Caractérisation Structurale et Biochimique de la Nucléoprotéine des virus grippaux de type A, B et D

Alice Tissot

▶ **To cite this version:**

Alice Tissot. Caractérisation Structurale et Biochimique de la Nucléoprotéine des virus grippaux de type A, B et D. Médecine humaine et pathologie. Université Grenoble Alpes, 2017. Français. NNT : 2017GREAV014 . tel-01686263

HAL Id: tel-01686263
https://theses.hal.science/tel-01686263

Submitted on 17 Jan 2018

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

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

Communauté
UNIVERSITÉ Grenoble Alpes

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITE GRENOBLE ALPES

Spécialité : **Virologie – Immunologie - Microbiologie**

Arrêté ministériel : 25 mai 2016

Présentée par

Alice TISSOT épouse LABARONNE

Thèse dirigée par le **Professeur Rob RUIGROK**,
codirigée par le **Docteur Guy SCHOEHN,** et
encadrée par le **Docteur Thibaut CREPIN**

préparée au sein du **l'Institut de Biologie Structurale (IBS)**

Ecole Doctorale Chimie et Sciences du vivant (EDCSV)

Caractérisation Structurale et Biochimique de la Nucléoprotéine des virus grippaux de type A, B et D

Thèse soutenue publiquement le **8 Juin 2017**
devant le jury composé de :

Docteur Yves Gaudin
DR, Institut de Biologie Intégrative de la cellule, Rapporteur
Docteur Roland Marquet
DR, Institut de Biologie Moléculaire et Cellulaire, Rapporteur
Docteur Mariette Ducatez
CR, Ecole Nationale Vétérinaire Toulouse, Examinateur
Professeur Pascal Poignard
PR, Institut de Biologie Structurale, Président du jury
Professeur Rob Ruigrok
PR, Institut de Biologie Structurale, Directeur de Thèse
Docteur Thibaut Crépin
CR, Institut de Biologie Structurale, Encadrant des Travaux de Thèse

A mon grand-père, André Legros

Remerciements

Je remercie très chaleureusement toutes les personnes qui m'ont accompagnée tout au long de cette aventure et plus particulièrement ***Rob Ruigrok*** *et* ***Guy Schoehn****, mes directeurs de thèse. J'ai eu plaisir à travailler avec eux tout au long de ma thèse et je les remercie vraiment de m'avoir fait confiance*.

Je tiens vraiment à remercier d'un immense MERCI ***Thibaut Crépin*** *sans qui tout ce travail n'aurait certainement pas été possible. Merci Thibaut d'avoir été là pour suivre et encadrer mon travail au jour le jour et aussi pour tout ce que tu m'as appris.*

Je remercie énormément ***Myriam Miloudi****, stagiaire de M2R qui a réalisé une grande partie des travaux portant sur la nucléoprotéine de la grippe de type D. Je remercie* ***Blandine*** *et* ***Elisa*** *qui ont contribué à ce travail.*

Bien évidemment je remercie ***Laura Tengo****,* ***Christopher Swale*** *et* ***Alexandre Monod*** *pour avoir été à mes côtés au laboratoire. Ils m'ont tous beaucoup apporté autant sur le plan professionnel que personnel.*

Un très grand merci aussi à ***Jean-Marie Bourhis*** *qui s'est beaucoup investi sur mes travaux de recherche. Merci aussi d'avoir partagé les galères de train Grenoble-Lyon !*

Je remercie toutes les autres personnes du laboratoire et plus particulièrement ***Caroline, Francine, Claire, Manon, Charles, Philippe****,* ***Nicolas, Pascal, Emilie, Corinne, Stéphanie*** *qui m'ont aidée dans mon travail ou bien avec qui j'ai simplement partagé de bons moments au cours de ces trois années de thèse. Merci aussi* ***à toute l'équipe de microscopie****.*

Merci à tous mes amis de M1/M2R pour leur soutien et plus particulièrement à ***Halima, Charlotte, Vanessa, Julien****.*

Merci à ***Marine*** *pour m'avoir soutenue et rassurée au moment où je cherchais une thèse et écoutée quand j'ai eu besoin pendant ces trois années. Merci à* ***Cécile, Charles, Juliette et Lucie*** *d'avoir supporté les conversations et blagues de biochimistes largement échangées entre Emmanuel et moi. Plus généralement merci aux* ***Scouts et Guides de France*** *d'avoir été la soupape d'évacuation de cette thèse.*

Merci à ***mes parents*** *grâce à qui j'ai pu faire de longues études sans me préoccuper des soucis de la vie quotidienne. Un merci tout particulier à* ***petit père*** *pour la relecture de ma thèse.*

Enfin je tiens à remercier plus spécialement ***Emmanuel Labaronne*** *qui est tous les jours à mes côtés depuis presque 5 ans et qui a toujours su trouver les mots pour m'aider à affronter les difficultés qui se sont présentées à moi notamment au cours de cette thèse. Merci aussi à* ***notre petite Charlotte*** *qui me redonne toujours le sourire quelque soit le moment ou la situation.*

Sommaire

REMERCIEMENTS *2*

SOMMAIRE *7*

CHAPITRE 1 : INTRODUCTION *11*

I. LE VIRUS DE LA GRIPPE **12**
1. UN VIRUS A ARN NEGATIF 12
2. UN VIRUS APPARTENANT A LA FAMILLE DES ORTHOMYXOVIRIDAE 14
3. ZOONOSE 15
4. ELEMENTS D'EPIDEMIOLOGIE 17
5. STRATEGIES ET TRAITEMENTS 20

II. STRUCTURE DU VIRUS DE LA GRIPPE **23**
1. STRUCTURE DE LA PARTICULE VIRALE 23
2. ORGANISATION DU GENOME VIRALE 24
3. LES PROTEINES VIRALES 26
4. LE CYCLE VIRAL 38

III. EVOLUTION DU VIRUS **42**
1. LA DERIVE GENETIQUE 42
2. LE REASSORTIMENT GENETIQUE 42

IV. LA RIBONUCLEOPROTEINE, MACHINERIE VIRALE DE SYNTHESE DE L'ARN **44**
1. STRUCTURE DE LA RIBONUCLEOPROTEINE 44
2. TRANSCRIPTION VERSUS REPLICATION 47
3. INTERACTION NUCLEOPROTEINE-ARN 50

V. CAS DES AUTRES VIRUS A ARN NEGATIF **52**
1. LA RIBONUCLEOPROTEINE DES VIRUS A ARN NEGATIF 52
2. NUCLEOPROTEINES DES VIRUS A ARN NEGATIF 53

VI. TRANSPORT NUCLEOCYTOPLASMIQUE DE LA NUCLEOPROTEINE DU VIRUS DE LA GRIPPE **59**
1. PROTEINES IMPLIQUEES DANS LE TRANSPORT ACTIF A TRAVERS LE COMPLEXE DU PORE NUCLEAIRE 59
2. TRANSPORT ACTIF DES PROTEINES CARGOS AU NOYAU 62
3. CAS DE LA NUCLEOPROTEINE DU VIRUS DE LA GRIPPE 63

OBJECTIFS DU PROJET DE THESE *66*

CHAPITRE 2 : PRINCIPES DES METHODES EXPERIMENTALES *72*

I. CHROMATOGRAPHIE D'EXCLUSION DE TAILLE COUPLEE AU MALLS **74**

II. MICROSCOPIE ELECTRONIQUE A TRANSMISSION **76**
1. PRINCIPE DE LA MICROSCOPIE ELECTRONIQUE A TRANSMISSION 76
2. PRINCIPE DE LA COLORATION NEGATIVE 78

III. ANISOTROPIE DE FLUORESCENCE **80**

IV. CRISTALLOGRAPHIE AUX RAYONS-X **83**
1. LA CRISTALLOGENESE 83
2. ACQUISITION ET INTEGRATION DES DONNEES 85
3. DETERMINATION DE LA PHASE 86

V. TEST DE STABILITE THERMIQUE (TSA) **89**

VI. DICHROÏSME CIRCULAIRE **90**

CHAPITRE 3 : ETUDE DE LA FIXATION DE L'ARN A NP A ET NP B *92*

NOTE D'INTRODUCTION AU CHAPITRE 3 **99**
Article 1: Binding of RNA by the nucleoproteins of influenza viruses A and B

CHAPITRE 4 : CARACTERISATION BIOCHIMIQUE ET STRUCTURALE DE NP D *98*

NOTE D'INTRODUCTION AU CHAPITRE 4 **100**
Article 2: Structure of influenza D virus nucleoprotein and function of NPTAIL

CONCLUSION *104*

LISTES DES ABBREVIATIONS *110*

BIBLIOGRAPHIE *114*

PUBLICATIONS *227*

STRUCTURAL CHARACTERIZATION OF RECOMBINANT IAV POLYMERASE REVEALS STABLE COMPLEX BETWEEN VIRAL PA-PB1 HETERODIMER AND HOST RANBP5 **229**
LEARNING FROM STRUCTURE-BASED DRUG DESIGN AND NEW ANTIVIRAL TARGETING THE RIBONUCLEOPROTEIN **257**

Chapitre 1 : Introduction

I. Le virus de la grippe

1. Un virus à ARN négatif

Les virus à ARN simple brin de polarité négative sont des virus enveloppés (figure 1) dont l'ARN viral (ARNv) est complémentaire de l'ARN messager (ARNm). Selon la classification du biologiste américain David Baltimore, les virus à ARN simple brin de polarité négative appartiennent au groupe V (Baltimore, 1971). Le groupe V comporte 7 familles divisées en deux ordres en fonction de la nature de leur génome (Tableau 1). Les virus possédant un unique brin d'ARNv appartiennent à l'ordre des *Mononegavirales.* Les familles des *Rhabdoviridae* (virus de la rage, virus de la stomatite vésiculaire), des *Paramyxoviridae* (virus de la rougeole, virus respiratoire syncytial, virus Hendra et Nipah), des *Filoviridae* (virus Ebola) et des *Bornaviridae* appartiennent à cet ordre. L'autre ordre inclut des virus dont le génome est segmenté. Les familles des *Arenaviridae (*2 segments d'ARN, virus de la fièvre de Lassa), des *Bunyaviridae* (3 segments d'ARN, virus de la fièvre de la vallée du Rift) et des *Orthomyxoviridae* (6 à 8 segments d'ARN, virus de la grippe) appartiennent à ce second ordre.

Les virus à ARN négatif infectent un large spectre d'hôtes (plantes, invertébrés, mammifères) et certains sont responsables de graves maladies chez l'homme. Le virus de la grippe et le virus respiratoire syncytial provoquent des affections aigües des voies respiratoires. Le virus Nipah et certains *Bunyaviridae* sont responsables d'encéphalites graves. D'autres virus, (virus Ebola, virus de la fièvre de Crimée Congo, virus de la fièvre de Lassa), sont à l'origine de fièvres hémorragiques.

Lors de l'infection d'une cellule hôte, la première activité de ces virus est de transcrire l'ARNv de polarité négative en ARNm qui est ensuite traduit en protéine. Les ARN viraux de ces virus ne sont jamais nus, ni dans les particules virales, ni dans les cellules infectées. Ils sont entièrement recouverts de nombreuses copies de nucléoprotéines et fixent la polymérase virale au sein d'un complexe ribonucléoprotéique. Chez les *Mononegavirales*, la polymérase est associée à la phosphoprotéine au sein de ce complexe (Chua et al., 2000).

