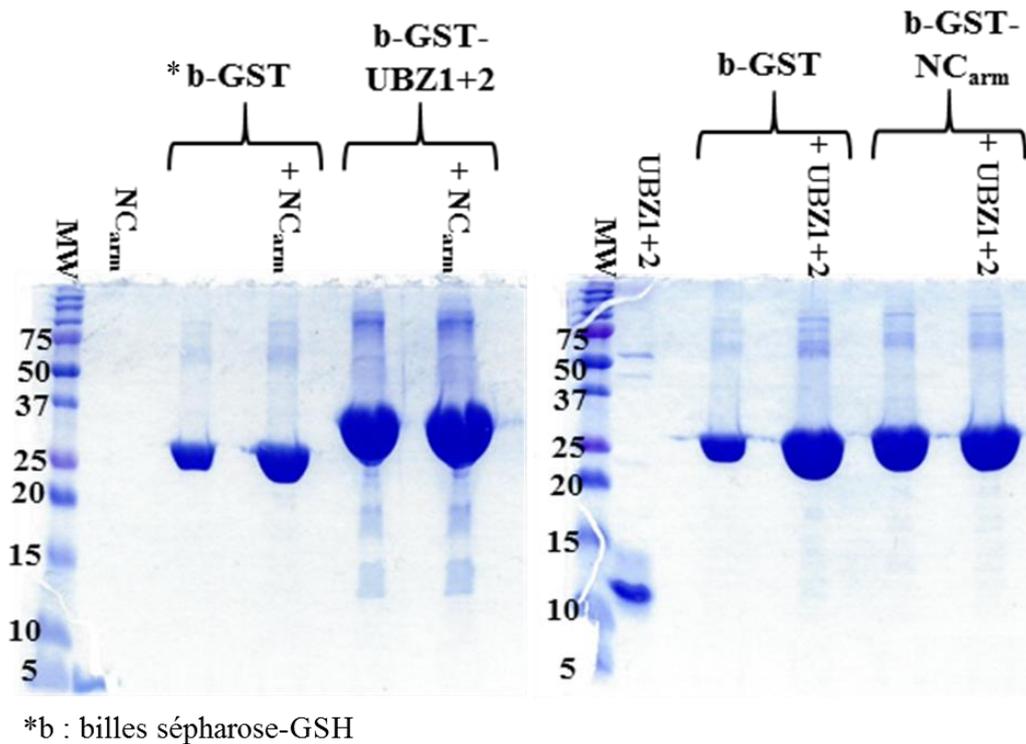
Pour être sûr de ne pas passer à côté d'une faible bande de protéines retenues, j'ai refait plusieurs fois les expériences mais en déposant cette fois 5 fois plus d'échantillons sur les gels pour espérer repérer une éventuelle bande correspondant à la taille de Tax1BP1(UBZ1+2) ou N(C arm). La **Figure 47** montre une de ces expériences de GST pulldown. Dans l'expérience présentée, j'ai utilisé uniquement 5 équivalents molaires des protéines proies car dans l'expérience précédente on voit une bande constante autour du poids de Tax1BP1(UBZ1+2) apparaître avec 5 équivalents. De plus, j'ai fait seulement 2 lavages dans cette expérience pour éviter de séparer les éventuels complexes.

Les résultats de la **Figure 47** ne montrent aucune trace d'une bande correspondant à la taille de N(C arm) co-précipitée avec les billes GSH-GST-Tax1BP1(UBZ1+2). Avec les billes GSH-GST-N(C arm), il y a une très faible bande qui pourrait peut-être correspondre à Tax1BP1(UBZ1+2), mais cette bande semble être aussi sur la piste avec les billes GSH-GST : par conséquent cela ne correspond pas à une fixation spécifique de Tax1BP1(UBZ1+2) sur GST-N(C arm).



**Figure 47 : Réplicats des expériences de GST pull-down de GST-Tax1BP1(UBZ1+2)—N(C arm) et GST- N(C arm)—Tax1BP1(UBZ1+2).** Les expériences ont été faites en mélangeant 100  $\mu$ M de billes sépharose-GSH-GST-\* (pour 50  $\mu$ L total) avec soit du tampon phosphate pH 6,5 ou plusieurs équivalents molaires de N(C<sub>arm</sub>) ou UBZ1+2 ainsi qu'une quantité suffisante de tampon pour 50  $\mu$ L totaux. Les pistes avec Tax1BP1(UBZ1+2) ou N(C arm) seules contiennent l'équivalent de 100  $\mu$ M de protéines. Ici, 5 équivalents molaires des protéines proies ont été déposées avec les billes GST\*, il y a eu seulement 2 lavages des billes et 5 fois plus d'échantillons ont été déposés sur gel.

Pour me rapprocher davantage des conditions initiales de GST pull-down du Dr Marie GALLOUX (INRA), j'ai aussi réalisé la même expérience que pour la **Figure 47** mais en utilisant un tampon 20 mM tris, 150 mM NaCl, pH 8,5 pour l'incubation des partenaires et les lavages de billes. Cependant, les gels des résultats ne laissent apparaître aucune bande correspondant à la taille de N(C arm) ou de Tax1BP1(UBZ1+2) avec leurs partenaires respectifs (**Figure 48**).



**Figure 48 : Réplicats des expériences de GST pulldown de GST-Tax1BP1(UBZ1+2)—N(C arm) et GST- N(C arm)—Tax1BP1(UBZ1+2).** Les expériences ont été faites en mélangeant 100  $\mu$ M de billes sépharose-GSH-GST-\* (pour 50  $\mu$ L total) avec soit du tampon 20 mM tris, 150 mM NaCl, pH 8,5 ou plusieurs équivalents molaires de N(Carm) ou UBZ1+2 ainsi qu'une quantité suffisante de tampon pour 50  $\mu$ L totaux. Les pistes avec Tax1BP1(UBZ1+2) ou N(C arm) seules contiennent l'équivalent de 100  $\mu$ M de protéines. Ici, 5 équivalents molaires des protéines proies ont été déposées avec les billes GST\*, il y a eu seulement 2 lavages des billes et 5 fois plus d'échantillons ont été déposés sur gel.

### 3.8. Conclusions

Les résultats préliminaires sur l'étude de l'interaction entre la nucléoprotéine du hRSV et la protéine cellulaire Tax1BP1 fait à l'INRA montrent bien une interaction directe entre les formes entières des deux protéines.

Cependant, l'interaction entre les domaines N(C arm) et Tax1BP1(UBZ1+2) semblent être soit de très faible affinités (> mM) soit les résultats observés par RMN sont probablement dus à des effets de crowding. De plus, les résultats de GST pulldown sont négatifs avec les domaines des 2 protéines partenaires isolées. Les résultats préliminaires ayant montré une interaction entre les domaines N(C arm) et Tax1BP1(UBZ1+2) avec les protéines partenaires respectives sous formes entières peuvent sûrement s'expliquer par le fait que :

- les protéines partenaires ont été utilisées sous formes entières et non tronquées en domaines,
- la construction UBZ1+2 a été très fortement surexprimée.
- le domaine N(C arm) n'a été détecté que par spectrométrie de masse (en faible quantité).

Par conséquent, il conviendrait de redéfinir les domaines d'interaction entre les protéines N et Tax1BP1.

# Chapitre 4. Caractérisation structurale de la phosphoprotéine (P) du hRSV

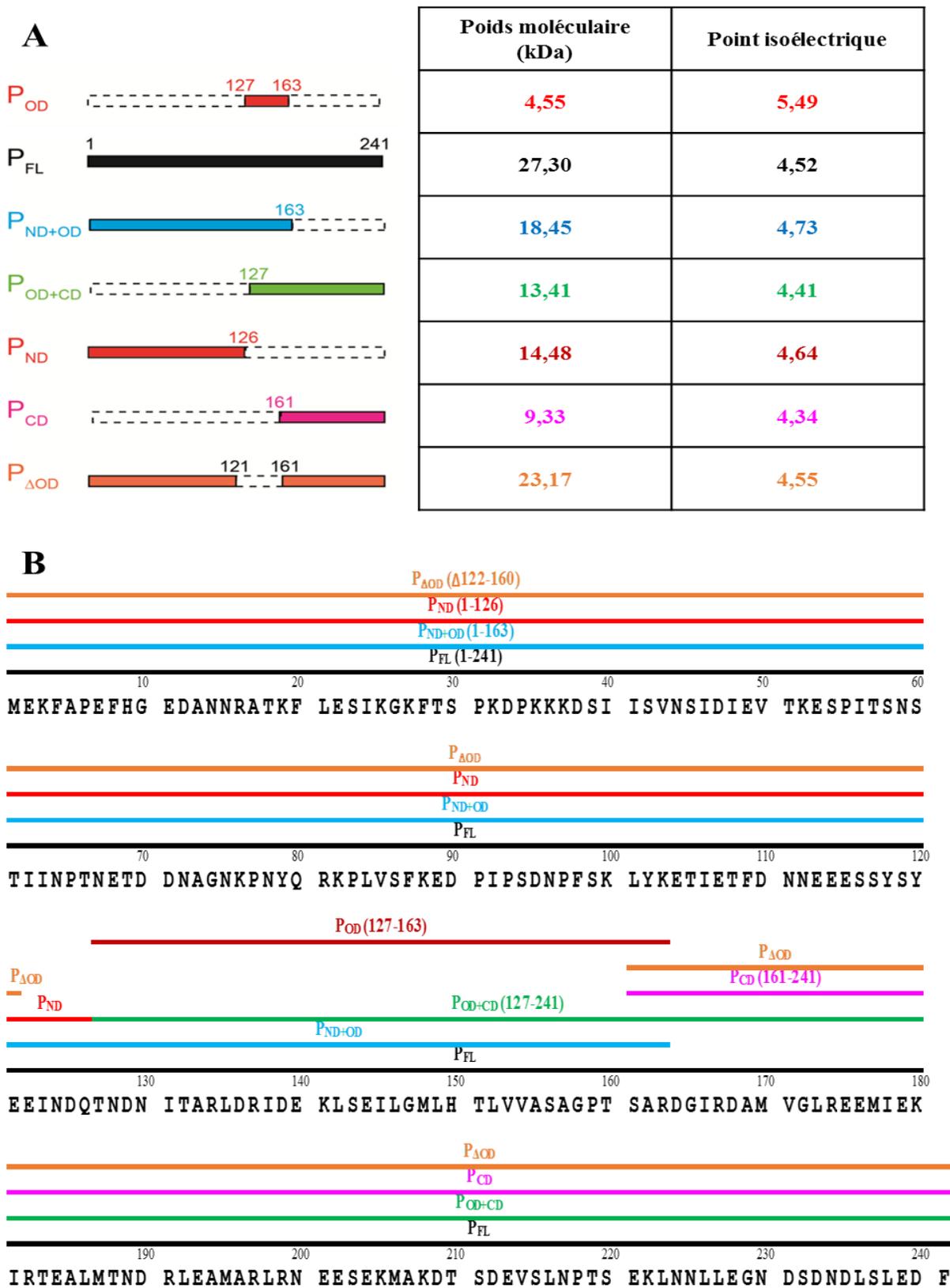
## 4.1. Intérêt de la caractérisation structurale de la protéine P du hRSV et stratégie

P étant le centre d'un réseau d'interactions avec des partenaires protéiques viraux et cellulaires, il est préférable de comprendre sa structuration et sa dynamique afin de mieux comprendre les interactions entre P et ses partenaires identifiés ou non d'un point de vue moléculaire.

Comme présenté dans l'introduction, nous savions initialement que la protéine P possède de longues régions désordonnées sur les domaines N- et C-terminaux et qu'au milieu de ces domaines se trouve un domaine d'oligomérisation prédit comme étant structuré en super hélice. A partir de ces informations, le choix a été fait au laboratoire d'étudier par domaines la protéine P et des fragments de cette dernière ont été produits au laboratoire. Il y a notamment :

- P entière ( $P_{FL}$ ),
- P[1-163] ( $P_{ND+OD}$ ) correspondant au domaine N-terminal et au domaine d'oligomérisation de P,
- P[127-241] ( $P_{OD+CD}$ ) correspondant au domaine au domaine d'oligomérisation et C-terminal de P,
- P[1-126] ( $P_{ND}$ ) correspondant au domaine N-terminal de P,
- P[ $\Delta$ 121-161] ( $P_{\Delta OD}$ ) Correspondant à P entière sans le domaine d'oligomérisation,
- P[161-241] ( $P_{CD}$ ) correspondant au domaine C-terminal de P.

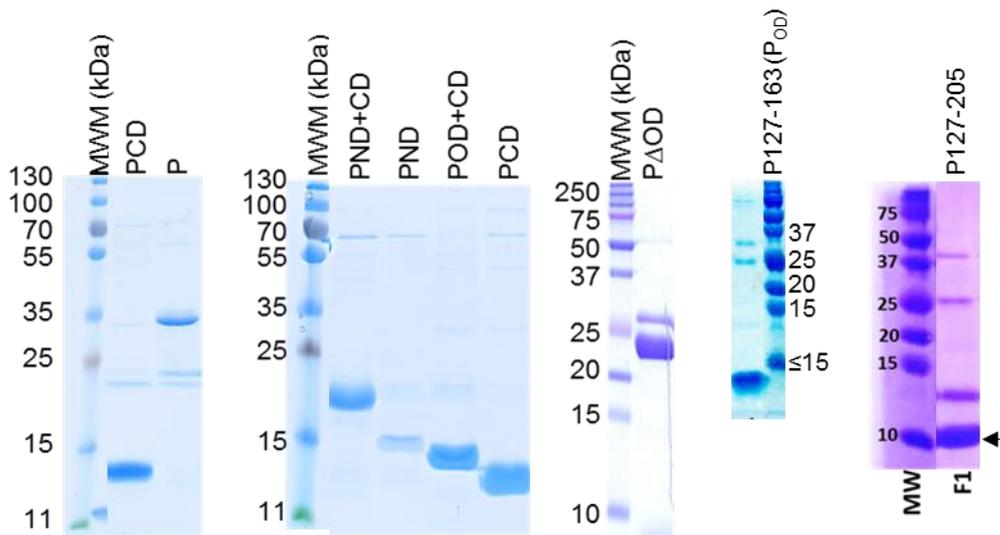
Tous ces fragments sont représentés sur la **Figure 49**.



**Figure 49 : Bornes pour les fragments de P et leurs séquences en acides aminés.** A) Schéma des différents fragments. Les régions absentes sont en pointillés. A droite sont indiqués les poids moléculaires des différents fragments et leurs points isoélectriques (pI). B, Séquence en acides aminés de P entière et indication des emplacements des fragments de P représentés par des lignes colorées au-dessus de la séquence primaire.

## 4.2. Production et purification de P entière ou en domaines séparés

La P entière et ses différents fragments ont été produits par voie recombinante dans *E. coli*, avec une étiquette de purification GST à partir de plasmides dérivés du vecteur pGEX. Cette étiquette est coupée grâce à un site de clivage à la thrombine. Plus de détails sont donnés dans le Chapitre 2 (Matériels et Méthodes). La solution tampon finale est adaptée à des mesures par RMN pour une détection des protons amides : 20 mM phosphate de sodium, pH 6,5. P et ses différents fragments sont stables dans des tampons avec 100 mM de NaCl. La pureté a été évaluée par SDS PAGE (**Figure 50**).



**Figure 50 :** Analyse par SDS PAGE de la protéine P entière et de différentes constructions de P purifiées. A gauche des pistes de chaque gel sont indiquées les tailles des marqueurs de poids moléculaires.

## 4.3. Caractérisation structurale de P entière

### 4.3.1. Principales données préliminaires

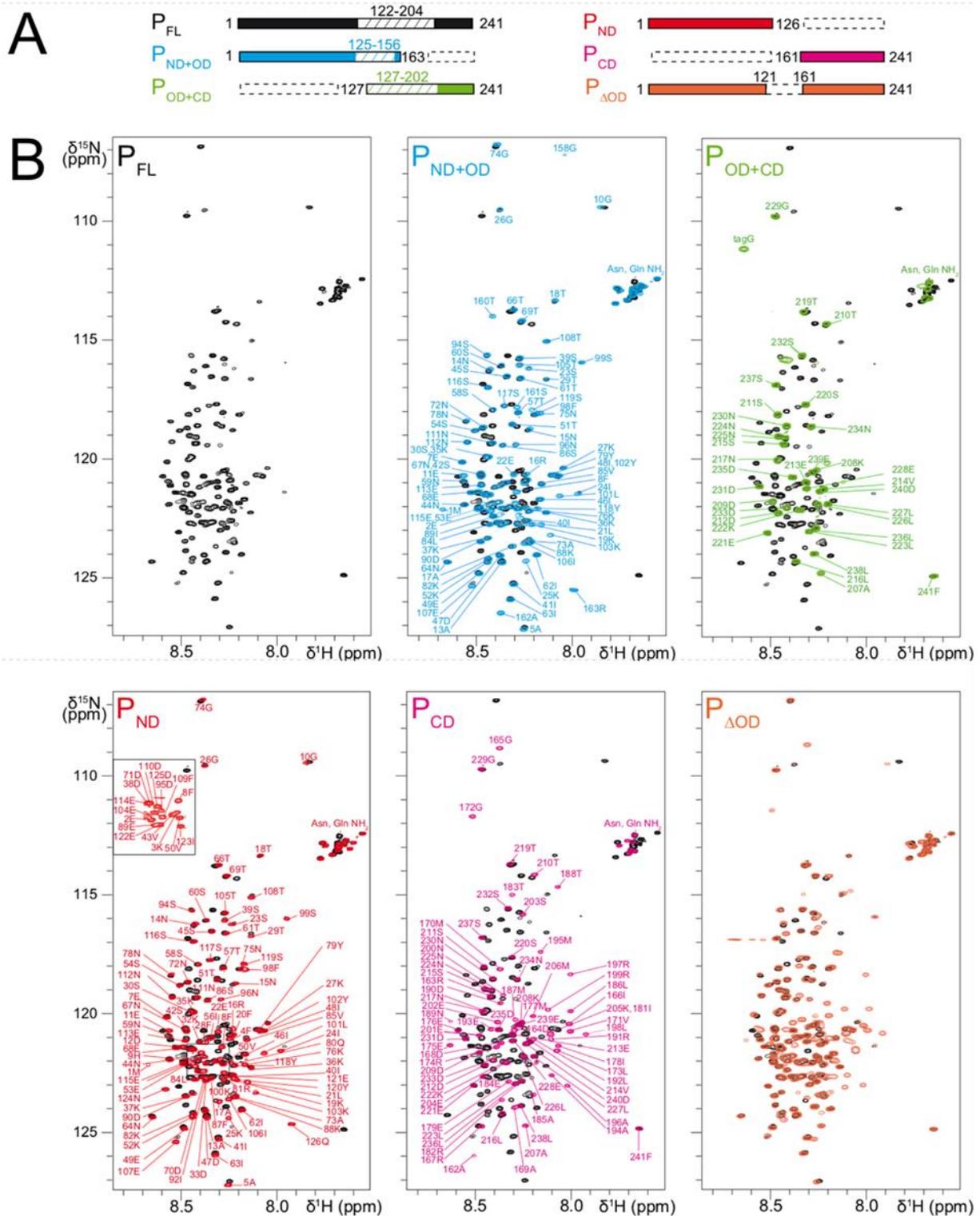
Au moment où j'ai démarré ma thèse, il n'y avait pas de données claires sur l'ensemble de la structure de la protéine P à cause, principalement, du désordre structural de cette dernière. Notre objectif a donc été dans un premier temps d'obtenir davantage d'informations sur la plasticité de la protéine P et d'en savoir plus sur d'éventuelles régions transitoirement structurées.

Les résultats présentés dans ce paragraphe et dans la **Figure 51** et la **Figure 52** qui vont suivre sont issus de travaux publiés au début de mon doctorat auxquels j'avais participé au cours de mon année de master 2 et de mes premiers mois de thèse (Pereira et al., 2016).

Avant mon arrivée au laboratoire, des spectres HSQC  $^1\text{H}$ - $^{15}\text{N}$  de P entière ( $P_{\text{FL}}$ ), P[1-163] ( $P_{\text{ND+OD}}$ ), P[127-241] ( $P_{\text{OD+CD}}$ ), P[1-126] ( $P_{\text{ND}}$ ), P[ $\Delta$ 121-161] ( $P_{\Delta\text{OD}}$ ) et P[161-241] ( $P_{\text{CD}}$ ) ont été enregistrés et les fréquences de résonances de leurs squelettes peptidiques ont été attribuées (**Figure 51**).

Sur le spectre de la protéine P entière, on remarque bien des signaux fins et une faible dispersion des signaux qui sont caractéristiques des protéines intrinsèquement désordonnées. Cependant, les signaux amides correspondant à la région centrale de P disparaissent au niveau des résidus 122 à 204. Cette région correspond à des élargissements de raies dus à des processus dynamiques et des échanges conformationnels à une échelle de temps de l'ordre des  $\mu\text{s}$ -ms. La région comprend le fragment  $P_{\text{OD}}$  (~40 résidus) et une cinquantaine de résidus en aval.

Les spectres HSQC  $^1\text{H}$ - $^{15}\text{N}$  de  $P_{\text{ND+OD}}$ ,  $P_{\text{OD+CD}}$ ,  $P_{\text{ND}}$ ,  $P_{\text{CD}}$  ont également des signaux amides fins dont les déplacements chimiques sont très proches de ceux de  $P_{\text{FL}}$ . Ils montrent un comportement indépendant des domaines N- et C-terminaux ainsi qu'un comportement particulier de  $P_{\text{OD}}$  qui n'est pas observable dans les formes  $P_{\text{ND+OD}}$  et  $P_{\text{OD+CD}}$ . De plus, il a été remarqué que les signaux amides des résidus 161 à 204 sont présents avec le fragment  $P_{\text{CD}}$  mais disparaissent en présence du domaine  $P_{\text{OD}}$  en amont.



**Figure 51 : Spectres HSQC  $^1\text{H}$ - $^{15}\text{N}$  de la phosphoprotéine du RSV marquée  $^{15}\text{N}$ .** (A) Représentation schématique des limites des fragments de protéine P utilisés en RMN. Les régions supprimées sont symbolisées par des cadres en pointillés. Les zones hachurées indiquent les régions de la protéine avec des attributions amides manquantes. (B) Les spectres HSQC  $^1\text{H}$ - $^{15}\text{N}$  de  $P_{FL}$ ,  $P_{ND+OD}$ ,  $P_{OD+CD}$ ,  $P_{ND}$ ,  $P_{CD}$ , et  $P_{\Delta OD}$  ont été acquis dans des conditions expérimentales identiques (50–100  $\mu\text{M}$ , 288 K, 14,1 T). Une couleur différente a été utilisée avec chaque construction pour représenter les contours des signaux et pour montrer les attributions. Pour tous les mutants de délétions, le spectre est superposé avec celui de

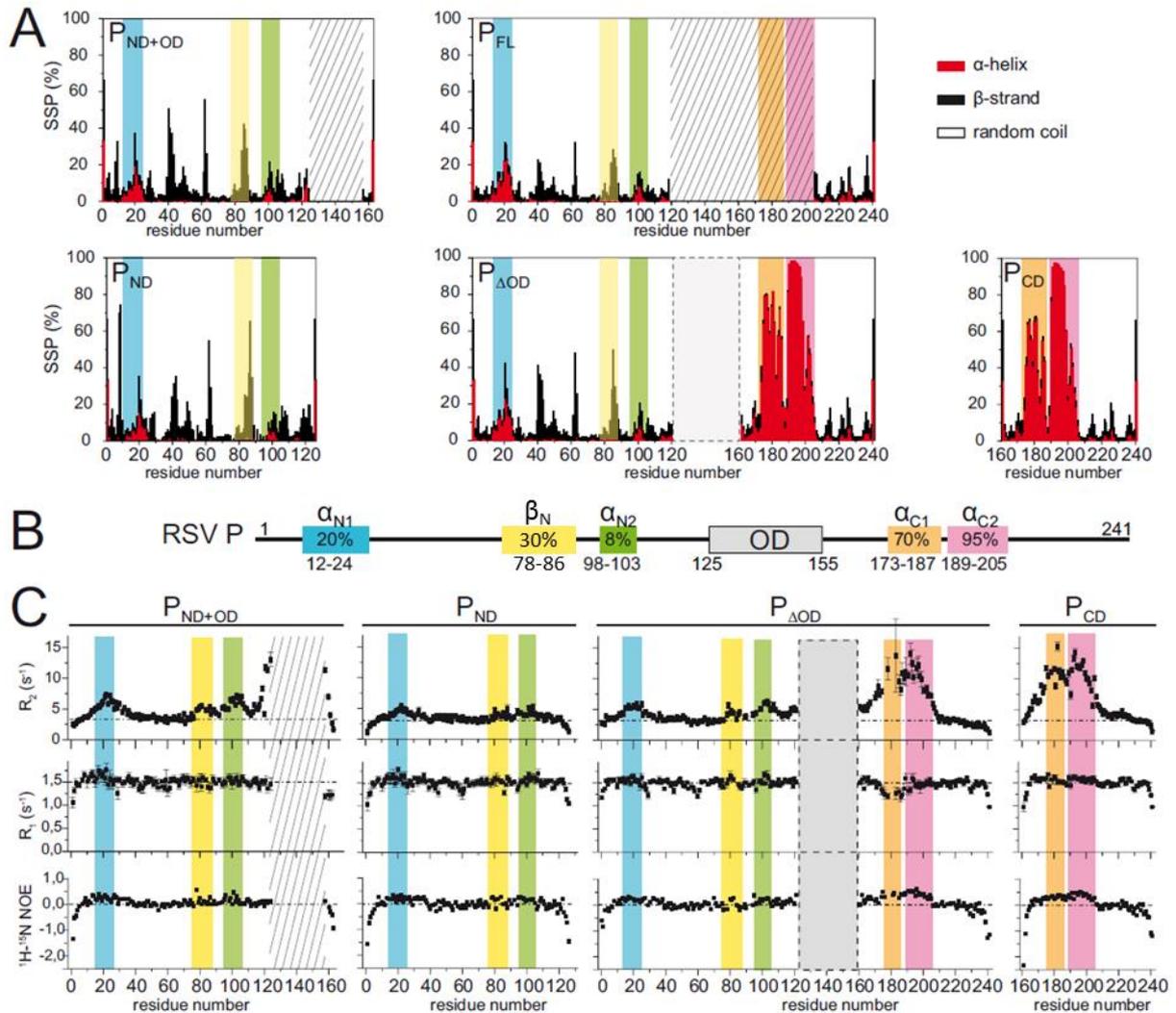
*P<sub>FL</sub>* pour comparer. Les attributions de résonance amide sont indiquées pour *P<sub>ND+OD</sub>*, *P<sub>OD+CD</sub>*, *P<sub>ND</sub>* et *P<sub>CD</sub>* (numéro de résidu et code à l lettre). Les signaux des NH<sub>2</sub> des chaînes latérales des asparagines et glutamines ne sont pas individuellement attribués. Pour *P<sub>ND</sub>* le cadre isolé en haut à gauche montre les attributions de la région centrale du spectre contenant un amas de signaux.

Avec les déplacements chimiques de fragments *P<sub>ND+OD</sub>*, *P<sub>ND</sub>*, *P<sub>FL</sub>*, *P<sub>ΔOD</sub>* et *P<sub>CD</sub>* les structures secondaires par résidus ont été prédites avec le logiciel TALOS+ (Shen et al., 2009) (**Figure 52A**). Les résultats de prédiction de structure secondaire montrent que seule la région 173-205 montre une propension significative de structure secondaire ( $\alpha_{C1}$  : résidus 173-187 avec < 70% d'hélice  $\alpha$  ;  $\alpha_{C2}$  : résidus 189-205 avec < 95% d'hélice  $\alpha$ ). Il existe également plusieurs régions dans le domaine N-terminal pour lesquelles de faibles propensions en hélices  $\alpha$  ou en brin  $\beta$  ont été trouvées ( $\alpha_{N1}$  : résidus 12-24 avec < 20% d'hélice  $\alpha$  ;  $\beta_N$  : résidus 78-86 avec < 30% de brin  $\beta$  ;  $\alpha_{N2}$  : résidus 98-103 avec < 8% d'hélice  $\alpha$ ).

Le fait que les résidus 161 à 204 disparaissent en présence de *P<sub>OD</sub>* pourrait être lié à l'augmentation du poids moléculaire lié à l'ajout de ce domaine par rapport aux constructions desquelles il a été enlevé. Il peut aussi signifier que *P<sub>OD</sub>* renforce un échange conformationnel de la région en aval des résidus 161 à 204, à l'échelle  $\mu$ s-ms, ou encore que *P<sub>OD</sub>* induit une stabilisation des hélices C-terminales, voire leur association. Cependant, il a été noté que la région  $\alpha_{C1/2}$  n'induisait pas d'oligomérisation de P en elle-même, car les signaux amides de cette région et *P<sub>CD</sub>* ont une largeur similaire au domaine *P<sub>ND</sub>* pour la construction *P<sub>ΔOD</sub>*. D'autre part, d'après Noval et al., 2016, des expériences de dénaturation thermique suivies par dichroïsme circulaire ont montré que la région C-terminale de P possède une température de transition plus basse que la région 119-161 comprenant *P<sub>OD</sub>* et cette même température de transition est augmentée de 25 à 37 °C pour la construction P[103-241] (quasiment l'équivalent de la construction *P<sub>OD+CD</sub>*), ce qui corrobore un renforcement de la stabilité des hélices  $\alpha_{C1/2}$  en présence du domaine d'oligomérisation de P sans pour autant s'y associer.

Dans la continuité, l'étude de la dynamique de relaxation nucléaire <sup>15</sup>N a été réalisée sur le fragment tétramérique *P<sub>ND+OD</sub>* et sur trois fragments monomériques (*P<sub>ND</sub>*, *P<sub>CD</sub>* et *P<sub>ΔOD</sub>*) de P (**Figure 52C**). Les régions intrinsèquement désordonnées (IDR) de P sans aucune structure secondaire prédite ont une valeur fluctuant autour de 0 concernant la valeur de l'effet Overhauser nucléaire hétéronucléaire (hetNOE), ce qui indique des fluctuations du squelette peptidique sur une échelle de temps ps-ns. La région  $\alpha_{C1/2}$  en particulier, mais aussi les régions  $\alpha_{N1}$ ,  $\alpha_{N2}$  et  $\beta_N$ , montrent des valeurs de NOE plus élevées (0-0,5), ce qui suggère qu'elles sont mieux ordonnées. Les vitesses de relaxation transverse <sup>15</sup>N ( $R_2$ ) sont plus hétérogènes le long de la séquence de P, mais cohérentes parmi les fragments de P. Les valeurs de  $R_2$  sont beaucoup plus élevées pour les régions prédites en hélice  $\alpha$  ou en brin  $\beta$  par rapport aux régions complètement désordonnées, pointant vers des mouvements de réorientation différents en solution et/ou un échange conformationnel entre états désordonnés et ordonnés, sur une échelle de temps  $\mu$ s-ms. Les vitesses de relaxation longitudinale <sup>15</sup>N ( $R_1$ ), qui sont insensibles aux échanges, sont

sensiblement uniformes le long de la séquence de P, mais soulignent la singularité structurale de la région  $\alpha_{C1}$ . Les paramètres de relaxation des IDR de P sont globalement indépendants de la longueur des fragments, ce qui suggère que leurs mouvements ne sont pas limités de façon significative et qu'ils ne sont associés de manière stable à aucune partie de la protéine.

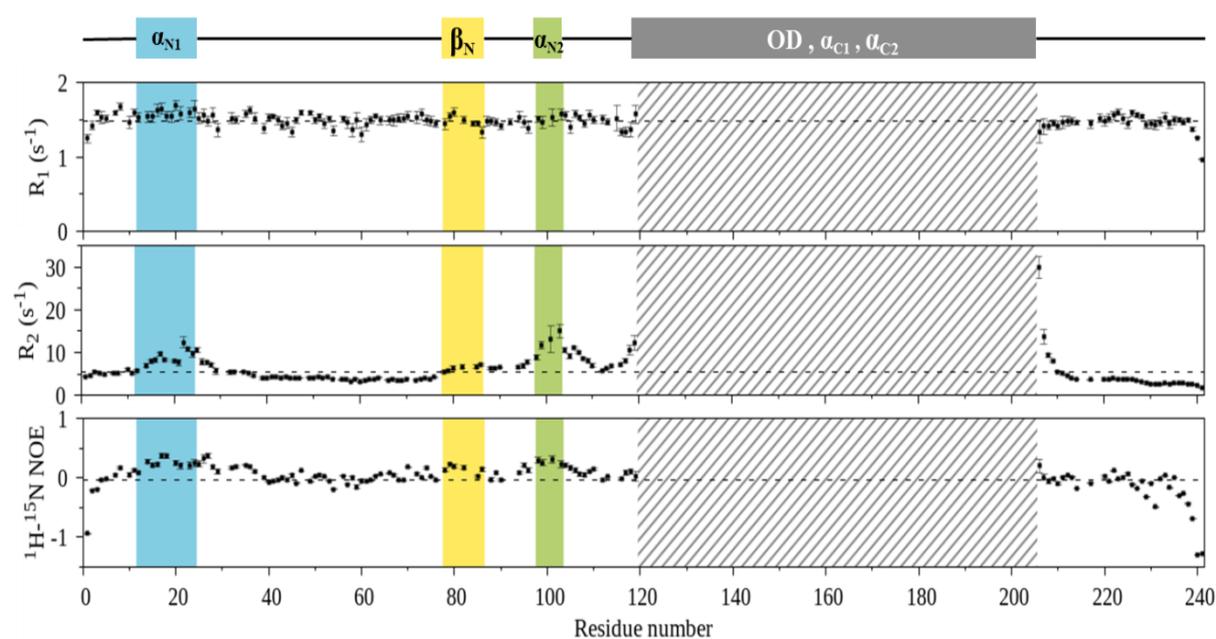


**Figure 52 : Caractérisation de l'hétérogénéité structurale de P par analyse des déplacements chimiques du squelette peptidique et expériences de relaxation  $^{15}N$ .** (A) Diagrammes empilés qui représentent les propensions de structures secondaires (hélices  $\alpha$ , brins  $\beta$ , pelote statistique) déterminées à partir des déplacements chimiques  $^1H_N$ ,  $^{15}N$ ,  $^1Ha$ ,  $^{13}C'$ ,  $^{13}Ca$  et  $^{13}C\beta$  avec le logiciel Talos+ sur différentes constructions de P. Des régions notables sont signalées par des couleurs de fonds. Les régions hachurées correspondent aux régions de la protéine où les signaux amides sont absents. (B) Organisation structurale de hRSV-P en lien avec les résultats montrés dans la partie A. Les hélices  $\alpha$  et le brin  $\beta$  transitoires sont représentés par des rectangles ayant le même code couleur que dans la partie A. La propension maximale des structures secondaires est indiquée à l'intérieur des rectangles. (C) Les temps de relaxation  $^{15}N-R_1$ ,  $^{15}N-R_2$  et NOE hétéronucléaires  $^1H-^{15}N$  ont été mesurés pour 4 fragments de P à 14.1 T et 288 K. Les barres d'erreur pour chaque point correspondent aux valeurs d'écart types calculées à partir des différences entre les modèles d'ajustement mono-exponentiel et les jeux de données expérimentaux correspondant. Les lignes horizontales en pointillés

correspondent aux valeurs moyennes en tenant uniquement en compte des résidus désordonnés. La localisation des régions en hélices  $\alpha$  et du brin  $\beta$  instables ainsi que du domaine d'oligomérisation est indiquée avec le même code couleur que dans les parties A et B. La région hachurée sur les graphiques correspond aux résidus dont les signaux  $^1\text{H}-^{15}\text{N}$  sont absents sur les spectres RMN correspondant; tandis que le rectangle grisé entouré de pointillés indique la séquence supprimée dans le fragment  $\text{P}_{\Delta\text{OD}}$ .

#### 4.3.2. Expériences de relaxation de spin nucléaire sur la protéine P entière

Dans la continuité des travaux montrés dans le paragraphe précédent (**Figure 52C**), nous avons souhaité poursuivre l'étude de la dynamique de la phosphoprotéine. En gardant les mêmes méthodes de production/purification et analyse par RMN de la protéine, nous avons simplement choisi de travailler avec la protéine P entière.



**Figure 53 : Caractérisation de l'hétérogénéité structurale de hRSV-P par analyse d'expériences de relaxation  $^{15}\text{N}$  du squelette peptidique.** Les temps de relaxation  $^{15}\text{N}-R_1$  et  $^{15}\text{N}-R_2$  et les NOE hétéronucléaires  $^1\text{H}-^{15}\text{N}$  ont été mesurés pour la protéine P entière à un champ magnétique de 14,1 T et à 288 K. Les barres d'erreur pour chaque point correspondent aux valeurs d'écart types calculées à partir des différences entre les modèles d'ajustement mono-exponentiel et les jeux de données expérimentaux correspondant. Les lignes horizontales en pointillés correspondent aux valeurs moyennes en tenant uniquement en compte des résidus désordonnés. La localisation des régions en hélices  $\alpha$  et du brin  $\beta$  instables ainsi que du domaine d'oligomérisation est indiquée au-dessus des graphiques. La région hachurée sur les graphiques correspond aux résidus dont les signaux  $^1\text{H}-^{15}\text{N}$  sont absents sur les spectres RMN correspondant.

Les résultats de la **Figure 53** montrent les différents paramètres de relaxation  $^{15}\text{N}-R_1$ ,  $^{15}\text{N}-R_2$  et NOE hétéronucléaires  $^1\text{H}-^{15}\text{N}$  pour P entière. Les résultats sont très similaires à ceux observés dans la **Figure 52C** pour le fragment  $\text{P}_{\Delta\text{OD}}$  en dehors du fait que les signaux amides des régions  $\alpha_{C1}$  et  $\alpha_{C2}$

disparaissent comme attendu. La vitesse de relaxation longitudinale  $R_1$ , insensible à l'échange, reste relativement uniforme tout le long de la séquence de la protéine P.

Concernant le paramètre  $R_2$ , les valeurs sont homogènes dans les IDR, mais augmentent sensiblement dans les régions structurées transitoirement. La vitesse de relaxation  $R_2$  augmente drastiquement sur les bords du domaine d'oligomérisation et semble atteindre environ  $30 \text{ s}^{-1}$  (contre  $3\text{-}4 \text{ s}^{-1}$  pour les IDR) pour le résidu 206 juste après la région  $\alpha_{C2}$ . Cela démontre la présence d'une région structurée et/ou en échange conformationnel dans laquelle les signaux sont probablement élargis par des mouvements de l'ordre de la microseconde à la milliseconde. Pour les régions  $\alpha_{N1}$ ,  $\alpha_{N2}$  et  $\beta_N$  on observe des valeurs  $R_2$  plus importantes, indiquant leur structuration partielle. Du reste, on observe encore une fois une homogénéité des valeurs de NOE hétéronucléaires fluctuant autour de 0 pour l'ensemble des IDR de P ainsi qu'une légère augmentation de la valeur pour les régions transitoirement repliées indiquant une dynamique du squelette peptidique de l'ordre de la picoseconde à la nanoseconde.

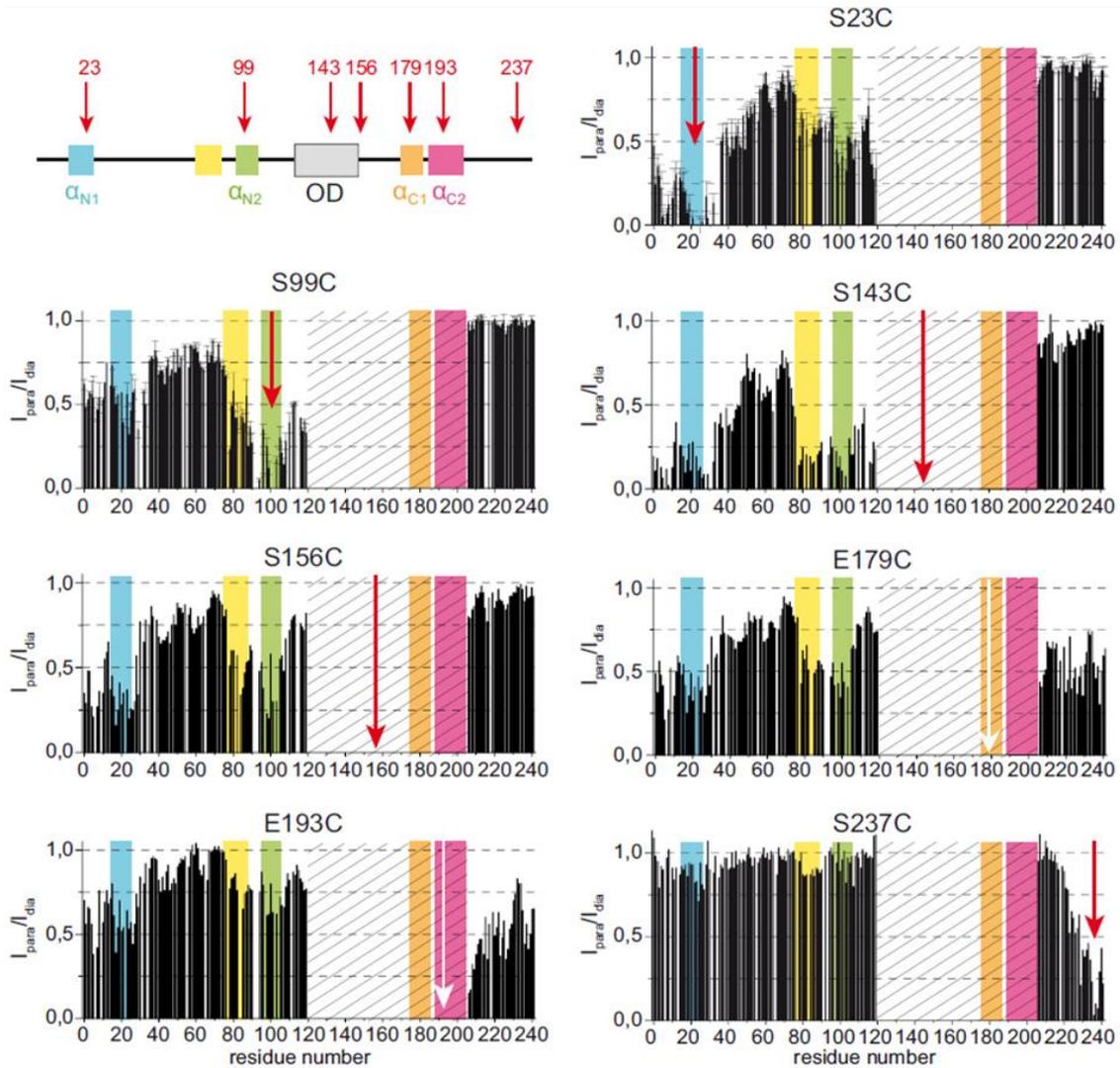
Globalement les valeurs de  $R_2$  sont plus grandes quand on compare les mêmes régions observées sur les fragments de P dans les mêmes conditions de température et de champ magnétique (**Figure 52C**), y compris pour les IDR. Cette augmentation pourrait être due à l'augmentation du poids moléculaire de  $P_{FL}$  et donc à une augmentation du temps de corrélation global par rapport aux fragments de P. Mais cet effet n'avait pas été observé sur les fragments tétramériques plus courts présentés dans le paragraphe précédent. Dans les parties structurées transitoirement, un effet de taille peut se combiner à des phénomènes d'échanges qui seraient exacerbés dans la protéine entière.

Quant à la région des hélices  $\alpha_{C1}$  et  $\alpha_{C2}$ , elle n'est pas observable dans le contexte de la protéine entière. Donc on ne peut pas conclure quant à leur comportement en comparaison avec le domaine d'oligomérisation à partir de ces données.

#### 4.3.3. Recherche de contacts à longue distance au sein de la protéine P entière

Dans l'objectif d'en apprendre plus sur la dynamique au sein de la protéine P, nous avons cherché à détecter d'éventuels contacts à longue distance en utilisant la RMN avec la technique d'augmentation de la relaxation par effet paramagnétique (PRE). Dans le cas d'une IDP, la technique permet de détecter des contacts à longue distance dans un rayon de  $15 \text{ \AA}$  autour d'une étiquette paramagnétique positionnée sur un des résidus de la protéine (Lietzow et al., 2002). J'ai utilisé comme sonde paramagnétique le radical libre 3-(2-iodoacetamido)-PROXYL (IAP) qui peut se fixer sur des cystéines libres (Gillespie et Shortle, 1997). La séquence de P ne contient pas de cystéine. Nous avons donc introduit des cystéines par mutation de résidus individuels répartis le long de la séquence de la protéine (pour minimiser les effets de variation de chaîne latérale, des sérines ont été choisies dans la mesure du possible) et je les ai marqués par l'IAP après production et purification. Les détails sont donnés dans le Chapitre 2 (Matériels

et Méthodes). Après réduction de l'étiquette paramagnétique, les mutants de P entière ne montraient pas de signaux dans la région  $\alpha_{C1/2}$  sur les spectres HSQC  $^1\text{H}$ - $^{15}\text{N}$  que j'ai enregistrés de la même manière que sur P sauvage. D'autre part, les déplacements chimiques en-dehors des régions mutées étaient les mêmes que dans la P sauvage. Ainsi, nous avons conclu que les mutations n'avaient pas d'impact sur l'oligomérisation de P. Les profils PRE (**Figure 54**) sont tous assez larges et compatibles avec des contacts diffus médiés par de grandes régions très flexibles. On observe une diminution progressive des PRE jusqu'à 40 résidus de part et d'autre de la zone de marquage dans les régions complètement désordonnées de P (positions 23, 99 et 237). De plus, toutes les étiquettes paramagnétiques, sauf en position 237, induisaient des PRE dans des régions plus distales. Des PRE plus forts ont été mesurés dans les régions comportant des éléments structuraux transitoires. Des contacts à longue distance ont été détectés dans la région désordonnée N-terminale de P en utilisant les mutants S23C et S99C, mais également entre cette région et le domaine d'oligomérisation en utilisant les mutants S143C et S156C, et même en utilisant les mutants E179C et E193C dans la région hélicoïdale  $\alpha_{C1/2}$  C-terminale. Seule l'extrémité C-terminale ne semble pas être impliquée dans des contacts avec les autres régions observables, car de forts PRE ne sont observés que pour la zone proximale de la zone de marquage paramagnétique. En outre, l'absence de PRE entre les régions désordonnées N- et C-terminales témoigne d'une organisation parallèle du tétramère du hRSV P. Ce travail a été publié également dans l'article de Pereira et al., 2016.



**Figure 54 :** Mise en évidence de contacts à longue distance dans la protéine P entière par des expériences d'augmentation de la relaxation par effet paramagnétique (PRE). Les PREs ont été mesurés avec un marquage radicalaire de type nitroxyde aux positions 23, 99, 143, 156, 179, 193 et 237 dans la séquence de P. Les zones hachurées indiquent les régions sans signal amide détectable dans l'état diamagnétique. Les barres d'erreur représentent les déviations quadratiques moyennes calculées avec deux jeux de données. Les positions de marquage de spin sont indiquées par des flèches pour chaque diagramme en barres. Les éléments de structure secondaire structure transitoire sont indiqués par des fond colorés. La figure est tirée de l'article de Pereira et al., 2016.

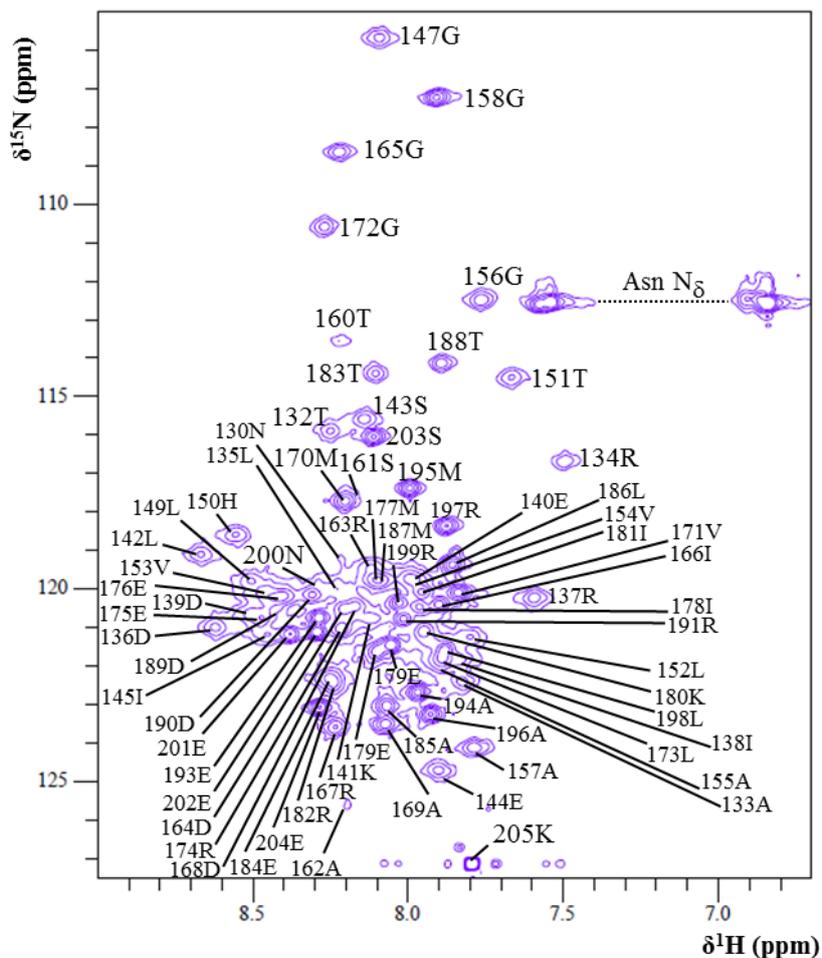
## 4.4. Caractérisation structurale de la partie centrale et des hélices C-terminales de P

### 4.4.1. Attribution des fréquences de résonance du squelette peptidique de P[127-205] et de P<sub>OD</sub>

#### 4.4.1.1. Attribution des fréquences de résonance du squelette peptidique de P[127-205]

Le fragment P[127-205] englobe le domaine d'oligomérisation et les régions  $\alpha_{C1/2}$  de P. Il a été produit, purifié et analysé par RMN. Pour réaliser l'attribution séquentielle des atomes du squelette peptidique et des  $^1\text{H}\alpha$  de P[127-205], des expériences de type HNC(O), HNCACO, HNCA, HN(CO)CA, HNCACB, CBCA(CO)NH, HBHANH, HNHA et NOESY-HSQC  $^1\text{H}$ - $^{15}\text{N}$  ont été enregistrées à 313 K (40°C) sur un spectromètre RMN de fréquence 14,1 T équipé d'une cryosonde. La température a été choisie de manière à augmenter le rapport signal sur bruit dans les spectres, en supposant que les largeurs de raie sont dominées par la relaxation de spin nucléaire intrinsèque et non des phénomènes d'échange. L'échantillon utilisé était P[127-205] marqué  $^{13}\text{C}$  et  $^{15}\text{N}$  dans un tampon Tris, NaCl 100 mM, 7,5% D<sub>2</sub>O, pH 6,5. La référence des déplacements chimiques dans la dimension proton a été obtenue avec du DSS dans les échantillons.

Un spectre HSQC  $^1\text{H}$ - $^{15}\text{N}$  du fragment P[127-205] marqué  $^{13}\text{C}$ - $^{15}\text{N}$ , avec indication des attributions des paires  $^1\text{H}$ - $^{15}\text{N}$  amides, est représenté dans la **Figure 55**. Il montre des signaux d'intensités hétérogènes, pas très fins, mais suffisamment bien éclatés au niveau des déplacements chimiques pour pouvoir piquer individuellement les différents signaux. Contrairement à ce qui a été observé avec la protéine P entière, on voit bien la quasi-totalité des signaux. Cette forme permet donc d'étudier à la fois les régions P<sub>OD</sub> et  $\alpha_{C1/2}$  dans le contexte d'un tétramère.



**Figure 55 : Spectre HSQC  $^1\text{H}$ - $^{15}\text{N}$  de P[127-205].** Le spectre a été enregistré sur un spectromètre RMN à 14.1 T à une température de 313 K (40°C). Les corrélations  $^1\text{H}$ - $^{15}\text{N}$  du squelette peptidique ont été indiquées par le numéro et la lettre du résidu auxquelles elles appartiennent. La ligne en trait pointillé sur le spectre représente la région des paires de protons  $^1\text{H}_\text{N}$  portées par les mêmes azotes sur les 2 chaînes latérales d'asparagines.

En dehors de 12 signaux amides pour la plupart situés en N-terminal, parmi lesquels on trouve la région GSGSGSGS de connexion entre l'étiquette GST qui a été coupée et la protéine, tous les signaux amides sont visibles et attribuables en dépit des superpositions sur le spectre HSQC  $^1\text{H}$ - $^{15}\text{N}$ . Les statistiques d'attribution de déplacements chimiques de P[127-205] sont données dans le **Tableau 5**.

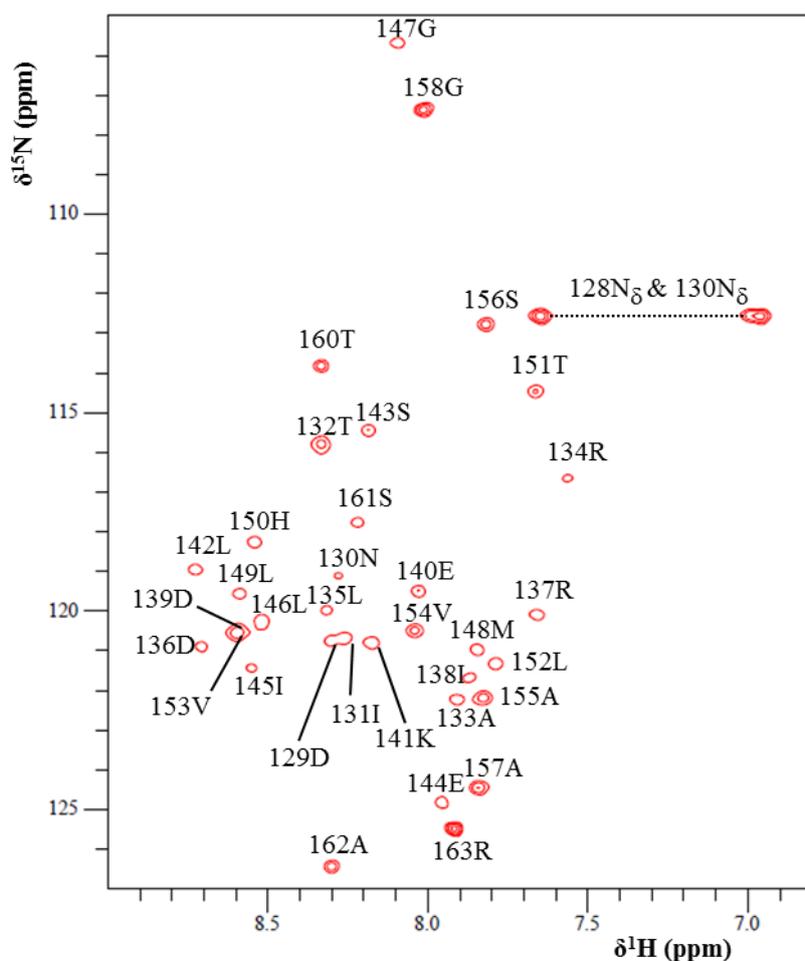
Catégorie	Disponibles	Attribués	Complétude (%)
Elément C	391	190	48,59
Elément H	490	114	23,27
Elément N	123	72	58,54
Amide	172	144	83,72
Squelette	261	204	78,16
H chaînes latérales	310	6	1,94
Type N	87	72	82,76
Type H	85	72	84,71
Type H $\alpha$	95	36	37,89
Type H $\beta$	135	6	4,44
Type C	87	66	75,86
Type C $\alpha$	87	66	75,86
Type C $\beta$	79	58	73,42

**Tableau 5 : Statistiques de l'attribution des déplacements chimiques du squelette peptidique de P[127-205].**

#### 4.4.1.2. Attribution des fréquences de résonance du squelette peptidique de P<sub>OD</sub>

De même, le fragment P<sub>OD</sub> correspondant uniquement au domaine d'oligomérisation de P a été produit, purifié et analysé par RMN. Pour réaliser l'attribution séquentielle du squelette peptidique et des chaînes latérales de P<sub>OD</sub> des expériences de type HNCO, HNCA, HN(CO)CA, HNCACB, HN(CO)CACB, HBHA(CO)NH et NOESY-HSQC <sup>1</sup>H-<sup>15</sup>N ont été enregistrées à 298 K (25°C) sur un spectromètre RMN à 950 MHz avec cryosonde. L'échantillon utilisé était P<sub>OD</sub> marqué <sup>13</sup>C et <sup>15</sup>N concentré à environ 600  $\mu$ M dans un tampon 20 mM phosphate, NaCl 100 mM, 7,5% D<sub>2</sub>O, pH 6,5. Le DSS a été utilisé pour référencer les déplacements chimiques dans la dimension proton.

Un spectre HSQC <sup>1</sup>H-<sup>15</sup>N de P<sub>OD</sub> marqué <sup>13</sup>C-<sup>15</sup>N attribuée est représenté dans la **Figure 56**. Il montre des signaux éclatés en termes de déplacements chimiques et homogènes en termes d'intensité.



**Figure 56 : Spectre HSQC  $^1\text{H}$ - $^{15}\text{N}$  de  $\text{P}_{0\text{D}}$ .** Le spectre a été enregistré sur un spectromètre RMN à 14.1 T à 298 K (25°C). Les corrélations  $^1\text{H}$ - $^{15}\text{N}$  du squelette peptidique ont été indiquées par le numéro et la lettre du résidu auxquelles elles appartiennent. La ligne en pointillés sur le spectre représente la région des paires de protons  $^1\text{H}_\text{N}$  portées par les mêmes azotes sur les chaînes latérales d'asparagines.

Seuls 10 signaux amides situés en N-terminal, dont la région GSGSGSGS de connexion entre l'étiquette GST qui a été coupée et la protéine, sont absents sur le spectre HSQC  $^1\text{H}$ - $^{15}\text{N}$  de  $\text{P}_{0\text{D}}$ . Les statistiques d'attribution de déplacements chimiques du squelette peptidique sont données dans le **Tableau 6**.

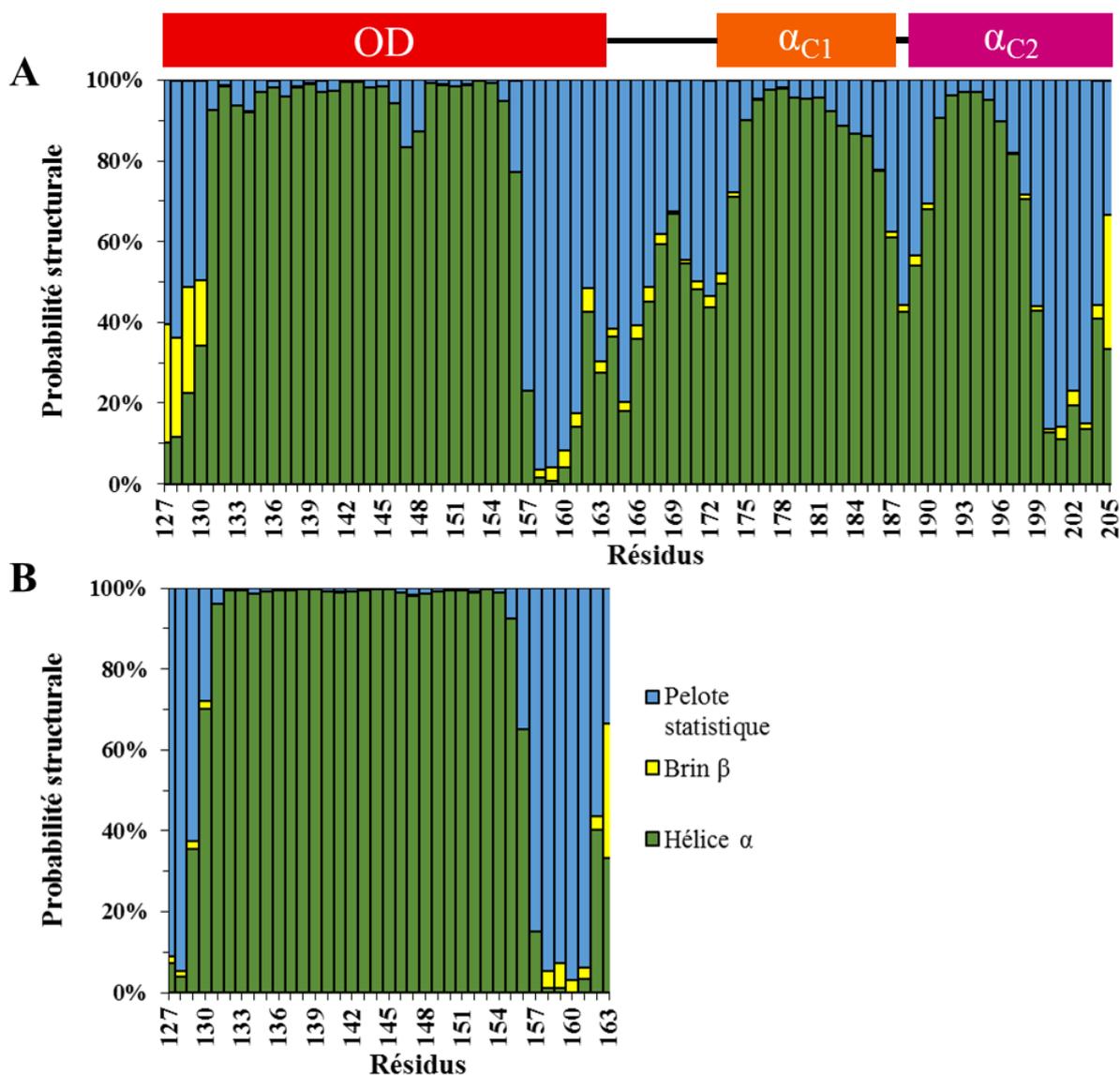
Catégorie	Disponibles	Attribués	Complétude (%)
Elément C	189	156	82,54
Elément H	232	203	87,50
Elément N	59	36	61,02
Amide	88	68	77,27
Squelette	135	107	79,26
H chaînes latérales	138	127	92,03
Type N	45	34	75,56
Type C	45	34	75,56
Type C $\alpha$	45	39	86,67
Type C $\beta$	39	36	92,31

**Tableau 6 : Statistiques de l'attribution des déplacements chimiques du squelette peptidique de P<sub>OD</sub>.**

#### 4.4.2. Prédiction de structures secondaires

A partir des attributions  $^1\text{H}_\text{N}$ ,  $^{15}\text{N}_\text{H}$ ,  $^{13}\text{C}_\alpha$ ,  $^{13}\text{C}_\beta$ ,  $^{13}\text{C}'$  et  $^1\text{H}_\alpha$  de P[127-205] marqué  $^{13}\text{C}$ - $^{15}\text{N}$  et P<sub>OD</sub> marqué  $^{13}\text{C}$ - $^{15}\text{N}$ , j'ai prédit les structures secondaires de chacun des fragments avec TALOS-N (**Figure 57**).

Les résultats pour P[127-205] font bien ressortir les 3 hélices attendues P<sub>OD</sub>,  $\alpha_{\text{C1}}$  et  $\alpha_{\text{C2}}$ . Cependant, l'hélice  $\alpha_{\text{C2}}$  semble écourtée par rapport à ce qui avait été déterminé dans les fragments P <sub>$\Delta$ OD</sub> et P<sub>CD</sub>. On constate également un enrichissement en propension en hélice  $\alpha$  de la région  $\alpha_{\text{C1}}$  par rapport à P <sub>$\Delta$ OD</sub> et P<sub>CD</sub> (avec des maxima d'environ 95% contre 70% respectivement) et une élongation vers le domaine d'oligomérisation. Ce renforcement d'hélice  $\alpha$  pourrait s'expliquer par l'absence de l'extrémité C-terminale (résidus 206-241), par la présence du domaine de tétramérisation en N-terminal ou par une stabilisation due à la température. En effet, les mesures ont été faites à 313 K pour P[127-205] et à 288 K pour P <sub>$\Delta$ OD</sub> et P<sub>CD</sub>.



**Figure 57 : Prédiction de structures secondaires pour P[127-205] et P<sub>OD</sub>.** Le calcul des probabilités de structures secondaires a été réalisé en utilisant le serveur TALOS-N (<https://spin.niddk.nih.gov/bax/software/TALOS-N/>) avec les déplacements chimiques  $^1H_N$ ,  $^{15}N_H$ ,  $^{13}C_\alpha$ ,  $^{13}C_\beta$ ,  $^{13}C'$  et  $^1H_\alpha$  des 2 fragments. (A) Diagramme empilé représentant les propensions de structures secondaires de P[127-205] à 313 K. Au-dessus du graphique est représenté le schéma des structures secondaires de la protéine selon Pereira et al., 2016. (B) Diagramme empilé représentant les propensions de structures secondaires de P<sub>OD</sub> à 298 K.

(Noval et al., 2016) avaient étudié la stabilité de la protéine P avec des expériences de dichroïsme circulaire. Ils ont observé une température de dénaturation pour un fragment équivalent à P<sub>CD</sub> vers 298 K et une première température de dénaturation à 303 K pour P entière. Ils ont interprété ces deux résultats comme reflétant la dénaturation des hélices C-terminales. Les résultats obtenus par RMN et montrés ici ne vont pas dans ce sens, puisque les hélices C-terminales sont bien présentes à 298 K. Dans (Noval et al., 2016) il manquait probablement un contrôle avec une construction équivalente à P<sub>ND+OD</sub> ou P[127-205], qui ne contiennent pas le domaine N-terminal, pour interpréter les données. Ils n'ont pas non plus

travaillé dans les mêmes conditions de tampon (Tris pH 7,5 et 50 mM NaCl) et avaient postulé l'existence d'hélices N-terminales prolongeant le domaine d'oligomérisation en N-terminal.

Concernant la prédiction de structure secondaire de P<sub>OD</sub> seul, les résultats sont très similaires à ceux de P[127-205] pour le domaine d'oligomérisation et montrent une hélice  $\alpha$  complètement formée et stable (100 % de propension) située entre les résidus 131 et 155.

Si on regarde les intensités relatives des signaux dans le spectre HSQC <sup>1</sup>H-<sup>15</sup>N de P[127-205], on s'aperçoit que les signaux correspondant à la partie du domaine d'oligomérisation sont globalement moins intenses que les signaux issus des régions  $\alpha_{C1}$  et  $\alpha_{C2}$ . Cela laisse à penser que les régions  $\alpha_{C1}$  et  $\alpha_{C2}$  sont arrangées différemment des hélices du domaine de tétramérisation et que par conséquent, elles ne sont pas associées entre elles, contrairement aux hélices du domaine d'oligomérisation.

#### 4.4.3. Détermination de la structure de P<sub>OD</sub>

Dans la lignée de tout ce qui a été fait pour comprendre la structure de P ainsi que la dynamique de ses différents éléments structuraux entre eux, je me suis intéressé dans cette partie à la détermination de la structure tridimensionnelle de la partie la plus rigide de la phosphoprotéine : P<sub>OD</sub>.

Dans cette partie, pour simplifier, je n'ai considéré que les 2 derniers résidus GS de l'étiquette (GS)<sub>x8</sub> de P<sub>OD</sub>. Les signaux amides des 6 autres résidus ne sont pas détectés, et les régions aliphatiques correspondantes sont très bruitées et peu informatives dans le spectre NOESY HSQC <sup>1</sup>H-<sup>13</sup>C. Par conséquent je n'ai pas tenu compte de ces résidus dans les calculs de structure et les statistiques d'attribution. J'assimile la séquence du fragment étudié au segment de 125 à 163, les 2 premiers résidus en N-terminal étant le G125-S126 restant de l'étiquette.

##### 4.4.3.1. Attribution complète des fréquences de résonance de P<sub>OD</sub>

Pour compléter l'attribution des fréquences de résonance du squelette peptidique par une attribution complète des chaînes latérales, un spectre TOCSY HCCH a été enregistré sur un spectromètre RMN de 800 MHz (avec cryosonde), toujours à 298 K. Pour cette expérience, l'échantillon était P<sub>OD</sub> marqué <sup>13</sup>C et <sup>15</sup>N et placé dans un tampon 20 mM phosphate, NaCl 100 mM, 100 % D<sub>2</sub>O, pH 6,5. Le résultat est donné dans le **Tableau 7**.

Catégorie	Disponibles	Attribués	Complétude (%)
Elément C	189	122	64,55
Elément H	232	165	71,12
Squelette	135	39	28,89
H chaînes latérales	138	123	89,13
Type H $\alpha$	51	42	82,35
Type H $\beta$	65	59	90,77
Type H $\gamma$	38	38	100,00
Type H $\delta$	28	23	82,14
Type H $\varepsilon$	7	3	42,86
Type C $\alpha$	45	39	86,67
Type C $\beta$	39	36	92,31
Type C $\gamma$	33	27	81,82
Type C $\delta$	21	18	85,71
Type C $\varepsilon$	3	2	66,67
Type C $\zeta$	3	0	0
Tous les résidus	45	39	86,67

Tableau 7 : Statistiques de l'attribution des déplacements chimiques de  $P_{OD}$  marqué  $^{13}C$ - $^{15}N$

#### 4.4.3.1. Contraintes d'angles dièdres

Concernant les contraintes d'angles dièdres, j'ai gardé celles déterminées par TALOS-N et considérées comme « fortes » (Tableau 8).

Numéro de résidu	Nom du résidu	Angle $\phi$ moyen	Angle $\psi$ moyen	Classification
126	S	-73,79	146,251	Dynamique
127	T	-79,757	145,87	Dynamique
128	N	-64,157	-32,15	Dynamique
129	D	-65,931	-36,845	Forte
130	N	-68,847	-35,783	Forte
131	I	-64,534	-43,859	Forte
132	T	-63,323	-42,217	Forte
133	A	-64,157	-41,407	Forte
134	R	-67,03	-40,286	Forte
135	L	-65,825	-40,647	Forte
136	D	-63,776	-40,196	Forte
137	R	-64,615	-43,473	Forte
138	I	-65,055	-44,496	Forte
139	D	-61,075	-42,496	Forte
140	E	-63,488	-43,011	Forte
141	K	-65,323	-42,348	Forte
142	L	-62,953	-42,284	Forte
143	S	-64,092	-40,996	Forte
144	E	-65,094	-43,07	Forte
145	I	-63,596	-43,349	Forte
146	L	-64,332	-38,483	Forte
147	G	-63,398	-41,511	Forte

148	M	-64,436	-44,947	Forte
149	L	-63,412	-41,647	Forte
150	H	-62,728	-43,556	Forte
151	T	-64,983	-39,921	Forte
152	L	-66,188	-42,314	Forte
153	V	-63,162	-42,861	Forte
154	V	-65,196	-39,477	Forte
155	A	-70,356	-27,674	Forte
156	S	-78,107	-22,844	Forte
157	A	-87,918	-20,989	Dynamique
158	G	98,067	-175,886	Dynamique
159	P	-63,448	146,244	Dynamique
160	T	-80,672	150,571	Dynamique
161	S	-92,321	155,153	Dynamique
162	A	99,784	-24,052	Dynamique

**Tableau 8 : Angles dièdres prédits par TALOS-N pour P<sub>OD</sub> marquée <sup>13</sup>C-<sup>15</sup>N dans l'eau. En vert sont indiqués les résidus dont les angles ont été gardés comme contraintes.**

#### 4.4.3.1. Calcul de structure d'un protomère isolé de P<sub>OD</sub>

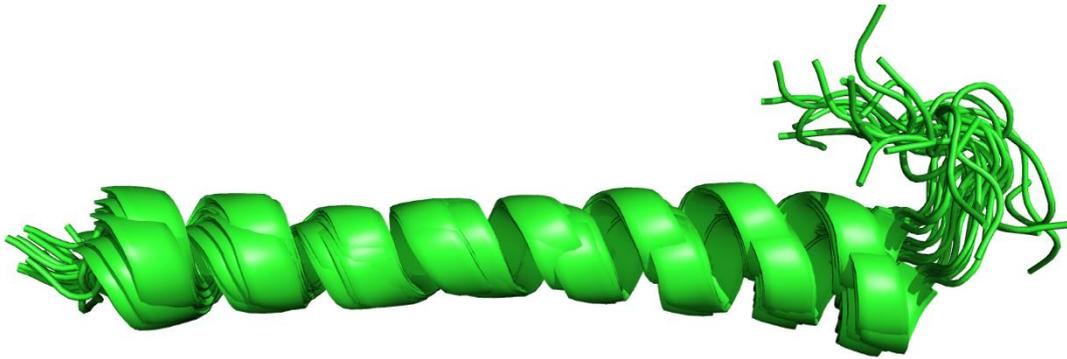
Avant de commencer à calculer la structure du tétramère de P<sub>OD</sub>, j'ai d'abord procédé au calcul de la structure d'un protomère isolé à l'aide du logiciel CYANA. Pour ce faire, j'ai utilisé les contraintes d'angle dièdre décrites dans le **Tableau 8** et des contraintes de distance issues du spectre NOESY HSQC <sup>1</sup>H-<sup>15</sup>N d'un échantillon de P<sub>OD</sub> marquée entièrement <sup>13</sup>C-<sup>15</sup>N dans l'eau, qui avait déjà été exploité au cours de l'attribution séquentielle pour vérifier les déplacements chimiques <sup>1</sup>H $\alpha$ , et d'un spectre NOESY HSQC <sup>1</sup>H-<sup>13</sup>C enregistré sur un échantillon de P<sub>OD</sub> marquée entièrement <sup>13</sup>C-<sup>15</sup>N dans 100% D<sub>2</sub>O sur un spectromètre RMN de 22,3 T (avec cryosonde), toujours à une température de 298 K, pour obtenir des informations de distances entre protons de chaînes latérales de P.

Un total de 1700 pics initiaux correspondant à des potentiels NOE ont été introduits dans le logiciel CYANA comme source de contraintes de distances (1158 provenant de la NOESY HSQC <sup>1</sup>H-<sup>13</sup>C et 542 de la NOESY HSQC <sup>1</sup>H-<sup>15</sup>N). Le logiciel ayant des difficultés à converger avec le jeu de pics utilisé, j'ai retiré 518 pics du jeu de données dans CYANA (382 de la NOESY HSQC <sup>1</sup>H-<sup>13</sup>C et 136 de la NOESY HSQC <sup>1</sup>H-<sup>15</sup>N) qui étaient positionnées soit dans le bruit des diagonales ou dans des zones avec beaucoup de superposition, ou avec des intensités/largeurs de raies de l'ordre du bruit, ou correspondant à des artefacts afin de faciliter la convergence des calculs. D'autre part, beaucoup de superpositions sont présentes dans les spectres et des contraintes pouvant correspondre à des contraintes inter protomères (visible notamment avec des NOE très intenses entre <sup>1</sup>H $\delta$  de leucines sur un même résidu ou H de résidu i avec H de résidu i+2/-2 par exemple) mélangées avec des contraintes intra protomère rendaient certainement la convergence du calcul pour obtenir une hélice  $\alpha$  droite assez difficile.

