UNIVERSITÉ DE TOURS

***ÉCOLE DOCTORALE* Santé, Sciences Biologiques et Chimie du Vivant**

EQUIPE Infectiologie et Immunité Innée des monogastriques (3Imo)

UNITE Infectiologie et Santé Publique

THÈSE présentée par : Alan WACQUIEZ

Soutenue le : **19 Octobre 2020**

Pour obtenir le grade de : **Docteur de l'université de Tours**

Discipline / Spécialité : Sciences de la vie / Virologie

Caractérisation moléculaire du domaine de liaison à l'ARN de la protéine non-structurale 1 (NS1) des virus *influenza* A lors de l'interaction avec un ARN double-brin de haute affinité contenant des séquences sélectionnées *in vitro*

THÈSE dirigée par :

M MARC Daniel	**Chargé de recherche**, HDR, INRAE, Nouzilly
M CASTAING Bertrand	**Directeur de recherche,** CBM, CNRS Orléans

RAPPORTEURS :

Mme ISEL-GRIFFITHS Catherine	**Chargée de recherche**, HDR, Institut Pasteur, Paris
M CREPIN Thibaut	**Directeur de recherche**, IBS, Grenoble

PRÉSIDENT DE JURY :

M TOUZE Antoine	**Professeur**, Université de Tours

JURY :

Mme DUCATEZ Mariette	**Directrice de recherche**, ENVT, Toulouse
Mme ISEL-GRIFFITHS Catherine	**Chargée de recherche**, HDR, Institut Pasteur, Paris
M CHEVALIER Christophe	**Chargé de recherche**, HDR, INRAE, Jouy-en-Josas
M CREPIN Thibaut	**Directeur de recherche**, IBS, Grenoble
M TOUZE Antoine	**Professeur**, Université de Tours
M MARC Daniel	**Chargé de recherche**, HDR, INRAE, Nouzilly
M CASTAING Bertrand	**Directeur de recherche,** CBM, CNRS Orléans
M COSTE Franck	**Ingénieur de recherche,** CBM, CNRS Orléans

À ma mère, pour ton soutien éternel...

REMERCIEMENTS

En tout premier lieu je remercie la région *Centre Val de Loire* pour avoir financé ma thèse. Je souhaite remercier l'*INRAE* de Nouzilly et le *Centre de Biophysique Moléculaire* d'Orléans, ainsi que tous les membres (techniques et administratifs) intervenant ces deux centres de recherche, pour leur accueil au cours mon doctorat au cours duquel de nombreuses expériences ont pu être réalisées.

Je tiens à remercier tout particulièrement Catherine ISEL-GRIFFITHS et Thibaut CREPIN pour avoir respectivement accepté d'être rapportrice et rapporteur de mon manuscrit de thèse. Je remercie tout autant Mariette DUCATEZ, Christophe CHEVALIER et Antoine TOUZÉ pour avoir accepté d'analyser mon travail de doctorat en tant que membres du jury. Je souhaite également remercier Laurence SERRE, Gwenaëlle ANDRE-LEROUX et Thibaut CREPIN pour avoir suivi continuellement le déroulement de ce travail de thèse.

Daniel, je souhaite te remercier pour la confiance accordée en réalisant cette thèse. Merci pour tes conseils et l'expérience que tu as pu m'apporter depuis notre première rencontre lors de mon stage de première année de Master. Ton implication dans mon projet de doctorat m'aura permis d'atteindre les objectifs fixés initialement. Un grand merci pour ton implication soutenue dans la rédaction et la correction des manuscrits de thèse et d'article réalisés au cours de ce travail. Merci aussi pour m'avoir permis de présenter mes travaux à plusieurs reprises sous forme de posters au *STUDIUM* d'Orléans (au cours duquel j'aurais eu l'opportunité de rencontrer Kiyoshi NAGAI et Chantal ABERGEL, spécialistes de renom sur le spliceosome et les virus géants, respectivement) et pendant les *Journées Francophones de Virologie* et pour la communication orale qui s'est tenue lors de la première réunion annuelle du réseau *RESAFLU* (au cours de laquelle j'aurais pu rencontrer et échanger avec de nombreux virologistes dont l'expertise porte sur les virus *influenza*, en particulier Nadia NAFFAKH qui a apporté son expertise dans mon projet de doctorat, Manuel ROSA-CALATRAVA, Olivier TERRIER, Bruno LINA, Caroline GOUJON et Olivier MONCORGE parmi lesquels des pistes professionnelles ont vu le jour). Les échanges qui en suivirent, en particulier avec Thibaut CREPIN, Rob RUIGROK, Ronan LE GOFFIC, Roland MARQUET, Christophe CHEVALIER, furent d'une grande inspiration intellectuelle pour moi, avec une remise en question plus conséquente autant sur le projet de thèse que sur l'avenir scientifique que j'envisage aujourd'hui. Merci aussi à l'unité de recherche *Infectiologie et Santé Publique* pour m'avoir permis de présenter l'avancée de mon projet de recherche chaque année.

Ce projet de recherche ayant été réalisé dans le cadre d'une collaboration avec le *Centre de Biophysique Moléculaire*, je souhaite remercier l'entièreté de l'équipe de recherche de Bertrand CASTAING pour son accueil. Bertrand, et je ne te le dirais jamais assez, un grand merci à toi pour ton implication autant personnelle que financière mais aussi et surtout pour tes conseils avisés sur le déroulement de mon doctorat. Tu as su écouter mes problèmes personnels lorsque j'étais dans le besoin. Merci pour tous ces moments de bonheur passés à vos côtés pendant ce doctorat. Un grand merci aussi

pour ton implication dans la réalisation des manuscrits de thèse et d'article réalisés au cours de mon doctorat. Je souhaite aussi te remercier pour m'avoir permis de participer aux sessions anuelles du réseau *BIOTECHNOCENTRE* au cours desquels j'ai pu échangé avec une multitude de scientifiques. Ces sessions m'auront également permis de rencontrer une multitude de jeunes chercheurs, collègues ou non à ce jour, de Tours et d'Orléans avec qui pour la plupart j'ai pu garder contact. Enfin, merci pour tous ces rires et ces bons moments passés parmi vous.

Franck, tu as joué un rôle crucial dans la réalisation de ce travail de recherche. Ainsi, je souhaite te remercier pour l'expérience et la « patience » dont tu as fait preuve avec moi. Arrivé sans grandes connaissances dans le domaine des protéines recombinantes et de la cristallographie, j'en sors grandi plus confirmé sur ces deux plans. J'ai adoré nos échanges sur le plan personnel même s'il est difficile de te faire parler de toi. Ayant peu d'intérêts (car peu de connaissances) sur les aspects structuraux des protéines et des interactions avec des acides nucléiques au démarrage de mon doctorat, j'ai acquis un goût prononcé pour ces domaines que je poursuis actuellement en post doctorat chez Mustapha SI-TAHAR au *Centre d'Etude des Pathologies Respiratoires* à Tours. Un grand merci à toi aussi pour ton encadrement, ton implication scientifique et ton expertise en tant que structuraliste dans la réalisation des expériences menées au *CBM* mais aussi dans la correction des manuscrits de thèse et d'article réalisés pendant mon doctorat. Un point négatif subsiste tout de même, je regrette que tu ne m'aies jamais laissé gagner au ping-pong le mercredi soir.

Virginie, un énorme merci à toi sans qui le projet de thèse aurait eu beaucoup plus de mal à atteindre cette finalité. Partagé sur les sites de Tours et d'Orléans, il mettait impossible physiquement parlant d'avancer sur les thématiques développés à Tours et à Orléans de façon simultanée. Un grand merci à Bertrand, à Franck et surtout à toi pour avoir poursuivi les aspects d'interaction protéine/ARN lorsque j'étais de retour à l'*INRAE* de Nouzilly. Tu as su me mettre à l'aise dès mon arrivée et j'ai énormément appris grâce à toi. Sur le plan professionnel, tu m'as transmis ta rigueur et ton savoir-faire. Sur le plan personnel, j'ai pu découvrir une belle personne toujours plein d'entrain. Merci aussi d'avoir pris du temps pour écouter les difficultés que j'ai rencontré en parallèle de mon doctorat. À Chaque instant tu étais disponible pour m'aider et me conseiller tant sur le plan technique que philosophique.

Manu, tu mérites toi aussi mes remerciements et ma plus grande sympathie. Ainsi, un grand merci à toi pour avoir participé (tout autant que Virginie a pu le faire sur Orléans) à la réalisation des travaux tourangeaux que j'ai engagé au cours de mon doctorat. Merci pour tous ces midis passés en ta compagnie, j'avais besoin de décompresser ! Merci de m'avoir transmis autant de choses sur le plan technique, je suis convaincu que ça me sera essentiel pour la suite de mon périple scientifique. J'avouerais qu'il est d'autant plus facile de parler avec toi puisqu'on partage de nombreuses idées. Un grand merci pour ta contribution technique, tu es le maitre des *western-blots* et je ne suis pas le seul à le dire !

Sascha, je te remercie pour ton suivi durant ces quatre années passées à vos côtés ainsi que pour tes suggestions lors de nos échanges qui m'auront permis d'entrevoir de nouvelles perspectives

professionnelles notamment en contactant des scientifiques de renom au *Centre de Recherche en virologie* d'Edinburgh. Merci aussi à Nathalie WINTER et à toi pour votre implication dans le déroulement de cette thèse. J'en profite pour remercier l'ensemble des membres des équipes qui m'ont accueillies au cours de cette thèse. Pour l'équipe *Infectiologie et Immunité Innée des monogastriques* (Tours), je remercie Rodrigo, Vanaique, Pascale, Evelyne (anciennement dans l'équipe), Ignacio, Isabelle, Claire, Sandrine, Michel, Raquel, Daniel et Sascha. Pour l'équipe *Réparations de l'ADN : Structure, Dynamique et Fonction* (Orléans), je remercie Stéphane (et pour tes échanges et les conseils dont tu as fait preuve pendant mon passage chez vous), Martine, Norbert (et pour ta participation scientifique en modélisation moléculaire sur le projet de recherche même si cette dernière n'a pas donné suite), Virginie, Julien, Franck et Bertrand. Je vous remercie tous aussi pour les nombreux pots d'anniversaire, les repas du midi, de pré-vacances et de Noël et pour les sorties organisées.

Ayant eu des difficultés par le passé à différencier les « copains » des vrais « amis », je pourrais dire aujourd'hui que j'ai pu me faire des vrais amis parmi les doctorant(e)s rencontré(e)s au cours de ce doctorat. Aurélien, Justine et Julien vous êtes tous trois concernés par ces remerciements. À Aurélien je dirais ceci : merci pour ton amitié infaillible, merci pour toutes ses soirées où nous avons « refait le monde », merci pour ta bonne humeur, merci pour les conneries que tu as pu faire et qui sont, même si constructives, toujours affichées sur ta paillasse, je parle ici des multiples *PCR* et *western-blots* que tu as raté. Merci pour ton soutien dans les moments difficiles que j'ai vécu tout au long de notre rencontre. Tu rayonnes de bonheur, tout le monde le voit et le ressent, au laboratoire comme en dehors. Je suis persuadé qu'une longue et belle amitié durable dans le temps est née. À Justine et Julien, *alias* les « jujus » sans distinction, je vous dis ceci : Un énorme merci à vous pour m'avoir accueilli parmi vous, autant chez vous car vous m'avez ouvert votre porte lorsque j'étais dans le besoin qu'au laboratoire, merci pour toutes ses soirées géniales passées à vos côtés et ces midis de pure déconnade. Tout comme je l'ai dit pour Aurélien, vous êtes à ce jour des amis et je suis persuadé que nous poursuivrons cette histoire dans le temps, et peut-être qui sait, en se suivant les uns les autres.

Un grand merci aussi aux différents membres de l'équipe *Biologie des Virus Aviaires*, en particulier KATIA et Sylvie pour vos écoutes, vos encouragements et bien évidemment pour vos rires tous ces midis. À Sylvie, je dirais ceci : s'imagine qu'il te sera maintenant difficile de regarder *Star Wars* sans faire le rapprochement avec la photo située au-dessus de ta paillasse de travail. À Katia, merci pour ton entrain et ta déconnade qui sont pour moi (comme pour d'autres j'en suis persuadé) infinis. Je remercie bien évidemment Caroline pour proposer de contribuer scientifiquement à leur dernière publication. Je remercie aussi David (enfin je ne suis pas certain qu'il faille te remercier) des nombreux échanges autant sur le plan personnel que professionnel. Tes remarques « pertinentes » sonnent le *glas* en moi.

Laetitia, je me dois de te dédier un paragraphe de remerciements élogieux tant sur ton implication dans mon projet que sur nos discussions, qui parfois étaient dénuées de sens reconnaissons-le. Tous les jeunes scientifiques rêveraient de réaliser leur thèse sous ta supervision. Excellente tant sur

le plan scientifique que sur plan amicale, il est très facile de se confier à toi. J'ai beaucoup appris à tes côtés, notamment sur la remise en question dont je devais faire preuve et j'ai d'ailleurs pu apprendre à gérer les choses avec une meilleure « diplomatie ». Merci pour ces bons moments passés avec toi, la session apéro devant la coupe du monde retransmis en directe à l'*INRAE*, l'après-midi « bataille d'eau », les repas du midi où nous parlions de tout sauf de sciences, et j'en passe ! Tu es géniale, ne perds pas cette qualité qui te vas si bien.

À Agnès WIEDMANN, je dirais ceci : un énorme merci pour ta franchise lors de notre dernier échange en juillet 2019. Sans prétention, j'ai compris que j'avais énormément évolué sur le plan scientifique depuis mon Master effectué à l'université de Tours. Tes remarques constructives sur ma condition tant d'individu que de jeune scientifique m'ont fait prendre conscience du potentiel que j'avais en moi.

J'ai rencontré bon nombre de doctorant(e)s au cours de ce périple entre Tours et Orléans. Il m'est important de les citer pour leur apport sur le plan décompression et festivité. À Tours, je remercierais activement Aurélien (et ses colocataires), Damien (PhD à ce jour), Julien, Ambre, Tiffany, Adeline, Corentin, Nicolas, Elise ainsi que les stagiaires que j'ai pu encadré pendant ma thèse, en particulier Adeline CEZARD avec qui travaille à ce jour. Du côté d'Orléans, je remercierais particulièrement Justine, Julien, Mohammed, Florian, Skander, Lauren, Saida et Anton.

Comme bon nombre d'autres doctorants arrivant à leur fin de leur cursus universitaire, la famille, *via* son soutien, joue un rôle essentiel dans le maintien tant physique que moral d'un jeune scientifique. Ainsi, je souhaite remercier toute ma famille, en particulier Sylvain TILLIER (et sa famille) et Prescillia DA SILVA pour leur soutien depuis ma plus tendre enfance. Le défi que je me suis lancé n'était pas gagné d'avance et je vous remercie sincèrement pour votre soutien indéfectible. À ma famille, je dirais ceci : merci pour tous ces bons moments conviviaux, pour tous ces partages d'idées qui m'ont fait grandir et évoluer sur les plans intellectuels et humains. Je souhaite remercier la famille PLAUD, GIRAUDET et VIAU pour leur soutien dans cette épreuve psychologique qui fût, parfois, difficile et très souvent incomprise. Quatre membres de ma famille ont joué un rôle crucial dans le bon acheminement de cette aventure. À ma cousine Alizée LAMBERT, qui tout comme moi revient de très loin, tu as été un soutien moral permanent; À mon oncle Laurent PLAUD, issu tout comme moi d'un quartier nantais difficile sans grande ambition professionnelle, tu m'as motivé à donner le meilleur de moi-même en toute circonstance. À ma petite sœur Inès, merci pour tous ces bons moments passés ensemble et pour ton soutien permanent. Alors que la jalousie de perdre ma place de fils obscurcissait aveuglément mon objectivité à ta naissance, je ne peux plus me passer de toi à ce jour. À ma mère Cécile, je dirais ceci : Ton soutien éternel dans les difficultés comme dans les bons moments a été décisif dans le choix de mon parcours professionnel et dans son aboutissement (même si ce n'est pas encore fini car des projets j'en ai plein la tête). Sans toi, sans ton indéfectible tendresse et sans ton soutien éternel, je ne serais pas la personne que je suis devenue aujourd'hui. MERCI de m'avoir protégé des « agressions » qu'aurait pu m'occasionné le monde extérieur. Merci de m'avoir soutenu tout au long de

ce périple difficile, merci d'avoir contredit ouvertement toutes les personnes environnantes qui ne croyaient ni en moi ni en mon avenir professionnel. Merci de toujours avoir été là pour moi prouvant ainsi que tu es la meilleure maman du monde.

Un regret personnel subsiste cependant au travers de cette thèse si pluridisciplinaire. J'aurais aimé avoir la possibilité d'effectuer quelques heures d'enseignement afin de savoir si l'intérêt « pédagogique » que j'ai pu développer au travers des discussions vulgarisées autour de mon projet de recherche et de la biologie en général avec mon entourage non-sensibilisé pouvait avoir un sens et un avenir par la suite. En dépit de ça, je souhaiterais remercier Nathalie WINTER pour m'avoir fait comprendre le concept *One Health* car, comme la situation pandémique survenue pendant la fin de mon doctorat le rappelle activement, il est essentiel de poursuivre les programmes de recherche sur les virus zoonotiques comme les virus *influenza*, *Ebola* ou les *coronavirus*, en tentant d'identifier de nouveaux virus environnants potentiellement pathogènes pour l'Homme et en tentant de comprendre les mécanismes moléculaires impliqués autant dans le cycle viral que sur le plan immunitaire. Ce travail de doctorat m'aura permis de valider mon choix de carrière professionnelle dans la recherche scientifique en biologie mais aussi et surtout d'attester mon goût déjà prononcé pour la virologie en général, avec un focus sur les virus à ARN. Passionné par la virologie animale et humaine et par le travail de rigueur en laboratoire de haute sécurité biologique (laboratoires P3 et P4, j'en profite pour remercier sincèrement Pierre GERMON pour m'avoir partagé son réseau professionnel américain, j'aurai ainsi pu échanger professionnellement avec Brandi LIMBAGO du Centre de Contrôle et de Prévention des Maladies (CDC) basé à Atlanta, USA), j'envisage par la suite de travailler sur les virus de chauve-souris dont le pouvoir pathogène pourrait présenter des risques sanitaires pour l'Homme. Sans aucune prétention, j'envisage également la construction d'une équipe de recherche qui s'axerait sur l'identification de nouvelles solutions thérapeutiques provenant de l'étude moléculaire du système immunitaire des chauve-souris qui hébergent, sereinement, de nombreux virus à risque. J'en profite pour remercier chaleureusement François MEURENS pour ses nombreux conseils et ces informations utiles pour démarrer un tel projet d'avenir. Je remercie tout autant Philippe ROINGEARD et Julien BURLAUD-GAILLARD pour leurs nombreuses discussions en virologie occasionnées lors mon stage de seconde année de Master effectué dans l'équipe tourangelle *Morphogénèse et Antigénicité du VIH et des Virus des Hépatites*.

LISTE DES ABRÉVIATIONS

Å : Ångström

aaRBD-A : RBD de NS1 d'allèle A (H5N1) contenant la double-substitution R38A-K41A

aaRBD-B : RBD de NS1 d'allèle B (H7N1) contenant la double-substitution R38A-K41A

ACADSB : Acyl-CoA Deshydrogenase Short/Branched chain (facteur chaine courte acyl-CoA deshydrogenase)

Ad : Adénine

ADAR : Adenosine Deaminase that act on RNA (adénosine désaminase agissant sur l'ARN ou riboadénosine désaminase)

ADN : Acide désoxyribonucléique

Akt (ou PKB): Protéine Kinase B

AMP^R : Résistance à l'ampicilline

ANSM : Agence Nationale de la Sécurité du Médicament

ARCN1 : Archaine 1

ARN : Acide ribonucléique

ARNc : ARN viraux complémentaires de polarité positive

ARNdb : ARN double-brin

ARNm : ARN messagers

ARNsb : ARN simple-brin

ARNt : ARN de transfert

ARNv : ARN viraux génomiques de polarité négative

ATP : Adénosine triphosphate

ATP5O : ATP synthase subunit O, mitochondrial (sous-unité O de l'ATP synthase)

ATP6V0D1 : ATPase H+ transporting V0 subunit D1 (sous-unité D1 de l'ATPase V0)

A549 : Adenocarcinomic human alveolar basal epithelial cells (cellules épithéliales d'adénocarcinome pulmonaire humain)

BamHI : Enzyme de restriction provenant de *Bacillus amyloliquefaciens*

BET : Bromure d'éthidium

BnAbs : Broadly neutralizing Antibodies (anticorps neutralisants à large spectre)

BSA : Bovine Serum Albumin (albumin de serum bovin)

BUB3 : Mitotic checkpoint protein (protéine de point de contrôle mitotique 3)

BXA : Baloxir Acid (acide baloxavir)

BXM : Baloxavir Marboxil (marboxil baloxavir)

C : Cystéine

$CaCl_2$: Calcium Chloride (chlorure de calcium)

CAMK2B : Calciul/calmodulin-dependent protein Kinase 2B (protéine kinase β-II dépendant de la calmoduline)

CAM^R : Résistance au Chloramphénicol

CAPRIN1 : Cell cycle Associated Protein 1 (caprine 1)

CD : Circular Dichroism (dichroïsme circulaire)

CDC : Center for Diseases Control and Prevention (Centre de Contrôle et de Prévention des Maladies)

CDC73 : Parafibromin (parafibromine)

CHD1 : Chromodomain-helicase-DNA-Binding protein 1 (protéine 1 liant l'ADN de la chromodomaine hélicase)

CHES : N-Cyclohexyl-2-aminoethanesulfonic acid (acide N-cyclohexyl-2-aminoéthanesulfonique)

CIC : Protein capicua homolog (protéine Capicua)

CMV : Cytomegalovirus

COPG : Coatomer subunit gamma (sous-unité gamma du facteur coatomer)

COX5A : Cytochrome C oxidase 5A (sous-unité 5A de la cytochrome C oxydase)

COX6A1 : Cytochrome C oxidase subunit 6A1 (sous-unité 6A1 de la cytochrome C oxydase)

CO_2 : Carbon Dioxide (dioxyde de carbone)

Co-IP : Co-immunoprécipitation

CPSF30 : 30 kDa Cleavage and Polyadenylation Specificity Factor (facteur spécifique de clivage et de polyadénylation de 30 kDa)

Crk(L) : Proto-Oncogene c-Crk (protéine adaptatrice crk)

CRM1 : Chromosome Region Maintenance 1 (exportine 1 ou XPO)

Cryo-EM : Cryo-Electro Microscopy (cryo-microscopie électronique)

cv : Column volume (volume de colonne à ajouter)

Cy : Cytosine

C2CD4C : C2 Calcium-Dependent Domain-containing protein 4C (facteur C2 calcium-dépendant contenant le domaine 4C)

D : Acide aspartique

DEAD : Diethylaminoethyl cellulose (cellulose diéthylaminoéthyle)

DLS : Diffusion light scattering (diffusion dynamique de la lumière)

DL1 : Stromal cells expressing Notch ligand Delta-Like 1 (cellules stromales exprimant le ligand delta-like 1 de Notch)

DMEM : Dulbecco's Modified Eagle's Medium (milieu utilisé pour la culture cellulaire et qui est enrichi en glucose et en acides aminés)

dNTP : Désoxyribonucléotide triphosphate quelconque

DPBA : 2,4-dioxo-4-phenylbutanoic acid (acide 2,4-dioxo-4-phenylbutanoique)

DRX : Diffraction des Rayons X

dsRBM : Double-stranded RNA-Binding Motif (Motif de liaison aux ARNdb)

dsRBP : Double-stranded RNA binding protein (protéines se liant à des ARN double-brin)

DTT : Dithiothreitol

E : Acide glutamique

ECB : Extrait cellulaire brut

EC_{50} : Half maximal Effictive Concentration (concentration efficace médiane)

EC_{50app} : EC_{50} apparent

ED : Effector Domain (domaine effecteur)

EDTA : Ethylenediaminetetraacetate (acide éthylènediaminetétraacétique)

EGCG : Epigallocatéchine Gallate

EIF4A2 : Eucaryotic initiation factor 4A2 (facteur eucaryotique 4A2 initiant la traduction)

EMEM : Eagle's Minimal Essential Medium (milieu utilisé pour la culture des cellules peu éxigentes)

EMSA : Electrophoretic Mobility Shift Assay (électrophorèse en retard sur gel)

F : Phénylalanine

FDA : Food and Drug Administration (administration américaine des denrées alimentaires et des médicaments)

FF-Luc : FireFly Luciferase (luciférase FireFly)

FGFR : Fibroblast Growth Factor Receptor (récepteur du facteur de croissance fibroblastique)

FGF-4 : Fibroblast Growth Factor 4 (facteur 4 de croissance des fibroblastes)

FluA-20 : Anticorps à large spectre ciblant la protéine HA de nombreux virus *influenza*

G : Glycine

GAPDH : Glycéraldéhyde-3-phosphate déshydrogénase

GBF1 : Golgi-specific brefeldin A-resistance guanine nucleotide exchange factor 1 (facteur d'échange de guanines résistant à la brefeldine A)

GDF11 : Growth Differentiation factor 11 (facteur 11 de croissance et de différenciation)

GRAMD2 : GRAM Domain containing 2A (facteur GRAM-domain containing 2A)

GSK3 : Glycogen Synthase Kinase 3 (glycogène kinase synthétase 3-β)

GST : Glutathione-S Transferase

GSTrap : Colonne de chromatographie permettant la purification de protéines recombinantes qui présentent une étiquette GST

GST-tag : Étiquette GST qui est additionnée à la protéine recombinante pour sa purification

GTP : Guanosine triphosphate

Gu : Guanine

HA : Hemagglutinin (hémagglutinine)

HA0 : Protéine HA immature

HA1 : Domaine N-terminal de HA qui héberge le peptide signal et le site de clivage

HA2 : Domaine C-terminal de HA qui héberge le peptide de fusion et le domaine transmembranaire

HBEC : Human Bronchial Epithelial Cells (cellules immortalisées d'épithélium bronchique humain)

HCl : Hydrochloric acid (acide chlorhydrique ou chlorure d'hydrogène)

HEK293 : Human Epithelial Kidney cells (cellules epithéliales de rein humain)

HEK293T : Immortalized Human Epithelial Kidney cells (cellules immortalisées d'épithélium rénale humain)

HEPES : 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (acide 4-(2-hydroxyéthyl)-1-pipérazine éthane sulfonique)

hFIP1 : Pre-mRNA 3'-end-processing factor FIP1 (facteur de maturation 3' des pré-ARNm)

His_6-tag : Étiquette de 6 histidines qui est additionnée à la protéine recombinante pour sa purification

HK68 : Souche virale A/Hong Kong/1/68 (H3N2)

HPAIV : Highly Pathogenic Avian *Influenza* Virus (virus *influenza* aviaire hautement pathogène)

HPLC : High Performance Liquid Chromatography (chromatographie en phase liquide à haute performance)

H1 : Histone 1

H2A : Histone 2A

H2B : Histone 2B

H_2O : Eau

H3 : Histone 3

H4 : Histone 4

I : Isoleucine

IAV : *Alphainfluenzavirus* (virus *influenza* A)

IBV : *Betainfluenzavirus* (virus *influenza* B)

ICV : *Gammainfluenzavirus* (virus *influenza* C)

IDV : *Deltainfluenzavirus* (virus *influenza* D)

IFITM : Interferon Induced-Transmembrane protein (protéines transmembranaires induites par l'interféron)

IFN : Interféron

IMAC : Immobilized Metal Affinity Chromatography (chromatographie d'affinité sur ions métaliques immobilisés)

In : Inositol

IRX3 : Iroquois-class homeodomain protein 3 (protéine homéodomaine iroquoise de classe 3)

ISG : Interferon-Stimulated Genes (gènes exprimés en réponse à l'interféron)

ITGB4BP : Human Intergrin Beta 4 Binding Protein (protéine de liaison à l'intégrine β-4)

JAK2 : Tyrosine-protein Janus kinase 2 (protéine Janus kinase 2)

K : Lysine

Kb : Kilobases

Kd : Dissociation constant (constante de dissociation)

kDa : Kilodalton

KRT14 : kératine 14

L : Leucine

LB : Lysogeny Broth medium ou Luria-Bertani medium (milieu de culture bactérienne de Luria-Bertani)

LPAIV : Low Pathogenic Avian *Influenza* Virus (virus *influenza* aviaire faiblement pathogène)

M : Methionine

MDCK : Madin-Darby canine kidney cells (cellules immortalisées de rein de chien)

$MgCl_2$: Magnesium Chloride (chlorure de magnésium)

mg : Milligramme

miARN : MicroARN

mL : Millilitre

mM : Millimolaire

MME : Monomethyl Ether (ether de monométhyle)

MOPS : 3-(N-morpholino)propanesulfonic acid (acide 3-morpholino-1-propanesulfonique)

MORC3 : MORC family CW-type zinc inger protein 3 (protéine MORC de doigt de zinc de type CW de la famille MORC3)

MOT : Micro-Organismes et Toxines

MRPL42 : Mitochondrial Ribosomal Protein L42 (protéine ribosomale mitochondriale 42)

MS : Mass Spectrometry (spectrométrie de masse)

MTOC : Microtubule Organizing Center (centre d'organisation des microtubules)

mutBCD : Triple mutant pour les sites B, C et D

mutB : Simple mutant pour le site B

mutC : Simple mutant pour le site C

mutD : Simple mutant pour le site D

M^7GTP : Analogue de structure aux acides sialiques

m.o.i : Multiplicity Of Infection (multiplicité des infections correspond au rapport des agents aux cibles d'infection)

N : Asparagine

NA : Neuraminidase

NaCl : Sodium chloride (chlorure de sodium)

NdeI : Enzyme de restriction provenant de *Neisseria denitrificans*

Neu5Ac5en : Analogue de structure aux acides sialiques

ng : Nanogramme

nL : Nanolitre

NLS : Nuclear Localization Signal (signal de localisation nucléaire)

nM : Nanomolaire

nm : Nanomètre

NP : Nucléoprotéine

NS : Segment génomique viral codant pour les protéines non-structurales 1 et 2

NS1 : Non-Structural Protein 1 (protéine non-structurale 1)

NS1-BP : NS1-Binding Protein (protéine liant la protéine NS1 des virus *influenza*)

NS2 (ou NEP) : Non-Structural Protein 2 (Nuclear Export Protein ou protéine d'export nucléaire)

NT : Nucléotides

NTP : Nucléosides triphosphates

NUP98 : Nucleoporin 98 and 96 precursor (protéine de pore nucléaire de 98 kDa)

NXF1 : Nuclear RNA export factor 1 (protéine d'export nucléaire 1 des ARNm)

NXT1 : NTF2-related export protein 1 (protéine d'export des ARNm liée au facteur NTF2)

OAS : 2',5'-Oligoadénylate Synthétases

OGM : Organismes Génétiquement Modifiés

OMS : Organisation Mondiale de la Santé

ORF : Open Reading Frame (cadre ouvert de lecture)

P : Proline

PA : Protéine Acide

PABP : Poly-A Binding Protein (protéine liant la queue poly-A des ARNm)

PABPII : Poly-A Binding Protein II (protéine 2 liant la queue poly-A des ARNm)

PACT : Protein ACTivator of the interferon-induced double-stranded RNA-dependent protein kinase (protéine activatrice de la Protéine Kinase R)

PAF1 : RNA polymerase II-associated protein 1 (protéine 1 associée à l'ARN polymérase II cellulaire)

PAP : Poly-A polymerase (protéine à activité Polymérase qui ajoute la queue poly-A aux ARNm)

PAGE : Polyacrylamide Gel Electrophoresis (electrophorèse en gel de polyacrylamide)

PA-C : Domaine C-terminal de PA

PA-endo : Domaine N-terminal de PA qui porte l'activité endonucléase

PA-N : Domaine N-terminal de PA

Pb : Paire de bases

PBS : Phosphate-Buffered Saline (tampon phosphate salin)

PB1 : Protéine Basique 1

PB1-C : Domaine C-terminal de PB1

PB1-N : Domaine N-terminal de PB1

PB2 : Protéine Basique 2

PB2-C : Domaine C-terminal de PB2

PB2-N : Domaine N-terminal de PB2

pcDNA : Vecteur plasmidique qui permet la transcription *via* Pol II des ARNm viraux

pCI : Vecteur plasmidique qui permet la transcription *via* Pol II des ARNm viraux

PCR : Polymerization Chain Reaction (réaction de polymérisation en chaine)

PDB : Protein Data Bank (banque de données sur les protéines dont la structure à été résolue)

PDZ : The postsynaptic protein PSD-95/SAP90, the *Drosophila* septate junction protein Discs-large, et the tight junction protein ZO-1 (protéine d'ancrage qui se lie aux protéines réceptrices dans la membrane aux composants du cytosquelette)

PEG : Polyethylene Glycol

pET15b(+) : Vecteur plasmidique qui permet l'expression de protéines recombinantes avec une étiquette His_6 en position N-terminale

PFU : Plaque Forming Units (plages de lyse)

Pfu : High-Fidelity DNA Polymerase (ADN polymérase de haute fidélité)

PGD : Phosphogluconate déshydrogénase

pGEX-5X-1 : Vecteur plasmidique qui permet l'expression de protéines recombinantes avec une étiquette GST en position N-terminale

pH : Potentiel hydrogène

PHB : Prohibitin (prohibitine)

PIAS3 : Protein inhibitor of Activated STAT3 (protéine inhibitrice de STAT3)

PIGN : Phosphatidylinositol Glycan anchor Biosynthesis class N (facteur de biosynthèse d'ancrage au phosphatidylinositol glycane, classe N)

PIP2 : Phosphatidylinositol-4,5-biphosphate

PIP3 : Phosphatidylinositol-(3,4,5)-triphosphate

PI3K : Phosphoinositide-3-Kinase

PKR : Protéine Kinase R

PLB : Passive Lysis Buffer (tampon de lyse passive)

PMSF : Phenylmethylsulfonyl Fluoride (fluorure de phénylméthylsulfonyle)

PNK : Polynucleotide Kinase T4

Pol I : ARN Polymérase cellulaire I

Pol II : ARN Polymérase cellulaire II

polyI:C : Polyinosinic :polycytidylic acid (acides polyinosinique et polycytidilique)

POROS HS20 : Colonne de chromatographie échangeuse de cations

pPolI : Vecteur plasmidique qui permet la transcription *via* Pol I des ARN viraux génomiques

PPP6C : Protein Phosphatase 6 Catalytic subunit (sous-unité catalytique sérine/thréonine-protéine phosphatase 6)

Pré-ARNm : Pré-ARN messagers cellulaires (immatures)

PR8 : Souche virale A/Puerto Rico/8/1934 (H1N1)

PS : Pénicilline/Streptomycine

PSMD11 : 26S proteasome non-ATPase regulatory subunit 11 (sous-unité 26S du protéasome non-ATPase 11)

P24 : Plaque 24 puits

p53 : Tumor Protein 53 (facteur tumorale 53)

P6 : Plaque 6 puits

P96 : Plaque 96 puits

Q : Glutamine

R : Arginine

Rab-5A : Ras-related in brain 5A (protéine Rab-5A liée à Ras)

Rab-10 : Ras-related in brain 10 (protéine Rab-10 liée à Ras)

Rab-11 : Ras-related in brain 11 (protéine Rab-11 liée à Ras)

RanBP2 : Ran-Binding Protein 2 (protéine de liaison Ran 2)

RanBP5 : Ran-Binding Protein 5 (importine 5 ou IPO5)

Ran-GTP : Ras-related nuclear protein (protéine Ran nucléaire liée à Ras qui fixe le GTP)

RBP : RNA-Binding Protein (protéine liant l'ARN)

RBD : RNA-Binding Domain (domaine de liaison à l'ARN)

RBD-A : RBD de NS1 d'allèle A

RBD-B : RBD de NS1 d'allèle B

RBS : Receptor-binding site (domaine de liaison au récepteur)

RdRp : RNA-dependent RNA polymerase (ARN polymérase ARN-dépendante)

RE : Réticulum endoplasmique

REG : Réticulum endoplasmique granuleux

Ren-Luc : Renilla Luciferase (luciférase Renilla)

RMN : Résonance magnétique nucléaire

RMSD : Root mean square deviation (déviation carrée moyenne de la racine)

RNP : Ribonucléoprotéine

RNPv : Ribonucléoprotéines virales

Rpm : Rapports par minute

RRM2 : RNA Recognition Motif 2 (Motif 2 de reconnaissance de l'ARN)

RT : Reverse Transcriptase (Transcriptase inverse)

S : Sérine

SA : Sialic acids (acides sialiques)

SCaMC2 : Calcium-binding Mitochondrial Carrier protein 2 (transporteur mitochondrial qui lie le calcium)

SDF2L1 : Stromal cell-derived factor 2-like protein 1 (facteur 2-like 1 dérivé de cellules stromales)

SDS : Sodium Dodecyl Sulfate (dodécylsulfate de sodium ou laurylsulfate de sodium)

SDS-PAGE : Sodium Dodecyl Sulfate Polyacrilamide Gel Electrophoresis (electrophorèse en gel de polyacrilamide en présence de SDS)

SELEX : Systematic Evolution of Ligands by Exponential enrichment (évolution systématique de ligands par enrichissement exponentiel)

SePOP : Selective Precipitation Optimized Process (technique optimisée de purification des ARN par précipitation sélective)

SFRS10 : Splicing Factor, Arginine/Serine-Rich 10 (facteur β-2 humain transformé)

SiRNA : Small interfering RNA (petits ARN interférents)

SLC35A1 : Solute Carrier Family 35 Member A1 (transporteur d'acide CMP-sialique)

snARN : Small Nuclear RNA (petits ARN nucléaires)

snRNP : small nuclear Ribonucleoprotein (petites sous-unités nucléaires de ribonucléoprotéines)

ssRBP : single-stranded RNA binding protein (protéines se liant à des ARN simple-brin)

STAT1 : Signal Transducer And Activator Of Transcription 1 (transducteur de signal et activateur de la transcription 1)

STAT3 : Signal Transducer And Activator Of Transcription 3 (transducteur de signal et activateur de la transcription 3)

Staufen : Double-stranded RNA-Binding protein Staufen homolog 1 (Protéine homolog 1 à Staufen liant les ARNdb)

SVF : Sérum de veau fœtal

svRNA : Small Viral RNA (petits ARN viraux qui correspondent à la région 5'-UTR des ARNv)

S200 : Colonne superdex de chromatographie d'exclusion-diffusion des protéines

S75 : Colonne superdex de chromatographie d'exclusion-diffusion des protéines

T : Thréonine

TAE : Tampon qui contient du Tris-HCl, de l'Acétate et de l'EDTA

TAPT1 : Transmembrane anterior posterior transformation protein 1 homolog (protéine transmembranaire antérieure postérieure 1)

TBE : Tampon contenant 89 mmol.L^{-1} Tris; 89 mmol.L^{-1} Borate; 2 mmol.L^{-1} EDTA

TBHQ : Tert-Butyl Hydroquinone

TCEP : tris(2-carboxyethyl)phosphine)

TEN : Tampon de purification et d'hybridation des ARN contenant 10 mmol.L^{-1} Tris-HCl; 1 mmol.L^{-1} EDTA; 500 mmol.L^{-1} NaCl

Th : Thymidine

TNBG : Tampon de dilution des protéines purifiées contenant 10 mmol.L^{-1} Tris-HCl; 1 mmol.L^{-1} EDTA, 100 mmol.L^{-1} NaCl; 0,1 % BSA; 10 % glycérol)

TNT : Tampon contenant 10 mmol.L^{-1} Tris-HCl; 150 mmol.L^{-1} NaCl; 0,001 % Tween 20

TPCK : L-(Tosylamido-2-Phenyl) ethyl Chloromethyl Ketone (tosyl phénylalanyl chlorométhyl cétone)

TRIM23 : Tripartite Motif containing 23 E3 ubiquitin-protein ligase (ligase E3 tripartite-motif 23)

TRIM25 : Tripartite Motif containing 25 E3 ubiquitin-protein ligase (ligase E3 tripartite-motif 25)

TWA : Mélange acétonitrile : eau : acide trifluoroacétique pour l'analyse par HPLC de la pureté des protéines recombinantes purifiées

T25 : Flasque de culture de 25 cm^2

T75 : Flasque de culture de 75 cm^2

T-705 : Favipiravir

U : unité

Udorn : Souche virale A/Udorn/1972 (H3N2)

Ur : Uracile

USA : United States (Etats-Unis d'Amérique)

UTR : Untranslated Region (région non-traduite)

U12 : Courte séquence ARN conservée dans la région 3'UTR des ARNv

U13 : Courte séquence ARN conservée dans la région 5'UTR des ARNv

U2OS : Human bone osteosarcoma epithelial cells (cellules immortalisées d'épithélium d'ostéocarcome humain)

V : Valine

vATPase : Vacuolar ATPase (ATPase vacuolaire)

VCP : Valosin-Containg Protein (ATPase de transition du ER)

VERO : Kidney Epithelial cells from an African green monkey (cellules épithéliales de rein de singes)

VIH : Virus de l'Immunodéfience Humaine

VN04 : Souche virale A/Vietnam/1203/04 (H5N1)

VX-787 : Pimodovir

W : Tryptophane

WDR33 : WD Repeat domain 33 (facteur de maturation 33 des pré-ARNm)

WSN : Wilson smith neurotropic virus (virus neurotrope Wilson Smith, A/WSN/1933 (H1N1))

WT : Wild-type (sauvage)

Y : Tyrosine

2D : Structure bidimensionnelle

2'OH : Groupement hydroxyle libre en position 2 du ribose de l'ARN (permet de différencier l'ARN de l'ADN)

3D : Structure tridimensionnelle

^{32}P : Phosphore 32 radioactif

4-HCCA : 4-hydroxy-α-cyano-cinnamique

4R-11a : Composé dérivé de monoterpènes ciblant la protéine HA

µg : Microgramme

µL : Microlitre

µM : Micromolaire

α : Protéine contenant des acides aminés qui se replient sous la forme d'une hélice alpha

β : Protéine contenant des acides aminés qui se replient sous la forme d'une brin (ou feuillet) beta

°C : Degrés Celsius

RÉSUMÉ

Les *alphainfluenzavirus* (IAV) sont un problème persistant au niveau mondial et notre arsenal thérapeutique actuel, limité à seulement deux molécules antivirales ciblant la neuraminidase virale et à un inhibiteur ciblant la polymérase, n'est pas à l'abri de résistances. Les IAV bloquent l'amorçage de la réponse antivirale en inhibant notamment l'induction du système interféron (IFN) *via* la protéine non-structurale 1 (NS1). Cette dernière est réputée pour ses interactions avec de nombreuses protéines cellulaires et avec l'ARN, en particulier avec des ARN double-brin (ARNdb) ainsi que des ARN viraux arborant des motifs conservés du virus. La double-substitution R38A-K41A, qui invalide le domaine de liaison à l'ARN (RBD) de NS1, est connue pour abolir la virulence et réduire fortement la multiplication virale, soulignant l'importance du RBD dans les fonctions de NS1. Essentielle au déroulement complet du cycle viral, NS1, et notamment son RBD, représente donc une cible thérapeutique prometteuse pour le développement de nouveaux antiviraux.

Mon travail de doctorat s'inscrit dans la continuité d'un projet antérieur ayant permis d'identifier par SELEX des aptamères d'ARN de haute affinité pour NS1 (Kd de l'ordre du nanomolaire) et qui arborent (i) les séquences 5'-AGCAAAAG-3' et 5'-UGAUUGAAG-3' qui semblent spécifiques du virus et (ii) le motif GUAAC qui pourrait correspondre à la séquence canonique d'un site donneur d'épissage d'ARNm cellulaires. Dans un premier temps, j'ai vérifié l'importance de la séquence 5'-UGAUUGAAG-3' hautement conservée dans le segment viral NS en la mutant spécifiquement et en évaluant son impact dans un système miniréplicon viral ainsi que directement dans le cycle viral en caractérisant le phénotype du virus mutant. Même si très conservée dans la région 3'-UTR des ARNm de NS1 justifiant un potentiel rôle régulateur dans la traduction de NS1, cette séquence virale n'est pas essentielle au virus puisque sa mutation n'impacte nullement ce dernier. La seconde partie de ce projet aura porté sur la caractérisation des éléments de l'ARN et du RBD de NS1 qui sont essentiels à la formation de complexes RBD/ARNdb de haute affinité. Ainsi, nous avons pu montré que de tels complexes se forment en gel retard lorsque le RBD rencontre un ARNdb (i) d'au moins 15 paires de bases et (ii) qui contient au moins un motif GUAAC accessible dans sa séquence. Nous avons également montré que la séquence virale 5'-AGCAAAAG-3' participait à la formation de complexes de haute affinité avec le RBD de NS1. Enfin, l'analyse structurale par cristallographie que nous avons mené sur un complexe RBD/ARNdb de haute affinité indique que le RBD de NS1 (i) interagit avec le squelette ribose phosphate de l'ARNdb mais sans reconnaître spécifiquement ses nucléobases et (ii) provoque une légère courbure de l'ARNdb lorsque ce dernier contient les motifs GUAAC dans sa séquence.

Mots-clés : *influenza* ; NS1 ; RBD ; ARNdb ; SELEX ; Cristallographie ; Courbure.

RÉSUMÉ EN ANGLAIS

Influenza A viruses (IAV) are a global public health problem and the available therapeutic options, limited to two anti-neuraminidase and one anti-viral polymerase, are amenable to viral resistance. IAVs block antiviral responses by inhibiting interferon (IFN) pathway by way of their non-structural protein 1 (NS1). NS1 is known to interact with several cellular proteins and RNA molecules, in particular with double-stranded RNA (dsRNA) and viral RNA harboring virus-specific motifs. The arginine 38 (R38)-lysine 41 (K41) double-substitution into alanine, which highly invalidates the RNA-Binding Domain (RBD) of NS1, is known to strongly affect the viral life cycle, indicating an essential role of the RBD in the functions of NS1. NS1, in particular its RBD domain, which is essential to the viral life cycle, could be a promising therapeutic target for the development of new anti-*influenza* compounds.

My PhD research project follows on from the identification of RNA aptamers with a high affinity to NS1 (Kd the nanomolar range) and which contain virus-specific motifs (5'-AGCAAAAG-3' and 5'-UGAUUGAAG-3') and canonical donor splice sites of cellular pre-mRNA (also called GUAAC motif). Hence, in the first part of my project, I focused on the 5'-UGAUUGAAG-3' viral sequence of the NS genomic segment, by mutating it and evaluated its impact in an in vitro viral minireplicon system and on the phenotype of mutant viruses rescued by reverse genetics. Even though the 5'-UGAUUGAAG-3' viral sequence is highly conserved in the 3'-untranslated region (UTR) of the NS1 mRNA, we showed that it is not required for the IAV viral life cycle. In the second part, we focused on the characterization, using gel-shift assays, of RBD/dsRNA complexes to identify key elements, in both partners, that are required to form high affinity complexes. According to our results, such complexes require a minimum of 15 base pairs and one accessible GUAAC motif in their sequence. We also showed that the AGCAAAAG viral sequence could participate to high affinity complex formation. To finish, our structural analyze of high affinity complex seems indicated that RBD of NS1 interact with the sugar-phosphate skeleton of RNA but without recognize the RNA basis. Moreover, our structural approach also showed that RBD of NS1 provoked a curvature in dsRNA when containing GUAAC motif in their sequence.

Key-words: *influenza* ; NS1; RBD; dsRNA; SELEX; X-ray crystallography; RNA curvature.

TABLE DES MATIERES

REMERCIEMENTS 4
LISTE DES ABRÉVIATIONS 10
RÉSUMÉ 21
RÉSUMÉ EN ANGLAIS 22
TABLE DES MATIERES 23
LISTE DES TABLEAUX 28
LISTE DES FIGURES 29
LISTE DES ANNEXES 32
INTRODUCTION 33
A. Classification, nomenclature, génomique, écologie et évolution des virus *influenza* A 34
A.I. Généralités 35
A.I.1. Classification 35
A.I.2. Morphologie et structure 35
A.I.3. Nomenclature 37
A.II. Génomique des virus *influenza* A 37
A.II.1. Organisation du génome viral 37
A.II.2. Le protéome viral 38
A.III. Ecologie et évolution 40
A.III.1. Ecologie 40
A.III.2. Evolution 41
A.III.2.a. La dérive antigénique ou *antigenic drift* 41
A.III.2.b. Cassure antigénique ou *antigenic shift* 41
A.III.3. Les virus grippaux humains et leurs réassortiments 42
B. Cycle viral des virus *influenza* A 44
B.I. Les grandes étapes du cycle viral 45
B.II. Attachement et entrée du virus dans la cellule hôte 45
B.III. Fusion membranaire de l'enveloppe virale et cellulaire, libération cytoplasmique et import nucléaire des ribonucléoprotéines virales 47
B.IV. Transcription et réplication du génome viral 49
B.IV.1. La ribonucléoprotéine virale 49
B.IV.2. L'ARN polymérase ARN-dépendante virale 50
B.IV.2.a. La protéine acide 51
B.IV.2.b. La protéine basique 1 51
B.IV.2.c. La protéine basique 2 52

B.IV.2.d. La nucléoprotéine ... 52
B.IV.3. La transcription virale ... 53
B.IV.3.a. Le vol de coiffe ou *cap-snatching* ... 54
B.IV.3.b. Initiation, élongation et polyadénylation ... 54
B.IV.4. La réplication virale ... 56
B.IV.5. Régulations des étapes de transcription et de réplication du génome viral : transition entre les deux états ... 56
B.V. Export nucléaire et traffic cytoplasmique des RNPv néo-synthétisées ... 57
B.VI. Empaquetage des RNPv, bourgeonnement et relargage des particules virales ... 59
C. Les moyens de lutte mis en place contre les virus *influenza* ... 61
C.I. Les méthodes préventives ... 62
C.I.1. La vaccination annuelle ... 62
C.I.1.a. Les vaccins antigrippaux inactivés ... 62
C.I.1.b. Les vaccins antigrippaux vivants atténués ... 63
C.I.2. Les vaccins de demain ... 63
C.I.2.a. Les vaccins antigrippaux « high-dose » ... 63
C.I.2.b. Développement d'un vaccin universel ... 64
C.II. Les méthodes curatives ... 64
C.II.1. Les inhibiteurs de la protéine M2 ... 64
C.II.2. Les inhibiteurs de la neuraminidase ... 65
C.II.3. Les traitements ciblant l'hémagglutinine ... 65
C.II.4. Les inhibiteurs de la RdRp virale ... 66
C.II.5. Les inhibiteurs de la protéine NP ... 68
C.III. Identification de nouveaux facteurs cellulaires essentiels au cycle viral pour le développement de traitements futurs pouvant efficacement limiter l'infection ... 69
D. La protéine non-structurale 1 (NS1) des virus *influenza* A ... 72
D.I. Généralités ... 73
D.I.1. Découverte ... 73
D.I.2. Polymorphisme ... 73
D.I.3. Modifications post-traductionnelles ... 74
D.II. Structure ... 75
D.II.1. Le domaine de liaison à l'ARN ... 77
D.II.2. Le domaine effecteur ... 77
D.II.3. La région linker ... 79
D.II.4. La queue C-terminale ... 80
D.III. La protéine NS1 et ses nombreuses fonctions biologiques dans la cellule infectée ... 80
D.III.1. Les fonctions de NS1 dépendantes du domaine RBD ... 80

D.III.1.a. Inhibition de la réponse interféron 80
D.III.1.b. Inhibition de l'activité des protéines antivirales 82
D.III.1.c. Réduction de l'expression des gènes cellulaires par inhibition de l'épissage des pré-ARNm cellulaires 83
D.III.1.d. Stimulation de la traduction des ARNm viraux 84
D.III.2. Les fonctions de NS1 dépendantes du domaine ED 85
D.III.2.a. Activité anti-apoptotique 85
D.III.2.b. Inhibition de la maturation des pré-ARNm cellulaires 86
D.III.2.c. Blocage du transport nucléocytoplasmique des ARNm cellulaires 87
D.III.2.d. Remodelage de l'architecture 3D de la chromatine par élongation de la transcription génomique 88
D.IV. Les inhibiteurs de NS1 89
D.IV.1. Les molécules ciblant le domaine RBD 89
D.IV.2. Les molécules ciblant le domaine ED 90
D.IV.3. Les inhibiteurs identifiés par *docking* virtuel 90
SITUATION ET PROBLEMATIQUE DE LA THESE 92
OBJECTIFS DE LA THESE 100
MATÉRIELS ET MÉTHODES 102
A. Culture cellulaire 103
A.I. Les bactéries 103
A.II. Les cellules de mammifères 103
B. Les plasmides 104
B.I. Vecteurs d'expression de protéines recombinantes 104
B.II. Plasmides de génétique inverse du virus WSN 105
B.III. Plasmides du système *miniréplicon viral* 106
B.IV. Plasmides du système *translation-only* 106
C. Constructions plasmidiques 107
C.I. Amplification et purification de vecteurs 107
C.II. Digestion d'ADN plasmidique et purification de fragments 107
C.III. Ligation et transformation 107
C.IV. Criblage de clones transformés par PCR 108
C.V. Mutagénèse dirigée 108
C.VI. Séquençage des constructions plasmidiques 109
D. Expression du système *miniréplicon viral* en cellules de mammifères 110
E. Production de virus par génétique inverse 111
E.I. Titrage des virus 112
E.II. Séquençage des génomes viraux 112

E.III. Phénotype des virus mutants 113
E.III.1. Cinétique de production de particules virales 113
E.III.2. Cinétique d'accumulation de NS1 dans les cellules infectées 114
F. Expression et purification de protéines recombinantes 115
F.I. Expression de protéines recombinantes en bactéries 115
F.II. Purification des protéines RBD de NS1 recombinantes 115
F.III. Purification des protéines de fusion GST-RBD et GST-NS1 recombinantes 116
F.IV. Analyse des fractions protéiques 117
G. Retards de migration sur gel 121
G.I. Radio-marquage et hybridation des ARNdb à évaluer 121
G.II. Détermination des constantes de dissociation (K_d) des complexes RBD/ARNdb 121
G.III. Compétition en gel-retard 122
H. Etudes cristallographiques du RBD-B de NS1 et de ces complexes avec l'ARNdb 123
H.I. Cristallogenèse 124
H.I.1. Cristallisation du RBD-B de NS1 124
H.I.2. Cristallisation des complexes RBD-B/ARNdb 124
H.II. Résolution des structures tridimensionnelles 125
H.II.1. Résolution de la structure du RBD-B 125
H.II.2. Résolution de la structure des complexes RBD-B/ARNdb 126
H.II.3. Représentation des structures 3D, analyse des paramètres hélicoïdaux des ARNdb et visualisation des interactions RBD/ARNdb 126
RÉSULTATS 127
A. Relevance biologique des séquences virales identifiées par SELEX 128
A.I. Des motifs ARN caractéristiques des ARN viraux de polarité positive 129
A.II. Choix des mutations 130
A.II.1. Impact des mutations dans le système *miniréplicon viral* 130
A.II.1.a. Expression de NS1 dans les cellules transfectées 131
A.II.1.b. Effet de NS1 sur l'expression du *miniréplicon viral* 132
A.II.1.c. Effet de NS1 sur la traduction des transcrits viraux 133
A.II.2. Phénotype des virus mutants 135
A.II.2.a. Courbe de croissance virale 135
A.II.2.b. Cinétique d'accumulation de NS1 dans les cellules infectées 135
B. Interaction de NS1 avec l'ARN double-brin 137
B.I. Identification et sélection d'une séquence minimale d'ARN double-brin de haute affinité pour NS1 inspirée de celles des aptamères DM01 et DM03 identifiées par la méthode SELEX 138
B.I.1. Sélection des protéines RBD et NS1 entières à purifier 138

B.I.2. Des préparations protéiques purifiées de haute qualité pour les interactions en gel retard et la cristallogenèse 139
B.I.3. Décryptage des déterminants structuraux et/ou fonctionnels des ARN double-brin DM01 et DM03 associés à une haute affinité pour NS1 141
B.I.3.a. Rôle potentiel du domaine ED de NS1 dans la fixation aux aptamères DM01-midi et DM03-midi 141
B.I.3.b. Longueur minimale de duplex d'ARN exigée par le RBD de NS1 pour former des complexes de haute affinité 143
B.I.3.c. Deux modes possibles de fixation du RBD de NS1 à l'ARN 145
B.I.3.d. Le motif viral AGCAAAAG, implication dans la formation de complexes 147
B.I.4. Conception d'une sonde ARN de haute affinité pour le RBD de NS1 utilisable dans une approche structurale par cristallographie 148
B.I.4.a. Le motif GUAAC identifiée par SELEX est une séquence de haute affinité pour le RBD de NS1 148
B.I.4.b. Conception de l'ARNdb AWFC01, une sonde modèle de haute affinité pour étudier les interactions RBD/ARN 149
B.I.4.c. AWFC01, un ligand de haute affinité pour le RBD de nombreuses NS1 151
B.I.4.d. Détermination des EC_{50app} des sondes par compétition contre AWFC01 153
B.II. Bases structurales de l'interaction de NS1 avec l'ARN 154
B.II.1. Cristallisation et résolution des structures 3D des protéines et des complexes RBD/ARNdb 155
B.II.1.a. Conservations structurales entre RBD-A et RBD-B 158
B.II.1.b. Comparaison entre RBD-B libre et RBD-B/AWFC01 159
B.II.1.c. Comparaison entre RBD-B/AWFC01 et aaRBD-B/AWFC01 162
B.II.1.d. Comparaison entre RBD-B/AWFC01 et RBD/CHENG 163
B.II.1.e. Comparaison entre RBD-B/ZKO* et RBD/CHENG 165
B.II.2. L'interaction du RBD-B avec AWFC01 induit une légère courbure de l'ARN 166
CONCLUSIONS GENERALES ET DISCUSSION 169
PERSPECTIVES 178
ANNEXES 207
PUBLICATION 1 : Structure and sequence determinants governing the interactions of RNAs with *influenza* A virus Non-structural protein NS1 220
PUBLICATION 2 : Major contribution of the RNA-binding domain of NS1 in the pathogenicity and replication potential of an avian H7N1 influenza virus in chickens 246

LISTE DES TABLEAUX

Tableau 1 : Le protéome viral des IAV 39
Tableau 2 : Structures de NS1 (RBD et/ou ED seuls) résolues à ce jour 76
Tableau 3 : Structures de NS1 (RBD et/ou ED seuls) associées à des partenaires..... 76
Tableau 4 : Conditions et paramètres de PCR pour l'amplification d'ADN d'intérêt et la recherche de clones bactériens transformés 108
Tableau 5 : Conditions et paramètres de PCR pour la mutagénèse dirigée 109
Tableau 6 : Conditions et paramètres de retro-transcription des ARN viraux en ADNc 113
Tableau 7 : Collecte et affinement statistiques des données de diffraction obtenues pour la structure RBD-B libre et pour les structures des complexes RBD-B/AWFC01, aaRBD-B/AWFC01 et RBD-B/ZKO* 156
Tableau 8 : Interactions directes et indirectes établis par le RBD de NS1 avec les ARNdb AWFC01 et CHENG 157

LISTE DES FIGURES

Figure 1 : Architecture des différentes formes d'IAV. 36
Figure 2 : Une particule virale d'IAV et son contenu génomique. 36
Figure 3 : Structure des segments d'ARN génomiques viraux d'IAV. 38
Figure 4 : Transmissions intra et inter-espèces des différents sous-types d'IAV. 40
Figure 5 : Réassortiment génomique du virus aviaire asiatique H7N9. 41
Figure 6 : Evolution et émergence des IAV humains depuis le début du 20ème siècle. 43
Figure 7 : Schéma simplifié du cycle viral des IAV 45
Figure 8 : Structure des acides sialiques 46
Figure 9 : Structure de l'hémagglutinine. 48
Figure 10 : La fusion de l'enveloppe virale et de la membrane endosomale. 49
Figure 11 : Une ribonucléoprotéine virale. 50
Figure 12 : Structure de l'ARN polymérase ARN-dépendante virale. 51
Figure 13 : Structure de la nucléoprotéine. 53
Figure 14 : Transcription des RNPv en ARNm viraux 55
Figure 15 : Réplication des RNPv 56
Figure 16 : Assemblage et traffic des RNPv néo-synthétisées 58
Figure 17 : Structure de la neuraminidase. 60
Figure 18 : Principe du réassortiment viral pour l'obtention de vaccins fragmentaires 62
Figure 19 : Les adamantanes, ciblant la protéine M2 65
Figure 20 : Les analogues d'acides sialiques, ciblant la protéine NA 65
Figure 21 : Les composés chimiques ciblant la protéine HA 66
Figure 22 : Les inhibiteurs de la RdRp virale 68
Figure 23 : Les inhibiteurs ciblant la protéine NP 69
Figure 24 : Polymorphisme de NS1. 74
Figure 25 : Modifications post-traductionnelles de NS1. 75
Figure 26 : Structure du domaine RBD de NS1. 77
Figure 27 : Structure du domaine ED de NS1. 78
Figure 28 : Structures des différentes conformations de NS1. 79
Figure 29 : Inhibition de la réponse IFN-I par NS1 82
Figure 30 : Inhibition de l'épissage des pré-ARNm cellulaires par NS1 84
Figure 31 : Stimulation de la voie Akt/PKB par NS1 86
Figure 32 : Inhibition de l'export nucléaire des ARNm cellulaires par NS1 88
Figure 33 : Inhibiteurs ciblant la protéine NS1 90
Figure 34 : Inhibiteurs potentiels de NS1 identifiés par *docking*. 91

Figure 35 : Modes de reconnaissance des ARNsb et des ARNdb par les RBP. 95
Figure 36 : Mécanisme de reconnaissance des ARNdb par le RBD de NS1 des IAV. 96
Figure 37 : Structure 2D des ARN U6m et U6atac qui interagissent avec le RBD de NS1. 96
Figure 38 : Les aptamères DM01 et DM03. 97
Figure 39 : Structure et conservation du site donneur d'épissage des introns. 98
Figure 40 : Carte du vecteur pET15b(+) et schéma des protéines RBD recombinantes 104
Figure 41 : Carte du vecteur pGEX-5X-1 et schéma des protéines de fusion à la GST 105
Figure 42 : Plasmides pcDNA et pPolI utilisés pour la génétique inverse. 106
Figure 43 : Construction de l'ARN pseudo-viral rapporteur. 106
Figure 44 : Principe du dosage des activités luciférases FireFly et Renilla. 110
Figure 45 : Principe de la production de virus par génétique inverse 111
Figure 46 : Purification des protéines recombinantes RBD, GST-RBD et GST-NS1 120
Figure 47 : Diagramme de phase. 123
Figure 48 : Obtention de cristaux et collecte des données de diffraction 125
Figure 49 : Localisation des motifs viraux identifiés par SELEX dans les ARNm de NS1. 129
Figure 50 : Mutations des motifs B, C et D dans la 3'-UTR de NS1. 130
Figure 51 : Principe du *miniréplicon viral.* 131
Figure 52 : Niveau d'expression de NS1 dans les cellules transfectées. 132
Figure 53 : Effet de NS1 (WT ou mutant R38A-K41A) sur l'expression du *miniréplicon.* 133
Figure 54 : Système *translation-only.* 133
Figure 55 : Effets de NS1 et de ses 5' et 3'-UTR sur la traduction des ARNm de NS1. 134
Figure 56 : Phénotype des virus WSN WT et mutants. 136
Figure 57 : Alignement de séquences peptidiques des RBD de NS1 étudiés. 139
Figure 58 : Homogénéité de la fraction purifiée du RBD-B. 140
Figure 59 : Interaction des protéines de fusion RBD-B et NS1-B avec DM01-midi et DM03-midi. . 143
Figure 60 : Reconnaissance par NS1 des dérivés d'aptamères DM01 et DM03. 144
Figure 61 : Titrage de DM01-midi et DM03-midi par le RBD-B. 146
Figure 62 : Interaction du RBD-B avec DM01-midi, DM01-Mod1 et DM01-Mod2. 148
Figure 63 : Interaction du RBD-B avec des aptamères présentant un ou deux sites GUAAC. 149
Figure 64 : Interaction du RBD-B avec DM01-midi, DM01-Mod1 et DM01-Mod2. 151
Figure 65 : Détermination des Kd des différents RBD purifiés pour le duplex AWFC01. 152
Figure 66 : Détermination par compétition des EC_{50} des différents aptamères et ARNdb pour la dissociation du complexe RBD/AWFC01. 154
Figure 67 : Structure cristallographique du RBD-B. 159
Figure 68 : Caractéristiques de la structure cristallographique du complexe RBD-B/AWFC01. 162
Figure 69 : Caractéristiques de la structure cristallographique du complexe aaRBD-B/AWFC01. 163
Figure 70 : Caractéristiques de la structure cristallographique du complexe RBD/CHENG. 164

Figure 71 : Caractéristiques de la structure cristallographique du complexe RBD-B/ZKO*..... 166

LISTE DES ANNEXES

Annexe 1 : Amorces nucléotidiques 208
Annexe 2 : Tampons utilisés lors de la purification de protéines recombinantes (RBD, GST-RBD et GST-NS1) 209
Annexe 3 : Comparaison des plages de lyse obtenues pour les virus WSN wt et mutants générés par génétique inverse 210
Annexe 4 : Comparaison des plages de lyse obtenues pour les virus WSN wt et mutants pour étudier la courbe de croissance virale 211
Annexe 5 : Cristallisation du RBD de NS1 pour le virus de chauve-souris H17N10 212
Annexe 6 : Comparaison des complexes formés par les différents RBD purifiés avec les sondes ARN DM01-midi et DM03-midi 213
Annexe 7 : Interactions indirectes impliquées dans les complexes RBD/CHENG, RBD-B/AWFC01, aaRBD-B/AWFC01 et RBD-B/ZKO* 214
Annexe 8 : Mouvements des nucléobases de l'ARN 215
Annexe 9 : Les paramètres hélicoïdaux de l'ARN 216
Annexe 10 : Les torsions du squelette ribose-phosphate de l'ARN 217
Annexe 11 : Largeur et profondeur des sillons de l'ARN 218
Annexe 12 : Liste des sondes oligonucléotidiques d'ARN utilisées 219
Annexe 13 : Structure and sequence determinants governing the interactions of RNAs with influenza A virus Non-structural protein NS1 221
Annexe 14 : Major contribution of the RNA-binding domain of NS1 in the pathogenicity and replication potential of an avian H7N1 influenza virus in chickens 247

INTRODUCTION

A. Classification, nomenclature, génomique, écologie et évolution des virus *influenza* A

A.I. Généralités

A.I.1. Classification

Les virus *influenza* appartiennent à la famille des *Orthomyxoviridae* et sont caractérisés par un génome segmenté associant de six à huit molécules d'ARN simple-brin de polarité négative (ARNv). Ces virus circulent chez de très nombreuses espèces animales – oiseaux et mammifères. Actuellement, 7 genres distincts ont été regroupés au sein de cette famille : les *Alphainfluenzavirus* (IAV), les *Betainfluenzavirus* (IBV), les *Gammainfluenzavirus* (ICV) et les *Deltainfluenzavirus* (IDV), les *Isavirus*, les *Quaranjavirus* et les *Thogotovirus*. Alors que le génome des IAV et des IBV comporte huit segments génomiques, les ICV et les IDV n'en possèdent que sept.

Les IAV infectent une grande variété d'oiseaux et de mammifères (pour revue, Kuchipudi et Nissly, 2018). Les IBV sont responsables de grippes saisonnières chez l'Homme mais ils ont aussi été isolés chez le phoque (Osterhaus *et al*., 2000). Les ICV, peu retrouvés dans le monde (peut-être parce qu'ils sont peu recherchés), sont responsables d'infections bénignes chez l'Homme et chez le porc. Enfin, les IDV seraient assujettis aux infections chez le bovin et le porc (Shaw et Palese, 2013; pour revue, Asha et Kumar, 2019).

En France, le réseau SENTINELLE (http://websenti.u707.jussieu.fr/sentiweb) estime que les IAV sont responsables de plusieurs millions d'infections respiratoires qui conduisent à environ 5000 décès chaque année en France. A l'échelle planétaire, l'Organisation Mondiale de la Santé (OMS) estime à plusieurs centaines de millions le nombre de cas d'infections grippales en période hivernale, conduisant lors d'infections respiratoires graves à un excès de mortalité compris entre 290 et 650 000.

Etant donné que mes travaux de recherche ont porté essentiellement sur la protéine non-structurale 1 (NS1) des virus *influenza* A, j'ai choisi de présenter dans mon introduction uniquement les IAV.

A.I.2. Morphologie et structure

Les IAV sont des virus pléomorphiques : les isolats naturels sont majoritairement sous forme filamenteuse (diamètre de 500 nm) alors que les isolats de laboratoire sont sous forme sphérique (diamètre de 100 nm) (Harris *et al*., 2006; Badham et Rossman, 2016; Dadonaite *et al*., 2016) (**Fig.1**).

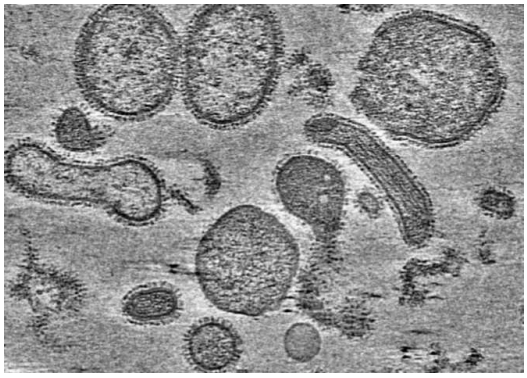


Figure 1 : Architecture des différentes formes d'IAV. *Les différentes formes de la souche virale A/Puerto Rico/8/1934 (H1N1) (PR8), observées par cryo-electron tomographie (Nayak et al., 2009).*

L'enveloppe lipidique qui compose l'enveloppe du virion est empruntée à la cellule lors du bourgeonnement viral. La surface de l'enveloppe virale porte environ 300 trimères d'hémagglutinine (HA) et 40 tétramères de neuraminidase (NA) (Harris *et al*., 2006). La face interne de l'enveloppe virale est tapissée par la protéine de matrice M1 qui interagit avec les ribonucléoprotéines virales (RNPv), les glycoprotéines HA et NA ainsi que la protéine M2. La protéine M2 est enchâssée dans la bicouche lipidique et forme un canal ionique. Les coupes fines de virus mettent en évidence la présence de huit segments génomiques sous la forme de RNPv de forme hélicoïdale. Chacune des huit RNP est composée d'un ARN génomique associé à de nombreuses nucléoprotéines (NP) et dont les extrémités 5' et 3' de l'ARN sont associées au complexe polymérase formé de la protéine acide (PA) et des protéines basiques 1 et 2 (respectivement PB1 et PB2) (**Fig.2**). Le plus petit segment génomique code pour les protéines non-structurales 1 et 2 (respectivement NS1 et NS2). NS1 n'est pas retrouvée dans le virion, ou alors en très faible quantité (Hutchinson *et al*., 2014), alors que de petites quantités de NS2 (plus connue sous le nom de NEP) y sont incorporées (pour revue, Hutchinson *et al*., 2014; Szewczyk *et al*., 2014).

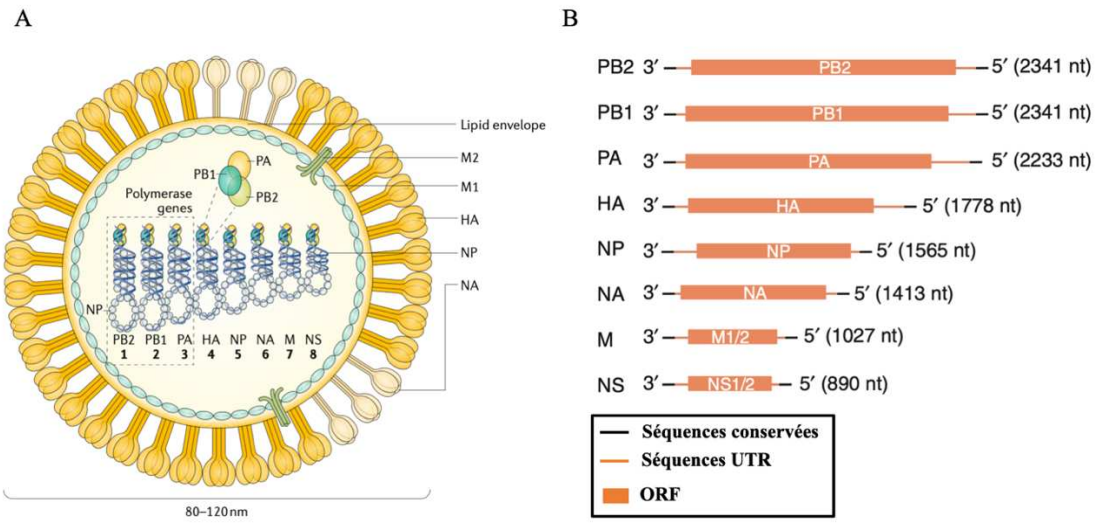


Figure 2 : Une particule virale d'IAV et son contenu génomique. *(**A**) Schéma d'une particule virale d'IAV (Krammer et al., 2018). (**B**) Représentation du génome viral des IAV (Fodor et Te Velthius, 2019).*

A.I.3. Nomenclature

Le système de nomenclature référence les IAV en fonction (i) du type viral (A, B, C ou D), (ii) de l'hôte d'origine dans le cas où l'échantillon n'aurait pas été isolé chez l'Homme, (iii) du lieu d'isolement, (iv) du numéro d'isolat, (v) de l'année d'isolement et (vi) des sous-types HA et NA du virus isolé. Ainsi, la souche virale référencée A/turkey/Italy/977/1999 (H7N1) provient d'un virus *influenza* A de sous-type H7N1 isolé en 1999 chez une dinde, en Italie (isolat N°977).

A.II. Génomique des virus *influenza* A

A.II.1. Organisation du génome viral

Le génome des IAV est composé de huit segments d'ARN simple brin de polarité négative (ARNv). La nomenclature officielle numérote les huit segments génomiques de 1 à 8, du plus grand (2341 nucléotides (nt) pour les segments PB2 et PB1) au plus petit segment (890 nt pour le segment NS) pour un total d'environ 13 kilobases (kb). Pour des raisons de commodité, les segments génomiques portent généralement le nom de la protéine majeure pour laquelle ils codent (**Tableau 1**). Plusieurs milliers de séquences sont disponibles dans la base de données *Influenza Virus Resource* (https://www.ncbi.nlm.nih.gov/genomes/FLU/Database/nph-select.cgi?go=database). Chaque segment présente le complémentaire d'une phase ouverte de lecture (ORF) bordée en 3' et 5' par de courtes régions transcrites non-traduites (UTR). Aux extrémités 3' et 5' des ARN viraux génomiques (ARNv) se trouvent respectivement les courtes séquences conservées U12 (12 nt) et U13 (13 nt) (**Fig.3A**), dont l'appariement en « manche de poêle » (panhandle) (**Fig.3B**) (Cheong *et al*., 1996; Fodor *et al*., 1994) constitue le promoteur de l'ARN polymérase ARN-dépendante virale (RdRp) (Pflug *et al*., 2014; Reich *et al*., 2014; pour revue, Te Velthuis et Fodor, 2016; Lo *et al*., 2018). Les courtes séquences U12-U13 peuvent aussi être repliées en forme de « tire-bouchon » (corkscrew) (**Fig.3B**) (Tomescu *et al*., 2014). Les modes de repliement « fourchette » (fork) et « crochet » (hook) (**Fig.3B**) des séquences U12-U13 ont également été proposés mais seule la forme *hook* est retrouvé dans les structures cristallographiques des complexes associant la RdRp virale aux ARNv ou aux ARN viraux complémentaires (ARNc) (Ferhadian *et al*., 2018).

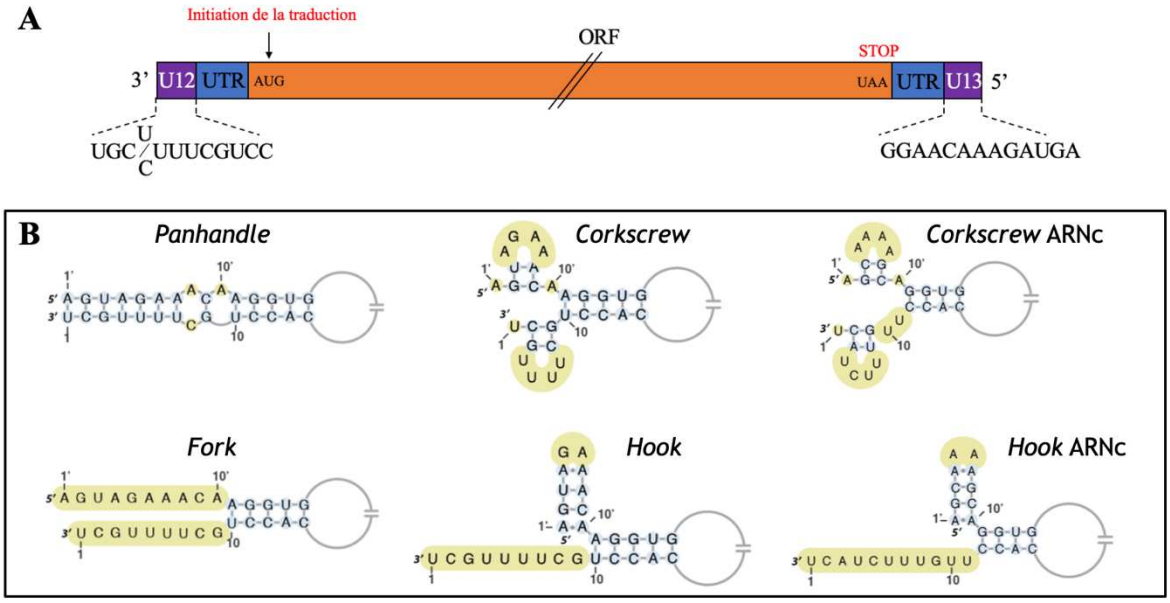


Figure 3 : Structure des segments d'ARN génomiques viraux d'IAV. *(**A**) Organisation d'un segment d'ARN génomique. L'ORF de chaque ARNv est bordée de séquences UTR ainsi que du promoteur viral formé par l'appariement des séquences strictement conservées U12 et U13. (**B**) Les différents repliements du promoteur viral des ARNv et des ARNc: manche de poêle (panhandle), tire-bouchon (corkscrew), fourchette (fork) et crochet (hook) (adapté de Ferhadian et al., 2018).*

A.II.2. Le protéome viral

Les IAV exploitent plusieurs mécanismes pour maximiser leur capacité codante. L'épissage alternatif, le décalage de cadre de lecture, le saut de phase de lecture ou encore la ré-initiation de traduction sont des mécanismes détournés par ces virus qui, à partir d'un segment d'ARN génomique viral, permettent d'obtenir plusieurs protéines (pour revue, Firth et Brierley, 2012). Tous les segments codent pour des protéines auxiliaires à l'exception des segments HA, NP et NA (Dubois *et al*., 2014). J'ai volontairement détaillé les grandes fonctions de NS1 dans le **Tableau 1**.

Tableau 1 : Le protéome viral des IAV

Segment	Longueur du segment (nt)	Protéine codée	ARN messager correspondant Mécanisme d'apparition	Longueur protéine (aa)	Fonction (s) principale(s)	Références
1-PB2	2341	PB2	ARNm non-épissé Traduction de AUG1	759	Interagit avec PA	Te Velthius et Fodor, 2016
					Fixe la coiffe des ARNm cellulaires	Guilligay et al., 2010
		PB2-S1	ARNm épissé (ARNm PB2) Traduction de AUG1 (délétion 1513-1894)	508	Inhibition des défenses cellulaires, adaptation des virus aviaires pour les mammifères	Yamayoshi et al., 2016
2-PB1	2341	PB1	ARNm non-épissé Traduction de AUG1	757	Interagit avec PA et PB2, assure l'activité l'ARN polymérase ARN-dépendante	Te Velthuis & Fodor, 2017
		PB1-F2	ARNm non-épissé Traduction de AUG4, décalage du cadre de lecture de PB1	90	Induit l'apoptose cellulaire	Cheng et al., 2017
					Active l'inflammation via la voie NLRP3	Pinar et al., 2016
					Inhibe la production d'IFN-I	Varga et al., 2012
					Influence l'activité ARN polymérase ARN-dépendante	Pena et al., 2012
		PB1-N40	ARNm non-épissé Traduction de AUG5 (glissement du ribosome)	718	Influence l'expression de PB1 et PB1-F2	Wise et al., 2009
					Interaction avec ARCN1	Wang et al., 2019
3-PA	2233	PA	ARNm non-épissé Traduction de AUG1	716	Clive la coiffe des ARNm cellulaires (endonucléase virale)	Dias et al., 2009 Yuan et al., 2009
		PA-X	ARNm non-épissé Traduction de AUG1, déphasage du ribosome (codon 190-193)	252	Activité endonucléase virale	Jagger et al., 2012
					Bloque l'activation de MAVS	Rigby et al., 2019
		PA-N155	ARNm non-épissé Traduction de AUG11, décalage du cadre de lecture	568	Interaction avec UBA52	Wang et al., 2018
		PA-N182	ARNm non-épissé Traduction de AUG13, décalage du cadre de lecture	535		
4-HA	1778	HA	ARNm non-épissé Traduction de AUG1	560	Fixation sur les acides sialiques et fusion des membranes	Blijleven et al., 2016
5-NP	1565	NP	ARNm non-épissé Traduction de AUG1	498	Composant majeur des RNP, contrôle le transport nucléo-cytoplasmique des RNP	Moreira et al., 2016
6-NA	1413	NA	ARNm non-épissé Traduction de AUG1	465	Clive la liaison acide sialique-HA et libère les virions infectieux	McAuley et al., 2019
7-M	1027	M1	ARNm non-épissé (ARNm1) Traduction de AUG1	252	Intervient dans la morphogénèse virale	Chen et al., 2008 Burleigh et al., 2005
		M2	ARNm épissé (ARNm2) Traduction de AUG1	97	Canal ionique à protons, permet l'acidification du contenu de la particule virale	To et Torres, 2019
		M42	ARNm épissé (ARNm4) Traduction de AUG1 de l'ARNm4	99	Assure les fonctions de M2 si virus délétés de M2	Wise et al., 2012
		NSP7	ARNm directement transcrit à partir du ARNv – (517 à 842)	?	?	Zhirnov, 2020
8-NS	890	NS1	ARNm non-épissé (ARNm1) Traduction de AUG1	217 à 230	Inhibe la réponse IFN-I	Schoggins, 2014
					Inhibe l'activité des ISG (PKR, OAS et ADAR)	Li et al., 2006 Min et Krug, 2006 Chassey et al., 2013
					Limite l'apoptose cellulaire	Ehrhardt et Ludwig, 2009 Ylösmaki et al., 2015 Terrier et al., 2013
					Réduit l'épissage des préARNm cellulaires	Nemeroff et al., 1992 Lu et al., 1994 Qiu et al., 1995
					Réduit la maturation des préARNm cellulaires	Hale et al., 2010 Dubois et al., 2019
					Réduit le transport nucléocytoplasmique des ARNm cellulaires	Satterly et al., 2007 Zhang et al., 2019
					Augmente la transcription et traduction virales	Enami et al., 1994 Park et Katze, 1995
					Interagit avec des ARN cellulaires	Yoshida et al., 1981 Lu et al., 1994 Someya et al., 2008
					Interagit avec des ARN viraux	Hatada et al., 1992 Hatada et Fukuda, 1992 Marc et al., 2013
		NS2 (NEP)	ARNm épissé (ARNm NS2) Traduction de AUG1	121	Assure le transport des RNP du noyau vers le cytoplasme	Paterson et Fodor, 2012
		NS3	ARNm épissé (ARNm NS3) (126-168) Traduction de AUG1	174	Adaptation à l'hôte et sauts de barrière d'espèces (hypothèse)	Selman et al., 2012
		NEG (NSP8)	ARNm directement transcrit à partir du ARN - (100 à 743)	167 (avant 1948) 216 (depuis 1948)	?	Clifford et al., 2009 Zhirnov et al., 2017 Zhirnov et Isaeva, 2019 Zhirnov, 2020

A.III. Ecologie et évolution

A.III.1. Ecologie

L'avifaune aquatique sauvage, notamment les Ansériformes (canards et oies) (Medina *et al*., 2011) et les Charadriiformes (pluviers et gravelots) (Munster *et al*., 2007), constitue le réservoir principal des IAV (Munster et Fouchier, 2009) (**Fig.4**). La promiscuité entre ces espèces dont beaucoup sont migratrices permet les échanges de segments génomiques entre deux virus infectant une même cellule hôte, conduisant ainsi à l'émergence de virus réassortants. Les flux d'oiseaux migrateurs participent activement à la circulation et à la transmission des IAV à l'échelle planétaire (Venkatesh *et al*., 2018). Pratiquement tous les sous-types théoriquement possibles ont été retrouvés chez les oiseaux sauvages (**Fig.4**) (Medina *et al*., 2011; Munster *et al*., 2007; 2009). Ces oiseaux sauvages aquatiques peuvent tout autant contaminer des oiseaux d'élevage que d'autres espèces animales comme le cheval, le porc ou l'Homme (**Fig.4**). Deux nouveaux sous-types viraux, H17N10 et H18N11, ont récemment été isolés chez des virus de chauve-souris (Tong *et al*., 2012; 2013).

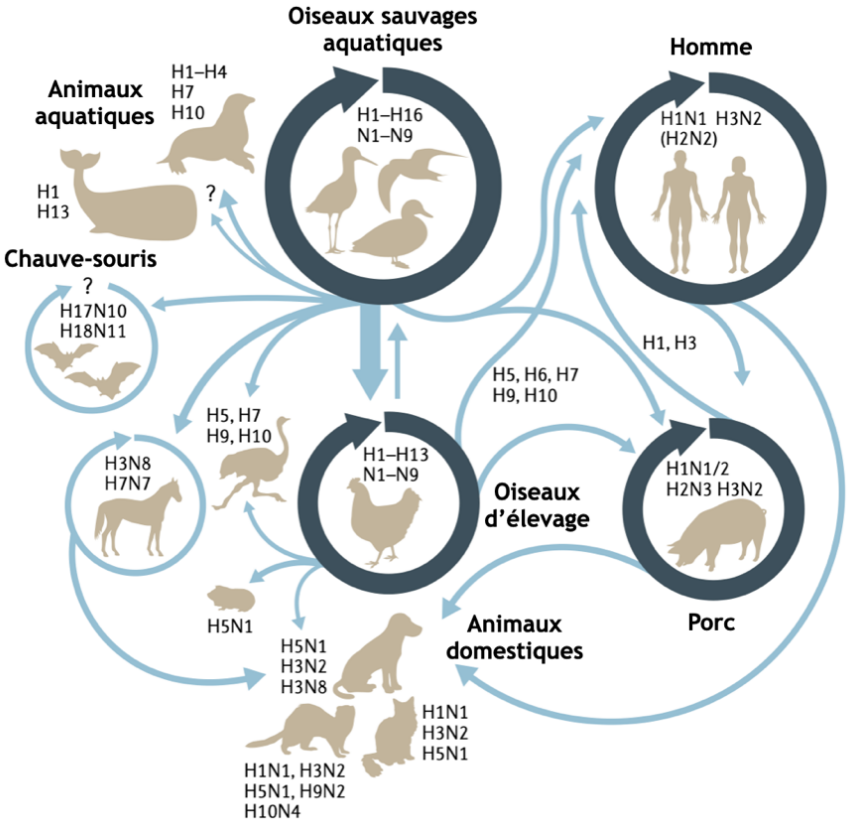


Figure 4 : Transmissions intra et inter-espèces des différents sous-types d'IAV. *Les flèches en bleu clair indiquent les transmissions directes avérées entre espèces animales. Les flèches en bleu foncé indiquent les transmissions au sein d'une même espèce (Homme, porc, oiseaux sauvages aquatiques et oiseaux d'élevage). Les points d'interrogation indiquent une origine de transmission inconnue (adapté de Long et al., 2019).*

A.III.2. Evolution

L'évolution des IAV est la résultante de deux mécanismes distincts : un taux de mutation élevé lié à la faible fidélité de la RdRp virale et les échanges fréquents de segments génomiques entre virus (réassortiments).

A.III.2.a. La dérive antigénique ou *antigenic drift*

Ne possédant ni activité de correction (*proof reading*) et ni activité de réparation des erreurs incorporées dans le génome viral lors de sa réplication, la RdRp virale introduit des mutations à chaque cycle de réplication du virus. Son taux de mutation, estimé entre 10^{-6} et 10^{-4} mutations par nucléotide et par cellule infectée (Pauly *et al*., 2017), entraîne la dérive antigénique (Hanada *et al*., 2004), notamment en raison de la pression de sélection induite par la réponse immunitaire de l'hôte. Les mutations dans l'une de ces deux protéines, qui modifient les épitopes immunogènes, permettent au virus d'échapper aux anticorps neutralisants produits en réponse à l'infection virale ou suite à une vaccination (Tewawong *et al*., 2015).

A.III.2.b. Cassure antigénique ou *antigenic shift*

La cassure antigénique ou *antigenic shift* se produit lorsque deux virus infectant une même cellule hôte s'échangent des segments génomiques. L'étape tardive d'assemblage des particules virales, au cours de laquelle des échanges d'un ou plusieurs segments d'ARN génomiques peuvent avoir lieu entre les deux virus, conditionne la production de virus réassortants (**Fig.5**). Alors que la dérive antigénique concerne tous les virus *influenza*, il semble que la cassure antigénique concerne uniquement les IAV. Le virus aviaire H7N9 apparu récemment en Asie est issu du réassortiment entre plusieurs virus aviaires d'oiseaux domestiques (canards et poules) et d'oiseaux sauvages (**Fig.5**).

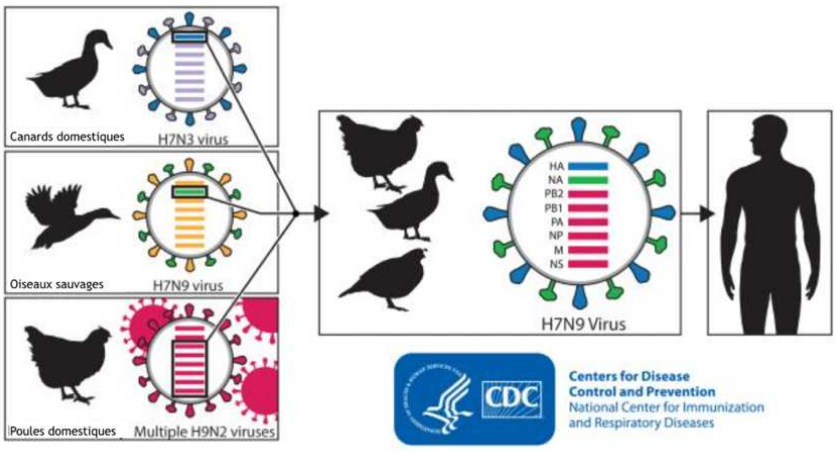


Figure 5 : Réassortiment génomique du virus aviaire asiatique H7N9. *Le virus H7N9 récemment apparu en Asie (2013) est issu du réassortiment entre plusieurs virus aviaires provenant d'oiseaux sauvages et domestiques. Le segment HA provient d'un virus aviaire H7N3 de canards domestiques, NA de virus aviaire H7N9 d'oiseaux sauvages et les six autres segments de virus aviaires H9N2 de poules domestiques (CDC).*

A.III.3. Les virus grippaux humains et leurs réassortiments

Une quinzaine de pandémies d'IAV ont été dénombrées chez l'Homme depuis la fin du 18^{e} siècle, dont quatre depuis le début du 20^{e} siècle (1918, 1957, 1968 et 2009). Les deux IAV qui, avec les IBV, causent les grippes saisonnières humaines actuelles, sont les virus A H1N1 et H3N2, descendant respectivement des virus pandémiques pdm09 et de la grippe de Hong-Kong (**Fig.6**).

La grippe Espagnole (H1N1)

La grippe Espagnole (*Spanish Flu*) de 1918, due à un virus H1N1, a été la pandémie la plus meurtrière de l'histoire avec entre 50 et 100 millions de morts (Johnson et Mueller, 2002). Les fragments de génomes viraux extraits des tissus de patients infectés décédés ont permis de reconstituer le génome du virus (Tumpey *et al*., 2005). Ce dernier aurait pour origine un virus aviaire ancestral (Rabadan *et al*., 2006), mais son origine exacte ainsi que sa date d'émergence chez l'Homme restent inconnues à ce jour (Smith *et al*., 2009). Ce virus H1N1 s'est par la suite maintenu pendant près de 40 ans au travers de grippes saisonnières avant de disparaitre en 1957 laissant place à un nouveau virus pandémique. Réapparu en 1977, le virus H1N1 a continué à circuler jusqu'en 2009 (**Fig.6**).

La grippe asiatique (H2N2)

Apparu en 1957 en Asie, le virus H2N2 a circulé sous forme de grippes saisonnières jusqu'en 1968. Ce virus est un descendant du virus H1N1, dont il a conservé cinq segments (PB2, PA, NP, M et NS) (Kawaoka *et al*., 1989), les trois autres (PB1, HA et NA) provenant d'un virus aviaire H2N2 inconnu (**Fig.6**).

La grippe de Hong-Kong (H3N2)

Responsable de la pandémie de Hong-Kong en 1968, le virus H3N2 est un descendant du virus H2N2 (donc du virus H1N1 de 1918), dont il a conservé six segments génomiques (PB2, PA, NA, NP, M et NS). Les segments PB1 et HA proviennent d'un virus H3 aviaire. Le virus H3N2 circule toujours actuellement sous forme de grippe saisonnière (**Fig.6**).

La grippe pandémique de 2009 (pdm2009, H1N1)

En 2009 apparaît un nouveau virus H1N1 pandémique (aussi appelé grippe porcine mexicaine), qui remplace l'ancien virus saisonnier H1N1 et circule désormais sous forme de grippe saisonnière. Le virus H1N1 (pdm2009), présente une origine multiple (Smith *et al*., 2009) : PB1 provient du virus humain H3N2, PB2 et PA de virus aviaires américains, NA et M de virus porcins eurasiatiques et NP, HA et NS du virus porcin américain. L'émergence du virus pandémique de 2009 est donc la conséquence du réassortiment de virus humain, porcins et aviaire (**Fig.6**).

Prochaines pandémies ?

Le triangle de transmission hommes – porcs – oiseaux illustre la dangerosité des virus aviaires dans l'émergence de nouveaux virus. Actuellement, on peut craindre que les virus aviaires H5N1 et H7N9 deviennent un jour pandémique. Depuis 2003, le virus H5N1 a contaminé plus de 830 personnes pour plus de 440 décès. Apparu en Asie en 2013, le virus H7N9 (**Fig.5**) a été responsable de plus de 1600 cas humains, entrainant plus de 600 décès (**Fig.6**).

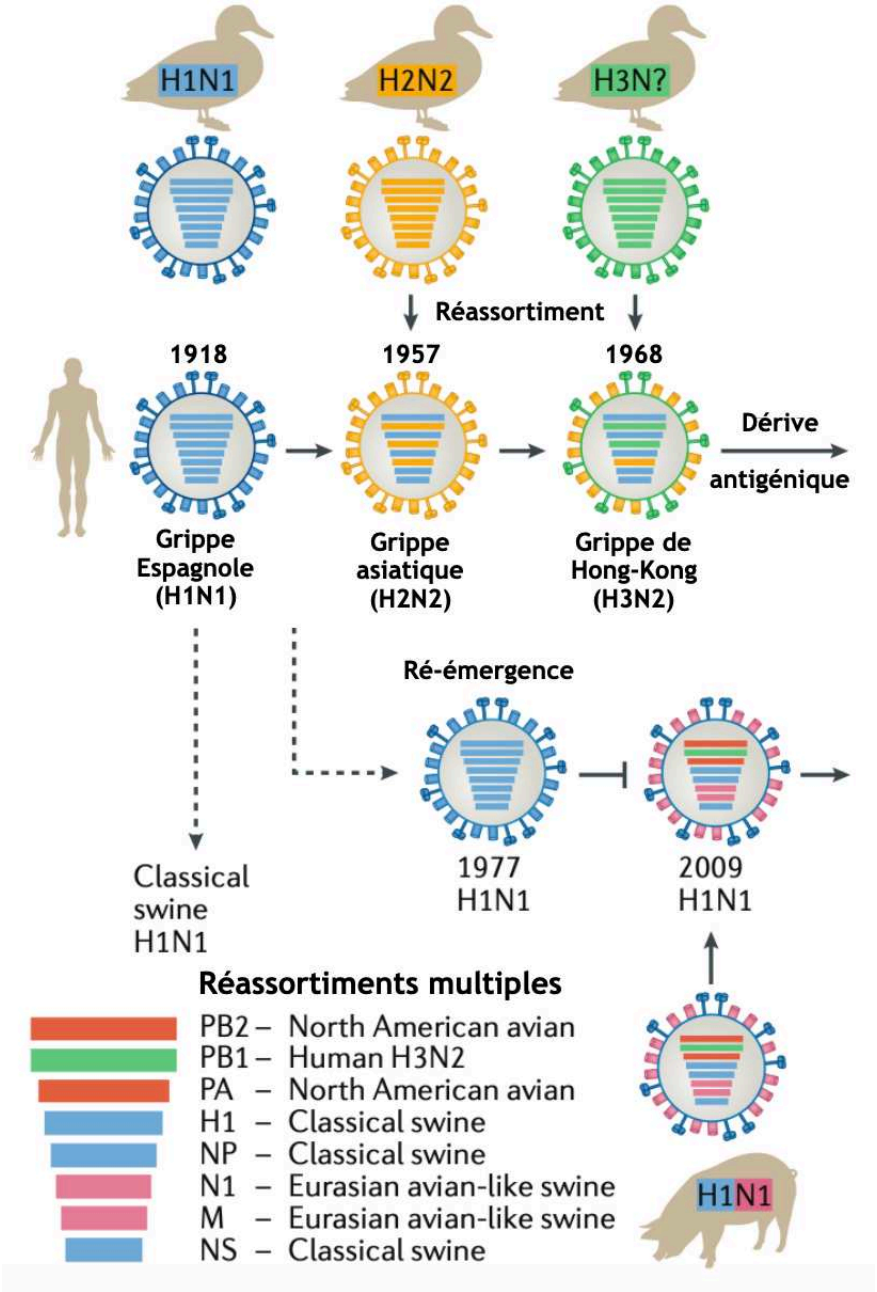


Figure 6 : Evolution et émergence des IAV humains depuis le début du 20ème siècle. *L'évolution de chacun des huit segments d'ARN génomiques est représentée. Les sous-types viraux ainsi que les dates d'apparition sont indiqués (adapté de Long et al., 2019).*

B. Cycle viral des virus *influenza* A

B.I. Les grandes étapes du cycle viral

Le cycle viral des IAV peut être résumé en six grandes étapes : l'attachement et l'entrée du virus dans la cellule, la libération cytoplasmique des ribonucléoprotéines virales (RNPv) et leur import nucléaire, la transcription et la réplication du génome viral, l'export nucléaire des RNPv néo-synthétisées et l'assemblage des particules virales (**Fig.7**). J'ai choisi de présenter dans le détail chacune de ces étapes.

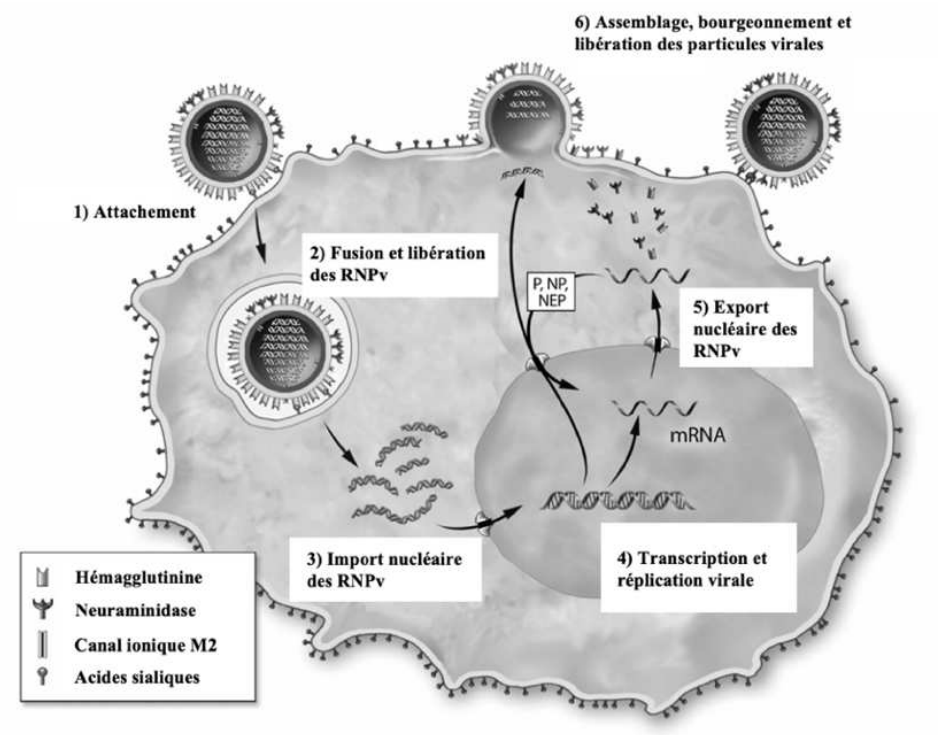


Figure 7 : Schéma simplifié du cycle viral des IAV *(adapté de Basler et Aguilar, 2008)*

B.II. Attachement et entrée du virus dans la cellule hôte

Les hémagglutinines (HA) des particules virales inhalées se fixent aux acides sialiques (SA) (N-acétylneuraminiques) présents en surface des cellules du tractus respiratoire. Les SA sont composés d'une longue chaine glycosylée terminée par un résidu de galactose, les deux étant reliés par une liaison de type alpha 2,3 (α-2,3) ou une liaison alpha 2,6 (α-2,6) (**Fig.8**).

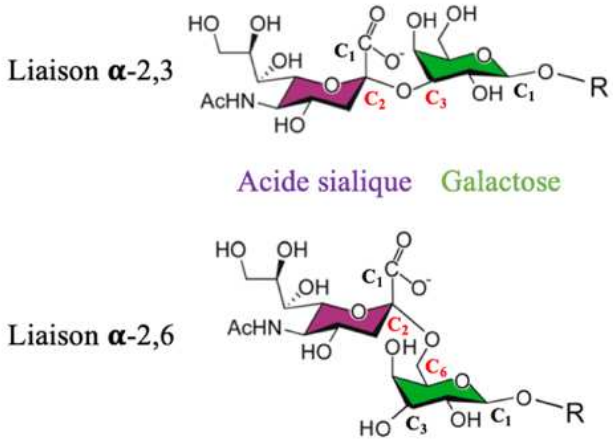


Figure 8 : Structure des acides sialiques *(adapté de Byrd-Leotis et al., 2017)*

La protéine HA est un déterminant majeur du spectre d'hôte et de la virulence des IAV. En premier lieu, sa tête globulaire porte le site d'attachement au récepteur (Receptor Binding Site, RBS), lequel détermine la préférence pour l'un ou l'autre des deux types d'acides sialiques. Les protéines HA de virus aviaires reconnaissent le motif α-2,3 (**Fig.9**), lequel est majoritairement retrouvé au niveau de l'épithélium intestinal des oiseaux sauvages et domestiques, et de façon plus minoritaire au niveau des bronchioles et des alvéoles pulmonaires chez l'Homme (voies respiratoires basses) (Shinya et Kawaoka, 2006). En revanche, les protéines HA de virus de mammifères reconnaissent le motif α-2,6 (**Fig.8**), lequel prédomine au niveau de l'épithélium respiratoire supérieur chez l'Homme (Nicholls *et al*., 2008). C'est certainement pour cette raison que l'Homme est moins susceptible à l'infection par des virus aviaires, même si le risque existe (Shinya *et al*., 2006). Le porc présente les deux motifs en quantités plus ou moins équivalentes au niveau de l'épithélium trachéal, le rendant susceptible à la fois aux virus porcins, humains et aviaires. D'ailleurs, il semble que la plupart des virus pandémiques du 20eme siècle ont émergé grâce à des réassortiments ayant eu lieu chez le porc (Smith *et al*., 2009).

Pour être fonctionnelle et permettre son changement de conformation, la protéine HA0 doit être clivée en HA1 et HA2. Ce sont généralement des protéases extracellulaires de type trypsine qui réalisent ce clivage, au site […KQTR|GLFGAIA…] (**Fig.9D**) que l'on peut qualifier de « monobasique ». Les virus aviaires de sous-type H5 et H7 peuvent présenter un site de clivage dit « polybasique » dans lequel plusieurs acides aminés basiques précèdent le site de clivage (par exemple Q**RERRRKKR**|GLFGAIA) (**Fig.9D**), assurant un clivage facilité de la protéine HA0 par de très nombreuses protéases intracellulaires, ce qui traduit la haute fonctionnalité de la protéine HA et augmente considérablement l'infectiosité de la particule virale. L'apparition de ce site dit « polybasique » survient lorsque la RdRp virale réplique le segment HA de ces virus aviaires où elle peut introduire des nucléotides supplémentaires, en particulier des adénines (A) qui constituent alors des codons additionnels pour des acides aminés basiques comme la lysine (K) (Nao *et al*., 2017). Les pathotypes viraux aviaires dits « Hautement Pathogène » (HPAIV), correspondant aux sous-types viraux H5 et H7, présentes un site

dit « polybasique » dans leur protéine HA alors que les « Faiblement Pathogène » (LPAIV) n'en possèdent pas.

Le virus entre dans la cellule en approximativement 10 minutes (Dou *et al*., 2017) par un mécanisme (i) d'endocytose clathrine dépendante (Matlin *et al*., 1981), (ii) d'endocytose cavéoline dépendante (Lakadamyali *et al*., 2004), (iii) d'endocytose indépendante de la clathrine et de la cavéoline (Sieczkarski et Whittaker, 2002) ou de (iv) macropinocytose (Rossman *et al*., 2012).

B.III. Fusion membranaire de l'enveloppe virale et cellulaire, libération cytoplasmique et import nucléaire des ribonucléoprotéines virales

L'acidification de l'endosome (pH~5) induit l'ouverture du canal ionique M2 ce qui entraine l'acidification du contenu de la particule virale (Cady *et al*., 2009), induisant d'un côté la dissociation des complexes associant les RNPv et les protéines de matrice M1 et d'un autre un changement conformationnel des trimères de HA (**Fig.9A**). Les protéines HA2 se déploient alors, ce qui permet à leurs peptides de fusion (indiqués en bleu sur la séquence peptidique, **Fig.9D**) de s'enchâsser dans la membrane de l'endosome (**Fig.9B et 9C**), entrainant la fusion de l'enveloppe virale et de la membrane endosomale suite à leur rapprochement physique et la formation d'un pore au travers duquel les RNPv sont transportées dans le cytoplasme de la cellule infectée (**Fig.10**).

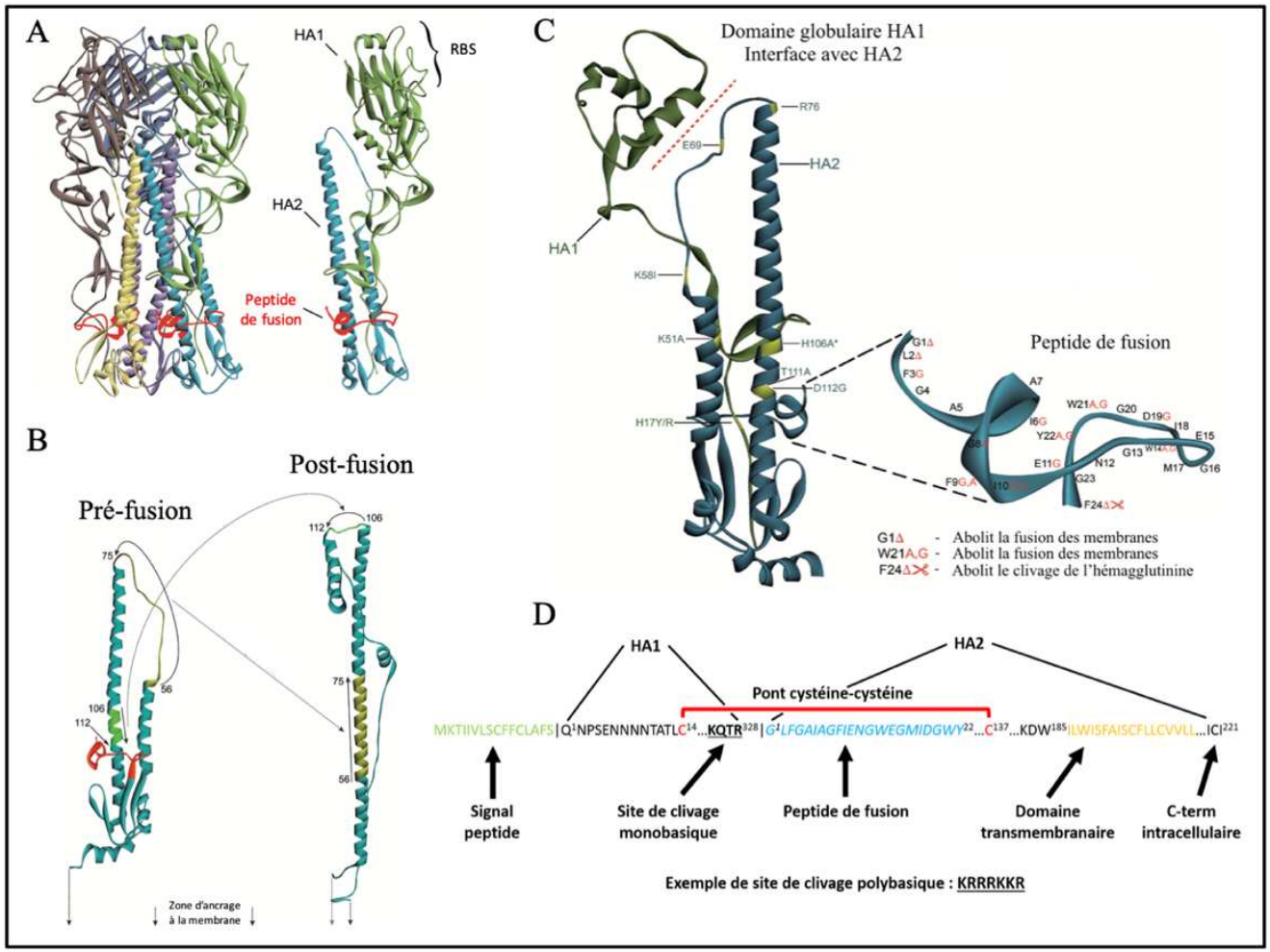


Figure 9 : Structure de l'hémagglutinine. *(**A**) Structure du trimère de HA. Le trimère globulaire de HA1 est représenté en vert, en violet et en marron. Le trimère tige de HA2 est représenté en jaune, en bleu et en rose. Le domaine globulaire essentiellement formé par HA1 porte le site de liaison au récepteur (RBS). Le peptide de fusion indiqué en rouge est constitué par l'extrémité N-terminale de HA2. (**B**) Structures de HA2 au cours des étapes pré et post-fusion. (**C**) Structure du peptide de fusion (adapté de Jakubcová et al., 2016). (**D**) Séquence peptidique de l'hémagglutinine. L'exemple cité ci-dessous correspond à la séquence peptidique de l'hémagglutinine du virus A/American black duck/New Brunswick/02650/2007 (H3N2). Les sites correspondant au peptide signal, au site de clivage monobasique, au peptide de fusion, au domaine transmembranaire ainsi qu'à la région C-terminale sont indiqués.*

Une durée d'environ 1 h semble nécessaire au transport nucléaire des RNPv (Dou *et al*., 2017) qui est possible grâce à la présence de signaux de localisation nucléaire (Nuclear Localisation Signal, NLS) dans les nucléoprotéines (NP) et dans PB2 des RNPv (Delaforge *et al*., 2015; Swale *et al*., 2016). Trop volumineuses (100 à 200 Å) pour diffuser de façon autonome au travers des pores nucléaires (9Å), les RNPv nécessitent l'intervention du complexe d'importines α et β pour y être transloquées. Ces complexes (RNPv/importines α et β) se dissocient lorsque l'importine β interagit avec la forme active de la protéine G cellulaire Ran (Ran-GTP) (**Fig.10**) (pour revue, Dou *et al*., 2018).