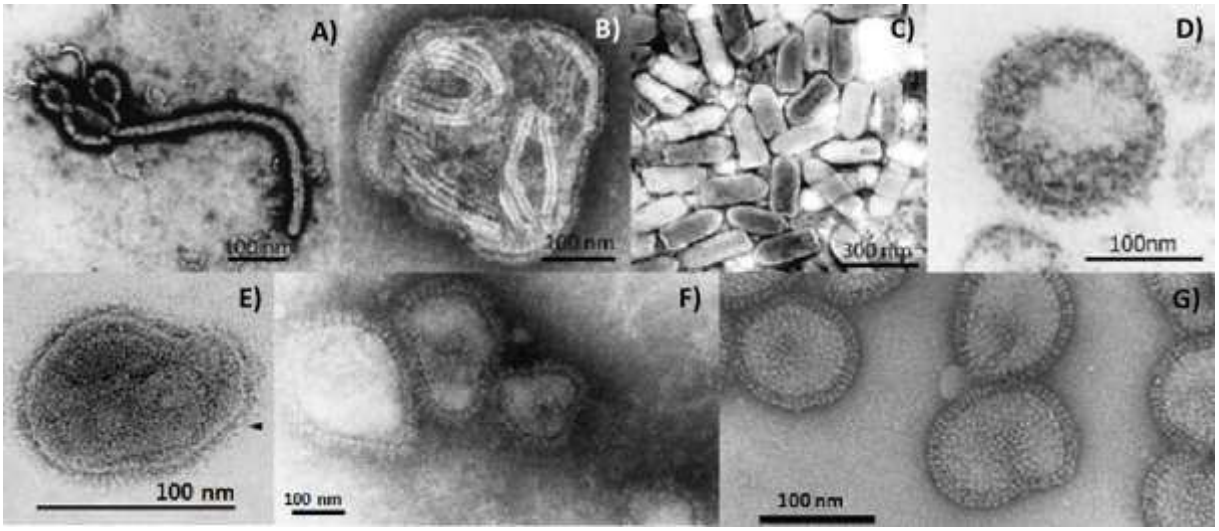


Figure 1: Image de microscopie électronique de virus à ARN négatif. A) Virus Ebola (*Filoviridae*). B) Virus Parainfluenza (*Paramyxoviridae*). C) Virus de la stomatite vésiculaire (*Rhabdoviridae*). D) Virus de la maladie de Borna (*Bornaviridae*). E) Virus de Lassa (*Arenaviridae*). F) Virus de la fièvre de la vallée du Rift (*Bornaviridae*).G) Virus de la grippe (*Orthomyxoviridae*) (Image adaptée de la Thèse d'Aurélie Albertini).

Ordre **Famille** Genre	Exemple de virus connus	Taille du génome (kb)
Mononegavirales		
Rhabdoviridae		**13-16**
Vesiculovirus	Virus de la stomatite vésiculaire	
Lyssavirus	Virus de la rage	
Filoviridae		**13**
Type Marburg	Virus de Marburg	
Type Ebola	Virus Ebola	
Paramyxoviridae		**16-20**
Respirovirus	Virus de Sendaï	
Morbillivirus	Virus de la rougeole	
Rubulavirus	Virus des oreillons	
Henipavirus	Virus Nipah, Virus Hendra	
Pneumovirus	Virus respiratoire syncytial	
Bornaviridae		
Bornavirus	Borna disease virus	**9**
Génome segmenté		
Orthomyxoviridae		**12-15 kb** dans :
Influenza A, B	Virus de la grippe A, B	8 segments
Influenza C, D	Virus de la grippe C, D	7 segments
Thogotovirus	Virus thogoto/Dhori	6/7 segments
Isavirus	Anémie infectieuse du saumon	8 segments
Bunyaviridae		**11-20 kb** dans 3 segments
Bunyavirus	Bunyamwera	
Hantavirus	Virus de séoul	
Phlebovirus	Virus de la vallée du Rift	
Nairovirus	Virus Congo-Crimée	
Tospovirus	Tomato Spotted Wilt Virus	
Arenaviridae		**10-14 kb** dans 2 segments
Arenavirus	Virus de la fièvre de Lassa	

Tableau 1 : Les virus à ARN négatif simple brin

2. Un virus appartenant à la famille des Orthomyxoviridae

Le virus de la grippe est un virus de la famille des *Orthomyxoviridae* (du grec *orthos* signifiant droit ou correct, et *myxa* signifiant mucus). La famille des *Orthomyxoviridae* comprend 7 genres différents de virus (International Committee on Taxonomy of Viruses, http://www.ictvonline.org/):

- *Influenza virus* A
- *Influenza virus* B
- *Influenza virus* C
- *Influenza virus* D
- *Thogotovirus* incluant *Thogoto virus* et *Dhori virus*
- *Isavirus* incluant le virus de l'anémie infectieuse du saumon (*Infectious Salmon Anemia Virus*, ISAV)
- *Quaranfilivirus* incluant *Johnston Atoll virus* et *Quaranfil virus*

Jusqu'à très récemment, les virus grippaux n'étaient classés qu'en 3 types, *influenza* A, B et C. En 2011, un nouveau virus infectant le porc et le bétail a été isolé aux Etats-Unis et a d'abord été apparenté au virus humain de type C. Cependant, ce virus ne présentant que 50 % d'identité globale de séquence et une absence de réactivité croisée entre ses anticorps et le virus humain de type C dans le cadre de l'inhibition de l'hémagglutination, il a très vite été classé en un nouveau type de virus grippal, *infuenza virus* D (Collin et al., 2014; Hause et al., 2014).

Les virus grippaux sont classés en fonction de leurs caractéristiques biologiques et épidémiologiques (Ritchey et al., 1976). Il est notamment possible de les distinguer via les différences antigéniques des protéines de surface et de la nucléoprotéine (Wright, P.F et al., 2013). Le virus de type A est classé en différents sous-types en fonction de la nature antigénique des deux glycoprotéines de surface, l'hémagglutinine (HA) et la neuraminidase (NA). Il existe actuellement 18 sous-types de HA (de H1 à H18), et 11 sous-types de NA (de N1 à N11). Les derniers sous-types H18 et N11 ont été identifiés récemment chez la chauve-souris (Tong et al., 2013). Concernant le type B, il existe un seul type de HA et de NA tandis que les types C et D ne possèdent qu'une seule glycoprotéine de surface multifonctionnelle, l'hémagglutinine-estérase de fusion (HEF) dotée à la fois des caractéristiques de HA et de NA (Hause et al., 2013; Nakada et al., 1984, 1985; Pekosz and Lamb, 1997). La nomenclature actuelle (Bulletin of the World Health Organization, 1980) permet de classifier les différentes souches grippales selon leur type (A, B, C ou D), l'espèce à partir de laquelle la particule virale a été isolée (exception faite des souches humaines), le lieu géographique de l'isolement, le numéro de l'isolat et son année d'isolation. Dans le cas de la grippe A, le sous-

type viral est aussi pris en compte, A HxNx. Les travaux présentés dans ce manuscrit de thèse portent sur les souches A/WSN/33(H1N1), B/Memphis/13/03 et D/bovine/France/2986/2012.

3. Zoonose

Le virus *influenza A*, est susceptible d'infecter un très large panel d'animaux incluant entre autres, les humains, les oiseaux, les porcs, les chevaux, les chiens et les chauves-souris (Chan et al., 2013; Vahlenkamp and Harder, 2006). Le type A peut aussi infecter d'autres animaux de manière plus sporadique, tels que les félins (chats, tigres, léopards), les chameaux, les baleines ou encore les phoques. Des études sérologiques montrent que la grippe A peut aussi toucher les ruminants, les reptiles et les amphibiens (Herfst et al., 2014; Reperant et al., 2009; Vahlenkamp and Harder, 2006) (figure 2). *Influenza* B infecte les humains mais le virus a aussi été isolé chez les phoques (Ohishi et al., 2002; Osterhaus et al., 2000). Le type C est retrouvé chez les humains, les porcs (Guo et al., 1983; Ohwada et al., 1987) et les chiens (Manuguerra and Hannoun, 1992; Ohwada et al., 1987). Le type D est présent chez les porcs, le bétail (Collin et al., 2014; Ducatez et al., 2015; Hause et al., 2014) et les petits ruminants tels que les chèvres et les moutons (Quast et al., 2015).

Les oiseaux aquatiques sauvages de l'ordre des ansériformes (canards, oies, cygnes) et de l'ordre des charadriiformes (mouettes, sternes ou hirondelles de mer et oiseaux de rivage) constituent le réservoir naturel de la grippe A (Kaplan and Webby, 2013; Yoon et al., 2014). Le virus *influenza* A a été isolé d'au moins 105 espèces aviaires appartenant à 26 familles différentes (Yoon et al., 2014). Quasiment tous les sous-types de HA et de NA ont été détectés chez le canard de surface ou barboteur (genre Anas, dont fait partie le canard colvert). Cependant, au milieu des années 1990, l'élevage intensif de volaille a conduit à l'émergence du virus H5N1, un virus aviaire réassorti (Shortridge, 1999; Shortridge et al., 1998) ayant décimé des élevages entiers. Pour la première fois, les épizooties chez la volaille se sont accompagnées d'une transmission d'un virus entièrement aviaire à l'homme avec des issues fatales (Claas et al., 1998a, 1998b; de Jong et al., 1997; Subbarao et al., 1998).

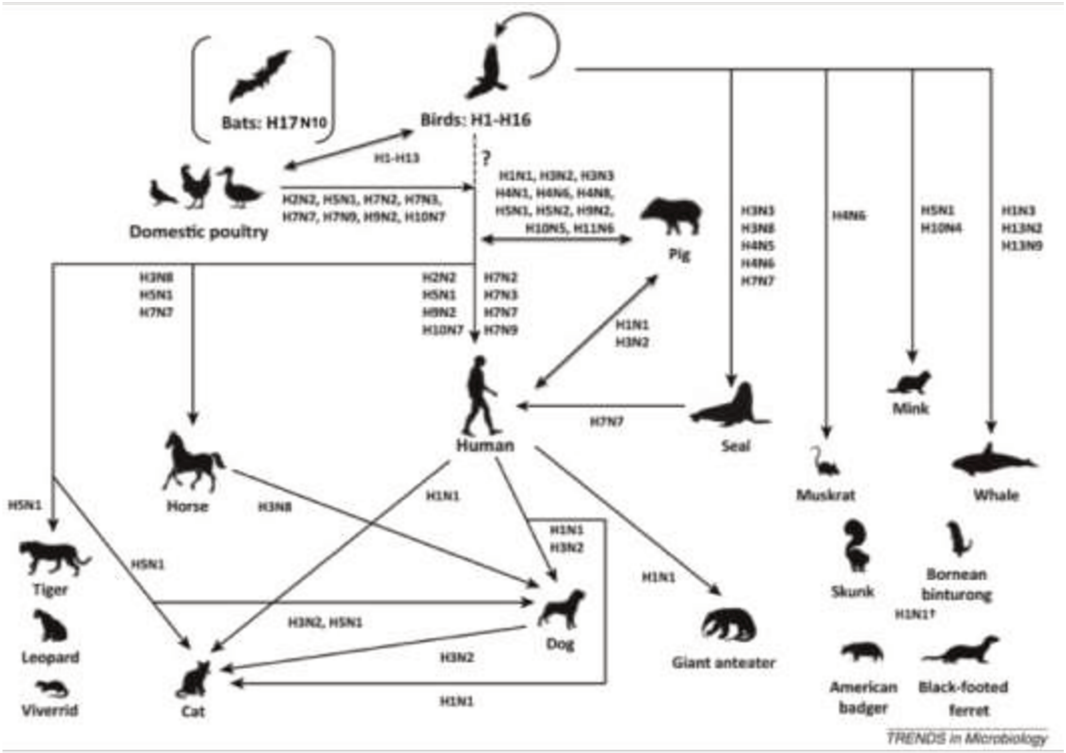


Figure 2 : Les différents hôtes d'*influenza* A. Le réservoir principal du virus de type A est constitué des oiseaux aquatiques sauvages. Le virus peut être transmis à la volaille, aux mammifères marins, aux cochons, aux chevaux, aux humains et à d'autres espèces de mammifères. La transmission peut aussi se faire entre les cochons et les humains et de la volaille aux humains. Plusieurs espèces (chevaux, humains, félins) peuvent transmettre le virus aux animaux domestiques comme le chat et le chien. (Adaptée de Chan et al., 2013).

Bien que moins fréquentes que la grippe de type A, *influenza* B peut être responsable d'épidémies de grippes saisonnières et sa virulence n'est pas négligeable. Au contraire, *influenza* C, qui infecte plus spécifiquement les enfants en bas âge, est rarement associée à des symptômes sévères (Moriuchi et al., 1991). La plupart des jeunes adultes (96 %) possèdent les anticorps contre le type C (O'Callaghan et al., 1980). Les virus A, B et C possèdent très certainement un ancêtre commun (Wright, P.F et al., 2013).