J'ai paramétré le calcul avec CYANA en utilisant une tolérance de 0,03 ppm pour les  $^1\text{H}$  des 2 dimensions d'acquisition directe et indirecte et 0,4 ppm pour les hétéronoyaux  $^{13}\text{C}$  et  $^{15}\text{N}$ . 100 structures ont été générées par itération et 20 au cycle final. Enfin j'ai utilisé la fonction de réattribution des pics NOE du logiciel. Les statistiques pour les différents cycles du calcul d'un protomère de  $P_{OD}$  avec CYANA sont données dans le **Tableau 9**. Le calcul génère bien une hélice  $\alpha$  qui est représentée dans la **Figure 58**.

Cycles	1	2	3	4	5	6	7	final
<b>Pics</b>								
<b>Sélectionnés</b>	1182	1182	1182	1182	1182	1182	1182	1182
<b>Attribués</b>	1112	1159	1132	1138	1135	1133	1135	
<b>Non attribués</b>	70	23	50	44	47	49	47	
<b>Avec attribution dans la diagonale</b>	0	0	0	0	0	0	0	
<b>Pics croisés</b>								
<b>Attribution hors diagonale</b>	1112	1159	1132	1138	1135	1133	1135	
<b>Avec attribution unique</b>	593	763	855	888	933	971	970	
<b>Avec attribution courte distance <math> i-j \leq 1</math></b>	862	871	854	853	847	845	846	
<b>Avec attribution distance moyenne <math>1&lt; i-j &lt;5</math></b>	250	288	278	285	288	288	289	
<b>Avec attribution longue distance <math> i-j \geq 5</math></b>	0	0	0	0	0	0	0	
<b>Comparaison avec les attributions de référence</b>								
<b>Pics croisés avec attribution de référence</b>	1180	1180	1180	1180	1180	1180	1180	
<b>Avec attribution identique à la référence</b>	560	743	814	836	862	881	884	
<b>Avec attribution compatible à la référence</b>	1054	1106	1065	1062	1042	1027	1030	
<b>Avec attribution incompatible à la référence</b>	58	51	66	75	92	105	104	
<b>Avec attribution en plus de la référence</b>	68	23	49	43	46	48	46	
<b>Avec attribution additionnelle</b>	0	2	1	1	1	1	1	
<b>Limites supérieures de distance</b>								
<b>Total</b>	672	665	613	605	598	592	589	586
<b>Courte distance <math> i-j \leq 1</math></b>	444	425	400	388	379	376	361	358
<b>Distance moyenne <math>1&lt; i-j &lt;5</math></b>	228	240	213	217	219	216	227	227
<b>Longue distance <math> i-j \geq 5</math></b>	0	0	0	0	0	0	1	1
<b>Nombre d'attribution moyenne/contrainte</b>	3.05	1.82	1.23	1.20	1.14	1.10	1.00	1.00
<b>Valeur moyenne de la fonction cible</b>	11.71	9.82	16.46	0.73	0.14	0.02	0.01	0.00
<b>RMSD (résidus 128-156)</b>								
<b>RMSD moyen du squelette peptidique (Å)</b>	0.72	0.40	0.16	0.29	0.22	0.28	0.25	0.18
<b>RMSD moyen des atomes lourds (Å)</b>	1.20	0.77	0.63	0.67	0.60	0.65	0.60	0.49

**Tableau 9** : Statistiques du calcul de la structure d'un protomère de  $P_{OD}$  avec le logiciel CYANA.



**Figure 58 : Résultat du calcul de structure de CYANA d'un protomère de P<sub>OD</sub>.** L'image représente l'ensemble des 20 conformères alignés en représentation cartoon. A gauche de l'image se situe l'extrémité N-terminale et à droite l'extrémité C-terminale.

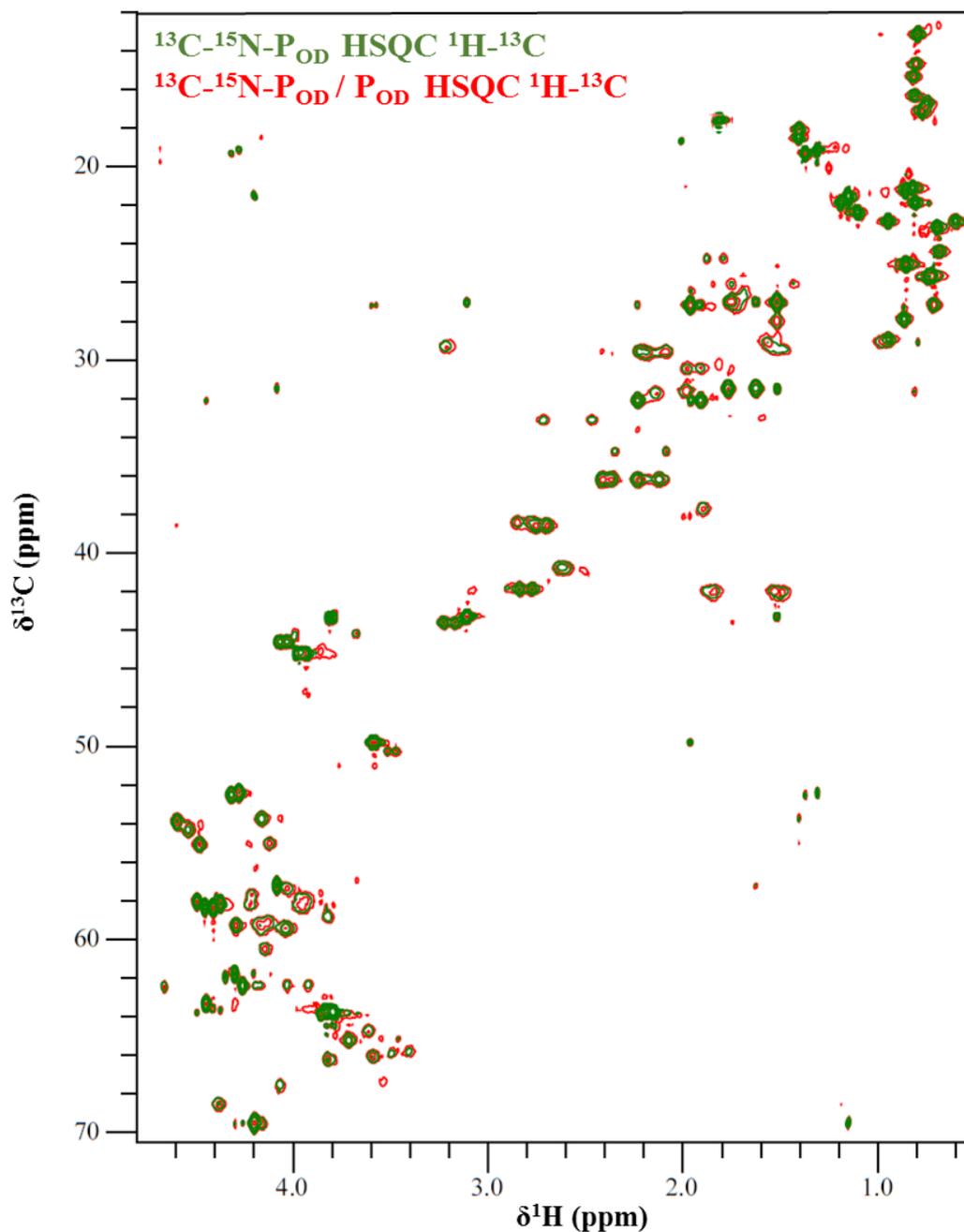
#### 4.4.3.2. Recherches de contraintes de distances intra- et inter-protomères

Calculer la structure d'une hélice  $\alpha$  isolée n'est pas très complexe, mais ici nous avons affaire à une région homotétramérique d'après les données publiées avec des expériences de pontage (Castagné et al., 2004). Le problème est que nous ne pouvons pas distinguer les NOE intra- et inter-protomères sur les spectres NOESY-HSQC. Il a donc fallu faire un mélange de protomères avec et sans marquage isotopique de P<sub>OD</sub> et enregistrer des expériences NOESY-HSQC  $^1\text{H}$ - $^{13}\text{C}$  filtrées et éditées pour « se débarrasser » des contraintes intra-protomères et ne garder que les contraintes inter-protomères (en ne gardant que les NOE entre  $^1\text{H}$  portés par des  $^{13}\text{C}$  et  $^1\text{H}$  portés par des  $^{12}\text{C}$ ). En comparant avec l'ensemble des NOE obtenus précédemment, on peut aussi espérer en déduire les contraintes intra-protomères dans les NOESY HSQC  $^1\text{H}$ - $^{13}\text{C}$  avec P<sub>OD</sub> complètement marquée  $^{13}\text{C}$ - $^{15}\text{N}$ .

Ainsi, j'ai produit et purifiée deux échantillons de P<sub>OD</sub> : le premier marqué  $^{15}\text{N}$ - $^{13}\text{C}$  et le deuxième sans marquage. J'ai mélangé les deux échantillons dans les proportions 50/50 de P<sub>OD</sub> marquée (P<sub>OD</sub><sup>\*</sup>) et non marquée (P<sub>OD</sub>). Puis, j'ai dénaturé les protéines à l'urée (8 M) et au chlorure de guanidinium (6 M), et enfin j'ai renaturé l'ensemble en enlevant les dénaturants par dialyse. J'ai terminé en échangeant la solution de P<sub>OD</sub> « mixte » dans le même tampon D<sub>2</sub>O que celui utilisé pour les expériences NOESY HSQC  $^1\text{H}$ - $^{13}\text{C}$  non filtrées. La solution finale était à 600  $\mu\text{M}$ .

L'hypothèse est que le domaine forme une structure en hélice surenroulée, comme celle prédite par Llorente et al., 2008, en particulier au niveau des résidus 119-160 correspondant au fragment Y\* dans cette publication. Les proportions statistiques attendues de répartition des deux formes P<sub>OD</sub><sup>\*</sup> et P<sub>OD</sub> dans des tétramères sont : 1/16<sup>e</sup> de (P<sub>OD</sub>)<sub>4</sub>, 1/16<sup>e</sup> de (P<sub>OD</sub><sup>\*</sup>)<sub>4</sub>, 4/16<sup>e</sup> de (P<sub>OD</sub>)<sub>1</sub>(P<sub>OD</sub><sup>\*</sup>)<sub>3</sub>, 4/16<sup>e</sup> de (P<sub>OD</sub>)<sub>3</sub>(P<sub>OD</sub><sup>\*</sup>)<sub>1</sub>, 2/16<sup>e</sup> de (P<sub>OD</sub>)(P<sub>OD</sub><sup>\*</sup>)(P<sub>OD</sub>)(P<sub>OD</sub><sup>\*</sup>) et 4/16<sup>e</sup> de (P<sub>OD</sub>)<sub>2</sub>(P<sub>OD</sub><sup>\*</sup>)<sub>2</sub>.

Les spectres HSQC  $^1\text{H}$ - $^{13}\text{C}$  de la protéine entièrement marquée et du tétramère mixte se superposent très bien (**Figure 59**). Ceci montre non seulement que la dénaturation/renaturation de  $\text{P}_{\text{OD}}$  a bien fonctionnée, mais aussi que le spectre de  $\text{P}_{\text{OD}}$  partiellement marquée peut être comparé à celui de  $\text{P}_{\text{OD}}$  complètement marquée pour faire un report d'attribution. J'ai reporté les attributions du spectre HSQC  $^1\text{H}$ - $^{13}\text{C}$  de  $\text{P}_{\text{OD}}$  entièrement marquée sur le spectre obtenu pour le tétramère mixte de  $\text{P}_{\text{OD}}$  et j'ai recentré légèrement les pics quand cela était nécessaire.



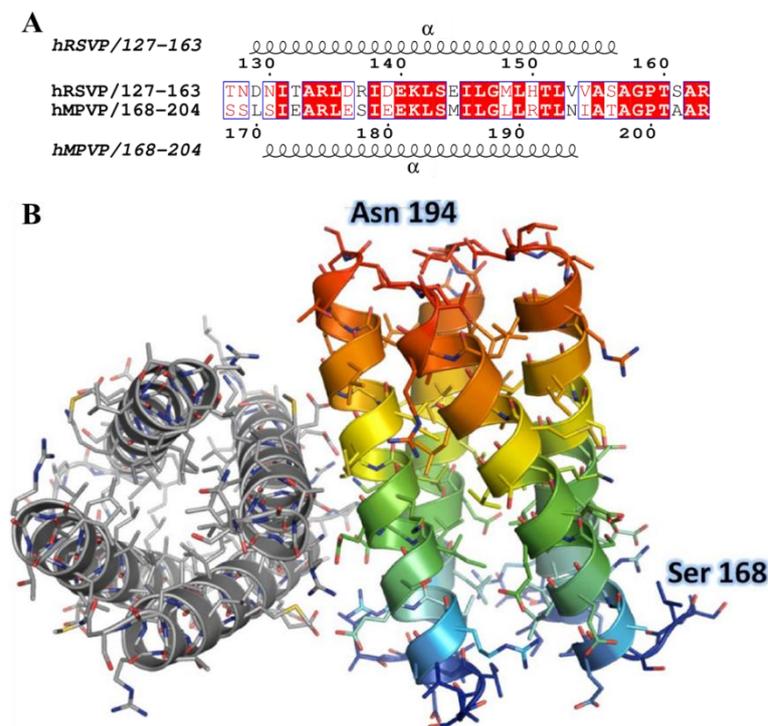
**Figure 59** : Superposition de spectres HSQC  $^1\text{H}$ - $^{13}\text{C}$  de tétramères entièrement marqués  $^{13}\text{C}^{15}\text{N}$  et de tétramères mixtes obtenus avec  $\text{P}_{\text{OD}}^*$  et  $\text{P}_{\text{OD}}$ . Les spectres ont été acquis dans un tampon 20 mM phosphate, 100 mM NaCl dans  $\text{D}_2\text{O}$  avec  $\text{P}_{\text{OD}}/\text{P}_{\text{OD}}^*$  concentré à 600  $\mu\text{M}$  à un champ magnétique de 22,3 T et à une température de 298 K.

#### 4.4.3.3. Calcul de structure du tétramère de P<sub>OD</sub>

Comme pour le calcul de protomère avec CYANA, j'ai utilisé la séquence de P<sub>OD</sub> sans prendre en compte 6 des 8 acides aminés restant après la coupure de l'étiquette GST en N-terminal (GSGSGS), prédits comme étant flexibles.

311 pics NOE ont été piqués sur la NOESY-HSQC <sup>1</sup>H-<sup>13</sup>C éditée et filtrée acquise sur le tétramère mixte. Les autres 1700 NOE avec P<sub>OD</sub> complètement marquée <sup>13</sup>C-<sup>15</sup>N ont été gardés pour le calcul.

J'ai cette fois utilisé le logiciel ARIA (Rieping et al., 2007) permettant de calculer des structures de multimères de protéines. En plus du fait que la protéine P est tétramérique et que dans le tétramère les protomères sont tous les 4 orientés dans le même sens (d'après les données de PRE, notamment sur les mutants P(S23C) et P(S237C) de la **Figure 54**), nous avons également supposé une symétrie de type C4 pour P<sub>OD</sub>. Cette hypothèse a été fondée en s'appuyant sur la forte identité de séquence avec le domaine d'oligomérisation de P du hMPV (~59,5% d'identité de séquence au niveau de la région 127-163 de hRSV-P avec la région correspondante de hMPV-P) qui a une structure avec une symétrie de type C4 (Leyrat et al., 2013 ; **Figure 60**). La structure publiée pour hMPV a été obtenue par cristallographie.



**Figure 60 : Structure du domaine d'oligomérisation de hMPV-P.** A, Alignement des séquences primaires de hRSV-P avec celle de hMPV-P à partir du logiciel Clustal Omega et préparé avec ESPRIPT 3.0. La position de l'hélice  $\alpha$  prédite par TALOS-N pour hRSV-P<sub>OD</sub> et la position de l'hélice  $\alpha$  pour le domaine de tétramérisation de hMPV-P sont indiquées en haut et en bas respectivement. B, La structure du domaine d'oligomérisation de hMPV-P obtenue par diffraction des rayons X est

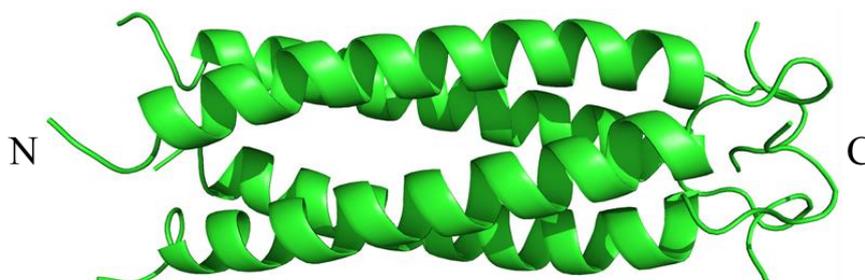
représentée avec vue du côté C-terminal en gris et avec vue latérale des faisceaux d'hélices en couleur arc-en-ciel, avec les extrémités C- et N-terminales en rouge et bleu respectivement (figure issue des travaux de Leyrat et al., 2013).

J'ai paramétré la version 2.3.2 de ARIA avec les 3 expériences NOESY-HSQC en utilisant :

- la fonction *use manual assignment* permettant d'utiliser les attributions faites « à la main » durant la première itération sans réattribuer les NOE (le calcul ne convergeait pas en réattribuant tout automatiquement) sans pour autant en forcer l'utilisation en tant que contraintes de distance ;
- la fonction *filter diagonal peaks* qui élimine les pics proches de la diagonale ;
- la fonction *only fully assigned peaks* ne gardant que les pics pré attribués ;
- les intensités et non les volumes des pics pour calculer les contraintes de distance.

Pour les pics issus des NOESY-HSQC avec P<sub>OD</sub> entièrement marquée <sup>13</sup>C-<sup>15</sup>N, dans le cadre de l'attribution automatique par ARIA, j'ai utilisé une tolérance <sup>1</sup>H de 0,03 ppm et de 0,3 ppm pour les hétéronoyaux tout en ajoutant, s'il y en a, les erreurs de déplacement chimique expérimentales (fonction *use experimental shift error*). De plus, j'ai défini ces 2 NOESY-HSQC comme fournissant des contraintes intra protomères. Pour les pics NOE de la NOESY-HSQC <sup>1</sup>H-<sup>13</sup>C éditée/filtrée, j'ai pris des valeurs plus stringentes de tolérance qui sont de 0,02 ppm pour les <sup>1</sup>H et 0,2 ppm pour les <sup>13</sup>C sans utiliser l'erreur de déplacement chimique expérimentale dans le but d'éviter des contraintes longues distances aberrantes. Pour les pics de cette dernière expérience, j'ai défini les contraintes comme inter protomère uniquement et j'ai également activé la fonction *enable structural rules* qui, si elle trouve des attributions de NOE à plus de 5 résidus de distance au sein des hélices  $\alpha$ , définit cette contrainte comme étant uniquement inter-protomères.

Du point de vue des autres paramètres, j'ai activé les symétries multimériques de type C4 et j'ai activé les contraintes de *packing* pour maintenir les protomères proches les uns des autres. J'ai utilisé comme structure initiale pour le recuit simulé la structure d'un tétramère calculé par ARIA et ayant convergé (**Figure 61**), mais avec seulement une partie des données NOE (NOESY-HSQC <sup>1</sup>H-<sup>15</sup>N absente et jeu de données NOESY-HSQC <sup>1</sup>H-<sup>13</sup>C non filtré/édité avec 382 pics retirés).



**Figure 61** : Structure obtenue par ARIA avec un jeu de données partiel et utilisé comme structure initiale pour un deuxième recuit simulé du tétramère de P<sub>OD</sub>. Les extrémités N- et C-terminales sont indiquées de part et d'autre de la représentation cartoon du tétramère.

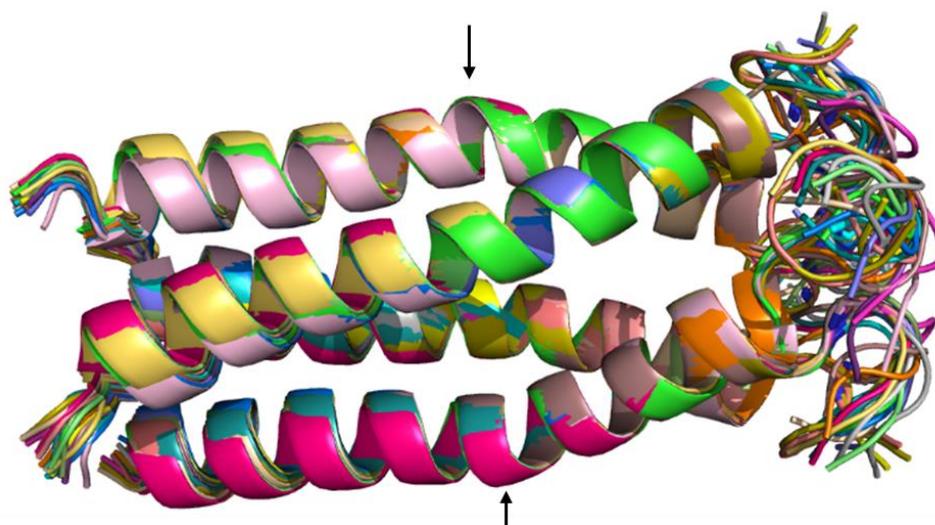
En effet, j'ai essayé de repartir de la structure étendue de tétramère de P<sub>OD</sub> générée par ARIA, mais sans succès concernant la convergence du calcul (distances d'une à plusieurs dizaines d'Å pour les NOE inter protomères selon ARIA et par conséquent topologies aberrantes de tétramères en résultant). De plus, j'ai utilisé 9 itérations au lieu de 8. A chaque itération, 50 structures (100 pour la 9<sup>ème</sup>) ont été générées et les 15 (20 pour la 9<sup>ème</sup>) meilleures (en termes du total d'énergie) ont été utilisées pour l'itération suivante. J'ai réduit légèrement le paramètre *ambiguity cutoff / weight\_threshold* entre les itérations 5 à 8 (0,95 ; 0,9 ; 0,8 ; 0,7 ; 0,7 ; dans l'ordre des itérations 5 à 9). Enfin, j'ai multiplié par 2 le facteur *TAD high temperature* (= 20 000) et j'ai augmenté à 30 000 les paramètres *Cool1/2 steps* pour éviter au calcul de « crasher ».

Rapport par itération	0	1	2	3	4	5	6	7	8	9	Affinement
<b>CONTRAINTES</b>											
Utilisées pendant le calcul	1540	1539	1537	1466	1443	1379	1169	1184	1185	1091	
Violées	0	0	1	28	22	30	182	126	97	88	
Convergées	467	467	466	448	442	422	372	375	377	376	
<b>NOESY-HSQC <sup>1</sup>H-<sup>13</sup>C éditée/filtrée (inter protomères)</b>											
Utilisées pendant le calcul	311	311	311	311	311	311	311	308	302	205	
Violées	0	0	0	0	0	0	0	0	0	2	
Non ambiguës	0	0	0	0	0	0	0	0	99	149	
Violées	0	0	0	0	0	0	0	0	0	1	
Ambiguës	311	311	311	311	311	311	311	308	203	56	
Violées	0	0	0	0	0	0	0	0	0	1	
Histogramme de contributions	2	0	0	0	0	23	297	306	202	56	
	3	310	310	310	310	287	13	2	1	0	
	4	0	0	0	0	0	1	0	0	0	
	5	0	0	0	0	0	1	0	0	0	
	6-10	1	1	1	1	1	0	0	0	0	
	>10	0	0	0	0	0	0	0	0	0	
<b>NOESY-HSQC <sup>1</sup>H-<sup>15</sup>N (intra protomères)</b>											
Utilisées pendant le calcul	494	494	494	469	460	437	354	361	365	366	
Violées	0	0	0	9	4	5	66	37	23	18	
Non ambiguës	494	494	494	469	460	437	354	361	365	366	
Violées	0	0	0	9	4	5	66	37	23	18	
<b>NOESY-HSQC <sup>1</sup>H-<sup>13</sup>C (intra protomères)</b>											
Utilisées pendant le calcul	735	734	732	686	672	631	504	515	518	520	
Violées	0	0	1	19	18	25	116	89	74	68	
Non ambiguës	735	734	732	686	672	631	504	515	518	520	
Violées	0	0	1	19	18	25	116	89	74	68	
<b>RMSD de l'ensemble (15 ou 20 structures) en Å</b>											
<b>Tous les résidus</b>											
Squelette (N,CA,C,O)	5,63	6,55	6,97	6,97	7,47	6,39	7,5	5,44	2,22	1,15	1,15
Atomes lourds	6,09	6,89	7,31	7,32	7,85	6,74	7,81	5,89	2,6	1,48	1,48
<b>Résidus structurés (128-156 ; 159 généralement) en Å</b>											
Squelette (N,CA,C,O)	5,56	6,16	6,79	6,74	7,25	5,82	7,04	5,11	1,76	0,53	0,50
Atomes lourds	6,09	6,38	7,04	6,99	7,57	6,1	7,31	5,49	2,01	0,63	0,60

Tableau 10 : Rapport des statistiques du calcul de structure du tétramère de hRSV-P<sub>OD</sub> avec le logiciel ARIA.

Le **Tableau 10** donne les statistiques par itérations du calcul de structure du tétramère de P<sub>OD</sub> avec ARIA. On voit qu'à la dernière itération environ 54% des pics totaux ont été gardés comme contraintes de distance mais aussi que le RMSD des parties structurées de P<sub>OD</sub> est inférieur à 1 Å et à 1,5 Å pour tous les résidus ce qui montre une convergence correcte du calcul.

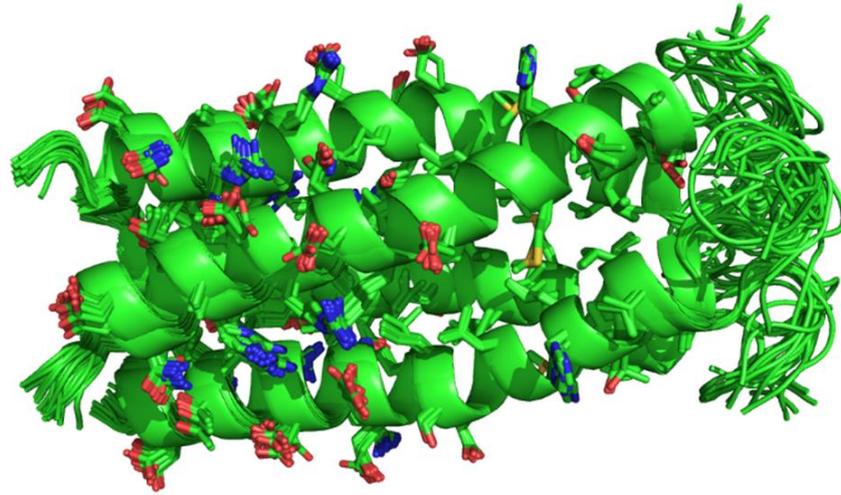
Le résultat du calcul du tétramère est montré sur la **Figure 62**. L'image montre que la symétrie C4 a correctement fonctionné avec 4 protomères presque identiques et orientés à 90° d'un protomère *i* à *i*+1, arrangés en faisceau d'hélices. Mais une contorsion des hélices est observée autour du résidu E144. L'origine peut éventuellement se trouver dans les données brutes au niveau des NOE mais *a priori* il n'y a pas de raison que cela soit une déformation « naturelle » de l'hélice.



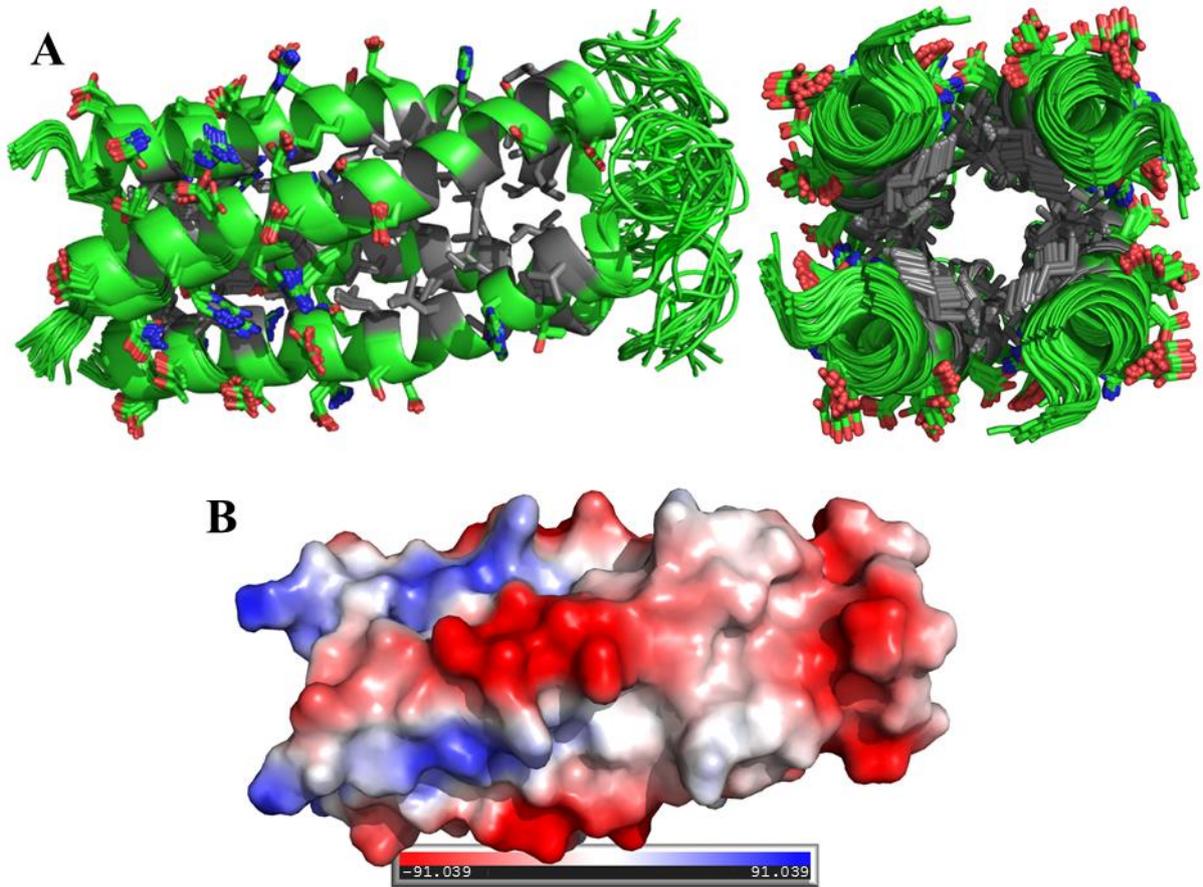
**Figure 62 :** Résultat du calcul de structure de ARIA du tétramère de P<sub>OD</sub>. L'image représente l'ensemble des 20 conformères alignés en représentation « cartoon ». A gauche de l'image se situe l'extrémité N-terminale et à droite l'extrémité C-terminale. Les flèches noires indiquent une contorsion locale des hélices autour du résidu E144.

Les chaînes latérales des résidus structurés en hélice  $\alpha$  (129-156) sont définies avec relativement peu de fluctuation spatiale que ce soit en surface de la super hélice ou au sein du cœur de P<sub>OD</sub> (**Figure 63**).

Sur la **Figure 64**, on distingue bien le cœur hydrophobe du tétramère et les résidus plus hydrophiles exposés en surface de la structure. Le cœur est composé de 5 leucines et 3 isoleucines. On trouve également des interfaces hydrophobes entre protomères côte à côte notamment avec la méthionine 148, et 2 valines du côté C-terminal.



**Figure 63 :** Superpositions des chaînes latérales du tétramère de  $P_{OD}$ . Seules les chaînes latérales des parties structurées en hélice (P[129-156]) sont indiquées.



**Figure 64 :** Représentation du cœur hydrophobe du tétramère de  $P_{OD}$ . A, Mise en valeur des résidus hydrophobes en gris. B, Représentation sous forme de surface électrostatique d'un conformère du tétramère avec en rouge les régions chargées négativement, en blanc les zones hydrophobes et en bleu les régions chargées positivement.

#### 4.4.3.4. Second calcul de structure du tétramère de P<sub>OD</sub>

Etant donné que le calcul de structure du tétramère de P<sub>OD</sub> a été réalisé en utilisant comme point de départ pour le recuit simulé une structure de tétramère en hélice super enroulée, j'ai souhaité relancer le calcul en utilisant des protomères de P<sub>OD</sub> totalement étendus (générés par CNS) pour éviter un biais quelconque.

Avant tout, j'ai retiré un grand nombre de pics sur les spectres qui correspondaient soit à du bruit, ou à des pics excentrés. Ainsi, il restait 275 pics NOE sur la NOESY-HSQC <sup>1</sup>H-<sup>13</sup>C éditée et filtrée, 356 et 624 pics gardés pour la NOESY-HSQC <sup>1</sup>H-<sup>15</sup>N et la NOESY-HSQC <sup>1</sup>H-<sup>13</sup>C respectivement.

Toujours avec le logiciel ARIA, j'ai réalisé le calcul de structure du tétramère de P<sub>OD</sub> en réutilisant :

- la fonction *use manual assignment* ;
- la fonction *filter diagonal peaks* ;
- la fonction *only fully assigned peaks* ;
- la symétrie C4.

Les pics de la NOESY-HSQC <sup>1</sup>H-<sup>13</sup>C éditée et filtrée ont été déclaré comme étant des contraintes de distance inter protomères, ceux de la NOESY-HSQC <sup>1</sup>H-<sup>15</sup>N ont été déclarés comme étant intra protomères. 114 pics de la NOESY-HSQC <sup>1</sup>H-<sup>13</sup>C sur 624 ont été définis comme inter protomères (après analyse des données en comparant les pics avec ceux de la NOESY-HSQC <sup>1</sup>H-<sup>13</sup>C éditée / filtrée) et les 510 pics restants ont été définis comme intra protomères.

Pour la réattribution automatique des pics par ARIA, j'ai défini une tolérance <sup>1</sup>H de 0,03 ppm et de 0,3 ppm pour les hétéronoyaux.

Cette fois j'ai utilisé les 8 itérations (au lieu de 9) avec les paramètres donnés par défaut dans ARIA, 50 structures ont été générées et les 15 meilleures (en termes du total d'énergie) ont été utilisées pour l'itération suivante. Enfin, j'ai gardé la valeur de *TAD high temperature* à 20 000 et j'ai modifié les paramètres de recuit simulé suivants :

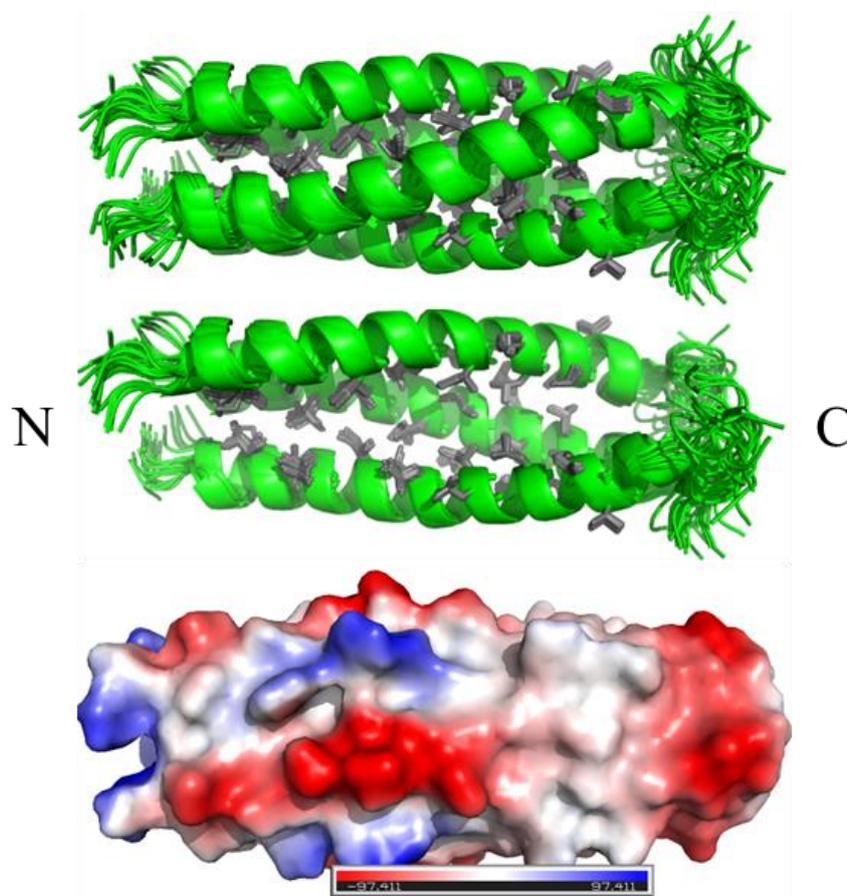
- *steps\_high* = 15 000 ;
- *steps\_refine* = 8 000 ;
- *steps\_cool1* = 30 000 ;
- *steps\_cool2* = 30 000.

Rapport par itération	0	1	2	3	4	5	6	7	8	Raffinement
<b>CONTRAINTES</b>										
Utilisées pendant le calcul	1012	1010	1006	958	950	949	855	837	826	
Violées	0	0	4	10	5	1	43	16	18	
Convergées	243	243	238	221	219	217	193	181	181	
<b>NOESY-HSQC <sup>1</sup>H-<sup>13</sup>C éditée/filtrée (inter protomères)</b>										
Utilisées pendant le calcul	242	201	336	327	323	323	282	285	287	
Violées	0	0	0	4	1	1	14	5	5	
Non ambiguës	0	0	336	327	323	323	282	285	287	
Violées	0	0	0	4	1	1	14	5	5	
Ambiguës	242	201	0	0	0	0	0	0	0	
Violées	0	0	0	0	0	0	0	0	0	
Histogramme de contributions	3	241	200	0	0	0	0	0	0	0
	6-10	1	1	0	0	0	0	0	0	0
<b>NOESY-HSQC <sup>1</sup>H-<sup>15</sup>N (intra protomères)</b>										
Utilisées pendant le calcul	336	473	206	211	211	211	217	209	196	
Violées	0	0	1	2	0	0	5	2	1	
Non ambiguës	336	361	0	0	0	0	0	0	43	
Violées	0	0	0	0	0	0	0	0	0	
Ambiguës	0	112	206	211	211	211	217	209	153	
Violées	0	0	1	2	0	0	5	2	1	
Histogramme de contributions	2	0	0	0	0	0	0	12	152	
	3	0	112	205	210	210	216	196	0	
	4	0	0	0	0	0	0	0	1	
	5	0	0	0	0	0	0	0	1	0
	6-10	0	0	1	1	1	1	1	0	0
<b>NOESY-HSQC <sup>1</sup>H-<sup>13</sup>C (intra protomères)</b>										
Utilisées pendant le calcul	434	336	464	420	416	415	356	343	343	
Violées	0	0	3	4	4	0	24	9	12	
Non ambiguës	362	336	361	345	344	343	299	287	324	
Violées	0	0	0	1	3	0	20	6	11	
Ambiguës	72	0	103	75	72	72	57	56	19	
Violées	0	0	3	3	1	0	4	3	1	
Histogramme de contributions	2	0	0	0	0	0	0	13	19	
	3	72	0	103	75	72	72	57	43	0
<b>RMSD de l'ensemble (15 ou 20 structures)</b>										
<b>Tous les résidus</b>										
Squelette (N,CA,C,O)	5,33	5,36	5,51	5,59	5,66	5,97	5,66	4,86	1,69	1,39
Atomes lourds	5,79	5,83	5,94	6,1	6,18	6,48	6,14	5,28	2,06	1,72
<b>Residus structurés (129-157 généralement)</b>										
Squelette (N,CA,C,O)	5,64	5,49	5,6	5,5	5,61	5,89	5,63	4,69	0,7	0,44
Atomes lourds	6,11	6,04	6,06	5,96	6,01	6,39	6,07	5,07	1,01	0,82

**Tableau 11 : Rapport des statistiques du second calcul de structure du tétramère de hRSV-P<sub>OD</sub> avec le logiciel ARIA.**

Le **Tableau 11** fournit les statistiques par itérations du calcul de structure du tétramère de P<sub>OD</sub> avec ARIA. Même si un grand nombre de pics ont été retirés des données brutes, on voit qu'ARIA a gardé plus de 80% des pics comme contraintes de distance à la dernière itération mais aussi que le

RMSD des parties structurées de P<sub>OD</sub> reste inférieur à 1 Å et à 1,8 Å pour tous les résidus confondus ce qui montre une convergence correcte du calcul.



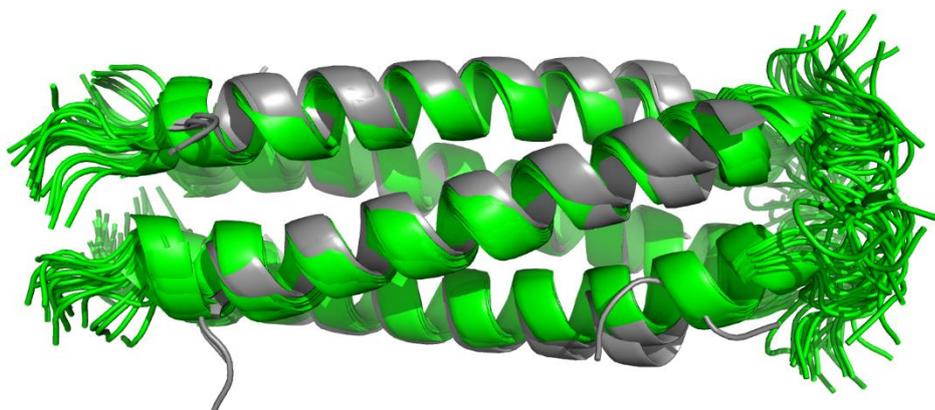
**Figure 65 :** Résultat du second calcul de structure de ARIA du tétramère de POD. Les 20 conformères alignés avec les chaînes latérales hydrophobes représentées en gris sont représentés sur les deux figures du haut. La seconde représentation de structure de P<sub>OD</sub> est montrée sans le protomère frontal qui était présent dans la structure du haut afin de visualiser le cœur hydrophobe. La structure au bas de la figure montre un tétramère sous forme de surface électrostatique.

Les résultats du second calcul de structure de P<sub>OD</sub> confirme la convergence de la structure du tétramère au niveau du squelette peptique et des chaînes latérales hydrophobes (**Figure 65**). D'autre part, on ne distingue pas de cassure dans les hélices comme c'était le cas dans le précédent calcul avec ARIA. Enfin, on retrouve à peu près la même dispersion des charges sur la surface électrostatique que lors du calcul précédent de la structure du tétramère de P<sub>OD</sub>.

L'ensemble de ces éléments montrent qu'en partant d'une structure étendue des protomères de P<sub>OD</sub> pour les recuits simulés utilisés pour le calcul de la structure du tétramère on obtient des résultats satisfaisants en termes de convergence des structures de conformères et on a même une disparition de la distorsion d'hélice  $\alpha$  au niveau du résidu 144 présente dans le calcul précédent.

#### 4.4.3.5. Comparaison de la structure de P<sub>OD</sub> du RSV avec celle de hMPV

En 2017, plusieurs structures du domaine d'oligomérisation du métapneumovirus humain (hMPV) plus résolues par cristallographie ont été publiées (Renner et al., 2017). J'ai superposé les unités asymétriques des structures publiées (codes pdb : 5OIX et 5OIIY) entre elles et j'ai ensuite superposé l'ensemble des conformères de hRSV-POD (**Figure 66**). On voit que les tétramères d'hélices se superposent bien (RMSD ~ 1 Å). Cependant, on observe que les hélices du tétramère de hMPV-P<sub>core</sub> sont légèrement raccourcies (malgré la forte homologie de séquence avec hRSV-POD) et qu'une légère distorsion des hélices de hRSV-POD cause un éloignement léger de l'extrémité C-terminale des hélices de hMPV-P<sub>core</sub>.



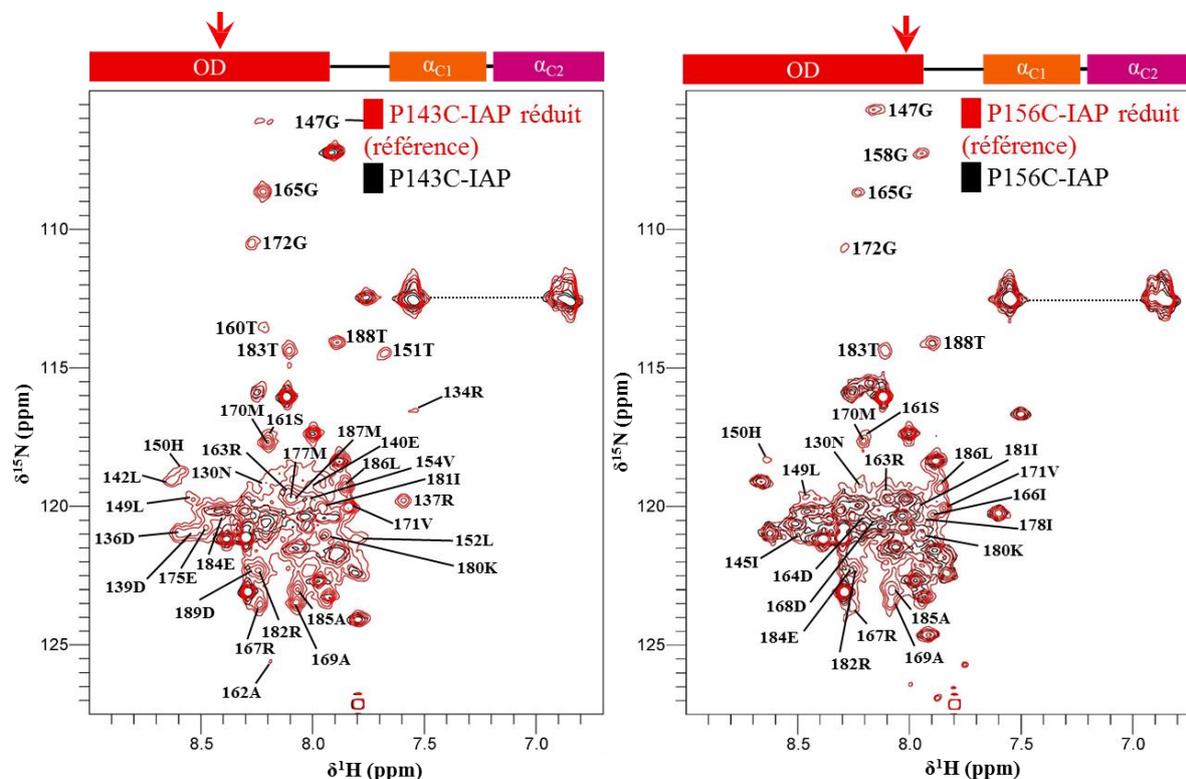
**Figure 66 :** *Superposition de la structure du tétramère de P<sub>OD</sub> du hRSV avec celle du hMPV. Les structures en vert correspondent à celles des conformères de hRSV-P<sub>OD</sub> calculés avec ARIA (2<sup>ème</sup> calcul) et en gris sont montrés les tétramères issus des structures de hMPV-P<sub>core</sub> (codes pdb : 5OIX et 5OIIY).*

#### 4.4.4. Recherche de contacts à longue distance au sein du fragment P[127-205]

Dans la continuité de mes travaux sur la recherche de contacts à longue distance au sein de la phosphoprotéine entière du hRSV (Pereira et al., 2016), j'ai entrepris de travailler sur la région P[127-205], invisible dans la protéine P entière, afin d'y rechercher des contacts à longue distance.

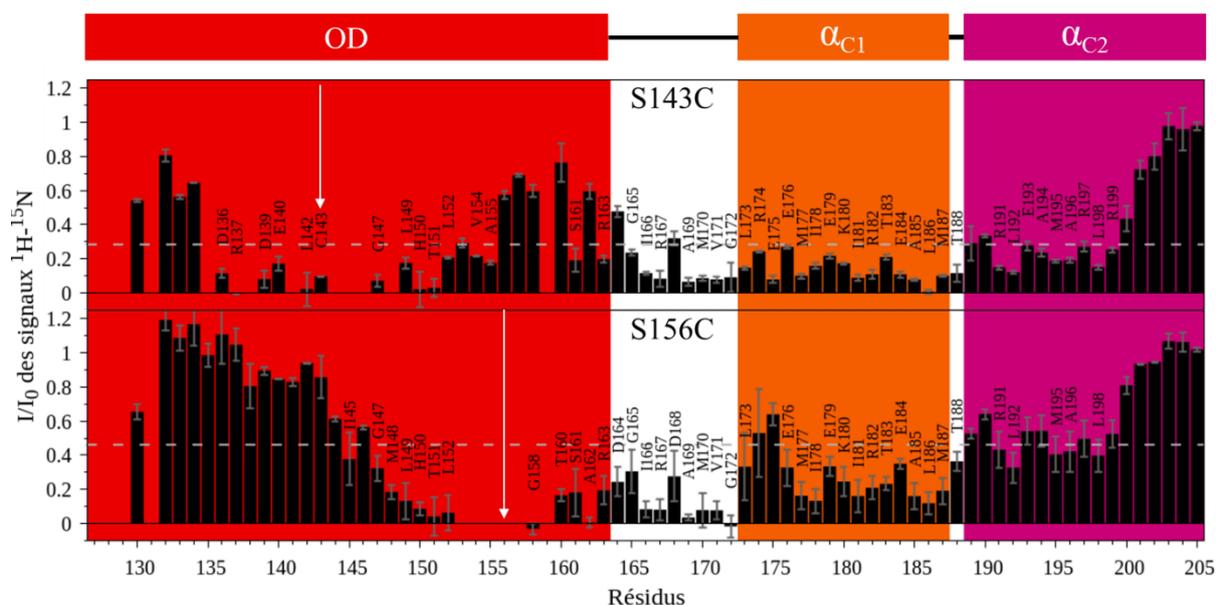
Pour ce faire, j'ai à nouveau utilisé des expériences de PRE sur deux mutants du peptide P[127-205], à savoir P[127-205](S143C) et P[127-205](S156C), en gardant le même protocole de production et de marquage paramagnétique que précédemment pour P pleine longueur. Les deux mutations sur P[127-205] sont toutes deux présentes dans le domaine d'oligomérisation de P. Grâce à la structure de P<sub>OD</sub> on peut dire que les chaînes latérales des résidus 143 et 156 sont exposées vers l'extérieur de l'hélice sur-enroulée et que la structure ne devrait pas être affectée par la mutation et l'ajout d'une sonde paramagnétique IAP. J'ai gardé la température des expériences à 313 K comme pour P[127-205] sauvage et j'ai travaillé sur un spectromètre à 600 MHz avec cryosonde. Enfin, j'ai utilisé des échantillons concentrés à 200 µM.

Les spectres HSQC  $^1\text{H}$ - $^{15}\text{N}$  des expériences PRE sur les mutants de P[127-205] (**Figure 67**) montrent tout d'abord qu'il n'y a pas de différences majeures dans les déplacements chimiques en comparaison avec le spectre de P[127-205] sauvage en-dehors des zones de mutation (**Figure 55**). J'en conclus que les mutations n'ont pas induit de différence majeure de structure.



**Figure 67 : Spectres HSQC  $^1\text{H}$ - $^{15}\text{N}$  des expériences de PRE avec les mutants cystéines de P[127-205].** Les PRE ont été mesurés après fixation de sondes radicalaires IAP aux positions 143 ou 156 sur les mutants cystéines de P[127-205]. Les expériences ont été réalisées sur les mutants P[127-205](S143C) et P[127-205](S156C) concentrés à 200  $\mu\text{M}$  dans un tampon 20 mM phosphate pH 6,5, 100 mM NaCl, 7,5%  $\text{D}_2\text{O}$  à 313 K et enregistrées sur un spectromètre RMN à 14,1 T avec cryosonde. Les positions de marquage sont indiquées par des flèches rouges sur les schémas de structures secondaires au-dessus de chaque spectre

Pour les deux mutants, une importante diminution des intensités dans les régions où la sonde paramagnétique est présente confirme la bonne fixation de cette dernière (**Figure 68**). On constate également la présence d'une perte d'intensité des signaux présents dans les deux régions  $\alpha_{\text{C1}}$  et  $\alpha_{\text{C2}}$  ainsi que dans la région charnière entre les résidus 164 et 172 inclus. Cela montre qu'en absence des domaines flanquants de P il y a certainement des contacts entre le domaine d'oligomérisation et les régions  $\alpha_{\text{C1}}$  et  $\alpha_{\text{C2}}$ . On remarque également que l'effet paramagnétique est d'autant plus fort quand on éloigne la sonde paramagnétique de la région  $\alpha_{\text{C1/2}}$ . Cela prouve d'autant plus qu'il y a un contact qui se fait à une grande distance qui est d'environ 50 résidus.

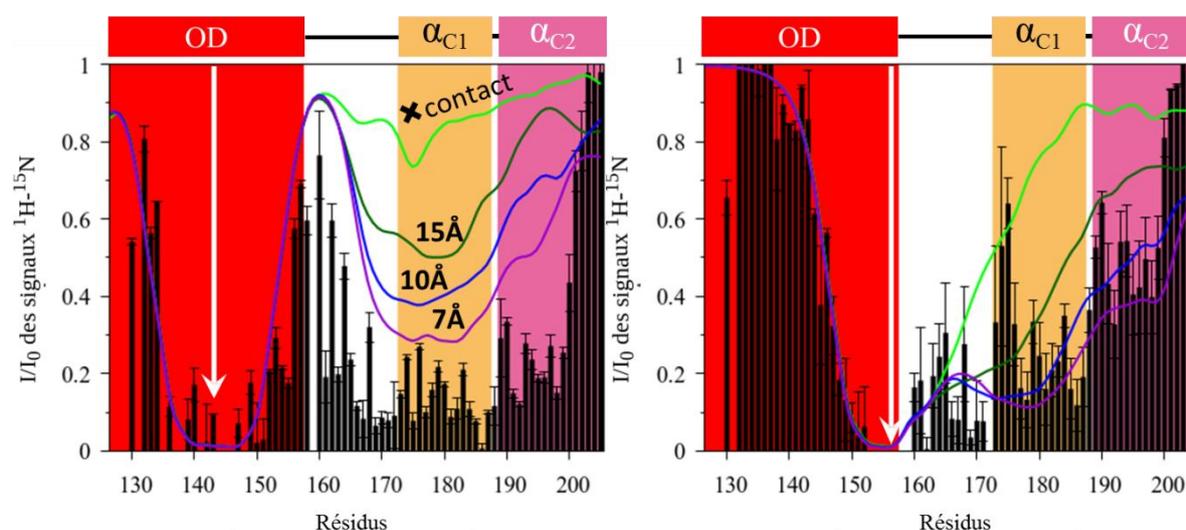


**Figure 68 :** Histogrammes des intensités normalisées pour les signaux des mutants de P[127-205] comportant une sonde paramagnétique. Les fonds colorés correspondent aux domaines et régions de la protéine signalés au-dessus des histogrammes. La flèche blanche indique la position de marquage par la sonde paramagnétique. La ligne pointillée correspond à la moyenne de toutes les intensités. Les barres d'erreur pour chaque intensité correspondent à l'écart type des ratios d'intensité en utilisant des réplicats des expériences.

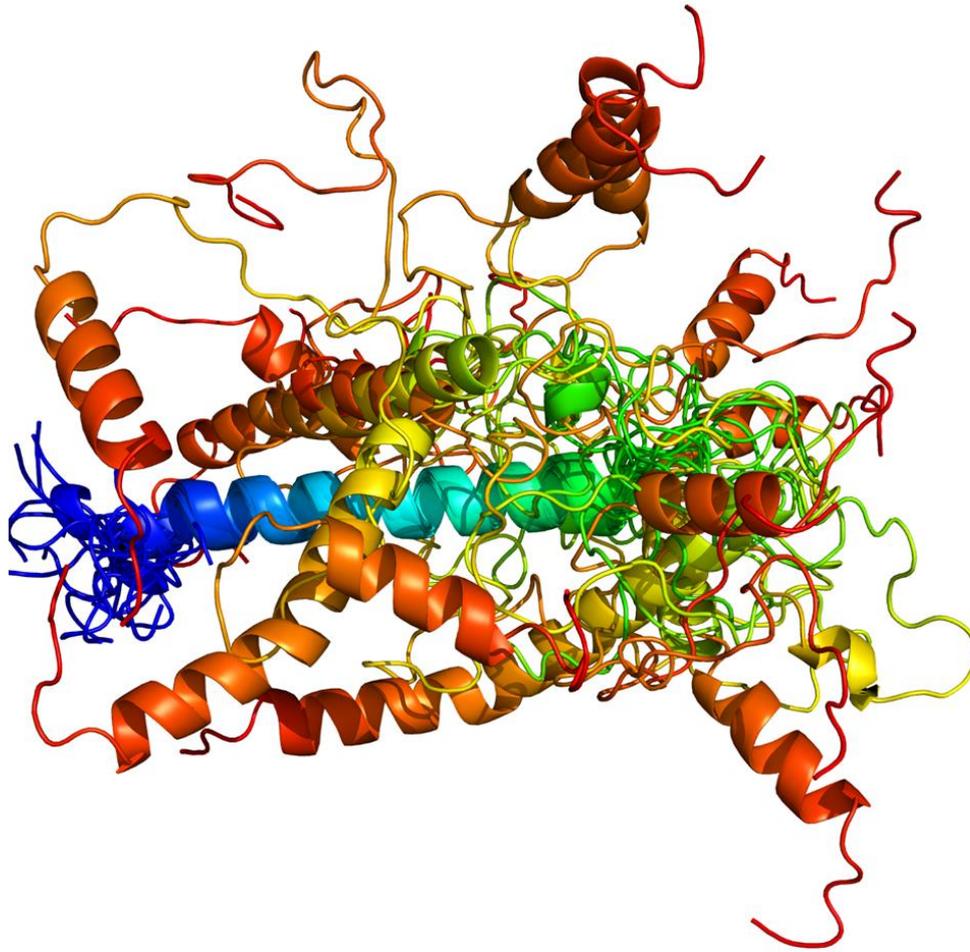
En vue d'en savoir plus sur la nature de ce contact entre la région  $\alpha_{C1/2}$  et  $P_{OD}$ , j'ai opté pour l'utilisation du logiciel Flexible-meccano (Ozenne et al., 2012) afin de rechercher un bon modèle s'ajustant à mes données PRE avec P[127-205]. Le logiciel permet en outre de rentrer la séquence d'une protéine en précisant éventuellement des structures secondaires, leurs propensions par acide aminé, les zones de contacts suspectées (par groupe de résidus) et de calculer un nombre N de conformères flexibles à partir des données rentrées. Après génération des conformères, le logiciel, selon le choix de l'utilisateur, est, entre autre, en capacité de calculer la moyenne des couplages dipolaires résiduels, des valeurs PRE par résidu pour plusieurs mutants indiqués par l'utilisateur et des constantes de couplages.

Pour cette étude, j'ai utilisé la séquence primaire de la protéine sauvage P[127-205], j'ai défini une hélice alpha permanente pour les résidus 130 à 156 et j'ai utilisé les valeurs de propension d'hélice  $\alpha$  par résidu ayant été déterminés par TALOS-N au niveau des résidus 167-187 et 189-198 (**Figure 57**). En revanche je n'ai pas tenu compte de la nature tétramérique de la protéine. Concernant la région de contact, j'ai défini la région des résidus 143 à 156 (où sont placées les mutations) en contact à différentes distances (15, 10 et 7 Å) de distance avec la région 160-199 en essayant de m'appuyer sur les baisses d'intensité des deux diagrammes de la **Figure 68**. Enfin, avec le logiciel Flexible-meccano j'ai lancé le calcul de PRE en signalant les 2 positions de mutations comme dans mes données expérimentales, le tout calculé à partir de 1000 conformères générés par simulation.

L'ajustement des données expérimentales de PRE de P[127-205] par le modèle de Flexible-meccano avec un contact de 7 Å (**Figure 69**) semble le plus optimal de par son allure générale et son relativement faible écart entre les valeurs expérimentales et les valeurs théoriques ici calculées. En conclusion et en dépit de la variabilité expérimentale, on peut raisonnablement définir un contact de l'ordre de 7 Å entre la région 143-156 et la région 160-199. Ce contact indique un rapprochement entre le domaine P<sub>OD</sub> et la région  $\alpha_{C1/2}$ . Ces contacts transitoires entre régions structurées pourraient expliquer en partie pourquoi on ne distingue pas la région  $\alpha_{C1/2}$  dans le cadre de P entière. Un échantillon de 20 conformères de la simulation avec un contact de 7 Å au sein de P[127-205] ont été sélectionnés aléatoirement et sont représentés dans la **Figure 70**. Sur cette dernière, on remarque bien un rapprochement spatial entre les régions en hélices  $\alpha$  correspondant à P<sub>OD</sub> et à  $\alpha_{C1/2}$ .



**Figure 69 :** Ajustement de données PRE de P[127-205] à partir de modèles générés par Flexible-meccano. Les diagrammes représentent les mêmes données que celles de la **Figure 68** mais en ajoutant les courbes théoriques qui ont chacune été calculées à partir de 1000 conformères générés par le logiciel. En vert clair est indiquée une simulation sans aucun contact à longue distance. En vert foncé, en bleu et en violet sont montrées les courbes des simulations avec des contacts à 15, 10 et 7 Å de distance entre la région 143-156 et 160-199.



*Figure 70 : Echantillon de 20 conformères de protomère de P[127-205] calculés par Flexible-meccano avec un contact de 7 Å entre  $P_{OD}$  et  $\alpha_{C1/2}$ . Les structures ont été sélectionnées aléatoirement parmi un ensemble de 1000 conformères calculés par le logiciel. Les conformères ont été alignés sur le domaine  $P_{OD}$  (ici au niveau des résidus 130-156). En bleu est indiqué le N-terminal et en rouge, le C-terminal.*

#### 4.1. Conclusions

La caractérisation structurale de la Phosphoprotéine du hRSV par RMN a permis de confirmer sa nature intrinsèquement désordonnée et a montré que P possède des régions transitoirement ou complètement structurées en contacts entre elles et impliquées dans des interactions protéines-protéines. De plus, l'obtention de la structure du domaine de tétramérisation de P a prouvé la formation d'une hélice surenroulée au niveau des résidus 129-157 qui ressemble fortement à la structure du domaine core de hMPV-P, son plus proche homologue connu en date.

# Chapitre 5. Etude des interactions entre protéines P—M2-1 et P—PP1 $\alpha$

## 5.1. Fonctions, interactions et phosphorylation de M2-1

Le facteur de transcription M2-1 du RSV, comme cité dans le Chapitre 1 (Introduction générale sur le RSV), a pour fonction principale d'augmenter la processivité du complexe ARN polymérase ARN dépendant (RdRp) du virus lors de l'étape de transcription de son ARN génomique en ARN messager. Cette fonction de M2-1 se déroule au sein de corps d'inclusion (inclusion bodies, IB) cytoplasmiques, qui sont des centres de répllication et de transcription du virus et où l'on trouve notamment les protéines L, P et N du RdRp post infection (García et al., 1993 ; Carromeu et al., 2007). De plus, la protéine M2-1 et les ARN messagers du virus se concentrent dans des sous-compartiments de ces corps d'inclusion appelés granules associées aux IB (IBAG) desquels N, P, L et l'ARN génomique du RSV sont absents (Rincheval et al., 2017).

Il est également connu que l'interaction du domaine core de M2-1 avec la protéine P est essentielle pour le recrutement du cofacteur de transcription du virus dans les IB (Blondot et al., 2012). Aussi, l'interaction compétitive qui a lieu entre la phosphoprotéine et les ARN sur la protéine M2-1 suggère une régulation cyclique de ces deux interactions (Richard et al., 2018).

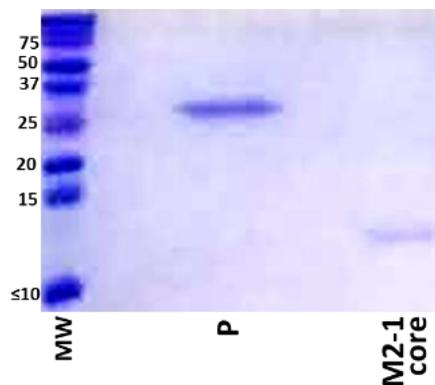
M2-1 existe dans différents états de phosphorylation (Routledge et al., 1987 ; Lambert et al., 1988 ; Hardy and Wertz, 2000). Les sérines 58 et 61 de M2-1 peuvent être phosphorylées et sont essentielles pour la fonction de la protéine, mais pas pour l'interaction avec la protéine P (Cartee and Wertz, 2001 ; Tran et al., 2009). La phosphorylation de M2-1 joue aussi un rôle dans la diminution de la capacité de fixation d'ARN sur la protéine. Dans les cellules infectées par le RSV ou lorsqu'elle est co-exprimée avec la protéine P dans des cultures cellulaires, la protéine M2-1 reste principalement non phosphorylée, alors que M2-1 est principalement phosphorylée lorsqu'elle est exprimée seule (Cuesta et al., 2000). Cet ensemble d'informations laisse à penser que M2-1 doit être phosphorylée de manière réversible et dynamique pour correctement exercer ses fonctions. C'est cette hypothèse de départ qui a conduit nos collaborateurs de l'INRA Jouy-en-Josas, notamment le Dr Charles-Adrien Richard (INRA), à étudier le mécanisme de déphosphorylation de M2-1 et l'implication de la protéine P dans ce processus.

Mon objectif dans ce projet a été de mettre en évidence et de cartographier des interactions entre protéines impliquées dans ce processus, notamment entre P et M2-1 dans un premier temps.

## 5.2. Etude de l'interaction M2-1—P

### 5.2.1. Production et purification des protéines M2-1<sub>core</sub>, P marquée <sup>15</sup>N et ses fragments marqués <sup>15</sup>N

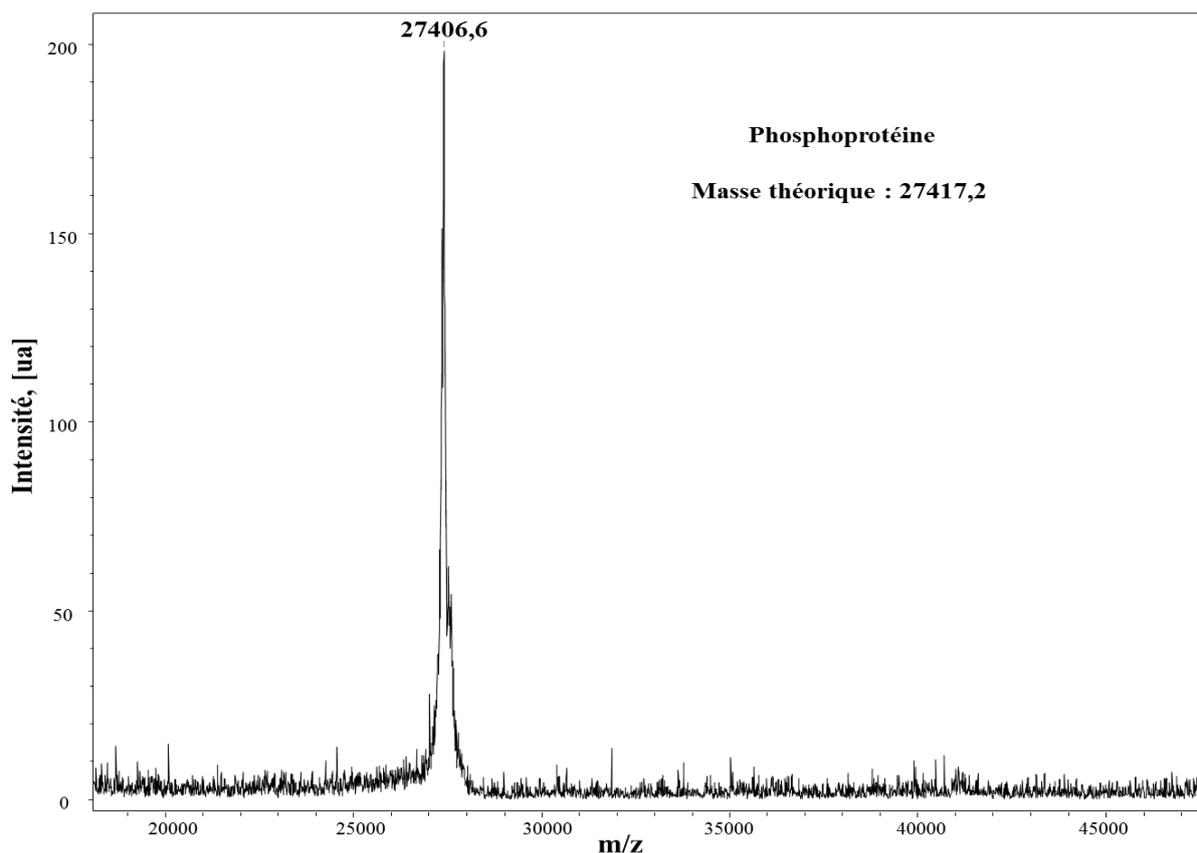
La protéine M2-1<sub>core</sub> (impliquée dans l'interaction avec P) ainsi que la protéine P et ses fragments marqués <sup>15</sup>N ont été exprimés et purifiés avec une étiquette GST en N-terminal et clivés par la thrombine comme indiqué dans le Chapitre 2 (Matériels et Méthodes). Les protéines <sup>15</sup>N-P et M2-1<sub>core</sub> purifiées sont représentées sur la **Figure 71**. Les protéines ont été dialysées dans un tampon 20 mM phosphate, 100 mM NaCl, 1 mM DTT (en effet M2-1<sub>core</sub> contient une cystéine libre dans sa séquence), pH 6,5.



**Figure 71** : SDS PAGE de <sup>15</sup>N-P et M2-1<sub>core</sub> purifiés. Les marqueurs de tailles et leurs masses en kDa sont indiqués dans la piste la plus à gauche.

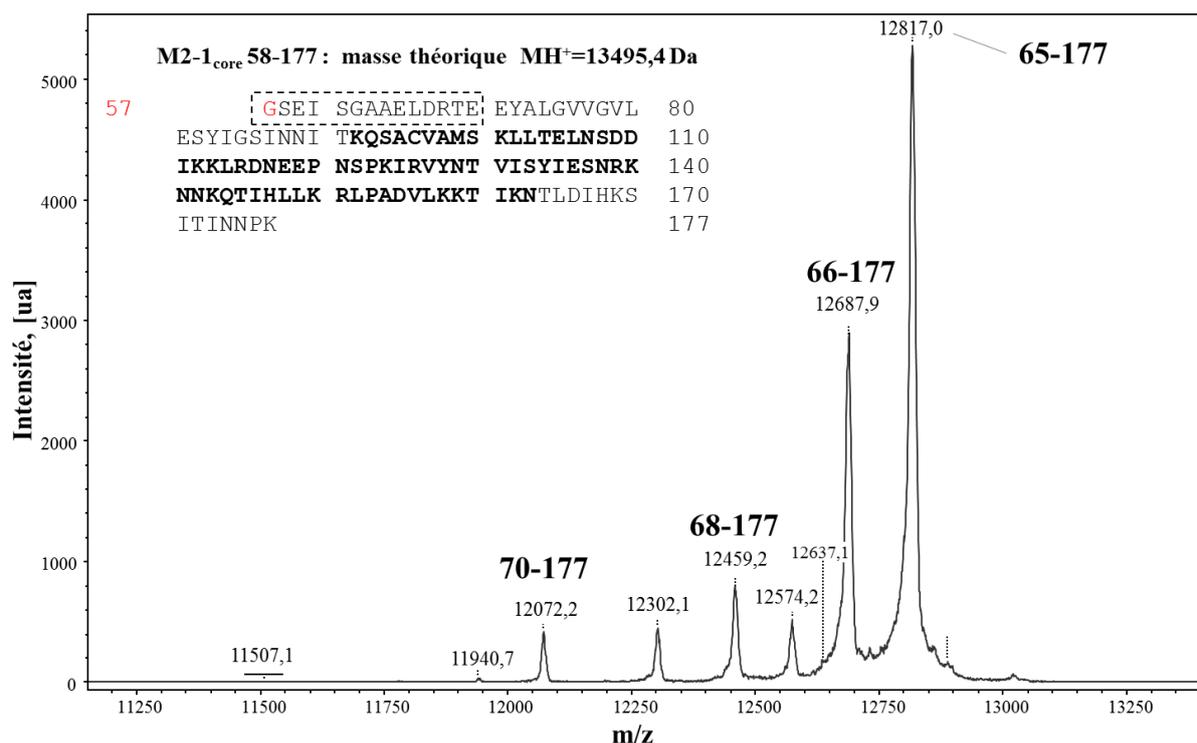
### 5.2.1. Analyse par spectrométrie de masse des protéines M2-1<sub>core</sub> et P marquée <sup>15</sup>N

Sachant que la phosphorylation de la thréonine 108 de la phosphoprotéine du RSV empêche la fixation de M2-1, nous avons vérifié voulu vérifier l'état de phosphorylation de la protéine P par spectrométrie de masse. Le spectre de masse de <sup>15</sup>N-P par MALDI-TOF montre que la masse moyenne du massif est très proche de la masse attendue pour <sup>15</sup>N-P non phosphorylée (**Figure 72**).



**Figure 72 :** Spectre MALDI-TOF de la protéine *P* marquée  $^{15}\text{N}$ . La masse mono isotopique attendue de la protéine est indiquée en haut à droite.

De même avec M2-1<sub>core</sub>, nous avons réalisé une analyse de l'état de la protéine par spectrométrie de masse (**Figure 73**). Les résultats montrent que le massif isotopique le plus abondant ne correspond pas à M2-1<sub>core</sub> en entier mais au domaine moins ses 8 premiers résidus. D'autres pics avec des intensités plus faibles apparaissent et correspondent à des fragments de plus en plus raccourcis du côté N-terminal de M2-1<sub>core</sub> (fragments 70-177 : fragment le plus écourté). Ce type de profil semble indiquer la présence de protéase contaminante qui aurait dégradé progressivement l'extrémité N-terminale non structurée de M2-1<sub>core</sub>. Il se trouve que l'échantillon de M2-1 utilisé était vieillissant. Pour éliminer toute trace de protéase, j'ai passé les échantillons de M2-1<sub>core</sub> en chromatographie d'exclusion. Toutefois, la dégradation N-terminale de M2-1<sub>core</sub> n'ont pas dû avoir trop d'impact étant donné que la région d'interaction de *P* attendue correspondant à la région 92-163 n'avait pas été atteinte et que les résidus 58 et 61 n'allaient pas être étudiés spécifiquement dans les expériences par RMN à venir.