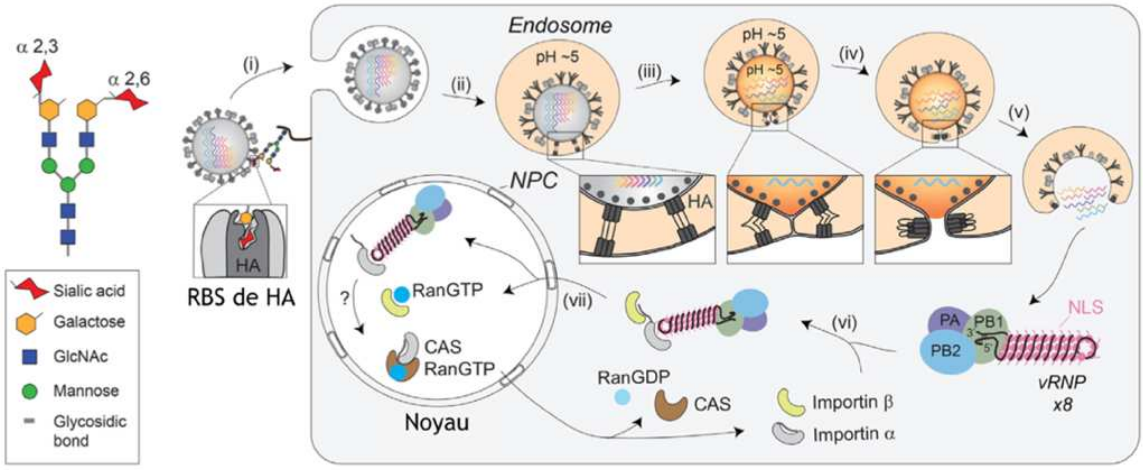


Figure 10 : La fusion de l'enveloppe virale et de la membrane endosomale. *L'acidification de l'endosome entraine le changement conformationnel de HA, exposant le peptide de fusion de HA2. HA2 se déplie et perfore l'enveloppe de l'endosome. HA2 commence à se replier ce qui rapproche la membrane endosomale de l'enveloppe virale. La membrane et l'enveloppe fusionnent sous la contrainte de HA2. La fusion complète conduit à la formation d'un pore (adapté de Dou et al., 2018).*

B.IV. Transcription et réplication du génome viral

B.IV.1. La ribonucléoprotéine virale

Dans la ribonucléoprotéine virale (RNPv), l'ARN polymérase ARN-dépendante (RdRp) virale est liée aux séquences conservées U12 et U13, dont l'appariement constitue le promoteur de la polymérase du virus (**Fig.11**) (Noda *et al*., 2012; pour revue, Fodor et Te Velthius, 2019).

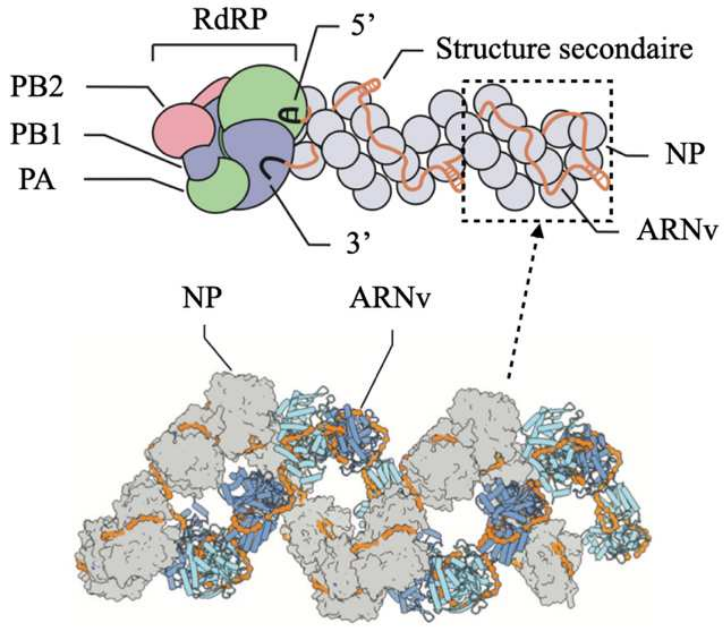


Figure 11 : Une ribonucléoprotéine virale. *Représentation schématique (haut) de la structure d'une RNPv à partir de la structure 3D obtenue par cryo-microscopie électronique (bas) (PDB : 4BBL, Arranz et al., 2012). L'ARN viral génomique (orange) est associé à une multitude de protéines NP et à un complexe RdRp viral (adapté de Fodor et te Velthius, 2019).*

B.IV.2. L'ARN polymérase ARN-dépendante virale

La RdRp virale est un hétérotrimère de 250 kDa composé de PA, PB1 et PB2 dans lequel l'activité polymérase est portée par PB1 qui siège au centre du complexe RdRp viral (**Fig.12A**) et est située au carrefour (i) de l'entrée de l'ARN matrice, (ii) de l'entrée des nucléosides triphosphates (NTP), (iii) de la sortie des produits de clivage et (iv) de la sortie de l'ARNv néo-synthétisé (**Fig.12B**) (pour revue, te Velthius et Fodor, 2016).

Reich *et al*. (2014) et Pflug *et al*. (2014; 2017) ont résolu les premières structures complètes du complexe RdRp viral d'un IAV (PDB : 4WSB) et d'un IBV (PDB : 5MSG et 4WRT) à partir de virus de chauve-souris H17N10 et de virus humain respectivement. Depuis, de nouvelles structures de la RdRp virale, y compris celles d'ICV et d'IDV, ont été résolues par diffraction des rayons X (DRX) et par cryo-microscopie électronique (cryo-EM) (Fan *et al*., 2019; Hengrung *et al*., 2015; Peng *et al*., 2019). L'analyse détaillée des structures de ces RdRP virales d'IAV a mis en évidence des changements de conformation dans le complexe, en particulier dans la région N-terminale de PA (PA-N) ainsi que dans la région C-terminale de PB2 (PB2-C). Ceux-ci pourraient survenir lorsque la RdRp virale passe de l'activité de transcription à l'activité de réplication. Le complexe RdRp viral montre globalement une structure en forme de U (115 x 100 x 75 Å) (**Fig.12**, ici la RdRp est penchée sur le côté) dont le domaine endonucléase de PA (PA-endo) et le domaine cap-binding (CAP-BD) de PB2 forment les deux bras, le corps étant en grande partie constitué de PB1 (Pflug *et al*., 2014) et du domaine C-terminal de PA (PA-C).

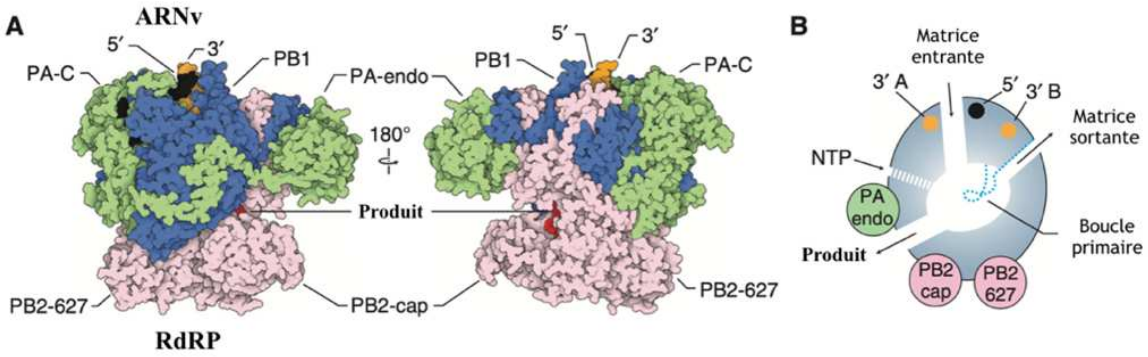


Figure 12 : Structure de l'ARN polymérase ARN-dépendante virale. *(A) Structure de la RdRp liée aux séquences terminales 5' et 3' de l'ARNv (PDB : 6RR7, Fan et al., 2019). PB1 est représentée en bleue, PB2 en rose et PA en vert. (**B**) Modèle schématique de la chambre catalytique de la RdRp virale, composée des canaux "entrée de la matrice", "sortie de la matrice", "entrée de NTP" et "sortie des produits de synthèse (adapté de Fodor et Te Velthius, 2019).*

B.IV.2.a. La protéine acide

La protéine acide (PA, 716 acides aminés) de la RdRp virale comporte deux domaines : le petit domaine N-terminal (PA-endo, acides aminés 1-196), portant l'activité endonucléase, qui est relié par un linker d'environ 60 acides aminés au large domaine C-terminal (PA-C, acides aminés 258-714), lequel est situé sous PB1 et forme le fond du U (**Fig.12A**). PA est l'endonucléase qui coupe à l'extrémité 5' de la coiffe des ARNm cellulaires, fournissant ainsi l'amorce nécessaire à la transcription des ARNm viraux.

La cavité du domaine PA-endo heberge l'activité endonucléase qui assure le clivage de la coiffe des ARNm cellulaires (**Fig.12A et 14**) (Dias *et al*., 2009; Yuan *et al*., 2009; Zhao *et al*., 2009). Dans cette poche, la fixation du phosphate et du ribose de l'ARN est assurée par plusieurs acides aminés hydrophiles et hydrophobes, respectivement. D'autres acides aminés de nature hydrophobes sont impliqués dans la reconnaissance directe des bases de l'ARN. Même si la région 50-72 du domaine PA-endo n'impacte ni le repliement la protéine ni son activité endonucléase (Dubois *et al*., 2012), elle semble néanmoins nécessaire à l'interaction avec PB1 et PB2 (Nilsson-Payant *et al*., 2018).

PA-C, qui est le domaine le plus large de PA, est subdivisé en deux régions, appelées la région *cerveau* et la région *bouche* (He *et al*., 2008; Obayashi *et al*., 2008). La première se compose de plusieurs acides aminés basiques alignés qui interagissent avec les ARN viraux. La seconde interagit à la façon d'une mâchoire avec les 15 premiers acides aminés de PB1, incluant le motif hydrophobe ^{7}LLFL^{10}. Contrairement à PA-N, PA-C a une conformation très stable dans le complexe, y compris lors des changements de conformation du complexe entier.

B.IV.2.b. La protéine basique 1

Longtemps restée inconnue en raison de sa faible stabilité en solution, la structure de la protéine basique 1 (PB1) a été finalement résolue en association avec PA et/ou à PB2 (He *et al*., 2008; Reich et al., 2014). La région centrale de PB1 (acides aminés 21-669) montre une organisation typique de

polymérase, avec des domaines *doigts*, *paume* et *pouce*. Les régions N- et C-terminales de PB1 entretiennent des interactions avec les deux autres protéines du complexe RdRp viral, en particulier le motif hydrophobe ^{7}LLFL^{10} de PB1-N qui est lié au domaine *bouche* hebergé par PA-C (Wunderlich *et al*., 2011) et par les trois hélices alpha (α) de la région C-terminale de PB1 (PB1-C, acides aminés 678-757) qui interagissent avec la région N-terminale de PB2 (PB2-N) (Sugiyama *et al*., 2009). L'activité polymérase virale de PB1 est directement dépendante des acides aminés D305/D445/D446 du site actif de la RdRp virale (Pflug *et al*., 2014) au sein duquel deux ions métalliques conditionnent l'activité polymérase. Le motif ^{406}GMMMGMF^{412} stabilise l'appariement des NTP entrants sur la matrice ARN. Les acides aminés basiques K480 et K481 de PB1, qui se lient aux NTP, sont stabilisés par des contacts avec PA. Un tunnel bordé d'acides aminés chargés, conduisant à la cavité interne de PB1, permettrait l'entrée des NTP au sein de PB1 alors qu'un second tunnel permettrait l'entrée de l'ARN matrice (**Fig.12B**).

B.IV.2.c. La protéine basique 2

La région N-terminale (PB2-N, acides aminés 1-247) de la protéine basique 2 (PB2, 759 acides aminés) qui interagit avec PB1 contient un domaine *couvercle* (acides aminés 155-212) qui recouvre les canaux de fixation de l'ARN matrice et des amorces coiffées (**Fig.12**). La région C-terminale (PB2-C, acides aminés 316-749) contient le site de fixation de la coiffe méthylée des ARNm (CAP-BD), positionné en face du domaine endonucléase de PA (**Fig.12**) (Liu *et al*., 2013; Tsurumura *et al*., 2013). Des essais d'interaction de PB2 avec la m^{7}GTP, un analogue de la coiffe des ARNs, montrent que cette coiffe méthylée se fixe à la base de cette structure protéique, en faisant intervenir plusieurs acides aminés chargés.

B.IV.2.d. La nucléoprotéine

La nucléoprotéine (NP, 498 acides aminés, 56 kDa) est l'une des protéines virales la plus abondamment exprimée dans la cellule infectée. Elle contient de très nombreux acides aminés basiques qui interagissent avec les phosphates et les riboses des ARN viraux génomiques et complémentaires (respectivement ARNv et ARNc). NP module le trafic nucléaire des RNPv (Digard *et al*., 1999) en interagissant avec les sous-unités PB1 - PB2 (Coloma *et al*., 2009) et PA (Arranz *et al*., 2012) contenues dans les RNPv.

L'ARNv s'enroule autour de NP à raison d'une molécule de NP tous les 15 à 24 nt (Ortega *et al*., 2000) en se liant à sa région N-terminale, localisée entre les domaines *tête* (head) et *corps* (body) de NP (Ng *et al*., 2008; Ye *et al*., 2006). NP s'oligomérise lorsqu'elle est associée aux ARNv (**Fig.13A, 13B et 13C**) (Labaronne *et al*., 2016), sa boucle flexible C-terminale interagissant avec des acides aminés conservés (E339 et R416) du domaine *corps* d'une NP adjacente (Arranz *et al*., 2012; Moeller *et al*., 2012). Une délétion, même partielle, de cette boucle (Δ 402-428), enraye le processus d'oligomérisation de NP (Ye *et al*., 2012).

Labaronne *et al*. (2016) ont montré que la protéine NP était sous forme monomérique en présence de faible concentration en sels (50 mmol.L^{-1} NaCl) et sous forme trimérique lorsque la concentration en NaCl augmente (150 et 300 mmol.L^{-1}) (**Fig.13D**). L'interaction avec des ARN de 8 nt entraine l'oligomérisation de NP lorsque celle-ci est purifiée avec 50 mmol.L^{-1} NaCl (**Fig.13D**). De façon surprenante, l'incubation de NP avec 150 mmol.L^{-1} NaCl et un ARN de 12 nt entraine la formation de structures similaires aux RNPv formées au cours du cycle viral (**Fig.13D**). Différemment, l'incubation de NP purifiée avec 300 mmol.L^{-1} NaCl et un ARN d'au moins 8 nt entraine la formation de dimères, de tétramères, de pentamères et de structures plus complexes.

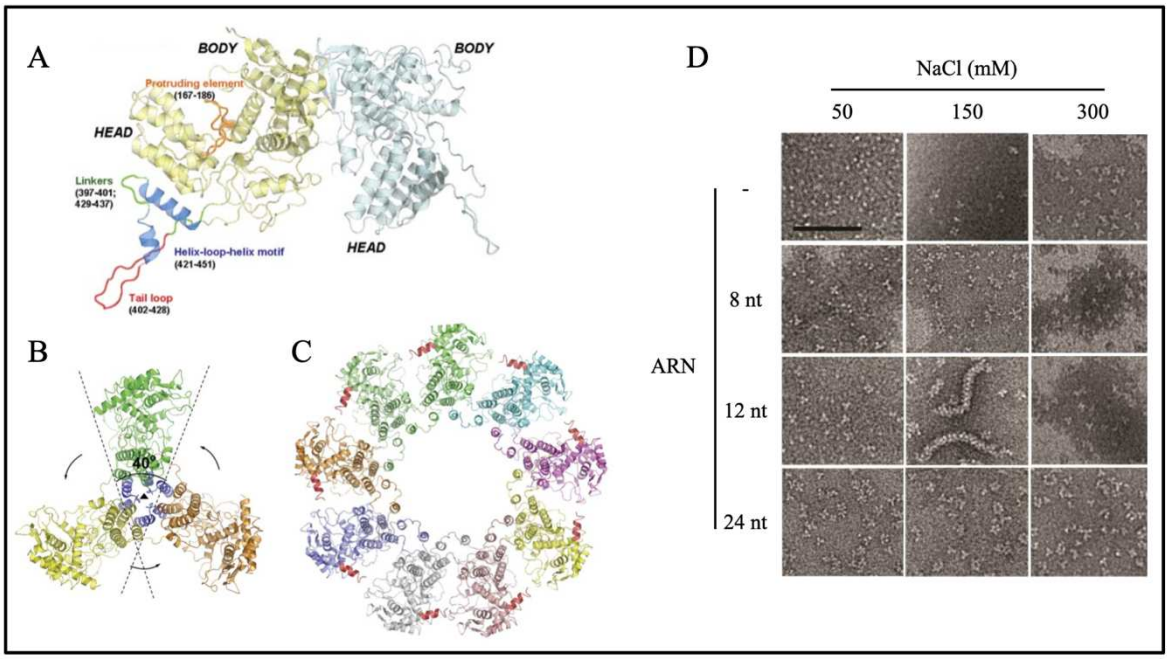


Figure 13 : Structure de la nucléoprotéine. *Représentation par PyMol de la structure de NP (PDB : 2Q06, Ng et al., 2008), en mode cartoon (homodimère, **A**; trimère, **B**; cercle associant 9 NP, **C**). (**A**) Un monomère de NP. Une unité asymétrique de NP contient deux monomères, où chacun d'eux est constitué des domaines tête (head) et corps (body) séparés par une région protubérante (orange). Un boucle externe (rouge) située en C-terminale de la protéine assure l'oligomérisation des protéines NP. NP peut former différents trimères de protéines en raison des différents contacts établis par les motifs hélice-boucle-hélice (bleus) entre les protomères de NP, l'ensemble étant possible grâce aux linkers (verts). (**B**) NP sous forme trimérique. Chaque NP présente dans le trimère protéique établit un contact avec les autres protomères par l'intermédiaire de la boucle C-terminale (bleue). (**C**) Modèle du cercle de NP associant 9 molécules protéiques (Martín-Benito et al., 2001). La région N-terminale de NP est représentée en rouge. (**D**) Clichés électromicroscopiques de NP (A/WSN/1933/H1N1) purifiées avec 50, 150 et 300 mmol.L^{-1} en NaCl en absence et en présence d'ARN de longueur variable (Labaronne et al., 2016).*

B.IV.3. La transcription virale

La transcription des ARNm viraux intervient en amont de l'étape de réplication du génome viral. Cette dernière nécessite le vol de coiffe ou *cap-snatching* des ARNm et/ou des pré-ARNm cellulaires pour permettre la synthèse des ARNm viraux.

B.IV.3.a. Le vol de coiffe ou *cap-snatching*

Le CAP-BD de PB2 fixe la coiffe méthylée des pré-ARNm cellulaires en cours de transcription par l'ARN polymérase II cellulaire (Pol II) (**Fig.14**) suite à l'interaction de PA-C avec la région C-terminale de Pol II (Bauer *et al*., 2018; Lukarska *et al*., 2017; De Vlugt *et al*., 2018; pour revue, Walker et Fodor, 2019). L'interaction de la RdRp virale avec Pol II entraine l'arrêt de la transcription des gènes cellulaires (Zhao *et al*., 2018). L'activité endonucléase de PA-N qui lui permet de cliver les pré-ARNm cellulaires et les transcrits maturés à environ 13 nt en aval de leur coiffe méthylée (aussi appelé vol de coiffe ou *cap-snatching*) (**Fig.14**), réduit la distance qui sépare PA-N de PB2-N (Reich *et al*., 2014) et semble également influencer la préférence du site de clivage ($AdGu^{13}Cy$) (Koppstein *et al*., 2015). Le *cap-snatching* s'opère aux dépens des ARNs les plus abondants, en particulier le petit snARN U2 qui est coupé en 11^{eme} ou 12^{eme} position ($CyUrCy^{11}$ ou $CyUrCyGu^{12}$) (Gu *et al*., 2015; Koppstein *et al*., 2015; Sikora *et al*., 2014; 2017 ; Wandzik *et al*., 2020).

B.IV.3.b. Initiation, élongation et polyadénylation

Suite au vol de la coiffe, PB2 change de conformation ce qui permet l'insertion de la courte séquence ARN coiffée dans le site actif de PB1 (**Fig.14**). Les bases Ad et/ou Gu précédant le site de clivage de l'amorce coiffée, s'apparient alors aux deux premières bases, Ur1 et Cy2 (ou Cy1), situées à l'extrémité 5' de l'ARNv matrice. PB1 initialise la synthèse des ARNm viraux. Un réalignement de l'amorce coiffée sur l'ARNv matrice peut avoir lieu (**Fig.14**) (Te Velthuis et Oymans, 2018). Les acides aminés K140 et R142 de PB2 maintiennent l'amorce dans le site actif de PB1 (Sikora *et al*., 2014).

En cours d'élongation, l'ARNv matrice et l'ARNm néo-synthétisé se désapparient progressivement en empruntant des voies de sorties différentes. Lorsque la RdRp virale atteint l'extrémité 5' de la matrice ARNv, le complexe RdRp viral rencontre une région riche en Ur (de 5 à 7 Ur qui correspondent au signal de polyadénylation des ARNm viraux) localisée 16 nucléotides en amont de l'extrémité 5' de l'ARNv matrice, ce qui favorise son bégaiement (*stuttering*) et provoque l'addition d'une queue poly-Ad à la fin d'ARNm viral néo-synthétisé (pour revue, Dou *et al*., 2018). Le remplacement de ce motif riche en Ur par un motif riche en Ad entraine la synthèse d'ARNm viraux présentant une queue poly-Ur ce qui a pour conséquence d'entrainer leur séquestration nucléaire rendant alors impossible leur traduction en protéine virale (Poon *et al*., 2000). L'appariement de l'extrémité 5' de l'ARNv matrice à la séquence 3' de l'ARNm viral néo-synthétisé est dissocié une fois la queue poly-A additionnée à l'extrémité 3' de l'ARNm viral. Des analyses structurales réalisées par cryo-EM ont permis de montrer des changements de conformation survenant dans la RdRP virale conduisant à (i) un relargage de l'ARNm viral néo-synthétisé et (ii) à la fixation de l'extrémité 3' de l'ARNv matrice sur un second site de liaison de l'extrémité 3' sur PB1. Ces changements de conformation de l'appariement des extrémités 3' et 5' de l'ARNv matrice induisent la formation d'un complexe promoteur de la transcription virale, permettant de ré-initier la transcription virale. Avec de tels changements de

conformation de la RdRP, un ARNv matrice permet de synthétiser plusieurs ARNm viraux (**Fig.14**) (Wandzik *et al*., 2020).

Les protéines virales sont synthétisées une fois que les ARNm viraux sont exportés dans le cytoplasme. Les protéines qui constituent la RdRp virale (PA, PB1 et PB2) ainsi que les protéines NP, NS1, NEP et M1, sont adressées au noyau. La karyophérine Ran-binding protein 5 (RanBP5) est directement impliquée dans l'import des protéines PA et PB1 néo-synthétisées (Swale *et al*., 2016). La résolution récente de sa structure et les essais de complémentation *in cellulo* basés sur l'émission d'une luciférase rapportrice semblent indiquer l'interaction directe de RanPB5 avec PB1. Certaines mutations localisées dans PB1, en particulier les simples mutations E299K – N439K – N439D – N486D – D539K et K574E, stimulent activement la formation des complexes PB1-RanBP5. A l'inverse, les simples mutations D362K – E405K – N486K – E489K – D490K – E528K et Q529K réduisent drastiquement la formation du complexe, limitant significativement l'import de PB1. Les simples mutations Q295K – E306K – E358K – Q443K – Q443E – et D447K réduisent faiblement cette interaction (Swale *et al*., 2020).

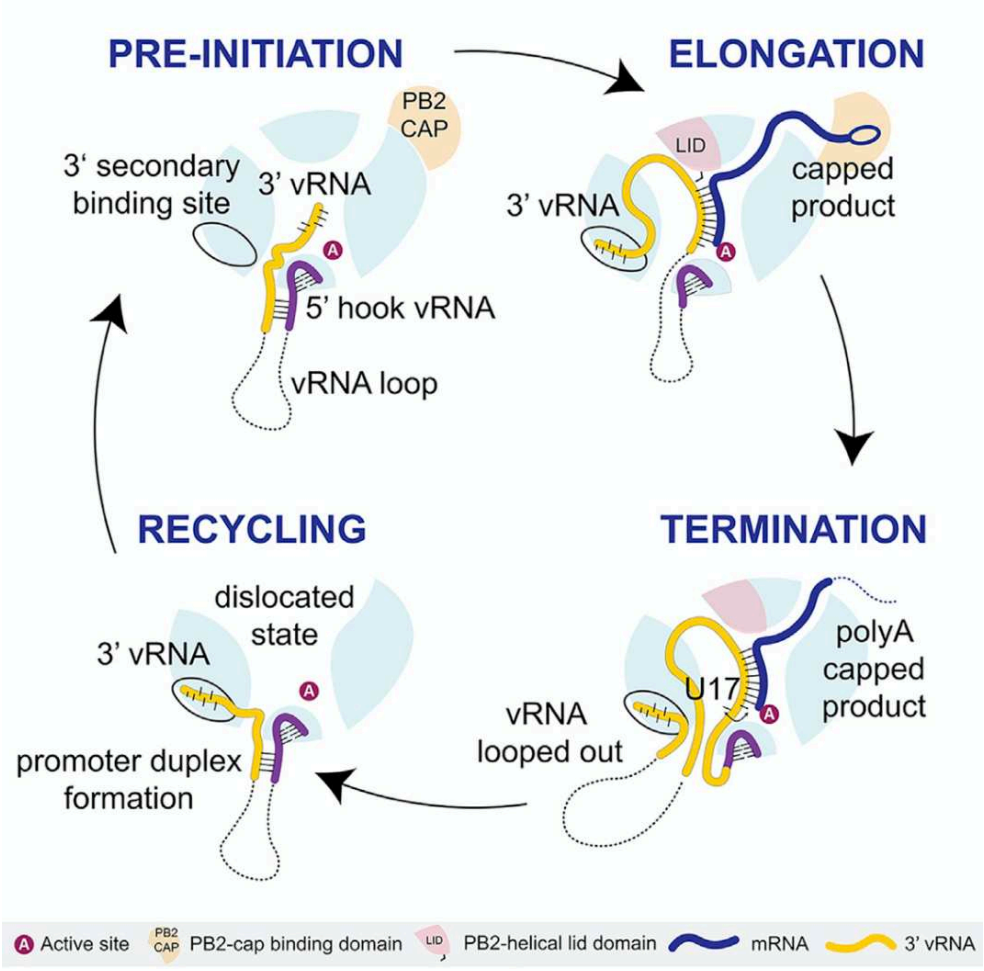


Figure 14 : Transcription des RNPv en ARNm viraux *(Wandzik et al., 2020)*

B.IV.4. La réplication virale

La réplication du génome, qui a lieu après la transcription, copie l'ARNv en ARN complémentaire (ARNc) de polarité positive, puis l'ARNc en ARNv. Même si l'ARNc est exactement complémentaire à l'ARNv génomique, il ne présente ni coiffe ni queue poly-A. Ils sont associés à PA, PB1, PB2 et NP pour former les RNPc qui serviront de matrice pour la synthèse de nouvelles RNPv (**Fig.15**).

Contrairement à la transcription virale qui nécessite l'utilisation d'une amorce, la réplication du génome viral dépend de l'insertion d'ATP et de GTP au sein du site actif de PB1 afin de réaliser le couplage des deux premiers nucléotides (pppAdGu) avec les bases (UrCy) de l'ARN matrice. L'initiation de la réplication de l'ARNv diffère de celle de l'ARNc. En effet, pour former de l'ARNc, le motif pppAdGu est d'abord synthétisé juste après son interaction avec les bases en position 4 et 5 de l'ARNv (**Fig.15**). A l'inverse, la synthèse d'ARNv nécessite l'ajout d'un nucléotide non-apparié (une purine généralement) à l'extrémité 3' de l'ARN matrice. On ignore encore à ce jour comment s'opère la transition entre les étapes de transcription et de réplication du virus. Les disponibilités en NTP et en amorces limitent l'avancée de ces deux étapes (pour revue, Resa-Infante *et al.*, 2011).

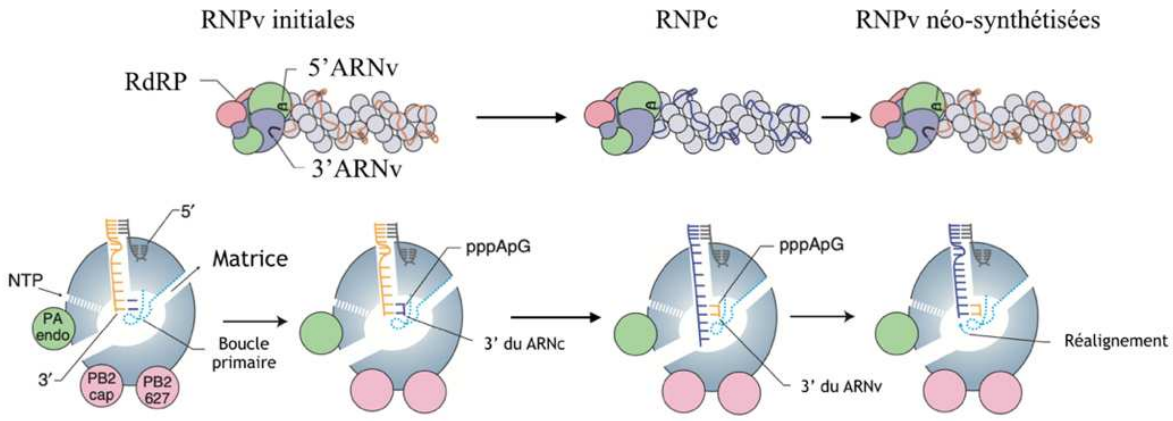


Figure 15 : Réplication des RNPv *(adapté de Fodor et Te Velthius, 2019)*

B.IV.5. Régulations des étapes de transcription et de réplication du génome viral : transition entre les deux états

Plusieurs études structurales ont tenté d'apporter des éléments de réponse quant aux mécanismes moléculaires gouvernant la transition entre la transcription et la réplication du génome viral. Des récentes études *in cellulo* et structurales réalisées sur un virus réassortant associant le segment PB2 de la souche Wilson Smith Neurotropic (WSN) au fond génétique du virus PR8 démontrent une atténuation du virus réassortant, qui peut être compensée, partiellement, par l'apparition de mutations dans les trois sous-unités de la RdRp virale (Chen *et al.*, 2019). Ils ont également pu observer une dimérisation de la RdRp virale, laquelle est modulée négativement par les mutations PA-349, PB2-74 et PB1-577 qui sont

cruciales pour la restauration d'un fitness proche du virus PR8 sauvage et qui sont toutes localisées dans la zone de dimérisation de la RdRp virale. D'autres mutations, en particulier PB2-701, PB1-195 et PA-349 augmentent l'activité réplicative du virus, souligné par l'augmentation du niveau d'ARNv et d'ARNc de NP, mais réduisent fortement la dimérisation du complexe RdRp du virus révertant (Isel *et al*., 2020). Tous ces résultats (y compris ceux d'autres études) suggèrent que l'état dimérique, trimérique, tétramérique voire même oligomérique du complexe RdRP viral pourrait moduler son passage de la transcription à la réplication. La protéine NP est également très étudiée sur le plan structural afin de comprendre son rôle dans la transition transcription-réplication du virus. D'après ses propriétés de liaison aux ARNv et ARNc, NP pourrait moduler l'état de transition entre transcription et réplication du génome viral en altérant la structure *panhandle* du promoteur viral (Fodor *et al*., 1994; Klumpp *et al*., 1997). En interagissant directement avec la RdRp virale, NP pourrait également bloquer le CAP-BD de PB2 et favoriser une transcription virale indépendante du *cap-snatching*. Plusieurs types d'ARN viraux sont produits au cours du cycle viral, notamment les ARNv et ARNc du virus mais aussi trois groupes de courts ARN viraux (DI RNA, mvRNA et svRNA; Fodor et te Velthius, 2019). Identifiés par Perez *et al*. (2010), les svRNA correspondent à la séquence 5'-UTR conservée des ARN viraux positifs (ARNc et ARNm viraux). Une fois synthétisés, ces svRNA peuvent s'associer au site actif de la RdRp virale et moduler ses activités. En effet, la dépletion des svRNA réduit drastiquement la synthèse d'ARNv (qui survient lors de l'étape de réplication du virus) mais présente peu d'effets sur la transcription virale (Perez *et al*., 2010). La protéine NEP régule activement le niveau de svRNA, modulant ainsi la transition transcription-réplication du génome viral par la RdRp virale (Perez *et al*., 2010).

B.V. Export nucléaire et traffic cytoplasmique des RNPv néo-synthétisées

Afin d'être empaquetés dans les virions en cours de formation, les RNPv néo-synthétisées sont exportées dans le cytoplasme par la voie de l'importine-β liée à l'exportine 1 aussi appelée CRM1. En effet, les RNPv sont séquestrées dans le noyau lorsque les cellules infectées sont traitées avec de la Leptomycine B, un inhibiteur spécifique de la voie CRM1 (Güttler *et al*., 2010). La protéine NEP assure le transport nucléo-cytoplasmique des RNPv en faisant l'intermédiaire entre M1 et CRM1, grâce à la présence de deux signaux d'export nucléaire (NES) dans la séquence de NEP (acides aminés 12-21 et 31-40) (**Fig.16**) (Huang *et al*., 2013).

La protéine NEP (121 acides aminés) est organisée en deux domaines : la région N-terminale (acides aminés 1-53) et la région C-terminale (acides aminés 54-121). Seule la structure de cette dernière a été résolue. Cette région, contenant de nombreux acides aminés hydrophobes en particulier l'acide aminé W78, forme une « épingle à cheveux » (hairpin) qui interagit avec M1 (Akarsu *et al*., 2003; Arzt

et al., 2004). Les protéines NEP, en interagissant d'un côté avec les complexes CRM1 - Ran-GTP, et de l'autre avec M1 qui interagit elle-même avec les protéines NP des RNPv, permettent l'export nucléo-cytoplasmique des complexes (RNPv-M1-NEP-CRM1-Ran-GTP) (pour revue, Paterson et Fodor, 2012) (**Fig.16**).

Les complexes (RNPv/M1/NEP/CRM1/Ran-GTP) se dissocient dans le cytoplasme, après hydrolyse du GTP de Ran. Les RNPv s'accumulent alors près du centre de contrôle des microtubules (MTOC), impliqué dans le transport intra-cytoplasmique de protéines (pour revue, Wu et Akhmanova, 2017) et décrit comme participant à l'infection virale (Amorim *et al*., 2011; Bruce *et al*., 2012; Zhang *et al*., 2018). Ils rejoignent ensuite la surface de la membrane plasmique pour y être empaquetées, grâce à l'interaction de M1 avec la GTPase Rab-11 (ras-related in brain), une protéine essentielle à l'infection virale et qui est associée au MTOC (**Fig.16**) (Bruce *et al*., 2010; pour revue, Amorim, 2018). PB2 a aussi été décrite comme interagissant avec Rab-11 (Avilov *et al*., 2012). Des études microscopiques ont permis de montrer que l'infection virale induisait des remodelages importants et irreguliers du réticulum endoplasmique (RE), en particulier du réticulum endoplasmique granuleux (REG), aussi appelées d'organelles liquides qui sont en étroite relation avec le RE. L'analyse par western-blot de ces structures irrégulières localisées tout autour du RE révèle la présence des protéines NP et PB2 ainsi que de structures RNPv-like, ce qui semble indiquer que les RNPv sont transportées à travers le cytoplasme par l'intermédiaire de ses structures irrégulières. Ce remaniement du RE provoqué par l'infection virale fait intervenir la protéine Rab-11 (de Castro-Martin *et al*., 2017; Alenquer *et al*., 2019).

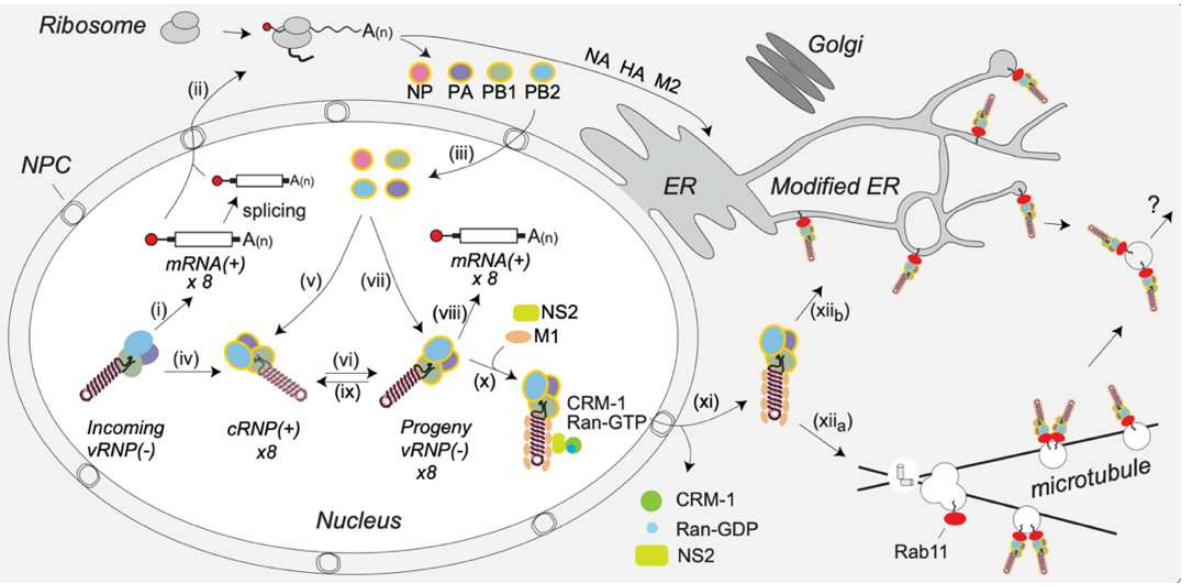


Figure 16 : Assemblage et traffic des RNPv néo-synthétisées *(Dou et al., 2018)*

B.VI. Empaquetage des RNPv, bourgeonnement et relargage des particules virales

Dans les virus, on observe des complexes macromoléculaires dans lesquels les RNPv sont organisées en 7+1 et il semble que ce soit réellement ces complexes qui sont acheminés dans les particules bourgeonnantes. Plusieurs études ont tenté d'éclairer les mécanismes permettant aux 8 RNPv d'être empactées. Ces complexes macromoléculaires seraient maintenus par des interactions ARN-ARN entre les différents segments génomiques *via* des séquences en cis présentes sur chacun des 8 RNPv pour un empaquetage spécifique d'un jeu complet des 8 RNPv (Gavazzi *et al*., 2013; Gerber *et al*., 2014; Miyamoto et Noda, 2019; pour revue, Isel *et al*., 2016). A titre d'exemple, le segment génomique PB1 interagit avec celui de NA (Gilbertson *et al*., 2016; Gerber *et al*., 2014) ainsi qu'avec ceux codant pour HA, M et NS (Miyamoto et Noda, 2019). De plus, la protéine NP contenue dans les RNPv semble être au cœur de l'interaction avec les ARNv qui composent ces mêmes RNPv (Bolte *et al*., 2019). La structure des RNPv qui est gouvernée par la capacité de NP à lier l'ARNv (et l'ARNc) tous les 24 nt ne protège pas toute la séquence de l'ARN. En effet, quelques structures d'ARN libres subsistent tout au long des RNPv (Miyamoto et Noda, 2020; pour revue, Noda *et al*., 2020).

Au cours du bourgeonnement viral, les glycoprotéines d'enveloppe HA et NA ainsi que le canal M2, s'associent avec les lipides des membranes du RE pour être acheminées vers la surface apicale de la membrane cellulaire. La protéine de matrice M1, associée au RNPv, est également liée aux radeaux lipidiques (Höfer *et al*., 2019; Bobone *et al*., 2017) ainsi qu'aux queues cytoplasmiques de HA (Siche *et al*., 2015), de NA (Ohkura *et al*., 2014) et de M2 (Chen *et al*., 2008). Même si on ignore encore avec exactitude comment les protéines virales provoquent le bourgeonnement, il semble que l'enchâssement de M1 dans la bicouche lipidique entraine son réarrangement (Martyna et Rossman, 2014). Différentes protéines virales et cellulaires sont embarquées dans la particule virale lors de l'étape de bourgeonnement viral (Hutchinson *et al*., 2014).

Une fois produites, les particules virales restent accrochées à la surface de la cellule hôte *via* l'interaction des HA avec les acides sialiques portés par les glycoprotéines de la cellule hôte. Ces acides sialiques sont coupés par l'activité sialydase (acylneuraminyl hydrolase) de la neuraminidase virale, libérant ainsi les particules (Glanz *et al*., 2018). Pour ce faire, les acides sialiques interagissent spécifiquement avec le domaine « tête » de NA (**Fig.17A**, **17B** et **17C**), dans lequel deux sites de liaison ont été identifiés (**Fig.17C**). Des analogues structuraux aux acides sialiques ont permis de développer des inhibiteurs dirigés contre la protéine NA (**Fig.20**), limitant la dissémination des virions bourgeonnant à la surface de la cellule infectée.

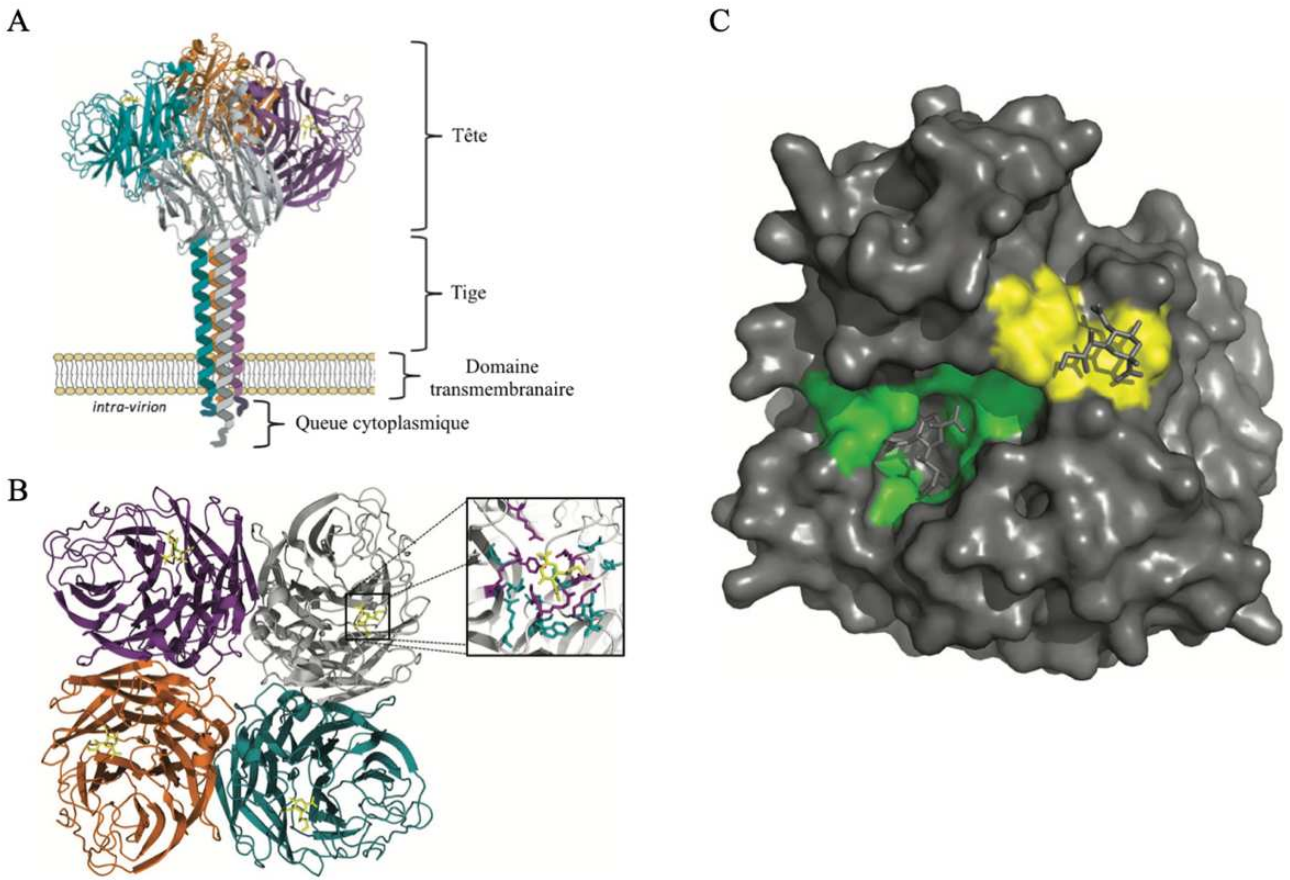


Figure 17 : Structure de la neuraminidase. *Représentation par PyMol du domaine "tête" de NA (PDB : 4GZX, Zhu et al., 2012). La protéine NA est un tétramère composé de quatre monomères identiques, où chacun d'eux constitue les domaines "tête", "tige", "transmembranaire" et "queue cytoplasmique". (**A**) Les monomères sont indiqués en vert, en gris, en violet et en orange. Les autres domaines de NA ("tige", "transmembranaire" et "queue cytoplasmique") sont représentés par une hélice. (**B**) Le domaine tête de NA présente une structure d'hélice à six pales, où chaque pale se compose de quatre feuillets B stabilisés par des ponts disulfures et connectés les uns aux autres par des boucles de longueur variable. Les acides sialiques (jaune) interagissent avec le site actif de NA, situé dans la tête de chaque monomère. L'interface d'interaction entre NA et l'acide sialique (acides aminés représentés en bleu et en violet) a été représentée. (**C**) Représentation par PyMol des deux sites de fixation des acides sialiques sur le domaine "tête" de NA (PDB : 1MWE, Varghese et al., 1997). Les acides aminés R118, D151, R152, R224, E276, R292, R371 et Y406 (verts) correspondent au site catalytique de NA, dans lequel le clivage des acides sialiques s'opère. Un second site de fixation des acides sialiques (jaune) est assuré par les acides aminés S367, S370, S372, N400, W403 et K432 (adapté de McAuley et al., 2019).*

C. Les moyens de lutte mis en place contre les virus *influenza*

C.I. Les méthodes préventives

C.I.1. La vaccination annuelle

Les vaccins antigrippaux permettent de contrôler l'impact des grippes saisonnières chez l'Homme. Toutefois, puisque ces virus évoluent constamment par dérive antigénique, il est nécessaire de réadapter chaque année la composition antigénique du vaccin. L'OMS adapte ainsi la composition vaccinale en tenant compte des caractéristiques antigéniques des virus circulants, lesquels sont collectés par une centaine de centres nationaux ainsi que des centres de références de la grippe répartis dans le monde. L'OMS émet chaque année en février ses recommandations pour la composition du vaccin utilisable à l'automne dans l'hémisphère Nord (http://www.euro.who.int/en/health-topics/communicable-diseases/Influenza).

C.I.1.a. Les vaccins antigrippaux inactivés

La souche virale choisie (HxNy) est d'abord co-inoculée dans la cavité allantoïque d'œufs embryonnés (fécondés) de poules avec la souche humaine PR8 (A/PuertoRico/8/1934 (H1N1)), afin d'obtenir des virus réassortants HxNy ayant un taux élevé de production virale en œufs embryonnés. Ces virus vaccinaux sont alors adressés aux laboratoires producteurs. Les virus récoltés sont inactivés par traitement chimique (utilisation du formaldéhyde ou du propionolactone) avant d'en extraire les glycoprotéines HA et NA (**Fig.18**). Entre 500 et 800 millions de doses vaccinales sont produites ainsi chaque année. Ces vaccins inactivés contiennent environ 15 µg de glycoprotéines HA de chacun des virus cibles qui, pour la saison 2019-2020 dans l'hémisphère Nord, sont les souches A/Brisbane/02/2018/H1N1 (pmd2009), A/Kansas/14/2017/H3N2, B/Colorado/06/2017 (lignée B/Victoria/2/87) et B/Phuket/3073/2013 (lignée Yamagata/16/88). D

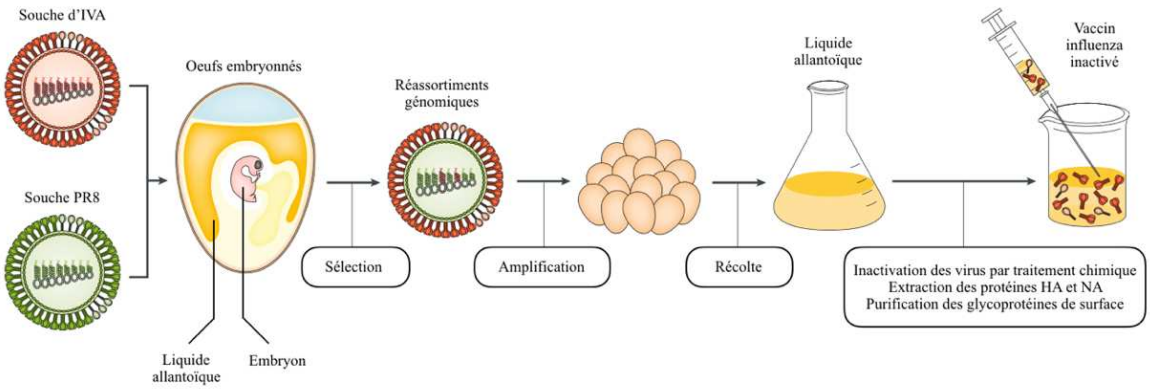


Figure 18 : Principe du réassortiment viral pour l'obtention de vaccins fragmentaires *(adapté de Krammer et al., 2018)*

C.I.1.b. Les vaccins antigrippaux vivants atténués

La vaccination antigrippale peut également être faite en utilisant des vaccins vivants atténués. Ces derniers, qui sont composés d'environ 10^7 particules virales infectieuses des quatre souches virales cibles, sont issus du réassortiment entre la souche sauvage produite sur œufs embryonnés et la souche atténuée produite sur cellules épithéliales de rein de singes (Kidney epithelial cells from an African green monkey, VERO). Ces vaccins atténués adaptés au froid sont (i) capables de se répliquer uniquement au niveau du tractus respiratoire supérieur (nasopharynx) à une température comprise entre 25 et 33 °C et (ii) thermosensibles car incapables de se répliquer au niveau du tractus respiratoire inférieur. Le vaccin vivant atténué tétravalent Fluenz (Tetra®), administrable par voie-intranasale, a vu le jour ces dernières années. Ces vaccins vivants atténués induisent dans le tractus respiratoire supérieur une réponse immunitaire locale.

La durée de production de vaccins inactivés ou vivants atténués est d'environ huit mois, ce qui est un inconvénient dans l'éventualité d'une pandémie. Des solutions alternatives basées sur la production de vaccins en cellules de mammifères, en particulier les cellules de rein de chien (Madin-Darby Canine Kidney Cell, MDCK) et les cellules VERO (pour revue, Pérez-Rubio et Eiros, 2018), ont été développées pour permettre de pallier aux contraintes de temps et aux besoins importants en œufs embryonnés, mais pour l'instant plus de 95 % des doses vaccinales sont produites en œufs embryonnés. Des études s'intéressent également à d'autres types cellulaires qui pourraient permettre la multiplication de souches vaccinales (Gränicher *et al*., 2019), voire leur production par génétique inverse (Koudstaal *et al*., 2009). La durée de production des vaccins anti-grippaux restant une limite importante, plusieurs équipes de recherche s'intéressent à produire les protéines de surface HA et NA à partir d'autres procédés. A ce jour, plusieurs essais cliniques de phase III ont été réalisés en utilisant un vaccin quadrivalent basé sur l'expression des protéines HA et NA directement à la surface de plantes recombinantes (Ward *et al*., 2020).

C.I.2. Les vaccins de demain

C.I.2.a. Les vaccins antigrippaux « high-dose »

Les vaccins high-dose sont obtenus selon le même protocole que pour les vaccins inactivés, cependant ils contiennent 60 µg de glycoprotéines HA de chacun des virus cibles soit près de quatre fois plus de protéines que dans les vaccins inactivés standarts. Un tel vaccin tétravalent, le Fluzone (ou Efuelda ou USA), a été autorisé aux USA. Les premiers essais de vaccination réalisés sur près de 32 000 volontaires américains agés de plus de 65 ans démontrent une meilleure immunogénicité de ces vaccins et une augmentation de la protection contre les virus grippaux circulants de 24,2 % par rapport aux vaccins inactivés standarts (DiazGranados *et al*., 2014).

C.I.2.b. Développement d'un vaccin universel

Alors que la protéine HA évolue constamment sous la pression de la réponse immunitaire (Meyer et Wilke, 2015), de nombreuses équipes tentent de développer des vaccins qui susciteraient des anticorps ciblant les épitopes les plus conservés du virus, pouvant ainsi potentiellement induire une réponse anticorps protectrice vis à vis de plusieurs sous-types viraux. L'ectodomaine de la protéine M2 pourrait constituer une cible intéressante (Padilla-Quirarte *et al*., 2019), mais la majorité des recherches actuelles se concentrent sur les épitopes conservés dans la tige de l'hémagglutinine, portés par son domaine HA2 (Avnir *et al*., 2014).

D'autres équipes s'intéressent aux variations de HA et de NA dans le but d'identifier les régions les plus conservées dans ces deux glycoprotéines (Julien *et al*., 2012; Laursen et Wilson, 2013). Plusieurs stratégies ont été développées, en particulier en immunisant avec HA2 seule (Graves *et al*., 1983) ou en réduisant l'immunodominance de HA1 (Chen *et al*., 2016; Krammer *et al*., 2013). Des études récentes portant sur le développement de vaccins ciblant spécifiquement la protéine NA ont permis d'identifier des anticorps à large spectre chez un patient infecté par le virus saisonnier H3N2 actuel. Ces anticorps, capables de neutraliser l'infection virale de plusieurs sous-types d'IAV et d'IVB, se lient à la région conservée du site actif de la neuraminidase (Stadlbauer *et al*., 2019). Des recherches sont actuellement en cours afin de développer un vaccin qui se base sur les motifs conservés du site actif de NA qui sont reconnus par ces anticorps.

C.II. Les méthodes curatives

De nombreuses approches ont été développées afin d'identifier des composés capables de perturber ou bloquer le cycle viral. La liste que j'ai choisi de présenter référence uniquement les molécules ciblant spécifiquement les protéines virales. J'ai également choisi de faire un listing des différentes protéines celulaires identifiés comme essentielles au cycle viral.

C.II.1. Les inhibiteurs de la protéine M2

Les adamantanes (**Fig.19**), premiers inhibiteurs développés contre le virus, se fixent sur la protéine M2 et empêchent l'ouverture du canal ionique. Ces molécules qui doivent être administrées immédiatement après infection, empêchent l'infection virale. Les virus saisonniers H3N2, puis H1N1, sont devenus majoritairement résistants aux adamantanes au milieu des années 2000, par acquisition de la substitution S31N dans M2 (Duwe, 2017). En conséquence, les adamantanes ne sont plus recommandés pour le traitement.

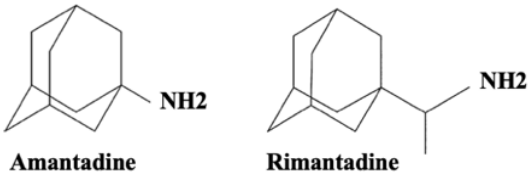


Figure 19 : Les adamantanes, ciblant la protéine M2

C.II.2. Les inhibiteurs de la neuraminidase

L'oseltamivir (Tamiflu®) et le zanamivir (Relenza®) (**Fig.20**) sont des analogues structuraux des SA (Neu5Ac5en, **Fig.20**) qui bloquent par compétition le site actif de NA (**Fig.17**). Même si plusieurs cas de résistances à l'oseltamivir (**Fig.20**) ont été répertoriés, cet antiviral reste l'inhibiteur le plus utilisé à ce jour. Le laninamivir (Inavir®) et le peramivir (**Fig.20**) ont récemment prouvé leur efficacité y compris face à des virus résistants à l'oseltamivir (Vavricka *et al*., 2011).

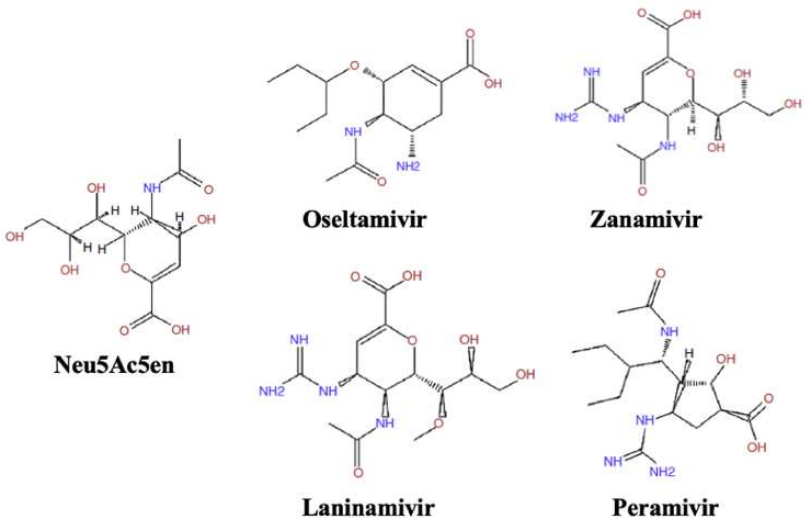


Figure 20 : Les analogues d'acides sialiques, ciblant la protéine NA

C.II.3. Les traitements ciblant l'hémagglutinine

Inhibiteurs peptidiques

Une première approche basée sur l'utilisation de peptides inhibiteurs réduisant partiellement la fusion des membranes a été développée. Un premier peptide de 20 acides aminés (RRKKAAVALLPAVLLALLAP) (Jones *et al*., 2006) dérivé du facteur de croissance 4 de fibroblastes humains (FGF-4) a été décrit comme capable de bloquer efficacement l'infection virale, et ce qu'il soit utilisé pré- ou post-infection, suggérant une utilisation en thérapeutique. Rajik *et al*. (2009) ont mis au point un second peptide, plus court (CNDFRSKTC), qui se lie spécifiquement à HA1 et entraine le blocage du déploiement du trimère de HA, inhibant ainsi le relargage cytoplasmique des RNPv.

Anticorps thérapeutiques

L'anticorps monoclonal C179 (Okuno *et al*., 1993), isolé à partir du sang de souris infectée avec un virus H2N2, protège de l'infection virale dans un modèle murin (Wohlbold *et al*., 2015; Sagawa *et al*., 1996), et son dérivé (CH65) (Whittle *et al*., 2011) neutralise l'infectiosité de plusieurs virus. Ekiert *et al*. (2012) ont identifié l'anticorps monoclonal C05 à large spectre qui reconnaît les structures conservées du RBS de HA. Des anticorps reconnaissant préférentiellement HA2, plus conservée que HA1 (Dreyfus *et al*., 2012; Ekiert *et al*., 2009), comme le CR6261 ont été isolés, parmi lesquels figure l'anticorps neutralisant à large spectre FluA-20 isolé dans le sang d'un donneur vacciné (Bangaru *et al*., 2019). Ces anticorps neutralisants à large spectre (bnAbs) pourraient constituer une nouvelle approche intéressante : des anticorps chimériques combinant plusieurs domaines de bnAbs ont montré leur efficacité dans un modèle souris (Laursen *et al*., 2018).

Composés chimiques

Plusieurs molécules ciblant préférentiellement HA2 ont aussi été identifiées, en particulier le tert-butyl hydroquinone (TBHQ) (**Fig.21**) qui se lie au peptide de fusion de HA2, le stabilisant ainsi dans une conformation non-dépliée (Russell *et al*., 2008). Ilyina *et al*., (2018) ont identifié plusieurs composés dérivés de monoterpènes ciblant HA, en particulier le composé (4R)-11a (**Fig.21**) qui est actif sur plusieurs virus mais reste néanmoins inefficace face à des virus résistants à l'oseltamivir.

Les bnAbs ciblant la tige HA2 de l'hémagglutinine (voir ci-dessus) ont aussi permis de concevoir de nouvelles molécules qui empêchent le déploiement de cette tige. Kadam *et al*. (2017) ont conçu des peptides cycliques combinant les fonctions chimiques de plusieurs bnAbs, alors que van Dongen *et al*. (2019) ont sélectionné par compétition avec des bnAbs des composés chimiques qui bloquent aussi le changement conformationnel de HA.

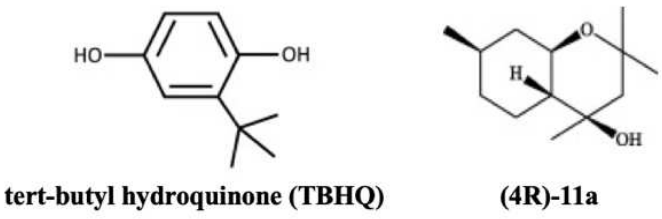


Figure 21 : Les composés chimiques ciblant la protéine HA

C.II.4. Les inhibiteurs de la RdRp virale

La RdRp est très conservée entre les sous-types viraux (Fan *et al*., 2019; Hengrung *et al*., 2015; Peng *et al*., 2019; Pflug *et al*., 2014; Reich *et al*., 2014). Un certain nombre d'antiviraux ciblant ce complexe sont des analogues de nucléotides dont la phosphorylation est bloquée. Une fois dans le site

actif de la RdRp, ces analogues structuraux bloquent la transcription et la réplication du génome viral. Constituée des protéines PA, PB1 et PB2, la RdRp virale peut être ciblée de plusieurs façons.

Ciblant la protéine PA

Tomassini *et al*. (1994) ont identifié plusieurs composés qui ciblent aussi le site actif de PA. Une vingtaine de dérivés, en particulier le DPBA (**Fig.22**), diminue l'activité endonucléase et réduit de façon dose-dépendante le titre viral *in cellulo* et *in vivo* (Hastings *et al*., 1996). L'analyse structurale du complexe associant le domaine PA-N au composé DPBA démontre que le composé séquestre les ions manganèses qui interviennent dans le site actif du domaine PA-N (Dias *et al*., 2009; Crépin *et al*., 2010; Kowalinski *et al*., 2012).

Approuvé par la FDA en octobre 2018, le marboxil baloxavir ou BXM (S-033188, Xofluza™) (**Fig.22**) se fixe au site actif de PA-N (Omoto *et al*., 2018) et empêche son activité endonucléase, ce qui réduit la réplication virale (pour revue, Ng, 2019). Le traitement par le BXA ou par sa forme active BXM (**Fig.22**), réduit la charge virale et permet la survie des souris infectées par des IAV et des IBV (Fukao *et al*., 2019; Taniguchi *et al*., 2019). Les premiers virus résistants au BXM, portant la substitution I38T dans PA, ont été récemment isolés chez des patients japonais (Imai *et al*., 2019).

Ciblant la protéine PB2

Clark *et al*. (2014) ont identifié le pimodovir (VX-787, JNJ63623872) (**Fig.22**) qui encombre la poche de fixation de la coiffe des ARNm de PB2 du complexe RdRp viral. D'ailleurs, ce composé, qui a montré une efficacité variable *in cellulo* et *in vivo* et qui est actuellement en phase clinique II, reste plus efficace que l'oseltamivir à dose équivalente.

Ciblant la protéine PB1

Déjà décrite comme réduisant la réplication de plusieurs virus à ARN (Furuta *et al*., 2013; Segura Guerrero *et al*., 2018), le favipiravir (T-705) (**Fig.22**), un analogue de nucléosides, présente une plus grande efficacité contre les virus *influenza*. Cette molécule, qui cible spécifiquement la sous-unité PB1 du complexe RdRp viral, demeure active sur des virus résistants à l'oseltamivir. A ce jour, seule la mutation en position V43I de PB1 a été décrite comme conférant une légère résistance à l'inhibiteur pour des virus humains H3N2 et H5N1 (Cheung *et al*., 2014).

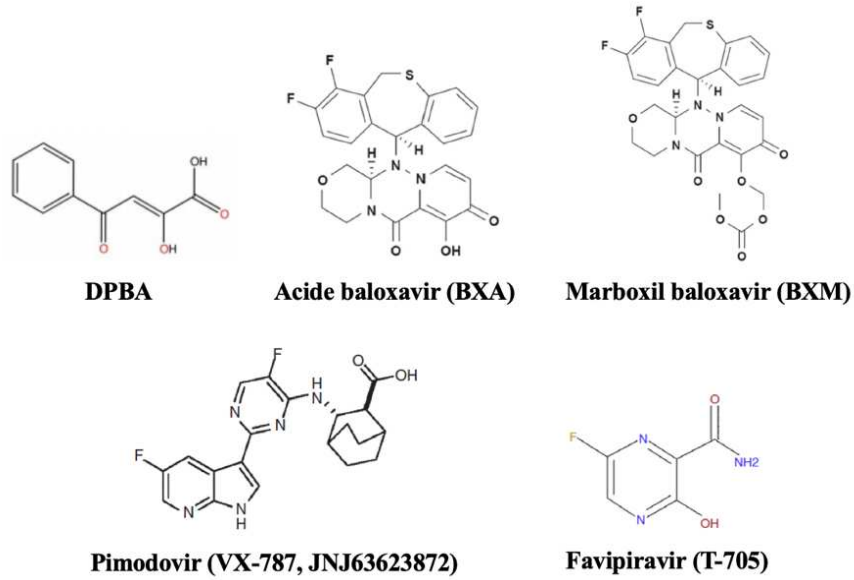


Figure 22 : Les inhibiteurs de la RdRp virale

C.II.5. Les inhibiteurs de la protéine NP

NP apparait comme une cible intéressante pour le développement de nouvelles molécules antivirales en raison de ses nombreux rôles dans le cycle viral et de sa haute conservation.

Inhibiteur peptidique

Ainsi, un premier peptide cyclique (QPTFSVQRNLCC), correspondant aux acides aminés 409 à 418 de la boucle de NP-C, a été développé dans le but de bloquer l'oligomérisation de NP. Toutefois, il a été rapidement abandonné en raison de son efficacité limitée (EC_{50} = 904 µmol.L^{-1}).

Composés chimiques

Des molécules chimiques ciblant le domaine core de NP ont aussi été identifiées (Gerritz *et al*., 2011), en particulier le nucleozin (**Fig.23**) qui cible les poches Y289/N309 et R382/R384 (Pang *et al*., 2016; Kao *et al*., 2010). Quatre molécules ont été sélectionnées pour leur interaction avec le domaine de liaison à l'ARN de NP : (i) le composé F66 (**Fig.23**), actif uniquement sur les IAV, a été abandonné en raison de son efficacité limitée chez la souris infectée (Fedichev *et al*., 2011); (ii) le naproxen (**Fig.23**), un anti-inflammatoire, a été sélectionné pour ses propriétés de liaison à NP (Lejal *et al*., 2013; Hu *et al*., 2017); (iii) le composé RK424 (**Fig.23**) provoque l'accumulation de NP dans le noyau de la cellule infectée (Kakisaka *et al*., 2015); et (iv) le composé PPQ581 (**Fig.23**) qui se lierait à l'acide aminé S377 (Lin *et al*., 2015), la substitution S377G réduisant l'efficacité du composé (Lin *et al*., 2015).

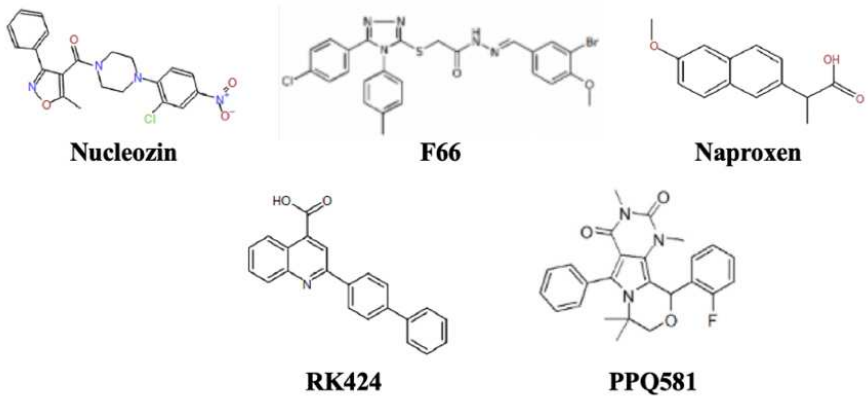


Figure 23 : Les inhibiteurs ciblant la protéine NP

C.III. Identification de nouveaux facteurs cellulaires essentiels au cycle viral pour le développement de traitements futurs pouvant efficacement limiter l'infection

A ce jour, la majorité des solutions thérapeutiques proposées pour lutter contre une infection virale avait pour cible une ou plusieurs protéines du virus. Cependant, l'utilisation de certains de ces composés antiviraux (voir chapitre C.II) entraine irrémédiablement l'apparition de virus résistants. Plusieurs travaux visant à limiter l'infection virale en ciblant spécifiquement des protéines cellulaires impliquées dans le cycle viral ont donné des résultats plus ou moins prometteurs dans le développement de nouvelles thérapies antivirales (Watanabe et Kawaoka, 2015). Plusieurs études de grande envergure impliquant l'utilisation de technologies de screening à haut débit sur plusieurs types cellulaires (les cellules épithéliales bronchiques humaines HBEC, les cellules d'ostéosarcome U2OS, les cellules d'adénocarcinome pulmonaire A549, les cellules embryonnaires de rein humain HEK293, les cellules stromales DL1 et les cellules de rein de chien MDCK) ont permis d'identifier de nombreux facteurs cellulaires essentiels au cycle des virus *influenza* (Hao *et al*., 2008; Brass *et al*., 2009; Shapira *et al*., 2009; Sui *et al*., 2009; König *et al*., 2010; Karlas *et al*., 2010; Ward *et al*., 2012; Su *et al*., 2013; Fournier *et al*., 2014; Watanabe *et al*., 2014; Han *et al*., 2018).

La première approche choisie pour identifier des protéines cellulaires essentielles au cycle viral des virus *influenza* consistait à utiliser des petits ARN interférents (siRNA) ciblant l'entiereté du génome cellulaire. Ainsi, l'utilisation de siRNA dans des cellules de drosophile D-Mel2 infectées par un virus de sous-type H5N1 ou H1N1 a permit de mettre en évidence l'importance de la sous-unité D1 de l'ATPase V0 (ATP6V0D1), de la sous-unité 6A1 de la cytochrome C oxydase (COX6A1) et de la protéine d'export nucléaire 1 des ARNm (NXF1) dans le cycle viral (Hao *et al*., 2008). Une seconde étude réalisée dans des MDCK infectées par le virus Udorn (A/Udorn/1972/H3N2) a permis à d'identifier près de 110 facteurs cellulaires essentiels à l'infection, en particulier la protéine ribosomale

mitochondriale 42 (MRPL42), la sous-unité 5A de la cytochrome C oxydase (COX5A), la protéine transmembranaire 1 (TAPT1) et le transporteur mitochondrial SCaMC2 qui lie le calcium (SLC25A25) (Sui *et al*., 2009). En utilisant la même approche, Brass *et al*., 2009) ont identifié près 130 facteurs cellulaires essentiels, en particulier la protéine NXF1, la protéine de pore nucléaire de 98 kDa (NUP98), le facteur de traduction 4A2 (EIF4A2), l'Archaine 1 (ARCN1), la sous-unité gamma du facteur coatomer (COPG), la phosphogluconate déshydrogénase (PGD), les protéines Rab-5A et Rab-10 affiliée à Ras (RAB5A et RAB10, respectivement) et le facteur de restriction transmembranaire induit par l'interféron (IFITM), dans des cellules U2OS infectées par le virus PR8. Beaucoup d'autres protéines ont également été identifiées dans des cellules A549 infectées par un virus WSN recombinant (l'ORF de HA ayant été remplacée par celle de la luciférase Renilla), notamment l'ATPase vacuolaire (vATPase), le récepteur du facteur de croissance fibroblastique (FGFR) et la glycogène kinase synthase 3-β (GSK3) qui régulent l'entrée du virus et la protéine kinase β-II dépendant de la calmoduline (CAMK2B) qui est impliquée dans les étapes tardives du cycle viral (König *et al*., 2010).

Watanabe *et al*. (2014) ont développé une seconde méthode qui consiste en premier lieu à identifier par spectrométrie de masse les interactants protéiques co-immunoprécipités liés aux différentes protéines virales, puis à vérifier leur implication dans le cycle viral en les inhibant par siRNA. Ils ont ainsi pu décrire, dans des cellules HEK293 infectées par le virus WSN, plus de 300 facteurs cellulaires essentiels au bon déroulement du cycle viral. Parmi les protéines identifiées, la prohibitine (PHB), l'intégrine β-4 binding protein (ITGB4BP), la sous-unité 26S du protéasome non-ATPase 11 (PSMD11), l'ATPase de transition du RE (VCP), la sous-unité O de l'ATP synthase (ATP5O) et l'exportine 1 (XPO1 ou CRM1) entrainent toutes la rétention nucléaire des RNPv lorsque ces dernières sont inhibées par siRNA. Ils ont également montré l'importance du facteur β-2 humain transformé (SFRS10), du facteur d'échange de nucléotides guanine résistant à la brefeldine A (GBF1), de la kératine 14 (KRT14), de la caprine 1 (CAPRIN1) et de la sous-unité catalytique sérine/thréonine-protéine phosphatase 6 (PPP6C) ainsi que des protéines PPP6C, de la protéine de point de contrôle mitotique 3 (BUB3), de GBF1, de KRT14 et du facteur 2-like 1 dérivé de cellules stromales (SDF2L1) dans l'acheminement intracellulaire des glycoprotéines HA et NA, respectivement.

Une troisième méthode, plus performante, a été développée peu après. En mettant au point un système CRISPR/Cas9 sur des cellules A549 infectées par le virus VN04 (A/Vietnam/1203/04 (H5N1)), Han *et al*. (2018) ont pu mettre en évidence près de 453 facteurs cellulaires essentiels (dont certains avaient déjà été décrits dans des études antérieures) qui régulent l'infection virale. L'analyse des screening successifs effectués sur ces cellules a permis d'identifier le transporteur d'acide CMP-sialique (SLC35A1), le facteur 11 de croissance et de différenciation (GDF11), la protéine homéodomaine 3 de classe iroquoise (IRX3), le facteur C2 calcium-dépendant contenant le domaine 4C (C2CD4C), le facteur tripartite-motif 23 (TRIM23), le facteur de biosynthèse d'ancrage au phosphatidylinositol glycane (PIGN), le facteur chaine courte acyl-CoA deshydrogenase (ACADSB), le facteur GRAM-

domain containing 2A (GRAMD2), la protéine Capicua (CIC), la protéine Janus kinase 2 (JAK2) et la protéine inhibitrice de STAT3 (PIAS3) comme régulant activement le déroulement du cycle viral. Ils ont également montré que l'inhibition des facteurs SLC35A1 et CIC, qui régulent respectivement le transport d'acides sialiques et la réponse antivirale cellulaire, impacte très fortement l'infection virale de plusieurs sous-types viraux, en particulier VN04 (H5N1), PR8 (H1N1) et HK68 (A/Hong Kong/1/68 (H3N2)) alors que l'inhibition de la protéine PIAS3 semble bloquer spécifiquement l'infection des cellules par les virus VN04 et PR8 sans réduire l'infection à HK68. Même si les acides sialiques jouent un rôle important dans les interactions cellulaires, moduler leur expression (Han *et al*., 2018) et/ou leur composition (Heaton *et al*., 2017) limite efficacement l'infection par différents sous-types viraux ce qui souligne l'intérêt de développer des thérapeutiques les ciblant.

D. La protéine non-structurale 1 (NS1) des virus *influenza* A

D.I. Généralités

D.I.1. Découverte

Peter Palese (1977) a mis en évidence sur gel d'électrophorèse la présence de huit segments d'ARN génomiques pour les IAV et les IBV contre sept pour les ICV. Le plus petit segment génomique code pour la protéine non-structurale 1 (NS1), laquelle est produite précocement et abondamment dans la cellule infectée mais n'est pas impliquée dans la structure de la particule virale, même si d'infimes quantités de NS1 sont embarquées dans le virion lors de l'étape de bourgeonnement viral avec de nombreuses autres protéines cellulaires (Hutchinson *et al*., 2014). Dix ans après sa découverte, plusieurs équipes de recherche ont pu démontrer son importance dans le cycle viral (Hatada *et al*., 1990). De nombreuses fonctions de NS1 ont été documentées depuis les années 1980.

La protéine NEP a été identifiée par la suite. Traduite à partir d'un transcrit épissé de l'ARNm codant pour NS1 (Lamb et Lai, 1980), NEP est impliquée dans le transport nucléaire des RNPv (O'Neill *et al*., 1998) ce qui explique son autre nom de Nuclear Export Protein (NEP). De faibles quantités de NEP sont retrouvées dans la particule virale (Hutchinson *et al*., 2014).

D.I.2. Polymorphisme

NS1 est codée par le plus petit segment d'ARN génomique des IAV et des IBV (Sham et Palese, 2013, Field Virology Sixth Edition). La longueur de NS1 est généralement de 281 acides aminés pour les IBV et de 230 acides aminés pour les IAV (246 et 243 acides aminés respectivement pour les ICV et les IDV, Nogales *et al*., 2019) avec parfois des variations de longueur en C-terminale de NS1 suite à l'apparition de codons stop précédant le codon usuel (Hale *et al*., 2008). La protéine NS1 est l'une des protéines virales des virus *influenza* qui présente le plus de variations protéiques (**Fig.24**).

Il existe deux allèles de NS1 chez les IAV (Ludwig *et al*., 1991) : l'allèle A (^{21}RFADQELG^{28}) (**Fig.24**), majoritaire, est retrouvé chez tous les virus de mammifères (avec de rares exceptions) (Zohari *et al*., 2008), ainsi que chez la majorité des virus aviaires; l'allèle B (^{21}LLSMRDMC^{28}) (**Fig.24**), minoritaire, est présent exclusivement chez des virus aviaires, même s'il peut être retrouvé chez de nombreuses autres espèces animales (Muñoz-Moreno *et al*., 2019; Turnbull *et al*., 2016). L'infection de petits singes avec un virus d'IAV recombinant arborant une NS1 d'allèle B est très limitée, suggérant que l'allèle A de NS1 est potentiellement mieux adapté aux virus de mammifères (Treanor *et al*., 1989).

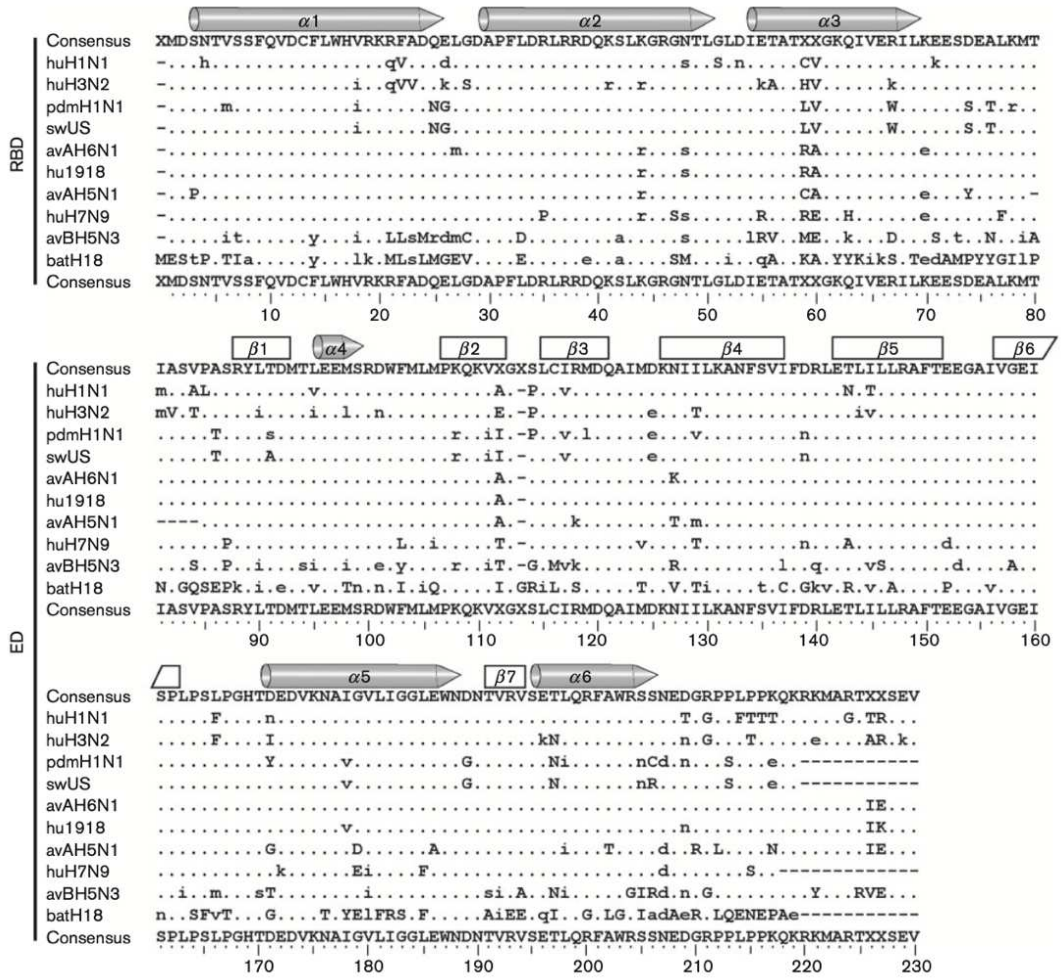


Figure 24 : Polymorphisme de NS1. *Alignement peptidique des séquences de NS1 représentatives des virus humains saisonniers H3N2 (huH3N2) et H1N1 avant 2009 (huH1N1), du virus pandémique de 2009 (pdmH1N1), de virus porcins américains (swUS), de virus aviaires d'allèle A (avAH6N1 et avAH5N1) et d'allèle B (avBH5N3), du virus pandémique de 1918 (hu1918), du virus humain H7N9 (huH7N9) et du virus de chauve-souris H18N10 (batH18). Les hélices alpha (α) et les brins beta (β) sont respectivement représentés par des tiges grises et des rectangles blancs. Les substitutions similaires et modifiant la charge des acides aminés sont respectivement représentées en minuscule et en majuscule. Les délétions sont indiquées par des tirets (Marc, 2014).*

D.I.3. Modifications post-traductionnelles

Plusieurs modifications post-traductionnelles de NS1 pouvant moduler le cycle viral ont été identifiées, en particulier des phosphorylations, des SUMOylations et des ISGylations (pour revue, Klemm *et al*., 2017). Ainsi, des phosphorylations ont été identifiées sur les acides aminés S42, S48, T49, T197 et T215. Certains acides aminés, notamment S42 et T49, ont été décrits comme importants pour le cycle viral puisque leur substitution par un acide aminé non-phosphorylable réduit le nombre de particules virales produites (Jiao *et al*., 2008; Kathum *et al*., 2016). Certaines lysines dans NS1, en particulier K70, K131 (Way *et al*., 2020), K219 et K221, peuvent être SUMOylées lors du cycle viral, alors que d'autres sont ISGylées comme K20, K41, K108, K126, K217 et K219 (**Fig.25**) (pour revue, Klemm *et al*., 2017 ; Rosário-Ferreira *et al*., 2020). Ma *et al*. (2020) ont démontré l'importance de l'acétylation de l'acide aminé K108 de NS1 dans la réplication et la virulence des virus influenza.

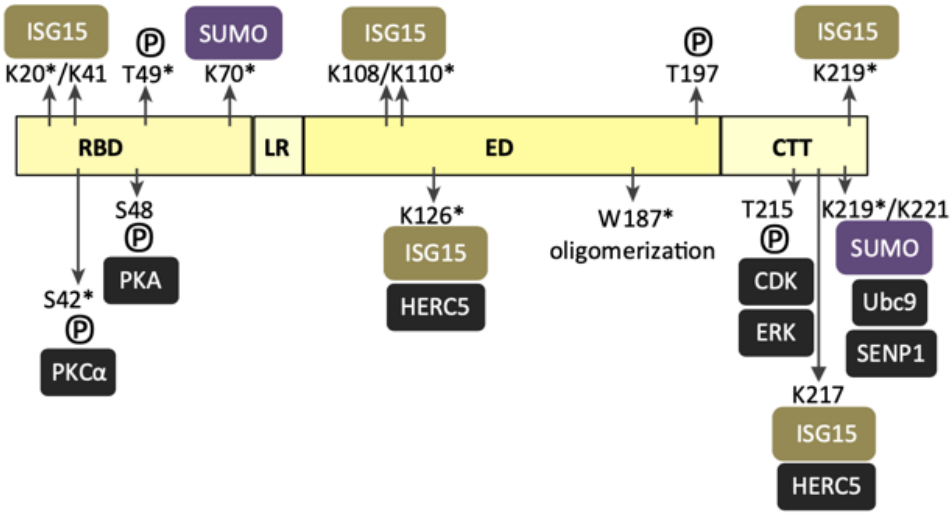


Figure 25 : Modifications post-traductionnelles de NS1. *Des modifications post-traductionnelles ont lieu sur NS1. La phosphorylation des acides aminés S42, S48, T49, T197 et T215, lesquelles sont assurées par plusieurs kinases. La SUMOylations des lysines K70, K219 et K221 est assurée notamment par la protéine Ubc9. La protéine SENP1 a été décrite comme assurant leur déconjugaison. Des ISGylations, médiées par Herc5, peuvent avoir lieu sur les lysines K20, K41, K108, K110, K126, K217 et K219. Les acides aminés hautement conservés dans NS1 sont indiqués par un astérisque (Klemm et al., 2017).*

D.II. Structure

Chaque monomère (26 kDa) est constitué d'un domaine N-terminal qui se lie à l'ARN (RNA-Binding Domain, RBD, acides aminés 1-73) et d'un domaine effecteur en C-terminal (Effector Domain, ED, acides aminés 84-207) reliés entre eux par un court linker d'une dizaine d'acides aminés (acides aminés 74-84). Une queue non-structurée de 15 à 30 acides aminés termine la protéine (CTT, **Fig.25**). NS1 peut être retrouvée dans tous les compartiments cellulaires (cytoplasme, noyau et nucléole), avec des variations pouvant être liées au moment du cycle viral, ainsi qu'au virus et au type cellulaire. NS1 est transloquée dans le noyau par l'importine α qui interagit avec la NLS de NS1 (acides aminés 35-41) (Melén *et al*., 2007). Les formes dimériques de NS1 (52 kDa) pourraient également traverser passivement les pores nucléaires en raison de leur petite taille (Timmey *et al*., 2016).

Plusieurs structures de RBD et ED seuls (**Tableau 2**) ou en complexes (**Tableau 3**) de différents sous-types viraux ont été résolues. Deux structures de NS1 entière d'un virus H5N1 et H6N6 seules et d'une NS1 de virus H1N1 en complexe ont été obtenues.

Tableau 2 : Structures de NS1 (RBD et/ou ED seuls) résolues à ce jour

Virus	Origine	RBD	ED	NS1 entière	Réference	PMID
Udorn (A/Udorn/307/1972 (H3N2))	Humain	1AIL 1NS1	3EE9 et 3EE8 2KKZ 3KWG et 3KWI		Liu et al., 1997 Chien et al., 1997 Xia et al., 2009 Aramini et al., 2009 Xia & Robertus, 2010	9360602 9360601 19153461 19995550
BM18 (A/Brevig Mission/1/1918 (H1N1))	Humain	2N74	6DGK		Jureka et al., 2015 Kleinpeter et al., 2018	26365801 30076219
WSN (A/WSN/1933 (H1N1))	Humain	2ZKO			Cheng et al., 2009	18813227
PR8 (A/Puerto Rico/8/1934 (H1N1))	Humain		3O9S, 3O9T et 3O9U 2GX9		Kerry et al., 2011 Bornholdt & Prasad, 2006	21464929 16715094
Cal09 (A/California/07/2009 (H1N1))	Humain	3M8A	3M5R		Brunzelle et al., 2010 Fremont et al., 2010	
VN1203 (A/Vietnam/1203/2004 (H5N1))	Humain	3F5T	3F5T	3F5T	Bornholdt & Prasad, 2008	18987632
KT03 (A/chicken/Kyoto/03/2004 (H5N1))	Aviaire	2Z0A	3P31, 3P38 et 3P39		Saijo et al., 2008 Chen et al., 2010	
Alb/76 (A/duck/Alberta/60/1976 (H12N5))	Aviaire		3D6R 3OA9		Hale et al., 2008 Kerry et al., 2011	18585749 21464929
MN993 (A/blue-winged teal/MN/993/1980 (H6N6))	Aviaire	4OPH et 4OPA	4OPH et 4OPA	4OPH et 4OPA	Carrillo et al., 2014	24478439
HL17 (A/little yellow-shouldered bat/Guatemala/153/2009 (H17N10)) HL17 (A/little yellow-shouldered bat/Guatemala/060/2010 (H17N10))	Chauve-souris	5BXZ 5H5N			Turkington et al., 2015 Zhao et al., 2016	26246567 27217257
HL18 (A/flat-faced bat/Peru/033/2010 (H18N11))	Chauve-souris	5BY1			Turkington et al., 2015	26246567
Influenza B (B/Lee/1940)	Humain	1XEQ			Yin et al., 2007	17475623

Tableau 3 : Structures de NS1 (RBD et/ou ED seuls) associées à des partenaires

Virus	Origine	RBD	ED	Queue C-terminale	Partenaire	Réference	PMID
WSN (A/WSN/1933 (H1N1))	Humain	2ZKO			dsRNA de 19 nt	Cheng et al., 2009	18813227
Udorn (A/Udorn/307/1972 (H3N2))	Humain	BMRB n°15117	2RHK	6O5W 4O45, 4O46, 4O42, 4NW2	dsRNA de 16 nt CPSF30 MORC3 WDR5	Yin et al., 2007 Das et al., 2008 Zhang et al., 2019 Qin et al., 2014	17475623 18725644 31006586 24853335
PR8 (A/Puerto Rico/8/1934 (H1N1))	Humain		3L4Q 5NT1, 5NT2, 6FLN et 6FLM	 5UL6 6ATV	PI3K (domaine P85β) TRIM25 CRKII (domaine N-terminal Src-homology 3)	Hale et al., 2010 Koliopoulos et al., 2018 Shen et al., 2017 Shen et al., 2018	20133840 29739942 28368102 29590589
WG47 (A/Wellington/47/1992 (H1N1))	Humain	6E5U	6E5U		NXF1 (domaines LRR et NTF2L)	Zhang et al., 2019	31263181
BM18 (A/Brevig Mission/1/1918 (H1N1))	Humain	2N74			RIG-I (domaine CARD)	Jureka et al., 2015	26365801
influenza B	Humain	3SDL, 3R66, 3RT3			ISG15	Guan et al., 2011	21808041

D.II.1. Le domaine de liaison à l'ARN

Chien *et al*. (1997) et Liu *et al*. (1997) ont résolu les premières structures du RBD seul respectivement par résonance magnétique nucléaire (RMN) et DRX. Le dimère de RBD est constitué de deux fois trois hélices α antiparallèles (α1 acides aminés 3-24, α2 acides aminés 30-50 et α3 acides aminés 54-69) (**Fig.26**). Les acides aminés R35 et R46 des deux hélices α2 et α2' du RBD de NS1 assurent la dimérisation de la protéine (Cheng *et al*., 2009; Lalime et Pekosz, 2014; Wang *et al*., 1999) alors que R38 est essentiellement impliqué dans la liaison à l'ARN (Nemeroff *et al*., 1995; Wang *et al*., 1999) (**Fig.26**). D'autres acides aminés chargés, en particulier D12/R19/D29/R35/K41/R46, participent aussi à l'interaction avec l'ARNdb. Wang *et al*. (2020) ont montré que l'acide aminé R19 joue un rôle essentiel dans les fonctions biologiques de NS1 puisque sa présence participe à la localisation nucléaire de NS1 et conditionne ses interactions avec l'ARN. La substitution R38A réduit considérablement l'interaction avec le ARNdb (Kerry *et al*., 2011). Kerry *et al*. (2011) ont aussi démontré l'implication du domaine effecteur, en particulier l'acide aminé W187, dans la formation de complexe RBD/ARN.

Les substitutions R38K et/ou K41R n'impactent par l'interaction avec l'ARNdb (Chien *et al*., 2004), indiquant que les acides aminés chargés positivement (R38 et K41) (**Fig.26**) entretiennent majoritairement des interactions de type électrostatique avec les groupements phosphates de l'ARN.

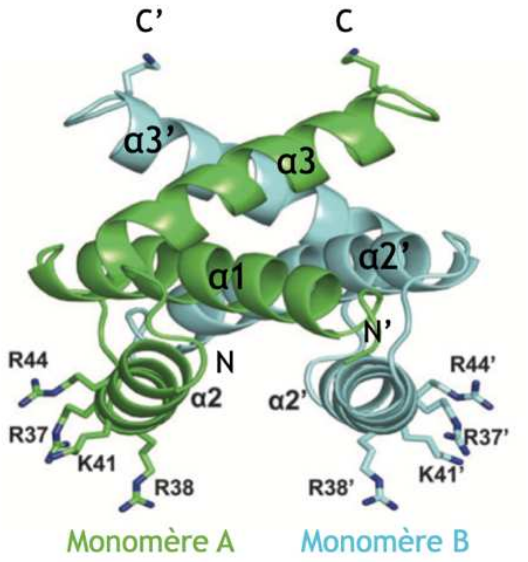


Figure 26 : Structure du domaine RBD de NS1. *Structure RMN du domaine RBD de NS1 (PDB : 1NS1, Chien et al., 1997). Les monomères A et B du RBD sont respectivement représentés en vert et en bleu. Les acides aminés très impliqués dans l'interaction avec les ARN sont indiqués. Les régions N- et C-terminale sont respectivement indiquées par N et C (adapté de Krug, 2014).*

D.II.2. Le domaine effecteur

Bornholdt et Prasad ont établi les premières structures du domaine ED seul puis de la NS1 entière (Bornholdt et Prasad, 2006; Bornholdt et Prasad, 2008). Le domaine ED est composé de trois

hélices α (α1 acides aminés 95-98, α2 acides aminés 171-187 et α3 acides aminés 196-201) et de sept brins β (β1 acides aminés 88-91, β2 acides aminés 107-112, β3 acides aminés 115-120, β4 acides aminés 127-137, β5 acides aminés 140-151, β6 acides aminés 156-162 et β7 acides aminés 191-194). Son homodimérisation est assurée par la présence de l'acide aminé W187 localisé en fin d'hélice α2 (Kerry *et al*., 2011) (**Fig.27**). Le motif $G^{179}xxxG^{183}G^{184}$, localisé dans l'hélice α2, se retrouve à l'interface de l'homodimère, assurant la stabilité et la flexibilité de ce dernier. Ce motif GxxxGG participe aussi à la dimérisation du ED (Steidle *et al*., 2010) et il permet aux hélices α2 d'être dynamiques l'une par rapport à l'autre (Kerry *et al*., 2011).

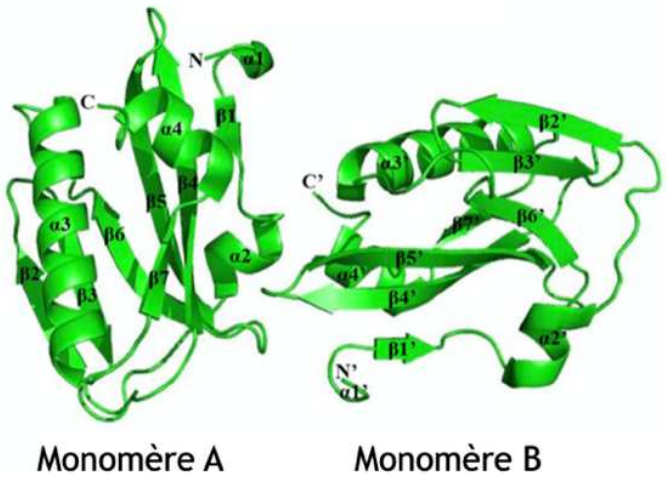


Figure 27 : Structure du domaine ED de NS1. *Structure crystallographique du domaine ED de NS1 du virus Udorn (PDB : 3EE8, Xia et al., 2009). Les régions N- et C-terminale sont respectivement indiquées par N et C (adapté de Han et al., 2019).*

Le RBD de NS1 est toujours assemblé sous la forme d'un dimère. Le domaine ED de NS1, moins compacte que son RBD, est également présent à l'état de dimère dans la cellule infectée mais dans des conformations variables. L'analyse des structures résolues par DRX semble indiquer la présence d'au moins trois conformations de NS1 (**Fig.28A**) : les conformations « ouverte » (**Fig.28B**), « semi-ouverte » (**Fig.28C**) et « fermée » (**Fig.28D**) (Kerry *et al*., 2011; Carrillo *et al*., 2014). La forme « ouverte » de NS1 (**Fig.28B**), stabilisée par l'interaction du RBD (acides aminés R59 et Q63) avec le domaine ED (acides aminés G168 et T170), serait susceptible d'interagir simultanément avec l'ARNdb et le facteur CPSF30 (Carrillo *et al*., 2014). L'analyse des modèles structuraux suggère aussi que ces changements de conformation varient en fonction du sous-type viral et de la concentration de la protéine (Kerry *et al*., 2011).

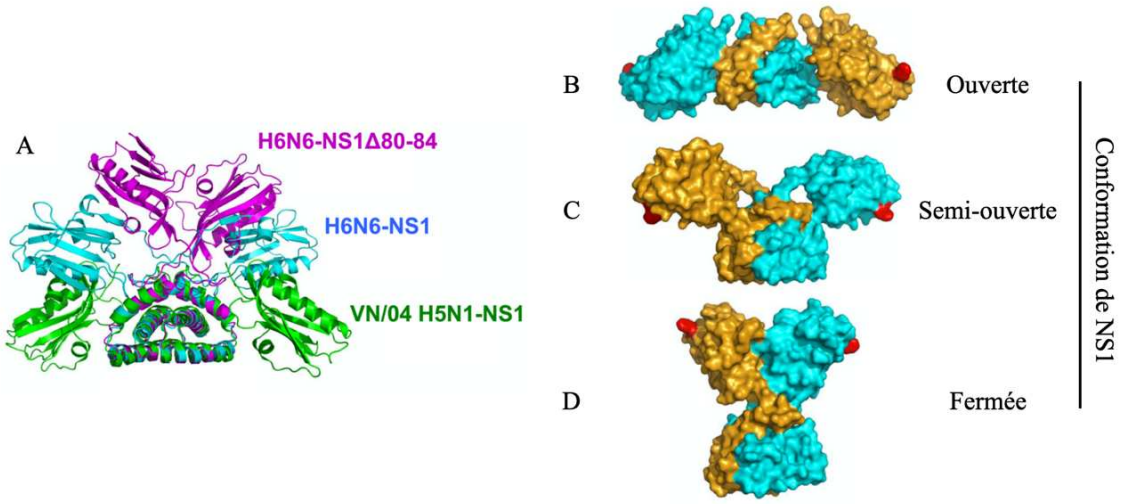


Figure 28 : Structures des différentes conformations de NS1. *(**A**) Superposition des différentes conformations de NS1 (H6N6 sauvage, bleue; H6N6 contenant la délétion 80-84, magenta; H5N1, vert; adapté de Carrillo et al., 2014). (**B**) Conformation ouverte du dimère de NS1 du virus VN04 (PDB : 3F5T, Bornholdt et Prasad, 2008) où les monomères de NS1 (RBD+ED) sont représentés en cyan et en jaune. Cette NS1 présente la déletion 80-84 située dans le linker ainsi que la mutation E71G. (**C**) Conformation semi-ouverte du dimère de NS1 d'un virus H6N6 (PDB : 4OPH, Carrillo et al., 2014) (coloration des monomères similaire à **B**). Cette conformation de NS1 est retrouvée chez de nombreux virus de sous-types non-H5N1. (**D**) Conformation fermée du dimère de NS1 d'un virus H6N6 contenant la déletion 80-84 dans le linker (PDB : 4OPA, Carrillo et al., 2014) (coloration des monomères similaire à **B**). L'acide aminé W187 (indiqué en rouge), impliqué dans la dimérisation des domaines ED de NS1, reste accessible dans les trois conformations.*

D.II.3. La région linker

Les virus aviaires H5N1 hautement pathogènes isolés depuis 2000 portent une courte délétion des acides aminés 80-84 dans le linker (acides aminés 74-84) situé entre le RBD et l'ED. La conservation dans le temps de cette délétion chez des virus H5N1 mais aussi chez d'autres sous-types viraux suggère un avantage sélectif pour le virus. D'ailleurs, cette même délétion du linker dans la NS1 d'un virus aviaire H1N1 (MZ) augmente significativement sa réplication et sa pathogénicité chez le poulet (Trapp *et al*., 2014). Cette courte délétion du linker semble également favoriser la conformation « fermée » de NS1 (**Fig.28D**) (Carrillo *et al*., 2014). Certaines substitutions d'acides aminés au sein du linker, en particulier L77A, affectent le fonctionnement et la localisation de NS1 *in cellulo* (Li *et al*., 2011). De récents isolats de virus de mammifères de sous-type H5N1 démontrent la présence des acides aminés 80-84 dans la région linker de NS1. Des analyses *in cellulo* et *in vivo* chez la souris démontrent que la présence de cette séquence 80-84 dans le linker de NS1 rend les cellules dendritiques plus susceptibles à l'infection par ces virus. La protéine NS1 de ces virus recombinants active l'expression forte de marqueurs immunitaires, comme la production d'interleukine-6, ce qui conduit au développement d'une « tempête cytokinique » dans les poumons de souris infectées qui pourrait être délétère pour l'Homme (Chen *et al*., 2020).

D.II.4. La queue C-terminale

La queue C-terminale de NS1 est flexible et non-structurée. L'alignement de multiples séquences de NS1 témoigne de la grande variabilité de séquence et de longueur de la queue C-terminale (Suarez et Perdue, 1998). Depuis les années 40 et jusque dans les années 80, la NS1 des virus humains H1N1, puis H2N2 et H3N2 portent une extension de 7 acides aminés (^{231}RRNKMAD^{237}), passant ainsi à 237 acides aminés. Une réversion dans la séquence nucléotidique de NS1 a restauré le codon stop initial, entrainant un retour à 230 acides aminés. Inversement, des délétions plus ou moins longues de la queue C-terminale ont aussi été répertoriées, en particulier la délétion des 11 derniers acides aminés de NS1 (^{220}RKMARTIESEV^{230}) qui conduit à une NS1 de 219 acides aminés. Cette délétion est naturellement retrouvée dans la NS1 de virus porcins (Hale *et al*., 2010), chez le virus pandémique de 2009 et chez des virus H7N9 hautement pathogènes isolés en Asie. Cette délétion a aussi été décrite comme exacerbant la réponse inflammatoire pulmonaire dans un model murin, occasionnant ainsi une plus grande mortalité (Tu *et al*., 2011).

L'extrémité C-terminale de NS1 présente aussi une séquence caractéristique des domaines PDZ de type I (the postsynaptic protein PSD-95/SAP90, the *Drosophila* septate junction protein Discs-large, et the tight junction protein ZO-1). Ces domaines interviennent dans de nombreuses interactions protéine-protéine et sont très retrouvés chez les végétaux, les mammifères ainsi que la drosophile (pour revue, Lee et Zheng, 2010). Le motif ^{227}ESEV^{230} contenu en fin de NS1 semble aussi influencer le déroulement du cycle viral puisque sa substitution en ^{227}GSEI^{230} ou en ^{227}EPEV^{230} augmente la virulence du virus alors les substitutions ^{227}RSKV^{230}, ^{227}RSEV^{230} et ^{227}PEQK^{230} la réduisent (Wang *et al*., 2012; Soubies *et al*., 2010).

D.III. La protéine NS1 et ses nombreuses fonctions biologiques dans la cellule infectée

D.III.1. Les fonctions de NS1 dépendantes du domaine RBD

D.III.1.a. Inhibition de la réponse interféron

La protéine RIG-I (Retinoic acid inducible-gene I) fait partie des RIG-Like Receptors (RLR), des senseurs cytoplasmiques qui détectent la présence d'ARN simple et double-brin étrangers, en particulier les ARN viraux (pour revue, Liu et Gu, 2011; Weber-Gerlach et Weber, 2016). La reconnaissance des ARNs viraux par RIG-I implique un changement de conformation de RIG-I (pour revue, Swatek et Komander, 2016), assuré par les ligases TRIM25 (Gack *et al*., 2007) et Riplet (Oshiumi *et al*., 2009), lesquelles permettent l'ubiquitination des lysines K27, K48 et K172 localisées dans les domaines de recrutement de protéases à cystéine (caspase activation and recruitment domains, CARD) (acides aminés 1-89 et 90-188) et des lysines K849 et K851 contenues dans le domaine de liaison à

l'ARN C-terminal (Gack *et al*., 2007; Oshiumi *et al*., 2009; Liu *et al*., 2017; Zhao *et al*., 2017). Le domaine CARD N-terminal de RIG-I interagit alors avec celui de la protéine MAVS (mitochondrial antiviral signaling protein) (Sanchez-Aparicio *et al*., 2017). Ce complexe entraine (i) le recrutement des facteurs de transcription 3 et 7 régulateurs de l'interféron (IFN regulatory factors, IRF3 et 7) (Sanchez-Aparicio *et al*., 2017) qui, une fois phosphorylés, induisent l'expression des gènes de l'IFN et (ii) la synthèse d'autres molécules comme le facteur NF- κβ (nuclear factor κβ) qui induit une réponse pro-inflammatoire locale. L'IFN de type I (IFN-I) produit par la cellule infectée interagit avec son récepteur transmembranaire (IFN α receptor, IFNAR) exprimé par les cellules voisines. La formation du complexe IFN-IFNAR transduit un signal intracellulaire par lequel les tyrosine kinases JAK1 (Janus protein tyrosine kinase 1) et TYK2 (Tyrosine kinase 2) phosphorylent les facteurs de transcription STAT1 et STAT2 (signal transducer and activator of transcription factors). Dès lors, le complexe hétérotrimérique qui associe STAT1 - STAT2 et IRF9 est transloqué au noyau et se fixe sur les promoteurs dotés d'éléments ISRE (IFN-stimulated response elements), induisant l'expression des gènes cibles de l'IFN, en particulier *via* l'expression de nombreuses protéines antivirales induites par l'IFN (IFN stimulated genes, ISG) (**Fig.29**) (pour revue, Sadler et Williams, 2008; Donnelly et Kotenko, 2010; Schoggins, 2014) lors de la fixation du complexe au niveau de l'élément de stimulation de réponse à l'IFN, ISRE).

NS1 réduit l'ubiquitination du domaine CARD de RIG-I (Gack *et al*., 2009) en interagissant avec les ligases TRIM25 (Rajsbaum *et al*., 2012; Koliopoulos *et al*., 2018) et Riplet (Oshiumi *et al*., 2013) (**Fig.29**). Certaines substitutions dans NS1 (en particulier R38A-K41A et E96A-E97A) limitent son interaction avec TRIM25, rétablissant ainsi l'ubiquitination de RIG-I. Une faible fraction de TRIM25 est également retrouvée dans le noyau où elle bloque la transcription du génome viral en interagissant directement avec les RNPv (Meyerson *et al*., 2017). L'interaction directe entre RIG-I et NS1 mise en évidence par co-immunoprécipitation (co-IP) (Guo *et al*., 2007; Mibayashi *et al*., 2007) a été confirmée sur le plan structurale par RMN (Jureka *et al*., 2015), révélant l'intervention de l'acide aminé R21 du RBD de NS1 dans l'interaction avec le second domaine CARD de RIG-I (Jureka *et al*., 2019).