4. Eléments d'épidémiologie

Le virus de la grippe est responsable chez l'homme d'infections respiratoires aiguës hautement contagieuses. La maladie se caractérise par un ensemble de symptômes non spécifiques tels qu'une forte fièvre, une toux sèche, des douleurs musculaires (courbatures) et articulaires accompagnées de céphalées et d'une intense asthénie. Dans les cas les plus sévères, l'affection peut s'accompagner de complications graves telles que des pneumonies avec surinfection bactérienne et déshydratation pouvant conduire au décès. Les personnes les plus sensibles au virus sont les enfants de moins de 15 ans, les personnes de plus de 65 ans et les personnes atteintes de maladies chroniques. Les symptômes persistent 2 à 7 jours en moyenne. Les symptômes grippaux sont dus à la réponse immunitaire suite à l'infection virale ainsi qu'à la mort des cellules infectées. Cette réponse est très personnelle et dépend à la fois de la virulence du virus et de l'exubérance du système immunitaire de l'hôte (Junge, 2011). Cette maladie, qui a traversé les siècles, reste aujourd'hui encore une cause importante de morbidité et de mortalité dans le monde.

L'histoire de la grippe s'articule autour de nombreuses épidémies et pandémies qui ont marqué différentes époques. En 412 avant J.-C., Hippocrate décrivait déjà une pathologie avec des symptômes caractéristiques du virus de la grippe (Hoehling, 1961). De même, au cours des 16^{ième} et 17^{ième} siècles, plusieurs épisodes épidémiques et pandémiques, dont les symptômes correspondent à ceux de la grippe, ont été décrits. En 1510, Tomassinno de'Bianchi décrit une épidémie de maladie de transpiration, « *Sudor Anglicus* », présentant des symptômes identiques à ceux de la grippe actuelle (Morens et al., 2010). Enfin, les épisodes pandémiques survenus au cours du 19^{ième} siècle, ont pu être attribués aux virus de la grippe grâce à des analyses séro-épidémiologiques et séro-archéologiques (Beveridge, 1977; Patterson, 1985).

Grâce aux progrès techniques en médecine et en infectiologie, le virologiste américain Richard Shope isole le premier virion grippal en 1930 (Lewis and Shope, 1931; Shope, 1931a, 1931b). Il s'agit d'un virus d'origine porcine (A/swine/iowa/1930). En 1933, Wilson Smith isole le premier virus humain (A/wilson-smith/1933) qui sera défini comme appartenant au type A. En 1940, un virus grippal aux propriétés antigéniques inconnues est isolé. Il sera défini comme appartenant au type B (Francis, 1940). Un virus de type C est isolé en 1949 (Taylor, 1949) et le premier virus de type D en 2011 (Hause et al., 2013).

Le virus de la grippe possède donc à la fois la capacité de sévir selon **un mode épidémique** saisonnier d'intensité variable et selon **un mode pandémique**. **La grippe saisonnière**, représentant la première cause d'infection des voies respiratoires, opère durant la période automno-hivernale de part et d'autre du globe. Elle touche l'hémisphère nord

d'octobre à avril et l'hémisphère sud d'avril à octobre (World Health Organization). Dans l'hémisphère nord, le pic d'infection survient généralement entre Janvier et Février. Dans les parties tropicales, il n'y a pas véritablement de saison bien marquée. Le virus circule donc tout au long de l'année (Russell et al., 2008; Viboud et al., 2006). C'est notamment le cas de l'Asie du sud-est, dont la localisation géographique fait d'elle une importante route commerciale avec une forte densité de population. Cela favorise le maintien d'un lien épidémique continu entre l'hémisphère nord et l'hémisphère sud. A l'inverse, un climat continental provoque d'importantes variations de température tout au long de l'année. Ces variations génèrent des pics de distribution du virus au cours de l'année en limitant ou favorisant la transmission du virus par voie aérienne (Noti et al., 2013) ainsi que les surfaces de viabilité du virus (McDevitt et al., 2010). Chez l'homme, la transmission du virus est surtout due à l'inhalation de gouttelettes issues des éternuements et de la toux des personnes infectées.

Chaque année, la grippe saisonnière touche 5 à 15 % de la population mondiale. Entre 3 et 5 millions de cas graves et 250 000 à 500 000 décès sont enregistrés par an dans le monde (World Health Organization). En France, ce sont 2,5 millions de personnes infectées et 1500 à 2000 décès répertoriés chaque année. Les décès touchent plus particulièrement les plus de 65 ans (90 % des décès)(Santé publique France, 2017).

Au-delà du fait que la grippe constitue un véritable problème de santé publique, les conséquences économiques dues au virus de la grippe sont elles aussi considérables. Les pertes financières sont évaluées à plusieurs milliards de dollars chaque année (Weycker et al., 2005). En France, selon l'intensité de l'épidémie, le coût annuel en médecine ambulatoire (hors hospitalisation et perte de production) peut atteindre jusqu'à 800 millions d'euros lors des manifestations les plus graves de la maladie (GEIG, Groupe d'Expertise et d'Information sur la Grippe). Il est cependant possible de prévenir les pertes économiques et humaines survenant au cours des épidémies de grippe saisonnière, notamment à l'aide de la vaccination (Cf. § I.5). A l'inverse, les pandémies sont difficilement contrôlables. Elles sont la conséquence d'une évolution du virus grippal due à des réassortiments viraux et/ou aux changements génétiques du virus (Cf. § III.1 et III.2). Les populations humaines ne possédant aucun anticorps vis-à-vis de ces nouveaux sous-types viraux, elles ne sont pas immunisées contre ces virus et les pandémies sont donc très meurtrières.

Une pandémie a pour caractéristique d'atteindre la totalité du globe en une courte période. Elle touche tous les sujets, même jeunes et en bonne santé, et présente un taux de mortalité très élevée. Jusqu'à ce jour, les pandémies ont toutes été provoquées par un virus *influenza* de type A. Plusieurs pandémies, assez exceptionnelles du fait de leur gravité et de leur extension, sont survenues au cours du 20ème siècle. La plus meurtrière reste la grippe

espagnole, A/H1N1 de 1918-1919. Cette pandémie, affectant particulièrement les jeunes adultes de 15 à 35 ans (Simonsen et al., 1998), a affecté 30 % de la population mondiale et est responsable d'au moins 50 millions de morts dans le monde (Johnson and Mueller, 2002; Patterson and Pyle, 1991) soit environ 5 fois plus que le nombre de morts dus à la première guerre mondiale (Clodfelter, 2002). Cependant la grippe espagnole est survenue dans un contexte particulier en infectant des organismes affaiblis par les 4 années de guerre, sujets à la famine ainsi que la malnutrition (Song, 2014) et la plupart des malades mourraient de surinfections d'origine bactérienne (Brundage, 2006; Chien et al., 2009; Taubenberger and Morens, 2006). Lors de la pandémie de grippe A/H1N1 de 2009, le spectre de 1918 a plané sur le monde. En 1957, le virus H2N2 fut responsable d'un nouvel épisode pandémique ayant pour origine le sud de la province de Guizhou en Chine. En seulement quelques mois il s'est propagé à Singapour, Hong-Kong puis finalement jusqu'au Japon. Provenant d'un réassortiment entre un virus humain et un virus aviaire, il n'était pas tellement virulent mais fut tout de même très meurtrier car les populations humaines ne possédaient aucune immunité préexistante contre les protéines de surface HA et NA. En 1968, le virus A/H3N2, originaire de l'Asie du sud-est, (Bedford et al., 2010; Chan et al., 2010; Russell et al., 2008) s'est rapidement répandu à travers le monde (Cockburn et al., 1969). Ces deux virus sont toujours en circulation chez l'animal et sont susceptibles de ré-émerger. En 1996, le virus H5N1 a émergé d'une ferme d'oie située dans une province chinoise (Claas et al., 1998b; Peiris et al., 2004; Sims et al., 2005). Depuis, ce virus dont la transmission est, la plupart du temps, provoquée par un contact direct avec les volailles/oiseaux contaminés, sévit en Asie du sud-est. De très rares cas de transmission d'homme à homme ont été rapportés (Normile, 2006; Ungchusak et al., 2005). Ce virus, bien que faiblement transmissible à l'homme présente un taux de mortalité de 60 %, ce qui le rend hautement pathogène. S'il venait à se réassortir avec un virus humain, il pourrait conduire à une nouvelle souche extrêmement virulente. Il faut aussi noter que plus récemment, la souche A/H7N9 a été identifiée comme infectant les humains avec une forte pathogénicité.

Depuis 1977 et la réémergence du virus H1N1, les souches saisonnières de type A/H1N1, A/H3N2 et de type B, B/yamagata et B/victoria circulent ensembles (World Health Organization). Depuis 2009, la souche saisonnière H1N1 a été largement remplacée par la souche pandémique A/H1N1 (Broor et al., 2012; Wright, P.F et al., 2013).

5. Stratégies et traitements

Du fait de son potentiel mutagène très élevé, le virus influenza est encore une véritable cible mouvante pour les scientifiques. Aujourd'hui, la protection la plus efficace contre le virus de la grippe reste la vaccination. Cependant, du fait de la variabilité des souches en circulation, le vaccin doit être réactualisé chaque année pour être efficace. Les vaccins proposés contiennent des virus inactivés ou des virus vivants atténués. La plupart du temps, ils contiennent deux virus de types A A/H3N2 et A/H1N1 et une ou deux souches de type B. Deux fois par an, en se basant sur l'observation des différentes souches de grippe saisonnière circulant dans l'hémisphère opposé, l'organisation mondiale de la santé donne les recommandations pour la préparation du nouveau vaccin.

Cependant, la vaccination saisonnière est inefficace contre les souches pandémiques. En effet, il est impossible de prédire la dérive antigénique/la souche virale à l'origine de la pandémie et il faut 6 mois pour concevoir un nouveau vaccin. Lors de ces périodes de conception, d'autres moyens de combattre le virus sont donc nécessaires. De nombreux efforts sont fournis par la communauté scientifique afin de développer des antiviraux. A ce jour, il y a seulement 4 inhibiteurs grippaux disponibles dans le commerce. Ces 4 inhibiteurs peuvent être divisés en deux groupes :

Les inhibiteurs du canal ionique M2 (Tableau 2) :

Développés au cours des années 60, ces inhibiteurs, ne présentant que peu d'effets secondaires, sont connus sous le nom d'Amantadine et de Rimantadine (Davies et al., 1964; Dolin et al., 1982). Ils sont commercialisés sous leur nom respectif de Symmetrel (Endo pharmaceuticals) et Flumadine (Forest pharmaceuticals). Ces composés sont des dérivés amines, de la classe des adamantanes, s'intercalant au sein du canal ionique M2 (Cf. § II.3) et bloquant le flux de protons nécessaire au relargage du contenu viral dans la cellule (Cf. § II.4). Malheureusement, ces composés efficaces uniquement contre les virus de types A (Belshe et al., 1989; Hayden and Hay, 1992) ont perdu en efficacité au milieu des années 2000. En effet, 99 % des souches saisonnières de H3N2 sont devenues résistantes à cette classe d'inhibiteurs (CDC, 2016).

Les inhibiteurs de Neuraminidase (Tableau 2) :

Connus sous les noms de Zanamivir et Oseltamivir, ils sont respectivement commercialisés sous les noms de Relenza (GlaxoSmithKline) et Tamiflu (Roche). Ces composés sont des analogues d'acide sialique ciblant le site actif de la neuraminidase. Par un mécanisme de compétition, ils empêchent le clivage de l'acide sialique présent sur la membrane des cellules infectées. Or ce clivage est indispensable au relargage des nouveaux virions (Cf. § II.4). Du fait du haut niveau de conservation du site actif de la neuraminidase, il s'agit d'inhibiteurs à large spectre étant efficaces sur la plupart des souches *influenza* A et B

(Yen et al., 2006). Depuis les années 90, ces médicaments ont été utilisés comme traitement prophylaxique et actif à la fois contre la grippe saisonnière et les souches pandémiques avec des degrés d'efficacité variable en fonction de la souche virale et des personnes infectées. Bien qu'ils soient bien plus efficaces que les autres, les inhibiteurs de neuraminidase présentent beaucoup plus d'effets secondaires tels que des vomissements, des douleurs abdominales et des nausées (Dutkowski et al., 2003). Quelques souches résistantes sont aussi apparues suite à l'utilisation de ces médicaments (Samson et al., 2013).