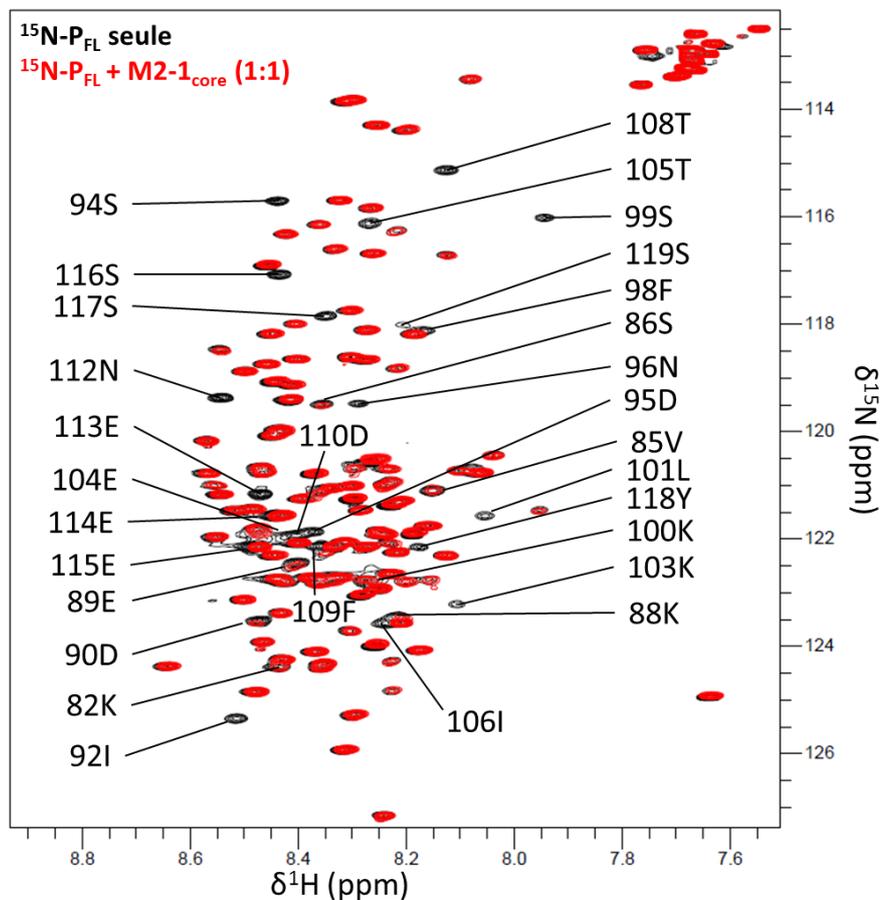
**Figure 73 : Spectre MALDI-TOF du domaine core de la protéine M2-1<sub>core</sub>.** La masse mono isotopique attendue de la protéine est indiquée en haut à gauche. Les intervalles de numéros en gras au-dessus de certains pics indiquent les fragments de M2-1 auxquels appartiennent les massifs. La séquence primaire de la protéine est également indiquée. Sur cette dernière est indiqué en rouge le résidu restant d'extension de l'étiquette GST après coupure à la thrombine et en gras la région d'interaction avec la protéine P (Blondot et al., 2012). Le cadre en pointillé autour de la séquence indique la partie partiellement ou complètement absente du spectre de masse.

### 5.2.2. Etude de l'interaction entre la protéine P marquée <sup>15</sup>N et le domaine core de M2-1 non marquée par RMN

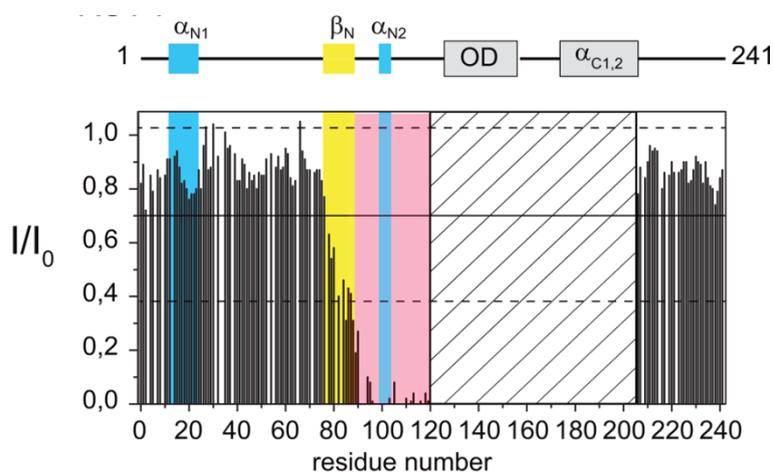
J'ai ensuite enregistré des spectres HSQC <sup>1</sup>H-<sup>15</sup>N de <sup>15</sup>N-P<sub>FL</sub> à ~50 μM en absence et en présence de M2-1<sub>core</sub> non marquée (1:1). J'ai reporté les attributions des déplacements chimiques de P entière sur les 2 spectres HSQC <sup>1</sup>H-<sup>15</sup>N. La superposition des spectres HSQC <sup>1</sup>H-<sup>15</sup>N (**Figure 74** et **Figure 75**) montre que la présence de M2-1<sub>core</sub> induit des diminutions d'intensité ou des élargissements de raie significatifs pour des signaux de la région couvrant les résidus 90–120, ce qui correspond notamment à la partie transitoirement structurée α<sub>N2</sub> en N-terminal de P juste en amont du domaine d'oligomérisation. Ces résultats sont en accord avec la localisation antérieure des résidus critiques, notamment la région englobant les résidus 100 à 120, pour la liaison de M2-1 sur P (Mason et al., 2003 ; Asenjo et al., 2006 ; Blondot et al., 2012).

La forte diminution (>60%) des intensités des signaux dans la région d'interaction de M2-1<sub>core</sub> est probablement due au fait que la région α<sub>N2</sub> voit son T<sub>2</sub> diminuer localement avec la fixation transitoire de M2-1<sub>core</sub>, ce qui augmente le temps de corrélation de réorientation en solution, ou à de l'échange

conformationnel entre la forme libre et liée de  $\alpha_{N2}$  ainsi que de l'échange entre l'état structuré en hélice de cette région à l'échelle  $\mu$ s-ms.



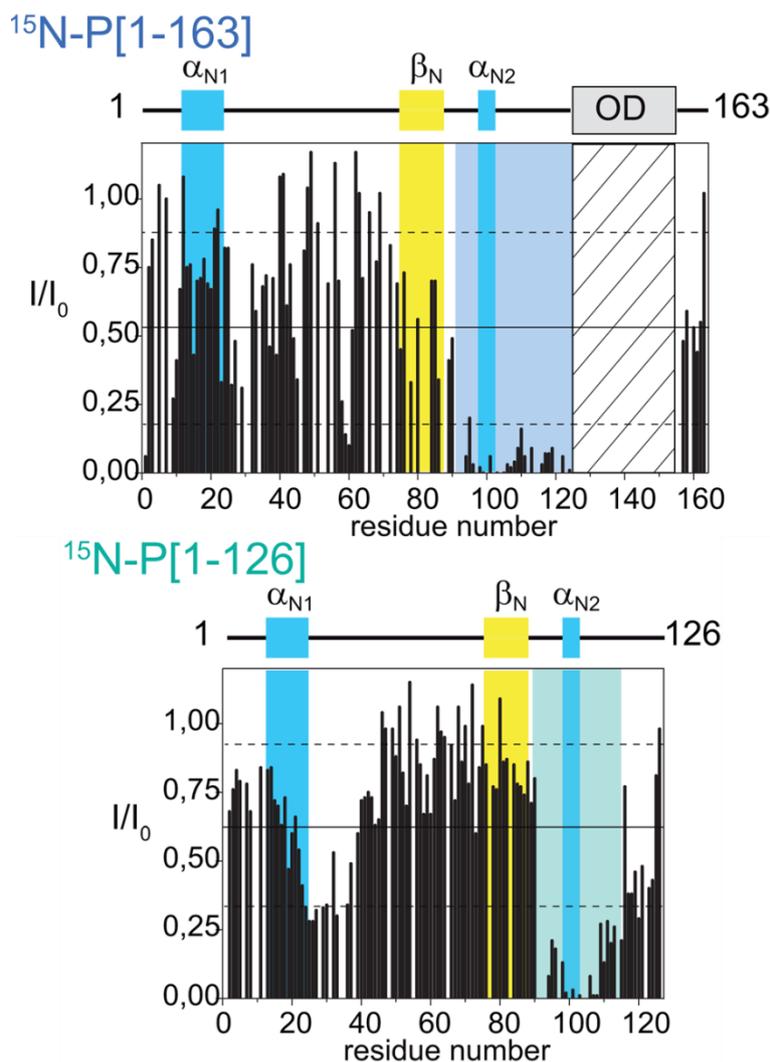
**Figure 74 :** Spectres HSQC  $^1\text{H}$ - $^{15}\text{N}$  superposés de  $^{15}\text{N}$ -P<sub>FL</sub> seule et en présence de M2-1<sub>core</sub> non marquée. Les numéros indiqués correspondent aux résidus dont l'intensité de leur signal est la plus affectée.



**Figure 75 :** Diagramme des intensités normalisées du spectre HSQC  $^1\text{H}$ - $^{15}\text{N}$  de  $^{15}\text{N}$ -P<sub>FL</sub> en présence de M2-1<sub>core</sub> non marquée. Le diagramme représenté a été tracé à partir des données d'intensité des signaux des spectres de la Figure 74. Les lignes droites et pointillées représentent les valeurs moyennes

et moyennes  $\pm$  RMSD (moyenne quadratique) sur tous les signaux. La position du domaine d'oligomérisation et des régions présentant une structure secondaire transitoire en hélice  $\alpha$  ou en brin  $\beta$  le long de la séquence de P est indiquée dans le schéma au-dessus du graphique. Le fond rouge indique la région autour de l'hélice transitoire  $\alpha_{N2}$  où M2-1<sub>core</sub> induit une diminution de l'intensité >60%. La zone hachurée correspond à la région de P englobant le domaine OD et  $\alpha_{C1/2}$  pour lesquelles les pics sont trop élargis et donc invisibles sur les spectres.

Nous avons également effectué des expériences d'interaction entre M2-1<sub>core</sub> et les deux fragments de P, P<sub>ND</sub> et P<sub>ND+OD</sub> marqués  $^{15}\text{N}$ . Les résultats sont présentés en **Figure 76**.



**Figure 76 : Diagrammes des intensités normalisées des spectres HSQC  $^1\text{H}$ - $^{15}\text{N}$  de  $^{15}\text{N}$ -P<sub>ND</sub> et  $^{15}\text{N}$ -P<sub>ND+OD</sub> en présence de M2-1<sub>core</sub> non marquée.** Les diagrammes ont été fait avec des spectres HSQC  $^1\text{H}$ - $^{15}\text{N}$  de P<sub>ND</sub> à 75  $\mu\text{M}$  et P<sub>ND+OD</sub> 35  $\mu\text{M}$  marqués  $^{15}\text{N}$ , seuls et en présence d'un équivalent molaire de M2-1<sub>core</sub>. Les régions autour de  $\alpha_{N2}$  avec des intensités normalisées inférieures à la moyenne quadratique sont mises en évidence avec des arrière-plans bleu et cyan pour P<sub>ND</sub> et P<sub>ND+OD</sub> respectivement.

Le fragment  $^{15}\text{N}$ -P<sub>ND+OD</sub> en présence de M2-1<sub>core</sub> montre un comportement similaire à celui de P<sub>FL</sub>, avec un élargissement de raie quasi total dans la région  $\alpha_{N2}$ . Avec  $^{15}\text{N}$ -P<sub>ND</sub>, il y a une seconde région perturbée, dans une moindre mesure, à l'extrémité N-terminale de P (résidus 23-37) à proximité de la

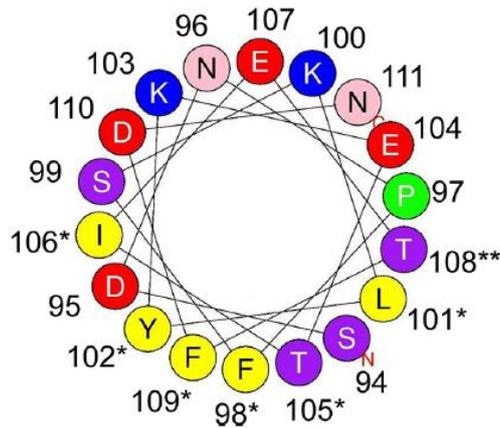
région  $\alpha_{N1}$ . Ce phénomène a été observé en sens inverse avec l'interaction de  $N_0$  se fixant sur  $\alpha_{N1}$  mais affectant les intensités de la région  $\alpha_{N2}$  (Pereira et al., 2016). Dans ce cadre, il serait risqué d'admettre un second site de fixation de M2-1 sur la région proximale à  $\alpha_{N1}$ . En effet, on pourrait corrélérer cet effet avec les contacts transitoires à longue distance entre  $\alpha_{N1}$  et  $\alpha_{N2}$  vus dans le Chapitre 4. Il est possible que ces contacts soient plus probables dans la forme monomère  $P_{ND}$  que dans le tétramère où des contacts entre mêmes régions sur des protomères différents,  $\alpha_{N1}-\alpha_{N1}$  ou  $\alpha_{N2}-\alpha_{N2}$  viennent en compétition avec des contacts  $\alpha_{N1}-\alpha_{N2}$ . La fixation de M2-1<sub>core</sub> sur  $\alpha_{N2}$  viendrait en compétition avec les contacts internes à P. On remarque également que les diminutions d'intensité supérieures à 60% s'étendent sur plus de signaux en aval de  $\alpha_{N2}$  quand le domaine d'oligomérisation est présent comme si la présence de ce dernier favorisait une extension de l'hélice  $\alpha_{N2}$  côté C-terminal, ce qui permettrait une plus importante surface d'interaction avec M2-1<sub>core</sub>.

### 5.2.3. Expériences complémentaires pour l'étude de l'interaction P—M2-1

Des expériences complémentaires ont été menées par nos collaborateurs pour l'étude de l'interaction P—M2-1 (Richard et al., 2018).

Des expériences de GST pulldown ont montré qu'un fragment de P contenant uniquement les résidus 93 à 110 fusionnés à une étiquette GST pouvait retenir M2-1 entière. Cela confirme l'hypothèse d'une région minimale d'interaction au niveau des résidus 90-112 de P. Aucune autre construction de P sans cette région n'a montré d'affinité pour M2-1, ce qui confirme également que la région proximale de  $\alpha_{N1}$  n'est pas un second site de fixation de M2-1.

Des expériences de transcription en minigénome avec les protéines P (sauvage ou mutée), L, N, M2-1 et une luciférase (schéma en **Figure 16A**) comme gène rapporteur ont été réalisées pour évaluer l'incidence de mutations au niveau de la région 85-115 de P. Les mutations des résidus F87, F98, L101, Y102, T105, I106 et F109 ont un effet drastiquement négatif sur l'activité de transcription du complexe RdRp et pourraient former une surface hydrophobe avec une hélice  $\alpha$  étendue de la région  $\alpha_{N2}$  (**Figure 77**). Une mutation phosphomimétique T108D diminuant complètement l'activité de transcription a quant à elle montré l'impact négatif d'une phosphorylation « permanente » dans cette région (Asenjo et al., 2006). Hormis le résidu F87, les résidus critiques pour la transcription sont tous localisés dans la région 93-110.

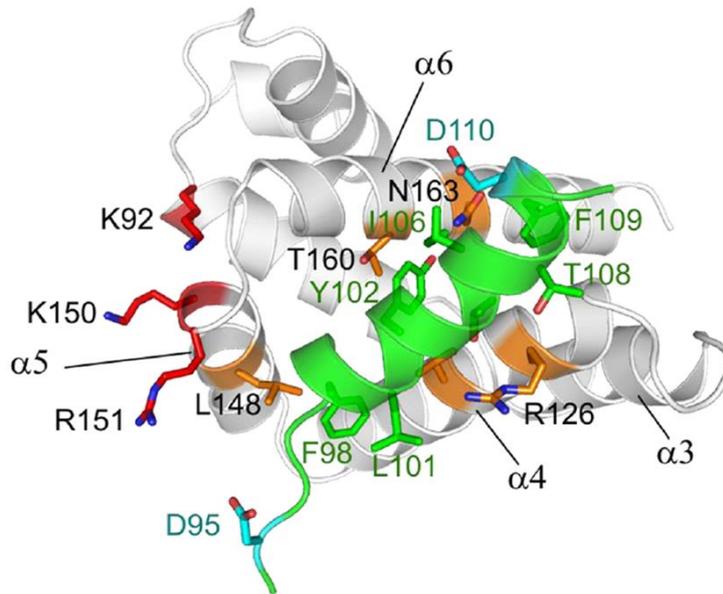


**Figure 77 : Roue hélicoïdale montrant les résidus de P critiques pour l'activité de transcription du RSV.** Représentation en roue hélicoïdale de l'hélice  $\alpha$  putative située entre les résidus S94 et N111 de P (serveur HeliQuest). Les résidus critiques pour la liaison M2-1 sont indiqués par une étoile et le résidu phosphorylable T108 est indiqué par deux étoiles. Les résidus chargés positivement sont en bleu, les résidus chargés négativement en rouge, les résidus phosphorylés putatifs en violet et les résidus hydrophobes en jaune.

De plus, des expériences de co-localisation par microscopie confocale de fluorescence avec M2-1-mCherry, N, P (sauvage ou mutée) et une immuno-détection indirecte de P ont été menées et ont montré que les mutants F98A, L101A, Y102A, T105A/D, I106A et F109A et T108D de P empêchaient le recrutement de M2-1 par P dans les IB contrairement à P<sub>FL</sub> sauvage. Cela expliquerait l'extinction de l'activité de transcription du RSV se déroulant dans les IB. Cependant, la mutation F87A affectant négativement la transcription n'a pas impacté la localisation de M2-1 en IB.

Par GST pulldown, il a été observé que les mutants F98A, L101A, Y102A, T105A/D, I106A et F109A et T108D de P couplés à la GST, abolissaient complètement l'interaction avec M2-1 contrairement à la mutation F87A.

En rassemblant les informations de résidus impliqués dans la fixation des deux partenaires (en prenant en compte les résidus impliqués de M2-1 dans Blondot et al., 2012) et en utilisant la structure de M2-1<sub>core</sub> (pdb : 4C3E) et une construction d'hélice  $\alpha$  dans P[94-112] (calculée avec CYANA à partir de contraintes d'angles  $\phi/\psi$  prédites avec les déplacements chimiques correspondants), il a été possible d'obtenir un modèle de complexe M2-1<sub>core</sub>—P[94-112] avec le serveur HADDOCK (de Vries et al., 2010 ; Wassenaar et al., 2012). Le résultat est représenté dans la **Figure 78**. Le modèle a obtenu le meilleur score dans HADDOCK et correspond à 101 conformères avec un RMSD de 0,7 Å sur 200 (1000 générées initialement). Le modèle montre des interactions de type hydrophobe ainsi qu'électrostatique et les énergies correspondantes dans HADDOCK montrent que l'énergie électrostatique contribue ~2 fois plus que celle des interactions hydrophobes (énergie en kcal/mol : Van der Waals : -40,4 ; désolvatation : -33,6 ; électrostatique : -155,5). Ainsi, il n'est pas surprenant que les mutations phosphomimétiques T105D et T108D dans P abolissent totalement l'interaction P—M2-1.



**Figure 78 : Modèle d'un complexe P—M2-1.** Le modèle a été obtenu en rapprochant l'hélice D95-F109 de P (en vert) sur un protomère de M2-1<sub>core</sub> (en gris, avec indication des numéros d'hélice). Les résidus de P et M2-1 critiques pour la fixation des deux sont indiqués (en orange pour M2-1). Les résidus impliqués dans le domaine de liaison à l'ARN au sein de M2-1 sont en rouge.

## 5.1. Déphosphorylation de M2-1 par PP1 $\alpha$ avec l'aide de P

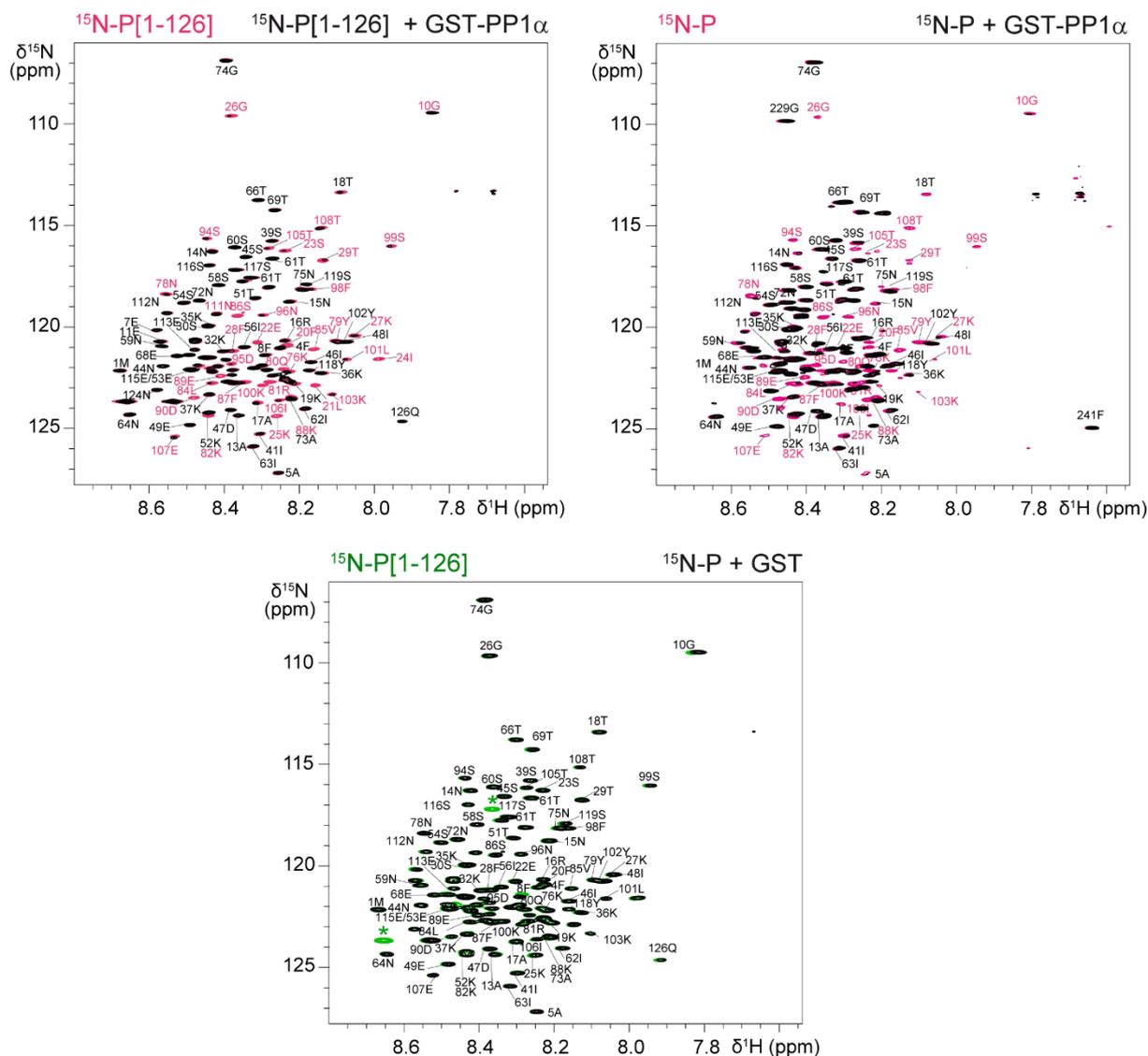
Pour clarifier les différences observées entre F87A et les mutants affectant l'interaction P—M2-1, des expériences d'immunoprécipitation de P (sauvage et mutée) à partir de cellules de mammifère ont été menées en présence de N et M2-1 avec détection de P et de M2-1 par le Dr Charles-Adrien Richard (INRA). L'analyse des résultats a montré que l'interaction P—M2 est primordiale pour la déphosphorylation de M2-1 et que le résidu F87 de P joue un rôle important dans cette dernière.

Ces données ont conduit à l'hypothèse d'une association entre P et une phosphatase cellulaire responsable de la déphosphorylation M2-1. Chez EBOV, il a été montré que l'homologue de M2-1, VP30 est déphosphorylé par la phosphatase PP1 et que la déphosphorylation de VP30 est critique pour la transcription virale (Ilinykh et al., 2014). PP1 ne reconnaît pas de séquences spécifiques sur sa protéine cible. Au lieu de cela, la liaison au substrat dépend de son association avec des protéines interagissant avec PP1 (PP1-Interacting Protein, PIP) qui fonctionnent comme des sous-unités de ciblage (Bollen et al., 2010). Une majorité des PIP connues contiennent un motif de liaison à PP1 court "RVxF" (R/K-K/R-x(0,1)-V/I-x-F/W/Y). Or, RSV P contient une séquence <sup>81</sup>RKPLVSF<sup>87</sup> conforme au motif "RVxF". De plus, la protéine P chez les *Pneumoviridae* montre une forte conservation du motif "R / x-x-x-V-T / S-F", se réduisant à "R-K-x-V-T/S-F" pour les pneumovirus. RSV-P abrite donc un motif "RVxF" dégénéré contenant le résidu F87 et est très probablement une PIP.

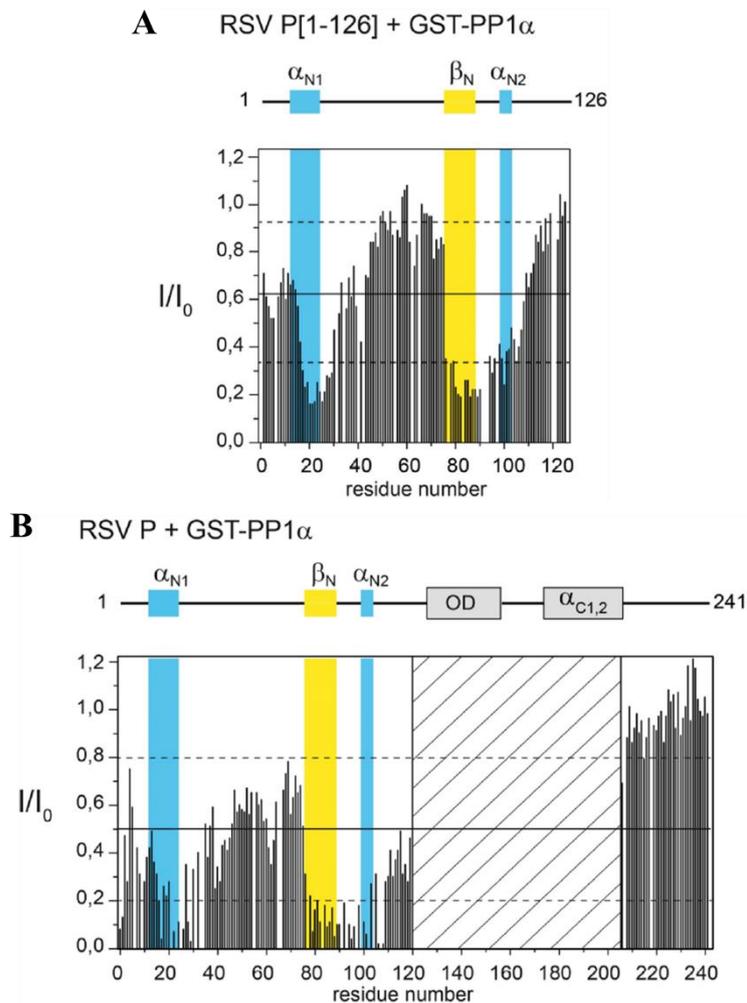
### 5.1.1. Démonstration d'une interaction directe entre PP1 $\alpha$ et P

Dans ce sens, une seconde expérience d'immunoprécipitation de P en présence de N dans des cellules de mammifère a été réalisée, toujours à l'INRA, et a permis de montrer une interaction entre P et la phosphatase PP1 $\alpha$  endogène alors que la même expérience avec le mutant F87A de P n'a pas révélé de présence de PP1 $\alpha$  co-précipitée.

Nous avons ensuite vérifié au laboratoire si P interagissait directement avec PP1 $\alpha$  par RMN. Nous avons, pour ce faire, enregistré des spectres  $^1\text{H}$ - $^{15}\text{N}$  BEST-TROSY de  $^{15}\text{N}$ -P<sub>ND</sub> et  $^{15}\text{N}$ -P seules et en présence de GST-PP1 $\alpha$  (produite et purifiée par le Dr Charles-Adrien Richard, INRA) ou GST non marquée (**Figure 79**). Un élargissement de raies induit par GST-PP1 $\alpha$  dans une région couvrant les résidus K76-D90 a été observé au sein de P<sub>ND</sub> (**Figure 80A**). Des expériences de contrôle en présence de GST seule n'ont révélé aucune perturbation d'intensité. La région 76-90 contient le motif "RVxF" dégénéré. Il est adjacent au site de liaison de M2-1, mais n'a pas été perturbé par la fixation de M2-1<sub>core</sub> dans P<sub>ND</sub> (**Figure 76**). La région 76-90 avait déjà attiré notre attention en raison d'une relaxation transversale  $^{15}\text{N}$  plus efficace que dans les IDR adjacentes et de sa propension en brin  $\beta$  de P (**Figure 52**). Comme les motifs "PPI RVxF" adoptent une conformation en brin  $\beta$  en étant en complexe avec PP1 $\alpha$  (Choy et al., 2012), nous avons émis l'hypothèse que la région 76-90 forme un site de liaison primaire à PP1 $\alpha$  de la même manière. PP1 $\alpha$ , comme M2-1<sub>core</sub> a induit une diminution des intensités de la région  $\alpha_{\text{N1}}$  et autour avec P<sub>ND</sub>. Contrairement à M2-1<sub>core</sub>, lorsque nous avons effectué des expériences d'interaction avec P entière (**Figure 80B**), des perturbations au voisinage de  $\alpha_{\text{N1}}$  ont encore été observées et nous ne pouvons pas exclure qu'elles reflètent cette fois-ci une liaison directe à un second site de liaison pour PP1 $\alpha$ .



**Figure 79 :** Spectres BEST-TROSY  $^1\text{H}$ - $^{15}\text{N}$  de  $^{15}\text{N}$ -P<sub>ND</sub>/ $^{15}\text{N}$ -P avec GST-PP1 $\alpha$ /GST. (Panneau supérieur) Spectres BEST-TROSY  $^1\text{H}$ - $^{15}\text{N}$  superposés de  $^{15}\text{N}$ -P<sub>ND</sub> ou de  $^{15}\text{N}$ -P à 25  $\mu\text{M}$  seuls (contours magenta) et en présence 2 équivalents molaires de GST-PP1 $\alpha$  (contours noirs) ; les attributions sont également indiquées. (Panneau inférieur) Spectres HSQC  $^1\text{H}$ - $^{15}\text{N}$  superposés de  $^{15}\text{N}$ -P<sub>ND</sub> à 10  $\mu\text{M}$  seul (contours verts) et en présence de 10 équivalents molaires de GST (contours noirs). Le tampon de l'échantillon  $^{15}\text{N}$ -P<sub>ND</sub> + GST contenait 25 mM de glutathion qui produisait des signaux d'abondance naturelle  $^{15}\text{N}$  indiqués par des étoiles.



**Figure 80 : Diagrammes des intensités normalisées des spectres BEST-TROSY  $^1\text{H}$ - $^{15}\text{N}$  de  $^{15}\text{N}$ -P<sub>ND</sub> et  $^{15}\text{N}$ -P<sub>FL</sub> en présence de PP1 $\alpha$  non marquée. Les rapports d'intensité ( $I/I_0$ ) ont été mesurés avec les spectres de  $^{15}\text{N}$ -P<sub>ND</sub> (A) ou  $^{15}\text{N}$ -P (B) entière à 25  $\mu\text{M}$  seules et en présence de 2 équivalents molaires de GST-PP1 $\alpha$ . Les lignes droites et pointillées indiquent les valeurs moyennes et moyennes  $\pm$  moyenne quadratique. Les régions présentant une structure secondaire transitoire en hélice  $\alpha$  ou en feuille  $\beta$  sont mises en surbrillance par un arrière-plan coloré avec le même code de couleur que celui de la **Figure 76**.**

Des expériences GST pulldown ont également confirmé la capacité de PP1 $\alpha$  à interagir directement avec P et l'importance primordiale du résidu F87A de P dans l'interaction avec PP1 $\alpha$ .

### 5.1.2. Expériences complémentaires avec P, PP1 $\alpha$ et M2-1

D'autres expériences de co-localisation par microscopie confocale de fluorescence avec un test de transcription en minigénome en exprimant la protéine PP1-GFP en supplément ont montré un recrutement de PP1 $\alpha$  dans les IBs par P sauvage mais pas avec le mutant F87A.

Des expériences supplémentaires de co-localisation par microscopie confocale de fluorescence en utilisant P (sauvage ou F87A), PP1 $\alpha$ , M2-1 et des ARNm détectés par hybridation *in situ* en

fluorescence ont confirmé la colocalisation de M2-1 non phosphorylée et d'ARNm néo-synthétisés dans les IBAGs en présence de P alors que seuls les ARNm se trouvent en IBAG quand M2-1 est phosphorylée. De plus, il a été montré dans ce travail que M2-1 n'est pas essentiel pour la transcription du premier ARNm du RSV ni pour la formation de ces agrégats d'ARN appelés IBAGs. Enfin, M2-1 phosphorylée ou non pourrait être impliquée dans la poly-adénylation d'ARNm, mais M2-1 phosphorylée ne semble pas se fixer aux ARN poly (A) néo-synthétisés.

Pour terminer, des expériences de transcription en minigénome du RSV ont été réalisées avec différents niveaux de surexpression de PP1 $\alpha$  dans les cellules et ont confirmé la déphosphorylation de M2-1 par PP1 $\alpha$ . L'expérience a aussi mis en évidence le fait qu'une diminution du ratio M2-1 phosphorylé / non phosphorylé défavorise le niveau de transcription du RdRp du RSV et prouve que qu'un équilibre dynamique de phosphorylation/déphosphorylation de M2-1 est essentiel à cette étape du virus.

## 5.2. Conclusions

En définitif, j'ai montré que la région  $\alpha_{N2}$  de P est impliquée dans l'interaction avec M2-1 et que la région  $\beta_N$ , proximale de  $\alpha_{N2}$ , sur P est impliquée dans l'interaction avec la phosphatase PP1 $\alpha$ .



# Chapitre 6. Synthèse, discussion et perspectives

Dans les chapitres précédents j'ai illustré mon apport dans l'étude structurale de deux protéines essentielles du complexe ARN polymérase ARN dépendante du hRSV, la phosphoprotéine (P) et la nucléoprotéine (N), en utilisant la biochimie et la RMN des protéines comme outils principaux. J'ai ainsi résolu la structure à l'échelle atomique de petits domaines de ces protéines et caractérisé des ensembles structuraux pour la protéine P qui est intrinsèquement désordonnée en-dehors de son domaine d'oligomérisation et de domaines isolés en hélices. J'ai aussi utilisé la RMN pour caractériser les interactions entre ces protéines et des protéines partenaires identifiées préalablement. Dans le cas de l'interaction entre le bras C-terminal de la protéine N et le domaine tandem UBZ de Tax1BP1, la RMN a montré que cette interaction n'était probablement pas pertinente. Dans le cas des interactions entre la protéine P et les protéines M2-1 et PP1 $\alpha$ , la RMN a permis de localiser les régions d'interaction et de montrer qu'elles se situaient dans des régions structurées transitoirement, soit en brin  $\beta$  soit en hélice  $\alpha$ . Les études RMN ont été combinées avec des études fonctionnelles qui ont permis de mieux comprendre le rôle des complexes formés par P pour la synthèse d'ARN viral du hRSV. Je résumerai et commenterai ci-dessous les principaux résultats.

## 6.1. Interaction entre la nucléoprotéine du hRSV et la protéine cellulaire Tax1BP1

Au-delà de son rôle dans l'encapsidation de l'ARN viral du RSV, la nucléoprotéine du RSV a été identifiée dans des interaction avec des protéines cellulaires telles que MDA5, RIG-I (Lifland et al., 2012), l'arginine méthyl-transférase, la protéine méthylosome, la protéine HSP70 (Oliveira et al., 2013) ainsi que la protéine Tax1BP1.

La protéine Tax1BP1 est impliquée dans la régulation négative de l'immunité innée et de l'apoptose, mais la protéine est aussi impliquée dans le recrutement de la machinerie pour l'autophagie. Après la découverte de l'interaction N—Tax1BP1 par le Dr Pierre-Olivier VIDALAIN (Institut Pasteur), le Dr Marie Galloux (INRA) a montré que, lors d'une infection par le RSV, Tax1BP1 pourrait avoir un rôle dans la diminution de l'autophagie qui fait partie de la réponse antivirale de l'hôte infecté, en l'occurrence la souris. De plus, l'absence de gène Tax1BP1 induit une diminution de la réplication virale, peut-être corrélée à une augmentation de la biogénèse d'autophagosome chez l'hôte infecté. L'augmentation de l'expression de Tax1BP1 en minigénome du RSV a aussi montré une augmentation significative de l'activité polymérase du RdRp *in cellula*. Cependant, aucun essai de surexpression de Tax1BP1 *in vivo* n'a encore été entrepris. Sur la base de ces expériences fonctionnelles

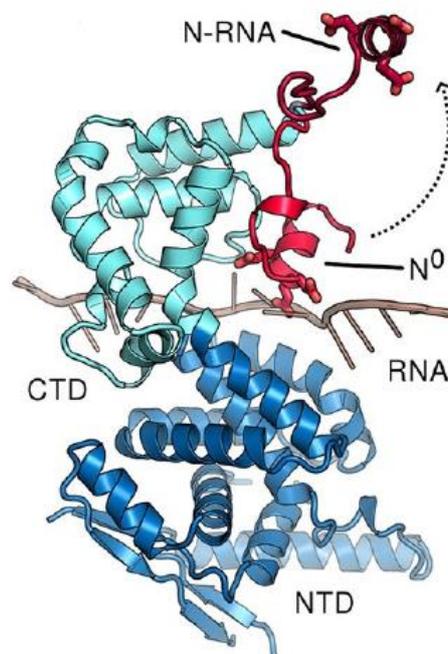
préliminaires et de preuves supplémentaires d'une interaction directe entre N et Tax1BP1 par GST pulldown et immunoprécipitation, je me suis lancé dans l'étude de l'interaction entre les deux partenaires. En effet, plusieurs expériences complémentaires de GST pulldown et d'immunoprécipitation de domaines isolés de Tax1BP1, puis de N, avaient été réalisées. Ainsi, les domaines N(C arm) et Tax1BP1(UBZ1 + 2) ont été proposés comme des interactants potentiels.

Le domaine Tax1BP1(UBZ1+2) avait été déjà étudié par RMN et cristallographie (Ceregido et al., 2014) et sa structure représentant deux doigts de zinc en tandem était disponible (pdb : 2m7q et 4bmj)

J'ai produit, puis purifié de manière isolée le domaine N(C arm) avec un marquage  $^{15}\text{N}$  et  $^{13}\text{C}$ . Après enregistrement d'expériences RMN, l'attribution des fréquences de résonance du squelette peptidique et des chaînes latérales a été réalisée. Une structure a été calculée à partir des données du spectre NOESY HSQC  $^1\text{H}$ - $^{15}\text{N}$  et des contraintes d'angle dièdre prédites avec les déplacements chimiques de N(C arm). Les contraintes d'angle dièdres ainsi que la structure générée de N(C arm) ont montré la formation d'une hélice  $\alpha$  stable au milieu du peptide entre les résidus 372 à 380 ainsi qu'une légère et plausible augmentation de rigidité de la région C-terminale (après le résidu 380) par rapport aux résidus avant l'hélice  $\alpha$  (résidus 361-371). J'ai aussi remarqué que l'augmentation de la concentration en N(C arm) semblait augmenter la propension de l'hélice  $\alpha$  372-380 au niveau de l'extrémité C-terminale.

Les titrages par RMN entre les domaines N(C arm) et Tax1BP1(UBZ1 + 2) marqués  $^{15}\text{N}$  tour à tour n'ont montré que de très faibles perturbations de déplacements chimiques (sans atteindre de plateau de saturation) et d'intensité des signaux, même en dépassant des ratios de protéines marquées sur protéines non marquées de l'ordre de 30 (ce qui correspond à plus de 3 mM pour le partenaire le plus concentré). Les différents contrôles ont montré que les légères perturbations des signaux RMN observées à forte concentration en partenaire pouvaient être dues à l'effet d'une augmentation de viscosité en solution, et/ou à un effet de rapprochement dans l'espace des protéines de plus en plus concentrées qui induirait une rigidification de certaines régions et/ou à de légères perturbations du pH (bien que celui-ci ait été contrôlé de manière drastique). Une autre explication de cette absence d'interaction entre les 2 domaines de N et Tax1BP1 pourrait être liée aux conditions expérimentales testées, notamment pour la concentration en sel ou le pH ou même la température à laquelle les expériences ont été faites. Aussi, il est important de souligner que Tax1BP1 est une protéine dimérique alors que le domaine UBZ1+2 isolé est monomérique. Par conséquent cela pourrait affecter la formation d'un complexe avec N(C arm) par absence de coopérativité. Cependant, si l'interaction dépendait uniquement de la dimérisation de UBZ1+2, alors le GST pulldown avec UBZ1+2 étiquetée à la GST aurait normalement pu permettre la co-précipitation de N(C arm) si l'on tient compte du fait que la GST est dimérique.

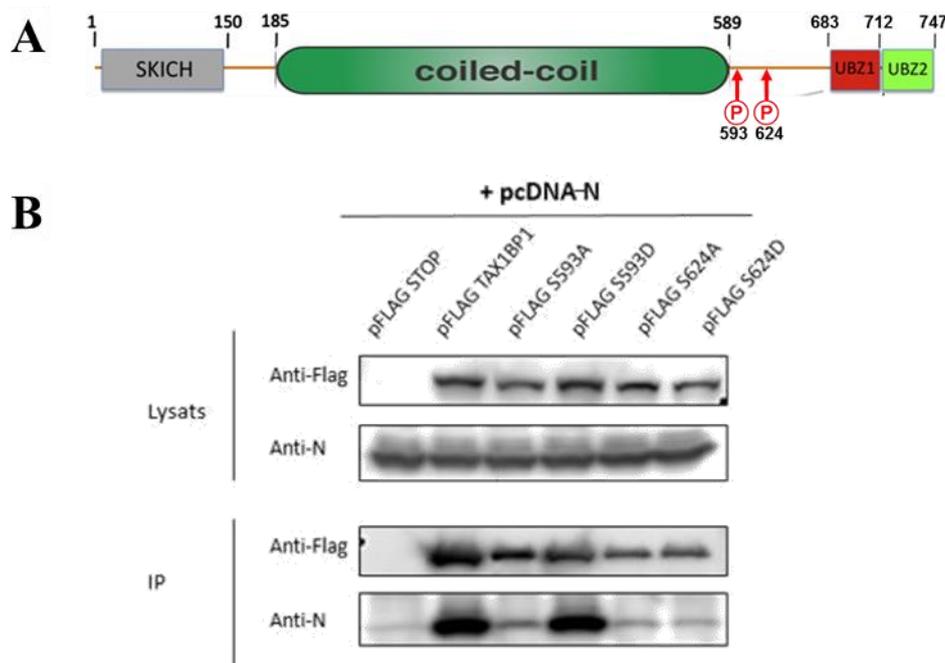
On peut aussi souligner que, même si les 10 acides aminés C-terminaux sont probablement libres, puisqu'ils sont absents des structures cristallographiques rapportées pour les anneaux de hRSV-N-ARN (pdb 2WJ8) et de hMPV-N-ARN (pdb 5FVC) ainsi que de la structure cristallographique de hMPV-N<sup>0</sup>-P[1-28] (pdb 5FVD), le bras C-terminal de N se replie partiellement sur le domaine C-terminal de N, ce qui pose la question de la disponibilité de cette région pour former des interactions. Mais même les résidus N-terminaux de N(C arm), hMPV-365-382 (équivalents à hRSV-363-380), ne sont pas associés de manière invariable au corps de la nucléoprotéine du HMPV (Renner et al., 2016) : cette partie effectue un grand mouvement dans le passage de la forme N<sup>0</sup> vers la forme N-ARN (**Figure 81**). Cette région de N(C arm) vient alors s'engager dans une liaison avec un protomère adjacent. On peut alors quand-même envisager que tout le N(C arm) puisse être disponible pour une interaction avec un autre partenaire protéique.



**Figure 81 : Changement conformationnel du bras C-terminal de la nucléoprotéine du hMPV.** Les domaines N et C-terminaux de N sont représentés en « cartoon » bleu et cyan respectivement. Le domaine N(C arm) est représenté en « cartoon » rouge : il est orienté vers le bas dans le complexe N<sup>0</sup>-P et vers le haut dans la structure d'anneaux N-ARN, comme indiqué sur la figure (Renner et al., 2016).

Néanmoins mes résultats montrent qu'il est nécessaire de redéfinir les domaines ou régions qui sont impliqués dans l'interaction entre les partenaires N et Tax1BP1. En plus du mutant Tax1BP1ΔSKITCH qui interagit très bien avec N mais dont la taille est trop importante pour l'étude RMN, les deux sérines phosphorylables 593 et 624 de l'isotype 2 de Tax1BP1, localisées juste après le domaine d'oligomérisation, ont été identifiées comme étant critiques dans l'interaction de Tax1BP1 avec la protéine N. A ce propos, Sarah MADRIERES (M2, 2016-2017) a mené des expériences d'immunoprécipitation pouvant suggérer que les résidus 593 sous forme phosphorylée 624 non

phosphorylé de Tax1BP1 sont nécessaires à l'interaction avec N (**Figure 82**). La détermination de l'état de phosphorylation de Tax1BP1 par spectrométrie de masse aurait ici un intérêt pour confirmer ou discuter ces résultats. Si l'impact de ces 2 sérines sur l'interaction Tax1BP1—N se confirme et que le linker sur lequel se situent ces résidus suffit à l'interaction avec N, il pourra être intéressant d'étudier par RMN une construction marquée  $^{13}\text{C}$ - $^{15}\text{N}$  du linker entre le domaine d'oligomérisation et le domaine UBZ1+2 de Tax1BP1 en interaction avec N (entière ou avec un/des domaine(s) spécifiques) avec éventuellement des mutations sur S593/S624 pour évaluer l'impact sur l'interaction avec N, puis tester des formes phosphorylées de ce fragment de Tax1BP1 si les essais sont concluants.



**Figure 82 : Etude du rôle de la phosphorylation de TAX1BP1 dans l'interaction avec N.** (A) Schéma des domaines de l'isoforme 2 de Tax1BP1. Les délimitations des domaines et les positions de phosphorylation sont indiquées par les numéros des résidus correspondants. (B) Immunoprécipitation à partir de cellules BSRT7 co-transfectées avec les plasmides codant pour pFLAG-TAX1BP1 (sauvage ou mutant) et pcDNA-N. Une immunoprécipitation en présence d'anticorps anti-FLAG a ensuite été réalisée sur les lysats cellulaires. Les protéines TAX1BP1 et N ont été révélées par western blot en présence d'anticorps anti-FLAG-HRP et anti-N (Rapport de stage de M2 Sarah MADRIERES).

Un autre stagiaire du Dr Marie Galloux (INRA), Antoine DOMENGER (AgroParisTech 2<sup>e</sup> année, 2017-2018), a montré par immunoprécipitation que la protéine P empêchait l'interaction N—Tax1BP1, donc que P entrerait en compétition avec TAX1BP1 pour l'interaction avec N. Ces données suggèrent que l'interaction N—TAX1BP1 devrait se passer hors des IB où se concentrent les protéines du complexe ARN polymérase, dont N et P, durant l'infection par le RSV.

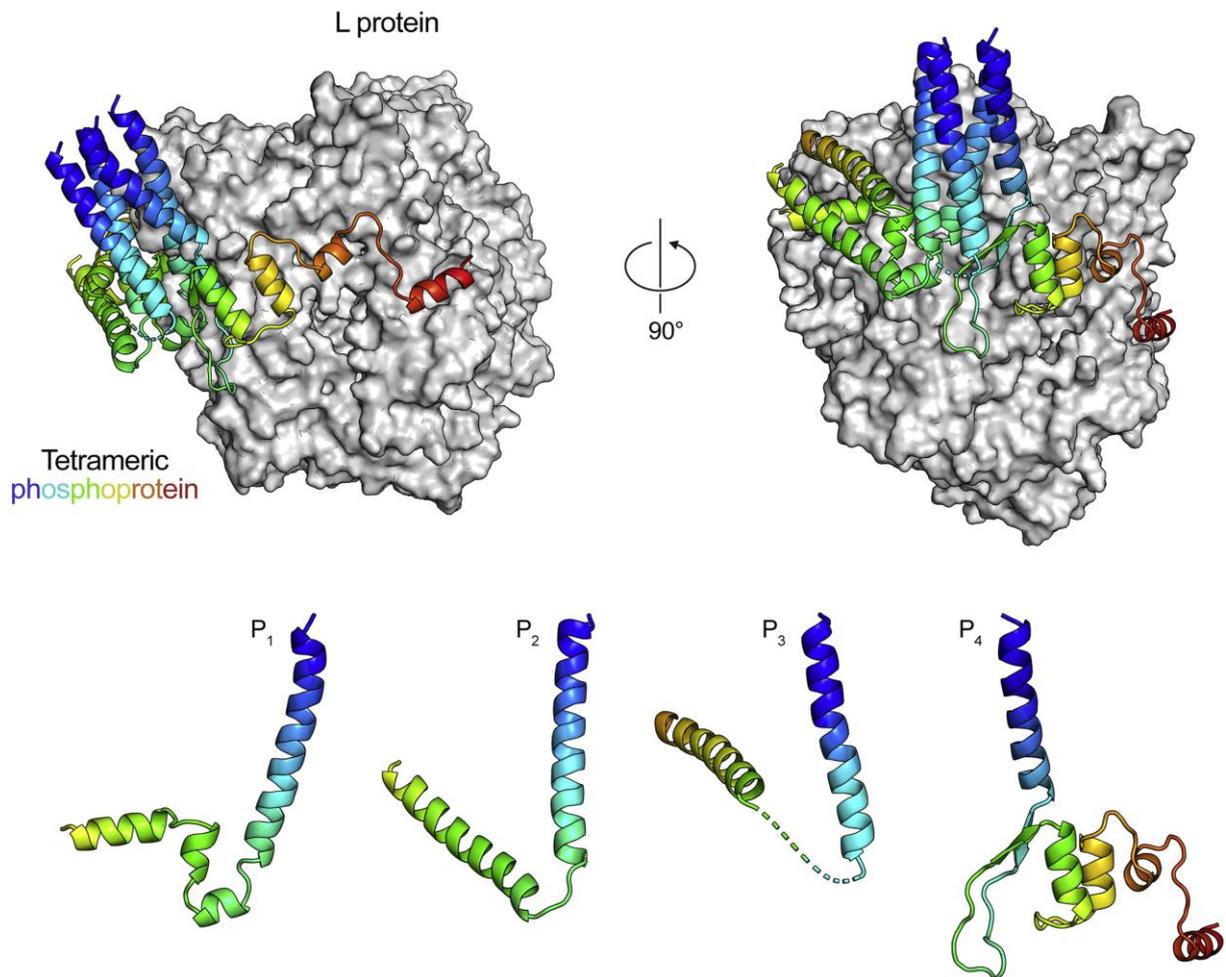
## 6.2. Caractérisation structurale de la phosphoprotéine du hRSV

La détermination de structure du domaine d'oligomérisation de P par RMN a été concluante et a prouvé la formation d'une hélice surenroulée comme attendu. La structure montre la formation d'hélice  $\alpha$  très stables entre les résidus 129 à 157 qui sont flanquées par des extrémités désordonnées. La structure de hRSV-P<sub>OD</sub> se superpose bien avec celle de hMPV-P<sub>core</sub> en dépit d'une longueur d'hélice plus importante et d'une légère distorsion C-terminale de hRSV-P<sub>OD</sub>.

Les 2 domaines N- et C-terminaux de P forment des structures secondaires assez courtes et transitoires en N-terminal ( $\alpha_{N1}$ ,  $\beta_N$ ,  $\alpha_{N2}$ ) et presque stables en C-terminal ( $\alpha_{C1}$ ,  $\alpha_{C2}$ ). Ces régions sont impliquées dans diverses interactions avec des protéines partenaires. Nous avons montré par des expériences de PRE que chacune de ces régions et P<sub>OD</sub> forment des contacts non permanents à longue distance à l'exception des extrémités de P. Ces contacts peuvent avoir une fonction d'écrantage contre des interactions non spécifiques, une fonction de protection des régions déstructurées contre des protéases cellulaires ou encore une fonction de rapprochement de partenaires de P entre eux, comme par exemple dans la déphosphorylation de M2-1 par PP1 $\alpha$  (Chapitre 5).

Au laboratoire nous avons déterminé ou confirmé que les régions  $\alpha_{N1}$ ,  $\beta_N$ ,  $\alpha_{N2}$  du domaine N-terminal de P étaient impliquées dans des interactions avec N<sub>0</sub>, PP1 $\alpha$  et M2-1 respectivement. Il a été montré très récemment que les régions  $\alpha_{C1}$  et  $\alpha_{C2}$  prédites en hélices  $\alpha$  (~90%) étaient impliquées dans l'interaction avec la grande sous unité catalytique L du RSV : la structure d'un complexe P—L a été obtenue par cryo-microscopie électronique (Gilman et al., 2019). La structure de P—L (**Figure 83**) a été reconstruite à partir de la structure de la protéine L du virus de stomatite vésiculaire (dont la séquence primaire est proche de celle de hRSV-L) comme structure initiale et en utilisant le domaine d'oligomérisation de hMPV à l'intérieur de la carte de microscopie électronique, puis en reconstruisant les parties manquantes « à la main ». Le complexe obtenu montre que des résidus entre 147 et 227, c'est-à-dire du milieu du domaine de tétramérisation de P jusqu'à avant la région de fixation de N en passant par les régions  $\alpha_{C1/2}$  sont impliqués dans l'interaction avec L. Des interactions électrostatiques, des ponts salins, des liaisons hydrogènes et des interactions hydrophobes semblent maintenir les 2 partenaires. La structure du complexe montre également que les 4 protomères de P adoptent des conformations différentes, mais bien définies. Cette structure explique l'intérêt d'avoir des hélices  $\alpha_{C1/2}$  pas complètement rigides et corrobore nos données de contacts à longue distance entre  $\alpha_{C1/2}$  et P<sub>OD</sub> puisqu'on observe une torsion entre les 3 régions des 4 protomères. En observant de plus près les 4 protomères de P, un contact transitoire entre  $\alpha_{C1}$  et  $\alpha_{C2}$  pourrait peut-être avoir également lieu au sein de la protéine P isolée : cela expliquerait que les prédictions de PRE faites avec flexible meccano ne se superposent pas complètement avec les données de PRE des 2 mutants de P[127-205]. Par ailleurs, une molécule au laboratoire semble avoir été identifiée comme ayant un effet de réarrangement des régions  $\alpha_{C1/2}$  (peut-

être une rigidification des hélices) qui pourrait être intéressant pour inhiber l'interaction L—P : il serait donc intéressant de comprendre l'effet de cette molécule par la suite.



**Figure 83 : Structure du complexe P—L du hRSV.** Le L apparaît en surface blanche et les monomères P en « cartoon ». Les 4 protomères sont colorés en arc-en-ciel (bleu à rouge correspond à N- vers C-terminus respectivement). Les images du haut montrent 2 orientations du complexe et en bas les 4 protomères sont représentés séparément. Figure tirée de (Gilman et al., 2019).

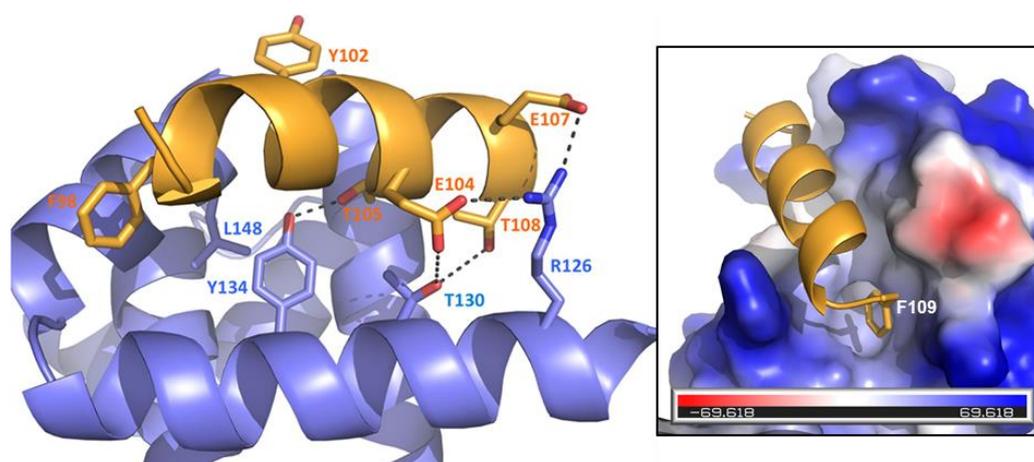
### 6.3. Interactions entre hRSV P et les protéines M2-1 et PP1 $\alpha$

Après avoir exploré les structures et la dynamique de la phosphoprotéine du RSV de manière isolée dans le Chapitre 4, j'ai présenté dans le Chapitre 5 une étude autour de l'interaction de la protéine P avec le cofacteur de transcription M2-1 du virus et avec la phosphatase cellulaire PP1 $\alpha$  ainsi que des aspects fonctionnels en découlant.

Notre travail a permis de prouver le recrutement de M2-1 par P au sein des IB qui est indispensable à l'activité du cofacteur de transcription au sein du complexe RdRp du RSV. Concernant la formation du complexe P—M2-1, nous avons identifié une région d'interaction minimale au niveau des résidus 93-110 où les résidus essentiels sont F98, L101, Y102, T105, I106, F109, dont certains résidus avaient

été identifiés précédemment (Mason et al., 2003). La phosphorylation de T108 de P a également été confirmée comme empêchant l'interaction P—M2-1 (Asenjo et al., 2006). En présence de M2-1, ces résidus distants en moyenne de 3 ou 4 positions forment très probablement un prolongement de l'hélice transitoire  $\alpha_{N2}$  initialement prédite pour les résidus 98 à 103 avec P isolée (Pereira et al., 2016 ; **Figure 77** et **Figure 78**). Il est également probable qu'une augmentation de propension d'hélice et un échange entre la forme libre et liée de P se fassent à une échelle de temps de l'ordre de  $\mu$ s-ms, ce qui expliquerait davantage la très forte diminution des intensités des signaux localement impliqués dans l'interaction étudiée par RMN (**Figure 75**).

Peu de temps après la sortie de notre article, une structure de P[90-110]—M2-1 obtenue par diffraction aux rayons X avec une résolution de 2,4 Å a été publiée (Selvaraj et al., 2018). Ce complexe (**Figure 84**) montre énormément de similitude avec le modèle obtenu par HADDOCK (**Figure 78**) à l'exception de quelques angles de torsion de chaînes latérales. On y voit la région 97-109 de P bien définie et la plupart des résidus semblent impliqués soit dans des interactions électrostatiques, hydrophobes ou des ponts salins à l'exception des résidus S99, K100 et K103. Les auteurs ont montré que les résidus R126 et L148 de M2-1 ainsi que le résidu F98 et Y102 de P semblent avoir le plus fort impact sur l'activité transcriptase du RSV. Néanmoins, le résidu F109 de P qui a montré un effet critique sur l'activité de transcription et sur la formation de complexe avec M2-1 dans Richard et al., 2018 semble clairement enfoui dans une poche hydrophobe de M2-1. Cette poche hydrophobe et la fente où se positionne l'hélice de P sont certainement des régions qui pourraient être une cible pharmacologique pour des peptidomimétiques optimisés qui rentreraient en compétition avec P sur M2-1 et qui pourraient devenir de potentiels antiviraux spécifiques contre le RSV. Cette dernière approche semble envisageable également du fait de la modeste affinité (de l'ordre de  $\sim 5 \mu$ M) de P pour M2-1 (Blondot et al., 2012).



**Figure 84** : Détails de l'interaction M2-1—P révélés par la structure cristallographique de M2-1—P[90–110] (pdb : 6G0Y). Panneau de gauche : Les structures de M2-1 et du peptide de P sont colorées en bleu et orange respectivement sur la structure en « cartoon ». Les résidus importants dans la surface

*d'interaction sont mis en évidence. Quelques potentielles interactions électrostatiques et ponts salins sont représentés par des tirets. Panneau de droite : La figure montre l'enfouissement du résidu hydrophobe F109 de P entre 2 hélices du domaine M2-1<sub>core</sub>, à l'intérieur d'une poche hydrophobe représentée en surface électrostatique. La figure de gauche est tirée de (Selvaraj et al., 2018).*

Nous avons également démontré que la déphosphorylation cyclique de M2-1, essentielle au fonctionnement de cette dernière, nécessite une interaction avec P (indépendante de l'état de phosphorylation de M2-1) qui fait intervenir la phosphatase PP1 $\alpha$ . Les régions de fixation de PP1 $\alpha$  et de M2-1 sur P sont situées sur les régions proximales  $\beta_N$  et  $\alpha_{N2}$  respectivement et il est très probable que les 2 partenaires forment un complexe ternaire avec P en se fixant sur différents protomères de la phosphoprotéine. Plus précisément, il a été montré que la phosphatase cellulaire PP1 $\alpha$  reconnaissait un motif de type « RVxF » dégénéré dans P, <sup>81</sup>RKPLVSF<sup>87</sup>, situé en amont du site de liaison M2-1 et dans lequel le résidu F87 de P est indispensable pour l'interaction avec PP1 $\alpha$ , son recrutement dans les IB cytoplasmiques et son activité sur M2-1. Nos données montrent que P peut être considérée comme un nouveau cofacteur de PP1 $\alpha$  pour la reconnaissance spécifique de substrat, en l'occurrence M2-1.

Parmi les *Mononegavirales* et en dehors des *Pneumoviridae*, la protéine VP30 des filovirus qui est la seule protéine homologue de M2-1. Elle est aussi déphosphorylée par PP1 (Ilinykh et al., 2014). Mais PP1 régule aussi la réponse immunitaire innée d'autres virus tels que le virus de la grippe ou un autre *Mononegavirales*, le virus de Sendai (Wies et al., 2013). Il a été démontré que la protéine V du virus de la rougeole interagissait directement avec PP1 $\alpha/\gamma$ , via un motif de liaison PP1 canonique « RIWY », empêchant la déphosphorylation de MDA5 par PP1, ce qui entraîne l'inhibition de l'activation des défenses immunitaires innées contre les virus à ARN par la voie MDA5 (Davis et al., 2014).

Il a été aussi démontré que la phosphorylation de M2-1 réduisait considérablement son affinité avec les ARN et impactait la polyadénylation des ARNm néo synthétisés par le RdRp ainsi que le recrutement de M2-1 dans les IBAG où se concentrent les nouveaux ARNm. L'affinité mesurée la plus élevée d'ARN avec M2-1 jusqu'alors est celle d'un oligomère de 13 adénosines avec un Kd d'environ 20 nM, c'est-à-dire presque 1000 fois plus affin que P. Juste après les séquences poly(A) en terme d'affinité pour M2-1, les séquences « gene end » d'ARNm positifs du RSV ont été titrés et un Kd de l'ordre de quelques dizaines à quelques centaines de nM en moyenne a été relevé. Les affinités des séquences d'ARNm positifs à l'extrémité 3' vis-à-vis de M2-1 sont donc supérieures à celle de P.

Lors de la découverte des IBAG (Rincheval et al., 2017) il a été indiqué que ce sont des structures dynamiques permettant le tri des ARNm viraux et leur transport vers le cytoplasme avec M2-1. L'ensemble de ces données a permis de proposer que la protéine P se lie à PP1 et à la forme phosphorylée de M2-1, permettant ainsi son recrutement dans les IB. M2-1 en plus de son activité d'anti-terminateur de transcription pourrait intervenir en tant que facteur de la transcriptase L pour la poly-adénylation. La déphosphorylation de M2-1 par PP1 pourrait permettre aux ARNm néo-synthétisés d'interagir avec M2-

l par les queues poly A et/ou les « gene end » afin de se regrouper dans les IBAG qui croîtraient jusqu'à libérer les ARNm attachés à M2-1 dans le cytosol où M2-1 pourrait peut-être jouer le rôle de transporteur. M2-1 pourrait aussi protéger les séquences poly(A) des ARNm de la dégradation. Enfin la phosphorylation de M2-1 entraînerait le détachement de M2-1 de la queue poly(A) de l'ARNm, avant que M2-1 soit recyclée par P et redirigée vers les IB pour un nouveau cycle de transcription.

## 6.4. Perspectives

Pour résumer, au vu des résultats que j'ai présentés, si je devais continuer mon projet de thèse, voici quelques axes que je développerais.

1. Etant donné le rôle essentiel de l'interaction entre P et PP1 $\alpha$  dans l'étape de transcription de l'ARN génomique du RSV, il me paraîtrait très judicieux d'essayer obtenir la structure du complexe de P—PP1 $\alpha$  avec un peptide qui correspond à la région de fixation minimale de P sur PP1 $\alpha$  en essayant de les co-cristalliser. Ensuite, l'idée serait de tenter une approche par petite molécule ou par peptidomimétique pour inhiber l'interaction entre les 2 partenaires en mimant par exemple le motif « RVxF » de P (si c'est une approche par peptidomimétique).
2. En utilisant la construction P[127-205], étudier la rigidification voire l'auto-association de la région  $\alpha_{C1/2}$  de P induite par la molécule testée en aveugle lors d'un criblage d'inhibiteurs de l'interaction N-P et éventuellement faire tester à l'INRA la molécule en minigénome du RSV ou sur sa réplication virale avec un virus recombinant.
3. Si des résultats sur les domaines d'interaction de Tax1BP1—N impliqués sont disponibles assez rapidement et que des expériences d'étude par RMN de l'interaction des nouveaux domaines sont envisageables, je pourrais effectuer un nouveau titrage des 2 interactants.
4. Tenter une approche par petite molécule ou peptidomimétique pour inhiber l'interaction M2-1—P[90-110] en ciblant la poche de F109 ou l'ensemble de la région de fixation de P sur M2-1 en utilisant la RMN puis la co-cristallisation si un composé ou un peptide rentre en compétition avec P pour M2-1 afin d'optimiser la molécule.



## Chapitre 7. Bibliographie

Amarasinghe, G.K., Ayllón, M.A., Bào, Y., Basler, C.F., Bavari, S., Blasdel, K.R., Briese, T., Brown, P.A., Bukreyev, A., Balkema-Buschmann, A., Buchholz, U.J., Chabi-Jesus, C., Chandran, K., Chiapponi, C., Crozier, I., de Swart, R.L., Dietzgen, R.G., Dolnik, O., Drexler, J.F., Dürrwald, R., Dundon, W.G., Duprex, W.P., Dye, J.M., Easton, A.J., Fooks, A.R., Formenty, P.B.H., Fouchier, R.A.M., Freitas-Astúa, J., Griffiths, A., Hewson, R., Horie, M., Hyndman, T.H., Jiāng, D., Kitajima, E.W., Kobinger, G.P., Kondō, H., Kurath, G., Kuzmin, I.V., Lamb, R.A., Lavazza, A., Lee, B., Lelli, D., Leroy, E.M., Lǐ, J., Maes, P., Marzano, S.-Y.L., Moreno, A., Mühlberger, E., Netesov, S.V., Nowotny, N., Nylund, A., Økland, A.L., Palacios, G., Pályi, B., Pawęska, J.T., Payne, S.L., Prosperi, A., Ramos-González, P.L., Rima, B.K., Rota, P., Rubbenstroth, D., Shī, M., Simmonds, P., Smither, S.J., Sozzi, E., Spann, K., Stenglein, M.D., Stone, D.M., Takada, A., Tesh, R.B., Tomonaga, K., Tordo, N., Towner, J.S., van den Hoogen, B., Vasilakis, N., Wahl, V., Walker, P.J., Wang, L.-F., Whitfield, A.E., Williams, J.V., Zerbini, F.M., Zhāng, T., Zhang, Y.-Z., Kuhn, J.H., 2019. Taxonomy of the order Mononegavirales: update 2019. *Arch. Virol.* 164, 1967–1980. <https://doi.org/10.1007/s00705-019-04247-4>

Asenjo, A., Calvo, E., Villanueva, N., 2006. Phosphorylation of human respiratory syncytial virus P protein at threonine 108 controls its interaction with the M2-1 protein in the viral RNA polymerase complex. *J. Gen. Virol.* 87, 3637–3642. <https://doi.org/10.1099/vir.0.82165-0>

Asenjo, A., Mendieta, J., Gómez-Puertas, P., Villanueva, N., 2008. Residues in human respiratory syncytial virus P protein that are essential for its activity on RNA viral synthesis. *Virus Res.* 132, 160–173. <https://doi.org/10.1016/j.virusres.2007.11.013>

Asenjo, A., Rodríguez, L., Villanueva, N., 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. *J. Gen. Virol.* 86, 1109–1120. <https://doi.org/10.1099/vir.0.80692-0>

Asenjo, A., Villanueva, N., 2016. Phosphorylation of the human respiratory syncytial virus P protein mediates M2-2 regulation of viral RNA synthesis, a process that involves two P proteins. *Virus Res.* 211, 117–125. <https://doi.org/10.1016/j.virusres.2015.10.011>

Asenjo, A., Villanueva, N., 2000. Regulated but not constitutive human respiratory syncytial virus (HRSV) P protein phosphorylation is essential for oligomerization. *FEBS Lett.* 467, 279–284. [https://doi.org/10.1016/s0014-5793\(00\)01171-6](https://doi.org/10.1016/s0014-5793(00)01171-6)

Bakker, S.E., Duquerroy, S., Galloux, M., Loney, C., Conner, E., Eléouët, J.-F., Rey, F.A., Bhella, D., 2013. The respiratory syncytial virus nucleoprotein-RNA complex forms a left-handed helical nucleocapsid. *J. Gen. Virol.* 94, 1734–1738. <https://doi.org/10.1099/vir.0.053025-0>

Berjanskii, M.V., Wishart, D.S., 2005. A Simple Method To Predict Protein Flexibility Using Secondary Chemical Shifts. *J. Am. Chem. Soc.* 127, 14970–14971. <https://doi.org/10.1021/ja054842f>

Bharat, T.A.M., Noda, T., Riches, J.D., Kraehling, V., Kolesnikova, L., Becker, S., Kawaoka, Y., Briggs, J.A.G., 2012. Structural dissection of Ebola virus and its assembly determinants using cryo-electron tomography. *Proc. Natl. Acad. Sci.* 109, 4275–4280. <https://doi.org/10.1073/pnas.1120453109>

Blondot, M.-L., Dubosclard, V., Fix, J., Lassoued, S., Aumont-Nicaise, M., Bontems, F., Eléouët, J.-F., Sizun, C., 2012. Structure and functional analysis of the RNA- and viral phosphoprotein-binding domain of respiratory syncytial virus M2-1 protein. *PLoS Pathog.* 8, e1002734. <https://doi.org/10.1371/journal.ppat.1002734>

Blount, R.E., Morris, J.A., Savage, R.E., 1956. Recovery of cytopathogenic agent from chimpanzees with coryza. *Proc. Soc. Exp. Biol. Med. Soc. Exp. Biol. Med. N. Y.* N 92, 544–549. <https://doi.org/10.3181/00379727-92-22538>

Bollen, M., Peti, W., Ragusa, M.J., Beullens, M., 2010. The extended PP1 toolkit: designed to create specificity. *Trends Biochem. Sci.* 35, 450–458. <https://doi.org/10.1016/j.tibs.2010.03.002>

Broadbent, L., Groves, H., Shields, M.D., Power, U.F., 2015. Respiratory syncytial virus, an ongoing medical dilemma: an expert commentary on respiratory syncytial virus prophylactic and therapeutic pharmaceuticals currently in clinical trials. *Influenza Other Respir. Viruses* 9, 169–178. <https://doi.org/10.1111/irv.12313>

Carroneu, C., Simabuco, F.M., Tamura, R.E., Farinha Arcieri, L.E., Ventura, A.M., 2007. Intracellular localization of human respiratory syncytial virus L protein. *Arch. Virol.* 152, 2259–2263. <https://doi.org/10.1007/s00705-007-1048-4>

Cartee, T.L., Wertz, G.W., 2001. Respiratory syncytial virus M2-1 protein requires phosphorylation for efficient function and binds viral RNA during infection. *J. Virol.* 75, 12188–12197. <https://doi.org/10.1128/JVI.75.24.12188-12197.2001>

Castagné, N., Barbier, A., Bernard, J., Rezaei, H., Huet, J.-C., Henry, C., Da Costa, B., Eléouët, J.-F., 2004. Biochemical characterization of the respiratory syncytial virus P-P and P-N protein complexes and localization of the P protein oligomerization domain. *J. Gen. Virol.* 85, 1643–1653. <https://doi.org/10.1099/vir.0.79830-0>

Ceregido, M.A., Spínola Amilibia, M., Buts, L., Rivera-Torres, J., Garcia-Pino, A., Bravo, J., van Nuland, N.A.J., 2014. The structure of TAX1BP1 UBZ1+2 provides insight into target specificity and adaptability. *J. Mol. Biol.* 426, 674–690. <https://doi.org/10.1016/j.jmb.2013.11.006>

Chanock, R., Finberg, L., 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. *Am. J. Hyg.* 66, 291–300.

Chanock, R., Roizman, B., Myers, R., 1957. Recovery from infants with respiratory illness of a virus related to chimpanzee coryza agent (CCA). I. Isolation, properties and characterization. *Am. J. Hyg.* 66, 281–290.

Choi, Y.B., Shembade, N., Parvatiyar, K., Balachandran, S., Harhaj, E.W., 2016. TAX1BP1 restrains virus-induced apoptosis by facilitating Itch-mediated degradation of the mitochondrial adaptor MAVS. *Mol. Cell. Biol.* <https://doi.org/10.1128/MCB.00422-16>

Choy, M.S., Hieke, M., Kumar, G.S., Lewis, G.R., Gonzalez-DeWhitt, K.R., Kessler, R.P., Stein, B.J., Hessenberger, M., Nairn, A.C., Peti, W., Page, R., 2014. Understanding the antagonism of retinoblastoma protein dephosphorylation by PNUITS provides insights into the PP1 regulatory code. *Proc. Natl. Acad. Sci. U. S. A.* 111, 4097–4102. <https://doi.org/10.1073/pnas.1317395111>

Choy, M.S., Page, R., Peti, W., 2012. Regulation of protein phosphatase 1 by intrinsically disordered proteins. *Biochem. Soc. Trans.* 40, 969–974. <https://doi.org/10.1042/BST20120094>

Collins, P.L., Hill, M.G., Camargo, E., Grosfeld, H., Chanock, R.M., Murphy, B.R., 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. *Proc. Natl. Acad. Sci. U. S. A.* 92, 11563–11567. <https://doi.org/10.1073/pnas.92.25.11563>

Collins, P.L., Melero, J.A., 2011. Progress in understanding and controlling respiratory syncytial virus: still crazy after all these years. *Virus Res.* 162, 80–99. <https://doi.org/10.1016/j.virusres.2011.09.020>

Cowton, V.M., Fearn, R., 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. *J. Virol.* 79, 11311–11322. <https://doi.org/10.1128/JVI.79.17.11311-11322.2005>

Cuesta, I., Geng, X., Asenjo, A., Villanueva, N., 2000. Structural phosphoprotein M2-1 of the human respiratory syncytial virus is an RNA binding protein. *J. Virol.* 74, 9858–9867. <https://doi.org/10.1128/jvi.74.21.9858-9867.2000>

Davis, M.E., Wang, M.K., Rennick, L.J., Full, F., Gableske, S., Mesman, A.W., Gringhuis, S.I., Geijtenbeek, T.B.H., Duprex, W.P., Gack, M.U., 2014. Antagonism of the phosphatase PP1 by the measles virus V protein is required for innate immune escape of MDA5. *Cell Host Microbe* 16, 19–30. <https://doi.org/10.1016/j.chom.2014.06.007>

de Vries, S.J., van Dijk, M., Bonvin, A.M.J.J., 2010. The HADDOCK web server for data-driven biomolecular docking. *Nat. Protoc.* 5, 883–897. <https://doi.org/10.1038/nprot.2010.32>

Esneau, C., Raynal, B., Roblin, P., Brûlé, S., Richard, C.-A., Fix, J., Eléouët, J.-F., Galloux, M., 2019. Biochemical characterization of the respiratory syncytial virus N0-P complex in solution. *J. Biol. Chem.* 294, 3647–3660. <https://doi.org/10.1074/jbc.RA118.006453>

Fearns, R., Collins, P.L., 1999. Role of the M2-1 transcription antitermination protein of respiratory syncytial virus in sequential transcription. *J. Virol.* 73, 5852–5864.