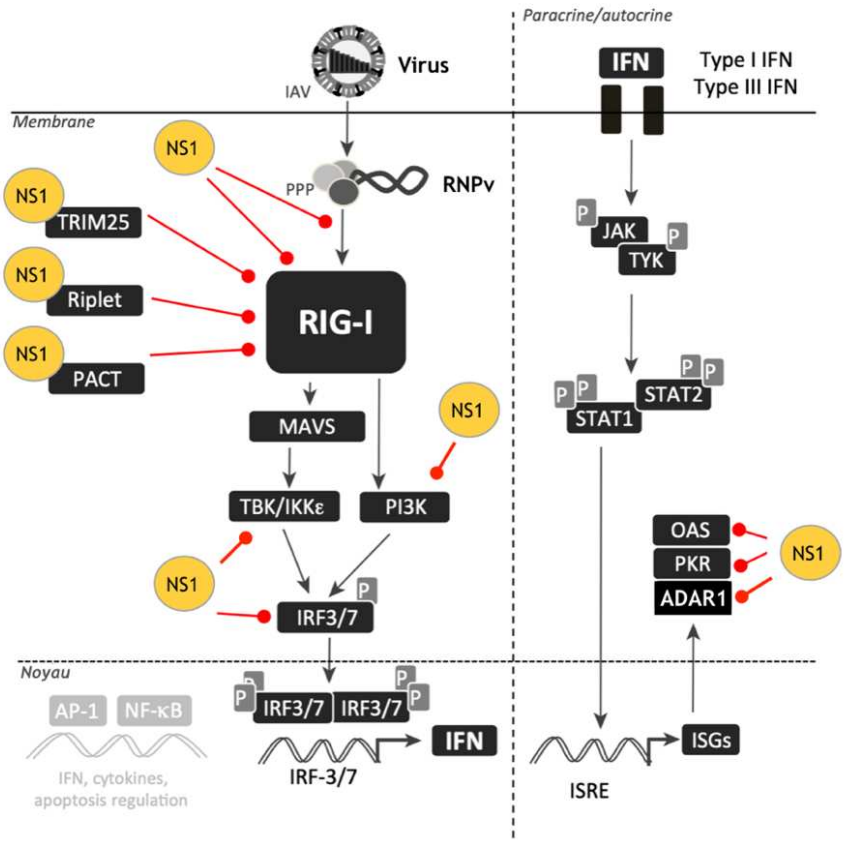


Figure 29 : Inhibition de la réponse IFN-I par NS1 *(adapté de Klemm et al., 2017)*

En outre, Zhang *et al*. (2018) ont montré que NS1 ciblait spécifiquement les introns du pré-ARNm codant pour RIG-I ainsi que ceux d'autres pré-ARNm impliquées dans la réponse antivirale. NS1 limite l'épissage du pré-ARNm de RIG-I, impactant la réponse antivirale de la cellule dans son ensemble.

D.III.1.b. Inhibition de l'activité des protéines antivirales

Près de 300 ISG sont exprimés en réponse à l'IFN. Certaines des protéines produites, en particulier la protéine kinase R (PKR), le complexe 2'-5' oligoadenylase synthetase (OAS)-Ribonucléase L (RNase L) et la riboadénosine désaminase de type 1 (ADAR1) perturbent le cycle viral, en particulier celui des virus *influenza*, en modifiant leur séquence nucléotidique et en les dégradant (pour revue, Schoggins, 2014). Plusieurs de ces protéines sont ciblées par NS1.

NS1 inhibe l'activité de la protéine kinase R (PKR)

La protéine kinase R (PKR) est activée sous forme de dimère par sa protéine activatrice PACT (PKR-activating protein) (pour revue, D'acquisto et Ghosh, 2001) et par l'interaction avec des ARNdb (Patel *et al*., 1995). La PKR activée phosphoryle le facteur eIF2α (Dar *et al*., 2005), provoquant alors le blocage du ribosome sur les ARNm en cours de traduction (pour revue, Gal-Ben-Ari *et al*., 2019). Li *et*

al. (2006) ont montré que le RBD de NS1 interagissait avec PACT *via* les acides aminés R35 et K46 du RBD (Lu *et al*., 1995; Schierhorn *et al*., 2017), empêchant l'activation de PKR.

NS1 inhibe l'activité des protéines oligoadénylates synthases (OAS)

Les protéines oligoadénylates synthétases 2'5' (OAS) sont des senseurs des ARNdb dans la cellule. Leur interaction avec un ARNdb provoque la formation de tétramères d'OAS fonctionnels, qui polymérisent en 2'5' des ARNdb deux à trois molécules d'ATP. Ces oligoadénylates entrainent la dimérisation et l'activation de la RNase L, laquelle dégrade alors les ARNdb (pour revue, Drappier et Michiels, 2015). De nombreux virus à ARN et à ADN limitent partiellement ou totalement le fonctionnement d'OAS en ciblant directement la protéine OAS et/ou la RNase L. Les virus *influenza* inhibent la réponse OAS en séquestrant les ARNdb viraux. Min et Krug (2006) ont montré que l'invalidation du RBD (R38A) entraine la reconnaissance et la dégradation des ARNs viraux par le complexe OAS-RNase L. Plus récemment, Zhang *et al*. (2018) ont montré que NS1 interagissait avec les introns de plusieurs centaines de pré-ARNm cellulaires, incluant ceux de OAS.

NS1 inhibe l'activité de la riboadénosine désaminase 1 (ADAR1)

Les protéines riboadénosines désaminases (Adenosine Desaminase that act on RNA, ADAR) convertissent certaines adénines (Ad) de l'ARN en Inosine (In) (pour revue, Wang *et al*., 2017). L'inosine ressemble de très près à une guanine (Gu) et la conversion Ad→In modifie donc la séquence et la structure de l'ARNm, ainsi que la séquence de la protéine traduite (pour revue, Wang *et al*., 2017). Ngamurulert *et al*. (2009) ont montré par double-hybride en levure l'interaction de la NS1 d'un virus H5N1 avec ADAR1. Une seconde étude menée par De Chassey *et al*. (2013) a confirmé cette interaction, y compris pour des NS1 de virus variés. En interagissant avec ADAR1 *via* son RBD, NS1 limiterait l'activité de conversion A→I d'ADAR1 qui serait en temps normal délétère pour le virus.

D.III.1.c. Réduction de l'expression des gènes cellulaires par inhibition de l'épissage des pré-ARNm cellulaires

Le spliceosome, un complexe ribonucléoprotéique composé des cinq snRNP (small nuclear Ribonucleoprotein) U1/U2/U4/U5 et U6 ainsi que d'autres facteurs (**Fig.30**) (pour revue, Plaschka *et al*., 2019), épisse les pré-ARNm de leurs introns. Ceci est réalisé à la suite de plusieurs étapes au cours desquelles (i) U1 se fixe sur le site donneur d'épissage situé en 5' de l'intron, (ii) U2 interagit avec le point de branchement, (iii) le complexe trimérique associant U4/U6 et U5 interagit avec U1 et U2 rapprochant alors les deux exons, (iv) U1, d'abord, puis U4 se détachent du complexe avant que (v) le complexe trimérique U2/U5/U6 délète l'intron sous la forme d'un lariat et joint les deux exons (**Fig.30**) (pour revue, Mittendorf *et al*., 2012).

Initialement décrite pour son interaction avec l'ARN U6 (Lu *et al*., 1994), NS1 déstabilise, majoritairement *via* son RBD, la formation des complexes U6-U2 et U6-U4 (Nemeroff *et al*., 1992; Qiu

et al., 1995), provoquant ainsi le blocage du processus d'épissage (**Fig.30**). NS1 pourrait également entrer en compétition avec U1c en reconnaissant le site donneur d'épissage. NS1 a été décrite comme inhibant l'épissage de la classe majeure de pré-ARNm qui est caractérisée respectivement par la présence des bases GuUr et AdGu au niveau des sites donneur et accepteur d'épissage. NS1 limite aussi l'épissage de classes mineures de pré-ARNm (interaction avec U6atac), qui est caractérisée respectivement par la présence des bases AdTh et AdCy au niveau des sites donneur et accepteur d'épissage (Wang et Krug, 1998). Différemment, NS1 stimule l'épissage de l'ARNm viral codant pour M1, modulant ainsi le rapport en protéines M1/M2. À l'inverse, NS1 ne semble pas agir sur l'épissage de son propre ARNm (Robb *et al*. 2010; Pereira *et al*., 2017; Robb et Fodor, 2012).

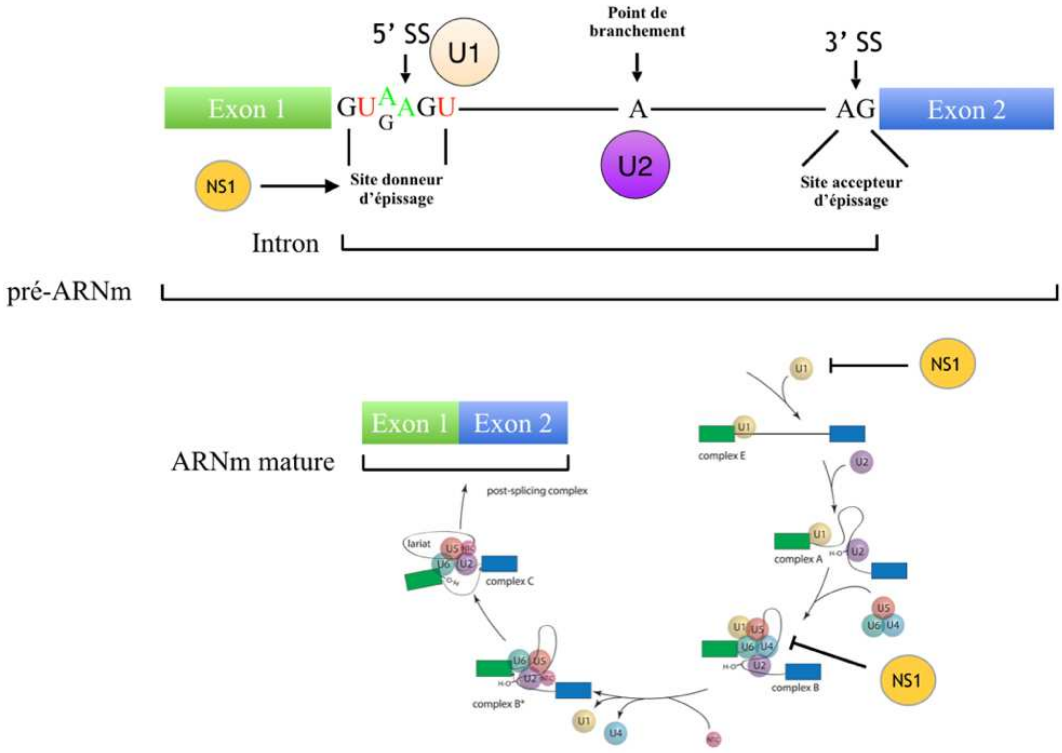


Figure 30 : Inhibition de l'épissage des pré-ARNm cellulaires par NS1 *(adapté de Mittendorf et al., 2012)*

D.III.1.d. Stimulation de la traduction des ARNm viraux

L'initiation de la traduction implique plusieurs facteurs cellulaires, notamment les protéines PABP et les facteurs d'initiation de la transcription (eIF). Les protéines PABP fixées sur l'ARNm recrutent plusieurs facteurs eIF, dont eIF4E qui reconnaît la coiffe des ARNm (pour revue, Batool *et al*., 2019). Le facteur eIF4G fait le pontage entre eIF4E et PABP. Trois autres facteurs, eIF4A, eIF4B et eIF4F, stabilisent le complexe protéine/ARNm. Un complexe multimérique associant eIF1, eIF3 et eIF5 à la petite sous-unité du ribosome se fixe en 5' de l'ARNm en amont du codon start AUG, recrutant à son tour la grande sous-unité du ribosome pour initier la traduction (pour revue, Sonenberg et Hinnebusch, 2009).

NS1 interagit directement avec le facteur eIF4G (Aragón *et al*., 2000; Burgui *et al*., 2003) ainsi qu'avec PABPII (Tu *et al*., 2011). PABPI a aussi été décrite comme un partenaire de NS1 (Burgui *et al*., 2003) mais cette interaction nécessite la formation de complexe NS1-ARN (Arias-Mireles *et al*., 2018). L'interaction de NS1 avec eIF4G, PABPI et PABPII stabilise le repliement des ARNm viraux ce qui favorise leur traduction (Dubois *et al*., 2019), en particulier pour les transcrits de M1 et de NP (Enami *et al*., 1994; Park et Katze, 1995). Panthu *et al*. (2017) ont démontré (i) un effet stimulant de NS1 sur la traduction de gènes rapporteurs de façon promoteur dépendant et (ii) que NS1 interagissait avec la petite sous-unité du ribosome (40S) ainsi qu'avec les ARN ribosomaux 18S et 28S, conduisant à la formation de polysomes (Anastasina *et al*., 2014; De la Luna *et al*., 1995). De part ces nombreux rôles dans la cellule infectée, la protéine NS1 régule la production de particules virales (Garaigorta *et al*., 2005). De Rozières et Joseph (2020) ont montré que le dimère de NS1 interagit avec PABPI lorsque ce dernier est dissocié de l'ARN.

D.III.2. Les fonctions de NS1 dépendantes du domaine ED

D.III.2.a. Activité anti-apoptotique

Les cellules infectées dont le système IFN est fonctionnel limitent rapidement la propagation du virus par l'intermédiaire de protéines antivirales (PKR, OAS et ADAR1). Toutefois et puisque l'induction du système IFN est bloquée en présence de NS1 entrainant l'absence d'une réponse antivirale efficace, la cellule n'a d'autre choix que d'entrainer sa propre mort afin de limiter la propagation de l'infection.

NS1 stimule la voie Akt/PKB

La mort cellulaire induite par apoptose fait intervenir de nombreuses protéines effectrices, en particulier les phosphoinositide-3-kinases (PI3K) (pour revue, Bilanges *et al*., 2019). Lorsqu'elles sont activées, les PI3K (composées des sous-unités p85 et p110) phosphorylent certains phospholipides de la face interne de la membrane plasmique de la cellule, notamment le phosphatidylinositol-4,5-biphosphate (PIP(4,5)2) (PIP2). Une fois phosphorylé, le PIP2 devient le PIP3, lequel induit l'activation de plusieurs cascades intracellulaires, en particulier la voie Akt/PKB qui stimule l'expression de gènes cellulaires impliqués dans la survie de la cellule, inhibant ainsi l'apoptose cellulaire (**Fig.31**) (pour revue, Manning et Toker, 2017). Plusieurs études ont décrit NS1 comme modulant l'apoptose cellulaire, en interagissant avec PI3K. Seul le domaine ED de NS1 est essentiel à l'interaction (Hale *et al*., 2010; pour revue, Ehrhardt et Ludwig, 2009). NS1 a aussi été montrée par co-immunoprécipitation (co-IP) comme interagissant avec Crk(L), la protéine adaptatrice de PI3K (Ylösmäki *et al*., 2015). L'interaction de NS1 avec PI3K conduit à la sur-activation de la voie Akt/PKB, ce qui inhibe l'apoptose cellulaire. Le développement d'une technologie basée sur l'utilisation d'une bibliothèque de distribution équitable

des ARNm cellulaires extraits du génome cellulaire de deux patients a permit d'identifier la protéine AKT1 comme partenaire de NS1 (Du *et al*., 2020).

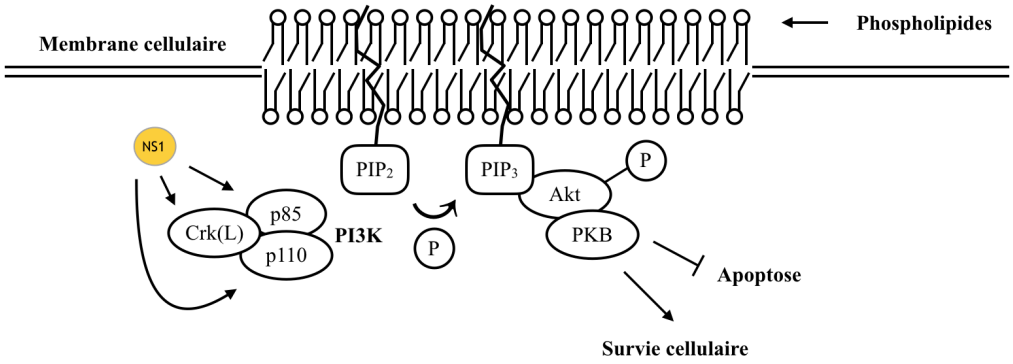


Figure 31 : Stimulation de la voie Akt/PKB par NS1

NS1 inhibe la voie p53

Le facteur p53 est un facteur de transcription clé qui est activé en présence de stress cellulaires intenses (déficience nutritive, hypoxie, dommages à l'ADN, détection d'oncogènes et stress oxydatif). Dans le cas de dommages à l'ADN, p53 provoque l'arrêt du cycle cellulaire et le recrutement de facteurs de réparation des dommages occasionnés dans l'ADN. Dans le cas où les dommages seraient irréversibles ou que les systèmes de réparation de l'ADN sont saturés, p53 induit alors l'apoptose cellulaire (pour revue, Niazi *et al*., 2018). L'interaction de NS1 avec p53, démontrée par co-IP et dans des cellules transfectées avec un système rapporteur (Wang *et al*., 2010; Terrier *et al*., 2013), pourrait inhiber p53 et empêcher l'apoptose des cellules infectées. NS1 a aussi été décrite comme limitant l'épissage des transcrits de p53 (Dubois *et al*., 2019) ce qui limite *in fine* les modes d'action de p53 dans la cellule infectée.

D.III.2.b. Inhibition de la maturation des pré-ARNm cellulaires

Les facteurs de clivage et de polyadénylation des pré-ARNm cellulaires (Cleavage and Polyadenylation specificity factor, CPSF) interviennent en amont de l'épissage des pré-ARNm (pour revue, Mandel *et al*., 2008). Une fois transcrits par Pol II, les pré-ARNm sont pris en charge par le complexe oligomérique CPSF à 6 composants associant les facteurs CPSF30 et WDR33, les facteurs CPSF160 et hFip1 et les facteurs CPSF73 et CPSF100. Après reconnaissance de la séquence UrUrAdUrUrUr par le complexe CPSF, le clivage s'opère en aval d'une région riche en Ur ou en Gu/Ur. Après clivage des pré-ARNm, la poly-A polymérase (PAP), modulée par la PABP (PAP-binding protein), ajoute 100 à 250 A (Clerici *et al*., 2018; Sun *et al*., 2018).

NS1 inhibe la maturation des pré-ARNm, en interagissant avec CPSF30 (Nemeroff *et al*., 1998). Plusieurs acides aminés du domaine ED de NS1 sont impliqués dans cette interaction, en particulier les

acides aminés 144-186 ainsi que F103 - M106 - K108 - D125 et D189 (Dankar *et al*., 2013; Das *et al*., 2008; Hale *et al*., 2010; Noah *et al*., 2003; Twu *et al*., 2006). NS1 interagit aussi avec CPSF2 et CPSF7 (Kuo *et al*., 2018), ainsi qu'avec la protéine PABPII impliquée dans la polyadénylation des pré-ARNm. Le blocage de la maturation des pré-ARNm cellulaires entraine leur rétention nucléaire (pour revue, Jurado *et al*., 2014).

D.III.2.c. Blocage du transport nucléocytoplasmique des ARNm cellulaires

Outre l'inhibition de la maturation des ARNm cellulaires, NS1 inhibe aussi leur export nucléaire en interagissant avec les protéines NXF1 (Nuclear RNA export Factor 1) et NXT1 (Nuclear Transport Factor 2) (**Fig.32**). Satterly *et al*. (2007) ont montré l'effet inhibiteur de NS1 sur le transport d'ARNm cellulaires dans des cellules MDCK infectées qui se traduit par leur accumulation dans le noyau des cellules infectées. Le blocage du transport des ARNm cellulaires par NS1 requiert ses deux domaines fonctionnels (Qian *et al*., 1994). Zhang *et al*. (2019) ont récemment mis en évidence l'interaction directe de NS1 avec NXF1, ce qui empêche la formation de complexes NXF1-NXT1 et provoque la séquestration des ARNm cellulaires dans le noyau (Fortes *et al*., 1994; Zhang *et al*., 2012; pour revue, Kuss *et al*., 2013). En séquestrant, indirectement, les ARNm cellulaires dans le noyau, la protéine NS1 permet à la RdRp virale d'en récupérer la coiffe méthylée, *via* le *cap-snatching*, pour la transcription des ARNm viraux (Walker *et al*., 2019). La formation du complexe NXF1-NS1 permet à NS1 de stimuler le transport des ARNm viraux (Alonso-Caplen et Krug, 1991; Alonso-Caplen *et al*., 1992; Larsen *et al*., 2014), en particulier pour les transcrits non-épissés de M1 (Pereira *et al*., 2017). Plus récemment, Zhang *et al*. (2018) ont également montré que NS1 régulait positivement l'export des ARNm viraux en interagissant avec le domaine Kelch de la protéine NS1-BP (Guo et Liu, 2018), cette dernière étant très impliquée dans la machinerie d'épissage et de transport des ARNm viraux codant pour la protéine M (Zhang *et al*., 2018). Le silencing de la protéine NS1-BP inhibe fortement et spécifiquement l'export des ARNm viraux de M sans altérer l'export des ARNm cellulaires (Esparza *et al*., 2020).

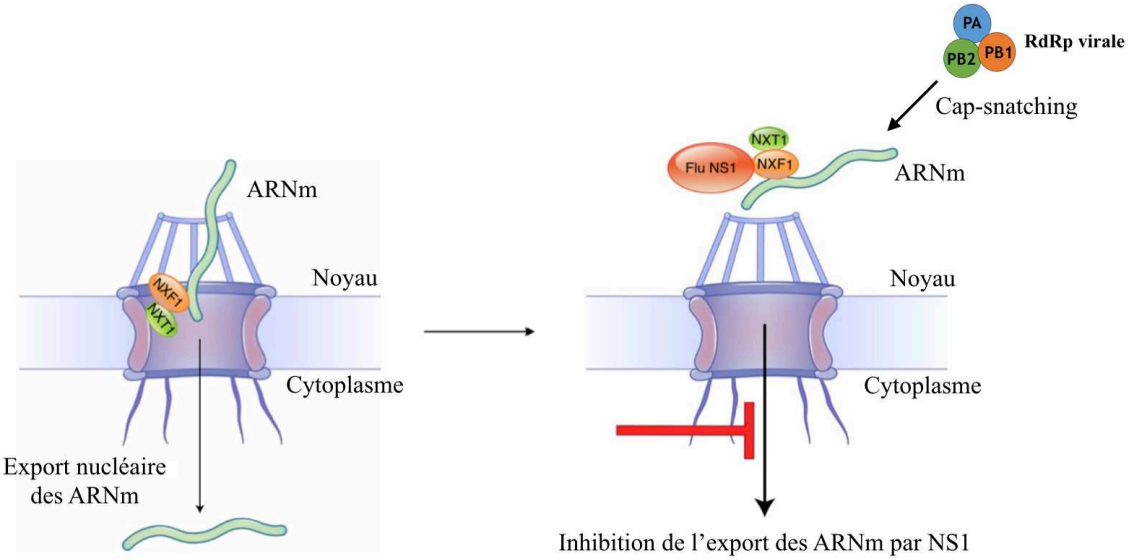


Figure 32 : Inhibition de l'export nucléaire des ARNm cellulaires par NS1 *(adapté de Zhang et al., 2019)*

D.III.2.d. Remodelage de l'architecture 3D de la chromatine par élongation de la transcription génomique

La chromatine, support de l'information génomique, est composée de nombreux nucléosomes associant une paire d'histones H2A, H2B, H3 et H4 autour desquels s'enroule 146 paires de bases, l'ensemble étant stabilisé par l'histone H1 (McGinty et Tan, 2015). L'état compacté de la chromatine, assuré par le rapprochement des nucléosomes les uns auprès des autres, limite la fixation de Pol II sur les promoteurs de gènes cellulaires, entrainant ainsi l'arrêt de la transcription des gènes cibles. Des études portant sur le niveau transcriptionnel de macrophages et de monocytes infectés par le virus VN04 (H5N1) ont permis de mettre en évidence le remodelage de l'architecture du génome cellulaire en cours de transcription, en particulier au niveau des gènes impliqués dans l'expression d'IFN, d'IFIT ou encore de STAT1 (Heinz *et al*., 2018). Le remodelage de la chromatine est directement imputable à NS1 puisque son absence (en utilisant un virus déficient pour NS1) rétablit le niveau transcriptionnel de ces gènes. La protéine NS1 des virus *influenza* entraine un remodelage de cette architecture en (i) recrutant le complexe d'initiation de la traduction *via* l'interaction avec les facteurs PABP et eIF4G et en (ii) bloquant la maturation des pré-ARNm cellulaires *via* l'interaction avec CPSF30 ainsi qu'avec les protéines du spliceosome, ce qui provoque la rétention nucléaire des transcrits cellulaires (matures et immatures). La transcription, qui est régulée activement par des facteurs inhibant l'interaction de Pol II avec les régions chromatiniennes accessibles, est maintenue en présence de NS1 car cette dernière réduit l'interaction des facteurs d'arrêt de la transcription avec la chromatine (Heinz *et al*., 2018). L'étude menée par Bauer *et al*. (2018), qui a donné des résultats similaires dans des cellules HEK293, MDCK et A549, a permis de montrer le rôle crucial joué par NS1 dans le maintien de la transcription des gènes mais qui ne semble pas reposer entièrement sur sa capacité à lier plus ou moins fortement le facteur CPSF30. *In fine*, ce remodelage chromatinien, qui assure la transcription en continue de pré-ARNm

cellulaires, est profitable au virus puisque le *pool* de transcrits cellulaires retenus dans le noyau est détourné par la RdRp virale (*via* le *cap-snatching*, voir paragraphe B.IV.3.a) pour permettre de transcrire l'intégralité du génome viral, ce qui limite l'expression des gènes impliqués dans l'immunité antivirale par la même occasion.

Plusieurs études ont également décrit la capacité de NS1 à interagir avec les histones de la chromatine, limitant ainsi la transcription des gènes cellulaires, en particulier ceux impliqués dans la réponse antivirale (Marazzi *et al*., 2013). La séquence ^{226}ARSK^{229} de la région C-terminale de NS1 des virus H3N2 humains est très proche de celle retrouvée en position N-terminale de l'histone H3 (^{1}ARTK^{4}), permettant ainsi à NS1 d'interagir avec plusieurs facteurs cellulaires qui interagissent spécifiquement avec H3, en particulier les facteurs du complexe de transcription-élongation PAF1C. NS1 déstabilise la transcription des gènes cellulaires en interagissant avec la protéine associée à Pol II (PAF1), la parafibromine (CDC73), le facteur associé à la matrice-actine dépendant SWI/SNF (SNF2I) et la protéine chromodomaine-hélicase 1 interagissant avec l'ADN (CHD1) (Qin *et al*., 2014). A l'inverse, le mutant K229R de NS1 limite son interaction avec les composants du complexe PAF1C (Marazzi *et al*., 2013). L'ATPase Microrchidia 3 (MORC3), qui régule le remodelage de la chromatine en interagissant avec H3, est également ciblée par NS1 (Zhu et Qin, 2019), permettant ainsi à NS1 de d'usurper l'identité de H3 et de moduler le mécanisme de transcription des gènes (Tarakhovsky et Prinjha, 2018).

D.IV. Les inhibiteurs de NS1

Des souris infectées avec un virus ΔNS1 ne présentent pas de symptômes et ont une réponse antivirale fonctionnelle (Garcia-Sastre *et al*., 1998). Ceci témoigne de l'importance de NS1 pour le bon déroulement du cycle viral, soulignant la nécessité d'identifier des composés capables d'inhiber les fonctions multiples de NS1, en particulier en ciblant le RBD de NS1.

D.IV.1. Les molécules ciblant le domaine RBD

Cibler la poche d'interaction du RBD avec l'ARN semble une approche pertinente (Yin *et al*., 2007). Plusieurs molécules ciblant cette interface ont été identifiées dans des banques de molécules, en particulier le gallate d'épigallocatéchine (EGCG) (**Fig.33**) (Cho *et al*., 2012; Maroto *et al*., 2008; You *et al*., 2011), une molécule extraite du thé vert qui était décrite pour ses effets anticancéreux (Sun *et al*., 2019) et anti-VIH (Li *et al*., 2011). Cette molécule est active sur tous les sous-types viraux à des concentrations de l'ordre du micromolaire.

D.IV.2. Les molécules ciblant le domaine ED

Le domaine ED interagit avec de très nombreuses protéines cellulaires, en particulier le facteur de clivage et de polyadénylation de 30 kDa (Cleavage and Polyadenylation specificity factor 30, CPSF30) (Nemeroff *et al*., 1998).

En utilisant un test de croissance de levures exprimant NS1 de façon inductible au galactose (Ward *et al*., 1994), Basu *et al*. (2009) ont identifié plusieurs molécules capables d'inhiber l'activité de la protéine NS1. *Parmi elles*, l'EGCG (**Fig.33**), déjà identifié auparavant, et le composé NSC125044 (**Fig.33**), réduisent la charge virale dans des MDCK infectées (Patnaik *et al*., 2013; 2019). Plusieurs dérivés du composé NSC125044 réduisant le titre viral ont été mis au point, en particulier les composés A9 (JJ3297) (EC_{50} = 0,8 µmol.L^{-1}) et A22 (EC_{50} = 0,08 µmol.L^{-1}) (**Fig.33**) (Jablonski *et al*., 2012; Walkiewicz *et al*., 2011) qui, en se fixant au niveau des brins β2 et β3 (acides aminés 106-110 et 114-118, respectivement) et de l'hélice α2 (acides aminés 171-186) du domaine ED de NS1 (Kleinpeter *et al*., 2018), entrent en compétition avec le facteur CPSF30 rétablissant ainsi le système de maturation des pré-ARNm cellulaires. Le composé HENM-18 (**Fig.33**), dérivé lui-aussi du composé NSC125044, bloque fortement la réplication virale (Barman *et al*., 2014). En criblant une large banque comprenant 270 000 composés (NIH Molecular Libraries-Small Molecule Repository, MLSMR), Patnaik *et al*. (2019) ont identifié près de 200 composés présentant un effet anti-NS1 *in cellulo*. L'une de ces molécules, la pyrazolopyridine (**Fig.33**), réduit fortement la réplication du virus PR8 dans des cellules MDCK.

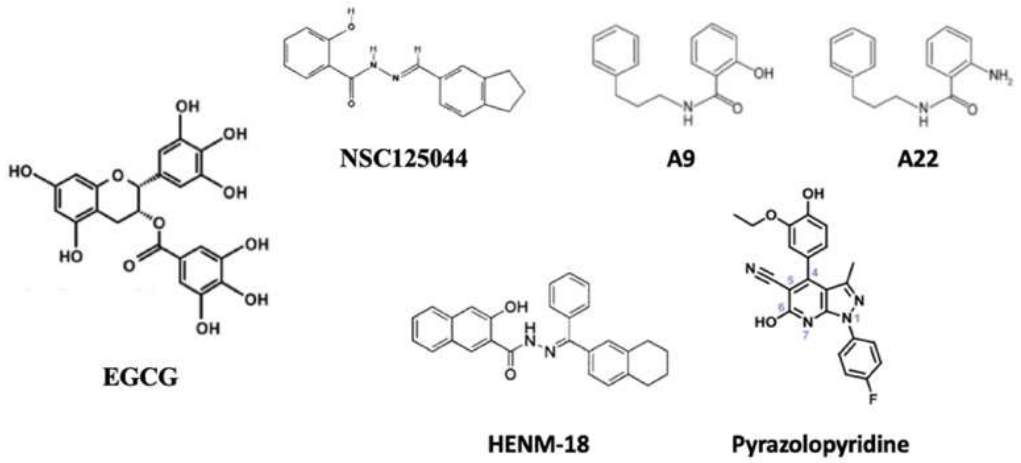


Figure 33 : Inhibiteurs ciblant la protéine NS1

D.IV.3. Les inhibiteurs identifiés par *docking* virtuel

L'oseltamivir et le zanamivir (**Fig.20**), deux inhibiteurs ciblant la protéine NA, ont été développées par une approche de *docking* virtuel, soulignant l'utilité des outils de prédiction informatique dans l'identification de nouveaux inhibiteurs. Chandra *et al*. (2017) ont identifié plusieurs

composés potentiellement inhibiteurs de NS1, en particulier les composés BAS 02862810, BAS 02912613, BAS 08978615 et BAS 01947678 (**Fig.34C**) qui seraient susceptibles d'interagir avec les acides aminés situés dans l'interface d'interaction avec l'ARN, en particulier l'acide aminé R38 (**Fig.26 et 34A**). Quatre molécules ciblant potentiellement le domaine ED ont été identifiées, à savoir les composés BAS 00337509, BAS 09627795, BAS 09627796 et BAS 09627799 (**Fig.34D**) qui pourraient interagir avec les acides aminés Q121 - G184 et N188 du domaine ED de NS1 (**Fig.34B**). Aucune étude n'a encore évalué l'efficacité de ces molécules identifiées par *docking*.

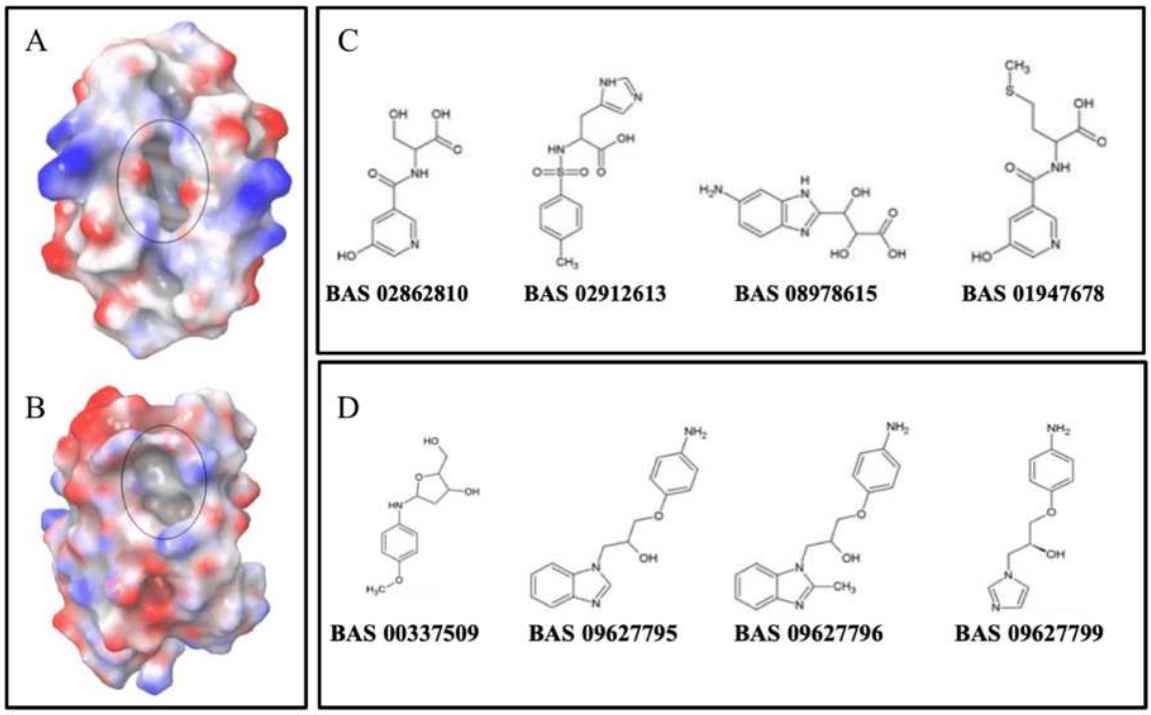


Figure 34 : Inhibiteurs potentiels de NS1 identifiés par *docking*. *(**A**) Poche du domaine RBD de NS1 utilisée pour le développement d'antiviraux. (**B**) Poche du domaine ED de NS1 utilisée pour le développement d'antiviraux. Les poches d'interaction avec des inhibiteurs potentiels sont encerclés (générés avec la suite de logiciels Schrodinger, Chandra et al., 2017). (**C**) Inhibiteurs potentiels du domaine RBD de NS1. (**D**) Inhibiteurs potentiels du domaine ED de NS1.*

SITUATION ET PROBLEMATIQUE DE LA THESE

L'ARN ou acide ribonucléique se compose de plusieurs bases hétérocycliques azotées (Adénine (Ad), Guanine (Gu), Cytosine (Cy) ou Uracile (Ur)) qui sont liées à un ou plusieurs groupements phosphates par l'intermédiaire d'un ribose. Etudier les mécanismes permettant de lier une protéine à de l'ARN peut avoir des applications biologiques dans la compréhension de processus physiologiques et/ou pathologiques. Les prédictions bio-informatiques réalisées sur l'ensemble du génome humain estiment à plusieurs centaines le nombre de protéines susceptibles de se lier à de l'ARN (RBP). Ces protéines sont au cœur de la régulation de l'expression génique, en particulier en maturant les pré-ARNm cellulaires et/ou en régulant la durée de vie, la localisation et la traduction des ARNm cellulaires. Certaines de ces RBPs peuvent aussi interagir avec leur propre ARNm régulant ainsi leur niveau d'expression (pour revue, Ganser *et al*., 2019). Les RBPs peuvent se lier à l'ARN de diverses façons, mais toujours *via* l'interaction d'un ou plusieurs domaines RBD avec l'ARN. L'étude approfondie de ces RBDs a permis d'identifier des motifs canoniques capables de reconnaitre l'ARN de façon séquence et/ou structure dépendante (Tang *et al*., 2019). Les méthodes analytiques de structures moléculaires développées depuis la fin du XX^{ème} siècle ont permis d'identifier plusieurs modes de reconnaissance des ARN simple-brin (ARNsb) et des ARN double-brin (ARNdb) (pour revue, Lunde *et al*., 2007). Les domaines de liaison aux ARNsb, à savoir le RNA-recognition motif (pour revue, Maris *et al*., 2005), les domaines à doigt de Zinc (ZnF) (pour revue, Klug, 2010), les domaines homologues à K (pour revue, Nicastro *et al*., 2015) et à Pumilio (pour revue, Rubinson et Eichman, 2012), les pentatricopeptides repeat proteins (Manna, 2015), les domaines OB-fold (Theobald *et al*., 2003) et les domaines DEAD-box (pour revue, Hardin *et al*., 2010), reconnaissent généralement une séquence de 2 à 4 bases dans l'ARN (**Fig.35A-D**) (pour revue, Auweter *et al*., 2006).

Contrairement à la reconnaissance des ARNss, l'interaction avec un ARNdb implique généralement la reconnaissance par le domaine dsRBD de sa structure indépendamment de sa séquence nucléotidique (pour revue, Tian *et al*., 2004). Les dsRBDs, qui sont les domaines les plus abondants après les ssRBDs RRM (pour revue, Maris *et al*., 2005) et ZnF (pour revue, Klug, 2010), sont retrouvés dans des protéines impliquées dans l'édition des pré-ARNm cellulaires comme la protéine ADAR et certaines protéines du spliceosome (Nagai, 1996), le transport nucléocytoplasmique d'ARNm cellulaires comme la protéine Staufen (pour revue, Banerjee et Barraud, 2014), la maturation des miARN comme la protéine Dicer (Burger et Gullerova, 2015) ou la réponse antivirale comme la protéine PKR (pour revue, Saunders et Barber, 2003). La résolution de la structure tridimensionnelle de plusieurs complexes dsRBP-ARNdb a mis en évidence des domaines canoniques et non-canoniques impliqués dans la reconnaissance des ARNdb.

Les dsRBDs canoniques, illustrés par les structures de la RNase III d'*E.coli* (Kharrat *et al*., 1995) et de la protéine Staufen de la drosophile (dStaufen) (Lazzaretti *et al*., 2018), ont une topologie αββββα (pour revue, Chang *et al*., 2005). A ce jour, plus d'une trentaine de structures tridimensionnelles de complexes dsRBP-ARNdb ont été résolues (Carlson *et al*., 2003; Masliah *et al*., 2018; Ryter et Schultz, 1998; Yamashita *et al*., 2011, pour revue, Chang *et al*., 2005). Dans les dsRBDs canoniques

(environ 70 acides aminés), des acides aminés aliphatiques et aromatiques, en particulier l'acide aminé E8 de l'hélice α1, le motif GPxH de la boucle L2 qui sépare les brins β1 et β2 et le motif KKxAK de l'hélice α2, assurent l'interaction avec le ARNdb (pour revue, Gleghorn et Maquat, 2014; Masliah *et al*., 2013) (**Fig.35E-G**).

Les dsRBDs non-canoniques des protéines à domaine ZnF, comme le facteur de transcription IIIA (TFIIIA) de *Xenopus laevis*, peuvent se lier aux ARNdb *via* l'intervention de motifs composés de deux cystéines et de deux histidines (C_2H_2) qui reconnaissent les phosphates exposés dans le grand sillon de l'ARN. Des acides aminés basiques peuvent également participer à l'interaction en interagissant avec les phosphates exposés dans le petit sillon de l'ARN (Lee *et al*., 2006; Lu *et al*., 2003). Certains dsRBDs non-canoniques peuvent aussi être composés uniquement d'hélices α (de deux à quatre hélices α) qui s'assemblent généralement en un homodimère, dans lequel des acides aminés basiques interagissent avec le squelette de l'ARNdb sans en reconnaître spécifiquement la séquence nucléotidique. Plusieurs structures de complexes associant un ARNdb avec un dsRBD de protéine virale ont été résolues à ce jour, en particulier pour la protéine B2 du virus Flock House (**Fig.35H**) (Chao *et al*., 2005; Lingel *et al*., 2005), pour la protéine NS3 du virus du riz à feuilles blanches (Yang *et al*., 2011) ou encore pour la protéine VP35 du virus Ebola (Bale *et al*., 2013; Zinzula *et al*., 2012). Dans tous les cas, des acides aminés basiques interagissent avec les charges négatives du squelette de l'ARNdb qui, en solution, adopte une conformation en double hélice de forme A ce qui favorise ainsi l'accessibilité de son petit sillon mais réduit celle de son grand sillon (pour revue, Stefl *et al*., 2005; Liu *et al*., 2019). La fonction –OH libre du carbone C2 (2'OH) du ribose interagit aussi avec certains acides aminés des protéines. Ainsi, le dsRBD reconnait préférentiellement le squelette ribose-phosphate plutôt que les bases du ARNdb.

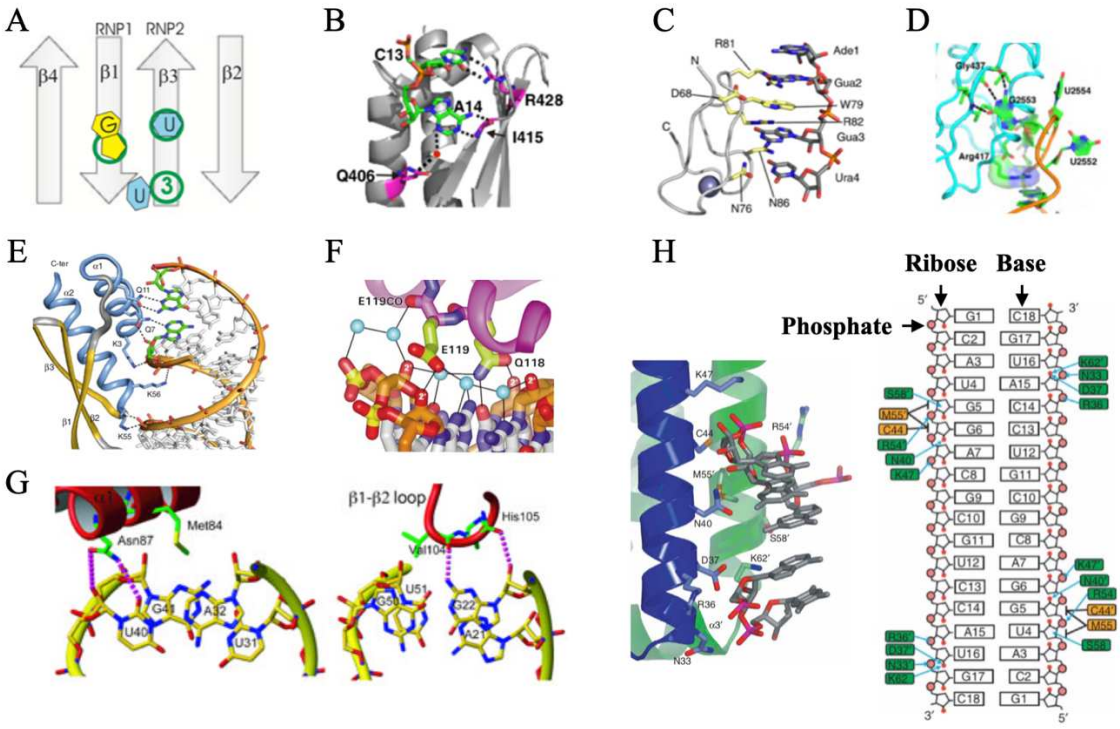


Figure 35 : Modes de reconnaissance des ARNsb et des ARNdb par les RBP. *(**A-D**) Reconnaissance des ARNsb par les ssRBP : (**A**) Domaine RRM2 de la protéine Sex-lethal Tra (PDB : 1B7F, Handa et al., 1999); (**B**) Domaine KH3 de la protéine NOVA2 (PDB : 1EC6, Lewis et al., 2000); (**C**) Protéine Ran-BP2 (PDB : 3G9Y, Nguyen et al., 2011); (**D**) Protéine YxiN de Bacillus subtilis (PDB : 2G0C, Wang et al., 2006). (**E-G**) Reconnaissance des ARNdb par les dsRBP à domaine canonique : (**E**) Protéine RNase III d'Aquifex aeolicus (PDB : 1YYW, Gan et al., 2006); (**F**) Protéine Xlrbpa (PDB : 1DI2, Ryter et Schultz, 1998); (**G**) Domaine dsRBM1 de la protéine ADAR 2 (PDB : 2L3J, Stefl et al., 2010). (**H**) Reconnaissance des ARNdb par le domaine non-canonique de la Protéine B2 du virus Flock House (PDB : 2AZ0, Chao et al., 2005).*

Les protéines NS1 des virus IAV et IBV interagissent avec des ARNdb d'origine cellulaire et virale *via* le domaine RBD au sein duquel plusieurs acides aminés basiques localisés dans les hélices α2 et α2' interviennent dans l'interaction avec l'ARN (**Fig.26 et 36**) (Yin *et al*., 2007; Cheng *et al*., 2009). Des changements de conformation de certains acides aminés situés au sein de l'interface d'interaction avec l'ARNdb, en particulier R38, peuvent avoir lieu à la suite de l'interaction avec l'ARN (**Fig.36D et 36E**).

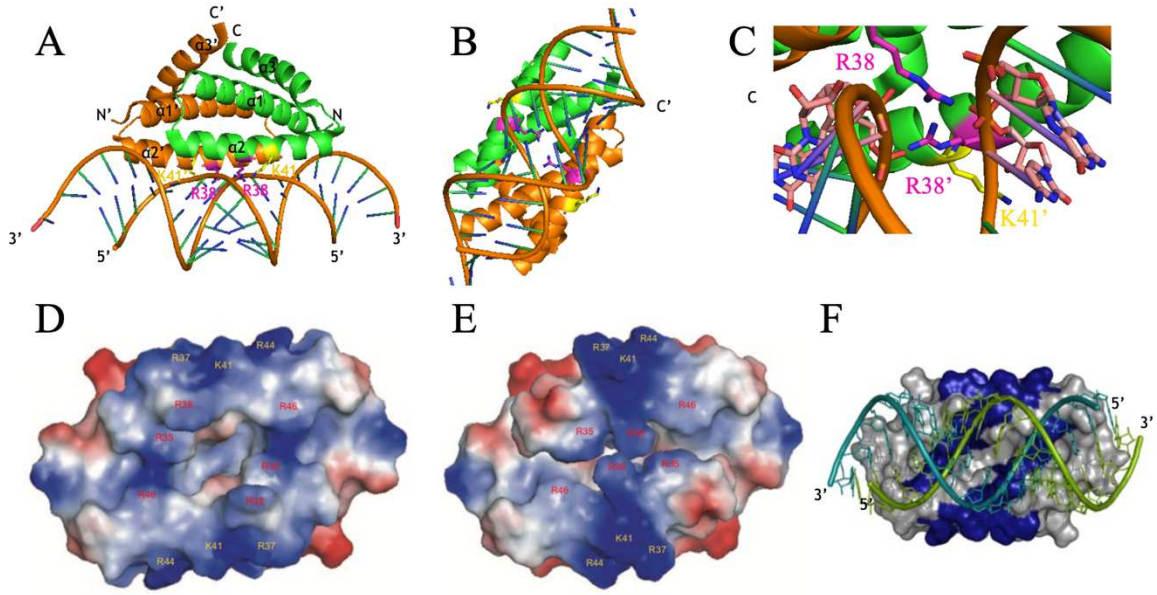


Figure 36 : Mécanisme de reconnaissance des ARNdb par le RBD de NS1 des IAV. *Représentations par PyMol de la structure du complexe cristallographique RBD/ARNdb (PDB : 2ZKO, Cheng et al., 2009), en mode cartoon (**A**, **B** et **C**). Les deux monomères sont représentés en vert et en orange. Les acides aminés R38 et K41 impliqués dans l'interaction avec l'ARNdb sont respectivement représentés en magenta et en jaune. (**A**) Vue de côté; (**B**) Vue du dessous; (**C**) Agrandissement de l'interface d'interaction RBD/ARNdb. (**D**) Surface d'interaction du RBD de NS1 en absence d'ARN (Yin et al., 2007). (**E**) Surface d'interaction du RBD de NS1 en présence du ARNdb de 16 paires de bases (Yin et al., 2007). (**F**) Modèle du complexe RBD/ARNdb dans lequel les acides aminés chargés positivement (lysines et arginines) sont représentées en bleu (Yin et al., 2007).*

NS1 inhibe l'activité du spliceosome (Agris *et al*., 1989; Nemeroff *et al*., 1992), en particulier avec l'ARN U6 (Lu *et al*., 1994; Qian *et al*., 1995) entrainant alors un blocage de la formation des complexes U6-U2 et U6-U4 nécessaires à la maturation des pré-ARNm cellulaires (**Fig.30**) (Qiu *et al*., 1995). L'interaction de NS1 avec U6 (Higashi *et al*., 2008; Someya *et al*., 2008) ou U6atac (Wang et Krug, 1998), fait intervenir plusieurs bases de l'ARN comme en témoigne la perturbation des nucléobases observées par RMN (**Fig.37**) (Someya *et al*., 2008).

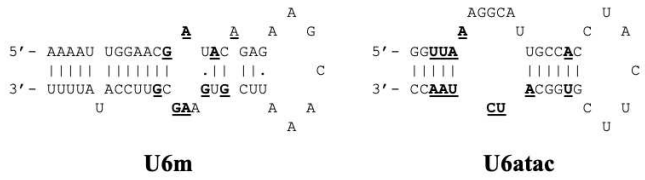


Figure 37 : Structure 2D des ARN U6m et U6atac qui interagissent avec le RBD de NS1. *Les complexes RBD/U6m et RBD/U6atac analysés par RMN (Someya et al., 2008) indiquent une perturbation de certaines bases (indiquées en gras) lors de l'interaction avec le RBD de NS1.*

D'autres travaux ont mis en évidence l'interaction de NS1 avec des ARNdb d'origine virale, en particulier avec les extrémités U12 et U13 des ARNv, dont l'appariement forme le promoteur de la RdRp virale (Hatada *et al*., 1992; 1997; Hatada et Fukuda, 1992). Des aptamères d'ARN se liant avec une haute affinité à NS1 ont également été identifiés par une approche SELEX (Marc *et al*., 2013). La

comparaison des séquences nucléotidiques des 6 familles d'aptamères révèle trois motifs plus ou moins partagés : (i) le motif 5' GUAAC 3', (ii) la séquence 5' AGCAAAAG 3' (motif A) qui est présente à l'extrémité 5' de tous les ARN viraux de polarité positive (ARNc et ARNm, la séquence complémentaire UCGUUUUC étant retrouvée dans les ARNv) et (iii) la séquence 5' UGAUUGAAG 3' (motif B) qui est essentiellement retrouvée dans la région 3'-UTR des ARNm de NS1 (**Fig.38**). Le motif A est essentiel à la réplication du virus. En effet, sa mutation ou son blocage par ARN d'interférence affecte considérablement la réplication du virus (Wu *et al*., 2008; Zhou *et al*., 2007). Moins fréquemment retrouvé que le motif A, le motif B est présent majoritairement dans les ARNm codant pour NS1 (Marc *et al*., 2013), sa présence pouvant suggérer un rôle dans le cycle viral par une interaction de NS1 avec son propre ARNm. Toutefois, aucune étude sur le sujet n'a été menée à ce jour.

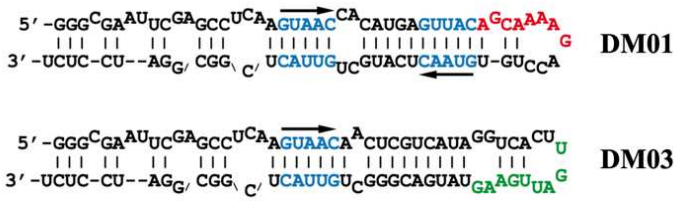


Figure 38 : Les aptamères DM01 et DM03. *La structure secondaire des aptamères a été représentée avec le programme Mfold (Zuker, 2003). Le motif indiqué en rouge est constamment retrouvé dans la région 5'-UTR de tous les ARNm viraux. Le motif indiqué en vert est retrouvé majoritairement dans la région 3'-UTR des ARNm codant pour NS1. Le motif GUAAC est indiqué en bleu.*

Plusieurs travaux ont montré le rôle joué par NS1 dans l'épissage de certains transcrits viraux, en particulier celui issu du segment M (Robb et Fodor, 2012). Le motif GUAAC qui est très proche du site canonique donneur d'épissage des pré-ARNm cellulaires pourrait représenter l'appariement du snRNA U1 sur le site donneur d'épissage situé en 5' de l'intron à épisser (**Fig.39A et 39C**) (Plaschka *et al*., 2019). Cette séquence du site donneur est très conservée chez de nombreuses espèces (Abril *et al*., 2005). Alors que les bases $Gu_{+1}Ur_{+2}$ y sont strictement conservées, la conservation est moins stricte pour les bases $Cy_{-3}Ad_{-2}Gu_{-1}$ (**Fig.39B**). Ceci suggère que NS1 pourrait interagir spécifiquement avec ce duplex, et entrer en compétition avec la protéine U1C du snRNP U1, laquelle se lie à ce duplex (Lu *et al*., 1994; Kondo *et al*., 2015; Singh et Singh, 2019).

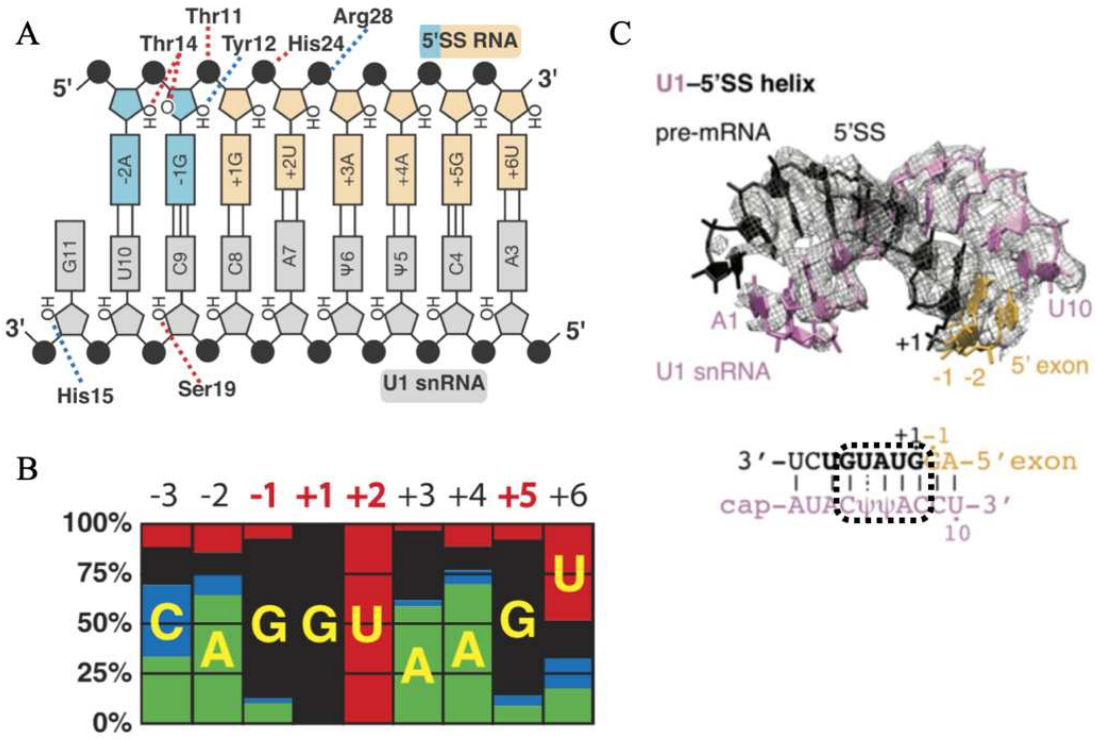


Figure 39 : Structure et conservation du site donneur d'épissage des introns. *(**A**) La protéine U1-C du spliceosome contacte le duplex entre l'extrémité 5' de l'ARN U1 et l'extrémité 5' du site d'épissage de l'intron présent dans le pré-ARNm. (**B**) La conservation des nucléotides est indiquée en pourcentage. La position des nucléotides du site donneur d'épissage est également représentée (Kondo et al., 2015). (**C**) Structure cryo-EM du duplex formé par le snRNA U1 et le site donneur d'épissage (Plaschka et al., 2019) (ψ : pseudouridine).*

Dans la cellule infectée, plusieurs fonctions de NS1 reposent sur la capacité de son domaine RBD à lier diverses molécules d'ARN. En effet, des mutations qui invalident son RBD réduisent considérablement ses interactions avec les ARN dont les ARNdb et impactent fortement le cycle viral, limitant sa réplication (Chen *et al*., 2018; Trapp *et al*., 2018 ; **Annexe 14**) tout en rétablissant un niveau d'expression d'IFN physiologique (Wu *et al*., 2014). Ces résultats démontrent le rôle crucial joué par le domaine RBD de NS1 dans le cycle viral puisque les mutations le ciblant spécifiquement sont plus délétères que les mutations touchant le domaine ED de NS1 (Trapp *et al*., 2018 ; **Annexe 14**). Chen et al. (2020) ont également démontré que la plateforme RBD/ARN module l'interaction du domaine ED de NS1 avec l'hélicase DHX30, démontrant ainsi l'importance de caractériser le mode de reconnaissance des ARNdb par le RBD de NS1 des IAV dans le développement de nouvelles solutions thérapeutiques ciblant spécifiquement la protéine NS1 et notamment son domaine RBD (Engel, 2013; Han *et al*., 2019). En effet, de nouvelles molécules antivirales ciblant l'interface RBD/ARNdb pourraient potentiellement bloquer la plupart des fonctions de NS1 qui dépendent du RBD (Chandra *et al*., 2017). A ce jour, seules deux structures de complexe associant le RBD de NS1 à un ARNdb ont été résolues, par RMN (Yin *et al*. 2007) puis par DRX (Cheng *et al*., 2009). Dans ces deux structures, plusieurs acides aminés du domaine RBD, notamment R35 et R38, interagissent avec le squelette de l'ARN, notamment avec les phosphates et la fonction 2'OH du ribose des nucléobases, ce qui semble

témoigner d'une reconnaissance non-spécifique de la séquence de l'ARN. Néanmoins et comme c'est souvent le cas dans la reconnaissance des ARNdb par les dsRBP, aucune base de l'ARN n'est reconnue spécifiquement par le RBD de NS1 (Cheng *et al*., 2009; Chien *et al*., 2004). De plus, ces ARNdb sont tous deux dépourvus de séquences virales, ce qui pourrait suggérer que ce mode de reconnaissance de l'ARN est plutôt non-spécifique. D'ailleurs, l'affinité du RBD pour de tels ARNdb reste plus modeste que pour les aptamères sélectionnés *in vitro* (Marc *et al*., 2013; Wacquiez *et al*., 2020). Le RBD de NS1 est capable de distinguer spécifiquement les ARNdb de l'ADN. Estimée entre 10 et 17 Å, la largeur de la surface d'interaction du RBD de NS1 hébergée par les hélices α2 et α2' lui permet d'interagir avec des ARNdb dont l'affinité serait plus modeste (Yin *et al*., 2007). L'absence de spécificité de séquence dans l'interaction RBD/ARNdb semble donc en contradiction avec la sélection *in vitro* d'aptamères dotés de motifs spécifiques, incluant des motifs propres au virus. Nous avons donc cherché à élucider quels sont les déterminants, dans l'ARN et/ou dans la protéine, qui pourraient rendre compte de cette affinité élevée, voire d'une certaine reconnaissance spécifique de séquences.

OBJECTIFS DE LA THESE

L'objectif de mon travail de thèse est de contribuer à l'identification des déterminants structuraux et/ou fonctionnels mis en jeu dans les interactions productives de NS1 avec les ARN double-brin originaires du virus ou de l'hôte. Pour ce faire, nous avons conduit en parallèle des approches fonctionnelles et structurales *in vitro* et *in cellulo*.

Dans un premier temps, mon travail de thèse s'est intéressé à évaluer la relevance biologique des motifs viraux contenus dans les aptamères identifiés par SELEX (Marc *et al*., 2013). Pour ce faire, l'importance des séquences virales hautement conservées dans les régions 5' et 3'-UTR des ARNm codant pour NS1 a été évaluée dans des cellules HEK293T transfectées avec un miniréplicon viral dans lequel les régions 5' et 3'-UTR de NS1 ont été couplées à un gène rapporteur. Ce test, qui repose sur la quantification du signal d'émission du gène rapporteur, sera utilisé pour nous renseigner sur le rôle biologique joué par NS1 dans les étapes de transcription, de réplication et de traduction de son propre segment génomique et sur l'importance de la 3'-UTR de NS1 dans ses trois étapes. Puisque le système miniréplicon viral nous renseigne sur un effet positif ou négatif global de NS1 sur les trois étapes sans discriminer séparément l'impact de NS1 sur chacune de ses trois étapes, nous avons développé un système biologique, basé lui-aussi sur un système rapporteur à la luciférase, qui nous permet d'étudier l'impact de NS1 et de ses régions 5' et 3'-UTR sur sa propre traduction. Enfin, l'importance de ses séquences virales hautement conservées a été évaluée dans le cycle viral. Nous avons cherché à caractériser le phénotype de virus mutants exempts de ces motifs viraux d'intérêts. Pour ce faire, nous avons étudié la cinétique de production virale et l'accumulation de NS1 à différents temps post-infection.

Dans un second temps, nous avons caractérisé sur les plans biochimique et structural le mode de reconnaissance du RBD de NS1 lorsque celui-ci est associé avec des aptamères de haute affinité présentant différents motifs séquentiels identifiés par l'approche SELEX (Marc *et al*., 2013). Une partie importante du projet de thèse aura été consacrée au clonage des différentes constructions de RBD et de NS1 ainsi qu'à l'expression et la purification de protéines recombinantes du RBD seul et de la NS1 entière. Une fois purifiées, les protéines ont été utilisées en retard de migration sur gel afin de déterminer les différentes caractéristiques (longueur, structure et séquence) de « l'ARNdb » mimé par les aptamères et leurs dérivés qui conditionnent la formation de complexes RBD/ARNdb de haute affinité. Les ARNdb de longueur minimale confectionnés à partir des aptamères les plus courts mais qui conservent une haute affinité en retard de migration sur gel pour les différentes protéines purifiées ont été utilisés pour la résolution de la structure tridimensionnelle de complexes RBD/ARNdb par remplacement moléculaire avec le modèle RBD/ARNdb existant (Cheng *et al*., 2009). Cette étape, qui est conditionnée par notre capacité à cristalliser les complexes RBD/ARNdb d'intérêts, devrait pouvoir nous renseigner sur l'existence potentielle d'un second mode de reconnaissance des ARNdb par le RBD de NS1, possiblement en relation avec une reconnaissance spécifique des nucléobases de l'ARNdb.

MATÉRIELS ET MÉTHODES

A. Culture cellulaire

A.I. Les bactéries

Les souches d'*Escherichia coli* XL1-Blue, DH5α et NEB®10-beta ont été utilisées pour produire les plasmides recombinants. Elles sont cultivées en milieu LB liquide, ou sur boîte de milieu LB gélosé contenant 17 $g.L^{-1}$ d'agar. L'expression de protéines recombinantes a été réalisée dans la souche d'*Escherichia coli* Rosetta 2, laquelle est dérivée de BL21-DE3 et héberge le plasmide pRARE2 (Cam^R, Novagen®) qui facilite l'expression de protéines recombinantes en exprimant sept ARNt correspondant aux codons rares AGG (Arg, R), AGA (R), CGG (R), GGA (Gly, G), AUA (Ile, I), CUA (Leu, L) et CCC (Pro, P). Tous les milieux de culture de bactéries sont stérilisés par autoclavage (20 minutes à 121 °C, 1 bar). Les bactéries hébergeant les constructions plasmidiques d'intérêts sont stockées à -80 °C en glycérol 50 %.

A.II. Les cellules de mammifères

Les cellules embryonnaires rénales humaines HEK293T sont cultivées en milieu DMEM (Dulbecco's modified Eagle's medium) (Lonza) additionné de 7,5 % de sérum de veau fœtal décomplémenté (SVF) (Gibco) et de 100 U/mL pénicilline/streptomycine (PS) (HyClone, GE Healthcare). Les cellules de rein de chien MDCK sont cultivées en milieu EMEM (Eagle's minimal essential medium) (Corning) avec les mêmes conditions que pour les HEK293T. Des cellules d'adénocarninome pulmonaire humain A549 sont cultivées en milieu DMEM (Lonza) additionné de 10% SVF, de 100 U/mL PS (HyClone, GE Healthcare) et de 1 $mmol.L^{-1}$ sodium-pyruvate (Sigma). Les cellules sont maintenues à 37 °C en présence de 5 % de CO_2.

Les cellules sont cultivées en flasque de 75 cm^2 (T75, Nunc™EasYFlasks™) (ThermoFisher Scientific) jusqu'à confluence. Le milieu de culture est alors enlevé, les cellules sont lavées en PBS 1x (Gibco) puis les cellules sont trypsinées avec 2 mL de trypsine-EDTA (0,25 %) (Gibco). Les cellules sont incubées à 37 °C pendant 5 minutes. La trypsine est neutralisée avec 8 mL de milieu de culture frais puis les cellules sont décrochées du fond de la flasque par grattage au râteau. Les cellules sont remises en suspension par aspirations/refoulements. Entre 1 et 2.10^6 cellules sont transférées dans une nouvelle flasque T75 additionnée de 12 mL de milieu de culture frais.

B. Les plasmides

B.I. Vecteurs d'expression de protéines recombinantes

Les ORF des RBD ont été amplifiées par PCR (couples d'amorces RBD-pdm09, RBD-H3N2, RBD-H7N9, RBD-Bat et RBD-H1N1, **Annexe 1**) puis clonées aux sites *NdeI* et *BamHI* du vecteur pET15b(+) (5708 pb, Amp^R, Novagen) (**Fig.40**). Les protéines RBD recombinantes exprimées à partir de pET-15b(+) portent un His_6-tag N-terminal, clivable à la thrombine (**Fig.40**). Les acides aminés R38 et K41 des constructions de RBD H5N1 et H7N1 ont été remplacés par une alanine par mutagénèse dirigée au moyen du kit QuikChange II XL site-Directed Mutagenesis (Agilent).

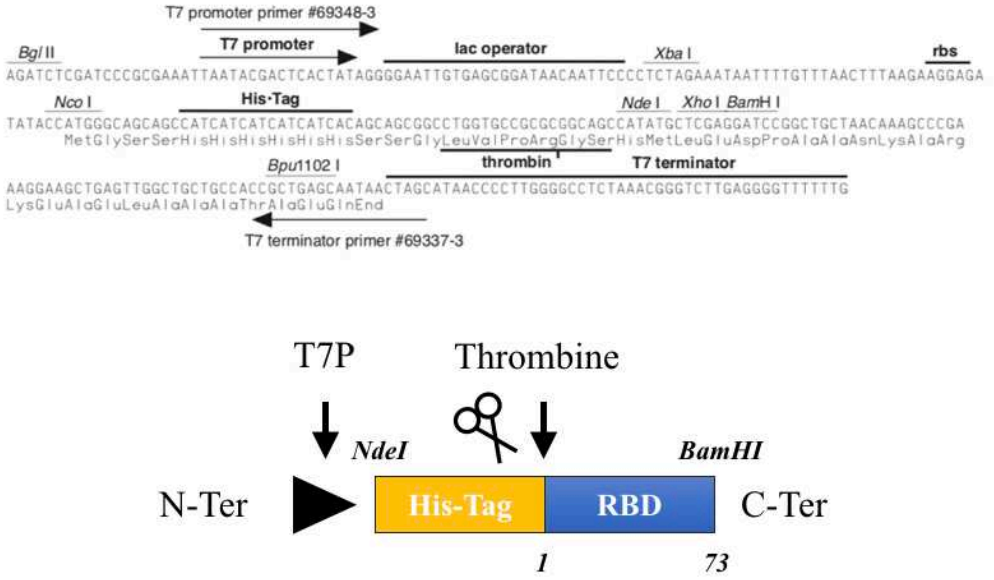


Figure 40 : Carte du vecteur pET15b(+) et schéma des protéines RBD recombinantes

En raison de la faible solubilité de la protéine NS1 entière en solution, nous avons choisi le vecteur pGEX-5X-1 (4972 pb, Amp^R, GenScript) (**Fig.41**) pour produire la protéine de fusion GST-NS1 du virus A/Turkey/Italy/977/1999/H7N1. Ce vecteur avait été utilisé par Qiu et Krug (1995) pour exprimer la NS1 du virus Udorn (Alonso-Caplen *et al.*, 1992; Lamb *et al.*, 1980) en vue d'évaluer son interaction avec les ARNm de M2 et NEP. La construction GST-NS1-H7N1 (pGEX-5X-1) a été commandée chez GenScript. Une construction GST-RBD a été réalisée par mutagénèse dirigée en utilisant le couple d'amorces GST-RBD (**Annexe 1**). Les protéines exprimées à partir de pGEX-5X-1 ont leur extrémité N-terminale fusionnée à un module Glutathione-S Transférase (GST-tag) de 218 acides aminés suivi d'un site de clivage par le facteur Xa (**Fig.41**).

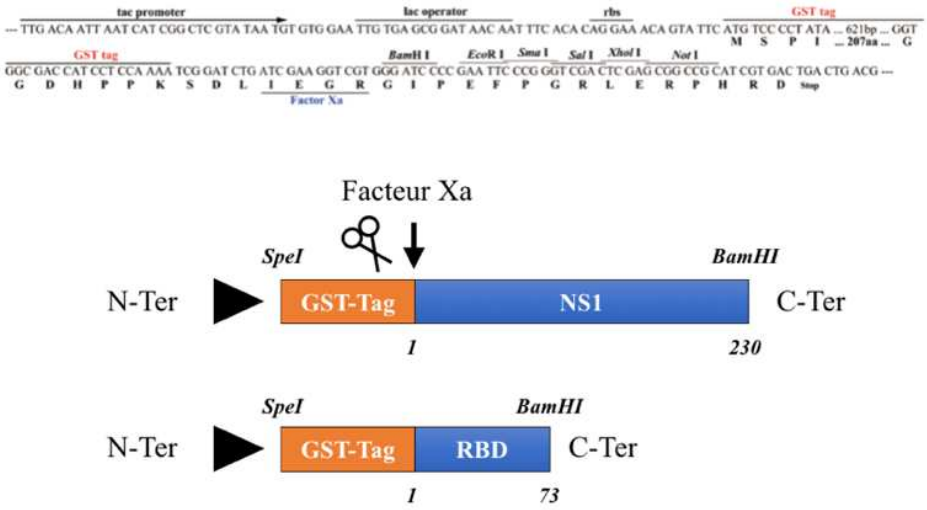


Figure 41 : Carte du vecteur pGEX-5X-1 et schéma des protéines de fusion à la GST

B.II. Plasmides de génétique inverse du virus WSN

La génétique inverse du virus WSN (A/WSN/1933/H1N1) repose sur le système à 12 plasmides de Fodor *et al*. (1999). Quatre plasmides recombinants (pcDNA3.1(+), Invitrogen) permettent l'expression des ARN messagers des protéines PB2, PB1, PA et NP de la souche virale WSN, sous le contrôle du promoteur CMV. Huit plasmides « pPol » dérivés de plasmides commerciaux pUC18 permettent l'expression des huit ARN génomiques viraux, sous contrôle du promoteur de Pol I. Une séquence du ribozyme de l'hépatite D à l'extrémité 3' des transcrits primaires issus des plasmides pPol permet de produire au nucléotide près des copies exactes de ces ARNv. Ces derniers sont transcrits et répliqués par la RdRp virale, initiant un cycle viral productif (**Fig.45**).

Les motifs viraux B, C et D ont été modifiés dans le plasmide de génétique inverse pPol-NS1 par mutagénèse dirigée en utilisant le kit QuikChange (Agilent). Nous avons ainsi produit les trois mutants simples B, C et D (respectivement mutB, mutC et mutD), ainsi que toutes leurs combinaisons à savoir les double-mutants BC, BD et CD et le triple mutant BCD (mutBCD). Les constructions ont été vérifiées par séquençage avec le couple d'amorces M13 (**Annexe 1**). Les plasmides utilisés pour la génétique inverse ont été vérifiés sur gel (**Fig.42**)

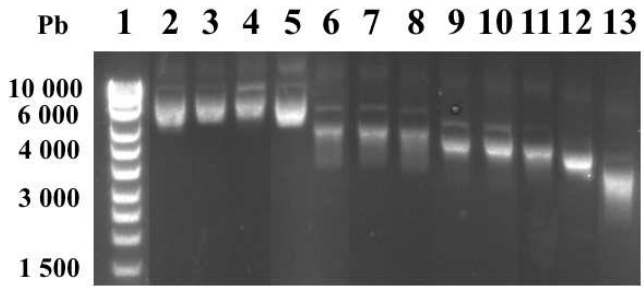


Figure 42 : Plasmides pcDNA et pPolI utilisés pour la génétique inverse. *Un volume de 5 µL de smart ladder 200 to 10 000 bp (Eurogentec) est déposé dans la piste 1. Cent nanogrammes de chaque plasmide de génétique inverse sont déposés sur gel. Les plasmides pcDNA PB2 (piste 2), PB1 (piste 3), PA (piste 4) et NP (piste 5) et les pPolI PB2 (piste 6), PB1 (piste 7), PA (piste 8), HA (piste 9), NP (piste 10), NA (piste 11), M (piste 12) et NS (piste 13) sont séparés par une migration de 30 minutes à 100 V sur gel d'agarose 1 % en tampon TAE 1X.*

B.III. Plasmides du système *miniréplicon viral*

Dans le système *miniréplicon viral*, l'ORF de NS1 dans le plasmide pPol-NS1 a été remplacée par l'ORF de la luciférase Renilla (Ren-Luc, 936 pb) (**Fig.43**). Cette construction a été réalisée sur le plasmide pPol-NS1-wt, ainsi que sur les trois simples mutants B, C et D. Ces quatre constructions contiennent donc en totalité la longue 3'-UTR de NS1, avec ou sans les mutations (**Fig.43**).

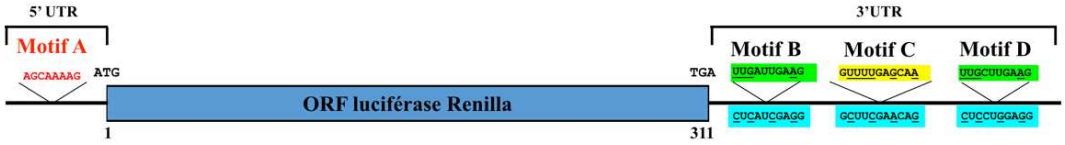


Figure 43 : Construction de l'ARN pseudo-viral rapporteur. *Les mutations des séquences virales de la 3'-UTR de NS1 sont représentées dans la construction.*

Dans ce système, la transfection des quatre plasmides pcDNA recombinants et du plasmide rapporteur permet de mesurer l'activité de la RdRp virale. Pour évaluer l'impact de NS1 sur l'activité de la RdRp virale, NS1 est exprimée transitoirement par la co-transfection d'un vecteur recombinant pCI (4006 bp, Amp^R, Promega) exprimant l'ARNm de NS1 ou du RBD, sous le contrôle du promoteur fort du cytomégalovirus (CMV).

B.IV. Plasmides du système *translation-only*

Pour évaluer la capacité de traduction des différents ARN, plusieurs vecteurs d'expression pCI recombinants ont été construits. Dans les cellules transfectées, ces derniers permettront l'expression transitoire de *viral-like mRNAs*, lesquels sont pourvus, ou non, de l'extrémité 5'-UTR des ARNm viraux, et se distinguent en outre par leur extrémité 3'-UTR correspondant à celle de NS1 (WT, mutB, mutC, ou mutD).

C. Constructions plasmidiques

C.I. Amplification et purification de vecteurs

Les bactéries XL1-Blue, DH5α ou NEB®10-beta transformées par les plasmides d'intérêts sont incubées sur la nuit en milieu LB additionné d'antibiotique. Les plasmides ont été purifiés avec le kit NucleoSpin Plasmid (Macherey-Nagel). Les bactéries récoltées par centrifugation (30 secondes à 13 000 g) sont resuspendues dans 250 µL de tampon A1 puis lysées avec 250 µL de tampon A2 pendant 5 minutes à température ambiante. Les lysats sont neutralisés par 300 µL de tampon A3 puis clarifiés par centrifugation (10 minutes à 11 000 g). Le surnageant résultant (800 µL), qui contient les ADN à purifier, est appliqué sur la colonne de silice. Après centrifugation (1 minute à 11 000 g), la silice est rincée une fois avec 600 µL de tampon A4 puis séchée par centrifugation (2 minutes à 11 000 g). L'ADN est élué dans 50 µL de H_2O. La qualité et la concentration des fragments d'ADN digérés sont évaluées par spectrophotométrie d'absorption (Nanodrop, Thermo Scientific). Les plasmides purifiés sont stockés à -20 °C.

C.II. Digestion d'ADN plasmidique et purification de fragments

Les plasmides contenant les séquences d'ADN d'intérêts sont digérés à 37 °C pendant au moins 2 h par les enzymes de restriction appropriées (New England Biolabs). Les produits de digestion sont séparés par migration sur gel d'agarose 1 % en tampon hydroxyméthyl aminométhane (Tris)-Acétate-Éthylènediaminetétraacétique (EDTA) 1x (TAE 1x) additionné de 0,5 $\mu g.mL^{-1}$ bromure d'éthidium (BET). Les bandes d'intérêts sont découpées puis les ADN purifiés avec le kit Gel and PCR clean-up (Macherey-Nagel). L'agarose est solubilisé par une incubation de 10 minutes à 50 °C en tampon NTI (200 µL x 100 mg gel). Le volume réactionnel est appliqué sur une colonne de silice. Après centrifugation (30 secondes à 11 000 g), la silice est rincée deux fois par 700 µL de tampon NT3 puis séchée par centrifugation de 1 minutes à 11 000 g. L'ADN est élué dans 30 µL de H_2O. La qualité et la concentration des fragments d'ADN sont évaluées par spectrophotométrie d'absorption (Nanodrop, Thermo Scientific).

C.III. Ligation et transformation

Les ligations ont été réalisées avec le kit T4 DNA ligase (Promega). Un mélange de 100 ng de vecteur linéarisé et de 40 ng de fragment purifié est incubé sur la nuit à 4 °C dans un volume de 10 µl contenant 1 U de T4 DNA ligase. 50 µL de bactéries thermo-compétentes (Hanahan, 1983) ou électro-compétentes (Cruz-Rodz et Gilmore, 1990) sont décongelés sur glace pendant 20 minutes. Elles sont

incubées en présence de 1 µL de la réaction de ligation. Dans le cas de bactéries thermo-compétentes, le mélange est incubé pendant 20 minutes sur glace puis pendant 2 minutes à 42 °C. Le mélange est ensuite replacé quelques instants dans la glace. Dans le cas de bactéries électro-compétentes, le mélange est incubé pendant 20 minutes sur glace puis placé au fond de cuves à électroporer de 0.1 cm (BioRad). Les bactéries du mélange sont électroporées par un pulse de 1,0 à 4,0 ms avec l'électroporateur MicroPulser (BioRad). Un volume de 500 µL de milieu LB est alors ajouté puis incubé à 37 °C sous agitation pendant 1 h. Le mélange est centrifugé 30 secondes à 5000 g. Une partie du volume est éliminé par aspiration. Le culot bactérien est ensuite repris dans le volume restant de 100 µL, et la suspension est étalée sur gélose de LB additionnée de l'antibiotique de sélection. Les géloses sont incubées à 37 °C sur la nuit.

C.IV. Criblage de clones transformés par PCR

Un minimum de 10 colonies, visibles à l'œil nu en surface des géloses incubées sur la nuit, est testé par une PCR utilisant des amorces judicieusement choisies (conditions de PCR décrites dans le **Tableau 4**).

Tableau 4 : Conditions et paramètres de PCR pour l'amplification d'ADN d'intérêt et la recherche de clones bactériens transformés

	Concentration stock	Concentration essai	Réaction
ADN matrice			2 µL
Amorce sens	10 µM	400 nM	1 µL
Amorce anti-sens	10 µM	400 nM	1 µL
Tampon Pfu	10x	1x	5 µL
$MgCl_2$	10 mM	600 nM	1,5 µL
Mix dNTP	10 mM	200 µM	0,5 µL
Taq pol pfu	2,5 U/µL	1,25 U	0,5 µL
H_2O			13,5 µL
Volume total			25 µL
Paramètres du thermocycleur			
Etapes n°	Objectif	Température (°C)	Durée
1	Dénaturation	95	4 minute
2	Dénaturation	95	20 secondes
3	Hybridation	59	30 secondes
4	Polymérisation	72	45 secondes
5	25 cycles des étapes 2 à 4		
6	Polymérisation finale	72	10 minutes
7	Stockage	10	Jusqu'à utilisation

C.V. Mutagénèse dirigée

Toutes les mutations ont été introduites par PCR avec le kit QuikChange II XL site-Directed Mutagenesis (Agilent) en suivant les instructions du fabricant (conditions de mutagénèse décrites dans le **Tableau 5**). Brièvement, la totalité du plasmide est amplifié par PCR en utilisant deux amorces

complémentaires portant la mutation (**Annexe 1**). L'ADN matrice, méthylé, est éliminé en incubant le mélange réactionnel résultant avec 10 U de l'endonucléase *DpnI* 1 h à 37 °C. Un total de 8 µL de mélange réactionnel est analysé sur gel d'agarose 1 % TAE 1x. 1 à 2 µl du mélange réactionnel sont utilisés pour transformer des bactéries supercompétentes XL1-blue fournies avec le kit de mutagénèse dirigée (Agilent).

Tableau 5 : Conditions et paramètres de PCR pour la mutagénèse dirigée

	Concentration stock	Concentration essai	Réaction
ADN matrice		10 à 25 ng	1 µL
Amorce sens	10 µM	250 nM	0,625 µL
Amorce anti-sens	10 µM	250 nM	0,625 µL
Tampon Pfu	10x	1x	2,5 µL
Mix dNTP	10 mM	200 µM	0,5 µL
Pfu ultra	2,5 U/µL	1,25 U	0,5 µL
H_2O			19,25 µL
Volume total			25 µL
Paramètres du thermocycleur			
Etapes n°	Objectif	Température (°C)	Durée
1	Dénaturation	95	1 minute
2	Dénaturation	95	30 secondes
3	Hybridation	55	1 minute
4	Polymérisation	68	6 minutes 30 secondes (1 minute/kb)
5	20 à 25 cycles des étapes 2 à 4		
6	Polymérisation finale	68	10 minutes
7	Stockage	10	Jusqu'à utilisation

C.VI. Séquençage des constructions plasmidiques

Toutes les constructions plasmidiques ont été vérifiées par séquençage avant utilisation. Un volume de 15 µL à 100 ng.µL^{-1} de chaque construction plasmidique a été adressé à Genewiz (https://www.genewiz.com//en/Public/Services/Next-Generation-Sequencing/Standalone-NGS-Solutions) ou à Eurofins (https://www.eurofinsgenomics.eu), pour un séquençage par la méthode de Sanger utilisant l'amorce T7 promoter (**Annexe 1**) (vecteurs pET et pCI) ou avec l'une des amorces universelles M13 (**Annexe 1**) (construction en vecteurs pPol I et pPR7).

D. Expression du système *miniréplicon viral* en cellules de mammifères

Des plaques 24 puits (P24) sont préparées la veille avec 1.10^5 HEK293T/puits. Un mélange de plasmides associant (i) les 4 pcDNA (100 ng pour PA, PB1 et PB2, et 200 ng NP), (ii) un vecteur permettant l'expression de la luciférase Firefly (FF-luc) sous contrôle du promoteur viral CMV (pCMV FF-luc 5 ng), (iii) un plasmide *miniréplicon* rapporteur Ren-luc et (iv) un plasmide pCI d'expression transitoire de NS1 (entière ou RBD) est transfecté avec du FuGENE (Promega) à un ratio ADN :FuGENE de 1 :3 dans les HEK293T. Après 24 h, les tapis cellulaires sont délicatement rincés en PBS 1x (Gibco) puis lysés pendant 15 minutes à température ambiante avec 100 µL de tampon PLB 1x (Dual-Luciferase® Reporter Assay System, Promega). Un volume de 20 µL de lysat préalablement centrifugé (30 secondes 10 000 g) est déposé sur plaque blanche 96 puits (P96) dans lequel l'activité des deux luciférases est quantifiée avec le kit Dual-Luciferase® Reporter Assay System (Promega) sur l'appareil GloMax-Multi Detection System (Promega) (**Fig.44**).

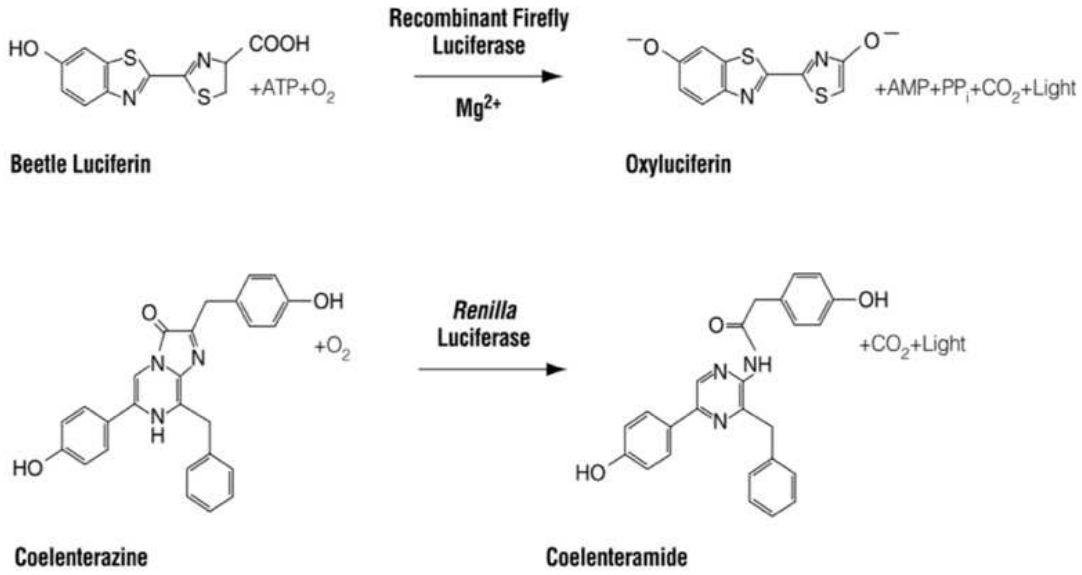


Figure 44 : Principe du dosage des activités luciférases FireFly et Renilla. *La luciférase Firefly catalyse la transformation de son substrat, la luciférine, en oxyluciférine en présence d'ATP, d'oxygène et de magnésium. La luciférase Renilla catalyse la transformation de son substrat, la coelenterazine, en coelenteramide en présence d'oxygène. Dans les deux cas, la réaction produit une émission de lumière (bioluminescence) quantifiable (http://www.thermofisher.com/order/catalog/product/16176).*

E. Production de virus par génétique inverse

Mise au point à la fin du 20e siècle (Fodor *et al*., 1999), la génétique inverse des virus *influenza* permet de générer des virus à ARN segmentés de polarité négative à partir d'ADN complémentaires clonés. Cette technique a permis de produire les premiers virus grippaux recombinants dans des cellules co-transfectées avec 8 ou 12 plasmides simultanément. Dans le cas d'un système à 8 plasmides bidirectionnels, les ADNc sont transcrits en ARN viraux génomiques (ARNv) par l'ARN polymérase I cellulaire (Pol I), et en ARNm par l'ARN polymérase II cellulaire (Pol II). Les ARNv sont transcrits et répliqués par la RdRp virale. Dans le cas d'un système à 12 plasmides, 4 plasmides monodirectionnels (PA, PB1, PB2 et NP) sont transcrits en ARNm viraux par Pol II, permettant la synthèse du complexe RdRp du virus. Les 8 autres plasmides (PA, PB1, PB2, NP, HA, NA, M et NS) sont transcrits par Pol I en ARNv avec une extrémité 3' exactement identique à celle des ARN viraux naturels, grâce à une séquence ribozyme. Ces derniers seront transcrits et répliqués par la RdRp virale. Dans les deux cas, la transfection des plasmides permet l'initiation d'un cycle viral complet (**Fig.45**).

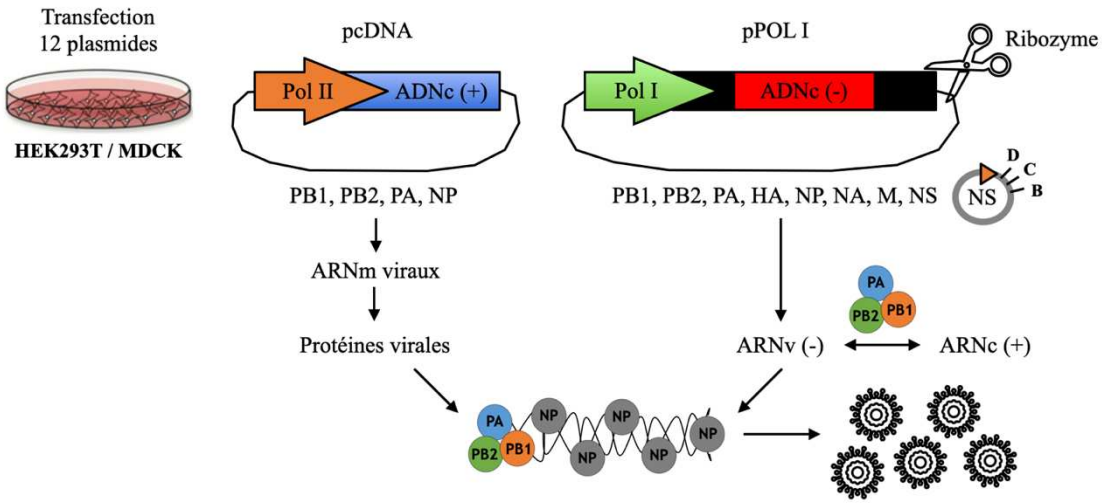


Figure 45 : Principe de la production de virus par génétique inverse

Un mélange de 4.10^5 cellules HEK293T et de 3.10^5 cellules MDCK est ensemencé la veille dans un puits de plaque 6 puits (P6). Un total de 12 plasmides monodirectionnels, dont (i) 4 pcDNA codant pour PA, PB1, PB2 et NP et (ii) 8 pPol I codant pour tous les segments viraux génomiques, est transfecté avec le kit FuGENE® HD Transfection Reagent (Promega) afin d'assurer la production puis l'amplification de virus par les cellules HEK293T et MDCK respectivement (Pleschka *et al*., 1996). Les

segments viraux génomiques utilisés sont originaires de la souche virale humaine WSN (A/WSN/1933 (H1N1)). Un mix contenant 500 ng de chacun des 12 plasmides est préparé extemporanément. Après ajout de 78 µL de milieu Opti-MEM (Gibco) et de 10 µL de FuGENE (Promega), le mélange (100 µL) est, incubé à température ambiante pendant 30 minutes, puis déposé goutte à goutte dans le milieu de culture des puits de P6. Les cellules sont alors incubées à 37 °C en présence de 5 % de CO_2. Le lendemain, le milieu de culture des cellules est remplacé par 3 mL de milieu EMEM complet frais additionné de 1 µg/mL de trypsine-TPCK (L-(tosylamido-2-phenyl) ethyl chloromethyl ketone) (Worthington®). Les cellules sont laissées à l'incubateur pendant 72 h, durée pendant laquelle les effets cytopathogènes du virus sont suivis au microscope optique. A l'issue de ces 72 h, 500 µL du surnageant de culture est prélevé puis appliqué sur un puits de P6 contenant 10^6 cellules MDCK. Après 1 h d'incubation, un volume de 2,5 mL de milieu EMEM complet frais additionné de 1 µg/mL de trypsine-TPCK (Worthington®) est ajouté aux cellules. Après 72 h de culture, le surnageant contenant le stock de virus amplifiés est récolté, clarifié par centrifugation et stocké à -80 °C.

E.I. Titrage des virus

Les virus produits puis amplifiés sont titrés en P6 sur cellules MDCK (10^6 par puits de P6). Des dilutions décimales successives du virus en milieu EMEM non-complet sans trypsine sont réalisées jusqu'à la dilution 10^{-8}. Pour chacune des dilutions de 10^{-3} à 10^{-8} un volume de 400 µL est appliqué en duplicate (deux puits de P6). Un volume de 2,6 mL de milieu EMEM complet frais contenant 1.2 % d'avicel (FMC BioPolymer) et 0,5 µg/mL de trypsine-TPCK (Worthington®) (Matrosovich *et al*., 2006) est ajouté au surnageant de culture. Les cellules sont rincées deux fois en PBS 1x (Gibco) après 72 h. Le tapis cellulaire est fixé et coloré par 15 minutes d'incubation dans 1 mL de solution de cristal violet (2 grammes par litre dans une solution contenant 20 % éthanol et 10 % formaldéhyde), puis rincé à l'eau. Le dénombrement des plages de lyse permet de calculer le titre viral, qui est exprimé en Unités Formant Plage par mL (Plaque Forming Units (PFU)/mL).

E.II. Séquençage des génomes viraux

Le surnageant de culture de cellules MDCK, infectées à la multiplicité d'infection de 0,001 PFU/cellule, est récoltée 48 à 72 h post-infection. Les ARN viraux sont purifiés avec le kit Qiamp Viral RNA (Qiagen) selon les instructions du fabricant, et finalement élués dans un volume de 60 µL de H_2O puis stockés à -80 °C. Un total de 100 ng d'ARN viral purifié est rétro-transcrit en ADN complémentaire (ADNc) avec le kit superscript™ II reverse transcriptase (Invitrogen™) en utilisant l'amorce universelle Uni12 (**Annexe 1**) (conditions de rétro-transcription indiquées dans le **Tableau 6**).

Tableau 6 : Conditions et paramètres de retro-transcription des ARN viraux en ADNc

Mix A			
Concentration	Stock	Essai	Réaction
ARN viral		100 µg	4 µL
Amorce Uni-12	100 µM	10 µM	1,2 µL
dNTP	10 mM	800 µM	1 µL
H_2O			5,5 µL
Total			12 µL
Incubation	5 minutes à 65°C puis à 4°C		
Mix B (superscript II)			
5X first strand		4 µL	
0.1 M DTT		2 µL	
Superscript II RT		0.5 µL	
H2O		1.5 µL	
Total		8 ul	
Incubation		50 minutes à 42°C puis 15 minutes à 70°C	

L'ADNc du segment NS est amplifié par PCR en utilisant deux amorces universelles du segment NS, l'une d'elle portant une extension correspondant à la séquence d'une amorce universelle M13 (Hoffmann *et al*., 2001). Les produits de PCR sont séquencés en utilisant cette amorce M13 (**Annexe 1**).

E.III. Phénotype des virus mutants

Les stocks de virus obtenus après 72 h ont d'abord subi une amplification sur MDCK afin d'atteindre un titre d'au moins 10^6 PFU/mL. Des flasques de 25 cm^2 (T25 Nunc™EasYFlasks™, ThermoFisher Scientific) préparées avec 5.10^5 cellules MDCK sont infectées à une m.o.i de 0.001 pfu/cellule avec 500 µL de dilution du stock viral initial. Les surnageants de culture (environ 10 mL) sont récoltés après 72 h, titrés puis stockés en cryo-tubes à -80 °C.

E.III.1. Cinétique de production de particules virales

Des flasques T25 sont ensemencées avec 5.10^5 MDCK. Après 24 h, le milieu de culture est éliminé et les cellules sont infectées avec une m.o.i de 0,001 pfu/cellule. Un volume de 500 µL de surnageant de culture de chaque T25 est prélevé à certains temps post-infection (0 h, 6 h, 12 h, 24 h, 48 h et 72 h). Le nombre de particules virales présentes dans les surnageants de culture est déterminé par titrage sur cellules MDCK.

E.III.2. Cinétique d'accumulation de NS1 dans les cellules infectées

Des puits de plaque 24 puits (P24) sont ensemencés avec 10^5 cellules A549. Après 24 h, les cellules sont lavées à deux reprises en PBS 1x (Gibco) puis infectées à une m.o.i de 1 PFU/cellule (200 µL/puits). Chaque puits est additionné de 300 µL de milieu DMEM complet frais additionné de 1 µg/mL de trypsine-TPCK (Worthington®). A différents temps (2 h, 4 h, 8 h et 12 h) post-infection, les cellules sont lavées avec du PBS 1x (Gibco) puis lysées dans 100 µL de tampon de Laemmli 2x pour une analyse par Western-blot. Les échantillons (10 à 20 µl de lysat par puits), chauffés 5 minutes à 95 °C, sont analysés par électrophorèse en mini-gel de polyacrylamide 12 %-SDS (SDS-PAGE, MiniProtean II, BioRad). A l'issue de la migration (90 minutes à 110 V), les protéines sont transférées sur membrane de nitrocellulose (Parablot NCL, Macherey-Nagel). Les membranes de nitrocellulose sont saturées pendant 1 h à température ambiante en tampon TNT (10 $mmol.L^{-1}$ Tris-HCl; 150 $mmol.L^{-1}$ NaCl; 0,001 % Tween 20) additionné de 3 % lait, puis incubées 1 h à température ambiante en présence d'anticorps primaires en tampon TNT additionné de 3 % lait. Après quatre lavages successifs de 10 minutes en tampon TNT additionné de 1 % lait, les membranes sont incubées 1 h à température ambiante en présence d'anticorps secondaires (anti-lapin ou anti-souris, selon les cas) couplés à la peroxydase en tampon TNT additionné de 3 % lait. Après quatre lavages, les membranes sont incubées en présence du substrat de la peroxydase (ECL chemiluminescence WesternBright, Advansta) et visualisées avec l'imageur Fusion Fx (Vilber-Lourmat).

La protéine NS1 (monomère de 26 kDa) est révélée en utilisant l'anticorps polyclonal de lapin produit au laboratoire (1/800^e). La protéine NP (56 kDa) est révélée en utilisant un anticorps polyclonal de lapin (Thermo Fischer Scientific, PA5-32242) (1/3000^e). La protéine GAPDH (38 kDa), utilisée pour contrôler la quantité déposée dans les gels, est révélée en utilisant un anticorps monoclonal murin (Invitrogen, mAB374) (1/1000^e).

F. Expression et purification de protéines recombinantes

F.I. Expression de protéines recombinantes en bactéries

Après repiquage d'un clone d'*E.coli* Rosetta 2 transformée (Rosetta 2) sur gélose contenant 25 µg.mL^{-1} de chloramphénicol et 100 µg.mL^{-1} d'ampicilline (sélectif), les bactéries transformées sont d'abord amplifiées à 37 °C sur la nuit en milieu MDG non-inducteur sélectif (Studier, 2005) (2 mmol.L-1 $MgSO_4$; 0,25 % aspartate; 0,5 % glucose; 25 mmol.L^{-1} NA_2HPO_4; 25 mmol.L^{-1} KH_2PO_4; 50 mmol.L^{-1} NH_4Cl; 5 mmol.L^{-1} Na_2SO_4; 0,2 x du stock 1000x de *trace elements* contenant 50 mmol.L^{-1} $FeCl_3$; 20 mmol.L^{-1} $CaCl_2$; 10 mmol.L^{-1} $MnCl_2$; 10 mmol.L^{-1} $ZnSO_4$; 2 mmol.L^{-1} $CoCl_2$; 2 mmol.L^{-1} $CuCl_2$; 2 mmol.L^{-1} $NiCl_2$; 2 mmol.L^{-1} Na_2MoO_4; 2 mmol.L^{-1} Na_2SeO_3; 2 mmol.L^{-1} H_3BO_3). Au total, 500 mL de milieu ZYM-5052 auto-inducteur sélectif (Studier, 2005) (10 g.L^{-1} tryptones; 5 g.L^{-1} extraits de levures; 25 mmol.L^{-1} NA_2HPO_4; 25 mmol.L^{-1} KH_2PO_4; 50 mmol.L^{-1} NH_4Cl; 5 mmol.L^{-1} Na_2SO_4; 0,5 % glycérol; 0,05 % glucose; 0,2 % lactose; 2 mmol.L^{-1} $MgSO_4$; 0.2 x du stock 1000x de *trace elements*) en fiole de culture avec déflecteurs sont ensemencés avec 0,5 mL de pré-culture. Les cultures sont placées à 37 °C sous agitation jusqu'à atteindre une turbidimétrie de 0,6 (mesurée par la densité optique à 600 nm). Les cultures sont alors placées à 18 °C sous agitation (180 à 220 rpm) pendant une durée de 16 à 20 h. Les bactéries sont ensuite centrifugées à froid pendant 15 minutes à 5 000 g (Super-lite® SLA-3000 rotor, Sorvall®), lavées avec 50 mL de tampon contenant 10 mmol.L^{-1} Tris-HCl pH 8; 150 mmol.L^{-1} NaCl puis transférées en tubes falcon de 50 mL avant leur stockage à -20 °C.

F.II. Purification des protéines RBD de NS1 recombinantes

Les culots de bactéries, préalablement décongelés sur glace sont resuspendus à froid dans 40 mL de tampon A (**Annexe 2**); Les suspensions bactériennes résultantes sont incubées pendant 30 minutes à 30 °C en présence de 0,75 mg.mL^{-1} de lysozyme. Trois cycles de congélation (azote liquide) - décongélation (bain-marie 30 °C) et 5 cycles de sonication à froid de 1 minute (sonde S&M-220, 40 % amplitude avec 5 secondes de pulse et 5 secondes de pause) sont pratiqués pour lyser les cellules. L'extrait cellulaire brut (ECB) (environ 45 mL) correspond au surnageant filtré à 0,45 µm d'une centrifugation de 80 minutes à 19 000 g. L'ECB contenant la protéine d'intérêt soluble, est appliqué sur une colonne de chromatographie d'affinité de 5 mL contenant une résine métallique au nickel (ou au cobalt) (IMAC) (HisTrap FF, GE Healthcare Life Sciences) pré-équilibrée en tampon A (**Annexe 2**). La colonne est lavée avec 10 volumes de colonnes (cv) de tampon A (**Annexe 2**), 10 cv de tampon B

(**annexe 2**) et 10 cv de tampon C (**Annexe 2**). Les protéines tagguées (His_6) sont éluées avec 10 cv de tampon D (**Annexe 2**). Chacune des fractions (*Flow through*, lavages et élution) est conservée pour une analyse par SDS-PAGE 12 %. Les fractions d'élution contenant les protéines d'intérêts sont regroupées et dialysées sur la nuit à 4 °C sous agitation avec du tampon E (**Annexe 2**) dilué au $1/10^e$ (50 $mmol.L^{-1}$ NaCl final). La fraction d'élution est appliquée sur une colonne de chromatographie échangeuse de cations (POROS^{TM} HS 20 µm; 4,6 x 50 mm, Applied Biosystems) préalablement équilibrée avec 10 % de tampon E (**Annexe 2**). Les protéines sont éluées par un gradient linéaire de concentration de 50 à 600 $mmol.L^{-1}$ NaCl contenus dans le tampon E (**Annexe 2**). Les fractions d'élution (collectées par 5 mL) sont analysées par SDS-PAGE 12 %. Les fractions enrichies en protéines d'intérêt sont concentrées jusqu'à un volume de 10 mL en tube Amicon®Ultra 10K (Millipore) par centrifugation successive de 15-20 minutes à 5 000 g à froid. Une fois les conditions de clivage de l'étiquette His_6 optimisées, le pool HS20 concentré contenant la protéine d'intérêt avec son tag His_6 est incubé à 4 °C pendant 16 h dans du tampon contenant 20 $mmol.L^{-1}$ Tris-HCl pH 8,4; 150 $mmol.L^{-1}$ NaCl; 2.5 $mmol.L^{-1}$ $CaCl_2$ avec un ratio molaire de 1:100 de thrombine (Millipore®) diluée dans du tampon contenant 50 $mmol.L^{-1}$ sodium citrate pH 6,5; 200 $mmol.L^{-1}$ NaCl; 0,1 % PEG 8000 et 50 % glycérol. Le protéolysat résultant est passé au travers d'une colonne benzamidine (GE Healthcare Life Sciences) pour éliminer la protéase et au travers d'une colonne IMAC équilibrée en tampon A pour éliminer les éventuelles traces de protéine d'intérêt ayant encore le tag His_6. Les fractions d'élution passées au travers de la colonne IMAC sont rassemblées et concentrées pour ensuite être soumises à une chromatographie d'exclusion-diffusion sur colonne S75 (Hiload, 16/600, 120 mL, GE Healthcare Life Sciences) pré-équilibrée avec du tampon F (**Annexe 2**). Les fractions d'élution enrichies en protéine d'intérêt sont rassemblées et concentrées à environ 10 $mg.mL^{-1}$ en tube Amicon®Ultra 10K (Millipore) avec du tampon G (ou du tampon H si la protéine RBD supporte bien une plus faible concentration en NaCl, **Annexe 2**). Un lot de protéines concentrées est conservé à 4 °C pour les essais de cristallographie, l'autre partie est aliquotée et stockée à -80 °C pour les essais d'interaction protéine/ARN par électrophorèse en retard sur gel (Electrophoretic mobility shift assay, EMSA). L'ensemble des étapes de purification des protéines recombinantes RBD est récapitulé dans la **Fig.46A**.

F.III. Purification des protéines de fusion GST-RBD et GST-NS1 recombinantes

Les culots de bactéries, préalablement décongelés sur glace sont resuspendus à froid dans 40 mL de tampon H (**Annexe 2**); Les suspensions bactériennes résultantes sont incubées pendant 30 minutes à 30 °C en présence de 0,75 $mg.mL^{-1}$ de lysozyme. Trois cycles de congélation (azote liquide) - décongélation (bain-marie 30 °C) et 5 cycles de sonication à froid de 1 minute (sonde S&M-220, 40 % amplitude avec 5 secondes de pulse et 5 secondes de pause) sont pratiqués pour lyser les cellules.

L'ECB (environ 45 mL) correspond au surnageant filtré à 0,45 µm d'une centrifugation de 80 minutes à 19 000 g. L'ECB, qui contient les protéines de fusion d'intérêt, est appliqué sur une colonne de chromatographie d'affinité au glutathion de 5 mL (GSTrap FF, GE Healthcare Life Sciences). La colonne est lavée avec 10 cv de tampon J (**Annexe 2**). Les protéines de fusion (GST) sont éluées avec 10 cv de tampon K (**Annexe 2**). Chacune des fractions (*Flow through*, lavage et élution) est conservée pour une analyse par SDS-PAGE 12 %. Les protéines sont concentrées le plus possible (en évitant la précipitation) en tube Amicon®Ultra 50K (Millipore).

Le clivage de la GST est effectué sur une moitié des protéines GST-NS1 concentrées (lot A) en présence d'un ratio molaire de 1:1 000 de facteur Xa (Sigma-Aldrich) pendant 20 h en présence de tampon L (**Annexe 2**). Les protéines clivées (lot A) sont passées au travers d'une colonne de benzamidine (GE Healthcare Life Sciences) liée à la colonne GSTrap pour éliminer la thrombine. Les protéines NS1 non-tagguées sont décrochées de la colonne GSTrap en utilisant 5 à 10 cv de tampon L (**Annexe 2**). Les protéines sont concentrées jusqu'à 2 mL en tube Amicon®Ultra 10K (Millipore) puis soumises à une chromatographie d'exclusion-diffusion sur colonne S75 (Hiload, 16/600, 120 mL, GE Healthcare Life Sciences) pré-équilibrée en tampon M (**Annexe 2**). Les protéines sont concentrées le plus possible en tube Amicon®Ultra 10K (Millipore), aliquotées puis stockées à -80 °C pour les essais d'interaction par électrophorèse en gel retard.

Les fractions d'élution contenant les protéines GST-NS1 non-clivées (lot B) sont concentrées (2 mL) et filtrées à 0,22 µm puis injectées (i) sur une colonne de chromatographie d'exclusion-diffusion (S75, GE Healthcare Life Sciences) pré-équilibrée en tampon M (**Annexe 2**) et (ii) sur une seconde colonne de chromatographie d'exclusion-diffusion sur colonne S200 (Hiload, 40 cm x 20 mm, 24 mL, GE Healthcare Life Sciences) pré-équilibrée en tampon N (**Annexe 2**). Les protéines sont concentrées en tube Amicon®Ultra 10K (Millipore), aliquotées puis stockées à -80 °C pour les essais d'interaction en retard de migration sur gel.