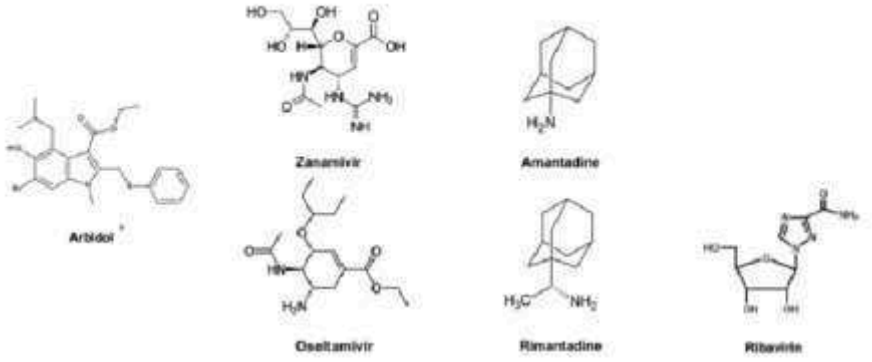
Il a été récemment reporté que des souches résistantes à la fois à l'Oseltamivir et l'Amantadine sont en train d'émerger. Du fait de la portée d'action relativement limitée de ces inhibiteurs et en considérant l'apparition rapide et inévitable des souches multi-résistantes, de nouvelles stratégies pour inhiber la prolifération virale sont en cours (Monod et al., 2015). De nouveaux inhibiteurs de NA tels que le Peramivir et le Lanimivir (Tableau 2), semblent assez prometteurs et sont efficaces contre les souches résistantes à l'Oseltamivir (Babu et al., 2000; Koyama et al., 2009; Samson et al., 2013). Plusieurs inhibiteurs actuellement testés sont déjà en phase clinique III aux Etats-Unis et ont déjà été approuvés cliniquement dans différents pays. D'autres molécules approuvées seulement dans certains pays sont aussi utilisées pour traiter la grippe. C'est le cas de certains antiviraux à large spectre comme l'Arbidol, un inhibiteur d'hémagglutinine, seulement approuvé en Chine et en Russie (Tableau 2) (Teissier et al., 2011). La Ribavirin (Tableau 2), un analogue de nucléoside, surtout utilisé contre l'hépatite C a montré son efficacité contre le virus *influenza*. Il est cependant responsable d'importants effets secondaires tels que des anémies hémolytiques. Le Favipiravir (Tableau 2) (Kiso et al., 2010), un analogue de nucléoside est approuvé au Japon pour le traitement de la grippe. Il a aussi été montré qu'il est efficace contre d'autres virus à ARN négatif notamment le virus Ebola (Oestereich et al., 2014).

L'objectif est désormais de développer de nouveaux antiviraux afin de cibler de multiples voies virales dans le but de limiter l'échappement du virus aux traitements antiviraux, échappement dû à son haut taux de mutation. Pour cela, de nombreux scientifiques tentent de comprendre toujours mieux le fonctionnement du virus notamment en étudiant les protéines impliquées dans les mécanismes réplicatif et transcriptionnel du virus de la grippe. Ce projet de thèse s'inscrit dans cette dynamique.

Protéines Cibles	Hemagglutinine	Neuraminidase	Canal ionique M2	ARN polymérase ARN dépendante
Fonctions ciblées	Reconnaissance/entrée dans la cellule hôte	Relargage des particules virales	Fusion des membranes durant l'endocytose	Réplication Transcription
Inhibiteurs	Anticorps IgG neutralisants Arbidol	Dérivés d'acides sialiques	Dérivés d'adamantanes	Analogues de nucléosides

Utilisation thérapeutique approuvée

*Seulement en Russie et en Chine



En phase III d'essai clinique

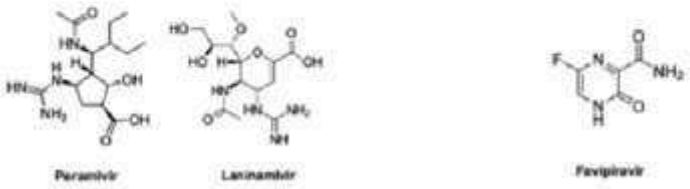


Tableau 2 : Les différents inhibiteurs utilisés pour le traitement du virus de la grippe (Représentation des composés adaptée de Monod et al., 2015).

II. Structure du virus de la grippe

1. Structure de la particule virale

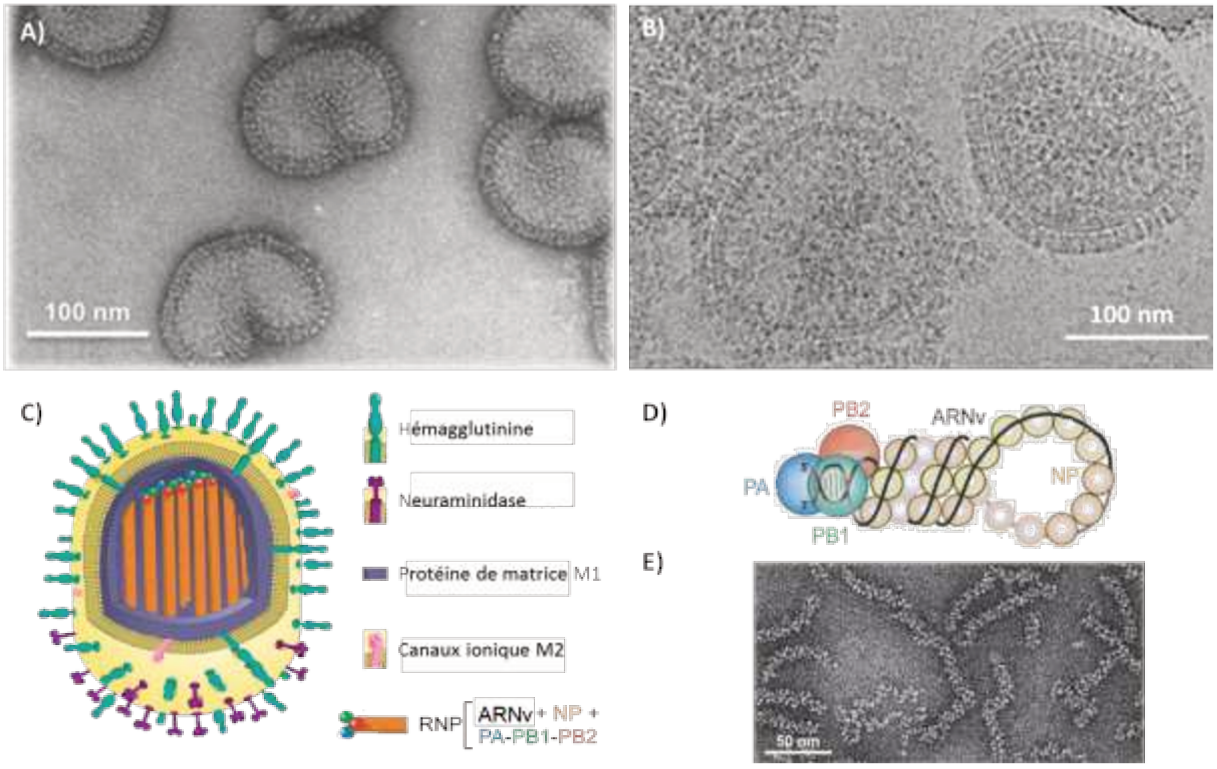


Figure 3 : La particule virale du virus de la grippe. A) Photo de microscopie électronique en coloration négative de virions de type A (Rob Ruigrok, non publiée). B) Photo de cryo-microscopie électronique de virions de type B (adaptée de Katz et al., 2014). C) Représentation schématique d'une particule virale de type A et de ses 8 segments d'ARNv organisés en complexes ribonucléoprotéiques (Claudia Dominguez, 2010). D) Représentation schématique d'une ribonucléoprotéine : l'ARNv est encapsidé par la NP et ses extrémités 3' et 5' sont fixées à la polymérase (Naffakh et al., 2008). E) Photo de microscopie électronique en coloration négative de ribonucléoprotéines purifiées à partir de virions de type A (Ruigrok et al., 2011).

Les virions de la grippe peuvent adopter des formes variables, ils sont pléomorphiques. Les particules virales de types A et B, quasiment indistinguables en microscopie électronique, sont généralement sphériques avec un diamètre d'environ 120 nm (Booy et al., 1985; Fujiyoshi et al., 1994; Katz et al., 2014) (figures 3A et 3B). Cependant, des virus de forme filamenteuse et de taille bien plus importante ont déjà été observés (Kilbourne and Murphy, 1960). Pour les virus de type C, les virions sont souvent trouvés à la surface des cellules infectées sous forme de structures allongées pouvant atteindre les 500 nm (Muraki et al., 2004; Nishimura et al., 1990).

Le virus de la grippe est un virus enveloppé. Son enveloppe dérive de la bicouche lipidique de la cellule hôte et est enrichie en cholestérol et en glycosphingolipides (Scheiffele et al., 1999). Au sein de la membrane des virions de type A, sont insérées deux glycoprotéines de surface, l'hémagglutinine et la neuraminidase, ainsi que le canal ionique M2 (figure 3C). HA et NA sont des protéines formant des spicules d'une dizaine de

nanomètres à la surface du virus. Les virions de type B possèdent 4 protéines de surface : les deux glycoprotéines HA et NA et les protéines BM2 et NB. Ces dernières sont des canaux ioniques (Betakova et al., 1996; Odagiri et al., 1999; Zebedee and Lamb, 1988). Pour *influenza* C et D, une seule glycoprotéine de surface regroupant les fonctions de HA et NA est retrouvée : l'hémagglutinine-estérase de fusion (Apostolov and Flewett, 1969; Song et al., 2016). La protéine CM2 des types C est analogue à M2 et NB (Nakada et al., 1984, 1985; Pekosz and Lamb, 1997). Pour tous les types, A, B, et C, la forme générale du virion est donnée par la protéine de matrice M1 qui est retrouvée à l'interface entre la membrane lipidique et le génome viral (Sha and Luo, 1997). En fonction du type viral, le virus de la grippe possède 7 à 8 brins d'ARN monocaténaires de polarité négative (Cf. § II.2). Chaque brin d'ARNv est lié à de multiples copies de nucléoprotéine (NP) et ses extrémités 5' et 3' sont liées au complexe de l'ARN-polymérase-ARN-dépendante (RdRp). La polymérase est une hétéroprotéine constituée de deux protéines basiques 1 et 2 (PB1 et PB2) et d'une protéine acide (PA). Cet ensemble macromoléculaire (ARNv/polymérase/NP) forme la particule ribonucléoprotéique (figures 3D et 3E). Au sein des virions purifiés se trouve aussi la protéine d'export nucléaire (NEP, précédemment décrite comme NS2) (O'Neill et al., 1998; Richardson and Akkina, 1991) (Cf. § II.3).

2. Organisation du génome virale

Au sein des virions, les virus de type A et B contiennent 8 segments d'ARNv (Palese and Schulman, 1976; Ritchey et al., 1976) tandis que les virus de type C et D n'en contiennent que 7 (Hause et al., 2013; Wright, P.F et al., 2013). Les segments sont numérotés en fonction de leur longueur et portent le nom de la protéine principale pour laquelle ils codent. Par exemple chez les virus de type A, les 8 brins sont numérotés de 1 à 8. Le segment 1 correspond au brin le plus long (2341 nucléotides) tandis que le segment 8 représente le plus court (890 nucléotides).

Tous les segments d'ARNv sont organisés selon le même schéma avec une phase de lecture (ORF) centrale bordée de courtes régions non traduites (UnTranslated Regions, UTR) (figure 4). Les régions non traduites sont différentes pour chaque segment d'ARNv. Par exemple pour les virus de type A, les UTR peuvent faire entre 9 et 58 nucléotides en fonction des segments. Aux extrémités 3' et 5' des régions UTR se trouvent deux séquences U12 et U13, respectivement de 12 et 13 nucléotides. Ces courtes séquences sont hautement conservées entre les segments d'ARNv d'un même type viral mais aussi entre les différentes souches virales d'un même type (Desselberger et al., 1980; Gultyaev et al., 2010). La complémentarité partielle qui existe entre les séquences U12 et U13 donne une structure particulière et unique à l'ARNv. Les extrémités conservées de chaque segment représentent

les séquences minimales à la transcription et à la réplication et constituent ainsi le promoteur de l'ARNv (Fodor et al., 1994; Luytjes et al., 1989).