Friesen, W.J., Wyce, A., Paushkin, S., Abel, L., Rappsilber, J., Mann, M., Dreyfuss, G., 2002. A novel WD repeat protein component of the methylosome binds Sm proteins. *J. Biol. Chem.* 277, 8243–8247. <https://doi.org/10.1074/jbc.M109984200>

Galloux, M., Gabiane, G., Sourimant, J., Richard, C.-A., England, P., Moudjou, M., Aumont-Nicaise, M., Fix, J., Rameix-Welti, M.-A., Eléouët, J.-F., 2015. Identification and characterization of the binding site of the respiratory syncytial virus phosphoprotein to RNA-free nucleoprotein. *J. Virol.* 89, 3484–3496. <https://doi.org/10.1128/JVI.03666-14>

Galloux, M., Tarus, B., Blazevic, I., Fix, J., Duquerroy, S., Eléouët, J.-F., 2012. Characterization of a viral phosphoprotein binding site on the surface of the respiratory syncytial nucleoprotein. *J. Virol.* 86, 8375–8387. <https://doi.org/10.1128/JVI.00058-12>

García, J., García-Barreno, B., Vivo, A., Melero, J.A., 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, 243–247. <https://doi.org/10.1006/viro.1993.1366>

García-Barreno, B., Delgado, T., Melero, J.A., 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. *J. Virol.* 70, 801–808.

Gillespie, J.R., Shortle, D., 1997. Characterization of long-range structure in the denatured state of staphylococcal nuclease. I. Paramagnetic relaxation enhancement by nitroxide spin labels. *J. Mol. Biol.* 268, 158–169. <https://doi.org/10.1006/jmbi.1997.0954>

Gilman, M.S.A., Liu, C., Fung, A., Behera, I., Jordan, P., Rigaux, P., Ysebaert, N., Tcherniuk, S., Sourimant, J., Eléouët, J.-F., Sutto-Ortiz, P., Decroly, E., Roymans, D., Jin, Z., McLellan, J.S., 2019. Structure of the Respiratory Syncytial Virus Polymerase Complex. *Cell*. <https://doi.org/10.1016/j.cell.2019.08.014>

Glezen, W.P., Taber, L.H., Frank, A.L., Kasel, J.A., 1986. Risk of primary infection and reinfection with respiratory syncytial virus. *Am. J. Dis. Child.* 1960 140, 543–546. <https://doi.org/10.1001/archpedi.1986.02140200053026>

Goldberg, J., Huang, H., Kwon, Y., Greengard, P., Nairn, A.C., Kuriyan, J., 1995. Three-dimensional structure of the catalytic subunit of protein serine/threonine phosphatase-1. *Nature* 376, 745–753. <https://doi.org/10.1038/376745a0>

Güntert, P., Mumenthaler, C., Wüthrich, K., 1997. Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J. Mol. Biol.* 273, 283–298. <https://doi.org/10.1006/jmbi.1997.1284>

Hanley, L.L., McGivern, D.R., Teng, M.N., Djang, R., Collins, P.L., Fearn, R., 2010. Roles of the respiratory syncytial virus trailer region: Effects of mutations on genome production and stress granule formation. *Virology* 406, 241–252. <https://doi.org/10.1016/j.virol.2010.07.006>

Hardy, R.W., Wertz, G.W., 2000. The Cys(3)-His(1) motif of the respiratory syncytial virus M2-1 protein is essential for protein function. *J. Virol.* 74, 5880–5885. <https://doi.org/10.1128/jvi.74.13.5880-5885.2000>

Herrmann, T., Güntert, P., Wüthrich, K., 2002. Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *J. Mol. Biol.* 319, 209–227. [https://doi.org/10.1016/s0022-2836\(02\)00241-3](https://doi.org/10.1016/s0022-2836(02)00241-3)

Heylen, E., Neyts, J., Jochmans, D., 2017. Drug candidates and model systems in respiratory syncytial virus antiviral drug discovery. *Biochem. Pharmacol.* 127, 1–12. <https://doi.org/10.1016/j.bcp.2016.09.014>

Honda, K., Takaoka, A., Taniguchi, T., 2006. Type I interferon [corrected] gene induction by the interferon regulatory factor family of transcription factors. *Immunity* 25, 349–360. <https://doi.org/10.1016/j.immuni.2006.08.009>

Ilinykh, P.A., Tigabu, B., Ivanov, A., Ammosova, T., Obukhov, Y., Garron, T., Kumari, N., Kovalskyy, D., Platonov, M.O., Naumchik, V.S., Freiberg, A.N., Nekhai, S., Bukreyev, A., 2014. Role of protein phosphatase 1 in dephosphorylation of Ebola virus VP30 protein and its targeting for the

inhibition of viral transcription. *J. Biol. Chem.* 289, 22723–22738.  
<https://doi.org/10.1074/jbc.M114.575050>

Janco, M., Kalyva, A., Scellini, B., Piroddi, N., Tesi, C., Poggesi, C., Geeves, M.A., 2012.  $\alpha$ -Tropomyosin with a D175N or E180G mutation in only one chain differs from tropomyosin with mutations in both chains. *Biochemistry* 51, 9880–9890. <https://doi.org/10.1021/bi301323n>

Jensen, M.R., Communie, G., Ribeiro, E.A., Martinez, N., Desfosses, A., Salmon, L., Mollica, L., Gabel, F., Jamin, M., Longhi, S., Ruigrok, R.W.H., Blackledge, M., 2011. Intrinsic disorder in measles virus nucleocapsids. *Proc. Natl. Acad. Sci. U. S. A.* 108, 9839–9844.  
<https://doi.org/10.1073/pnas.1103270108>

Kay, L.E., Ikura, M., Tschudin, R., Bax, A., 1990. Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins. *J. Magn. Reson.* 1969 89, 496–514.  
[https://doi.org/10.1016/0022-2364\(90\)90333-5](https://doi.org/10.1016/0022-2364(90)90333-5)

Kiss, G., Holl, J.M., Williams, G.M., Alonas, E., Vanover, D., Lifland, A.W., Gudheti, M., Guerrero-Ferreira, R.C., Nair, V., Yi, H., Graham, B.S., Santangelo, P.J., Wright, E.R., 2014. Structural analysis of respiratory syncytial virus reveals the position of M2-1 between the matrix protein and the ribonucleoprotein complex. *J. Virol.* 88, 7602–7617. <https://doi.org/10.1128/JVI.00256-14>

Kombe, I.K., Munywoki, P.K., Baguelin, M., Nokes, D.J., Medley, G.F., 2019. Model-based estimates of transmission of respiratory syncytial virus within households. *Epidemics* 27, 1–11.  
<https://doi.org/10.1016/j.epidem.2018.12.001>

Lambert, D.M., Hambor, J., Diebold, M., Galinski, B., 1988. Kinetics of synthesis and phosphorylation of respiratory syncytial virus polypeptides. *J. Gen. Virol.* 69 ( Pt 2), 313–323.  
<https://doi.org/10.1099/0022-1317-69-2-313>

Leyrat, C., Renner, M., Harlos, K., Grimes, J.M., 2013. Solution and crystallographic structures of the central region of the phosphoprotein from human metapneumovirus. *PLoS One* 8, e80371.  
<https://doi.org/10.1371/journal.pone.0080371>

Lietzow, M.A., Jamin, M., Dyson, H.J., Wright, P.E., 2002. Mapping long-range contacts in a highly unfolded protein. *J. Mol. Biol.* 322, 655–662. [https://doi.org/10.1016/s0022-2836\(02\)00847-1](https://doi.org/10.1016/s0022-2836(02)00847-1)

Lifland, A.W., Jung, J., Alonas, E., Zurla, C., Crowe, J.E., Santangelo, P.J., 2012. Human respiratory syncytial virus nucleoprotein and inclusion bodies antagonize the innate immune response mediated by MDA5 and MAVS. *J. Virol.* 86, 8245–8258. <https://doi.org/10.1128/JVI.00215-12>

Liljeroos, L., Krzyzaniak, M.A., Helenius, A., Butcher, S.J., 2013. Architecture of respiratory syncytial virus revealed by electron cryotomography. *Proc. Natl. Acad. Sci. U. S. A.* 110, 11133–11138. <https://doi.org/10.1073/pnas.1309070110>

Ling, L., Goeddel, D.V., 2000. TRAF6, a TRAF6-interacting protein involved in IL-1 signaling. *Proc. Natl. Acad. Sci. U. S. A.* 97, 9567–9572. <https://doi.org/10.1073/pnas.170279097>

Llorente, M.T., García-Barreno, B., Calero, M., Camafeita, E., López, J.A., Longhi, S., Ferrón, F., Varela, P.F., Melero, J.A., 2006. Structural analysis of the human respiratory syncytial virus phosphoprotein: characterization of an alpha-helical domain involved in oligomerization. *J. Gen. Virol.* 87, 159–169. <https://doi.org/10.1099/vir.0.81430-0>

Llorente, M.T., Taylor, I.A., López-Viñas, E., Gomez-Puertas, P., Calder, L.J., García-Barreno, B., Melero, J.A., 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, 946–958. <https://doi.org/10.1002/prot.21988>

Madeira, F., Park, Y.M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A.R.N., Potter, S.C., Finn, R.D., Lopez, R., 2019. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* 47, W636–W641. <https://doi.org/10.1093/nar/gkz268>

Mason, S.W., Aberg, E., Lawetz, C., DeLong, R., Whitehead, P., Liuzzi, M., 2003. Interaction between human respiratory syncytial virus (RSV) M2-1 and P proteins is required for reconstitution of M2-1-dependent RSV minigenome activity. *J. Virol.* 77, 10670–10676.

Mink, M.A., Stec, D.S., Collins, P.L., 1991. Nucleotide sequences of the 3' leader and 5' trailer regions of human respiratory syncytial virus genomic RNA. *Virology* 185, 615–624. [https://doi.org/10.1016/0042-6822\(91\)90532-g](https://doi.org/10.1016/0042-6822(91)90532-g)

Navarro, J., López-Otín, C., Villanueva, N., 1991. Location of phosphorylated residues in human respiratory syncytial virus phosphoprotein. *J. Gen. Virol.* 72 ( Pt 6), 1455–1459. <https://doi.org/10.1099/0022-1317-72-6-1455>

Noval, M.G., Esperante, S.A., Molina, I.G., Chemes, L.B., Prat-Gay, G. de, 2016. Intrinsic Disorder to Order Transitions in the Scaffold Phosphoprotein P from the Respiratory Syncytial Virus RNA Polymerase Complex. *Biochemistry* 55, 1441–1454. <https://doi.org/10.1021/acs.biochem.5b01332>

Obando-Pacheco, P., Justicia-Grande, A.J., Rivero-Calle, I., Rodríguez-Tenreiro, C., Sly, P., Ramilo, O., Mejías, A., Baraldi, E., Papadopoulos, N.G., Nair, H., Nunes, M.C., Kragten-Tabatabaie, L., Heikkinen, T., Greenough, A., Stein, R.T., Manzoni, P., Bont, L., Martín-Torres, F., 2018.

Respiratory Syncytial Virus Seasonality: A Global Overview. *J. Infect. Dis.* 217, 1356–1364. <https://doi.org/10.1093/infdis/jiy056>

Oliveira, A.P., Simabuco, F.M., Tamura, R.E., Guerrero, M.C., Ribeiro, P.G.G., Libermann, T.A., Zerbini, L.F., Ventura, A.M., 2013. Human respiratory syncytial virus N, P and M protein interactions in HEK-293T cells. *Virus Res.* 177, 108–112. <https://doi.org/10.1016/j.virusres.2013.07.010>

Otomo, C., Metlagel, Z., Takaesu, G., Otomo, T., 2013. Structure of the human ATG12~ATG5 conjugate required for LC3 lipidation in autophagy. *Nat. Struct. Mol. Biol.* 20, 59–66. <https://doi.org/10.1038/nsmb.2431>

Quizougun-Oubari, M., Pereira, N., Tarus, B., Galloux, M., Lassoued, S., Fix, J., Tortorici, M.A., Hoos, S., Baron, B., England, P., Desmaële, D., Couvreur, P., Bontems, F., Rey, F.A., Eléouët, J.-F., Sizun, C., Slama-Schwok, A., Duquerroy, S., 2015. A Druggable Pocket at the Nucleocapsid/Phosphoprotein Interaction Site of Human Respiratory Syncytial Virus. *J. Virol.* 89, 11129–11143. <https://doi.org/10.1128/JVI.01612-15>

Ozenne, V., Bauer, F., Salmon, L., Huang, J.-R., Jensen, M.R., Segard, S., Bernadó, P., Charavay, C., Blackledge, M., 2012. Flexible-meccano: a tool for the generation of explicit ensemble descriptions of intrinsically disordered proteins and their associated experimental observables. *Bioinforma. Oxf. Engl.* 28, 1463–1470. <https://doi.org/10.1093/bioinformatics/bts172>

Pereira, N., Cardone, C., Lassoued, S., Galloux, M., Fix, J., Assrir, N., Lescop, E., Bontems, F., Eléouët, J.-F., Sizun, C., 2016. New Insights into Structural Disorder in Human Respiratory Syncytial Virus Phosphoprotein and Implications for Binding of Protein Partners. *J. Biol. Chem.* jbc.M116.765958. <https://doi.org/10.1074/jbc.M116.765958>

Renner, M., Bertinelli, M., Leyrat, C., Paesen, G.C., Saraiva de Oliveira, L.F., Huiskonen, J.T., Grimes, J.M., 2016. Nucleocapsid assembly in pneumoviruses is regulated by conformational switching of the N protein. *eLife* 5, e12627. <https://doi.org/10.7554/eLife.12627>

Renner, M., Paesen, G.C., Grison, C.M., Granier, S., Grimes, J.M., Leyrat, C., 2017. Structural dissection of human metapneumovirus phosphoprotein using small angle x-ray scattering. *Sci. Rep.* 7. <https://doi.org/10.1038/s41598-017-14448-z>

Richard, C.-A., Rincheval, V., Lassoued, S., Fix, J., Cardone, C., Esneau, C., Nekhai, S., Galloux, M., Rameix-Welti, M.-A., Sizun, C., Eléouët, J.-F., 2018. RSV hijacks cellular protein phosphatase 1 to regulate M2-1 phosphorylation and viral transcription. *PLoS Pathog.* 14, e1006920. <https://doi.org/10.1371/journal.ppat.1006920>

Rieping, W., Habeck, M., Bardiaux, B., Bernard, A., Malliavin, T.E., Nilges, M., 2007. ARIA2: automated NOE assignment and data integration in NMR structure calculation. *Bioinforma. Oxf. Engl.* 23, 381–382. <https://doi.org/10.1093/bioinformatics/btl589>

Rima, B., Collins, P., Easton, A., Fouchier, R., Kurath, G., Lamb, R.A., Lee, B., Maisner, A., Rota, P., Wang, L., ICTV Report Consortium, 2017. ICTV Virus Taxonomy Profile: Pneumoviridae. *J. Gen. Virol.* 98, 2912–2913. <https://doi.org/10.1099/jgv.0.000959>

Rincheval, V., Lelek, M., Gault, E., Bouillier, C., Sitterlin, D., Blouquit-Laye, S., Galloux, M., Zimmer, C., Eleouet, J.-F., Rameix-Welti, M.-A., 2017. Functional organization of cytoplasmic inclusion bodies in cells infected by respiratory syncytial virus. *Nat. Commun.* 8, 563. <https://doi.org/10.1038/s41467-017-00655-9>

Robert, X., Gouet, P., 2014. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* 42, W320–W324. <https://doi.org/10.1093/nar/gku316>

Rostad, C.A., 2019. Respiratory Syncytial Virus: Spectrum of Clinical Manifestations and Complications in Children. *Pediatr. Ann.* 48, e349–e353. <https://doi.org/10.3928/19382359-20190815-01>

Routledge, E.G., Willcocks, M.M., Morgan, L., Samson, A.C., Scott, R., Toms, G.L., 1987. Heterogeneity of the respiratory syncytial virus 22K protein revealed by Western blotting with monoclonal antibodies. *J. Gen. Virol.* 68 ( Pt 4), 1209–1215. <https://doi.org/10.1099/0022-1317-68-4-1209>

Saitô, H., 1986. Conformation-dependent <sup>13</sup>C chemical shifts: A new means of conformational characterization as obtained by high-resolution solid-state <sup>13</sup>C NMR. *Magn. Reson. Chem.* 24, 835–852. <https://doi.org/10.1002/mrc.1260241002>

Selvaraj, M., Yegambaram, K., Todd, E.J.A.A., Richard, C.-A., Dods, R.L., Pangratiou, G.M., Trinh, C.H., Moul, S.L., Murphy, J.C., Mankouri, J., Éléouët, J.-F., Barr, J.N., Edwards, T.A., 2018. The Structure of the Human Respiratory Syncytial Virus M2-1 Protein Bound to the Interaction Domain of the Phosphoprotein P Defines the Orientation of the Complex. *mBio* 9. <https://doi.org/10.1128/mBio.01554-18>

Shembade, N., Pujari, R., Harhaj, N.S., Abbott, D.W., Harhaj, E.W., 2011. The kinase IKK $\alpha$  inhibits activation of the transcription factor NF- $\kappa$ B by phosphorylating the regulatory molecule TAX1BP1. *Nat. Immunol.* 12, 834–843. <https://doi.org/10.1038/ni.2066>

Shen, Y., Bax, A., 2013. Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J. Biomol. NMR* 56, 227–241. <https://doi.org/10.1007/s10858-013-9741-y>

Shen, Y., Delaglio, F., Cornilescu, G., Bax, A., 2009. TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR* 44, 213–223. <https://doi.org/10.1007/s10858-009-9333-z>

Shi, T., McAllister, D.A., O'Brien, K.L., Simoes, E.A.F., Madhi, S.A., Gessner, B.D., Polack, F.P., Balsells, E., Acacio, S., Aguayo, C., Alassani, I., Ali, A., Antonio, M., Awasthi, S., Awori, J.O., Azziz-Baumgartner, E., Baggett, H.C., Baillie, V.L., Balmaseda, A., Barahona, A., Basnet, S., Bassat, Q., Basualdo, W., Bigogo, G., Bont, L., Breiman, R.F., Brooks, W.A., Broor, S., Bruce, N., Bruden, D., Buchy, P., Campbell, S., Carosone-Link, P., Chadha, M., Chipeta, J., Chou, M., Clara, W., Cohen, C., de Cuellar, E., Dang, D.-A., Dash-Yandag, B., Deloria-Knoll, M., Dherani, M., Eap, T., Ebruke, B.E., Echavarria, M., de Freitas Lázaro Emediato, C.C., Fasce, R.A., Feikin, D.R., Feng, L., Gentile, A., Gordon, A., Goswami, D., Goyet, S., Groome, M., Halasa, N., Hirve, S., Homaira, N., Howie, S.R.C., Jara, J., Jroundi, I., Kartasasmita, C.B., Khuri-Bulos, N., Kotloff, K.L., Krishnan, A., Libster, R., Lopez, O., Lucero, M.G., Lucion, F., Lupisan, S.P., Marcone, D.N., McCracken, J.P., Mejia, M., Moisi, J.C., Montgomery, J.M., Moore, D.P., Moraleta, C., Moyes, J., Munywoki, P., Mutyara, K., Nicol, M.P., Nokes, D.J., Nymadawa, P., da Costa Oliveira, M.T., Oshitani, H., Pandey, N., Paranhos-Baccalà, G., Phillips, L.N., Picot, V.S., Rahman, M., Rakoto-Andrianarivelo, M., Rasmussen, Z.A., Rath, B.A., Robinson, A., Romero, C., Russomando, G., Salimi, V., Sawatwong, P., Scheltema, N., Schweiger, B., Scott, J.A.G., Seidenberg, P., Shen, K., Singleton, R., Sotomayor, V., Strand, T.A., Sutanto, A., Sylla, M., Tapia, M.D., Thamthitawat, S., Thomas, E.D., Tokarz, R., Turner, C., Venter, M., Waicharoen, S., Wang, J., Watthanaworawit, W., Yoshida, L.-M., Yu, H., Zar, H.J., Campbell, H., Nair, H., RSV Global Epidemiology Network, 2017. Global, regional, and national disease burden estimates of acute lower respiratory infections due to respiratory syncytial virus in young children in 2015: a systematic review and modelling study. *Lancet Lond. Engl.* 390, 946–958. [https://doi.org/10.1016/S0140-6736\(17\)30938-8](https://doi.org/10.1016/S0140-6736(17)30938-8)

Simabuco, F.M., Asara, J.M., Guerrero, M.C., Libermann, T.A., Zerbini, L.F., Ventura, A.M., 2011. Structural analysis of human respiratory syncytial virus P protein: identification of intrinsically disordered domains. *Braz. J. Microbiol.* 42, 340–345. <https://doi.org/10.1590/S1517-83822011000100043>

Slack, M.S., Easton, A.J., 1998. Characterization of the interaction of the human respiratory syncytial virus phosphoprotein and nucleocapsid protein using the two-hybrid system. *Virus Res.* 55, 167–176. [https://doi.org/10.1016/s0168-1702\(98\)00042-2](https://doi.org/10.1016/s0168-1702(98)00042-2)

Spera, S., Bax, A., 1991. Empirical correlation between protein backbone conformation and C.alpha. and C.beta. <sup>13</sup>C nuclear magnetic resonance chemical shifts. *J. Am. Chem. Soc.* 113, 5490–5492. <https://doi.org/10.1021/ja00014a071>

Tanida, I., Ueno, T., Kominami, E., 2008. LC3 and Autophagy. *Methods Mol. Biol.* Clifton NJ 445, 77–88. [https://doi.org/10.1007/978-1-59745-157-4\\_4](https://doi.org/10.1007/978-1-59745-157-4_4)

Tanner, S.J., Ariza, A., Richard, C.-A., Kyle, H.F., Dods, R.L., Blondot, M.-L., Wu, W., Trincão, J., Trinh, C.H., Hiscox, J.A., Carroll, M.W., Silman, N.J., Eléouët, J.-F., Edwards, T.A., Barr, J.N., 2014. Crystal structure of the essential transcription antiterminator M2-1 protein of human respiratory syncytial virus and implications of its phosphorylation. *Proc. Natl. Acad. Sci. U. S. A.* 111, 1580–1585. <https://doi.org/10.1073/pnas.1317262111>

Tawar, R.G., Duquerroy, S., Vornrhein, C., Varela, P.F., Damier-Piolle, L., Castagné, N., MacLellan, K., Bedouelle, H., Bricogne, G., Bhella, D., Eléouët, J.-F., Rey, F.A., 2009. Crystal structure of a nucleocapsid-like nucleoprotein-RNA complex of respiratory syncytial virus. *Science* 326, 1279–1283. <https://doi.org/10.1126/science.1177634>

Tayyari, F., Marchant, D., Moraes, T.J., Duan, W., Mastrangelo, P., Hegele, R.G., 2011. Identification of nucleolin as a cellular receptor for human respiratory syncytial virus. *Nat. Med.* 17, 1132–1135. <https://doi.org/10.1038/nm.2444>

Tran, T.-L., Castagné, N., Bhella, D., Varela, P.F., Bernard, J., Chilmonczyk, S., Berkenkamp, S., Benhamo, V., Grznarova, K., Grosclaude, J., Nespoulos, C., Rey, F.A., Eléouët, J.-F., 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. *J. Gen. Virol.* 88, 196–206. <https://doi.org/10.1099/vir.0.82282-0>

Tran, T.-L., Castagné, N., Dubosclard, V., Noinville, S., Koch, E., Moudjou, M., Henry, C., Bernard, J., Yeo, R.P., Eléouët, J.-F., 2009. The respiratory syncytial virus M2-1 protein forms tetramers and interacts with RNA and P in a competitive manner. *J. Virol.* 83, 6363–6374. <https://doi.org/10.1128/JVI.00335-09>

Tumbarello, D.A., Manna, P.T., Allen, M., Bycroft, M., Arden, S.D., Kendrick-Jones, J., Buss, F., 2015. The Autophagy Receptor TAX1BP1 and the Molecular Motor Myosin VI Are Required for Clearance of *Salmonella Typhimurium* by Autophagy. *PLOS Pathog.* 11, e1005174. <https://doi.org/10.1371/journal.ppat.1005174>

Verstrepen, L., Verhelst, K., Carpentier, I., Beyaert, R., 2011. TAX1BP1, a ubiquitin-binding adaptor protein in innate immunity and beyond. *Trends Biochem. Sci.* 36, 347–354. <https://doi.org/10.1016/j.tibs.2011.03.004>

Villanueva, N., Hardy, R., Asenjo, A., Yu, Q., Wertz, G., 2000. The bulk of the phosphorylation of human respiratory syncytial virus phosphoprotein is not essential but modulates viral RNA transcription and replication. *J. Gen. Virol.* 81, 129–133. <https://doi.org/10.1099/0022-1317-81-1-129>

Vranken, W.F., Boucher, W., Stevens, T.J., Fogh, R.H., Pajon, A., Llinas, M., Ulrich, E.L., Markley, J.L., Ionides, J., Laue, E.D., 2005. The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* 59, 687–696. <https://doi.org/10.1002/prot.20449>

Wassenaar, T.A., van Dijk, M., Loureiro-Ferreira, N., van der Schot, G., de Vries, S.J., Schmitz, C., van der Zwan, J., Boelens, R., Giachetti, A., Ferella, L., Rosato, A., Bertini, I., Herrmann, T., Jonker, H.R.A., Bagaria, A., Jaravine, V., Güntert, P., Schwalbe, H., Vranken, W.F., Doreleijers, J.F., Vriend, G., Vuister, G.W., Franke, D., Kikhney, A., Svergun, D.I., Fogh, R.H., Ionides, J., Laue, E.D., Spronk, C., Jurkša, S., Verlato, M., Badoer, S., Dal Pra, S., Mazzucato, M., Frizziero, E., Bonvin, A.M.J.J., 2012. WeNMR: Structural Biology on the Grid. *J. Grid Comput.* 10, 743–767. <https://doi.org/10.1007/s10723-012-9246-z>

Wies, E., Wang, M.K., Maharaj, N.P., Chen, K., Zhou, S., Finberg, R.W., Gack, M.U., 2013. Dephosphorylation of the RNA sensors RIG-I and MDA5 by the phosphatase PP1 is essential for innate immune signaling. *Immunity* 38, 437–449. <https://doi.org/10.1016/j.immuni.2012.11.018>

Wishart, D.S., Sykes, B.D., Richards, F.M., 1991. Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J. Mol. Biol.* 222, 311–333.

Yang, Y., Wang, G., Huang, X., Du, Z., 2015. Crystallographic and modelling studies suggest that the SKICH domains from different protein families share a common Ig-like fold but harbour substantial structural variations. *J. Biomol. Struct. Dyn.* 33, 1385–1398. <https://doi.org/10.1080/07391102.2014.951688>

## Chapitre 8. Annexes

### 8.1. Déplacements chimiques attribués de N(C arm)

Les déplacements chimiques attribués qui suivent correspondent à ceux fait avec le peptide N(C arm) concentré à 1 mM dans un tampon 20 mM phosphate, 100 mM NaCl, 1 mM TCEP et pH 6,5.

Résidu	Atome attribué	Déplacement chimique	Ecart type (ppm)
360Ser	Cb	63,86	0,02
360Ser	Ca	58,33	0,02
360Ser	Hba	3,90	0,00
360Ser	Hbb	3,90	0,00
360Ser	Ha	4,53	0,01
360Ser	C	174,87	0,00
361Gly	Haa	3,98	0,01
361Gly	Hab	3,98	0,01
361Gly	Ca	45,19	0,02
361Gly	C	173,77	0,00
361Gly	H	8,56	0,01
361Gly	N	111,16	0,02
362Val	Hga*	0,89	0,00
362Val	Hgb*	0,90	0,00
362Val	Hb	2,02	0,01
362Val	Ha	4,10	0,01
362Val	Cb	32,69	0,05
362Val	Ca	62,23	0,02
362Val	C	176,02	0,00
362Val	H	8,01	0,01
362Val	N	119,66	0,03
363Ile	Hd1*	0,63	0,00
363Ile	Hg2*	0,79	0,01
363Ile	Hg1a	1,08	0,00
363Ile	Hg1b	1,41	0,00
363Ile	Hb	1,69	0,01
363Ile	Ha	4,06	0,01
363Ile	Cb	38,69	0,01
363Ile	Ca	60,81	0,02
363Ile	C	175,28	0,00
363Ile	H	8,22	0,00
363Ile	N	125,18	0,05
364Asn	Hba	2,67	0,01
364Asn	Hbb	2,79	0,01
364Asn	Ha	4,72	0,00
364Asn	Cb	38,87	0,03
364Asn	Ca	52,51	0,04
364Asn	C	174,62	0,00
364Asn	H	8,36	0,00
364Asn	N	123,18	0,02
365Tyr	Hd*	7,09	0,00
365Tyr	He*	6,82	0,00
365Tyr	Hba	2,90	0,01
365Tyr	Hbb	3,07	0,00
365Tyr	Ha	4,50	0,00

Résidu	Atome attribué	Déplacement chimique	Ecart type (ppm)
365Tyr	Cb	38,72	0,05
365Tyr	C	175,75	0,00
365Tyr	Ca	58,31	0,07
365Tyr	H	8,10	0,00
365Tyr	N	121,22	0,01
366Ser	Ca	58,33	0,06
366Ser	Hba	3,85	0,01
366Ser	Hbb	3,85	0,01
366Ser	Ha	4,47	0,00
366Ser	Cb	63,80	0,01
366Ser	C	174,72	0,00
366Ser	N	116,74	0,01
366Ser	H	8,30	0,00
367Val	Hga*	0,94	0,00
367Val	Hgb*	0,94	0,00
367Val	Hb	2,18	0,00
367Val	Ha	4,16	0,00
367Val	Cb	32,18	0,04
367Val	Ca	62,44	0,02
367Val	C	176,07	0,00
367Val	H	7,96	0,00
367Val	N	120,19	0,03
368Leu	Hg	1,60	0,00
368Leu	Hda*	0,89	0,00
368Leu	Hdb*	0,89	0,00
368Leu	Ca	55,25	0,05
368Leu	Cb	42,35	0,05
368Leu	Hba	1,56	0,01
368Leu	Hbb	1,64	0,01
368Leu	Ha	4,32	0,00
368Leu	C	176,83	0,00
368Leu	H	7,97	0,00
368Leu	N	123,22	0,03
369Asp	Cb	40,90	0,04
369Asp	Hba	2,57	0,00
369Asp	Hbb	2,77	0,01
369Asp	Ha	4,64	0,01
369Asp	Ca	54,08	0,07
369Asp	C	175,76	0,00
369Asp	H	8,19	0,01
369Asp	N	120,40	0,02
370Leu	Hg	1,68	0,00
370Leu	Hda*	0,85	0,00
370Leu	Hdb*	0,86	0,01
370Leu	Hba	1,58	0,01
370Leu	Hbb	1,65	0,01
370Leu	Ha	4,56	0,01
370Leu	Cb	43,26	0,00
370Leu	Ca	54,76	0,03
370Leu	C	177,89	0,00
370Leu	H	7,81	0,01
370Leu	N	122,01	0,02
371Thr	Hg2*	1,34	0,00
371Thr	Ha	4,36	0,00
371Thr	Hb	4,63	0,01
371Thr	Ca	61,40	0,09
371Thr	Cb	70,97	0,03

Résidu	Atome attribué	Déplacement chimique	Ecart type (ppm)
371Thr	H	8,75	0,01
371Thr	N	113,48	0,01
371Thr	C	175,14	0,00
372Ala	Hb*	1,46	0,00
372Ala	Ha	4,11	0,01
372Ala	Ca	55,09	0,02
372Ala	Cb	18,16	0,02
372Ala	C	180,20	0,00
372Ala	H	8,71	0,00
372Ala	N	123,53	0,01
373Glu	Hga	2,34	0,00
373Glu	Hgb	2,34	0,00
373Glu	Hba	2,02	0,01
373Glu	Hbb	2,03	0,01
373Glu	Ha	4,08	0,00
373Glu	Ca	59,46	0,02
373Glu	Cb	29,32	0,01
373Glu	C	179,22	0,00
373Glu	H	8,67	0,01
373Glu	N	117,79	0,02
374Glu	Hga	2,33	0,01
374Glu	Hgb	2,33	0,00
374Glu	Hba	1,99	0,01
374Glu	Hbb	2,28	0,00
374Glu	Ha	4,08	0,00
374Glu	Ca	58,64	0,04
374Glu	Cb	30,34	0,06
374Glu	C	178,92	0,00
374Glu	N	121,59	0,02
374Glu	H	7,90	0,00
375Leu	Hg	1,60	0,00
375Leu	Hda*	0,87	0,01
375Leu	Hdb*	0,87	0,01
375Leu	Hba	1,58	0,01
375Leu	Hbb	1,77	0,01
375Leu	Ha	4,08	0,00
375Leu	Ca	57,38	0,04
375Leu	Cb	41,73	0,05
375Leu	H	8,37	0,01
375Leu	N	120,71	0,03
375Leu	C	179,11	0,00
376Glu	Hga	2,31	0,01
376Glu	Hgb	2,31	0,01
376Glu	Hbb	2,07	0,01
376Glu	Hba	2,07	0,01
376Glu	Ha	4,04	0,01
376Glu	Cb	29,37	0,07
376Glu	Ca	58,50	0,03
376Glu	C	178,00	0,00
376Glu	H	8,02	0,01
376Glu	N	119,31	0,07
377Ala	Hb*	1,50	0,01
377Ala	Ha	4,19	0,00
377Ala	Cb	18,21	0,03
377Ala	Ca	54,47	0,03
377Ala	C	180,08	0,00
377Ala	H	7,69	0,01

Résidu	Atome attribué	Déplacement chimique	Ecart type (ppm)
377Ala	N	121,17	0,02
378Ile	Cb	38,38	0,02
378Ile	Ca	63,34	0,02
378Ile	Hg2*	0,86	0,00
378Ile	Hd1*	0,86	0,00
378Ile	Hg1a	1,19	0,00
378Ile	Hg1b	1,60	0,01
378Ile	Hb	1,91	0,01
378Ile	Ha	3,93	0,00
378Ile	C	177,70	0,00
378Ile	H	7,71	0,00
378Ile	N	117,73	0,02
379Lys	Cb	32,38	0,01
379Lys	Ca	58,32	0,05
379Lys	Hga	1,34	0,01
379Lys	Hgb	1,46	0,00
379Lys	Hda	1,61	0,01
379Lys	Hdb	1,61	0,01
379Lys	Hea	2,90	0,00
379Lys	Heb	2,90	0,00
379Lys	Hba	1,78	0,00
379Lys	Hbb	1,78	0,00
379Lys	Ha	4,04	0,00
379Lys	C	177,74	0,00
379Lys	H	7,86	0,00
379Lys	N	121,04	0,02
380His	Cb	29,29	0,01
380His	Ca	56,41	0,02
380His	Hbb	3,32	0,00
380His	Hba	3,16	0,01
380His	Ha	4,61	0,00
380His	C	175,20	0,00
380His	H	7,98	0,00
380His	N	116,66	0,02
381Gln	Ca	56,43	0,01
381Gln	Cb	29,13	0,05
381Gln	Hga	2,39	0,01
381Gln	Hgb	2,39	0,01
381Gln	Hba	2,11	0,01
381Gln	Hbb	2,11	0,01
381Gln	Ha	4,26	0,00
381Gln	C	176,11	0,00
381Gln	H	7,94	0,00
381Gln	N	119,50	0,02
382Leu	Cb	42,44	0,02
382Leu	Ca	55,31	0,03
382Leu	Hg	1,68	0,00
382Leu	Hda*	0,89	0,00
382Leu	Hdb*	0,89	0,00
382Leu	Hbb	1,67	0,00
382Leu	Hba	1,56	0,01
382Leu	Ha	4,31	0,01
382Leu	C	176,81	0,00
382Leu	N	121,38	0,01
382Leu	H	8,07	0,00
383Asn	Cb	38,96	0,03
383Asn	Ca	51,22	0,11

Résidu	Atome attribué	Déplacement chimique	Ecart type (ppm)
383Asn	Hba	2,68	0,01
383Asn	Hbb	2,84	0,01
383Asn	Ha	5,00	0,00
383Asn	H	8,30	0,00
383Asn	N	119,52	0,05
384Pro	Ca	63,54	0,01
384Pro	Cb	32,13	0,02
384Pro	Hda	3,74	0,00
384Pro	Hdb	3,74	0,00
384Pro	Hga	1,98	0,00
384Pro	Hgb	1,98	0,00
384Pro	Hbb	2,28	0,00
384Pro	Hba	1,95	0,01
384Pro	Ha	4,40	0,00
384Pro	C	177,05	0,00
385Lys	Ca	56,25	0,04
385Lys	Cb	32,87	0,03
385Lys	Hga	1,44	0,00
385Lys	Hgb	1,44	0,00
385Lys	Hea	3,01	0,00
385Lys	Heb	3,01	0,00
385Lys	Hda	1,71	0,00
385Lys	Hdb	1,71	0,00
385Lys	Hba	1,76	0,01
385Lys	Hbb	1,84	0,01
385Lys	Ha	4,31	0,01
385Lys	C	176,46	0,00
385Lys	H	8,30	0,00
385Lys	N	120,33	0,05
386Asp	Hba	2,61	0,00
386Asp	Hbb	2,73	0,01
386Asp	Ha	4,58	0,01
386Asp	Cb	41,10	0,00
386Asp	C	175,93	0,00
386Asp	Ca	54,54	0,03
386Asp	H	8,21	0,00
386Asp	N	120,92	0,02
387Asn	Hba	2,73	0,01
387Asn	Hbb	2,83	0,01
387Asn	Ha	4,70	0,00
387Asn	Cb	39,18	0,01
387Asn	Ca	53,31	0,04
387Asn	C	174,81	0,00
387Asn	H	8,28	0,00
387Asn	N	118,32	0,01
388Asp	Hba	2,63	0,01
388Asp	Hbb	2,71	0,01
388Asp	Ha	4,62	0,01
388Asp	Cb	41,06	0,00
388Asp	Ca	54,50	0,05
388Asp	C	175,96	0,00
388Asp	H	8,34	0,00
388Asp	N	120,76	0,02
389Val	Hgb*	0,93	0,01
389Val	Hga*	0,93	0,01
389Val	Hb	2,08	0,00
389Val	Ha	4,13	0,00

Résidu	Atome attribué	Déplacement chimique	Ecart type (ppm)
389Val	Cb	33,01	0,04
389Val	Ca	62,07	0,04
389Val	C	175,75	0,00
389Val	H	7,91	0,00
389Val	N	119,15	0,06
390Glu	Hga	2,27	0,00
390Glu	Hgb	2,27	0,00
390Glu	Hba	1,92	0,00
390Glu	Hbb	2,06	0,01
390Glu	Cb	30,29	0,06
390Glu	Ha	4,32	0,00
390Glu	Ca	56,35	0,17
390Glu	H	8,39	0,00
390Glu	N	125,15	0,01
391Leu	Ca	56,63	0,00
391Leu	Hba	1,58	0,00
391Leu	Hbb	1,58	0,00
391Leu	Ha	4,18	0,00
391Leu	Cb	43,37	0,00
391Leu	C	175,28	0,00
391Leu	H	7,92	0,00
391Leu	N	129,83	0,01

## 8.2. Titrage de $^{13}\text{C}^{15}\text{N}$ -N(C arm) par Tax1BP1(UBZ1+2) non marquée

Etant donné les faibles amplitudes et le régime linéaire des valeurs de perturbations j'ai seulement mis les valeurs du premier, intermédiaire et du dernier point du titrage.

Résidus	Référence			15 équivalents en UBZ1+2			30 équivalents en UBZ1+2				
	$\delta^1\text{H}$ (ppm)	$\delta^{15}\text{N}$ (ppm)	Intensité	$\delta^1\text{H}$ (ppm)	$\delta^{15}\text{N}$ (ppm)	Intensité	$\delta^1\text{H}$ (ppm)	$\delta^{15}\text{N}$ (ppm)	Intensité	$\Delta\delta^1\text{H}^{15}\text{N}$ (ppm)	$I/I_0$ $^{13}\text{C}^{15}\text{N}$
V362	8,00	119,68	4,85E+04	8,00	119,66	4,51E+04	8,01	119,63	5,61E+04	0,020	1,16
I363	8,21	125,24	6,57E+04	8,21	125,17	6,16E+04	8,21	125,12	6,94E+04	0,045	1,06
N364	8,36	123,20	1,82E+04	8,36	123,20	1,71E+04	8,36	123,19	2,10E+04	0,001	1,16
Y365	8,10	121,26	3,31E+04	8,10	121,24	2,88E+04	8,10	121,22	3,47E+04	0,015	1,05
S366	8,29	116,77	2,40E+04	8,30	116,72	2,42E+04	8,30	116,69	2,64E+04	0,029	1,10
V367	7,95	120,22	4,74E+04	7,96	120,19	4,01E+04	7,97	120,17	4,37E+04	0,023	0,92
L368	7,97	123,34	5,32E+04	7,96	123,24	4,96E+04	7,97	123,14	4,87E+04	0,076	0,91
D369	8,19	120,45	5,39E+04	8,19	120,42	4,93E+04	8,19	120,39	4,87E+04	0,024	0,90
L370	7,82	122,07	5,25E+04	7,81	122,03	4,84E+04	7,81	122,01	4,79E+04	0,025	0,91
T371	8,74	113,50	6,05E+04	8,75	113,49	5,38E+04	8,76	113,49	5,30E+04	0,017	0,88
A372	8,69	123,58	4,29E+04	8,70	123,56	4,24E+04	8,71	123,55	4,06E+04	0,024	0,95
E373	8,65	117,85	4,87E+04	8,66	117,83	4,88E+04	8,66	117,80	5,15E+04	0,022	1,06
E374	7,91	121,58	4,87E+04	7,91	121,58	4,51E+04	7,91	121,59	4,53E+04	0,004	0,93
E376	8,01	119,38	6,27E+04	8,01	119,36	6,17E+04	8,02	119,33	6,23E+04	0,023	0,99
A377	7,68	121,12	4,60E+04	7,68	121,09	4,20E+04	7,68	121,09	4,45E+04	0,010	0,97
I378	7,68	117,47	5,90E+04	7,69	117,53	5,10E+04	7,70	117,59	5,31E+04	0,047	0,90
K379	7,83	121,30	5,30E+04	7,84	121,24	4,54E+04	7,86	121,17	4,72E+04	0,058	0,89
H380	7,96	117,09	1,79E+04	7,97	117,03	1,41E+04	7,98	116,98	2,06E+04	0,047	1,15
Q381	7,94	119,56	2,12E+04	7,94	119,49	2,06E+04	7,93	119,43	2,16E+04	0,052	1,02
L382	8,06	121,49	4,00E+04	8,05	121,36	3,85E+04	8,05	121,26	3,94E+04	0,089	0,99
N383	8,31	119,63	2,87E+04	8,29	119,54	2,67E+04	8,29	119,47	3,07E+04	0,064	1,07
K385	8,31	120,43	5,03E+04	8,31	120,44	4,67E+04	8,31	120,43	5,10E+04	0,005	1,01
D386	8,21	120,95	4,29E+04	8,21	120,95	4,26E+04	8,22	120,94	4,53E+04	0,006	1,06
N387	8,28	118,35	3,47E+04	8,28	118,38	3,29E+04	8,28	118,39	4,14E+04	0,018	1,19
V389	7,91	119,17	6,73E+04	7,91	119,18	6,62E+04	7,91	119,18	6,90E+04	0,007	1,03
E390	8,39	125,18	6,86E+04	8,39	125,18	6,98E+04	8,40	125,17	7,00E+04	0,006	1,02
L391	7,91	129,86	6,42E+04	7,92	129,86	6,51E+04	7,92	129,85	7,01E+04	0,007	1,09

### 8.3. Titrage de <sup>15</sup>N-Tax1BP1(UBZ1+2) par N(C arm) non marquée

Etant donné les faibles amplitudes et le régime linéaire des valeurs de perturbations j'ai seulement mis les valeurs du premier, intermédiaire et du dernier point du titrage.

Résidus	Référence			25 équivalents en N(C arm)			45 équivalents en N(C arm)				
	$\delta^1\text{H}$ (ppm)	$\delta^{15}\text{N}$ (ppm)	Intensité	$\delta^1\text{H}$ (ppm)	$\delta^{15}\text{N}$ (ppm)	Intensité	$\delta^1\text{H}$ (ppm)	$\delta^{15}\text{N}$ (ppm)	Intensité	$\Delta\delta^1\text{H}^{15}\text{N}$ (ppm)	I/I <sub>0</sub> <sup>1</sup> H <sup>15</sup> N
725Asp	8,58	122,60	8,66E+04	8,59	122,58	8,53E+04	8,59	122,59	8,79E+04	0,007	1,02
726Val	8,02	118,61	1,61E+05	8,02	118,60	1,36E+05	8,01	118,60	1,62E+05	0,011	1,00
727His	8,32	122,26	1,08E+05	8,32	122,22	1,04E+05	8,32	122,16	1,03E+05	0,033	0,95
728Lys	8,56	122,64	8,37E+04	8,56	122,63	7,60E+04	8,55	122,64	6,04E+04	0,018	0,72
729Lys	8,45	125,14	4,74E+04	8,45	125,08	4,16E+04	8,44	125,04	3,39E+04	0,039	0,71
730Cys	9,02	129,74	-2,32E+04	9,02	129,72	-2,27E+04	9,02	129,67	-2,21E+04	0,024	0,95
732Leu	9,38	121,89	1,87E+04	9,37	121,84	1,39E+04	9,37	121,72	1,03E+04	0,061	0,55
733Cys	7,90	119,58	2,53E+04	7,90	119,54	2,26E+04	7,90	119,51	1,88E+04	0,026	0,74
734Glu	8,16	114,48	2,66E+04	8,15	114,49	2,51E+04	8,14	114,51	2,55E+04	0,028	0,96
735Leu	8,45	124,47	1,81E+04	8,44	124,41	1,46E+04	8,44	124,43	1,25E+04	0,022	0,69
736Met	7,76	120,94	3,11E+04	7,76	120,97	2,52E+04	7,76	121,00	2,73E+04	0,025	0,88
737Phe	9,00	122,57	5,25E+04	8,99	122,54	4,55E+04	8,98	122,49	3,54E+04	0,036	0,67
740Asn	8,36	113,66	3,80E+04	8,37	113,65	3,53E+04	8,37	113,66	2,75E+04	0,002	0,72
741Tyr	7,70	121,48	5,65E+04	7,70	121,49	4,94E+04	7,70	121,50	4,49E+04	0,015	0,79
742Asp	7,89	126,27	2,37E+04	7,89	126,26	2,16E+04	7,88	126,26	2,20E+04	0,022	0,93
743Gln	8,92	128,02	-7,59E+04	8,92	128,00	-6,72E+04	8,92	127,99	-5,48E+04	0,017	0,72
744Ser	8,48	115,45	1,13E+05	8,48	115,45	1,01E+05	8,47	115,45	8,41E+04	0,013	0,75
745Lys	7,78	123,19	1,06E+05	7,78	123,20	8,26E+04	7,78	123,20	7,71E+04	0,008	0,73
746Phe	8,43	123,76	7,04E+04	8,44	123,75	5,63E+04	8,43	123,73	5,97E+04	0,015	0,85
747Glu	8,53	118,84	9,03E+04	8,54	118,85	7,69E+04	8,53	118,87	6,60E+04	0,013	0,73
748Glu	7,99	118,88	1,18E+05	8,00	118,87	9,62E+04	7,99	118,82	7,73E+04	0,022	0,66
749His	7,82	118,23	8,69E+04	7,82	118,22	7,36E+04	7,82	118,21	6,82E+04	0,011	0,78
750Val	8,22	122,41	8,20E+04	8,22	122,37	6,46E+04	8,21	122,36	6,15E+04	0,023	0,75
751Glu	7,99	116,02	9,79E+04	8,00	115,97	7,83E+04	8,00	115,94	6,20E+04	0,027	0,63
752Ser	7,60	115,70	8,14E+04	7,60	115,71	7,70E+04	7,60	115,71	6,86E+04	0,006	0,84
753His	7,31	118,66	8,25E+04	7,31	118,64	7,56E+04	7,31	118,62	5,96E+04	0,017	0,72
754Trp	7,97	121,17	7,30E+04	7,98	121,19	5,94E+04	7,98	121,20	4,90E+04	0,013	0,67
755Lys	8,62	123,00	6,93E+04	8,61	122,94	5,81E+04	8,60	122,87	3,90E+04	0,056	0,56
756Val	8,35	122,52	9,26E+04	8,35	122,47	8,67E+04	8,34	122,41	7,75E+04	0,042	0,84

Résidus	Référence			25 équivalents en N(C arm)			45 équivalents en N(C arm)				
	$\delta^1\text{H}$ (ppm)	$\delta^{15}\text{N}$ (ppm)	Intensité	$\delta^1\text{H}$ (ppm)	$\delta^{15}\text{N}$ (ppm)	Intensité	$\delta^1\text{H}$ (ppm)	$\delta^{15}\text{N}$ (ppm)	Intensité	$\Delta\delta^1\text{H}^{15}\text{N}$ (ppm)	$I/I_0$ $^1\text{H}^{15}\text{N}$
757Cys	9,17	130,39	-4,96E+04	9,18	130,32	-4,92E+04	9,17	130,31	-3,87E+04	0,031	0,78
759Met	9,09	119,24	7,23E+04	9,09	119,16	5,17E+04	9,09	119,11	4,26E+04	0,046	0,59
760Cys	8,11	118,53	6,26E+04	8,11	118,47	5,66E+04	8,10	118,42	4,30E+04	0,042	0,69
761Ser	7,94	113,55	5,91E+04	7,94	113,55	4,63E+04	7,94	113,58	4,23E+04	0,013	0,72
762Glu	8,52	125,06	6,54E+04	8,52	125,04	6,16E+04	8,51	125,02	5,37E+04	0,021	0,82
763Gln	8,05	121,92	7,24E+04	8,05	121,93	6,10E+04	8,05	121,97	4,65E+04	0,015	0,64
764Phe	9,34	122,86	6,34E+04	9,33	122,88	6,02E+04	9,33	122,84	4,86E+04	0,022	0,77
767Asp	8,04	113,39	7,27E+04	8,05	113,41	5,82E+04	8,05	113,43	5,24E+04	0,014	0,72
768Tyr	7,83	122,45	5,31E+04	7,83	122,43	4,20E+04	7,83	122,38	3,96E+04	0,025	0,75
769Asp	8,05	125,39	3,87E+04	8,06	125,36	4,30E+04	8,05	125,34	3,32E+04	0,022	0,86
770Gln	8,90	127,01	7,30E+04	8,90	126,95	7,13E+04	8,89	126,90	5,75E+04	0,040	0,79
771Gln	8,45	117,71	1,08E+05	8,45	117,71	9,60E+04	8,45	117,70	7,68E+04	0,010	0,71
772Val	7,88	122,16	1,12E+05	7,88	122,14	8,89E+04	7,88	122,11	5,60E+04	0,021	0,50
773Phe	8,30	122,09	6,42E+04	8,30	122,07	6,17E+04	8,29	122,05	5,40E+04	0,023	0,84
774Glu	8,64	118,75	8,41E+04	8,63	118,77	9,10E+04	8,63	118,78	6,84E+04	0,014	0,81
775Arg	8,03	119,07	1,04E+05	8,03	119,02	6,89E+04	8,03	118,98	4,82E+04	0,031	0,46
776His	7,94	119,57	9,01E+04	7,94	119,53	8,72E+04	7,94	119,52	8,08E+04	0,020	0,90
777Val	8,16	121,33	8,42E+04	8,16	121,31	6,24E+04	8,16	121,27	4,93E+04	0,022	0,59
778Gln	7,76	115,26	9,95E+04	7,78	115,27	8,28E+04	7,78	115,28	6,92E+04	0,008	0,69
779Thr	7,68	112,28	7,50E+04	7,68	112,28	7,83E+04	7,68	112,28	5,70E+04	0,012	0,76
780His	7,25	117,40	7,97E+04	7,25	117,37	6,99E+04	7,25	117,38	5,50E+04	0,012	0,69
781Phe	7,62	118,20	8,78E+04	7,62	118,11	6,81E+04	7,62	118,05	7,29E+04	0,053	0,83
782Asp	8,06	121,00	1,16E+05	8,06	120,98	8,52E+04	8,06	120,96	8,40E+04	0,016	0,72
783Gln	8,29	120,01	1,54E+05	8,30	119,99	1,21E+05	8,30	119,97	1,10E+05	0,015	0,72
784Asn	8,43	119,10	1,47E+05	8,43	119,07	1,31E+05	8,43	119,03	1,21E+05	0,026	0,82
785Val	7,92	119,66	1,89E+05	7,92	119,61	1,67E+05	7,92	119,57	1,89E+05	0,032	1,00
786Leu	8,14	124,77	1,95E+05	8,14	124,70	1,67E+05	8,13	124,63	1,59E+05	0,053	0,82
787Asn	8,07	118,98	1,86E+05	8,07	118,95	1,64E+05	8,07	118,94	1,87E+05	0,019	1,00
788Phe	8,23	121,22	1,90E+05	8,23	121,21	1,78E+05	8,23	121,20	1,77E+05	0,012	0,93
789Asp	7,99	126,79	1,73E+05	7,99	126,76	1,68E+05	7,99	126,76	1,72E+05	0,012	1,00

## 8.4. Données des expériences de $^{15}\text{N}$ -hetNOE sur P entière

Num.	Rés.	$\delta^1\text{H}$ (ppm)	$\delta^{15}\text{N}$ (ppm)	$I_{\text{insat}(0)}$	$I_{\text{sat}}$	$I_{\text{sat}}/I_{\text{insat}}$	$\Delta I/I_0$
1	Met	8,46	123,9	4,09E+04	-3,85E+04	-0,94	1,6E-03
2	Glu	8,32	122,7	6,81E+04	-1,45E+04	-0,21	1,3E-02
3	Lys	8,25	121,8	5,77E+04	-1,10E+04	-0,19	1,6E-02
4	Phe	8,24	121,0	5,61E+04	-2,21E+03	-0,04	1,9E-02
5	Ala	8,24	127,2	5,92E+04	-1,53E+03	-0,03	1,8E-02
7	Glu	8,57	120,2	7,63E+04	3,82E+03	0,05	1,5E-02
8	Phe	8,27	121,3	5,41E+04	8,88E+03	0,16	2,4E-02
10	Gly	7,86	109,6	3,75E+04	1,98E+03	0,05	3,1E-02
11	Glu	8,55	120,8	7,04E+04	8,41E+03	0,12	1,8E-02
12	Asp	8,49	121,4	7,30E+04	6,33E+03	0,09	1,7E-02
13	Ala	<b>8,36</b>	<b>124,3</b>	<b>1,06E+05</b>	<b>1,14E+04</b>	<b>0,11</b>	-
14	Asn	8,42	116,3	5,37E+04	1,50E+04	0,28	2,6E-02
15	Asn	8,21	118,8	3,82E+04	8,24E+03	0,22	3,5E-02
16	Arg	8,23	120,7	4,05E+04	8,95E+03	0,22	3,3E-02
17	Ala	8,30	123,7	3,87E+04	1,43E+04	0,37	3,9E-02
18	Thr	8,08	113,4	4,54E+04	1,65E+04	0,36	3,3E-02
19	Lys	<b>8,19</b>	<b>122,8</b>	<b>3,01E+04</b>	<b>4,98E+03</b>	<b>0,17</b>	-
20	Phe	8,22	120,9	3,64E+04	8,95E+03	0,25	3,8E-02
21	Leu	8,15	122,7	2,74E+04	5,56E+03	0,20	4,9E-02
22	Glu	<b>8,29</b>	<b>120,7</b>	<b>3,00E+04</b>	<b>8,09E+03</b>	<b>0,27</b>	-
23	Ser	8,21	116,2	2,32E+04	4,88E+03	0,21	5,8E-02
24	Ile	7,95	121,4	2,38E+04	5,84E+03	0,24	5,8E-02
25	Lys	8,22	124,3	2,44E+04	5,78E+03	0,24	5,6E-02
26	Gly	8,37	109,6	2,27E+04	7,29E+03	0,32	6,5E-02
27	Lys	8,03	120,4	3,56E+04	1,30E+04	0,37	4,3E-02
28	Phe	8,36	121,1	3,74E+04	6,79E+03	0,18	3,5E-02
29	Thr	8,12	116,7	2,66E+04	2,88E+03	0,11	4,6E-02
30	Ser	<b>8,44</b>	<b>120,0</b>	<b>6,58E+04</b>	<b>1,20E+04</b>	<b>0,18</b>	-
32	Lys	8,39	121,2	5,82E+04	9,93E+03	0,17	2,2E-02
33	Asp	8,36	122,8	6,83E+04	1,29E+04	0,19	1,9E-02
35	Lys	8,43	120,0	1,02E+05	2,20E+04	0,22	1,3E-02
36	Lys	8,13	122,3	6,31E+04	1,15E+04	0,18	2,1E-02
37	Lys	8,43	123,4	4,86E+04	5,34E+03	0,11	2,5E-02
38	Asp	<b>8,44</b>	<b>121,6</b>	<b>2,20E+05</b>	<b>1,24E+04</b>	<b>0,06</b>	-
39	Ser	8,26	115,8	5,13E+04	-6,28E+01	0,00	2,2E-02
40	Ile	8,23	122,7	7,92E+04	-5,60E+03	-0,07	1,3E-02
41	Ile	8,29	125,3	8,48E+04	-5,51E+03	-0,06	1,2E-02
42	Ser	8,46	120,8	6,61E+04	-2,54E+03	-0,04	1,6E-02
43	Val	8,31	122,1	7,83E+04	7,55E+02	0,01	1,4E-02
44	Asn	8,55	122,0	5,05E+04	-2,11E+03	-0,04	2,1E-02
45	Ser	8,33	116,6	4,48E+04	1,74E+03	0,04	2,6E-02
46	Ile	8,16	121,8	1,02E+05	-9,56E+03	-0,09	9,9E-03
47	Asp	8,37	124,1	1,12E+05	1,38E+04	0,12	1,1E-02
48	Ile	<b>8,05</b>	<b>120,8</b>	<b>1,28E+05</b>	<b>1,10E+04</b>	<b>0,09</b>	-
49	Glu	8,48	124,9	1,16E+05	-7,06E+03	-0,06	9,0E-03
50	Val	8,27	122,2	9,98E+04	2,13E+03	0,02	1,1E-02
51	Thr	8,30	118,6	6,96E+04	3,34E+03	0,05	1,7E-02
52	Lys	8,43	124,2	7,75E+04	1,27E+03	0,02	1,5E-02
53	Glu	8,48	122,2	1,37E+05	-6,43E+03	-0,05	7,7E-03
54	Ser	8,50	118,9	7,84E+04	-1,54E+04	-0,20	1,1E-02
56	Ile	8,34	121,1	9,21E+04	2,71E+03	0,03	1,2E-02
57	Thr	8,27	118,1	5,53E+04	-6,00E+03	-0,11	1,8E-02
58	Ser	8,40	118,0	3,28E+04	-6,15E+01	0,00	3,4E-02
59	Asn	8,54	121,0	2,56E+04	-3,86E+03	-0,15	3,7E-02

Num.	Rés.	$\delta^1\text{H}$ (ppm)	$\delta^{15}\text{N}$ (ppm)	$I_{\text{insat}(0)}$	$I_{\text{sat}}$	$I_{\text{sat}}/I_{\text{insat}}$	$\Delta I/I_0$
60	Ser	8,36	116,1	3,05E+04	-1,41E+03	-0,05	3,5E-02
61	Thr	8,26	116,7	5,71E+04	-2,84E+03	-0,05	1,8E-02
62	Ile	8,17	124,1	1,06E+05	-3,97E+03	-0,04	1,0E-02
63	Ile	8,31	125,9	8,80E+04	2,46E+03	0,03	1,3E-02
64	Asn	8,64	124,4	8,70E+04	5,38E+03	0,06	1,4E-02
66	Thr	8,30	113,8	1,14E+05	9,29E+03	0,08	1,0E-02
67	Asn	8,47	120,7	6,48E+04	2,53E+03	0,04	1,8E-02
68	Glu	8,52	121,5	9,25E+04	-4,06E+03	-0,04	1,1E-02
69	Thr	8,25	114,3	9,65E+04	-4,30E+03	-0,04	1,1E-02
70	Asp	8,38	122,7	1,19E+05	2,13E+04	0,18	1,1E-02
71	Asp	<b>8,42</b>	<b>121,6</b>	<b>1,71E+05</b>	<b>2,36E+04</b>	<b>0,14</b>	-
72	Asn	8,46	118,7	9,04E+04	5,86E+03	0,06	1,3E-02
73	Ala	8,21	123,6	1,05E+05	2,72E+02	0,00	1,1E-02
74	Gly	8,38	106,9	7,76E+04	1,33E+04	0,17	1,7E-02
75	Asn	8,18	118,2	6,84E+04	1,17E+03	0,02	1,7E-02
76	Lys	8,22	122,3	7,20E+04	-2,96E+03	-0,04	1,5E-02
78	Asn	8,54	118,5	4,01E+04	5,32E+03	0,13	3,1E-02
79	Tyr	8,10	120,7	6,56E+04	1,48E+04	0,23	2,1E-02
80	Gln	8,23	122,1	4,41E+04	8,35E+03	0,19	3,0E-02
81	Arg	<b>8,26</b>	<b>122,8</b>	<b>4,48E+04</b>	<b>4,72E+03</b>	<b>0,11</b>	-
82	Lys	8,43	124,4	4,83E+04	8,40E+03	0,17	2,7E-02
84	Leu	<b>8,43</b>	<b>122,8</b>	<b>6,96E+04</b>	<b>-5,60E+03</b>	<b>-0,08</b>	-
85	Val	8,15	121,1	5,30E+04	1,66E+03	0,03	2,2E-02
86	Ser	8,35	119,5	3,41E+04	4,72E+03	0,14	3,7E-02
87	Phe	<b>8,33</b>	<b>122,7</b>	<b>8,05E+04</b>	<b>-6,28E+03</b>	<b>-0,08</b>	-
88	Lys	8,21	123,4	4,09E+04	-1,15E+03	-0,03	2,6E-02
89	Glu	8,40	122,4	5,49E+04	4,74E+03	0,09	2,2E-02
90	Asp	8,47	123,5	6,36E+04	-2,14E+03	-0,03	1,7E-02
92	Ile	<b>8,35</b>	<b>122,8</b>	<b>9,87E+04</b>	<b>1,23E+04</b>	<b>0,12</b>	-
94	Ser	8,44	115,7	4,54E+04	4,05E+03	0,09	2,7E-02
95	Asp	8,37	121,9	4,34E+04	9,18E+03	0,21	3,1E-02
96	Asn	8,29	119,5	3,11E+04	3,86E+03	0,12	4,0E-02
98	Phe	8,16	118,1	2,92E+04	8,72E+03	0,30	4,9E-02
99	Ser	7,94	116,0	2,95E+04	7,57E+03	0,26	4,7E-02
100	Lys	<b>8,28</b>	<b>122,8</b>	<b>4,42E+04</b>	<b>9,32E+03</b>	<b>0,21</b>	-
101	Leu	8,05	121,6	2,30E+04	7,21E+03	0,31	6,3E-02
102	Tyr	<b>8,07</b>	<b>120,8</b>	<b>1,40E+05</b>	<b>1,27E+04</b>	<b>0,09</b>	-
103	Lys	8,10	123,2	2,36E+04	5,50E+03	0,23	5,8E-02
104	Glu	8,43	122,0	4,26E+04	9,03E+03	0,21	3,2E-02
105	Thr	8,27	116,1	3,02E+04	5,03E+03	0,17	4,3E-02
106	Ile	8,24	123,6	2,95E+04	3,90E+03	0,13	4,3E-02
107	Glu	8,51	125,4	3,66E+04	2,16E+03	0,06	3,2E-02
108	Thr	8,12	115,1	4,29E+04	2,32E+03	0,05	2,7E-02
109	Phe	8,36	122,1	5,36E+04	6,09E+03	0,11	2,3E-02
110	Asp	8,40	121,9	6,57E+04	1,03E+04	0,16	2,0E-02
111	Asn	<b>8,40</b>	<b>119,4</b>	<b>1,24E+05</b>	<b>-2,44E+03</b>	<b>-0,02</b>	-
112	Asn	8,54	119,4	6,65E+04	-2,45E+03	-0,04	1,6E-02
113	Glu	8,47	121,2	8,32E+04	1,59E+03	0,02	1,4E-02
114	Glu	<b>8,45</b>	<b>121,6</b>	<b>1,17E+05</b>	<b>2,67E+03</b>	<b>0,02</b>	-
115	Glu	<b>8,51</b>	<b>121,9</b>	<b>1,36E+04</b>	<b>2,58E+03</b>	<b>0,19</b>	-
116	Ser	8,43	117,1	5,13E+04	-9,57E+02	-0,02	2,1E-02
117	Ser	8,35	117,8	3,54E+04	2,97E+03	0,08	3,4E-02
118	Tyr	8,17	122,1	2,94E+04	2,88E+03	0,10	4,2E-02
119	Ser	8,20	118,0	1,44E+04	4,01E+02	0,03	7,9E-02
206	Met	8,19	120,2	1,12E+04	2,23E+03	0,20	1,2E-01
207	Ala	8,22	124,8	2,22E+04	1,74E+02	0,01	5,0E-02
208	Lys	8,24	120,5	3,15E+04	-1,66E+03	-0,05	3,3E-02

Num.	Rés.	$\delta^1\text{H}$ (ppm)	$\delta^{15}\text{N}$ (ppm)	$I_{\text{insat}(0)}$	$I_{\text{sat}}$	$I_{\text{sat}}/I_{\text{insat}}$	$\Delta I/I_0$
209	Asp	8,47	121,8	5,33E+04	-7,12E+02	-0,01	2,1E-02
209	Asp	8,47	121,8	5,33E+04	-7,12E+02	-0,01	2,1E-02
210	Thr	8,20	114,4	5,61E+04	-5,15E+03	-0,09	1,8E-02
211	Ser	8,45	118,2	6,79E+04	9,45E+02	0,01	1,7E-02
212	Asp	8,44	122,3	8,69E+04	3,61E+03	0,04	1,3E-02
213	Glu	8,30	121,0	1,05E+05	1,51E+03	0,01	1,1E-02
214	Val	8,21	121,3	1,03E+05	-1,85E+04	-0,18	8,8E-03
215	Ser	<b>8,42</b>	<b>119,4</b>	<b>1,28E+05</b>	<b>-3,67E+03</b>	<b>-0,03</b>	-
216	Leu	<b>8,34</b>	<b>124,3</b>	<b>1,06E+05</b>	<b>1,40E+04</b>	<b>0,13</b>	-
217	Asn	8,45	120,1	7,96E+04	-7,23E+03	-0,09	1,3E-02
219	Thr	<b>8,31</b>	<b>113,8</b>	<b>1,10E+05</b>	<b>5,89E+03</b>	<b>0,05</b>	-
220	Ser	8,30	117,7	5,70E+04	1,51E+03	0,03	2,0E-02
221	Glu	8,50	123,1	7,87E+04	-5,14E+03	-0,07	1,3E-02
222	Lys	8,33	122,1	9,96E+04	1,21E+04	0,12	1,3E-02
223	Leu	8,28	123,0	9,71E+04	-2,24E+03	-0,02	1,1E-02
224	Asn	8,44	119,1	8,58E+04	3,44E+02	0,00	1,3E-02
225	Asn	8,41	119,1	8,44E+04	5,45E+03	0,06	1,4E-02
226	Leu	8,24	121,9	1,13E+05	-1,27E+04	-0,11	8,7E-03
227	Leu	8,18	121,9	1,27E+05	-2,25E+04	-0,18	7,2E-03
228	Glu	8,30	121,3	1,22E+05	-6,37E+03	-0,05	8,6E-03
229	Gly	8,46	109,8	1,19E+05	-3,85E+04	-0,32	6,3E-03
230	Asn	8,40	118,7	8,50E+04	-7,42E+03	-0,09	1,2E-02
231	Asp	8,54	121,2	1,02E+05	-4,98E+04	-0,49	5,6E-03
232	Ser	8,32	115,7	1,14E+05	-2,16E+03	-0,02	9,6E-03
233	Asp	8,40	122,1	1,23E+05	5,42E+03	0,04	9,4E-03
234	Asn	8,27	118,7	1,05E+05	-1,75E+04	-0,17	8,8E-03
235	Asp	8,37	120,8	1,45E+05	9,22E+02	0,01	7,7E-03
236	Leu	8,25	122,9	1,23E+05	-3,66E+04	-0,30	6,3E-03
237	Ser	8,46	116,9	1,34E+05	-3,58E+04	-0,27	6,1E-03
238	Leu	8,26	124,0	1,17E+05	-5,08E+04	-0,43	5,4E-03
239	Glu	8,27	120,6	1,37E+05	-9,46E+04	-0,69	2,5E-03
240	Asp	8,23	121,4	1,34E+05	-1,73E+05	-1,29	-2,4E-03
241	Phe	7,64	124,9	1,96E+05	-2,49E+05	-1,27	-1,5E-03

Les résidus en gras correspondent aux résidus qui n'ont pas été représentés sur les diagrammes en barres à cause de superpositions de pics sur les spectres  $^1\text{H}^{15}\text{N}$ -HSQC.