Les formes fusion GST-RBD et GST-NS1 entière ont également préparées en omettant dans le protocole de purification ci-dessus l'étape de digestion au Facteur Xa. L'ensemble des étapes de purification des protéines de fusion GST-RBD et GST-NS1 est récapitulé dans la **Fig.47B**.

F.IV. Analyse des fractions protéiques

Les mélanges protéiques sont analysés par SDS-PAGE de 10 à 12 %. Après électrophorèse, les protéines sont colorées avec le kit InstantBlue™ Coomassie Protein Stain (Expedeon®). L'homogénéité des protéines purifiées est vérifiée par spectrométrie de masse (**Fig.46A**). Une concentration de 2 µmol.L^{-1} de protéines purifiées est mélangée à une solution d'acide 4-hydroxy-α-cyano-cinnamique (4-HCCA) saturée dans un mélange acétonitrile : eau : acide trifluoroacétique 33,3:66,6:0,1 (TWA). Le mélange est déposé sur une cible plaquée or selon la méthode en couche ultramince (Gabant et Cadene, 2008). Les dépôts sont rincés en TFA 0,1 % avant l'acquisition des données à l'aide d'un MALDI-TOF/TOF

LIFT UltraFlex I (Bruker Daltonics). Les spectres ont été acquis en mode linéaire positif dans la gamme m/z 1000-20000 avec le logiciel Flexcontrol 3.3 (Bruker Daltonics). Un mélange de cytochrome C et d'apomyoglobine sert de calibrage externe lors de l'analyse et de l'exploitation des données à l'aide du logiciel Flexanalysis 3.3 (Bruker Daltonics).

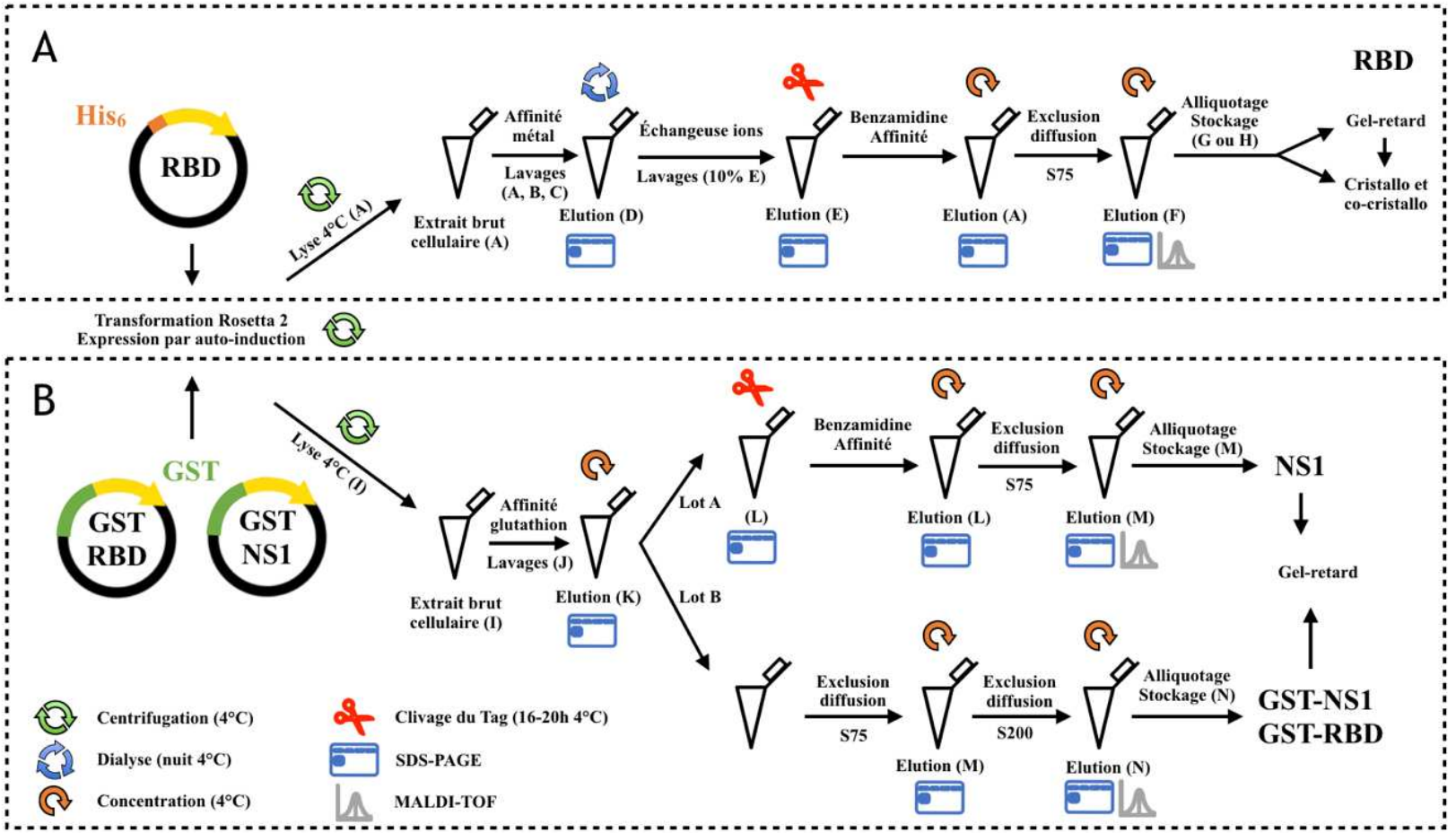


Figure 46 : Purification des protéines recombinantes RBD, GST-RBD et GST-NS1. *(A) Purification des protéines recombinantes RBD. (**B**) Purification des protéines de fusion GST-RBD et GST-NS1.*

G. Retards de migration sur gel

Dans ce type d'expérience, un complexe entre un ARN (aptamère ou ARNdb) radiomarqué en 5' au phosphore 32 (^{32}P) et la protéine est identifié après électrophorèse et autoradiographie du gel par un retard de migration électrophorétique de l'ARN. Ainsi, l'ARN libre migre plus loin dans le gel que lorsqu'il est associé à la protéine. On parle d'un effet retard indiquant la formation d'un complexe ARN/Protéine. Ce type de technique permet de caractériser une interaction entre une protéine et un acide nucléique d'un point de vue qualitatif (formation d'un ou plusieurs complexes) et quantitatif (détermination de la constante de dissociation des complexes) (Ryder *et al*., 2008). Les ARN utilisés dans le but d'évaluer l'affinité d'interaction avec les RBD purifiés ont été synthétisés à l'échelle 40 nmol chez Eurogentec avec une qualité SePOP (Selective Precipitation Optimized Process).

G.I. Radio-marquage et hybridation des ARNdb à évaluer

Les ARN sont préalablement purifiés par électrophorèse en présence de formamide sur un grand gel de polyacrylamide à 12 % contenant de l'urée puis élués en H_2O au travers d'une colonne de cellulose diéthylaminoéthyle (DEAE). Dix picomoles d'ARN sont radiomarqués au ^{32}P (PerkinElmer Health Science) en incubant le mélange à 37 °C pendant 30 minutes avec 1 U de polynucléotide kinase T4 (PNK) (Thermo Scientific). Les ARN radiomarqués sont dessalés par chromatographie d'exclusion sur colonne micro-spin G25 (50-150 µm, BioRad) pré-équilibrée avec de l'eau milliQ. Les ARN radiomarqués résultants sont séchés par évaporation au speedvac™ (Savant) puis resuspendus dans 20 à 40 µL de tampon contenant 10 mmol.L^{-1} Tris-HCl; 1 mmol.L^{-1} EDTA; 500 mmol.L^{-1} NaCl (TEN). Les ARN arborant un hairpin sont dénaturés 2 minutes à 100 °C puis hybridés à 4 °C sur la nuit. Les ARNdb palindromiques sont concentrés puis hybridés pendant 2 minutes dans un bécher d'eau sous-agitation porté à ébullition puis refroidi à 4 °C sur la nuit.

G.II. Détermination des constantes de dissociation (K_d) des complexes RBD/ARNdb

Des aliquots concentrés de protéines purifiées, préalablement décongelés sur glace, sont dilués en série avec un tampon contenant 10 mmol.L^{-1} Tris-HCl; 1 mmol.L^{-1} EDTA; 100 mmol.L^{-1} NaCl; 0,1 % BSA; 10 % glycérol (TNBG). 100 pmol.L^{-1} d'aptamère ARN ou d'ARNdb radiomarqué sont incubés à 4 °C pendant 30 minutes avec des dilutions de la protéine dans un volume final de 40 µL contenant 4 µmol.L^{-1} d'ARN de transfert (ARNt) de levure (Sigma-Aldrich). Les complexes RBD/ARN (10 µL) contenus dans les mélanges réactionnels sont séparés de la sonde ARN par électrophorèse (80 minutes à 110 V) en gels de polyacrylamide 10 % (dimension des plaques : 8,3 x 7,3 cm, épaisseur des spacers :

1,5 mmol.L^{-1}, acrylamide/Bis-acrylamide, 29:1) (MiniProtean, BioRad) en tampon contenant 89 mmol.L^{-1} Tris; 89 mmol.L^{-1} borate; 2 mmol.L^{-1} EDTA (TBE). Les gels sont ensuite séchés sur papier whatmann et exposés sur la nuit en cassettes d'exposition (Molecular Dynamics™). Les ARN libres et les complexes RBD/ARN sont révélés avec le programme phosphorimaging de l'imageur Typhoon FLA 9500 (GE Healthcare Life Sciences). L'intensité de chaque bande est quantifiée avec le logiciel ImageQuant (GE Healthcare Life Sciences). Les constantes de dissociation Kd (nmol.L^{-1}) sont déterminées avec le logiciel GraphPad en utilisant la fonction logistique de régression non-linéaire d'équation $Y = \frac{A2 + (A1 - A2)}{(1 + (x/x0)p)}$, dans laquelle A_1 correspond à l'intensité du signal obtenue avec une concentration de protéines pour laquelle l'ARN est totalement libre et A_2 à l'intensité du signal obtenue avec une concentration de protéines pour laquelle tout l'ARN est sous forme complexée. x représente la concentration de protéine (nM) variant dans notre essai. x0 correspond à la concentration en protéine qui nécessaire pour former 50 % de complexe protéine/ARN, c'est-à-dire à la valeur du Kd. P (power) représente indirectement la vitesse de formation des complexes protéine/ARN. Plus sa valeur est élevée et plus les complexes se forment rapidement.

G.III. Compétition en gel-retard

Cent pmol.L^{-1} de l'ARNdb AWFC01 radiomarqué sont incubés à 4 °C pendant 15 minutes avec une gamme de concentrations croissantes en ARN compétiteur non-radiomarqué. A cet instant, 1 nmol.L^{-1} de protéine RBD purifiée préalablement diluée en série en tampon TNBG additionné de 4 µmol.L^{-1} d'ARNt de levure (Sigma-Aldrich) est ajouté au mélange des deux ARN. A l'équilibre (incubation à 4 °C pendant 15 minutes), les complexes RBD/ARNdb (10 µL) contenus dans les mélanges réactionnels sont séparés de la sonde ARN par électrophorèse dans les mêmes conditions que précédemment (voir plus haut). Dans ces conditions expérimentales, l'EC50 de l'ARN compétiteur correspond à la concentration en compétiteur nécessaire pour déplacer la moitié (soit 25 %) des complexes RBD/AWFC01 initialement formés. Les valeurs d'EC50 sont déterminées comme ci-dessus.

H. Etudes cristallographiques du RBD-B de NS1 et de ces complexes avec l'ARNdb

La cristallisation d'une macromolécule biologique en solution consiste à diminuer sa solubilité pour atteindre un état de sursaturation pouvant conduire, dans le meilleur des cas, à la formation de cristaux. Dans la technique de cristallisation par diffusion de vapeur, une goutte contenant un mélange d'agent précipitant et de protéine concentrée est scellé dans une chambre (paraffine) en présence d'un précipitant pur. La vapeur d'eau diffuse alors jusqu'à osmolarité entre la goutte et le précipitant pur. La déshydratation de la goutte provoque une lente concentration de protéines et de précipitant jusqu'à ce qu'un équilibre soit atteint, idéalement situé dans la zone de nucléation des cristaux du diagramme de phase (**Fig.47**). Étant donné le caractère empirique de la cristallisation des macromolécules, il est nécessaire de réaliser de multiples essais de cristallisation en faisant varier le pH, la température, la nature et la concentration des sels et/ou de l'agent précipitant ou encore l'utilisation de solvants organiques.

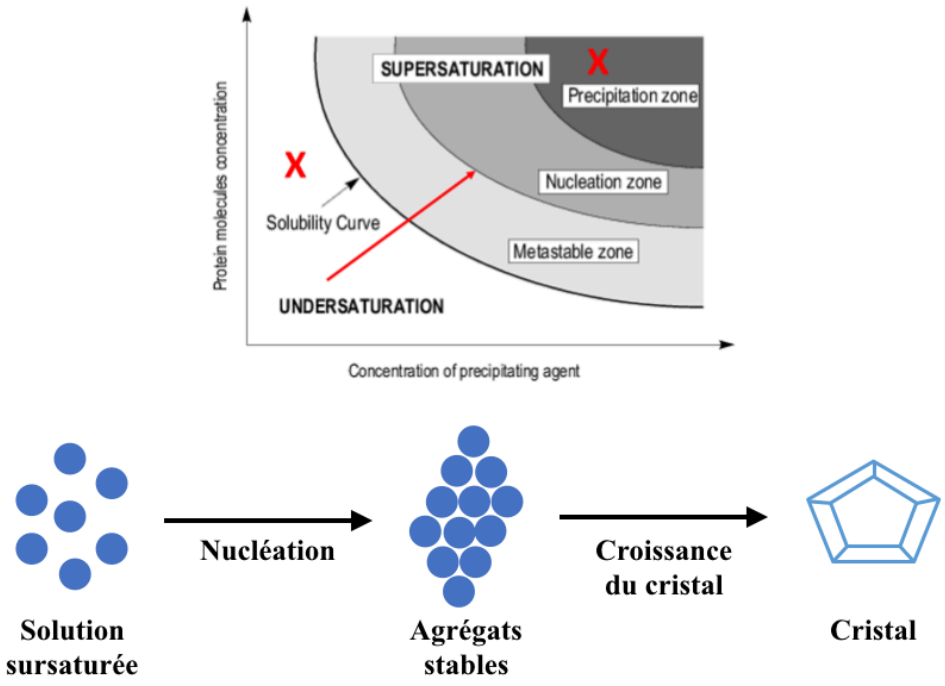


Figure 47 : Diagramme de phase. *Les macromolécules biologiques solubles en solution sont concentrées en présence d'agents cristallisants (sel, polymères et/ou alcool) qui réduisent la solubilité de la macromolécule biologiques jusqu'à atteindre un état de sursaturation. La nucléation s'initie via la formation d'amas moléculaires stables qui, lorsqu'ils dépassent une taille critique, donnent naissance au cristal (d'après Giegé et Sauter, 2010).*

H.I. Cristallogenèse

H.I.1. Cristallisation du RBD-B de NS1

Les RBD purifiés sont concentrés (Amicon 10 kDa, Millipore) à 10 mg.mL^{-1} dans du tampon contenant 20 mmol.L^{-1} HEPES pH 7,6; 250 mmol.L^{-1} NaCl. La détermination des conditions de cristallisation est menée par diffusion de vapeur en goutte assise (*sitting drop vapor diffusion*). Les plaques de cristallisation (Swissci 96-well 2-drop MRC) sont pré-remplies avec l'automate de pipetage Evo75 (Tecan) avec les kits de cristallisation Morpheus, JCSG+, Wizard Classic et Wizard Cryo (Molecular Dimensions, Ltd.) et le kit de cristallisation $AmSO_4$ (Qiagen). Les gouttes de cristallisation sont réalisées à l'aide du robot mosquito® (TTP Labtech) en mélangeant 100 nL de protéine concentrée et 100 nL des solutions contenues dans les puits. Les plaques sont ensuite scellées avec un film plastique, identifiées puis placées en chambre régulée à 20 °C. L'apparition de cristaux dans chacune des conditions de la plaque est suivie à l'aide d'une loupe binoculaire (**Fig.48**). Les cristaux du RBD H7N1 (RBD-B) ont été obtenus après plusieurs jours d'incubation dans la condition 2.5 du kit Morpheus (Molecular Dimensions) (100 mmol.L^{-1} HEPES/MOPS pH 7,5; 8 % ethylene glycol; 30 % P550MME-PEG 20K) avant d'être congelés rapidement puis stockés en azote liquide jusqu'à leur analyse sur la ligne de lumière PROXIMA 2 (Synchrotron SOLEIL).

H.I.2. Cristallisation des complexes RBD-B/ARNdb

Les RBD purifiés sont concentrés à 10 mg.mL^{-1} comme indiqués précédemment. Les ARN synthétiques purifiés (HPLC) et dessalés (Dharmacon) arrivés lyophilisés sont resuspendus concentrés (3 mmol.L^{-1}, pour optimiser la formation d'ARNdb car réaction bimoléculaire pour obtenir notre duplex) dans 10 mmol.L^{-1} Tris; 1 mmol.L^{-1} EDTA pH 8.0; 500 mmol.L^{-1} NaCl. Le duplex d'ARN formé est ensuite hybridé à 100 °C pendant 2 minutes puis refroidi à 4 °C sur la nuit (la proportion de duplex hybridé étant vérifiée lors des essais par retard de migration sur gel). Les complexes RBD/ARNdb sont réalisés à 4 °C dans un volume de 100 µL en incubant un ratio molaire de 1:1,1 de protéine RBD-B fraichement préparée et d'ARNdb synthétique à une concentration de 5 à 10 mg.mL^{-1}. Les gouttes de cristallisation sont réalisées comme précédemment en mélangeant 100 nL des complexes RBD/ARNdb et 100 nL des solutions contenues dans les puits des plaques de cristallisation (Swissci 96-well 2-drop MRC) qui contiennent les mêmes kits de cristallisation que pour le RBD-B. Les cristaux (i) du complexe RBD RBD-B/AWFC01 ont été obtenus dans la condition 2.26 du kit Wizard Classic (100 mmol.L^{-1} CHES pH 9,5; 30 % PEG 400), (ii) du complexe RBD aaRBD-B/AWFC01 (mutant R38A-K41A, indiqué par « aa ») dans la condition E3 du kit Wizard Cryo (100 mmol.L^{-1} sodium phosphate/acide citrique pH 4,2; 200 mmol.L^{-1} ammonium sulfate; 40 % ethylene glycol) et (iii) du complexe RBD RBD-B/ZKO (ARNdb qui dérive de celui utilisé par Cheng *et al*., 2009) de qualité suffisante pour faire diffracter les rayons X ont été obtenus après optimisation de la condition 10 du kit $AmSO_4$ (Qiagen) (2

mol.L^{-1} ammonium sulfate; 200 mmol.L^{-1} ammonium nitrate) par diffusion de vapeur en goutte suspendue (*hanging drop vapor diffusion*).



Figure 48 : Obtention de cristaux et collecte des données de diffraction

H.II. Résolution des structures tridimensionnelles

H.II.1. Résolution de la structure du RBD-B

Les cristaux du RBD-B préalablement récupérés à l'aide d'une boucle de diamètre adapté (Hampton Research) sont trempés (au besoin) dans une solution cryo-protectrice, congelés dans un bain d'azote liquide puis placés en tubes cryocap (Hampton Research) préalablement refroidi en azote liquide. Chaque cristal obtenu est analysé pour ses propriétés de diffraction des rayons X. Les données de diffraction sont collectées à 100 Kelvins à une longueur d'onde de 0.98 Å sur les lignes de lumière PROXIMA-2 du synchrotron SOLEIL (**Fig.48**). Des clichés de diffraction sont pris dans 4 orientions différentes (0 ; 90 ; 180 ; 270°) pour juger de la qualité des données de diffraction. Un total de 3600 images correspondant à un pas de 0,1° autour de l'axe ω du goniomètre sont enregistrées pour les cristaux montrant de bonnes statistiques au niveau de leurs données de diffraction. Des données numériques, extraites des clichés de diffraction à l'aide du logiciel *XDS* (Kabsch, 2010), permettent de déterminer les paramètres de maille a, b et c (en Å) et les angles α, β, γ (en °) du cristal ainsi que son groupe d'espace. Les intensités diffractées sont ensuite triées et mises à l'échelle grâce à la fonction *aimless* de la suite *CCP4* (Winn *et al*., 2011; Evans et Murshudov, 2013). La structure tridimensionnelle (3D) du RBD-B a été résolue par remplacement moléculaire avec la fonction *phaser* (Winn *et al*., 2011;

McCoy, 2007) de la suite *Phenix* (Adams et al., 2010; Afonine *et al*., 2012) en utilisant la structure PDB *1AIL* (Liu *et al*., 1997) comme structure de référence.

H.II.2. Résolution de la structure des complexes RBD-B/ARNdb

Chacun des cristaux de nos complexes a été préalablement récupéré à l'aide d'une boucle de diamètre adapté (Hampton Research), trempé (au besoin) dans une solution cryo-protectrice, congelé dans un bain d'azote liquide puis placé en tubes cryocap (Hampton Research) préalablement refroidi en azote liquide. Les données de diffraction recueillies pour chacun des complexes étudiés ainsi que l'analyse de leur structure ont été réalisées dans des conditions similaires à celles utilisées pour le RBD-B, mais en utilisant la structure PDB *2ZKO* (Cheng *et al*., 2009) comme structure de référence.

H.II.3. Représentation des structures 3D, analyse des paramètres hélicoïdaux des ARNdb et visualisation des interactions RBD/ARNdb

L'ensemble des structures 3D a été réalisé avec le logiciel *UCSF Chimera* (Pettersen *et al*., 2004). L'analyse des paramètres hélicoïdaux des ARNdb complexés dans nos complexes a été faite avec le logiciel *Curves+* (Blanchet *et al*., 2011). La cartographie des interactions protéine/ARNdb a été réalisée avec la fonction *SNAP* de la suite *3DNA* (Lu et Olson, 2008).

RÉSULTATS

A. Relevance biologique des séquences virales identifiées par SELEX

A.I. Des motifs ARN caractéristiques des ARN viraux de polarité positive

Marc *et al*. (2013) ont utilisé la technique de SELEX pour obtenir des aptamères d'ARN de haute affinité pour la protéine NS1. Cette technique repose sur une sélection systématique de ligands par enrichissement exponentiel à partir de banques combinant des oligonucléotides synthétiques. La SELEX fait intervenir plusieurs cycles où (i) les oligonucléotides, qui contiennent des séquences invariables en bords de séquence, sont incubés une première fois avec la protéine d'intérêt ; (ii) les protéines sont lavées à plusieurs reprises pour éliminer les oligonucléotides qui se seraient fixés de manière non-covalente et (iii) les oligonucléotides restés accrochés à la protéine sont élués puis amplifiés par PCR par l'intermédiaire des séquences invariables. Un nouveau cycle démarre jusqu'à ce que l'on obtienne des oligonucléotides (aptamères) de très haute affinité. C'est ainsi que près de 63 aptamères de haute affinité pour NS1 ont pu être identifiés, dont certains (les aptamères DM01 et DM03) présentent des séquences d'intérêts.

Les aptamères identifiés par l'approche SELEX (Marc *et al*., 2013) ont été classés en plusieurs groupes parmi lesquels figurent les aptamères DM01 et DM03 qui présentent respectivement les motifs 5'-AGCAAAAG-3' et 5'-UGAUUGAAG-3' (motif A et motif B, respectivement, **Fig.49**) caractéristiques des ARNm viraux des virus *influenza*. Le motif A est strictement conservé en 5' de tous les ARN viraux de polarité positive (ARNc et ARNm viraux) alors que le motif B est retrouvé essentiellement dans les ARNm viraux codant pour NS1 (**Fig.49**).

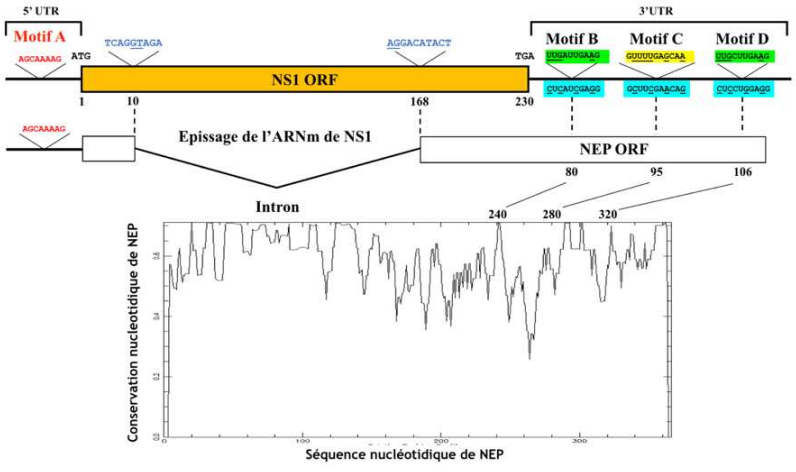


Figure 49 : Localisation des motifs viraux identifiés par SELEX dans les ARNm de NS1. *Les motifs viraux A (indiqué en rouge) et B (indiqué en vert) sont respectivement localisés dans les régions 5' et 3'-UTR qui bordent l'ORF de NS1 (jaune). Les mutations des motifs viraux B, C et D sont représentées en noir. Les sites donneur et accepteur d'épissage de l'ARNm de NS1 sont représentés en bleu, avec les dinucléotides GT et AG aux extrémités 5' et 3' de l'intron. Les codons start (atg) et stop (tga) sont également représentés. La conservation nucléotidique de NEP pour les virus H1N1, H2N2, H3N2 et H1N1 pdm09 a été réalisée sur une fenêtre glissante de 7 nucléotides en utilisant la fonction plotcon du logiciel mEMBOSS (version 6.6.0.0 ; Rice et al., 2000). Les motifs viraux B, C et D sont très conservés dans la séquence nucléotidique de NS1 et de NEP.*

L'identification, dans les aptamères se liant à NS1, de motifs spécifiques des ARN viraux de polarité positive pouvait suggérer que NS1 interagit au cours du cycle viral avec les ARN messagers viraux et/ou les ARN complémentaires du virus. Afin d'examiner cette hypothèse, nous avons choisi de modifier, par mutagénèse dirigée, les motifs identifiés dans la séquence du segment génomique NS. Pour ce faire, nous avons introduit des mutations dans le plasmide pPol I du segment NS utilisé dans la génétique inverse afin de produire des virus mutants. Le cadre de lecture de NS1 a été exactement substitué par la séquence codant pour la luciférase Renilla. Ainsi, ce plasmide sera transcrit en un *viral-like genomic RNA*, lequel pourra être répliqué et transcrit par la RdRp virale.

A.II. Choix des mutations

Le motif A (AGCAAAAG) correspond au promoteur de la RdRp virale, et nous avons pensé qu'il n'était pas judicieux de le cibler dans cette étude du fait de son importance dans les étapes de transcription et de réplication du génome viral. Sa mutation pourrait fortement impacter le cycle du virus. En revanche, nous avons ciblé le motif B ^{733}UUGAUUGAAG^{742}, ainsi que deux autres sites dans la région 3'-UTR de l'ARNm de NS1, que nous avons appelés motif C et motif D. Ces deux motifs (C ^{777}GUUUUGAGCAA^{787} et D ^{814}UUGCUUGAAG^{823}) ont été sélectionnés pour leur haute conservation dans la séquence codante de NEP (**Fig.49**). Le motif D est très proche du motif B, dont il se distingue par une unique mutation (**Fig.49 et 50**). Les motifs C et/ou D seront utilisés comme contrôles (**Fig.50**). Les mutations dans ces trois motifs ont été conçues de manière à respecter la séquence peptidique de NEP et à privilégier, dans la mesure du possible, les transitions purine > purine ou pyrimidine > pyrimidine (**Fig.49 et 50**).

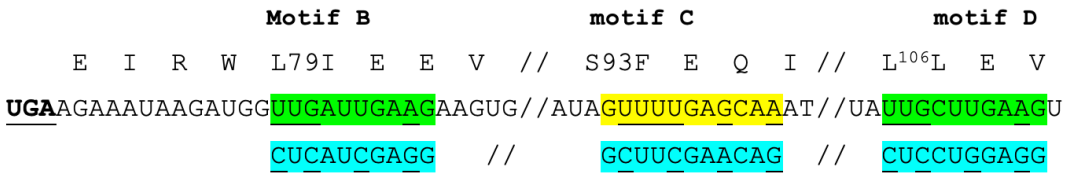


Figure 50 : Mutations des motifs B, C et D dans la 3'-UTR de NS1. *Les motifs viraux B-D et C sont respectivement représentés en vert et en jaune. Les mutations de bases indiquées en bleu sont soulignées. La 3'-UTR de NS1 se superpose à l'ORF de NEP.*

A.II.1. Impact des mutations dans le système *miniréplicon viral*

Selon notre hypothèse de travail, NS1 en interagissant avec le motif B (voire aussi avec le motif D puisque ce dernier est très proche du motif B), pourrait moduler la transcription et la traduction des propres transcrits viraux. Afin de tester cette hypothèse, nous avons mis au point un système *miniréplicon viral* dans lequel l'ORF de NS1 a été remplacée par l'ORF de la luciférase Renilla (Ren-Luc). La 5'-UTR de NS1 a été conservée intacte alors que sa 3'-UTR a été mutée ou non aux sites B, C

ou D (**Fig.49 et 50**). Dans ce système, des cellules HEK293T sont transfectées par un minimum de cinq plasmides, dont quatre permettent l'expression transitoire du complexe RdRp viral (PA, PB1, PB2, NP) alors que le cinquième assure l'expression transitoire du *miniréplicon* chimérique (**Fig.51**). La mesure de l'activité Ren-Luc permet d'évaluer l'activité de la RdRp virale (transcription et réplication du *miniréplicon*), ainsi que le niveau de traduction de l'ARNm transcrit à partir du *miniréplicon*. Nous avons appliqué ce système à notre approche, afin de rechercher un éventuel effet de NS1 sur la transcription et la traduction des ARNm viraux (voire sur la réplication), et aussi afin d'examiner l'impact des mutations des motifs B, C et D sur le *miniréplicon*. Quatre variantes du *miniréplicon* ont ainsi pu être comparées (WT, mutB, mutC et mutD).

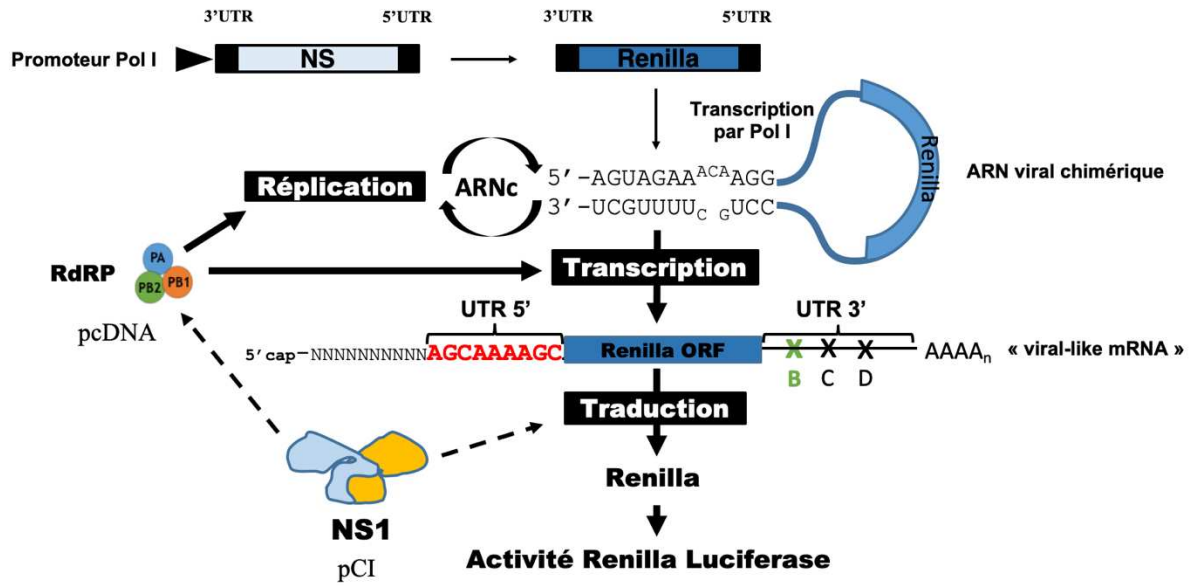


Figure 51 : Principe du *miniréplicon viral*. *Des cellules HEK293T sont transfectées avec les quatre vecteurs d'expression des sous-unités de la RdRp virale (PA, PB1, PB2 et NP), et avec un vecteur permettant la transcription (PolI dépendante) d'un ARN viral chimérique. NS1 est co-exprimée transitoirement, afin de rechercher ses effets éventuels (flèches rouges pointillées) sur l'activité de la RdRp virale (réplication et transcription) et sur la traduction des « viral-like mRNA ». La comparaison de l'activité luciférase entre les conditions « sans NS1 » et « avec NS1 » permet d'évaluer ces effets. L'activité Ren-Luc est normalisée relativement à l'activité FF-Luc (non représentée ici).*

A.II.1.a. Expression de NS1 dans les cellules transfectées

Avant d'évaluer l'effet de NS1 dans ce système *miniréplicon viral*, nous avons vérifié le niveau d'expression transitoire de NS1 dans les cellules transfectées pour cette approche. Comme le montre la **Fig.52**, les deux variants de la protéine NS1 (WT et R38A) sont bien exprimés dans les cellules HEK293T transfectées (la présence d'une bande dans la condition mock correspond à un artefact de migration lié au puits voisin qui contient la condition WT), et nous pourrons donc évaluer leur effet dans le système *miniréplicon*.



Figure 52 : Niveau d'expression de NS1 dans les cellules transfectées. *Des cellules HEK293T, transfectées avec 100 ng de vecteur pCI-vide, pCI-NS1 ou pCI-NS1 R38A, ont été lysées après 24 h. Les protéines du lysat cellulaire ont été séparées par électrophorèse, puis les protéines NS1 (26 kDa) et GAPDH (38 kDa) ont été révélées par western-blot en utilisant les anticorps polyclonaux et monoclonaux comme indiqué dans la section Matériels et Méthodes.*

A.II.1.b. Effet de NS1 sur l'expression du *miniréplicon viral*

Dans ce système nous avons comparé quatre régions, qui se distinguent par les mutations présentes ou absentes dans la région correspondant à la 3'-UTR de NS1 : wt, mutB, mutC et mutD. Dans tous les cas, l'activité Ren-Luc mesurée est normalisée par rapport à l'activité Luciferase Firefly (FF-Luc) afin de réduire d'éventuels biais qui seraient liés notamment à l'efficacité de transfection.

Nous avons observé que le *miniréplicon* « mutB » produit un signal environ 20 fois inférieur à celui mesuré pour les *miniréplicons* WT, mutC et mutD. Toutefois, nous avons par la suite identifié une mutation très délétère dans le vecteur du *miniréplicon* mutB : une délétion d'environ un tiers du promoteur de PolI, laquelle très probablement réduit considérablement l'expression du *miniréplicon*. Compte tenu des niveaux équivalents produits par les trois autres *miniréplicons*, nous avons alors considéré que les comparaisons les plus pertinentes devaient se faire entre les différentes conditions d'un même *miniréplicon* (absence de NS1 / présence de NS1 / présence de NS1 mutante).

De façon générale, on observe que NS1 (points rouges, **Fig.53**) accroît considérablement l'expression du *miniréplicon*, multipliant le signal par un facteur 4 à 6, quelle que soit la 3'UTR. Ceci montre l'effet majeur de NS1, et suggère que cet effet de NS1 sur le *miniréplicon* est indépendant de la 3'-UTR. Il faut souligner que dans tous les cas la 3'-UTR demeure fonctionnelle à l'égard de la réplication et de la transcription, puisque la conservation des séquences U12 et U13 assure la fonctionnalité du promoteur de la RdRp virale. La protéine NS1 mutante (points bleus, **Fig.53**) augmente aussi l'expression de NS1, mais de façon inférieure à la protéine WT, suggérant que le RBD n'est pas le seul domaine de NS1 impliqué dans cette activité. ***NS1 semble donc moduler positivement l'expression du miniréplicon, en agissant sur au moins une des principales étapes du système miniréplicon : transcription, réplication du miniréplicon et traduction du « viral-like mRNA ». Le motif B (5' UGAUUGAAG 3') ne semble pas jouer de rôle ici puisque sa mutation ne modifie pas l'effet de NS1 sur le système, et il en est de même pour les motifs C et D (si l'un de ces motifs avait un rôle important, le facteur d'amplification induit par NS1 serait significativement modifié).***

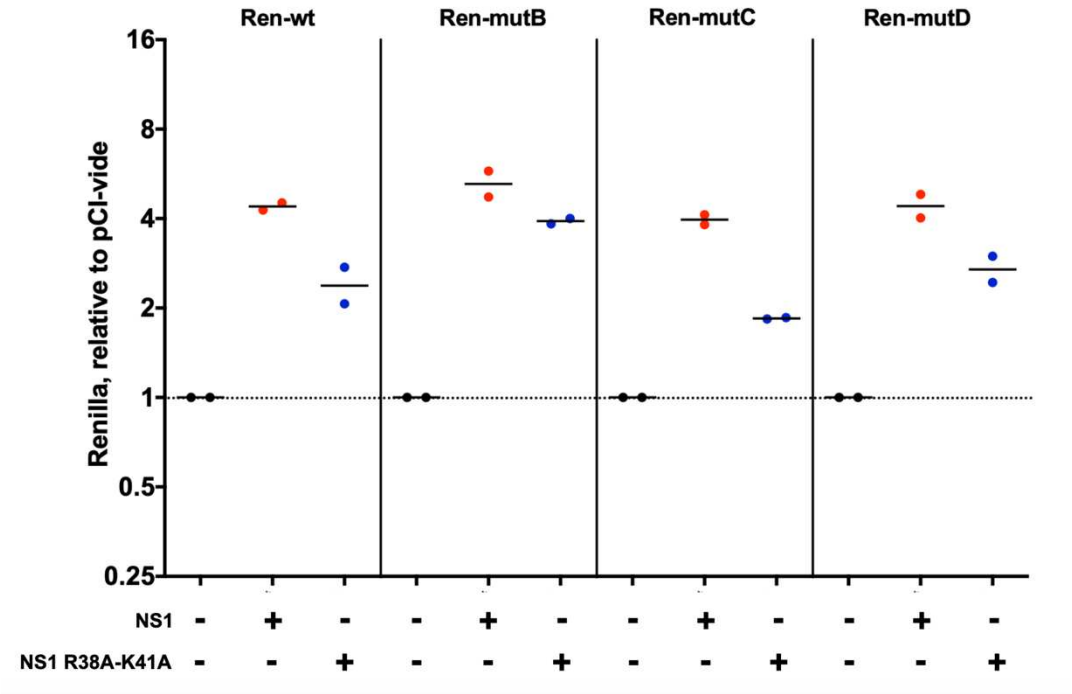


Figure 53 : Effet de NS1 (WT ou mutant R38A-K41A) sur l'expression du *miniréplicon*. *Pour chacun des quatre miniréplicons (WT, mutB, mutC et mutD), les signaux normalisés (Ren-Luc/FF-Luc) sont exprimés relativement au signal normalisé de la condition pCI-vide (absence de NS1, points noirs) de la même construction. Deux expériences indépendantes ont été réalisées, chaque point représente la moyenne géométrique du triplicat technique d'une expérience. Points rouges : pCI NS1 wt; Points bleus : pCI NS1 mutante.*

A.II.1.c. Effet de NS1 sur la traduction des transcrits viraux

Le système *miniréplicon* ne permet toutefois pas de déterminer si NS1 agit sur la réplication ou la transcription du *miniréplicon*, ou bien sur la traduction du *viral-like mRNA*, ou bien encore sur plusieurs de ces trois étapes. Afin de rechercher plus précisément un effet éventuel de NS1 sur la traduction de son propre ARNm, nous avons conçu un second système, que nous avons dénommé *translation-only* (**Fig.54**).

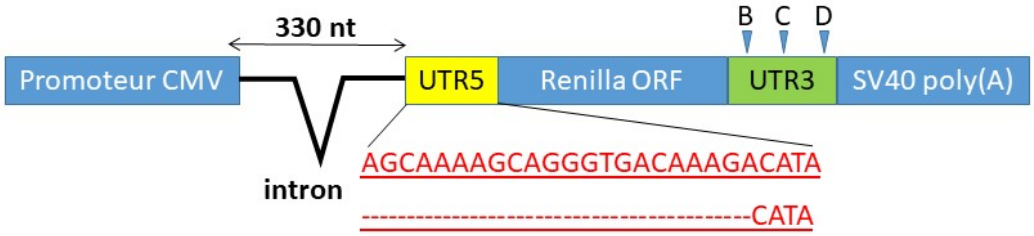


Figure 54 : Système *translation-only*. *Les quatre variants du segment chimérique « Ren-NS » (WT, mutB, mutC et mutD) ont été clonés dans le vecteur d'expression eucaryote pCI (carte simplifiée), soit en conservant l'intégralité de la région 5'-UTR (26 nt incluant le motif AGCAAAAGCA), soit en enlevant la plus grande part de cette région (absence du motif AGCAAAAG). Des cellules HEK293T ont été transfectées par l'un ou l'autre de ces huit vecteurs, en présence ou non du vecteur d'expression de NS1. 24 heures après transfection, l'activité Ren-Luc a été mesurée dans le lysat des cellules, afin de rechercher un éventuel effet de NS1 sur la traduction des différents ARNm. La carte schématise la structure des ARNm produits.*

Des cellules HEK293T ont été transfectées par les différents vecteurs d'expression, accompagnés ou non (pCI-NS ou pCI-vide) du vecteur d'expression de NS1 et du vecteur d'expression de la Firefly luciférase utilisé comme normalisateur. Par cette approche nous cherchons à mettre en évidence un éventuel effet de NS1 sur la traduction des différents variants de l'ARNm chimérique.

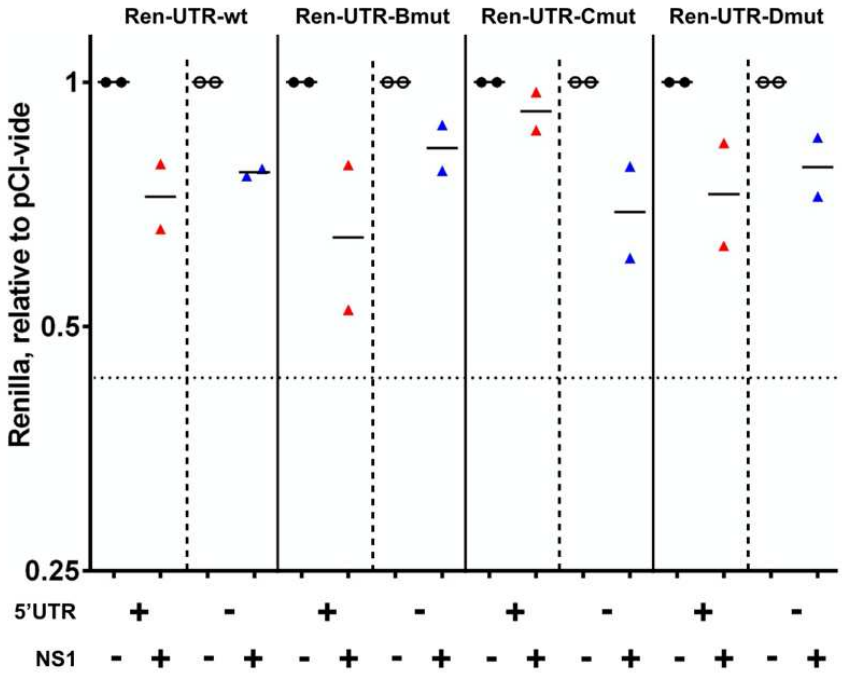


Figure 55 : Effets de NS1 et de ses 5' et 3'-UTR sur la traduction des ARNm de NS1. *Dans chacune des conditions, l'activité Ren-Luc normalisée (activité Ren-Luc/activité FF-Luc) mesurée lorsque NS1 est exprimée est rapportée à l'activité normalisée en l'absence de NS1. Deux expériences indépendantes ont été réalisées, chaque point représente la moyenne géométrique d'un triplicat technique. Rouge : présence de la 5'-UTR de NS1 sur les transcrits viraux; Bleu : absence de la 5'-UTR de NS1 sur les transcrits viraux.*

Nous observons que la présence de NS1 (triangles rouges) réduit d'environ 20 à 25 % le signal produit par les constructions contenant la 5'-UTR de NS1, relativement au signal mesuré en absence de NS1 (points noirs) (**Fig.55**). Le signal en présence de NS1 vaut environ 75 % de celui observé en son absence. La réduction du signal est d'environ 5 à 10 % pour le transcrit portant la mutation C (**Fig.55**). Bien qu'avec des niveaux légèrement différents, l'expression de NS1 réduit également l'expression des transcrits dépourvus de la région 5'-UTR de NS1 (triangles bleus comparés aux cercles noirs). NS1 semble donc réduire modérément l'expression des ARN chimériques, quelles que soient les particularités des régions 5'-UTR (présence ou absence de la 5'-UTR de NS1) ou 3'-UTR (wt ou portant l'une ou l'autre des mutations testées) (**Fig.55**). En combinant cette observation avec les résultats de l'approche *miniréplicon*, il semble donc que l'effet positif de NS1 dans le système *miniréplicon* repose essentiellement sur une amélioration par NS1 de l'activité de la RdRp virale (transcription et/ou réplication).

Toutefois, il importe de remarquer que les transcrits chimériques dont nous avons testé l'expression en présence ou en absence de NS1 sont en réalité très éloignés des ARNm viraux : (i) ils sont transcrits à partir de plasmide par l'ARN polymérase cellulaire Pol II, et non à partir des ARNv par la RdRp virale, (ii) ils sont polyadénylés comme le sont les ARNm cellulaires, et non comme le sont les ARNm viraux, (iii) par la construction utilisée (**Fig.54**), ils sont épissés comme le sont les ARNm cellulaires et enfin (iv) ils portent en 5' et en 3' des extensions artificielles que ne portent pas les vrais ARNm viraux. Pour l'ensemble de ces raisons, il faut reconnaître que notre système *translation-only* ne permet pas de tester fidèlement l'effet de NS1 sur les ARNm viraux.

A.II.2. Phénotype des virus mutants

Les approches *miniréplicon* et *translation-only* n'ont pas permis de déceler un effet des mutations des motifs B, C, et D. Cependant, des effets subtils non observables dans les systèmes artificiels pourraient avoir un impact sur le phénotype viral. Il était donc nécessaire d'examiner le phénotype des virus portant ces mutations. Nous avons introduit chacune des trois mutations (mutB, mutC, mutD) ainsi que toutes leurs combinaisons possibles, dans le plasmide NS du système de génétique inverse du virus WSN. Tous les virus ont été produits par ce système (wt, les trois mutants simples, les trois double-mutants et le triple mutant), et les premières caractérisations n'ont pas révélé de changement notable du phénotype viral. Aussi nous avons choisi de décrire plus précisément les trois mutants simples, ainsi que le triple mutant. Les différents virus (le virus WT, les virus contenant les simples mutations B, C et D et le virus contenant la triple mutation BCD) générés par génétique inverse (voir la section *Matériels et Méthodes*) ont d'abord été amplifiés sur des cellules MDCK avant d'évaluer leur phénotype. Les plages de lyse présentées en **annexe 3** n'indiquent pas de différences de phénotypes entre les différents virus générés suite à la génétique inverse (72h post-amplification des surnageants de culture provenant de la généntique inverse).

A.II.2.a. Courbe de croissance virale

Afin d'évaluer la cinétique de croissance virale, des cellules MDCK ont été infectées à une m.o.i de 0.001 pfu/cellule (m.o.i de 10^{-3} équivaut à 1 virus pour 1000 cellules MDCK) avec les différents virus WSN. Les surnageants de culture contenant les virus produits sont récoltés à différents temps post-infection. Tous les virus mutants sont viables et présentent une cinétique de croissance identique à celle du virus WSN WT, avec environ 10^5 pfu/mL à 12 h post-infection suivi par un plateau de production virale atteignant environ 10^8 pfu/mL à 24 h (**Fig.56**). Les plages de lyse présentées en **annexe 4** n'indiquent pas de différences de phénotypes viraux.

A.II.2.b. Cinétique d'accumulation de NS1 dans les cellules infectées

Pour compléter ces observations, nous avons évalué dans les cellules infectées la cinétique d'accumulation de NS1 au cours du cycle viral de ces mêmes virus. Les cellules MDCK étant connues

pour être partiellement déficientes pour l'activité de Mx dans la réponse antivirale (Seitz *et al*., 2010), nous avons choisi d'évaluer cette accumulation dans des cellules A549 qui présentent une réponse antivirale fonctionnelle. Comme le montre la **Fig.56B**, NS1 et NP sont décelables à 4 h et facilement détectables à 8 h post-infection. Nous n'observons pas de différence notable entre les trois virus (WT, mutB et mutBCD) quant à la cinétique d'accumulation des deux protéines virales. A 24 h post-infection nous n'avons pas non plus observé de différences notables entre les différents virus. L'ensemble de ces résultats montre que les différents virus, porteurs de mutations aux sites B, C et D, ont un phénotype identique à celui du virus WSN WT, pour le potentiel réplicatif ainsi que pour la cinétique d'accumulation des protéines virales. Dans nos conditions expérimentales, aucune des mutations de la 3'-UTR de NS1 ne semble donc impacter le phénotype viral, ce qui invalide notre hypothèse de travail quant au rôle du motif UGAUUGAAG dans une interaction avec NS1 qui serait importante dans le cycle viral.

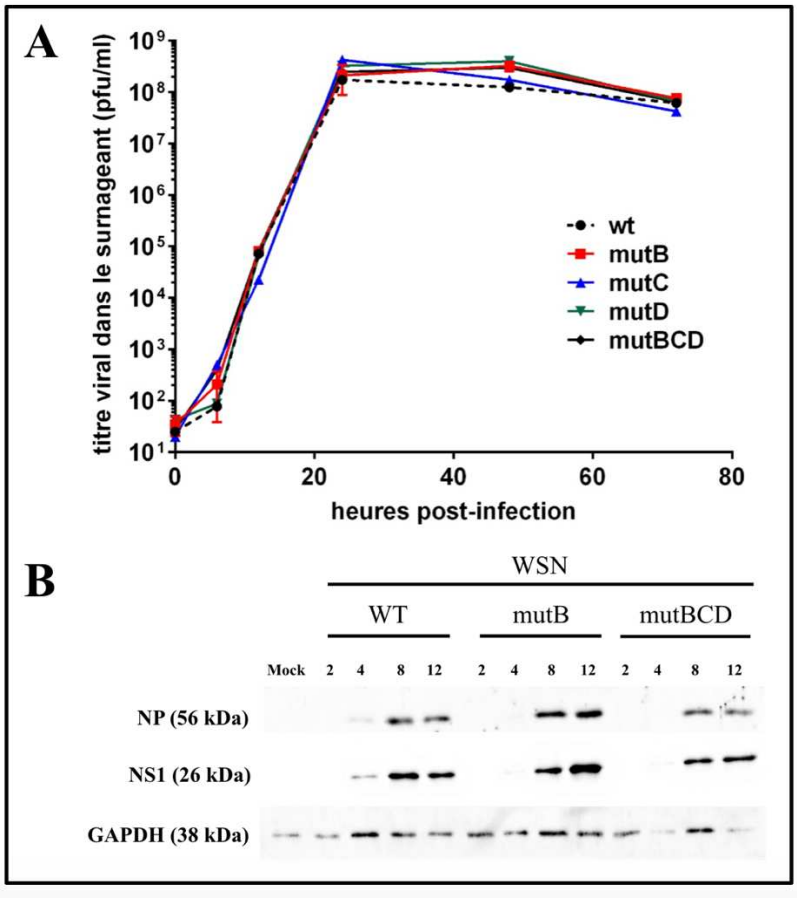


Figure 56 : Phénotype des virus WSN WT et mutants. *(**A**) Cinétique de production virale. Des cellules A549 ont été infectées à une m.o.i de 10^{-3}. Les surnageants, prélevés aux temps indiqués, ont été titrés sur cellules MDCK comme indiqué dans la section Matériels et Méthodes. (**B**) Cinétique d'accumulation de NS1 et NP dans les cellules infectées. Des cellules A549 ont été infectées avec les virus WSN WT, mutB et mutBCD à une m.o.i de 1 puis lysées aux temps 2 h, 4 h, 8 h et 12 h post-infection. Les protéines dans le lysat ont été séparées par électrophorèse, puis les protéines NS1, NP et GAPDH ont été révélées par western-blot. La GAPDH, dont la masse est supérieure à celle de NS1, est placée en bas afin de faciliter la comparaison.*

B. Interaction de NS1 avec l'ARN double-brin

B.I. Identification et sélection d'une séquence minimale d'ARN double-brin de haute affinité pour NS1 inspirée de celles des aptamères DM01 et DM03 identifiées par la méthode SELEX

Mes travaux de doctorat font suite à la découverte des aptamères DM01 et DM03 (Marc *et al*., 2013). L'objectif essentiel des travaux présentés dans ce paragraphe a consisté à décliner les aptamères DM01 et DM03 (ils présentent tous deux des séquences caractéristiques du virus) sous une forme la plus simple et la plus courte possible pour les études fines d'interaction NS1/ARN et pour l'approche structurale par cristallographie. L'ensemble des sondes utilisées dans ce travail de doctorat est présenté en **annexe 12**.

B.I.1. Sélection des protéines RBD et NS1 entières à purifier

En raison de l'évolution rapide des IAV, nous avons fait le choix de produire des protéines recombinantes (NS1 entière ou RBD seul) représentatives de plusieurs virus grippaux d'intérêts en santé humaine. Nos choix se sont portés sur la protéine NS1 de virus humains saisonniers H3N2 et pdm09 (H1N1), de virus aviaires émergents H5N1 et H7N9, de virus de chauve-souris H17N10 et de virus aviaire H7N1 (allèle B de NS1) (**Fig.57**).

Les virus aviaires H5N1 et H7N9 ne sont pas encore pandémiques, mais ils pourraient le devenir dans un futur proche. Toutefois, étudier la protéine NS1 de ces deux virus pose un problème pratique : en effet, ils sont tous deux considérés comme des MOT (Micro-organismes et Toxines) et donc soumis à une règlementation très stricte et une surveillance pointilleuse par l'Agence Nationale de la Sécurité du Médicament (ANSM). Les Organismes Génétiquement Modifiés (OGM) qui en sont dérivés sont considérés comme des MOT s'ils contiennent au moins 500 nt de la séquence issus d'un MOT. La séquence d'une NS1 entière (230 acides aminés) qui avoisine 690 nt correspond à un MOT par conséquent. Mais, le RBD seul (73 acides aminés, 219 nt) n'est plus un MOT. Etudier le RBD de NS1 provenant de ces deux virus ne représente pas un problème, contrairement à l'étude de NS1 entière. Pour pallier à cette difficulté, nous avons fait des constructions dérivées du virus aviaires non-MOT H1N1 A/Mallard/Marquenterre/Z237/83 (MZ) puisque la séquence peptidique de sa NS1 est très proche de celle de virus *influenza* aviaires H5N1.

La NS1 du virus H17N10, nouvellement découverte (Turkington *et al*., 2015; Zhao *et al*., 2016), pourrait aussi être intéressante en raison de son originalité de séquence (**Fig.57**). D'ailleurs, c'est à partir du complexe RdRp d'un virus de chauve-souris (H17N10) que Pflug *et al*. (2014) ont pu résoudre la structure de la RdRp virale par DRX.

La NS1 provenant des virus H3N2, pmd2009, H5N1 et H7N9 est d'allèle A. Nous avons choisi d'inclure la NS1 d'un virus aviaire H7N1 représentant l'allèle B de NS1. Même si l'allèle B de NS1 est généralement considéré comme « non adapté » aux mammifères et minoritaire y compris chez les oiseaux, il n'empêche pas l'infection de mammifères pour autant (Munõz-Moreno *et al*., 2019; Turnbull

et al., 2016). De plus, l'allèle B présente aussi l'avantage de constituer un variant exotique (**Fig.57**), tout comme la NS1 de virus de chauve-souris H17N10, qui pourrait avoir des propriétés différentes intrinsèquement liées à la séquence peptidique.

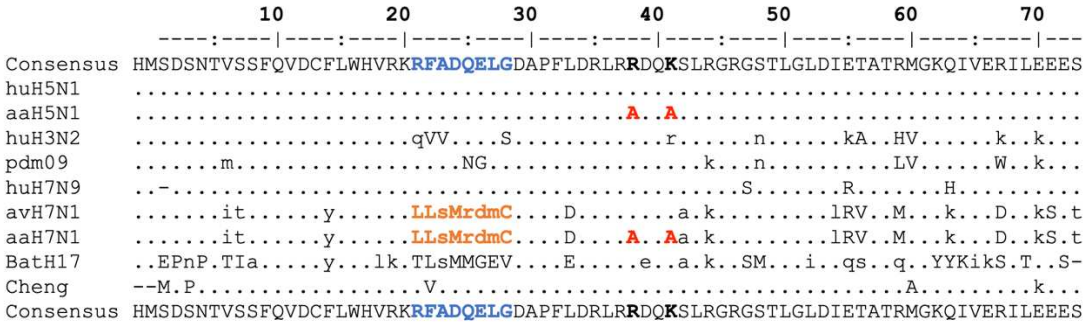


Figure 57 : Alignement de séquences peptidiques des RBD de NS1 étudiés. *Les séquences peptidiques des différentes NS1 utilisées dans notre étude ont été alignées avec les fonctions emma puis showalign du logiciel mEMBOSS (version 6.6.0.0 ; Rice et al., 2000). La NS1 de virus humain H5N1 est indiquée comme séquence de référence. Les variations peptidiques indiquées en minuscule correspondent à des acides aminés similaires en termes de charge. Les variations peptidiques indiquées en majuscule correspondent à des substitutions d'acides aminés de charge différente. Les délétions d'acides aminés sont indiquées par des tirets. Une sérine additionnelle est insérée en position $_{+2}$ de la protéine afin de respecter la N-end rule (Varshavsky, 1996). La NS1 dénommée Cheng est celle qui a été utilisée pour résoudre la structure cristallographique du complexe RBD/ARNdb (Cheng et al., 2009). Les séquences de NS1 d'allèles A et B sont respectivement représentées en bleu et en orange. La double-substitution R38A-K41A est indiquée en rouge.*

B.I.2. Des préparations protéiques purifiées de haute qualité pour les interactions en gel retard et la cristallogenèse

Les protéines recombinantes (RBD, GST-RBD et GST-NS1) utilisées dans mes travaux de doctorat ont été exprimées chez *E. coli* et purifiées jusqu'à homogénéité (voir la section Matériels et méthodes). Ainsi, nous avons purifié selon le même protocole les domaines RBD des protéines NS1 du virus aviaire H7N1 (sauvage et son double mutant R38A-K41A notés RBD-B et aaRBD-B, respectivement), des virus humains H5N1 (sauvage et son double mutant R38A-K41A notés RBD-A et aaRBD-A, respectivement), H3N2, pdm09 (SOIV) et H7N9 et du virus de chauve-souris Bat H17N10. De plus, le domaine RBD de NS1 ainsi que la forme NS1 entière du virus aviaire H7N1 ont été produites sous la forme de protéine fusionnées à la Glutathione-S-Transférase (GST).

Des versions de RBD portant la double mutation R38A-K41A connue pour invalider l'interaction de NS1 à l'ARN double-brin ont également été purifiées afin de servir de contrôle dans nos expériences d'interaction avec l'ARN. Pour des raisons pratiques, le RBD de NS1 du virus aviaire H7N1 ou RBD de NS1 d'allèle B, sera surnommé RBD-B (son mutant R38A-K41A étant indiqué par aaRBD-B). Le RBD de NS1 du virus humain H5N1 sera surnommé RBD-A (son mutant R38A-K41A étant indiqué par aaRBD-A).

En bref et à titre d'exemple, la procédure de purification des protéines RBD recombinantes des ECB repose sur 4 étapes essentielles : (i) une chromatographie d'affinité, (ii) une chromatographie

échangeuse d'ions, (iii) un clivage de l'étiquette His_6 et (iv) une chromatographie par exclusion-diffusion (**Fig.46**). La fraction finale de purification est concentrée et soit utilisée extemporanément pour les essais de cristallisation de la protéine seule ou en complexe avec l'ARN, soit aliquotée et stockée à -80 °C pour les expériences d'interaction avec l'ARN. Tout au long de la purification, l'homogénéité de chaque fraction de purification est estimée par SDS-PAGE (**Fig.58A**). En fin de purification, le contenu de l'échantillon final est analysé par spectrométrie MALDI-TOF (**Fig.58B**). Dans le cas présenté en **Fig.58B**, la masse moyenne observée pour le produit majeur est de 8802,5 Da et correspond parfaitement à la masse théorique d'un monomère du RBD-B. On peut remarquer sur ce spectre la détection d'une très faible quantité de dimère qui n'a pas été détruit lors de la préparation de l'échantillon pour la spectrométrie MALDI-TOF (sensée être dénaturante pour les protéines) et la présence d'un contaminant (moins de 10%) de masse apparente de 9244,9 Da (**Fig.58B**). La présence de dimère après traitement des échantillons pour l'analyse MALDI-TOF témoigne de la robustesse de ce dimère. Cette procédure générale a été suivie pour l'ensemble des protéines purifiées (à noter que pour la NS1 entière d'allèle B et son RBD fusionnés à la GST, la résine d'affinité est une GSTrap à la place de la TALON et l'étiquette GST est conservée pour les expériences comparatives d'interaction avec l'ARN car la forme entière de la protéine précipite rapidement sans son étiquette GST). Les procédures de purification mises en place sont compatibles avec l'analyse structurale par cristallographie du RBD seul ou complexé à un ARN double-brin. Autre détail, nous avons très facilement cristallisé le RBD Bat H17N10 (**Annexe 5**) et obtenu la structure tridimensionnelle (3D) de ce dernier. La structure résolue au laboratoire (et qui n'est pas présentée dans ce manuscrit) correspond en tout point à la structure résolue par Turkington *et al.* (2015; PDB : 5BXZ), attestant ainsi de la qualité de nos préparations protéiques pour les essais d'interaction en gel retard et la cristallogenèse des protéines et des complexes RBD/ARN.

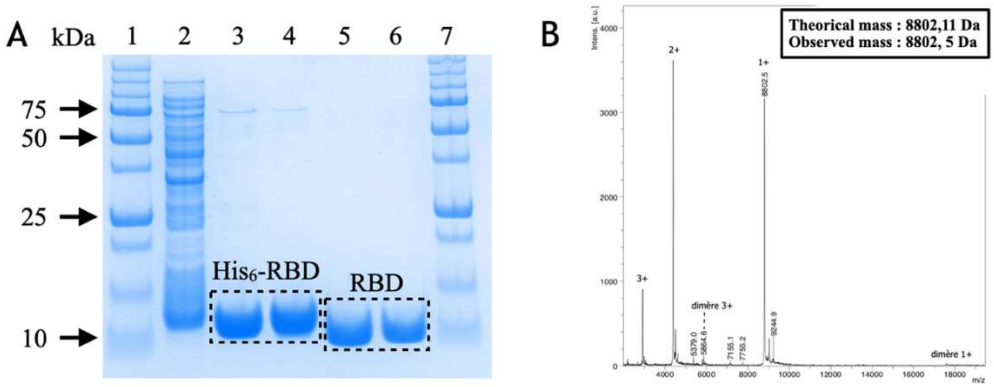


Figure 58 : Homogénéité de la fraction purifiée du RBD-B. *(A) Analyse par SDS-PAGE des différentes fractions de purification : Pistes 1 et 7 : marqueurs de poids moléculaires (Precision plus protein™ All blue Prestained Protein standards, BioRad); Piste 2 : extrait cellulaire brut; Piste 3 : fraction d'élution contenant la protéine après chromatographie d'affinité métallique; Piste 4 : fraction d'élution contenant la protéine après chromatographie échangeuse de cations POROS HS20; Piste 5 : fraction d'élution contenant la protéine après clivage du His_6-tag par la thrombine; Piste 6 : fraction de purification après la chromatographie d'exclusion-diffusion. Les conditions de migration et de révélation des protéines sont indiquées dans la section Matériel et Méthodes. (**B**) Analyse par spectrométrie de masse MALDI-TOF de la fraction S75.*

B.I.3. Décryptage des déterminants structuraux et/ou fonctionnels des ARN double-brin DM01 et DM03 associés à une haute affinité pour NS1

Les aptamères ARN DM03 et DM01 identifiés par SELEX (Marc *et al*., 2013) forment des complexes de haute affinité avec les protéines recombinantes NS1 (**Fig.38**). Comme déjà énoncé précédemment, ces deux ARNs contiennent respectivement 1 et 2 motifs double-brin GUAAC. Ce(s) motif(s) GUAAC est associé à une séquence virale correspondant soit à la région 5'-UTR retrouvée dans tous les ARNm viraux (le motif A dans DM01) soit à la région 3'-UTR de l'ARNm codant pour NS1 (le motif B dans DM03). Toutefois, nous avons privilégié de réduire la longueur de ces aptamères de 80 nucléotides (nt) pour envisager sereinement les études systématiques d'interaction avec l'ARN et celles par cristallographie des complexes NS1/ARN.

B.I.3.a. Rôle potentiel du domaine ED de NS1 dans la fixation aux aptamères DM01-midi et DM03-midi

Ma *et al*. (2016) ont montré la présence d'un second dsRBD dans le domaine ED de NS1 d'un IBV (B/Lee/1940). A ce jour et à preuve du contraire, le domaine ED de NS1 des IAV n'a pas été décrit comme participant directement à l'interaction avec l'ARN, y compris lorsque la protéine NS1 se trouve dans une conformation ouverte ou semi-ouverte (**Fig.28**) (Carrillo *et al*., 2014; Jerry *et al*., 2011). Les délétions localisées dans la région linker séparant le domaine RBD du domaine ED influencent le repliement de la protéine NS1, en particulier la conformation du domaine ED et sa proximité vis à vis du domaine RBD (Carrillo *et al*., 2014; Jerry *et al*., 2011). En solution, la protéine NS1 adopte majoritairement la conformation *ouverte* (**Fig.28B**) (Mitra *et al*., 2019) qui, en rapprochant le domaine ED du domaine RBD, pourrait permettre au domaine ED de participer à l'interaction avec l'ARN, et pourquoi pas à la reconnaissance des motifs viraux identifiés par SELEX. Cette hypothèse est d'autant plus plausible que l'expérience de SELEX a été réalisée sur un mélange de NS1 entières provenant d'un virus humain H5N1 (allèle A de NS1) et d'un virus aviaire H7N1 (allèle B de NS1).

Pour évaluer cette hypothèse, nous avons examiné par électrophorèse en gel retard, l'interaction de la forme entière de NS1 d'allèle B avec les aptamères DM01-midi et DM03-midi qui dérivent respectivement des aptamères DM01 et DM03. A cette fin, nous avons choisi de purifier la forme entière de NS1 fusionnée à une étiquette GST en position N-terminale (GST-NS1), en raison de la forte insolubilité de NS1 non-fusionnée. Sans cette étiquette, nous avons vérifié que la protéine NS1 entière est produite dans *E.coli* sous forme insoluble. De plus, l'élimination de l'étiquette GST de la protéine de fusion par clivage au facteur Xa provoque la précipitation irréversible de NS1. Afin d'avoir une référence avec les expériences utilisant le domaine RBD purifié et pour atténuer un effet potentiel de l'étiquette GST, nous avons comparé l'interaction de la GST-NS1 et de la GST-RBD avec les sondes DM01-midi et DM03-midi (**Fig.59**).

Le RBD sous sa forme fusion GST-RBD forme un complexe unique C1 avec la sonde DM03-midi et deux complexes C1 et C2 avec la sonde DM01-midi (triangles bleus, **Fig.59**). Nous avons conduit la même expérience avec la GST-NS1. Cette dernière est capable de former deux complexes C1' et C2' apparemment similaires avec les deux sondes (triangles verts, **Fig.59**) ce qui est significativement différent de l'expérience avec la GST-RBD. Plusieurs hypothèses peuvent alors être émises : (i) le domaine ED présent dans la forme entière de la protéine pourrait être à l'origine de la capacité de celle-ci à former le complexe C2'. Si cette interprétation est correcte, alors le complexe C2 observé uniquement avec la GST-RBD et la sonde DM01-midi n'est pas de même nature que C2' observé avec GST-NS1; (ii) les complexes C2 et C2' sont de même nature et ne feraient pas intervenir directement le domaine ED dans l'interaction. En effet, on peut noter que C2' est formé plus facilement avec DM01-midi qu'avec DM03-midi (effet sélectif pour une des deux sondes comme précédemment décrit avec le GST-RBD). En effet, pour 1 000 nmol.L^{-1} de GST-NS1, le complexe C1' obtenu avec DM01-midi est complètement titré dans le complexe C2' (piste 22, **Fig.59**) alors que dans les mêmes conditions, seulement 50 % de C1' est titré en C2' avec DM03-midi (piste 11, **Fig.59**); (iii) Etant donnée la propension de la GST à dimériser, elle pourrait favoriser la dimérisation de GST-NS1 (Fabrini *et al*., 2009). La présence du domaine ED pourrait favoriser la dimérisation *via* la GST. Dans le cadre de cette hypothèse, le complexe C2' observé avec la GST-NS1 pourrait correspondre à la fixation à l'ARN d'un dimère $(GST\text{-}NS1)_2$. Cette dernière hypothèse est cependant peu probable car il semble bien que, dans nos expériences, C2' se forme au détriment de C1'; (iv) Il est également possible que le domaine ED module la structure du RBD, lui permettant alors de former deux complexes avec DM03-midi. ***Même si ce résultat n'a pas été décrit dans le cas d'une protéine NS1 d'IAV mais pour une NS1 de virus influenza B (Ma et al., 2016), il semble que le domaine ED de NS1A puisse contribuer à la formation de complexes RBD/ARN selon un mécanisme qui reste cependant à élucider.***

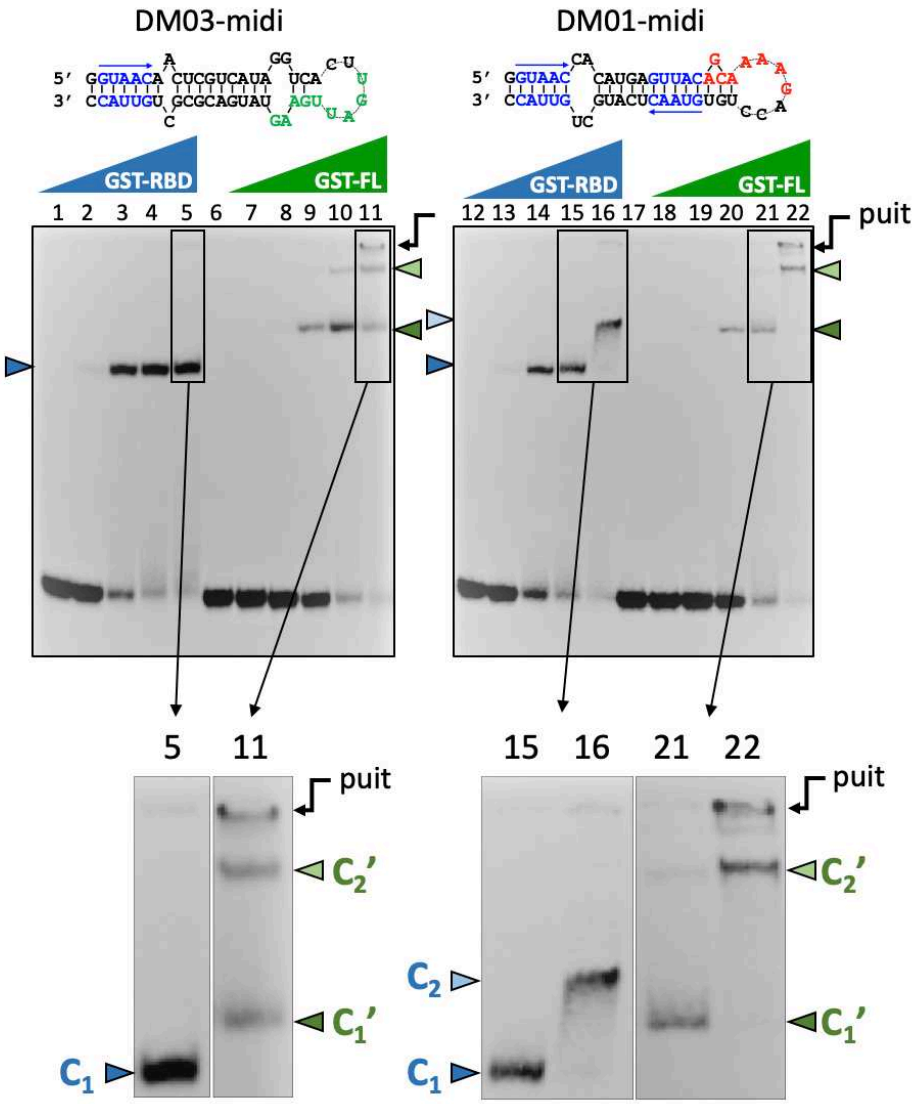


Figure 59 : Interaction des protéines de fusion RBD-B et NS1-B avec DM01-midi et DM03-midi. *0,1 nmol.L^{-1} d'aptamère DM01-midi (piste 17) et DM03-midi (piste 6) radiomarqué ont été incubées en présence de 0.1, 1, 10, 100, 1000 nmol.L^{-1} de GST-RBD (pistes 1-5 et 12-16) ou de GST-NS1 (pistes 7-11 et 18-22). Les mélanges réactionnels ont été analysés par électrophorèse en gel retard. Les autoradiographies des gels sont présentées.*

B.I.3.b. Longueur minimale de duplex d'ARN exigée par le RBD de NS1 pour former des complexes de haute affinité

Après avoir décrit une potentielle participation du domaine ED de NS1 dans l'interaction avec l'ARN, nous avons choisi de porter notre étude sur le domaine RBD de NS1 uniquement sans utiliser de protéines de fusion à la GST. Dans un premier temps, nous avons cherché à évaluer la taille minimale du duplex d'ARN permettant toujours d'observer une forte interaction avec le domaine RBD du virus aviaire H7N1 (allèle B de NS1, RBD-B). Pour cela, nous avons défini des sondes d'ARN tronquées à partir des extrémités 3' et 5' des structures en « épingle à cheveux » (*hairpin*) des aptamères DM01 et DM03 (**Fig.60**).

La capacité du RBD-B à reconnaitre avec une haute affinité les aptamères tronqués dérivant de DM01 et DM03 a été évaluée par électrophorèse en gel retard. Alors que des faibles concentrations de RBD de NS1 (GST-RBD) permettent de former des complexes de haute affinité avec les aptamères DM01-midi (deux complexes observés, **Fig.59**) et DM03-midi (un seul complexe observé, **Fig.59**), ce n'est pas le cas avec les courts aptamères tronqués DM01-sh et DM03-sh où aucun complexe stable n'est visible pour une concentration de protéines inférieure à 2,5 µmol.L^{-1}. Ces résultats indiquent que les troncations associées à ces sondes sont trop importantes pour permettre la fixation stable du RBD et ce quel que soit la présence des motifs A ou B et du motif GUAAC. D'ailleurs, les séquences virales A et B seules (DM01-MotA et DM03-MotB, **Fig.60**) ne permettent plus d'obtenir des complexes stables y compris en présence de concentrations en RBD supérieures à 2,5 µmol.L^{-1}. D'après ces résultats, une structure d'ARN bicaténaire d'une vingtaine de nt semble nécessaire pour permettre une interaction forte avec le RBD de NS1 (**Fig.60**). ***Outre la présence de séquences particulières, il semble donc que la longueur de l'ARN soit un élément essentiel pour observer des complexes de haute affinité avec la protéine.***

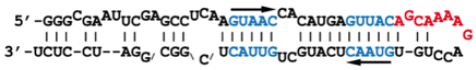
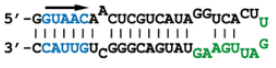
Séquence/structure 2D	Nom	Longueur (nt)	Interaction avec NS1
	DM01	81	++++
	DM01-midi	50	++++
	DM01-sh	30	+/---
	DM01-MotA	8	-
	DM03	81	++++
	DM03-midi	50	++++
	DM03-sh	32	+/---
	DM03-MotB	9	-

Figure 60 : Reconnaissance par NS1 des dérivés d'aptamères DM01 et DM03. *Les motifs GUAAC sont indiqués en bleu et les motifs viraux A et B en rouge et en vert, respectivement. Les expériences d'interaction ont été réalisées par gel retard avec le domaine RBD-B. (++++) indique une interaction forte, (+/---) une très faible interaction et (-) pas d'interaction mis en évidence par gel retard.*

B.I.3.c. Deux modes possibles de fixation du RBD de NS1 à l'ARN

Il nous est possible de titrer quantitativement les aptamères DM01-midi et DM03-midi par le RBD de NS1 sous la forme de complexes stables parfaitement visibles en gel retard (**Fig.61A et 61B**). Dans nos conditions expérimentales, la totalité des deux aptamères est sous forme de complexes avec seulement 20 nmol.L^{-1} de RBD ce qui doit correspondre à une forte affinité du RBD pour ces deux aptamères (**Fig.61A**). Contrairement à l'aptamère DM03-midi qui est titré sous la forme d'un seul complexe C1 (**Fig.61A**), l'aptamère DM01-midi forme au moins deux complexes discrets avec le RBD (C1 et C2) (résultats similaires avec les protéines de fusion GST-RBD et GST-NS1, **Fig.59;** Nous avons également vérifié que tous RBD purifiés au laboratoire étaient capable de former respectivement 1 et 2 complexes avec les sondes ARN DM03-midi et DM01-midi (**Annexe 6**)) ainsi qu'un complexe intermédiaire C1' à partir de 60 nmol.L^{-1} de RBD qui présente un retard de migration non-constant s'échelonnant entre celui de C1 et celui de C2 jusqu'à être confondu avec le complexe C2 en présence de plus hautes concentrations en protéine (**Fig.61A**). Il semble raisonnable de penser que le complexe C1 observé avec les deux aptamères correspond à un complexe RBD/ARN de nature similaire dans lequel une molécule de RBD est fixée à une molécule d'ARN. Entre 1 et 20 nmol.L^{-1} de RBD, il semble que les complexes C1 et C2 observés uniquement avec l'aptamère DM01-midi se forment *quasi* simultanément (avec un léger décalage dans le temps pour C2) avant même que l'ARN DM01-midi libre soit complètement titrée dans le complexe C1. Dans cette même gamme de concentrations, l'abondance relative de C2 par rapport à C1 augmente de façon *quasi*-exponentielle pour finalement atteindre un plateau entre 20 et 500 nmol.L^{-1} de RBD (**Fig.61A**). L'apparition *quasi*-exponentielle de C2 suggère qu'il se forme de façon coopérative probablement à partir de C1 et qu'il pourrait correspondre au titrage de C1 par une 2ème molécule de RBD. L'existence de cet extrémum pour l'abondance relative de C2 n'est cependant pas simple à expliquer d'autant plus qu'il correspond à la disparition de C1 (courbe en *cloche*, **Fig.61B**) et à l'apparition progressive de C1' au-delà de 20 nmol.L^{-1} de RBD.

Tout laisse à penser que C1' se forme à partir de C1. En l'état actuel de nos connaissances, la nature physique de C1' et C2 reste floue. Deux hypothèses pourraient être formulées : (i) l'aptamère DM01-midi et/ou le complexe C1 est hétérogène impliquant la formation de complexes RBD/ARN (rapport 2/1) de nature différente (C1' et C2) et (ii) le complexe C1' correspondrait en fait à une fraction instable du complexe C2 qui se dissocierait en cours d'électrophorèse. Dans le cadre de cette seconde hypothèse, la grande largeur de la bande de C1' pourrait signifier que cette dissociation est progressive au cours de l'électrophorèse et dépendrait de la concentration en RBD. Au-delà de 500 nmol.L^{-1} de RBD, la sonde est complètement titrée dans le seul complexe migrant à la même position que C2. Les hypothèses (i) et (ii) pourraient alors être vraies toutes les deux. Si notre interprétation est correcte, l'affinité du RBD pour le complexe C1 est probablement bien inférieure à celle du RBD pour l'ARN seul. Néanmoins, la fixation d'une première molécule de RBD à l'ARN favoriserait la fixation d'une seconde molécule selon un phénomène coopératif.

Son origine reste cependant indéterminée à l'heure actuelle. L'interprétation la plus simple de ces données serait que l'aptamère DM01-midi contiendrait dans sa structure et/ou sa séquence un déterminant fonctionnel permettant la possibilité de former le complexe C2, déterminant qui serait absent dans l'aptamère DM03-midi. ***Si l'on fait l'hypothèse que les complexes C1 observés avec les deux aptamères sont de nature identique, il est probable que la séquence virale contenant le motif A (retrouvée dans tous les ARNm viraux), présente uniquement dans l'aptamère DM01-midi, soit associée à la fixation d'une seconde molécule de RBD au complexe C1 formant ainsi le complexe C2. Ces dernières expériences semblent révéler deux modes possibles d'interaction de NS1 pour l'ARN.***

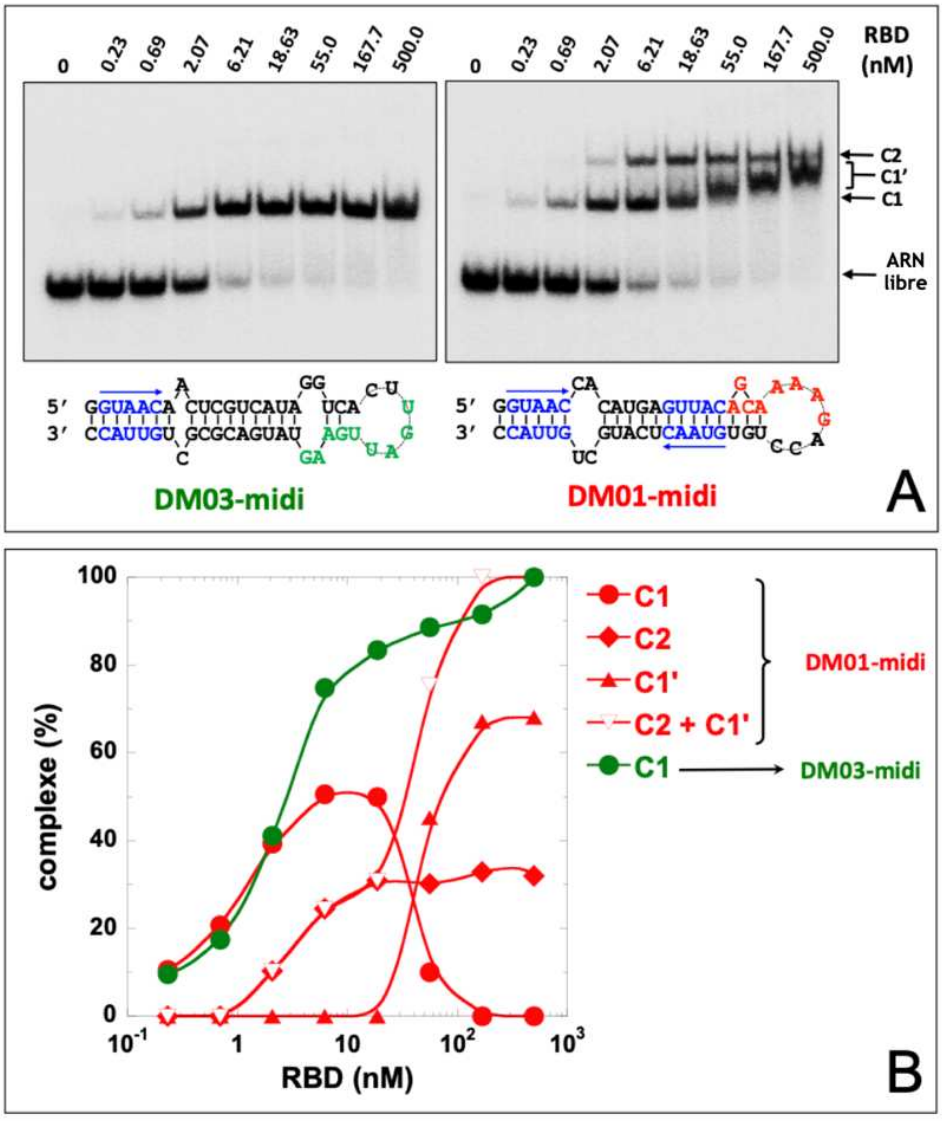


Figure 61 : Titrage de DM01-midi et DM03-midi par le RBD-B. *0,1 nmol.L^{-1} d'ARNs radiomarqués indiquées ont été incubées séparément en présence de concentrations croissantes de RBD (indiquées en nmol.L^{-1} au-dessus de chaque piste). Les mélanges d'incubation ont été ensuite analysés par électrophorèse en gel retard comme décrit dans la section Matériels et Méthodes. Les complexes stables sont indiqués par des flèches. (**A**) Autoradiographies des gels (**B**) Quantification des gels présentés en (**A**).*

B.I.3.d. Le motif viral AGCAAAAG, implication dans la formation de complexes

Il existe cependant dans les deux aptamères un site de haute affinité pour le RBD, probablement indépendant des séquences virales et associé à la formation d'un complexe de haute affinité C1. Ce motif commun aux deux aptamères pourrait être le motif GUAAC. Pour tenter d'apporter de nouveaux éléments soutenant cette interprétation, l'aptamère DM01-midi a été décliné en deux nouveaux ARN de même taille (DM01-Mod1 et DM01-Mod2, **Fig.62**). L'ARN DM01-Mod1 correspond à l'aptamère DM01-midi dans laquelle seul le motif A a été conservé tandis que les deux motifs double-brin GUAAC ont été remplacés. L'ARN DM01-Mod2 correspond à l'aptamère DM01-midi dans laquelle seul le motif viral A a été modifié.

L'interaction du RBD-B avec ces deux nouveaux ARN a été comparée à celle du RBD avec l'aptamère DM01-midi en gel retard dans la même gamme de concentrations en protéine (**Fig.62**). Des concentrations de 1 à 10 nmol.L^{-1} de RBD-B suffisent pour former des complexes stables de haute affinité (bande de complexe discrète) avec les aptamères DM01-midi et DM01-Mod2 alors qu'aucun complexe discret n'apparaît dans cette même gamme de concentrations pour l'ARN DM01-Mod1. On remarque aussi que l'aptamère DM01-midi forme deux complexes stables (C1 et C2, **Fig.62**) avec le RBD de NS1 dès 1 nmol.L^{-1} en protéine. Le complexe C1 disparaît pour laisser place à un complexe intermédiaire (C1') à partir de 100 nmol.L^{-1} de RBD et qui semble se stabiliser en complexe C2 en présence de 500 nmol.L^{-1} de RBD (**Fig.62**). L'ARN DM01-Mod2 forme un unique complexe avec 1 nmol.L^{-1} de RBD, puis un complexe C2-like moins stable en présence de concentrations plus élevées en RBD (**Fig.62**). Contrairement aux deux autres ARN, DM01-Mod1 forme un seul complexe avec le RBD de NS1 de type C2-like pour des concentrations en RBD supérieures à 100 nmol.L^{-1}. A une concentration de 500 nmol.L^{-1} de RBD, les ARN sont titrés quantitativement par la protéine sauf l'ARN DM01-Mod1 qui subsiste encore partiellement sous forme libre. Il semble que le complexe C2 soit formé seulement lorsque la séquence virale AGCAAAAG est présente dans l'ARN (**Fig.62**). Il semble aussi que la présence des motifs double-brin GUAAC dans l'aptamère DM01-midi conditionne au moins la formation du complexe C1 de haute affinité, puisque son dérivé DM01-Mod1 qui ne possède plus ces deux motifs de haute affinité dans sa structure ne forme plus de complexe C1 stable et des concentrations beaucoup plus élevées en RBD sont nécessaires pour former le complexe C2 avec cet ARN. ***Il semble que les motifs double-brin GUAAC contenus dans l'aptamère DM01-midi et son dérivé DM01-Mod2 soient responsables de la formation du complexe C1 avec le RBD de NS1. A l'inverse, la séquence virale AGCAAAAG semble à elle-seule supporter la formation du complexe C2 lorsque l'ARN qui la contient (DM01-midi) est incubé avec de faibles concentrations en protéine. En l'état actuelle de nos investigations, il reste difficile de conclure quant au rôle précis joué par la structure de la tige-boucle dans la fixation d'une deuxième molécule de RBD à l'aptamère DM01-midi. Cependant, la présence d'au moins un motif GUAAC en position distale par rapport à la boucle est clairement associée à la formation d'un complexe de haute affinité entre l'ARN et une molécule de RBD.***

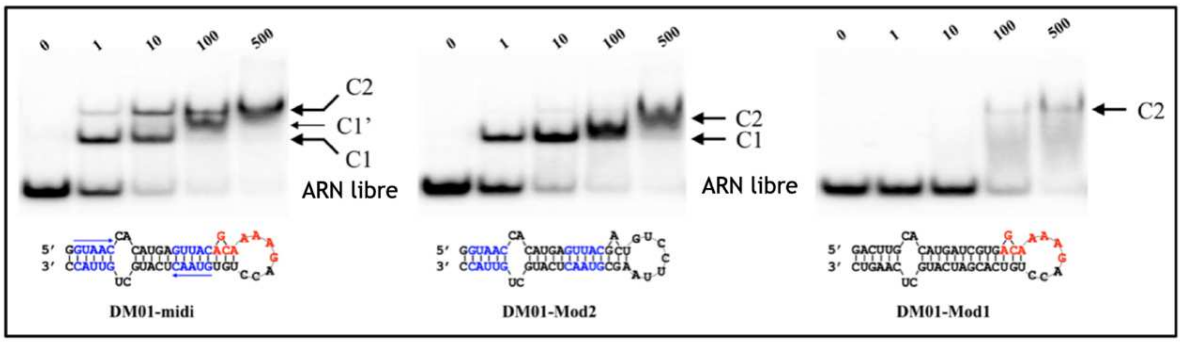


Figure 62 : Interaction du RBD-B avec DM01-midi, DM01-Mod1 et DM01-Mod2. *Les conditions expérimentales utilisées sont les mêmes que celles indiquées dans la figure 66. Les complexes stables sont indiqués par des flèches.*

B.I.4. Conception d'une sonde ARN de haute affinité pour le RBD de NS1 utilisable dans une approche structurale par cristallographie

B.I.4.a. Le motif GUAAC identifiée par SELEX est une séquence de haute affinité pour le RBD de NS1

Nos derniers résultats indiquent que la formation de complexe RBD/ARN de haute affinité (i.e. de type C1) nécessite un ARNdb de longueur minimale (une vingtaine de paires de bases) ainsi que la présence d'au moins un motif GUAAC. Toutefois, le rôle joué par cette séquence présente en un et deux exemplaires dans les aptamères DM03-midi et DM01-midi respectivement reste à évaluer. C'est ce que nous avons essayé de faire en évaluant l'affinité d'interaction de sondes ARN double-brin synthétiques dérivant en partie des aptamères DM01-midi et DM03-midi. Pour étudier plus précisément le rôle de la séquence GUAAC, les séquences virales contenues dans la boucle de la structure en tige-boucle des aptamères DM01-midi et DM03-midi et celle de la pentaboucle CAUCA de l'ARN sh16 (Chien *et al*., 2004; Yin *et al*., 2007; Marc *et al*., 2013) ont été respectivement éliminées et remplacées par la tétraboucle GAGA. Cette dernière, issue de la famille GNRA, est très fréquemment retrouvée dans les structures en tige-boucle des ARN procaryotes et eucaryotes et est connue pour être très structurée et pour stabiliser fortement les structures d'ARN en tige-boucle (Antao *et al*., 1991; Jucker *et al*., 1996). Cet artifice nous permet d'éviter une interférence de la nature de la boucle de la tige-boucle de nos sondes ARN.

Les quatre aptamères synthétiques shDM02-GAGA, shDM03-GAGA, shDM05-GAGA et shDM06-GAGA ont été conçus pour essayer de décrypter plus précisément le rôle joué par le motif GUAAC dans l'interaction avec le RBD. A l'exception de la sonde shDM05-GAGA qui n'est pas reconnue dans la gamme de concentrations de RBD utilisée, toutes les autres sondes sont bien reconnues (**Fig.63**). A 1 nmol.L^{-1} de RBD, les sondes shDM02-GAGA et shDM03-GAGA, qui contiennent deux motifs GUAAC dans leur séquence, sont titrées quantitativement par le RBD tandis que l'ARN

shDM06-GAGA est encore très faiblement titrées à cette concentration en RBD. Un minimum de 5 à 10 nmol.L^{-1} de RBD est nécessaire pour titrer quantitativement la sonde shDM06-GAGA (**Fig.63**). La présence de deux motifs GUAAC dans l'aptamère séparés d'une structure bicaténaire continue de 7 nt (shDM02-GAGA) ou interrompue par une bulle de 2 nt mésappariés (shDM03-GAGA) semble donc être un déterminant fort de l'interaction avec le RBD. Néanmoins, la présence de la bulle de 2 nt présente dans shDM03-GAGA (qui est aussi retrouvée dans DM01-midi) ne semble pas être particulièrement favorable à l'interaction. La présence d'une seule séquence GUAAC en position distale par rapport à la boucle GAGA dans shDM06-GAGA n'impacte pas significativement l'interaction avec le RBD. Par contre, la présence d'une seule séquence GUAAC en position proximale par rapport à la boucle GAGA dans la sonde shDM05-GAGA est très défavorable à une interaction avec le RBD de NS1 (**Fig.63**). ***Le RBD-B semble reconnaitre avec une haute affinité le motif GUAAC. Toutefois, même si la présence de deux motifs GUAAC semble plus favorable à l'interaction, une seule version de ce motif en position distale par rapport à la boucle des sondes en tige-boucle est suffisante pour observer une interaction forte. Le complexe C1 observé entre le RBD et DM01-midi (2 motifs GUAAC) ou DM03-midi (1 motif GUAAC en position distal par rapport à la boucle de la tige-boucle) doit probablement correspondre à la reconnaissance de cette séquence par le RBD.***

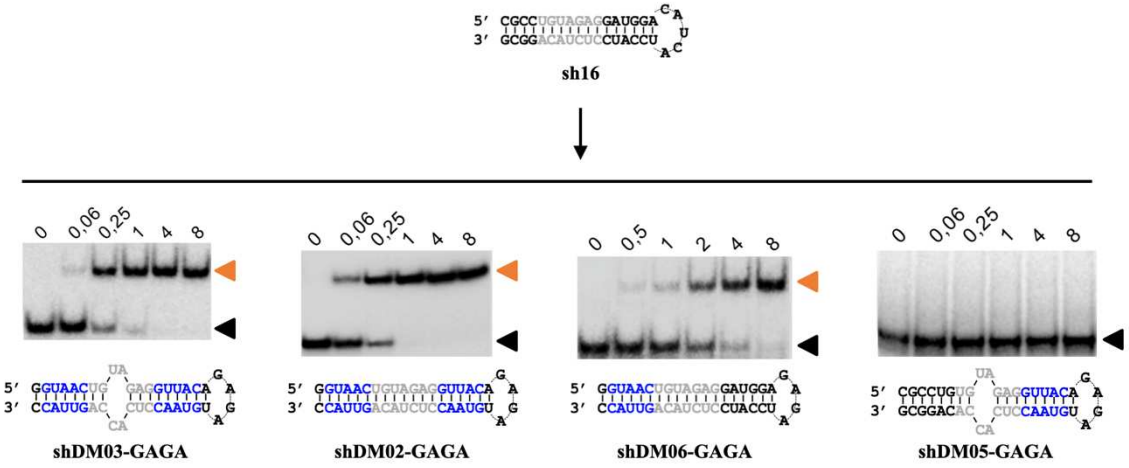


Figure 63 : Interaction du RBD-B avec des aptamères présentant un ou deux sites GUAAC. *Comme indiqué, les sondes shDM02-GAGA, shDM03-GAGA, shDM05-GAGA et shDM06-GAGA ont été incubées en présence de concentrations croissantes de RBD (les nombres indiqués au-dessus de chaque piste indiquent les concentrations utilisées en nmol.L^{-1}) et les mélanges réactionnels résultants analysés par électrophorèse en gel retard. Les conditions expérimentales sont détaillées dans le Matériels et Méthodes. Une autoradiographie des gels représentative de chaque titrage est montrée. L'ARN libre est indiqué par un triangle orange, les complexes formés par un triangle noire.*

B.I.4.b. Conception de l'ARNdb AWFC01, une sonde modèle de haute affinité pour étudier les interactions RBD/ARN

Pour aller plus loin dans la caractérisation de l'interaction de NS1 avec l'ARNdb, nous avons conçu une nouvelle sonde nommée AWFC01 tentant de rassembler les critères pour une « bonne

interaction » établis par nos expériences précédentes et tenant compte aussi de celles de la littérature. AWFC01 est un duplex ARN de 19 paires de bases sans tige-boucle (c'est-à-dire ouvert des deux côtés de l'ARN) contenant deux motifs double-brin GUAAC séparés de 7 paires de bases. Pour des raisons de commodité et de coût, nous avons choisi un duplex palindromique sur toute sa longueur interrompue uniquement en son centre par la paire de base *Hoogsteen* [U.U] (**Fig.64**). Notre choix semble pertinent d'un point de vue littérature puisque Cheng *et al*. (2009) avait aussi choisi un duplex palindromique pour résoudre la première structure cristallographique du complexe RBD/ARN (**Fig.64**). Afin de compléter au mieux notre étude, nous avons également conçu l'ARNdb AWFC02, un second duplex d'ARN sans tige-boucle contenant un seul motif GUAAC dans sa séquence (**Fig.64**).

Nous avons vérifié que AWFC01 et AWFC02 sont bien reconnus par le RBD-B comme des sondes de haute affinité (**Fig.64**). Sans surprise, une concentration de 10 $nmol.L^{-1}$ en RBD-B suffit pour titrer quantitativement les ARNdb AWFC01 et AWFC02 en un complexe stable de type C1 (**Fig.64A et 64B**). Comme attendu car déjà décrit à nombreuses reprises dans la littérature comme étant très affecté dans son interaction avec l'ARNdb (Wang *et al*., 1999), le RBD-B mutant (R38A-K41A, aaRBD-B) ne forme aucun complexe visible dans ces conditions (pistes 3, **Fig.64B**). Dans les mêmes conditions, l'ARNdb CHENG n'est reconnu ni par le RBD-B ni par le aaRBD-B, démontrant ainsi l'affinité médiocre du RBD pour cette sonde et validant le choix de nos sondes ARN. Avec les sondes palindromiques, on peut noter qu'elles apparaissent hétérogènes (deux espèces), les espèces migrant le plus vite pendant l'électrophorèse (AWFC01/2 et CHENG/2) correspondent probablement à un ARNsb replié en tige-boucle résultant d'une hybridation monomoléculaire de l'ARN (auto-appariement résultant du caractère palindromique). Dans les conditions expérimentales choisies, ces courtes tige-boucle d'ARN ne sont pas reconnus par le RBD. Une observation très similaire avait été faite en raccourcissant les aptamères DM01-midi et DM03-midi (**Fig.60**) qui indique le caractère déterminant de la longueur du duplex d'ARN dans l'interaction avec le RBD. Le fait de ne pas observer de complexe avec 10 $nmol.L^{-1}$ de RBD ne signifie pas pour autant que le RBD est complètement incapable de se lier à des ARN très courts, mais indique justement que son affinité pour ces ligands est beaucoup plus faible que celle pour les ARNdb AWFC01 et AWFC02. A de très hautes concentrations en RBD-B, il est possible de titrer complètement l'ARNdb CHENG dans de multiples complexes instables qui apparaissent diffus dans toute la piste, traduisant directement d'une instabilité de ces complexes formés au cours de l'électrophorèse.

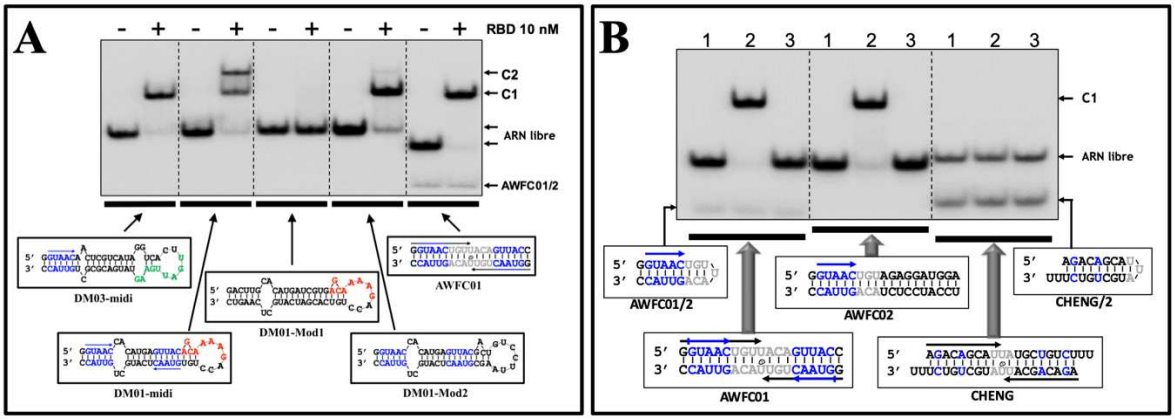


Figure 64 : Interaction du RBD-B avec DM01-midi, DM01-Mod1 et DM01-Mod2. *(**A**) Comparaison avec DM01-midi et DM03-midi. Dans les mêmes conditions expérimentales, l'interaction du RBD de type sauvage a été évaluée par électrophorèse en gel retard. Les complexes C1 et C2 ont été définis plus haut dans le texte. (**B**) Comparaison avec CHENG. 0,1 nmol.L^{-1} des sondes ARN radiomarquées indiquées (pistes 1) ont été incubées séparément avec 10 nmol.L^{-1} de RBD de type sauvage et mutant R38A-K41A (pistes 2 et 3, respectivement). Les mélanges réactionnels ont été ensuite analysés par électrophorèse en gel retard. Comme indiqué, les espèces AWFC01/2 et CHENG/2 de type tige-boucle résultent d'une hybridation monomoléculaire d'un seul brin de AWFC01 et CHENG.*

B.I.4.c. AWFC01, un ligand de haute affinité pour le RBD de nombreuses NS1

Afin de préciser un peu mieux ce que l'on qualifie de « haute affinité » et de vérifier que le modèle AWFC01 est un ligand de haute affinité pour les protéines NS1 de plusieurs sous-types viraux, nous avons déterminé les constantes de dissociation (Kd) des complexes RBD/AWFC01 pour toutes les protéines purifiées. En bref, l'ARNdb AWFC01 radiomarqué a été incubé avec une gamme de concentrations croissantes en RBD (provenant du virus aviaire H7N1 (sauvage et son mutant R38A-K41A, indiqués RBD-B et aaRBD-B, respectivement), des virus humains H5N1 (sauvage et son mutant R38A-K41A, indiqués RBD-A et aaRBD-A, respectivement), H3N2, pdm09 (SOIV) et H7N9 et du virus de chauve-souris H17N10) de telle sorte que l'on puisse atteindre des plateaux correspondant à 0 % (ARN libre) et à 100 % de complexes formés. Trois expériences indépendantes ont été réalisées pour chaque protéine RBD afin de déterminer les Kd.

Nos résultats d'interaction en gel retard indiquent que tous les RBD sauvages évalués reconnaissent très bien AWFC01, avec des Kd compris entre 1 et 4 nmol.L^{-1} (**Fig.65C**). Même si ces Kd sont très proches les uns des autres, il semble que les RBD-A et RBD-B présentent une meilleure affinité pour AWFC01, ceci pouvant être expliqué par le fait que Marc *et al*. (2013) avaient utilisé la forme entière des protéines NS1 de ces mêmes virus (A/HK/213/03/H5N1); A/Turkey/Italy/977/1999/H7N1) pour la sélection des aptamères. Au-delà de 100 nmol.L^{-1} de RBD, il est possible d'observer plusieurs complexes discrets (au moins trois) entre AWFC01 et le aaRBD-A de NS1 (**Fig.65A**) dont le plus abondant est plus retardé que le complexe RBD-A/AWFC01. ^{Ces} résultats indiquent que, même très affecté dans sa capacité de liaison à l'ARN, ce mutant garde une affinité modeste (environ 400 nmol.L^{-1}) pour celui-ci (**Fig.65C**). Cette affinité n'est pas évaluable par titrage en gel retard pour le aaRBD-B car les complexes formés au-delà du micromolaire ne sont pas assez stables

pour les quantifier avec précision (**Fig.65B**). Contrairement au aaRBD-A qui est modérément affecté dans son interaction avec AWFC01, le aaRBD-B est très invalidé par la double-substitution R38A-K41A (**Fig.65A et 65B**). La constante d'affinité du aaRBD-B pour AWFC01 a été estimée à plus de 50 µmol.L^{-1}. ***AWFC01 forme un complexe stable et de haute affinité (Kd de 1 à 4 nmol.L^{-1}) avec toutes les protéines RBD purifiées au laboratoire y compris lorsque certaines protéines présentent la double-substitution R38A-K41A (aaRBD-A) qui est connue pour invalider le RBD de NS1 et abolir l'interaction avec l'ARN double-brin (Hale et al., 2008). Alors que cette double-substitution est très délétère pour le aaRBD-B, elle l'est beaucoup moins pour le aaRBD-A puisque ce dernier présente un Kd d'environ 400 nmol.L^{-1}, ce qui correspond encore à une bonne affinité. Dans le cas présent, l'ARNdb AWFC01 est titré par le aa-RBD-A sous la forme de plusieurs complexes. En l'état d'avancement de nos investigations, la nature de ces différents complexes reste à élucider.***

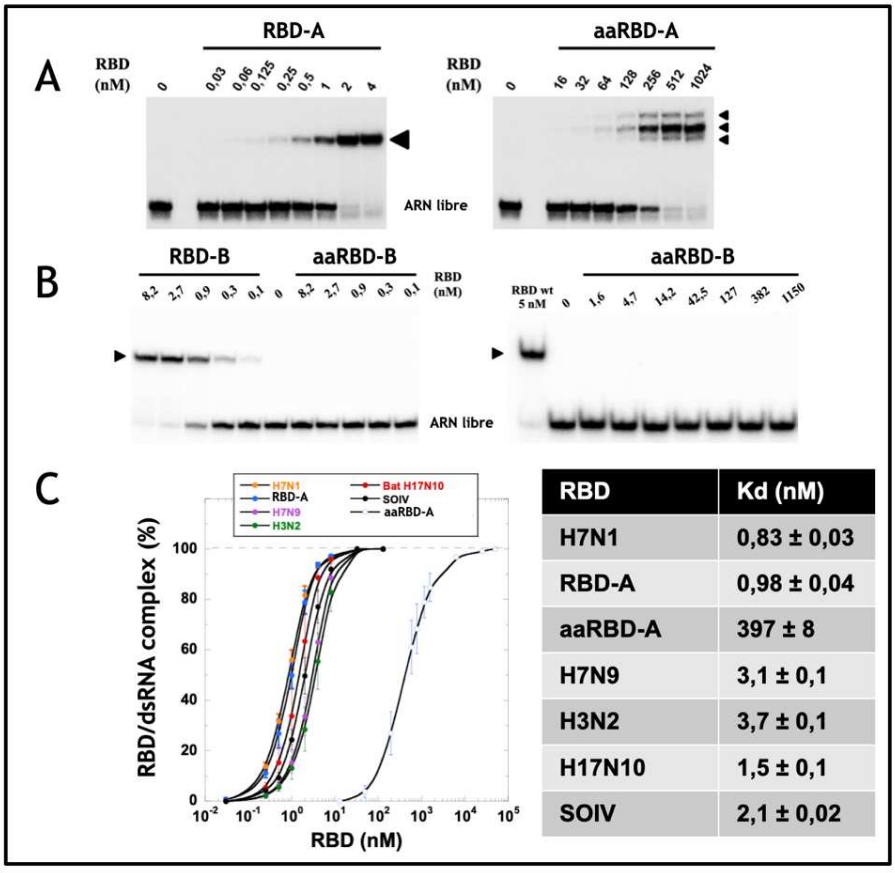
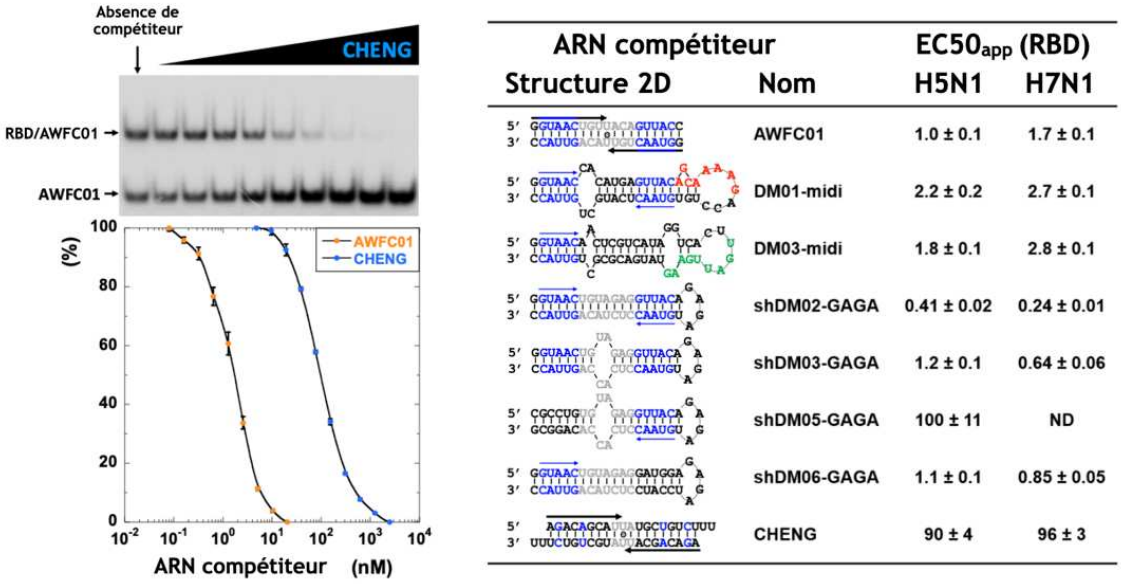


Figure 65 : Détermination des Kd des différents RBD purifiés pour le duplex AWFC01. *0,1 nmol.L^{-1} de AWFC01 est incubé en présence de concentrations croissantes des différents RBD indiqués. A l'équilibre, les mélanges réactionnels sont analysés par électrophorèse en gel retard. (**A**) Autoradiographies représentatives du titrage de AWFC01 par le RBD-A et le aaRBD-A. (**B**) Autoradiographies représentatives du titrage de AWFC01 par le RBD-B et le aaRBD-B. (**C**) Quantification des autoradiographies. Le pourcentage de complexe formé est analysé en fonction de la concentration en RBD. Dans le cas du aaRBD-A, le pourcentage de complexe correspond à la somme des différents complexes observés en (**A**). Les données représentent la moyenne obtenue pour 3 expériences indépendantes. A partir des courbes de titrage, la constante de dissociation Kd pour chaque couple RBD/AWFC01 correspond à la concentration de RBD permettant la formation de 50 % de complexe (voir Matériels et Méthodes »).*

B.I.4.d. Détermination des EC_{50app} des sondes par compétition contre AWFC01

Nous avons montré que l'ARNdb AWFC01 est très affin pour l'ensemble des RBD testés alors que leurs séquences peptidiques varient sensiblement d'une protéine à l'autre (**Fig.57**). Il n'est cependant pas possible d'évaluer le Kd pour toutes les ARN testés dans cette étude par des expériences de titrage en gel retard en raison des multiples complexes formés ou des bandes des complexes diffuses pour certains de ces ARN. Pour faire néanmoins une évaluation des différentes sondes d'ARN avec notre modèle de référence AWFC01, nous avons comparé par électrophorèse en gel retard la capacité de celles-ci à déplacer le complexe stable RBD/AWFC01 pour les RBD-A et RBD-B. Ces expériences en gel retard nous permettent de déterminer l'EC_{50} apparente (EC_{50app}) pour chaque sonde que nous pouvons comparer avec l'EC_{50app} de AWFC01 contre elle-même (**Fig.66**). Pour ce faire, 0,1 nmol.L^{-1} de sonde AWFC01 radiomarqué est incubé avec 1 nmol.L^{-1} de RBD-A ou de RBD-B. A l'équilibre, des concentrations croissantes de sondes compétitrices non radiomarquées sont ajoutées au mélange qui est analysé par électrophorèse en gel retard après un temps d'incubation supplémentaire. Dans ces conditions expérimentales où on se place au Kd (valeurs de Kd des complexes RBD/AWFC01 indiquées dans la **Fig.65C**) c'est-à-dire à 50 % de complexe RBD/AWFC01, l'EC_{50app} de l'ARN compétiteur correspond à la concentration en compétiteur nécessaire pour déplacer la moitié (soit 25 %) des complexes initialement formés.