Figure 4 : Organisation d'un segment d'ARN viral. La région codante (cadre bleu) est flanquée de séquences non codantes spécifiques à chaque segment d'ARNv. Aux extrémités terminales 5' et 3' les séquences U12 et U13 (cadres jaunes) sont hautement conservées entre tous les segments des différents types viraux. (Schéma adapté de la thèse d'Alexandre Monod, 2014).

Chaque segment code pour au moins une protéine. Certains segments codent à la fois pour une protéine majeure exprimée avec l'ORF principale et une ou plusieurs autres protéines. Ces protéines sont exprimées soit par épissage alternatif soit par le biais d'une seconde ORF. En moyenne, le génome viral code pour une dizaine de protéines retrouvées dans tous les types A et B. Dans certains types viraux des protéines accessoires sont aussi retrouvées. C'est notamment le cas dans certaines souches de type A où par exemple, la protéine PB1-F2 (impliquée dans l'échappement du virus au système immunitaire) (Chen et al., 2001) et la protéine PB1-N40 (ayant potentiellement un rôle dans la prolifération virale) (Tauber et al., 2012; Wise et al., 2009) sont exprimées à partir d'une ORF alternative au sein du gène codant pour PB1.

Les trois segments les plus longs, 1, 2 et 3 codent respectivement pour PB2, PB1 et PA, les trois sous-unités de la polymérase (Tableau 3). Pour les types C et D, P3 est l'homologue de PA et n'est pas chargée à pH neutre (Yamashita et al., 1989). Pour les types A et B, le segment 4 code pour la protéine HA alors qu'il code pour la protéine HEF pour les types C (Herrler et al., 1988) et D (Ducatez et al., 2015; Hause et al, 2013). Dans tous les cas, le segment 5 code pour la NP. Pour *influenza* A, le segment 6 code pour NA tandis qu'il code pour NA et NB pour *influenza* B (Racaniello and Palese, 1979). Pour *influenza* C et D, le segment 6 code pour la protéine de matrice M et le canal ionique CM2 (Pekosz and Lamb, 1997). Pour les virus de type A et B, le segment 7 code pour la protéine M1 et respectivement M2 et BM2 pour *influenza* A et B. Les protéines NS1 et NEP/NS2 sont codées par le segment 8 des types A et B tandis qu'il s'agit du segment 7 pour les types C (Alamgir et al., 2000; Nakada et al., 1985) et D (Ducatez et al., 2015; Hause et al., 2013).

Segment	Type viral	Taille (nucléotides)	Protéines codées	Abréviation	Taille (acides aminés)	Poids moléculaire
1	A, B, C et D	2341	Protéine basique 2	PB2	759	85,7 kDa
2	A, B, C et D	2341	Protéine basique 1	PB1	757	86,5 kDa
3	A et B	2233	Protéine acide	PA	716	84,2 kDa
	C et D		Protéine 3	P3		
4	A et B	1778	Hémagglutinine	HA	566	61,5 kDa
	C et D		Hémagglutinine-estérase de fusion	HEF		
5	A, B, C et D	1565	Nucléoprotéine	NP	498	56,1 kDa
6	A et B	1413	Neuraminidase	NA	454	50,1 kDa
	B			NB		
	C et D		Canal ionique	CM2		
			Protéine de matrice	M		
7	A et B	1027	Protéine de matrice	M1	252	27,8 kDa
	A		Canal ionique	M2	97	11 kDa
	B		Canal ionique	BM2		
	C et D		Protéine non structurale 1	NS1		
			Protéine d'export nucléaire	NEP/NS2		
8	A et B	1027	Protéine non structurale 1	NS1	230	26,8 kDa
		890	Protéine d'export nucléaire	NEP/NS2	121	14,2 kDa

Tableau 3 : Protéines principales codées par le génome des virus grippaux A, B, C et D. La plupart des segments sont transcrits en une seule protéine. Cependant certains segments codent pour plusieurs protéines. La taille du génome, de la protéine et son poids moléculaire sont indiqués à titre indicatif pour un virus de type A (A/PR/8/34). De légères différences sont observées pour ces trois paramètres au sein des différents types viraux. (Adaptée de Wright, P.F. et al (2013).

3. Les protéines virales

Selon les protéines, la traduction et plus particulièrement la maturation s'effectuent à des endroits différents. Les glycoprotéines (HA, NA, HEF) et les canaux ioniques (M2, NB, BM2, CM2) sont synthétisés par les ribosomes liés au réticulum endoplasmique (RE). Ces protéines transitent ensuite par l'appareil de golgi où certaines sont glycosylées (HA, NA et HEF). Toutes ces protéines de surface sont finalement insérées au sein de la membrane plasmique. Les autres protéines (PB1, PB2, PA, P3, NP, NS1, NEP/NS2, M1 et M), sont synthétisées par des ribosomes libres présents dans le cytoplasme.

Les 10 protéines principales du virus de la grippe sont décrites dans les paragraphes suivants. Cette description est faite uniquement pour les protéines de virions de type A exception faite de la nucléoprotéine. Cependant, il faut noter que des ressemblances structurales et fonctionnelles importantes sont retrouvées entre toutes les protéines des différents types viraux.

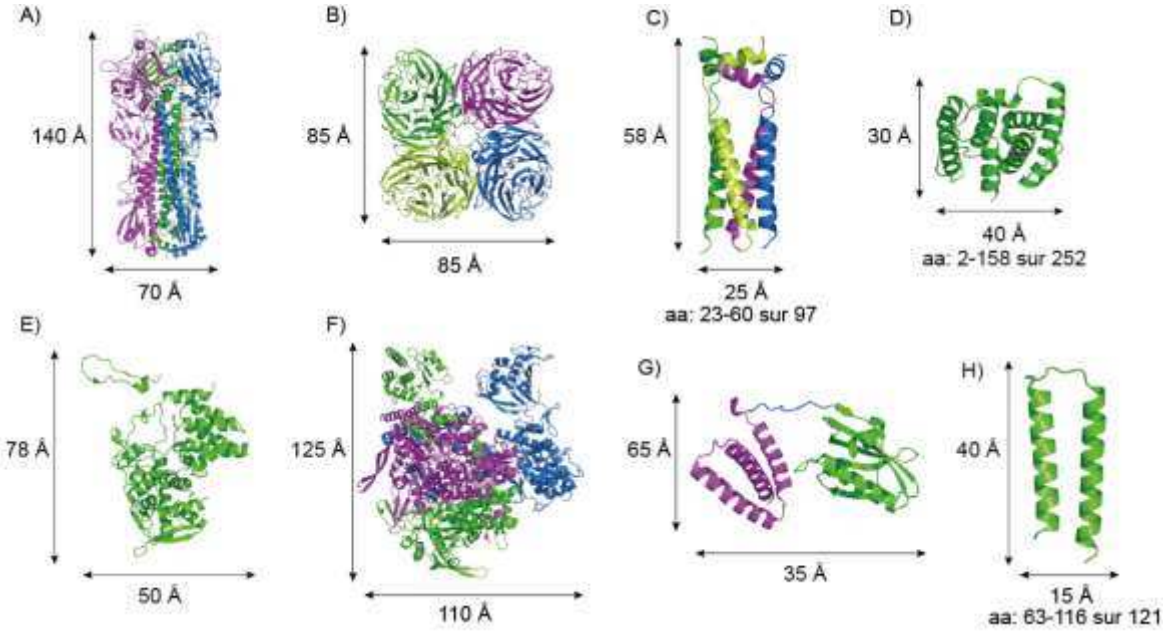


Figure 5 : Structure 3D des protéines du virus de la grippe de type A. A) Structure cristallographique du trimère de l'hémagglutinine à 3 Å de résolution (Lin et al., 2012) (PDB ID 2YPG).B) Structure cristallographique du tétramère de la neuraminidase à 2.2 Å de résolution (Zhu et al., 2012) (PDB ID 4GZP). C) Structure RMN du canal ionique M2 (Schnell and Chou, 2008)(PDB ID 2RLF). D) Structure cristallographique de la protéine matricielle M1 à 2.3 Å de résolution (Arzt et al., 2001) (PDB ID 1EA3). E) Structure cristallographique de la nucléoprotéine (NP) à 3.2 Å de résolution (Ye et al., 2006)(PDB ID 2IQH). F) Structure cristallographique à 2.65 Å de résolution de la polymérase composée des 3 sous-unités : PB1 en vert, PB2 en bleu et PA en magenta (Pflug et al., 2014)(PDB ID 4WSB). G) Structure cristallographique de la protéine non structurale (NS1) résolue à 3.16 Å (Carrillo et al., 2014a)(PDB ID 4OPH). H) Structure cristallographique du domaine carboxy-terminal de la protéine d'export nucléaire (NEP/NS2) à 2.6 Å de résolution (Akarsu et al., 2003) (PDB ID 1PD3).

L'hémagglutinine

Il s'agit de la protéine de surface la plus abondante du virus de la grippe. Elle permet la reconnaissance et l'attachement à la cellule hôte. HA est une protéine de fusion virale de classe I, tout comme les protéines de fusion GP41 du VIH et GP2 du virus Ebola (Wright, P.F. et al., 2013, Baquero et al., 2013).

De nombreuses structures de HA sont disponibles. Toutes ces structures révèlent une tige filamenteuse faite d'un enroulement en triple brins d'hélice-α correspondant à la partie HA2. Cette tige est surmontée d'une tête globulaire à trois domaines correspondant à la partie HA1 (figure 5A). La tête globulaire permet la reconnaissance et l'interaction avec les acides sialiques des cellules infectées (Cf. § II.4). Au cours de sa synthèse, HA subit de nombreuses modifications structurales. La protéine est d'abord synthétisée sous forme d'un précurseur (HA0) au niveau du RE. Ce précurseur s'associe en trimère. La protéine est ensuite clivée en deux sous-unités par les protéases cellulaires : l'ectodomaine HA1 (domaine amino-terminal de HA0) et l'ectodomaine HA2 (domaine carboxy-terminal de

HA0). Lors de l'endocytose, l'acidification de l'endosome va entraîner des changements conformationnels majeurs de HA2 permettant alors l'exposition du peptide de fusion hydrophobe. Ce peptide va venir se nicher dans la membrane endosomale et permettre le rapprochement puis la fusion des membranes (Cf. § II.4) (Danieli et al., 1996; Harrison, 2008). Les ribonucléoprotéines (RNP) peuvent alors être libérées dans le cytoplasme à travers des pores de fusion (Helenius, 1992).

HA est aussi l'un des deux principaux antigènes de surface du virus *influenza*. Durant l'infection, le domaine globulaire HA1 stimule la génération d'anticorps neutralisants par les lymphocytes B. Ces anticorps peuvent conduire à la sélection de variants qui vont échapper au système immunitaire de l'hôte. Accumulés dans le temps ces variants peuvent mener à une dérive antigénique (Cf. § III.1) (Smith et al., 2004; Wan et al., 2014).

Finalement, HA est, avec NA, directement impliquée dans le processus de bourgeonnement, à travers la coalescence des radeaux lipidiques très favorables aux zones de bourgeonnement (Schmitt and Lamb, 2005).

La Neuraminidase

NA est une protéine tétramérique retrouvée à la surface des virions. Majoritairement globulaire, elle possède 4 têtes ancrées au sein des radeaux lipidiques par leur extrémité amino-terminale (figure 5B) (Blok et al., 1982). Ces têtes globulaire hébergent une poche permettant la reconnaissance et le clivage du motif acide sialique (Colman et al., 1983). Lors du relargage des virions, NA clive les acides sialiques et évite ainsi que les virions restent collés à la surface cellulaire à cause des interactions HA/acides sialique (Cf. § II.4) (Schmitt and Lamb, 2005; Gamblin and Skehel, 2010). L'activité de la neuraminidase est aussi requise pour couper les acides sialiques des particules virales afin d'éviter qu'elles ne s'agglutinent entres elles et diminuent la prolifération virale (Palese et al., 1974; Palese and Compans, 1976). Cette protéine est donc impliquée dans la propagation de l'infection virale.