## 8.5. Données des expériences de relaxation $^{15}\text{N}$ $R_1$ sur P entière

Rés.	Num.	$\delta^1\text{H}$ (ppm)	$\delta^{15}\text{N}$ (ppm)	$I_0$	$I_1$	$I_2$	$I_3$	$I_{3(\text{bis})}$	$I_4$	$I_5$	$I_6$	$I_7$	$I_8$	$R_1$ (s $^{-1}$ )	$\Delta R_1$ (s $^{-1}$ )
Met	1	8,47	123,9	1,98E+05	1,85E+05	1,71E+05	1,62E+05	1,67E+05	1,18E+05	8,99E+04	6,94E+04	4,76E+04	1,62E+04	1,25	5,9E-02
Glu	2	8,32	122,7	2,52E+05	2,36E+05	2,16E+05	1,99E+05	2,06E+05	1,41E+05	1,04E+05	7,83E+04	4,90E+04	1,28E+04	1,42	5,9E-02
Lys	3	8,25	121,8	2,18E+05	2,03E+05	1,82E+05	1,59E+05	1,65E+05	1,14E+05	8,35E+04	5,87E+04	3,35E+04	1,00E+04	1,59	3,3E-02
Phe	4	8,24	121,0	1,92E+05	1,82E+05	1,61E+05	1,50E+05	1,56E+05	1,03E+05	7,40E+04	5,50E+04	3,18E+04	9,37E+03	1,53	7,5E-02
Ala	5	8,24	127,2	2,26E+05	2,14E+05	1,96E+05	1,76E+05	1,78E+05	1,23E+05	9,01E+04	6,56E+04	3,85E+04	1,08E+04	1,51	4,7E-02
Glu	7	8,57	120,2	2,56E+05	2,43E+05	2,17E+05	1,90E+05	1,93E+05	1,36E+05	9,79E+04	7,23E+04	3,75E+04	1,05E+04	1,59	2,9E-02
Phe	8	8,27	121,3	1,87E+05	1,70E+05	1,55E+05	1,37E+05	1,40E+05	9,33E+04	6,71E+04	4,60E+04	2,71E+04	5,72E+03	1,68	4,8E-02
Gly	10	7,86	109,5	8,99E+04	8,41E+04	7,37E+04	7,00E+04	7,30E+04	4,89E+04	3,68E+04	2,60E+04	1,75E+04	3,45E+03	1,46	7,8E-02
Glu	11	8,55	120,8	2,09E+05	1,92E+05	1,73E+05	1,58E+05	1,59E+05	1,09E+05	7,72E+04	5,61E+04	3,27E+04	8,67E+03	1,59	5,3E-02
Asp	12	8,49	121,4	2,01E+05	1,88E+05	1,71E+05	1,54E+05	1,62E+05	1,06E+05	7,95E+04	5,59E+04	3,32E+04	1,01E+04	1,53	6,5E-02
<b>Ala</b>	<b>13</b>	<b>8,36</b>	<b>124,3</b>	<b>3,46E+05</b>	<b>3,22E+05</b>	<b>2,90E+05</b>	<b>2,59E+05</b>	<b>2,59E+05</b>	<b>1,82E+05</b>	<b>1,33E+05</b>	<b>9,87E+04</b>	<b>5,83E+04</b>	<b>1,86E+04</b>	<b>1,55</b>	<b>3,3E-02</b>
Asn	14	8,42	116,3	1,32E+05	1,20E+05	1,08E+05	1,01E+05	1,04E+05	6,98E+04	5,05E+04	3,67E+04	2,18E+04	4,03E+03	1,54	7,5E-02
Asn	15	8,21	118,8	9,21E+04	8,56E+04	7,74E+04	6,80E+04	7,37E+04	4,81E+04	3,68E+04	2,52E+04	1,68E+04	1,72E+03	1,54	7,0E-02
Arg	16	8,23	120,7	1,06E+05	1,05E+05	9,11E+04	8,28E+04	8,48E+04	5,74E+04	3,88E+04	2,60E+04	1,67E+04	7,02E+03	1,63	8,6E-02
Ala	17	8,30	123,7	9,83E+04	9,19E+04	8,08E+04	7,48E+04	7,56E+04	4,97E+04	3,61E+04	2,47E+04	1,46E+04	4,26E+03	1,65	6,8E-02
Thr	18	8,08	113,4	1,06E+05	9,66E+04	9,01E+04	8,07E+04	8,11E+04	5,29E+04	3,98E+04	2,95E+04	2,00E+04	4,56E+03	1,55	6,5E-02
Lys	19	8,19	122,8	7,37E+04	6,67E+04	5,94E+04	5,60E+04	5,73E+04	3,80E+04	2,77E+04	2,20E+04	1,17E+04	1,69E+03	1,54	7,5E-02
Phe	20	8,22	120,9	9,22E+04	8,70E+04	8,04E+04	6,82E+04	7,03E+04	4,80E+04	3,01E+04	2,37E+04	1,42E+04	4,34E+03	1,69	6,4E-02
Leu	21	8,15	122,7	6,37E+04	5,65E+04	5,18E+04	4,84E+04	4,83E+04	3,41E+04	2,57E+04	1,52E+04	7,22E+03	4,72E+03	1,58	8,9E-02
<b>Glu</b>	<b>22</b>	<b>8,29</b>	<b>120,6</b>	<b>6,72E+04</b>	<b>5,99E+04</b>	<b>5,81E+04</b>	<b>5,33E+04</b>	<b>5,11E+04</b>	<b>3,73E+04</b>	<b>2,82E+04</b>	<b>2,18E+04</b>	<b>9,81E+03</b>	<b>2,69E+03</b>	<b>1,45</b>	<b>6,0E-02</b>
Ser	23	8,21	116,2	6,13E+04	5,56E+04	4,76E+04	4,43E+04	4,45E+04	3,13E+04	2,38E+04	1,56E+04	9,30E+03	2,00E+03	1,6	6,5E-02
Ile	24	7,95	121,5	5,38E+04	5,06E+04	4,05E+04	4,04E+04	3,86E+04	2,64E+04	1,91E+04	1,32E+04	9,93E+03	2,98E+03	1,64	1,1E-01
Lys	25	8,22	124,3	6,20E+04	5,58E+04	5,09E+04	4,58E+04	4,93E+04	3,26E+04	2,44E+04	1,76E+04	9,89E+03	3,42E+03	1,52	6,8E-02
Gly	26	8,37	109,6	5,73E+04	5,18E+04	4,67E+04	4,40E+04	4,52E+04	2,89E+04	2,17E+04	1,68E+04	7,16E+03	4,17E+03	1,56	8,9E-02
Lys	27	8,04	120,4	9,40E+04	8,50E+04	7,81E+04	7,29E+04	7,34E+04	5,01E+04	3,79E+04	2,70E+04	1,65E+04	4,49E+03	1,48	6,0E-02
Phe	28	8,36	121,1	8,15E+04	7,81E+04	7,06E+04	6,28E+04	6,82E+04	4,19E+04	3,28E+04	2,04E+04	1,47E+04	5,92E+03	1,56	9,9E-02
Thr	29	8,12	116,7	7,19E+04	6,60E+04	5,95E+04	5,63E+04	6,08E+04	3,99E+04	2,96E+04	2,40E+04	1,47E+04	4,85E+03	1,36	8,6E-02
<b>Ser</b>	<b>30</b>	<b>8,44</b>	<b>120,0</b>	<b>1,30E+05</b>	<b>1,22E+05</b>	<b>1,11E+05</b>	<b>9,71E+04</b>	<b>1,01E+05</b>	<b>6,84E+04</b>	<b>5,03E+04</b>	<b>3,57E+04</b>	<b>2,26E+04</b>	<b>6,51E+03</b>	<b>1,56</b>	<b>4,5E-02</b>
Lys	32	8,39	121,2	1,65E+05	1,51E+05	1,36E+05	1,25E+05	1,33E+05	8,87E+04	6,69E+04	4,35E+04	2,58E+04	1,01E+04	1,52	8,0E-02
Asp	33	8,36	122,8	1,86E+05	1,71E+05	1,57E+05	1,41E+05	1,47E+05	9,93E+04	7,64E+04	5,26E+04	3,16E+04	8,46E+03	1,5	5,3E-02

Rés.	Num.	$\delta^1\text{H}$ (ppm)	$\delta^{15}\text{N}$ (ppm)	I <sub>0</sub>	I <sub>1</sub>	I <sub>2</sub>	I <sub>3</sub>	I <sub>3(bis)</sub>	I <sub>4</sub>	I <sub>5</sub>	I <sub>6</sub>	I <sub>7</sub>	I <sub>8</sub>	R <sub>1</sub> (s <sup>-1</sup> )	$\Delta\text{R}_1$ (s <sup>-1</sup> )
Lys	35	8,43	120,0	2,95E+05	2,71E+05	2,46E+05	2,21E+05	2,29E+05	1,56E+05	1,10E+05	8,01E+04	4,68E+04	1,30E+04	1,58	5,4E-02
Lys	36	8,13	122,3	1,94E+05	1,78E+05	1,62E+05	1,46E+05	1,46E+05	9,92E+04	7,22E+04	5,04E+04	2,78E+04	8,66E+03	1,63	4,6E-02
Lys	37	8,43	123,4	1,31E+05	1,21E+05	1,09E+05	1,00E+05	1,04E+05	6,84E+04	5,34E+04	3,63E+04	2,27E+04	4,19E+03	1,52	6,8E-02
<b>Asp</b>	<b>38</b>	<b>8,44</b>	<b>121,6</b>	<b>6,29E+05</b>	<b>5,87E+05</b>	<b>5,31E+05</b>	<b>4,77E+05</b>	<b>4,86E+05</b>	<b>3,32E+05</b>	<b>2,44E+05</b>	<b>1,78E+05</b>	<b>1,05E+05</b>	<b>2,99E+04</b>	<b>1,55</b>	<b>4,3E-02</b>
Ser	39	8,26	115,8	1,82E+05	1,67E+05	1,54E+05	1,43E+05	1,49E+05	1,01E+05	7,55E+04	5,73E+04	3,67E+04	1,19E+04	1,39	6,6E-02
Ile	40	8,23	122,7	2,90E+05	2,71E+05	2,46E+05	2,24E+05	2,27E+05	1,54E+05	1,15E+05	8,31E+04	4,66E+04	1,59E+04	1,53	5,0E-02
Ile	41	8,29	125,3	2,88E+05	2,69E+05	2,43E+05	2,15E+05	2,16E+05	1,54E+05	1,14E+05	8,26E+04	4,76E+04	1,46E+04	1,54	2,3E-02
Ser	42	8,47	120,8	2,16E+05	1,99E+05	1,78E+05	1,68E+05	1,68E+05	1,17E+05	8,50E+04	6,38E+04	3,76E+04	1,21E+04	1,49	5,8E-02
Val	43	8,31	122,1	2,81E+05	2,60E+05	2,39E+05	2,19E+05	2,24E+05	1,60E+05	1,18E+05	8,49E+04	5,42E+04	1,67E+04	1,41	4,9E-02
Asn	44	8,55	122,0	1,59E+05	1,40E+05	1,31E+05	1,23E+05	1,26E+05	8,77E+04	6,33E+04	4,72E+04	2,77E+04	6,79E+03	1,45	7,5E-02
Ser	45	8,33	116,6	1,41E+05	1,33E+05	1,19E+05	1,16E+05	1,18E+05	8,24E+04	6,30E+04	4,68E+04	2,85E+04	1,01E+04	1,33	6,8E-02
Ile	46	8,16	121,8	3,71E+05	3,52E+05	3,21E+05	2,86E+05	2,90E+05	2,07E+05	1,49E+05	1,11E+05	6,33E+04	1,81E+04	1,49	3,6E-02
Asp	47	8,37	124,1	3,61E+05	3,37E+05	3,07E+05	2,70E+05	2,74E+05	1,89E+05	1,37E+05	1,01E+05	5,46E+04	1,49E+04	1,59	3,2E-02
<b>Ile</b>	<b>48</b>	<b>8,06</b>	<b>120,8</b>	<b>4,04E+05</b>	<b>3,84E+05</b>	<b>3,51E+05</b>	<b>3,10E+05</b>	<b>3,12E+05</b>	<b>2,30E+05</b>	<b>1,68E+05</b>	<b>1,26E+05</b>	<b>6,79E+04</b>	<b>1,77E+04</b>	<b>1,47</b>	<b>2,4E-02</b>
Glu	49	8,48	124,9	3,41E+05	3,20E+05	2,88E+05	2,51E+05	2,56E+05	1,80E+05	1,30E+05	9,58E+04	5,06E+04	1,57E+04	1,59	2,4E-02
Val	50	8,27	122,2	3,53E+05	3,35E+05	3,07E+05	2,76E+05	2,77E+05	1,99E+05	1,45E+05	1,06E+05	5,77E+04	1,33E+04	1,49	3,9E-02
Thr	51	8,30	118,6	2,43E+05	2,25E+05	2,02E+05	1,83E+05	1,85E+05	1,29E+05	9,39E+04	6,80E+04	4,18E+04	1,21E+04	1,54	4,3E-02
Lys	52	8,43	124,2	2,27E+05	2,09E+05	1,91E+05	1,78E+05	1,84E+05	1,23E+05	9,06E+04	6,87E+04	4,32E+04	1,06E+04	1,45	6,9E-02
Glu	53	8,48	122,2	4,16E+05	3,92E+05	3,59E+05	3,23E+05	3,25E+05	2,26E+05	1,67E+05	1,20E+05	7,21E+04	2,06E+04	1,51	4,3E-02
Ser	54	8,50	118,9	2,46E+05	2,32E+05	2,12E+05	2,03E+05	2,07E+05	1,42E+05	1,08E+05	7,89E+04	4,93E+04	1,86E+04	1,35	7,0E-02
Ile	56	8,33	121,1	3,50E+05	3,30E+05	3,00E+05	2,70E+05	2,75E+05	1,91E+05	1,40E+05	1,02E+05	5,92E+04	1,70E+04	1,51	4,3E-02
Thr	57	8,27	118,1	2,00E+05	1,85E+05	1,69E+05	1,59E+05	1,62E+05	1,08E+05	8,08E+04	5,96E+04	3,70E+04	1,08E+04	1,46	7,1E-02
Ser	58	8,40	118,0	8,54E+04	7,59E+04	7,08E+04	6,81E+04	6,98E+04	4,73E+04	3,61E+04	2,64E+04	1,87E+04	3,96E+03	1,36	8,4E-02
Asn	59	8,55	121,0	5,11E+04	4,83E+04	4,13E+04	4,10E+04	4,26E+04	2,79E+04	2,10E+04	1,45E+04	8,04E+03	2,73E+03	1,49	1,0E-01
Ser	60	8,36	116,1	8,56E+04	7,64E+04	7,09E+04	6,83E+04	7,35E+04	5,18E+04	3,64E+04	2,67E+04	1,93E+04	5,72E+03	1,3	9,7E-02
Thr	61	8,26	116,7	1,88E+05	1,74E+05	1,55E+05	1,49E+05	1,53E+05	1,02E+05	8,01E+04	5,61E+04	3,44E+04	1,11E+04	1,43	7,6E-02
Ile	62	8,17	124,1	3,28E+05	3,07E+05	2,81E+05	2,55E+05	2,59E+05	1,80E+05	1,32E+05	9,51E+04	5,80E+04	1,72E+04	1,49	4,7E-02
Ile	63	8,31	125,9	3,28E+05	3,07E+05	2,82E+05	2,46E+05	2,48E+05	1,77E+05	1,32E+05	9,09E+04	5,06E+04	1,49E+04	1,55	2,5E-02
Asn	64	8,64	124,4	2,41E+05	2,28E+05	2,07E+05	1,86E+05	1,93E+05	1,32E+05	9,68E+04	7,11E+04	4,36E+04	1,12E+04	1,49	5,2E-02
Thr	66	8,30	113,8	3,95E+05	3,69E+05	3,38E+05	3,11E+05	3,16E+05	2,16E+05	1,58E+05	1,14E+05	6,96E+04	1,91E+04	1,49	5,9E-02
Asn	67	8,47	120,7	2,08E+05	1,92E+05	1,73E+05	1,63E+05	1,66E+05	1,12E+05	8,09E+04	5,85E+04	4,04E+04	1,05E+04	1,49	7,3E-02
Glu	68	8,52	121,5	3,00E+05	2,76E+05	2,50E+05	2,30E+05	2,38E+05	1,58E+05	1,17E+05	8,64E+04	5,03E+04	1,43E+04	1,52	6,2E-02

Rés.	Num.	$\delta^1\text{H}$ (ppm)	$\delta^{15}\text{N}$ (ppm)	I <sub>0</sub>	I <sub>1</sub>	I <sub>2</sub>	I <sub>3</sub>	I <sub>3(bis)</sub>	I <sub>4</sub>	I <sub>5</sub>	I <sub>6</sub>	I <sub>7</sub>	I <sub>8</sub>	R <sub>1</sub> (s <sup>-1</sup> )	$\Delta\text{R}_1$ (s <sup>-1</sup> )
Thr	69	8,25	114,3	3,82E+05	3,58E+05	3,25E+05	2,97E+05	3,02E+05	2,05E+05	1,51E+05	1,08E+05	6,64E+04	1,75E+04	1,52	5,6E-02
Asp	70	8,37	122,7	4,02E+05	3,72E+05	3,39E+05	3,08E+05	3,17E+05	2,14E+05	1,57E+05	1,10E+05	6,24E+04	1,92E+04	1,55	5,7E-02
<b>Asp</b>	<b>71</b>	<b>8,43</b>	<b>121,6</b>	<b>5,88E+05</b>	<b>5,49E+05</b>	<b>5,03E+05</b>	<b>4,72E+05</b>	<b>4,82E+05</b>	<b>3,29E+05</b>	<b>2,39E+05</b>	<b>1,74E+05</b>	<b>1,04E+05</b>	<b>2,92E+04</b>	<b>1,46</b>	<b>7,1E-02</b>
Asn	72	8,46	118,7	2,88E+05	2,68E+05	2,44E+05	2,26E+05	2,34E+05	1,51E+05	1,10E+05	8,05E+04	5,10E+04	1,35E+04	1,53	7,9E-02
Ala	73	8,21	123,6	3,68E+05	3,40E+05	3,07E+05	2,80E+05	2,91E+05	1,89E+05	1,39E+05	1,01E+05	5,97E+04	1,68E+04	1,57	6,8E-02
Gly	74	8,39	106,9	2,51E+05	2,34E+05	2,13E+05	1,96E+05	2,03E+05	1,34E+05	9,84E+04	7,18E+04	4,52E+04	1,27E+04	1,5	6,9E-02
Asn	75	8,18	118,2	2,07E+05	1,88E+05	1,76E+05	1,62E+05	1,66E+05	1,11E+05	8,31E+04	5,91E+04	3,78E+04	8,90E+03	1,48	7,0E-02
Lys	76	8,22	122,3	2,82E+05	2,61E+05	2,38E+05	2,19E+05	2,23E+05	1,54E+05	1,16E+05	8,28E+04	5,21E+04	1,65E+04	1,45	5,0E-02
Asn	78	8,54	118,5	1,01E+05	9,63E+04	8,70E+04	7,88E+04	8,46E+04	5,49E+04	4,10E+04	3,21E+04	2,00E+04	5,53E+03	1,44	7,4E-02
Tyr	79	8,10	120,7	1,93E+05	1,78E+05	1,58E+05	1,44E+05	1,48E+05	1,02E+05	7,25E+04	5,31E+04	3,36E+04	1,10E+04	1,54	5,5E-02
Gln	80	8,23	122,1	1,21E+05	1,08E+05	9,91E+04	8,68E+04	9,33E+04	6,22E+04	4,47E+04	3,17E+04	1,86E+04	4,47E+03	1,6	6,2E-02
<b>Arg</b>	<b>81</b>	<b>8,27</b>	<b>122,8</b>	<b>1,05E+05</b>	<b>9,63E+04</b>	<b>8,79E+04</b>	<b>8,72E+04</b>	<b>8,54E+04</b>	<b>6,03E+04</b>	<b>4,49E+04</b>	<b>3,03E+04</b>	<b>1,94E+04</b>	<b>6,82E+03</b>	<b>1,41</b>	<b>8,5E-02</b>
Lys	82	8,43	124,4	1,30E+05	1,24E+05	1,12E+05	1,00E+05	1,02E+05	7,16E+04	5,37E+04	3,68E+04	2,38E+04	4,04E+03	1,5	4,5E-02
Leu	84	8,43	122,8	1,87E+05	1,75E+05	1,61E+05	1,42E+05	1,48E+05	1,03E+05	7,78E+04	5,73E+04	3,44E+04	9,60E+03	1,45	3,8E-02
Val	85	8,15	121,1	1,57E+05	1,47E+05	1,35E+05	1,20E+05	1,22E+05	9,01E+04	6,63E+04	4,80E+04	2,91E+04	6,96E+03	1,44	2,8E-02
Ser	86	8,35	119,5	9,35E+04	8,74E+04	7,91E+04	7,57E+04	7,81E+04	5,38E+04	4,02E+04	3,27E+04	2,02E+04	4,64E+03	1,33	6,9E-02
Phe	87	8,33	122,7	2,22E+05	2,05E+05	1,85E+05	1,71E+05	1,74E+05	1,19E+05	8,79E+04	6,39E+04	4,25E+04	1,21E+04	1,48	5,8E-02
Lys	88	8,21	123,5	1,34E+05	1,25E+05	1,16E+05	1,05E+05	1,06E+05	7,16E+04	5,64E+04	3,80E+04	2,39E+04	8,73E+03	1,48	5,4E-02
Glu	89	8,40	122,5	1,60E+05	1,50E+05	1,36E+05	1,22E+05	1,27E+05	8,90E+04	6,29E+04	4,92E+04	2,94E+04	8,34E+03	1,47	4,9E-02
Asp	90	8,47	123,5	1,81E+05	1,70E+05	1,56E+05	1,44E+05	1,43E+05	1,04E+05	7,68E+04	5,49E+04	3,56E+04	1,14E+04	1,41	4,4E-02
Ile	92	8,35	122,8	2,67E+05	2,51E+05	2,29E+05	2,08E+05	2,09E+05	1,50E+05	1,09E+05	8,03E+04	4,72E+04	1,14E+04	1,47	4,0E-02
Ser	94	8,44	115,7	1,24E+05	1,13E+05	1,03E+05	9,26E+04	9,78E+04	6,50E+04	4,60E+04	3,44E+04	2,44E+04	5,19E+03	1,53	7,2E-02
Asp	95	8,37	121,9	1,10E+05	1,07E+05	9,59E+04	8,91E+04	9,25E+04	6,17E+04	4,58E+04	3,28E+04	1,97E+04	4,20E+03	1,47	7,8E-02
Asn	96	8,29	119,5	8,60E+04	8,00E+04	7,28E+04	6,79E+04	7,22E+04	4,72E+04	3,59E+04	2,71E+04	1,80E+04	4,88E+03	1,39	7,9E-02
Phe	98	8,16	118,1	7,90E+04	7,40E+04	6,89E+04	6,03E+04	6,25E+04	4,41E+04	3,22E+04	2,29E+04	1,25E+04	2,87E+03	1,51	4,3E-02
Ser	99	7,94	116,0	5,91E+04	5,71E+04	4,84E+04	4,35E+04	4,69E+04	3,19E+04	2,48E+04	1,61E+04	1,25E+04	3,25E+03	1,47	7,9E-02
<b>Lys</b>	<b>100</b>	<b>8,28</b>	<b>122,8</b>	<b>1,11E+05</b>	<b>1,01E+05</b>	<b>9,08E+04</b>	<b>8,36E+04</b>	<b>8,47E+04</b>	<b>5,70E+04</b>	<b>4,35E+04</b>	<b>2,82E+04</b>	<b>1,73E+04</b>	<b>4,34E+03</b>	<b>1,58</b>	<b>6,2E-02</b>
Leu	101	8,05	121,6	5,06E+04	4,96E+04	4,03E+04	3,87E+04	3,98E+04	2,67E+04	1,96E+04	1,30E+04	1,18E+04	1,58E+03	1,53	1,1E-01
<b>Tyr</b>	<b>102</b>	<b>8,07</b>	<b>120,8</b>	<b>3,39E+05</b>	<b>3,24E+05</b>	<b>2,89E+05</b>	<b>2,48E+05</b>	<b>2,45E+05</b>	<b>1,79E+05</b>	<b>1,31E+05</b>	<b>9,81E+04</b>	<b>5,20E+04</b>	<b>1,46E+04</b>	<b>1,59</b>	<b>2,4E-02</b>
Lys	103	8,10	123,2	4,72E+04	4,26E+04	3,93E+04	3,63E+04	3,48E+04	2,52E+04	1,60E+04	1,26E+04	9,15E+03	2,02E+03	1,57	7,8E-02
Glu	104	8,43	122,0	7,82E+04	6,68E+04	6,56E+04	5,67E+04	5,73E+04	4,09E+04	2,99E+04	2,10E+04	1,17E+04	3,28E+03	1,56	5,8E-02
Thr	105	8,27	116,1	7,90E+04	7,40E+04	6,48E+04	6,40E+04	6,57E+04	4,35E+04	3,25E+04	2,57E+04	1,42E+04	5,99E+03	1,4	8,7E-02

Rés.	Num.	$\delta^1\text{H}$ (ppm)	$\delta^{15}\text{N}$ (ppm)	I <sub>0</sub>	I <sub>1</sub>	I <sub>2</sub>	I <sub>3</sub>	I <sub>3(bis)</sub>	I <sub>4</sub>	I <sub>5</sub>	I <sub>6</sub>	I <sub>7</sub>	I <sub>8</sub>	R <sub>1</sub> (s <sup>-1</sup> )	$\Delta\text{R}_1$ (s <sup>-1</sup> )
Ile	106	8,24	123,6	8,08E+04	7,27E+04	6,54E+04	5,81E+04	6,01E+04	4,01E+04	3,24E+04	2,15E+04	1,24E+04	4,03E+03	1,58	5,4E-02
Glu	107	8,51	125,4	9,09E+04	8,30E+04	7,77E+04	6,73E+04	6,76E+04	4,85E+04	3,47E+04	2,65E+04	1,61E+04	3,65E+03	1,53	2,9E-02
Thr	108	8,12	115,1	1,24E+05	1,16E+05	1,05E+05	9,63E+04	9,74E+04	6,99E+04	4,99E+04	3,83E+04	2,37E+04	5,13E+03	1,45	4,5E-02
Phe	109	8,36	122,1	1,36E+05	1,31E+05	1,20E+05	1,06E+05	1,05E+05	7,09E+04	5,67E+04	3,58E+04	2,34E+04	5,70E+03	1,56	5,9E-02
Asp	110	8,40	122,0	1,93E+05	1,82E+05	1,69E+05	1,52E+05	1,52E+05	1,05E+05	7,70E+04	6,00E+04	3,22E+04	9,98E+03	1,49	4,7E-02
<b>Asn</b>	<b>111</b>	<b>8,41</b>	<b>119,4</b>	<b>4,43E+05</b>	<b>4,13E+05</b>	<b>3,80E+05</b>	<b>3,54E+05</b>	<b>3,62E+05</b>	<b>2,54E+05</b>	<b>1,88E+05</b>	<b>1,40E+05</b>	<b>8,41E+04</b>	<b>2,58E+04</b>	<b>1,39</b>	<b>5,7E-02</b>
Asn	112	8,54	119,4	2,05E+05	1,92E+05	1,72E+05	1,59E+05	1,60E+05	1,11E+05	8,24E+04	5,66E+04	3,66E+04	9,36E+03	1,51	5,4E-02
Glu	113	8,47	121,2	2,49E+05	2,37E+05	2,16E+05	1,92E+05	1,98E+05	1,41E+05	1,02E+05	7,59E+04	4,31E+04	1,33E+04	1,47	3,9E-02
<b>Glu</b>	<b>114</b>	<b>8,45</b>	<b>121,6</b>	<b>3,47E+05</b>	<b>3,23E+05</b>	<b>2,91E+05</b>	<b>2,50E+05</b>	<b>2,53E+05</b>	<b>1,74E+05</b>	<b>1,31E+05</b>	<b>9,35E+04</b>	<b>5,49E+04</b>	<b>1,49E+04</b>	<b>1,63</b>	<b>2,7E-02</b>
Glu	115	8,50	121,9	2,52E+04	2,48E+04	2,31E+04	2,13E+04	2,01E+04	1,34E+04	1,29E+04	6,78E+03	4,54E+03	-2,67E+03	1,52	1,7E-01
Ser	116	8,43	117,1	1,48E+05	1,40E+05	1,26E+05	1,15E+05	1,22E+05	8,48E+04	6,57E+04	4,98E+04	3,30E+04	9,57E+03	1,33	5,1E-02
Ser	117	8,35	117,9	9,65E+04	9,03E+04	8,19E+04	7,70E+04	8,06E+04	5,53E+04	4,26E+04	3,18E+04	2,09E+04	5,34E+03	1,34	6,4E-02
Tyr	118	8,17	122,2	8,04E+04	7,53E+04	7,21E+04	6,98E+04	6,79E+04	4,35E+04	3,52E+04	2,62E+04	1,62E+04	5,51E+03	1,37	9,8E-02
Ser	119	8,20	118,0	3,78E+04	3,29E+04	2,99E+04	2,66E+04	3,01E+04	1,89E+04	1,21E+04	1,07E+04	6,71E+03	2,88E+03	1,57	1,2E-01
Met	206	8,20	120,2	2,16E+04	1,94E+04	1,92E+04	1,69E+04	1,83E+04	9,81E+03	1,09E+04	6,16E+03	5,91E+03	1,02E+03	1,34	1,5E-01
Ala	207	8,22	124,8	5,30E+04	4,74E+04	4,43E+04	4,29E+04	4,51E+04	2,94E+04	2,33E+04	1,49E+04	8,37E+03	4,66E+03	1,41	1,1E-01
Lys	208	8,24	120,5	9,16E+04	8,46E+04	7,98E+04	7,15E+04	7,42E+04	5,32E+04	3,82E+04	2,66E+04	1,83E+04	4,65E+03	1,42	5,8E-02
Asp	209	8,48	121,8	1,42E+05	1,32E+05	1,20E+05	1,09E+05	1,12E+05	7,82E+04	5,82E+04	4,36E+04	2,55E+04	6,83E+03	1,45	4,2E-02
Thr	210	8,20	114,4	1,92E+05	1,82E+05	1,67E+05	1,54E+05	1,57E+05	1,07E+05	8,11E+04	5,94E+04	3,73E+04	1,21E+04	1,42	6,0E-02
Ser	211	8,45	118,2	2,10E+05	1,91E+05	1,76E+05	1,64E+05	1,70E+05	1,12E+05	8,30E+04	6,10E+04	4,16E+04	1,21E+04	1,46	7,7E-02
Asp	212	8,45	122,3	2,88E+05	2,64E+05	2,46E+05	2,24E+05	2,30E+05	1,57E+05	1,15E+05	8,53E+04	5,03E+04	1,58E+04	1,48	5,7E-02
Glu	213	8,30	121,0	4,17E+05	3,88E+05	3,56E+05	3,23E+05	3,30E+05	2,32E+05	1,69E+05	1,25E+05	6,71E+04	2,11E+04	1,48	4,7E-02
Val	214	8,21	121,3	4,07E+05	3,84E+05	3,55E+05	3,12E+05	3,17E+05	2,30E+05	1,70E+05	1,25E+05	7,12E+04	1,98E+04	1,46	2,6E-02
<b>Ser</b>	<b>215</b>	<b>8,42</b>	<b>119,4</b>	<b>3,92E+05</b>	<b>3,68E+05</b>	<b>3,33E+05</b>	<b>2,94E+05</b>	<b>2,97E+05</b>	<b>2,09E+05</b>	<b>1,55E+05</b>	<b>1,15E+05</b>	<b>6,88E+04</b>	<b>1,99E+04</b>	<b>1,52</b>	<b>2,9E-02</b>
<b>Leu</b>	<b>216</b>	<b>8,35</b>	<b>124,3</b>	<b>3,36E+05</b>	<b>3,13E+05</b>	<b>2,91E+05</b>	<b>2,69E+05</b>	<b>2,74E+05</b>	<b>1,91E+05</b>	<b>1,41E+05</b>	<b>1,04E+05</b>	<b>6,10E+04</b>	<b>1,73E+04</b>	<b>1,43</b>	<b>6,0E-02</b>
Asn	217	8,45	120,1	2,72E+05	2,61E+05	2,32E+05	2,15E+05	2,22E+05	1,50E+05	1,12E+05	8,27E+04	5,25E+04	1,67E+04	1,44	6,3E-02
Thr	219	8,31	113,8	3,80E+05	3,54E+05	3,20E+05	2,91E+05	2,99E+05	2,02E+05	1,49E+05	1,09E+05	6,56E+04	1,87E+04	1,52	5,2E-02
Ser	220	8,30	117,7	1,89E+05	1,73E+05	1,55E+05	1,45E+05	1,50E+05	1,03E+05	7,45E+04	5,48E+04	3,34E+04	9,81E+03	1,48	6,3E-02
Glu	221	8,50	123,1	2,64E+05	2,47E+05	2,20E+05	2,04E+05	2,05E+05	1,39E+05	1,04E+05	7,58E+04	4,42E+04	1,25E+04	1,52	5,3E-02
Lys	222	8,33	122,1	3,25E+05	3,03E+05	2,79E+05	2,54E+05	2,55E+05	1,72E+05	1,28E+05	8,80E+04	5,03E+04	1,31E+04	1,56	5,8E-02
Leu	223	8,28	123,0	3,50E+05	3,28E+05	2,94E+05	2,63E+05	2,65E+05	1,83E+05	1,31E+05	9,43E+04	5,37E+04	1,77E+04	1,6	3,9E-02
Asn	224	8,44	119,1	2,63E+05	2,48E+05	2,23E+05	2,04E+05	2,10E+05	1,43E+05	1,03E+05	7,45E+04	4,74E+04	1,14E+04	1,51	6,0E-02

Rés.	Num.	$\delta^1\text{H}$ (ppm)	$\delta^{15}\text{N}$ (ppm)	I <sub>0</sub>	I <sub>1</sub>	I <sub>2</sub>	I <sub>3</sub>	I <sub>3(bis)</sub>	I <sub>4</sub>	I <sub>5</sub>	I <sub>6</sub>	I <sub>7</sub>	I <sub>8</sub>	R <sub>1</sub> (s <sup>-1</sup> )	$\Delta\text{R}_1$ (s <sup>-1</sup> )
Asn	225	8,41	119,1	1,95E+05	1,80E+05	1,63E+05	1,53E+05	1,59E+05	1,07E+05	7,85E+04	5,79E+04	3,70E+04	9,40E+03	1,45	7,3E-02
Leu	226	8,24	121,9	4,03E+05	3,76E+05	3,42E+05	3,04E+05	3,07E+05	2,11E+05	1,54E+05	1,08E+05	6,33E+04	1,56E+04	1,59	3,9E-02
Leu	227	8,18	121,9	4,19E+05	3,91E+05	3,54E+05	3,16E+05	3,15E+05	2,28E+05	1,67E+05	1,15E+05	6,24E+04	1,55E+04	1,56	3,1E-02
Glu	228	8,30	121,3	5,01E+05	4,67E+05	4,27E+05	3,75E+05	3,80E+05	2,71E+05	1,95E+05	1,43E+05	8,16E+04	2,30E+04	1,54	2,5E-02
Gly	229	8,46	109,8	4,22E+05	3,90E+05	3,59E+05	3,29E+05	3,33E+05	2,34E+05	1,74E+05	1,28E+05	7,99E+04	2,69E+04	1,43	4,6E-02
Asn	230	8,40	118,7	3,23E+05	3,01E+05	2,73E+05	2,53E+05	2,60E+05	1,76E+05	1,33E+05	9,82E+04	6,15E+04	1,99E+04	1,44	6,0E-02
Asp	231	8,54	121,2	4,13E+05	3,86E+05	3,52E+05	3,29E+05	3,37E+05	2,29E+05	1,70E+05	1,26E+05	8,10E+04	2,10E+04	1,43	6,6E-02
Ser	232	8,32	115,7	4,57E+05	4,27E+05	3,86E+05	3,58E+05	3,68E+05	2,49E+05	1,86E+05	1,33E+05	8,41E+04	2,49E+04	1,47	6,2E-02
Asp	233	8,40	122,1	4,72E+05	4,37E+05	3,99E+05	3,67E+05	3,77E+05	2,53E+05	1,85E+05	1,31E+05	7,66E+04	1,98E+04	1,53	6,5E-02
Asn	234	8,27	118,7	4,48E+05	4,19E+05	3,84E+05	3,57E+05	3,66E+05	2,51E+05	1,84E+05	1,35E+05	8,18E+04	2,32E+04	1,44	6,4E-02
Asp	235	8,37	120,8	5,20E+05	4,80E+05	4,40E+05	4,03E+05	4,12E+05	2,84E+05	2,08E+05	1,52E+05	8,66E+04	2,52E+04	1,49	5,4E-02
Leu	236	8,25	122,9	4,89E+05	4,58E+05	4,21E+05	3,76E+05	3,78E+05	2,69E+05	1,98E+05	1,45E+05	8,22E+04	2,31E+04	1,5	3,1E-02
Ser	237	8,46	116,9	4,87E+05	4,56E+05	4,19E+05	3,80E+05	3,86E+05	2,70E+05	1,98E+05	1,46E+05	8,54E+04	2,58E+04	1,47	4,5E-02
Leu	238	8,26	124,0	4,90E+05	4,61E+05	4,20E+05	3,71E+05	3,74E+05	2,68E+05	1,98E+05	1,45E+05	8,74E+04	2,51E+04	1,49	2,5E-02
Glu	239	8,27	120,6	5,91E+05	5,55E+05	5,17E+05	4,62E+05	4,66E+05	3,43E+05	2,59E+05	1,95E+05	1,14E+05	3,54E+04	1,37	2,3E-02
Asp	240	8,23	121,4	6,45E+05	6,13E+05	5,69E+05	5,10E+05	5,14E+05	3,90E+05	3,03E+05	2,33E+05	1,46E+05	5,38E+04	1,26	1,5E-02
Phe	241	7,64	124,9	7,67E+05	7,40E+05	7,00E+05	6,38E+05	6,35E+05	5,26E+05	4,30E+05	3,55E+05	2,42E+05	1,11E+05	0,97	4,7E-03

Les valeurs I<sub>0</sub>, I<sub>1</sub>, I<sub>2</sub>, I<sub>3</sub>, I<sub>3(bis)</sub>, I<sub>4</sub>, I<sub>5</sub>, I<sub>6</sub>, I<sub>7</sub> et I<sub>8</sub> correspondent respectivement aux valeurs d'intensité des pics de spectres ayant des délais de relaxation de 8, 55, 108, 208, 208, 408, 608, 808, 1208 et 2008 ms.

Les résidus en gras correspondent aux résidus qui n'ont pas été représentés sur les diagrammes en barres à cause de superpositions de pics sur les spectres <sup>1</sup>H<sup>15</sup>N-HSQC.

## 8.6. Données des expériences de relaxation $^{15}\text{N}$ $R_2$ sur P entière

Rés.	Num.	$\delta^1\text{H}$ (ppm)	$\delta^{15}\text{N}$ (ppm)	$I_0$	$I_1$	$I_2$	$I_3$	$I_3(\text{bis})$	$I_4$	$I_5$	$I_6$	$I_7$	$I_8$	$R_2$ (s $^{-1}$ )	$\Delta R_2$ (s $^{-1}$ )
Met	1	8,46	123,9	1,99E+05	1,74E+05	1,31E+05	9,22E+04	9,13E+04	7,10E+04	5,48E+04	3,95E+04	1,92E+04	7,23E+03	4,42	7,9E-02
Glu	2	8,32	122,7	2,19E+05	2,11E+05	1,57E+05	1,07E+05	1,04E+05	7,63E+04	5,59E+04	4,21E+04	2,13E+04	3,66E+03	4,51	2,2E-01
Lys	3	8,25	121,8	1,70E+05	1,59E+05	1,04E+05	6,80E+04	7,08E+04	4,65E+04	3,46E+04	2,33E+04	1,10E+04	2,96E+03	5,41	2,3E-01
Phe	4	8,24	121,0	1,52E+05	1,42E+05	9,90E+04	6,67E+04	6,28E+04	4,55E+04	3,06E+04	2,31E+04	1,55E+04	1,89E+03	5,11	2,2E-01
Ala	5	8,24	127,2	1,66E+05	1,64E+05	1,08E+05	7,82E+04	7,79E+04	5,00E+04	4,02E+04	3,05E+04	1,52E+04	4,29E+03	4,81	2,7E-01
Glu	7	8,57	120,1	2,01E+05	1,79E+05	1,28E+05	8,67E+04	8,47E+04	6,19E+04	4,07E+04	3,05E+04	1,67E+04	5,21E+03	5,06	1,3E-01
Phe	8	8,26	121,3	1,50E+05	1,35E+05	9,91E+04	6,47E+04	6,06E+04	4,55E+04	2,83E+04	2,43E+04	1,01E+04	5,20E+03	5,15	2,0E-01
Gly	10	7,86	109,6	6,43E+04	5,86E+04	4,09E+04	2,27E+04	2,13E+04	1,60E+04	9,97E+03	6,16E+03	5,59E+03	2,25E+03	6,00	3,7E-01
Glu	11	8,55	120,8	1,57E+05	1,53E+05	1,03E+05	7,04E+04	6,65E+04	4,74E+04	3,63E+04	2,44E+04	1,18E+04	1,39E+03	5,07	2,7E-01
Asp	12	8,48	121,4	1,71E+05	1,54E+05	1,05E+05	6,38E+04	6,00E+04	4,96E+04	2,98E+04	2,33E+04	9,08E+03	1,51E+03	5,68	2,6E-01
<b>Ala</b>	<b>13</b>	<b>8,36</b>	<b>124,3</b>	<b>2,84E+05</b>	<b>2,53E+05</b>	<b>1,71E+05</b>	<b>1,18E+05</b>	<b>1,18E+05</b>	<b>8,69E+04</b>	<b>6,11E+04</b>	<b>4,43E+04</b>	<b>2,57E+04</b>	<b>7,66E+03</b>	<b>5,10</b>	<b>1,2E-01</b>
Asn	14	8,42	116,3	9,95E+04	9,02E+04	5,31E+04	3,15E+04	2,91E+04	1,75E+04	1,25E+04	8,90E+03	4,85E+03	1,23E+03	6,93	3,7E-01
Asn	15	8,21	118,8	7,08E+04	6,06E+04	3,51E+04	1,78E+04	1,68E+04	1,11E+04	6,90E+03	2,55E+03	2,92E+03	-5,10E+02	7,88	3,9E-01
Arg	16	8,23	120,7	7,75E+04	6,48E+04	3,59E+04	2,09E+04	1,78E+04	1,07E+04	4,15E+03	5,65E+03	1,41E+03	-9,33E+02	8,13	3,6E-01
Ala	17	8,30	123,7	7,10E+04	5,50E+04	2,75E+04	1,17E+04	1,44E+04	7,11E+03	4,20E+03	2,22E+03	1,90E+03	1,95E+03	9,60	3,6E-01
Thr	18	8,08	113,4	7,19E+04	5,80E+04	3,21E+04	1,79E+04	1,65E+04	1,19E+04	6,82E+03	2,47E+03	-5,59E+02	2,97E+02	8,16	2,6E-01
<b>Lys</b>	<b>19</b>	<b>8,19</b>	<b>122,8</b>	<b>4,81E+04</b>	<b>4,71E+04</b>	<b>2,15E+04</b>	<b>1,56E+04</b>	<b>1,50E+04</b>	<b>9,46E+03</b>	<b>7,35E+03</b>	<b>1,24E+03</b>	<b>2,54E+03</b>	<b>1,25E+03</b>	<b>7,12</b>	<b>6,6E-01</b>
Phe	20	8,22	120,9	7,01E+04	5,22E+04	3,13E+04	1,82E+04	1,89E+04	1,11E+04	6,07E+03	2,19E+03	-6,67E+02	1,38E+03	7,94	1,9E-01
Leu	21	8,15	122,7	4,33E+04	3,89E+04	1,62E+04	1,28E+04	1,23E+04	7,41E+03	5,22E+03	4,65E+03	1,72E+03	-7,16E+02	7,62	6,4E-01
Glu	22	8,29	120,6	4,10E+04	3,39E+04	1,09E+04	4,81E+03	2,94E+03	2,09E+03	2,96E+02	6,65E+02	-1,14E+03	6,66E+02	12,32	1,3E+00
Ser	23	8,21	116,2	3,83E+04	2,87E+04	1,24E+04	5,92E+03	5,81E+03	4,36E+03	6,89E+02	-1,07E+02	1,75E+03	-7,36E+02	10,71	5,4E-01
Ile	24	7,95	121,5	3,27E+04	2,66E+04	1,16E+04	6,30E+03	7,31E+03	1,82E+03	2,09E+03	2,52E+03	4,10E+02	1,65E+02	9,56	6,0E-01
Lys	25	8,22	124,3	4,23E+04	3,28E+04	1,30E+04	7,56E+03	8,24E+03	3,23E+03	2,01E+03	5,52E+02	-1,40E+03	-2,49E+02	10,42	5,6E-01
Gly	26	8,37	109,6	3,70E+04	3,44E+04	1,80E+04	1,06E+04	1,03E+04	2,75E+03	4,99E+03	2,02E+03	-5,04E+02	2,16E+03	7,82	7,1E-01
Lys	27	8,03	120,4	6,90E+04	5,96E+04	3,33E+04	1,69E+04	1,84E+04	1,33E+04	6,92E+03	5,21E+03	1,71E+03	-3,54E+02	7,58	3,6E-01
Phe	28	8,36	121,1	6,67E+04	6,54E+04	3,48E+04	2,20E+04	2,12E+04	1,25E+04	7,44E+03	4,48E+03	6,96E+02	1,40E+03	6,98	5,4E-01
Thr	29	8,12	116,7	4,97E+04	5,66E+04	3,01E+04	1,95E+04	2,19E+04	1,40E+04	8,80E+03	7,38E+03	3,49E+03	1,29E+03	5,63	6,5E-01
<b>Ser</b>	<b>30</b>	<b>8,44</b>	<b>120,0</b>	<b>9,39E+04</b>	<b>7,75E+04</b>	<b>4,83E+04</b>	<b>2,65E+04</b>	<b>2,64E+04</b>	<b>1,65E+04</b>	<b>1,28E+04</b>	<b>6,64E+03</b>	<b>5,10E+03</b>	<b>3,88E+02</b>	<b>7,12</b>	<b>2,3E-01</b>
Lys	32	8,39	121,2	1,32E+05	1,32E+05	8,40E+04	5,75E+04	5,76E+04	3,68E+04	2,74E+04	2,06E+04	5,28E+03	3,47E+03	5,28	3,5E-01
Asp	33	8,36	122,8	1,46E+05	1,34E+05	8,97E+04	5,98E+04	5,74E+04	4,14E+04	2,47E+04	1,97E+04	7,93E+03	5,15E+03	5,50	2,2E-01

Rés.	Num.	$\delta^1\text{H}$ (ppm)	$\delta^{15}\text{N}$ (ppm)	I <sub>0</sub>	I <sub>1</sub>	I <sub>2</sub>	I <sub>3</sub>	I <sub>3(bis)</sub>	I <sub>4</sub>	I <sub>5</sub>	I <sub>6</sub>	I <sub>7</sub>	I <sub>8</sub>	R <sub>2</sub> (s <sup>-1</sup> )	$\Delta\text{R}_2$ (s <sup>-1</sup> )
Lys	35	8,43	120,0	2,35E+05	2,06E+05	1,43E+05	9,34E+04	9,00E+04	6,50E+04	4,49E+04	3,17E+04	1,70E+04	4,67E+03	5,45	1,3E-01
Lys	36	8,12	122,3	1,51E+05	1,40E+05	9,52E+04	6,70E+04	6,25E+04	4,17E+04	3,09E+04	2,33E+04	7,83E+03	7,31E+02	5,25	2,1E-01
Lys	37	8,42	123,4	1,16E+05	1,13E+05	8,13E+04	5,54E+04	5,30E+04	3,97E+04	3,06E+04	2,39E+04	1,31E+04	1,50E+03	4,55	2,3E-01
<b>Asp</b>	<b>38</b>	<b>8,43</b>	<b>121,6</b>	<b>5,62E+05</b>	<b>5,36E+05</b>	<b>3,78E+05</b>	<b>2,65E+05</b>	<b>2,60E+05</b>	<b>1,86E+05</b>	<b>1,41E+05</b>	<b>9,87E+04</b>	<b>5,29E+04</b>	<b>1,81E+04</b>	<b>4,67</b>	<b>1,9E-01</b>
Ser	39	8,26	115,8	1,47E+05	1,45E+05	1,08E+05	8,09E+04	7,25E+04	5,88E+04	4,49E+04	3,49E+04	2,18E+04	9,50E+03	3,95	2,0E-01
Ile	40	8,22	122,6	2,27E+05	2,22E+05	1,68E+05	1,22E+05	1,15E+05	8,97E+04	7,03E+04	5,20E+04	2,86E+04	1,19E+04	4,02	1,8E-01
Ile	41	8,29	125,3	2,36E+05	2,11E+05	1,53E+05	1,12E+05	1,12E+05	8,27E+04	6,41E+04	4,97E+04	2,80E+04	9,73E+03	4,34	7,0E-02
Ser	42	8,46	120,8	1,83E+05	1,85E+05	1,32E+05	9,64E+04	8,96E+04	6,70E+04	5,15E+04	4,06E+04	2,37E+04	5,02E+03	4,27	2,5E-01
Val	43	8,31	122,1	2,25E+05	2,29E+05	1,61E+05	1,25E+05	1,16E+05	8,78E+04	7,10E+04	5,71E+04	3,00E+04	9,86E+03	3,98	2,2E-01
Asn	44	8,55	122,0	1,29E+05	1,31E+05	8,98E+04	6,50E+04	6,28E+04	4,94E+04	3,47E+04	2,75E+04	1,74E+04	6,05E+03	4,31	2,5E-01
Ser	45	8,33	116,6	1,20E+05	1,23E+05	9,11E+04	6,39E+04	6,23E+04	5,01E+04	3,94E+04	2,89E+04	1,84E+04	3,57E+03	3,93	2,5E-01
Ile	46	8,15	121,8	3,03E+05	2,87E+05	2,11E+05	1,64E+05	1,58E+05	1,18E+05	9,18E+04	7,19E+04	3,90E+04	1,25E+04	3,97	1,2E-01
Asp	47	8,36	124,1	3,08E+05	2,77E+05	2,16E+05	1,60E+05	1,57E+05	1,21E+05	8,83E+04	6,70E+04	3,70E+04	1,19E+04	4,05	7,8E-02
<b>Ile</b>	<b>48</b>	<b>8,05</b>	<b>120,8</b>	<b>2,63E+05</b>	<b>2,43E+05</b>	<b>2,03E+05</b>	<b>1,57E+05</b>	<b>1,55E+05</b>	<b>1,23E+05</b>	<b>9,18E+04</b>	<b>7,48E+04</b>	<b>4,26E+04</b>	<b>1,16E+04</b>	<b>3,40</b>	<b>1,2E-01</b>
Glu	49	8,47	124,9	2,89E+05	2,55E+05	2,01E+05	1,45E+05	1,43E+05	1,09E+05	8,19E+04	6,09E+04	3,52E+04	1,07E+04	4,13	7,2E-02
Val	50	8,27	122,1	2,83E+05	2,72E+05	1,98E+05	1,53E+05	1,49E+05	1,11E+05	8,36E+04	6,37E+04	3,48E+04	1,15E+04	4,03	1,5E-01
Thr	51	8,30	118,6	1,97E+05	1,89E+05	1,40E+05	1,02E+05	9,74E+04	7,11E+04	5,52E+04	3,95E+04	2,17E+04	8,76E+03	4,28	1,8E-01
Lys	52	8,42	124,2	2,07E+05	1,96E+05	1,47E+05	1,05E+05	9,98E+04	8,32E+04	6,01E+04	4,83E+04	2,89E+04	1,03E+04	4,06	1,6E-01
Glu	53	8,47	122,1	3,36E+05	3,33E+05	2,27E+05	1,75E+05	1,65E+05	1,22E+05	9,77E+04	7,39E+04	4,05E+04	1,55E+04	4,26	2,1E-01
Ser	54	8,49	118,9	2,15E+05	2,02E+05	1,63E+05	1,16E+05	1,11E+05	8,90E+04	6,53E+04	5,15E+04	3,08E+04	1,21E+04	3,86	1,5E-01
Ile	56	8,33	121,0	2,92E+05	2,82E+05	2,13E+05	1,63E+05	1,62E+05	1,29E+05	1,00E+05	8,09E+04	4,91E+04	1,92E+04	3,57	1,1E-01
Thr	57	8,27	118,1	1,66E+05	1,74E+05	1,26E+05	9,30E+04	8,64E+04	7,02E+04	5,34E+04	4,22E+04	2,74E+04	1,06E+04	3,85	2,6E-01
Ser	58	8,40	118,0	6,88E+04	7,33E+04	5,70E+04	4,26E+04	3,94E+04	3,33E+04	2,92E+04	2,09E+04	1,36E+04	5,71E+03	3,28	2,4E-01
Asn	59	8,55	121,0	4,67E+04	4,23E+04	3,35E+04	2,65E+04	2,32E+04	1,82E+04	1,54E+04	1,08E+04	6,16E+03	3,75E+03	3,82	1,5E-01
Ser	60	8,35	116,1	7,09E+04	7,40E+04	5,93E+04	4,53E+04	4,16E+04	3,56E+04	2,88E+04	2,50E+04	1,65E+04	7,74E+03	3,05	2,0E-01
Thr	61	8,25	116,7	1,51E+05	1,60E+05	1,20E+05	9,11E+04	8,78E+04	6,92E+04	5,67E+04	4,68E+04	3,11E+04	1,35E+04	3,37	2,3E-01
Ile	62	8,17	124,1	2,82E+05	2,67E+05	2,15E+05	1,61E+05	1,55E+05	1,25E+05	9,21E+04	7,34E+04	4,35E+04	1,89E+04	3,62	1,3E-01
Ile	63	8,31	125,9	2,61E+05	2,47E+05	1,81E+05	1,45E+05	1,44E+05	1,10E+05	8,67E+04	6,98E+04	3,94E+04	1,50E+04	3,67	9,7E-02
Asn	64	8,64	124,4	2,11E+05	1,98E+05	1,51E+05	1,06E+05	1,02E+05	8,33E+04	6,22E+04	5,02E+04	3,01E+04	1,13E+04	4,04	1,5E-01
Thr	66	8,30	113,8	3,40E+05	3,24E+05	2,60E+05	1,93E+05	1,87E+05	1,55E+05	1,16E+05	9,22E+04	5,88E+04	2,43E+04	3,52	1,2E-01
Asn	67	8,46	120,7	1,79E+05	1,77E+05	1,38E+05	9,85E+04	9,26E+04	7,35E+04	5,62E+04	4,68E+04	2,70E+04	1,44E+04	3,81	2,1E-01
Glu	68	8,51	121,5	2,56E+05	2,67E+05	2,06E+05	1,55E+05	1,48E+05	1,20E+05	9,09E+04	7,58E+04	4,93E+04	2,07E+04	3,42	2,1E-01

Rés.	Num.	$\delta^1\text{H}$ (ppm)	$\delta^{15}\text{N}$ (ppm)	I <sub>0</sub>	I <sub>1</sub>	I <sub>2</sub>	I <sub>3</sub>	I <sub>3(bis)</sub>	I <sub>4</sub>	I <sub>5</sub>	I <sub>6</sub>	I <sub>7</sub>	I <sub>8</sub>	R <sub>2</sub> (s <sup>-1</sup> )	$\Delta\text{R}_2$ (s <sup>-1</sup> )
Thr	69	8,25	114,3	3,23E+05	3,14E+05	2,57E+05	1,85E+05	1,79E+05	1,50E+05	1,13E+05	9,64E+04	6,47E+04	2,85E+04	3,38	1,6E-01
Asp	70	8,37	122,7	3,42E+05	3,37E+05	2,55E+05	1,87E+05	1,79E+05	1,43E+05	1,10E+05	8,61E+04	4,91E+04	1,99E+04	3,83	1,7E-01
<b>Asp</b>	<b>71</b>	<b>8,42</b>	<b>121,6</b>	<b>4,22E+05</b>	<b>4,37E+05</b>	<b>3,62E+05</b>	<b>2,82E+05</b>	<b>2,67E+05</b>	<b>2,18E+05</b>	<b>1,73E+05</b>	<b>1,37E+05</b>	<b>8,08E+04</b>	<b>3,52E+04</b>	<b>3,11</b>	<b>2,0E-01</b>
Asn	72	8,45	118,7	2,57E+05	2,51E+05	2,02E+05	1,46E+05	1,36E+05	1,15E+05	8,68E+04	7,21E+04	4,49E+04	2,32E+04	3,54	1,9E-01
Ala	73	8,20	123,6	3,11E+05	2,95E+05	2,33E+05	1,65E+05	1,60E+05	1,32E+05	1,00E+05	7,69E+04	4,97E+04	2,07E+04	3,77	1,5E-01
Gly	74	8,38	106,9	2,03E+05	1,92E+05	1,52E+05	1,06E+05	9,90E+04	8,41E+04	6,13E+04	5,01E+04	3,27E+04	1,29E+04	3,88	1,8E-01
Asn	75	8,18	118,2	1,71E+05	1,73E+05	1,32E+05	9,47E+04	8,88E+04	7,21E+04	5,62E+04	4,52E+04	2,74E+04	1,26E+04	3,77	2,2E-01
Lys	76	8,21	122,3	2,24E+05	2,20E+05	1,59E+05	1,17E+05	1,12E+05	8,47E+04	6,32E+04	5,07E+04	3,08E+04	9,86E+03	4,15	1,9E-01
Asn	78	8,54	118,5	8,32E+04	7,89E+04	5,41E+04	3,40E+04	3,24E+04	2,46E+04	1,65E+04	1,14E+04	5,93E+03	9,54E+02	5,37	2,8E-01
Tyr	79	8,09	120,7	1,42E+05	1,23E+05	9,04E+04	5,20E+04	5,17E+04	3,89E+04	2,10E+04	1,80E+04	1,02E+04	1,56E+03	5,69	2,6E-01
Gln	80	8,23	122,1	8,73E+04	8,44E+04	4,62E+04	3,13E+04	3,05E+04	1,79E+04	1,36E+04	1,26E+04	-6,96E+02	-4,46E+02	6,39	4,8E-01
<b>Arg</b>	<b>81</b>	<b>8,27</b>	<b>122,8</b>	<b>8,03E+04</b>	<b>8,04E+04</b>	<b>4,71E+04</b>	<b>3,31E+04</b>	<b>2,97E+04</b>	<b>2,33E+04</b>	<b>1,61E+04</b>	<b>9,33E+03</b>	<b>4,87E+03</b>	<b>3,20E+03</b>	<b>5,62</b>	<b>4,0E-01</b>
Lys	82	8,43	124,4	9,82E+04	9,06E+04	4,98E+04	3,36E+04	3,36E+04	2,07E+04	1,29E+04	7,66E+03	3,41E+03	-2,35E+02	6,64	3,5E-01
<b>Leu</b>	<b>84</b>	<b>8,42</b>	<b>122,8</b>	<b>1,49E+05</b>	<b>1,37E+05</b>	<b>8,83E+04</b>	<b>5,84E+04</b>	<b>6,00E+04</b>	<b>3,82E+04</b>	<b>2,76E+04</b>	<b>1,78E+04</b>	<b>7,79E+03</b>	<b>3,08E+03</b>	<b>5,63</b>	<b>2,2E-01</b>
Val	85	8,15	121,1	1,13E+05	1,05E+05	5,95E+04	3,89E+04	3,86E+04	2,10E+04	1,20E+04	1,28E+04	2,26E+03	2,00E+03	6,66	3,9E-01
Ser	86	8,35	119,5	7,55E+04	6,35E+04	4,14E+04	2,25E+04	2,21E+04	1,48E+04	6,09E+03	4,11E+03	1,43E+03	5,70E+01	7,10	3,0E-01
<b>Phe</b>	<b>87</b>	<b>8,33</b>	<b>122,7</b>	<b>1,92E+05</b>	<b>1,74E+05</b>	<b>1,15E+05</b>	<b>7,19E+04</b>	<b>7,11E+04</b>	<b>4,96E+04</b>	<b>3,39E+04</b>	<b>2,42E+04</b>	<b>9,00E+03</b>	<b>2,16E+03</b>	<b>5,78</b>	<b>2,1E-01</b>
Lys	88	8,21	123,5	1,04E+05	9,33E+04	6,09E+04	3,37E+04	3,33E+04	2,50E+04	1,43E+04	8,94E+03	5,66E+03	7,04E+02	6,39	3,1E-01
Glu	89	8,40	122,4	1,16E+05	1,08E+05	6,70E+04	4,39E+04	4,13E+04	2,58E+04	1,70E+04	1,09E+04	3,41E+03	2,22E+03	6,19	3,0E-01
Asp	90	8,47	123,5	1,49E+05	1,27E+05	8,16E+04	5,22E+04	5,07E+04	3,41E+04	1,75E+04	1,20E+04	5,72E+03	-3,02E+02	6,41	1,7E-01
<b>Ile</b>	<b>92</b>	<b>8,35</b>	<b>122,8</b>	<b>2,23E+05</b>	<b>2,07E+05</b>	<b>1,40E+05</b>	<b>9,78E+04</b>	<b>9,07E+04</b>	<b>7,34E+04</b>	<b>5,06E+04</b>	<b>3,93E+04</b>	<b>2,25E+04</b>	<b>8,80E+03</b>	<b>4,92</b>	<b>2,0E-01</b>
Ser	94	8,43	115,7	9,37E+04	8,44E+04	5,00E+04	3,39E+04	3,06E+04	1,92E+04	1,41E+04	7,29E+03	3,16E+03	3,51E+02	6,50	2,8E-01
Asp	95	8,37	121,9	8,64E+04	8,41E+04	4,77E+04	3,00E+04	2,71E+04	1,59E+04	1,14E+04	7,22E+03	-1,10E+03	2,62E+02	6,79	5,2E-01
Asn	96	8,29	119,5	6,70E+04	5,79E+04	3,07E+04	1,90E+04	1,76E+04	9,34E+03	7,10E+03	4,26E+03	2,32E+03	-2,15E+03	7,74	3,8E-01
Phe	98	8,16	118,1	5,00E+04	4,15E+04	2,00E+04	1,05E+04	1,04E+04	5,74E+03	5,65E+03	5,51E+03	1,05E+03	8,07E+02	8,70	5,7E-01
Ser	99	7,94	116,0	4,02E+04	2,92E+04	1,44E+04	5,16E+03	3,80E+03	6,11E+02	3,33E+01	2,52E+03	-3,17E+02	2,17E+03	11,66	8,9E-01
<b>Lys</b>	<b>100</b>	<b>8,28</b>	<b>122,8</b>	<b>8,86E+04</b>	<b>8,37E+04</b>	<b>4,19E+04</b>	<b>2,59E+04</b>	<b>2,46E+04</b>	<b>1,53E+04</b>	<b>1,10E+04</b>	<b>7,03E+03</b>	<b>1,63E+03</b>	<b>-1,29E+03</b>	<b>7,41</b>	<b>5,4E-01</b>
Leu	101	8,05	121,6	4,32E+04	1,82E+04	1,90E+04	9,56E+02	2,44E+03	6,20E+03	-1,49E+03	-7,52E+01	3,02E+03	-1,03E+03	13,13	3,0E+00
<b>Tyr</b>	<b>102</b>	<b>8,07</b>	<b>120,8</b>	<b>2,48E+05</b>	<b>2,05E+05</b>	<b>1,33E+05</b>	<b>8,73E+04</b>	<b>8,77E+04</b>	<b>6,08E+04</b>	<b>4,53E+04</b>	<b>2,82E+04</b>	<b>1,54E+04</b>	<b>1,59E+03</b>	<b>5,96</b>	<b>1,0E-01</b>
Lys	103	8,10	123,2	3,10E+04	1,84E+04	6,41E+03	3,31E+03	3,54E+03	-4,15E+02	-2,68E+03	1,50E+03	4,03E+02	-2,97E+03	14,97	1,7E+00
Glu	104	8,43	122,0	4,76E+04	3,67E+04	1,60E+04	7,71E+03	9,95E+03	2,01E+03	1,70E+03	-8,34E+02	-4,61E+02	-4,68E+02	10,47	5,8E-01
Thr	105	8,27	116,1	4,83E+04	4,04E+04	1,70E+04	1,07E+04	1,10E+04	5,56E+03	3,40E+03	2,52E+03	9,03E+02	1,13E+03	9,08	5,4E-01

Rés.	Num.	$\delta^1\text{H}$ (ppm)	$\delta^{15}\text{N}$ (ppm)	I <sub>0</sub>	I <sub>1</sub>	I <sub>2</sub>	I <sub>3</sub>	I <sub>3(bis)</sub>	I <sub>4</sub>	I <sub>5</sub>	I <sub>6</sub>	I <sub>7</sub>	I <sub>8</sub>	R <sub>2</sub> (s <sup>-1</sup> )	$\Delta\text{R}_2$ (s <sup>-1</sup> )
Ile	106	8,24	123,6	6,14E+04	4,28E+04	2,15E+04	7,63E+03	8,51E+03	5,31E+03	1,96E+03	1,16E+03	2,36E+02	-8,10E+02	11,08	3,2E-01
Glu	107	8,51	125,4	6,31E+04	4,53E+04	2,44E+04	1,07E+04	1,26E+04	4,61E+03	2,45E+03	1,59E+02	9,83E+01	-1,16E+03	9,97	2,8E-01
Thr	108	8,12	115,1	8,87E+04	7,03E+04	3,88E+04	2,00E+04	2,04E+04	1,05E+04	6,85E+03	2,07E+03	-5,16E+02	-8,00E+02	8,64	2,5E-01
Phe	109	8,36	122,1	1,01E+05	8,67E+04	4,59E+04	2,81E+04	2,57E+04	1,50E+04	6,68E+03	1,28E+03	-6,90E+02	-1,53E+03	8,07	4,3E-01
Asp	110	8,40	122,0	1,54E+05	1,40E+05	7,82E+04	4,96E+04	5,05E+04	2,42E+04	2,06E+04	1,02E+04	5,39E+03	1,84E+03	6,97	3,8E-01
<b>Asn</b>	<b>111</b>	<b>8,40</b>	<b>119,4</b>	<b>3,12E+05</b>	<b>3,02E+05</b>	<b>2,43E+05</b>	<b>1,71E+05</b>	<b>1,61E+05</b>	<b>1,32E+05</b>	<b>9,55E+04</b>	<b>7,57E+04</b>	<b>4,29E+04</b>	<b>1,29E+04</b>	<b>3,87</b>	<b>2,0E-01</b>
Asn	112	8,54	119,4	1,69E+05	1,54E+05	1,04E+05	6,52E+04	6,65E+04	4,39E+04	3,13E+04	1,94E+04	1,28E+04	3,83E+03	5,61	2,1E-01
Glu	113	8,47	121,2	2,07E+05	1,86E+05	1,17E+05	7,56E+04	7,29E+04	4,94E+04	3,17E+04	1,94E+04	1,02E+04	1,45E+03	6,14	2,1E-01
Glu	114	8,45	121,6	3,56E+05	3,00E+05	1,82E+05	1,10E+05	1,12E+05	7,09E+04	4,74E+04	2,69E+04	1,40E+04	1,38E+03	6,83	1,5E-01
<b>Glu</b>	<b>115</b>	<b>8,48</b>	<b>122,0</b>	<b>2,26E+04</b>	<b>1,33E+04</b>	<b>8,97E+03</b>	<b>2,52E+03</b>	<b>4,91E+03</b>	<b>1,92E+03</b>	<b>-1,90E+03</b>	<b>1,94E+03</b>	<b>-1,70E+03</b>	<b>-1,22E+02</b>	<b>10,70</b>	<b>1,4E+00</b>
Ser	116	8,43	117,1	1,12E+05	9,98E+04	5,83E+04	3,56E+04	3,20E+04	2,04E+04	1,26E+04	6,66E+03	3,54E+03	7,58E+02	7,09	3,4E-01
Ser	117	8,35	117,9	7,05E+04	6,21E+04	3,53E+04	1,85E+04	1,54E+04	1,04E+04	4,54E+03	3,93E+03	-9,71E+02	-3,85E+02	8,07	5,2E-01
Tyr	118	8,17	122,2	4,44E+04	3,88E+04	1,43E+04	1,01E+04	5,25E+03	2,24E+03	7,25E+02	8,97E+02	-2,72E+03	-3,06E+01	10,55	1,1E+00
Ser	119	8,20	118,0	2,15E+04	1,55E+04	5,08E+03	3,31E+03	4,66E+03	-9,47E+02	-1,38E+03	6,31E+02	-1,20E+03	-2,67E+03	12,28	1,8E+00
Met	206	8,19	120,2	1,18E+04	4,28E+03	1,07E+03	-2,29E+03	-2,86E+02	1,42E+03	-1,52E+02	-1,25E+03	2,37E+03	2,35E+03	29,95	2,6E+00
Ala	207	8,22	124,8	3,46E+04	2,08E+04	1,20E+04	1,12E+03	1,54E+03	1,44E+03	-1,04E+03	3,35E+03	-1,76E+02	-5,20E+02	13,66	1,6E+00
Lys	208	8,24	120,5	5,95E+04	4,85E+04	2,16E+04	1,16E+04	1,24E+04	6,44E+03	4,11E+03	3,67E+03	2,20E+03	1,41E+03	9,32	5,0E-01
Asp	209	8,47	121,8	1,06E+05	9,06E+04	4,96E+04	2,78E+04	2,83E+04	1,55E+04	5,20E+03	4,48E+03	1,10E+03	6,04E+02	8,00	3,8E-01
Thr	210	8,20	114,4	1,53E+05	1,38E+05	9,43E+04	6,07E+04	5,73E+04	4,33E+04	2,99E+04	1,92E+04	9,42E+03	-8,08E+02	5,51	2,0E-01
Ser	211	8,44	118,2	1,71E+05	1,62E+05	1,13E+05	7,20E+04	7,14E+04	5,16E+04	3,41E+04	2,53E+04	1,41E+04	4,60E+03	5,17	2,5E-01
Asp	212	8,44	122,3	2,34E+05	2,26E+05	1,55E+05	1,11E+05	1,06E+05	7,97E+04	6,03E+04	4,10E+04	2,33E+04	6,44E+03	4,67	2,0E-01
Glu	213	8,30	121,0	3,36E+05	3,21E+05	2,46E+05	1,83E+05	1,80E+05	1,33E+05	1,05E+05	7,97E+04	4,55E+04	1,40E+04	3,91	1,4E-01
Val	214	8,20	121,3	3,23E+05	3,01E+05	2,31E+05	1,83E+05	1,79E+05	1,35E+05	1,09E+05	8,37E+04	4,64E+04	1,87E+04	3,66	8,6E-02
<b>Ser</b>	<b>215</b>	<b>8,42</b>	<b>119,4</b>	<b>2,95E+05</b>	<b>2,64E+05</b>	<b>1,71E+05</b>	<b>1,08E+05</b>	<b>1,07E+05</b>	<b>7,78E+04</b>	<b>5,52E+04</b>	<b>3,79E+04</b>	<b>1,84E+04</b>	<b>4,71E+03</b>	<b>5,76</b>	<b>2,0E-01</b>
<b>Leu</b>	<b>216</b>	<b>8,35</b>	<b>124,3</b>	<b>3,02E+05</b>	<b>2,92E+05</b>	<b>2,30E+05</b>	<b>1,69E+05</b>	<b>1,64E+05</b>	<b>1,32E+05</b>	<b>9,71E+04</b>	<b>7,49E+04</b>	<b>4,53E+04</b>	<b>2,15E+04</b>	<b>3,72</b>	<b>1,6E-01</b>
Asn	217	8,44	120,1	2,38E+05	2,33E+05	1,81E+05	1,34E+05	1,26E+05	1,04E+05	7,74E+04	6,34E+04	3,98E+04	1,82E+04	3,66	1,7E-01
<b>Thr</b>	<b>219</b>	<b>8,31</b>	<b>113,8</b>	<b>3,20E+05</b>	<b>3,03E+05</b>	<b>2,25E+05</b>	<b>1,54E+05</b>	<b>1,51E+05</b>	<b>1,18E+05</b>	<b>8,79E+04</b>	<b>6,46E+04</b>	<b>4,01E+04</b>	<b>1,47E+04</b>	<b>4,34</b>	<b>1,7E-01</b>
Ser	220	8,30	117,7	1,60E+05	1,64E+05	1,26E+05	9,25E+04	8,84E+04	6,74E+04	5,34E+04	4,11E+04	2,41E+04	1,02E+04	3,74	2,3E-01
Glu	221	8,50	123,1	2,25E+05	2,22E+05	1,78E+05	1,29E+05	1,24E+05	9,95E+04	7,60E+04	5,90E+04	3,57E+04	1,61E+04	3,63	1,8E-01
Lys	222	8,33	122,1	2,73E+05	2,74E+05	1,97E+05	1,53E+05	1,44E+05	1,12E+05	8,83E+04	6,66E+04	3,84E+04	1,67E+04	3,88	1,9E-01
Leu	223	8,28	123,0	2,91E+05	2,73E+05	2,14E+05	1,64E+05	1,58E+05	1,25E+05	9,63E+04	7,67E+04	4,51E+04	1,74E+04	3,65	9,1E-02
Asn	224	8,44	119,1	2,40E+05	2,26E+05	1,82E+05	1,30E+05	1,28E+05	1,02E+05	7,50E+04	5,98E+04	3,86E+04	1,56E+04	3,75	1,3E-01

Rés.	Num.	$\delta^1\text{H}$ (ppm)	$\delta^{15}\text{N}$ (ppm)	$I_0$	$I_1$	$I_2$	$I_3$	$I_{3(\text{bis})}$	$I_4$	$I_5$	$I_6$	$I_7$	$I_8$	$R_2$ (s <sup>-1</sup> )	$\Delta R_2$ (s <sup>-1</sup> )
Asn	225	8,41	119,1	1,73E+05	1,67E+05	1,30E+05	9,36E+04	8,80E+04	7,53E+04	5,54E+04	4,39E+04	2,92E+04	1,36E+04	3,74	1,7E-01
Leu	226	8,24	121,9	3,28E+05	3,16E+05	2,39E+05	1,83E+05	1,76E+05	1,41E+05	1,10E+05	8,82E+04	5,25E+04	1,87E+04	3,67	1,2E-01
Leu	227	8,18	121,9	3,41E+05	3,20E+05	2,53E+05	1,98E+05	1,93E+05	1,51E+05	1,22E+05	9,51E+04	5,79E+04	2,16E+04	3,46	8,0E-02
Glu	228	8,29	121,3	4,10E+05	3,97E+05	3,10E+05	2,50E+05	2,45E+05	1,96E+05	1,64E+05	1,31E+05	8,32E+04	3,65E+04	3,14	8,7E-02
Gly	229	8,45	109,8	3,41E+05	3,68E+05	2,89E+05	2,29E+05	2,23E+05	1,80E+05	1,49E+05	1,29E+05	8,86E+04	4,43E+04	2,84	1,9E-01
Asn	230	8,40	118,7	2,84E+05	2,98E+05	2,53E+05	2,00E+05	1,88E+05	1,62E+05	1,30E+05	1,09E+05	7,92E+04	4,13E+04	2,65	1,7E-01
Asp	231	8,54	121,2	3,45E+05	3,75E+05	3,13E+05	2,53E+05	2,42E+05	2,02E+05	1,73E+05	1,45E+05	1,05E+05	5,99E+04	2,47	1,8E-01
Ser	232	8,32	115,7	3,89E+05	4,03E+05	3,38E+05	2,75E+05	2,62E+05	2,29E+05	1,92E+05	1,63E+05	1,15E+05	6,25E+04	2,46	1,3E-01
Asp	233	8,39	122,1	4,11E+05	4,37E+05	3,38E+05	2,77E+05	2,66E+05	2,19E+05	1,87E+05	1,59E+05	1,05E+05	5,33E+04	2,78	1,7E-01
Asn	234	8,27	118,7	3,88E+05	3,98E+05	3,40E+05	2,70E+05	2,60E+05	2,22E+05	1,81E+05	1,49E+05	1,06E+05	5,63E+04	2,62	1,4E-01
Asp	235	8,37	120,8	4,40E+05	4,48E+05	3,56E+05	2,86E+05	2,76E+05	2,36E+05	1,98E+05	1,68E+05	1,17E+05	5,79E+04	2,74	1,2E-01
Leu	236	8,24	122,9	4,06E+05	3,99E+05	3,30E+05	2,66E+05	2,66E+05	2,15E+05	1,77E+05	1,46E+05	9,81E+04	4,56E+04	2,81	9,5E-02
Ser	237	8,45	116,9	4,28E+05	4,21E+05	3,33E+05	2,78E+05	2,71E+05	2,27E+05	1,90E+05	1,60E+05	1,03E+05	5,05E+04	2,78	8,3E-02
Leu	238	8,25	124,0	4,26E+05	4,00E+05	3,34E+05	2,74E+05	2,71E+05	2,32E+05	1,90E+05	1,57E+05	1,09E+05	5,50E+04	2,68	3,7E-02
Glu	239	8,26	120,6	4,87E+05	4,64E+05	3,89E+05	3,26E+05	3,25E+05	2,73E+05	2,32E+05	1,91E+05	1,37E+05	6,85E+04	2,51	3,8E-02
Asp	240	8,22	121,4	5,39E+05	5,22E+05	4,36E+05	3,78E+05	3,68E+05	3,17E+05	2,77E+05	2,38E+05	1,71E+05	9,49E+04	2,26	4,3E-02
Phe	241	7,63	124,9	6,77E+05	6,36E+05	5,69E+05	4,95E+05	4,96E+05	4,42E+05	3,87E+05	3,50E+05	2,79E+05	1,70E+05	1,77	1,9E-02

Les valeurs  $I_0$ ,  $I_1$ ,  $I_2$ ,  $I_3$ ,  $I_{3(\text{bis})}$ ,  $I_4$ ,  $I_5$ ,  $I_6$ ,  $I_7$  et  $I_8$  correspondent respectivement aux valeurs d'intensité des pics de spectres ayant des délais de relaxation de 34, 68, 136, 204, 204, 271, 339, 407, 543 et 814 ms.

Les résidus en gras correspondent aux résidus qui n'ont pas été représentés sur les diagrammes en barres à cause de superpositions de pics sur les spectres  $^1\text{H}^{15}\text{N}$ -HSQC.

## 8.7. Déplacements chimiques attribués de P<sub>OD</sub>

Les déplacements chimiques attribués qui suivent correspondent à ceux fait avec le domaine P<sub>OD</sub> purifié et concentré à 1 mM dans de l'eau avec un tampon 20 mM phosphate, 100 mM NaCl et pH 6,5.