Hormis l'aptamère shDM05-GAGA et l'ARNdb CHENG, toutes les sondes évaluées sont de très bons compétiteurs de AWFC01 et ceux pour les deux protéines testées. L'ARN shDM02-GAGA semble être le meilleur compétiteur avec des EC_{50} de 0,24 et 0,41 nmol.L^{-1} pour les RBD-B et RBD-A respectivement (**Fig.66**). La présence d'une structure tige de type AWFC01 avec une boucle GAGA (shDM02-GAGA) semble augmenter l'affinité respective des RBD-A et RBD-B d'un facteur 2 et 7 (**Fig.66**). Comme attendu par les expériences précédentes, shDM05-GAGA est un compétiteur plus modeste indiquant sa faible affinité pour les RBD-A et RBD-B. Il est intéressant de noter que l'EC_{50app} de CHENG (comprise entre 90 et 95 nmol.L^{-1} en fonction de la protéine utilisée, **Fig.66**) est comparable à celle de shDM05-GAGA ce qui démontre que l'ARNdb utilisé par Cheng *et al*. (2009) est un ligand médiocre pour le RBD. Le RBD est environ 60 à 100 fois moins affin pour la sonde CHENG que pour la sonde AWFC01. ***Notre ARNdb AWFC01 est un très bon ligand pour le RBD de NS1. Comme nous l'avions montré précédemment, la présence minimale d'un motif GUAAC, et surtout son accessibilité (shDM06-GAGA), conditionne la formation de complexe de haute affinité. L'analyse des EC_{50app} démontre que le RBD de NS1 présente une bien meilleure affinité pour la sonde AWFC01 que pour la sonde utilisée par Cheng et al. (2009) pour résoudre la structure cristallographique du complexe RBD/ARNdb. Avec une telle différence d'affinité, il devient intéressant d'initier une étude cristallographique sur le complexe RBD-B/AWFC01.***



ARN compétiteur		$EC50_{app}$ (RBD)	
Structure 2D	Nom	H5N1	H7N1
	AWFC01	1.0 ± 0.1	1.7 ± 0.1
	DM01-midi	2.2 ± 0.2	2.7 ± 0.1
	DM03-midi	1.8 ± 0.1	2.8 ± 0.1
	shDM02-GAGA	0.41 ± 0.02	0.24 ± 0.01
	shDM03-GAGA	1.2 ± 0.1	0.64 ± 0.06
	shDM05-GAGA	100 ± 11	ND
	shDM06-GAGA	1.1 ± 0.1	0.85 ± 0.05
	CHENG	90 ± 4	96 ± 3

Figure 66 : Détermination par compétition des EC_{50} des différents aptamères et ARNdb pour la dissociation du complexe RBD/AWFC01. *0,1 nmol.L^{-1} de duplex AWFC01 radiomarqué est pré-incubé avec 1 nmol.L^{-1} de RBD (H5N1 ou H7N1). A l'équilibre (15 min), des concentrations croissantes en ARN compétiteurs sont ajoutées au mélange. Après 20 minutes d'incubation additionnelle, les mélanges résultants sont analysés par électrophorèse en gel retard. Les EC_{50app}, qui correspondent à la concentration en ARN compétiteur pour laquelle 50 % du complexe RBD/AWFC01 est déplacé vers la forme libre de l'ARN, sont extraits des courbes de compétition obtenues à partir de 3 expériences indépendantes.*

B.II. Bases structurales de l'interaction de NS1 avec l'ARN

Toutes les structures tridimensionnelles de RBD et de NS1 entières (non-complexées à de l'ARN) disponibles dans Protein Data Bank (PDB) proviennent d'IAV et d'IBV (**Tableaux 2 et 3**). Parmi les IAV, seules des structures de RBD (et de ED) de NS1 d'allèle A ont été résolues alors qu'aucune structure d'allèle B ne l'a été. Concernant la structure de complexes associant le RBD de NS1 à un duplex d'ARN, seules deux structures tridimensionnelles ont été résolues. La première, qui a été résolue par Yin *et al*. (2007) en utilisant la RMN, associait le RBD de NS1 du virus Udorn (allèle A de NS1) à un ARNdb de 16 paires de bases dérivant de longs ARN transcrits à partir du plasmide pGEM1 (Chien *et al*., 2004; Qian *et al*., 1995; Qiu *et al*., 1995). La seconde structure, qui a été résolue par Cheng *et al*. (2009) en utilisant la DRX, associait le RBD de NS1 du virus WSN (A/WSN/1933/H1N1, allèle A de NS1) à un ARNdb de 21 bases (dont 19 paires sont appariées) dérivés d'ARN interférents issus de la plante (Bucher *et al*., 2004; Delgadillo *et al*., 2004). Nos résultats précédents qui démontrent que le RBD de NS1 (qu'ils soient d'allèles A, B ou C) est moins affin pour l'ARNdb CHENG que pour AWFC01 (**Fig.66**) valident l'utilisation de AWFC01 pour la co-cristallisation. De plus, nous avons montré que AWFC01 se liait encore fortement au aaRBD-A avec un Kd de 400 nmol.L^{-1}. Les différences d'affinité existantes entre les ARNdb CHENG et AWFC01 peuvent suggérer un mode de reconnaissance alternatif de l'ARN qui se baserait sur une reconnaissance spécifique des nucléobases, en particulier celles du motif GUAAC. L'approche cristallographique du complexe RBD/AWFC01 qui va suivre, et

en particulier la résolution des structures tridimensionnelles (3D) et leurs affinements, n'aurait pu voir le jour sans une contribution essentielle de Franck COSTE. Même si je me suis initié au traitement des données de diffraction pour résoudre les structures tridimensionnelles (3D), le temps limité de la thèse associé à son partage entre mes deux laboratoires d'accueil entre Tours et Orléans ne m'a pas permis d'acquérir une autonomie dans la résolution des structures 3D par cristallographie.

B.II.1. Cristallisation et résolution des structures 3D des protéines et des complexes RBD/ARNdb

Plusieurs structures 3D ont été résolues au laboratoire, apportant de nouvelles informations structurales sur le RBD de NS1 et sur sa capacité à lier des ARNdb. Les informations concernant la collecte et l'affinement statistique des données de diffraction a été regroupé dans le **tableau 7**. Les interactions directes et indirectes médiées par des molécules d'eau ont été regroupées dans le **tableau 8**.

Tableau 7 : Collecte et affinement statistiques des données de diffraction obtenues pour la structure RBD-B libre et pour les structures des complexes RBD-B/AWFC01, aaRBD-B/AWFC01 et RBD-B/ZKO*

Collecte statistique des données de diffraction				
	RBD-B	**RBD-B/AWFC01**	**aaRBD-B/AWFC01**	**RBD-B/ZKO***
Source de radiation	SOLEIL PROXIMA 2	SOLEIL PROXIMA 2	SOLEIL PROXIMA 2	SOLEIL PROXIMA 2
Longueur d'onde (Å)	0.98009	0.98009	0.98009	0.98010
Groupe d'espace	$P3_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P4_12_12$
Paramètres de maille: *a, b, c* (Å)	83.83, 83.83, 68.53	49.78, 60.40, 87.40	40.69, 102.04, 102.70	103.60, 103.60, 102.21
Résolution (Å)	49.83-1.93 (1.97-1.93)	43.26-1.75 (1.78-1.75)	45.87-1.90 (1.94-1.90)	46.33-2.30 (2.38-2.30)
Observations totales	239136 (9353)	177791 (10027)	179050 (11970)	336777 (34004)
Reflexions uniques	21170 (1030)	27291 (1481)	34557 (2202)	25358 (2431)
Completeness (%)	100.0 (100.0)	99.9 (100.0)	99.9 (100.0)	100.0 (100.0)
Multiplicity	11.3 (9.1)	6.5 (6.8)	5.2 (5.4)	13.3 (14.0)
R_{merge} [a] (%)	7.8 (109.0)	4.9 (141.0)	4.3 (106.5)	8.6 (225.3)
Average $I/\sigma(I)$	15.9 (1.6)	16.7 (1.1)	16.3 (1.3)	18.8 (1.1)
$CC_{1/2}$ (%)	99.9 (70.4)	99.9 (57.1)	99.9 (48.1)	99.9 (55.3)
Affinement statistique des données				
Résolution (Å)	49.83-1.93	43.26-1.75	37.80-1.90	46.33-2.30
Nombre de réflexions utilisées	21170	27229	34486	25302
R_{work} [b] / R_{free} [c] (%)	17.6/19.3	17.3/20.4	16.8/19.2	19.1/22.1
Moyenne B values (Å^2)				
Total d'atomes	54.36	42.07	50.05	67.29
Atomes : Protéine, chaine A	51.50	39.10	39.74	63.15
Atomes : Protéine, chaine B	56.67	39.54	43.45	61.32
Atomes : ARN, chaine C	-	43.28	60.93	73.73
Atomes : ARN, chaine D	-	43.77	63.08	-
Atomes d'Ethane-1,2-diol	86.20	-	56.67	-
Atomes de Di(hydroxyethyl)ether	66.33	-	63.74	-
Atomes Sulfate	-	-	-	137.81
Atomes de Nitrate	-	-	-	94.87
Atomes d'eau	51.16	45.56	49.77	58.57
Root mean square deviation from ideality				
Bond lengths (Å)	0.020	0.005	0.005	0.004
Bond angles (°)	1.500	0.818	0.778	0.667
Analyse de Ramachandran (% de résidus)				
Favoured regions / Allowed regions / Outliers	99.3/0.7/0.0	100.0/0.0/0.0	100.0/0.0/0.0	99.3/0.7/0.0
Nombre d'atomes				
Protéine, chaine A	591	561	579	595
Protéine, chaine B	587	572	569	570
ARN, chaine C	-	388	399	399
ARN, chaine D	-	391	399	399
Ethane-1,2-diol	12	-	36	-
Di(hydroxyethyl)ether	14	-	67	-
Sulfate	-	-	-	10
Nitrate	-	-	-	16
Eau	66	156	158	56
Code PDB	6SW8	6SX0	6SX2	6ZLC

Tableau 8 : Interactions directes et indirectes établis par le RBD de NS1 avec les ARNdb AWFC01 et CHENG

Complexe RBD/CHENG : ARN palindromique (résolution 1.7 Å)		
RBD	**ARNdb**	**Interaction**
D2	Ur^{17} ribose, $Ur^{17'}$ ribose et phosphate	Indirecte
T5	Gu^{16} ribose, Ur^{17} ribose	Indirecte
D29	Gu^{16} ribose, Ur^{17} ribose	Indirecte
D29'	$Ur^{15'}$ ribose, $Gu^{16'}$ ribose	Indirecte
A30	Gu^{16} ribose	Indirecte
P31	Gu^{6} ribose / Ur^{15} nucléobase	Directe / indirecte
D34	Gu^{6} nucléobase, Cy^{7} ribose, Ur^{15} nucléobase	Indirecte
R35	Gu^{6} ribose / Ad^{5} nucléobase, Gu^{6} ribose	Directe / indirecte
R38	Cy^{7} phosphate / Cy^{7} phosphate	Directe / indirecte
R38'	Cy^{7} phosphate / Ad^{8} phosphate, $Gu^{6'}$ phosphate	Directe / indirecte
D39	Cy^{7} ribose et phosphate, $Gu^{6'}$ phosphate	Indirecte
K41	Ad^{8} phosphate	Indirecte
S42	Ad^{5} phosphate	Directe
T49	Cy^{4} ribose / Ad^{3} ribose, Cy^{4} ribose	Directe / indirecte
Complexe RBD-B/AWFC01 : ARN palindromique (résolution 1.75 Å)		
RBD	**ARNdb**	**Interaction**
D2	Ad^{17} ribose	Indirecte
T5	Ur^{16} ribose, Ad^{17} ribose	Indirecte
D29	Ur^{16} ribose, Ad^{17} ribose	Indirecte
A30	Ur^{16} ribose	Indirecte
P31	Cy^{6} ribose / Ur^{15} nucléobase	Directe / indirecte
D34	Cy^{6} nucléobase, Ur^{7} ribose, Gu^{14} nucléobase, Ur^{15} nucléobase	Indirecte
R35	Cy^{6} ribose / Ad^{5} ribose	Directe / indirecte
R37	$Ur^{7'}$ ribose	Directe et indirecte
R38	Cy^{6} phosphate, Ur^{7} phosphate, $Ur^{7'}$ phosphate	Directe et indirecte
D39	Cy^{6} phosphate	Indirecte
K41	Ad^{4} phosphate, Ur^{7} ribose, Gu^{8} phosphate	Indirecte
R46	Ad^{5} ribose	Indirecte
T49	Ad^{4} ribose / Ur^{3} ribose, Ad^{4} ribose, Ad^{5} ribose	Directe / indirecte
Complexe aaRBD-B/AWFC01 : ARN palindromique (résolution 1.90 Å)		
RBD	**ARNdb**	**Interaction**
D2	Ad^{17} ribose	Indirecte
T5	Ur^{16} ribose, Ad^{17} ribose	Indirecte
D29	Ur^{16} ribose, Ad^{17} ribose	Indirecte
A30	Ur^{16} ribose	Indirecte
P31	Ad^{4} ribose, Cy^{6} ribose / Ur^{15} ribose	Directe / indirecte
D34	Gu^{14} nucléobase, Ur^{15} ribose, Cy^{6} ribose, Ur^{7} ribose	Indirecte
R35	Ad^{4} ribose, Cy^{6} ribose / Ad^{5} ribose	Directe / indirecte
A38	Gu^{8} ribose	Indirecte
D39	Cy^{6} phosphate	Indirecte
K41	Ad^{4} phosphate, Ur^{7} ribose, Gu^{8} phosphate	Indirecte
R46	Ad^{5} ribose	Indirecte
T49	Ad^{4} ribose, $Ad^{4'}$ ribose / Ur^{3} ribose, Cy^{2} ribose	Directe / indirecte

Complexe RBD-B/ZKO* : ARN palindromique (résolution 2.30 Å)		
RBD	**ARNdb**	**Interaction**
H-1	Ur^{19} phosphate	Directe
S1	Ur^{17} ribose	Directe
D2	Ad^{17} ribose	Indirecte
T5	Ur^{16} ribose, Ad^{17} ribose	Indirecte
D29	Ur^{16} ribose, Ad^{17} ribose	Indirecte
A30	Ur^{16} ribose	Indirecte
P31	Gu^{6} ribose, $Gu^{6'}$ ribose / Ur^{15} nucléobase	Directe / indirecte
D34	Gu^{14} nucléobase, Ur^{15} nucléobase, Cy^{6} nucléobase, Ur^{7} ribose	Indirecte
R35	Gu^{6} ribose, $Gu^{6'}$ ribose / Ad^{5} ribose	Directe / indirecte
R37	Ur^{7} ribose	Indirecte
R38	Gu^{6} phosphate, Cy^{7} phosphate, $Cy^{7'}$ phosphate	Directe et indirecte
R38'	Cy^{7} phosphate, $Gu^{6'}$ phosphate, $Cy^{7'}$ phosphate	Directe et indirecte
D39	Cy^{6} phosphate	Indirecte
R46	Ad^{5} ribose	Indirecte
T49	Cy^{4} ribose, $Cy^{4'}$ ribose / Ur^{3} ribose, Ad^{4} ribose, Ad^{5} ribose	Directe / indirecte

B.II.1.a. Conservations structurales entre RBD-A et RBD-B

Dans le but de résoudre la structure du RBD-B libre, nous avons utilisé les kits de cristallisation qui sont décrits dans la section Matériel et Méthodes. Les cristaux (**Fig.67A**) formés en présence de 0,1 mol.L^{-1} HEPES pH 7,5; 8 % EG (v/v); 30 % P550MME-PEG 20K pH 7,5 ont permis de résoudre la structure 3D de la protéine libre. Des conditions de cristallisation très proches avaient d'ailleurs permis de cristalliser et de résoudre la structure tridimensionnelle du RBD de NS1 des virus de chauve-souris H17N10 et H18N11 (Turkington *et al*., 2015).

Le RBD-B de NS1 (PDB : 6SW8) est un homodimère dans lequel chaque monomère (indiquées respectivement en orange et en vert, **Fig.67B**) est composé de trois hélices α antiparallèles d'une vingtaine d'acides aminés chacune. La base des hélices α2 et α2' est majoritairement composée d'acides aminés basiques (surface cationique représentée en bleue, **Fig.67C et 67D**) qui assure l'interaction avec l'ARN. Même si l'on remarque du polymorphisme entre les RBD-A et RBD-B (**Fig.57**), la structure 3D du RBD-B reste très similaire à celles des autres RBD-A résolues à ce jour avec des RMSD (Root Mean Square Deviation) de 0,47; 0,63; 1,05 Å calculé sur les carbones Cα de la protéine en comparaison des RBD H3N2, H1N1 et H5N1, respectivement. La structure du RBD-B est cependant plus éloignée de la structure des RBD BatH17N10 et BatH18N11 qui présentent respectivement un RMSD de 1,42 Å et 2,18 Å calculés sur les carbones Cα de la protéine (Wacquiez *et al*., 2020; **Annexe 13**).

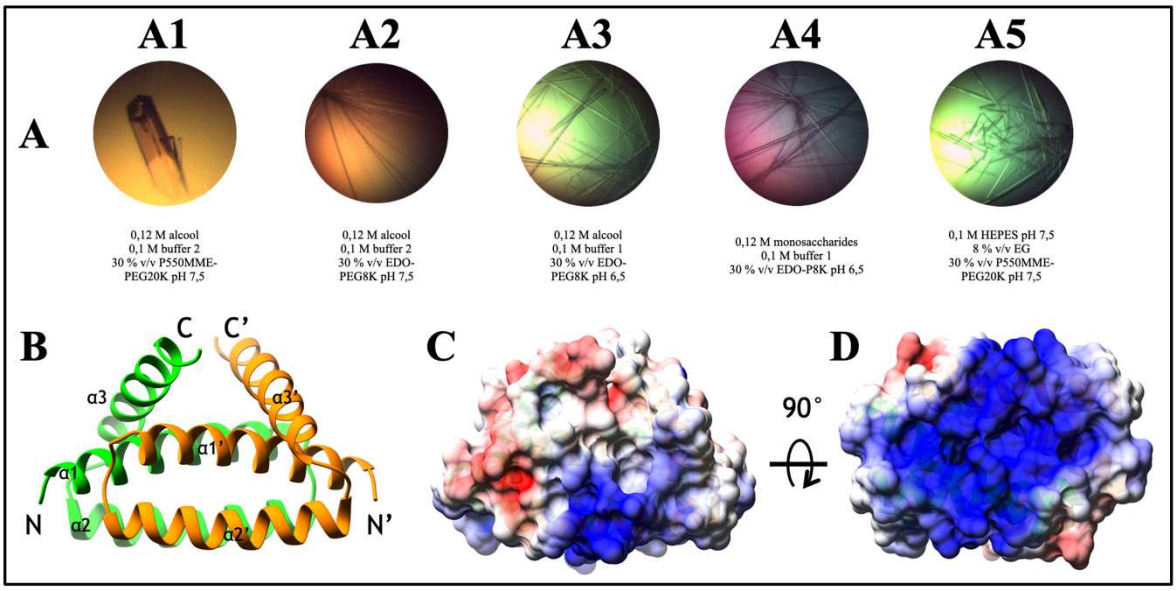


Figure 67 : Structure cristallographique du RBD-B. *(**A**) Les conditions de cristallisation du RBD-B sont indiquées en dessous de chacun des clichés de cristaux obtenus avec les kits Morpheus (A1-A4) et Classic (A5) (voir Matériel et Méthodes). Les monomères du RBD de NS1 sont représentés en orange et en vert respectivement (**B**). Les acides aminés chargés négativement sont représentés en rouge alors que ceux chargés positivement sont bleus (**C** et **D**). La structure du RBD-B (PDB : 6SW8) a été représentée d'après la conformation cartoon avec le logiciel Chimera.*

B.II.1.b. Comparaison entre RBD-B libre et RBD-B/AWFC01

Pour obtenir des cristaux du complexe RBD-B/AWFC01 (**Fig.68B** et **68C**), nous avons utilisé les kits de cristallisation qui sont décrits dans la section Matériel et Méthodes. Les cristaux obtenus avec 0,1 M CHES pH 9,5; 30 % PEG 400 (**Fig.68A**) ont été utilisés pour acquérir un jeu de données de diffraction suffisant pour résoudre la structure 3D du complexe RBD-B/AWFC01.

Les interactions directes et indirectes établies dans les différents complexes structuraux étudiés sont respectivement présentées dans le **tableau 8** et en **annexe 7**. L'analyse de la structure du complexe RBD-B/AWFC01 (PDB : 6SX0), résolue par remplacement moléculaire en utilisant la structure de Cheng *et al*. (2009) (PDB : 2ZKO) comme structure de référence, démontre que les hélices α2 et α2' du dimère de RBD forment la surface d'interaction avec le grand sillon de l'ARN, interface au sein de laquelle la double hélice d'ARN interagit de façon *quasi* antiparallèle à l'axe des hélices α2 et α2' (Yin *et al*., 2007; Cheng *et al*., 2009). En solution, l'hélice d'ARN adopte une conformation de type A qui, lors de l'interaction avec le RBD de NS1, présente peu de courbure particulière le long de son axe central. Cependant, des courbures supérieures (environ 40°) sont observées sur les bords de la double hélice facilitant ainsi l'interaction de l'ARN avec la protéine. Une poche profonde située au niveau des hélices α2 et α2' du RBD est dissimulée au niveau de l'interface d'interaction protéine-ARN, dont la base de la poche est assurée par les hélices α1 et α1' et le domaine *couvercle* de la poche par les acides aminés conservés R38 présents sur chacun des deux monomères.

L'analyse des structures démontre l'importance des acides aminés contenus dans les hélices α2 et α2' dans l'interaction avec l'ARN. Les acides aminés chargés positivement contenus dans ces deux

hélices parallèles, en particulier R35, R37, R38 et K41, sont tournés vers le centre de l'interface de liaison à l'ARN, ce qui permet au RBD d'établir des interactions électrostatiques et des ponts hydrogènes avec les deux brins d'ARN simultanément. Parmi ces acides aminés, seul R38 pénètre au cœur de la région centrale de la double hélice d'ARN (Yin *et al*., 2007; Cheng *et al*., 2009), en formant des interactions de type hydrogène avec les phosphates des nucléobases Cy^6, Ur^7 et $Ur^{7'}$ de l'ARNdb AWFC01 (**Fig.68E**). Les bases Gu^2Ur^3 du motif $Gu^2Ur^3AdAdCy^6$ (notre séquence de haute affinité) ne semblent pas non plus reconnues par le RBD de NS1, suggérant l'absence d'un effet de bord de l'ARN. L'acide aminé R38 permet d'ancrer l'ARNdb à la base de la poche d'interaction du RBD *via* la formation d'un réseau de liaisons hydrogènes entretenues par les deux monomères symétriques, empêchant ainsi l'ARN de glisser le long de la surface d'interaction des hélices α2 et α2'. L'acide aminé R38 est véritablement essentiel à la formation de complexe RBD/ARNdb. C'est certainement pour cette raison que le mutant R38A, à l'inverse du mutant K41A, réduit significativement (mais pas totalement) l'interaction avec l'ARN, impactant ainsi très fortement le cycle viral (Min *et al*., 2006). Les acides aminés R35 et R35', qui sont impliqués dans la dimérisation du RBD en interagissant respectivement avec R46 et R46', participent aussi à l'interaction avec l'ARN de façon symétrique, stabilisant ainsi la double hélice d'ARN sur les hélices α2 et α2' *via* des liaisons hydrogènes entretenues avec le ribose des bases Cy^6 et $Cy^{6'}$ (**Fig.68E**). Les acides aminés T49 et T49' du RBD interagissent avec le ribose des bases Ad^4 et $Ad^{4'}$. Tout comme T49, l'acide aminé S42 (qui n'existe pas dans le RBD-B mais est présent dans le complexe RBD/CHENG) joue un rôle important dans la formation de complexe RBD/ARNdb comme l'indique la réduction d'un facteur 10 en termes d'affinité d'interaction pour ces deux mutants (Cheng *et al*., 2009). Les acides aminés P31 et P31' et R37 interagissent respectivement avec le ribose des bases Cy^6 et Ur^7 de l'ARN (**Fig.68E**).

En absence d'ARN, les acides aminés localisés proches de R38 restent globalement dans une conformation inchangée, laissant ainsi la poche profonde du RBD apparente comme déjà décrit par Yin *et al*. (2007) et par Cheng *et al*. (2009). A l'inverse et contrairement à ce qui a été décrit par Chien *et al*. (2004), l'interaction avec l'ARN provoque des changements de conformation pour certains de ces acides aminés. L'effet le plus marquant concerne l'acide aminé R38 (et R38') dont l'orientation est modifiée de telle sorte que ce dernier soit tourné vers l'interface RBD/ARNdb (**Fig.68D**), facilitant l'interaction des charges positives de R38 avec les phosphates du grand sillon de l'ARN. Ce réarrangement structural entraine alors un masquage de la poche profonde du RBD de NS1 (Cheng *et al*., 2009). L'orientation des acides aminés R37', K41' et K44' semble aussi modifiée lorsque le RBD-B est associé à AWFC01 (**Fig.68D**), cependant aucun d'eux ne contacte directement l'ARNdb (**tableau 8**). Ces résultats suggèrent que tous les changements de conformation intervenant dans le RBD-B ne modulent pas forcément tous l'interaction avec l'ARN.

Certains acides aminés du RBD de NS1 peuvent interagir de façon directe et/ou indirecte avec l'ARNdb (**Fig.68E** ; **Annexe 7** ; **tableau 8**). C'est le cas des acides aminés P31, R35, R37, R38 et T49 qui sont capables d'interagir à la fois directement avec le squelette phosphate de certaines nucléobases

de l'ARN mais aussi indirectement avec ces mêmes bases (et d'autres aussi) *via* des interactions médiées par des molécules d'eau. A titre d'exemple, l'acide aminé R38 interagit de façon directe et indirecte avec les phosphates des nucléobases Cy^6, Ur^7 et $Ur^{7'}$ de AWFC01. A l'inverse, P31 est capable d'interagir directement avec le ribose de la nucléobase Cy^6 et indirectement avec la nucléobase Ur^{15}. De façon similaire, R35 établit de façon directe des liaisons hydrogènes avec le ribose de la nucléobase Cy^6 et de façon indirecte avec le ribose de la nucléobase Ad^5 de l'ARN (**Annexe 7**). D'autres acides aminés sont uniquement retrouvés dans un profil d'interactions indirectes avec AWFC01. En effet, les acides aminés D2, T5, D29, A30, D34, D39, K41, R46 et T49 interagissent avec l'ARNdb uniquement par le biais de molécules d'eau. L'acide aminé D34 est capable d'établir des contacts intermoléculaires de façon symétrique avec au moins quatre molécules d'eau, stabilisant ainsi la formation du complexe RBD/ARNdb. D'autres acides aminés comme A30, D39 et R46 contactent une unique nucléobase de l'ARNdb *via* une seule molécule d'eau. Au contraire, T5, D29 et K41 interagissent avec une molécule d'eau qui est capable de ponter deux nucléobases (**Annexe 7**). Globalement, seules les bases des deux motifs GUAAC sont ciblées par les acides aminés du RBD, aucunes des bases centrales ne semblent reconnues par le RBD de NS1 (**tableau 8** ; **Annexe 7**). ***L'analyse de cette nouvelle structure RBD/ARN montre une reconnaissance non-spécifique de la séquence de l'ARN par le RBD de NS1 puisqu'aucunes nucléobases de l'ARN n'est reconnue. Cependant et de façon surprenante, aucunes des bases de l'ARN ne sont directement reconnues par la protéine puisque le RBD de NS1 contacte l'ARN par l'intermédiaire du squelette ribose-phosphate des nucléobases de ce dernier. Si nos résultats d'interaction en gel retard et après avoir démontré la haute affinité de notre sonde AWFC01 pouvaient indiquer une reconnaissance spécifique des nucléobases de l'ARN, en particulier celles des motifs GUAAC, l'analyse structurale du complexe RBD-B/AWFC01 démontre l'absence de reconnaissance directe des bases de l'ARN par la protéine.***

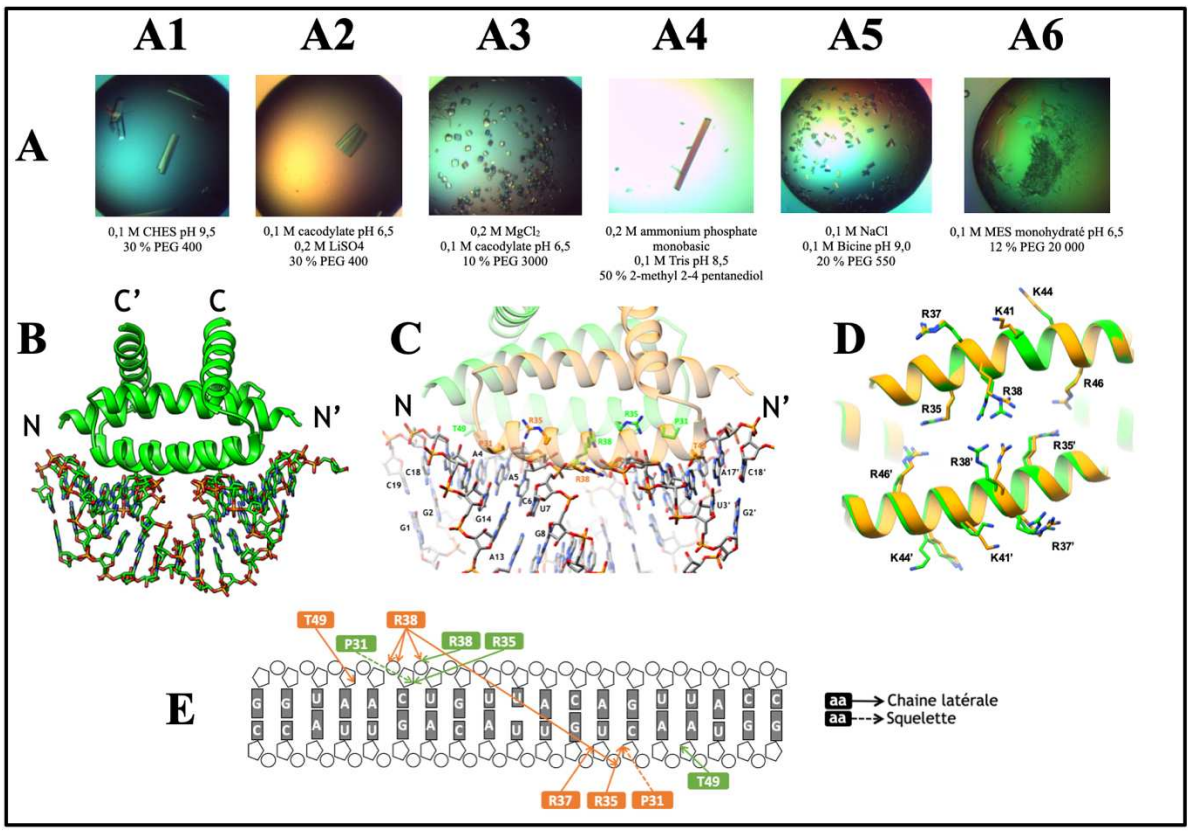


Figure 68 : Caractéristiques de la structure cristallographique du complexe RBD-B/AWFC01. *(**A**) Les conditions de cristallisation du complexe RBD-B/AWFC01 sont indiquées en dessous de chacun des clichés de cristaux obtenus avec les kits Wizard (A1-A3) et HR (A4-A6) (voir Matériel et Méthodes). (**B**) Structure du complexe RBD-B/AWFC01 (PDB : 6SX0) représentée d'après la conformation cartoon avec le logiciel Chimera. (**C**) Structure annotée du complexe RBD-B/AWFC01. (**D**) Représentation de la conformation des acides aminés du RBD par superposition des structures du RBD-B libre (jaune) et du complexe RBD-B/AWFC01 (vert). (**E**) Interactions directes des acides aminés du RBD-B avec AWFC01.*

B.II.1.c. Comparaison entre RBD-B/AWFC01 et aaRBD-B/AWFC01

Pour obtenir des cristaux du complexe aaRBD-B/AWFC01, nous avons utilisé les kits de cristallisation qui sont décrits dans la section Matériel et Méthodes. Les cristaux obtenus avec 0,1 M Na phosphate/acide citrique pH 4,2; 0,2 mol.L^{-1} ammonium sulfate; 40 % éthylène glycol (**Fig.69A**) ont été utilisés pour acquérir un jeu de données de diffraction suffisant pour résoudre la structure 3D du complexe RBD-B/AWFC01.

La superposition de la structure du complexe aaRBD-B/AWFC01 et du RBD-B indique des réorientations de certains acides aminés, en particulier R37 et R37' (**Fig.69B**). Cependant, ni l'un ni l'autre de ces deux acides aminés n'interagit directement ou indirectement avec AWFC01 (**Fig.69D** ; **Annexe 7**). Les acides aminés P31 et R35 du RBD-B interagissent avec le ribose de la nucléobase Ad[4] alors qu'ils étaient décrits comme interagissant avec le ribose de la nucléobase Cy[6] dans le cas du RBD-B sauvage. Dans ce cas précis, il semble que l'absence de R38 conduit à des réorientations d'acides aminés pouvant moduler les interactions directes avec l'ARN. Les structures des complexes RBD-B/AWFC01 et aaRBD-B/AWFC01 superposées représentées dans la **Fig.69C** indiquent la conservation des complexes structuraux.

Contre toute attente et alors que la double-substitution R38A-K41A réduit considérablement l'interaction du aaRBD-B avec l'ARN double-brin, nous avons obtenu la structure du complexe aaRBD-B/AWFC01 (**Fig.69C**), démontrant que cette double-mutation n'impacte pas la formation du complexe à haute concentration et sa cristallisation. La cristallisation favorise les complexes de plus haute affinité correspondant alors à l'espèce la plus énergétiquement stable. A titre d'exemple, Bornhold et Prasad (2008) avaient résolu la structure d'une NS1 contenant la double-substitution R38A-K41A à partir d'un virus humain H5N1 (PDB : 3F5T, **Tableau 2**). L'analyse de cette nouvelle structure démontre que l'interface d'interaction, et non sa surface électrostatique, reste inchangée même en absence de R38 (et de K41) (RMSD de 0,78 Å).

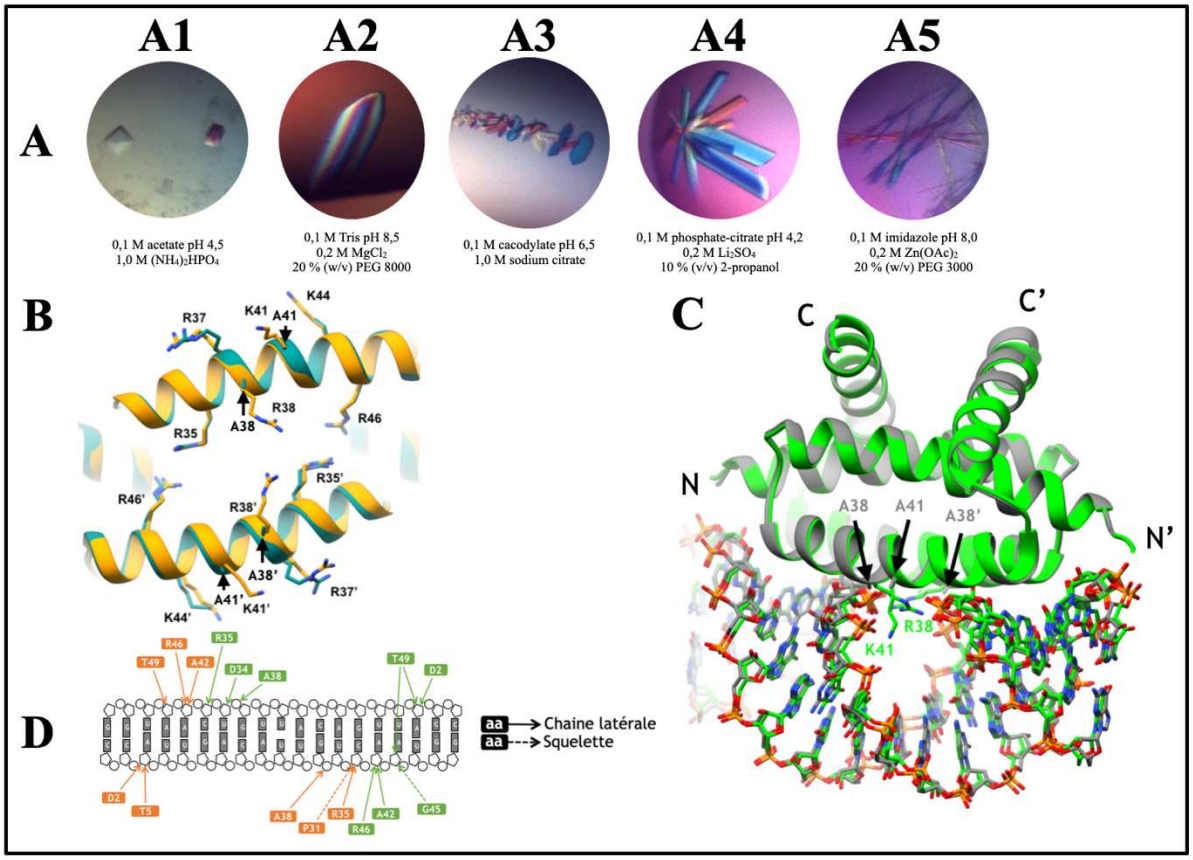


Figure 69 : Caractéristiques de la structure cristallographique du complexe aaRBD-B/AWFC01. *(**A**) Les conditions de cristallisation du complexe aaRBD-B/AWFC01 sont indiquées à droite de chacun des clichés de cristaux obtenus avec le kit Wizard cryo (A1-A5) (voir Matériel et Méthodes). (**B**) Représentation de la conformation des acides aminés du RBD par superposition des structures du RBD-B libre (jaune) et du complexe aaRBD-B/AWFC01 (bleu). (**C**) Superposition des structures des complexes RBD-B/AWFC01 (vert; PDB : 6SX0) et aaRBD-B/AWFC01 (gris; PDB : 6SX2). (**D**) Interactions directes des acides aminés du aaRBD-B avec AWFC01.*

B.II.1.d. Comparaison entre RBD-B/AWFC01 et RBD/CHENG

Ce sont globalement les mêmes profils d'interactions directes et indirectes qui interviennent dans les complexes RBD-B/AWFC01 et RBD/CHENG, même si des différences subtiles dues au polymorphisme entre les deux RBD utilisés apparaissent. Que ce soit dans le cas du complexe RBD-

B/AWFC01 ou du complexe RBD/CHENG, seul le squelette ribose-phosphate de l'ARNdb est reconnu par le RBD de NS1 ce qui témoigne d'une reconnaissance non-spécifique de l'ARN passant par l'établissement de nombreuses interactions avec la fonction 2'OH du ribose de certaines nucléobases. Toutefois, aucune reconnaissance des bases de l'ARN n'a été observée. L'analyse structurale des complexes RBD-B/AWFC01 et RBD/CHENG démontre de grandes similitudes entre les complexes, comme en témoigne la superposition des deux complexes qui est présentée en **Figure 70**. Autre détail remarquable, la superposition des complexes RBD-B/AWFC01 (RMSD de 0,82 Å) et aaRBD-B/AWFC01 (RMSD de 0,78 Å) vis à vis du complexe RBD/CHENG, démontre la conservation de l'interface d'interaction d'un modèle structural à l'autre. ***Comme déjà décrit dans la littérature, ce sont les mêmes acides aminés qui contactent indirectement les nucléobases de l'ARNdb, en interagissant avec leur squelette ribose-phosphate. A l'heure actuelle, l'analyse seule des interactions directes et indirectes engagées par les acides aminés du RBD de NS1 envers l'ARNdb ne nous permet ni de mettre en évidence un mode reconnaissance spécifique des bases de l'ARN (en particulier concernant la reconnaissance du motif GUAAC) ni d'expliquer clairement les différences d'affinité du RBD de NS1 pour ces deux ARNdb.***

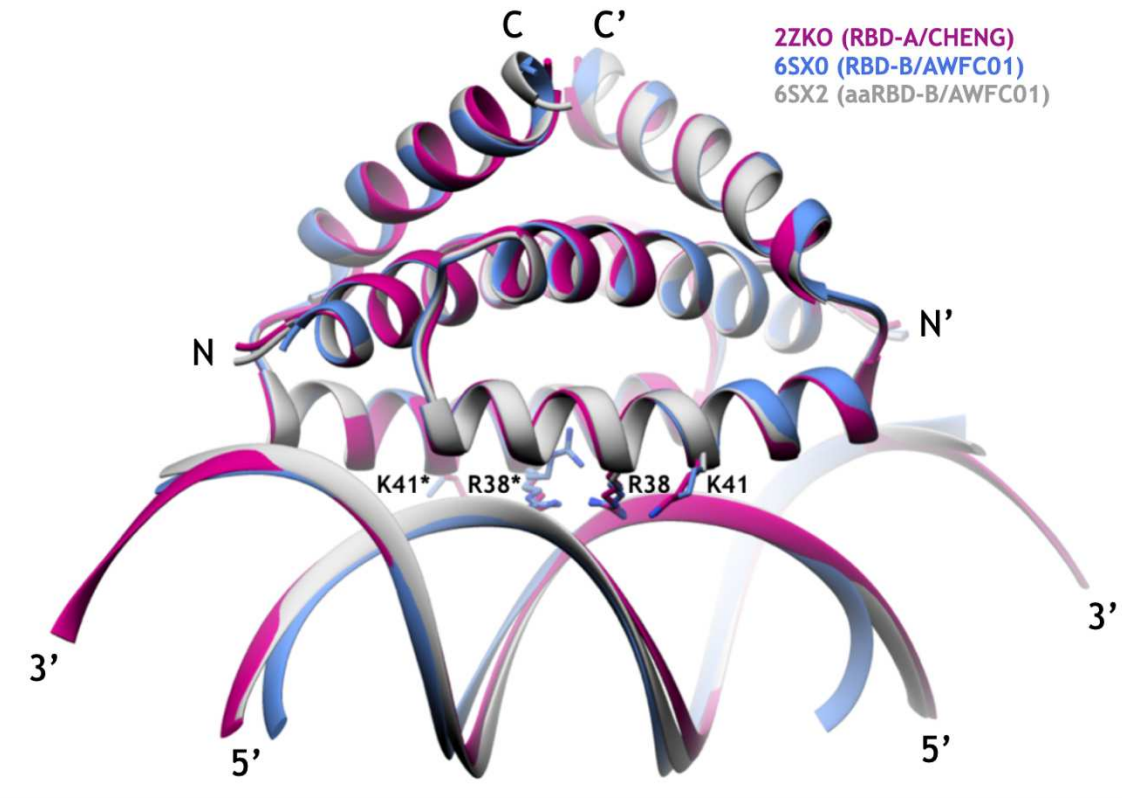


Figure 70 : Caractéristiques de la structure cristallographique du complexe RBD/CHENG. *Superposition des structures des complexes RBD-B/AWFC01 (bleu, PDB : 6SX0), aaRBD-B/AWFC01 (gris, PDB : 6SX2) et RBD/CHENG (violet, PDB : 2ZKO).*

B.II.1.e. Comparaison entre RBD-B/ZKO* et RBD/CHENG

Nous avons résolu la structure du complexe RBD-B/ZKO*, un ARN qui dérive de CHENG mais qui ne présente pas les bases UrUr à ses extrémités, afin de savoir si le polymorphisme du RBD-B pouvait induire d'éventuelles changements structuraux dans l'ARN afin d'expliquer les différences d'affinité flagrantes du RBD de NS1 pour les sondes AWFC01 et CHENG. Pour obtenir des cristaux du complexe RBD-B/ZKO*, nous avons utilisé les kits de cristallisation qui sont décrits dans la section Matériel et Méthodes. Les cristaux issus de l'optimisation des conditions de cristallisation avec 2 $mol.L^{-1}$ ammonium sulfate; 0,2 $mol.L^{-1}$ ammonium nitrate ont été utilisés pour acquérir un jeu de données structurales suffisant pour résoudre la structure 3D du complexe RBD-B/ZKO*. Par ailleurs, les conditions de cristallisation du complexe RBD/CHENG résolu par Cheng *et al*. (2009) (0,02 $mol.L^{-1}$ magnesium acetate, 0,05 $mol.L^{-1}$ 4-morpholineethanesulfonic acid (MES) pH 5,2; 2,6 $mol.L^{-1}$ ammonium sulfate; 0,01 $mol.L^{-1}$ adenosine-5-triphosphate disodium) divergent de nos conditions de co-cristallisation.

L'analyse des interactions directes intervenant dans le complexe RBD-B/ZKO* indique l'intervention des acides aminés H-1 et S1 qui interagissent respectivement avec le ribose de la nucléobase Ur^{17} et le phosphate de Ur^{19} de ZKO* (**Fig.71A**). L'absence de l'acide aminé S42 (remplacé par A42 dans la séquence du RBD-B) ne semble pas impacter l'interface d'interaction. La paire d'acide aminé R38 forme tout de même plus d'interactions directes avec ZKO* qu'avec CHENG (**Fig.71A** et **71B**). Il semble aussi que P31, de part sa flexibilité puisque localisé en N-terminal des hélices $\alpha 2$ et $\alpha 2'$ (Abi Hussein *et al*., 2020), interagit de façon symétrique avec le ribose de la nucléobase Gu^{6} (et $Gu^{6'}$) de ZKO* (**Fig.71B**) alors qu'une seule interaction est retrouvée pour le complexe RBD/CHENG (**Fig.71A**).

L'identification d'un maximum de molécules d'eau intervenant dans un complexe protéine-ARN dépend directement de la résolution d'une structure 3D. Ainsi, la résolution limitée de 2,30 Å obtenue pour le complexe RBD-B/ZKO* ne permet pas d'identifier clairement la majorité des molécules d'eau intervenant dans le complexe. Même si moins nombreuses car moins résolutive que les autres structures, les molécules d'eau intervenant dans ce complexe sont confirmées. Le complexe RBD-B/ZKO* présente globalement les mêmes interactions directes et indirectes que celles retrouvées dans les complexes RBD/CHENG, RBD-B/AWFC01 et aaRBD-B/AWFC01 (**tableau 8**). Tous les complexes semblent se comporter de façon plus ou moins similaire comme en témoigne la superposition structurale du complexe RBD-B/AWFC01 avec (i) aaRBD-B/AWFC01 (**Fig.69C**) et (ii) RBD-CHENG (**Fig.70**) et également entre les complexes RBD-B/ZKO* et RBD/CHENG (**Fig.71**). Dans le cas du complexe RBD-B/ZKO*, seule l'orientation des acides aminés R37 et K41 semble modifiée par rapport à la structure RBD/CHENG (**Fig.71D**), cependant ni l'un ni l'autre n'est impliqué dans une interaction directe avec l'ARN (**Fig.71A** et **71B**). A l'inverse, l'acide aminé R37, qui n'était pas retrouvé dans le complexe RBD/CHENG, interagit avec le ribose de la nucléobase Ur^{7} *via* une molécule d'eau (**Annexe**

7). ***Le polymorphisme existant entre les RBD-A et RBD-B de ces deux structures ne modifie pas les profils d'interactions directes et indirectes intervenant dans le complexe RBD/ARN. D'après nos résultats, il semble que la symétrie imposée par le palindrome (CHENG, ZKO* et AWFC01) n'impacte pas ces profils d'interactions.***

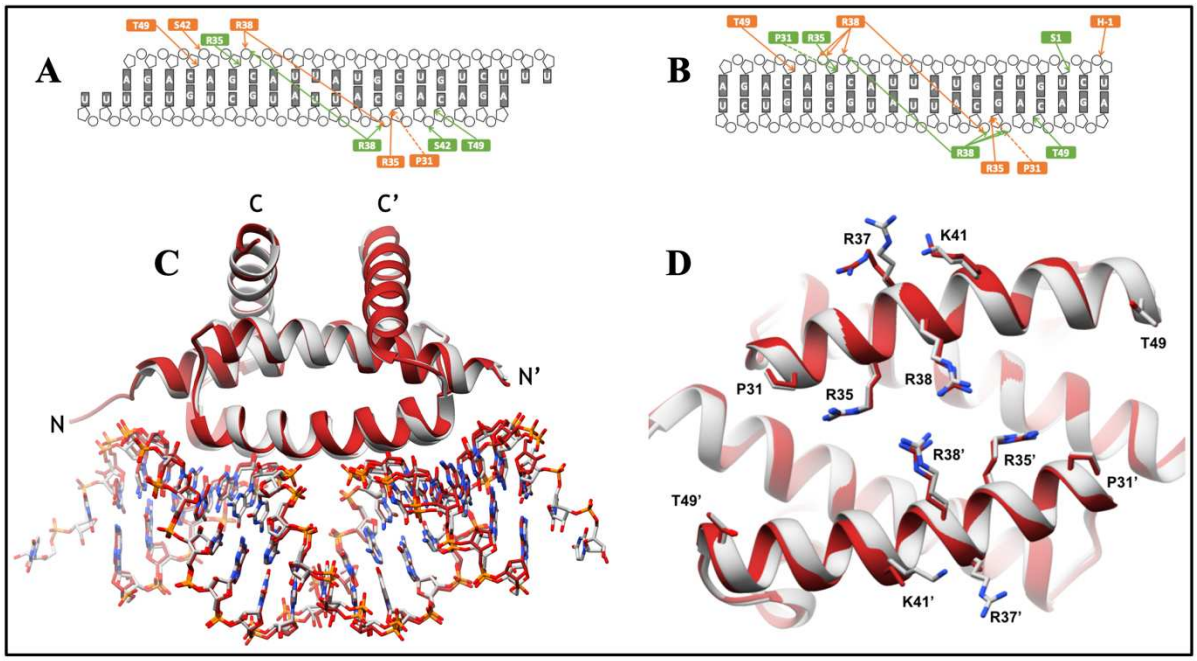


Figure 71 : Caractéristiques de la structure cristallographique du complexe RBD-B/ZKO*. *(**A**) Interactions directes des acides aminés du RBD avec CHENG. (**B**) Interactions directes des acides aminés du RBD-B avec ZKO*. (**C**) Superposition des structures RBD/CHENG (gris) et RBD-B/ZKO* (rouge). (**D**) Représentation de la conformation des acides aminés du RBD par superposition des structures RBD/CHENG (gris) et RBD-B/ZKO* (rouge).*

B.II.2. L'interaction du RBD-B avec AWFC01 induit une légère courbure de l'ARN

Même si nous avons précédemment montré que le RBD de NS1 était environ 50 à 100 fois plus affin pour AWFC01 que pour CHENG, l'analyse des structures, médiées par le détail des interactions directes et indirectes intervenant dans les différentes structures 3D, ne montre pas de différences flagrantes entre les complexes. A ce jour, il nous est difficile d'apporter des explications approfondies concernant les différences d'affinité existantes entre les ARNdb observées en gel retard. Pour tenter d'aller un peu plus loin, nous avons réalisé des analyses fines détaillant les paramètres hélicoïdaux de l'ARN complexé ou non au RBD de NS1 (CHENG, ARN canonique nous servant de contrôle). L'objectif est de mettre en évidence d'éventuelles modifications de la double-hélice de l'ARN qui surviendraient lors de l'interaction avec le RBD de NS1 du fait de sa flexibilité. En complexe, l'ARNdb peut présenter une certaine élasticité pouvant être caractérisée par des torsions, des flexions et des étirements de ce dernier. Nous avons donc analysé les mouvements des nucléobases (selon l'axe

centrale, intra- et inter-bases, la courbure de l'ARN), la torsion de leurs squelettes ribose-phosphate ainsi que la profondeur et la largeur des sillons de l'ARN en complexe vis à vis de l'ARN libre (CHENG).

La classification internationale établie par Dickerson (1989) indique que les mouvements intra- et inter-bases de l'ARN peuvent être évalués selon plusieurs paramètres en suivant ces derniers sur les axes x, y et z (**Annexe 8**). Les paramètres hélicoïdaux de l'ARN, à savoir les déplacements des bases sur les axes x, y et z, les glissements des bases selon les axes x et y, les angles de pivotements, de dévers et d'inclinaison des bases et le basculement, l'inclinaison, le décalage, l'étirement, le cisaillement, l'ouverture, la dislocation et la pliure des bases sont globalement similaires entre les différents complexes étudiés (**Annexe 9**). Seule le paramètre de courbure de l'ARN semble diverger entre les différents complexes. En effet, le RBD-B induit une légère courbure de l'axe de AWFC01 au niveau des paires de bases 7 et 8 qui est remarquable. Alors que la conformation de type A de l'ARN provoque une réduction de la largeur du grand sillon de l'ARN et au contraire une augmentation de sa profondeur et inversement pour le petit sillon de l'ARN, la courbure de AWFC01 induite par le RBD-B s'accompagne d'un léger rétrécissement et une diminution de la profondeur du petit sillon de l'ARN qui sont compensés par un effet inverse sur le grand sillon de AWFC01, indiquant clairement que l'interaction du RBD-B avec AWFC01 provoque des changements dans la structure de l'ARN (Wacquiez *et al*., 2020).

Pour compléter notre analyse, nous avons évalué la torsion du squelette ribose-phosphate des nucléobases de l'ARN (**Annexe 10**) qui pourrait être modifiée lors de l'interaction avec le RBD de NS1. Analyser la torsion du squelette ribose-phosphate revient à évaluer les angles de torsion α, β, γ, δ, ε et ζ du squelette ribose-phosphate, les angles endocycliques du ribose ($\nu_{0\rightarrow4}$) et l'angle de rotation de la liaison glycosidique qui sépare le ribose de la base (Chi, χ). Les six angles (α, β, γ, δ, ε et ζ) sont décrits par les torsions α (03'-P-O5'-C5'), β (P-O5'-C5'-C4'), γ (O5'-C5'-C4'-C3'), δ (C5'-C4'-C3'-O3'), ε (C4'-C3'-O3'-P) et ζ (C3'-O3'-P-O5'), auxquels s'ajoutent la torsion de la fonction 2'OH libre du ribose (C1'-C2'-O2'-H2') et l'angle glycosidique (Chi**,** χ) (O4'-C1'-N1-C2 pour des pyrimidines et O4'-C1'-N9-C4 pour des purines) qui sépare le ribose de la nucléobase. Les valeurs des angles endocycliques du ribose ($\nu_{0\rightarrow4}$) de l'ARN sont déterminées en utilisant jusqu'à quatre paramètres de pseudorotation, en particulier les paramètres de *Phase* et d'*amplitude* qui sont les plus utilisés. Alors que l'*amplitude* décrit le degré de plissement du ribose, la *phase* indique les atomes les plus déplacés par rapport au plan médian du ribose. Ces deux paramètres sont évalués en tenant compte des angles dihèdres du ribose, à savoir ν_1 (C1'-C2'-C3'-C4') jusqu'à ν_5 (O4'-C1'-C2'-C3') (Westhof et Sundaralingam, 1983). L'angle de rotation de la liaison glycosidique (Chi, χ), qui sépare le ribose de la base azotée, correspond à la rotation de la base par rapport au ribose. Deux conformations possibles sont définies en fonction de la valeur de l'angle χ: la conformation *anti* (χ = 180 +/- 90°) et la conformation *syn* (χ = 180 +/- 90°).

Notre analyse, dont les résultats sont présentés en **annexes 8-11**, ne montre pas de différences marquantes entre les complexes structuraux étudiés. Sur ce plan et à la différence de la légère différence observée au niveau de la courbure de l'ARN induite par le RBD-B sur AWFC01, nos complexes résolus au laboratoire sont très similaires entre-eux, y compris pour le aa-RBD-B/AWFC01, et globalement similaires au complexe résolu par Cheng *et al*. (2009). ***L'analyse détaillée des paramètres hélicoïdaux de l'ARN indique une légère courbure de AWFC01 induite par le RBD-B. Cependant, l'ensemble des autres paramètres structuraux reste très similaire entre les différents complexes analysés. D'après nos observations, il semble que l'interaction du RBD-B avec l'ARNdb suit de près les mêmes schémas que les autres RBD déjà caractérisés.***

CONCLUSIONS GENERALES ET DISCUSSION

Importance biologique de l'interaction NS1-ARN

Dans cette partie, nous n'avons examiné en réalité que l'interaction de NS1 avec le motif UGAUUGAAG (motif B) qui est présent dans la région 3'-UTR de l'ARNm de NS1. Cela a été exploré en utilisant essentiellement le système miniréplicon viral ainsi que la génération de virus WSN mutants. L'ensemble de nos observations expérimentales peut se résumer en quelques éléments essentiels.

(i) NS1 accroit l'activité de la RdRp virale. Nous avons observé que NS1 augmente considérablement l'activité du miniréplicon (facteur 4 à 6), ce qui pourrait résulter d'une amélioration par NS1 de l'activité de la RdRp virale lors des étapes de réplication et de transcription virales, mais aussi d'une amélioration de la traduction du *viral-like mRNA*. Ceci est aussi observé, mais de manière moins importante, avec la protéine NS1 portant la double-substitution R38A-K41A qui abolit l'interaction avec l'ARN. Ceci suggère que cette activité, qui ne repose pas en totalité sur la capacité du RBD à se lier aux ARN, pourrait être moduler par le domaine ED seul. Cette modulation positive de l'activité de la RdRp virale repose probablement sur une interaction indirecte de NS1 avec la RdRp virale. Marion *et al*. (1997) avaient déjà montré l'interaction de NS1 avec le complexe RdRp du virus *in vivo*, et précisé que cette interaction ne reposait pas sur l'ARN viral qui se lierait aux deux partenaires (RdRp et NS1). Robb *et al*. (2011) ont également montré que dans des cellules infectées NS1 interagissait spécifiquement avec la protéine NP qui est le composant majeur des RNPv et des RNPc. En fractionnant le contenu de cellules infectées (cytoplasme, noyau et chromatine) par une lyse chimique et mécanique, Robb *et al*. (2011) ont montré que NS1 co-immunoprécipitait avec les protéines PA et NP (et PB2 lorsque les auteurs utilisent un virus WSN PB2-streptavidine recombinant) dans les trois compartiments, avec une importante accumulation au niveau de la chromatine. Même si NS1 n'interagit pas directement avec le complexe RdRp viral, son interaction avec NP (cette dernière interagit avec PB2) lui permet de se retrouver en étroite relation avec la RdRp virale. NS1 a été décrite comme interagissant avec plusieurs types d'ARN viraux (Hatada *et al*., 1992; 1997; Hatada et Fukuda, 1992; Park et Katze, 1995). L'interaction directe de NS1 avec ces ARN viraux ainsi qu'avec NP situés au sein des RNPv et des RNPc semble l'hypothèse la plus probable selon laquelle NS1 stimule les étapes de transcription et de réplication du génome viral. Les résultats du système *translation-only* suggèrent que, dans ce système, NS1 n'accroît pas la traduction des transcrits viraux. Toutefois, nous devons remarquer que ce système très artificiel reflète mal la traduction des ARNm viraux dans la cellule infectée. En effet, les transcrits chimériques se distinguent des ARNm viraux : (i) leur transcription et leur polyadénylation sont assurées par l'ARN Polymerase II cellulaire, et non par l'ARN polymérase ARN-dépendante virale; (ii) ils subissent un épissage et (iii) ils portent en 5' et en 3' des extensions artificielles qui ne sont pas retrouvées dans les ARNm viraux. De plus, nous observons dans nos conditions expérimentales, un effet réducteur de NS1 sur la traduction des ARNm, alors que plusieurs autres travaux avaient démontré un effet stimulateur de NS1 sur la traduction des transcrits viraux (Anastasina *et al*., 2014; De la Luna *et al*., 1995; Marion *et al*., 1997; Panthu *et al*.,

2018) et cellulaires (Anastasina *et al*., 2014; Salvadore *et al*., 2002). En utilisant un système RRL (Rabbit Reticulocyte Lysate), Anastasina *et al*. (2014) ont montré l'effet stimulateur de NS1 sur la traduction de transcrits cellulaires rapporteurs de la luciférase. Cependant, le système RRL dans lequel la fraction ribosomique est fonctionnelle ne requiert pas d'étape de transcription contrairement à nos conditions expérimentales où l'évaluation de l'activité modulatrice de NS1 peut se faire uniquement lorsque les transcrits chimériques ont été synthétisés à partir du vecteur rapporteur transfecté en cellules HEK293T. Panthu *et al*. (2017) ont également mis en évidence l'activité stimulatrice de la traduction par NS1 dans des cellules A549 transfectées avec des ARNm transcrits *in cellulo* et qui co-expriment ou non NS1. Ils ont aussi pu mettre en évidence cette activité de NS1 dans le système RRL développé par Anastasina *et al*. (2014). L'effet stimulant de NS1 sur la traduction a été mis en évidence dans deux modèles distincts, renforçant l'idée que NS1 module positivement cette étape-là. Nos résultats, qui divergent de ceux obtenus par plusieurs autres équipes, pourraient suggérer que nos ARNm chimériques ressemblent fortement à un ARNm cellulaire. De ce fait, il est plausible que l'effet de NS1 sur les transcrits chimériques résulte de l'activité inhibitrice de NS1 sur la maturation et l'export nucléo-cytoplasmique des transcrits cellulaires. **Finalement, en combinant nos propres résultats à ceux décrits dans la littérature, il est vraisemblable que l'effet de NS1 observé sur le système miniréplicon repose à la fois sur une augmentation de la traduction des ARNm viraux et sur une activité augmentée de la RdRp virale, sans que nous puissions évaluer précisément l'impact de NS1 sur ces deux mécanismes.**

(ii) L'éventuelle interaction de NS1 avec le motif B dans son propre ARN messager n'est pas biologiquement relevant. Les mutations introduites dans le motif B (C et D) ont consisté à éviter de muter l'ORF de NEP dans son domaine de NEP organisé par la dimérisation des 2 hélices C1 et C2 repliées en épingle à cheveux. Ces éléments de structure de NEP sont localisés dans la partie C-terminale de NEP (résistante au protéase, résidus 56 à 116). Les aa du motif B se trouvent dans l'hélice C1, ceux du motif C dans la boucle de l'hairpin connectant C1 à C1 et les aa du motif D dans l'hélices C2. Les aa de NEP contenus dans les motifs B, C et D sont tous impliqués dans des interactions tertiaires intramoléculaires ou quaternaires intermoléculaires entre C1 et C2. On peut donc s'attendre à ce que des mutations dans NEP au niveau de ces motifs pourraient être associées à une déstabilisation de la forme dimérique de NEP et une déstructuration locale de NEP. Toute déstructuration de NEP en particulier dans le motif B positionné immédiatement après le résidu W78 (résidu hydrophobe exposé dans C1) pourrait aussi impacter l'interaction de NEP avec M1 (Akarsu *et al*., 2003), interaction essentielle pour permettre le transport nucléocytoplasmique des RNPv néo-synthétisées (voir introduction). Il est possible à la vue de nos résultats que la contrainte de sélection du motif B (et peut-être aussi des motifs C et D) s'exerce plutôt sur la conservation de la séquence de NEP.

L'activité de NS1 dans les systèmes *miniréplicon* et *translation-only* reste inchangée lorsque les motifs B, C ou D dans la 3'-UTR de NS1 sont mutés. Donc ces motifs (et particulièrement le motif B

qui nous intéressait) n'ont aucun impact dans les deux systèmes *in cellulo*. Par conséquent, l'interaction de NS1 sur le motif B dont nous avions fait l'hypothèse n'a pas d'importance biologique. Même si nos résultats dans ces deux systèmes ne sont pas en faveur d'une interaction relevante sur le plan biologique de NS1 avec le motif B dans la 3'UTR de son propre ARN messager, il demeurait important de tester cette hypothèse dans le contexte viral. Lorsque les motifs B, C ou D (ou les 3) ont été modifiés dans le génome du virus WSN reconstitué par génétique inverse, le phénotype viral demeure inchangé, confirmant ainsi nos précédents résultats obtenus dans les systèmes artificiels. Les mutations, seules ou associées dans le génome viral, sont sans effet ni sur la cinétique virale ni sur la cinétique d'accumulation de NS1 dans les cellules infectées. **En bref, on ne peut formellement exclure une interaction de NS1 avec le motif B dans son ARN messager, même si nos expériences ne nous permettent pas de certifier le contraire.**

Concernant le motif A, nous n'avons évalué son importance ni dans le *miniréplicon* ni dans le virus lors de la génétique inverse car toute mutation dans ce site perturbe considérablement le promoteur de la RdRp virale, et toute modification dans un tel système serait directement imputable à une invalidation de ce promoteur, sans que l'on puisse distinguer un effet qui serait propre à NS1. Il est en effet bien reconnu que des mutations dans les séquences U12 et U13, propres à chaque genre de virus *influenza*, réduisent considérablement la réplication et la transcription du génome viral (Crescenzo-Chaigne *et al*. 2013). De ce fait et puisque que la séquence virale AGCAAAAG est transcrite à partir de la région U12 du promoteur viral, il semble que cette séquence soit réellement essentielle au virus. Plusieurs travaux ont montré, plus spécifiquement, que NS1 favorise la traduction des ARNm viraux et que cette activité requiert la présence de la 5'UTR contenant le motif AGCAAAAG (De la Luna et al., 1995; Enami et al., 1994; Hatada et al., 1990; Marión et al., 1997; Park et Katze, 1995). **Sans que les détails moléculaires en soient précisément connus, il est raisonnable de penser que cette activité repose sur une interaction de NS1 avec la séquence virale AGCAAAAG**, comme le suggèrent les travaux menés par Robb et Fodor (2012). Cette interaction, qui pourrait faciliter l'export nucléo-cytoplasmique des ARNm viraux, favoriserait aussi l'initiation de leur traduction, notamment en recrutant les facteurs eIF4GI et PABPI, lesquels à leur tour induisent le recrutement de la petite sous-unité 40S du ribosome (Burgui et al., 2003). En absence de cette séquence virale, NS1 pourrait ne plus se lier (ou que très faiblement) aux transcrits viraux au cours du cycle viral mais cela reste encore à démontrer.

(iii) Le RBD de NS1 forme des complexes de haute affinité avec les ARN double-brin mais cette interaction est conditionnée par plusieurs paramètres. La méthode SELEX (Marc *et al*., 2013) a permis d'identifier plusieurs familles d'aptamères de haute affinité pour la protéine NS1, en particulier les aptamères DM01 et DM03 qui sont caractérisés par la présence de séquences d'intérêts (les séquences virales AGCAAAAG (motif A, **Fig.49**) et UGAUUGAAG (motif B, **Fig.49** et **Fig.50**) ainsi que le motif GUAAC) et par un Kd de l'ordre du nanomolaire. Avant de pouvoir réaliser les différents

essais d'interaction protéine/ARN, nous avons dû produire et purifier toutes les protéines RBD d'intérêts. Ces dernières étaient globalement bien surexprimées et relativement solubles lors de la purification. De façon surprenante, des quantités de protéines plus importantes ont été purifiées pour les double-mutants R38A-K41A. C'est peut-être pour cette raison que Carrillo *et al*. (2014) avaient choisi de produire et de purifier ce même double-mutant pour la protéine NS1 entière d'un virus aviaire H6N6 pour en réaliser la structure tridimensionnelle. (1) **La formation de complexes NS1/ARN de haute affinité dépend directement de la longueur de l'ARN**. Nous avons décliné à plusieurs reprises ces deux aptamères en aptamères de plus petite taille afin (i) d'identifier les déterminants essentiels à l'interaction et (ii) une fois avoir obtenu un ARN de longueur minimale qui conserve encore une bonne affinité pour la protéine de l'utiliser pour la résolution de nouvelles structures moléculaires du complexe RBD/ARNdb. Les différents essais d'interaction effectués en gel retard nous ont permis de montrer que NS1, *via* son RBD, interagissait fortement avec des ARN lorsque ces derniers présentent une structure double-brin d'au moins 18 nt. A l'inverse, un aptamère de longueur inférieure entraine une diminution de la stabilité et de l'affinité des complexes. (2) **Certaines séquences dans l'ARN semblent participer de façon coopérative à la formation de complexes NS1/ARN de haute affinité**. La déclinaison de ces deux aptamères, qui contiennent respectivement les séquences AGCAAAAG ou UGAUUGAAG dans la partie tige-boucle de leur structure, a permis de mettre en évidence la formation de complexes NS1/ARN qui se distinguent d'un ARN à l'autre. Dans nos conditions expérimentales, le RBD de NS1 forme deux complexes de haute affinité avec l'aptamère DM01-midi, suggérant une interaction coopérative avec le motif AGCAAAAG. D'ailleurs, et même s'il n'est pas certain que la séquence virale AGCAAAAG contenue dans notre aptamère présente une structure similaire à celle retrouvée dans les ARNm viraux (+), son absence dans l'ARN provoque la disparition de ce second complexe en gel retard, confortant l'hypothèse selon laquelle le RBD de NS1 interagit spécifiquement avec cette séquence virale. Nous avons également montré que ce second complexe était formé avec toutes les protéines purifiées dans notre étude, y compris avec la protéine issue du virus de chauve-souris H17N10 dont la séquence peptidique est significativement différente des autres. Ceci suggère que NS1 interagit spécifiquement avec cette séquence contenue dans les ARNm viraux (+) lors du cycle viral. Néanmoins, la seule présence de cette séquence dans l'ARN ne suffit pas pour former des complexes de haute affinité avec NS1. Nous avons également montré que la séquence virale UGAUUGAAG (présente dans l'aptamère DM03-midi), qui est très conservée dans la région 3'-UTR des ARNm codant pour NS1, ne semble pas impliquée dans la formation de complexes NS1/ARN auxiliaires, soulignant la non-relevance biologique de cette séquence dans l'interaction de NS1 avec l'ARN. Les essais biologiques que nous avons réalisé *in vitro* semblent aller dans le même sens, confirmant l'absence d'une reconnaissance spécifique de cette séquence ARN par NS1. Toutefois, aucune des deux séquences virales seules, y compris la séquence AGCAAAAG, ne suffit à la formation de complexes NS1/ARN, suggérant que ni l'une ni l'autre ne suffit à la formation de complexes avec NS1. (3) **La présence du motif GUAAC dans l'ARN est un élément essentiel à l'obtention de complexes de haute affinité**.

Les aptamères DM03 et DM01 obtenus *in vitro* présentent respectivement un et deux motifs double-brin GUAAC dans leur séquence nucléotidique. Nous avons montré que cette séquence, qui pourrait correspondre à un site canonique donneur d'épissage des ARNm (Marc et al., 2013), est suffisant pour la formation de complexes NS1/ARN stables et de haute affinité. Nos essais d'interaction en gel retard nous ont permis de montrer qu'un seul motif GUAAC était nécessaire et suffisant pour permettre forte interaction du RBD de NS1 à l'ARN à partir du moment où ce motif reste « accessible » à la protéine en étant situé à la position distale (le plus éloigné de la partie tige-boucle de l'ARN). Dans les structures d'ARN en tige-boucle, la localisation du motif GUAAC est un paramètre déterminant dans l'interaction. Arias-Mireles *et al*. (2018) ont également montré que NS1 interagit plus fortement avec des ARN riches en bases Gu/Cy. La plupart de nos aptamères et ARNdb sont composés par de nombreuses bases Gu/Cy ce qui pourrait contribuer à une meilleure interaction avec NS1. Le motif GUAAC identifié *in vitro* pourrait correspondre au site donneur d'épissage des pré-ARNm lorsque l'ARN U1 est apparié à la séquence de l'intron (Kondo *et al*., 2015; Plaschka *et al*., 2019). Plusieurs travaux antérieurs ont permis d'identifier les mécanismes mis en jeu par NS1 pour bloquer partiellement voire totalement la maturation des pré-ARNm cellulaires. L'interaction de NS1 avec les complexes snRNP U6-U2 et U6-U4 du spliceosome (Qiu et al., 1995) entraine un blocage de l'épissage des pré-ARNm cellulaires qui s'accumulent dès lors dans le noyau cellulaire. NS1 pourrait également entrer en compétition avec la protéine U1c qui se fixe au niveau du site donneur d'épissage lorsque l'ARN U1 est lié à la séquence de l'intron. Observée d'abord en solution (gel retard), la forte liaison de NS1 à l'ARN qui contient au moins un motif GUAAC pourrait avoir une pertinence biologique dans la cellule infectée. En effet, Zhang *et al*. (2018) ont montré par des essais de pontages covalents chimiques et physiques (UV) que NS1 interagissait avec près de 607 ARN dont 538 étaient des pré-ARNm. Après un enrichissement des ARN précipités, ils ont pu mettre en évidence que NS1 ciblait majoritairement les transcrits cellulaires codant pour RIG-I, soulignant les multiples mécanismes mis en jeu par NS1 pour limiter la réponse antivirale cellulaire. Ils ont aussi montré que NS1 se liait préférentiellement aux introns des pré-ARNm cellulaires. L'analyse des ARN liées à NS1 n'a cependant pas révélé de séquences similaires au motif GUAAC. (4) **Le domaine ED de NS1 semble participer de façon indirecte à l'interaction avec l'ARN**. Les conformations *ouverte*, *semi-ouverte* et *fermée* de NS1 sont liées à l'orientation relative de son domaine ED par rapport au domaine RBD (Carrillo *et al*., 2014; Kerry *et al*., 2011). Des contacts entre le domaine ED et l'ARN lié au RBD pourraient avoir lieu lorsque NS1 est dans une conformation *ouverte* en raison de la proximité des domaines protéiques. Nos essais d'interaction en gel retard avec la protéine de fusion GST-NS1 semblent indiquer que le domaine ED de NS1 participerait au moins indirectement à la fixation de l'ARN. Cependant et contre toute attente, le caractère coopératif du domaine ED est observé par la présence d'un second complexe stable uniquement avec l'aptamère DM03-midi, alors que ce même ARN ne formait qu'un seul complexe avec le RBD de NS1. Kerry *et al*. (2011) avaient auparavant montré que le domaine ED de NS1 pouvait contribuer à l'interaction RBD/ARNdb. En effet, certains mutants du domaine ED, en particulier W187A, sont connus pour abolir la dimérisation hélice-hélice de

ce domaine, réduisant ainsi fortement la fixation du RBD de NS1 à un ARNdb synthétique (polyI:C) (Kerry *et al*., 2011). Cependant et même si l'oligomérisation de NS1 pourrait assurer une plus forte interaction avec l'ARN afin d'empêcher ce dernier d'être détecter par les RLR cellulaires, il n'est pas encore acté que la dimérisation du domaine ED entraine de façon directe la stabilisation de l'interaction du RBD de NS1 avec l'ARN ou de façon indirecte l'oligomérisation de NS1 provoquant alors une meilleure interaction avec l'ARN (Jerry *et al*., 2011). L'inverse, à savoir que la liaison du RBD à l'ARN module l'interaction du domaine ED avec l'hélicase DHX30, a été décrit récemment par Chen *et al*. (2020). Carrillo *et al*. (2014) propose que les différentes conformations du domaine ED (**Fig.29**) puissent être considérées comme un domaine d'oligomérisation de NS1 sur l'ARN. Néanmoins avec des sondes d'ARN d'une vingtaine de nucléotides comme celles que nous avons conçues et évaluées, on voit mal comment d'après ce modèle un dimère de NS1 *via* le domaine ED pourrait se stabiliser. Un ARN beaucoup plus grand, ou tout du moins au moins deux fois plus grand, serait nécessaire pour vérifier cette hypothèse. La boucle (sa structure et/ou sa séquence) présente dans nos sondes pourrait également être impliquée dans le fait que la protéine de fusion GST-NS1 forme plus facilement du complexe C2' avec DM01-midi.

(iv) **L'interaction du RBD-B de NS1 avec la sonde AWFC01 induit une légère courbure de l'ARN**. L'objectif de la caractérisation des paramètres de l'ARN était de pouvoir concevoir une ou plusieurs sondes de longueur minimale pour la co-cristallographie qui présenteraient une meilleure affinité pour NS1 que les ARN utilisés précédemment (Cheng et al., 2009; Yin et al., 2007). (1) **L'ARNdb synthétique AWFC01 conçu au laboratoire est un ligand de haute affinité pour NS1**. Après avoir défini les paramètres d'interaction des complexes NS1/ARNdb, nous avons pu concevoir l'ARNdb palindromique AWFC01 qui contient deux motifs double-brin GUAAC bordant de part et d'autre sa séquence de 19 nt. Nous avons montré que le duplex AWFC01 présentait une des meilleures affinités entre les ARN évalués en gel retard (évaluées par compétition contre notre duplex synthétique). Cette forte affinité du RBD de NS1 de H7N1 est une caractéristique partagée par tous les RBD purifiés et testés au laboratoire (Kd d'environ 1 nmol.L^{-1}). Alors que le double-mutant R38A-K41A du RBD-A n'est pas complètement invalidé dans son interaction avec l'ARN (Kd d'environ 400 nmol.L^{-1}), le RBD-B mutant est très fortement impacté par la double-mutation (Kd d'au moins 50 µmol.L^{-1}). Cette dernière observation indique que la grande variabilité observée en particulier du RBD de NS1 selon son origine virale peut être un déterminant important dans l'interaction avec l'ARN. (2) **L'ARNdb AWFC01 est un bien meilleur ligand que l'ARNdb CHENG**. Etant impossible d'évaluer directement l'affinité d'interaction du RBD de NS1 pour l'ARNdb CHENG par des expériences de titrage en gel retard (car formant des complexes RBD/ARNdb instables en solution), nous avons dû faire cette estimation indirectement par compétition contre notre duplex AWFC01. Nous avons ainsi pu montrer que le RBD de NS1 est 50 à 90 fois plus affin pour l'ARNdb AWFC01 que pour l'ARNdb CHENG utilisé pour obtenir la structure cristallographique du complexe RBD/ARNdb. (3) **Les analyses structurales menées au laboratoire ont permis de montrer une légère courbure de l'ARNdb AWFC01 qui**

survient lors de l'interaction avec le RBD de NS1. En raison des différences d'affinité de NS1 pour les ARNdb AWFC01 et CHENG, nous avons envisagé que les modes d'interaction du RBD avec ces ARNs devaient être très différents. Pour essayer d'en établir les bases structurales, nous avons cristallisé et résolu les structures 3D du RBD-B de NS1 libre ou en complexe avec les deux duplexes d'ARN. L'analyse de nos structures 3D démontre une reconnaissance non-spécifique de l'ARN par le RBD de NS1, qui se fait majoritairement par la reconnaissance des fonctions 2'OH du ribose et *via* l'interaction avec les phosphates de l'ARN. Tout comme Cheng *et al*. (2009) et avant eux Chien *et al*. (2004) et Yin *et al*. (2007), les acides aminés basiques contenus dans les hélices α2 et α2' du dimère de RBD sont au cœur de l'interface RBD/ARN. Globalement, on retrouve les mêmes profils d'interactions directes et indirectes entre les différents complexes résolus au laboratoire mais aussi vis à vis du complexe résolu par Cheng *et al*. (2009), indiquant que le RBD-B de NS1 n'échappe pas à la règle. A titre d'exemple, les acides aminés R38 et D34 sont respectivement retrouvés au cœur des interactions directes et indirectes dans tous les complexes structuraux. Alors que tous nos résultats d'interactions RBD/ARN réalisées par gel retard tendaient à indiquer une reconnaissance spécifique par le RBD de NS1 des nucléobases de l'ARN, en particulier des motifs double-brin GUAAC dans AWFC01, il n'en est rien et ce résultat inexpliqué à ce jour reste remarquable. Cependant, l'analyse fine des paramètres hélicoïdaux de l'ARN en complexe avec la protéine a permis de montrer une légère courbure de AWFC01 lors de son interaction avec le RBD de NS1 (Wacquiez *et al*., 2020). (4) **Le RBD de NS1 contacte l'ARNdb de façon spécifique indépendamment de sa séquence et cette caractéristique est partagée par d'autres protéines**. La reconnaissance spécifique de l'ARN indépendamment de sa séquence nucléotidique par le RBD de NS1 a déjà été décrite pour d'autres protéines, y compris des protéines virales. La protéine C du virus *West Nile* (Ivanyi-Nagy *et al*., 2008), la protéine B2 du virus *Flock House* (Chao *et al*., 2005; Lingel *et al*., 2005), la protéine NS3 du virus du riz à feuilles blanches (Yang *et al*., 2011) ou encore la protéine VP35 du virus *Ebola* (Leung *et al*., 2010; Zinzula *et al*., 2012) partagent des caractéristiques communes avec la protéine NS1 des virus *influenza*. La présence de trois à quatre hélices α qui s'associent en un dimère fonctionnel, la présence d'une surface d'interaction avec l'ARN qui est composée majoritairement d'acides aminés basiques et la reconnaissance du squelette ribose-phosphate de l'ARNdb caractérisent toutes ces protéines. De façon similaire au RBD de NS1, le RBD de VP35 (acides aminés 221-340) est constituée de quatre hélices α (pour revue, Leung *et al*., 2010). De part ses nombreuses fonctions dans la cellule infectée, notamment en servant de cofacteur à la RdRp virale (Mühlberger *et al*., 1998) et en inhibant la réponse IFN (Hartman *et al*., 2004; Prins *et al*., 2009), le RBD de VP35 interagit avec des ARNdb viraux (Leung *et al*., 2010) par l'intermédiaire de nombreuses arginines et lysines (R305/K309/R312/K319/R322/K339). Il semblerait également que les RBD de NS1 et de VP35 partagent une homologie de séquences (RBD de VP35 : ^{304}P**R**AC**QKSLR**PV^{314}; RBD de NS1 : ^{36}L**R**RD**QKSLR**GR^{46}) (Krug, 2014). Certaines substitutions dans la séquence du RBD de VP35 (notamment R305A, K309A et K312A) réduisent fortement voire totalement l'interaction avec les ARNdb, soulignant ainsi leur importance dans l'interaction avec l'ARN (Zinzula *et al*., 2012). Cette

caractéristique est également partagée par le RBD de NS1 puisque des substitutions R38A, R38A-K41A et R35-R46 limite l'interaction avec l'ARN. Toutefois et contrairement au RBD de NS1, le RBD de VP35 reconnait les phosphates de part d'autre de l'ARN, y compris ceux situés aux extrémités ouvertes de l'ARN (Bale *et al*., 2013; Kimberlin *et al*., 2010).

(v) **Les limites expérimentales de notre étude**. (1) **SELEX**. Même si la technique SELEX a permis d'obtenir des aptamères d'ARN de haute affinité pour la protéine NS1, cette dernière reste tout du moins discutable dans sa conception. (A) Un mélange de protéines NS1 provenant d'un virus humain H5N1 (allèle A de NS1) et d'un virus aviaire H7N1 (allèle B de NS1) a été utilisé pour réaliser cette sélection d'aptamères de haute affinité pour la protéine de NS1. Sans éléments complémentaires à ce jour, il est probable que les deux aptamères DM01 et DM03 aient pu se lier à l'une et/ou à l'autre des deux protéines virales, en sachant que leurs séquences peptidiques (d'allèle A et B respectivement) étaient homologues à environ 80 % (Marc, 2014); (B) Les protéines NS1 recombinantes ont été purifiées en présence de 2 M d'urée afin de solubiliser un maximum de protéines présentes dans les corps d'inclusion bactériens. Toutefois, ce traitement chaotropique dénature fortement les protéines et, même si le dimère de RBD est relativement stable et compact et qu'il est susceptible de se replier correctement lors de sa purification, il n'est pas certain qu'il en soit ainsi pour le domaine ED qui est beaucoup flexible; (C) Les oligos choisis pour la SELEX (Mercey *et al*., 2006) sont à l'origine de l'introduction de la séquence GUAAC du premier motif GUAAC (position distale) dans les aptamères DM01 et DM03. (2) **Purification**. (A) Même si nous avons mis en évidence la formation de complexes supplémentaires avec l'ARN en présence du domaine ED de NS1, la persistance de l'étiquette GST pourrait limiter les changements de conformation de NS1, en particulier ceux du domaine ED. Si tel est le cas, le bon repliement de la NS1 entière dans une conformation *ouverte* pourrait être incomplet, limitant ainsi la participation du domaine ED de NS1 dans l'interaction avec l'ARN. (B) Nos protéines de fusion GST-RBD et GST-NS1 d'allèle B ont été purifiées en présence 1 % de triton X-100 (détergeant non-ionique) qui à cette concentration pourrait avoir également modifiée les conformations de NS1, en particulier celle du domaine ED de NS1. La technique de diffusion dynamique de la lumière (DLS), qui permet d'obtenir le rayon hydrodynamique des molécules en solution, pourrait permettre de comparer le repliement des protéines NS1 entières purifiées ou non en présence de triton X-100. Le dichroïsme circulaire (CD), qui permet d'obtenir des informations sur la structure secondaire des protéines, pourrait également être utile.

PERSPECTIVES

Plusieurs pistes mériteraient d'être explorées suite à ce travail

(i) On est en droit de se demander si la séquence virale AGCAAAAG qui interagit de façon coopérative avec le RBD de NS1 est directement reconnue par ce dernier. Bons nombres de résultats publiés auparavant (ainsi que nos propres observations) témoignent d'une interaction spécifique de NS1 avec cette séquence, et notamment au cours du cycle viral afin de moduler les étapes de réplication et de transcription du génome viral. Mais est-ce que le RBD de NS1 reconnaît directement les nucléobases de la séquence virale AGCAAAAG ? Si ce n'est pas le cas, de quelle nature est le complexe C2 formé entre le RBD et l'aptamère DM01-midi ? D'après les structures cristallographiques, la fixation d'une première molécule de RBD sur la partie double-brin de l'ARN s'accompagne d'une occupation de *quasi* 20 paires de bases. Le mode de fixation de la seconde molécule qui est probablement très différent du mode de fixation de la première est donc de nature inconnue à ce jour. Il peut impliquer en même temps des interactions avec la séquence virale et entre les 2 protéines de RBD. Résoudre la structure du complexe C2 avec une sonde ARN de type DM01-midi pourrait permettre de savoir si NS1 reconnaît spécifiquement les nucléobases de cette séquence virale. Dans l'éventualité où notre hypothèse viendrait à être confirmée, nous pourrions mettre en évidence un nouveau mode de reconnaissance des ARN par le RBD de NS1 (avec ou sans la participation du domaine ED de NS1) qui ne se limiterait pas seulement à une reconnaissance spécifique du squelette ribose-phosphate de l'ARN.

La déclinaison du ARNdb CHENG par ajout successif des motifs double-brin GUAAC dans sa séquence nucléotidique pourrait aussi permettre de vérifier si leur présence est associée à la formation de complexes RBD/ARNdb plus stables et plus affins.

Utiliser la technique de cross-link (Darnell *et al*., 2010) sur des extraits de cellules transfectées avec le *miniréplicon viral* ou infectées, pourrait permettre d'identifier de nouveaux ARNs qui se lieraient à NS1, et notamment des ARN viraux qui contiendraient la séquence virale AGCAAAAG dans leur séquence. C'est notamment avec cette technique que Zhang *et al*. (2018) ont pu montrer que la protéine NS1 ciblait des séquences d'introns de pré-ARNm cellulaires.

(ii) Anastasina *et al*. (2016) ont également montré que NS1 se liait directement à l'ADN de certains gènes impliqués dans la réponse antivirale, suggérant que dans nos conditions, NS1 pourrait limiter, en se liant à certaines séquences du vecteur transfecté, la formation des transcrits d'intérêts dans notre système *translation-only*. Même si cette hypothèse reste peu probable (Salvadore *et al*., 2002), la quantification dans notre système du nombre de transcrits produits en absence et en présence de NS1 pourrait permettre d'infirmer ou de confirmer cette dernière.

(iii) Nous avons montré que les complexes structuraux résolus au laboratoire étaient très proches de ceux existants. Afin de confirmer ce résultat, il est presque indispensable de tenter de résoudre la structure cristallographique d'un complexe associant le RBD-A à l'ARNdb AWFC01 car tous nos modèles structuraux ont tous été réalisés avec le RBD-B. Alors que la résolution d'un tel complexe par DRX ne semble pas être le plus judicieux pour apporter les réponses à nos interrogations, nous pourrions tenter d'analyser la structure de ces mêmes complexes par cryo-EM. Cette technologie pourrait

permettre de s'affranchir des conditions inhérentes à la cristallographie. De plus, elle pourrait permettre d'étudier les changements structuraux de ces complexes ayant lieu lors de l'interaction avec l'ARNdb alors que la DRX permet d'étudier des structures dans un état concentré et figé. Puisque l'analyse de nos structures à l'état figé reste limitée en terme de mise en évidence de nouveaux éléments structuraux à l'exception d'une légère courbure de l'ARN AWFC01 survenant lors de l'interaction avec le RBD de NS1, nous avons également démarré une approche de dynamique moléculaire dans laquelle des calculs d'énergie libre et d'enthalpie (ce second paramètre reflétant la force d'interaction) dans le but d'obtenir des informations théoriques sur les mécanismes mis en jeux par le RBD de NS1 lors des interactions avec nos différents ARN. S'ils donnent des résultats prometteurs, nos essais pourraient s'inscrire dans la continuité des travaux initiés par Pan *et al*. (2011) au cours desquels ces derniers avaient pu démontrer l'importance des acides aminés R35/R38/S42/R46/T49, déjà décrits auparavant (Cheng et al., 2009; Yin et al., 2007), dans l'interface d'interaction du RBD de NS1 avec l'ARNdb. Ils avaient également mis en évidence des différences d'interaction (enthalpie) entre plusieurs mutants (notamment les simples mutants R38A – S42A – T49A et le double mutant R35A-R46A), avec un impact plus fort pour le simple mutant R38A (très impliqué dans l'interaction avec l'ARN) et le double mutant R35A-R46A qui bloque la dimérisation du RBD de NS1 (Wang *et al*., 1999).

(iv) L'ARNdb AWFC01 est un ligand de très forte affinité pour tous les RBD purifiés au laboratoire (Kd d'environ 1 nmol.L^{-1}). Il pourrait potentiellement être utilisé comme molécule « poison » pour cibler spécifiquement le RBD de NS1 et enrailler l'infection virale. En réalisant une SELEX sur la protéine NS1 du virus Udorn, Woo *et al*. (2019) ont démontré l'efficacité de cette approche puisque leurs aptamères d'ARN réduisent significativement l'infection virale, entrainant ainsi un rétablissement du niveau d'expression d'IFN. Cette approche a aussi été utilisée dans le cadre d'autres infections virales (pour revue, Wandtke *et al*., 2015). En utilisant cette même approche, nous pourrions cibler le RBD de NS1 d'IAV et d'IBV en raison de leur conservation sur le plan structural (Yin *et al*., 2007). L'utilisation de AWFC01 comme molécule « poison » pourrait être une piste thérapeutique prometteuse. De plus et même si de telles affinités (Kd nanomolaire) restent néanmoins discutables dans la cellule infectée (d'après le nombre de ligands ARN potentiels qui se lient au RBD de NS1), AWFC01 pourrait permettre de séquestrer spécifiquement le RBD de NS1 et d'inhiber les fonctions de NS1 qui en résultent. L'interface d'interaction du RBD de NS1 avec l'ARNdb héberge la séquence NLS majeure de NS1 (localisée entre les acides aminés R35 à K41 du RBD, voir la section *Introduction*). Le blocage du RBD de NS1 par AWFC01 pourrait (A) empêcher son interaction avec les ARN viraux au cours de leur transcription et/ou leur réplication, (B) rétablir l'expression des gènes cellulaires qui est fortement impactée par NS1 (inhibition de la maturation des pré-ARNm cellulaires), (C) permettre leur détection par le système antiviral cellulaire et assurer l'activation d'une réponse antivirale effectrice (la réponse IFN étant la cible majeure de NS1) et enfin (D) réduire voire même empêcher totalement la multiplication du virus dans la cellule. Nous avons lancé des premiers essais d'inhibition de AWFC01 *in cellulo* dans le système *miniréplicon viral*. Dans ce système où NS1 stimule la transcription et/ou la

réplication des ARN viraux, nous avons cherché à mettre en évidence l'effet de notre duplex sur NS1. Dans l'éventualité d'un blocage même partiel du RBD, nous pourrions probablement visualiser des différences de stimulations du système par la RdRp virale. Nos premiers essais, qui comprenaient aussi l'ARNdb CHENG en tant que contrôle puisque moins affin pour la protéine, ne nous ont cependant pas permis de mettre en évidence des différences notables entre les conditions. Toutefois, nos conditions expérimentales manquent probablement d'optimisation avant d'évaluer un potentiel effet positif inhibiteur de AWFC01. Plusieurs biais expérimentaux peuvent en être à l'origine :

(E) Un minimum de 8 h est nécessaire aux cellules HEK293T transfectées avec le système *miniréplicon viral* afin de produire un signal Ren-Luc suffisamment élevé pour être détecté et quantifié (résultats non-montrés). Toutefois et même si nous choisissons de détecter l'activité Ren-Luc après 24 h, notre ARNdb est ajouté 8 h après la transfection du *miniréplicon viral*. Durant ce laps de temps (à savoir 16 h), la transfection d'ARN double-brin étrangers dans la cellule pourrait initier une réponse antivirale qui serait susceptible de limiter l'activité inhibitrice de notre duplex en séquestrant et/ou en dégradant ce dernier (voir les paragraphes qui concernent la réponse antivirale cellulaire dans la section *Introduction*).

(F) Les ARNdb AWFC01 et CHENG sont tous deux des palindromes, il est nécessaire de les hybrider le plus concentré possible (au moins 1 $mmol.L^{-1}$) tout en les conservant dans la glace pour empêcher la déshybridation et éviter la réaction monomoléculaire aboutissant à la formation de petits hairpins. Toutefois, une fois transfecté dans les cellules, notre duplex ARN pourrait se déshybrider en raison de la température d'incubation des cellules (37 °C).

(G) Même si la protéine NS1 est la protéine la plus abondamment produite après la protéine NP, on ignore la quantité exacte de NS1 qui est produite au cours du temps dans les cellules transfectées avec notre système *miniréplicon viral*. Dans l'idéal, un ARNdb compétiteur doit cibler un dimère de RBD. Toutefois, en raison de la réponse antivirale et des conditions d'incubation modifiant l'hybridation des ARN, il n'est pas certain que nos ARNdb transfectés puissent agir efficacement sur toutes les protéines NS1 produites. Dans l'éventualité où une partie, même infime, de NS1 resterait libre dans la cellule, on observerait encore l'activité stimulatrice de NS1 sur le *miniréplicon viral*.

(H) Dans l'objectif d'optimiser notre molécule poison, nous pourrions y ajouter un hairpin pour stabiliser notre duplex (une tétraboucle GAGA par exemple puisque sa présence ne perturbe pas l'affinité d'interaction avec le RBD de NS1) et éviter la formation de structures monobrins d'ARN (puisque notre duplex est un palindrome). L'addition d'un fluorophore en 5' ou en 3' de l'ARNdb pourrait apporter plusieurs caractéristiques intéressantes comme son suivi et sa localisation dans la cellule transfectée et infectée (témoignant ainsi de l'entrée de notre ARN) et son interaction avec NS1. Déjà utilisée par Woo *et al*. (2019), cette modification pourrait également apporter une meilleure protection de l'ARNdb contre sa dégradation par des exonucléases cellulaires 5'→3' et 3'→5'.

(v) Une grande partie de mon travail de thèse aura été de résoudre la structure de nouveaux complexes NS1/ARNdb renouvelant les connaissances actuelles sur le sujet, l'objectif *in fine* étant de

développer une approche *drug design* ciblant spécifiquement la poche située à l'interface RBD/ARN. Si l'approche, initiée par Cheng *et al*. (2009) puis par Engel (2013) et plus récemment par Abi Hussein *et al*. (2020) se poursuit, elle pourrait être réalisée à la fois sur le complexe détaillé par Cheng et al. (2009) mais aussi sur les nôtres en raison de leur similitude. Qu'ils soient structuraux ou non, tous les travaux réalisés jusque-là sur l'interaction de NS1 avec l'ARN ont permis de montrer l'importance des acides aminés R35/R38/K41 et R46. Le RBD a été décrit comme essentiel aux fonctions de NS1 dans la cellule infectée. En effet, la délétion complète du domaine ED de NS1 impacte moins fortement le cycle viral que celle du RBD (Trapp *et al*., 2018 ; **Annexe 14**). Cibler le RBD, plutôt que le domaine ED, semble donc une approche rationnelle. *Via* une approche de *docking* virtuel, Chandra *et al*. (2017) ont identifié des inhibiteurs potentiels susceptibles d'interagir avec la poche du RBD, en ciblant préférentiellement les acides aminés R35 et R38.

BIBLIOGRAPHIE

A

Abril, J. F., R. Castelo and R. Guigo (2005). "Comparison of splice sites in mammals and chicken." Genome Res 15(1): 111-119.

Adams, P.D., P.V. Afonine, G. Bunkóczi, V.B. Chen, I.W. Davis, N. Echols, J.J. Headd, L.W. Hung, G.J. Kapral, R.W.Grosse-Kunstleve, A.J. McCoy, N.W. Moriarty, R. Oeffner, R.J. Read, D.C. Richardson, J.S. Richardson, T.C. Terwilligera and P.H. Zwart (2010). "PHENIX: A Comprehensive Python-based System for Macromolecular Structure Solution." Acta Crystallogr D Biol Crystallogr 66(Pt 2):213-21.

Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H. and P.D. Adams (2012). "Towards automated crystallographic structure refinement with phenix.refine." Acta Crystallogr D Biol Crystallogr 68: 352-67.

Agris, C. H., M. E. Nemeroff and R. M. Krug (1989). "A block in mammalian splicing occurring after formation of large complexes containing U1, U2, U4, U5, and U6 small nuclear ribonucleoproteins." Mol Cell Biol 9(1): 259-267.

Akarsu, H., W. P. Burmeister, C. Petosa, I. Petit, C. W. Muller, R. W. Ruigrok and F. Baudin (2003). "Crystal structure of the M1 protein-binding domain of the influenza A virus nuclear export protein (NEP/NS2)." EMBO J 22(18): 4646-4655.

Alenquer, M., S. Vale-Costa., T.A. Etibor., F. Ferreira., A.L. Sousa and M.J. Amorim (2019). "Influenza A virus ribonucleoproteins form liquid organelles at endoplasmic reticulum exit sites." Nat Commun 10(1): 1629.

Alonso-Caplen, F. V. and R. M. Krug (1991). "Regulation of the extent of splicing of influenza virus NS1 mRNA: role of the rates of splicing and of the nucleocytoplasmic transport of NS1 mRNA." Mol Cell Biol 11(2): 1092-1098.

Alonso-Caplen, F. V., M. E. Nemeroff, Y. Qiu and R. M. Krug (1992). "Nucleocytoplasmic transport: the influenza virus NS1 protein regulates the transport of spliced NS2 mRNA and its precursor NS1 mRNA." Genes Dev 6(2): 255-267.

Amorim, M. J. (2018). "A Comprehensive Review on the Interaction Between the Host GTPase Rab11 and Influenza A Virus." Front Cell Dev Biol 6: 176.

Amorim, M. J., E. A. Bruce, E. K. Read, A. Foeglein, R. Mahen, A. D. Stuart and P. Digard (2011). "A Rab11- and microtubule-dependent mechanism for cytoplasmic transport of influenza A virus viral RNA." J Virol 85(9): 4143-4156.

Anastasina, M., I. Terenin, S. J. Butcher and D. E. Kainov (2014). "A technique to increase protein yield in a rabbit reticulocyte lysate translation system." Biotechniques 56(1): 36-9.

Anastasina, M., N. Le May, A. Bugai, Y. Fu, S. Söderholm, L. Gaelings, T. Ohman, J. Tynell, S. Kyttänen, M. Barboric, T. A. Nyman, S. Matikainen, I. Julkunen, S. J. Butcher, J. M. Egly and D. E. Kainov (2016). "Influenza virus NS1 protein binds cellular DNA to block transcription of antiviral genes." Biochim Biophys Acta 1859(11): 1440-1448.

Antao, V. P., S. Y. Lai and I Jr. Tinoco (1991)." A thermodynamic study of unusually stable RNA and DNA hairpin." Nucleic Acids Res. 19(21): 5901-5.

Aragón, T., S. De la Luna, I. Novoa, L. Carrasco, J. Ortin and A. Nieto (2000). "Eukaryotic translation initiation factor 4GI is a cellular target for NS1 protein, a translational activator of influenza virus." Mol Cell Biol 20(17): 6259-6268.

Aragón, T., S. De la Luna, I. Novoa, L. Carrasco, J. Ortín and A. Nieto (2000). "Eukaryotic translation initiation factor 4GI is a cellular target for NS1 protein, a translational activator of influenza virus." Mol Cell Biol. 20(17): 6259-68.

Arias-Mireles, B. H., C. M. de Rozieres, K. Ly and S. Joseph (2018). "RNA Modulates the Interaction between Influenza A Virus NS1 and Human PABP1." Biochemistry 57(26): 3590-3598.

Arranz, R., R. Coloma, F. J. Chichon, J. J. Conesa, J. L. Carrascosa, J. M. Valpuesta, J. Ortin and J. Martin-Benito (2012). "The structure of native influenza virion ribonucleoproteins." Science 338(6114): 1634-1637.

Arzt, S., I. Petit, W. P. Burmeister, R. W. Ruigrok and F. Baudin (2004). "Structure of a knockout mutant of influenza virus M1 protein that has altered activities in membrane binding, oligomerisation and binding to NEP (NS2)." Virus Res 99(2): 115-119.

Asha, K. and B. Kumar (2019). "Emerging influenza D virus threat: What we know so far!" J Clin Med 8(2). Pii: E192.

Auweter, S. D., Oberstrass, F. C. and F. H. Allain (2006). "Sequence-specific binding of single-stranded RNA: is there a code for recognition?" Nucleic Acids Res. 34(17): 4943-59.

Avilov, S. V., D. Moisy, N. Naffakh and S. Cusack (2012). "Influenza A virus progeny vRNP trafficking in live infected cells studied with the virus-encoded fluorescently tagged PB2 protein." Vaccine 30(51): 7411-7417.

Avnir, Y., Tallarico, A. S., Zhu, Q., Bennett, A. S., Connelly, G., Sheehan, J., Sui, J., Fahmy, A., Huang, C. Y., Cadwell, G., Bankston, L. A., McGuire, A. T., Stamatatos, L., Wagner, G., Liddington, R. C. and W. A. Marasco (2014). "Molecular signatures of hemagglutinin stem-directed heterosubtypic human neutralizing antibodies against influenza A virus." PLoS Pathog 10(5):e1004103.

B

Badham, M. D. and J. S. Rossman (2016). "Filamentous Influenza Viruses." Curr Clin Microbiol Rep 3(3): 155-161.

Bale, S., J. P. Julien, Z. A. Bornholdt, A. S. Krois, I. A. Wilson and E. O. Saphire (2013). "Ebolavirus VP35 coats the backbone of double-stranded RNA for interferon antagonism." J Virol 87(18): 10385-10388.

Baltz, A. G., M. Munschauer, B. Schwanhausser, A. Vasile, Y. Murakawa, M. Schueler, N. Youngs, D. Penfold-Brown, K. Drew, M. Milek, E. Wyler, R. Bonneau, M. Selbach, C. Dieterich and M. Landthaler (2012). "The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts." Mol Cell 46(5): 674-690.

Banerjee, S. and P. Barraud (2014). "Functions of double-stranded RNA-binding domains in nucleocytoplasmic transport." RNA Biol 11(10): 1226-1232.

Bangaru, S., S. Lang, M. Schotsaert, H. A. Vanderven, X. Zhu, N. Kose, R. Bombardi, J. A. Finn, S. J. Kent, P. Gilchuk, I. Gilchuk, H. L. Turner, A. Garcia-Sastre, S. Li, A. B. Ward, I. A. Wilson and J. E. Crowe, Jr. (2019). "A Site of Vulnerability on the Influenza Virus Hemagglutinin Head Domain Trimer Interface." Cell 177(5): 1136-1152 e1118.

Barman, S., L. You, R. Chen, V. Codrea, G. Kago, R. Edupuganti, J. Robertus, R. M. Krug and E. V. Anslyn (2014). "Exploring naphthyl-carbohydrazides as inhibitors of influenza A viruses." Eur J Med Chem 71: 81-90.

Basler, C. F. and P. V. Aguilar (2008). "Progress in identifying virulence determinants of the 1918 H1N1 and the Southeast Asian H5N1 influenza A viruses." Antiviral Res 79(3): 166-178.

Basu, D., M. P. Walkiewicz, M. Frieman, R. S. Baric, D. T. Auble and D. A. Engel (2009). "Novel influenza virus NS1 antagonists block replication and restore innate immune function." J Virol 83(4): 1881-1891.

Batool, A., S. Aashaq and K. I. Andrabi (2019). "Eukaryotic initiation factor 4E (eIF4E): A recap of the cap-binding protein." J Cell Biochem 120(9): 14201-14212.

Bauer, D. L. V., M. Tellier, M. Martinez-Alonso, T. Nojima, N. J. Proudfoot, S. Murphy and E. Fodor (2018). "Influenza Virus Mounts a Two-Pronged Attack on Host RNA Polymerase II Transcription." Cell Rep 23(7): 2119-2129 e2113.

Bilanges, B., Y. Posor and B. Vanhaesebroeck (2019). "PI3K isoforms in cell signalling and vesicle trafficking." Nat Rev Mol Cell Biol 20(9): 515-534.

Blanchet, C., M. Pasi, K. Zakrzewska and R. Lavery (2011). « CURVES+ web server for analyzing and visualizing the helical, backbone and groove parameters of nucleic acid structures.” Nucleic Acides Res. 39, W68-73.

Blijleven, J. S., S. Boonstra, P. R. Onck, E. van der Giessen and A. M. van Oijen (2016). "Mechanisms of influenza viral membrane fusion." Semin Cell Dev Biol 60: 78-88.

Bobone, S., M. Hilsch, J. Storm, V. Dunsing, A. Herrmann and S. Chiantia (2017). "Phosphatidylserine Lateral Organization Influences the Interaction of Influenza Virus Matrix Protein 1 with Lipid Membranes." J Virol 91(12).

Bolte, H., Rosu, M.E., Hagelauer, E., García-Sastre, A. and M. Schwemmle (2019). “Packaging of the Influenza Virus Genome Is Governed by a Plastic Network of RNA- and Nucleoprotein-Mediated Interactions.” J Virol. 93(4). pii: e01861-18. doi: 10.1128/JVI.01861-18.

Bornholdt, Z. A. and B. V. Prasad (2006). "X-ray structure of influenza virus NS1 effector domain." Nat Struct Mol Biol 13(6): 559-560.

Bornholdt, Z. A. and B. V. Prasad (2008). "X-ray structure of NS1 from a highly pathogenic H5N1 influenza virus." Nature 456(7224): 985-988.

Brass, A.L., I.C. Huang, Y. Benita, S.P. John, M.N. Krishnan, E.M. Feeley, B. Ryan, J.L. Weyer, L. van der Weyden, E. Fikrig, D.J. Adams, R.J. Xavier, M. Farzan and S.J. Elledge (2009). “IFITM Proteins Mediate the Innate Immune Response to Influenza A H1N1 Virus, West Nile Virus and Dengue Virus.” Cell 139(7): 1243–1254.

Brown, R. S. (2005). "Zinc finger proteins: getting a grip on RNA." Curr Opin Struct Biol 15(1): 94-98.