Le canal à protons

Le canal à protons M2, est obtenu par épissage alternatif du segment 7. Il s'agit d'une protéine tétramérique présente à la surface du virion (figure 5C) (Holsinger and Lamb, 1991; Sugrue and Hay, 1991). M2 peut être divisée en plusieurs domaines: le domaine externe amino-terminal doté de propriétés antigéniques, le domaine transmembranaire central organisé en hélice-α et finalement le domaine C-terminal (Pielak and Chou, 2011). Ce dernier

est important dans le processus d'assemblage des nouveaux virions, notamment à travers l'interaction avec M1 qui sert d'interface entre la membrane virale et le génome (McCown and Pekosz, 2006).

Durant l'endocytose, M2 permet le passage des protons pour l'acidification de la particule virale (Cf.§II.4) (Lamb et al., 1985; Sugrue and Hay, 1991). De plus, M2 joue un rôle dans la formation des virions filamenteux (Rossman et al., 2010)

La protéine de matrice M1

La protéine M1 représente l'élément le plus abondant au sein des particules virales. (Schulze, 1972). Il s'agit d'une protéine structurale qui participe à l'architecture du virion. Elle tapisse en effet toute la surface interne de l'enveloppe virale. Les observations de microscopie électronique en coloration négative décrivent M1 comme un bâtonnet de 6 nm de long avec une face contre la membrane et une autre face en regard de la partie interne du virion (Ruigrok et al., 1989, 2000). La protéine est structurellement divisée en deux domaines globulaires composés d'hélices-α et reliés par une région sensible aux protéases (figure 5D) (Arzt et al., 2001; Harris et al., 2001; Sha and Luo, 1997). Chacun des domaines possède des propriétés biochimiques propres.

M1 sert d'interface entre les ribonucléoparticules du génome viral et la membrane où sont ancrées les protéines de surface. Sa partie amino-terminale interagit avec la membrane lipidique via des interactions électrostatiques (Ruigrok et al., 2000), tandis que le domaine carboxy-terminal interagit avec les ribonucléoparticules. Cette interaction ne se fait pas via l'ARNv, ce qui suggère une potentielle interaction de M1 avec la NP (Baudin et al., 2001). M1 pourrait jouer un rôle capital lors du recrutement et de l'assemblage des éléments de la particule virale (Nayak et al., 2009; Schmitt and Lamb, 2005). Pour cela elle interagirait avec NEP, permettant l'export nucléaire des ribonucléoprotéines (Cros and Palese, 2003), et la queue cytoplasmique de M2 (Chen and Lamb, 2008) en formant un pont entre les composants internes et les protéines membranaires de l'enveloppe virale. De plus, M1 joue un rôle actif dans la séparation terminale de la membrane lors de la phase de bourgeonnement (Nayak and Hui, 2004).

La nucléoprotéine

La NP est depuis longtemps connue pour être le composant majeur des ribonucléoprotéines (Pons et al., 1969). Elle est impliquée dans les processus de transcription et de réplication du génome (Huang et al., 1990; Kimura et al., 1992; de la Luna et al., 1993) et pourrait avoir un rôle clé dans la balance entre ces deux processus (Skorko et al., 1991). NP joue de nombreux rôles au sein des RNP. C'est notamment elle qui structure les RNP sous la forme d'une double hélice surenroulée en recouvrant la quasi-totalité de l'ARNv (Heggeness et al., 1982; Jennings et al., 1983; Ruigrok and Baudin, 1995) (Cf. § IV.1). NP interagit aussi avec la polymérase fixée aux extrémités 5' et 3' de l'ARNv (Albo et al., 1995; Biswas et al., 1998; Turan et al., 2004) et possède de nombreux partenaires cellulaires comme les protéines de transport nucléaire (Cf. § IV.3) ou encore l'actine (Mayer et al., 2007).

Quelque soit la souche virale, NP est toujours codée par le segment 5 du génome mais sa taille peut légèrement varier en fonction du type viral. Les NP d'*influenza* A (NP A) et d'*influenza* B (NP B) possèdent respectivement 498 et 560 acides aminés. Pour *influenza* C (NP C) NP possède 565 acides aminés et pour *influenza* D (NP D), elle en possède 552. Les NP des différents virus de la grippe présentent de fortes identités de séquences. Ainsi NP A présente 37 % d'identité avec NP B et environ 20 % d'identité avec NP C et NP D tandis que NP B possède respectivement 23 et 24 % pour NP C et NP D. Ces deux dernières possèdent 40 % d'identité de séquence et sont donc très similaires.

Au cours des 10 dernières années, 4 structures cristallographiques de NP ont été obtenues dont 3 pour *influenza* A (Chenavas et al., 2013a; Ng et al., 2008; Ye et al., 2006) et 1 pour *influenza* B (Ng et al., 2012). Toutes ces structures, obtenues sans ARN, ont permis de mettre en évidence la ressemblance architecturale entre les NP de différentes souches virales. De manière intéressante, les protéines ont été cristallisées sous la forme de trimères pour *influenza* A et de tétramères pour *influenza* B (figure 6). Cela souligne la tendance inhérente de NP à oligomériser en l'absence d'ARN.

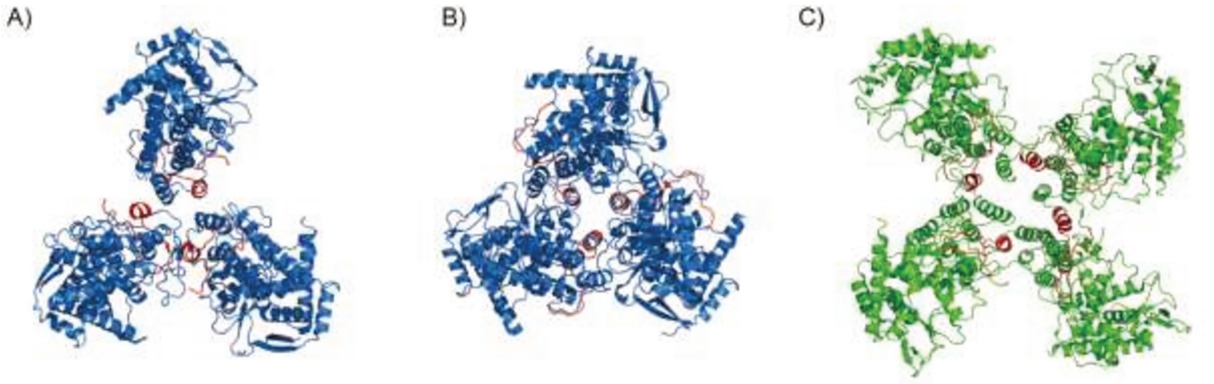


Figure 6 : Structures cristallographiques des nucléoprotéines d'*influenza* A et B. A) Trimère de NP A (PDB ID 2IQH) (Ye et al., 2006). B) Trimère du mutant monomérique de NP A (PDB ID 3ZDP) (Chenavas et al., 2013a). C) Tétramère de NP B (PDB ID 3TJ0) (Ng et al., 2012). Dans chacune des structures la boucle d'oligomérisation est représentée en rouge. Dans le cas de A et C) chaque NP lie une molécule voisine via sa boucle d'oligomérisation. Dans le cas de B), la boucle d'oligomérisation ne lie pas la NP adjacente mais est repliée au niveau du sillon de liaison à l'ARN de la même molécule.

Les structures obtenues permettent aussi de souligner les différences entre les NP des virus grippaux et de faire quelques hypothèses quant aux structures de NP C et NP D. Pour NP A, les 21 premiers acides aminés ne figurent pas dans la densité électronique ce qui suggère que son extrémité amino-terminale est désordonnée. Chez *influenza* B, ce sont les 72 premiers résidus qui sont désordonnés. A la différence de NP A, NP B possède donc une longue extrémité N-terminale. De plus, l'alignement de séquences des NP d'*influenza* A, B, C et D (Article 2) permet de mettre en évidence quelques différences majeures. NP A et NP B possèdent une extrémité N-terminale alors que NP C et NP D semblent posséder une longue extrémité C-terminale (figure 7). L'origine de ces différences n'a pas encore été établie. Cela pourrait être une conséquence de l'adaptation de chaque virus à son hôte, notamment en ce qui concerne la liaison aux différents partenaires cellulaires présents dans chaque hôte.

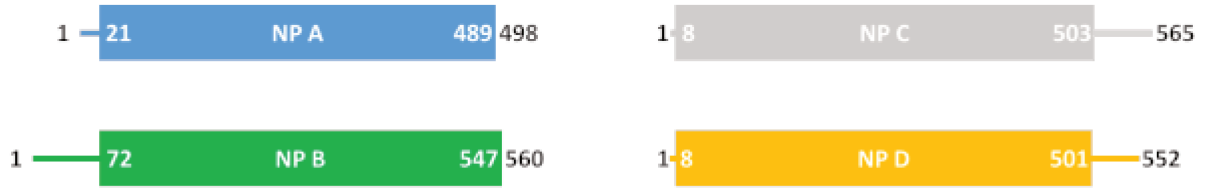


Figure 7 : Taille en acides aminés des nucléoprotéines des grippes de types A, B, C et D. Le rectangle représente le corps de la NP qui est conservé entre les différents types viraux. Les extrémités amino et carboxy-terminales sont représentées par un trait fin plus ou moins long en fonction du nombre d'acides aminés. NP A et NP B possèdent une extrémité N-terminale tandis que NP C et NP D possèdent une extrémité C-terminale.

Concernant sa structure, la NP est une protéine composée majoritairement d'hélices-α et repliée en deux domaines principaux : le domaine tête (résidus 150-272 ; 438-552 pour NP A et 212-323 ; 496-509 pour NP B) et le domaine corps (résidus 21-149 ; 273-396 ; 453-489 et 72-208 ; 336-404 ; 521-551 respectivement pour NP A et NP B - figure 8A). Les deux domaines étant constitués de régions discontinues de la chaîne polypeptidique, ils sont peu flexibles l'un par rapport à l'autre. Les structures obtenues permettent aussi de mettre en

évidence la présence d'un sillon basique entre le domaine tête et le domaine corps. Ce sillon représente le site potentiel de fixation de l'ARNv (Cf. § III.3). La tendance que possède la NP à oligomériser facilement, même en l'absence d'ARN, est due à la boucle d'oligomérisation (résidus 402-428 pour NP A et 459-486 pour NP B) permettant la liaison à la protéine adjacente. La différence majeure entre toutes les structures connues est l'orientation de cette boucle qui varie même entre différentes souches du type A (Chenavas et al., 2013a). Deux résidus sont absolument indispensables au processus d'oligomérisation : chez NP A il s'agit de l'arginine 416 située sur la boucle d'un premier monomère et du glutamate 339 situé au niveau du domaine corps d'un second monomère (figure 8B). Dans le cas du mutant monomérique, où l'Arginine 416 est mutée en alanine, il y a perte totale d'oligomérisation quelques soient les conditions physico-chimiques (Chenavas et al., 2013a). Les acides aminés correspondants à ceux de la boucle d'oligomérisation sont alors fixés sur le même monomère en établissant des interactions très similaires à celles observées dans le cas du trimère de NP A.