Num.	Res.	H	N	ND	HA	HB	HG	HD	CO	CA	CB	CG	CD	CE
125	G	-	-	-	3,8	-	-	-	-	43,62	-	-	-	-
126	S	-	-	-	4,49	3,86	-	-	-	58,17	64,09	-	-	-
127	T	-	-	-	4,26	4,2	1,15	-	-	62,53	69,59	21,66	-	-
128	N	-	-	112,58	4,59	2,70 ; 2,75	-	7,55 ; 6,87	173,59	54,06	38,68	-	-	-
129	D	8,21	120,76	-	4,48	2,61 ; 2,62	-	-	174,63	55,13	40,8	-	-	-
130	N	8,19	119,08	112,55	4,53	2,78 ; 2,85	-	6,90 ; 7,56	174,64	54,44	38,53	-	-	-
131	I	8,17	120,68	-	3,61	1,89	0,77 ; 0,98 ; 1,57	0,79	175,48	64,76	37,75	29,06 ; 17,11	13,12	-
132	T	8,24	115,77	-	3,53	4,07	1,1	-	173,61	67,46	67,62	22,42	-	-
133	A	7,82	122,22	-	4,12	1,4	-	-	179,1	55,14	18,11	-	-	-
134	R	7,47	116,64	-	3,94	1,72	1,69	3,08 ; 3,08	177,82	58,16	27,16	25,61	42,06	-
135	L	8,23	119,98	-	3,96	1,20 ; 1,96	1,83	0,71 ; 0,59	176,41	57,91	40,78	26,04	22,76 ; 27,16	-
136	D	8,62	120,88	-	4,21	2,64 ; 2,83	-	-	176,84	57,78	39,47	-	-	-
137	R	7,57	120,1	-	4,12	1,97 ; 1,90	1,75	3,22 ; 3,17	177,15	59,22	30,49	27,02	43,73	-
138	I	7,78	121,66	-	3,49	1,97	0,91 ; 0,91 ; 0,80	0,8	175,21	65,95	38,15	30,38 ; 16,26	14,66	-
139	D	8,5	120,49	-	4,21	2,71 ; 2,69	-	-	177,15	58,17	41,48	-	-	-
140	E	7,94	119,46	-	4,04	2,09 ; 2,22	2,23 ; 2,41	-	178,13	59,51	29,6	36,19	-	-
141	K	8,08	120,83	-	4,16	1,59 ; 2,11	1,74 ; 1,42	1,52 ; 1,48	177,54	59,42	33,04	26,05	29,56	42,14
142	L	8,63	118,95	-	3,96	1,15 ; 2,19	1,96	0,68 ; 0,94	177,1	58,16	41,6	26,49	24,35 ; 28,92	-
143	S	8,09	115,43	-	4,17	3,93 ; 4,03	-	-	175,17	62,51	62,37	-	-	-
144	E	7,86	124,8	-	4,04	2,12 ; 2,18	2,37 ; 2,12	-	177,72	59,48	29,66	36,16	-	-
145	I	8,46	121,41	-	3,4	2	0,74 ; 0,67 ; 0,67	0,81	175,35	65,81	38,15	16,70 ; 31,01	15,3	-
146	L	8,43	120,23	-	3,82	1,84 ; 1,52	1,52	0,85	177,09	58,9	42,01	27,86	25,06	-
147	G	8	105,66	-	3,94 ; 3,85	-	-	-	174,93	47,41	-	-	-	-
148	M	7,75	120,95	-	4,17	1,77 ; 2,19	2,71 ; 2,46	-	176,7	59,35	34,22	33,14	-	17,61
149	L	8,5	119,55	-	3,93	1,87 ; 1,49	1,71	0,86 ; 0,73	176,51	57,94	42,05	26,8	27,85 ; 25,65	-
150	H	8,45	118,26	-	4,14	3,21	-	-	175,21	60,56	29,27	-	-	-
151	T	7,57	114,47	-	3,82	4,38	1,19	-	174,34	66,35	68,71	21,92	-	-
152	L	7,7	121,33	-	4,02	1,50 ; 1,84	1,69	0,72 ; 0,69	177,2	57,44	42,15	26,58	23,11 ; 25,64	-
153	V	8,51	120,57	-	3,59	2,13	0,86 ; 0,94	-	176,71	66,16	31,78	22,79 ; 21,24	-	-
154	V	7,95	120,48	-	3,71	1,98	0,82 ; 0,81	-	176,05	65,33	31,68	21,10 ; 21,88	-	-

<b>Num.</b>	<b>Res.</b>	<b>H</b>	<b>N</b>	<b>ND</b>	<b>HA</b>	<b>HB</b>	<b>HG</b>	<b>HD</b>	<b>CO</b>	<b>CA</b>	<b>CB</b>	<b>CG</b>	<b>CD</b>	<b>CE</b>
155	A	7,73	122,16	-	4,16	1,4	-	-	176,83	53,81	18,55	-	-	-
156	S	7,73	112,77	-	4,29	3,81	-	-	172,34	59,38	63,83	-	-	-
157	A	7,75	124,44	-	4,32	1,37	-	-	175,88	52,61	19,33	-	-	-
158	G	7,92	107,37	-	4,07 ; 4,03	-	-	-	-	44,71	-	-	-	-
159	P	-	-	-	4,45	2,23 ; 1,91	1,96	3,57 ; 3,60	175,66	63,34	32,11	27,18	49,8	-
160	T	8,24	113,83	-	4,3	4,2	1,15	-	172,67	61,91	69,67	21,47	-	-
161	S	8,13	117,76	-	4,37	3,83 ; 3,80	-	-	171,94	58,24	63,88	-	-	-
162	A	8,21	126,46	-	4,28	1,31	-	-	174,57	52,51	19,22	-	-	-
163	R	7,83	125,48	-	4,08	1,63 ; 1,77	1,52	3,11	-	57,3	31,49	27,04	43,41	-

## 8.8. Données des expériences de PRE sur P[127-205](S143C)

Num.	Rés.	$\delta^1\text{H}$ (ppm)	$\delta^{15}\text{N}$ (ppm)	I	$\Delta\text{I}$	$\text{I}_0$	$\Delta\text{I}_0$	$\text{I}/\text{I}_0$	$\Delta\text{I}/\text{I}_0$
130	Asn	8,22	119,2	9,07E+03	2,7E+03	1,68E+04	5,2E+03	0,54	8,7E-03
132	Thr	8,24	115,9	2,16E+04	1,7E+03	2,68E+04	9,4E+02	0,81	3,4E-02
133	Ala	7,81	122,4	2,09E+04	7,6E+02	3,71E+04	2,3E+03	0,56	1,4E-02
134	Arg	7,53	116,6	9,00E+03	1,4E+03	1,39E+04	2,1E+03	0,65	1,2E-04
136	Asp	8,60	121,0	4,33E+03	1,6E+03	3,73E+04	5,4E+03	0,11	2,6E-02
137	Arg	7,60	119,8	-1,61E+02	6,4E+01	3,14E+04	4,7E+03	-0,01	1,3E-03
139	Asp	8,55	120,9	3,40E+03	2,5E+03	4,05E+04	4,9E+03	0,08	5,1E-02
140	Glu	8,02	119,2	3,65E+03	1,2E+03	2,11E+04	2,0E+03	0,17	4,3E-02
142	Leu	8,63	119,0	6,84E+02	2,6E+03	2,61E+04	2,5E+03	0,02	9,8E-02
143	Cys	8,10	118,8	2,08E+03	2,3E+00	2,18E+04	1,6E+02	0,10	5,9E-04
147	Gly	8,24	106,1	1,01E+03	4,3E+02	1,46E+04	1,4E+03	0,07	3,7E-02
149	Leu	8,54	119,8	3,28E+03	7,7E+02	1,85E+04	1,1E+03	0,18	3,1E-02
150	His	8,58	118,7	6,52E+02	2,4E+03	2,22E+04	3,5E+03	0,02	1,0E-01
151	Thr	7,68	114,5	8,90E+02	1,5E+03	2,56E+04	5,9E+03	0,03	5,3E-02
152	Leu	7,80	121,2	4,90E+03	4,5E+02	2,38E+04	2,8E+03	0,21	5,3E-03
153	Val	8,45	120,2	1,64E+04	4,7E+03	5,56E+04	1,1E+04	0,29	2,8E-02
154	Val	8,03	119,8	1,08E+04	1,6E+03	5,02E+04	6,7E+03	0,22	4,1E-03
155	Ala	7,89	121,9	2,03E+04	3,4E+03	1,15E+05	1,2E+04	0,18	1,2E-02
156	Ser	7,76	112,5	3,62E+04	8,0E+03	6,27E+04	1,1E+04	0,57	2,6E-02
157	Ala	7,80	124,1	4,22E+04	1,0E+04	6,11E+04	1,4E+04	0,69	1,1E-02
158	Gly	7,91	107,2	8,57E+04	2,0E+04	1,43E+05	2,5E+04	0,60	3,6E-02
160	Thr	8,22	113,5	1,43E+04	4,5E+03	1,85E+04	3,1E+03	0,76	1,1E-01
161	Ser	8,20	117,5	6,19E+03	2,4E+03	3,22E+04	1,4E+03	0,19	6,7E-02
162	Ala	8,19	125,6	7,35E+03	1,4E+01	1,24E+04	8,6E+02	0,60	4,3E-02
163	Arg	8,11	119,5	1,08E+04	2,9E+03	5,40E+04	9,2E+03	0,20	1,9E-02
164	Asp	8,21	120,5	5,31E+04	1,1E+04	1,10E+05	1,6E+04	0,48	3,3E-02
165	Gly	8,22	108,7	1,25E+04	3,2E+03	5,27E+04	9,9E+03	0,24	1,6E-02
166	Ile	7,90	120,4	1,62E+04	1,2E+03	1,40E+05	2,4E+04	0,12	1,1E-02
167	Arg	8,24	123,6	4,22E+03	2,0E+03	5,29E+04	7,6E+03	0,08	4,9E-02
168	Asp	8,17	120,4	2,95E+04	9,1E+03	9,13E+04	1,7E+04	0,32	3,9E-02
169	Ala	8,08	123,5	5,62E+03	2,6E+03	8,50E+04	1,2E+04	0,06	2,2E-02
170	Met	8,20	117,7	5,78E+03	1,9E+02	6,79E+04	9,4E+03	0,09	1,5E-02
171	Val	7,84	120,0	1,05E+04	7,7E+02	1,35E+05	2,2E+04	0,08	1,8E-02
172	Gly	8,27	110,6	2,86E+03	3,0E+03	2,85E+04	6,2E+03	0,09	8,6E-02
173	Leu	7,89	121,7	2,57E+04	6,1E+03	1,74E+05	3,1E+04	0,15	8,8E-03
174	Arg	8,21	121,0	2,08E+04	2,8E+03	8,55E+04	9,7E+03	0,24	5,8E-03
175	Glu	8,48	120,8	3,87E+03	1,8E+03	4,86E+04	9,2E+03	0,08	2,1E-02
176	Glu	8,42	120,2	2,23E+04	4,3E+03	8,33E+04	1,9E+04	0,27	8,3E-03
177	Met	8,11	119,7	1,00E+04	3,5E+03	9,82E+04	2,0E+04	0,10	1,5E-02
178	Ile	7,96	120,4	1,71E+04	3,3E+02	1,09E+05	1,5E+04	0,16	1,9E-02
179	Glu	8,09	121,5	1,84E+04	4,3E+03	8,43E+04	1,4E+04	0,22	1,5E-02
180	Lys	7,95	121,1	1,37E+04	1,3E+03	7,94E+04	1,0E+04	0,17	6,1E-03
181	Ile	7,96	120,0	9,15E+03	3,5E+03	1,02E+05	1,7E+04	0,09	2,0E-02
182	Arg	8,24	122,4	6,79E+03	8,0E+02	6,36E+04	8,6E+03	0,11	2,7E-02
183	Thr	8,11	114,4	7,11E+03	1,3E+03	3,37E+04	3,6E+03	0,21	1,6E-02
184	Glu	8,26	122,3	7,10E+03	3,7E+02	6,67E+04	8,0E+03	0,11	1,9E-02
185	Ala	8,07	123,0	5,24E+03	2,5E+02	6,85E+04	9,6E+03	0,08	7,1E-03
186	Leu	7,86	119,3	4,75E+02	9,4E+02	8,59E+04	1,3E+04	0,01	1,2E-02
187	Met	8,09	119,7	9,61E+03	1,3E+03	9,61E+04	1,6E+04	0,10	3,6E-03
188	Thr	7,89	114,1	4,70E+03	2,2E+03	3,99E+04	2,3E+03	0,12	4,9E-02
189	Asn	8,41	120,3	1,12E+04	5,0E+03	3,77E+04	4,3E+03	0,29	1,0E-01
190	Asp	8,30	120,2	2,70E+04	5,1E+03	8,06E+04	1,3E+04	0,33	9,7E-03
191	Arg	8,03	120,8	1,81E+04	3,6E+03	1,21E+05	1,5E+04	0,15	1,2E-02

Num.	Rés.	$\delta^1\text{H}$ (ppm)	$\delta^{15}\text{N}$ (ppm)	I	$\Delta\text{I}$	$\text{I}_0$	$\Delta\text{I}_0$	I/I <sub>0</sub>	$\Delta\text{I}/\text{I}_0$
192	Leu	8,05	121,5	2,01E+04	3,2E+03	1,66E+05	1,8E+04	0,12	6,1E-03
193	Glu	8,29	120,8	5,67E+04	1,3E+04	2,03E+05	2,7E+04	0,28	2,7E-02
194	Ala	7,97	122,7	3,06E+04	6,9E+03	1,29E+05	1,6E+04	0,24	2,4E-02
195	Met	8,00	117,4	2,15E+04	4,3E+03	1,14E+05	1,7E+04	0,19	1,0E-02
196	Ala	7,93	123,3	1,97E+04	4,0E+03	1,03E+05	1,3E+04	0,19	1,5E-02
197	Arg	7,88	118,4	2,98E+04	6,7E+03	1,09E+05	1,3E+04	0,27	2,9E-02
198	Leu	7,89	121,6	2,77E+04	6,4E+03	1,84E+05	2,3E+04	0,15	1,6E-02
199	Arg	8,03	120,4	3,93E+04	7,3E+03	1,54E+05	1,9E+04	0,25	1,5E-02
200	Asn	8,29	119,8	1,61E+04	3,9E+03	3,67E+04	2,8E+03	0,44	7,4E-02
201	Glu	8,38	121,2	9,51E+04	2,0E+04	1,31E+05	1,9E+04	0,72	5,2E-02
202	Glu	8,30	121,1	2,26E+05	4,8E+04	2,81E+05	3,3E+04	0,80	7,6E-02
203	Ser	8,12	116,0	1,55E+05	3,5E+04	1,57E+05	2,4E+04	0,98	7,8E-02
204	Glu	8,29	123,1	1,55E+05	3,9E+04	1,60E+05	2,0E+04	0,96	1,3E-01
205	Lys	7,80	127,1	1,19E+06	1,8E+05	1,22E+06	1,5E+05	0,98	2,3E-02

Les valeurs I et I<sub>0</sub> correspondent respectivement aux moyennes des valeurs d'intensité des signaux HN provenant de la protéine oxydée à l'IAP et de la protéine oxydée à l'IAP puis réduite. Plusieurs résidus n'ont pas été affichés car les signaux de corrélation HN n'étaient pas visible sur les spectres.

## 8.9. Données des expériences de PRE sur P[127-205](S156C)

Num.	Rés.	$\delta^1\text{H}$ (ppm)	$\delta^{15}\text{N}$ (ppm)	I	$\Delta\text{I}$	$\text{I}_0$	$\Delta\text{I}_0$	$\text{I}/\text{I}_0$	$\Delta\text{I}/\text{I}_0$
130	Asn	8,21	119,3	1,14E+04	2,0E+03	1,74E+04	1,9E+03	0,65	4,4E-02
132	Thr	8,26	115,9	4,77E+04	5,9E+03	3,99E+04	2,9E+03	1,19	6,1E-02
133	Ala	7,83	122,4	5,42E+04	7,9E+03	4,97E+04	3,9E+03	1,09	7,5E-02
134	Arg	7,50	116,7	3,41E+04	3,9E+03	2,92E+04	2,7E+02	1,17	1,2E-01
135	Leu	8,23	120,0	4,05E+04	6,7E+03	4,09E+04	3,9E+03	0,99	7,0E-02
136	Asp	8,62	121,0	4,81E+04	8,2E+03	4,34E+04	6,1E+02	1,11	1,7E-01
137	Arg	7,60	120,2	5,46E+04	8,2E+03	5,20E+04	3,1E+03	1,05	9,4E-02
138	Ile	7,85	121,8	3,88E+04	5,6E+03	4,95E+04	1,5E+04	0,81	1,3E-01
139	Asp	8,51	120,6	4,58E+04	8,1E+03	5,10E+04	7,7E+03	0,90	2,4E-02
140	Glu	8,01	119,7	5,12E+04	6,3E+03	6,06E+04	7,4E+03	0,85	1,3E-03
141	Lys	8,12	120,7	4,78E+04	5,3E+03	5,77E+04	8,1E+03	0,83	2,4E-02
142	Leu	8,66	119,1	3,38E+04	1,9E+03	3,59E+04	1,8E+03	0,94	4,2E-03
143	Ser	8,18	115,6	3,50E+04	6,8E+03	4,07E+04	1,9E+03	0,86	1,3E-01
144	Glu	7,91	124,6	2,48E+04	3,0E+03	4,03E+04	4,0E+03	0,61	1,4E-02
145	Ile	8,50	120,9	1,43E+04	3,1E+03	4,32E+04	2,6E+04	0,38	1,5E-01
146	Leu	8,47	120,2	2,60E+04	2,7E+03	4,61E+04	3,9E+03	0,56	1,1E-02
147	Gly	8,15	105,7	9,26E+03	8,0E+02	2,97E+04	9,3E+03	0,32	7,5E-02
148	Met	7,72	120,7	2,59E+03	1,3E+03	1,53E+04	1,1E+04	0,18	4,2E-02
149	Leu	8,47	119,6	2,61E+03	1,5E+03	2,28E+04	7,3E+03	0,13	1,1E-01
150	His	8,63	118,3	1,94E+03	1,7E+03	2,03E+04	1,0E+04	0,09	3,8E-02
151	Thr	7,69	114,3	1,43E+03	2,5E+03	1,83E+04	1,2E+04	0,04	1,1E-01
152	Leu	7,73	120,8	4,08E+02	9,4E+02	1,55E+04	1,1E+04	0,06	1,0E-01
157	Ala	7,98	126,2	1,34E+03	1,6E+03	1,06E+04	2,5E+03	0,15	1,8E-01
158	Gly	7,95	107,3	-9,72E+02	4,6E+02	5,57E+04	4,6E+04	-0,03	3,4E-02
160	Thr	8,20	113,4	1,91E+03	6,6E+02	1,24E+04	6,8E+03	0,16	3,7E-02
161	Ser	8,21	117,4	2,74E+03	1,1E+03	1,78E+04	7,3E+03	0,18	1,4E-01
162	Ala	8,17	125,3	1,17E+02	3,2E+02	9,89E+03	4,6E+03	0,00	3,0E-02
163	Arg	8,11	119,4	6,34E+03	7,9E+02	3,70E+04	2,0E+04	0,19	8,3E-02
164	Asp	8,21	120,5	1,71E+04	3,3E+03	7,73E+04	4,1E+04	0,24	8,5E-02
165	Gly	8,20	108,9	1,03E+04	3,5E+03	3,98E+04	2,8E+04	0,31	1,3E-01
166	Ile	7,90	120,3	7,91E+03	1,6E+03	1,19E+05	8,5E+04	0,08	4,6E-02
167	Arg	8,25	123,7	2,50E+03	2,3E+02	4,25E+04	3,0E+04	0,08	6,1E-02
168	Asp	8,17	120,5	1,52E+04	6,0E+02	6,50E+04	3,7E+04	0,28	1,5E-01
169	Ala	8,09	123,5	1,88E+03	4,7E+02	6,71E+04	4,9E+04	0,03	1,8E-02
170	Met	8,20	117,6	2,44E+03	2,8E+03	5,86E+04	4,1E+04	0,08	1,0E-01
171	Val	7,85	120,1	6,73E+03	2,5E+02	1,16E+05	8,2E+04	0,08	5,2E-02
172	Gly	8,29	110,7	1,73E+02	1,5E+03	2,91E+04	2,1E+04	-0,02	6,3E-02
173	Leu	7,89	121,7	4,54E+04	8,5E+03	1,57E+05	6,8E+04	0,33	2,0E-01
174	Arg	8,21	120,8	4,18E+04	1,1E+03	8,89E+04	4,1E+04	0,53	2,6E-01
175	Glu	8,48	120,6	2,52E+04	1,0E+04	4,05E+04	2,0E+04	0,64	6,6E-02
176	Glu	8,43	120,2	2,33E+04	2,9E+03	7,65E+04	3,4E+04	0,33	1,1E-01
177	Met	8,11	119,7	1,31E+04	1,9E+03	9,47E+04	5,8E+04	0,16	7,8E-02
178	Ile	7,96	120,4	9,68E+03	7,1E+02	8,75E+04	5,3E+04	0,13	7,1E-02
179	Glu	8,10	121,5	2,28E+04	5,3E+03	7,07E+04	2,8E+04	0,33	5,6E-02
180	Lys	7,95	121,0	1,64E+04	2,7E+03	7,33E+04	3,7E+04	0,25	8,7E-02
181	Ile	7,96	120,0	1,18E+04	7,5E+02	8,96E+04	5,6E+04	0,16	9,2E-02
182	Arg	8,25	122,3	1,36E+04	3,7E+03	7,19E+04	4,1E+04	0,21	6,8E-02
183	Thr	8,11	114,4	7,39E+03	2,8E+03	3,34E+04	1,8E+04	0,23	3,9E-02

Num.	Rés.	$\delta^1\text{H}$ (ppm)	$\delta^{15}\text{N}$ (ppm)	I	$\Delta\text{I}$	$\text{I}_0$	$\Delta\text{I}_0$	$\text{I}/\text{I}_0$	$\Delta\text{I}/\text{I}_0$
184	Glu	8,27	122,3	1,48E+04	1,7E+03	4,25E+04	8,3E+03	0,35	2,9E-02
185	Ala	8,08	123,0	1,10E+04	1,6E+03	8,12E+04	5,0E+04	0,16	7,9E-02
186	Leu	7,86	119,3	9,05E+03	1,2E+03	9,48E+04	6,4E+04	0,12	6,7E-02
187	Met	8,08	119,7	1,56E+04	3,7E+03	9,32E+04	5,7E+04	0,19	7,8E-02
188	Thr	7,89	114,1	1,61E+04	5,6E+03	4,60E+04	2,3E+04	0,36	5,7E-02
189	Asn	8,40	120,2	2,27E+04	7,0E+03	4,35E+04	1,6E+04	0,53	2,9E-02
190	Asp	8,31	120,2	5,48E+04	1,6E+04	8,59E+04	2,8E+04	0,64	3,0E-02
191	Arg	8,02	120,8	5,38E+04	1,1E+04	1,32E+05	5,8E+04	0,43	1,0E-01
192	Leu	8,06	121,4	5,28E+04	1,3E+04	1,73E+05	8,7E+04	0,33	8,9E-02
193	Glu	8,29	120,7	1,14E+05	2,6E+04	2,17E+05	8,0E+04	0,54	7,9E-02
194	Ala	7,97	122,7	6,87E+04	1,7E+04	1,31E+05	5,4E+04	0,54	9,3E-02
195	Met	8,00	117,4	4,50E+04	1,1E+04	1,18E+05	5,7E+04	0,40	1,0E-01
196	Ala	7,93	123,2	4,23E+04	9,6E+03	1,07E+05	5,3E+04	0,42	1,2E-01
197	Arg	7,88	118,3	5,87E+04	1,5E+04	1,24E+05	5,7E+04	0,50	1,1E-01
198	Leu	7,89	121,5	6,13E+04	2,2E+04	1,65E+05	9,5E+04	0,40	9,4E-02
199	Arg	8,03	120,3	7,43E+04	1,7E+04	1,46E+05	5,4E+04	0,52	8,1E-02
200	Asn	8,30	119,9	3,25E+04	1,5E+03	4,03E+04	4,4E+03	0,81	5,1E-02
201	Glu	8,38	121,2	1,27E+05	3,0E+04	1,36E+05	3,2E+04	0,93	2,3E-03
202	Glu	8,30	121,1	2,63E+05	6,2E+04	2,78E+05	6,5E+04	0,95	4,1E-03
203	Ser	8,12	116,0	1,67E+05	4,6E+04	1,56E+05	3,7E+04	1,07	4,3E-02
204	Glu	8,29	123,1	1,69E+05	4,5E+04	1,58E+05	3,4E+04	1,06	5,6E-02
205	Lys	7,80	127,1	1,21E+06	2,0E+05	1,19E+06	2,1E+05	1,02	1,5E-02

Les valeurs I et  $\text{I}_0$  correspondent respectivement aux moyennes des valeurs d'intensité des signaux HN provenant de la protéine oxydée à l'IAP et de la protéine oxydée à l'IAP puis réduite. Plusieurs résidus n'ont pas été affichés car les signaux de corrélation HN n'étaient pas visible sur les spectres.

## 8.10. Articles scientifiques publiés

1) Richard C-A, Rincheval V, Lassoued S, Fix J, **Cardone C**, Esneau C, Nekhai S, Galloux M, Rameix-Welti M-A, Sizun C\*, Eléouët J-F. RSV hijacks cellular protein phosphatase 1 to regulate M2-1 phosphorylation and viral transcription. *PLoS Pathog* (2018) 14(2):e1006920.

DOI: [10.1371/journal.ppat.1006920](https://doi.org/10.1371/journal.ppat.1006920).

2) Pereira N, **Cardone C**, Lassoued S, Galloux M, Fix J, Assrir N, Lescop E, Bontems F, Eléouët JF, Sizun C\*. New Insights into Structural Disorder in Human Respiratory Syncytial Virus Phosphoprotein and Implications for Binding of Protein Partners. *J Biol Chem* 292(6):2120-31 (2017).

DOI: [10.1074/jbc.M116.765958](https://doi.org/10.1074/jbc.M116.765958).

RESEARCH ARTICLE

# RSV hijacks cellular protein phosphatase 1 to regulate M2-1 phosphorylation and viral transcription

Charles-Adrien Richard<sup>1</sup>, Vincent Rincheval<sup>2</sup>, Safa Lassoued<sup>3</sup>, Jenna Fix<sup>1</sup>, Christophe Cardone<sup>3</sup>, Camille Esneau<sup>1</sup>, Sergei Nekhai<sup>4</sup>, Marie Galloux<sup>1</sup>, Marie-Anne Rameix-Welti<sup>2,5</sup>, Christina Sizun<sup>3\*</sup>, Jean-François Eléouët<sup>1\*</sup>

**1** Unité de Virologie et Immunologie Moléculaires (UR892), INRA, Université Paris-Saclay, Jouy-en-Josas, France, **2** UMR1173, INSERM, Université de Versailles St. Quentin, Montigny le Bretonneux, France, **3** Institut de Chimie des Substances Naturelles, CNRS, Université Paris-Saclay, Avenue de la Terrasse, Gif-sur-Yvette, France, **4** Center for Sickle Cell Disease and Department of Medicine, Howard University, Washington, D. C., United States of America, **5** AP-HP, Laboratoire de Microbiologie, Hôpital Ambroise Paré, Boulogne-Billancourt, France

\* [jean-francois.eleouet@inra.fr](mailto:jean-francois.eleouet@inra.fr) (J-FE); [christina.sizun@cnrs.fr](mailto:christina.sizun@cnrs.fr) (CS)



**OPEN ACCESS**

**Citation:** Richard C-A, Rincheval V, Lassoued S, Fix J, Cardone C, Esneau C, et al. (2018) RSV hijacks cellular protein phosphatase 1 to regulate M2-1 phosphorylation and viral transcription. *PLoS Pathog* 14(3): e1006920. <https://doi.org/10.1371/journal.ppat.1006920>

**Editor:** Matthias Johannes Schnell, Thomas Jefferson University, UNITED STATES

**Received:** October 28, 2017

**Accepted:** February 4, 2018

**Published:** February 28, 2018

**Copyright:** © 2018 Richard et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** This work was supported by the French Agence Nationale de la Recherche, specific programs ANR Blanc 2011 « Bronchioliteaser » [ANR-11-BSV8-0024] and ANR Blanc 2013 « Respi syncy » [ANR-13-ISV3-0007], and by NIH Research Grant U19AI109664 (to SN). Doctoral fellowships from Région Ile-de-France DIM Malinf to SL, CC and CC and access to high field NMR

## Abstract

Respiratory syncytial virus (RSV) RNA synthesis occurs in cytoplasmic inclusion bodies (IBs) in which all the components of the viral RNA polymerase are concentrated. In this work, we show that RSV P protein recruits the essential RSV transcription factor M2-1 to IBs independently of the phosphorylation state of M2-1. We also show that M2-1 dephosphorylation is achieved by a complex formed between P and the cellular phosphatase PP1. We identified the PP1 binding site of P, which is an RVxF-like motif located nearby and upstream of the M2-1 binding region. NMR confirmed both P-M2-1 and P-PP1 interaction regions in P. When the P-PP1 interaction was disrupted, M2-1 remained phosphorylated and viral transcription was impaired, showing that M2-1 dephosphorylation is required, in a cyclic manner, for efficient viral transcription. IBs contain substructures called inclusion bodies associated granules (IBAGs), where M2-1 and neo-synthesized viral mRNAs concentrate. Disruption of the P-PP1 interaction was correlated with M2-1 exclusion from IBAGs, indicating that only dephosphorylated M2-1 is competent for viral mRNA binding and hence for a previously proposed post-transcriptional function.

## Author summary

Respiratory syncytial virus (RSV) is the leading cause of lower respiratory tract illness in infants. Since no vaccine and no potent antivirals are available against RSV, it is essential to better understand the mechanisms of viral replication to develop new antiviral strategies. Here we have investigated the mechanisms by which two essential components of the viral RNA polymerase machinery, the phosphoprotein P and the M2-1 transcription factor, interact and function. We identified the amino acid residues of P critical for this interaction and showed that they are required for P recruiting M2-1 to

spectrometers financially supported by IR-RMN-THC FR3050 CNRS are kindly acknowledged. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

cytoplasmic inclusions, where viral polymerase complex proteins concentrate and viral RNA synthesis occurs. We also showed that M2-1 dephosphorylation, required for viral transcription, is achieved by a complex formed between P and the cellular phosphatase PP1. The region of P binding to PP1 is located nearby and upstream of the M2-1 binding domain. This is the first report showing that the phosphoprotein of a negative strand RNA virus can hijack a cellular phosphatase to modulate the phosphorylation state of its partners. These two P regions interacting with M2-1 and PP1 are also potential targets for future antiviral therapy.

## Introduction

Human respiratory syncytial virus (RSV) is the leading cause of severe respiratory tract infections in infants worldwide and the primary cause of infant hospitalization for respiratory infections [1]. In addition, RSV is increasingly recognized as a significant cause of disease in the elderly population and can often be fatal for patients with a compromised immune system [2]. The virus belongs to the *Orthopneumovirus* genus of the *Pneumoviridae* family, order *Mononegavirales* [3]. The RSV genome is a single strand, negative sense RNA of about 15 kb that is packaged by the nucleoprotein (N) and maintained as a left-handed helical N-RNA ribonucleoprotein complex (RNP) [4–6]. This RNP is the template for two distinct activities: RNA replication that generates genomic and antigenomic RNA, which is encapsidated by N immediately after synthesis [7, 8], and RNA transcription that generates 10 capped and poly-adenylated mRNAs, which are not encapsidated by N. Both activities are carried out by the viral RNA-dependent RNA polymerase complex (RdRp) [9]. The viral N, P (phosphoprotein) and L (large polymerase) proteins are the essential components of the RdRp. RSV P is the main cofactor of the large polymerase L protein. In particular, by interacting with L and the RNP, P is essential to properly position the L protein for RNA synthesis [10].

RSV transcription is dependent on a fourth viral protein, M2-1 [11]. The transcriptase complex first engages promoter sequences that lie at the 3' end of the genome [12]. Transcription proceeds through sequential stop-and-restart events, in which the RdRp recognizes gene start (GS) and gene end (GE) sequences, that flank each gene and direct initiation and termination of transcription, respectively [13]. The transcriptase complex has the propensity to dissociate from the RNP template, but cannot reinitiate at a downstream gene after a premature termination. This leads to a decreasing gradient of mRNA from the 3' to the 5' end of the genome [14]. In contrast, the highly processive replicase bypasses GS and GE signals to produce complete genomic and antigenomic RNAs [15]. The exact mechanism of how RdRp differentiates between transcription and replication still remains unknown.

By increasing the processivity of the RdRp complex, RSV transcription antiterminator protein M2-1 prevents premature transcription termination [16, 17]. For this activity, M2-1 has to be recruited to cytoplasmic inclusion bodies (IBs), which contain other components of the RdRp complex, notably N, P and L, and [18–20], [21],[22]. Moreover we recently showed that IBs are a place of viral RNA synthesis and that M2-1 and viral mRNAs concentrate in IB substructures called IB associated granules (IBAGs), from which N, P, L and genomic RNA are absent [20]. M2-1 is a 22 kDa basic protein that forms stable tetramers in solution [23, 24]. Each protomer features an N-terminal zinc finger domain, an  $\alpha$ -helical tetramerization motif, and a C-terminal  $\alpha$ -helical core domain [24]. M2-1 is an RNA binding protein [25] that binds preferentially to RSV mRNAs and A-rich sequences [26]. RSV M2-1 can also interact with RSV P. We showed previously that substitution of M2-1 residues involved in the M2-1–P interaction

prevented the recruitment of M2-1 to IBs, suggesting that formation of a P–M2-1 complex is critical for M2-1 recruitment to IBs [19]. Since M2-1 interacts with RNA and P in a competitive manner through the core domain [19, 23], it is expected that these interactions are regulated in a cyclic manner.

In RSV-infected cells M2-1 exists in different phosphorylation states, resulting in its migration as two major bands in SDS-PAGE [27–29]. The slower-migrating species contains phosphorylated M2-1, whereas the faster-migrating species lacks significant phosphorylation [27]. In RSV infected cells or when co-expressed with P, M2-1 protein remains mainly unphosphorylated, whereas M2-1 is mainly phosphorylated when expressed alone [25]. Using recombinant (unphosphorylated) M2-1 produced in *E. coli*, it has also been shown that M2-1 can be phosphorylated *in vitro* by casein kinase I on serines S58 and S61 [26]. Abolishing phosphorylation of these residues by alanine substitution impaired the antitermination function of M2-1 [26]. However, the P–M2-1 interaction appears to be independent of the phosphorylation status of M2-1 [23]. On the other hand, phosphorylated M2-1 has reduced RNA binding capacities [25]. All these data point to the critical role of dynamic and reversible M2-1 phosphorylation for its function in transcription.

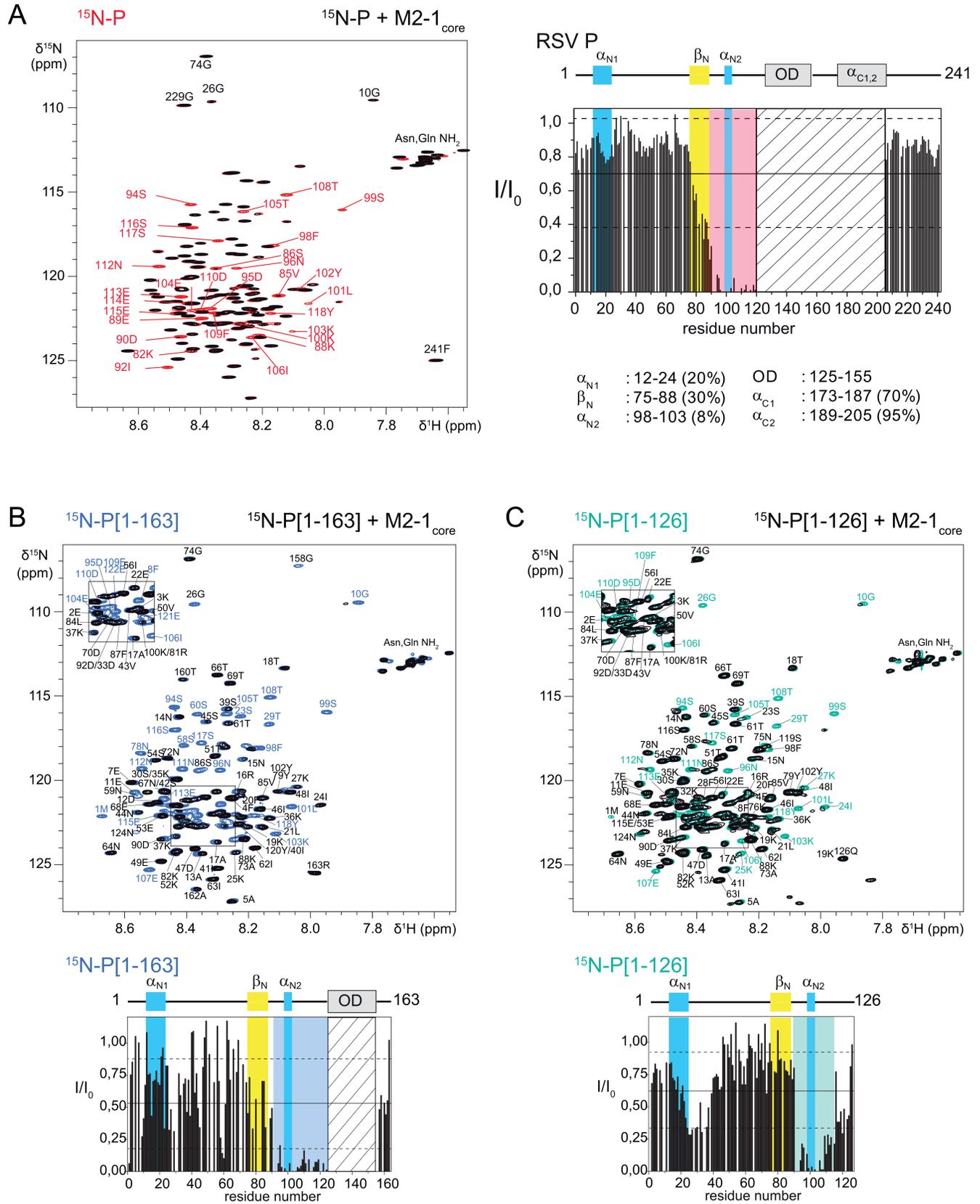
Neither P nor M2-1 produced in *E. coli* are phosphorylated [23, 30, 31]. These two unphosphorylated proteins interact together and have been used previously to study the P–M2-1 interactions *in vitro* [23, 32]. In P, the region encompassing residues 100–120 and more specifically residues L101, Y102, and F109 were reported to be critical for the M2-1–P interaction and for efficient transcription [33]. It was also shown that P residue T105 is probably involved in M2-1 binding [34]. P can be phosphorylated on several serine and threonine residues with different turnover rates [34–39]. In particular phosphorylation of T108, which occurs with a high-turnover, would prevent M2-1 binding. However, the role of P phosphorylation remains unclear, since phosphorylation is not required for viral transcription or replication [38, 40, 41]. Here we further investigated structural and functional aspects of the M2-1–P interaction. We show that a region encompassing residues 93–110 of P is required for the presence of M2-1 in IBs. We further identified another element, located upstream of this region which is responsible for M2-1 dephosphorylation. We show that this region is involved in the binding of the cellular protein phosphatase-1 (PP1) to P, but not in M2-1 binding, and that the complex formed by P and PP1 is responsible for M2-1 dephosphorylation, a key process for efficient viral transcription.

## Results

### Mapping of M2-1<sub>core</sub> interaction regions on P by NMR

We previously mapped the interaction surface of RSV P on the core domain of M2-1 (M2-1<sub>core</sub>, residues 58–177) by NMR [19], by observing spectral perturbations, at a residue-specific level, induced by unlabeled P on <sup>15</sup>N-labeled M2-1<sub>core</sub>. RSV P forms highly stable tetramers [31, 42, 43], with large N- and C-terminal intrinsically disordered regions (IDRs) flanking the oligomerization domain (residues 126–163) [31, 32]. We recently analyzed the propensities of these IDRs to form transient secondary structures and to transiently associate, either with each other or with the N protein, also by NMR [44]. Here we proceeded to determine the M2-1 binding region in P.

Both P and M2-1, produced in *E. coli*, were unphosphorylated. We compared <sup>1</sup>H-<sup>15</sup>N spectra of <sup>15</sup>N-labeled P in the absence and presence of unlabeled M2-1<sub>core</sub>. The presence of M2-1<sub>core</sub> induced a significant decrease of NMR signal intensities, or line broadening, in a region spanning residues 90–120 in the N-terminal IDR of P, upstream of the oligomerization domain (Fig 1A). This finding is in agreement with the previous localization of residues critical for M2-1



**Fig 1. Mapping of the M2-1<sub>core</sub> interaction regions on P by NMR.** (A) Superimposed <sup>1</sup>H-<sup>15</sup>N HSQC spectra of 35 μM <sup>15</sup>N-labeled P, alone (red contours) and in the presence of 1 molar equivalent of M2-1<sub>core</sub> (black contours). Residue-specific assignments are represented by red labels for peaks with an intensity decrease by more than 60% on addition of M2-1<sub>core</sub>. The bar diagram represents the intensity ratios (I/I<sub>0</sub>) measured for each peak. Straight and broken lines represent the mean and mean ± rmsd values over all signals. The position of the oligomerization domain and regions with transient α-helical or β-sheet secondary structure along the sequence of P, previously determined by NMR [44], is indicated in the cartoon on top. The boundaries and maximal propensity of these secondary structures [44] are indicated below. The red background indicates the region around the transient helix α<sub>N2</sub> where M2-1<sub>core</sub> induces >60% intensity decrease. The hatched area corresponds to a P region encompassing the OD and C-terminal helices for which peaks cannot be observed due to severe line broadening without M2-1<sub>core</sub>. (B) and (C) Superimposed <sup>1</sup>H-<sup>15</sup>N HSQC spectra of 35 μM <sup>15</sup>N-labeled P[1–126] and 75 μM P[1–163], alone (cyan and blue contours) and in the presence of 1 molar equivalent of M2-1<sub>core</sub> (black contours). Residue-specific assignments are given for all peaks. Insets on the top left give assignments for the crowded central region of the spectra. These spectra were acquired with less resolution in the <sup>15</sup>N dimension than the spectrum shown in A. Bar diagrams below represent the intensity ratios. The regions around α<sub>N2</sub> with I/I<sub>0</sub> < (mean-rmsd) are highlighted with light cyan and blue backgrounds for P[1–126] and P[1–163], respectively.

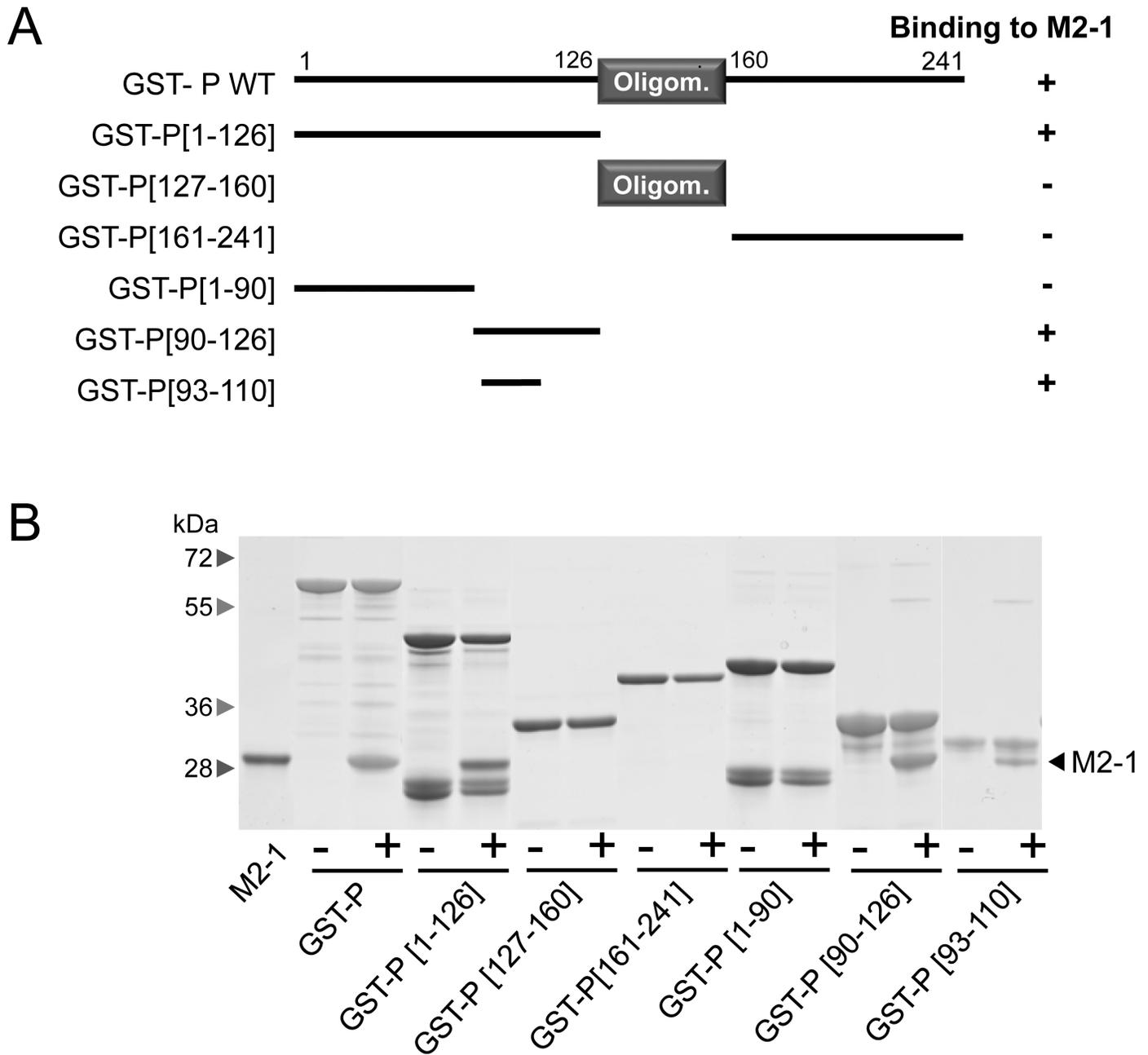
<https://doi.org/10.1371/journal.ppat.1006920.g001>

binding [33, 34]. The binding region includes an extremely transient α-helix detected in free P (α<sub>N2</sub> in Fig 1A) [44]. It is expected that in the M2-1–P complex, the α<sub>N2</sub> region adopts the molecular tumbling properties of the globular M2-1<sub>core</sub>, which results in increased transverse relaxation and hence line broadening. But this cannot explain that the signal is fully broadened out. A K<sub>d</sub> of 3 μM was previously determined for the M2-1–P complex by isothermal titration calorimetry (ITC) [19]. It is compatible with exchange on a μs–ms timescale between free and M2-1<sub>core</sub>-bound P, which contributes to line broadening. Additional broadening probably arises from conformational exchange taking place between partially folded states of the α<sub>N2</sub> helix, which may all contribute to M2-1<sub>core</sub> binding.

We also carried out NMR interaction experiments between M2-1<sub>core</sub> and two <sup>15</sup>N-labeled P fragments, P[1–126] and P[1–163], which were designed for previous characterization and interaction experiments with N by NMR [44]. They both comprise the 90–120 region. P[1–126] is a monomeric fragment that represents the N-terminal IDR of P. P[1–163] additionally contains the oligomerization domain of P and displays a similar behavior to full-length P, with nearly complete line broadening in the α<sub>N2</sub> region (Fig 1B). Surprisingly, in P[1–126] a second region is perturbed at the N-terminus of P (residues 23–37) at a lesser extent. This region overlaps with the N<sup>0</sup>-binding site of P, which contains another transient α-helix (α<sub>N1</sub>) [44, 45]. Notably, in N<sup>0</sup>-P interaction experiments using RNA-free N, we had observed the symmetrical scenario: N not only induced line broadening in the N<sup>0</sup>-binding region, but also to a lesser extent in the M2-1-binding region [44]. We cannot rule out that a direct interaction takes place in both cases at a second binding site. However, since we evidenced transient contacts between these two regions in free P [44], the perturbations observed at the second site may indirectly arise from breaking of internal contacts to expose the primary binding site [44]. Finally, from the interaction experiment with P[1–126], it appears that the primary interaction region is probably restricted to residues 90–112, since residues 113–120 are no longer completely broadened out (Fig 1C). The presence of the oligomerization domain in full-length P and P[1–163] can affect nuclear relaxation in this region by restricting motions due to steric hindrance or by promoting an extension of either α<sub>N2</sub> or the coiled-coil helices of the oligomerization domain.

### P fragment P[93–110] is sufficient for binding M2-1 *in vitro*

To validate our NMR results, we performed GST pulldown assays. Full-length and truncated forms of RSV P were produced as GST-fusion proteins (Fig 2A) and incubated with recombinant His-tagged M2-1 protein. All proteins were expressed in *E. coli*. After extensive wash, the complexes were analyzed by SDS-PAGE and Coomassie blue staining. As shown in Fig 2B, M2-1 and P constructs were purified to > 90% homogeneity, except for GST-P[1–126] and GST-P[1–90], for which faster migrating bands were observed in a reproducible manner. Mass spectrometry analysis revealed that these bands correspond to degradation of these P fragments.



**Fig 2. Identification of the M2-1 binding domain on P by GST pull-down.** (A) Schematic illustration of the full-length and truncated forms of GST-P used for M2-1 pull-downs in this study. The oligomerization domain of P is represented as a grey box, and numbers indicate amino acid positions. (B) GST-P proteins were purified on glutathione-Sepharose beads and incubated in the presence of M2-1. After extensive washing, the binding of M2-1 was determined by SDS-PAGE and Coomassie blue staining. For each deletion mutant, the ability to interact with M2-1 is summarized on the right in A.

<https://doi.org/10.1371/journal.ppat.1006920.g002>

The results clearly show that M2-1 binding was retained for GST-P, GST-P[1c126], GST-P [90-126] and GST-P[93-110]. However, no M2-1 binding was seen for GST-P[161-241], GST-P [127-160] or GST-P[1-90]. These results, which are in agreement with previous data by Mason et al. [33], demonstrate that P region P93-D110 is sufficient for binding M2-1 *in vitro*.

### Effect of targeted P gene substitutions on M2-1 controlled transcription

In order to identify the P residues critical for P–M2-1 interaction, we used a functional assay based on a firefly luciferase (Luc) reporter RSV minigenome. In this system transcription of the second gene coding for Luc is absolutely dependent on the P–M2-1 interaction [19, 23]. Site-directed mutagenesis was performed on residues V85 to E115, encompassing the P93–D110 region. Ala substitution of seven mainly hydrophobic residues, F87, F98, L101, Y102, T105, I106 and F109, had a drastic effect on RSV transcription (Fig 3A). Western blot shows that the drop in transcription efficiency was not due to a defect in expression of the P variants (Fig 3B). Among these, L101, Y102, T105 and F109 had been shown previously to be involved in the P–M2-1 interaction [33, 34]. The phosphomimetic T108D variant impaired transcription, in contrast to T108A. This is in line with the results of Asenjo et al. who had suggested that phosphorylation of T108 could negatively regulate the P–M2-1 interaction [34]. Our results are thus in agreement with previous data, but highlight the critical role of newly identified residues F87, F98 and I106.

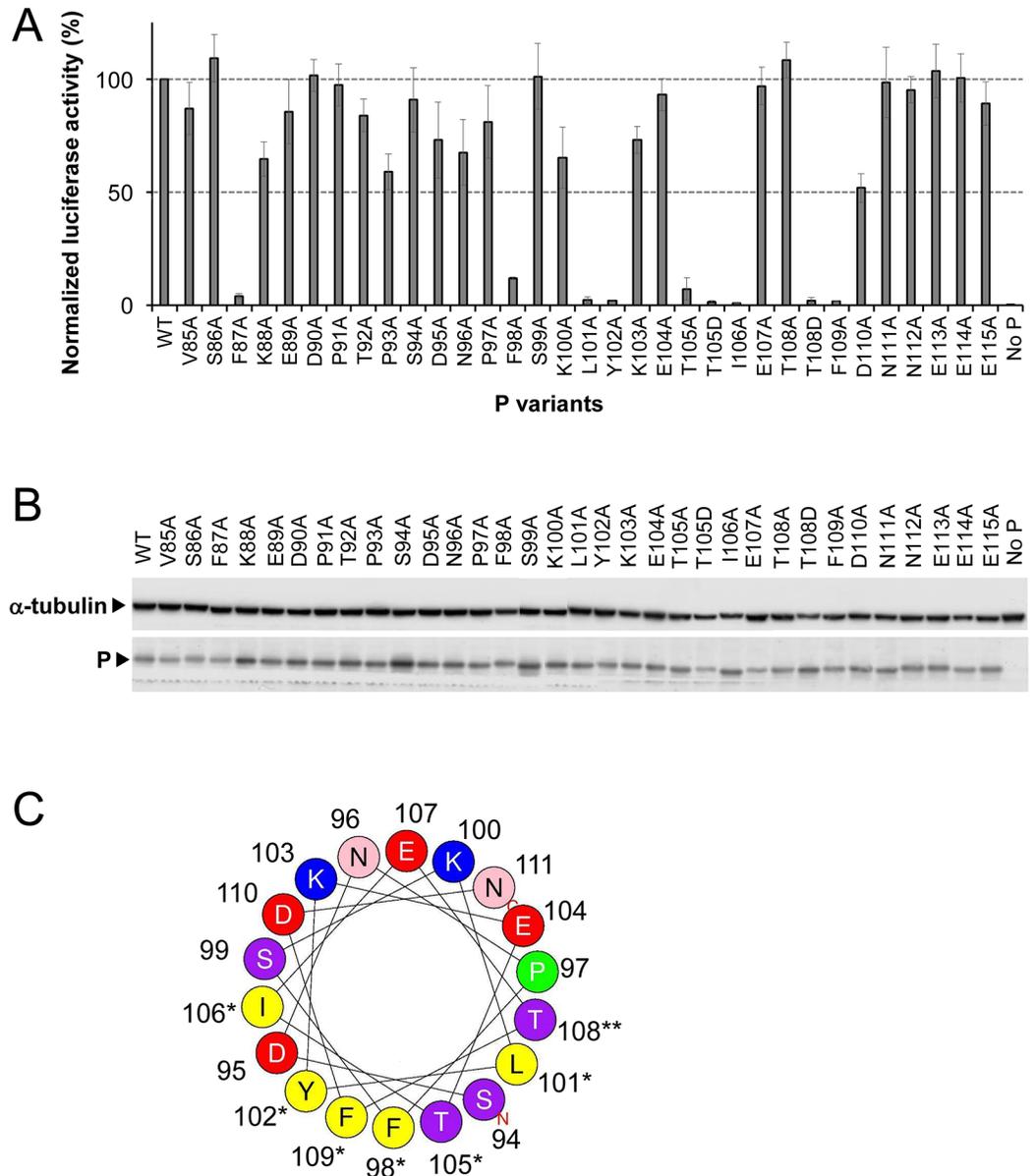
Mutations of P impairing transcription are located in the P93–D110 region, except for F87. The (*i, i+3; i, i+4*) periodicity of the critical residues F98–F109 of P identified here suggests that M2-1 binding stabilizes the transient  $\alpha$ -helix formed in free P [44]. Helical representation of this domain reveals that critical residues F98, L101, Y102, T105, I106 and F109 are located on the same side of the putative helix and form a contiguous surface (Fig 3C).

### Impact of mutations of P on the intracellular localization of M2-1

When co-expressed in the absence of other viral proteins, RSV P and N proteins induce the formation of cytoplasmic IBs similar to those observed during RSV infection, where N and P co-localize [18]. When co-expressed with P and N, M2-1 also localizes preferentially in these IBs, which has been linked to its interaction with P [19]. We thus analyzed the impact of the P mutations identified as critical for RdRp activity on the intracellular localization of M2-1 by fluorescence microscopy. Cells were cotransfected with expression vectors encoding P (WT and variants), N and M2-1-mCherry. All tested P variants were able to induce the formation of IBs (Fig 4). The M2-1-mCherry fusion protein accumulated in IBs in the presence of wild type (WT) P. M2-1 was also present in the IBs in the presence of P variant T108A, which did not impact transcription in the minigenome assay. In contrast, M2-1 was absent from the IBs for P variants that were defective for RdRp activity, F98A, L101A, Y102A, T105A, T105D, I106A, T108D and F109A, with the notable exception of F87A. It must be specified that mCherry fusion at the C-terminus of M2-1 does not affect the polymerase activity in the context of the minigenome, since ~70% of activity was recovered, as compared to WT M2-1 (S1 Fig). These results show that there is a good correlation between the presence of M2-1 in IBs and RdRp activity, except for F87A. They also point to the potential role of T108 phosphorylation for the presence of M2-1 in IBs.

### Effects of targeted P gene substitutions on the P–M2-1 interaction single out residue F87

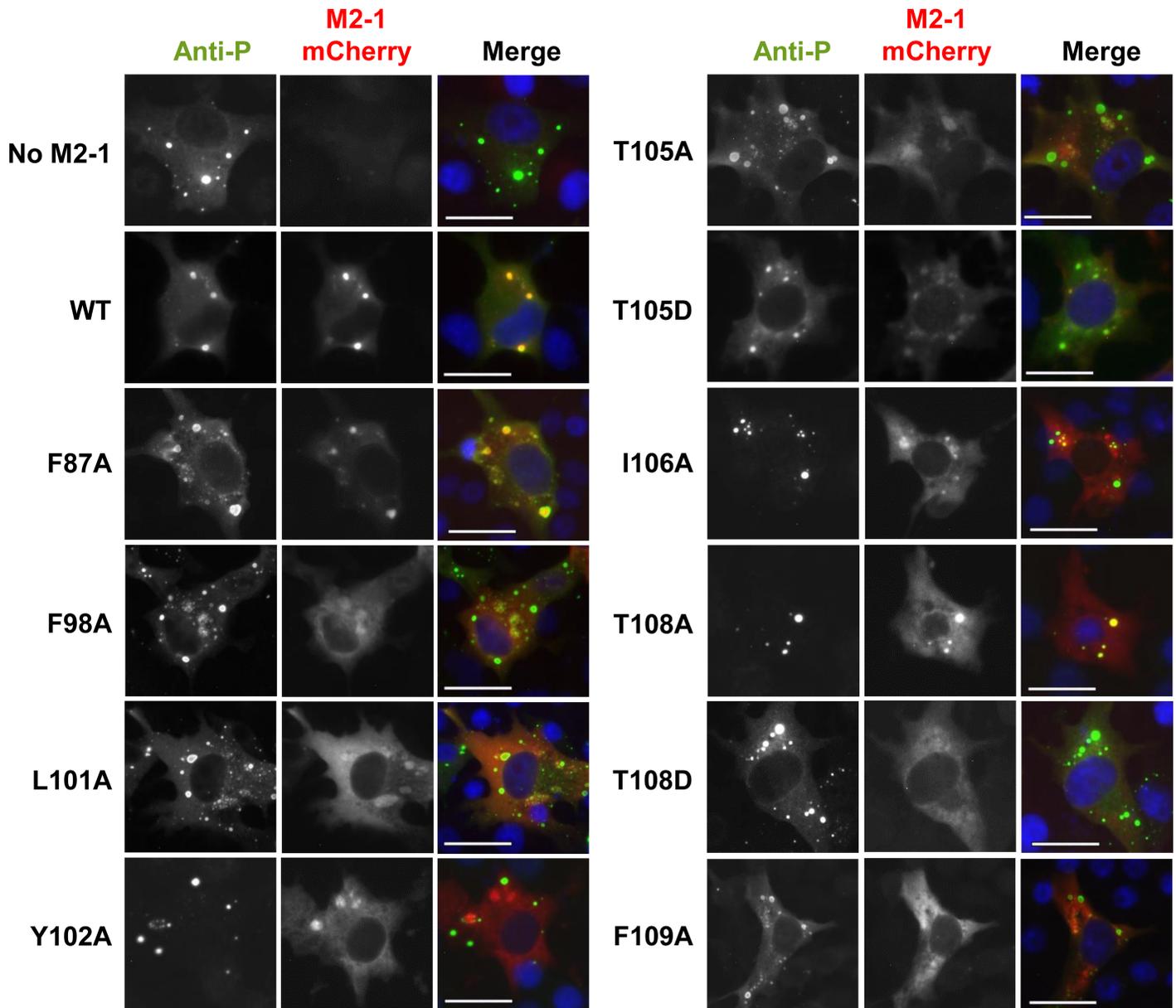
To determine whether point mutations affecting RSV transcription and M2-1 localization directly impact the P–M2-1 interaction, we performed GST pulldown assays using recombinant, non-phosphorylated GST-P and M2-1 proteins produced in *E. coli*. Fig 5A shows that substitutions F98A, L101A, Y102A, T105A, T105D, T108D and F109A were sufficient to impair P–M2-1 interaction *in vitro*, while M2-1 still interacted with the T108A and F87A variants.



**Fig 3. Identification of residues of P critical for RSV polymerase activity.** (A) Polymerase activity assay in the presence of P mutants. BSRT7/5 cells were transfected with the RSV minigenome composed of plasmids encoding the WT N, M2-1, and L proteins, the pMT/Luc minigenome, and WT or mutated P proteins, together with pCMV- $\beta$ Gal for transfection standardization. Viral RNA synthesis was quantified by measuring the Luc activity after cell lysis 24 h post-transfection. Each luciferase minigenome activity value was normalized based on  $\beta$ -galactosidase expression and is the average of results from three independent experiments performed in triplicate. Error bars represent standard deviations. (B) Western blot analysis showing efficient expression of P variant proteins in BSRT7/5 cells. (C) Helical wheel representation of the putative  $\alpha$ -helix located between residues S94 and N111 of P (HeliQuest online program). Residues critical for M2-1-binding are indicated by a star and the phosphorylatable residue T108 is indicated by two stars. Positively charged residues are in blue, negatively charged residues in red, putative phosphorylatable residues in purple and hydrophobic residues in yellow.

<https://doi.org/10.1371/journal.ppat.1006920.g003>

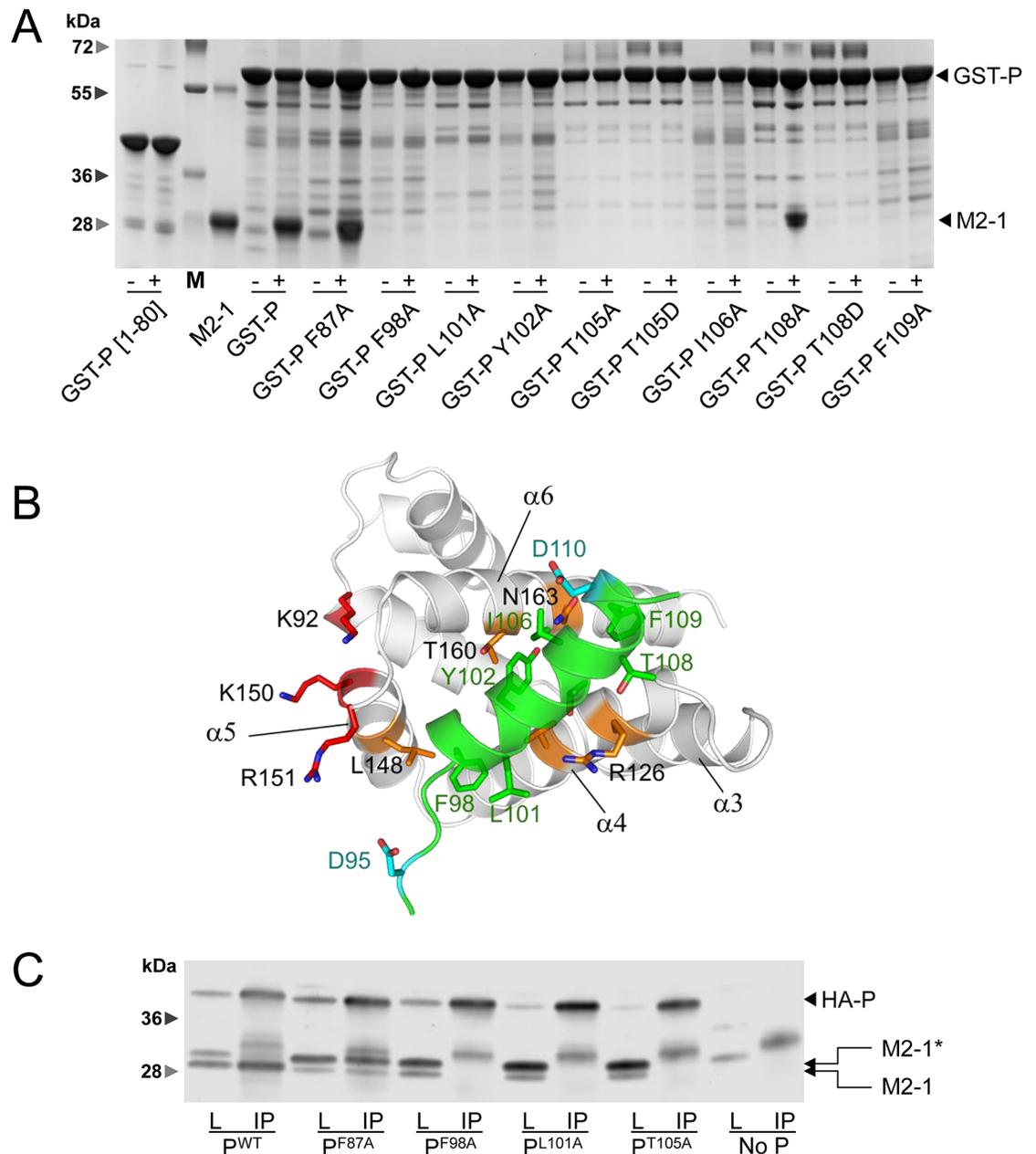
We hypothesized that the M2-1 binding region in P, which displays helical propensity in solution folds into a stable helix upon M2-1—P complex formation, and we docked a structural model of the P helix (Fig 5B, green) onto the structure of M2-1<sub>core</sub> under constraints obtained by mutagenesis and NMR interaction data. We have shown previously that the P



**Fig 4. Effects of point mutations targeting the M2-1-binding domain of P on the recruitment of M2-1 to cytoplasmic IBs.** BSRT7/5 cells were transfected with pP (WT and variants), pN, and pM2-1-mCherry plasmids. Cells were fixed 24 h after transfection, labeled with anti-P antibody (green), and colocalization of P and M2-1-mCherry (red) was analyzed by fluorescence microscopy. Nuclei were stained with DAPI. Scale bars, 20  $\mu$ m.

<https://doi.org/10.1371/journal.ppat.1006920.g004>

binding site on M2-1 forms a groove between helices  $\alpha$ 4 and  $\alpha$ 6 and that the interaction surface on M2-1 is composed of hydrophobic residues (V127, L152, V156), neutral (N129, T130, S133) and basic residues (R126, R151) [19]. M2-1 R126D and L148A variants completely impaired P binding as well as its recruitment to cytoplasmic IBs. Critical P residues determined in the present study are rather hydrophobic or neutral. However a 50% decrease in mini-genome activity was observed for P variants D95A and D110A. These two acidic residues are at the edge of the P region involved in the interaction with M2-1 and could thus also play a role in this interaction (see Fig 5B). Altogether our observations strongly suggest that the binding surfaces of P and M2-1 involve both hydrophobic and electrostatic interactions.



**Fig 5. Effects of P substitutions on P-M2-1 interaction and M2-1 phosphorylation.** (A) GST-P (WT and variants) and M2-1 with a C-terminal 6xHis-tag were expressed in *E. coli* and purified separately. Beads coated with GST-P (variants and WT), as well as GST-P[1-80] used as a negative control, were saturated with 3% BSA, and incubated alone (-) or in the presence of M2-1 (+), and washed. Complexes were resolved by SDS-PAGE and stained with Coomassie blue. M, molecular weight ladder. (B) Model of a P-M2-1 complex obtained by docking the P D95-F109 helix (green ribbon) onto a M2-1 protomer (grey ribbon, with indication of helix numbers). P residues critical for M2-1 binding are shown in sticks. M2-1 residues critical for P binding are in orange sticks, while residues involved in RNA-binding domain are in red sticks. (C) Analysis of the M2-1-P interaction in mammalian cells by immunoprecipitation. BSRT7/5 cells were transfected with plasmids encoding for N, HA-P (WT and variants) and M2-1. Immunoprecipitations (IP) from cell lysates (L) were performed 24 h post-transfection using an anti-HA antibody. P and M2-1 were revealed by Western blotting using anti-P and anti-M2-1 rabbit antisera. The star indicates the phosphorylated form of M2-1.

<https://doi.org/10.1371/journal.ppat.1006920.g005>

To clarify the differences observed between F87A and the mutations affecting the P-M2-1 interaction, we performed immunoprecipitation after co-transfecting BSRT7 cells with plasmids

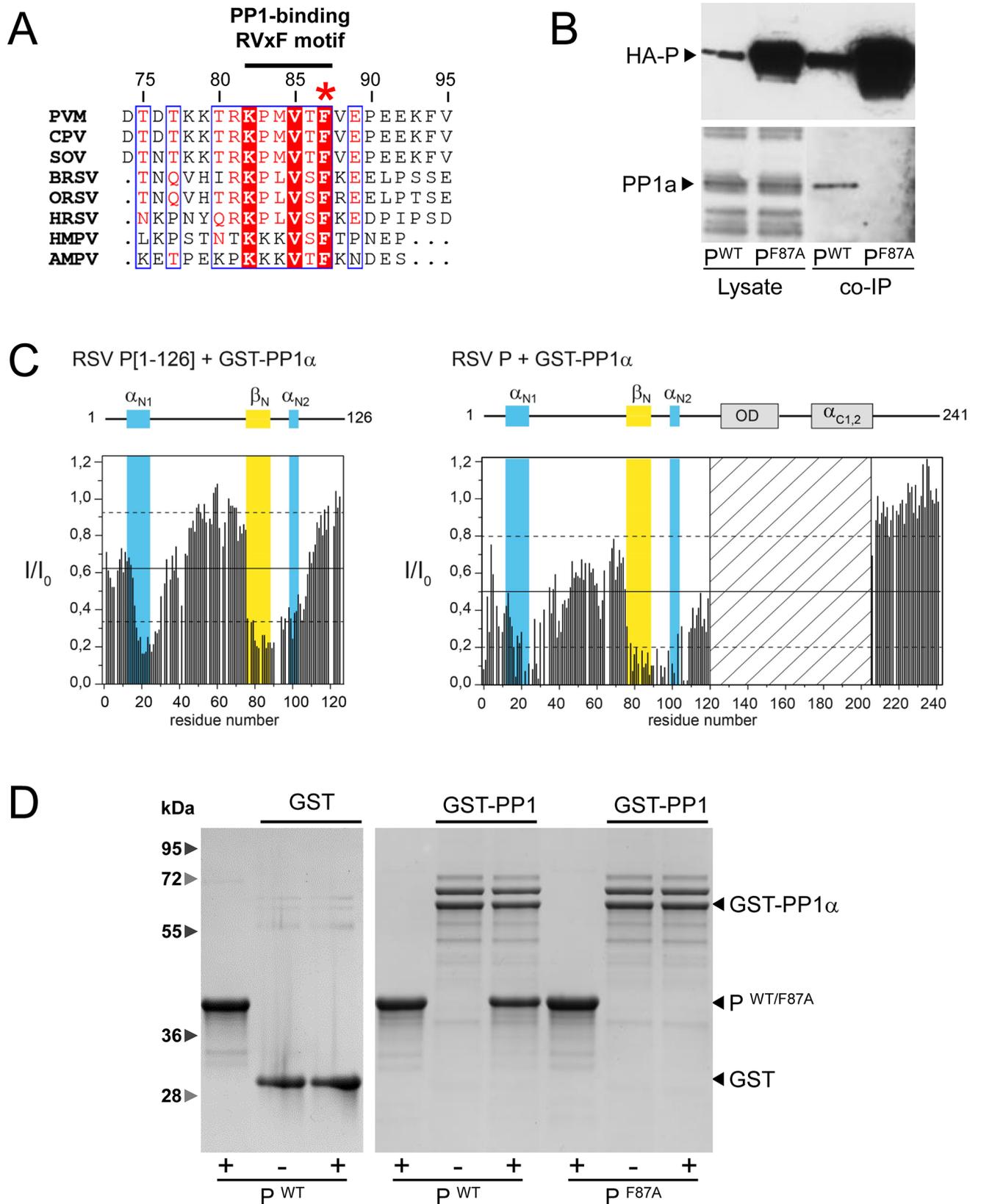
encoding N, M2-1 and HA-P (WT and variants impairing *in vitro* transcription) and by precipitating with an anti-HA monoclonal antibody. The precipitated complexes were analyzed by Western blot using anti-P and anti-M2-1 polyclonal sera. As expected [29] [28], two main bands, corresponding to phosphorylated (upper band) and unphosphorylated (lower band) M2-1, were observed in cell extracts (Fig 5C). However, whereas M2-1 in the complex with WT P was mainly unphosphorylated, M2-1 was almost exclusively phosphorylated in the presence of all the P variants tested. Even more surprisingly, whereas only the unphosphorylated form of M2-1 co-precipitated with WT P, the F87A mutant was able to precipitate phosphorylated M2-1 as well. M2-1 remained phosphorylated with F98A, L101A and T105A P variants, which do not bind any form of M2-1, suggesting that the M2-1-P interaction is involved in M2-1 dephosphorylation. Taken together, these observations indicate that both unphosphorylated and phosphorylated forms of M2-1 can interact with P, and that the interaction between P and M2-1 favors M2-1 dephosphorylation. They also reveal that the F87A mutation impairs M2-1 dephosphorylation without impeding P-M2-1 binding.

### RSV P binds the cellular protein phosphatase 1 (PP1) by a "RVxF"-like motif

Since RSV P is required for M2-1 dephosphorylation, we hypothesized that P could associate with a cellular phosphatase, that in turn would be responsible for M2-1 dephosphorylation. For Ebola virus, which belongs to the *Filoviridae* family in the *Mononegavirales* order, it was recently shown that VP30, which shares functional and structural similarities with RSV M2-1, is dephosphorylated by PP1 [46], and that VP30 dephosphorylation is critical for viral transcription. PP1 does not recognize specific sequences on its target protein. Instead, substrate binding depends on its association with PP1-interacting proteins (PIPs) that function as targeting subunits [47]. A majority of known PIPs contain a short PP1-binding motif "RVxF" (R/K-K/R-x(0,1)-V/I-x-F/W/Y). RSV P contains a <sup>81</sup>RKPLVSF<sup>87</sup> sequence, which conforms to the "RVxF" motif. Sequence alignment of *Pneumoviridae* P shows high conservation of a "R/x-K-x-x-V-T/S-F" motif, which reduces to "R-K-P-x-V-T/S-F" for pneumoviruses (Fig 6A). RSV P thus harbors a degenerate "RVxF" motif containing the residue F87 and is most likely a PIP.

To determine whether P could associate with endogenous PP1, cells were co-transfected with pHA-P and p-N, and P was immunoprecipitated from cell lysates using an anti-HA antibody. The presence of cellular phosphatases was then analyzed by Western blot. The presence of PP1 was clearly revealed in the precipitated products using WT P (Fig 6B). With the P F87A variant, PP1 was no longer precipitated (Fig 6B), emphasizing the role of this residue. Of note, for unclear reasons, the F87A P variant was overexpressed as compared to WT P.

We then investigated whether RSV P could directly interact with PP1 by using NMR interaction experiments. We observed the perturbations in <sup>1</sup>H-<sup>15</sup>N correlation spectra (BEST-TROSY) of <sup>15</sup>N-labeled P[1-126] in the presence of unlabeled GST-PP1 $\alpha$ . GST-PP1 $\alpha$  induced line broadening of NMR signals, i.e. a decrease of intensity, notably in a region spanning residues K76-D90 (Fig 6C and S2 Fig). Control experiments in the presence of GST alone did not reveal any intensity perturbation (S2 Fig). The 76-90 region contains the degenerate "RVxF" motif identified above. It is adjacent to the M2-1 binding site, but was not perturbed by M2-1<sub>core</sub> in P [1-126] (Fig 1C). The 76-90 region had already drawn our attention owing to more efficient <sup>15</sup>N transverse relaxation than in adjacent IDRs and to its  $\beta$ -strand propensity in free P [44]. Since PPI "RVxF" motifs adopt an extended  $\beta$ -strand conformation in complex with PP1, it was tempting to hypothesize that 76-90 region forms a primary PP1 binding site [48]. Similarly to M2-1<sub>core</sub>, GST-PP1 $\alpha$  also perturbed the N<sup>0</sup>-binding region in P [1-126]. In contrast to M2-1<sub>core</sub>, when we performed interaction experiments with full-length P (Fig 6C and S2 Fig),



**Fig 6. Phosphatase PP1 binds to the "RVxF" motif of RSV P.** (A) Sequence alignment of *Pneumoviridae* P showing the conservation of the "RVxF" motif. Human RSV A strain (HRSV), bovine RSV (BRSV), ovine RSV (ORSV), pneumonia virus of mice (PVM), canine pneumovirus (CPV), swine orthopneumovirus (SOV), human metapneumovirus (HMPV) and avian metapneumovirus (AMPV) (accession codes AAX23990.1, NP\_048051.1, Q83956.1, Q5MKM7.1, AHF88957.1, ANO40516.1, YP\_012606.1, and AAF05910.1, respectively) P sequences were aligned by Clustal Omega and prepared with ESPript 3. Numbers correspond to RSV P amino acid residues. (B) BSRT7/5 cells were co-transfected with pN and pHA-P or p-HA-P [F87A]. HA-P (WT and variant) was immunoprecipitated from cell lysates and the presence of PP1 in the precipitate was analyzed by Western blot. L, cell lysate; IP, immunoprecipitated products. (C) NMR analysis of GST-PP1 $\alpha$  binding to RSV P. Intensity ratios ( $I/I_0$ ), measured for each peak in  $^1\text{H}$ - $^{15}\text{N}$  BEST-TROSY spectra of 25  $\mu\text{M}$   $^{15}\text{N}$ -labeled P[1–126] or  $^{15}\text{N}$ -labeled full-length P, alone and in the presence of 2 molar equivalents of GST-PP1 $\alpha$ , are represented in the bar diagrams. Straight and broken lines indicate mean and mean  $\pm$  rmsd values. Regions with transient  $\alpha$ -helical or  $\beta$ -sheet secondary structure [44] are highlighted by a colored background with the same color code as in Fig 1. The hatched area corresponds to a P region encompassing the OD and C-terminal helices for which peaks cannot be observed in the free form. (D) Proteins GST-PP1, M2-1, P WT and F87A mutant were expressed in *E. coli* and purified separately. Beads coated with GST or GST-PP1 were saturated with 3% BSA, and incubated with, P WT or F87A mutant and washed. Complexes were pulled down, resolved by SDS-PAGE and stained with Coomassie blue.