Bruce, E. A., P. Digard and A. D. Stuart (2010). "The Rab11 pathway is required for influenza A virus budding and filament formation." J Virol 84(12): 5848-5859.

Bruce, E. A., A. Stuart, M. W. McCaffrey and P. Digard (2012). "Role of the Rab11 pathway in negative-strand virus assembly." Biochem Soc Trans 40(6): 1409-1415.

Bucher, E., Hemmes, H., de Haan, P., Goldbach, R. and M. Prins (2004). “The influenza A virus NS1 protein binds small interfering RNAs and suppresses RNA silencing in plants.” J Gen Virol 85: 983-991.

Burger, K. and M. Gullerova (2015). "Swiss army knives: non-canonical functions of nuclear Drosha and Dicer." Nat Rev Mol Cell Biol 16(7): 417-430.

Burgui, I., T. Aragón, J. Ortin and A. Nieto (2003). "PABP1 and eIF4GI associate with influenza virus NS1 protein in viral mRNA translation initiation complexes." J Gen Virol 84(Pt 12): 3263-3274.

Burleigh, L. M., L. J. Calder, J. J. Skehel and D. A. Steinhauer (2005). "Influenza A viruses with mutations in the M1 helix six domain display a wide variety of morphological phenotypes." J Virol 79(2): 1262-1270.

Byrd-Leotis, L., R. D. Cummings and D. A. Steinhauer (2017). "The Interplay between the Host Receptor and Influenza Virus Hemagglutinin and Neuraminidase." Int J Mol Sci 18(7).
Byrd-Leotis, L., N. Jia, S. Dutta, J. F. Trost, C. Gao, S. F. Cummings, T. Braulke, S. Muller-Loennies, J. Heimburg-Molinaro, D. A. Steinhauer and R. D. Cummings (2019). "Influenza binds phosphorylated glycans from human lung." Sci Adv 5(2): eaav2554.

C

Cady, S. D., W. Luo, F. Hu and M. Hong (2009). "Structure and function of the influenza A M2 proton channel." Biochemistry 48(31): 7356-7364.
Carlson, C. B., O. M. Stephens and P. A. Beal (2003). "Recognition of double-stranded RNA by proteins and small molecules." Biopolymers 70(1): 86-102.
Carrillo, B., J. M. Choi, Z. A. Bornholdt, B. Sankaran, A. P. Rice and B. V. Prasad (2014). "The influenza A virus protein NS1 displays structural polymorphism." J Virol 88(8): 4113-4122.
Chandra, I., Behera, A. and S. S. Cherian (2017). "Identification of Potential Inhibitors against the Human Influenza A Virus Targeting the CPSF30 and RNA Binding Domains of the NS1 Protein: An E-Pharmacophore approach." Indian Journal of Pharmaceutical Education and Research doi:10.5530/ijper.51.1.5
Chang, K. Y. and A. Ramos (2005). "The double-stranded RNA-binding motif, a versatile macromolecular docking platform." FEBS J 272(9): 2109-2117.
Chao, J. A., J. H. Lee, B. R. Chapados, E. W. Debler, A. Schneemann and J. R. Williamson (2005). "Dual modes of RNA-silencing suppression by Flock House virus protein B2." Nat Struct Mol Biol 12(11): 952-957.
Chen, B. J., G. P. Leser, D. Jackson and R. A. Lamb (2008). "The influenza virus M2 protein cytoplasmic tail interacts with the M1 protein and influences virus assembly at the site of virus budding." J Virol 82(20): 10059-10070.
Chen, C., W. Fan, J. Li, W. Zheng, S. Zhang, L. Yang, D. Liu, W. Liu and L. Sun (2018). "A Promising IFN-Deficient System to Manufacture IFN-Sensitive Influenza Vaccine Virus." Front Cell Infect Microbiol 8: 127.
Chen, C. J., M. E. Ermler, G. S. Tan, F. Krammer, P. Palese and R. Hai (2016). "Influenza A Viruses Expressing Intra- or Intergroup Chimeric Hemagglutinins." J Virol 90(7): 3789-3793.
Chen, G., L.C. Ma, S. Wang, R.L. Woltz, E.M. Grasso, G.T. Montelione and R.M. Krug (2020). "A double-stranded RNA platform is required for the interaction between a host restriction factor and the NS1 protein of influenza A virus." Nucleic Acids Research 48(1):304-315.
Chen, K. Y, E. D. Santos Afonso, V. Enouf, C. Isel and N. Naffakh (2019). "Influenza virus polymerase subunits co-evolve to ensure proper levels of dimerization of the heterotrimer." PLoS Pathog 15(10): e1008034.
Chen, S., X. Miao, D. Huandfu, X. Zhao, M. Zhang, T. Qin, D. Peng and X. Liu (2020). "H5N1 avian influenza virus without 80-84 amino acid deletion t the NS1 protein hijacks the innate immune system of dendritic cells for an enhanced mammalian pathogenicity." Transbound Emerg Dis.
Cheng, A., S. M. Wong and Y. A. Yuan (2009). "Structural basis for dsRNA recognition by NS1 protein of influenza A virus." Cell Res 19(2): 187-195.
Cheng, Y. Y., S. R. Yang, Y. T. Wang, Y. H. Lin and C. J. Chen (2017). "Amino Acid Residues 68-71 Contribute to Influenza A Virus PB1-F2 Protein Stability and Functions." Front Microbiol 8: 692.
Cheong, H. K., C. Cheong and B. S. Choi (1996). "Secondary structure of the panhandle RNA of influenza virus A studied by NMR spectroscopy." Nucleic Acids Res 24(21): 4197-4201.
Cheung, P. P., S. J. Watson, K. T. Choy, S. Fun Sia, D. D. Wong, L. L. Poon, P. Kellam, Y. Guan, J. S. Malik Peiris and H. L. Yen (2014). "Generation and characterization of influenza A viruses with altered polymerase fidelity." Nat Commun 5: 4794.
Chien, C. Y., R. Tejero, Y. Huang, E. Zimmerman, C. B. Ríos, R. M. Krug and G. T. Montelione (1997). "A novel RNA-binding motif in influenza A virus non-structural protein 1." Nat StructBiol 4(11): 891-5.
Chien, C. Y., Y. Xu, R. Xiao, J. M. Aramini, P. V. Sahasrabudhe, R. M. Krug and G. T. Montelione (2004). "Biophysical characterization of the complex between double-stranded RNA and the N-terminal domain of the NS1 protein from influenza A virus: evidence for a novel RNA-binding mode." Biochemistry 43(7): 1950-62.
Cho, E. J., S. Xia, L. C. Ma, J. Robertus, R. M. Krug, E. V. Anslyn, G. T. Montelione and A. D. Ellington (2012). "Identification of influenza virus inhibitors targeting NS1A utilizing fluorescence polarization-based high-throughput assay." J Biomol Screen 17(4): 448-459.
Clark, M. P., M. W. Ledeboer, I. Davies, R. A. Byrn, S. M. Jones, E. Perola, A. Tsai, M. Jacobs, K. Nti-Addae, U. K. Bandarage, M. J. Boyd, R. S. Bethiel, J. J. Court, H. Deng, J. P. Duffy, W. A. Dorsch, L. J. Farmer, H. Gao, W. Gu, K. Jackson, D. H. Jacobs, J. M. Kennedy, B. Ledford, J. Liang, F. Maltais, M. Murcko, T. Wang, M. W. Wannamaker, H. B. Bennett, J. R. Leeman, C. McNeil, W. P. Taylor, C. Memmott, M. Jiang, R. Rijnbrand, C. Bral, U. Germann, A. Nezami, Y. Zhang, F. G. Salituro, Y. L. Bennani and P. S. Charifson (2014). "Discovery of a novel, first-in-class, orally bioavailable azaindole inhibitor (VX-787) of influenza PB2." J Med Chem 57(15): 6668-6678.

Clerici, M., M. Faini, L. M. Muckenfuss, R. Aebersold and M. Jinek (2018). "Structural basis of AAUAAA polyadenylation signal recognition by the human CPSF complex." Nat Struct Mol Biol 25(4): 355.
Clifford, M., J. Twigg and C. Upton (2009). "Evidence for a novel gene associated with human influenza A viruses." Virol J 6: 198.
Coloma, R., J. M. Valpuesta, R. Arranz, J. L. Carrascosa, J. Ortin and J. Martin-Benito (2009). "The structure of a biologically active influenza virus ribonucleoprotein complex." PLoS Pathog 5(6): e1000491.
Cordier, C., L. Marcourt, M. Petitjean and G. Dodin (1999). "Conformational variation of the central CG site in d(ATGACGTCAT)2 and d(GAAAACGTTTTC)2. An NMR, molecular modelling and 3D-homology investigation." Eur J Biochem 261(3): 722-733.
Cordin, O., J. Banroques, N. K. Tanner and P. Linder (2006). "The DEAD-box protein family of RNA helicases." Gene 367: 17-37.
Crépin, T., A. Dias, A. Palencia, C. Swale, S. Cusack and R.W.H. Ruigrok (2010). "Mutational and metal binding analysis of the endonuclease domain of the influenza virus polymerase PA subunit." J Virol 84(18): 9096-104.
Crescenzo-Chaigne, B., C. Barbezange and S. van der Werf (2013). "The panhandle formed by influenza A and C virus NS non-coding regions determines NS segment expression." PLoS One 8(11): e81550.
Cruz-Rodz, A. L. and M. S. Gilmore (1990). "High efficiency introduction of plasmid DNA into glycine treated Enterococcus faecalis by electroporation." Mol Gen Genet 224(1): 152-4.

D

D'Acquisto, F. and S. Ghosh (2001). "PACT and PKR: turning on NF-kappa B in the absence of virus." Sci STKE 2001(89): re1.
Dadonaite, B., Vijayakrishnan, S., Fodor, E., Bhella, D. and E.C Hutchinson (2016). "Filamentous influenza viruses." J Gen Virol 97(8): 1755-1764.
Dankar, S. K., E. Miranda, N. E. Forbes, M. Pelchat, A. Tavassoli, M. Selman, J. Ping, J. Jia, and E. G. Brown (2013). "Influenza A/Hong Kong/156/1997(H5N1) virus NS1 gene mutations F103L and M106I both increase IFN antagonism, virulence and cytoplasmic localization but differ in binding to RIG-I and CPSF30." Virol J 10: 243.
Dar, A. C., T. E. Dever and F. Sicheri (2005). "Higher-order substrate recognition of eIF2alpha by the RNA-dependent protein kinase PKR." Cell 122(6): 887-900.
Darnell, R. B. (2010). "HITS-CLIP: panoramic views of protein-RNA regulation in living cells." Wiley Interdiscip Rev RNA 1(2): 266-286.
Das, K., L. C. Ma, R. Xiao, B. Radvansky, J. Aramini, L. Zhao, J. Marklund, R. L. Kuo, K. Y. Twu, E. Arnold, R. M. Krug and G. T. Montelione (2008). "Structural basis for suppression of a host antiviral response by influenza A virus." Proc Natl Acad Sci U S A 105(35): 13093-13098.
De Castro-Martin, I.F., G. Fournier., M. Sachse., J. Pizarro-Cerda., C. Risco and N. Naffakh (2017). "Influenza virus genome reaches the plasma membrane via a modified endoplasmic reticulum and Rab11-dependent vesicles." Nat Commun 8(1): 1396.
De Chassey, B., A. Aublin-Gex, A. Ruggieri, L. Meyniel-Schicklin, F. Pradezynski, N. Davoust, T. Chantier, L. Tafforeau, P. E. Mangeot, C. Ciancia, L. Perrin-Cocon, R. Bartenschlager, P. Andre and V. Lotteau (2013). "The interactomes of influenza virus NS1 and NS2 proteins identify new host factors and provide insights for ADAR1 playing a supportive role in virus replication." PLoS Pathog 9(7): e1003440.
De la Luna, S., P. Fortes, A. Beloso and J. Ortin (1995). "Influenza virus NS1 protein enhances the rate of translation initiation of viral mRNAs." J Virol 69(4): 2427-2433.
De Rozière, C.M. and S. Joseph (2020). "Influenza A virus NS1 protein binds as a dimer to RNA-free PABP1 but not to the PABP1-poly(A) RNA complex." Biochemistry.
De Vlugt, C., D. Sikora and M. Pelchat (2018). "Insight into Influenza: A Virus Cap-Snatching." Viruses 10(11).
Delaforge, E., S. Milles, G. Bouvignies, D. Bouvier, S. Boivin, N. Salvi, D. Maurin, A. Martel, A. Round, E. A. Lemke, M. R. Jensen, D. J. Hart and M. Blackledge (2015). "Large-Scale Conformational Dynamics Control H5N1 Influenza Polymerase PB2 Binding to Importin alpha." J Am Chem Soc 137(48): 15122-15134.
Delgadillo, M. O., Sáenz, P., Salvador, B., García, J. A. and C. Simón-Mateo (2004). "Human influenza virus NS1 protein enhances viral pathogenicity and acts as an RNA silencing suppressor in plants." J Gen Virol 85: 993-999.
Dias, A., D. Bouvier, T. Crepin, A. A. McCarthy, D. J. Hart, F. Baudin, S. Cusack and R. W. Ruigrok (2009). "The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit." Nature 458(7240): 914-918.
DiazGranados, C.A., A.J. Dunning, M. Kimmel, D. Kirby, J. Treanor, A. Collins, R. Pollak, J. Christoff, J. Earl, V. Landolfi, E. Martin, S. Gurunathan, R. Nathan, D.P. Greenberg, N.G. Tornieporth, M.D. Decker and H.K. Talbot (2014). "Efficacy of high-dose versus strandart-dose influenza vaccine in older adults." N Engl J Med 371(7): 635-45.
Dickerson, R. E. (1989). "Definitions and nomenclature of nucleic acid structure parameters." J Biomol Struct Dyn 6(4): 627-634.

Digard, P., D. Elton, K. Bishop, E. Medcalf, A. Weeds and B. Pope (1999). "Modulation of nuclear localization of the influenza virus nucleoprotein through interaction with actin filaments." J Virol 73(3): 2222-2231.
Donnelly, R. P. and S. V. Kotenko (2010). "Interferon-lambda: a new addition to an old family." J Interferon Cytokine Res 30(8): 555-564.
Dou, D., I. Hernandez-Neuta, H. Wang, H. Ostbye, X. Qian, S. Thiele, P. Resa-Infante, N. M. Kouassi, V. Sender, K. Hentrich, P. Mellroth, B. Henriques-Normark, G. Gabriel, M. Nilsson and R. Daniels (2017). "Analysis of IAV Replication and Co-infection Dynamics by a Versatile RNA Viral Genome Labeling Method." Cell Rep 20(1): 251-263.
Dou, D., R. Revol, H. Östbye, H. Wang and R. Daniels (2018). "Influenza A virus cell entry, replication, virion assembly and movement." Front Immunol 9: 1581.
Drappier, M. and T. Michiels (2015). "Inhibition of the OAS/RNase L pathway by viruses." Curr Opin Virol 15: 19-26.
Dreyfus, C., N. S. Laursen, T. Kwaks, D. Zuijdgeest, R. Khayat, D. C. Ekiert, J. H. Lee, Z. Metlagel, M. V. Bujny, M. Jongeneelen, R. van der Vlugt, M. Lamrani, H. J. Korse, E. Geelen, O. Sahin, M. Sieuwerts, J. P. Brakenhoff, R. Vogels, O. T. Li, L. L. Poon, M. Peiris, W. Koudstaal, A. B. Ward, I. A. Wilson, J. Goudsmit and R. H. Friesen (2012). "Highly conserved protective epitopes on influenza B viruses." Science 337(6100): 1343-1348.
Du, Y., J.F. Hultquist, Q. Zhou, A. Olson, Y. Tseng, T.H. Zhang, M. Hong, K. Tang, L. Chen, X. Meng, M.J. McGregor, L. Dai, D. Gong, L. Martin-Sancho, S. Chanda, X. Li, S. Bensenger, N.J. Krogan and R. Sun (2020). "mRNA display with library of even-distribution reveals cellular interactors of influenza virus NS1." Nat Commun 11(1): 2449.
Dubois, J., O. Terrier and M. Rosa-Calatrava (2014). "Influenza viruses and mRNA splicing: doing more with less." mBio 5(3): e00070-14.
Dubois, J., A. Traversier, T. Julien, B. Padey, B. Lina, J. C. Bourdon, V. Marcel, G. Boivin, M. Rosa-Calatrava and O. Terrier (2019). "The Nonstructural NS1 Protein of Influenza Viruses Modulates TP53 Splicing through Host Factor CPSF4." J Virol 93(7).
DuBois, R. M., P. J. Slavish, B. M. Baughman, M. K. Yun, J. Bao, R. J. Webby, T. R. Webb and S. W. White (2012). "Structural and biochemical basis for development of influenza virus inhibitors targeting the PA endonuclease." PLoS Pathog 8(8): e1002830.
Duwe, S (2017). "influenza virus – antiviral therapy and resistance." GMS Infect Dis 5: Doc04.
Ehrhardt, C. and S. Ludwig (2009). "A new player in a deadly game: influenza viruses and the PI3K/Akt signalling pathway." Cell Microbiol 11(6): 863-871.

E

Ekiert, D. C., G. Bhabha, M. A. Elsliger, R. H. Friesen, M. Jongeneelen, M. Throsby, J. Goudsmit and I. A. Wilson (2009). "Antibody recognition of a highly conserved influenza virus epitope." Science 324(5924): 246-251.
Ekiert, D. C., A. K. Kashyap, J. Steel, A. Rubrum, G. Bhabha, R. Khayat, J. H. Lee, M. A. Dillon, R. E. O'Neil, A. M. Faynboym, M. Horowitz, L. Horowitz, A. B. Ward, P. Palese, R. Webby, R. A. Lerner, R. R. Bhatt and I. A. Wilson (2012). "Cross-neutralization of influenza A viruses mediated by a single antibody loop." Nature 489(7417): 526-532.
Emsley, P. and K. Cowtan (2004). "Coot: model-building tools for molecular graphics". Acta Crystallogr D Biol Crystallogr. 60: 2136-32.
Enami, K., T. A. Sato, S. Nakada and M. Enami (1994). "Influenza virus NS1 protein stimulates translation of the M1 protein." J Virol 68(3): 1432-1437.
Engel, D. A (2013). "The influenza virus NS1 protein as a therapeutic target." Antiviral Res 99(3): 409-16.
Esparza, M., A. Mor, H. Niederstrasser, K. White, A. White, K. Zhang, S. Gao, J. Wang, J. Liang, S. Sho, R. Sakthivel, A.A. Sathe, C. Xing, R. Muñoz-Moreno, J.W. Shay, A. Garcia-Sastre, J. Ready, B. Posner and B.M.A. Fontoura (2020). "Chemical intervention of influenza virus mRNA nuclear export." PLoS Pathog 16(4): e1008407.
Essere, B., M. Yver, C. Gavazzi, O. Terrier, C. Isel, E. Fournier, F. Giroux, J. Textoris, T. Julien, C. Socratous, M. Rosa-Calatrava, B. Lina, R. Marquet and V. Moules (2013). "Critical role of segment-specific packaging signals in genetic reassortment of influenza A viruses." Proc Natl Acad Sci U S A 110(40): E3840-3848.
Evans, P.R. and G.N. Murshudov (2013). "How good are my data and what is the resolution?." Acta Crystallogr D Biol Crystallogr. 69(Pt 7): 1204-14.

F

Fabrini, R., A. De Luca, L. Stella, G. Mei, B. Orioni, S. Ciccone, G. Federici, M. Lo Bello and G. Ricci (2009). "Monomer-dimer equilibrium n glutathione transferases: a critical re-examination." Biochemistry 48(43): 10473-82.
Falcón, A. M., P. Fortes, R. M. Marion, A. Beloso and J. Ortin (1999). "Interaction of influenza virus NS1 protein and the human homologue of Staufen in vivo and in vitro." Nucleic Acids Res 27(11): 2241-2247.

Fan, H., A. P. Walker, L. Carrique, J. R. Keown, I. Serna Martin, D. Karia, J. Sharps, N. Hengrung, E. Pardon, J. Steyaert, J. M. Grimes and E. Fodor (2019). "Structures of influenza A virus RNA polymerase offer insight into viral genome replication." Nature 573(7773): 287-290.
Fedichev, P., R. Timakhov, T. Pyrkov, E. Getmantsev and A. Vinnik (2011). "Structure-based drug design of a new chemical class of small molecules active against influenza A nucleoprotein in vitro and in vivo." PLoS Curr 3: RRN1253.
Ferhadian, D., M. Contrant, A. Printz-Schweigert, R.P. Smyth, J.C. Paillart and R. Marquet (2018). "Structural and functional motifs in influenza virus RNAs." Front Microbiol 9:559.
Firth, A. E. and I. Brierley (2012). "Non-canonical translation in RNA viruses." J Gen Virol 93(Pt 7): 1385-1409.
Fodor, E., L. Devenish, O. G. Engelhardt, P. Palese, G. G. Brownlee and A. Garcia-Sastre (1999). "Rescue of influenza A virus from recombinant DNA." J Virol 73(11): 9679-9682.
Fodor, E., D. C. Pritlove and G. G. Brownlee (1994). "The influenza virus panhandle is involved in the initiation of transcription." J Virol 68(6): 4092-4096.
Fodor, E. and A. J. W. Te Velthius (2019). "Structure and function of the influenza virus transcription and replication machinery." Cold Spring Harb Perspect Med. pii: a038398.
Fournier, G., C. Chiang, S. Munier, A. Tomoiu, C. Demeret, P.O. Vidalain, Y. Jacob and N. Naffakh (2014). "Recruitment of RED-SMU1 complex by influenza A virus RNA polymerase to control viral mRNA splicing." PLoS Pathog 10(6): e1004164.
Fortes, P., A. Beloso and J. Ortin (1994). "Influenza virus NS1 protein inhibits pre-mRNA splicing and blocks mRNA nucleocytoplasmic transport." EMBO J 13(3): 704-712.
Fukao, K., T. Noshi, A. Yamamoto, M. Kitano, Y. Ando, T. Noda, K. Baba, K. Matsumoto, N. Higuchi, M. Ikeda, T. Shishido and A. Naito (2019). "Combination treatment with the cap-dependent endonuclease inhibitor baloxavir marboxil and a neuraminidase inhibitor in a mouse model of influenza A virus infection." J Antimicrob Chemother 74(3): 654-662.
Furuta, Y., B. B. Gowen, K. Takahashi, K. Shiraki, D. F. Smee and D. L. Barnard (2013). "Favipiravir (T-705), a novel viral RNA polymerase inhibitor." Antiviral Res 100(2): 446-454.

G

Gabant, G. and M. Cadene (2008). "Mass spectrometry of full-length integral membrane proteins to define functionally relevant structural features." Methods 46(2): 54-61.
Gack, M. U., Y. C. Shin, C. H. Joo, T. Urano, C. Liang, L. Sun, O. Takeuchi, S. Akira, Z. Chen, S. Inoue and J. U. Jung (2007). "TRIM25 RING-finger E3 ubiquitin ligase is essential for RIG-I-mediated antiviral activity." Nature 446(7138): 916-920.
Gal-Ben-Ari, S., I. Barrera, M. Ehrlich and K. Rosenblum (2018). "PKR: A Kinase to Remember." Front Mol Neurosci 11: 480.
Gan, J., J.E. Tropea, B.P. Austin, D.L. Court, D.S. Waugh and X. Ji (2005). "Intermediate states of ibonuclease III in complex with double-stranded RNA." Structure 13(10): 1435-42.
Ganser, L. R., M. L. Kelly, D. Herschlag and H. M. Al-Hashimi (2019). "The roles of structural dynamics in the cellular functions of RNAs." Nat Rev Mol Cell Biol 20(8): 474-489.
Gao, R., B. Cao, Y. Hu, Z. Feng, D. Wang, W. Hu, J. Chen, Z. Jie, H. Qiu, K. Xu, X. Xu, H. Lu, W. Zhu, Z. Gao, N. Xiang, Y. Shen, Z. He, Y. Gu, Z. Zhang, Y. Yang, X. Zhao, L. Zhou, X. Li, S. Zou, Y. Zhang, X. Li, L. Yang, J. Guo, J. Dong, Q. Li, L. Dong, Y. Zhu, T. Bai, S. Wang, P. Hao, W. Yang, Y. Zhang, J. Han, H. Yu, D. Li, G. F. Gao, G. Wu, Y. Wang, Z. Yuan and Y. Shu (2013). "Human infection with a novel avian-origin influenza A (H7N9) virus." N Engl J Med 368(20): 1888-1897.
Garaigorta, U., A. M. Falcon and J. Ortin (2005). "Genetic analysis of influenza virus NS1 gene: a temperature-sensitive mutant shows defective formation of virus particles." J Virol 79(24): 15246-15257.
Garcia-Sastre, A., A. Egorov, D. Matassov, S. Brandt, D. E. Levy, J. E. Durbin, P. Palese and T. Muster (1998). "Influenza A virus lacking the NS1 gene replicates in interferon-deficient systems." Virology 252(2): 324-330.
Garten, R. J., C. T. Davis, C. A. Russell, B. Shu, S. Lindstrom, A. Balish, W. M. Sessions, X. Xu, E. Skepner, V. Deyde, M. Okomo-Adhiambo, L. Gubareva, J. Barnes, C. B. Smith, S. L. Emery, M. J. Hillman, P. Rivailler, J. Smagala, M. de Graaf, D. F. Burke, R. A. Fouchier, C. Pappas, C. M. Alpuche-Aranda, H. Lopez-Gatell, H. Olivera, I. Lopez, C. A. Myers, D. Faix, P. J. Blair, C. Yu, K. M. Keene, P. D. Dotson, Jr., D. Boxrud, A. R. Sambol, S. H. Abid, K. St George, T. Bannerman, A. L. Moore, D. J. Stringer, P. Blevins, G. J. Demmler-Harrison, M. Ginsberg, P. Kriner, S. Waterman, S. Smole, H. F. Guevara, E. A. Belongia, P. A. Clark, S. T. Beatrice, R. Donis, J. Katz, L. Finelli, C. B. Bridges, M. Shaw, D. B. Jernigan, T. M. Uyeki, D. J. Smith, A. I. Klimov and N. J. Cox (2009). "Antigenic and genetic characteristics of swine-origin 2009 A(H1N1) influenza viruses circulating in humans." Science 325(5937): 197-201.
Gaucherand, L., B. K. Porter, R. E. Levene, E. L. Price, S. K. Schmaling, C. H. Rycroft, Y. Kevorkian, C. McCormick, D. A. Khaperskyy and M. M. Gaglia (2019). "The Influenza A Virus Endoribonuclease PA-X Usurps Host mRNA Processing Machinery to Limit Host Gene Expression." Cell Rep 27(3): 776-792 e777.

Gavazzi, C., M. Yver, C. Isel, R.P. Smyth, M.R. Calatrava, B. Lina, V. Moulès and Roland Marquet (2013). “A Functional Sequence-Specific Interaction Between Influenza A Virus Genomic RNA Segments.“ Proc Natl Acad Sci110(41):16604-9.
Gerber, M., C. Isel, V. Moules and R. Marquet (2014). "Selective packaging of the influenza A genome and consequences for genetic reassortment." Trends Microbiol 22(8): 446-455.
Gerritz, S. W., C. Cianci, S. Kim, B. C. Pearce, C. Deminie, L. Discotto, B. McAuliffe, B. F. Minassian, S. Shi, S. Zhu, W. Zhai, A. Pendri, G. Li, M. A. Poss, S. Edavettal, P. A. McDonnell, H. A. Lewis, K. Maskos, M. Mortl, R. Kiefersauer, S. Steinbacher, E. T. Baldwin, W. Metzler, J. Bryson, M. D. Healy, T. Philip, M. Zoeckler, R. Schartman, M. Sinz, V. H. Leyva-Grado, H. H. Hoffmann, D. R. Langley, N. A. Meanwell and M. Krystal (2011). "Inhibition of influenza virus replication via small molecules that induce the formation of higher-order nucleoprotein oligomers." Proc Natl Acad Sci U S A 108(37): 15366-15371.
Giege, R. and C. Sauter (2010). "Biocrystallography: past, present, future." HFSP J 4(3-4): 109-121.
Gilbertson, B., Zheng, T., Gerber, M., Printz-Schweigert, A., Ong, C., Marquet, R., Isel, C., Rockman, S. and L. Brown (2016). “Influenza NA and PB1 Gene Segments Interact during the Formation of Viral Progeny: Localization of the Binding Region within the PB1 Gene.” Viruses 8(8). pii: E238. doi: 10.3390/v8080238.
Glanz, V. Y., V. A. Myasoedova, A. V. Grechko and A. N. Orekhov (2018). "Inhibition of sialidase activity as a therapeutic approach." Drug Des Devel Ther 12: 3431-3437.
Gleghorn, M. L. and L. E. Maquat (2014). "'Black sheep' that don't leave the double-stranded RNA-binding domain fold." Trends Biochem Sci 39(7): 328-340.
Gränicher, G., Coronel, J., Pralow, A., Marichal-Gallardo, P., Wolff, M., Rapp, E., Karlas, A., Sandig, V., Genzel, Y. and U. Reichl (2019). "Efficient influenza A virus production in high cell density using the novel porcine suspension cell line PBG.PK2.1." Vaccine 37(47): 7019-7028.
Graves, P. N., J. L. Schulman, J. F. Young and P. Palese (1983). "Preparation of influenza virus subviral particles lacking the HA1 subunit of hemagglutinin: unmasking of cross-reactive HA2 determinants." Virology 126(1): 106-116.
Gu, W., G. R. Gallagher, W. Dai, P. Liu, R. Li, M. I. Trombly, D. B. Gammon, C. C. Mello, J. P. Wang and R. W. Finberg (2015). "Influenza A virus preferentially snatches noncoding RNA caps." RNA 21(12): 2067-2075.
Guan, R., L. C. Ma, P. G. Leonard, B. R. Amer, H. Sridharan, C. Zhao, R. M. Krug and G. T. Montelione (2011). "Structural basis for the sequence-specific recognition of human ISG15 by the NS1 protein of influenza B virus." Proc Natl Acad Sci 108(33): 13468-73.
Guilligay, D., F. Tarendeau, P. Resa-Infante, R. Coloma, T. Crepin, P. Sehr, J. Lewis, R. W. Ruigrok, J. Ortin, D. J. Hart and S. Cusack (2008). "The structural basis for cap binding by influenza virus polymerase subunit PB2." Nat Struct Mol Biol 15(5): 500-506.
Guo, L. and Y. Liu (2018). “Crystal structure of the kelch domain of human NS1-binding protein t 1.98 Å resolution.” Acta Crystallogr F Struct Biol Commun 74(Pt 3): 174-178.
Guo, Z., L. M. Chen, H. Zeng, J. A. Gomez, J. Plowden, T. Fujita, J. M. Katz, R. O. Donis and S. Sambhara (2007). "NS1 protein of influenza A virus inhibits the function of intracytoplasmic pathogen sensor, RIG-I." Am J Respir Cell Mol Biol 36(3): 263-269.
Guttler, T., T. Madl, P. Neumann, D. Deichsel, L. Corsini, T. Monecke, R. Ficner, M. Sattler and D. Gorlich (2010). "NES consensus redefined by structures of PKI-type and Rev-type nuclear export signals bound to CRM1." Nat Struct Mol Biol 17(11): 1367-1376.

H

Handa, N., O. Nureki, K. Kurimoto, I. Kim, H. Sakamoto, Y. Shimura, Y. Muto and S. Yokoyama (1999). “Structural basis for recognition of the tra mRNA precursor by the sex-lethal protein.” Nature 398(6728): 579-85.
Hale, B. G., R. E. Randall, J. Ortin and D. Jackson (2008). "The multifunctional NS1 protein of influenza A viruses." J Gen Virol 89(Pt 10): 2359-2376.
Hale, B. G., J. Steel, B. Manicassamy, R. A. Medina, J. Ye, D. Hickman, A. C. Lowen, D. R. Perez and A. Garcia-Sastre (2010). "Mutations in the NS1 C-terminal tail do not enhance replication or virulence of the 2009 pandemic H1N1 influenza A virus." J Gen Virol 91(Pt 7): 1737-1742.
Hale, B. G., J. Steel, R. A. Medina, B. Manicassamy, J. Ye, D. Hickman, R. Hai, M. Schmolke, A. C. Lowen, D. R. Perez and A. Garcia-Sastre (2010). "Inefficient control of host gene expression by the 2009 pandemic H1N1 influenza A virus NS1 protein." J Virol 84(14): 6909-6922.
Han, J., J.T. Perez, C. Chen, Y. Li, A. Benitez, M. Kandasamy, Y. Lee, J. Andrade, B. tenOever and B. Manicassamy (2018). “Genome-wide CRISPR/Cas9 Screen Identifies Host Factors Essential for Influenza Virus Replication.” Cell Report 23(2):596-607.
Hanada, K., Y. Suzuki and T. Gojobori (2004). "A large variation in the rates of synonymous substitution for RNA viruses and its relationship to a diversity of viral infection and transmission modes." Mol Biol Evol 21(6): 1074-1080.
Hanahan, D. (1983). "Studies on transformation of Escherichia coli with plasmids." J Mol Biol 166(4): 557-80.

Handa, N., O. Nureki, K. Kurimoto, I. Kim, H. Sakamoto, Y. Shimura, Y. Muto and S. Yokoyama (1999). "Structural basis for recognition of the tra mRNA precursor by the sex-lethal protein." Nature 398(6728): 579-85.
Hao, L., A. Sakurai, T. Watanabe, E. Sorensen, C.A Nidom, M.A. Newton, P. Ahlquist and Y. Kawaoka (2008). "Drosophila RNAi Screen Identifies Host Genes Important for Influenza Virus Replication. » Nature 454(7206): 890-3.
Harris, A., G. Cardone, D. C. Winkler, J. B. Heymann, M. Brecher, J. M. White and A. C. Steven (2006). "Influenza virus pleiomorphy characterized by cryoelectron tomography." Proc Natl Acad Sci U S A 103(50): 19123-19127.
Hartman, A. L., J. S. Towner and S. T. Nichol (2004). "A C-terminal basic amino acid motif of Zaire ebolavirus VP35 is essential for type I interferon antagonism and displays high identity with the RNA-binding domain of another interferon antagonist, the NS1 protein of influenza A virus." Virology 328(2): 177-184.
Hastings, J. C., H. Selnick, B. Wolanski and J. E. Tomassini (1996). "Anti-influenza virus activities of 4-substituted 2,4-dioxobutanoic acid inhibitors." Antimicrob Agents Chemother 40(5): 1304-1307.
Hatada, E. and R. Fukuda (1992). "Binding of influenza A virus NS1 protein to dsRNA in vitro." J Gen Virol 73 (Pt 12): 3325-3329.
Hatada, E., M. Hasegawa, K. Shimizu, M. Hatanaka and R. Fukuda (1990). "Analysis of influenza A virus temperature-sensitive mutants with mutations in RNA segment 8." J Gen Virol 71 (Pt 6): 1283-1292.
Hatada, E., S. Saito, N. Okishio and R. Fukuda (1997). "Binding of the influenza virus NS1 protein to model genome RNAs." J Gen Virol 78 (Pt 5): 1059-1063.
Hatada, E., T. Takizawa and R. Fukuda (1992). "Specific binding of influenza A virus NS1 protein to the virus minus-sense RNA in vitro." J Gen Virol 73 (Pt 1): 17-25.
He, X., J. Zhou, M. Bartlam, R. Zhang, J. Ma, Z. Lou, X. Li, J. Li, A. Joachimiak, Z. Zeng, R. Ge, Z. Rao and Y. Liu (2008). "Crystal structure of the polymerase PA(C)-PB1(N) complex from an avian influenza H5N1 virus." Nature 454(7208): 1123-1126.
Heaton, B.E., E.M. Kennedy, R.E. Dumm, A.T. Harding, M.T. Sacco, D. Sachs and N.S. Heaton (2017). "A CRISPR Activation Screen Identifies a Pan-avian Influenza Virus Inhibitory Host Factor." Cell Report 20(7):1503-1512.
Heinz, S., L. Texari, M. G. B. Hayes, M. Urbanowski, M. W. Chang, N. Givarkes, A. Rialdi, K. M. White, R. A. Albrecht, L. Pache, I. Marazzi, A. García-Sastre, M. L. Shaw and C. Benner (2018). "Transcription elongation can affect genome 3D structure." Cell 174(6): 1522-1536.
Hengrung, N., K. El Omari, I. Serna Martin, F. T. Vreede, S. Cusack, R. P. Rambo, C. Vonrhein, G. Bricogne, D. I. Stuart, J. M. Grimes and E. Fodor (2015). "Crystal structure of the RNA-dependent RNA polymerase from influenza C virus." Nature 527(7576): 114-117.
Higashi, K., Y. Terui, E. Inomata, D. Katagiri, Y. Nomura, T. Someya, K. Nishimura, K. Kashiwagi, G. Kawai and K. Igarashi (2008). "Selective structural hange of bulged-out region of double-stranded RNA containing bulged nucleotides by spermidine." Biochem Biophys Res Commun. 370(4): 572-7.
Hofer, C. T., S. Di Lella, I. Dahmani, N. Jungnick, N. Bordag, S. Bobone, Q. Huang, S. Keller, A. Herrmann and S. Chiantia (2019). "Structural determinants of the interaction between influenza A virus matrix protein M1 and lipid membranes." Biochim Biophys Acta Biomembr 1861(6): 1123-1134.
Hoffmann, E., J. Stech, Y. Guan, R. G. Webster and D. R. Perez (2001). "Universal primer set for the full-length amplification of all influenza A viruses." Arch Virol 146(12): 2275-2289.
Hu, Y., H. Sneyd, R. Dekant and J. Wang (2017). "Influenza A Virus Nucleoprotein: A Highly Conserved Multi-Functional Viral Protein as a Hot Antiviral Drug Target." Curr Top Med Chem 17(20): 2271-2285.
Huang, S., J. Chen, Q. Chen, H. Wang, Y. Yao, J. Chen and Z. Chen (2013). "A second CRM1-dependent nuclear export signal in the influenza A virus NS2 protein contributes to the nuclear export of viral ribonucleoproteins." J Virol 87(2): 767-778.
Hutchinson, E. C. and E. Fodor (2013). "Transport of the influenza virus genome from nucleus to nucleus." Viruses 5(10): 2424-2446.
Hutchinson, E. C., P. D. Charles, S. S. Hester, B. Thomas, D. Trudgian, M. Martínez-Alonso and E. Fodor (2014). "Conserved and host-specific features of influenza virion architecture." Nat Commun. 5: 4816.

I

Ilyina, I. V., V. V. Zarubaev, I. N. Lavrentieva, A. A. Shtro, I. L. Esaulkova, D. V. Korchagina, S. S. Borisevich, K. P. Volcho and N. F. Salakhutdinov (2018). "Highly potent activity of isopulegol-derived substituted octahydro-2H-chromen-4-ols against influenza A and B viruses." Bioorg Med Chem Lett 28(11): 2061-2067.
Imai, M., Yamashita, M., Sakai-Tagawa, Y., Iwatsuki-Horimoto, K., Kiso, M., Murakami, J., Yasuhara, A., Takada, K., Ito, M., Nakajima, N., Takahashi, K., Lopes, T. J. S., Dutta, J., Khan, Z., Kriti, D., van Bakel, H., Tokita, A., Hagiwara, H., Izumida, N., Kuroki, H., Nishino, T., Wada, N., Koga, M., Adachi, E., Jubishi, D., Hasegawa, H. and Y. Kawaoka. (2019). "Influenza A variants with reduced susceptibility to baloxavir isolated

from Japanese patients are fit and transmit through respiratory droplets." *Nat Microbiol* doi: 10.1038/s41564-019-0609-0.
Isel, C.G., Munier, S. and N. Naffakh (2016). « Experimental Approaches to Study Genome Packaging of Influenza A Viruses." Viruses 8(8) pii: E218.
Isel, C. G., K. Y. Chen and N. Naffakh (2020). "Dimerisation, a novel feature of the influenza virus RNA polymerase." Med Sci 36(4): 332-335.
Ivanyi-Nagy, R., J. P. Lavergne, C. Gabus, D. Ficheux and J. L. Darlix (2008). "RNA chaperoning and intrinsic disorder in the core proteins of Flaviviridae." Nucleic Acids Res 36(3): 712-725.

J

Jablonski, J. J., D. Basu, D. A. Engel and H. M. Geysen (2012). "Design, synthesis, and evaluation of novel small molecule inhibitors of the influenza virus protein NS1." Bioorg Med Chem 20(1): 487-497.
Jakubcova, L., J. Holly and E. Vareckova (2016). "The role of fusion activity of influenza A viruses in their biological properties." Acta Virol 60(2): 121-135.
Jiao, P., G. Tian, Y. Li, G. Deng, Y. Jiang, C. Liu, W. Liu, Z. Bu, Y. Kawaoka and H. Chen (2008). "A single-amino-acid substitution in the NS1 protein changes the pathogenicity of H5N1 avian influenza viruses in mice." J Virol. 82(3): 1146-54.
Johnson, N. P. and J. Mueller (2002). "Updating the accounts: global mortality of the 1918-1920 "Spanish" influenza pandemic." Bull Hist Med 76(1): 105-115.
Jones, J. C., E. A. Turpin, H. Bultmann, C. R. Brandt and S. Schultz-Cherry (2006). "Inhibition of influenza virus infection by a novel antiviral peptide that targets viral attachment to cells." J Virol 80(24): 11960-11967.
Jucker, F. M., H. A. Heus, P. F. Yip, E. H. Moors and A. Pardi (1996)." A network of heterogeneous hydrogen bonds in GNRA tetraloops." J Mol Biol 264(5): 968-80.
Julien, J. P., P. S. Lee and I. A. Wilson (2012). "Structural insights into key sites of vulnerability on HIV-1 Env and influenza HA." Immunol Rev 250(1): 180-198.
Jurado, A. R., D. Tan, X. Jiao, M. Kiledjian and L. Tong (2014). "Structure and function of pre-mRNA 5'-end capping quality control and 3'-end processing." Biochemistry 53(12): 1882-1898.
Jureka, A. S., A. B. Kleinpeter, G. Cornilescu, C. C. Cornilescu and C. M. Petit (2015). "Structural Basis for a Novel Interaction between the NS1 Protein Derived from the 1918 Influenza Virus and RIG-I." Structure 23(11): 2001-2010.
Jureka, A., A. B. Kleinpeter, J. L. Tipper, K.S. Harrod and C. M. Petit (2019). "The influenza NS1 protein modulates RIG-I activation via a strain-specific direct interaction with the second CARD of RIG-I." J Biol Chem. pii: jbc. RA119.011410.

K

Kabsch, W. (2010). "XDS." Acta Cryst D66: 125-132.
Kadam, R. U., J. Juraszek, B. Brandenburg, C. Buyck, W. B. G. Schepens, B. Kesteleyn, B. Stoops, R. J. Vreeken, J. Vermond, W. Goutier, C. Tang, R. Vogels, R. H. E. Friesen, J. Goudsmit, M. J. P. van Dongen, I. A. Wilson (2017). "Potent peptidic fusion inhibitors of influenza virus. "Science 358(6362): 496-502.
Kakisaka, M., Y. Sasaki, K. Yamada, Y. Kondoh, H. Hikono, H. Osada, K. Tomii, T. Saito and Y. Aida (2015). "A Novel Antiviral Target Structure Involved in the RNA Binding, Dimerization, and Nuclear Export Functions of the Influenza A Virus Nucleoprotein." PLoS Pathog 11(7): e1005062.
Kao, R. Y., D. Yang, L. S. Lau, W. H. Tsui, L. Hu, J. Dai, M. P. Chan, C. M. Chan, P. Wang, B. J. Zheng, J. Sun, J. D. Huang, J. Madar, G. Chen, H. Chen, Y. Guan and K. Y. Yuen (2010). "Identification of influenza A nucleoprotein as an antiviral target." Nat Biotechnol 28(6): 600-605.
Karlas, A., N. Machuy, Y. Shin, K.P. Pleissner, A. Artarini, D. Heuer, D. Becker, H. Khalil, L.A. Ogilvie, S. Hess, A.P. Mäurer, E. Müller, T. Wolff, T. Rudel and T.F. Meyer (2010). "Genome-wide RNAi Screen Identifies Human Host Factors Crucial for Influenza Virus Replication" Nature 463(7282):818-22.
Kathum, O. A., T. Schräder, D. Anhlan, C. Nordhoff, S. Liedmann, A. Pande, A. Mellmann, C. Ehrhardt, V. Wixler and S. Ludwig (2016). "Phosphorylation of influenza A virus NS1 protein at threonine 49 suppresses its interferon antagonistic activity." Cell Microbiol. 18(6): 784-91.
Kawaoka, Y., S. Krauss and R. G. Webster (1989). "Avian-to-human transmission of the PB1 gene of influenza A viruses in the 1957 and 1968 pandemics." J Virol 63(11): 4603-4608.
Kerry, P. S., J. Ayllon, M. A. Taylor, C. Hass, A. Lewis, A. Garcia-Sastre, R. E. Randall, B. G. Hale and R. J. Russell (2011). "A transient homotypic interaction model for the influenza A virus NS1 protein effector domain." PLoS One 6(3): e17946.
Kharrat, A., M. J. Macias, T. J. Gibson, M. Nilges and A. Pastore (1995). "Structure of the dsRNA binding domain of E. coli RNase III." EMBO J 14(14): 3572-3584.
Kimberlin, C. R., Z. A. Bornholdt, S. Li, V. L. Woods, Jr., I. J. MacRae and E. O. Saphire (2010). "Ebolavirus VP35 uses a bimodal strategy to bind dsRNA for innate immune suppression." Proc Natl Acad Sci U S A 107(1): 314-319.

Kleinpeter, A. B., A. S. Jureka, S. M. Falahat, T. J. Green and C. M. Petit (2018). "Structural analyses reveal the mechanism of inhibition of influenza virus NS1 by two antiviral compounds." J Biol Chem 293(38): 14659-14668.
Klemm, C., Y. Boergeling, S. Lugwig and C. Ehrhardt (2017). “Immunomodulatory non-structural proteins of influenza A virus.” Trends Microbiol. 26(7): 624-636.
Klug, A. (2010). "The discovery of zinc fingers and their applications in gene regulation and genome manipulation." Annu Rev Biochem 79: 213-231.
Klumpp, K., R.W. Riogrok and F. Baudin (1997). “Roles of the influenza virus polymerase and nucleoprotein in forming a functional RNP structure.” EMBO J 16(6): 1248-57.
Koliopoulos, M. G., M. Lethier, A. G. Van der Veen, K. Haubrich, J. Hennig, E. Kowalinski, R.V. Stevens, S. R. Martin, C. Reis E Sousa, S. Cusack and K. Rittinger (2018). "Molecular mechanism of influenza A NS1-mediated TRIM25 recognition and inhibition." Nat Commun 9(1): 1820.
Kondo, Y., C. Oubridge, A. M. van Roon and K. Nagai (2015). "Crystal structure of human U1 snRNP, a small nuclear ribonucleoprotein particle, reveals the mechanism of 5' splice site recognition." elife 4.
König, R., S. Stertz, Y. Zhou, A. Inoue, H.H. Hoffmann, S. Bhattacharyya, J.G. Alamares, D.M. Tscherne, M.B. Ortigoza, Y. Liang, Q. Gao, S.E. Andrews, S. Bandyopadhyay, P.D. Jesus, B.P. Tu, L. Pache, C. Shih, A. Orth, G. Bonamy, L. Miraglia, T. Ideker, A. García-Sastre, J.A.T. Young, P. Palese, M.L. Shaw and S.K. Chanda (2010). “Human Host Factors Required for Influenza Virus Replication.” Nature 463(7282): 813–817.
Koppstein, D., J. Ashour and D. P. Bartel (2015). "Sequencing the cap-snatching repertoire of H1N1 influenza provides insight into the mechanism of viral transcription initiation." Nucleic Acids Res 43(10): 5052-5064.
Koudstaal, W., Hartgroves, L., Havenga, M., Legastelois, I., Ophorst, C., Sieuwerts, M., Zuijdgeest, D., Vogels, R., Custers, J., de Boer-Luijtze, E., de Leeuw, O., Cornelissen, L., Goudsmit, J. and W. Barclay (2009). "Suitability of PER.C6 cells to generate epidemic and pandemic influenza vaccine strains by reverse genetics." Vaccine 27(19): 2588-93.
Kowalinski, E., C. Zubieta, A. Wolkerstorfer, O. H. Szolar, R. W. Ruigrok and S. Cusack (2012). "Structural analysis of specific metal chelating inhibitor binding to the endonuclease domain of influenza pH1N1 (2009) polymerase." PLoS Pathog 8(8): e1002831.
Krammer, F., N. Pica, R. Hai, I. Margine and P. Palese (2013). "Chimeric hemagglutinin influenza virus vaccine constructs elicit broadly protective stalk-specific antibodies." J Virol 87(12): 6542-6550.
Krammer, F., Smith, G. J. D., Fouchier, R. A. M., Peiris, M., Kedzierska, K., Doherty, P. C., Palese, P., Shaw, M. L., Treanor, J., Webster, R. G. and A. García-Sastre (2018). ”Influenza.” Nat Rev Dis Primers 4(1):3.
Krug, R. M. (2014). "Viral proteins that bind double-stranded RNA: countermeasures against host antiviral responses." J Interferon Cytokine Res 34(6): 464-468.
Kuchipudi, S. V. and R. H. Nissly (2018). "Novel Flu Viruses in Bats and Cattle: "Pushing the Envelope" of Influenza Infection." Vet Sci 5(3).
Kuo, R. L., C. J. Chen, E. H. Tam, C. G. Huang, L. H. Li, Z. H. Li, P. C. Su, H. P. Liu and C. C. Wu (2018). "Interactome Analysis of NS1 Protein Encoded by Influenza A H7N9 Virus Reveals an Inhibitory Role of NS1 in Host mRNA Maturation." J Proteome Res 17(4): 1474-1484.
Kuss, S. K., Mata, M. A., Zhang, L. and Fontoura, B. M (2013). "Nuclear imprisonment: viral strategies to arrest host mRNA nuclear export."Viruses 5(7): 1824-49.

L

Labaronne, A., C. Swale, A. Monod, G. Schoehn, T. Crepin and R. W. Ruigrok (2016). "Binding of RNA by the Nucleoproteins of Influenza Viruses A and B." Viruses 8(9).
Lakadamyali, M., M. J. Rust and X. Zhuang (2004). "Endocytosis of influenza viruses." Microbes Infect 6(10): 929-936.
Lalime, E. N. and A. Pekosz (2014). The R35 residue of the influenza A virus NS1 protein has minimal effects on nuclear localization but alters virus replication through disrupting protein dimerization. Virology 458-459: 33– 42.
Lamb, R. A., P. W. Choppin, R. M. Chanock and C. J. Lai (1980). "Mapping of the two overlapping genes for polypeptides NS1 and NS2 on RNA segment 8 of influenza virus genome." Proc Natl Acad Sci U S A 77(4): 1857-1861.
Lamb, R. A. and C. J. Lai (1980). "Sequence of interrupted and uninterrupted mRNAs and cloned DNA coding for the two overlapping nonstructural proteins of influenza virus." Cell 21(2): 475-485.
Larsen, S., S. Bui, V. Perez, A. Mohammad, H. Medina-Ramirez and L. L. Newcomb (2014). "Influenza polymerase encoding mRNAs utilize atypical mRNA nuclear export." Virol J 11: 154.
Laursen, N. S. and I. A. Wilson (2013). "Broadly neutralizing antibodies against influenza viruses." Antiviral Res 98(3): 476-483.
Laursen, N. S., R. H. E. Friesen, X. Zhu, M. Jongeneelen, S. Blokland, J. Vermond, A. van Eijgen, C. Tang, H. van Diepen, G. Obmolova, M. van der Neut Kolfschoten, D. Zuijdgeest, R. Straetemans, R. M. B. Hoffman, T. Nieusma, J. Pallesen, H. L. Turner, S. M. Bernard, A. B. Ward, J. Luo, L. L. M. Poon, A. P. Tretiakova, J. M.

Wilson, M. P. Limberis, R. Vogels, B. Brandenburg, J. A. Kolkman and I. A. Wilson. "Universal protection against influenza infection by a multidomain antibody to influenza hemagglutinin." Science 362(6414): 598-602.
Lazzaretti, D., L. Bandholz-Cajamarca, C. Emmerich, K. Schaaf, C. Basquin, U. Irion and F. Bono (2018). "The crystal structure of Staufen1 in complex with a physiological RNA sheds light on substrate selectivity." Life Sci Alliance 1(5): e201800187.
Lee, B. M., J. Xu, B. K. Clarkson, M. A. Martinez-Yamout, H. J. Dyson, D. A. Case, J. M. Gottesfeld and P. E. Wright (2006). "Induced fit and "lock and key" recognition of 5S RNA by zinc fingers of transcription factor IIIA." J Mol Biol 357(1): 275-291.
Lee, H. J. and J. J. Zheng (2010). "PDZ domains and their binding partners: structure, specificity, and modification." Cell Commun Signal 8: 8.
Lejal, N., B. Tarus, E. Bouguyon, S. Chenavas, N. Bertho, B. Delmas, R.H.W. Ruigrok, C. Di Primo and A. Slama-Schwok (2013). "Structure-based discovery of the novel antiviral properties of naproxen against the nucleoprotein of influenza A virus." Antimicrob Agents Chemother 57(5): 2231-42.
Lewis, H. A., K. Musunuru, K. B. Jensen, C. Edo, H. Chen, R. B. Darnell and K. Burley (2000). "Sequence-specific RNA binding by a Nova KH domain: implications for paraneoplastic disease and the fragile X syndrome." Cell 100(3): 323-32.
Leung, D. W., K. C. Prins, D. M. Borek, M. Farahbakhsh, J. M. Tufariello, P. Ramanan, J. C. Nix, L. A. Helgeson, Z. Otwinowski, R. B. Honzatko, C. F. Basler and G. K. Amarasinghe (2010). "Structural basis for dsRNA recognition and interferon antagonism by Ebola VP35." Nat Struct Mol Biol 17(2): 165-172.
Li, S., T. Hattori and E. N. Kodama (2011). "Epigallocatechin gallate inhibits the HIV reverse transcription step." Antivir Chem Chemother 21(6): 239-243.
Li, S., J. Y. Min, R. M. Krug and G. C. Sen (2006). "Binding of the influenza A virus NS1 protein to PKR mediates the inhibition of its activation by either PACT or double-stranded RNA." Virology 349(1): 13-21.
Li, W., J. W. Noah and D. L. Noah (2011). "Alanine substitutions within a linker region of the influenza A virus non-structural protein 1 alter its subcellular localization and attenuate virus replication." J Gen Virol 92(Pt 8): 1832-1842.
Lin, M. I., B. H. Su, C. H. Lee, S. T. Wang, W. C. Wu, P. Dangate, S. Y. Wang, W. I. Huang, T. J. Cheng, O. A. Lin, Y. S. Cheng, Y. J. Tseng and C. M. Sun (2015). "Synthesis and inhibitory effects of novel pyrimido-pyrrolo-quinoxalinedione analogues targeting nucleoproteins of influenza A virus H1N1." Eur J Med Chem 102: 477-486.
Lingel, A., B. Simon, E. Izaurralde and M. Sattler (2005). "The structure of the flock house virus B2 protein, a viral suppressor of RNA interference, shows a novel mode of double-stranded RNA recognition." EMBO Rep 6(12): 1149-1155.
Liu, F. and J. Gu (2011). "Retinoic acid inducible gene-I, more than a virus sensor." Protein Cell 2(5): 351-357.
Liu, J., P. A. Lynch, C. Y. Chien, G. T. Montelione, R. M. Krug and H. M. Berman (1997). "Crystal structure of the unique RNA-binding domain of the influenza virus NS1 protein." Nat Struct Biol 4(11): 896-899.
Liu, J. H., K. Xi, X. Zhang, L. Bao, X. Zhang and Z. J. Tan (2019). "Structural Flexibility of DNA-RNA Hybrid Duplex: Stretching and Twist-Stretch Coupling." Biophys J 117(1): 74-86.
Liu, Y., G. Meng, M. Luo and X. Zheng (2013). "Crystallization and X-ray crystallographic analysis of the cap-binding domain of influenza A virus H1N1 polymerase subunit PB2." Acta Crystallogr Sect F Struct Biol Cryst Commun 69(Pt 3): 280-283.
Liu, Y., D. Olagnier and R. Lin (2017). "Host and viral modulation of RIG-I-mediated antiviral immunity." Front Immunol 3(7): 662.
Lo, C. Y., Y. S. Tang and P. C. Shaw (2018). "Structure and Function of Influenza Virus Ribonucleoprotein." Subcell Biochem 88: 95-128.
Long, J.S., B. Mistry, S.M. Haslam and W.S. Barclay (2019). "Host and viral determinants of influenza A virus species specificity." Nat Rev Microbiol 17(2): 67-81.
Lu, D., M. A. Searles and A. Klug (2003). "Crystal structure of a zinc-finger-RNA complex reveals two modes of molecular recognition." Nature 426(6962): 96-100.
Lu, X.J. and K.W. Olson (2008). "3DNA: a versatile, integrated software system for the analysis, rebuilding and visualization of three-dimensional nucleic-acid structures." Nat Protoc 3, 1213-1227.
Lu, Y., X. Y. Qian and R. M. Krug (1994). "The influenza virus NS1 protein: a novel inhibitor of pre-mRNA splicing." Genes Dev 8(15): 1817-1828.
Lu, Y., M. Wambach, M. G. Katze and R. M. Krug (1995). "Binding of the influenza virus NS1 protein to double-stranded RNA inhibits the activation of the protein kinase that phosphorylates the eIF-2 translation initiation factor." Virology 214(1): 222-8.
Ludwig, S., U. Schultz, J. Mandler, W. M. Fitch and C. Scholtissek (1991). "Phylogenetic relationship of the nonstructural (NS) genes of influenza A viruses." Virology 183(2): 566-577.
Ludwig, S., U. Vogel and C. Scholtissek (1995). "Amino acid replacements leading to temperature-sensitive defects of the NS1 protein of influenza A virus." Arch Virol 140(5): 945-950.

Lukarska, M., G. Fournier, A. Pflug, P. Resa-Infante, S. Reich, N. Naffakh and S. Cusack (2017). "Structural basis of an essential interaction between influenza polymerase and Pol II CTD." Nature 541(7635): 117-121.
Lunde, B. M., C. Moore and G. Varani (2007). "RNA-binding proteins: modular design for efficient function." Nat Rev Mol Cell Biol 8(6): 479-490.

M

Ma, J., R. Wu, G. Xu, Y. Cheng, Z. Wang, H. Wang, Y. Yan, J. Li and J. Sun (2020). "Acetylation at K108 of the NS1 protein is important for the replication and virulence of influenza virus." Vet Res 51(1): 20.
Ma, L. C., R. Guan, K. Hamilton, J. M. Aramini, L. Mao, S. Wang, R. M. Krug and G. T. Montelione (2016). "A Second RNA-Binding Site in the NS1 Protein of Influenza B Virus." Structure 24(9): 1562-1572.
Mandel, C. R., Y. Bai and L. Tong (2008). "Protein factors in pre-mRNA 3'-end processing." Cell Mol Life Sci 65(7-8): 1099-1122.
Manna, S. (2015). "An overview of pentatricopeptide repeat proteins and their applications." Biochimie 113: 93-9.
Manning, B. D. and A. Toker (2017). "AKT/PKB Signaling: Navigating the Network." Cell 169(3): 381-405.
Marazzi, I., J.S.Y. Ho, J. Kim, B. Manicassamy, S. Dewell, R.A. Albrecht, C.W. Seibert, U. Schaefer, K.L. Jeffrey, R.K. Prinjha, K. Lee, A. García-Sastre, R.G. Roeder and A. Tarakhovsky (2013) "Suppression of the antiviral response by an influenza histone mimic." Nature 483(7390):428-33.
Marc, D (2014). "Influenza virus non-structural protein NS1: interferon antagonism and beyond." J Gen Virol 95(12): 2594-611.
Marc, D., S. Barbachou and D. Soubieux (2013). "The RNA-binding domain of influenzavirus non-structural protein-1 cooperatively binds to virus-specific RNA sequences in a structure-dependent manner." Nucleic Acids Res 41(1): 434-449.
Marion, R. M., T. Aragón, A. Beloso, A. Nieto and J. Ortin (1997). "The N-terminal half of the influenza virus NS1 protein is sufficient for nuclear retention of mRNA and enhancement of viral mRNA translation." Nucleic Acids Res 25(21): 4271-4277.
Marion, R. M., T. Zurcher, S. De la Luna and J. Ortin (1997). "Influenza virus NS1 protein interacts with viral transcription-replication complexes in vivo." J Gen Virol 78 (Pt 10): 2447-2451.
Maris, C., C. Dominguez and F. H. Allain (2005). "The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression." FEBS J 272(9): 2118-2131.
Maroto, M., Y. Fernandez, J. Ortin, F. Pelaez and M. A. Cabello (2008). "Development of an HTS assay for the search of anti-influenza agents targeting the interaction of viral RNA with the NS1 protein." J Biomol Screen 13(7): 581-590.
Martín-Benito, J., E. Area, J. Ortega, O. Llorca, J. M. Valpuesta, J. L. Carrascosa and J. Ortín (2001). "Three-dimensional reconstruction of a recombinant influenza virus ribonucleoprotein particle." EMBO rep. 2(4): 313-7.
Martyna, A. and J. Rossman (2014). "Alterations of membrane curvature during influenza virus budding." Biochem Soc Trans 42(5): 1425-1428.
Masliah, G., P. Barraud and F. H. Allain (2013). "RNA recognition by double-stranded RNA binding domains: a matter of shape and sequence." Cell Mol Life Sci 70(11): 1875-1895.
Masliah, G., C. Maris, S. L. König, M. Yulikov, F. Aeschimann, A. L. Malinowska, J. Mabille, J. Weiler, A. Holla, J. Hunziker, N. Meisner-Kober, B. Schuler, G. Jeschke and F. H. Allain (2018). "Structural basis of siRNA recognition by TRBP double-stranded RNA binding domains." EMBO J 37(6).
Matlin, K. S., H. Reggio, A. Helenius and K. Simons (1981). "Infectious entry pathway of influenza virus in a canine kidney cell line." J Cell Biol 91(3 Pt 1): 601-613.
Matrosovich, M., T. Matrosovich, W. Garten and H. D. Klenk (2006). "New low-viscosity overlay medium for viral plaque assays." Virol J 3: 63.
McAuley, J. L., B. P. Gilbertson, S. Trifkovic, L. E. Brown and J. L. McKimm-Breschkin (2019). "Influenza Virus Neuraminidase Structure and Functions." Front Microbiol 10: 39.
McCoy, A.J. (2007). "Solving structures of protein complexes by molecular replacement with Phaser." Acta Crystallograph D Biol Crystallograph. 63, 32-41.
McGinty R.K. and S. Tan (2015). "Nucleosome Structure and Function. « Chem Rev 115(6): 2255–2273.
Medina, R. A. and A. Garcia-Sastre (2011). "Influenza A viruses: new research developments." Nat Rev Microbiol 9(8): 590-603.
Melén, K., L. Kinnunen, R. Fagerlund, N. Ikonen, K.Y. Twu, R.M. Krug and I. Julkunen (2007). "Nuclear and nucleolar targeting of influenza A virus NS1 protein: Striking differences between different virus subtypes." J Virol 81(11): 5995-6006.
Mercey, R., Lantier, I., Maurel, M. C., Grosclaude, J., Lantier, F. and D. Marc (2006). "Fast, reversible interaction of prion protein with RNA aptamers containing specific sequence patterns." Arch. Virol. 151: 2197–2214.
Meyer, A. G. and C. O. Wilke (2015). "Geometric Constraints Dominate the Antigenic Evolution of Influenza H3N2 Hemagglutinin." PLoS Pathog 11(5): e1004940.

Meyerson, N. R., L. Zhou, Y. R. Guo, C. Zhao, Y. J. Tao, R. M. Krug and S. L. Sawyer (2017). “Nuclear TRIM25 specifically targets influenza virus ribonucleoproteins to block the onset of RNA hain elongation.” Cell Host Microbe 22(5): 627-638.e7.
Mibayashi, M., L. Martinez-Sobrido, Y. M. Loo, W. B. Cardenas, M. Gale, Jr. and A. Garcia-Sastre (2007). "Inhibition of retinoic acid-inducible gene I-mediated induction of beta interferon by the NS1 protein of influenza A virus." J Virol 81(2): 514-524.
Min, J. Y. and R. M. Krug (2006). "The primary function of RNA binding by the influenza A virus NS1 protein in infected cells: Inhibiting the 2'-5' oligo (A) synthetase/RNase L pathway." Proc Natl Acad Sci U S A 103(18): 7100-7105.
Mitra, S., D. Kumar, L. Hu, B. Sankaran, M. M. Moosa, A. P. Rice, J. C. Ferreon, A. C. M. Ferreon and B. V. V. Prasad (2019). "Influenza A virus protein NS1 exhibits strain-independent conformational plasticity." J Virol.
Miyamoto, S. and T. Noda (2019). “In vitro vRNA-vRNA interactions in the H1N1 influenza A virus genome.” Microbiol Immunol. doi: 10.1111/1348-0421.12766.
Mittendorf, K. F., C. L. Deatherage, M. D. Ohi and C. R. Sanders (2012). "Tailoring of membrane proteins by alternative splicing of pre-mRNA." Biochemistry 51(28): 5541-5556.
Moeller, A., R. N. Kirchdoerfer, C. S. Potter, B. Carragher and I. A. Wilson (2012). "Organization of the influenza virus replication machinery." Science 338(6114): 1631-1634.
Monto, A. S. and K. Fukuda (2019). "Lessons from Influenza Pandemics of the last 100 Years." Clin Infect Dis.
Moreira, E. A., A. Weber, H. Bolte, L. Kolesnikova, S. Giese, S. Lakdawala, M. Beer, G. Zimmer, A. Garcia-Sastre, M. Schwemmle and M. Juozapaitis (2016). "A conserved influenza A virus nucleoprotein code controls specific viral genome packaging." Nat Commun 7: 12861.
Mühlberger, E., B. Lotfering, H. D. Klenk and S. Becker (1998). "Three of the four nucleocapsid proteins of Marburg virus, NP, VP35, and L, are sufficient to mediate replication and transcription of Marburg virus-specific monocistronic minigenomes." J Virol 72(11): 8756-8764.
Muñoz-Moreno, R., C. Martínez-Romero, D. Blanco-Melo, C. V. Forst, R. Nachbagauer, A. A. Benitez, I. Mena, S. Aslam, V. Balasubramaniam, I. Lee, M. Panis, J. Ayllón, D. Sachs, M. S. Park, F. Krammer, B. R. tenOever and A. García-Sastre (2019).
”Viral Fitness Landscapes in Diverse Host Species Reveal Multiple Evolutionary Lines for the NS1Gene of Influenza A Viruses. Cell Rep 29(12): 3997-4009.
Munster, V. J., C. Baas, P. Lexmond, J. Waldenstrom, A. Wallensten, T. Fransson, G. F. Rimmelzwaan, W. E. Beyer, M. Schutten, B. Olsen, A. D. Osterhaus and R. A. Fouchier (2007). "Spatial, temporal, and species variation in prevalence of influenza A viruses in wild migratory birds." PLoS Pathog 3(5): e61.
Munster, V. J. and R. A. Fouchier (2009). "Avian influenza virus: of virus and bird ecology." Vaccine 27(45): 6340-6344.

N

Nagai, K. (1996). "RNA-protein complexes." Curr Opin Struct Biol 6(1): 53-61.
Nao, N., J. Yamagishi, H. Miyamoto, M. Igarashi, R. Manzoor, A. Ohnuma, Y. Tsuda, W. Furuyama, A. Shigeno, M. Kajihara, N. Kishida, R. Yoshida and A. Takada (2017). "Genetic Predisposition To Acquire a Polybasic Cleavage Site for Highly Pathogenic Avian Influenza Virus Hemagglutinin." MBio 8(1).
Nayak, D. P., R. A. Balogun, H. Yamada, Z. H. Zhou and S. Barman (2009). "Influenza virus morphogenesis and budding." Virus Res 143(2): 147-161.
Nemeroff, M. E., S. M. Barabino, Y. Li, W. Keller and R. M. Krug (1998). "Influenza virus NS1 protein interacts with the cellular 30 kDa subunit of CPSF and inhibits 3'end formation of cellular pre-mRNAs." Mol Cell 1(7): 991-1000.
Nemeroff, M. E., X. Y. Qian and R. M. Krug (1995). "The influenza virus NS1 protein forms multimers in vitro and in vivo." Virology 212(2): 422-428.
Nemeroff, M. E., U. Utans, A. Kramer and R. M. Krug (1992). "Identification of cis-acting intron and exon regions in influenza virus NS1 mRNA that inhibit splicing and cause the formation of aberrantly sedimenting presplicing complexes." Mol Cell Biol 12(3): 962-970.
Ng, A. K., H. Zhang, K. Tan, Z. Li, J. H. Liu, P. K. Chan, S. M. Li, W. Y. Chan, S. W. Au, A. Joachimiak, T. Walz, J. H. Wang and P. C. Shaw (2008). "Structure of the influenza virus A H5N1 nucleoprotein: implications for RNA binding, oligomerization, and vaccine design." FASEB J 22(10): 3638-3647.
Ng, K. E. (2019). "Xofluza (Baloxavir Marboxil) for the Treatment Of Acute Uncomplicated Influenza." P T 44(1): 9-11.
Ngamurulert, S., T. Limjindaporn and P. Auewaraku (2009). "Identification of cellular partners of Influenza A virus (H5N1) non-structural protein NS1 by yeast two-hybrid system." Acta Virol 53(3): 153-159.
Nguyen, C. D., R. E. Mansfield, W. Leung, P. M. Vaz, F. E. Loughlin, R. P. Grant and J. P. Mackay (2011). “Characterization of a family of RanBP2-type zinc fingers that can recognize single-stranded RNA.” J Mol Biol 407(2): 273-83.

Niazi, S., M. Purohit and J. H. Niazi (2018). "Role of p53 circuitry in tumorigenesis: A brief review." Eur J Med Chem 158: 7-24.
Nicastro, G., I. A. Taylor and A. Ramos (2015). "KH-RNA interactions: back in the groove." Curr Opin Struct Biol 30: 63-70.
Nicholls, J. M., R. W. Chan, R. J. Russell, G. M. Air and J. S. Peiris (2008). "Evolving complexities of influenza virus and its receptors." Trends Microbiol 16(4): 149-157.
Nilsson-Payant, B. E., J. Sharps, N. Hengrung and E. Fodor (2018). "The Surface-Exposed PA(51-72)-Loop of the Influenza A Virus Polymerase Is Required for Viral Genome Replication." J Virol 92(16).
Noah, D. L., K. Y. Twu and R. M. Krug (2003). "Cellular antiviral responses against influenza A virus are countered at the posttranscriptional level by the viral NS1A protein via its binding to a cellular protein required for the 3' end processing of cellular pre-mRNAS." Virology 307(2): 386-395.
Noda. T. (2020). "Selective genome packaging mechanisms of influenza A viruses." Cold Spring Harb Perspect Med a038497.
Noda, T., Y. Sugita, K. Aoyama, A. Hirase, E. Kawakami, A. Miyazawa, H. Sagara and Y. Kawaoka (2012). "Three-dimensional analysis of ribonucleoprotein complexes in influenza A virus." Nat Commun 3: 639.
Nogales, A., T. Aydillo, G. Ávila-Pérez, A. Escalera, K. Chiem, R. Cadagan, M. L. DeDiego, F. Li, A. García-Sastre and L. Martínez-Sobrido (2019). "Functional characterization and direct comparison of influenza A, B, C and D NS1 proteins in vitro and in vivo." Front Microbiol. 10:2862.

O

Obayashi, E., H. Yoshida, F. Kawai, N. Shibayama, A. Kawaguchi, K. Nagata, J.R.H. Tame and S.Y. Park (2008). "The structural basis for an essential subunit interaction in influenza virus RNA polymerase." Nature 454(7208): 1127-31.
Ohkura, T., F. Momose, R. Ichikawa, K. Takeuchi and Y. Morikawa (2014). "Influenza A virus hemagglutinin and neuraminidase mutually accelerate their apical targeting through clustering of lipid rafts." J Virol 88(17): 10039-10055.
Okuno, Y., Y. Isegawa, F. Sasao and S. Ueda (1993). "A common neutralizing epitope conserved between the hemagglutinins of influenza A virus H1 and H2 strains." J Virol 67(5): 2552-2558.
Olson, W. K., Bansal, M., Burley, S. K., Dickerson, R. E., Gerstein, M., Harvey, S. C., Heinemann, U., Lu, X. J., Neidle, S., Shakked, Z., Sklenar, H., Suzuki, M., Tung, C. S., Westhof, E., Wolberger, C and H. M. Berman (2001). "A standard reference frame for the description of nucleic acid base-pair geometry." J Mol Biol 313(1): 229-37.
Omoto, S., V. Speranzini, T. Hashimoto, T. Noshi, H. Yamaguchi, M. Kawai, K. Kawaguchi, T. Uehara, T. Shishido, A. Naito and S. Cusack (2018). "Characterization of influenza virus variants induced by treatment with the endonuclease inhibitor baloxavir marboxil." Sci Rep 8(1): 9633.
O'Neill, R. E., J. Talon and P. Palese (1998). "The influenza virus NEP (NS2 protein) mediates the nuclear export of viral ribonucleoproteins." EMBO J 17(1): 288-296.
Ortega, J., J. Martin-Benito, T. Zurcher, J. M. Valpuesta, J. L. Carrascosa and J. Ortin (2000). "Ultrastructural and functional analyses of recombinant influenza virus ribonucleoproteins suggest dimerization of nucleoprotein during virus amplification." J Virol 74(1): 156-163.
Oshiumi, H., M. Matsumoto, S. Hatakeyama and T. Seya (2009). "Riplet/RNF135, a RING finger protein, ubiquitinates RIG-I to promote interferon-beta induction during the early phase of viral infection." J Biol Chem 284(2): 807-817.
Oshiumi, H., M. Miyashita, M. Matsumoto and T. Seya (2013). "A distinct role of Riplet-mediated K63-Linked polyubiquitination of the RIG-I repressor domain in human antiviral innate immune responses." PLoS Pathog 9(8): e1003533.
Osterhaus, A. D., G. F. Rimmelzwaan, B. E. Martina, T. M. Bestebroer and R. A. Fouchier (2000). "Influenza B virus in seals." Science 288(5468): 1051-1053.

P

Padilla-Quirarte, H. O., Badillo-Godinez, O., Gutierrez-Xicotencatl, L., Acevedo-Betancur, Y., Luna-Andon, J. D., Montiel-Hernandez, J. L., Lopez-Guerrero, D. V. and F. Esquivel-Guadarrama (2019). "Targeting M2 to DEC-205 induces an enhanced serum antibody-dependent heterosubtypic protection against influenza A virus infection." Vaccine 37(19): 2624-2633.
Palese, P. (1977). "The genes of influenza virus." Cell 10(1): 1-10.
Pan, D., Sun, H., Shen, Y., Liu, H. and X. Yao (2011). "Exploring the molecular basis of dsRNA recognition by NS1 protein of influenza A virus using molecular dynamics simulation and free energy calculation." Antiviral Res. 92(3): 424-33.
Pang, B., N. N. Cheung, W. Zhang, J. Dai, R. Y. Kao, H. Zhang and Q. Hao (2016). "Structural Characterization of H1N1 Nucleoprotein-Nucleozin Binding Sites." Sci Rep 6: 29684.

Panthu, B., O. Terrier, C. Carron, A. Traversier, A. Corbin, L. Balvay, B. Lina, M. Rosa-Calatrava and T. Ohlmann (2017). "The NS1 Protein from Influenza Virus Stimulates Translation Initiation by Enhancing Ribosome Recruitment to mRNAs." J Mol Biol 429(21): 3334-3352.
Park, Y. W. and M. G. Katze (1995). "Translational control by influenza virus. Identification of cis-acting sequences and trans-acting factors which may regulate selective viral mRNA translation." J Biol Chem 270(47): 28433-28439.
Patel, R. C., P. Stanton, N. M. McMillan, B. R. Williams and G. C. Sen (1995). "The interferon-inducible double-stranded RNA-activated protein kinase self-associates in vitro and in vivo." Proc Natl Acad Sci U S A 92(18): 8283-8287.
Paterson, D. and E. Fodor (2012). "Emerging roles for the influenza A virus nuclear export protein (NEP)." PLoS Pathog 8(12): e1003019.
Patnaik, S., D. Basu, N. Southall, S. Dehdashti, K. K. Wan, W. Zheng, M. Ferrer, M. Taylor, D. A. Engel and J. J. Marugan (2019). "Identification, design and synthesis of novel pyrazolopyridine influenza virus nonstructural protein 1 antagonists." Bioorg Med Chem Lett 29(9): 1113-1119.
Patnaik, S., D. Basu, S. Dehdashti, W. Zheng, M. Ferrer, N. Southall, M. Taylor, D. A. Engel and J. J. Marugan (2013). “Discovery of small molecule influenza virus NS1 antagonist.” Probe reports from the NIH molecular libraries program (internet).
Pauly, M.D., M.C. Procario and A.S. Lauring (2017). “A novel twelve class fluctuation test reveals higher than expected mutation rates for influenza A viruses.” Elife 6:e26437.
Peiris, J. S., W. C. Yu, C. W. Leung, C. Y. Cheung, W. F. Ng, J. M. Nicholls, T. K. Ng, K. H. Chan, S. T. Lai, W. L. Lim, K. Y. Yuen and Y. Guan (2004). "Re-emergence of fatal human influenza A subtype H5N1 disease." Lancet 363(9409): 617-619.
Pena, L., A. L. Vincent, C. L. Loving, J. N. Henningson, K. M. Lager, W. Li and D. R. Perez (2012). "Strain-dependent effects of PB1-F2 of triple-reassortant H3N2 influenza viruses in swine." J Gen Virol 93(Pt 10): 2204-2214.
Pena, L., A. L. Vincent, C. L. Loving, J. N. Henningson, K. M. Lager, A. Lorusso and D. R. Perez (2012). "Restored PB1-F2 in the 2009 pandemic H1N1 influenza virus has minimal effects in swine." J Virol 86(10): 5523-5532.
Peng, Q., Y. Liu, R. Peng, M. Wang, W. Yang, H. Song, Y. Chen, S. Liu, M. Han, X. Zhang, P. Wang, J. Yan, B. Zhang, J. Qi, T. Deng, G. F. Gao and Y. Shi (2019). "Structural insight into RNA synthesis by influenza D polymerase." Nat Microbiol.
Pereira, C. F., E. K. C. Read, H. M. Wise, M. J. Amorim and P. Digard (2017). "Influenza A Virus NS1 Protein Promotes Efficient Nuclear Export of Unspliced Viral M1 mRNA." J Virol 91(15).
Perez Rubio, A. and J. M. Eiros (2018). "Cell culture-derived flu vaccine: Present and future." Hum Vaccin Immunother 14(8): 1874-1882.
Perez, J.T., A. Varble., R. Sachidanandam., I. Zlatev., M. Manoharan., A. Garcia-Sastre and B.R. TenOever (2010). “Influenza A virus-generated small RNAs regulate the switch from transcription to replication.” Proc Natl Acad Sci 107(25):11525-30.
Pettersen, E.F., T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng and T.E. Ferrin (2004). “UCSF Chimera visualization system for exploratory research and analysis.” J Comput Chem 25, 1605-1612.
Pflug, A., D. Guilligay, S. Reich and S. Cusack (2014). "Structure of influenza A polymerase bound to the viral RNA promoter." Nature 516(7531): 355-360.
Pinar, A., J. K. Dowling, N. J. Bitto, A. A. Robertson, E. Latz, C. R. Stewart, G. R. Drummond, M. A. Cooper, J. L. McAuley, M. D. Tate and A. Mansell (2017). "PB1-F2 Peptide Derived from Avian Influenza A Virus H7N9 Induces Inflammation via Activation of the NLRP3 Inflammasome." J Biol Chem 292(3): 826-836.
Plaschka, C., A. J. Newman and K. Nagai (2019). "Structural Basis of Nuclear pre-mRNA Splicing: Lessons from Yeast." Cold Spring Harb Perspect Biol 11(5).
Pleschka, S., R. Jaskunas, O. G. Engelhardt, T. Zürcher, P. Palese and A. García-Sastre (1996). "A plasmid-based reverse genetics system for influenza A virus." J Virol 70(6): 4188-92.
Poon, L. L., E. Fodor and G. G. Brownlee (2000). "Polyuridylated mRNA synthesized by a recombinant influenza virus is defective in nuclear export." J Virol 74(1): 418-427.
Prins, K. C., W. B. Cardenas and C. F. Basler (2009). "Ebola virus protein VP35 impairs the function of interferon regulatory factor-activating kinases IKKepsilon and TBK-1." J Virol 83(7): 3069-3077.

Q

Qian, X. Y., F. Alonso-Caplen and R. M. Krug (1994). “Two functional domains of the influenza virus NS1 protein are required for regulation of nuclear export of mRNA.” J. Virol. 68(4): 2433-41.
Qian, X. Y., C. Y. Chien, Y. Lu, G. T. Montelione and R. M. Krug (1995). "An amino-terminal polypeptide fragment of the influenza virus NS1 protein possesses specific RNA-binding activity and largely helical backbone structure." RNA 1(9): 948-56.

Qin, S., Y. Liu, W. Tempel, M. S. Eram, C. Bian, K. Liu, G. Senisterra, L. Crombet, M. Vedadi and J. Min (2014). "Structural basis for histone mimicry and hijacking of host proteins by influenza virus protein NS1". Nat Commun 5: 3952.
Qiu, Y. and R. M. Krug (1994). "The influenza virus NS1 protein is a poly(A)-binding protein that inhibits nuclear export of mRNAs containing poly(A)." J Virol 68(4): 2425-2432.
Qiu, Y., M. Nemeroff and R. M. Krug (1995). "The influenza virus NS1 protein binds to a specific region in human U6 snRNA and inhibits U6-U2 and U6-U4 snRNA interactions during splicing." RNA 1(3): 304-316.

R

Rabadan, R., A. J. Levine and M. Krasnitz (2008). "Non-random reassortment in human influenza A viruses." Influenza Other Respir Viruses 2(1): 9-22.
Rabadan, R., A. J. Levine and H. Robins (2006). "Comparison of avian and human influenza A viruses reveals a mutational bias on the viral genomes." J Virol 80(23): 11887-11891.
Rajik, M., F. Jahanshiri, A. R. Omar, A. Ideris, S. S. Hassan and K. Yusoff (2009). "Identification and characterisation of a novel anti-viral peptide against avian influenza virus H9N2." Virol J 6: 74.
Rajsbaum, R., R. A. Albrecht, M. K. Wang, N. P. Maharaj, G. A. Versteeg, E. Nistal-Villan, A. Garcia-Sastre and M. U. Gack (2012). "Species-specific inhibition of RIG-I ubiquitination and IFN induction by the influenza A virus NS1 protein." PLoS Pathog 8(11): e1003059.
Reich, S., D. Guilligay, A. Pflug, H. Malet, I. Berger, T. Crepin, D. Hart, T. Lunardi, M. Nanao, R. W. Ruigrok and S. Cusack (2014). "Structural insight into cap-snatching and RNA synthesis by influenza polymerase." Nature 516(7531): 361-366.
Reich, S., D. Guilligay and S. Cusack (2017). “An in vitro fluorescence based study of initiation of RNA synthesis by influenza B polymerase.” Nucleic Acid Res 45(6): 3353-3368.
Resa-Infante, P., N. Jorba, R. Coloma and J. Ortin (2011). "The influenza virus RNA synthesis machine: advances in its structure and function." RNA Biol 8(2): 207-215.
Rice, P., I. Longden and A. Bleasby (2000). “EMBOSS: The European Molecular Biology Open Software Suite.” Trends Genet 16(6):276-7.
Rigby, R. E., H. M. Wise, N. Smith, P. Digard and J. Rehwinkel (2019). "PA-X antagonises MAVS-dependent accumulation of early type I interferon messenger RNAs during influenza A virus infection." Sci Rep 9(1): 7216.
Robb, N. C., G. Chase, K. Bier, F. T. Vreede, P. C. Shaw, N. Naffakh, M. Schwemmle and E. Fodor (2011). "The influenza A virus NS1 protein interacts with the nucleoprotein of viral ribonucleoprotein complexes." J Virol 85(10): 5228-5231.
Robb, N. C. and E. Fodor (2012). "The accumulation of influenza A virus segment 7 spliced mRNAs is regulated by the NS1 protein." J Gen Virol 93(Pt 1): 113-118.
Robb, N. C., D. Jackson, F. T. Vreede and E. Fodor (2010). "Splicing of influenza A virus NS1 mRNA is independent of the viral NS1 protein." J Gen Virol 91(Pt 9): 2331-2340.
Rosário-Ferreira, N., A.J. Preto, R. Melo, I.S. Moreira and R.M.M. Brito (2020). “The central role of non-structural protein 1 (NS1) in influenza biology and infection.” Int J Mol Sci 21(4): 1511.
Rossman, J. S., G. P. Leser and R. A. Lamb (2012). "Filamentous influenza virus enters cells via macropinocytosis." J Virol 86(20): 10950-10960.
Rubinson, E. H. and B. F. Eichman (2012). “Nucleic acid recognition by tandem helical repeats.” Curr Opin Struct Biol. 22(1): 101-9.
Russell, R. J., P. S. Kerry, D. J. Stevens, D. A. Steinhauer, S. R. Martin, S. J. Gamblin and J. J. Skehel (2008). "Structure of influenza hemagglutinin in complex with an inhibitor of membrane fusion." Proc Natl Acad Sci U S A 105(46): 17736-17741.
Ryder, S. P., Recht, M. I. and J. R. Williamson (2008). "Quantitative analysis of protein-RNA interactions by gel mobility shift." Methods Mol Biol 488: 99-115.
Ryter, J. M. and S. C. Schultz (1998). "Molecular basis of double-stranded RNA-protein interactions: structure of a dsRNA-binding domain complexed with dsRNA." EMBO J 17(24): 7505-7513.

S

Sadler, A. J. and B. R. Williams (2008). "Interferon-inducible antiviral effectors." Nat Rev Immunol 8(7): 559-568.
Sagawa, H., A. Ohshima, I. Kato, Y. Okuno and Y. Isegawa (1996). "The immunological activity of a deletion mutant of influenza virus haemagglutinin lacking the globular region." J Gen Virol 77 (Pt 7): 1483-1487.
Salvadore, M., C. F. Basler, J. P. Parisien, C. M. Horvath, S. Bourmakina, H. Zheng, T. Muster, P. Palese and A. García-Sastre (2002). “Effects of influenza A virus NS1 protein on protein expression: the NS1 protein enhances translation and is not required for shutoff of host protein synthesis.” J Virol. 76(3): 1206-12.
Sanchez-Aparicio, M. T., J. Ayllon, A. Leo-Macias, T. Wolff and A. Garcia-Sastre (2017). "Subcellular Localizations of RIG-I, TRIM25, and MAVS Complexes." J Virol 91(2).

Satterly, N., P. L. Tsai, J. van Deursen, D. R. Nussenzveig, Y. Wang, P. A. Faria, A. Levay, D. E. Levy and B. M. Fontoura (2007). "Influenza virus targets the mRNA export machinery and the nuclear pore complex." Proc Natl Acad Sci U S A 104(6): 1853-1858.
Saunders, L. R. and G. N. Barber (2003). "The dsRNA binding protein family: critical roles, diverse cellular functions." FASEB J 17(9): 961-983.
Schierhorn, K. L., F. Jolmes, J. Bespalowa, S. Saenger, C. Peteranderl, J. Dzieciolowski, M. Mielke, M. Budt, S. Pleschka, A. Herrmann, S. Herold and T. Wolff (2017). "Influenza A Virus Virulence Depends on Two Amino Acids in the N-Terminal Domain of Its NS1 Protein To Facilitate Inhibition of the RNA-Dependent Protein Kinase PKR." J Virol 91(10).
Schoggins, J. W. (2014). "Interferon-stimulated genes: roles in viral pathogenesis." Curr Opin Virol 6: 40-46.
Segura Guerrero, N. A., S. Sharma, J. Neyts and S. J. F. Kaptein (2018). "Favipiravir inhibits in vitro Usutu virus replication and delays disease progression in an infection model in mice." Antiviral Res 160: 137-142.
Seitz, C., T. Frensing, D. Hoper, G. Kochs and U. Reichl (2010). "High yields of influenza A virus in Madin-Darby canine kidney cells are promoted by an insufficient interferon-induced antiviral state." J Gen Virol 91(Pt 7): 1754-1763.
Selman, M., S. K. Dankar, N. E. Forbes, J. J. Jia and E. G. Brown (2012). "Adaptive mutation in influenza A virus non-structural gene is linked to host switching and induces a novel protein by alternative splicing." Emerg Microbes Infect 1(11): e42.
Shapira S.D., I. Gat-Viks, B.O.V. Shum, A. Dricot, M.M. de Grace, L. Wu, P.B. Gupta, T. Hao, S.J. Silver, D.E. Root, D.E. Hill, A. Regev and N. Hacohen (2009). "A Physical and Regulatory Map of Host-Influenza Interactions Reveals Pathways in H1N1 Infection" Cell 139(7): 1255-67.
Shen, Q., D. Zeng, B. Zhao, V. S. Bhatt, P. Li and J. H. Cho (2017). "The molecular mechanisms underlying the hijack of host proteins by the 1918 spanish influenza virus." ACS Chem Biol 12(5); 1199-1203.
Shimizu, K., H. Handa, S. Nakada and K. Nagata (1994). "Regulation of influenza virus RNA polymerase activity by cellular and viral factors." Nucleic Acids Res 22(23): 5047-5053.
Shinya, K., M. Ebina, S. Yamada, M. Ono, N. Kasai and Y. Kawaoka (2006). "Avian flu: influenza virus receptors in the human airway." Nature 440(7083): 435-436.
Shinya, K. and Y. Kawaoka (2006). "[Influenza virus receptors in the human airway]." Uirusu 56(1): 85-89.
Siche, S., K. Brett, L. Moller, L. V. Kordyukova, R. R. Mintaev, A. V. Alexeevski and M. Veit (2015). "Two Cytoplasmic Acylation Sites and an Adjacent Hydrophobic Residue, but No Other Conserved Amino Acids in the Cytoplasmic Tail of HA from Influenza A Virus Are Crucial for Virus Replication." Viruses 7(12): 6458-6475.
Sieczkarski, S. B. and G. R. Whittaker (2002). "Influenza virus can enter and infect cells in the absence of clathrin-mediated endocytosis." J Virol 76(20): 10455-10464.
Sikora, D., L. Rocheleau, E. G. Brown and M. Pelchat (2014). "Deep sequencing reveals the eight facets of the influenza A/HongKong/1/1968 (H3N2) virus cap-snatching process." Sci Rep 4: 6181.
Sikora, D., L. Rocheleau, E. G. Brown and M. Pelchat (2017). "Influenza A virus cap-snatches host RNAs based on their abundance early after infection." Virology 509: 167-177.
Singh, R. N. and N. N. Singh (2019). "A novel role of U1 snRNP: Splice site selection from a distance." Biochim Biophys Acta Gene Regul Mech 1862(6): 634-642.
Skorko, R., D. F. Summers and J. M. Galarza (1991). "Influenza A virus in vitro transcription: roles of NS1 and NP proteins in regulating RNA synthesis." Virology 180(2): 668-677.
Smith, B. J., P. M. Colman, M. Von Itzstein, B. Danylec and J. N. Varghese (2001). "Analysis of inhibitor binding in influenza virus neuraminidase." Protein Sci 10(4): 689-696.
Smith, G. J., D. Vijaykrishna, J. Bahl, S. J. Lycett, M. Worobey, O. G. Pybus, S. K. Ma, C. L. Cheung, J. Raghwani, S. Bhatt, J. S. Peiris, Y. Guan and A. Rambaut (2009). "Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic." Nature 459(7250): 1122-1125.
Someya, T., K. Hosono, K. Morimura, H. Takaku and G. Kawai (2008). "Recognition of a bulged RNA by peptides derived from the influenza NS1 protein." J Biochem 143(3): 339-347.
Sonenberg, N. and Hinnebusch, A.G (2009). "Rgulation of translation initiation in eucaryotes: mechanisms and biological targets." Cell 136(4): 731-45.
Soubies, S. M., C. Volmer, G. Croville, J. Loupias, B. Peralta, P. Costes, C. Lacroux, J. L. Guerin and R. Volmer (2010). "Species-specific contribution of the four C-terminal amino acids of influenza A virus NS1 protein to virulence." J Virol 84(13): 6733-6747.
Stadlbauer, D., Zhu, X., McMahon, M., Turner, J. S., Wohlbold, T. J., Schmitz, A. J., Strohmeier, S., Yu, W., Nachbagauer, R., Mudd, P. A., Wilson, I. A., Ellebedy, A. H. and F. Krammer (2019). "Broadly protective human antibodies that target the active site of influenza virus neuraminidase." Science 366(6464): 499-504.
Stefl. R., F.C. Oberstrass, J.L. Hood, M. Jourdan, M. Zimmermann, L. Skrisovska, C. Maris, L. Peng, C. Hofr, R.B. Emeson and F. H. Allain (2010). "The solution structure of the ADAR2 dsRBM-RNA complex reveals a sequence-specific read out of the RNA minor groove." Cell 143(2): 225-237.

Stefl, R., L. Skrisovska and F. H. Allain (2005). "RNA sequence- and shape-dependent recognition by proteins in the ribonucleoprotein particle." EMBO Rep 6(1): 33-38.
Steidle, S., L. Martinez-Sobrido, M. Mordstein, S. Lienenklaus, A. Garcia-Sastre, P. Staheli and G. Kochs (2010). "Glycine 184 in nonstructural protein NS1 determines the virulence of influenza A virus strain PR8 without affecting the host interferon response." J Virol 84(24): 12761-12770.
Studier, F. W. (2005). "Protein production by auto-induction in high density shaking cultures." Protein Expr Purif 41(1): 207-234.
Su, W.C., Y.C. Chen, C.H. Tseng, P.W.C. Hsu, K.F. Tung, K.S. Jeng and M.M.C. Lai (2013). "Pooled RNAi Screen Identifies Ubiquitin Ligase Itch as Crucial for Influenza A Virus Release From the Endosome During Virus Entry." Proc Natl Acad Sci 110(43):17516-21.
Suarez, D. L. and M. L. Perdue (1998). "Multiple alignment comparison of the non-structural genes of influenza A viruses." Virus Res 54(1): 59-69.
Sugiyama, K., E. Obayashi, A. Kawaguchi, Y. Suzuki, J. R. Tame, K. Nagata and S. Y. Park (2009). "Structural insight into the essential PB1-PB2 subunit contact of the influenza virus RNA polymerase." EMBO J 28(12): 1803-1811.
Sui, B., D. Bamba, K. Weng, H. Ung, S. Chang, J.V. Dyke, M. Goldblatt, R. Duan, M.S. Kinch and W.B. Li (2009). "The Use of Random Homozygous Gene Perturbation to Identify Novel Host-Oriented Targets for Influenza" Virology 387(2):473-81.
Sun, X., J. Song, E. Li, H. Geng, Y. Li, D. Yu and C. Zhong (2019). "Epigallocatechin3gallate inhibits bladder cancer stem cells via suppression of sonic hedgehog pathway." Oncol Rep 42(1): 425-435.
Sun, Y., Zhang, Y., Hamilton, K., Manley, J. L., Shi, Y., Walz, T. and L. Tong (2018). "Structural basis for the recognition of the human AAUAAA polyadenylation signal." Proc Natl Acad Sci 115(7): E1419-E1428.
Suttie, A., Y. M. Deng, A. R. Greenhill, P. Dussart, P. F. Horwood and E. A. Karlsson (2019). "Inventory of molecular markers affecting biological characteristics of avian influenza A viruses." Virus Genes.
Swale, C., A. Monod, L. Tengo, A. Labaronne, F. Garzoni, J. M. Bourhis, S. Cusack, G. Schoehn, I. Berger, R. W. Ruigrok and T. Crepin (2016). "Structural characterization of recombinant IAV polymerase reveals a stable complex between viral PA-PB1 heterodimer and host RanBP5." Sci Rep 6: 24727.
Swale, C., B. Da Costa, L. Sedano, F. Garzoni, A. A. McCarthy, I. Berger, C. Bieniossek, R. W. H. Ruigrok, B. Delmas and T. Crepin (2020). "X-ray structure of the human karyopherin RanBP5, an essential factor for influenza polymerase nuclear trafficking." J Mol Biol 432(10): 3353-3359.
Swatek, K. N. and D. Komander (2016). "Ubiquitin modifications." Cell Res 26(4): 399-422.
Szewczyk, B., K. Bienkowska-Szewczyk and E. Krol (2014). "Introduction to molecular biology of influenza A viruses." Acta Biochim Pol 61(3): 397-401.

T

Tang, T.T.L., J.A.W. Stowell, C.H. Hill and L.A. Passmore (2019). "The intrinsic structure of poly(A) RNA determines the specificity of Pan2 and Caf1 deadenylases." Nat Struct Mol Biol 26(6): 433-442.
Taniguchi, K., Y. Ando, H. Nobori, S. Toba, T. Noshi, M. Kobayashi, M. Kawai, R. Yoshida, A. Sato, T. Shishido, A. Naito, K. Matsuno, M. Okamatsu, Y. Sakoda and H. Kida (2019). "Inhibition of avian-origin influenza A(H7N9) virus by the novel cap-dependent endonuclease inhibitor baloxavir marboxil." Sci Rep 9(1): 3466.
Tarakhovsky, A. and R.K. Prinjha (2018). "Drawing on disorder: How viruses use histone mimicry to their advantage." J Exp Med 215(7):1777-1787.
Te Velthuis, A. J. and E. Fodor (2016). "Influenza virus RNA polymerase: insights into the mechanisms of viral RNA synthesis." Nat Rev Microbiol 14(8): 479-493.
Te Velthuis, A. J. W. and J. Oymans (2018). "Initiation, Elongation, and Realignment during Influenza Virus mRNA Synthesis." J Virol 92(3).
Terrier, O., A. Diederichs, J. Dubois, G. Cartet, B. Lina, J. C. Bourdon and M. Rosa-Calatrava (2013). "Influenza NS1 interacts with p53 and alters its binding to p53-responsive genes, in a promoter-dependent manner." FEBS Lett 587(18): 2965-2971.
Timmey, B.L., B. Raveh, R. Mironska, J.M. Trivedi, S.J. Kim, D. Russel, S.R. Wente, A. Sali and M.P. Rout (2016). "Simple rules for passive diffusion through the nuclear pore complex." J Cell Biol 215(1): 57-76.
Theobald, D. L., Mitton-Fry, R. M. and D. S. Wuttke (2003). "Nucleic acid recognition by OB-fold proteins." Annu Rev Biophys Biomol Struct. 32: 115-33.
Tewawong, N., S. Prachayangprecha, P. Vichiwattana, S. Korkong, S. Klinfueng, S. Vongpunsawad, T. Thongmee, A. Theamboonlers and Y. Poovorawan (2015). "Assessing Antigenic Drift of Seasonal Influenza A(H3N2) and A(H1N1)pdm09 Viruses." PLoS One 10(10): e0139958.
Tian, B., P. C. Bevilacqua, A. Diegelman-Parente and M. B. Mathews (2004). "The double-stranded-RNA-binding motif: interference and much more." Nat Rev Mol Cell Biol 5(12): 1013-1023.
To, J. and J. Torres (2019). "Viroporins in the Influenza Virus." Cells 8(7).

Tomassini, J., H. Selnick, M. E. Davies, M. E. Armstrong, J. Baldwin, M. Bourgeois, J. Hastings, D. Hazuda, J. Lewis, W. McClements and *et al*. (1994). "Inhibition of cap (m7GpppXm)-dependent endonuclease of influenza virus by 4-substituted 2,4-dioxobutanoic acid compounds." Antimicrob Agents Chemother 38(12): 2827-2837.
Tomescu, A. I., N. C. Robb, N. Hengrung, E. Fodor and A. N. Kapanidis (2014). "Single-molecule FRET reveals a corkscrew RNA structure for the polymerase-bound influenza virus promoter." Proc Natl Acad Sci U S A 111(32): E3335-3342.
Tong, S., Y. Li, P. Rivailler, C. Conrardy, D. A. Castillo, L. M. Chen, S. Recuenco, J. A. Ellison, C. T. Davis, I. A. York, A. S. Turmelle, D. Moran, S. Rogers, M. Shi, Y. Tao, M. R. Weil, K. Tang, L. A. Rowe, S. Sammons, X. Xu, M. Frace, K. A. Lindblade, N. J. Cox, L. J. Anderson, C. E. Rupprecht and R. O. Donis (2012). "A distinct lineage of influenza A virus from bats." Proc Natl Acad Sci U S A 109(11): 4269-4274.
Tong, S., X. Zhu, Y. Li, M. Shi, J. Zhang, M. Bourgeois, H. Yang, X. Chen, S. Recuenco, J. Gomez, L. M. Chen, A. Johnson, Y. Tao, C. Dreyfus, W. Yu, R. McBride, P. J. Carney, A. T. Gilbert, J. Chang, Z. Guo, C. T. Davis, J. C. Paulson, J. Stevens, C. E. Rupprecht, E. C. Holmes, I. A. Wilson and R. O. Donis (2013). "New world bats harbor diverse influenza A viruses." PLoS Pathog 9(10): e1003657.
Trapp, S., D. Soubieux, A. Lidove, E. Esnault, A. Lion, V. Guillory, A. Wacquiez, E. Kut, P. Quere, T. Larcher, M. Ledevin, V. Nadan, C. Camus-Bouclainville and D. Marc (2018). "Major contribution of the RNA-binding domain of NS1 in the pathogenicity and replication potential of an avian H7N1 influenza virus in chickens." Virol J 15(1): 55.
Trapp, S., D. Soubieux, H. Marty, E. Esnault, T. W. Hoffmann, M. Chandenier, A. Lion, E. Kut, P. Quere, T. Larcher, M. Ledevin, S. Munier, N. Naffakh and D. Marc (2014). "Shortening the unstructured, interdomain region of the non-structural protein NS1 of an avian H1N1 influenza virus increases its replication and pathogenicity in chickens." J Gen Virol 95(Pt 6): 1233-1243.
Treanor, J. J., M. H. Snyder, W. T. London and B. R. Murphy (1989). "The B allele of the NS gene of avian influenza viruses, but not the A allele, attenuates a human influenza A virus for squirrel monkeys." Virology 171(1): 1-9.
Tsurumura, T., H. Qiu, T. Yoshida, Y. Tsumori, D. Hatakeyama, T. Kuzuhara and H. Tsuge (2013). "Conformational polymorphism of m7GTP in crystal structure of the PB2 middle domain from human influenza A virus." PLoS One 8(11): e82020.
Tu, J., J. Guo, A. Zhang, W. Zhang, Z. Zhao, H. Zhou, C. Liu, H. Chen and M. Jin (2011). "Effects of the C-terminal truncation in NS1 protein of the 2009 pandemic H1N1 influenza virus on host gene expression." PLoS One 6(10): e26175.
Tumpey, T. M., Basler, C. F., Aguilar, P. V., Zeng, H., Solórzano, A., Swayne, D. E., Cox, N. J., Katz, J. M., Taubenberger, J. K. Palese, P. and A. García-Sastre (2005). "Characterization of the reconstructed 1918 Spanich influenza pandemic virus." Science 310(5745): 77-80.
Turkington, H. L., M. Juozapaitis, P. S. Kerry, T. Aydillo, J. Ayllon, A. Garcia-Sastre, M. Schwemmle and B. G. Hale (2015). "Novel Bat Influenza Virus NS1 Proteins Bind Double-Stranded RNA and Antagonize Host Innate Immunity." J Virol 89(20): 10696-10701.
Turnbull, M. L., H. M. Wise, M. Q. Nicol, N. Smith, R. L. Dunfee, P. M. Beard, B. W. Jagger, Y. Ligertwood, G. R. Hardisty, H. Xiao, D. J. Benton, A. M. Coburn, J. A. Paulo, S. P. Gygi, J. W. McCauley, J. K. Taubenberger, S. J. Lycett, M. P. Weekes, B. M. Dutia and P. Digard (2016). "Role of the B Allele of Influenza A Virus Segment 8 in Setting Mammalian Host Range and Pathogenicity." J Virol 90(20): 9263-9284.
Twu, K. Y., D. L. Noah, P. Rao, R. L. Kuo and R. M. Krug (2006). "The CPSF30 binding site on the NS1A protein of influenza A virus is a potential antiviral target." J Virol 80(8): 3957-3965.

V

Van Dongen, M. J. P., R. U. Kadam, J. Juraszek, E. Lawson, B. Brandenburg, F. Schmitz, W. B. G. Schepens, B. Stoops, H. A. van Diepen, M. Jongeneelen, C. Tang, J. Vermond, A. van Eijgen-Obregoso Real, S. Blokland, D. Garg, W. Yu, W. Goutier, E. Lanckacker, J. M. Klap, D. C. G. Peeters, J. Wu, C. Buyck, T. H. M. Jonckers, D. Roymans, P. Roevens, R. Vogels, W. Koudstaal, R. H. E. Friesen, P. Raboisson, D. Dhanak, J. Goudsmit, I. A. Wilson (2019). “A small-molecule fusion inhibitor of influenza virus is orally active in mice.” Science 363(6431). pii: eaar6221.
Varga, Z. T., A. Grant, B. Manicassamy and P. Palese (2012). "Influenza virus protein PB1-F2 inhibits the induction of type I interferon by binding to MAVS and decreasing mitochondrial membrane potential." J Virol 86(16): 8359-8366.
Varghese, J. N., P. M. Colman, A. van Donkelaar, T. J. Blick, A. Sahasrabudhe and J. L. McKimm-Breschkin (1997). “Structural evidence for a second sialic acid binding site in avian influenza virus neuraminidases.” Proc Natl Acad Sci 94(22): 11808-12.
Varshavsky, A. (1995). "The N-end rule." Cold Spring Harb Symp Quant Biol 60: 461-478.

Vavricka, C. J., Q. Li, Y. Wu, J. Qi, M. Wang, Y. Liu, F. Gao, J. Liu, E. Feng, J. He, J. Wang, H. Liu, H. Jiang and G. F. Gao (2011). "Structural and functional analysis of laninamivir and its octanoate prodrug reveals group specific mechanisms for influenza NA inhibition." PLoS Pathog 7(10): e1002249.
Venkatesh, D., M. J. Poen, T. M. Bestebroer, R. D. Scheuer, O. Vuong, M. Chkhaidze, A. Machablishvili, J. Mamuchadze, L. Ninua, N. B. Fedorova, R. A. Halpin, X. Lin, A. Ransier, T. B. Stockwell, D. E. Wentworth, D. Kriti, J. Dutta, H. van Bakel, A. Puranik, M. J. Slomka, S. Essen, I. H. Brown, R. A. M. Fouchier and N. S. Lewis (2018). "Avian Influenza Viruses in Wild Birds: Virus Evolution in a Multihost Ecosystem." J Virol 92(15).

W

Walker, A. P. and E. Fodor (2019). "Interplay between Influenza Virus and the Host RNA Polymerase II Transcriptional Machinery." Trends Microbiol 27(5): 398-407.
Walkiewicz, M. P., D. Basu, J. J. Jablonski, H. M. Geysen and D. A. Engel (2011). "Novel inhibitor of influenza non-structural protein 1 blocks multi-cycle replication in an RNase L-dependent manner." J Gen Virol 92(Pt 1): 60-70.
Wandtke, T., J. Wozniak and P. Kopinski (2015). “Aptamers in diagnostics and treatment of viral infections.” Viruses 7(2): 751-80.
Wandzik, J.M., T. Kouba, M. Karuppasamy, A. Pflug, P. Drncova, J. Provaznik, N. Azevedo and S. Cusack (2020). “A structure-based model of the complex transcription cycle of influenza polymerase.” Cell 181(4): 877-893.
Wang, J., X. Qi and C. Lu (2012). "Mutations in the C-terminal tail of NS1 protein facilitate the replication of classical swine H1N1 influenza A virus in mice." Folia Microbiol (Praha) 57(3): 169-175.
Wang, Q., Q. Li, T. Liu, G. Chang, Z. Sun, Z. Gao, F. Wang, H. Zhou, R. Liu, M. Zheng, H. Cui, G. Chen, H. Li, X. Yuan, J. Wen, D. Peng and G. Zhao (2018). "Host Interaction Analysis of PA-N155 and PA-N182 in Chicken Cells Reveals an Essential Role of UBA52 for Replication of H5N1 Avian Influenza Virus." Front Microbiol 9: 936.
Wang, Q., R. Liu, Q. Li, F. Wang, B. Zhu, M. Zheng, H. Cui, J. Wen and G. Zhao (2019). "Host cell interactome of PB1 N40 protein of H5N1 influenza A virus in chicken cells." J Proteomics 197: 34-41.
Wang, S., Y. Hu, M. T. Overgaard, F. V. Farginov, O. C. Uhlenbeck and D. B. McKay (2006). ”The domain of the Bacillus substilis DEAD-box helicase YxiN that is responsible for specific binding of 23S rRNA has an RNA recognition motif fold.” RNA 12(6): 959-67.
Wang, W. and R. M. Krug (1998). "U6atac snRNA, the highly divergent counterpart of U6 snRNA, is the specific target that mediates inhibition of AT-AC splicing by the influenza virus NS1 protein." RNA 4(1): 55-64.
Wang, W., K. Riedel, P. Lynch, C. Y. Chien, G. T. Montelione and R. M. Krug (1999). "RNA binding by the novel helical domain of the influenza virus NS1 protein requires its dimer structure and a small number of specific basic amino acids." RNA 5(2): 195-205.
Wang, X., L. Lin, Y. Yu, Y. Yan, N.K. Ojha and J. Zhou (2020). “The N-terminal residual arginine 19 of influenza A virus NS1 protein is required for its nuclear localization and RNA binding.” Vet Microbiol 251:108895.
Wang, X., Y. Shen, Y. Qiu, Z. Shi, D. Shao, P. Chen, G. Tong and Z. Ma (2010). "The non-structural (NS1) protein of influenza A virus associates with p53 and inhibits p53-mediated transcriptional activity and apoptosis." Biochem Biophys Res Commun 395(1): 141-145.
Wang, Y., Y. Zheng and P. A. Beal (2017). "Adenosine Deaminases That Act on RNA (ADARs)." Enzymes 41: 215-268.
Ward, A. C., A. A. Azad and I. G. Macreadie (1994). "Expression and characterisation of the influenza A virus non-structural protein NS1 in yeast." Arch Virol 138(3-4): 299-314.
Ward, B.J., A. Makarkov, A. Séguin, S. Pillet, S. Trépanier, J. Dhaliwall, M.D. Libman, T. Vesikari and N Landry (2020). “Efficacy, immunogenicity, and safety of a plant-derived, quadrivalent, virus-like particle influenza vaccine in adults (18-64 years) and older adults (≥65 years): two multicentre, randomised phase 3 trials.” Lancet 396(10261): 1491-1503.
Ward, S.E., H.S. Kim, K. Komurov, S. Mendirata, P.L. Tsai, M. Schmolke, N. Satterly, B. Manicassamy, C.V. Forst, M.G. Roth, A. García-Sastre, K.M. Blazewska, C.E. McKenna, B.M. Fontoura and M.A. White (2012). “Host modulators of H1N1 cytopathogenicity.“ PLoS One 7(8): e39284.
Watanabe, T., E. Kawakami, J.E. Shoemaker, T.J.S. Lopes, Y. Matsuoka, Y. Tomita, H. Kozuka-Hata, T. Gorai, T. Kuwahara, E. Takeda, A. Nagata, R. Takano, M. Kiso, M. Yamashita, Y. Sakai-Tagawa, H. Katsura, N. Nonaka, H. Fujii, K. Fujii, Y. Sugita, T. Noda, H. Goto, S. Fukuyama, S. Watanabe, G. Neumann, M. Oyama, H. Kitano and Y. O. Kawaoka (2014). “Influenza virus-host interactome screen as a platform for antiviral drug development.“ Cell Host Microbe 16(6): 795-805.
Watanabe, T. and Y. Kawaoka (2017). “Influenza virus-host interactomes as a basis for antiviral drug development.“ Curr Opin virol 14: 71-78.
Way, G., Z. Xiong, G. Wang, H. Dai, S. Zheng, A. Garcia-Sastre and J. Liao (2020). “A novel SUMOylation ite in the influenza A virus NS1 protein identified with a highly sensitive FRET assay.” J Biotechnol 323: 121-127.