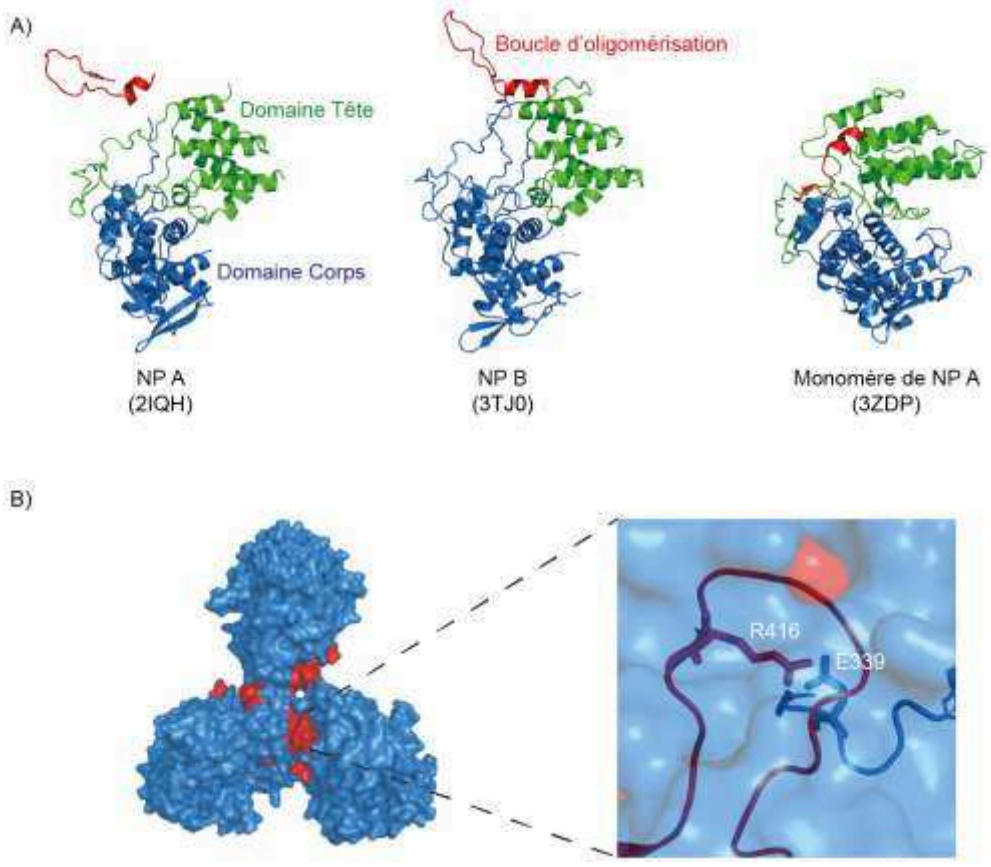


Figure 8 : Les différents domaines structuraux de la nucléoprotéine du virus de la grippe. A) Comparaison des structures de NP A et NP B. Les nucléoprotéines des virus grippaux présentent toutes un repliement similaire en deux domaines principaux, le domaine tête et le domaine corps. La boucle d'oligomérisation est orientée différemment en fonction des souches. Pour le mutant monomérique, la boucle d'oligomérisation est repliée sur le même monomère et ne lie donc pas de NP voisine. B) Représentation schématique de l'oligomérisation entre monomères de NP. La boucle d'oligomérisation (en rouge) vient fixer le domaine corps d'une NP adjacente. Les résidus R416 et E339 de NP A sont indispensables au processus d'oligomérisation.

La propriété d'oligomérisation a été décrite comme dynamique et influencée par la concentration en sel ainsi que par la température (Ruigrok and Baudin, 1995; Tarus et al., 2012; Ye et al., 2012). A température ambiante et faible concentration en sel (10 mM NaCl), la protéine est monomérique alors qu'elle oligomérise lorsque la concentration en sel augmente (300 mM NaCl) (figure 9). Cependant, à 4°C, pour une concentration en sel élevée (300 mM), la protéine a tendance à rester monomérique. Le processus d'oligomérisation peut aussi être initié par la fixation de la protéine à l'ARN (Cf. § IV.3).

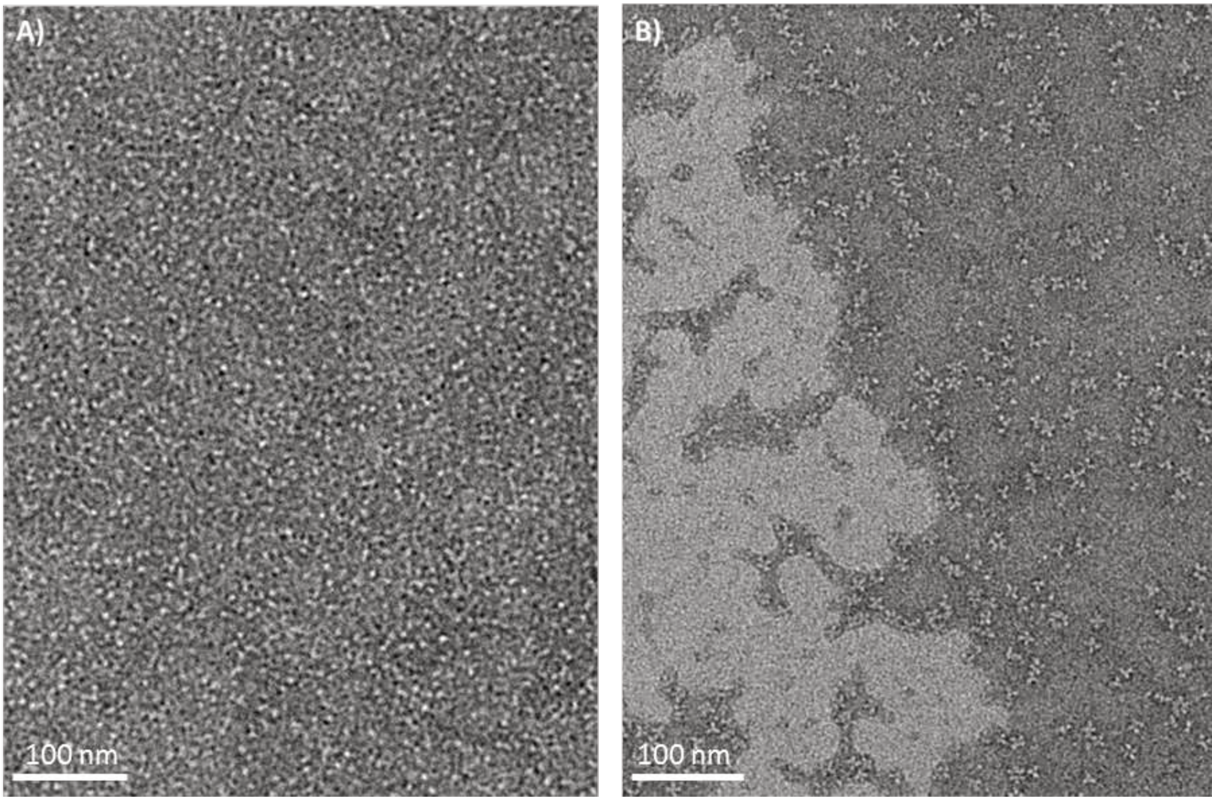


Figure 9 : Image de microscopie électronique par coloration négative de la nucléoprotéine du virus *influenza* A. A) Image du monomère de nucléoprotéine à 50 mM NaCl. B) Image des trimères et tétramères de la nucléoprotéine à 300 mM NaCl. (Images prises sur la plateforme de microscopie de l'IBS par Alice Labaronne).

Un autre élément de régulation potentiel de l'oligomérisation est la phosphorylation de la NP. Des études montrent différents types de phosphorylation de la NP en fonction des souches et du moment d'infection (Kistner et al., 1985, 1989). Plus récemment il a été montré que la mutation de certains résidus phosphorylés (S165A, S407A, S486A chez NP A) pouvait interrompre à la fois l'oligomérisation et la fixation à l'ARNv (Chenavas et al., 2013a; Mondal et al., 2015). De même, la phosphorylation induite lors de la culture cellulaire peut promouvoir l'oligomérisation des NP purifiées (Mondal et al., 2015).

Le complexe polymérase du virus de la grippe

L'ARN-polymérase du virus de la grippe est une enzyme multifonctionnelle nécessaire à la transcription et la réplication au cours du cycle viral. Elle est composée de 3 sous-unités: PB1 et PB2, des protéines basiques, et PA, une protéine acide. L'activité polymérase est répartie au sein de ces trois sous-unités. PB1 est indispensable au processus de transcription et de réplication (Mahy, 1983) : elle possède l'activité polymérase (Digard et al., 1989). PB2 joue un rôle essentiel dans l'initiation de la transcription (Mahy, 1983; Plotch et al., 1981). C'est cette sous-unité qui reconnaît la coiffe m7Gppp située en 5' des ARNm cellulaires (Blaas et al., 1982; Ulmanen et al., 1981). PB2 est aussi impliquée dans le processus de réplication (Gastaminza et al., 2003; Jorba et al., 2009). Concernant PA, son activité endonucléase lui permet de couper la coiffe des ARNm cellulaires (Gastaminza et al., 2003; Jorba et al., 2009). Tout comme les deux autres sous-unités, PA est essentielle au processus de réplication (Honda et al., 2002; Krug et al., 1975; Mahy, 1983) et de transcription (Fodor et al., 2002; Regan et al., 2006).

Les trois sous-unités protéiques de la polymérase s'assemblent tête-bêche dans l'ordre PA-PB1-PB2. L'extrémité carboxy-terminale d'une sous-unité interagit avec l'extrémité amino-terminale de la suivante.

Récemment, une structure à haute résolution des polymérases d'*influenza* B (Reich et al., 2014) et d'*influenza* A H17N10 (Pflug et al., 2014) ont été obtenues par cristallographie. Les deux polymérases possèdent une architecture globale très similaire. La RdRp du virus de la grippe adopte une configuration en forme de U avec deux pouces, correspondant respectivement au domaine endonucléase de PA et au domaine de fixation de la coiffe de PB2. Les deux domaines se font face et constituent l'entrée d'un tunnel important où ont lieu la réplication et la transcription. Les deux structures ont été obtenues en présence des séquences promotrices du 5' et du 3' de l'ARNv.

Au sein de ce complexe (figure 10), PB1 se présente sous la forme d'une main droite humaine repliée successivement en « doigt, paume, pouce ». PB2 est repliée en trois domaines répartis sur l'ensemble du complexe de la polymérase. La première région amino-terminale consiste en une série d'hélices-α et de feuillets-β imbriqués avec le domaine carboxy-terminal de PB1. PB2 interagit aussi avec le domaine endonucléase de PA via son extrémité amino-terminale. Le second domaine de PB2 est impliqué dans la fixation de la coiffe et repose sur le dernier domaine, situé du côté carboxy-terminal de la protéine. Au sein de la polymérase, PA est étalée sur la totalité du complexe RdRp. PA est repliée en deux domaines éloignés l'un de l'autre : un petit domaine amino-terminal portant l'activité

endonucléase et un plus gros domaine qui est enroulé contre PB1. Ces deux régions sont reliées par un long domaine de liaison appelé domaine charnière.