<https://doi.org/10.1371/journal.ppat.1006920.g006>

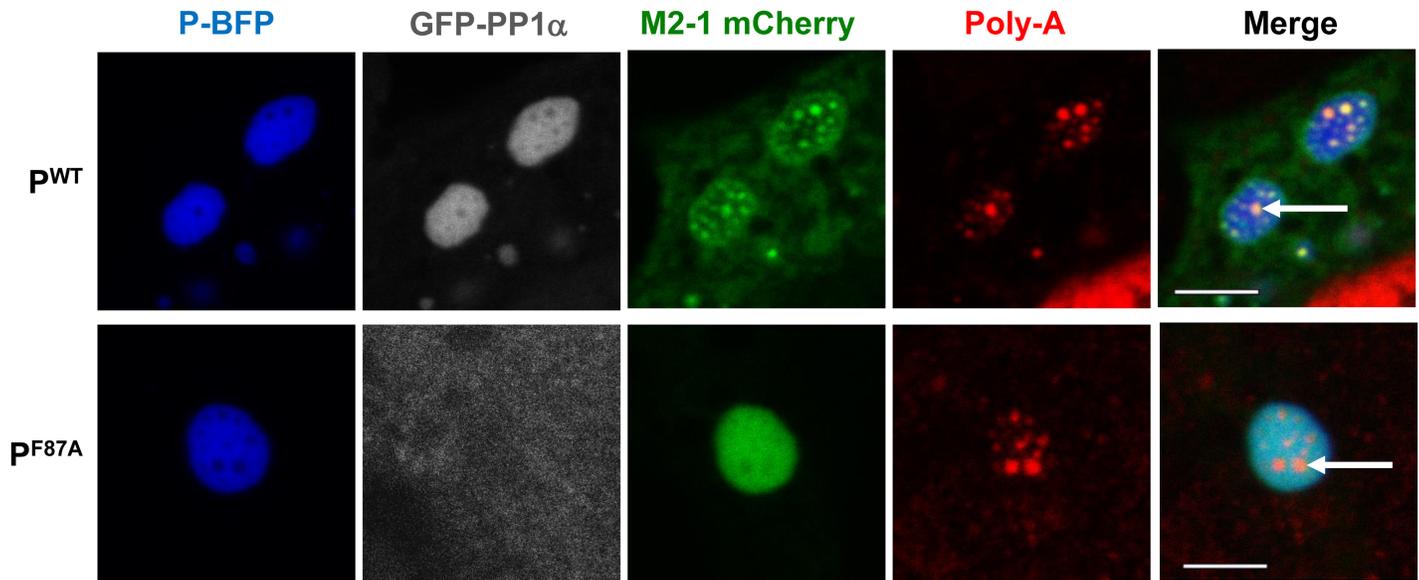
perturbations in the vicinity of  $\alpha_{\text{N1}}$  were still observed, so that we cannot rule out that they reflect direct binding to a second binding site [48]. The presence of GSH in the buffer did not affect intensities.

The ability of PP1 to interact specifically with P was thus investigated by GST pulldown using GST-PP1 $\alpha$  and recombinant P, WT or F87A mutant, all produced in *E. coli*. As shown in Fig 6D, only WT P was efficiently pulled down by GST-PP1. In summary, these results demonstrate that the P–PP1 interaction is direct and that residue F87 of P plays a pivotal role in this interaction.

### RSV P recruits the cellular protein phosphatase 1 (PP1) to IBs, which is required for M2-1 accumulation in IBAGs

The presence of PP1 in IBs was analyzed by confocal microscopy. BSRT7/5 cells were co-transfected with plasmids expressing PP1-GFP, L, P-BFP, N, M2-1–mCherry proteins together with the pM/Luc vector expressing a firefly luciferase-reporter RSV minigenome, then fixed 24 hours post-transfection. Of note, BFP insertion between residues 73–74 of P, in a naturally disordered and poorly conserved region among *Pneumoviridae* [44], only moderately affected the polymerase activity in the context of the minigenome, since  $\sim 80\%$  of activity was recovered, as compared to WT P (S1 Fig). As shown in Fig 7, PP1 fluorescence clearly overlapped with the WT P fluorescence, indicating that PP1 was significantly targeted to the IBs. In contrast, PP1 was no longer detected in IBs when expressed in the presence of the F87A P variant. These results revealed that RSV P recruits PP1 to IBs and that F87 plays a critical role in this process.

By studying the ultrastructure of IBs, we recently found that M2-1 colocalizes with viral neo-synthesized mRNA in IBAGs, revealed by a poly(dT) probe, and from which genomic viral RNA, N, P and L proteins are excluded [20]. We thus wondered if the absence of a P-PP1 complex could affect the localization of M2-1 in IBs. As the formation of IBAGs requires viral mRNA synthesis, we first verified that viral transcription can occur in the absence of M2-1 activity in our system. For that purpose, we engineered a new minigenome in which the first gene was replaced by the *Gussia* luciferase coding sequence, upstream from the firefly luciferase gene. As shown in S3 Fig, whereas no Firefly luciferase activity was detected in the absence of M2-1, the *Gussia* luciferase activity was still detectable although reduced to  $\sim 15\%$ . This confirmed that M2-1 is not absolutely required for expression of the first gene but increases expression of this gene. Based on this result, transfected cells were prepared for FISH using poly(dT) probes and analyzed by confocal microscopy. Fig 7 shows that, in contrast to what was observed with WT P, in the presence of the F87A P mutant M2-1 was present throughout the IBs, where it colocalized with P, but absent from IBAGs as revealed by a poly(dT) probe. These results indicate that phosphorylated M2-1, which is still competent for P-binding, cannot associate with neo-synthesized viral poly-adenylated mRNAs.



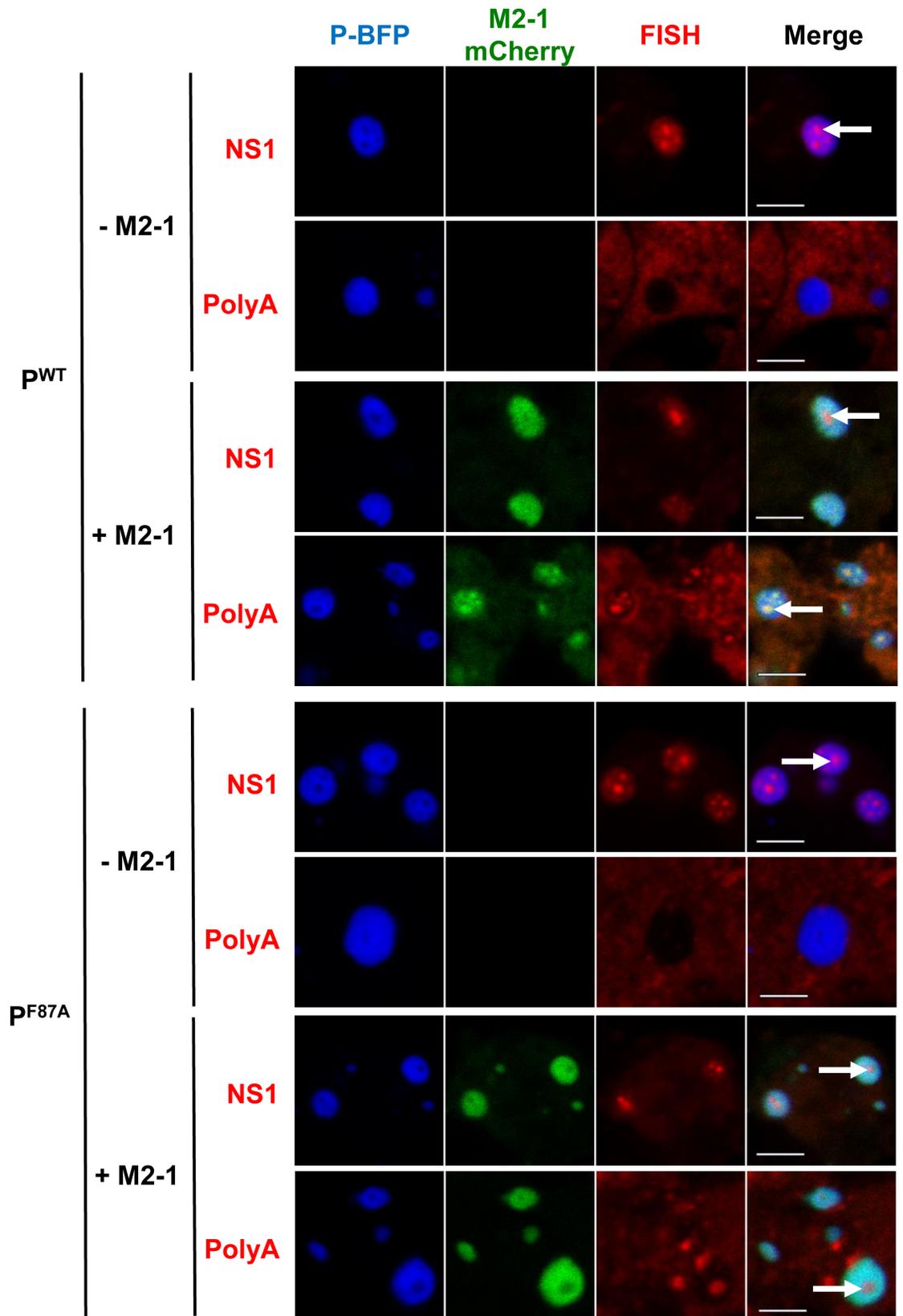
**Fig 7. P-PP1 interaction allows recruitment of PP1 to IBs and M2-1 localization in IBAGs.** BSRT7/5 cells were transfected with plasmids encoding the N, L and M2-1-mCherry proteins, the M/Luc RSV minigenome, and either wild type (WT) or F87A mutant P-BFP. Tagged proteins were expressed instead of the corresponding wild type as indicated on the pictures. FISH analyses were performed to detect poly(A) RNA (in red). The expressed tagged proteins are visualized thanks to their spontaneous fluorescence. IBs are delimited by the P fluorescence. Arrows point to IBAGs. Representative images from 3 independent experiments are shown. Images were taken under a Leica SP8 confocal microscope. Scale bars 5µm.

<https://doi.org/10.1371/journal.ppat.1006920.g007>

In a previous report we observed that IBAGs are not detected when using an oligo(dT) probe in the absence of M2-1 [20]. These results suggested that either M2-1 is needed for the formation of IBAGs, or that M2-1 binds to mRNAs and is carried to the IBAGs as a passenger, not as a required chaperone. To clarify this point, we used two different probes, an oligo(dT) and probes targeting the first transcription unit of the M/Luc minigenome (the NS1-M chimeric mRNA), and compared the presence of IBAGs in the absence or presence of either P WT or P F87A. Some pictures representing the trend of what we observed are shown in Fig 8; in the absence of M2-1, some IBAGs were detected when using NS1/M probes but not with an oligo(dT) probe. Similar results were observed with either WT or F87A P. Thus these results indicate that (i) transcription of the first gene can occur either in the absence of M2-1 or in presence of phosphorylated M2-1, as revealed by NS1 probes; (ii) poly-adenylation is affected in the absence of M2-1 but not in the presence of phosphorylated M2-1; and (iii) IBAGs revealed by either NS1 or oligo(dT) probes can form in the absence of M2-1 or in the presence of phosphorylated M2-1 which is excluded from them. In conclusion these results suggest that M2-1 is not essential for the formation of RNA aggregates called IBAGs, and could be involved in mRNA poly-adenylation. They also suggest that a defect in M2-1 dephosphorylation does not affect mRNA poly-adenylation but binding of M2-1 to poly(A) RNAs after their synthesis.

### PP1 overexpression dephosphorylates M2-1 *in cellula*

A previous report suggested a requirement for cyclic phosphorylation/dephosphorylation of M2-1 for efficient antitermination function [24]. To further demonstrate that PP1 is involved in M2-1 dephosphorylation *in cellula*, PP1 was overexpressed in BSRT7/5 cells in the context of the RSV minigenome. Fig 9A shows that RSV RdRp activity, as revealed by Luc activity, significantly decreased in a dose-dependent manner with pEGFP-PP1 plasmid addition, while P and N expression levels were not affected (Fig 9B). However, although the expression level of



**Fig 8. Effect of M2-1 on RSV mRNA poly-adenylation either in the presence or absence of PP1 in IBs.** BSRT7/5 cells were transfected with plasmids encoding the N, L and M2-1-mCherry proteins, the M/Luc RSV minigenome, and either WT or F87A mutant P-BFP. FISH analyses were performed to detect poly(A) or NS1 RNAs (in red). The expressed tagged proteins are visualized thanks to their spontaneous fluorescence. IBs are delimited by the P fluorescence. Arrows point to IBAGs. Representative images from 3 independent experiments are shown. Images were taken under a Leica SP8 confocal microscope. Scale bars 5µm.

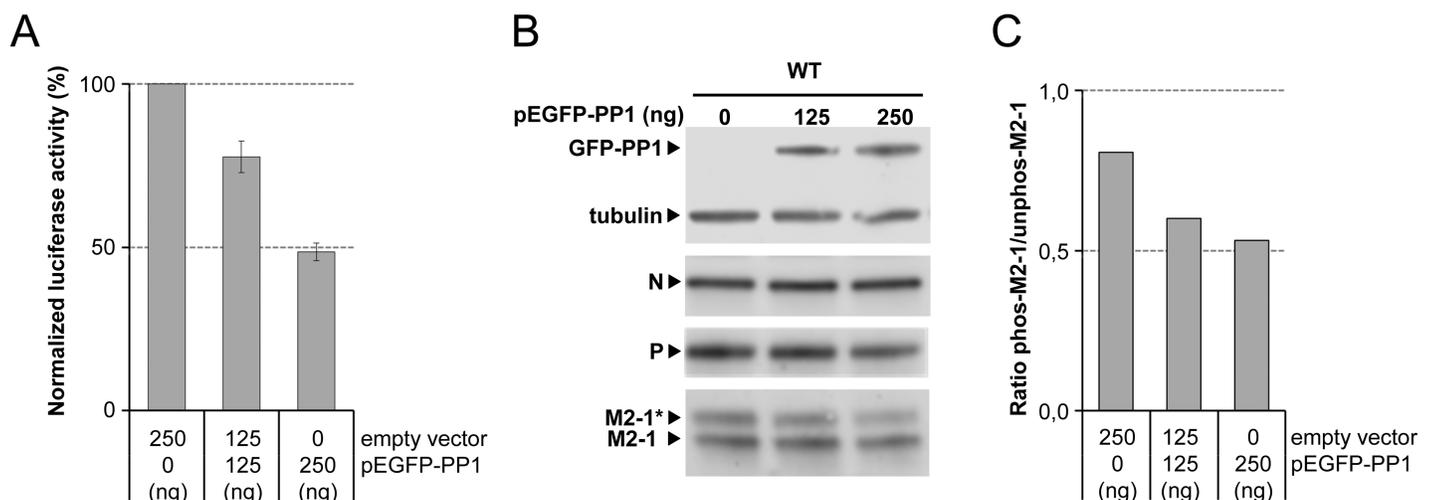
<https://doi.org/10.1371/journal.ppat.1006920.g008>

unphosphorylated M2-1 was not or poorly affected (Fig 9B), overexpression of PP1 significantly induced a decrease in phosphorylated M2-1, and thus a decrease of the ratio of phosphorylated versus unphosphorylated M2-1 (Fig 9B and 9C), confirming that PP1 is involved in M2-1 dephosphorylation in living cells.

## Discussion

### RSV P recruits M2-1 to IBs

The efficient activity of the RSV RdRp complex depends on regulated and highly specific protein-protein interactions, which are potential targets for antiviral therapy [49]. P plays a pivotal role through multiple interactions with L, N and M2-1, mediated by its high structural plasticity linked to its disordered regions [44, 50]. The M2-1 binding region has been mapped previously to P residues 100–120, using proteins from two RSV strains (A2 and Long), and the role of specific residues L101, Y102, F109, T105 and T108 was highlighted [33], [34]. By combining NMR, biochemical and functional approaches based on an RSV minigenome and microscopy, we precisely mapped the M2-1 binding region of P to a stretch encompassing residues 93–110. We identified 7 residues of P that are directly involved in this interaction, summarized in Fig 10A. These include two newly identified residues F98 and I106 and validate the major role of hydrophobic P residues in the P–M2-1 interaction [33].



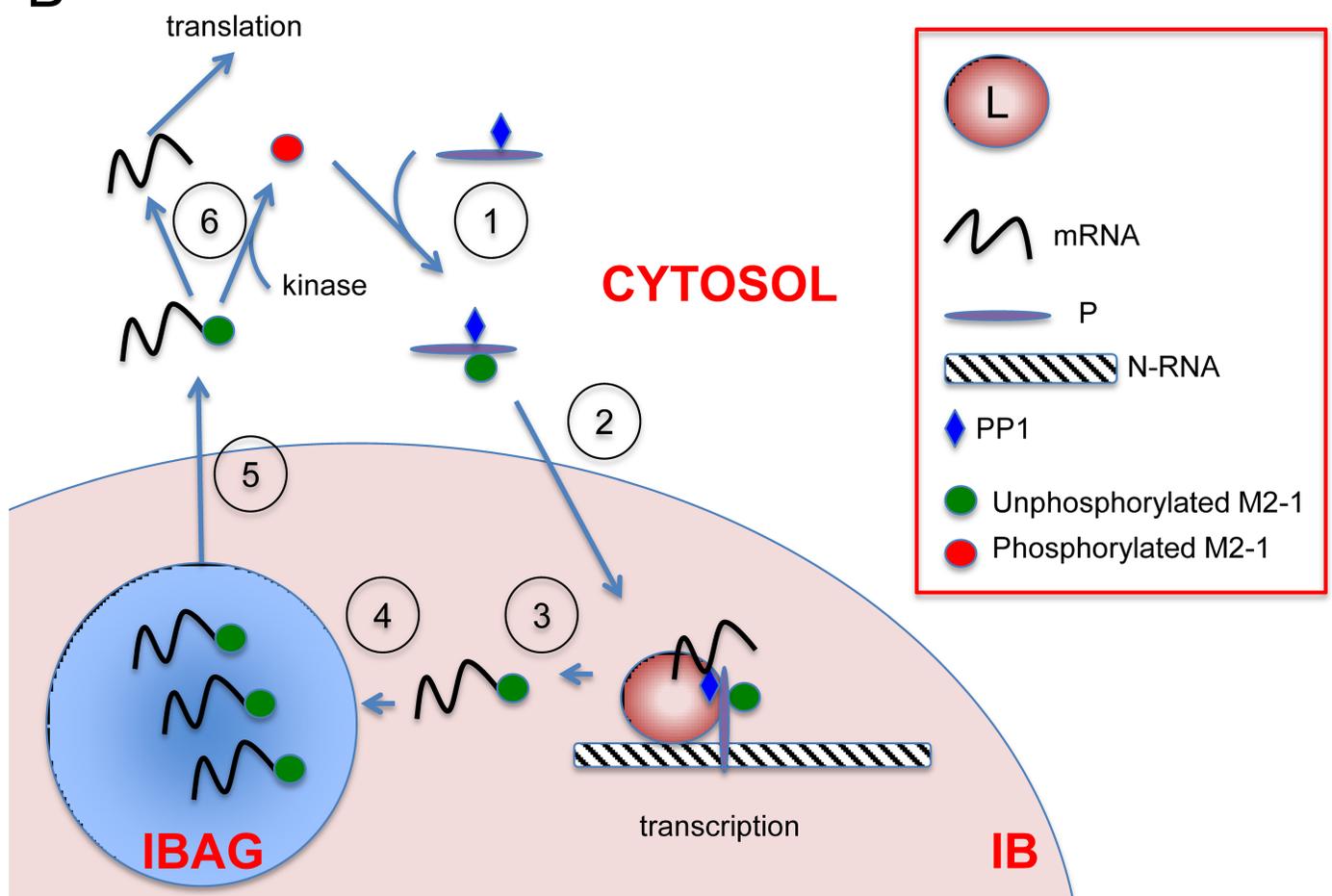
**Fig 9. Effect of PP1 overexpression on RSV RNA polymerase activity and M2-1 phosphorylation.** (A) Inhibition of RSV RNA polymerase activity by PP1 overexpression. BSRT7/5 cells were transfected with the RSV minigenome, and various amounts of pEGFP-PP1 vector. Viral RNA synthesis was quantified by measuring the Luc activity after cell lysis 24 h post-transfection. Each Luc activity value was normalized based on β-galactosidase expression and is the average of results from three independent experiments performed in triplicate. Error bars represent standard deviations. (B) Expression of phosphorylated (\*) or unphosphorylated M2-1, N and P in BSR-T7/5 cells transfected with the RSV minigenome and with increasing quantities of pGFP-PP1 vector. Cell extracts were resolved by SDS-PAGE 24 h post-transfection and analyzed by Western-blot using rabbit polyclonal anti-M2-1 or anti-P, and mouse monoclonal anti-α-tubulin antibodies. (C) Ratio of phosphorylated (Phos-M2-1) versus unphosphorylated (Unphos-M2-1) M2-1, normalized to levels of α-tubulin, in the presence of increasing amounts of pGFP-PP1 vector. Signals were quantified using a Chemidoc Touch Imaging System (Bio-Rad, France).

<https://doi.org/10.1371/journal.ppat.1006920.g009>

A

P mutations	PP1-binding region	M2-1 binding region	Polymerase activity (%)	GST Pull down	IP of M2-1 by P variants	Recruitment of M2-1 into IBs by IF
	80 85 90	95 100 105 110 115				
	QRKPLVS	FKEDPTPSDNPFSKLYKETIETFDNNEEE				
F87A	-----A-----		4	+	+	+
F98A		-----A-----	12	-	-	-
L101A		-----A-----	3	-	-	-
Y102A		-----A-----	2	-	-	-
T105A		-----A-----	8	-	-	-
T105D		-----D-----	2	-	ND	-
I106A		-----A-----	1	-	-	-
T108A		-----A-----	108	+	ND	+
T108D		-----D-----	2	-	ND	-
F109A		-----A-----	1	-	-	-

B



**Fig 10. Model for phosphorylation turnover of M2-1 taking into account the P–M2-1 and P–PP1 interactions.** (A) The primary sequence of the 80–115 region of P is indicated at the top. Amino acid residues previously identified as critical for M2-1 binding [33, 34] are in green, newly identified residues F98 and I106 are in red. Residue F87 critical for M2-1 dephosphorylation and efficient transcription is highlighted in red font on a yellow background. The left-hand column shows amino acid substitutions. Right-hand columns summarize the impact of P mutations on (i) the polymerase activity, (ii) *in vitro* interaction between P and M2-1 using recombinant proteins, (iii) interaction between P and M2-1 assessed by co-immunoprecipitation from transfected BSRT7/5 cells, and (iv) on the recruitment of M2-1 to IBs; ND, not determined. (B) Diagram of M2-1 phosphorylation turnover. (1) M2-1 is phosphorylated (at Ser 58 and 61), binds to P in the cytosol and is dephosphorylated by PP1 in the P–PP1 complex. (2) The P–PP1–M2-1 complex is directed to IBs (pink) where transcription takes place. (3) M2-1 binds to neo-synthesized mRNAs at the end of transcription and (4) M2-1–mRNAs complexes are concentrated in IBAGs (blue) before being released in the cytosol (5). M2-1 is then phosphorylated by a cellular kinase and parts with viral mRNAs which will be used for translation of viral proteins (6).

<https://doi.org/10.1371/journal.ppat.1006920.g010>

Furthermore, we found a perfect correlation between the capacity of P to interact with M2-1, RdRp activity and the presence of M2-1 in IBs. These results indicate that M2-1 cannot reach IBs without P, P acting as a recruiter for M2-1, and are in agreement with previous observations obtained with M2-1 variants unable to interact with P [19].

### Dephosphorylation of M2-1 is mediated by a P-PP1 complex

M2-1 is mainly unphosphorylated in the context of an RSV infection, whereas the phosphorylated state is predominant when M2-1 is expressed alone in cells, S58 and S61 being the main sites of phosphorylation [26]. Previous experiments showed that phosphoablatant substitutions S58A/S61A as well as phosphomimetic substitutions S58D/S61D reduced RSV transcription activity to less than 20% and to ~40% as compared to WT M2-1, respectively [24, 26]. These data suggest that a cyclic turnover of phosphorylation-dephosphorylation of M2-1 is required for efficient RSV transcription.

Here we observed that unphosphorylated M2-1 was predominant in cell lysates when co-expressed with WT P. A reverse situation was observed when M2-1 was co-expressed with P variants F98A, L101A, T105A that do not interact with M2-1, M2-1 being mainly phosphorylated, suggesting that P binding induces M2-1 dephosphorylation (Fig 5C). The finding that M2-1 is predominantly phosphorylated in the presence of another P variant F87A, which efficiently pulled down M2-1, confirmed that P is capable of interacting with both phosphorylated and unphosphorylated forms of M2-1 [23]. Altogether these results indicated that (i) P mediates M2-1 dephosphorylation or prevents M2-1 phosphorylation and that (ii) a P region adjacent to the M2-1 binding region comes into play for this activity.

Previous studies suggested that RSV P could be a target of cellular phosphatases PP1 and PP2A in cultured cells [38, 51]. PP1 is a well-characterized and conserved Ser/Thr phosphatase holoenzyme. It is composed of a variable regulatory subunit that determines the localization, activity, and substrate specificity of the phosphatase and of one of three highly homologous catalytic phosphatase subunits PP1 $\alpha$ , PP1 $\gamma$ , or PP1 $\beta/\delta$  (reviewed in [47, 48]). PP1 is the most widely expressed and abundant Ser/Thr phosphatase and is estimated to catalyze about one third of all protein dephosphorylations in eukaryotic cells. It dephosphorylates hundreds of key biological targets by associating with nearly 200 regulatory proteins to form highly specific holoenzymes. Of note, many of the > 200 established PIPs are predicted to be intrinsically disordered like RSV P.

The defect in M2-1 dephosphorylation observed for the P F87A variant mainly argued in favor of the hypothesis of the recruitment by P of the PP1 cellular phosphatase through a RVxF-like motif, upstream of the M2-1 binding site. This hypothesis was consolidated by several complementary approaches showing that (i) WT P, but not the F87A variant, could bind PP1 *in vitro* and *in cellula*; (ii) PP1 colocalized with WT P, but not with the F87A variant, in IBs; and (iii) overexpression of PP1 increased the unphosphorylated/phosphorylated M2-1 ratio. Altogether, our data show that RSV P can be considered as a PP1-interacting protein (PIP), targeting PP1 to the M2-1 substrate.

## Mode of action of PP1 for M2-1 dephosphorylation

The degenerate "RVxF" motif in RSV P, <sup>81</sup>RKPLVSF<sup>87</sup>, is well conserved among *Pneumoviridae*, with a consensus KxxVxF (Fig 6A). The lysine residue is surrounded by basic residues, with an arginine just upstream for orthopneumoviruses, and two lysines downstream for metapneumoviruses. Thus, it is likely that the P proteins of *Pneumoviridae* share the property of interacting with PP1 to regulate viral protein phosphorylation. Notably, M2-1 protein is unique to *Pneumoviridae* and present in all members of this virus family. It was shown that residues S57 and S60 of hMPV M2-1 protein, which are equivalent to RSV M2-1 S58 and S61, are also critical for virus replication, consistent with the critical role of cyclic phosphorylation/dephosphorylation of hMPV M2-1 for efficient RNA synthesis [52]. It is thus expected that hMPV and RSV present the same mechanism of regulation of M2-1 dephosphorylation.

It is noteworthy that PP1 was previously shown to play a significant role in the replication of several viruses, including papovavirus, adenovirus, human immunodeficiency virus 1 (HIV-1) and 2, Ebola virus (EBOV), and Rift Valley Fever virus [46, 53, 54]. Within *Mononegavirales*, the VP30 protein from filoviruses is the only protein that shares structural and functional similarity with the M2-1 protein from *Pneumoviridae*. It was shown that the phosphorylation status of VP30 was also regulated by PP1 [46]. It has also been demonstrated that PP1 regulates the innate immune responses for numerous RNA viruses, such as influenza virus, Sendai virus, dengue virus, and picornavirus [55]. The V protein of measles virus was shown to interact directly with PP1 $\alpha/\gamma$ , via a canonical PP1-binding motif, <sup>288</sup>RIWY<sup>291</sup>, preventing PP1-mediated dephosphorylation of MDA5, a cytosolic sensor crucial for innate immune defense against various RNA viruses, thereby impairing its activation [56].

## Potential role of cyclic M2-1 phosphorylation/dephosphorylation

Previous reports indicated that the phosphorylation state of M2-1 regulates the affinities with its partners [25]. The highest affinities of RSV M2-1 (produced as recombinant protein in *E. coli*, i.e. unphosphorylated) for RNA were found with poly(A) (Kd~20 nM) and sequences present on viral mRNAs that are complementary to GE signals (Kd ~ 46 nM) [19, 24]. We previously determined that the phosphomimetic S58D/S61D substitution decreased the RNA binding affinity as compared to WT M2-1 [24]. This is consistent with the crystal structure of full-length tetrameric RSV M2-1 [24], where S58 and S61 are located on a flexible loop, facing the RNA binding region located in M2-1<sub>core</sub>. Addition of a negative charge on these residues by phosphorylation could thus affect the interaction with RNA. Recombinant tetrameric P and M2-1 proteins (produced in *E. coli*, i.e. unphosphorylated) form a complex with a 1:1 stoichiometry and a Kd of ~ 8.1 nM [32]. The affinities of P and RNA for M2-1 are therefore comparable, which would be in line with a switch of M2-1 between the two M2-1–RNA and M2-1–P complexes.

This assumption was confirmed by comparing the organization of IBs in the presence of WT or F87A P variant in our minigenome assay. We previously observed a distinct compartment in IBs that contains concentrated viral mRNAs and M2-1 (IBAG). The rest of the IBs holds all the other proteins of the RSV polymerase complex together with the genomic RNA [20]. When the P F87A variant was used, PP1 was absent from IBs, and M2-1, which was found mainly in a phosphorylated form (see Fig 5C), was no longer associated with IBAGs. The M/Luc minigenome we used for fluorescence microscopy studies has two transcription units (see [materials and Methods](#)); the first one can be transcribed in the absence of M2-1, although with a 5 to 7 fold reduction (see S3 Fig), while the second one coding for Firefly luciferase fully depends on a functional M2-1 [17, 23, 57]. It is thus likely that the poly(A) RNA seen in IBs when P F87A was used represents transcription from the first gene. These results

also indicate that IBAGs form in the absence of dephosphorylated, competent for mRNA binding, M2-1. Finally, by using two different RNA probes, i.e. NS1 and poly(dT), we observed different effects on viral RNA poly-adenylation when M2-1 was either absent or co-expressed with P F87A and thus in the absence of PP1; in the first case some residual transcription was detected with the NS1 probe but not with the poly(dT) probe; in the second case (phosphorylated M2-1 only), both probes could detect some IBAGs. Together these results suggest that (i) poly-adenylation of the first transcription unit is impaired in the absence of M2-1 and (ii) phosphorylated M2-1 can help poly-adenylation, although it is excluded from IBAGs after transcription.

To summarize, RSV P can be considered as a newly identified PIP, targeting PP1 to the M2-1 substrate. To our knowledge, this is the first report showing such a mechanism of regulation of *Pneumoviridae* transcription. Recently, we have shown that, in RSV-infected cells, viral RNA synthesis occur in IBs, M2-1 and viral mRNAs concentrating in IB sub-compartments called IBAGs [20]. The data indicated that IBAGs are dynamic structures, allowing the sorting of viral mRNAs and their transport to the cytoplasm together with M2-1. Accordingly, we propose the following model (Fig 10B): in the cytoplasm of infected cells the P protein binds PP1 and the phosphorylated form of M2-1 allowing their recruitment into IBs. M2-1 works as an anti-terminator of transcription in IBs and could intervene as a poly-adenylation factor. Once dephosphorylated by PP1, M2-1 has a higher affinity for RNA, in particular for GE and poly(A) sequences at the end of the transcripts. Terminated mRNAs concentrate in IBAGs and drag M2-1 along. M2-1 could act in a manner similar to the cellular poly(A) binding protein, protecting mRNA from degradation and perhaps playing a role in the transport of mRNAs from IBAGs to the cytosol in order to translate viral mRNAs, and possibly playing an active role in translation. M2-1 phosphorylation could then occur in the cytosol, resulting in the detachment of M2-1 from the poly(A) tail of mRNA, before being recycled by P and redirected to IBs for a new round of transcription.

## Materials and methods

### Plasmid constructs

All viral sequences were derived from the human RSV strain Long. The full length or segments of the P gene were PCR amplified by using *Pfu* DNA polymerase (Stratagene, Les Ulis, France) and cloned into pGEX-4T-3 bacterial expression vector (GE Healthcare) at BamHI-SmaI sites to engineer the pGEX-P and derived plasmids. The M2-1 cDNA [23] was subcloned into pET-22b+ (Novagen) to allow bacterial expression of full-length M2-1 with a C-terminal poly-histidine tag. Plasmids for eukaryotic expression of the human RSV (Long strain) proteins N, P, M2-1, and L, designated pN, pP, pM2-1, and pL, respectively, and pM/Luc have been described previously [22, 23]. The pM/Luc minigenome has two transcription units; the first ORF is a chimera between NS1 (327 first nucleotides) and M genes (138 last nucleotides), and transcription termination depends on a N Gene End sequence. The second ORF codes for firefly luciferase and ends with a SH Gene End sequence. The first and second mRNA expressed by this minigenome are 698 and 1897 nucleotides in length, excluding the poly(A) tail, respectively. A second minigenome containing *Gaussia* and *Firefly* luciferases was engineered and synthesized by Genscript. This p*Gaussia*/*Firefly* minigenome is similar to the pM/Luc minigenome excepted that it contains the *Gaussia* luciferase gene upstream in place of the NS1-M chimeric gene (complete sequences available on demand). Point mutations were introduced in the P sequence by site-directed mutagenesis using the QuikChange site-directed mutagenesis kit (Stratagene). To generate the plasmid pHA-P, complementary oligonucleotides encoding a hemagglutinin (HA) tag epitope (MYPYDVPDYA) were annealed and inserted at the BamHI

restriction site in frame at the 5' end of the P gene in pP. For the eukaryotic expression of a M2-1-mCherry fusion protein, the mCherry gene was amplified by PCR from the pmCherry vector (Clontech) and cloned in frame at the 3' end of the M2-1 gene at BglII-XhoI sites in pM2-1. To engineer the pP-BFP vector, first a NheI site was introduced at nucleotides 741–746 of the P sequence; then, the BFP gene was PCR amplified from the pTagBFP-Tubulin vector (Evrogen) and inserted at the NheI restriction site, between P residues 73–74. The eukaryotic expression vector pEGFP(C1)-PP1 $\alpha$  was purchased from Addgene. The PP1A insert (residues 7–300) was subcloned into pGEX-4T-3 at BamHI-XhoI sites. Sequence analysis was carried out to check the integrity of all the constructs. All the oligonucleotide sequences are available on request.

### Bacterial expression of recombinant proteins and purification

For M2-1 expression, *E. coli* BL21(DE3) (Novagen) bacteria were transformed with pET-M2-1 plasmid, and bacteria were grown at 37°C for 8 h in Luria-Bertani medium (LB) containing 100  $\mu$ g/ml ampicillin. Protein expression was induced by adding one volume of fresh LB medium, 400  $\mu$ M isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) and 50  $\mu$ M ZnSO<sub>4</sub> for 16 h at 28°C. Cultures were centrifuged at 5,000 g for 15 min and the pellet was resuspended in lysis buffer (20 mM Tris-HCl pH 7.4, 150 mM NaCl, 50 mM imidazole, 0.1% Triton X-100 and 1 mg/ml lysozyme). Benzonase (Novagen) was then added to the lysates (final concentration 5U/mL), which were further incubated for 1 h at room temperature under rotation. NaCl was then added to reach a final concentration of 1M and lysates were clarified by centrifugation at 10,000 g for 1 h at 4°C. The supernatant was loaded onto a 5-mL HiTrap IMAC column (GE Healthcare) charged with 0.2 M ZnSO<sub>4</sub> and equilibrated with low-imidazole buffer and high salted buffer (50 mM imidazole, 20 mM Tris-HCl [pH 7.4], 1M NaCl) using a 50-mL Superloop. Then a linear gradient of 1–0.15M NaCl was applied to reduce the concentration in NaCl. Finally, a linear gradient of 80–800 mM imidazole in the same buffer was applied to elute M2-1 fractions containing the His-tagged proteins. After equilibration with 20 mM Tris-HCl [pH 7.4], 150 mM NaCl buffer, M2-1 was further purified by size exclusion chromatography on a HiLoad Superdex-200 column with a 120-mL total bed volume (GE Healthcare). Appropriate fractions were pooled. M2-1 protein was confirmed RNA-free by spectrophotometry (OD 260/280 ratio) and stored at 4°C.

For P expression, *E. coli* BL21(DE3) bacteria transformed with pGEX-P derived plasmids were grown as described above. Bacterial pellets were resuspended in lysis buffer (20 mM Tris/HCl pH 7.4, 60 mM NaCl, 1 mM EDTA, 1 mg/mL lysozyme, 1 mM DTT, 0.1% Triton X-100) supplemented with complete protease inhibitor cocktail (Roche) for 1 h on ice. Benzonase was then added and the lysate was incubated for 1 h at ambient temperature under rotation. The lysates were centrifuged at 4°C for 30 min at 10,000 g. Glutathione-Sepharose 4B beads (GE Healthcare) were added to the clarified supernatants and the mixtures were incubated overnight at 4°C under rotation. The beads were washed with lysis buffer, three times with 1X PBS and then stored at 4°C in an equal volume of PBS.

For PP1 expression, chaperone competent cells pGro7/BL21 (Takara) were transformed with pGEX-PP1 and grown as previously described [58]. Briefly, an overnight starter culture was grown at 37°C in LB medium supplemented with antibiotics and 1 mM MnCl<sub>2</sub>. PP1 production was initiated by inoculating 1/2 liter of LB supplemented with 1 mM MnCl<sub>2</sub> with 35 ml of starter culture. The bacteria were grown at 30°C to an OD of ~0.5, then arabinose was added (2 g/l) to induce the expression of the GroEL/GroES chaperone. When OD was ~1, the temperature was lowered to 10°C (ice-bath) and the expression of PP1 was induced with 0.1 mM IPTG for ~20 h. Culture was centrifuged at 5,000 g for 15 min and bacterial pellets were

resuspended in lysis buffer (20 mM Tris/HCl pH 8.0, 150 mM NaCl, 1 mM MnCl<sub>2</sub>, 1 mg/mL lysozyme, 1 mM DTT, 0.1% Triton X-100) supplemented with complete protease inhibitor cocktail (Roche) for 1 h on ice and treated as described above except that NaCl was then added to reach a final concentration of 700mM.

For NMR experiments <sup>15</sup>N-labeled P and two N-terminal fragments of P, P[1–126] (residues 1–126) and P[1–163] (residues 1–163), were produced in M9 medium supplemented with <sup>15</sup>N-labeled NH<sub>4</sub>Cl (Eurisotop) and glucose and purified following the same protocol as for full-length P. The core domain of M2-1 (residues 58–177) was prepared as described previously [19]. Proteins were cleaved from GST with biotinylated thrombin (Novagen). Thrombin was later removed using streptavidin resin (Novagen). The samples were subsequently concentrated on centrifugation filter units (Amicon Ultra) to 50–150 μM and dialyzed against NMR buffer (20 mM sodium phosphate pH 6.8, 100 mM NaCl). GST-PP1α was eluted from GSH-sepharose beads by using 50 mM glutathione, concentrated and dialyzed against NMR buffer. Purity was assessed by SDS-PAGE and by mass spectrometry.

### GST pulldown assays

GST pull-downs were performed by incubating 50 μl of a 50% slurry of Glutathione-Sepharose 4B beads (GE Healthcare) containing either GST-PP1 or GST-P (WT and mutants) at 25 μM in 20mM Tris/HCl [pH 7.4], 150mM NaCl (TN buffer) with a 3-fold molar excess of M2-1 or P and BSA 3%. After 1 h at 20°C under agitation, the beads were washed extensively with TN buffer, boiled in 25 μl of Laemmli buffer and analyzed by SDS-PAGE and Coomassie blue staining.

### Cell culture

BSRT7/5 [59] cells were maintained in Eagle's minimum essential medium and Dulbecco's modified Eagle's medium, respectively, supplemented with 10% fetal calf serum, 2 mM L-glutamine, and penicillin–streptomycin solution. The cells were grown at 37°C in 5% CO<sub>2</sub>. Cytotoxicity measurements were performed using the CellTiter-Glo Luminescent cell viability assay (Promega).

### Minigenome experiments

BSRT7/5 cells at 90% confluence in 24-well dishes were transfected using Lipofectamine 2000 (Invitrogen) with a plasmid mixture containing 0.25 μg of pM/Luc minigenome, 0.25 μg of pN, 0.25 μg of pP (WT and mutants), 0.125 μg of pL, and 0.06 μg of pM2-1, as well as 0.06 μg of pSV-β-Gal (Promega) to normalize transfection efficiencies as previously described [23]. Transfections were done in triplicate and each independent transfection was performed three times. Cells were harvested at 24 h post-transfection and lysed in luciferase lysis buffer (30 mM Tris [pH 7.9], 10 mM MgCl<sub>2</sub>, 1 mM dithiothreitol [DTT], 1% [vol/vol] Triton X-100, and 15% [vol/vol] glycerol). Luciferase activities were measured for each cell lysate after injection of lysis buffer supplemented with ATP and D-luciferin (final concentrations 1mM each) with an Infinite 200 Pro (Tecan, Männedorf, Switzerland) and normalized to β-Gal expression levels.

### Fluorescence microscopy

BSRT7/5 cells grown on coverslips were transfected with pN, pM2-1-mCherry and pP or p-BFP (WT or variants) using Lipofectamine2000 (Invitrogen). At 24 h post-transfection, samples were fixed in 4% paraformaldehyde (PFA) for 30 min, and permeabilized in PBS containing 0.1% Triton X-100 and 3% BSA. Coverslips were incubated for 1h at room temperature

with primary antibodies, washed, and then incubated for an additional hour with Alexa Fluor 488 goat anti-mouse IgG (Invitrogen). For P detection the mouse anti-P monoclonal antibody 021/2P [18] was used. Coverslips were mounted with ProLong Gold Antifade reagent containing DAPI (Life Technologies). Cells were observed with a Nikon TE200 inverted microscope equipped with a Photometrics CoolSNAP ES2 camera. Images were processed using MetaVue software (Molecular Devices).

Confocal microscopy was used to study IB ultrastructure. Z-stack image acquisitions of multi-labeled (P-BFP, N, M2-1-Cherry, FISH) cells were performed under a 63x apochromatic lens and a numerical zoom comprised between 1x and 15x (LSAF acquisition software) under the WLL Leica SP8 microscope and representative pictures were taken.

### Fluorescent in situ hybridization (FISH)

Infected cells were fixed 24 h p.i. and permeabilized as described above. Endogenous biotin was blocked in PBS-BSA 1% (w/v) supplemented with free streptavidin (4  $\mu$ g/ml) for 1h. Coverslips were rinsed 3 times with PBS, post-fixed 10 min at 4°C in formaldehyde 4% (v/v), rinsed 2 times with PBS, and incubated in hybridization mix (2x SSC (1x SSC is 150 mM NaCl and 15 mM sodium citrate), dextran 10% (w/v), formamide 20% or 50% (v/v) for oligo(dT) and NS1 probes, respectively, 1 mg/ml herring sperm DNA; mRNAs were detected by using either 3'-biotinylated poly(dT) or a pool of NS1 oligonucleotides (for sequences see ref. [20] at a final concentration of 100  $\mu$ M in a humidified chamber at 37°C for 3h. Next, cells were washed 2 times at 42°C with the following 3 solutions: 2x SSC plus formamide 20% (v/v), 2x SSC, 1x SSC and finally PBS at room temperature. Probes were then detected by incubating cells with streptavidin-Alexa Fluor 647 conjugate (8  $\mu$ g/ml) in PBS-BSA 1% (w/v) during 1h prior to 3 washes with PBS. Cells were then submitted to immunofluorescent staining and confocal microscopy.

### Immunoblotting

Cells were lysed for 30 min at 4°C in lysis buffer (20 mM Tris [pH 7.4], 150 mM NaCl, 0.1% Triton X-100) supplemented with a complete protease inhibitor cocktail (Roche). Cell lysates were spun for 10 min at 10,000 g; supernatants were recovered, mixed with Laemmli buffer, and boiled. Proteins were resolved by SDS-PAGE and transferred onto nitrocellulose membranes. The membranes were incubated in blocking solution (1X PBS, 0.05% Tween 20 supplemented with 5% milk) for 1 h. Blots were incubated with primary antibodies in blocking solution: rabbit anti-P and anti-M2-1 antisera [31], mouse monoclonal anti- $\alpha$ -tubulin antibody (Sigma), rabbit polyclonal anti-PP1A antibody (Abcam) and rat monoclonal anti-HA-peroxidase antibody (Roche). The membranes were rinsed with PBS containing 0.05% Tween 20 and incubated for 1 h with the appropriate HRP-conjugated secondary antibodies diluted in blocking solution. The membranes were rinsed, and immunodetection was performed by using an enhanced chemiluminescence (ECL) substrate (BioRad, France).

### Coimmunoprecipitation assays

BSRT7/5 cells were cotransfected with PHA-P (WT and variants), pN and pM2-1. After 24 h, transfected cells were lysed for 30 min at 4°C in ice-cold lysis buffer (20 mM Tris HCl [pH 7.4], 150 mM NaCl, 0.1% Triton X100, 20 $\mu$ M RNase A and 15% glycerol) with a complete protease inhibitor cocktail (Roche). Cell lysates were centrifuged at 4°C for 10 min at 10,000 g and incubated overnight at 4°C with a rat anti-HA monoclonal antibody (Roche cl. 3F10) coupled to magnetic beads (Invitrogen). The beads were then washed 3 times with lysis buffer and 1

time with PBS, proteins were boiled in Laemmli buffer for 5 min and samples were subjected to SDS-PAGE and immunoblotting as described above.

### NMR spectroscopy

$^1\text{H}$ - $^{15}\text{N}$  correlation spectra, Heteronuclear Single Quantum Correlation (HSQC) or BEST-TROSY, were measured at a temperature of 288 K on Bruker Avance III 800 or 950 MHz spectrometers equipped with cryogenic TCI probes. 7%  $\text{D}_2\text{O}$  was added to the samples to lock the magnetic field. Spectra were processed with Topspin 3.2 (Bruker Biospin) and analyzed with CCPNMR 2.2 software [60].

### Modelling

Flexible docking was carried out on the guru interface of the Haddock Webserver [61, 62], using the X-ray structure of M2-1<sub>core</sub> extracted from PDB 4C3E, chain B. Models of a P helix spanning residues 94–112 were built using CYANA 3.2 [63] under torsion angle constraints obtained from backbone chemical shifts of P. Following active M2-1 (126, 127, 129, 130, 133, 148, 152, 156, 160, 163) and P residues (98, 101, 102, 105, 106, 108 and 109) were identified by mutagenesis experiments or by NMR interaction experiments, in this work or in [19]. Passive residues were automatically defined around active residues. 1000 initial structures were generated. 200 final structures were refined in water and clustered according to RMSD criterion. More than 50% structures clustered in the same cluster with the best Haddock score. Statistics are shown in [S1 Table](#).

### Supporting information

**S1 Fig. Effects of M2-1-mCherry and P-BFP for RSV polymerase activity.** BSRT7/5 were transfected with pMT/Luc, pP or p-P-BFP, pL, pN and either pM2-1 or pM2-1-cherry, and Luc reporter activity was measured.

(DOCX)

**S2 Fig. Perturbations of GST-PP1 $\alpha$  in NMR spectra of P.** (Upper panel) Superimposed  $^1\text{H}$ - $^{15}\text{N}$  BEST-TROSY spectra of 25  $\mu\text{M}$   $^{15}\text{N}$ -labeled P[1–126] or  $^{15}\text{N}$ -labeled P, alone (magenta contours) and in the presence of 2 molar equivalents of GST-PP1 $\alpha$  (black contours) are shown with residue-specific assignments. (Lower panel) Superimposed  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of 10  $\mu\text{M}$   $^{15}\text{N}$ -labeled P[1–126], alone (green contours) and in the presence of 10 molar equivalents of GST (black contours). The buffer of the P[1–126]+GST sample contained 25 mM glutathione that yields natural abundance  $^{15}\text{N}$  signals indicated by stars.

(DOCX)

**S3 Fig. Comparative normalized activities of *Gaussia* and Firefly luciferase in the absence (-) or presence (+) of M2-1.** BSRT7/5 were transfected with p*Gaussia*/Firefly minigenome vector, pP pL, pN and either pM2-1 or an empty vector pGEM3 and luciferase activities were measured 24 hours post-transfection.

(DOCX)

**S1 Table. Statistics for P-M2-1 docking and clustering with HADDOCK2.2 Webserver (S1 References 1,2).** The top cluster is the most reliable according to HADDOCK.

(DOCX)

**S1 References. References for S1 Table.**

(DOCX)

## Acknowledgments

We are grateful to José A. Melero (Madrid, Spain) for providing 021/2P anti-P monoclonal antibody, Damien Vitour (ANSES Maisons-Alfort) for providing rabbit N antiserum, Didier Chevret and Sophie Liuu (UMR1319 Micalis, Plateforme d'Analyse Protéomique de Paris Sud-Ouest, INRA Jouy-en-Josas, France) for mass spectrometry analysis, Nadine Assrir for discussion about PP1 production, Marin Bouchez for technical help, and Monika Bajorek for critical reading of the manuscript.

## Author Contributions

**Conceptualization:** Sergei Nekhai, Christina Sizun, Jean-François Eléouët.

**Formal analysis:** Christina Sizun, Jean-François Eléouët.

**Funding acquisition:** Christina Sizun, Jean-François Eléouët.

**Investigation:** Charles-Adrien Richard, Vincent Rincheval, Safa Lassoued, Jenna Fix, Christophe Cardone, Camille Esneau, Christina Sizun.

**Methodology:** Sergei Nekhai, Christina Sizun, Jean-François Eléouët.

**Project administration:** Jean-François Eléouët.

**Supervision:** Marie Galloux, Marie-Anne Rameix-Welti, Christina Sizun, Jean-François Eléouët.

**Writing – original draft:** Charles-Adrien Richard, Christina Sizun, Jean-François Eléouët.

**Writing – review & editing:** Sergei Nekhai, Marie Galloux, Marie-Anne Rameix-Welti, Christina Sizun, Jean-François Eléouët.

## References

1. Nair H, Nokes DJ, Gessner BD, Dherani M, Madhi SA, Singleton RJ, et al. Global burden of acute lower respiratory infections due to respiratory syncytial virus in young children: a systematic review and meta-analysis. *Lancet*. 2010; 375(9725):1545–55. Epub 2010/04/20. [https://doi.org/10.1016/S0140-6736\(10\)60206-1](https://doi.org/10.1016/S0140-6736(10)60206-1) PMID: 20399493; PubMed Central PMCID: PMC2864404.
2. Collins PL, Melero JA. Progress in understanding and controlling respiratory syncytial virus: Still crazy after all these years. *Virus Res*. 2011; 162(1–2):80–99. Epub 2011/10/04. <https://doi.org/10.1016/j.virusres.2011.09.020> PMID: 21963675.
3. Afonso CL, Amarasinghe GK, Banyai K, Bao Y, Basler CF, Bavari S, et al. Taxonomy of the order Mononegavirales: update 2016. *Arch Virol*. 2016; 161(8):2351–60. Epub 2016/05/25. <https://doi.org/10.1007/s00705-016-2880-1> PMID: 27216929; PubMed Central PMCID: PMC4947412.
4. Bakker SE, Duquerroy S, Galloux M, Loney C, Conner E, Eleouet JF, et al. The respiratory syncytial virus nucleoprotein-RNA complex forms a left-handed helical nucleocapsid. *J Gen Virol*. 2013; 94(Pt 8):1734–8. Epub 2013/05/17. <https://doi.org/10.1099/vir.0.053025-0> PMID: 23677789; PubMed Central PMCID: PMC3749527.
5. Liljeroos L, Krzyzaniak MA, Helenius A, Butcher SJ. Architecture of respiratory syncytial virus revealed by electron cryotomography. *Proc Natl Acad Sci U S A*. 2013; 110(27):11133–8. Epub 2013/06/19. <https://doi.org/10.1073/pnas.1309070110> PMID: 23776214; PubMed Central PMCID: PMC3703984.
6. Tawar RG, Duquerroy S, Vonrhein C, Varela PF, Damier-Piolle L, Castagne N, et al. Crystal structure of a nucleocapsid-like nucleoprotein-RNA complex of respiratory syncytial virus. *Science*. 2009; 326(5957):1279–83. Epub 2009/12/08. doi: 326/5957/1279 [pii] <https://doi.org/10.1126/science.1177634> PMID: 19965480.
7. Grosfeld H, Hill MG, Collins PL. 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. *J Virol*. 1995; 69(9):5677–86. Epub 1995/09/01. PMID: 7637014; PubMed Central PMCID: PMC189426.

8. Yu Q, Hardy RW, Wertz GW. Functional cDNA clones of the human respiratory syncytial (RS) virus N, P, and L proteins support replication of RS virus genomic RNA analogs and define minimal trans-acting requirements for RNA replication. *J Virol.* 1995; 69(4):2412–9. Epub 1995/04/01. PMID: [7884888](#).
9. Morin B, Kranzusch PJ, Rahmeh AA, Whelan SP. The polymerase of negative-stranded RNA viruses. *Curr Opin Virol.* 2013; 3(2):103–10. Epub 2013/04/23. <https://doi.org/10.1016/j.coviro.2013.03.008> PMID: [23602472](#); PubMed Central PMCID: PMC4159711.
10. Yabukarski F, Leyrat C, Martinez N, Communie G, Ivanov I, Ribeiro EA Jr., et al. Ensemble Structure of the Highly Flexible Complex Formed between Vesicular Stomatitis Virus Unassembled Nucleoprotein and its Phosphoprotein Chaperone. *J Mol Biol.* 2016; 428(13):2671–94. Epub 2016/04/25. <https://doi.org/10.1016/j.jmb.2016.04.010> PMID: [27107640](#).
11. Collins PL, Hill MG, Camargo E, Grosfeld H, Chanock RM, Murphy BR. 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. *Proc Natl Acad Sci U S A.* 1995; 92(25):11563–7. Epub 1995/12/05. PMID: [8524804](#).
12. Mink MA, Stec DS, Collins PL. Nucleotide sequences of the 3' leader and 5' trailer regions of human respiratory syncytial virus genomic RNA. *Virology.* 1991; 185(2):615–24. Epub 1991/12/01. PMID: [1840712](#).
13. Kuo L, Grosfeld H, Cristina J, Hill MG, Collins PL. Effects of mutations in the gene-start and gene-end sequence motifs on transcription of monocistronic and dicistronic minigenomes of respiratory syncytial virus. *J Virol.* 1996; 70(10):6892–901. Epub 1996/10/01. PMID: [8794332](#); PubMed Central PMCID: PMC190738.
14. Barik S. Transcription of human respiratory syncytial virus genome RNA in vitro: requirement of cellular factor(s). *J Virol.* 1992; 66(11):6813–8. Epub 1992/11/01. PMID: [1404620](#); PubMed Central PMCID: PMC240184.
15. Collins PL, Crowe JE. Respiratory Syncytial Virus and Metapneumovirus. In: Knipe DM, Howley PM, editors. *Fields Virology*. 5th ed. Philadelphia: Lippincott Williams & Wilkins; 2007. p. 1601–46.
16. Collins PL, Hill MG, Cristina J, Grosfeld H. Transcription elongation factor of respiratory syncytial virus, a nonsegmented negative-strand RNA virus. *Proc Natl Acad Sci U S A.* 1996; 93(1):81–5. Epub 1996/01/09. PMID: [8552680](#); PubMed Central PMCID: PMC40182.
17. Fearn R, Collins PL. Role of the M2-1 transcription antitermination protein of respiratory syncytial virus in sequential transcription. *J Virol.* 1999; 73(7):5852–64. Epub 1999/06/11. PMID: [10364337](#).
18. Garcia J, Garcia-Barreno B, Vivo A, Melero JA. 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.* 1993; 195(1):243–7. Epub 1993/07/01. <https://doi.org/10.1006/viro.1993.1366> PMID: [8317099](#).
19. Blondot ML, Dubosclard V, Fix J, Lassoued S, Aumont-Nicaise M, Bontems F, et al. Structure and Functional Analysis of the RNA and Viral Phosphoprotein Binding Domain of Respiratory Syncytial Virus M2-1 Protein. *PLoS Pathogens.* 2012; 8(5):e1002734. <https://doi.org/10.1371/journal.ppat.1002734> PMID: [22675274](#)
20. Rincheval V, Lelek M, Gault E, Bouillier C, Sitterlin D, Blouquit-Laye S, et al. Functional organization of cytoplasmic inclusion bodies in cells infected by respiratory syncytial virus. *Nat Commun.* 2017; 8(1):563. Epub 2017/09/17. <https://doi.org/10.1038/s41467-017-00655-9> PMID: [28916773](#).
21. Carromeu C, Simabuco FM, Tamura RE, Farinha Arcieri LE, Ventura AM. Intracellular localization of human respiratory syncytial virus L protein. *Arch Virol.* 2007; 152(12):2259–63. Epub 2007/08/19. <https://doi.org/10.1007/s00705-007-1048-4> PMID: [17703289](#).
22. Fix J, Galloux M, Blondot ML, Eleouet JF. The insertion of fluorescent proteins in a variable region of respiratory syncytial virus L polymerase results in fluorescent and functional enzymes but with reduced activities. *Open Virol J.* 2011; 5:103–8. Epub 2011/10/04. <https://doi.org/10.2174/1874357901105010103> PMID: [21966341](#); PubMed Central PMCID: PMC3178903.
23. Tran TL, Castagne N, Dubosclard V, Noinville S, Koch E, Moudjou M, et al. The respiratory syncytial virus M2-1 protein forms tetramers and interacts with RNA and P in a competitive manner. *J Virol.* 2009; 83(13):6363–74. Epub 2009/04/24. doi: JVI.00335-09 [pii] <https://doi.org/10.1128/JVI.00335-09> PMID: [19386701](#).
24. Tanner SJ, Ariza A, Richard CA, Kyle HF, Dods RL, Blondot ML, et al. Crystal structure of the essential transcription antiterminator M2-1 protein of human respiratory syncytial virus and implications of its phosphorylation. *Proc Natl Acad Sci U S A.* 2014; 111(4):1580–5. Epub 2014/01/18. <https://doi.org/10.1073/pnas.1317262111> PMID: [24434552](#).

25. Cuesta I, Geng X, Asenjo A, Villanueva N. Structural phosphoprotein M2-1 of the human respiratory syncytial virus is an RNA binding protein. *J Virol.* 2000; 74(21):9858–67. Epub 2000/10/12. PMID: [11024112](https://pubmed.ncbi.nlm.nih.gov/11024112/); PubMed Central PMCID: PMC102022.
26. Cartee TL, Wertz GW. Respiratory syncytial virus M2-1 protein requires phosphorylation for efficient function and binds viral RNA during infection. *J Virol.* 2001; 75(24):12188–97. Epub 2001/11/17. <https://doi.org/10.1128/JVI.75.24.12188-12197.2001> PMID: [11711610](https://pubmed.ncbi.nlm.nih.gov/11711610/); PubMed Central PMCID: PMC116116.
27. Hardy RW, Wertz GW. The Cys(3)-His(1) motif of the respiratory syncytial virus M2-1 protein is essential for protein function. *J Virol.* 2000; 74(13):5880–5. PMID: [10846068](https://pubmed.ncbi.nlm.nih.gov/10846068/).
28. Lambert DM, Hambor J, Diebold M, Galinski B. Kinetics of synthesis and phosphorylation of respiratory syncytial virus polypeptides. *J Gen Virol.* 1988; 69 Pt 2):313–23. Epub 1988/02/01. <https://doi.org/10.1099/0022-1317-69-2-313> PMID: [3339328](https://pubmed.ncbi.nlm.nih.gov/3339328/).
29. Routledge EG, Willcocks MM, Morgan L, Samson AC, Scott R, Toms GL. Heterogeneity of the respiratory syncytial virus 22K protein revealed by Western blotting with monoclonal antibodies. *J Gen Virol.* 1987; 68 Pt 4):1209–15. Epub 1987/04/01. <https://doi.org/10.1099/0022-1317-68-4-1209> PMID: [3572360](https://pubmed.ncbi.nlm.nih.gov/3572360/).
30. Mazumder B, Adhikary G, Barik S. Bacterial expression of human respiratory syncytial viral phosphoprotein P and identification of Ser237 as the site of phosphorylation by cellular casein kinase II. *Virology.* 1994; 205(1):93–103. Epub 1994/11/15. <https://doi.org/10.1006/viro.1994.1623> PMID: [7975241](https://pubmed.ncbi.nlm.nih.gov/7975241/).
31. Castagne N, Barbier A, Bernard J, Rezaei H, Huet JC, Henry C, et al. Biochemical characterization of the respiratory syncytial virus P-P and P-N protein complexes and localization of the P protein oligomerization domain. *J Gen Virol.* 2004; 85(Pt 6):1643–53. <https://doi.org/10.1099/vir.0.79830-0> PMID: [15166449](https://pubmed.ncbi.nlm.nih.gov/15166449/).
32. Esperante SA, Paris G, de Prat-Gay G. 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.* 2012; 51(41):8100–10. Epub 2012/09/18. <https://doi.org/10.1021/bi300765c> PMID: [22978633](https://pubmed.ncbi.nlm.nih.gov/22978633/).
33. Mason SW, Aberg E, Lawetz C, DeLong R, Whitehead P, Liuzzi M. Interaction between human respiratory syncytial virus (RSV) M2-1 and P proteins is required for reconstitution of M2-1-dependent RSV minigenome activity. *J Virol.* 2003; 77(19):10670–6. <https://doi.org/10.1128/JVI.77.19.10670-10676.2003> PMID: [12970453](https://pubmed.ncbi.nlm.nih.gov/12970453/).
34. Asenjo A, Calvo E, Villanueva N. Phosphorylation of human respiratory syncytial virus P protein at threonine 108 controls its interaction with the M2-1 protein in the viral RNA polymerase complex. *J Gen Virol.* 2006;87(Pt 12):3637–42. PMID: [17098979](https://pubmed.ncbi.nlm.nih.gov/17098979/).
35. Navarro J, Lopez-Otin C, Villanueva N. Location of phosphorylated residues in human respiratory syncytial virus phosphoprotein. *J Gen Virol.* 1991;72 Pt 6):1455–9. Epub 1991/06/01. <https://doi.org/10.1099/0022-1317-72-6-1455> PMID: [2045795](https://pubmed.ncbi.nlm.nih.gov/2045795/)
36. Mazumder B, Barik S. Requirement of casein kinase II-mediated phosphorylation for the transcriptional activity of human respiratory syncytial viral phosphoprotein P: transdominant negative phenotype of phosphorylation-defective P mutants. *Virology.* 1994; 205(1):104–11. Epub 1994/11/15. <https://doi.org/10.1006/viro.1994.1624> PMID: [7975205](https://pubmed.ncbi.nlm.nih.gov/7975205/).
37. Sanchez-Seco MP, Navarro J, Martinez R, Villanueva N. C-terminal phosphorylation of human respiratory syncytial virus P protein occurs mainly at serine residue 232. *J Gen Virol.* 1995; 76 Pt 2):425–30. Epub 1995/02/01. <https://doi.org/10.1099/0022-1317-76-2-425> PMID: [7844563](https://pubmed.ncbi.nlm.nih.gov/7844563/).
38. Asenjo A, Rodriguez L, Villanueva N. Determination of phosphorylated residues from human respiratory syncytial virus P protein that are dynamically dephosphorylated by cellular phosphatases: a possible role for serine 54. *J Gen Virol.* 2005; 86(Pt 4):1109–20. Epub 2005/03/24. <https://doi.org/10.1099/vir.0.80692-0> PMID: [15784905](https://pubmed.ncbi.nlm.nih.gov/15784905/).
39. Asenjo A, Mendieta J, Gomez-Puertas P, Villanueva N. Residues in human respiratory syncytial virus P protein that are essential for its activity on RNA viral synthesis. *Virus Res.* 2008; 132(1–2):160–73. Epub 2008/01/09. <https://doi.org/10.1016/j.virusres.2007.11.013> PMID: [18179840](https://pubmed.ncbi.nlm.nih.gov/18179840/).
40. Villanueva N, Hardy R, Asenjo A, Yu Q, Wertz G. The bulk of the phosphorylation of human respiratory syncytial virus phosphoprotein is not essential but modulates viral RNA transcription and replication. *J Gen Virol.* 2000; 81(Pt 1):129–33. Epub 2000/01/21. <https://doi.org/10.1099/0022-1317-81-1-129> PMID: [10640550](https://pubmed.ncbi.nlm.nih.gov/10640550/).
41. Lu B, Ma CH, Brazas R, Jin H. The major phosphorylation sites of the respiratory syncytial virus phosphoprotein are dispensable for virus replication in vitro. *J Virol.* 2002; 76(21):10776–84. Epub 2002/10/09. <https://doi.org/10.1128/JVI.76.21.10776-10784.2002> PMID: [12368320](https://pubmed.ncbi.nlm.nih.gov/12368320/); PubMed Central PMCID: PMC136636.

42. Llorente MT, Taylor IA, Lopez-Vinas E, Gomez-Puertas P, Calder LJ, Garcia-Barreno B, et al. 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*. 2008; 72(3):946–58. Epub 2008/02/27. <https://doi.org/10.1002/prot.21988> PMID: 18300250.
43. Asenjo A, Villanueva N. Regulated but not constitutive human respiratory syncytial virus (HRSV) P protein phosphorylation is essential for oligomerization. *FEBS Lett*. 2000; 467(2–3):279–84. Epub 2000/02/17. PMID: 10675554.
44. Pereira N, Cardone C, Lassoued S, Galloux M, Fix J, Assrir N, et al. New Insights into Structural Disorder in Human Respiratory Syncytial Virus Phosphoprotein and Implications for Binding of Protein Partners. *J Biol Chem*. 2016. Epub 2016/12/30. <https://doi.org/10.1074/jbc.M116.765958> PMID: 28031463.
45. Galloux M, Gabiane G, Sourimant J, Richard CA, England P, Moudjou M, et al. Identification and characterization of the binding site of the respiratory syncytial virus phosphoprotein to RNA-free nucleoprotein. *J Virol*. 2015; 89(7):3484–96. Epub 2015/01/09. <https://doi.org/10.1128/JVI.03666-14> PMID: 25568210; PubMed Central PMCID: PMC4403424.
46. Ilinykh PA, Tigabu B, Ivanov A, Ammosova T, Obukhov Y, Garron T, et al. Role of protein phosphatase 1 in dephosphorylation of Ebola virus VP30 protein and its targeting for the inhibition of viral transcription. *J Biol Chem*. 2014; 289(33):22723–38. Epub 2014/06/18. <https://doi.org/10.1074/jbc.M114.575050> PMID: 24936058; PubMed Central PMCID: PMC4132779.
47. Bollen M, Peti W, Ragusa MJ, Beullens M. The extended PP1 toolkit: designed to create specificity. *Trends Biochem Sci*. 2010; 35(8):450–8. Epub 2010/04/20. <https://doi.org/10.1016/j.tibs.2010.03.002> PMID: 20399103; PubMed Central PMCID: PMC3131691.
48. Choy MS, Page R, Peti W. Regulation of protein phosphatase 1 by intrinsically disordered proteins. *Biochem Soc Trans*. 2012; 40(5):969–74. Epub 2012/09/20. <https://doi.org/10.1042/BST20120094> PMID: 22988849; PubMed Central PMCID: PMC3502941.
49. Cox R, Plemper RK. Structure-guided design of small-molecule therapeutics against RSV disease. *Expert Opin Drug Discov*. 2016; 11(6):543–56. Epub 2016/04/06. <https://doi.org/10.1517/17460441.2016.1174212> PMID: 27046051.
50. Jamin M, Yabukarski F. Nonsegmented Negative-Sense RNA Viruses-Structural Data Bring New Insights Into Nucleocapsid Assembly. *Adv Virus Res*. 2017; 97:143–85. Epub 2017/01/07. <https://doi.org/10.1016/bs.aivir.2016.09.001> PMID: 28057258.
51. Bitko V, Barik S. Persistent activation of RelA by respiratory syncytial virus involves protein kinase C, underphosphorylated I $\kappa$ B $\beta$ , and sequestration of protein phosphatase 2A by the viral phosphoprotein. *J Virol*. 1998; 72(7):5610–8. Epub 1998/06/17. PMID: 9621019; PubMed Central PMCID: PMC110221.
52. Cai H, Zhang Y, Lu M, Liang X, Jennings R, Niewiesk S, et al. Phosphorylation of Human Metapneumovirus M2-1 Protein Upregulates Viral Replication and Pathogenesis. *J Virol*. 2016; 90(16):7323–38. Epub 2016/06/03. <https://doi.org/10.1128/JVI.00755-16> PMID: 27252537; PubMed Central PMCID: PMC4984647.
53. Zeng Q, Huang Y, Zeng L, Lan X, He S, Zhang H. Effect of IPP5, a novel inhibitor of PP1, on apoptosis and the underlying mechanisms involved. *Biotechnol Appl Biochem*. 2009; 54(4):231–8. Epub 2009/10/31. <https://doi.org/10.1042/BA20090168> PMID: 19874272.
54. Nekhai S, Jerebtsova M, Jackson A, Southerland W. Regulation of HIV-1 transcription by protein phosphatase 1. *Curr HIV Res*. 2007; 5(1):3–9. Epub 2007/02/03. PMID: 17266553.
55. Wies E, Wang MK, Maharaj NP, Chen K, Zhou S, Finberg RW, et al. Dephosphorylation of the RNA sensors RIG-I and MDA5 by the phosphatase PP1 is essential for innate immune signaling. *Immunity*. 2013; 38(3):437–49. Epub 2013/03/19. <https://doi.org/10.1016/j.immuni.2012.11.018> PMID: 23499489; PubMed Central PMCID: PMC3616631.
56. Davis ME, Wang MK, Rennick LJ, Full F, Gableske S, Mesman AW, et al. Antagonism of the phosphatase PP1 by the measles virus V protein is required for innate immune escape of MDA5. *Cell Host Microbe*. 2014; 16(1):19–30. Epub 2014/07/11. <https://doi.org/10.1016/j.chom.2014.06.007> PMID: 25011105; PubMed Central PMCID: PMC4120867.
57. Hardy RW, Harmon SB, Wertz GW. Diverse gene junctions of respiratory syncytial virus modulate the efficiency of transcription termination and respond differently to M2-mediated antitermination. *J Virol*. 1999; 73(1):170–6. Epub 1998/12/16. PMID: 9847319; PubMed Central PMCID: PMC103820.
58. Kelker MS, Page R, Peti W. Crystal structures of protein phosphatase-1 bound to nodularin-R and tauromycin: a novel scaffold for structure-based drug design of serine/threonine phosphatase inhibitors. *J Mol Biol*. 2009; 385(1):11–21. Epub 2008/11/11. <https://doi.org/10.1016/j.jmb.2008.10.053> PMID: 18992256; PubMed Central PMCID: PMC2652124.
59. Buchholz UJ, Finke S, Conzelmann KK. Generation of bovine respiratory syncytial virus (BRSV) from cDNA: BRSV NS2 is not essential for virus replication in tissue culture, and the human RSV leader

region acts as a functional BRSV genome promoter. *J Virol.* 1999; 73(1):251–9. Epub 1998/12/16. PMID: [9847328](https://pubmed.ncbi.nlm.nih.gov/9847328/); PubMed Central PMCID: PMC103829.

60. Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, et al. The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins.* 2005; 59(4):687–96. Epub 2005/04/09. <https://doi.org/10.1002/prot.20449> PMID: [15815974](https://pubmed.ncbi.nlm.nih.gov/15815974/).
61. de Vries SJ, van Dijk M, Bonvin AM. The HADDOCK web server for data-driven biomolecular docking. *Nat Protoc.* 2010; 5(5):883–97. Epub 2010/05/01. <https://doi.org/10.1038/nprot.2010.32> PMID: [20431534](https://pubmed.ncbi.nlm.nih.gov/20431534/).
62. Wassenaar TA, van Dijk M, Loureiro-Ferreira N, van der Schot G, De Vries SJ, Schmitz C, et al. WeNMR: Structural Biology on the Grid. *J Grid Comp.* 2012; 10(4):743–67.
63. Guntert P. Automated NMR structure calculation with CYANA. *Methods Mol Biol.* 2004; 278:353–78. Epub 2004/08/20. <https://doi.org/10.1385/1-59259-809-9-353> PMID: [15318003](https://pubmed.ncbi.nlm.nih.gov/15318003/).

# New Insights into Structural Disorder in Human Respiratory Syncytial Virus Phosphoprotein and Implications for Binding of Protein Partners\*

Received for publication, November 2, 2016, and in revised form, December 21, 2016. Published, JBC Papers in Press, December 28, 2016, DOI 10.1074/jbc.M116.765958

Nelson Pereira<sup>‡1</sup>, Christophe Cardone<sup>‡1</sup>, Safa Lassoued<sup>‡1</sup>, Marie Galloux<sup>§</sup>, Jenna Fix<sup>§</sup>, Nadine Assrir<sup>‡</sup>, Ewen Lescop<sup>‡</sup>, François Bontems<sup>‡</sup>, Jean-François Elouët<sup>§</sup>, and  Christina Sizon<sup>‡2</sup>

From the <sup>‡</sup>Institut de Chimie des Substances Naturelles, UPR2301, Centre National de la Recherche Scientifique, Université Paris Saclay, 91190 Gif-sur-Yvette and the <sup>§</sup>Unité de Virologie et Immunologie Moléculaires, UR892, Institut National de la Recherche Agronomique, 78350 Jouy-en-Josas, France

Edited by Charles E. Samuel

Phosphoprotein is the main cofactor of the viral RNA polymerase of *Mononegavirales*. It is involved in multiple interactions that are essential for the polymerase function. Most prominently it positions the polymerase complex onto the nucleocapsid, but also acts as a chaperone for the nucleoprotein. *Mononegavirales* phosphoproteins lack sequence conservation, but contain all large disordered regions. We show here that N- and C-terminal intrinsically disordered regions account for 80% of the phosphoprotein of the respiratory syncytial virus. But these regions display marked dynamic heterogeneity. Whereas almost stable helices are formed C terminally to the oligomerization domain, extremely transient helices are present in the N-terminal region. They all mediate internal long-range contacts in this non-globular protein. Transient secondary elements together with fully disordered regions also provide protein binding sites recognized by the respiratory syncytial virus nucleoprotein and compatible with weak interactions required for the processivity of the polymerase.