Weber-Gerlach, M. and F. Weber (2016). "Standing on three legs: antiviral activities of RIG-I against influenza viruses." Curr Opin Immunol 42: 71-75.
Westhof, E. and M. Sundaralingam (1983). "A method for the analysis of puckering disorder in five-membered rings: the relative mobilities of furanose and proline rings and their effects on polynucleotide and polypeptide backbone flexibility." J Am Chem Soc 105: 970–976.
Whittle, J. R., Zhang, R., Khurana, S., King, L. R., Manischewitz, J., Golding, H., Dormitzer, P. R., Haynes, B. F., Walter, E. B., Moody, M. A., Kepler, T. B., Liao, H. X. and S. C. Harrison (2011). "Broadly neutralizing human antibody that recognizes the receptor-binding pocket of influenza virus hemagglutinin." Proc Natl Acad Sci U S A 108(34): 14216-14221.
Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A. K. S. Wilson (2011). "Overview of the CCP4 suite and current developments". Acta. Cryst. D67: 235-242.
Wise, H. M., Barbezange, C., Jagger, B. W., Dalton, R. M., Gog, J. R., Curran, M. D., Taubenberger, J. K., Anderson, E. C. and P. Digard (2011). "Overlapping signals for translational regulation and packaging of influenza A virus segment 2." Nucleic Acids Res 39(17): 7775-7790.
Wise, H. M., Hutchinson, E. C., Jagger, B. W., Stuart, A. D., Kang, Z. H., Robb, N., Schwartzman, L. M., Kash, J. C., Fodor, E., Firth, A. E., Gog, J. R., Taubenberger, J. K. and P. Digard (2012). "Identification of a novel splice variant form of the influenza A virus M2 ion channel with an antigenically distinct ectodomain." PLoS Pathog 8(11): e1002998.
Wohlbold, T. J., Nachbagauer, R., Margine, I., Tan, G. S., Hirsh, A. and F. Krammer (2015). "Vaccination with soluble headless hemagglutinin protects mice from challenge with divergent influenza viruses." Vaccine 33(29): 3314-3321.
Woo, H. M., K. S. Kim, J. M. Lee, H. S. Shim, S. J. Cho, W. K. Lee, H. W. Ko, Y. S. Keum, S. Y. Kim, P. Pathinayake, C. J. Kim and Y. J. Jeong (2013). “Single-stranded DNA aptamer that specifically inds to the influenza NS1 protein suppresses interferon antagonism.“ Antiviral Res. 100(2): 337-45.
Woo, H. M., J. M. Lee, C. J. Kim, L. S. Lee and Y. J. Jeong (2019). "Recovery of TRIM25-mediated RIG-I ubiquitination through suppression of NS1 by RNA aptamers.” Mol Cells 42(10): 721-728.
Wu, J. and A. Akhmanova (2017). "Microtubule-Organizing Centers." Annu Rev Cell Dev Biol 33: 51-75.
Wu, N. C., Young, A. P., Al-Mawsawi, L. Q., Olson, C. A., Feng, J., Qi, H., Luan, H. H., Li, X., Wu, T. T. and R. Sun (2014). "High-throughput identification of loss-of-function mutations for anti-interferon activity in the influenza A virus NS segment." J Virol 88(17): 10157-10164.
Wu, Y., Zhang, G., Li, Y., Jin, Y., Dale, R., Sun, L. Q. and M. Wang (2008). "Inhibition of highly pathogenic avian H5N1 influenza virus replication by RNA oligonucleotides targeting NS1 gene." Biochem Biophys Res Commun 365(2): 369-374.
Wunderlich, K., Juozapaitis, M., Ranadheera, C., Kessler, U., Martin, A., Eisel, J., Beutling, U., Frank, R. and M. Schwemmle (2011). "Identification of high-affinity PB1-derived peptides with enhanced affinity to the PA protein of influenza A virus polymerase." Antimicrob Agents Chemother 55(2): 696-702.
Xia, S., Monzingo, A. F. and J. D. Robertus (2009). "Structure of NS1A effector domain from the influenza A/Udorn/72 virus." Acta Crystallogr D Biol Crystallogr 65(Pt 1): 11-7.
Xia, S. and J. D. Robertus (2010). "X-ray structures of NS1 effector domain mutants." Arch Biochem Biophys 494(2): 198-204.

Y

Yamashita, S., T. Nagata, M. Kawazoe, C. Takemoto, T. Kigawa, P. Guntert, N. Kobayashi, T. Terada, M. Shirouzu, M. Wakiyama, Y. Muto and S. Yokoyama (2011). "Structures of the first and second double-stranded RNA-binding domains of human TAR RNA-binding protein." Protein Sci 20(1): 118-130.
Yamayoshi, S., M. Watanabe, H. Goto and Y. Kawaoka (2016). "Identification of a Novel Viral Protein Expressed from the PB2 Segment of Influenza A Virus." J Virol 90(1): 444-456.
Yang, X., S. H. Tan, Y. J. Teh and Y. A. Yuan (2011). "Structural implications into dsRNA binding and RNA silencing suppression by NS3 protein of Rice Hoja Blanca Tenuivirus." RNA 17(5): 903-911.
Ye, Q., R. M. Krug and Y. J. Tao (2006). "The mechanism by which influenza A virus nucleoprotein forms oligomers and binds RNA." Nature 444(7122): 1078-1082.
Yin, C., J. A. Khan, G. V. Swapna, A. Ertekin, R. M. Krug, L. Tong and G. T. Montelione (2007). "Conserved surface features form the double-stranded RNA binding site of non-structural protein 1 (NS1) from influenza A and B viruses." J Biol Chem 282(28): 20584-20592.
Ylösmäki, L., C. Schmotz, E. Ylösmäki and K. Saksela (2015). "Reorganization of the host cell Crk(L)-PI3 kinase signaling complex by the influenza A virus NS1 protein." Virology 484: 146-152.

You, L., E. J. Cho, J. Leavitt, L. C. Ma, G. T. Montelione, E. V. Anslyn, R. M. Krug, A. Ellington and J. D. Robertus (2011). "Synthesis and evaluation of quinoxaline derivatives as potential influenza NS1A protein inhibitors." Bioorg Med Chem Lett 21(10): 3007-3011.
Yuan, P., M. Bartlam, Z. Lou, S. Chen, J. Zhou, X. He, Z. Lv, R. Ge, X. Li, T. Deng, E. Fodor, Z. Rao and Y. Liu (2009). “Crystal structure of an avian influenza polymerase PA(N) reveals an endonuclease active site.” Nature 458(7240): 909-13.

Z

Zhang, K., G. Shang, A. Padavannil, J. Wang, R. Sakthivel, X. Chen, M. Kim, M.G. Thompson, A. Garcia-Sastre, K.W. Lynch, Z.J. Chen, Y.M. Chook and B.M.A. Fontoura (2018). "Structural-functional interactions of NS1-BP protein with the splicing and mRNA export machineries for viral and host gene expression.” Proc Natl Acad Sci 115(52):E12218-E12227.
Zhang, K., Y. Xie, R. Munoz-Moreno, J. Wang, L. Zhang, M. Esparza, A. Garcia-Sastre, B. M. A. Fontoura and Y. Ren (2019). "Structural basis for influenza virus NS1 protein block of mRNA nuclear export." Nat Microbiol.
Zhang, L., P. Das, M. Schmolke, B. Manicassamy, Y. Wang, X. Deng, L. Cai, B. P. Tu, C. V. Forst, M. G. Roth, D. E. Levy, A. Garcia-Sastre, J. de Brabander, M. A. Phillips and B. M. Fontoura (2012). "Inhibition of pyrimidine synthesis reverses viral virulence factor-mediated block of mRNA nuclear export." J Cell Biol 196(3): 315-326.
Zhang, L., J. Wang, R. Munoz-Moreno, M. Kim, R. Sakthivel, W. Mo, D. Shao, A. Anantharaman, A. Garcia-Sastre, N. K. Conrad and B. M. A. Fontoura (2018). "Influenza Virus NS1 Protein-RNA Interactome Reveals Intron Targeting." J Virol 92(24).
Zhang, L. J., L. Xia, S. L. Liu, E. Z. Sun, Q. M. Wu, L. Wen, Z. L. Zhang and D. W. Pang (2018). "A "Driver Switchover" Mechanism of Influenza Virus Transport from Microfilaments to Microtubules." ACS Nano 12(1): 474-484.
Zhao, C., M. Jia, H. Song, Z. Yu, W. Wang, Q. Li, L. Zhang, W. Zhao and X. Cao (2017). "The E3 Ubiquitin Ligase TRIM40 Attenuates Antiviral Immune Responses by Targeting MDA5 and RIG-I." Cell Rep 21(6): 1613-1623.
Zhao, C., Z. Lou, Y. Guo, M. Ma, Y. Chen, S. Liang, L. Zhang, S. Chen, X. Li, Y. Liu, M. Bartlam and Z. Rao (2009). "Nucleoside monophosphate complex structures of the endonuclease domain from the influenza virus polymerase PA subunit reveal the substrate binding site inside the catalytic center." J Virol 83(18): 9024-9030.
Zhao, N., V. Sebastiano, N. Moshkina, N. Mena, J. Hultquist, D. Jimenez-Morales, Y. Ma, A. Rialdi, R. Albrecht, R. Fenouil, M. T. Sanchez-Aparicio, J. Ayllon, S. Ravisankar, B. Haddad, J. S. Y. Ho, D. Low, J. Jin, V. Yurchenko, R. K. Prinjha, A. Tarakhovsky, M. Squatrito, D. Pinto, K. Allette, M. Byun, M. L. Smith, R. Sebra, E. Guccione, T. Tumpey, N. Krogan, B. Greenbaum, H. van Bakel, A. Garcia-Sastre and I. Marazzi (2018). "Influenza virus infection causes global RNAPII termination defects." Nat Struct Mol Biol 25(9): 885-893.
Zhao, X., B. Tefsen, Y. Li, J. Qi, G. Lu, Y. Shi, J. Yan, H. Xiao and G. F. Gao (2016). "The NS1 gene from bat-derived influenza-like virus H17N10 can be rescued in influenza A PR8 backbone." J Gen Virol 97(8): 1797-1806.
Zhirnov, P.O (2020). “Unique Bipolar Gene Architecture in the RNA Genome of Influenza A Virus.” Biochemistry (Mosc.) 85(3):387-392.
Zhirnov, O. P. and E. I. Isaeva (2019). "NSP Protein Encoded in Negative NS RNA Strand of Influenza A Virus Induces Cellular Immune Response in Infected Animals." Dokl Biochem Biophys 486(1): 201-205.
Zhirnov, P.O., K.A. Akulich, A.V. Lipatova and E.V. Usachev (2017). “Negative-sense virion RNA of segment 8 (NS) of influenza a virus is able to translate in vitro a new viral protein.” Dokl Biochem Biophys 473(1):122-127.
Zhou, H., M. Jin, Z. Yu, X. Xu, Y. Peng, H. Wu, J. Liu, H. Liu, S. Cao and H. Chen (2007). "Effective small interfering RNAs targeting matrix and nucleocapsid protein gene inhibit influenza A virus replication in cells and mice." Antiviral Res 76(2): 186-193.
Zhu, L. and J. Qin (2019). “A Viral Protein Mimics Histone to Hijack Host MORC3.” Structure 27(6):883-885.
Zhu, X., R. McBride, C. M. Nycholat, W. Yu, J. C. Paulson and I. A. Wilson (2012). “Influenza virus neuraminidases with reduced enzymatic activity that avidly bind sialic Acid receptors.” J Virol. 86(24): 13371-83.
Zinzula, L., F. Esposito, D. Pala and E. Tramontano (2012). "dsRNA binding characterization of full length recombinant wild type and mutants Zaire ebolavirus VP35." Antiviral Res 93(3): 354-363.
Zohari, S., P. Gyarmati, P. Thoren, G. Czifra, C. Brojer, S. Belak and M. Berg (2008). "Genetic characterization of the NS gene indicates co-circulation of two sub-lineages of highly pathogenic avian influenza virus of H5N1 subtype in Northern Europe in 2006." Virus Genes 36(1): 117-125.
Zuker, M. (2003). “Mfold web server for nucleic acid folding and hybridization prediction.” Nucleic Acids Res. 31: 3406-3415.

Chapitre dans un ouvrage imprimé

Shaw ML, Palese P (2013). Orthomyxoviridae: The viruses and their replication. Fields Virology, eds Knipe DM, Howley PM (Wolters Kluwer Health/Lippincott Williams et Wilkins). Philadelphia, PA, pp 1691–1740. 6th Ed.eds Knipe DM, Howley PM.

Wright PF, Neumann G, Kawaoka Y (2013). Orthomyxoviruses. Fields Virology, eds Knipe DM, Howley PM (Wolters Kluwer Health/Lippincott Williams et Wilkins). Philadelphia, PA, pp 1186–1243. 6th Ed.eds Knipe DM, Howley PM.

ANNEXES

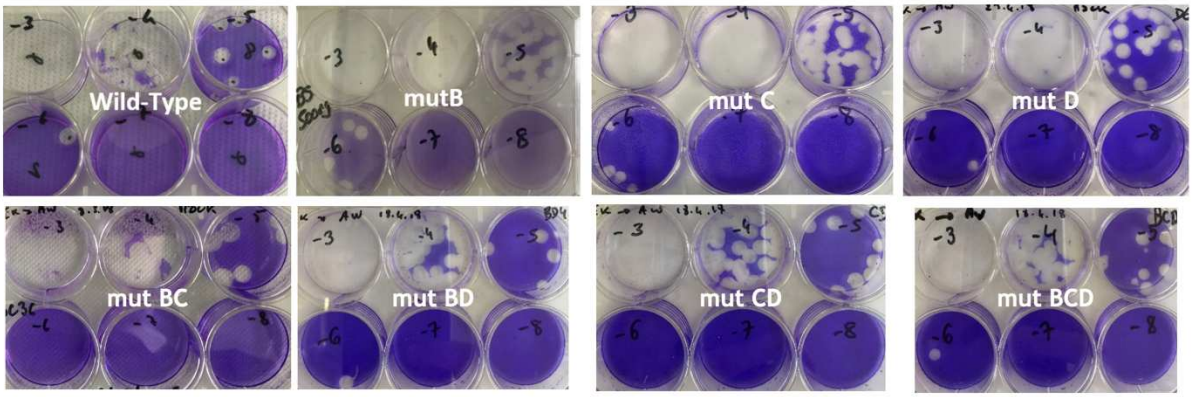
Annexe 1 : Amorces nucléotidiques

Amorce	Sequence sens (5' vers 3')	Sequence anti-sens (3' vers 5')
NS-H7N1	GAGTTGAGTCAGAAGTGTAATAGGATCCGGCTGCTAAC	GTTAGCAGCCGGATCCTATTACACTTCTGACTCAACTC
RBD-pdm2009	CTTGAAAGAGGAATCCTGATAGGATCCGGCTGCTAAC	GTTAGCAGCCGGATCCTATCAGGATTCCTCTTTCAAG
RBD-H3N2	CTGAAAGAAGAATCTTGATAGGATCCGGCTGCTAAC	GTTAGCAGCCGGATCCTATCAAGATTCTTCTTTCAG
RBD-H7N1	CCTGAAGAGTGAAACATAATAGGATCCGGCTGCTAAC	GTTAGCAGCCGGATCCTATTATGTTTCACTCTTCAGG
RBD-H7N9	CTGGAGGAAGAATCCTAATAGGATCCGGCTGCTAAC	GTTAGCAGCCGGATCCTATTAGGATTCTTCCTCCAGAATC
RBD-Bat	CCATTACCGAGGAATCCTAATAGGATCCGGCTGCTAAC	GTTAGCAGCCGGATCCTATTAGGATTCCTCGGTAATGG
WSN-mutB	GAAGAAATAAGATGGCTCATCGAGGAAGTGAGACACAG	CTGTGTCTCACTTCCTCGATGAGCCATCTTATTTCTTC
WSN-mutC	GATAACAGAGAATAGCTTCGAACAGATAACATTTATGCAAG	CTTGCATAAATGTTATCTGTTCGAAGCTATTCTCTGTTATC
WSN-mutD	GCAAGCCTTACAACTACTCCTGGAGGTGGAGCAAGAGATAAG	CTTATCTCTTGCTCCACCTCCAGGAGTAGTTGTAAGGCTTGC
FF-UTR	CAAGAAGGGCGGCAAGATCGCCGTGTGAAGAAATAAGATGGTTGATTGAAGAAG	CTTCTTCAATCAACCATCTTATTTCTTCACACGGCGATCTTGCCGCCCTTCTTG
NS3-RenLuc/Bsa	CACACAGGTCTCAGGCCAGCAAAAGCAGGGTGACAAAGACATAATGGCTTCCAAGGTGTACGACCCCGAGC	
Ren-UTR	GAGCGCGTGCTGAAGAACGAGCAGTGAAGAAATAAGATGGTTGATTGAAGAAG	CTTCTTCAATCAACCATCTTATTTCTTCACTGCTCGTTCTTCAGCACGCGCTC
Ren-NS	AGGATCCTCGAGCAGCAAAAGCAGGGTGACAAAGACATAATGG	AAGCGGCCGCATTTTTTATTATTAAATAAGCTGAAACGAGAAAGTTCTTATCTCTTGCTCC
NS-SpeI/BamHI	TCAACTAGTGGGGGCAGCAGCCATCATCATC	GCTTTGTTAGCAGCCGGATC
PolNS-Bbs	GGAACCAGGTCTTCGTATAGATATCGAAGACCTGGAACCAACTCTTTTGCGGACATGCCAAAG	GGTTCCAGGTCTTCGATATCTATACGAAGACCTGGTTCCGGGACAATTGTTGGCGAAATTTCAC
NS-RenFus	CACACAGGTCTCATTCCGCTTCCAAGGTGTACGACCCCGAG	CACACAGGTCTCAAACCCTGCTCGTTCTTCAGCACGCGCTC
M13F-NS		ATGTAAAACGACGGCCAGTGAATAGTAGAAACAAGGGTGTTTT
AGCBridge	CTATAGGCTAGCCTCGAGCATAATGGCTTCCAAGGTG	CACCTTGGAAGCCATTATGCTCGAGGCTAGCCTATAG
Bbs-RBD	TGATATCGAAGACCTGGAAGCCTCATCAGATTCTTCCTTCAG	AGATATCGAAGACTCGGTTTGTCTGTACCTGCATCGCGCTACC
BM-NS	ATATCGTCTCGTATTAGTAGAAACAAGGGTGTTTT	ATCAGGAAACAGCTATGACTGCAGCAAAAGCAGGGTG
T7	ATCGATAATACGACTCACTATAGG	TAATACGACTCACTATAGGG
M13	GTTTTCCCAGTCACGAC	CAGGAAACAGCTATGAC
Uni-12	AGCAAAAGCAGG	

Annexe 2 : Tampons utilisés lors de la purification de protéines recombinantes (RBD, GST-RBD et GST-NS1)

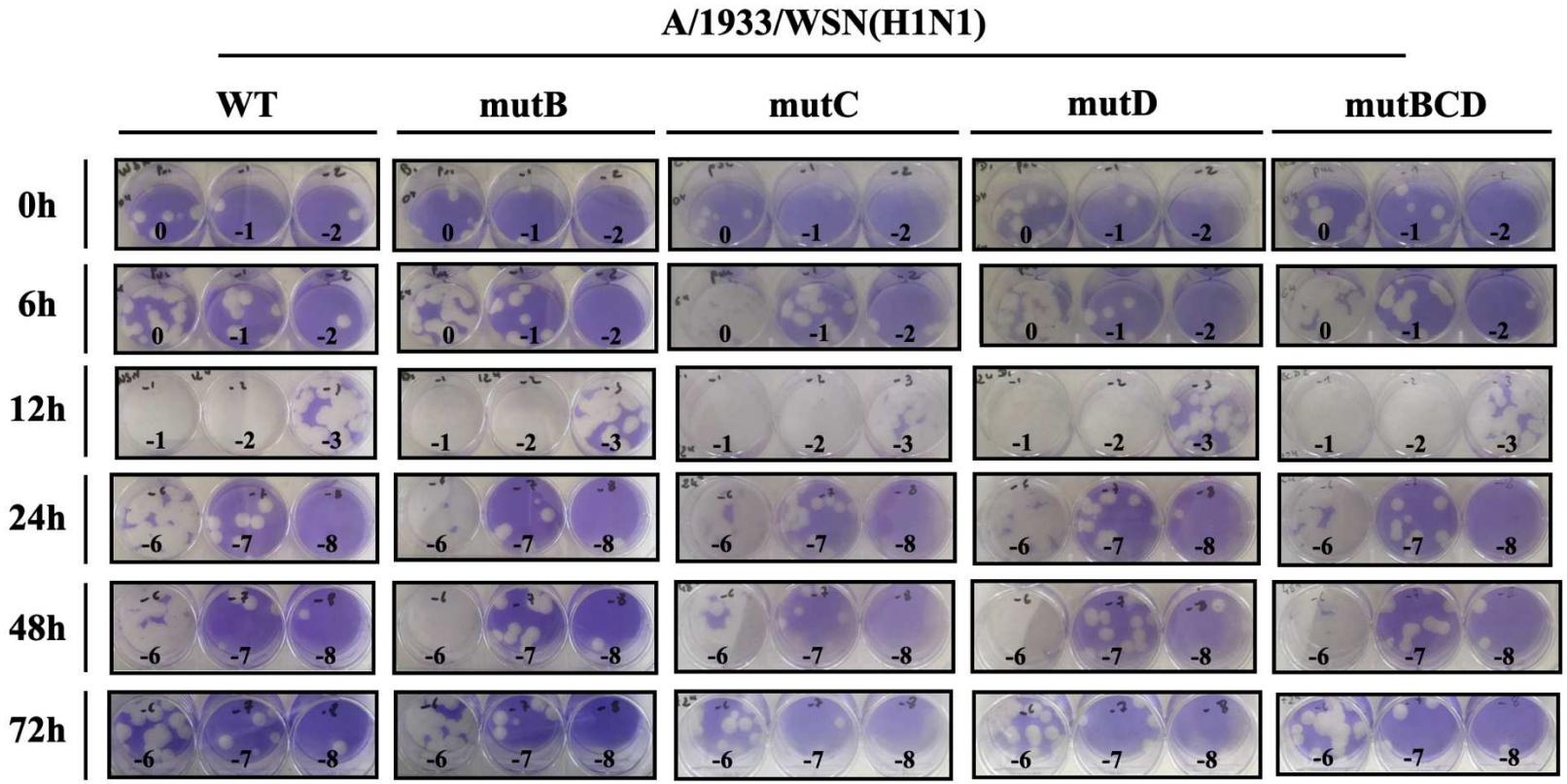
Tampon	Composition
A	20 mmol.L^{-1} HEPES pH 7,6; 500 mmol.L^{-1} NaC l; 5 mmol.L^{-1} Imidazole
B	20 mmol.L^{-1} HEPES pH 7,6; 500 mmol.L^{-1} NaCl; 20 mmol.L^{-1} Imidazole
C	20 mmol.L^{-1} HEPES pH 7,6; 2 mol.L^{-1} NaCl
D	20 mmol.L^{-1} HEPES pH 7,6; 500 mmol.L^{-1} NaCl; 200 mmol.L^{-1} Imidazole
E	20 mmol.L^{-1} HEPES pH 7,6
F	20 mmol.L^{-1} HEPES pH 7,6; 1 mol.L^{-1} NaCl
G	20 mmol.L^{-1} HEPES pH 7,6; 500 mmol.L^{-1} NaCl
H	20 mmol.L^{-1} HEPES pH 7,6; 250 mmol.L^{-1} NaCl
I	1x PBS; 1 % Triton X100; 1 mmol.L^{-1} PMSF
J	1x PBS
K	50 mmol.L^{-1} Tris pH 8; 10 mmol.L^{-1} Glutathion réduit
L	25 mmol.L^{-1} Tris HCl pH 7,6; 150 mmol.L^{-1} NaCl; 2 mmol.L^{-1} $CaCl_2$
M	25 mmol.L^{-1} Tris HCl pH 7,6; 500 mmol.L^{-1} NaCl; 1 mmol.L^{-1} TCEP
N	25 mmol.L^{-1} Tris HCl pH 7,6; 5 % Glycérol; 1 mol.L^{-1} NaCl; 0,1 mmol.L^{-1} TCEP

Annexe 3 : Comparaison des plages de lyse obtenues pour les virus WSN wt et mutants générés par génétique inverse



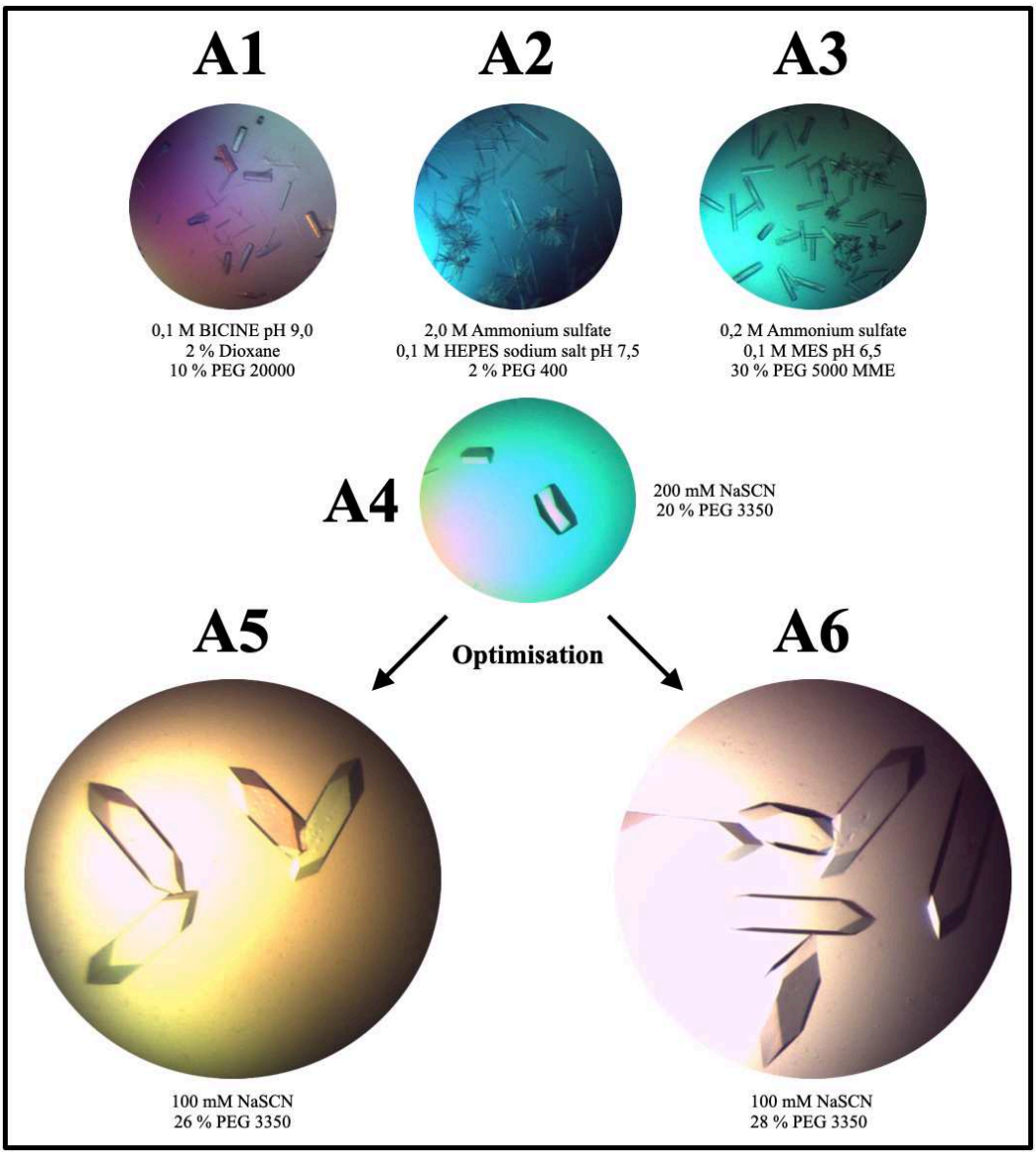
Comparaison des plages de lyse observées entre les différents virus (wt, mutB, mutC, mutD, mutBC, mutBD, mutCD et mutBCD) au temps 72h post-amplification des surnageants de culture récoltés suite à la génétique inverse. Les plages de lyse obtenues pour dilutions des stocks viraux de 10^{-3} à 10^{-8} sont représentées.

Annexe 4 : Comparaison des plages de lyse obtenues pour les virus WSN wt et mutants pour étudier la courbe de croissance virale



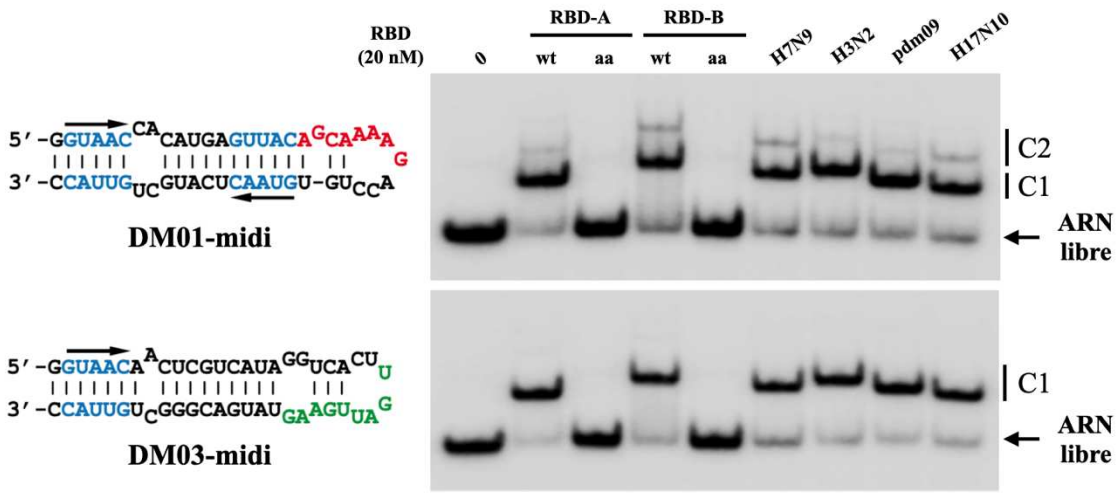
Comparaison des plages de lyse observées entre les différents virus (wt, mutB, mutC, mutD et mutBCD) à différents temps post-infection. L'ensemble des plages de lyse obtenues est représenté.

Annexe 5 : Cristallisation du RBD de NS1 pour le virus de chauve-souris H17N10



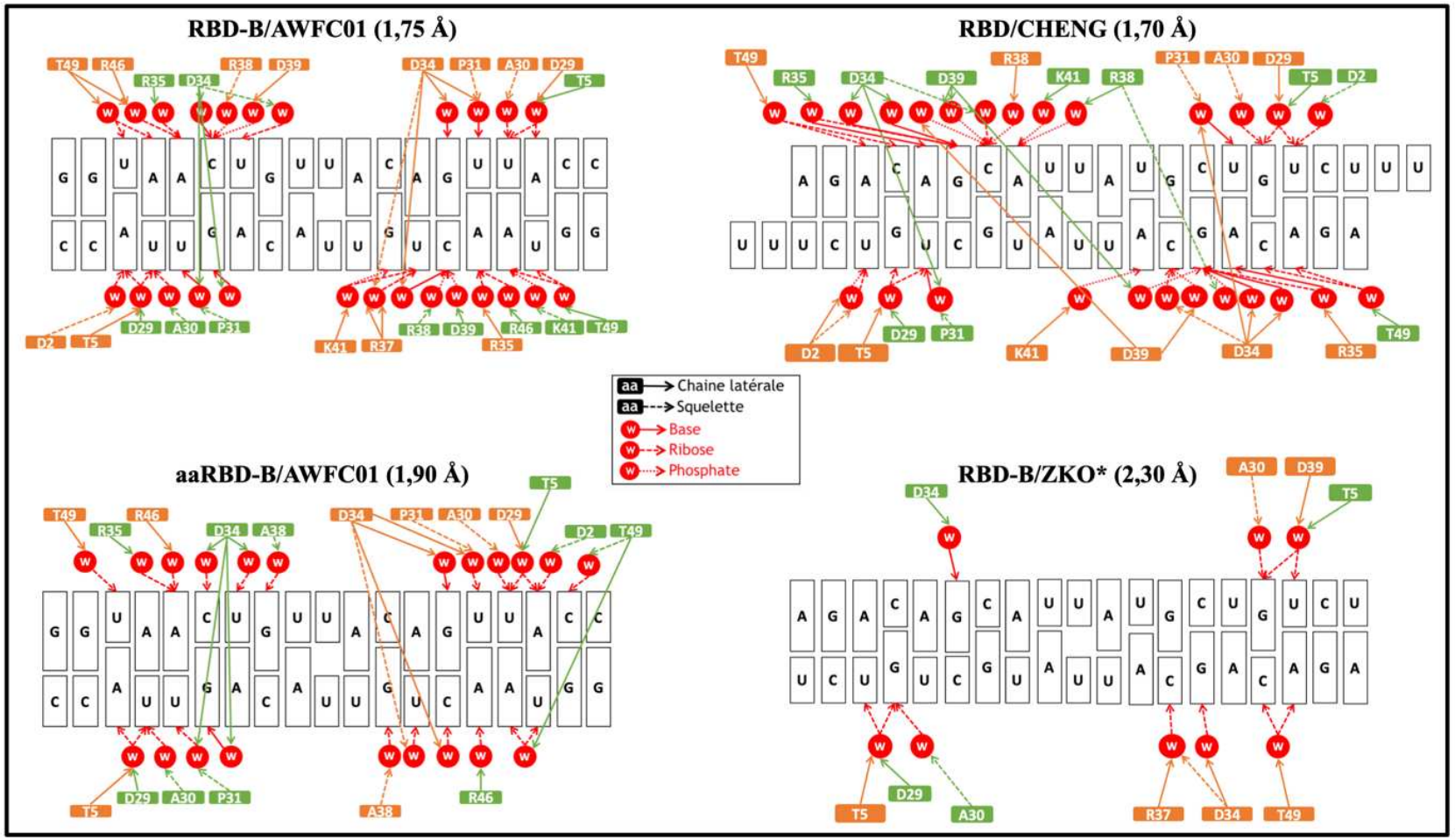
Les conditions de cristallisation du RBD BatH17 sont indiquées en dessous de chaque cliché des cristaux obtenus avec les kits JCSG (A1 et A2) et Classic (A3 et A4). L'optimisation (voir Matériel et Méthodes) de la condition A4 a permis d'obtenir des cristaux diffractants suffisamment les rayons X pour résoudre la structure de ce RBD.

Annexe 6 : Comparaison des complexes formés par les différents RBD purifiés avec les sondes ARN DM01-midi et DM03-midi

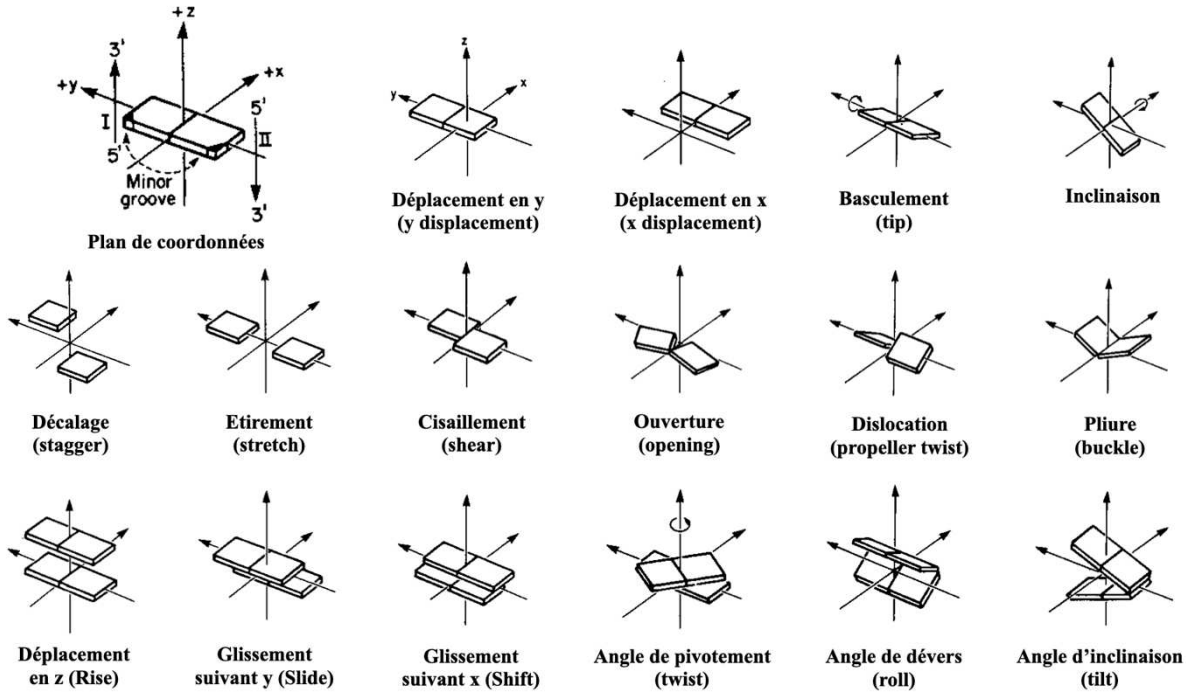


Tous les RBD purifiés au laboratoire sont capable de formés des complexes C1 et C2 avec la sonde ARN DM01-midi alors qu'un unique complexe C1 est observé en gel retard avec la sonde ARN DM03-midi. Dans nos conditions d'interaction, ni le aaRBD-A ni le aaRBD-B ne sont pas capables de former de complexes stables avec les deux sondes ARN.

Annexe 7 : Interactions indirectes impliquées dans les complexes RBD/CHENG, RBD-B/AWFC01, aaRBD-B/AWFC01 et RBD-B/ZKO*

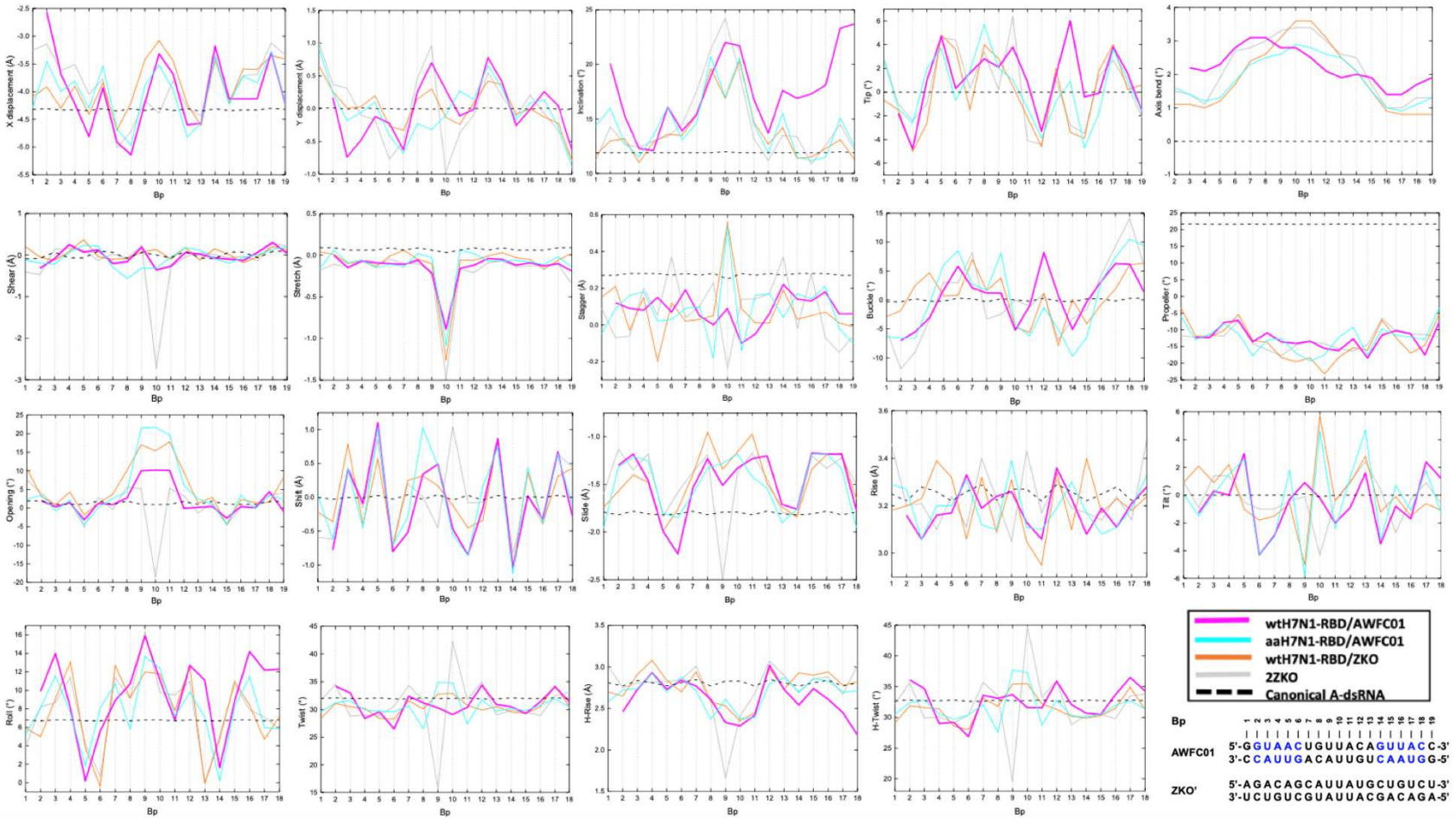


Annexe 8 : Mouvements des nucléobases de l'ARN



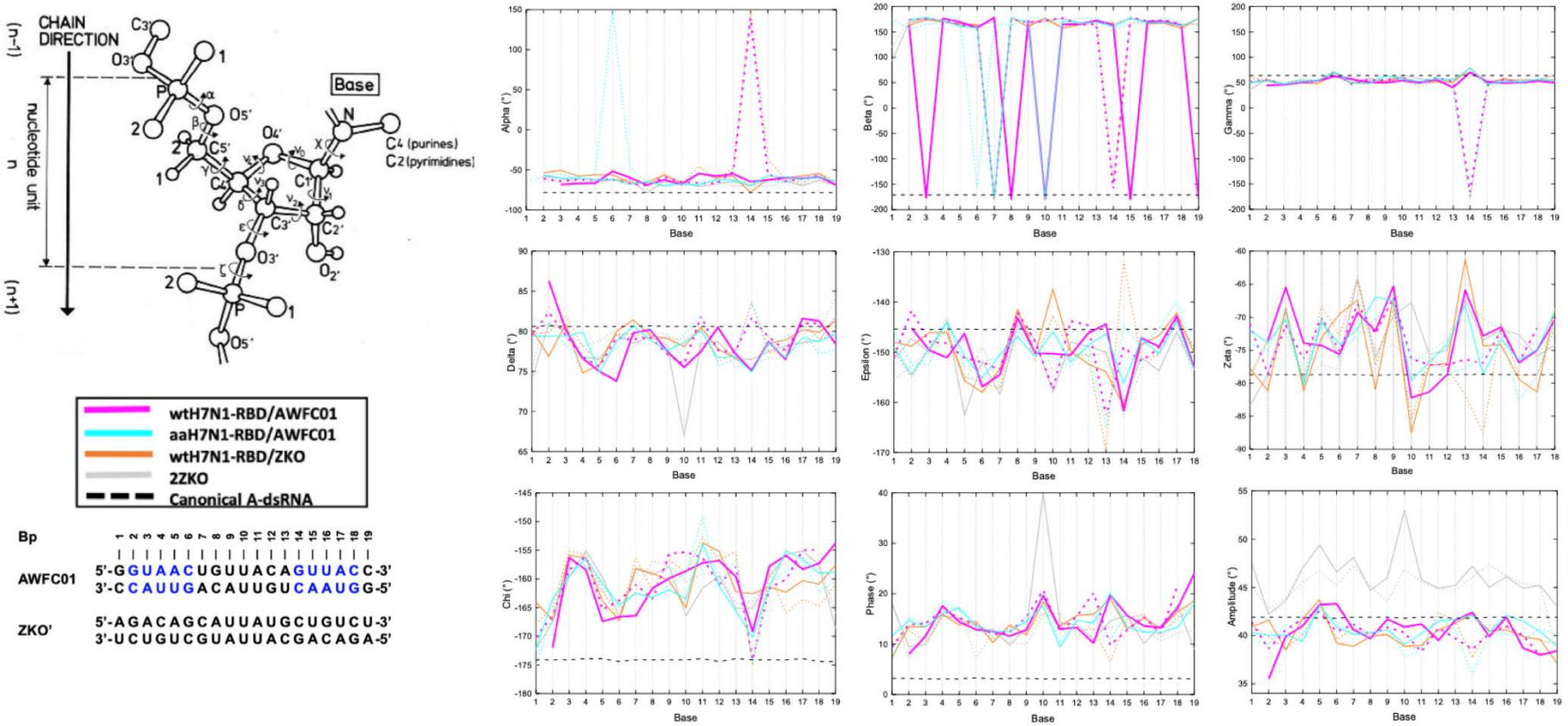
Les mouvements des nucléobases de l'ARN, qui permettent de renseigner sur la conformation générale de ce dernier, sont définis selon la convention internationale établie par Dickerson (1989).

Annexe 9 : Les paramètres hélicoïdaux de l'ARN

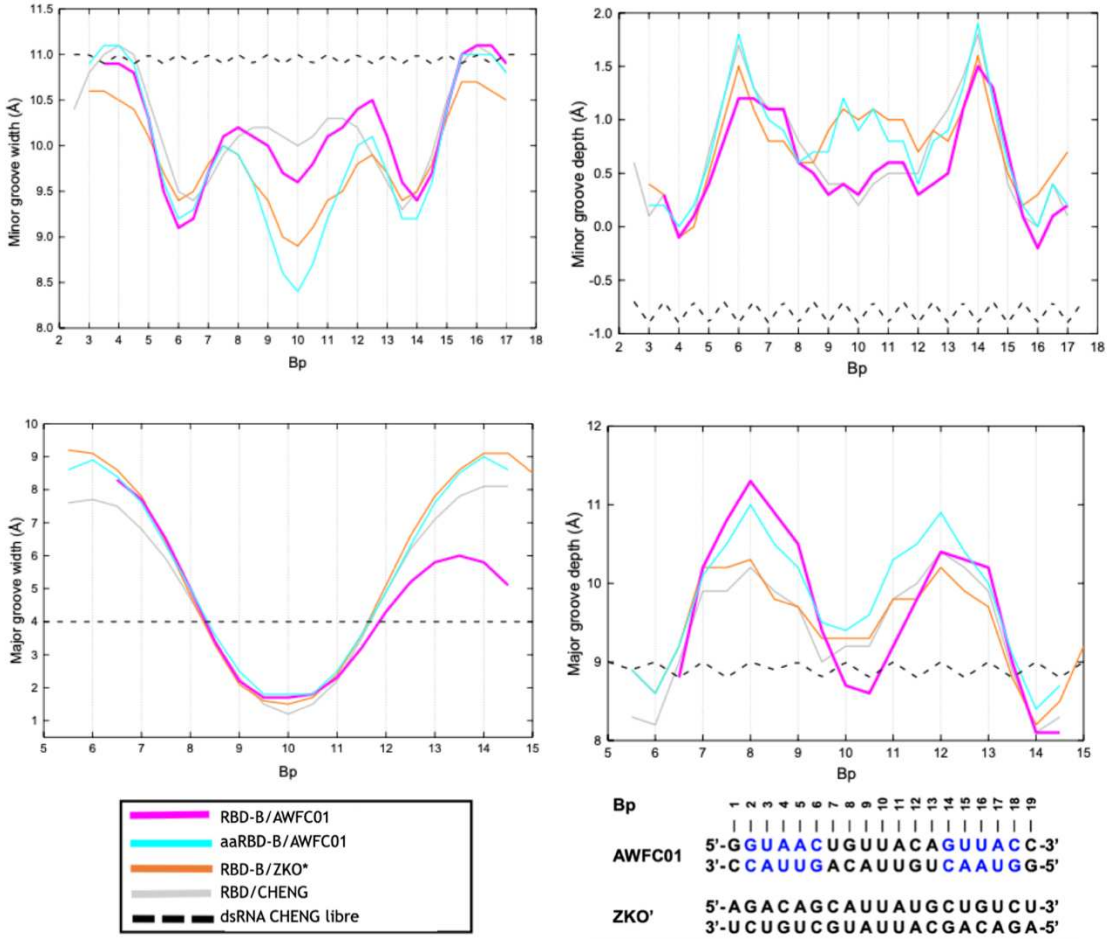


L'analyse des différents paramètres hélicoïdaux de l'ARN CHENG libre (noire) et des complexes RBD-B/AWFC01 (rose), aaRBD-B/AWFC01 (bleu), RBD-B/ZKO (orange) et RBD/CHENG (2ZKO, gris), évaluée sur l'axe des paires de bases, sur les paramètres intra- et inter-bases et sur la courbure de l'ARN, a été réalisée avec le logiciel Curves+. Les différents mouvements des nucléobases de l'ARN sont regroupées en annexe 5.*

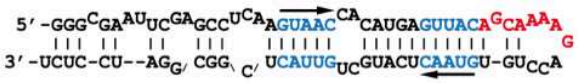
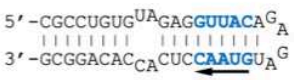
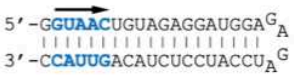
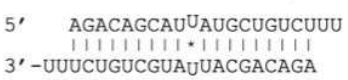
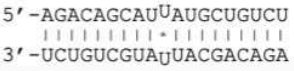
Annexe 10 : Les torsions du squelette ribose-phosphate de l'ARN



Annexe 11 : Largeur et profondeur des sillons de l'ARN



Annexe 12 : Liste des sondes oligonucléotidiques d'ARN utilisées

Structure 2D	Nom	Interaction avec le RBD
	DM01	++++
	DM01-midi	++++
	DM01-short	+/-
	DM01-motA	-
	DM03	++++
	DM03-midi	++++
	DM03-short	+/-
	DM01-motB	-
	shDM02-GAGA	++++
	shDM03-GAGA	++++
	shDM05-GAGA	+
	shDM06-GAGA	++++
	AWFC01	++++
	ZKO	+
	ZKO*	+

PUBLICATION 1 : Structure and sequence determinants governing the interactions of RNAs with *influenza* A virus Non-structural protein NS1

Annexe 13 : Structure and sequence determinants governing the interactions of RNAs with influenza A virus Non-structural protein NS1

viruses

MDPI

Article

Structure and Sequence Determinants Governing the Interactions of RNAs with Influenza A Virus Non-Structural Protein NS1

Alan Wacquiez [1,2,3], Franck Coste [3], Emmanuel Kut [1,2], Virginie Gaudon [3], Sascha Trapp [1,2], Bertrand Castaing [3,*] and Daniel Marc [1,2,*]

1 Equipe 3IMo, UMR1282 Infectiologie et Santé Publique, INRAE, F-37380 Nouzilly, France; alanwacquiez@gmail.com (A.W.); emmanuel.kut@inrae.fr (E.K.); sascha.trapp@inrae.fr (S.T.)
2 UMR1282 Infectiologie et Santé Publique, Université de Tours, F-37000 Tours, France
3 Centre de Biophysique Moléculaire, UPR4301 CNRS, rue Charles Sadron, CEDEX 02, 45071 Orléans, France; franck.coste@cnrs-orleans.fr (F.C.); virginie.gaudon@cnrs-orleans.fr (V.G.)
* Correspondence: bertrand.castaing@cnrs-orleans.fr (B.C.); daniel.marc@inrae.fr (D.M.)

Received: 13 July 2020; Accepted: 25 August 2020; Published: 27 August 2020

check for updates

Abstract: The non-structural protein NS1 of influenza A viruses is an RNA-binding protein of which its activities in the infected cell contribute to the success of the viral cycle, notably through interferon antagonism. We have previously shown that NS1 strongly binds RNA aptamers harbouring virus-specific sequence motifs (Marc et al., Nucleic Acids Res. 41, 434–449). Here, we started out investigating the putative role of one particular virus-specific motif through the phenotypic characterization of mutant viruses that were genetically engineered from the parental strain WSN. Unexpectedly, our data did not evidence biological importance of the putative binding of NS1 to this specific motif (UGAUUGAAG) in the 3′-untranslated region of its own mRNA. Next, we sought to identify specificity determinants in the NS1-RNA interaction through interaction assays in vitro with several RNA ligands and through solving by X-ray diffraction the 3D structure of several complexes associating NS1's RBD with RNAs of various affinities. Our data show that the RBD binds the GUAAC motif within double-stranded RNA helices with an apparent specificity that may rely on the sequence-encoded ability of the RNA to bend its axis. On the other hand, we showed that the RBD binds to the virus-specific AGCAAAAG motif when it is exposed in the apical loop of a high-affinity RNA aptamer, probably through a distinct mode of interaction that still requires structural characterization. Our data are consistent with more than one mode of interaction of NS1's RBD with RNAs, recognizing both structure and sequence determinants.

Keywords: influenza A virus; non-structural NS1; RNA; RNA-protein interaction; 3D structure

1. Introduction

With 3–5 million severe cases and an excess mortality of 290–650 thousand deaths per year [1,2], influenza viruses remain one of the major infectious threats to human health. In addition to this seasonal disease burden, the sporadic human cases of infection with avian influenza viruses (mainly H5N1 and H7N9) are a reminder of the pandemic potential of these emerging viruses if ever they acquired the ability to transmit between humans. Due to the limits of vaccination strategies targeting the mutation-prone influenza viruses, antiviral therapies are still a valid option, notably for the treatment of severe influenza-related respiratory infections, in both seasonal and pandemic influenza.

While the first generation of influenza antivirals, adamantanes, have become obsolete due to the acquired resistance of the currently circulating viruses [3], neither the neuraminidase inhibitors nor the recently approved viral-polymerase inhibitor baloxavir marboxil [4,5] are safe from the development

of resistances [6]. As a consequence, there is a pressing need to enrich our assortment of antivirals, both to improve our preparedness against future pandemics and to design combination therapies that are less prone to the emergence of resistances.

Several viral proteins could be targeted by novel antiviral compounds that could be used alone or in combination. One of the most promising targets is the influenza virus non-structural protein 1 (NS1) [7]. Encoded by the smallest of the eight viral genome segments, NS1 is a multifunctional RNA-binding protein, which is considered the main weapon of the virus to antagonize the innate immune response of the host and to ensure viral replication in a hostile anti-viral environment [8]. This 230-residue protein is highly expressed in the infected cell, where it exhibits several pro-viral activities [9,10]. Its homodimeric RNA-binding domain (RBD, amino-acids 1–73) interacts with several RNAs, including viral mRNAs [11–15], while its Effector Domain (ED, amino-acids 80–202) is a non-obligate dimer that interacts with several cellular proteins. A linker region connecting the two structured domains [16,17] confers plasticity to the quaternary structure, while the C-terminal tail (amino-acids 202–230) has the properties of an intrinsically disordered peptide region [17]. While the effector domain is to some extent dispensable for the biological activities of NS1 [14,18], the RBD plays an essential role in several activities that require the interaction of NS1 with viral and non-viral RNAs. This is best illustrated by the total loss of pathogenicity conferred by the alanine substitution of Arg38 and Lys41, the two main residues involved in binding of RNAs by NS1 [19].

In agreement with its cellular protein partners that have been identified in systematic searches [20–23], NS1 inhibits several pathways involved in the interferon activation and antiviral state [24] and also blocks the processing and nucleo-cytoplasmic export of host mRNAs [21,25]. However, when performing its RNA-related biological activities, NS1 may be selective, either towards some virus-derived RNAs or towards some subset of host cell RNAs, depending on the presence of sequence or structure determinants. Indeed, a systematic crosslink-based search revealed that NS1 preferentially binds introns in a subset of mRNAs, including that of the retinoic-acid induced gene RIG-I [26]. On the other hand, the dsRNA-binding property of NS1 was shown to be required for its inhibition of helicase DHX30 in that it is mediated by a dsRNA that is bound by both proteins [27]. The sequence and structure elements that control the interaction or determine its biological consequences are for the most part unknown. In a previous work [13] we have shown that NS1 binds in vitro with high affinity to RNA aptamers harboring virus-specific RNA motifs. Two conspicuous virus-specific motifs were identified: AGCAAAAG, which is strictly conserved at the 5′end of the virus-derived positive-strand RNAs, and UGAUUGAAG, which is highly conserved in the 3′untranslated region (3′UTR) of NS1's own mRNA. We also showed that NS1 strongly interacts with double-stranded (ds) RNAs containing the GUAAC motif.

In the present work, we first investigated the role of one particular putative NS1-binding site in the 3′UTR of its own mRNA through the phenotypic characterization of mutant viruses that were genetically engineered from the parental strain WSN. In a complementary approach, we explored the hypothesis that NS1 may recognize the virus-specific motifs or the GUAAC motif through specific modes of interaction that might be distinct from what was described for the published structures of NS1-dsRNA complexes [28,29]. To identify such a specific NS1-RNA interaction that could represent a promising target of novel antiviral drugs, we purified a large variety of RNA-binding domains (RBDs) that were representative of NS1's sequence diversity. Through in vitro interaction assays, we analyzed the contribution of structure and sequence determinants in a series of aptamer-derived RNAs. Finally, we solved the crystal structure of several complexes, associating NS1's RBD with dsRNAs of high and low affinity.

2. Materials and Methods

2.1. Rescue of Recombinant Influenza Viruses

Wild-type (WT) and NS mutants of virus A/WSN/1933 (H1N1) were generated by reverse genetics using the 12-plasmid reverse genetics system [30] kindly provided by Ervin Fodor (University of Oxford, Oxford, UK). Mutations in the NS segment-plasmid were introduced by site-directed mutagenesis using the QuikChange II site-directed mutagenesis kit (Agilent, Santa Clara, CA, USA) according to the manufacturer's protocol. The recombinant viruses were rescued by reverse genetics as described previously [31]. The presence of the desired sequence alterations in the amplified viruses was verified by sequencing their NS gene segment using reverse transcription (RT)-PCR.

2.2. Cells, Infections, Virus Titration, and Multicycle Growth Kinetics

HEK293T cells were maintained in Dulbecco's modified Eagle medium supplemented with 10% fetal calf serum, 2 mM L-glutamine, 100 IU/mL penicillin, and 100 mg/mL streptomycin. Madin-Darby canine kidney (MDCK) cells and A549 cells were grown in Eagle's minimal essential medium (EMEM) supplemented with 7.5% fetal calf serum. For virus growth kinetics, subconfluent monolayers of MDCK in 75cm^2-flasks were virus-infected at a multiplicity of infection of 0.001 PFU/cell. Following 1 h of adsorption at 37 °C, the cells were further incubated in serum-free DMEM containing 1 µg of TPCK-treated trypsin (Worthington Biochemicals, Lakewood, NJ, USA). Samples of supernatants that were taken at the indicated times were subsequently titrated by plaque assays on MDCK cells [32].

2.3. Immunodetection of NS1 and NP

Subconfluent monolayers of A549 cells in 24-well plates were virus-infected at a multiplicity of infection of one PFU/cell. At the indicated times, cells were lyzed in 100 µl of Laemmli buffer and an aliquot was used for electrophoretic separation and immunoblot detection using the NS1- and NP-specific polyclonal rabbit antisera (anti-NP PA5-32242, Thermo Fischer Scientific, Waltham, MA, USA) and the ECL-system detection (Advansta, San Jose, CA, USA).

2.4. Minigenome Assay

Eukaryotic expression vector pCIwt-NS1 encoding wt-NS1 was constructed by sub-cloning its coding sequence between the *Xho*I and *Not*I sites of the pCI plasmid (Promega, Madison, WI, USA). In order to prevent production of spliced mRNAs, splice-donor and splice-acceptor sites were both invalidated by point mutations [33]. Substitutions within NS1 (R38A-K41A) were introduced using the QuikChange II site-directed mutagenesis kit. Subconfluent HEK293T cells in 24-well plates (one technical triplicate for each condition) were transfected using the FuGENE (Promega, Madison, WI, USA) reagent with the expression vectors of the four viral polymerase subunits (PB1, PB2, PA, and NP) and the pPolI-NS-Renilla encoding the chimeric minigenome (consisting of the NS genome segment with NS1's Open Reading Frame (ORF) replaced by the Renilla Luciferase ORF), along with the pCI-NS expression vector (or empty vector as a control). The pCMV-Firefly plasmid (Promega) was used as control for transfection efficiency. Twenty-four hours post-transfection, cells were lysed and the activities of the two luciferases were measured using the Dual-luciferase reporter assay system (Promega) and a GloMax-Multi microplate luminometer (Promega). The minireplicon-driven Renilla-luciferase activity was normalized with respect to the activity of the Firefly luciferase, which was used as a transfection control.

2.5. RNA Probes and Proteins

All RNA probes used in binding experiments were purchased from Eurogentec (Belgium). AWFC01 and ZKO* RNA probes used in crystal structure investigations were purchased from *Dharmacon* (Open Biosystem, Horizon Discovery, Waterbeach, UK) (Figure S1).

The nucleotide sequences of RBDs (aa 1-73) were either PCR-amplified with appropriate primers from the cDNAs of the corresponding viruses or purchased as synthetic genes (Eurogentec, Liege City, Belgium), then cloned into the *Nde*I-*BamH*I sites of the expression vector pET15b(+) DNA (Novagen, Merck, Germany). The R38A-K41A double substitution was introduced using the QuikChange kit (Agilent, Santa Clara, CA, USA). In addition, the full-length NS1 sequence of the avian H7N1 virus was purchased as a synthetic gene cloned into the *Spe*I-*BamH*I sites of expression vector pGEX-5X-1 DNA (Genscript, Piscataway, NJ, USA).

Recombinant proteins were overproduced in Rosetta2-(DE3) bacteria in auto-induction medium [34]. Bacteria were grown at 37 °C until reaching an OD600 value of 0.6, then 20 h at 18 °C with agitation (200 rpm). Bacteria were harvested by centrifugation and stored at −20 °C as a dry pellet in a 50 mL polypropylene tube.

Bacterial pellets were gently thawed on ice and resuspended in 45 mL of buffer containing 20 mM Hepes pH 7.6, 500 mM NaCl, and 5 mM Imidazole. The bacteria were lyzed by a 30 min-incubation at 30 °C with 0.75 mg/mL of lysozyme, followed by three freeze/thawing cycles and five 1 min sonication cycles (amplitude 0.4, five pulses per second). The lysate was clarified by a 80 min centrifugation (19,000× *g*, 4 °C). The cleared lysate was applied onto a metal affinity-chromatography column (HisTrap FF, GE Healthcare, Chicago, IL, USA). Untagged proteins were eliminated through successive washes with (i) 10 column-volumes (cv) of wash buffer (20 mM Hepes pH 7.6, 500 mM NaCl) containing 5 mM Imidazole, (ii) 10 cv of wash buffer containing 20 mM Imidazole, and (iii) 10 cv with (20 mM Hepes pH 7.6, 2 M NaCl). His-tagged proteins were eluted by slow application of 10 cv of wash buffer containing 200 mM Imidazole. Fractions containing recombinant proteins (as verified by SDS-PAGE) were loaded onto a POROS HS20-cation exchange chromatography column (Applied Biosystems, Foster City, CA, USA) and eluted by a linear NaCl gradient (20 mM Hepes pH 7.6, 1 M NaCl). Elution fractions containing the protein of interest were concentrated (Amicon®Ultra 10K, Millipore, Burlington, MA, USA) to a 10 mL final volume with successive cold-centrifugations (15–20 min at 5.000 g) before His-tag cleavage at 4 °C during 16 h in 20 mM Tris-HCl pH 8.4, 150 mM NaCl, and 2.5 mM CaCl2 by 1:100 molar ratio of thrombin (Millipore, Burlington, MA, USA). The resulting proteolysate was applied onto a benzamidin column (GE Healthcare) to eliminate thrombin. Recombinant proteins were applied onto a HisTrap FF (GE Healthcare) and eluted with 5 cv of 20 mM Hepes pH 7.6, 500 mM NaCl, and 5 mM Imidazole. Untagged proteins were separated by size-exclusion chromatography onto a Superdex 75 column (GE Healthcare) equilibrated in 20 mM Hepes pH 7.6 and 500 mM NaCl. Fractions of the elution peak of the protein of interest were pooled and concentrated as described previously and stored at −80 °C.

Bacterial pellets containing the fusion proteins GST-RBD and GST-NS1 (H7N1) were resuspended in 45 mL of PBS containing 1% Triton X-100 and 1 mM PMSF and the bacteria were lyzed as described above. The cleared lysate was applied onto a glutathione-affinity chromatography column (GSTrap FF, GE Healthcare). Unbounded proteins were eliminated by washing with 10 cv of PBS. The fusion protein was eluted with 10 cv of 50 mM Tris pH 8 and 10 mM reductive form of Glutathione. Elution fractions containing GST-fusion proteins (as checked by SDS-PAGE) were pooled and concentrated as described previously. The concentrated GSTrap elution fraction was then loaded onto a size-exclusion chromatography column (Superdex S200 or S75 for GST-NS1 and GST-RBD, respectively, GE Healthcare) equilibrated in 25 mM Tris HCl pH 7.6, 5% Glycerol, 1 M NaCl, 0.1 mM TCEP. The elution fractions containing the fusion protein were concentrated as described previously and stored at −80 °C. The homogeneity of all purified proteins was assessed by SDS-PAGE and by MALDI-TOF mass spectrometry.

2.6. Electrophoretic Mobility Shift Assays (EMSA)

5′-[^{32}P]-labelled RNA probe (0.1 nM) was incubated at 4 °C for 30 min with indicated amounts of protein in a 20 µL-reaction mixture (10 mM Tris-HCl, 1 mM EDTA, 0.1% bovine serum albumin, 10% glycerol, 50 mM NaCl, pH 8, 4 µM of yeast tRNA (*Sigma-Aldrich*)). RNA species were

separated by electrophoresis (14V/cm at 4 °C) through a non-denaturing 10% polyacrylamide gel (29:1 acrylamide:bisacrylamide) in TBE buffer. Gels were subsequently dried and exposed for autoradiography. RNA species (free and bound to protein) were quantified using the Typhoon-FLA 9500 imager (GE Healthcare) and *ImageQuant* software, version TL v.8.1. Triplicate EMSA titration experiments were used to extract the dissociation constant K_D [35]. The binding curves were fitted using a non-linear regression logistics function (Hill's equation, $Y = [A2 + (A1-A2)]/[1 + (x/x_0)\ p]$ with software Origins, version 9.0.0 (OriginLab, Northampton, MA, USA).

For competition experiments with unlabeled RNA competitors, radiolabeled AWFC01 (0.1 nM) was first incubated for 30 min with a limiting concentration of the indicated RBD (close to its K_D value). Subsequently, increasing concentrations of competitor were added and the mixture was incubated for an additional 30 min at 4 °C. The reaction mixtures were then separated by EMSA. Triplicate EMSA competition experiments were used to extract the apparent half-maximal effective concentration (EC50app) for the dissociation of the preformed AWFC01/RBD complex. After EMSA quantification, EC50app was determined by fitting the dose-response curves using the same equation as above.

2.7. X-ray 3D Structure Determination

Prior to crystallization, H7N1 NS1-RBD protein (wtH7N1-RBD) was concentrated to 10 mg/mL using Amicon 10 K ultrafiltration devices (Millipore) in last purification buffer. Initial crystallization screening was performed with a Mosquito liquid handler (STP Labtech Ltd., Royston, UK) using Morpheus, JCSG+, Wizard Classic, and Wizard Cryo kits from Molecular Dimensions Ltd. and the AmS04 kit from Qiagen. Diffraction-quality crystals appeared in Morpheus condition 2.5 (100 mM Hepes/MOPS pH 7.5, 8% ethylene glycol, and 30% P550MME-PEG20K) after few days and were flash-frozen directly into liquid nitrogen. Then, 100 K X-ray diffraction data were collected at PROXIMA 2 beamline (SOLEIL, Paris, France) and processed using XDS [36] and AIMLESS [37]. The 3D structure was determined at 1.93Å resolution by molecular replacement with Phaser [38] of the Phenix suite [39] and using PDBid 1AIL as a search model. The atomic model was refined using phenix.refine and manually improved using COOT [40]. Concerning the structures of H7N1-RBD/RNA complexes, double strand RNA (dsRNA) were prepared in 10 mM Tris, 1 mM EDTA pH 8.0, and 500 mM NaCl by mixing the two complementary single strand RNA synthesized and purified at Dharmacon, heating for 2 min at 90 °C and then being cooled down overnight to 4 °C. Freshly purified H7N1-RBD was mixed with dsRNA at a molar ratio of 1:1.1 and incubated for 30 min at 4 °C. Initial crystallization screening was realized as indicated for the H7N1-RBD protein alone. Crystals of the complex of wild type H7N1 NS1-RBD protein and AWFC01 dsRNA (wtH7N1-RBD/AWFC01) were obtained in Wizard Classic condition 2.26 (100 mM CHES pH 9.5 and 30% PEG400) and double mutant R38AK41A H7N1 NS1-RBD protein with AWFC01 dsRNA complex (aaH7N1-RBD/AWFC01 RNA) crystallized in Wizard Cryo condition E3 (100 mM sodium phosphate/citric acid pH 4.2, 200 mM ammonium sulfate and 40% ethylene glycol). After little optimization of AmSO4 kit condition 10, crystals of wild type H7N1 NS1-RBD protein bound to ZKO* dsRNA (wtH7N1-RBD/ZKO*) grew in 2 M ammonium sulfate and 200 mM ammonium nitrate. Then, 100 K X-ray diffraction data were collected at PROXIMA 2 beamline (SOLEIL, France) and processed using XDS [36] AIMLESS [37]. The 3D structures of wtH7N1-RBD/AWFC01, aaH7N1-RBD/AWFC01, and wtH7N1-RBD/ZKO* were determined at 1.75Å, 1.90 Å, and 2.30 Å resolution, respectively, by molecular replacement with Phaser [38] of the Phenix suite [39] and using PDBid 2ZKO as a search model. The atomic model was refined using phenix.refine and manually improved using COOT [40]. The 3D structures described here were deposited at the Protein Data Bank under access numbers 6SW8 (wtH7N1-RBD), 6SX0 (wtH7N1-RBD/AWFC01), 6SX2 (R38A-K41A-H7N1-RBD/AWFC01) and 6ZLC (wtH7N1-RBD/ZKO*). Data collections and refinement statistics are listed in Table S1. For the apo H7N1-RBD structure, residues 1 to 73 of the protein construct were visible in the electron density maps as well as three 1,2-ethanediol and one partial polyethylene glycol molecules. For the wtH7N1-RBD/AWFC01 complex, RBD residues E72 and T73 and RNA basepair G1:C19 could not be positioned by lack of electron density due to mobility, but

six partial polyethylene glycol molecules and one CHES molecule could be located. Concerning the aaH7N1-RBD/AWFC01 complex, only residue T73 was missing and nine 1,2-ethanediol molecules were assigned. Lastly, RBD chain A residues −3 to 73 and RBD chain B residues 1 to 72, the whole double stranded RNA, as well as three sulfate ions and four nitrate ions were visible in electron density maps for the wtH7N1-RBD/ZKO* complex. Molecular graphics images were produced using UCSF Chimera [41]. RNA structure analysis was carried out using Curves+ [42]. RNA-Protein interactions were analyzed using the SNAP program of the 3DNA package [43].

3. Results

3.1. Biological Relevance of the UGAUUGAAG Motif

Our previous SELEX approach [13] identified two virus-specific RNA motifs in the NS1-binding aptamers, corresponding to short sequence motifs that are highly conserved in the virus-derived RNAs of positive polarity. The AGCAAAAG motif (here named motif A, Figure 1), which in the complementary RNA is part of the U12 sequence that makes up the viral polymerase promoter, is strictly conserved at the 5′end of all influenza A viruses complementary RNAs. At the same time, the UGAUUGAAG motif (here named motif B) is highly conserved in the 3′UTR of NS1's mRNA (Figure 1). We first attempted to assess the biological relevance of the second motif, considering that for the first motif its well-established, prominent role in the viral cycle precluded a proper evaluation of its NS1-related role by mutagenesis of the viral genome. The UGAUUGAAG motif (or more generally UGRUUGAAG) is highly conserved in the 3′UTR of NS1's mRNA, 14 nucleotides downstream of NS1's stop codon. It corresponds to nucleotides 236–244 of Nuclear Export Protein (NEP)'s Open Reading Frame (ORF). In addition, a similar UGCUUGAAG motif (or less stringently, URYUUGAAG) is also present at the 3′end of the 3′UTR (nt 317–325 of NEP ORF). Strikingly, with about 2–4 occurrences in each NS-segment positive-strand RNA and a ratio of observed/expected frequency of 13.82, UUGAAG is the second most frequent six-letter word (AAUGGA is the first, at 14.93) in the nucleotide sequences of influenza A virus NS segments (~74,000 nucleotide sequences downloaded from the Influenza Virus Resource, available online: https://www.ncbi.nlm.nih.gov/genomes/FLU/Database/nph-select.cgi?go=database, accessed on 12 august 2020), representing about 63 million six-letter words).

On the basis of (i) the identification of the B motif by the SELEX approach, combined with (ii) the high conservation of its two copies (B and D) in the 3′UTR of NS1's mRNA, we reasoned that NS1 could specifically interact with these motifs in its own mRNA and thereby regulate its translation. Accordingly, mutations in one of these motifs would be expected to alter both the synthesis of NS1 and the viral phenotype. We therefore introduced mutations in the NS-segment of virus A/WSN/33 and rescued the mutant viruses by reverse genetics. We modified either of the two UGRUUGAAG and URYUUGAAG motifs (i.e., B and D, respectively) in the 3′UTR of NS1, along with a third site, named C, that we used as a control. We were cautious as to not change the amino-acid sequence of NEP since motifs B, C, and D encompass the codons of amino-acids that belong to the α-helices C1 and C2 of the C-terminal part of NEP and are critical to its tertiary and quaternary structure as well as to its interactions with matrix protein M1 [44].

All seven possible combinations of mutants were rescued, along with the wild-type virus. They all grew to similar titers on MDCK and their replication potential was similar to that of the wild-type virus, as shown by a multicycle growth assay in MDCK cells (Figure 2a). Their plaque phenotype was also undistinguishable from that of the wt virus (Figure S2). Further, we monitored the accumulation of viral proteins NS1 and NP in A549 cells that were infected with either the wt virus, single mutant B, or triple mutant BCD and noticed no difference between the three viruses regarding the accumulation of the two viral proteins (Figure 2b).

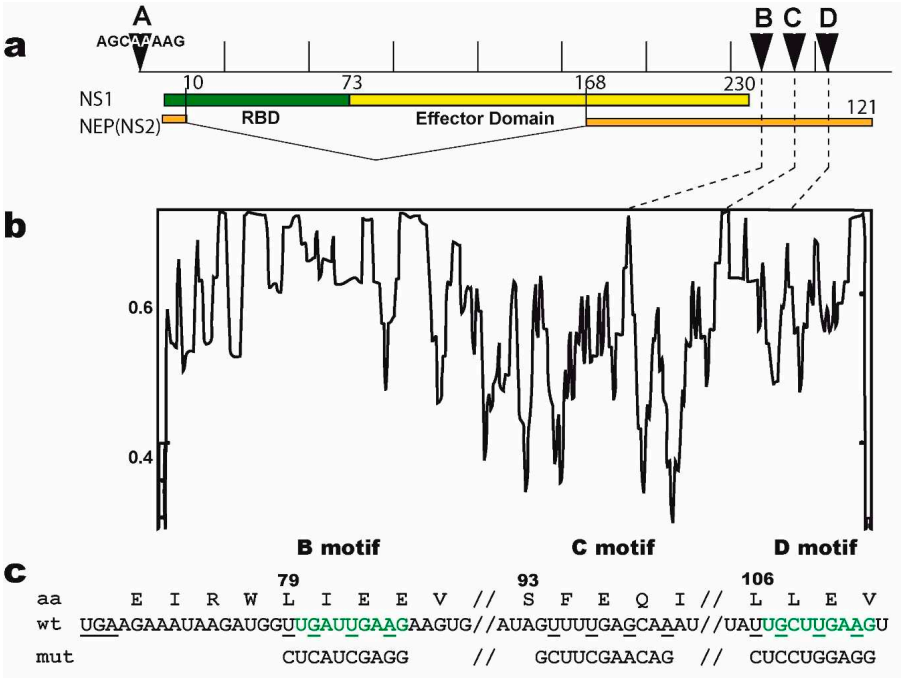


Figure 1. The non-structural (NS) segment and the mutated B, C, and D motifs. (**a**) The NS segment, showing the NS1 and NEP ORFs (NEP in orange; NS1 in green and yellow for the RBD and ED, respectively), with arrows indicating the positions of the A, B, C, and D motifs. Vertical bars = 100 nucleotides. (**b**) nucleotide conservation of NEP's Open Reading Frame (ORF). The open reading frames of NEP from all human H1, H2, and H3 viruses were aligned and the nucleotide conservation (sliding window of 7 nucleotides) was plotted along the length of the ORF. The dotted lines highlight the relatively high conservation of motifs B, C, and D. (**c**) nucleotide sequence of the wt and mutated motifs, with the B and D motifs in green. The underlined UGA is the stop codon of NS1. The three sites are highlighted, with the mutated nucleotides underlined. Below are the three mutated motifs. The amino-acid sequence of NEP (shown above with L79, S93, and L106) remains unchanged. All possible combinations of mutants were rescued (i.e., single mutants B, C, and D; double mutants BC, BD, and CD; and the triple mutant BCD).

As an additional attempt to examine whether NS1 could somehow modulate the fate of the NS-viral genomic segment, we set up a "minireplicon" assay where NS1's Open Reading Frame (ORF) in the NS-segment of WSN was replaced by that of the Renilla Luciferase. In this system based on the transfection of HEK293T cells, the normalized Renilla Luciferase signal reflects the replication, transcription, and translation of the chimeric "NS-Renilla"segment. We compared four such NS-Renilla segments, where the 3'UTR of NS1 corresponded to the wt sequence or to either of the three single-mutant sequences with the B, C, or D mutations. More specifically, we examined to what extent the transient co-expression of (wild-type) NS1 could impact the reporter signal (i.e., the replication, transcription, or translation). As shown in Figure 3, we repeatedly observed that transient co-expression of NS1 increased the minireplicon-signal by four- to six-fold compared to the signal measured with the same chimeric NS-Renilla segment in the absence of NS1. Similar foldchange ratios were observed with the four NS-Renilla segments, i.e., the foldchange was independent of the particular 3'UTR. Taken together, our data indicate firstly that NS1 dramatically increases the

minireplicon efficiency, i.e., at least one of the three steps consisting of replication, transcription, or translation. Secondly, because this effect was observed regardless of the presence of the putative NS1-binding motifs in the 3'UTR of NS1 mRNA, it is unlikely that the putative binding, if any, of NS1 to the 3'UTR of its own mRNA has any relevance in the viral cycle.

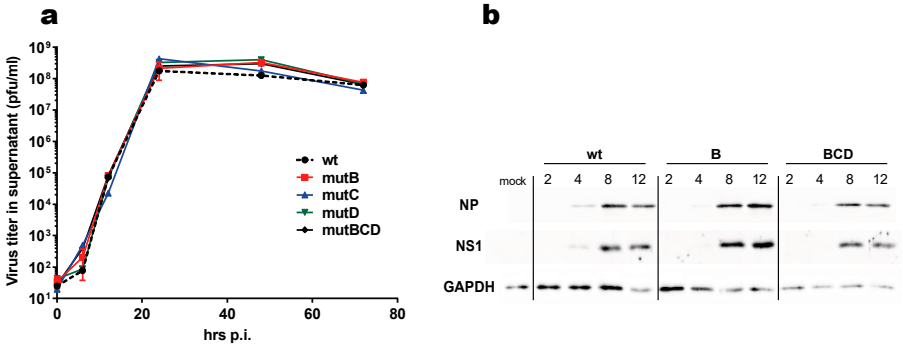


Figure 2. Multicycle growth curve of the three single mutants and of the triple mutant (**a**) MDCK cells that were infected with the indicated viruses (multiplicity of infection = 10^{-3} PFU/cell). Viral titers were measured in the supernatants that were collected at 6 h, 12 h, 24 h, 48 h, and 72 h post-infection. (**b**) Accumulation of NS1 and NP in virus-infected cells. A549 cells in a 24-well plate were infected (one PFU/cell), then lysed at the indicated times post-infection. Proteins in the lysates (2% of the 100 µL-lysate) were separated through SDS-PAGE, then transferred to a nylon membrane. NS1 and NP were revealed through immunoblot, using polyclonal rabbit antisera (anti-NP PA5-32242, *Thermo Fischer Scientific*).

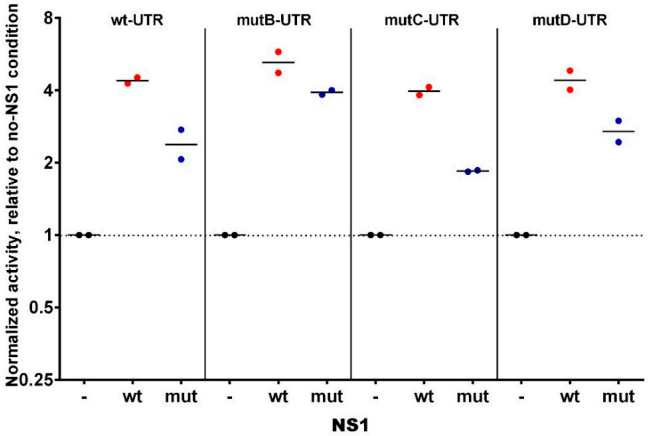


Figure 3. Modulation by NS1 of the activity of wt and mutated minigenomes HEK293T cells were transfected with the expression vectors of the polymerase subunits and the chimeric minigenome-expression plasmid, along with the NS1-expression vector (or empty vector as a control (-), or mutated-NS1-expression vector (mut)). Further, 24 h post-transfection, the activity of the Renilla Luciferase in the transfected cell lysate was measured and normalized relative to that of the Firefly luciferase activity. The normalized renilla activities were then expressed relative to that measured in empty-vector transfected cells (no NS1-condition). Each dot is the geometric mean of a technical triplicate in a given experiment (n = two independent experiments).

3.2. Functional and/or Structural Insights into the NS1-RNA Interactions

3.2.1. Minimal Structure of Selected Aptamers for Efficient Recognition by NS1

SELEX-identified aptamers DM01 and DM03 [13] have been selected for their interaction with a mixture of recombinant NS1 proteins representative of its two alleles A and B (from H5N1 and H7N1 viruses, respectively). These aptamers harbour one or two copies of the double-stranded motif GUAAC, along with a virus-specific sequence motif that is present in the apical loop of the long double-stranded hairpin, i.e., the AGCAAAAG motif in DM01 and the UGAUUGAAG motif in DM03 (corresponding to motifs A and B, respectively, in Figure 1). In order to identify the minimal RNA determinants required for binding, we first sought to identify shorter aptamer derivatives that still kept a high affinity for the RBD. Towards that aim, we designed two medium-length synthetic RNAs, DM01-midi and DM03-midi, that were derived from the 80-nt RNA-aptamers DM01 and DM03, respectively, along with shorter RNAs (DM01-short and DM03-short, respectively), consisting of the terminal hairpin of these aptamers encompassing the virus-specific motif. Their ability to form stable complexes with RBD were assessed through an electrophoretic mobility shift assay (EMSA) with the purified allele B-RBD (H7N1) (Table 1 and Figure S3).

Table 1. Binding of RBD of NS1 from H7N1 to native and truncated RNA aptamers.

Name	nt Sequence / 2D Structure	Length (nt)	RBD Binding	Reference
DM01	5′-GGGCGAAUUCGAGCCUCAAGUAACCACAUGAGUUACAGCAAAAG 3′-UCUC-CU--AGG/CGG\C/UCAUUGUCGUACUCAAUGU-GUCCA	81	++++	[13]
DM01-midi	5′-GGUAACCACAUGAGUUACAGCAAAAG 3′-CCAUUGUCGUACUCAAUGU-GUCCA	50	++++	this work
DM01-short	5′-GGAGUUACAGCAAAAG 3′-CCUCAAUGU-GUCCA	30	+ / - - -	this work
DM01-mot A	5′-AGCAAAAG	8	-	this work
DM03	5′-GGGCGAAUUCGAGCCUCAAGUAACACUCGUCAUAGGUCACUU 3′-UCUC-CU--AGG/CGG\C/UCAUUGUCGGGCAGUAUGAAGUUAG	81	++++	[13]
DM03-midi	5′-GGUAACACUCGUCAUAGGUCACUU 3′-CCAUUGUCGGGCAGUAUGAAGUUAG	50	++++	this work
DM03-short	5′-CCGUCAUAGGUCACUU 3′-GGCAGUAUGAAGUUAG	32	+ / - - -	this work
DM03-mot B	5′-UGAUUGAAG	9	-	this work

While the medium-size dsRNAs DM01-midi and DM03-midi kept their binding properties (Table 1 and see below), the short RNAs DM01-short and DM03-short, which lacked the terminal GUAAC motif close to the free blunt end of their dsRNA structure, interacted poorly with the RBD. RNA-protein complexes with a shifted mobility were observed only at very high concentrations of the RBD (>2 µM), indicating a very low affinity (Figure S2). The two very short single-stranded RNAs consisting of only the 8- or 9-nt virus-specific sequence motifs (motifs A and B in DM01 and DM03, respectively) yielded no observable complex, even at the highest concentration of the RBD. Taken together, these data suggest that the duplex part of the hairpin RNA must extend at least over about 20 base pairs and must contain at least one GUAAC/CAUUG motif close to the blunt end of the dsRNA. Further, at least under the stringent binding conditions that we used, none of the virus-specific motifs that were identified by the SELEX approach showed a measurable, nanomolar-range affinity for the RBD in their isolated form.

3.2.2. Major Contribution of the RBD in the Recognition of DM01-Midi and DM03-Midi

We next assessed the interaction of DM01-midi and DM03-midi with the full-length protein NS1 and we chose to perform these experiments with the allele B NS1 from the avian H7N1 virus. Because the mature full-length protein was observed to precipitate even at moderate concentrations, these assays were performed with the fusion protein GST-NS1 or with GST-RBD as a control. Up to 100 nM of protein GST-RBD was observed to quantitatively titrate the two RNAs in a C1 complex, which likely corresponds to the binding of one GST-RBD molecule per RNA molecule (Figure 4). With DM01-midi at 1000 nM of GST-RBD, C1 was replaced by a second complex C2 (Figure 4a), likely of 2:1 stoichiometry, while under the same conditions, no additional complex was observed with DM03-midi (Figure 4b). This suggests that DM01-midi contains an additional determinant (not present in DM03-midi) that is recognized by a second GST-RBD, dependent on the binding of the first molecule. When GST-RBD was replaced by GST-NS1 in similar binding experiments, we observed that both RNAs formed a C'1 complex. In this case, however, not only for DM01-midi but also for DM03-midi, the C'1 complex was itself partially or completely replaced by a C2' complex in the presence of higher concentrations of GST-NS1. This indicates that the formation of the low-mobility complexes (C2 and/or C'2), likely of 2:1 stoechiometry, required the presence of NS1's effector domain for DM03-midi, while the RBD alone was sufficient in the case of DM01-midi.

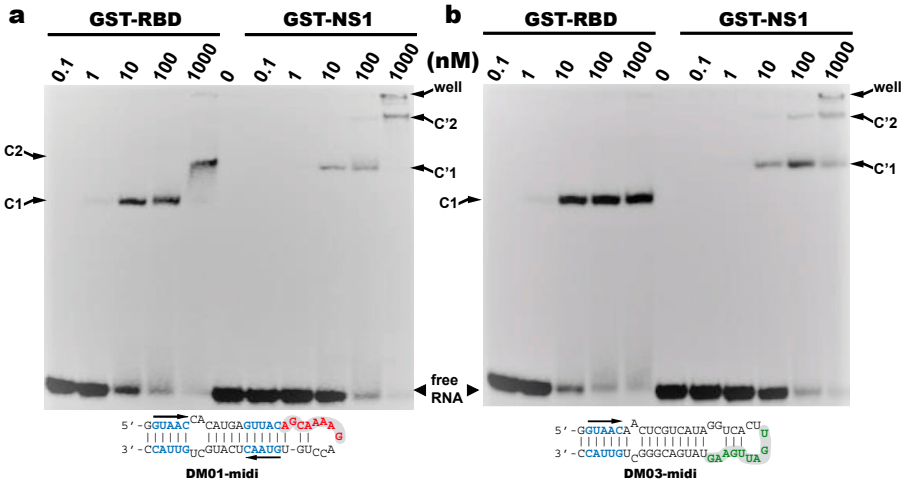


Figure 4. Binding of the RBD and full-length NS1 to DM01-midi and DM03-midi Radiolabeled DM01-midi (**a**) or DM03-midi (**b**) were incubated with increasing concentrations of GST-RBD or GST-NS1 full-length (H7N1). Free-dsRNAs and complexes were separated as described in Materials and Methods. C1 and C2 indicate the RBD:RNA complexes of 1:1 and 2:1 stoechiometry, respectively, while similarly, C'1 and C'2 indicate the NS1:RNA complexes of 1:1 and 2:1 stoechiometry.

GST is known to be able to dimerize [45] and this property could favor the binding of several GST-NS1 molecules to the DM01-midi and DM03-midi RNAs. We therefore repeated our EMSAs using the native, untagged H7N1-RBD. The two RNAs were almost completely titrated by low nanomolar amounts of RBD (Figure 5). As was observed with GST-RBD, untagged RBD formed a unique C1 complex with DM03-midi, while at least two discrete complexes C1 and C2 were formed with DM01-midi. It is likely that the C1 complex observed with the two RNA probes corresponds to a 1:1 complex, associating one RBD dimer and one RNA molecule. The additional C2 complex that was observed only with DM01-midi was formed as soon as the RBD concentration reached 2 nM, even before the free RNA probe was completely titrated in C1. The quantification of EMSA suggests that C2 results

from the facilitated binding of a second RBD to the 1:1 RBD:RNA C1 complex (Figure 5c). Above 20 nM of RBD, the apparent abundance of C2 remained relatively constant (around 30%), while C1 was progressively replaced by a C2′ complex, of which its diffuse and intermediate electrophoretic mobility gradually shifted towards that of the C2 complex. While its reduced electrophoretic mobility is indicative of a RBD:RNA stoichiometry above 1:1, the smeared appearance of C2′, along with the fact its position (its relative shift) is concentration-dependent, suggests that the C2′ and C1 complexes are in dynamic equilibrium during electrophoresis, with their relative amounts being reflected by the position of C2′. Taken together, these data suggest that as soon as the RBD concentration reaches 2 nM, a second RBD molecule binds to the preformed C1 to yield two distinct complexes of 2:1 RBD:RNA stoechiometry. Among these complexes, about 30% consist of the stable C2 complex, while the unstable one C2′ is in dynamic equilibrium with the 1:1 complex C1.

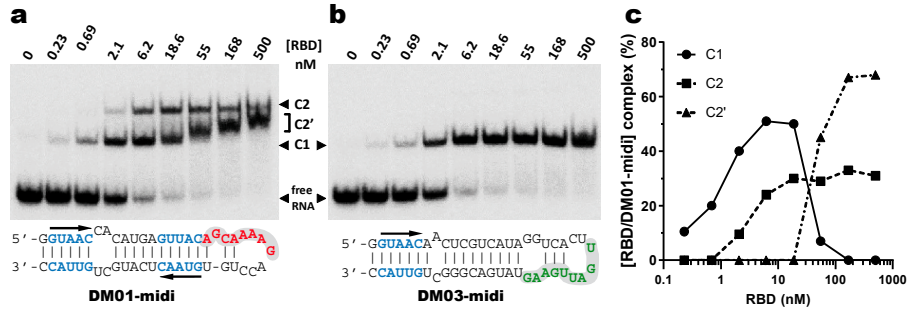


Figure 5. Interaction of DM01-midi and DM03-midi with the native H7N1-RBD Indicated radiolabeled RNA probes were incubated with increasing concentrations of RBD and the resulting reaction mixtures were analyzed by EMSA as described in Materials and Methods. (**a**,**b**) Representative gel autoradiographs obtained with DM01-midi and DM03-midi, respectively. (**c**) The relative intensities of the bands corresponding to the C1, C2, and C2′ complexes formed with DM01-midi were quantified and plotted as a function of the RBD concentration.

3.2.3. Contribution of the AGCAAAAG and GUAAC Motifs to the Interaction

The different behaviors of the RBD towards DM01-midi and DM03-midi suggest that while both RNAs harbor a high affinity binding site that allows the formation of the 1:1 RBD:RNA ratio complexes C1, DM01-midi harbors an additional determinant that uniquely allows the formation of complexes C2 and C2′. The two RNAs share the long double-stranded structure as well as the terminal GUAAC motif, which together could form the high-affinity binding site. In addition, DM01-midi harbors a second GUAAC motif, close to the apical loop that also contains the virus-specific motif A (AGCAAAAG). In order to investigate the role played by each of these features, we designed and assessed the binding of DM01-mod1 and DM01-mod2, two hairpin RNAs derived from DM01-midi (Figure 6). DM01-mod1 conserved the virus-specific A motif while its dsRNA structure no longer contained the two GUAAC motifs. Conversely, DM01-mod2 kept the two GUAAC motifs but its A motif was replaced by an unrelated sequence.

RBD formed C2-like complexes with DM01-mod1, which harbors the AGCAAAAG motif, albeit only at high concentrations, while there was no discrete band corresponding to the C1 complex. Instead of C1, we observed a large smear (Figure 6). This suggests that RBD forms, with DM01-mod1 which harbors the AGCAAAAG motif, unstable C1-like complexes that dissociate during electrophoresis. Very differently, with DM01-mod2, which harbors the two GUAAC motifs, RBD at low concentrations (1–10 nM) formed C1 complexes, while at the 100 and 500 nM concentrations, unstable C2-like complexes of reduced mobility were observed. This last observation suggests that the presence of a loop of which its 2D structure is close to that present in DM01-midi and DM01-mod1 could facilitate

the fixation of a second molecule on the preformed complex C1. However, because it lacks the viral sequence AGCAAAAG, this loop cannot form a stable C2 complex in the presence of moderate concentrations of RBD. Additionally, higher concentrations of the RBD induce the formation of unstable C2′ complexes that steadily dissociate during electrophoresis and yield the fuzzy band.

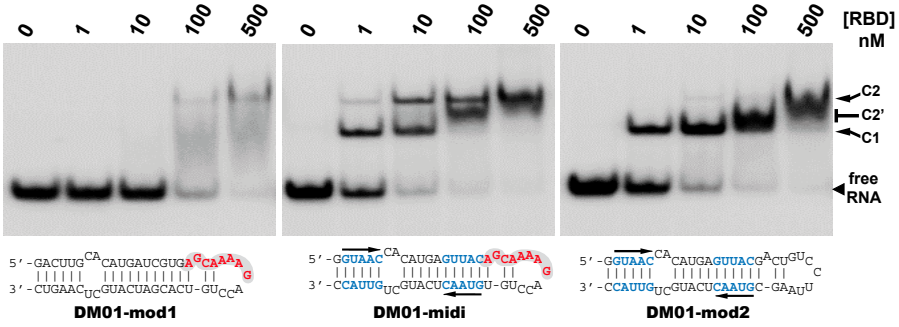


Figure 6. Comparative binding of RBD to DM01-midi, DM01-mod1, and DM01-mod2. The indicated radiolabeled RNA probes were incubated with increasing concentrations of H7N1-RBD and the incubation mixtures were analyzed by EMSA as described in Materials and Methods.

Noteworthily, the ability to form a C2 complex with the DM01-midi probe was shared by all the NS1 RBD proteins that we tested in this study, whilst only the C1 complex was observed with DM03-midi (Figure S4). The distinct electrophoretic mobility of the complexes are consistent with the divergent isoelectric points (pI) of the RBDs: for instance, the complexes made with the H5N1-RBD (pI 6.11) are more acidic and therefore less shifted than those formed with the H7N1-RBD (pI = 8.00). As expected, the double substitution R38A/K41A, either in the H7N1 or H5N1-origin RBD, abolished their binding to both RNAs.

Consistent with all the observations, we can propose the following scenario for the binding of RBD to DM01-midi. At subnanomolar RBD concentrations, RNA probes containing at least one GUAAC/GUUAC motif in their double stranded part are stably recognized by a first RBD molecule to form the C1 complex. In the 1–10 nM concentration range, a second RBD molecule binds to the preformed C1 complex in a cooperative-like manner that likely involves the recognition of the viral sequence (2:1 RBD:RNA ratio complex C2). Thus, the formation of a stable C2 complex intrinsically depends on a stable C1 complex and is favored by the presence of the viral sequence AGCAAAAG in the apical loop. The cooperative-like interaction with the AGCAAAAG motif is a property shared by all NS1 RBDs.

3.2.4. AWFC01, a High-Affinity dsRNA Containing Two GUAAC Motifs

Our observations are consistent with the strong and preferential binding of the RBD to double-stranded RNAs harboring at least one copy of the double-stranded motif GUAAC. In order to establish the importance of this motif and of its location, we designed a series of short hairpin RNAs harboring either zero, one, or two copies of the GUAAC duplex. In order to avoid the binding of a second RBD molecule to the 1:1 complex of RBD with RNA, all RNA probes were designed with a GAGA apical loop that is known to stabilize the double stranded structure of RNA hairpins [46].

Nanomolar concentrations of RBD formed a unique 1:1 RBD:RNA ratio complex (C1-like) with shDM02-GAGA and shDM03-GAGA, both harboring two GUAAC motifs (Figure 7). Of the two short RNAs that contained only one GUAAC motif, a strong interaction was observed only with shDM06-GAGA, which harbored this motif at the blunt end of the hairpin RNA. In contrast, shDM05-GAGA, which harbors the unique GUAAC motif in the immediate vicinity of the GAGA loop, was not recognized by RBD in the nanomolar concentration range. Altogether, these data show

that the RBD binds the GUAAC motif provided that the latter is readily accessible, being positioned at the free blunt end of the dsRNA, while contrariwise, its proximity with the apical loop probably hinders its accessibility or unfavorably alters its shape. Based on this rationale, we designed a model double-stranded RNA made up by the annealing of two synthetic RNA molecules of the same sequence. This synthetic palindromic dsRNA, named AWFC01 (Figure 7), harbors one GUAAC motif at each of its blunt ends and is perfectly symmetric thanks to a central uracil nucleotide that stabilizes the RNA duplex by forming a non-canonical U*U Hoogsteen base pair [47].

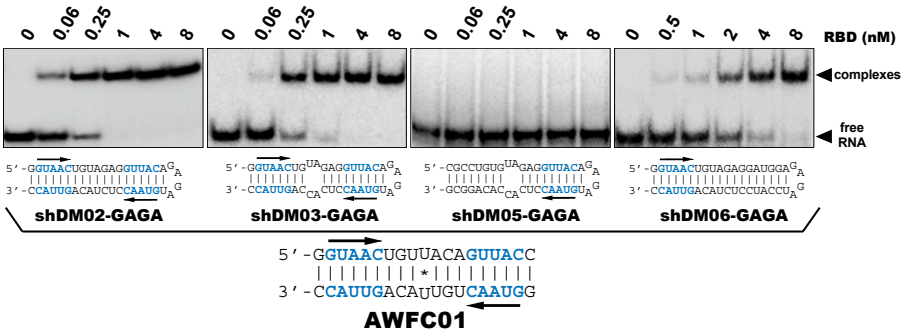


Figure 7. Importance of the number and position of the GUAAC motifs. The indicated radiolabeled RNAs were incubated with increasing concentrations of the H7N1-RBD, then analyzed by EMSA as described in Materials and Methods. Bottom: secondary structure of AWFC01, with its central U*U Hoogsteen base pair. The double-stranded GUAAC motif (blue) is emphasized and oriented by the horizontal arrow.

As expected, all wild type RBDs tested in this study recognized AWFC01 and formed a stable C1-like complex in a nanomolar range of protein concentrations (Figure 8a,b for H5N1- and H7N1-RBD, respectively). The dissociation constants extracted from EMSA titration experiments (Figure S5 and Table 2), with K_D values between 0.83 and 3.7 nM, place AWFC01 as the NS1 best dsRNA ligand described to date.

Table 2. Dissociation constant of various RBDs for AWFC01 K_D values were extracted from three independent EMSA titration experiments (Figure S5). NA is for "non-applicable".

RBD.	K_D (nM)
H7N1	0.83 ± 0.03
H7N1 R38A-K41A	NA (>50 µM)
H5N1	0.98 ± 0.04
H5N1 R38A-K41A	NA (>100)
H7N9	3.1 ± 0.1
H3N2	3.7 ± 0.1
H17N10	1.5 ± 0.1
pdmH1N1	2.05 ± 0.02

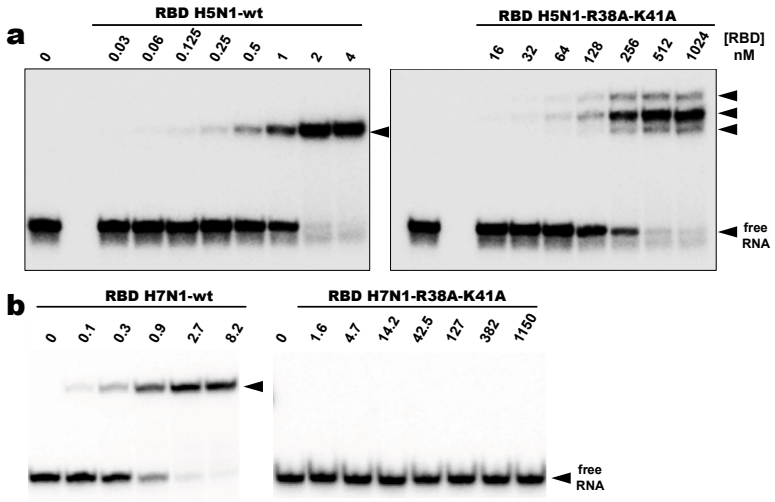
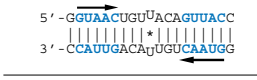
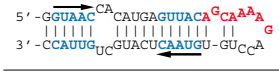
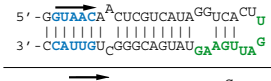
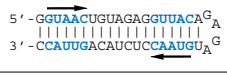
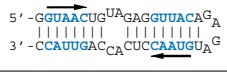
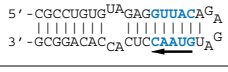
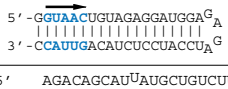
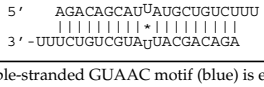


Figure 8. Binding of RBD to AWFC01 (**a**,**b**) Comparative EMSA titration experiments of AWFC01 by the wild type (wt) and the double mutant R38A-K41A RBD from H5N1 and H7N1, respectively. Radiolabeled AWFC01 was incubated in standard conditions with increasing concentrations of wild type or double mutant RBD in standard conditions and analyzed by EMSA. Representative gel autoradiographs are shown.

Surprisingly, the negative impacts of the double substitution (R38A-K41A) were very different for the H5N1 and H7N1 RBDs. Under the stringent binding conditions that we used, the mutant H7N1-RBD was unable to form a stable complex with AWFC01 (Figure 8b, right panel). At a ~50 μM RBD concentration, 50% of the RNA probe was titrated in unstable complexes that yielded a large smear by EMSA (data not shown). On the other hand, while the double substitution of the H5N1-RBD also dramatically reduced its affinity for RNA, a 400 nM concentration of protein was sufficient to titrate 50% of AWFC01 in several distinct complexes (Figure 8a, right panel). These were even more shifted than that formed with the wild type RBD. Their reduced mobility cannot result solely from the different isoelectric points of the wt (pI 6.11) and mutated (pI 5.23) proteins and we surmise that the complexes made with the mutant RBD are less compact. Further, the existence of three distinct complexes could originate either from a structural heterogeneity of the mutant protein or from distinct modes of interaction.

We also compared AWFC01 to the dsRNAs that were evaluated above (i.e., shDM02-GAGA, shDM03-GAGA, shDM05-GAGA, shDM06-GAGA, DM01-midi, and DM03-midi) as regards their affinity for the RBD. We also included in this comparison the model dsRNA that we named "ZKO", the palindromic RNA duplex that was used to solve the crystal structure (PDBid 2ZKO) of H1N1-RBD bound to RNA [28]. To that end, we set up "competition EMSAs", where increasing amounts of the unlabeled competitor RNAs were added after the incubation of radiolabeled AWFC01 with the protein (H5N1 or H7N1 RBD). After quantification of the free and bound fractions of the labeled AWFC01, we calculated the apparent half-maximal Effective Concentration ($EC50_{app}$), expressed as the concentration of competitor required to shift the fraction of bound AWFC01 from 50% to 25%. As indicated by its very low $EC50_{app}$ values (Figure S6 and Table 3), AWFC01 is among the best ligands for the two RBDs that we used. Although three of our dsRNAs also showed similarly low $EC50_{app}$ values (shDM03, shDM02, and shDM06), they do not have the high symmetry of AWFC01. As for ZKO, its high EC50 value suggests a far lower affinity for the RBD.

Table 3. EC50app for the dissociation of [AWFC01/RBD] by cold RNA competitors. EC50app were determined from dose-response curves as described in Materials and Methods. A representative experiment is shown in Figure S6.

RNA Competitor		EC50app (nM)	
2D-structure	**Name**	**H5N1**	**H7N1**
	AWFC01	1.0 ± 0.1	1.7 ± 0.1
	DM01-midi	2.2 ± 0.2	2.7 ± 0.1
	DM03-midi	1.8 ± 0.1	2.8 ± 0.1
	shDM02-GAGA	0.41 ± 0.02	0.24 ± 0.01
	shDM03-GAGA	1.2 ± 0.1	0.64 ± 0.06
	shDM05-GAGA	100 ± 11	ND
	shDM06-GAGA	1.1 ± 0.1	0.85 ± 0.05
	ZKO-RNA	90 ± 4	96 ± 3

The double-stranded GUAAC motif (blue) is emphasized and oriented by the horizontal arrow. The virus specific motifs A and B are shown in red and green, respectively. * depicts the central U*U Hoogsteen base pair in AWFC01 and ZKO-RNA.

3.3. Structural Basis of NS1's Interaction with Model dsRNA AWFC01

In spite of the fact that AWFC01 harbored neither of the virus-specific motifs, both its symmetry and its high affinity for the RBDs prompted us to choose it for our subsequent structural characterization. While essentially all of the RBDs that we purified yielded high quality crystals when crystallized alone, we chose to solve the structure of the avian H7N1 RBD alone and in complex with AWFC01. In parallel and for the sake of comparison, we also crystallized and solved the structure of two other complexes: AWFC01, with the mutant H7N1 RBD harboring the double substitution R38A-K41A and the slightly modified ZKO*-RNA, with the wt RBD. Table S1 (Suppl. Data) summarizes the X-ray data collection and refinement statistics.