Human respiratory syncytial virus (hRSV),<sup>3</sup> a member of the family *Pneumoviridae* (1) and order *Mononegavirales* (MNV), is the main viral cause of lower respiratory tract illness worldwide, and the main agent responsible for bronchiolitis and pneumonia in infants (2). All children have been

infected by the age of two, requiring hospitalization in ~5% cases (3). Elderly and immunocompromised adults are also at increased risk. No efficient treatment is presently available for hRSV (4), and vaccination is challenging due to complex immunogenicity (5). The search for hRSV antiviral drugs directed toward specific viral functions is therefore still ongoing (6).

The hRSV RNA-dependent RNA complex (RdRp) constitutes a virus-specific target with specific protein-protein interactions that have not all been investigated in detail (7). It uses the nonsegmented single-stranded negative sense RNA genome of hRSV as a template. In infected cells, the viral RdRp is found in specific inclusion bodies (8), which have been shown to be transcription and replication centers for other *Mononegavirales*, e.g. rabies (9) and vesicular stomatitis viruses (10). The apo RdRp complex is composed *a minima* of the large catalytic subunit (L) and its essential cofactor, the phosphoprotein (P) (11, 12). The P protein plays a central role in the RdRp by interacting with all main RdRp components. During transcription and replication it tethers the L protein to the nucleocapsid (NC), consisting of the genomic RNA packaged by the nucleoprotein (N), by direct interaction with N (13–16). hRSV P also binds to the transcription antitermination factor M2-1 (17–19). Phosphorylation of P has been proposed to regulate these interactions, although it is not essential for replication (20–22). P also acts as a chaperone for neo-synthesized N by forming an N<sup>0</sup>·P complex that preserves N in a monomeric and RNA-free state (23). We have shown previously that formation of hRSV NC·P and N<sup>0</sup>·P complexes proceeds via two distinct binding sites on P (14, 24).

Bioinformatic and biochemical investigations have established that hRSV P is tetrameric and contains large disordered N- and C-terminal regions (25–27). Fragment Y\* (Table 1) was described as a minimal oligomerization domain (OD) with predicted helical coiled-coil structure (28). However, a clear picture of the overall structure of P is still lacking, mainly because of its structural disorder. Our aim was to get a deeper insight into the structural plasticity of P and to explore the role of transiently ordered regions for interactions with hRSV RdRp proteins, here with N, by using NMR spectroscopy.

\* This work was supported by Agence Nationale de la Recherche Grants ANR-11-BSV8-024-02 and ANR-13-ISV3-0007-04 and doctoral fellowships from Institut de Chimie des Substances Naturelles (to N. P.) and Région Ile-de-France DIM Malinf (to S. L. and C. C.). The authors declare that they have no conflicts of interest with the contents of this article.

Data were deposited at the Biological Magnetic Resonance Bank with accession numbers 26902, 26903, 26904, 26905, and 26906 for  $P_{ND}$ ,  $P_{\Delta OD}$ ,  $P_{ND+OD}$ ,  $P_{FL-S23C}$ , and  $P_{CD}$ , respectively.

<sup>1</sup> These authors contributed equally to the results of this work.

<sup>2</sup> To whom correspondence should be addressed: ICSN, CNRS, 1 Avenue de la Terrasse, 91190 Gif-sur-Yvette, France. Tel.: 33-169823764; Fax: 33-169823784; E-mail: christina.sizon@cnrs.fr.

<sup>3</sup> The abbreviations used are: hRSV, human respiratory syncytial virus; CSP, chemical shift perturbation; HSQC, heteronuclear single quantum spectroscopy; IDP/R, intrinsically disordered protein/region; MPV, *Metapneumovirus*; N, nucleoprotein; NCBD, nucleocapsid binding domain; NC, nucleocapsid; OD, oligomerization domain; P, phosphoprotein; RdRp, RNA-dependent RNA polymerase; RSV, respiratory syncytial virus; SSP, secondary structure propensity; PRE, paramagnetic relaxation enhancement; hMPV, human *Metapneumovirus*; IAP, 3-(2-iodoacetamido)-PROXYL.

## Results

**Extent of Intrinsically Disordered Regions in hRSV Phosphoprotein**—To probe the structural organization of hRSV phosphoprotein at the single residue level by NMR, we first used full-length P protein ( $P_{FL}$ ). The two-dimensional  $^1H$ - $^{15}N$  HSQC spectrum of  $P_{FL}$  exhibits sharp amide signals with narrow  $^1H$  chemical shift dispersion in the 7.5–8.5 ppm range (Fig. 1). This is the signature of intrinsically disordered proteins (IDPs) and regions (IDRs) (29). Experimental conditions were adjusted for the detection of IDP amide protons, low temperature (288 K) and acidic pH (6.5), to reduce the contribution of water exchange to line widths. Only 60% (140 of 229) of expected  $P_{FL}$  amide signals were observed. 40% of amide signals were too broad to be detected and correspond to protein regions undergoing dynamic processes and conformational exchange at the  $\mu s$ -ms time scale. They comprise the OD (fragment  $Y^*$ ,  $\sim 40$  residues), but also extensive additional regions ( $\sim 50$  residues).

Due to the structural heterogeneity of P, we resorted to protein fragments to delineate domains. Several fragments had been produced before (13, 24, 25, 30). Constructs are detailed in Table 1. In particular,  $P_{ND+OD}$  and  $P_{OD+CD}$  correspond to the N- and C-terminal domains with OD.  $P_{ND}$  and  $P_{CD}$  are their counterparts without OD. The  $^1H$ - $^{15}N$  HSQC spectra of P fragments exhibit sharp lines and narrow  $^1H$  chemical shift dispersion, similarly to  $P_{FL}$  (Fig. 1). All signals superimpose well, showing that the fragments are representative of the corresponding domains in  $P_{FL}$ . For instance, overlay of  $P_{ND+OD}$  and  $P_{OD+CD}$  spectra reproduces the spectrum of  $P_{FL}$  (Fig. 1). Comparison of  $P_{ND}$  and  $P_{ND+OD}$  indicates that the OD signals are missing for  $P_{ND+OD}$ . Remarkably, fragment  $P_{CD}$  displays more signals than  $P_{OD+CD}$ , revealing that the C-terminal domain of P contains residues that are not completely disordered when attached to the OD (Fig. 1).

Sequential assignment of backbone chemical shifts was carried out separately for all fragments. The signals of the 120 N-terminal residues and 40 C-terminal residues could be observed for all constructs, including  $P_{FL}$ , indicating that they form two independent N- and C-terminal IDRs in P. The signals of the Asp<sup>125</sup>-Thr<sup>160</sup> region, equivalent to fragment  $Y^*$  (Table 1), were missing for all constructs containing the OD. Residues Ser<sup>161</sup>-Glu<sup>204</sup> were also missing in the spectra of  $P_{OD+CD}$  and  $P_{FL}$ , but were present in the spectrum of  $P_{CD}$ .

**Determination of Transient Secondary Structure Elements in hRSV Phosphoprotein**—On closer inspection, many NMR signals display local heterogeneity in intensity and line width. Taking advantage of the sensitivity of  $^{13}C$  backbone and  $^1H\alpha$  chemical shifts to protein dihedral angles, we determined residue-specific secondary structure propensities (SSPs) in  $P_{ND}$ ,  $P_{ND+OD}$ ,  $P_{CD}$ ,  $P_{\Delta OD}$ , and  $P_{FL}$ , using Talos+ (31) (Fig. 2A). Outside the Ser<sup>161</sup>-Glu<sup>204</sup> region no significant secondary structure was detected. This confirms that the N and C termini of P are fully disordered. Still, weak  $\alpha$ -helical propensity is observed in the N-terminal IDR for residues Asp<sup>12</sup>-Ile<sup>24</sup> and Phe<sup>98</sup>-Lys<sup>103</sup>. The latter define two transient helices  $\alpha_{N1}$  and  $\alpha_{N2}$  (Fig. 2B).

In contrast, residues Leu<sup>173</sup>-Lys<sup>205</sup>, which can only be observed in P fragments devoid of the OD, display high  $\alpha$ -hel-

ical propensity and define two helices,  $\alpha_{C1}$  (Leu<sup>173</sup>-Met<sup>187</sup>) and  $\alpha_{C2}$  (Asn<sup>189</sup>-Lys<sup>205</sup>), with up to 70 and 95% propensity, respectively (Fig. 2, A and B). Because SSPs depend on the model used to extract them, we also extracted SSPs with  $\delta 2D$  (32). We obtained lower  $\alpha$ -helical propensities than with Talos+ (up to 20% for  $\alpha_{C1}$  and 70% for  $\alpha_{C2}$ ), indicating that these helices are also transient, with  $\alpha_{C2}$  being almost stable. The  $\alpha_{C1/2}$  region does not induce oligomerization on its own, because  $P_{\Delta OD}$  and  $P_{CD}$  display the same line widths as  $P_{ND}$  (Fig. 1).  $\alpha_{C1}$  and  $\alpha_{C2}$  thus form a second C-terminal IDR with high  $\alpha$ -helical propensity, in addition to the fully disordered C terminus. This domain likely accounts for the thermal transition observed in P at physiological temperature (27), indicating that these helices do not tightly associate in the P tetramer. Signal broadening of  $\alpha_{C1/2}$  in tetrameric P fragments may be explained by the increased molecular size, which affects overall dynamics in solution, or by interactions between these transient helices and possibly with the OD.

**Investigation of the Dynamics of hRSV Phosphoprotein by  $^{15}N$  Nuclear Relaxation**—We measured  $^{15}N$  relaxation parameters for a tetrameric ( $P_{ND+OD}$ ) and three monomeric ( $P_{ND}$ ,  $P_{CD}$ , and  $P_{\Delta OD}$ ) fragments of P (Fig. 2C) to analyze the dynamic behavior of the different regions of P. P IDRs display overall homogeneous, negative or near zero heteronuclear  $^1H$ - $^{15}N$  nuclear Overhauser effects (NOEs), indicative of large amplitude backbone fluctuations on a ps-ns time scale. The  $\alpha_{C1/2}$  region displays higher NOE values (0–0.5), suggesting that it is more ordered. This is also the case for the transient  $\alpha_{N1}$  helix.  $^{15}N$  transverse relaxation rates ( $R_2$ ) are more heterogeneous along the sequence of P, but consistent among P fragments.  $R_2$  values are much higher for regions with  $\alpha$ -helical SSP as compared with completely random regions, denoting differential tumbling in solution and conformational exchange between disordered and ordered conformational states, on a  $\mu s$ -ms time scale. A fifth region (Asn<sup>78</sup>-Ser<sup>86</sup>) without clear SSP displays similar behavior (Fig. 2C), suggesting that exchange broadening also arises from internal interactions.  $^{15}N$  longitudinal relaxation rates ( $R_1$ ), which are not sensitive to exchange, are nearly uniform along the sequence of P, but underline the structural singularity of  $\alpha_{C1}$ . Relaxation parameters of P IDRs are globally independent of the length of the fragments, suggesting that their motions are not significantly restricted and that they are not stably associated with any part of the protein.

**Detection of Long-range Contacts in hRSV P by Paramagnetic Relaxation Enhancement**—Next, we used paramagnetic relaxation enhancement (PRE) to investigate the spatial organization of P. Line broadening due to PRE in a  $\sim 15$  Å radius around a paramagnetic spin label can be used to measure long-range distances by NMR in globular proteins, but also to probe long-range contacts in IDPs (33, 34). The sequence of  $P_{FL}$  does not contain any cysteine. We therefore introduced cysteines by mutating individual residues distributed along the sequence of  $P_{FL}$  and labeled them with IAP free radical (35). The  $\alpha_{C1/2}$  region remained undetectable in the  $^1H$ - $^{15}N$  HSQC spectra of all Cys mutants in their diamagnetic state, similarly to wild type  $P_{FL}$ . We therefore concluded that the mutations did not impact P oligomerization. Hence PREs could only be measured outside the OD and  $\alpha_{C1/2}$  region.



**TABLE 1**  
Definition of hRSV phosphoprotein fragments

Designation of P fragments	Boundaries
P <sub>FL</sub>	M1-F241
P <sub>ND+OD</sub>	M1-R163
P <sub>ND</sub>	M1-Q126
P <sub>OD+CD</sub>	T127-F241
P <sub>CD</sub>	S161-F241
P <sub>ΔOD</sub>	M1-E121+S161-F241
P <sub>161-229</sub>	S161-G229
P <sub>40</sub>	M1-I40
Fragment X (41)	E104-R163
Fragment Y* (28)	S119-T160

stronger PREs were measured in regions with transient structural elements and in the Asn<sup>78</sup>-Ser<sup>86</sup> region, identified by <sup>15</sup>N relaxation measurements. Long-range contacts were detected within the N-terminal IDR of P using the S23C and S99C mutants, but also between this region and the OD, using the S143C and S156C mutants, and even the C-terminal  $\alpha$ -helical IDR using E179C and E193C. Only the C terminus does not appear to be involved in any specific contact, because strong PREs are only observed for proximal spin labels. The absence of PREs between the C- and N-terminal IDRs moreover provides evidence for a parallel organization of the hRSV P tetramer.

**Interactions of hRSV P with Nucleocapsid Analogs**—Because transiently structured regions of IDPs are potential molecular recognition elements (36, 37), we carried out NMR interaction experiments to investigate the impact of transient structures within P on hRSV nucleoprotein binding. We first tested N in the form of N-RNA rings, which mimic the hRSV nucleocapsid (38). As the nucleocapsid binding domain of P (P<sub>NCBD</sub>) had been assigned to its 9 C-terminal residues (13), we worked with <sup>15</sup>N-labeled P<sub>CD</sub> instead of P<sub>FL</sub>. In the presence of N-RNA rings, the signals of the eight last residues are completely broadened out in the <sup>1</sup>H-<sup>15</sup>N HSQC spectrum of P<sub>CD</sub> (Fig. 4, A and B), confirming that they are involved in P binding to the NC. An overall 15% intensity loss seems to arise from increased viscosity, as assessed by a control experiment with excess BSA (Fig. 4B).

Direct observation of N-bound P residues by <sup>1</sup>H-<sup>15</sup>N HSQC could not be achieved due to the large size of N-RNA rings. We therefore proceeded with the monomeric N-terminal domain of N (N<sub>NTD</sub>), which was shown to be relevant for NC·P binding (14, 39). By titrating N<sub>NTD</sub> into <sup>15</sup>N-P<sub>CD</sub>, linear chemical shift perturbations (CSPs) were observed for P<sub>NCBD</sub> (e.g. Asn<sup>234</sup> and Ser<sup>237</sup> in Fig. 4C). A titration end point was reached with 5 eq of N<sub>NTD</sub> (Fig. 4C). By fitting the CSP data with a two-site fast exchange model, a  $K_d$  of 25–50  $\mu$ M was determined, in agreement with a  $K_d$  of 30  $\mu$ M previously determined for the N<sub>NTD</sub>·P<sub>CD</sub> complex by isothermal titration calorimetry (39).

Notably, CSPs and line broadening occurred concomitantly, pointing to moderately fast exchange between free and N<sub>NTD</sub>-

bound P. However, at the titration end point lines are broader than expected for a 35-kDa complex (Fig. 4C). Surprisingly, no signal was recovered for the very C-terminal Phe<sup>241</sup> residue, which in X-ray structures appears to be the main structural determinant for NC·P complex formation, by tightly inserting into a hydrophobic pocket at the surface of N<sub>NTD</sub> (39). Our results thus suggest the possibility of additional binding modes corresponding to weaker interactions. Indeed, the large line broadening observed for  $\alpha_{C1}$  (Fig. 4D) would be in favor of a secondary binding site on  $\alpha_{C1}$ . This effect is also observed with the P<sub>161-229</sub> fragment deleted of P<sub>NCBD</sub>, although attenuated (Fig. 4D), suggesting that binding to  $\alpha_{C1}$  could be promoted by binding to P<sub>NCBD</sub>.

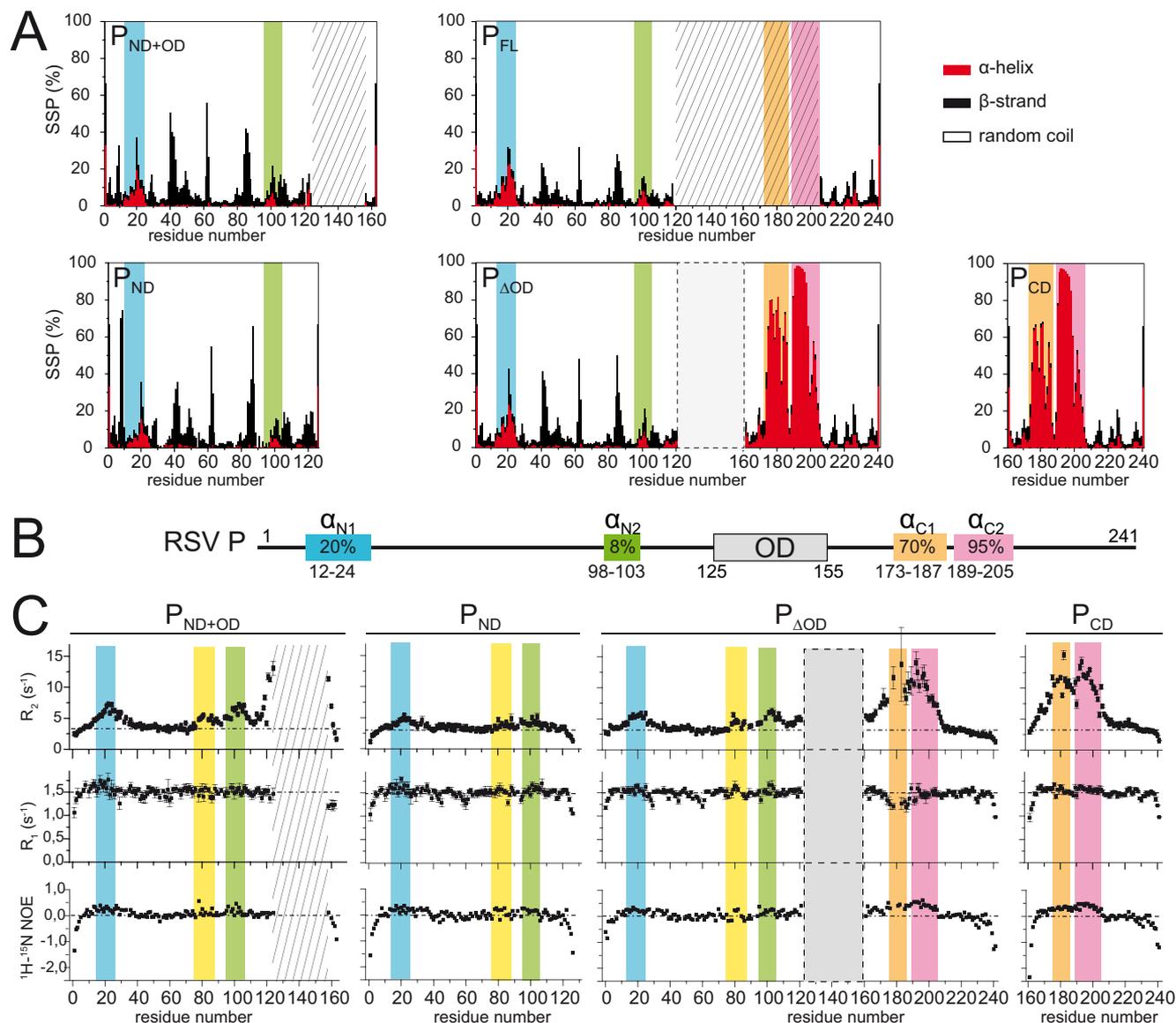
**Investigation of the hRSV N<sup>0</sup>-P Binding Mode**—In a last part we investigated the N<sup>0</sup>-P binding mode by using N<sup>mono</sup>, a monomeric N mutant impaired for RNA binding (24). N<sup>mono</sup> leads to line broadening at the  $\alpha_{N1}$  site in P<sub>ND+OD</sub> (Fig. 4, E and F), in agreement with previous results that showed that the N-terminal P<sub>40</sub> peptide was able to pull down N<sup>mono</sup> (24). NMR interaction experiments with N<sup>mono</sup> and P<sub>FL</sub> show that N<sup>mono</sup> is competent for both NC·P and N<sup>0</sup>-P binding modes, via P<sub>NCBD</sub> and  $\alpha_{N1}$ , respectively (Fig. 4F). Because both sites in P<sub>FL</sub> can be occupied with only 1 eq of N<sup>mono</sup>, either none of the complexes is very tight or the two sites are not mutually exclusive.

Unexpectedly, a third region,  $\alpha_{N2}$ , was perturbed (Fig. 4, E and F). As previous experiments with the P<sub>60-126</sub> fragment showed that this region does not bind to N<sup>mono</sup> (24), line broadening of  $\alpha_{N2}$  signals would not be explained by direct binding to N. However, we showed by PRE that  $\alpha_{N2}$  transiently associates with  $\alpha_{N1}$ . Formation of an N<sup>0</sup>· $\alpha_{N1}$  complex could displace the equilibrium between free and  $\alpha_{N1}$ -bound  $\alpha_{N2}$ .

As a control, we performed additional interaction experiments with <sup>15</sup>N-labeled N-terminal P fragments and N<sub>NTD</sub>. Surprisingly, we observed perturbations (Fig. 5, A and B). However, the intensity ratio patterns are different from N<sup>mono</sup>. N<sup>mono</sup> perturbed a large region (Met<sup>1</sup>-Ser<sup>30</sup>), whereas N<sub>NTD</sub> affects only a few residues around Lys<sup>25</sup> (Figs. 4F and 5B), suggesting a difference in binding. We carried out complementary experiments by measuring spectra of <sup>15</sup>N-N<sub>NTD</sub> in the presence of P<sub>40</sub>. Line broadening was induced in different N<sub>NTD</sub> regions (Fig. 5, C and D) delineating a contiguous surface on N (Fig. 5E). This surface is on the inside of the nucleocapsid as opposed to the binding site of the C terminus of P in the NC·P complex (Fig. 5E). It partly overlaps with the interaction surface of the N-terminal arm of the adjacent protomer in N-RNA rings, but is shifted with respect to the interaction surface recently published for the human *Metapneumovirus* (hMPV) N<sup>0</sup>-P complex, where the N-terminal P<sub>1-28</sub> peptide obstructs the binding sites of the N- and most prominently C-terminal arms of adjacent N protomers (40).

**FIGURE 1. <sup>1</sup>H-<sup>15</sup>N HSQC spectra of <sup>15</sup>N-labeled RSV phosphoprotein.** A, schematic representation of the boundaries of RSV P and P fragments measured by NMR. Deleted regions are symbolized by boxes in broken line. Hatched areas indicate protein regions with missing amide assignments. B, <sup>1</sup>H-<sup>15</sup>N HSQC spectra of P<sub>FL</sub>, P<sub>ND+OD</sub>, P<sub>OD+CD</sub>, P<sub>ND</sub>, P<sub>CD</sub>, and P<sub>ΔOD</sub> were acquired under identical experimental conditions (50–100  $\mu$ M concentration, 288 K temperature, 14.1 T magnetic field). A different color was used for each construct to plot contours and show assignments, with full-length P (P<sub>FL</sub>) in black. For all deletion mutants, the spectrum is superimposed onto that of P<sub>FL</sub> for comparison. Amide resonance assignments are indicated for P<sub>ND+OD</sub>, P<sub>OD+CD</sub>, P<sub>ND</sub>, and P<sub>CD</sub> (residue number and amino acid type in single-letter code). Asn and Gln NH<sub>2</sub> side chain signals are not individually assigned. For P<sub>ND</sub> the inset shows assignments of the crowded central region of the spectrum.

## Linking Interactions to Disorder in RSV Phosphoprotein



**FIGURE 2. Characterization of structural heterogeneity in hRSV phosphoprotein by backbone chemical shift analysis and <sup>15</sup>N relaxation.** *A*, stack plots represent secondary structure propensities (α-helix, β-strand, random coil) of P constructs determined from <sup>1</sup>H<sub>N</sub>, <sup>15</sup>N, <sup>1</sup>Hα, <sup>13</sup>C', <sup>13</sup>Cα, and <sup>13</sup>Cβ chemical shift analysis with Talos+ (31). Notable regions are highlighted by background colors. *Hatched* areas indicate protein regions with missing amide signal. *B*, structural organization of hRSV P as determined from *A*. Transient α-helices are represented as boxes with the same color code as in *A*. Maximal α-helical propensity indicated inside. *C*, <sup>15</sup>N R<sub>1</sub> and R<sub>2</sub> relaxation rates and heteronuclear <sup>1</sup>H-<sup>15</sup>N NOEs were measured for four P fragments. *Error bars* are S.D. produced by covariance matrix analysis. *Dotted horizontal lines* indicate mean values taking only disordered residues into account. The location of transient α-helices is indicated with the same color code as in *B*. An additional region with dynamics different from IDRs, but not identified in *A*, is highlighted in yellow. The hatched area corresponds to the OD. The box in broken line indicates the deleted sequence in P<sub>ΔOD</sub>.

## Discussion

**Domain Organization in hRSV P**—hRSV phosphoprotein is characterized by extensive structural disorder that hampers high resolution structural characterization. The only structured part appears to be the tetramerization domain, which has been investigated by bioinformatic tools (26), resistance to protease digestion (28, 41), and deletion series (13, 25, 42). Although fragment Y\*, which exhibits high stability and homogeneity (25, 27, 28), has been acknowledged as the core of the tetramerization domain, the OD of hRSV P is often represented by fragment X (Table 1), longer than Y\* by 15 residues at its N terminus. Predicted structural models of fragment X present this stretch as a second short coiled-coil domain (28, 42). We

show here that this stretch has no significant SSP and is highly dynamic in solution, at least in the absence of protein partners. This is also in agreement with the observation that P<sub>ND</sub> was not able to form oligomers (25).

In contrast to *Mononegavirales* N proteins, which share a conserved fold (43), P proteins have largely diverged, hRSV P having the shortest sequence (44). However, several regions of P are highly conserved among *Pneumoviridae*, more particularly the OD (Fig. 6A). Its structure was solved by X-ray diffraction for hMPV P (45). It displays a coiled-coil helical conformation for a region equivalent to the hRSV fragment Y\*, indicating that a short OD is specific of this family. The protomers are arranged in a parallel orientation, consistently with the PRE

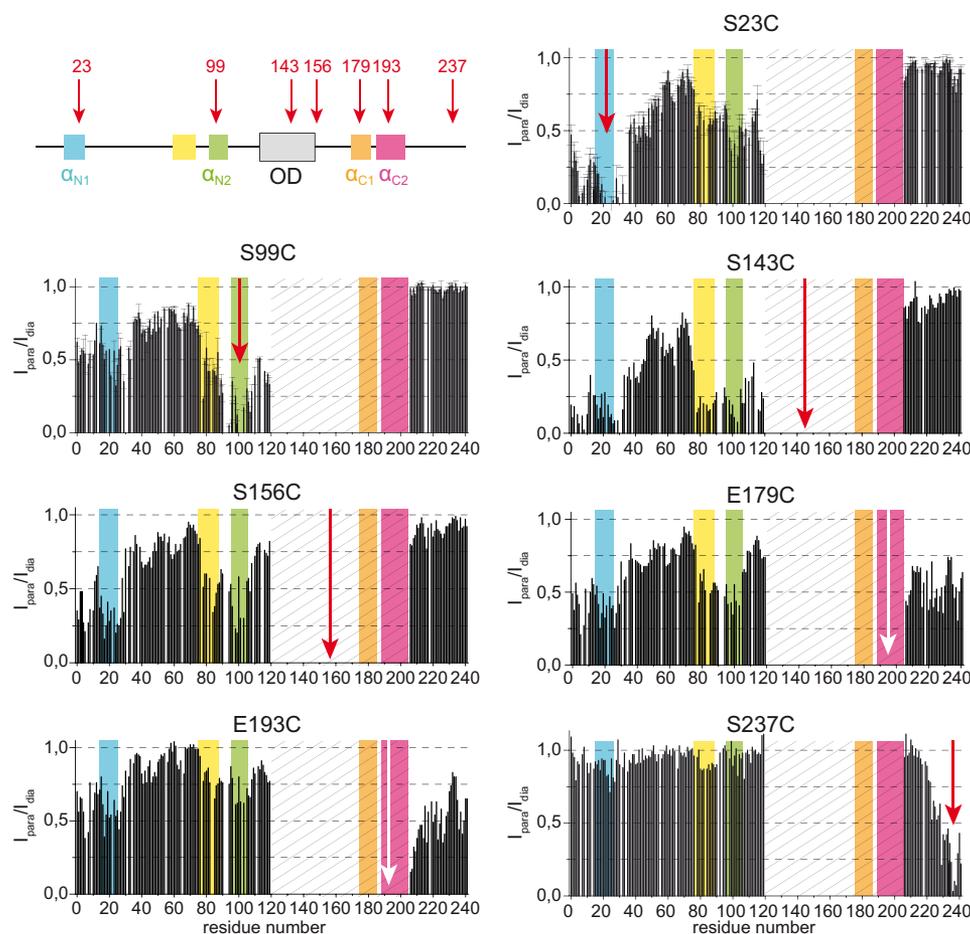


FIGURE 3. Evidence for long-range contacts in hRSV P by paramagnetic relaxation enhancement. PREs were measured with nitroxide spin labels at positions 23, 99, 143, 156, 179, 193, and 237 in the sequence of P. Hatched areas indicate regions with missing amide signal in the diamagnetic state. Error bars represent the root mean square deviations calculated from two data sets. The position of the spin label is indicated by an arrow on each bar diagram. Transient secondary structure elements are indicated by colored background.

results for hRSV P, and not as a dimer of anti-parallel dimers like in Mumps virus P (46, 47).

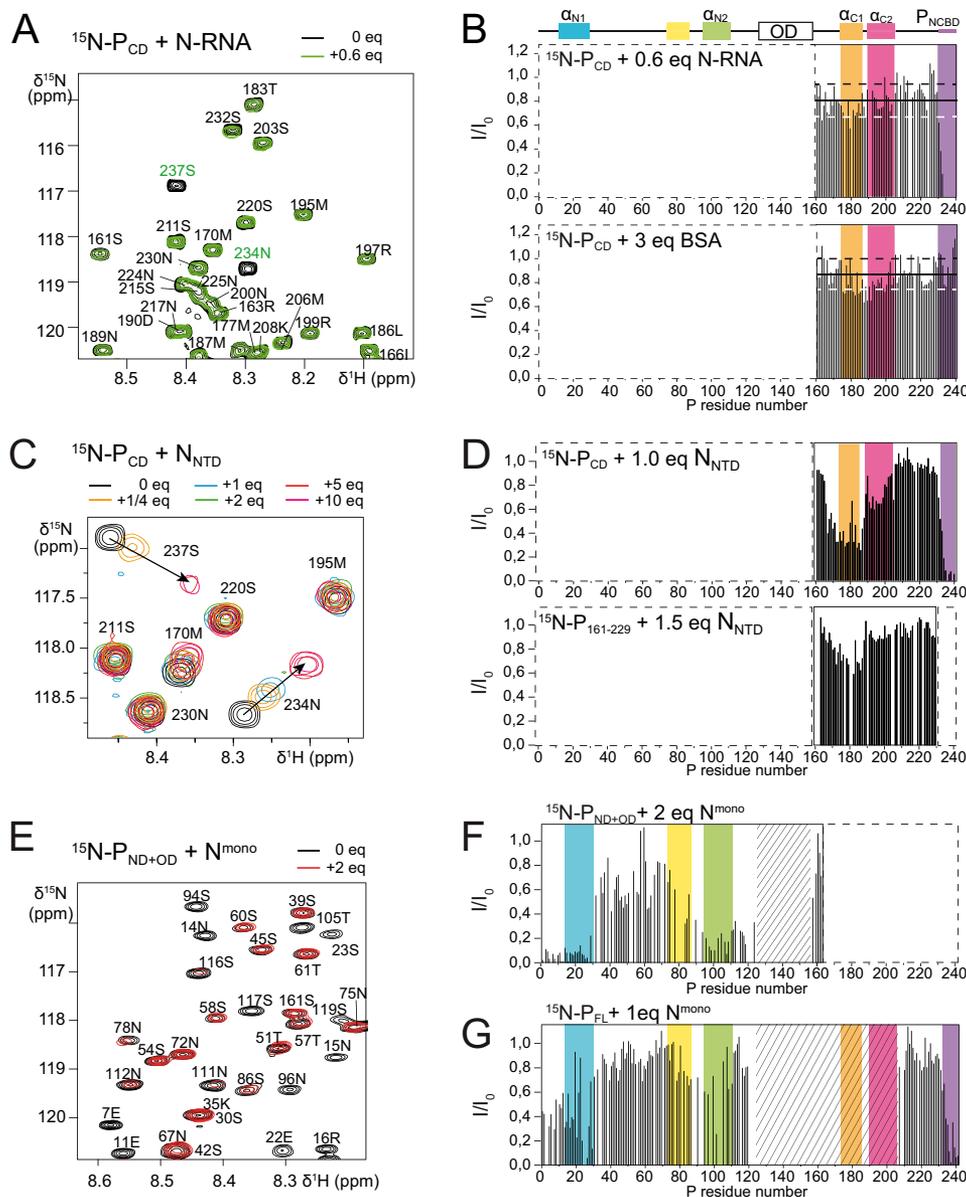
Finally, we were able to identify a C-terminal domain with high  $\alpha$ -helical propensity, whose dynamics are distinct from those of the OD and the fully disordered C terminus. The presence of transient C-terminal helices was also proposed for hMPV, based on small angle X-ray scattering data (45), showing that they constitute another hallmark of *Pneumoviridae*. But whereas  $\alpha_{C1}$  is partly conserved,  $\alpha_{C2}$  appears to be specific of the *Orthopneumovirus* genus (Fig. 6A). A tentative structural model of P in its most disordered state, summarizing the structural information determined by NMR, is given in Fig. 6B.

**Functional Relevance of Transiently Structured Regions in hRSV P**—The central role of the phosphoprotein in hRSV replication is associated to its role as a hub inside the RdRp complex, mediating interactions with both viral and cellular proteins (18, 24, 48). We show here that two regions identified before as binding regions for N<sup>mono</sup> (24) and M2-1 (17, 18),  $\alpha_{N1}$  and  $\alpha_{N2}$ , respectively, display weak  $\alpha$ -helical propensity. These transient helices might fold completely upon binding, as related for other protein-protein interactions (36, 37). Such a hypothesis would be supported by the large number ( $\sim 30$ ) of N-terminal P residues involved in the interaction with N<sup>mono</sup> (Fig. 4). Our definition of  $\alpha_{N1}$  matches with the second helix formed by

hMPV P<sub>1–28</sub> in the X-ray structure of the N<sup>0</sup>-P complex (40). The sequence of  $\alpha_{N1}$  is rather well conserved among *Pneumoviridae* (Fig. 6A), suggesting that  $\alpha_{N1}$  constitutes a molecular recognition element for N<sup>0</sup>.

Contradictory data are available for the interaction properties of the  $\alpha$ -helical C-terminal domain of P. Immunoprecipitation assays showed that deletion of residues 160–180 impaired N binding in bRSV P (49), but not in hRSV P (50), despite high sequence conservation of N and P between hRSV and bRSV. This internal region partly overlaps with  $\alpha_{C1}$  and was also perturbed in NMR interaction experiments with N<sub>NTD</sub> (Fig. 4), which prompted us to propose it as a secondary N binding site. Moreover, temperature-sensitive mutations of P were reported in the same region (50). G172S and E176G do not support replication at 37 °C, which was linked to decreased interaction between N and P. The triple mutant R174A/E175A/E176A proved non-functional in a minigenome assay, but was still able to bind to N. Finally, a recombinant E176G RSV virus reverted to Asp<sup>176</sup> (50). These data underline the functional importance of residue Glu<sup>176</sup>, which contributes to a negative cluster with Glu<sup>175</sup> and Glu<sup>179</sup>, exacerbated in the structural context of the  $\alpha_{C1}$  helix (Fig. 6B). If we link these results with the lability of  $\alpha_{C1}$ , it appears that structural modulation of this helix could have a direct impact on RSV replication. In contrast,

## Linking Interactions to Disorder in RSV Phosphoprotein



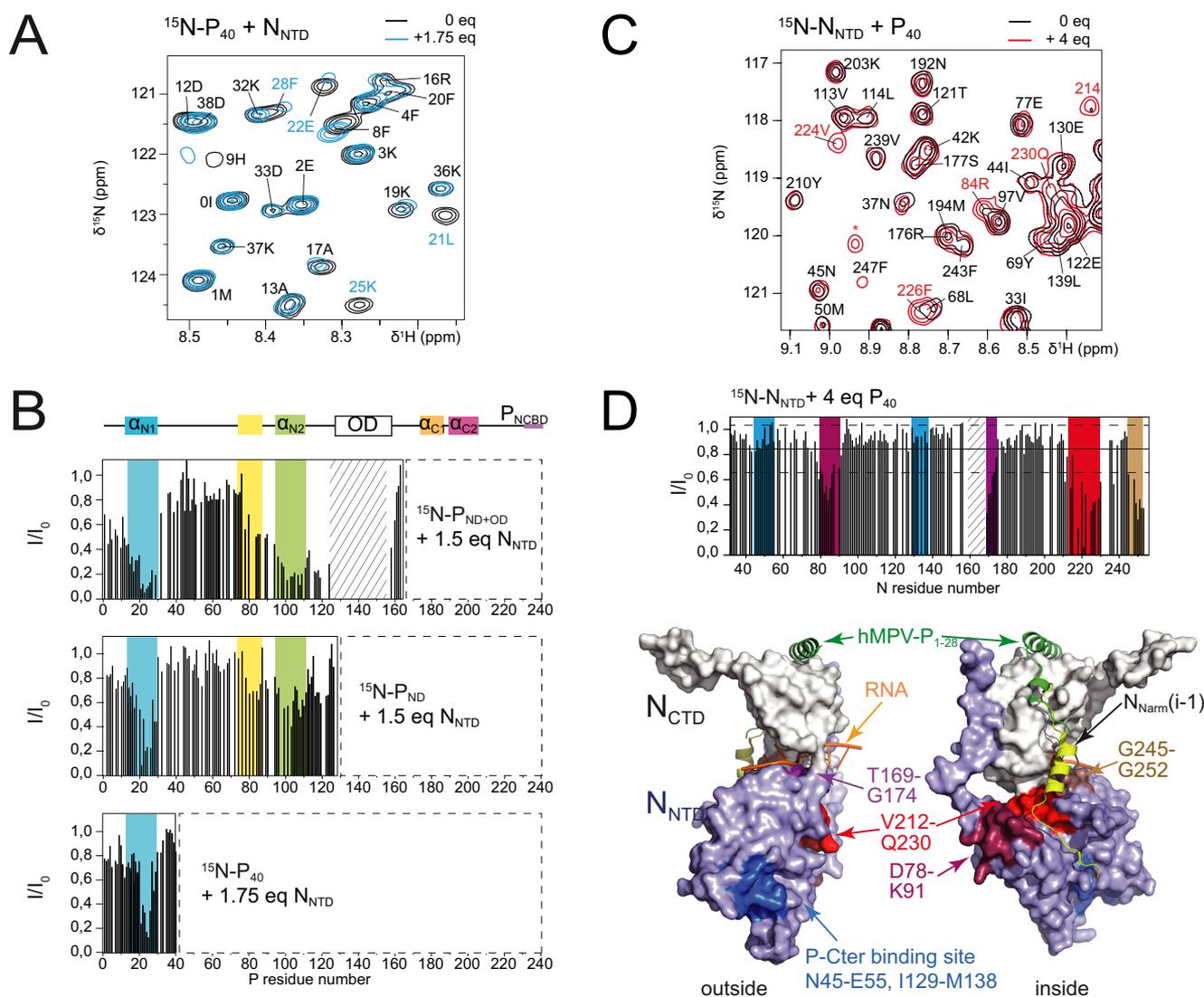
**FIGURE 4. Probing N/P interactions by NMR using  $^{15}\text{N}$ -labeled P constructs.** *A*, selected region of the  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum of  $^{15}\text{N}$ -P<sub>CD</sub> before and after addition of N-RNA rings. *B*, intensity ratios ( $I/I_0$ ) for  $^{15}\text{N}$ -P<sub>CD</sub> in the presence of N-RNA rings measured from the  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra shown in *A*. A control experiment was carried out with 3 eq of BSA. Mean  $\pm$  S.D. are indicated with *full* and *broken horizontal lines*, respectively. Background colors indicate transient secondary structure elements and the nucleocapsid binding domain (P<sub>NCBD</sub>). *C*, titration of  $^{15}\text{N}$ -P<sub>CD</sub> by N<sub>NTD</sub> followed by chemical shift and intensity variations. *D*, intensity ratios measured for  $^{15}\text{N}$ -P<sub>CD</sub> and  $^{15}\text{N}$ -P<sub>161-229</sub> in the presence of N<sub>NTD</sub>. *E*, selected region of the  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum of  $^{15}\text{N}$ -P<sub>ND+OD</sub> before and after addition of N<sup>mono</sup>. *F*, intensity ratios for  $^{15}\text{N}$ -P<sub>ND+OD</sub> in the presence of N<sup>mono</sup>. *G*, intensity ratios measured for  $^{15}\text{N}$ -P<sub>FL</sub> in the presence of N<sup>mono</sup>.

Leu<sup>198</sup>-Asn<sup>217</sup> has been reported to be a “negative N-binding region,” the deletion of which enhanced N binding (16). This region partly overlaps with  $\alpha_{C2}$  and was not affected by N<sub>NTD</sub> in our experiments (Fig. 4).

*Disordered Regions in hRSV P Mediate Diffuse as Well as Specific Interactions*—In light of the X-ray structure of the hMPV N<sup>0</sup>-P complex (40), the hRSV P<sub>40</sub>/N<sub>NTD</sub> interaction does not seem to be relevant for the N<sup>0</sup>-P complex. However, it may shed light on the binding properties of  $\alpha_{N1}$ . The interaction surfaces on N<sub>NTD</sub> and N<sup>0</sup> are proximal and connected by the interaction surface with the N-terminal arm of an adjacent protomer in N-RNA rings (Fig. 5E). The P<sub>40</sub>/N<sub>NTD</sub> interaction might therefore correspond to an intermediate state on the

binding/folding pathway of  $\alpha_{N1}$ , whereby hydrophobic interactions play a role, as shown by the hydrophobic face in fully formed  $\alpha_{N1}$  (Fig. 6B), involved in N<sup>0</sup> binding (40).  $\alpha_{N1}$  appears to be a sticky helix that is able to mediate external (N/P) as well as internal interactions, the latter being favored by the tetrameric organization of P.

More generally, we identified transient internal long-range contacts in hRSV by PRE, mainly mediated by regions with SSPs. These seemingly unspecific interactions may play several roles. Compaction of the structure of P can be achieved, preventing unspecific interactions with other proteins in the cell. The interplay between  $\alpha_{N1}$  and  $\alpha_{N2}$ , which both expose hydrophobic faces when stabilized (Fig. 6B), suggests that upon bind-



**FIGURE 5. Interaction between hRSV  $\text{N}_{\text{NTD}}$  and the N terminus of P.** *A*, selected region of the  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum of  $^{15}\text{N-P}_{40}$  in the presence or absence of  $\text{N}_{\text{NTD}}$ . *B*, intensity ratios ( $I/I_0$ ) measured from  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of  $^{15}\text{N}$ -labeled N-terminal fragments of P in the presence of  $\text{N}_{\text{NTD}}$ . The hatched area indicates missing signals in the free forms. The background is colored according to transient secondary structures as indicated above. *C*, selected region of the  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum of  $^{15}\text{N-N}_{\text{NTD}}$  in the presence or absence of  $\text{P}_{40}$ . *D*, intensity ratios for  $^{15}\text{N-N}_{\text{NTD}}$  in the presence of  $\text{P}_{40}$  measured from *C*. Full and broken horizontal lines indicate the mean  $\pm$  S.D., respectively. Regions with significant line broadening ( $I/I_0 < \text{mean} - 1 \text{ S.D.}$ ) are highlighted by red to brown backgrounds and the P C terminus binding region by blue color (39). The hatched area corresponds to unassigned  $\text{N}_{\text{NTD}}$  residues Ile<sup>160</sup>-Val<sup>167</sup>. The same color code is used to map these regions on the three-dimensional structure of an N protomer of N-RNA rings (PDB 2WJ8). The two views in surface representation are rotated by 180° and show the outside and inside of the N-RNA ring. The N-terminal arm (Ala<sup>2</sup>-Gly<sup>30</sup>) of the preceding N protomer in the ring (i-1) is in yellow schematic. The N-terminal P<sub>1-28</sub> peptide of the hMPV N<sup>0</sup>-P complex (PDB 5FVD) (40) in green schematic was superimposed onto hRSV N.

ing to one site, another site could become competent for binding. Plasticity of the structure of P can also help fulfilling the requirement of simultaneous binding at the C and N terminus to a same protein partner, *e.g.* to N. At the same time diffuse interactions may play a role by retaining RdRp relevant proteins in hRSV inclusion bodies (8).

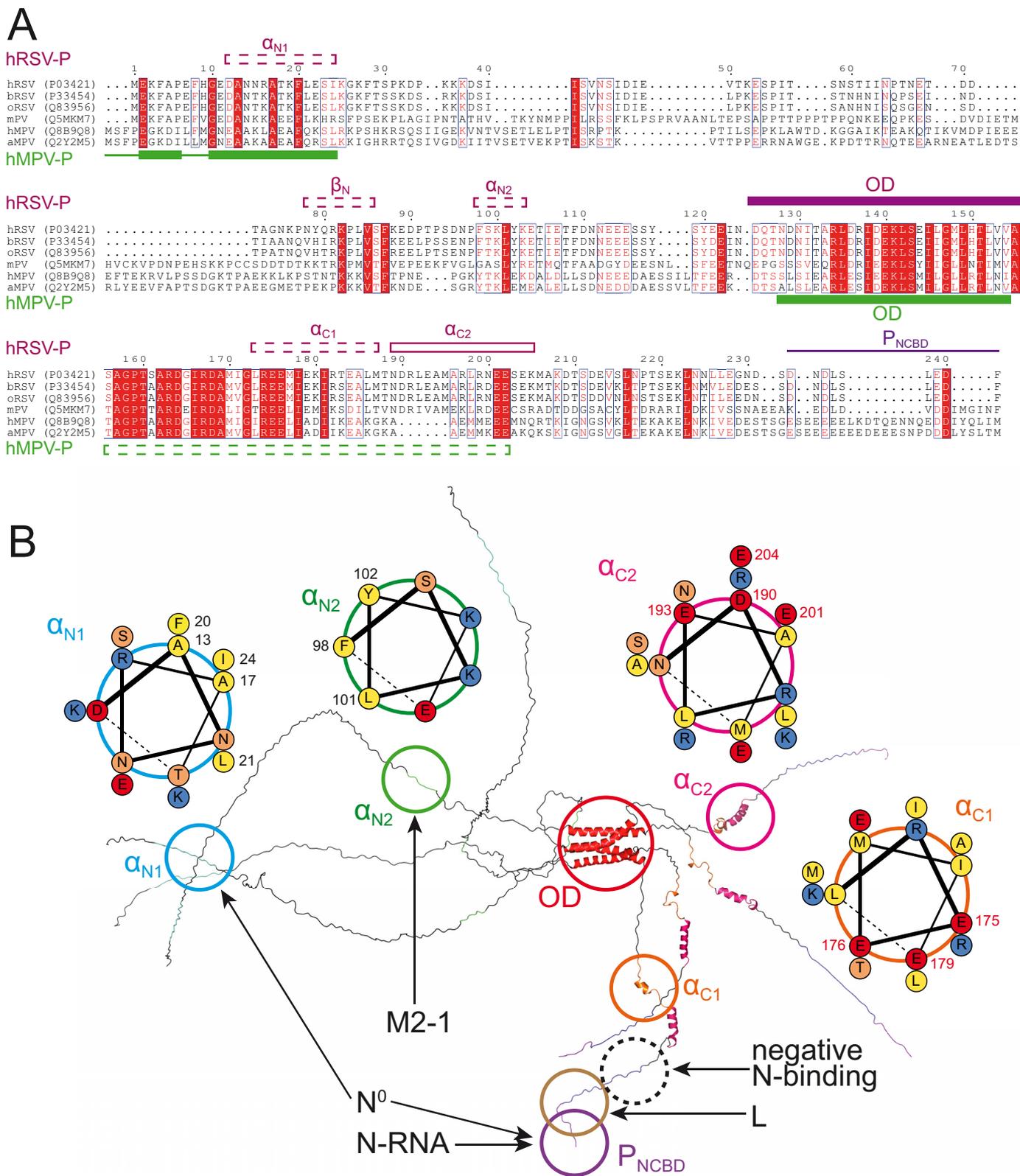
The hRSV L binding site was recently reported to span residues Pro<sup>218</sup>-Glu<sup>239</sup> (Fig. 6B), and it was proposed that this region might fold into a helix (30). Under our experimental conditions this region did not display any significant SSP, conformational exchange, or internal contacts, similarly to  $\text{P}_{\text{NCBD}}$ . Analogy with  $\text{P}_{\text{NCBD}}$  suggests that this region might not fold upon binding. Except for Phe<sup>241</sup>,  $\text{P}_{\text{NCBD}}$  is disordered in its bound form, as shown by X-ray crystal structures of  $\text{N}_{\text{NTD}}$  complexed to C-terminal P peptides (39). Although Phe<sup>241</sup> is essen-

tial for NC:P binding, fuzzy electrostatic interactions mediated by acidic residues as well as phosphorylation of serines in  $\text{P}_{\text{NCBD}}$  significantly contribute to the affinity of P (14, 39, 51). Our NMR data suggest that even Phe<sup>241</sup>, the linchpin for P binding to the NC, could explore different binding sites. Interactions that come into play in the NC:P complex are thus based on disorder and on a balance between recognition and moderate affinity, required for the processivity of the polymerase. This could be a more general scheme for P interactions with other RdRp components.

### Experimental Procedures

**Plasmids**—Plasmids for expression of recombinant hRSV proteins in *E. coli* were described previously for N terminally GST-fused hRSV phosphoprotein and P fragments listed in

# Linking Interactions to Disorder in RSV Phosphoprotein



**FIGURE 6. Location of structural elements and binding sites in hRSV P.** *A*, sequence alignment of *Pneumoviridae* phosphoproteins for human, bovine, and ovine respiratory syncytial virus, murine pneumonia virus, and avian metapneumovirus, generated with Clustal Omega (57) and displayed with ESPrict (58). Uniprot accession numbers are shown in parentheses. On top are indicated hRSV P secondary structure elements determined by NMR, with almost stable and transient elements in full and broken lines, respectively. At the bottom are indicated known secondary structure elements for hMPV P (40, 45). *B*, structural model of the hRSV P tetramer. The location of regions with  $\alpha$ -helical propensity is shown and their sequences are given in helical wheel representation. The location of RdRp protein binding sites is also shown, for regions with and without significant SSPs.

Table 1 (13, 24, 25, 30) and C terminally His-tagged hRSV nucleoprotein (13), N<sub>NTD</sub> (N residues 31–252) (14), and the K170A/R185A N<sup>mono</sup> mutant (24).

**Expression and Purification of Proteins**—All proteins were expressed in *E. coli* BL21(DE3). <sup>15</sup>N- and <sup>15</sup>N,<sup>13</sup>C-labeled P protein samples for NMR experiments were produced in minimum M9 medium supplemented with 1 g liter<sup>-1</sup> <sup>15</sup>NH<sub>4</sub>Cl (Eurisotop), 4 or 3 g liter<sup>-1</sup> unlabeled or <sup>13</sup>C-labeled glucose (Eurisotop), and 100 μg ml<sup>-1</sup> of ampicillin, using a protocol adapted from cultures in rich medium (30). Bacteria were disrupted (Constant Systems Ltd) in 50 mM Tris, pH 7.8, 60 mM NaCl, 1 mM EDTA, 2 mM β-mercaptoethanol, 0.2% Triton X-100 lysis buffer. After clarification by ultracentrifugation the supernatant was mixed with 2 ml of glutathione-Sepharose beads (GE Healthcare) per liter of culture and incubated for 15 h at 4 °C. The resin beads were then washed with thrombin cleavage buffer (20 mM Tris, pH 8.4, 150 mM NaCl, 2 mM β-mercaptoethanol, 2.5 mM CaCl<sub>2</sub>) before addition of 5 units of biotinylated thrombin (Novagen). The beads were incubated for 16 h at 4 °C. Thrombin was removed with streptavidin resin (Novagen) according to the manufacturer's instructions. Purification of N protein was carried out as described previously for N<sup>mono</sup> (24), N<sub>NTD</sub> (39), and N-RNA rings (13). His tags were not removed. The quality of protein samples was assessed by SDS-PAGE. Samples were subsequently dialyzed into NMR buffer (20 mM sodium phosphate, pH 6.5, 100 mM NaCl) and concentrated to 50–300 μM on Amicon Ultra centrifugal filters (10 kDa cut-off, Merck-Millipore).

**Paramagnetic Spin Labeling**—Individual residues of hRSV P, preferentially serines (Ser<sup>23</sup>, Ser<sup>99</sup>, Ser<sup>143</sup>, Ser<sup>156</sup>, Glu<sup>179</sup>, Glu<sup>193</sup>, Ser<sup>237</sup>), were mutated into cysteines using the Quik-Change mutagenesis kit (Stratagene). Mutant proteins were expressed and purified like wild type. Buffers contained 5 mM dithiothreitol (DTT) as a reducing agent. Protein samples were completely reduced by addition of another 5 mM DTT at room temperature for 2 h. DTT was then removed by passing twice through a Biospin desalting column (Bio-Rad) equilibrated in 50 mM Tris, pH 8.0, 200 mM NaCl. Protein samples were reacted overnight in the dark at 15 °C with 10 molar eq of 3-(2-iodoacetamido)-PROXYL radical (IAP, Sigma) in a 45 mM solution in ethanol. Unreacted IAP was removed by applying the samples three times onto Biospin desalting columns equilibrated in NMR buffer.

**NMR Spectroscopy**—NMR measurements were carried out on a Bruker Avance III spectrometer at a magnetic field of 14.1 T (600 MHz <sup>1</sup>H frequency) equipped with a cryogenic TCI probe. The magnetic field was locked with 7% <sup>2</sup>H<sub>2</sub>O. The temperature was 288 K if not indicated otherwise. Spectra were processed with Topspin 3.1 (Bruker Biospin) and analyzed with CCPNMR 2.2 (52).

Sequential backbone assignment of the hRSV phosphoprotein constructs was carried out with BEST versions of HNCA, HN(CO)CA, HNCACB, HN(CO)CACB, and HNCO triple resonance experiments (53). <sup>1</sup>Hα chemical shifts were obtained from HNHA or HBHA(CO)NH experiments. <sup>1</sup>H and <sup>13</sup>C chemical shifts were referenced to 4,4-dimethyl-4-silapentane-1-sulfonic acid. Assignment completeness was 95% amides, 87% <sup>13</sup>C', 98% <sup>13</sup>Cα, 60% <sup>13</sup>Cβ, and 86% <sup>1</sup>Hα for P<sub>ND</sub>; 74%

amides, 69% <sup>13</sup>C', 79% <sup>13</sup>Cα, 79% <sup>13</sup>Cβ, and 76% <sup>1</sup>Hα for P<sub>ND+OD</sub>; 98% amides, 95% <sup>13</sup>C', 98% <sup>13</sup>Cα, 99% <sup>13</sup>Cβ, and 93% <sup>1</sup>Hα for P<sub>CD</sub>; 95% amides, 93% <sup>13</sup>C', 99% <sup>13</sup>Cα, 99% <sup>13</sup>Cβ, 81% <sup>1</sup>Hα, and 78% <sup>1</sup>Hβ for P<sub>ΔOD</sub>; 61% amides, 59% <sup>13</sup>C', 64% <sup>13</sup>Cα, 64% <sup>13</sup>Cβ, 56% <sup>1</sup>Hα, and 57% <sup>1</sup>Hβ for P<sub>FL</sub> (S23C mutant).

**NMR Interaction Experiments**—NMR interaction experiments were carried out at a magnetic field of 14.1 T by acquiring <sup>1</sup>H-<sup>15</sup>N HSQC spectra of <sup>15</sup>N-labeled P constructs in a 30–100 μM concentration range in the presence of 0.25 to 10 molar eq of unlabeled N protein in the form of N<sub>NTD</sub>, N<sup>mono</sup>, or N-RNA rings. Samples were prepared by mixing concentrated protein solutions. Line broadening was analyzed by measuring the intensity ratios of amide signals between the spectra with and without N protein in CCPNMR 2.2. Dissociation constants for the P<sub>CD</sub>·N<sub>NTD</sub> complex were determined for each perturbed residue by assuming a two-site fast exchange model with a 1:1 stoichiometry and by fitting <sup>1</sup>H and <sup>15</sup>N chemical shift differences with Origin 7 (OriginLab) according to Equation 1,

$$(\delta - \delta_{\text{free}}) = \frac{1}{2}(\delta_{\text{bound}} - \delta_{\text{free}})(A + r - \sqrt{(A + r)^2 - 4r}) \quad (\text{Eq. 1})$$

where  $r = [N]_{\text{tot}}/[P_{\text{CD}}]_{\text{tot}}$ ,  $K_d = (A - 1) \times [P_{\text{CD}}]_{\text{tot}}$  and  $[P_{\text{CD}}]_{\text{tot}} = 50 \mu\text{M}$ .

Complementary experiments were performed with <sup>15</sup>N-N<sub>NTD</sub> (50 μM in 20 mM MES, pH 6.5, 250 mM NaCl buffer) at 293 K, by adding lyophilized P<sub>40</sub> with 1:1 to 12:1 ratios (solubility limit).

**<sup>15</sup>N Relaxation Measurements**—<sup>15</sup>N relaxation data were recorded at a magnetic field of 14.1 T and a temperature of 288 K. R<sub>1</sub> and R<sub>2</sub> relaxation rates were measured for 50–200 μM <sup>15</sup>N-labeled P<sub>ND</sub>, P<sub>ND+OD</sub>, P<sub>CD</sub>, and P<sub>ΔOD</sub> with a pseudo-three-dimensional version recorded in an interleaved manner with a recycling delay of 4 s. Relaxation delays were 5, 50, 100, 200(\*2), 400, 600, 800, 1200, and 2000 ms for R<sub>1</sub> measurements and 34, 68, 136, 204(\*2), 271, 339, 407, 543, and 814 ms for R<sub>2</sub>. Heteronuclear <sup>1</sup>H-<sup>15</sup>N NOEs were measured by recording two interleaved spectra with on- and off-resonance <sup>1</sup>H saturation during the recycling delay. Peak intensities were extracted in CCPNMR 2.2. Relaxation curves were fitted to a monoexponential model and errors estimated from covariance matrix analysis in CCPNMR 2.2.

**PRE Measurements**—PREs were determined as the ratios between <sup>1</sup>H-<sup>15</sup>N HSQC peak intensities in the paramagnetic and diamagnetic state ( $I_{\text{para}}/I_{\text{dia}}$ ). Measurements were carried out at 14.1 T and 288 K. The diamagnetic state was obtained by incubating the spin-labeled sample with 10 molar eq of ascorbic acid (Sigma) from a 45 mM solution at pH 6.5 for 3–4 h at 303 K.

**Modeling**—Homology modeling of OD residues Asp<sup>129</sup>-Leu<sup>152</sup> was done with Modeler 9v16 (54) using hMPV Leu<sup>170</sup>-Leu<sup>193</sup> (PDB 4BXT) as a template. P IDRs were built in CYANA 3.97 (55) by using dihedral angle restraints generated from P<sub>CD</sub> backbone chemical shifts with Talos+ (31).

**Illustrations**—Figures were prepared using PyMOL (56). Helical wheel diagrams were generated with DrawCoil 1.0.

**Author Contributions**—C. S. conceived and coordinated the study and wrote the paper. E. L., F. B., J. F. E., M. G., and N. A. participated in early experiments and provided assistance in sample preparation or NMR spectroscopy. C. C., C. S., J. F., M. G., N. P., and S. L. conducted and analyzed experiments. All authors reviewed the results and approved the final version of the manuscript.

### References

- Afonso, C. L., Amarasinghe, G. K., Banyai, K., Bao, Y., Basler, C. F., Bavari, S., Bejerman, N., Blasdel, K. R., Briand, F. X., Briese, T., Bukreyev, A., Calisher, C. H., Chandran, K., Cheng, J., Clawson, A. N., *et al.* (2016) Taxonomy of the order mononegavirales: update 2016. *Arch. Virol.* **161**, 2351–2360
- Collins, P. L., and Melero, J. A. (2011) Progress in understanding and controlling respiratory syncytial virus: still crazy after all these years. *Virus Res.* **162**, 80–99
- Nair, H., Nokes, D. J., Gessner, B. D., Dherani, M., Madhi, S. A., Singleton, R. J., O'Brien, K. L., Roca, A., Wright, P. F., Bruce, N., Chandran, A., Theodoratou, E., Sutanto, A., Sedyaningih, E. R., Ngama, M., *et al.* (2010) Global burden of acute lower respiratory infections due to respiratory syncytial virus in young children: a systematic review and meta-analysis. *Lancet* **375**, 1545–1555
- Mazur, N. I., Martín-Torres, F., Baraldi, E., Fauroux, B., Greenough, A., Heikkinen, T., Manzoni, P., Mejias, A., Nair, H., Papadopoulos, N. G., Polack, F. P., Ramilo, O., Sharland, M., Stein, R., Madhi, S. A., Bont, L., and Respiratory Syncytial Virus Network (ReSViNET) (2015) Lower respiratory tract infection caused by respiratory syncytial virus: current management and new therapeutics. *Lancet Respir. Med.* **3**, 888–900
- Collins, P. L., and Karron, R. A. (2013) Respiratory Syncytial Virus and Metapneumovirus. in *Fields Virology* (Knipe, D. M., and Howley, P. M., eds) 6th Ed., pp. 1086–1123, Lippincott Williams & Wilkins, Wolters Kluwer, Philadelphia, PA
- Cox, R., and Plemper, R. K. (2016) Structure-guided design of small-molecule therapeutics against RSV disease. *Expert Opin. Drug Discov.* **11**, 543–556
- Cox, R., and Plemper, R. K. (2015) The paramyxovirus polymerase complex as a target for next-generation anti-paramyxovirus therapeutics. *Front. Microbiol.* **6**, 459
- García, J., García-Barreno, B., Vivo, A., and Melero, J. A. (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**, 243–247
- Lahaye, X., Vidy, A., Pomier, C., Obiang, L., Harper, F., Gaudin, Y., and Blondel, D. (2009) Functional characterization of Negri bodies (NBs) in rabies virus-infected cells: evidence that NBs are sites of viral transcription and replication. *J. Virol.* **83**, 7948–7958
- Heinrich, B. S., Cureton, D. K., Rahmeh, A. A., and Whelan, S. P. (2010) Protein expression redirects vesicular stomatitis virus RNA synthesis to cytoplasmic inclusions. *PLoS Pathog.* **6**, e1000958
- Grosfeld, H., Hill, M. G., and Collins, P. L. (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. *J. Virol.* **69**, 5677–5686
- Yu, Q., Hardy, R. W., and Wertz, G. W. (1995) Functional cDNA clones of the human respiratory syncytial (RS) virus N, P, and L proteins support replication of RS virus genomic RNA analogs and define minimal transacting requirements for RNA replication. *J. Virol.* **69**, 2412–2419
- Tran, T. L., Castagné, N., Bhella, D., Varela, P. F., Bernard, J., Chilmonczyk, S., Berkenkamp, S., Benhamo, V., Grznarova, K., Grosclaude, J., Nespoulos, C., Rey, F. A., and Eléouët, J. F. (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. *J. Gen. Virol.* **88**, 196–206
- Galloux, M., Tarus, B., Blazevic, I., Fix, J., Duquerroy, S., and Eléouët, J. F. (2012) Characterization of a viral phosphoprotein binding site on the surface of the respiratory syncytial nucleoprotein. *J. Virol.* **86**, 8375–8387
- García-Barreno, B., Delgado, T., and Melero, J. A. (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. *J. Virol.* **70**, 801–808
- Slack, M. S., and Easton, A. J. (1998) Characterization of the interaction of the human respiratory syncytial virus phosphoprotein and nucleocapsid protein using the two-hybrid system. *Virus Res.* **55**, 167–176
- Mason, S. W., Aberg, E., Lawetz, C., DeLong, R., Whitehead, P., and Liuzzi, M. (2003) Interaction between human respiratory syncytial virus (RSV) M2-1 and P proteins is required for reconstitution of M2-1-dependent RSV minigenome activity. *J. Virol.* **77**, 10670–10676
- Blondot, M. L., Dubosclard, V., Fix, J., Lassoued, S., Aumont-Nicaise, M., Bontems, F., Eléouët, J. F., and Sizun, C. (2012) Structure and functional analysis of the RNA- and viral phosphoprotein-binding domain of respiratory syncytial virus M2-1 protein. *PLoS Pathog.* **8**, e1002734
- Tran, T. L., Castagné, N., Dubosclard, V., Noinville, S., Koch, E., Moudjou, M., Henry, C., Bernard, J., Yeo, R. P., and Eléouët, J. F. (2009) The respiratory syncytial virus M2-1 protein forms tetramers and interacts with RNA and P in a competitive manner. *J. Virol.* **83**, 6363–6374
- Villanueva, N., Hardy, R., Asenjo, A., Yu, Q., and Wertz, G. (2000) The bulk of the phosphorylation of human respiratory syncytial virus phosphoprotein is not essential but modulates viral RNA transcription and replication. *J. Gen. Virol.* **81**, 129–133
- Asenjo, A., Calvo, E., and Villanueva, N. (2006) Phosphorylation of human respiratory syncytial virus P protein at threonine 108 controls its interaction with the M2-1 protein in the viral RNA polymerase complex. *J. Gen. Virol.* **87**, 3637–3642
- Asenjo, A., and Villanueva, N. (2016) Phosphorylation of the human respiratory syncytial virus P protein mediates M2-2 regulation of viral RNA synthesis, a process that involves two P proteins. *Virus Res.* **211**, 117–125
- Curran, J., Marq, J. B., and Kolakofsky, D. (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. *J. Virol.* **69**, 849–855
- Galloux, M., Gabiane, G., Sourimant, J., Richard, C. A., England, P., Moudjou, M., Aumont-Nicaise, M., Fix, J., Rameix-Welti, M. A., and Eléouët, J. F. (2015) Identification and characterization of the binding site of the respiratory syncytial virus phosphoprotein to RNA-free nucleoprotein. *J. Virol.* **89**, 3484–3496
- Castagné, N., Barbier, A., Bernard, J., Rezaei, H., Huet, J. C., Henry, C., Da Costa, B., and Eléouët, J. F. (2004) Biochemical characterization of the respiratory syncytial virus P-P and P-N protein complexes and localization of the P protein oligomerization domain. *J. Gen. Virol.* **85**, 1643–1653
- Simabuco, F. M., Asara, J. M., Guerrero, M. C., Libermann, T. A., Zerbini, L. F., and Ventura, A. M. (2011) Structural analysis of human respiratory syncytial virus p protein: identification of intrinsically disordered domains. *Braz. J. Microbiol.* **42**, 340–345
- Noval, M. G., Esperante, S. A., Molina, I. G., Chemes, L. B., and Prat-Gay, G. (2016) Intrinsic disorder to order transitions in the scaffold phosphoprotein P from the respiratory syncytial virus RNA polymerase complex. *Biochemistry* **55**, 1441–1454
- Llorente, M. T., Taylor, I. A., López-Viñas, E., Gomez-Puertas, P., Calder, L. J., García-Barreno, B., and Melero, J. A. (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**, 946–958
- Kosol, S., Contreras-Martos, S., Cedeño, C., and Tompa, P. (2013) Structural characterization of intrinsically disordered proteins by NMR spectroscopy. *Molecules* **18**, 10802–10828
- Sourimant, J., Rameix-Welti, M. A., Gaillard, A. L., Chevret, D., Galloux, M., Gault, E., and Eléouët, J. F. (2015) Fine mapping and characterization of the L-polymerase-binding domain of the respiratory syncytial virus phosphoprotein. *J. Virol.* **89**, 4421–4433

31. Shen, Y., Delaglio, F., Cornilescu, G., and Bax, A. (2009) TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR* **44**, 213–223
32. Camilloni, C., De Simone, A., Vranken, W. F., and Vendruscolo, M. (2012) Determination of secondary structure populations in disordered states of proteins using nuclear magnetic resonance chemical shifts. *Biochemistry* **51**, 2224–2231
33. Clore, G. M., and Iwahara, J. (2009) Theory, practice, and applications of paramagnetic relaxation enhancement for the characterization of transient low-population states of biological macromolecules and their complexes. *Chem. Rev.* **109**, 4108–4139
34. Lietzow, M. A., Jamin, M., Dyson, H. J., and Wright, P. E. (2002) Mapping long-range contacts in a highly unfolded protein. *J. Mol. Biol.* **322**, 655–662
35. Gillespie, J. R., and Shortle, D. (1997) Characterization of long-range structure in the denatured state of staphylococcal nuclease: I. paramagnetic relaxation enhancement by nitroxide spin labels. *J. Mol. Biol.* **268**, 158–169
36. Wright, P. E., and Dyson, H. J. (2009) Linking folding and binding. *Curr. Opin. Struct. Biol.* **19**, 31–38
37. Mohan, A., Oldfield, C. J., Radivojac, P., Vacic, V., Cortese, M. S., Dunker, A. K., and Uversky, V. N. (2006) Analysis of molecular recognition features (MoRFs). *J. Mol. Biol.* **362**, 1043–1059
38. Bakker, S. E., Duquerroy, S., Galloux, M., Loney, C., Conner, E., Eléouët, J. F., Rey, F. A., and Bhella, D. (2013) The respiratory syncytial virus nucleoprotein-RNA complex forms a left-handed helical nucleocapsid. *J. Gen. Virol.* **94**, 1734–1738
39. Ouizougoun-Oubari, M., Pereira, N., Tarus, B., Galloux, M., Lassoued, S., Fix, J., Tortorici, M. A., Hoos, S., Baron, B., England, P., Desmaele, D., Couvreur, P., Bontems, F., Rey, F. A., *et al.* (2015) A druggable pocket at the nucleocapsid/phosphoprotein interaction site of human respiratory syncytial virus. *J. Virol.* **89**, 11129–11143
40. Renner, M., Bertinelli, M., Leyrat, C., Paesen, G. C., Saraiva de Oliveira, L. F., Huiskonen, J. T., and Grimes, J. M. (2016) Nucleocapsid assembly in pneumoviruses is regulated by conformational switching of the N protein. *Elife* **5**, e12627
41. Llorente, M. T., García-Barreno, B., Calero, M., Camafeita, E., López, J. A., Longhi, S., Ferrón, F., Varela, P. F., and Melero, J. A. (2006) Structural analysis of the human respiratory syncytial virus phosphoprotein: characterization of an  $\alpha$ -helical domain involved in oligomerization. *J. Gen. Virol.* **87**, 159–169
42. Asenjo, A., Mendieta, J., Gómez-Puertas, P., and Villanueva, N. (2008) Residues in human respiratory syncytial virus P protein that are essential for its activity on RNA viral synthesis. *Virus Res.* **132**, 160–173
43. Ruigrok, R. W., Crépin, T., and Kolakofsky, D. (2011) Nucleoproteins and nucleocapsids of negative-strand RNA viruses. *Curr. Opin. Microbiol.* **14**, 504–510
44. Karlin, D., and Belshaw, R. (2012) Detecting remote sequence homology in disordered proteins: discovery of conserved motifs in the N-termini of Mononegavirales phosphoproteins. *PLoS ONE* **7**, e31719
45. Leyrat, C., Renner, M., Harlos, K., and Grimes, J. M. (2013) Solution and crystallographic structures of the central region of the phosphoprotein from human metapneumovirus. *PLoS ONE* **8**, e80371
46. Cox, R., Green, T. J., Purushotham, S., Deivanayagam, C., Bedwell, G. J., Prevelige, P. E., and Luo, M. (2013) Structural and functional characterization of the mumps virus phosphoprotein. *J. Virol.* **87**, 7558–7568
47. Cox, R., Pickar, A., Qiu, S., Tsao, J., Rodenburg, C., Dokland, T., Elson, A., He, B., and Luo, M. (2014) Structural studies on the authentic mumps virus nucleocapsid showing uncoiling by the phosphoprotein. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 15208–15213
48. Munday, D. C., Wu, W., Smith, N., Fix, J., Noton, S. L., Galloux, M., Touzelet, O., Armstrong, S. D., Dawson, J. M., Aljabr, W., Easton, A. J., Rameix-Welti, M. A., de Oliveira, A. P., Simabuco, F. M., Ventura, A. M., *et al.* (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. *J. Virol.* **89**, 917–930
49. Khattar, S. K., Yunus, A. S., and Samal, S. K. (2001) Mapping the domains on the phosphoprotein of bovine respiratory syncytial virus required for N-P and P-L interactions using a minigenome system. *J. Gen. Virol.* **82**, 775–779
50. Lu, B., Brazas, R., Ma, C. H., Kristoff, T., Cheng, X., and Jin, H. (2002) Identification of temperature-sensitive mutations in the phosphoprotein of respiratory syncytial virus that are likely involved in its interaction with the nucleoprotein. *J. Virol.* **76**, 2871–2880
51. Shapiro, A. B., Gao, N., O'Connell, N., Hu, J., Thresher, J., Gu, R. F., Overman, R., Hardern, I. M., and Sproat, G. G. (2014) Quantitative investigation of the affinity of human respiratory syncytial virus phosphoprotein C-terminus binding to nucleocapsid protein. *Virol. J.* **11**, 191
52. Vranken, W. F., Boucher, W., Stevens, T. J., Fogh, R. H., Pajon, A., Llinas, M., Ulrich, E. L., Markley, J. L., Ionides, J., and Laue, E. D. (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687–696
53. Lescop, E., Schanda, P., and Brutscher, B. (2007) A set of BEST triple-resonance experiments for time-optimized protein resonance assignment. *J. Magn. Reson.* **187**, 163–169
54. Sali, A., and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815
55. Guntert, P. (2004) Automated NMR structure calculation with CYANA. *Methods Mol. Biol.* **278**, 353–378
56. Schrödinger, L. (2010) *The AxPyMOL Molecular Graphics Plugin for Microsoft PowerPoint*, version 1.0, Schrödinger, LLC, New York
57. Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D., and Higgins, D. G. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539
58. Gouet, P., Courcelle, E., Stuart, D. I., and Metzoz, F. (1999) ESPript, analysis of multiple sequence alignments in PostScript. *Bioinformatics* **15**, 305–308