The X-ray structure of recombinant wild type H7N1 NS1 RBD (residues 1–73) was solved to 1.93 Å resolution. While it is the first structure that is representative of the RBD of allele B-NS1, its organization is identical to that of the other RBD structures available, forming a symmetric C2 homodimer shaped by three interlocking α-helices from each monomer. Helices α1 (residues S3-M24) and α2 (A30 to L50) are oriented in an antiparallel fashion, whereas helix α3 (L54-K70) sits perpendicular to helix α1 (Figure 9a). The dimerization interface, with an area of 1189Å2, involves 43% of the monomer residues, which are engaged in 15 H-bonds, including 4 salt bridges (PISA server). Superimposition of all available models shows that the six helices are positioned the same way (Figure S7a) and that only small structural divergences occur at the ends of the polypeptide chains. Most of the interfacing residues are located in helices α1 and α3 and in the turn between α1 and α2 (Figure S7b). Three highly conserved residues of α2 helix (F32, R35, and R46) are involved in the stabilization of the dimer, with

R35 being engaged in a salt bridge with D12. Invariant residues R35, R37, R38, and K41, which are involved in RNA binding, adopt different conformations when comparing all the apo RBD homodimers (Figure S8a). Nonetheless (as seen in Figure 9b), the RNA binding interface formed by helices α2 and α2* of the homodimer seems to be well preserved on H7N1. This 3D structure also shows several ligands bound to the RBD, originating from the crystallization buffers. Three 1,2-ethanediol and one partial polyethylene glycol molecules have been assigned in a groove formed by helices α3 and α3*. Interestingly, another partial polyethylene glycol molecule is located in a nonpolar pocket found underneath the RNA binding channel at the RBD molecular two-fold axis (between helices α2 and α2*), making an H-bond with the side chain of R46(A). This arginine residue was also found to contact a malonate ion and a succinic acid molecule in the structure of H1N1 (PDB 3M8T) and H5N1 RBD (3F5T, [48]), respectively.

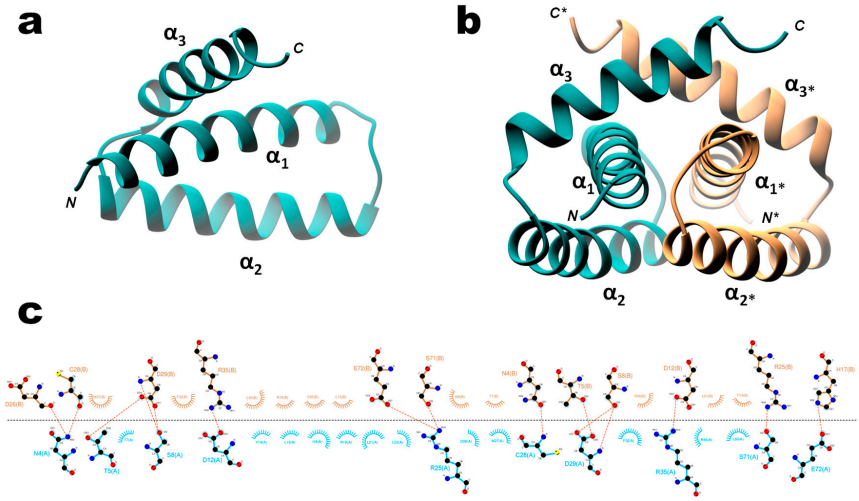


Figure 9. 3D X-ray structure of wild type H7N1 NS1 RBD Ribbon representation of H7N1 NS1 RBD monomer (**a**) and homodimer (**b**), with chain A in dark cyan; chain B in salmon. (**c**) Ligplot representation of the dimer interfacing residues (with chains A and B colored as in b, and atoms colored in black, red and blue for carbon, oxygen and nitrogen atoms, respectively).

To get insight into the sequence-specific recognition of the double stranded RNA by the RBD, we solved, at 1.75 Å resolution, the crystal structure of the complex associating the wild type H7N1-RBD to the AWFC01 duplex (wtH7N1-RDB/AWFC01, Figure 10). The overall structure of the RBD in the complex is identical to that found in the apo form, with minor side chain rearrangements at the binding interface of helix α2 residues R37, R38, and K41 (Figure S7). In RBD chain A, the sidechain of R38 adopted two alternative conformations that were clearly distinguished in electron density maps. In the first conformation ("a" in Table 4), R38 makes two hydrogen bonds with the side chain of D39(A) and a third hydrogen bond with one strand of the RNA, between atoms R38(A).N(η2) and U7(C).OP1. In the second conformation ("b" in Table 4), the side chain of R38 sits in the middle of the RNA minor groove; it interacts with its symmetrically related residue and simultaneously contacts the phosphate groups of U7 nucleotides from the two strands (U7(C).OP1 and U7(D).OP2 atoms) and the phosphate group (OP1 atom) of C6 nucleotide from chain D (Table 4, columns 2 and 3). In RBD chain B, R37 side chain also adopts two conformations: in the first one, it interacts only with D34, while in the second one, it sits closer to the RNA, making H-bonds with the O2′ atom of U7(D) and with D34. The H7N1 RBD sits almost parallel to the dsRNA helix axis and overlays the minor and

the major grooves of the duplex. The assembly buries 1173.8 Å2 of the RNA solvent-accessible area (17.1% of the total) and 1156.7 Å2 for the RBD (14.1% of the total). In the complex, the RNA keeps its A-form conformation but a curvature of about 20° is observed, stretching along base pairs 8 and 9 at one end of the duplex (Figure 10b). As shown in Table 4, all direct interactions between the H7N1 RBD and AWFC01 occur via phosphate/amino acid and sugar/amino acid H-bonds and none through base recognition. The sequence-specificity of the H7N1 RBD for AWFC01, which we observed in the preceding biochemistry experiments, could not easily be explained by the 3D structure of the complex. Of note, like in the apo-structure, we also identified in the structure of the wtH7N1-RBD/AWFC01 complex a partial PEG molecule that almost completely filled the pocket found underneath the RNA binding channel.

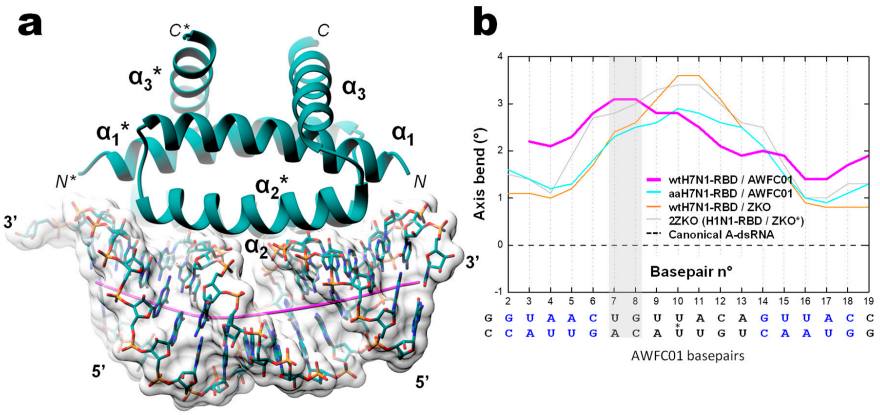


Figure 10. Overall structure of the complex between the wild type H7N1 NS1 RBD protein and the AWFC01 RNA (**a**) 3D structure of the wtH7N1-RBD/AWFC01 complex. The dsRNA axis calculated using Curves+ is represented as an orchid stick. (**b**) Plot of dsRNA axis bend versus base pair, calculated for wtH7N1-RBD/AWFC01 (pink), aaH7N1-RBD/AWFC01 (cyan), wtH7N1-RBD/zko* (orange), and 2ZKO (grey) complexes, relative to the canonical A-form dsRNA (dashed line). The secondary structure of AWFC01 is depicted below.

Table 4. Hydrogen bonds between protein and RNA.

	wt-RBD / AWFC01				aa-RBD / AWFC01				wt-RBD / ZKO *				2ZKO			
	RBD chain A		RBD chain B		RBD chain A		RBD chain B		RBD chain A		RBD chain B		RBD chain A		RBD chain B	
Atom 1	**dist.**	Atom 2	**dist.**	Atom 2	**dist.**	Atom 2	**dist.**	Atom 2	**dist.**	Atom 2	**dist.**	Atom 2	**dist.**	Atom 2	**dist.**	Atom 2
Amino-acid / phosphate H-bonds																
H(-1) N									**3.37**	U19(C)OP1						
R38 N(ε)	**2.92 (b)**	C6(D)OP1	**2.76**	C6(C)OP1					**3.18**	G6(D)OP1	**2.98**	G6(C)OP1	**3.47**	G6(D)OP1	**3.42**	G6(C)OP1
R38 N(η1)	**3.23 (b)**	C6(D)OP														
R38 N(η1)	**2.99 (b)**	U7(D)OP	**2.88**	U7(D)OP1					**2.59**	C7(C)OP1	**2.77**	C7(D)OP1	**2.76**	C7(C)OP1	**2.76**	C7(D)OP1
R38 N(η2)			**3.26**	C6(C)OP1					**3.18**	G6(D)OP1	**3.16**	G6(C)OP1	**3.43**	G6(D)OP1	**3.35**	G6(C)OP1
R38 N(η2)	**2.32 (a)** **2.27 (b)**	U7(C)OP	**3.06**	U7(C)OP2					**2.96**	C7(D)OP2	**2.77**	C7(C)OP2	**2.83**	C7(D)OP2	**2.87**	C7(C)OP2
S42 O(γ)													**2.87**	A5(D)O3′	**2.91**	A5(C)O3′
Amino-acid / ribose H-bonds																
S1 O(γ)									**2.49**	U17(C)O2′						
P31 O	**3.71**	C6(C)O2′	**3.61**	C6(D)O2′	**3.59**	C6(C)O2′	**3.54**	C6(D)O2′	**3.63**	G6(C)O2′	**3.69**	G6(D)O2′			**3.67**	G6(D)O2′
R35 N(η1)	**3.54**	C6(C)O2′	**3.66**	C6(D)O2′	**3.55**	C6(C)O2′	**3.51**	C6(D)O2′	**3.44**	G6(C)O2′	**3.38**	G6(D)O2′	**3.14**	G6(C)O2′	**3.22**	G6(D)O2′
R37 N(η)			**3.43 (a)**	U7(D)O2′												
T49 O(γ)	**3.19**	A4(D)O2′	**3.17**	A4(C)O2′	**3.00**	A4(D)O2′	**3.11**	A4(C)O2′	**2.81**	C4(D)O2′	**3.10**	C4(C)O2′	**2.82**	C4(D)O2′	**2.80**	C4(C)O2′
T49 O(γ)	**3.01**	A4(D)O4′	**3.06**	A4(C)O4′	**3.08**	A4(D)O4′	**2.96**	A4(C)O4′	**2.99**	C4(D)O4′	**2.94**	C4(C)O4′	**3.03**	C4(D)O4′	**3.03**	C4(C)O4′

*: a and b (second and fourth columns) refer to the two alternative conformations that were observed for the R38 and R37 side chains. The shaded cells highlight the distinctive asymmetry in the structure of the complex wt-RBD/AWFC01. C or D after the nucleotide number refer to nucleotide chains C and D.

In order to search for complementary clues for the apparent sequence specificity of the H7N1 RBD for AWFC01, we solved the 3D structures of two other complexes: (i) R38A-K41A double mutant H7N1 RDB with AWFC01 (aaH7N1-RBD/AWFC01) and (ii) wild type H7N1 RBD, with the slightly modified nonspecific RNA "ZKO*" (wtH7N1-RBD/ZKO*). The dsRNA that Cheng et al. used to solve their structure [28] was slightly modified by removing its terminal, unpaired nucleotides. While the two complexes crystallized in different conditions and space groups, noteworthy similarities were observed (Figures S8 and S9) with low rmsd values of 0.33 Å for the proteins and 0.71 Å for the RNA. When comparing by superimposition the structures of the three RBD-RNA complexes, subtle structural differences were identified, mainly for the RNA molecule. Relative to the wtH7N1-RBD/AWFC01 complex, the RNAs in complexes aaH7N1-RBD/AWFC01 and wtH7N1-RBD/ZKO* differ by rmsd values of 2.00 Å and 2.21 Å, respectively, whereas the difference is much smaller as regards the RBD (rmsd of 0.26 Å and 0.33 Å, respectively, Figure S10). We observed the same difference when comparing the wtH7N1-RBD/AWFC01 complex to the 2ZKO complex, with rmsd values of 2.10 Å for the RNA phosphodiester backbone atoms and 0.45 Å for the protein main chain atoms. Since the main differences occurred at the RNA level, we used CURVES+ [42] to dissect and compare the nucleic acid conformations in the different complexes (Figures S11 and S12). Surprisingly, this revealed that in the high affinity dsRNA (AWFC01), the RNA axis is bent in a non-symmetrical fashion, whereas in the two other complexes (aaH7N1-RBD/AWFC01 and wtH7N1-RBD/ZKO*), the RNA axis bend is distributed symmetrically from the centre of the duplex (Figure 10b). For AWFC01, axis bending is at a maximum between the seventh and eighth basepair at one end of the duplex, right after the GUAAC sequence. At this position, the minor groove is less wide and deep than that at the other side of the dsRNA, while the opposite is seen for the major groove. This peculiar asymmetry results from a corresponding asymmetry in the network of amino-acid/nucleotides hydrogen bonds. As highlighted in Table 4 (shaded cells), several hydrogen bonds in the wt-RBD/AWFC01 complex are distinctively asymmetric, compared to their equivalents in the other structures: for instance, the η1 nitrogen atom of R38(chain A) interacts with the OP2 atom of U7(chain D) and contrary to the expected symmetrical relation that is observed in the other structures, the η1 nitrogen atom of R38(chain B) also interacts with U7 of the same chain D, this time with the OP1 atom. A similar relation is observed with the η2 nitrogen of R38 and U7 of chain C. As a consequence, the seventh nucleotide of both chains of the dsRNA are pulled by these interactions, bending its axis and skewing its shape.

Closer inspection of the protein-RNA interactions did not allow us to reveal sequence-related features that would account for the specificity and high-affinity of the RBD-AWFC01 interaction. As shown in Table 4, essentially four amino-acids formed a network of hydrogen bonds with either the phosphate groups or the ribose of the RNA backbone, with no sequence-specific bonds involving the nucleobases. Apart from the peculiarities of the hydrogen bonds involving the three nitrogen atoms of R38 (discussed above), the H-bond network is symmetric and involves the oxygen atom of P31, the γ oxygen atom of T49, and the η1 nitrogen of R35, which consistently make hydrogen bonds with O2′ of nucleotide 6 (P31), O2′ and O4′ of nucleotide 4 (T49), and O2′ of nucleotide 6 (R35 N(η1)). Interestingly, according to molecular dynamics simulations [49], P31 at the N-terminus of alpha-helix 2 is at the centre of the most flexible region in the RBD.

4. Discussion

In this work, we set out to investigate the importance of three distinct sequence and structure motifs in the interaction of influenza A virus NS1 with RNAs. These motifs were previously identified through an in vitro SELEX approach in which two recombinant proteins NS1, representative of its two alleles, were used to iteratively select high-affinity synthetic RNAs [13]. They included the double-stranded motif of sequence GUAAC, along with two sequence motifs of undetermined structure that are strictly or highly conserved in influenza virus A-derived positive strand RNAs: AGCAAAAG, which is strictly conserved at the 5′-end of all complementary RNAs and UGAUUGAAG, which is highly conserved in the 3′-UTR of NS1 mRNA, fourteen nucleotides downstream of NS1's stop codon.

We first investigated the biological relevance of one of the two virus-specific motifs in various steps of the virus cycle. We considered it unrealistic to attempt to rescue mutant viruses with a mutated AGCAAAAG motif since it would be impossible to disentangle the putative NS1-related impacts from the expected impacts any such mutation in the viral polymerase promoter would have on the viral cycle. We therefore limited our in vivo approach to the study of viruses harboring mutations in the conserved motif UGAUUGAAG within the 3'UTR region of NS1's mRNA. In addition and for the sake of comparison, we also mutated two other highly conserved motifs in the 3'UTR region of NS1's mRNA, including the very similar sequence UGCUUGAAG, which is present at the distal end of this long 3'UTR. These mutations did not change the peptide sequence of NEP. All three single-mutant WSN viruses harboring either of these three mutated motifs (mutB, mutC, and mutD) were readily rescued and their phenotype in cultured cells was undistinguishable from that of the wt-virus. Furthermore, viruses harboring all three mutated sites also exhibited a wild type phenotype, as assessed (i) by the kinetics of multicycle virus growth and (ii) by the kinetics of viral protein (NS1 and NP) accumulation in mutant virus-infected cells. We also attempted to evaluate whether NS1 could modulate the activity of the viral polymerase on a "minireplicon" consisting of the NS genome segment where NS1's ORF was replaced by that of the Renilla luciferase and whether mutations in the B, C, or D motifs could alter the replication, transcription, or translation of this minigenome. Indeed, we consistently observed that NS1 dramatically increased the activity of the viral polymerase (4-fold to 6-fold increase). However, this polymerase-enhancing activity of NS1 was equally observed, with the minireplicons harboring either of the B, C, or D mutation, thus ruling out any major role of these motifs in this observed activity. Moreover, even when the minireplicon system was complemented with the R38A-K41A NS1 rather than with wt-NS1, the enhancing activity was still observed, albeit to a lower level. This suggests that the polymerase-enhancing activity does not rely solely on the RNA-binding ability of the RBD. Indeed, several studies had shown that NS1 interacts with the viral polymerase [50], most likely through a direct interaction of its RBD with the viral nucleoprotein [51]. It should be highlighted, however, that the read-out of the minireplicon system (i.e., the Renilla luciferase activity) results from at least three activities combined: (i) replication and (ii) transcription of the minigenome by the viral polymerase, along with (iii) translation of the viral-like mRNA. Therefore, it is likely that the observed effects of NS1 on that system combine the enhanced activity of the viral polymerase with the increased translation of the viral-like mRNA, which has been demonstrated by several studies [14,52,53]. Taken together, although our data confirm the important activity of NS1 in the viral cycle (notably in relation with the viral polymerase and the translation of viral mRNAs), they provide no evidence for our initial hypothesis that NS1's binding to the 3'UTR of its own mRNA could be critical to the viral cycle.

In spite of the apparently negligible relevance of the putative interaction of NS1 with the B motif, we stress that we could not assess, in the context of the viral cycle, the biological relevance of its interaction with motif A, i.e., the AGCAAAAG motif at the 5'end of all virus-derived RNAs of positive-polarity. These two virus-specific motifs probably represent only a fraction of the putative NS1-binding motifs in viral or cellular RNAs and indeed, a systematic search showed that NS1 preferentially binds intronic sequence motifs in a subset of host cell mRNAs including that of the retinoic-acid induced gene RIG-I [26], resulting in decreased processing of pre-mRNAs and a reduced expression. This illustrates the view that the high concentration of NS1 in the infected cell [19], combined with its presence in several subcellular compartments, allows it to interact, with a broad range of affinities, with several RNAs that could represent the targets of its biological activities. We therefore chose to study in depth the determinants that allow a given RNA to interact with NS1, starting from the RNA aptamers that had been selected previously through the SELEX approach.

We first showed that the RBD preferentially interacted with unbranched dsRNA structures with a minimal length of about 20 base pairs. While we observed no high-affinity interaction of the RBD with short single-stranded RNAs consisting only of the two 8 nt- or 9 nt- virus-specific sequence motifs, our data show that the RBD, by itself, was able to interact with the AGCAAAAG motif with no obvious involvement of the effector domain. This sequence-specific interaction was observed to

occur cooperatively on a preformed 1:1 RBD:RNA ratio complex, associating one RBD bound to the dsRNA structure of aptamer derivative DM01-midi. We hypothesize that this first interaction induces a structural change in the dsRNA that propagates to its apical loop, thereby facilitating the binding of a second RBD to this apical loop containing the AGCAAAAG motif. Such a behavior was not observed in the other aptamer derivate, but in that case, we observed that the effector domain was required for a cooperative-like behavior and we surmise that the ED specifically recognizes the second virus-specific motif, UGAUUGAAG, or contributes at its RBD recognition. The structural dynamic of full-length NS1, especially regarding the relative orientation of its two domains [16,17], could also be involved in this cooperative RNA binding between NS1 molecules and the RBD in the solution could exist in equilibrium between different conformations that are non-equivalent for the cooperative binding to RNA.

We then chose to focus on the interaction of NS1 with the double-stranded structure. Altogether, our experiments had shown that an optimal interaction with the RBD required the presence of the GUAAC motif positioned close to the free blunt end of the dsRNA. We especially addressed the question, what mechanism could allow the RBD to specifically recognize this sequence, given the non-specificity that is generally observed in dsRNA-binding proteins? Such a lack of specificity is notably observed in the B2 protein of Flock House virus [54], which like NS1, interacts with dsRNA through two antiparallel alpha-helices of the same sequence (PDB 2AZ2). Indeed, there is a notorious paradox in the sequence recognition within double-stranded RNAs since the nucleobases are not easily accessible [55,56] and most of the interactions involve either the phosphodiester backbone or the 2′-hydroxyl groups of the ribose moiety. In order to elucidate this point, we designed AWFC01, a model double-stranded RNA of perfect symmetry containing a central U*U Hoogsteen base pair. We showed the very high affinity of this model dsRNA for the RBD and determined the structure of three RBD-dsRNA complexes, including two with AWFC01. Unexpectedly, the high symmetry of both the secondary structure of AWFC01 and the 3D structure of the RBD did not result in a corresponding symmetry of the dsRNA 3D structure in the complex. Instead, we observed a pronounced asymmetry in the network of hydrogen bonds that links the dsRNA to the RBD, owing to the position of residue R38 in polypeptide chain A. This skewed network pulls the seventh nucleotide of the two RNA chains, thereby locally bending its axis and skewing the overall shape of the dsRNA. This probably constitutes an important determinant in the sequence-specific recognition of the dsRNA. The A-form helix of the dsRNA can undergo subtle conformation changes that alter its shape and it is believed that sequence specificity in dsRNAs often relies on recognition of shape [56,57], which in itself can be dependent on the underlying sequence of the dsRNA. We hypothesize that the affinity of the dsRNA-RBD interaction is dramatically increased by the altered shape of the RNA or more precisely, by the sequence-dependent predisposition to adopt this given shape.

Besides allowing us to propose an explanation to the sequence specificity within dsRNAs, our structure and the distinctive bending of the dsRNA axis suggests that the RBD interaction can indeed alter the conformation of the dsRNA (although we did not solve the structure of the RNA alone). This conformational change likely can propagate in RNAs with more complex structures and as discussed above, this could result in aptamer-derived DM01-midi in the cooperative binding to the virus-specific sequence in its apical loop, through a necessarily distinct binding mode. We stress that our study was performed with a bacterially-expressed NS1, while several studies have shown that NS1 can be modified post-translationally, notably by phosphorylation at several sites including serine 48 and threonine 49 in its RNA-binding interface [58–61]. It was also shown that phosphorylation of threonine 49 abolished the anti-interferon activity of NS1 and slightly reduced its RNA-binding activity [60]. Phosphorylation of Ser 48 and Thr 49 in the infected cell could therefore provide an optional switch that modulates the RNA-binding capacity of NS1, adding a layer of complexity in the dynamic network of NS1-RNA interactions in the infected cells.

While keeping in mind that the high-affinity ligands are not always optimal for the biological function [62], we acknowledge that our data do not provide direct evidence as to the biological

relevance of these sequence-specific interactions of NS1 with RNAs and do not support the importance of the putative interaction of NS1 with its own mRNA. However, our structural and biochemical data, combined with the known abundance of NS1 in the infected cell, support the view that NS1, through its RBD, can interact by more than one single mode of binding and with a broad range of affinities, with several distinct RNAs in the infected cell. As we have hypothesized previously [13], NS1 could compete with spliceosomal protein U1C in its interaction with the U1-mRNA duplex at the splice donor site [63,64], owing to its preferential recognition of the GUAAC motif. Further, the specific recognition of the AGCAAAAG motif most likely plays an important role in the viral cycle [14,52], although we found no suitable system to evaluate it in the viral context. Nevertheless, because of the biological importance of this interaction, we believe that it would be worthwhile to solve the structure of NS1 interacting with the viral motif AGCAAAAG. Indeed, in spite of its low affinity, the interaction of the RBD that we observed with aptamer-derived DM01-short (Figure S3) is a promising hint as to the possibility of solving the structure of such a complex and hopefully unveiling a new mode of interaction of NS1 with RNAs.

Supplementary Materials: The following are available online at http://www.mdpi.com/1999-4915/12/9/947/s1, Figure S1: RNA probes used in this study, Figure S2: plaque phenotype of the mutant viruses; Figure S3: Comparative titration of DM01-short, DM01-MotA, DM03-short and DM03-MotB by RBD; Figure S4: Comparative binding properties of selected RBDs with DM01-midi and DM03-midi; Figure S5: Titration curves obtained for the recognition of AWFC01 by various RBDs; Figure S6: EC50app for the displacement of the preformed RBD/AWFC01 by other RNA probes; Figure S7: Comparison of H7N1 NS1 RBD with allele A RBDs; Figure S8: RNA binding interface of NS1 RBD; Figure S9: Conserved positively charged residues of the RNA binding interface; Figure S10: Comparison of the protein backbones in the structure of the three RBD-RNA complexes; Figures S11 and S12: RNA structural parameters from Curves+; Table S1: X-ray data collection and refinement statistics.

Author Contributions: Conceptualization, D.M., B.C., and F.C.; methodology, D.M., B.C., and F.C.; investigation, A.W., E.K., and V.G.; resources, S.T.; writing—original draft preparation, D.M. and B.C.; writing—review and editing, D.M., B.C., F.C., S.T., and A.W.; supervision, D.M., B.C., and F.C.; funding acquisition, D.M., B.C., and S.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by REGION CENTRE-Val de Loire, France (RNApur project, convention n 2015-00099396). A.W. was financially supported by REGION CENTRE-Val de Loire (convention n 201600109193).

Acknowledgments: We are grateful to Ervin Fodor for the kind gift of the 12-plasmids WSN reverse genetics system. We thank the staff of the SOLEIL synchrotron facility (Gif-sur-Yvette, France) for the use of PROXIMA-2 beamline and Région Centre-Val de Loire for its financial support. B.C. thanks the Centre National pour la Recherche Scientifique (CNRS).

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Iuliano, A.D.; Roguski, K.M.; Chang, H.H.; Muscatello, D.J.; Palekar, R.; Tempia, S.; Cohen, C.; Gran, J.M.; Schanzer, D.; Cowling, B.J.; et al. Estimates of global seasonal influenza-associated respiratory mortality: A modelling study. *Lancet* **2017**, *391*, 1285–1300. [CrossRef]
2. World Health Organization. *Global Influenza Strategy 2019–2030*; World Health Organization: Geneva, Switzerland, 2019; p. 31. ISBN 978-92-4-151532-0.
3. Duwe, S. Influenza viruses–Antiviral therapy and resistance. *GMS Infect. Dis.* **2017**, *5*. [CrossRef]
4. Hussain, M.; Galvin, H.D.; Haw, T.Y.; Nutsford, A.N.; Husain, M. Drug resistance in influenza a virus: The epidemiology and management. *Infect. Drug Resist.* **2017**, *10*, 121–134. [CrossRef] [PubMed]
5. O'Hanlon, R.; Shaw, M.L. Baloxavir marboxil: The new influenza drug on the market. *Curr. Opin. Virol.* **2019**, *35*, 14–18. [CrossRef] [PubMed]
6. Abed, Y.; Boivin, G. A review of clinical influenza A and B infections with reduced susceptibility to both Oseltamivir and Zanamivir. *Open Forum Infect. Dis.* **2017**, *4*, 105. [CrossRef]
7. Engel, D.A. The influenza virus NS1 protein as a therapeutic target. *Antivir. Res.* **2013**, *99*, 409–416. [CrossRef]

8. Newby, C.M.; Sabin, L.; Pekosz, A. The RNA Binding Domain of Influenza A Virus NS1 Protein Affects Secretion of Tumor Necrosis Factor Alpha, Interleukin-6, and Interferon in Primary Murine Tracheal Epithelial Cells. *J. Virol.* **2007**, *81*, 9469–9480. [CrossRef]
9. Hale, B.G.; Randall, R.E.; Ortín, J.; Jackson, D. The multifunctional NS1 protein of influenza a viruses. *J. Gen. Virol.* **2008**, *89*, 2359–2376. [CrossRef]
10. Marc, D. Influenza virus non-structural protein NS1: Interferon antagonism and beyond. *J. Gen. Virol.* **2014**, *95*, 2594–2611. [CrossRef]
11. Hatada, E.; Saito, S.; Okishio, N.; Fukuda, R. Binding of the influenza virus NS1 protein to model genome RNAs. *J. Gen. Virol.* **1997**, *78*, 1059–1063. [CrossRef]
12. Lu, Y.; Qian, X.Y.; Krug, R.M. The influenza virus NS1 protein: A novel inhibitor of pre-mRNA splicing. *Genes Dev.* **1994**, *8*, 1817–1828. [CrossRef]
13. Marc, D.; Barbachou, S.; Soubieux, D. The RNA-binding domain of influenzavirus non-structural protein-1 cooperatively binds to virus-specific RNA sequences in a structure-dependent manner. *Nucleic Acids Res.* **2012**, *41*, 434–449. [CrossRef] [PubMed]
14. Marión, R.M.; Aragón, T.; Beloso, A.; Nieto, A.; Ortín, J. The N-terminal half of the influenza virus NS1 protein is sufficient for nuclear retention of mRNA and enhancement of viral mRNA translation. *Nucleic Acids Res.* **1997**, *25*, 4271–4277. [CrossRef] [PubMed]
15. Qiu, Y.; Krug, R.M. The influenza virus NS1 protein is a poly(A)-binding protein that inhibits nuclear export of mRNAs containing poly(A). *J. Virol.* **1994**, *68*, 2425–2432. [CrossRef]
16. Carrillo, B.; Choi, J.-M.; Bornholdt, Z.A.; Sankaran, B.; Rice, A.P.; Prasad, B.V.V. The Influenza A Virus Protein NS1 Displays Structural Polymorphism. *J. Virol.* **2014**, *88*, 4113–4122. [CrossRef] [PubMed]
17. Hale, B.G. Conformational plasticity of the influenza a virus NS1 protein. *J. Gen. Virol.* **2014**, *95*, 2099–2105. [CrossRef] [PubMed]
18. Zhou, B.; Ma, J.; Liu, Q.; Bawa, B.; Wang, W.; Shabman, R.S.; Duff, M.; Lee, J.; Lang, Y.; Cao, N.; et al. Characterization of Uncultivable Bat Influenza Virus Using a Replicative Synthetic Virus. *PLoS Pathog.* **2014**, *10*, e1004420. [CrossRef] [PubMed]
19. Trapp, S.; Soubieux, D.; Lidove, A.; Esnault, E.; Lion, A.; Guillory, V.; Wacquiez, A.; Kut, E.; Quere, P.; Larcher, T.; et al. Major contribution of the RNA-binding domain of NS1 in the pathogenicity and replication potential of an avian H7N1 influenza virus in chickens. *Virol. J.* **2018**, *15*, 55. [CrossRef]
20. De Chassey, B.; Aublin-Gex, A.; Ruggieri, A.; Meyniel-Schicklin, L.; Pradezynski, F.; Davoust, N.; Chantier, T.; Tafforeau, L.; E Mangeot, P.; Ciancia, C.; et al. The interactomes of influenza virus NS1 and NS2 proteins identify new host factors and provide insights for ADAR1 playing a supportive role in virus replication. *PLoS Pathog.* **2013**, *9*, e1003440. [CrossRef]
21. Satterly, N.; Tsai, P.-L.; Van Deursen, J.; Nussenzveig, D.R.; Wang, Y.; Faria, P.A.; Levay, A.; Levy, D.E.; Fontoura, B.M.A. Influenza virus targets the mRNA export machinery and the nuclear pore complex. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 1853–1858. [CrossRef]
22. Shapira, S.D.; Gat-Viks, I.; Shum, B.O.; Dricot, A.; De Grace, M.M.; Wu, L.; Gupta, P.B.; Hao, T.; Silver, S.J.; Root, D.E.; et al. A Physical and Regulatory Map of Host-Influenza Interactions Reveals Pathways in H1N1 Infection. *Cell* **2009**, *139*, 1255–1267. [CrossRef]
23. Watanabe, T.; Kawakami, E.; Shoemaker, J.E.; Lopes, T.J.S.; Matsuoka, Y.; Tomita, Y.; Kozuka-Hata, H.; Gorai, T.; Kuwahara, T.; Takeda, E.; et al. Influenza virus-host interactome screen as a platform for antiviral drug development. *Cell Host Microbe* **2014**, *16*, 795–805. [CrossRef] [PubMed]
24. Ayllon, J.; García-Sastre, A. The NS1 Protein: A Multitasking Virulence Factor. *Inducible Lymphoid Organs* **2014**, *386*, 73–107. [CrossRef]
25. Zhang, K.; Xie, Y.; Muñoz-Moreno, R.; Wang, J.; Zhang, L.; Esparza, M.; García-Sastre, A.; Fontoura, B.M.A.; Ren, Y. Structural basis for influenza virus NS1 protein block of mRNA nuclear export. *Nat. Microbiol.* **2019**, *4*, 1671–1679. [CrossRef] [PubMed]
26. Zhang, L.; Wang, J.; Muñoz-Moreno, R.; Kim, M.; Sakthivel, R.; Mo, W.; Shao, D.; Anantharaman, A.; García-Sastre, A.; Conrad, N.K.; et al. Influenza Virus NS1 Protein-RNA Interactome Reveals Intron Targeting. *J. Virol.* **2018**, *92*, 92. [CrossRef] [PubMed]
27. Chen, G.; Ma, L.-C.; Wang, S.; Woltz, R.L.; Grasso, E.M.; Montelione, G.T.; Krug, R.M. A double-stranded RNA platform is required for the interaction between a host restriction factor and the NS1 protein of influenza a virus. *Nucleic Acids Res.* **2020**, *48*, 304–315. [CrossRef]

28. Cheng, A.; Wong, S.M.; Yuan, Y.A. Structural basis for dsRNA recognition by NS1 protein of influenza a virus. *Cell Res.* **2008**, *19*, 187–195. [CrossRef]
29. Yin, C.; Khan, J.A.; Swapna, G.V.T.; Ertekin, A.; Krug, R.M.; Tong, L.; Montelione, G.T. Conserved Surface Features Form the Double-stranded RNA Binding Site of Non-structural Protein 1 (NS1) from Influenza A and B Viruses. *J. Boil. Chem.* **2007**, *282*, 20584–20592. [CrossRef]
30. Fodor, E.; Devenish, L.; Engelhardt, O.G.; Palese, P.; Brownlee, G.G.; García-Sastre, A. Rescue of Influenza A Virus from Recombinant DNA. *J. Virol.* **1999**, *73*, 9679–9682. [CrossRef]
31. Da Costa, B.; Sausset, A.; Munier, S.; Ghounaris, A.; Naffakh, N.; Le Goffic, R.; Delmas, B. Temperature-Sensitive Mutants in the Influenza A Virus RNA Polymerase: Alterations in the PA Linker Reduce Nuclear Targeting of the PB1-PA Dimer and Result in Viral Attenuation. *J. Virol.* **2015**, *89*, 6376–6390. [CrossRef]
32. Matrosovich, M.; Matrosovich, T.; Garten, W.; Klenk, H.-D. New low-viscosity overlay medium for viral plaque assays. *Virol. J.* **2006**, *3*, 63. [CrossRef] [PubMed]
33. Talon, J.; Horvath, C.M.; Polley, R.; Basler, C.F.; Muster, T.; Palese, P.; García-Sastre, A. Activation of Interferon Regulatory Factor 3 Is Inhibited by the Influenza A Virus NS1 Protein. *J. Virol.* **2000**, *74*, 7989–7996. [CrossRef] [PubMed]
34. Studier, F.W. Protein production by auto-induction in high-density shaking cultures. *Protein Expr. Purif.* **2005**, *41*, 207–234. [CrossRef] [PubMed]
35. Ryder, S.P.; Recht, M.I.; Williamson, J.R. Quantitative Analysis of Protein-RNA Interactions by Gel Mobility Shift. *Methods Mol. Biol.* **2008**, *488*, 99–115. [CrossRef]
36. Kabsch, W. Xds. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 125–132. [CrossRef]
37. vans, P.R.; Murshudov, G.N. How good are my data and what is the resolution? *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2013**, *69*, 1204–1214. [CrossRef]
38. McCoy, A.J. Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr. Sect. D Boil. Crystallogr.* **2006**, *63*, 32–41. [CrossRef]
39. Adams, P.D.; Afonine, P.V.; Bunkóczi, G.; Chen, V.B.; Davis, I.W.; Echols, N.; Headd, J.J.; Hung, L.-W.; Kapral, G.J.; Grosse-Kunstleve, R.W.; et al. PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D Boil. Crystallogr.* **2010**, *66*, 213–221. [CrossRef]
40. Emsley, P.; Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr. Sect. D Boil. Crystallogr.* **2004**, *60*, 2126–2132. [CrossRef]
41. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng, E.C.; Ferrin, T.E. UCSF Chimera–a visualization system for exploratory research and analysis. *J. Comput. Chem* **2004**, *25*, 1605–1612. [CrossRef]
42. Blanchet, C.; Pasi, M.; Zakrzewska, K.; Lavery, R. CURVES+ web server for analyzing and visualizing the helical, backbone and groove parameters of nucleic acid structures. *Nucleic Acids Res.* **2011**, *39*, W68–W73. [CrossRef] [PubMed]
43. Lu, X.-J.; Olson, W.K. 3DNA: A versatile, integrated software system for the analysis, rebuilding and visualization of three-dimensional nucleic-acid structures. *Nat. Protoc.* **2008**, *3*, 1213–1227. [CrossRef] [PubMed]
44. Akarsu, H.; Burmeister, W.P.; Petosa, C.; Petit, I.; Müller, C.W.; Ruigrok, R.W.; Baudin, F. Crystal structure of the M1 protein-binding domain of the influenza A virus nuclear export protein (NEP/NS2). *EMBO J.* **2003**, *22*, 4646–4655. [CrossRef] [PubMed]
45. Fabrini, R.; De Luca, A.; Stella, L.; Mei, G.; Orioni, B.; Ciccone, S.; Federici, G.; Bello, M.L.; Ricci, G. Monomer−dimer equilibrium in glutathione transferases: A critical re-examination. *Biochemistry* **2009**, *48*, 10473–10482. [CrossRef]
46. Orita, M.; Nishikawa, F.; Shimayama, T.; Taira, K.; Endo, Y.; Nishikawa, S. High-resolution NMR study of a synthetic oligoribonucleotide with a tetranucleotide GAGA loop that is a substrate for the cytotoxic protein, ricin. *Nucleic Acids Res.* **1993**, *21*, 5670–5678. [CrossRef] [PubMed]
47. Sheng, J.; Gan, J.; Soares, A.S.; Salon, J.; Huang, Z. Structural insights of non-canonical U•U pair and Hoogsteen interaction probed with Se atom. *Nucleic Acids Res.* **2013**, *41*, 10476–10487. [CrossRef]
48. Bornholdt, Z.A.; Prasad, B.V.V. X-ray structure of NS1 from a highly pathogenic H5N1 influenza virus. *Nature* **2008**, *456*, 985–988. [CrossRef]

49. Hussein, H.A.; Geneix, C.; Cauvin, C.; Marc, D.; Flatters, D.; Camproux, A.-C. Molecular Dynamics Simulations of Influenza A Virus NS1 Reveal a Remarkably Stable RNA-Binding Domain Harboring Promising Druggable Pockets. *Viruses* **2020**, *12*, 537. [CrossRef]
50. Mari, R.M.; Ort, J.; de la Luna, S. Influenza virus NS1 protein interacts with viral transcription-replication complexes in vivo. *J. Gen. Virol.* **1997**, *78*, 2447–2451. [CrossRef]
51. Robb, N.C.; Chase, G.; Bier, K.; Vreede, F.T.; Shaw, P.-C.; Naffakh, N.; Schwemmle, M.; Fodor, E. The Influenza a Virus NS1 Protein Interacts with the Nucleoprotein of Viral Ribonucleoprotein Complexes. *J. Virol.* **2011**, *85*, 5228–5231. [CrossRef]
52. De La Luna, S.; Fortes, P.; Beloso, A.; Ortín, J. Influenza virus NS1 protein enhances the rate of translation initiation of viral mRNAs. *J. Virol.* **1995**, *69*, 2427–2433. [CrossRef]
53. Panthu, B.; Terrier, O.; Carron, C.; Traversier, A.; Corbin, A.; Balvay, L.; Lina, B.; Rosa-Calatrava, M.; Ohlmann, T. The NS1 Protein from Influenza Virus Stimulates Translation Initiation by Enhancing Ribosome Recruitment to mRNAs. *J. Mol. Boil.* **2017**, *429*, 3334–3352. [CrossRef] [PubMed]
54. Monsion, B.; Incarbone, M.; Hleibieh, K.; Poignavent, V.; Ghannam, A.; Dunoyer, P.; Daeffler, L.; Tilsner, J.; Ritzenthaler, C. Efficient Detection of Long dsRNA in Vitro and in Vivo Using the dsRNA Binding Domain from FHV B2 Protein. *Front. Plant Sci.* **2018**, *9*. [CrossRef] [PubMed]
55. Corley, M.; Burns, M.C.; Yeo, G.W. How RNA-Binding Proteins Interact with RNA: Molecules and Mechanisms. *Mol. Cell* **2020**, *78*, 9–29. [CrossRef] [PubMed]
56. Masliah, G.; Barraud, P.; Allain, F.H.-T. RNA recognition by double-stranded RNA binding domains: A matter of shape and sequence. *Cell. Mol. Life Sci.* **2012**, *70*, 1875–1895. [CrossRef]
57. Stefl, R.; Skrisovska, L.; Allain, F.H.-T. RNA sequence- and shape-dependent recognition by proteins in the ribonucleoprotein particle. *EMBO Rep.* **2005**, *6*, 33–38. [CrossRef]
58. Hsiang, T.-Y.; Zhou, L.; Krug, R.M. Roles of the Phosphorylation of Specific Serines and Threonines in the NS1 Protein of Human Influenza A Viruses. *J. Virol.* **2012**, *86*, 10370–10376. [CrossRef]
59. Hutchinson, E.C.; Denham, E.M.; Thomas, B.; Trudgian, D.C.; Hester, S.S.; Ridlova, G.; York, A.; Turrell, L.; Fodor, E. Mapping the Phosphoproteome of Influenza A and B Viruses by Mass Spectrometry. *PLoS Pathog.* **2012**, *8*, e1002993. [CrossRef]
60. Kathum, O.A.; Schrader, T.; Anhlan, D.; Nordhoff, C.; Liedmann, S.; Pande, A.; Mellmann, A.; Ehrhardt, C.; Wixler, V.; Ludwig, S. Phosphorylation of influenza A virus NS1 protein at threonine 49 suppresses its interferon antagonistic activity. *Cell. Microbiol.* **2016**, *18*, 784–791. [CrossRef]
61. Santos, A.; Pal, S.; Chacón, J.; Meraz, K.; Gonzalez, J.; Prieto, K.; Rosas-Acosta, G. SUMOylation Affects the Interferon Blocking Activity of the Influenza A Nonstructural Protein NS1 without Affecting Its Stability or Cellular Localization. *J. Virol.* **2013**, *87*, 5602–5620. [CrossRef]
62. Shultzaberger, R.K.; Schneider, T.D. Using sequence logos and information analysis of Lrp DNA binding sites to investigate discrepancies between natural selection and SELEX. *Nucleic Acids Res.* **1999**, *27*, 882–887. [CrossRef] [PubMed]
63. Kondo, Y.; Oubridge, C.; Van Roon, A.-M.M.; Nagai, K. Crystal structure of human U1 snRNP, a small nuclear ribonucleoprotein particle, reveals the mechanism of 5′ splice site recognition. *eLife* **2015**, *4*, 4. [CrossRef] [PubMed]
64. Krummel, D.A.P.; Oubridge, C.; Leung, A.; Li, J.; Nagai, K. Crystal structure of human spliceosomal U1 snRNP at 5.5 Å resolution. *Acta Crystallogr. Sect. A Found. Crystallogr.* **2009**, *65*, 475–480. [CrossRef]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

PUBLICATION 2 : Major contribution of the RNA-binding domain of NS1 in the pathogenicity and replication potential of an avian H7N1 influenza virus in chickens

Annexe 14 : Major contribution of the RNA-binding domain of NS1 in the pathogenicity and replication potential of an avian H7N1 influenza virus in chickens

Trapp *et al. Virology Journal* (2018) 15:55
https://doi.org/10.1186/s12985-018-0960-4

Virology Journal

RESEARCH Open Access

Major contribution of the RNA-binding domain of NS1 in the pathogenicity and replication potential of an avian H7N1 influenza virus in chickens

Sascha Trapp[1,2], Denis Soubieux[1,2], Alexandra Lidove[1,2], Evelyne Esnault[1,2], Adrien Lion[1,2], Vanaique Guillory[1,2], Alan Wacquiez[1,2], Emmanuel Kut[1,2], Pascale Quéré[1,2], Thibaut Larcher[3,4], Mireille Ledevin[3,4], Virginie Nadan[5], Christelle Camus-Bouclainville[6] and Daniel Marc[1,2*]

Abstract

Background: Non-structural protein NS1 of influenza A viruses harbours several determinants of pathogenicity and host-range. However it is still unclear to what extent each of its two structured domains (i.e. RNA-binding domain, RBD, and effector domain, ED) contribute to its various activities.

Methods: To evaluate the respective contributions of the two domains, we genetically engineered two variants of an H7N1 low pathogenicity avian influenza virus harbouring amino-acid substitutions that impair the functionality of either domain. The RBD- and ED-mutant viruses were compared to their wt- counterpart in vivo and in vitro, notably in chicken infection and avian cell culture models.

Results: The double substitution R38A-K41A in the RBD dramatically reduced the pathogenicity and replication potential of the virus, whereas the substitution A149V that was considered to abrogate the IFN-antagonistic activity of the effector domain entailed much less effects. While all three viruses initiated the viral life cycle in avian cells, replication of the R38A-K41A virus was severely impaired. This defect was associated with a delayed synthesis of nucleoprotein NP and a reduced accumulation of NS1, which was found to reach a concentration of about 30 micromol.L^{-1} in wt-infected cells at 8 h post-infection. When overexpressed in avian lung epithelial cells, both the wt-NS1 and 3841AA-NS1, but not the A149V-NS1, reduced the poly(I:C)-induced activation of the IFN-sensitive chicken Mx promoter. Unexpectedly, the R38A-K41A substitution in the recombinant RBD did not alter its in vitro affinity for a model dsRNA. When overexpressed in avian cells, both the wt- and A149V-NS1s, as well as the individually expressed wt-RBD to a lesser extent, enhanced the activity of the reconstituted viral RNA-polymerase in a minireplicon assay.

Conclusions: Collectively, our data emphasized the critical importance and essential role of the RNA-binding domain in essential steps of the virus replication cycle, notably expression and translation of viral mRNAs.

Keywords: Influenza A, NS1, Viral replication, Chicken

* Correspondence: daniel.marc@inra.fr
[1]Equipe PIA, UMR1282-ISP Infectiologie et Santé Publique, INRA, 37380 Nouzilly, France
[2]UMR1282 Infectiologie et Santé Publique, Université de Tours, F-37000 Tours, France
Full list of author information is available at the end of the article

BioMed Central

© The Author(s). 2018 **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.

Background

Influenza A viruses (IAVs) pose a permanent threat to human and animal health, as illustrated by the 2009 H1N1 pandemic and the frequent occurrence of epizootics with highly pathogenic avian influenza viruses (e.g. H5N1 and H5N8) across the world. Their genome consists of eight single stranded RNAs of negative polarity, which collectively code for up to fifteen proteins [1–4]. Most of these genes harbour determinants that modulate the fitness, replication potential, pathogenicity and the host range of the virus, in birds and mammals [5–7].

The non-structural protein NS1 of influenza viruses is considered as the major interferon antagonist of the virus. This ~ 230-residue protein is highly expressed in the infected cell, where it exhibits several pro-viral activities [8, 9]. Its homodimeric RNA-binding domain (RBD, amino-acids 1-73) interacts with several RNAs, including viral mRNAs [10–14], while its Effector Domain (ED, amino-acids 80-202) is a non-obligate dimer that interacts with several cellular proteins. A linker region connecting the two structured domains [15, 16] confers plasticity to the quaternary structure, while the C-terminal tail (CTT, amino-acids 202-230) has the properties of an intrinsically disordered region [16].

Several studies have shown that substitutions of critical amino-acids in NS1, as well as various truncations, result in viruses with dramatically reduced pathogenicity. These notably include alanine replacement of residues within the RNA-binding domain that directly interact with RNA, and C-terminal truncations that totally or partially remove the effector domain [17–20]. In addition to those rationally designed mutations that were introduced into genetically-engineered viruses, some substitutions were demonstrated to cause the distinctive attenuated phenotype of some field-isolated avian influenza viruses [21–23]. Although most of these attenuated viruses have been characterized in depth in vivo and in vitro, to our knowledge the respective contributions of the RNA-binding- and Effector-domains to the activities of NS1 and to the pathogenicity of the virus have not been characterized.

Our objective in the present study was to elucidate the contribution of each of NS1's two domains (i.e. RNA-binding domain and Effector domain) to its activities and to the viral phenotype. To this end, we rescued a virus harbouring a well-characterized double-substitutions in its RBD (R38A-K41A) that is known to invalidate the function of the RBD, along with a virus harbouring substitution A149V in NS1's ED that was previously shown to abrogate the inhibitory activity of NS1 towards type I-IFN production in avian cells [22, 24]. The phenotypes of the mutant viruses were compared to that of the parental strain, in vivo after challenge of chickens, and in vitro using avian cell culture and chicken egg infection models. Our results showed that the double substitution in the RBD dramatically reduced the pathogenicity and replication potential of the virus, while the A149V substitution in the ED had only a modest effect.

Methods

Viruses

The low-pathogenicity avian influenza (LPAI) virus A/Turkey/Italy/977/1999 (H7N1) [25] was a kind gift of Ilaria Capua (Istituto Zooprofilattico Sperimentale Delle Venezie, Legnaro, Italy). Site-directed mutagenesis was performed on the NS segment using the QuikChange II kit (Agilent Technologies) according to the manufacturer's protocol. The double substitution R38A-K41A was introduced using a primer pair harbouring two modified codons (codon 38 AGA > GCC and codon 41 AAG > GCG), while substitution A149V was introduced by changing codon GCA > GTG. The three recombinant IAVs (wt, 3841AA and A149V) were rescued by reverse genetics as described previously [26]. Virus stocks were prepared by one round of amplification in 10 day-old embryonated chicken eggs at a multiplicity of infection (m.o.i.) of 2 plaque-forming units (PFU) per egg. The identity of the amplified viruses was verified by sequencing their NS gene segment using reverse transcription (RT)-PCR.

Chicken experiments

All animals were kept and treated in strict compliance with the Good Animal Practice rules as defined by the relevant national and local animal welfare authorities, and all animal experimental interventions were approved by the local ethics committee (approval number 2010/7, Comité d'Ethique pour l'Expérimentation Animale, Région Centre-Val de Loire).

Four-week-old specific-pathogen-free (SPF) histocompatible B13/B13 White Leghorn chickens were housed in biosafety level 3 cabinets under negative pressure with HEPA-filtered air. Briefly, four groups of 24 (wt), 22 (R38A-K41A), 23 (A149V) and 10 (mock) birds were inoculated intra-tracheally with 5×10^5 EID_{50} (0.1 ml) of each virus, along with 10^6 EID_{50} (0.2 ml) in the choanal cleft, while in the mock group the virus suspension was replaced by PBS. Birds were carefully monitored daily during the course of the trial (one week). Four birds at days 1, 2, 3 and 4 post-inoculation (p.i.) and 1-3 birds at day 7 p.i. for each virus-inoculated group were euthanized and necropsied, along with 2 mock-inoculated birds at each of the indicated times. Swabs and tissue samples from lung, kidney, and brain were collected from each animal and either frozen at − 80 °C until use for further downstream analyses (viral RNA quantification) or fixed in 10% neutral buffered formalin for subsequent histopathological evaluation.

Quantification of viral RNAs

Tissue samples (50 mg) were transferred in 1 ml of PBS, and dissociated mechanically using a tissue grinder (Retsch,

Germany). The QiaAmp viral RNA mini kit (Qiagen) was used to prepare viral RNAs (vRNAs) from 140 μl of tissue homogenates, according to the manufacturer's recommendations. M-segment vRNAs were quantified through real-time RT-PCR, using the Superscript III Platinum SYBR Green one-step quantitative RT-PCR kit (Invitrogen) and a Chromo 4 thermocycler (Bio-Rad). The sequences of the primer pairs targeting the M-segment as well as the qRT-PCR conditions were previously described [27]. qRT-PCR values below the detection threshold, i.e., 200 copies, were arbitrarily replaced by 1.

Histopathological evaluation

Formalin-fixed tissues were embedded in paraffin wax and 6 μm-thick sections were cut and routinely stained with the hematoxylin-eosin-saffron procedure. Histologic observation was then performed and each lesion recorded. For lung tissue, semi-quantitative assessment of broncho-interstitial lesions was performed and the corresponding individual score attributed as follows: 0 = no lesion; 1 = focal presence of bronchial lesions associated with moderate afflux of mononuclear cells in the surrounding parenchyma; 2 = multifocal presence of bronchial lesions with marked peripheral afflux of inflammatory cells occupying 10 to 30% of the lung parenchyma; 3 = multifocal bronchial lesions with severe infiltration of more than 30% of the parenchyma by inflammatory cells.

Cells, infections and virus titration

MDCK and chicken lung epithelial CLEC213 cells [28] were grown in Eagle's minimal essential medium (EMEM) and Dulbecco's modified Eagle's medium (DMEM-F12), respectively, supplemented with 7.5% fetal calf serum. All infections were performed with DMEM supplemented with 0.2% bovine serum albumin. TPCK-treated trypsin (Worthington, 0.4 μg/ml) was added in the case of multiple-cycle growth assays.

Multicycle growth kinetics

Subconfluent monolayers of MDCK or CLEC213 cells were prepared in 75cm^2-flasks. The cells were virus-infected at an m.o.i. of 0.001 PFU/cell. Following 1 h of adsorption at 37 °C, the cells were further incubated in serum-free DMEM containing 0.4 μg/ml of TPCK-treated trypsin. Supernatant samples were harvested at the indicated times post-infection and subsequently titrated by plaque assays on MDCK cells as described previously [29].

Immunodetection of NS1

Subconfluent monolayers of CLEC213 cells in 24-well plates (grown on glass coverslips for immunofluorescence) were virus-infected at an m.o.i. of 0.5 PFU/cell. At the indicated times, cells were washed twice with PBS, fixed for 15 min in 4% paraformaldehyde and permeabilized with PBS 0.5% TritonX100. NS1 was revealed with a NS1-specific rabbit antiserum, using Alexa-conjugated secondary goat anti-rabbit antibody (1/3000 dilution). Fluorescence imaging was done using an AxioVision (Zeiss) Digital Image System. Alternatively, cells were lysed in 100 μl of Laemmli buffer and an aliquot was used for electrophoretic separation and immunoblot detection using the NS1-specific antiserum and ECL-system detection (Advansta, USA). Monoclonal antibody 9G8 (Santa Cruz Biotechnology) was used to detect NP.

Plasmids, transfections and luciferase reporter assays

A eukaryotic expression vector encoding wt-NS1 (pCIwt-NS1) was constructed by sub-cloning the coding sequence between the *Xho*I and *Not*I sites of the pCI plasmid (Promega). In order to prevent production of spliced mRNAs, the splice-donor and splice-acceptor sites were both invalidated by point mutations [30]. Substitutions within NS1, as well as N-terminal and C-terminal truncations that were designed to express the RBD and ED were introduced using the QuikChange II kit.

For the chMX-Luc reporter assay, subconfluent monolayers of CLEC213 cells in 24-well plates were co-transfected, using the Lipofectamine 2000 transfection reagent (Life Technologies), with (i) 100 ng/well of pCI (empty vector control) or pCI-NS1 harboring the different variants of NS1, together with (ii) 100 ng/well of pGL3-chMx-Luc encoding the Firefly luciferase under the control of the chicken Mx promoter [31] (kindly provided by Dr. Nicolas Ruggli, IVI, Mittelhäusern, Switzerland), and (iii) 10 ng/well of pRL-TK (Promega) encoding the Renilla luciferase as a transfection control. Twenty-four hours post-transfection, cells were either treated with 1 μg/ml of the RIG-like receptor agonist LyoVec-poly(I:C) (L-poly(I:C)) (Invivogen), or left untreated ($n = 3$ wells for each condition). Twenty-four hours after L-poly(I:C) treatment, cells were lysed and activation of the chicken Mx-promoter was revealed by using the Dual-luciferase reporter assay system (Promega) and a GloMax-Multi microplate luminometer (Promega). For each L-poly(I:C)-treated well, the calculated foldchange is the ratio of the normalized Mx-Luc signal to the geometric mean of the normalized signals in the three untreated wells.

For the viral polymerase assays, we used the pHW-derived plasmids for the PA, PB1, PB2 and NP segments of the virus A/tk/Italy/977/1999 (H7N1) that allow the CMV promoter-driven expression of intronless mRNAs. Plasmid pgHH21-vNA-Luc [32] directs, under the control of a chicken RNA PolI promoter, the expression of a viral-like RNA containing the Firefly Luciferase ORF flanked by the extremities of the NA segment of influenza virus A/WSN/33. Subconfluent CLEC213 cells in 24-well plates (one technical triplicate for each condition) were co-transfected using the Fugene reagent (Promega) with

(i) 150 ng/well of each of the pHW-PA, pHW-PB1, pHW-PB2 plasmids, (ii) 200 ng/well of the pHW-NP, pCI-NS harboring the different variants of NS1 and pgHH21-vNA-Luc plasmids, together with (iii) 5 ng/well of pRL-CMV (transfection control) encoding the Renilla luciferase under the control of the CMV promoter.

Gel-shift assay

The RBD (aa 1-73) of A/turkey/Italy/977/1999 (H7N1), after PCR-amplification with appropriate primers, was ligated to the *Nde*I-*BamH*I digested pET15b(+) DNA (Novagen). The R38A-K41A double substitution was introduced using the QuikChange kit. Proteins were overexpressed in Rosetta2-(DE3) bacteria and purified through Nickel affinity (Histrap HP, GE Healthcare) followed by ion exchange chromatography (POROS® HS, Applied Biosystems™). After thrombin cleavage, the proteins were further separated by gel filtration (Superdex S75, GE Healthcare). Synthetic RNA shDM03 (Eurogentec) that was 5'end-labeled with ^{32}P was heated (5 min. at 85 °C), then snap-cooled and incubated 30 min. on ice with increasing concentrations of protein in a 10 μl-reaction mixture (10 mM Tris-Cl, 1 mM EDTA, 0.1% bovine serum albumin, 10% glycerol, 50 mM NaCl, pH 8). RNA and RNA-protein complexes were separated as described [12]. Gels were then dried and revealed using a Typhoon-Trio laser scanner.

Statistical analyses

All statistical analyses were performed using Graphpad Prism 6.05. Kruskall-Walis non parametric tests and Dunn's post-tests were used to analyze in vivo data (Figs. 1 and 2). For the in vitro assays in transfected cells, groups of values (Figs. 5c and 7) were compared through a one-way Anova and subsequently compared pairwise with the wt- or the vector-control group by a Dunnett's test.

Results

Viral replication and pathogenicity in inoculated chickens

The Low-Pathogenicity avian influenza (LPAI) virus A/Turkey/Italy/977/1999 (H7N1), a direct precursor of High Pathogenicity viruses [25, 33], expresses a NS1 protein of allele B [34, 35]. We genetically engineered

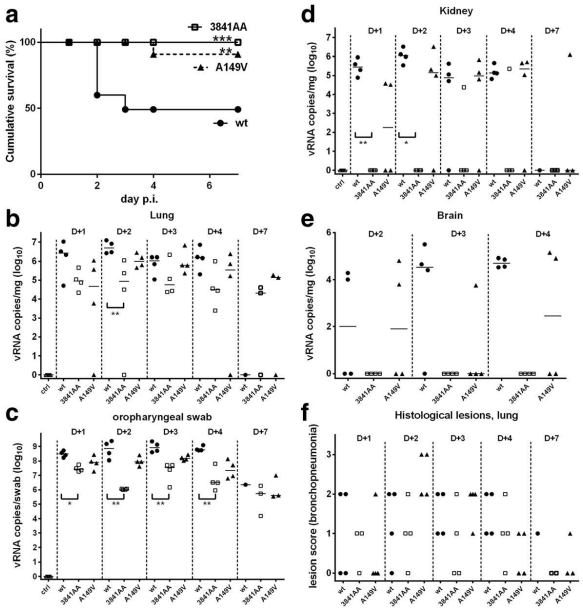


Fig. 1 Virus pathogenicity. Four-week old SPF chickens were inoculated with the three wild-type (wt) and mutant viruses. (**a**) Deaths were recorded daily, and a survival curve was established. *** $P < 0.001$, ** $P < 0.01$ in a Mantel-Cox log-rank test comparing the considered mutant with the wt-virus. (**b-e**) At the indicated times (D + 1 to D + 7 p.i.), M-segment viral RNA was quantified in the indicated tissues or swabs, and expressed as log10 of vRNA copies per mg of tissue or per swab. Wt, filled circles; 3841AA, open squares; A149V, black triangles. Null qRT-PCR values were arbitrarily assigned a + 1 value (10^0). ** $P < 0.01$, * $P < 0.05$ in pairwise Dunn's post-tests comparing the considered mutant with the wt-virus in panels b, c and d. (**f**) Histological lesions of interstitial broncho-pneumonia in the lungs of the virus-infected birds were scored

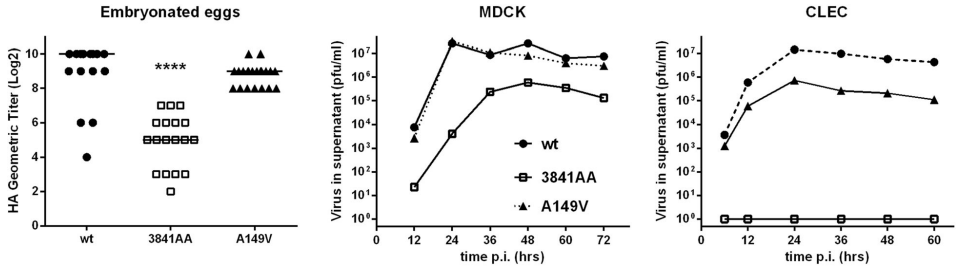


Fig. 2 Growth properties of the three viruses. (*Left panel*) The three indicated viruses were inoculated in embryonated chicken eggs to determine virus titres (EID_{50}) and the HA titres in the allantoic fluids of individual eggs were recorded. **** $P < 0.0001$ in pairwise Dunn's post-tests comparing the considered mutant with the wt-virus. (*Middle and right panel*) The three indicated viruses were subjected to a multicycle growth assay after infection (m.o.i. = 0.001 PFU/cell) of MDCK and avian CLEC213 cells. Supernatants that were collected at the indicated times p.i. were titrated in MDCK

three variants of this virus: the recombinant wild-type (wt) virus harboured a segment 8 identical to that of the parental isolate, the mutant virus dubbed 3841AA harboured a modified segment 8 encoding an RBD-mutated NS1 with the double substitution R38A-K41A [17, 21]; and the virus named A149V encoded an NS1 protein with a substitution in the ED that was shown to critically alter the IFN-antagonist activity of NS1 [22, 24]. We then compared the replication and pathogenic properties of the three viruses in four-week-old SPF histocompatible B13/B13 White Leghorn chickens that were inoculated intra-tracheally with 5×10^5 EID_{50} of each virus, along with 10^6 EID_{50} in the choanal cleft. Infection with the wt virus caused severe respiratory symptoms and was associated with significant mortality: of the 24 wt-inoculated animals, eight died at day 2 and two at day 3. The clinical symptoms were much less severe in the A149V group, of which only one animal died at day 4. In the 3841AA group, no clinical symptoms were observed and no deaths were recorded (Fig. 1a).

At days 1, 2, 3, 4, and 7 p.i., M-segment viral RNAs (M-vRNAs) in the tissue samples were quantified by RT-qPCR in order to evaluate viral replication and shedding [36]. All mock-inoculated chickens were negative for the presence of viral RNAs. High loads of M-vRNAs were recorded in the lungs of wt-inoculated chickens from d1 to d4 p.i. (Fig. 1b). The highest loads, close to 10^7 copies of vRNA per mg of lung tissue, were recorded at d1 and d2. High loads in the lungs were paralleled by high amounts of M-vRNAs excreted in the upper respiratory tract (up to 10^9 copies of M-vRNA per oro-pharyngeal swab at d2 and d3, Fig. 1b, c). The kinetics of viral replication in the lung was slightly delayed with the A149V virus, for which the highest loads were observed at days 2 and 3 p.i. At day 2 p.i., M-vRNA loads in the lungs of A149V-inoculated chickens were about tenfold lower than those of wt-inoculated chickens. In addition, virus excretion in the upper respiratory tract was reduced by about tenfold in A149V-inoculated chickens, when compared to wt-inoculated animals. M-vRNA loads in the lungs and upper respiratory tract of 3841AA-inoculated animals were dramatically reduced, by ten to 100-fold relative to wt-inoculated birds. The wt-virus, and to a slightly lesser extent the A149V-virus, were found in the kidneys and brains, indicating limited systemic dissemination (Fig. 1d, e), while in the 3841AA-group, only two kidneys, and no brain tissues, were found vRNA-positive.

Lung lesions were observed only in virus-inoculated chickens. Macroscopic lesions were mainly observed at d2 and d3 in wt- and A149V-inoculated chickens, and were not recorded in 3841AA-inoculated birds. Histological examination of the lungs of wt- and A149V-inoculated chickens revealed severe lesions of peribronchiolar interstitial pneumonia, associated with intraluminal presence of a fibrinous exudate and some degenerating heterophils, typical of exudative bronchitis and parabronchitis (Fig. 1f). This exudate obstructed airways in severe cases. No lesions were observed in the other examined tissues. In spite of the symptomless infection, mild histological lesions were observed in 3841AA-inoculated animals, consisting essentially of foci of infiltrating mononuclear inflammatory cells scattered in the lung parenchyma around airways.

NS1 mutants replicate to lower levels, both in embryonated eggs and in cultured cells

We reasoned that the distinct replication efficiencies and pathogenicities of the three viruses could be paralleled by similar properties in embryonated chicken eggs or in infected cells. In embryonated eggs, the A149V virus grew to HA titers that were slightly lower than those of the wt virus, while the 3841AA-mutant grew to dramatically reduced levels, about 5 $\log_2$ units lower than those of the wt virus (Fig. 2).

In multicycle growth assays in MDCK cells, the wt- and A149V-viruses showed virtually identical replication kinetics reaching 10^7 PFU/ml in the supernatant as early as 24 h post infection, while the growth of the 3841AA-virus was clearly delayed, with a maximum titer that remained at least tenfold lower than those of the two other viruses. In CLEC213 cells, a chicken lung epithelial cell line that we have previously described [28], the growth of the A149V virus was reduced by about tenfold as compared to its wt-counterpart, while no virus was found in the supernatant of the 3841AA-inoculated cells, in spite of repeated attempts.

Mutations alter the accumulation and subcellular distribution of NS1 in infected cells

To better understand by which mechanisms substitutions in NS1 might hamper the viral replication cycle, we first measured the accumulation of NS1 in cells infected with the three viruses. Avian CLEC213 cells were infected with the three viruses at an m.o.i. of 0.5 PFU per cell and the accumulation of NS1 in the infected cells during the course of the viral cycle was monitored by a Western-blot assay (Fig. 3). On the same gel we loaded increasing calibrated amounts of recombinant NS1 from the same virus, thus allowing us to estimate the amount of NS1 in the lysates and its cellular concentration, assuming that confluent cells in the well of a 24-well plate represent a total intracellular volume of about one microliter. From several independent experiments, we estimated that, as early as 2 h p.i., NS1 reaches a concentration of 2 to 5 micromol.L^{-1} in wt-infected cells. At later time-points, NS1 concentration continues to increase, up to about 30 micromol.L^{-1}, a surprisingly high concentration that is of the same order of magnitude as those of the most abundant cellular proteins, such as actin [37, 38]. As shown in Fig. 3, NS1 expression was markedly reduced (by about 4 to 8-fold) in 3841AA-infected cells and slightly reduced in A149V-infected cells, compared to the levels observed in wt-infected cells. Of note, if one assumes that NS1 is homogeneously distributed in the infected cell, it is expected to passively diffuse into the budding virion: assuming a cytoplasmic NS1 concentration of 30 micromol.L^{-1}, a virion of 100 nm in diameter ($\sim 5 \times 10^{-19}$ l in volume) could contain about nine molecules of NS1, which is consistent with the published estimates [39]. Using the same infected cell-lysates, we observed that the accumulation of the viral nucleoprotein NP was considerably delayed in 3841AA-infected cells, where it was detected only at 12 h p.i., while it was detected as early as 4 h p.i. in wt-infected cells (Fig. 3, lower panel).

The distinct phenotypes of the three viruses likely resulted from several distinct biological properties of the NS1 variants, including their subcellular distribution which we analyzed in infected cells. Avian CLEC213 cells were infected with the three viruses at an m.o.i. of 0.5 PFU/cell, then fixed and immunostained with a polyclonal anti-NS1 rabbit serum at 12 and 24 h p.i. (Fig. 4). The three NS1 variants exhibited distinct patterns of subcellular distribution. The wt-NS1 formed crystal-like cytoplasmic inclusions, similar to those previously described for some avian influenza viruses [40]. In stark contrast, 3841AA showed an intense staining that filled the whole cytoplasm. A149V produced cytoplasmic aggregates that did not exhibit the crystal-like structures observed with the wt-NS1. All three viruses produced a cytopathic effect, as judged by the low number of adherent cells at 24 h p.i.. Furthermore, at 12 h p.i. we observed that most cells infected with the wt and A149V viruses exhibited a disfigured shape of nuclei, a feature that was much less pronounced with the 3841AA-virus.

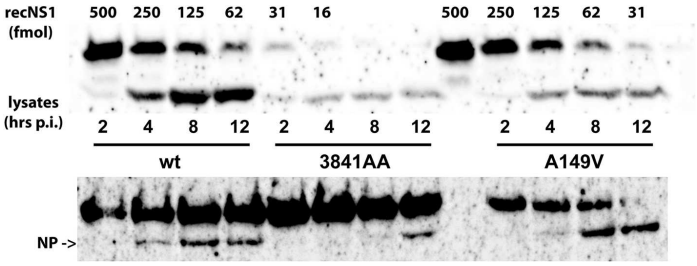


Fig. 3 Accumulation of NS1 in virus-infected CLEC213 cells. Subconfluent CLEC213 cells in a 24-well plate were infected (0.5 PFU/cell) and lysed at the indicated times post-infection. Proteins in the lysates (0.8% of the 100 µl-lysate) were separated through SDS-PAGE, then transferred to a nylon membrane. For NS1 quantification, increasing amounts of recombinant NS1 protein (16 to 500 femtomoles) were loaded on the same gel 20 min after the lysates. Except for a C-terminal $(His)_6$ extension, the recNS1 is identical to the viral NS1; its apparently slower migration profile only results from the delayed loading. NS1 was revealed through immunoblot, using a polyclonal rabbit serum. Lower panel: the same lysates were used for the western-blot detection of NP (the upper band corresponds to an unidentified cellular protein that was also detected in non-infected cells)

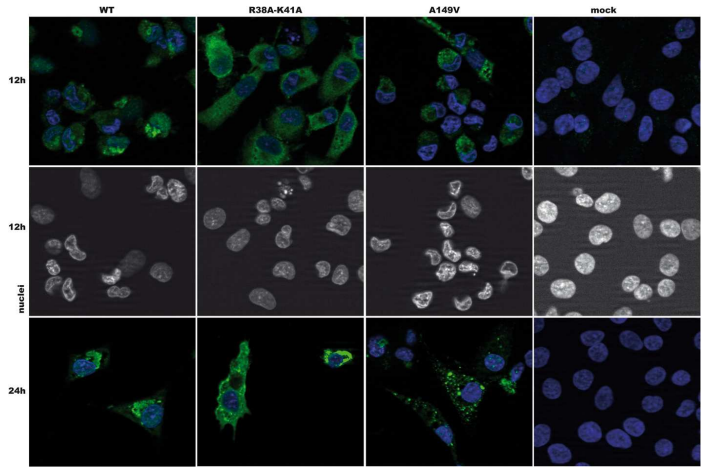


Fig. 4 Subcellular distribution of NS1 in virus-infected CLEC213 cells. CLEC213 cells were mock-infected or infected with the indicated viruses (0.5 PFU/cell), then fixed and immunostained with an NS1-specific rabbit antiserum at 12 h or 24 h p.i. Images in the top and middle panels are from the same selected fields (middle panel shows the DAPI-stained nuclei at 12 h p.i.)

The three NS1 variants moderately impact the type I IFN response in avian CLEC213 cells

We then sought to determine whether the variants of NS1 might differentially regulate the antiviral interferon response. We surmised that the antiviral response in virus-infected cells could be biased by the differing replication efficiencies of the wild-type and NS-mutant viruses. Hence we opted to measure the LyoVec-poly(I:C)-induced antiviral response in cells that transiently expressed the three NS1 variants [31]. CLEC213 cells were transfected with plasmid pGL3-chMx-Luc encoding the firefly luciferase under the control of the IFN-inducible chicken Mx promoter, along with a "normalizer" Renilla Luciferase-encoding plasmid as well as expression plasmids encoding wt- or substituted NS1s, or the empty plasmid pCI as a control. The three variants of NS1 were expressed to similar levels, as verified by a western-blot assay using a rabbit anti-NS1 polyclonal serum (Fig. 5d). Transfected cells were stimulated by treatment with the MDA5-agonist Lyovec-poly(I:C) and the onset of an IFN-I response was measured 24 h after treatment by a dual-luciferase assay. Compared to the corresponding untreated cells, the L-poly(I:C) treatment of empty vector-transfected cells induced a strong activation of the Mx promoter, that varied from 10 to 28-fold depending on the experiment. Therefore we arbitrarily set a 100% value for the fold-change in empty pCI-transfected cells in order to combine all data from five independent experiments (Fig. 5c). Transient expression of wt-NS1 diminished the L-poly(I:C)-induced Mx-promoter activation, while A149V-NS1 slightly increased the activation. Unexpectedly, the 3841AA mutant reduced the activation by about 50% relative to empty-vector transfected cells. Interestingly, in these experiments we consistently observed a dramatic decrease in the activity of the two types of luciferase in cells expressing NS1 with an intact effector domain (Fig. 5a, b): in wt-NS1-transfected cells the activities of both Firefly and Renilla luciferases were 6-8% of those recorded in empty vector-transfected cells, while in cells transfected with the 3841AA variant they were further reduced to 2-3%. Substitution A149V in the effector domain completely abolished this effect.

Mutated RBD still binds double-stranded RNAs

It is generally considered that the attenuated phenotype of viruses harboring the double substitution R38A-K41A results from the fact that the mutated RBD has lost its ability to bind double-stranded RNAs [17, 41]. In order to confirm this issue, we compared the in vitro RNA-binding properties of the wt- and mutated RBDs. Both RBDs were purified from overexpressing bacteria, and used in a gel-shift experiment using a model double-stranded RNA that we had designed previously [12]. As shown in Fig. 6, the wt-RBD strongly bound the model dsRNA with an apparent K_D close to 1 nM. Unexpectedly, the double substitution did not substantially alter the affinity of the RBD for this model dsRNA, as shown by the virtually unchanged K_D. However, the distinct migration profiles observed with the two variant RBDs suggested that the RBD-RNA complexes were structurally different.

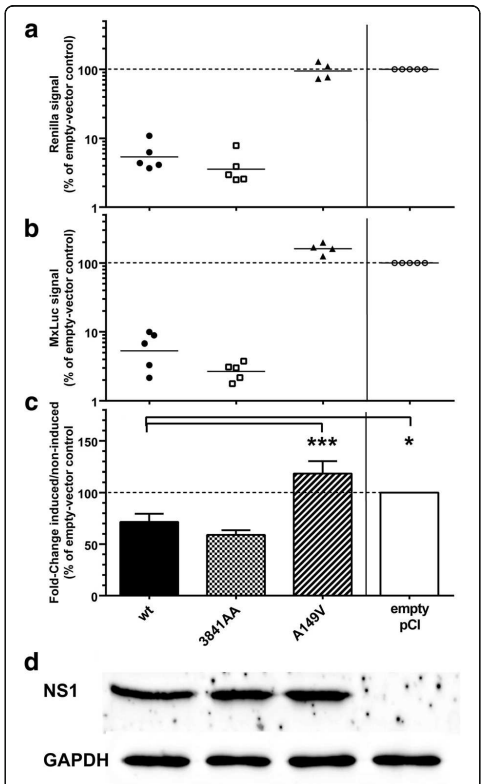


Fig. 5 NS1 variants differentially modulate type I IFN response. CLEC213 cells in 24-well plates were transfected with the two plasmids used for the dual-luciferase assay for IFN-induced chicken Mx promoter activity, along with recombinant expression vectors encoding the indicated NS1-variants. Cells were stimulated at 24 h post-transfection by a L-poly(IC) treatment, and subsequently lysed 24 h later (48 h post-transfection). The chMx promoter-driven Firefly activity (FFL) was measured in the lysates (**b**) and normalized relative to the Tk-promoter driven Renilla-luciferase activity (**a**). (**c**) Fold-change ratios (induced/non-induced) are expressed relative to that calculated from the respective empty-vector control value set at 100%; * $P < 0.05$; *** $P < 0.001$ in a Dunnett's test comparing each set to the wt-NS1 condition. (**d**) Western-blot detection of NS1 in transfected cells. Proteins in the lysates (1/8th of the 100 μl-lysate) were separated through SDS-PAGE, then transferred to a nylon membrane and revealed as in Fig. 3. Data are from five independent experiments, each point in (**a**) and (**b**) representing the geometric mean of a technical triplicate

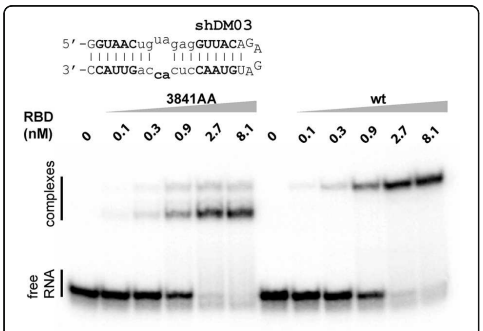


Fig. 6 In vitro interaction of wt- and mutated RBDs with model dsRNA (electrophoretic mobility shift assay). ^{32}P-labeled, synthetic RNA shDM03 (depicted above as MFold-predicted structure) at a final concentration of 0.1 nMol.L^{-1} was incubated with increasing concentrations (as indicated) of the dimeric wt-RBD (right panel) or mutated-RBD (left panel). RNA-protein complexes were electrophoretically separated in a 10% non-denaturing polyacrylamide gel, and were subsequently revealed using a Typhoon-Trio laser scanner

NS1 modulates the activity of the viral polymerase

Both the in vitro and in vivo data pointed to a dramatically reduced replicative efficiency of the 3841AA-virus, while the replication fitness of the A149V virus appeared to be moderately affected. This suggested that NS1, and notably its RBD, might modulate the activity of the viral polymerase complex. To address this issue we set up a minireplicon assay: CLEC213 cells were transfected with the four pHW2000-derived expression vectors allowing the reconstitution of the viral polymerase complex (PA, PB1, PB2, NP), along with a Firefly-luciferase-encoding minireplicon and NS1 expression vector. The Firefly plasmid reports the expression of a viral RNA mimic in which the Luciferase ORF is flanked by the extremities of the NA segment of influenza virus A/WSN/33. Transcription of this vRNA mimic by the viral polymerase produces an artificial viral mRNA encoding the Firefly-luciferase, whose activity was measured by a dual luciferase assay. We observed that the expression of the distinct variants of NS1 (truncations or substitutions) increased the apparent polymerase activity to various degrees (Fig. 7): (i) the wt-ED induced a ~twofold increase of the apparent polymerase activity, while the wt-RBD had a lesser impact (~ 38% increase); (ii) substitutions in either individually expressed domain (3841AA-RBD or A149V-ED) reduced or nearly abolished the positive effect observed with the wt-domain; (iii) the full-length NS1 induced a twofold increase of polymerase activity; (iv) the double substitution in the RBD, but not the A149V substitution in the ED, reduced the effect of the full-length protein. The latter observation could result from the fact that the effect of the A149V substitution was compensated for by the presence of the wt-RBD. Similarly to what we observed in the Mx-Luc assay, all three conditions where cells expressed a wt-ED (wt and 3841AA full-length NS1, wt-ED) resulted in a

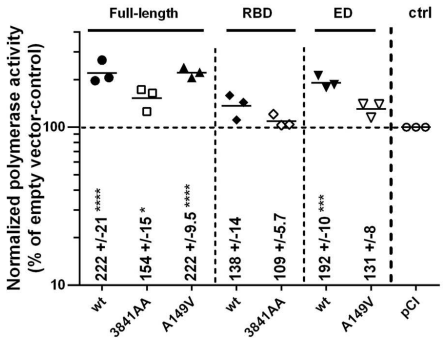


Fig. 7 Complementation of the viral polymerase activity by NS1. CLEC213 cells in 24-well plates were transfected with the plasmid set of the pHW-based H7-minireplicon system (see Methods section), along with an NS1-expression vector or an empty vector control. Twenty-four hours later, cells were lyzed and the Firefly-luciferase activity of the lysates, resulting from translation of the polymerase-transcribed mRNA, was measured and normalized relative to the Renilla-luciferase through a dual luciferase assay. Each point represents the normalized polymerase activity (geometric mean of a technical triplicate) relative to the empty-vector control in a given experiment (numbers are the mean and s.e.m from three independent experiments). The "no-PB1" control values (not shown) were less than 0.1%. * $P < 0.05$; *** $P < 0.001$; **** $P < 0.0001$ in a Dunnett's test comparing each set to the control (empty pCI) condition

~twofold reduction in the activity of the normalizer Renilla Luciferase. This probably resulted from a negative effect of the wt-ED on the expression of cellular mRNAs, and could also partly bias our data and their interpretation (as discussed below).

Discussion

In an attempt to delineate the respective roles of the two structured domains of NS1, we genetically engineered three variants of the Low Pathogenicity Avian influenza virus A/tk/Italy/977/99(H7N1). On the one hand, we introduced in the RNA-binding domain the double substitution R38A-K41A, which is considered to abrogate its RNA-binding properties [17, 20, 41]; on the other hand, the ED was modified by substitution A149V, which was previously shown to alter the IFN-antagonist activity of NS1 in avian cells [22, 24]. In a well-established chicken model of avian influenza virus infection [42, 43], the wt-virus induced severe respiratory symptoms and deaths while the symptoms were much less severe in the A149V-inoculated birds and infection remained silent in the 3841AA-inoculated birds. The different pathological outcomes were reflected both by the M-vRNA loads in tissues and in the histopathological findings, while the distinct replication efficiencies of the three viruses in vivo were correlated to similar phenotypes in vitro: replication of the A149V mutant was slightly reduced relative to that of the wt-virus in both embryonated eggs and in the chicken lung epithelial cell line CLEC213, while that of the 3841AA-virus was severely reduced in eggs and non-measurable in the avian cell line.

In these cells, all three viruses initiated a replication cycle, as evidenced by the accumulation of viral protein NS1, and all three induced a cytopathic effect (Fig. 4). Nevertheless, the 3841AA virus seemed unable to produce infectious viral progeny in CLEC213 cells, as demonstrated by the multi-cycle growth assay. At least three mechanisms could explain such a defect: (i) a reduced replication rate, either inherent to the virus or resulting from an increased antiviral defense, (ii) a reduced synthesis of viral proteins, or (iii) a defect in the late steps of the viral cycle (assembly and release). Reduced IFN-antagonistic activities of NS1 may account for an increased antiviral response in cells infected with the two NS-mutant viruses. Indeed we observed that the transiently-expressed wt-NS1 reduced the L-poly(I:C)-induced antiviral response. Substitution A149V in the ED abolished this effect, in agreement with the critical role of the effector domain in dampening the IFN response [22, 24, 44]. However, the R38A-K41A substitution did not abrogate this effect; on the contrary, it further reduced the activity of the IFN-responsive Mx promoter. This observation contrasts with the findings from previous studies with mutant influenza A viruses harbouring the same amino-acid substitutions in NS1's RBD. Specifically, transient expression of wt-NS1 from the WSN virus was shown to prevent the Sendai virus-induced IFN-beta response in 293 T cells, while WSN-NS1 harbouring the double substitution R38A-K41A was unable to reduce this IFN-beta response [17]. The discrepancy between our data and the literature could originate from at least four distinct experimental parameters: (i) NS1 allele and proteotype; (ii) the cell-type and their species origin; (iii) the interferon-induction treatment; and (iv) the assay used to evaluate the interferon response. Firstly, distinct proteotypes of NS1, both within and between its two alleles, can differ by their biological activities [34]; while the WSN virus harbors allele A of NS1, our study focused on allele B NS1 from the avian H7N1 virus. Secondly, while we investigated, in avian lung epithelial cells, the L-poly(I:C)-induced activation of the MDA5 pathway by measuring the activity of the IFN-induced chicken Mx gene-promoter, Donelan et al. [17] evaluated the IFN-response at a broader level by measuring, in Sendai virus-treated mammalian 293 T cells, the activity of the IFN-beta promoter. In addition, we speculate that the R38A-K41A NS1 may still bind the synthetic dsRNA of L-poly(I:C) and thereby prevent the activation of the RIG-like receptor pathway. On the one hand, this could be facilitated by the diffuse cytoplasmic distribution of this mutant NS1 that we observed in

infected cells. On the other hand, we observed that the double substitution R38A-K41A did not substantially alter the binding of the recombinant RBD to a model dsRNA. Furthermore, and contrary to the common belief, this double substitution likely does not totally invalidate the RBD. Indeed Schierhorn et al. recently showed that NS1's RBD was endowed with at least two distinct activities that each critically involved partially overlapping sets of amino acids. While R38, like R35 and R46, is critical for binding of the PR/8 virus-derived NS1 to poly-IC, R35 and R46, unlike R38, also have a critical role in binding to PKR and preventing its phosphorylation [45]. Our data and this discrepancy clearly highlight the need for a better understanding of the multiple pathways of the interferon/antiviral response in avian cells, including their modulation by NS1 as well as comparisons with the interferon response in mammalian cells.

Our data showed that the three NS1 variants differentially modulated the activity of the viral polymerase. Using a minireplicon assay in transfected cells, we showed that NS1 by itself increased by nearly twofold the activity of the transiently-expressed viral polymerase. The polymerase enhancing property of NS1 was reduced by the double substitution R38A-K41A, and virtually unaltered by the A149V substitution. Although each individually expressed domain also increased the apparent polymerase activity, we observed that the expression of a protein harbouring a wt-ED (wt-ED, wt-NS1 and 3841AA-NS1) consistently reduced to ~ 50% the activity of the normalizer Renilla Luciferase. As a result, the twofold increase of the normalized polymerase activity that we observed in these three cases could result in a large part from this bias. Therefore, if we disregard these potentially biased observations, our data indicated that (i) the individually-expressed wt-RBD, but not its 3841AA mutant, induced a ~ 38% increase in the polymerase activity; (ii) the individually-expressed A149V-ED induced a ~ 30% increase in the polymerase activity; (iii) when adjoined (i.e. in A149V-NS1), these two domains acted synergistically, resulting in a more than twofold increase in the polymerase activity. In short, the wt-RBD enhanced the activity of the viral polymerase, an effect that was further increased when it was linked to the ED. This is in line with several independent studies showing that NS1 interacts with the viral polymerase and can regulate its activity [46–48]. More specifically, this interaction was shown to rely on the binding of NS1's RBD to the viral nucleoprotein NP, a binding that was abolished by the double substitution R38A-K41A [49]. If NS1 directly enhances the activity of the viral polymerase through this interaction, it may also act indirectly through interacting with host factors such as DDX21 and hnRNP-U that were shown to modulate the polymerase activity [50, 51].

The inability of the 3841AA virus to produce infectious viral progeny in the avian cells likely also resulted from a reduced synthesis of viral proteins. Indeed, while NS1 reached a very high concentration in wt virus-infected cells, its concentration in 3841AA-infected cells was reduced by about 4 to 8-fold relative to that in wt-infected cells, and slightly reduced in A149V-infected cells. We also observed a severely reduced yield of the viral nucleoprotein NP in the 3841AA-infected cells (Fig. 3), which likely was indicative of a reduced or delayed synthesis of all viral proteins. Indeed NS1 was shown to specifically enhance the translation of viral mRNAs [13, 52] notably through binding of its N-terminal region to the strictly conserved sequence motif at the 5'end of viral mRNAs and increasing their translation-initiation rate [13, 53]. We hypothesize that the 3841AA NS1 no longer recognizes the viral-specific motif [12], and as a consequence has lost its translation-enhancing activity.

In addition to the defects discussed above, other mechanisms that we have not investigated at the late steps of the viral cycle may also account for the dramatically attenuated phenotype of the 3841AA mutant virus. First, it was recently shown that NS1, unlike its R38A-K41A mutant, binds both NXF1 and the late viral mRNAs (notably HA-mRNA and the unspliced viral M1 mRNA), thereby directing the latter to the nuclear export pathway [54]. The R38A-K41A NS1 mutant of the PR/8 virus was shown to exhibit a severe defect in the nuclear export of HA- and M1-mRNAs. Independently from its requirement in the expression of the late viral genes, NS1 could also be involved in some steps of the virion formation: an in-depth analysis of a NS1-mutant virus identified a block in the production of virus particles, pointing to a possible involvement of NS1 in viral assembly [55]. Incidentally, the 3841AA virus has the properties required for a live-attenuated vaccine candidate [56].

As for the ED, our data show that it is involved in two activities that are both abolished by the A149V substitution. Firstly, in agreement with previous reports [22, 24], we observed that wt-NS1, unlike its A149V variant, decreased the IFN-response of L-poly(I:C)-stimulated avian cells, as measured by the activity of the chMx-promoter. Secondly, the expression of either wt-ED or full-length NS1 with an intact ED consistently reduced the expression of the luciferase mRNA that we used to normalize our luciferase-based assays. Compared to that measured in empty-vector control-cells, the activity of the normalizer Renilla luciferase in all the conditions where cells expressed a wild-type ED was reduced to ~ 50% at 24 h post-transfection (minireplicon assay) and to a few percent at 48 h post-transfection (Mx-Luc assay). Substitution A149V consistently abolished this effect. This emphasizes the major role of the effector domain in inhibiting the maturation and nuclear export of non-viral mRNAs, relying on its interaction with Cleavage and Polyadenylation Specificity Factor [57, 58]. Given that alanine 149 is deeply

buried within the hydrophobic core of the ED [59], we surmise that its replacement with the bulky valine residue destabilizes the structure of the effector domain [22].

Conclusions

Our data clearly showed that the double substitution R38A-K41A severely impacted the viral phenotype, much more than the A149V substitution. This severe attenuation stems from at least three, and possibly four combined major defects: (i) a lower activity of the viral polymerase, which is no longer enhanced by the RBD-impaired NS1 (present data); (ii) a reduced translation of viral mRNAs (present data); (iii) a blocked nuclear export of late viral mRNAs [54]; (iv) possibly an impaired virion assembly [55]. The severely altered phenotype of the 3841AA-mutant, as compared to the slighter impact of the A149V substitution, emphasized the essential role of the RBD and the relative dispensability of the effector domain, which have been discussed previously [13] and are further supported by the unaltered pathogenicity of a modified bat-influenza virus expressing an RBD-only NS1 [60]. Altogether, our data and the published literature suggest that NS1 acts much more at the core of the viral cycle, through interactions of its RBD with virus-derived RNAs and proteins, than at its periphery through its ED being involved in the interactions with the host cell.

Acknowledgements
We thank Jennifer Doudna for the kind gift of plasmid pgHH21-vNA-Luc, Romain Volmer, Erich Hoffmann and Robert Webster for the gift of plasmid pHW2000 and its derivatives, and Nicolas Ruggli for providing the pGL3-chMx-Luc reporter plasmid. We also thank Antoine Marc for his help in the quantification of NS1 protein concentration.

Funding
This work was supported by grants from Institut Carnot Santé Animale (Influenza2011 and FlumLi) and from Région Centre-Val de Loire (RNApur).

Availability of data and materials
The data that support the findings of this study are available from the corresponding author upon request.

Authors' contributions
ST, DS, VG and DM did the animal experiments. EE, DS, AlxL, AL and EK performed the cell-based in vitro experiments. CCB did the reverse genetics to rescue the recombinant viruses. AW and VN purified the recombinant proteins and did the gelshift experiments. TL and ML did the histopathological analysis. ST, PQ and DM conceived the experiments, analyzed the data and wrote the paper. All authors read and approved the final manuscript.

Ethics approval and consent to participate
Animal experiments were approved by the local ethics committee (approval number 2010/7, Comité d'Ethique pour l'Expérimentation Animale, Région Centre-Val de Loire).

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
[1]Equipe PIA, UMR1282-ISP Infectiologie et Santé Publique, INRA, 37380 Nouzilly, France. [2]UMR1282 Infectiologie et Santé Publique, Université de Tours, F-37000 Tours, France. [3]INRA UMR 703, APEX, Oniris-La Chantrerie, F-44307 Nantes, France. [4]LUNAM Université, École Nationale Vétérinaire, agro-alimentaire et de l'alimentation Nantes-Atlantique (Oniris), Nantes, France. [5]Centre de Biophysique Moléculaire, CNRS, Orléans, France. [6]IHAP, UMR1225, Université de Toulouse, INRA, ENVT, Toulouse, France.

Received: 26 January 2018 Accepted: 11 March 2018
Published online: 27 March 2018

References

1. Jagger BW, Wise HM, Kash JC, Walters KA, Wills NM, Xiao YL, Dunfee RL, Schwartzman LM, Ozinsky A, Bell GL, et al. An overlapping protein-coding region in influenza a virus segment 3 modulates the host response. Science. 2012;337:199–204.
2. Palese P, Shaw M. Orthomyxoviridae: the viruses and their replication. In: Knipe D, Howley P, editors. Fields Virology. Philadelphia: Lippincott, Williams and Wilkins; 2007.
3. Wise HM, Foeglein A, Sun J, Dalton RM, Patel S, Howard W, Anderson EC, Barclay WS, Digard P. A complicated message: identification of a novel PB1-related protein translated from influenza a virus segment 2 mRNA. J Virol. 2009;83:8021–31.
4. Muramoto Y, Noda T, Kawakami E, Akkina R, Kawaoka Y. Identification of novel influenza a virus proteins translated from PA mRNA. J Virol. 2013;87:2455–62.
5. Perez-Cidoncha M, Killip MJ, Oliveros JC, Asensio VJ, Fernandez Y, Bengoechea JA, Randall RE, Ortin J. An unbiased genetic screen reveals the polygenic nature of the influenza virus anti-interferon response. J Virol. 2014;88:4632–46.
6. Maines TR, Chen LM, Belser JA, Van Hoeven N, Smith E, Donis RO, Tumpey TM, Katz JM. Multiple genes contribute to the virulent phenotype observed in ferrets of an H5N1 influenza virus isolated from Thailand in 2004. Virology. 2011;413:226–30.
7. Munoz O, De Nardi M, van der Meulen K, van Reeth K, Koopmans M, Harris K, von Dobschuetz S, Freidl G, Meijer A, Breed A, et al. Genetic adaptation of influenza a viruses in domestic animals and their potential role in interspecies transmission: a literature review. EcoHealth. 2016;13:171–98.
8. Hale BG, Randall RE, Ortin J, Jackson D. The multifunctional NS1 protein of influenza a viruses. J Gen Virol. 2008;89:2359–76.
9. Marc D. Influenza virus non-structural protein NS1: interferon antagonism and beyond. J Gen Virol. 2014;95:2594–611.
10. Hatada E, Saito S, Okishio N, Fukuda R. Binding of the influenza virus NS1 protein to model genome RNAs. J Gen Virol. 1997;78(Pt 5):1059–63.
11. Lu Y, Qian XY, Krug RM. The influenza virus NS1 protein: a novel inhibitor of pre-mRNA splicing. Genes Dev. 1994;8:1817–28.
12. Marc D, Barbachou S, Soubieux D. The RNA-binding domain of influenzavirus non-structural protein-1 cooperatively binds to virus-specific RNA sequences in a structure-dependent manner. Nucleic Acids Res. 2013;41:434–49.
13. Marion RM, Aragon T, Beloso A, Nieto A, Ortin J. The N-terminal half of the influenza virus NS1 protein is sufficient for nuclear retention of mRNA and enhancement of viral mRNA translation. Nucleic Acids Res. 1997;25:4271–7.
14. Qiu Y, Krug RM. The influenza virus NS1 protein is a poly(a)-binding protein that inhibits nuclear export of mRNAs containing poly(a). J Virol. 1994;68:2425–32.
15. Carrillo B, Choi JM, Bornholdt ZA, Sankaran B, Rice AP, Prasad BV. The influenza a virus protein NS1 displays structural polymorphism. J Virol. 2014;88:4113–22.
16. Hale BG. Conformational plasticity of the influenza a virus NS1 protein. J Gen Virol. 2014;95:2099–105.
17. Donelan NR, Basler CF, Garcia-Sastre A. A recombinant influenza a virus expressing an RNA-binding-defective NS1 protein induces high levels of beta interferon and is attenuated in mice. J Virol. 2003;77:13257–66.
18. Egorov A, Brandt S, Sereinig S, Romanova J, Ferko B, Katinger D, Grassauer A, Alexandrova G, Katinger H, Muster T. Transfectant influenza a viruses with

long deletions in the NS1 protein grow efficiently in Vero cells. J Virol. 1998; 72:6437–41.
19. Kochs G, Koerner I, Thiel L, Kothlow S, Kaspers B, Ruggli N, Summerfield A, Pavlovic J, Stech J, Staeheli P. Properties of H7N7 influenza a virus strain SC35M lacking interferon antagonist NS1 in mice and chickens. J Gen Virol. 2007;88:1403–9.
20. Newby CM, Sabin L, Pekosz A. The RNA binding domain of influenza a virus NS1 protein affects secretion of tumor necrosis factor alpha, interleukin-6, and interferon in primary murine tracheal epithelial cells. J Virol. 2007;81:9469–80.
21. Jiao P, Tian G, Li Y, Deng G, Jiang Y, Liu C, Liu W, Bu Z, Kawaoka Y, Chen H. A single-amino-acid substitution in the NS1 protein changes the pathogenicity of H5N1 avian influenza viruses in mice. J Virol. 2008;82:1146–54.
22. Li Z, Jiang Y, Jiao P, Wang A, Zhao F, Tian G, Wang X, Yu K, Bu Z, Chen H. The NS1 gene contributes to the virulence of H5N1 avian influenza viruses. J Virol. 2006;80:11115–23.
23. Zhu Q, Yang H, Chen W, Cao W, Zhong G, Jiao P, Deng G, Yu K, Yang C, Bu Z, et al. A naturally occurring deletion in its NS gene contributes to the attenuation of an H5N1 swine influenza virus in chickens. J Virol. 2008;82:220–8.
24. Liniger M, Moulin HR, Sakoda Y, Ruggli N, Summerfield A. Highly pathogenic avian influenza virus H5N1 controls type I IFN induction in chicken macrophage HD-11 cells: a polygenic trait that involves NS1 and the polymerase complex. Virol J. 2012;9:7.
25. Dundon WG, Milani A, Cattoli G, Capua I. Progressive truncation of the non-structural 1 gene of H7N1 avian influenza viruses following extensive circulation in poultry. Virus Res. 2006;119:171–6.
26. Soubies SM, Volmer C, Croville G, Loupias J, Peralta B, Costes P, Lacroux C, Guerin JL, Volmer R. Species-specific contribution of the four C-terminal amino acids of influenza a virus NS1 protein to virulence. J Virol. 2010;84: 6733–47.
27. Munier S, Larcher T, Cormier-Aline F, Soubieux D, Su B, Guigand L, Labrosse B, Cherel Y, Quere P, Marc D, et al. A genetically engineered waterfowl influenza virus with a deletion in the stalk of the neuraminidase has increased virulence for chickens. J Virol. 2010;84:940–52.
28. Esnault E, Bonsergent C, Larcher T, Bed'hom B, Vautherot JF, Delaleu B, Guigand L, Soubieux D, Marc D, Quere P. A novel chicken lung epithelial cell line: characterization and response to low pathogenicity avian influenza virus. Virus Res. 2011;159:32–42.
29. Matrosovich M, Matrosovich T, Garten W, Klenk HD. New low-viscosity overlay medium for viral plaque assays. Virol J. 2006;3:63.
30. Talon J, Horvath CM, Polley R, Basler CF, Muster T, Palese P, Garcia-Sastre A. Activation of interferon regulatory factor 3 is inhibited by the influenza a virus NS1 protein. J Virol. 2000;74:7989–96.
31. Liniger M, Summerfield A, Zimmer G, McCullough KC, Ruggli N. Chicken cells sense influenza a virus infection through MDA5 and CARDIF signaling involving LGP2. J Virol. 2012;86:705–17.
32. Mehle A, Doudna JA. An inhibitory activity in human cells restricts the function of an avian-like influenza virus polymerase. Cell Host Microbe. 2008;4:111–22.
33. Monne I, Fusaro A, Nelson MI, Bonfanti L, Mulatti P, Hughes J, Murcia PR, Schivo A, Valastro V, Moreno A, et al. Emergence of a highly pathogenic avian influenza virus from a low pathogenic progenitor. J Virol. 2014;88: 4375–88.
34. Munir M, Zohari S, Metreveli G, Baule C, Belak S, Berg M. Alleles a and B of non-structural protein 1 of avian influenza a viruses differentially inhibit beta interferon production in human and mink lung cells. J Gen Virol. 2011;92:2111–21.
35. Adams S, Xing Z, Li J, Mendoza K, Perez D, Reed K, Cardona C. The effect of avian influenza virus NS1 allele on virus replication and innate gene expression in avian cells. Mol Immunol. 2013;56:358–68.
36. Guionie O, Guillou-Cloarec C, Courtois D, Bougeard BS, Amelot M, Jestin V. Experimental infection of Muscovy ducks with highly pathogenic avian influenza virus (H5N1) belonging to clade 2.2. Avian Dis. 2010;54:538–47.
37. Beck M, Schmidt A, Malmstroem J, Claassen M, Ori A, Szymborska A, Herzog F, Rinner O, Ellenberg J, Aebersold R. The quantitative proteome of a human cell line. Mol Syst Biol. 2011;7:549.
38. Matsumoto M, Matsuzaki F, Oshikawa K, Goshima N, Mori M, Kawamura Y, Ogawa K, Fukuda E, Nakatsumi H, Natsume T, et al. A large-scale targeted proteomics assay resource based on an in vitro human proteome. Nat Methods. 2017;14:251–8.
39. Hutchinson EC, Charles PD, Hester SS, Thomas B, Trudgian D, Martinez-Alonso M, Fodor E. Conserved and host-specific features of influenza virion architecture. Nat Commun. 2014;5:4816.
40. Terrier O, Moules V, Carron C, Cartet G, Frobert E, Yver M, Traversier A, Wolff T, Riteau B, Naffakh N, et al. The influenza fingerprints: NS1 and M1 proteins contribute to specific host cell ultrastructure signatures upon infection by different influenza a viruses. Virology. 2012;432:204–18.
41. Wang W, Riedel K, Lynch P, Chien CY, Montelione GT, Krug RM. RNA binding by the novel helical domain of the influenza virus NS1 protein requires its dimer structure and a small number of specific basic amino acids. RNA. 1999;5:195–205.
42. Hoffmann TW, Munier S, Larcher T, Soubieux D, Ledevin M, Esnault E, Tourdes A, Croville G, Guerin JL, Quere P, et al. Length variations in the NA stalk of an H7N1 influenza virus have opposite effects on viral excretion in chickens and ducks. J Virol. 2012;86:584–8.
43. Soubies SM, Hoffmann TW, Croville G, Larcher T, Ledevin M, Soubieux D, Quere P, Guerin JL, Marc D, Volmer R. Deletion of the C-terminal ESEV domain of NS1 does not affect the replication of a low-pathogenic avian influenza virus H7N1 in ducks and chickens. J Gen Virol. 2013;94:50–8.
44. Wang BX, Brown EG, Fish EN. Residues F103 and M106 within the influenza a virus NS1 CPSF4-binding region regulate interferon-stimulated gene translation initiation. Virology. 2017;508:170–9.
45. Schierhorn KL, Jolmes F, Bespalowa J, Saenger S, Peteranderl C, Dzieciolowski J, Mielke M, Budt M, Pleschka S, Herrmann A, et al. Influenza a virus virulence depends on two amino acids in the N-terminal domain of its NS1 protein to facilitate inhibition of the RNA-dependent protein kinase PKR. J Virol. 2017;91
46. Kuo RL, Krug RM. Influenza a virus polymerase is an integral component of the CPSF30-NS1A protein complex in infected cells. J Virol. 2009;83:1611–6.
47. Marion RM, Zurcher T, de la Luna S, Ortin J. Influenza virus NS1 protein interacts with viral transcription-replication complexes in vivo. J Gen Virol. 1997;78(Pt 10):2447–51.
48. Min JY, Li S, Sen GC, Krug RM. A site on the influenza a virus NS1 protein mediates both inhibition of PKR activation and temporal regulation of viral RNA synthesis. Virology. 2007;363:236–43.
49. Robb NC, Chase G, Bier K, Vreede FT, Shaw PC, Naffakh N, Schwemmle M, Fodor E. The influenza a virus NS1 protein interacts with the nucleoprotein of viral ribonucleoprotein complexes. J Virol. 2011;85:5228–31.
50. Chen G, Liu CH, Zhou L, Krug RM. Cellular DDX21 RNA helicase inhibits influenza a virus replication but is counteracted by the viral NS1 protein. Cell Host Microbe. 2014;15:484–93.
51. Pichlmair A, Kandasamy K, Alvisi G, Mulhern O, Sacco R, Habjan M, Binder M, Stefanovic A, Eberle CA, Goncalves A, et al. Viral immune modulators perturb the human molecular network by common and unique strategies. Nature. 2012;487:486–90.
52. de la Luna S, Fortes P, Beloso A, Ortin J. Influenza virus NS1 protein enhances the rate of translation initiation of viral mRNAs. J Virol. 1995;69: 2427–33.
53. Park YW, Katze MG. Translational control by influenza virus. Identification of cis-acting sequences and trans-acting factors which may regulate selective viral mRNA translation. J Biol Chem. 1995;270:28433–9.
54. Pereira CF, Read EKC, Wise HM, Amorim MJ, Digard P. The influenza a virus NS1 protein promotes efficient nuclear export of unspliced viral M1 mRNA. J Virol. 2017;91
55. Garaigorta U, Falcon AM, Ortin J. Genetic analysis of influenza virus NS1 gene: a temperature-sensitive mutant shows defective formation of virus particles. J Virol. 2005;79:15246–57.
56. Du Y, Xin L, Shi Y, Zhang TH, Wu NC, Dai L, Gong D, Brar G, Shu S, Luo J, et al. Genome-wide identification of interferon-sensitive mutations enables influenza vaccine design. Science. 2018;359:290–6.
57. Das K, Ma LC, Xiao R, Radvansky B, Aramini J, Zhao L, Marklund J, Kuo RL, Twu KY, Arnold E, et al. Structural basis for suppression of a host antiviral response by influenza a virus. Proc Natl Acad Sci U S A. 2008;105:13093–8.
58. Nemeroff ME, Barabino SM, Li Y, Keller W, Krug RM. Influenza virus NS1 protein interacts with the cellular 30 kDa subunit of CPSF and inhibits 3'end formation of cellular pre-mRNAs. Mol Cell. 1998;1:991–1000.
59. Bornholdt ZA, Prasad BV. X-ray structure of influenza virus NS1 effector domain. Nat Struct Mol Biol. 2006;13:559–60.
60. Zhou B, Ma J, Liu Q, Bawa B, Wang W, Shabman RS, Duff M, Lee J, Lang Y, Cao N, et al. Characterization of uncultivable bat influenza virus using a replicative synthetic virus. PLoS Pathog. 2014;10:e1004420.

Alan WACQUIEZ

Caractérisation moléculaire du domaine de liaison à l'ARN de la protéine NS1 des virus *influenza* A lors de l'interaction avec un ARN double-brin de haute affinité contenant des séquences sélectionnées *in vitro*

Résumé

Le domaine de liaison à l'ARN de la protéine non-structurale 1 des virus *influenza* A interagit avec des ARN double-brin. Ce travail fait suite à l'identification par SELEX d'aptamères ARN de haute affinité pour NS1 et démontre par gel retard l'interaction de NS1 avec la séquence virale AGCAAAAG, promoteur de la RdRp virale et le motif GUAAC (site canonique donneur d'épissage des pré-ARNm cellulaires). L'analyse par cristallographie de l'ARNdb AWFC01 (qui contient deux motifs GUAAC qui maximisent l'affinité d'interaction) confirme l'importance des hélices α2 et α2' du RBD dans l'interaction avec le squelette ribose-phosphate de l'ARN et révèle une légère courbure de AWFC01 sans en reconnaître spécifiquement les bases.

Mots-clés : *influenza*; NS1; RBD; ARNdb; SELEX; Cristallographie; courbure.

Résumé en anglais

The RNA-Binding Domain of Non-structural protein 1 of *Influenza* A viruses interacts with double-stranded RNA. Our work follows on discover of high affinity's RNA aptamers to NS1 by SELEX and demonstrates by gel shift assays interaction of NS1 with AGCAAAAG's viral sequence, the viral polymerase promoter and the double-strand GUAAC motif (canonical donor splice sites of cellular pre-mRNA). X-Ray crystallography with AWFC01's dsRNA (which contains two GUAAC motifs which confer high affinity) confirms the importance of α2/α2' helical propellers of RBD which interact with ribose-phosphates skeleton of dsRNA and reveals a small curvature in AWFC01's dsRNA without recognize directly basis of dsRNA.

Key-words: *influenza*; NS1; RBD; dsRNA; SELEX; Crystallography; RNA curvature.