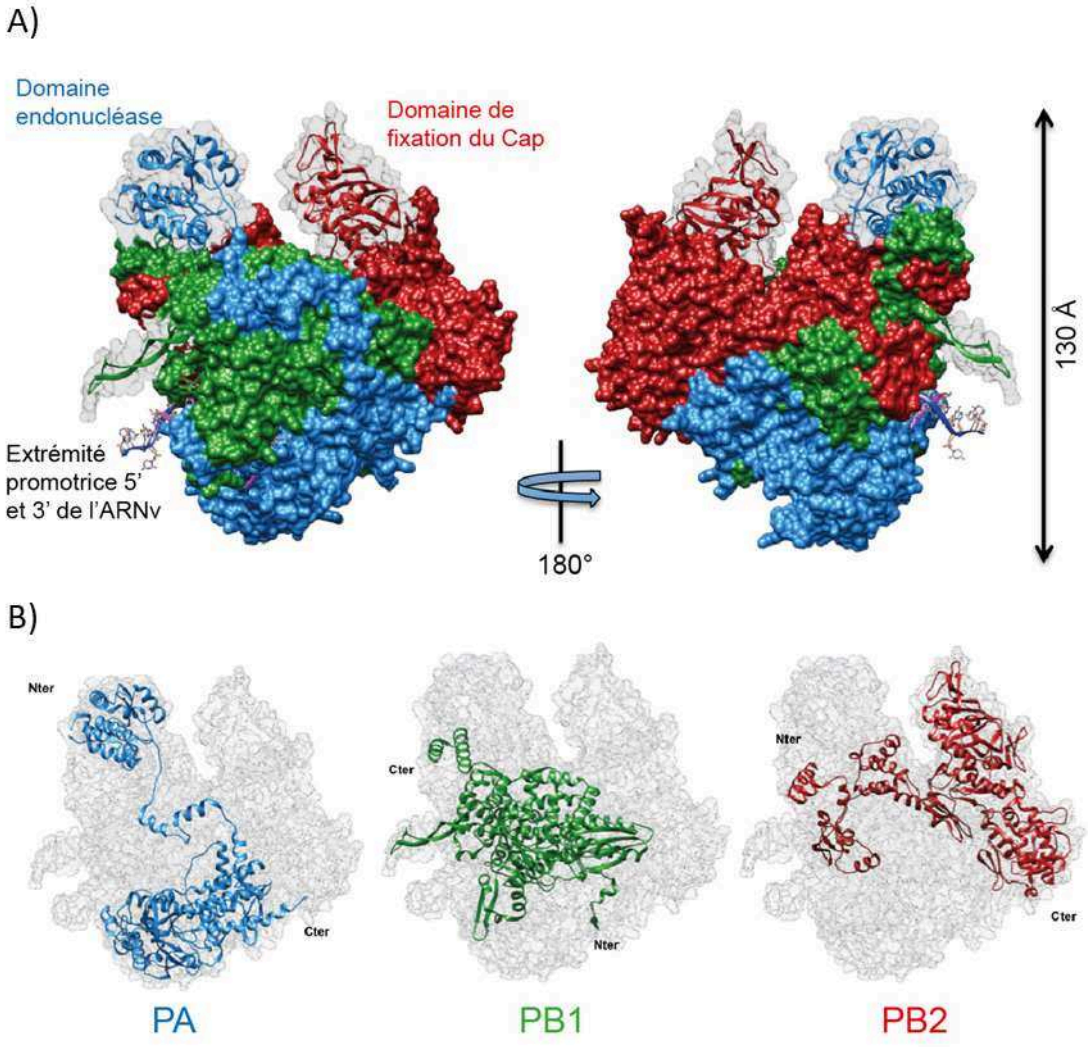


Figure 10 : Structure cristallographique de la polymérase du virus de la grippe A/H17N10. A) Surface de la polymérase représentée avec PA en bleu, PB1 en vert et PB2 en rouge. Le domaine endonucléase de PA et le domaine de fixation du cap de PB2 sont représentés en cartoon. Le segment 5' de l'ARNv est représenté en magenta et le segment 3' en bleu. B) Représentation Cartoon des différentes sous-unités au sein du complexe de la polymérase. (PDB ID 4WSA) (Adaptée de la thèse de Christopher Swale, 2015).

La protéine non structurale NS1

NS1 est une protéine qui n'est pas présente au sein de la particule virale mais qui est seulement exprimée dans les cellules infectées. Il s'agit d'une protéine dimérique (figure 5G) (Hale et al., 2008) pouvant être divisée en deux domaines séparés par une région non structurée (Carrillo et al., 2014). Le domaine amino-terminal est dédié à la reconnaissance et à la fixation de l'ARN double brin (Cheng et al., 2009; Liu et al., 1997). Des études biochimiques (Nemeroff et al., 1995) et structurales (Liu et al., 1997; Cheng et al., 2009) ont

montré qu'un procédé de dimérisation est requis pour la fixation à l'ARN. Le domaine carboxy-terminal de NS1, est impliqué dans la fixation des facteurs cellulaires et stabilise le domaine de fixation à l'ARN (Bornholdt and Prasad, 2006; Hale et al., 2008, 2010a; Kerry et al., 2011). L'interaction fonctionnelle entre les deux domaines de NS1 n'est pas encore bien comprise. Chacun des domaines a tendance à dimériser à travers différentes interfaces, suggérant une multitude d'états d'oligomérisation (Kerry et al., 2011).

NS1 est une protéine multifonctionnelle impliquée dans de nombreux mécanismes qui favorisent la prolifération virale. Un de ses rôles essentiels concerne l'inhibition de la réponse immunitaire de l'hôte. Pour se faire, elle agit de différentes façons. Par exemple, en inhibant la réponse à l'interféron de type 1 (IFN-α/IFN-β) (Hale et al., 2010b). Durant l'infection, un des mécanismes de défense de l'organisme, est la production de cytokine (Type I IFN). Ce mécanisme est induit par des capteurs moléculaires tels que les récepteurs Toll-like (TLRs) reconnaissant les ARNv exogènes (Kawai and Akira, 2011). Pour la grippe, il s'agit plus spécifiquement du récepteur RIG-I like (Yoneyama et al., 2004). Après activation de ces récepteurs, la réponse IFN entraîne la sur-activation de plus de 300 gènes antiviraux. (Haller et al., 2009; Randall and Goodbourn, 2008). Pour inhiber cette cascade de cytokines, NS1 est exprimée en très grande quantité au sein du cytoplasme et rentre en compétition avec RIG-I en fixant les ARN doubles brins (Liu et al., 1997). Il ne s'agit ici que d'un des nombreux mécanismes de défense de NS1 pour contrer la réponse de l'hôte.

Au sein du noyau, NS1 module le transcriptome des gènes cellulaires en inhibant la maturation des ARNm néo-synthétisés de l'hôte. Pour cela elle interagit avec le complexe « cleavage and polyadenylation specificity factor », CPSF (Nemeroff et al., 1998) ainsi qu'avec la « poly A binding protein II » (PABP II) (Chen et al., 2001). La protéine NS1 serait également impliquée dans l'augmentation de la traduction des ARNm viraux sans pour autant influencer celle des ARNm cellulaires (Enami et al., 1994; de la Luna et al., 1995).

La protéine d'export nucléaire NEP

NEP, aussi connue sous le nom de NS2, est transcrite à partir de l'épissage alternatif du segment 8. La protéine NEP est présente en plusieurs exemplaires dans la particule virale (Richardson and Akkina, 1991; Yasuda et al., 1993). NEP possède deux domaines : un domaine amino-terminal prédit comme majoritairement désordonné et un domaine carboxy-terminal (figure 5H) qui a été cristallisé (Akarsu et al., 2003). Ce dernier est replié en 2 hélices-α antiparallèles. Chaque hélice génère une structure amphiphile qui pourrait interagir avec M1 (Akarsu et al., 2003).

NEP est impliquée dans l'export nucléaire des RNPs (O'Neill et al., 1998; Neumann et al., 2000; Cros and Palese, 2003). Des études récentes ont aussi mis en évidence le potentiel rôle de NEP dans la régulation de la transcription et de la réplication virale (Bullido et al., 2001; Mänz et al., 2012; Paterson and Fodor, 2012). Dans le cytoplasme, NEP contribue au bourgeonnement à travers l'interaction avec l'ATPase cellulaire (Gorai et al., 2012).

Les protéines accessoires

D'autres protéines peuvent être retrouvées dans la particule virale, notamment chez *influenza* A. Elles ne sont pas présentes dans tous les types A, B, C et D, ni même dans toutes les souches d'un même type. Ces protéines ont des rôles accessoires dont certains ne sont pas encore connus. Pour les virus de type A, elles peuvent réprimer la transcription des ARNm par la polymérase II cellulaire (PA-X) (Jagger et al., 2012) ou encore être impliquées dans l'apoptose (PB1-F2) permettant alors au virus d'échapper au système immunitaire (Chen et al., 2001; Henkel et al., 2010). D'autres protéines virales peuvent aussi favoriser la réplication virale (NS3) et avoir un lien avec les mécanismes d'adaptation à l'hôte (Selman et al., 2012).

4. Le cycle viral

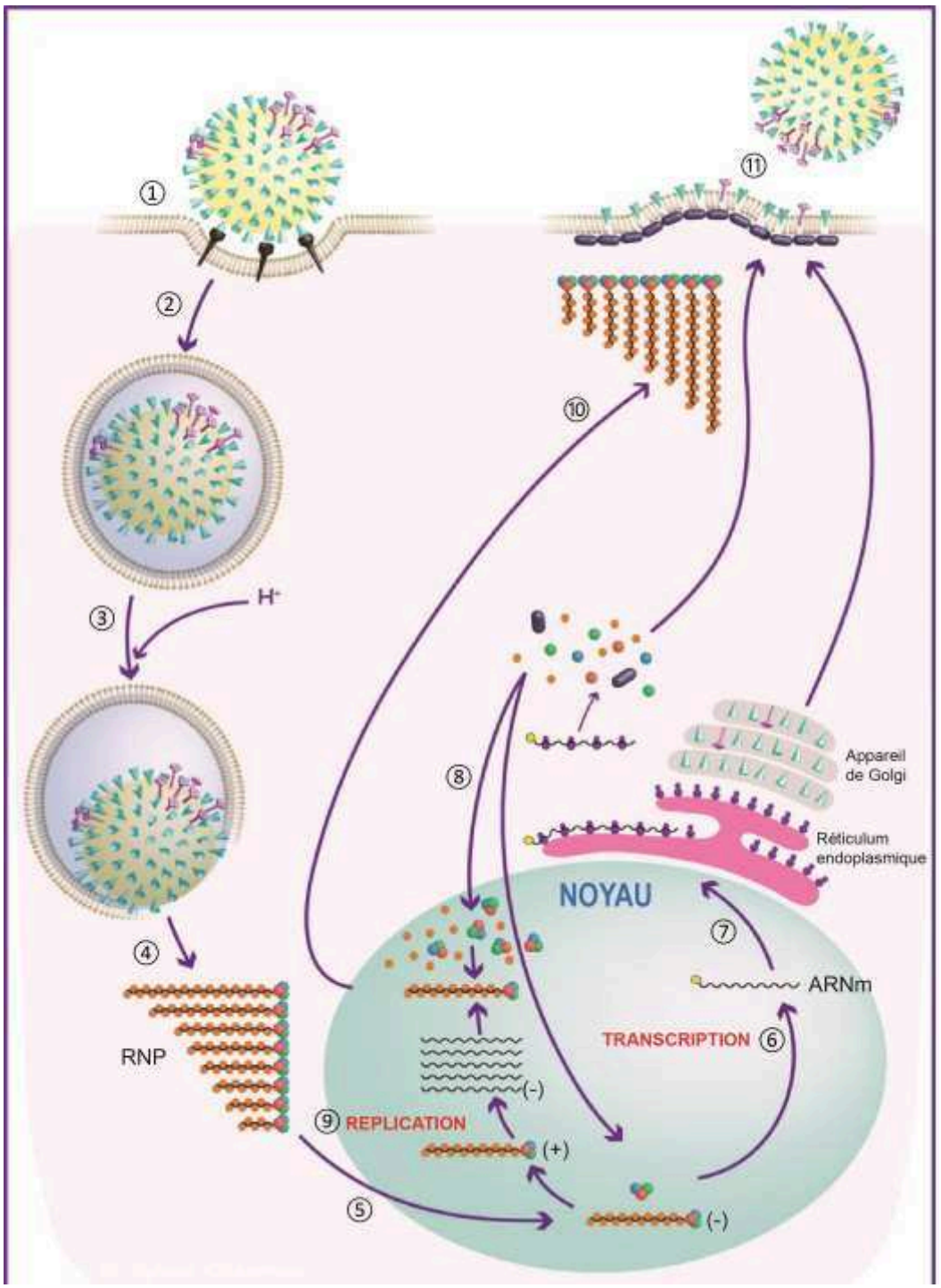


Figure 11 : Le cycle viral du virus de la grippe. ① La Liaison de la particule virale à la cellule, médiée par la reconnaissance des acides sialiques par HA, conduit à l'internalisation de l'endosome ②. L'acidification de l'endosome ③ provoque la fusion des membranes virale et endosomale, ce qui permet le relargage des RNP dans le cytosol ④. Les RNP sont transportées jusqu'au noyau ⑤ où ont lieu la transcription ⑥ et la réplication ⑨. Les ARNm sont exportés dans le cytoplasme pour être traduits en protéine par la machinerie cellulaire ⑦. L'ARN-polymérase et la NP sont importées au noyau pour former de nouvelles RNP ⑧. Les RNP néo-synthétisées sont exportées jusqu'à la membrane plasmique ⑩ où ont lieu l'assemblage et le bourgeonnement ⑪. (Adaptée de (Guilligay et al., 2010).

Le cycle viral de la grippe peut être divisé en 5 étapes principales (figure 11).

L'entrée du virus: attachement, endocytose, fusion et libération du génome

Le virus de la grippe a pour cible les cellules épithéliales des systèmes respiratoire et digestif. Le cycle d'infection débute par la liaison spécifique de la particule virale à la cellule ①. HA se lie aux acides sialiques des glycoprotéines ou glycolipides cellulaires. En fonction des souches virales elle ne reconnait pas les mêmes acides sialiques Saα2,6 (pour les souches humaines) et/ou Saα2,3 (pour les souches aviaires) (Connor et al., 1994; Viswanathan et al., 2010). La faible interaction (de l'ordre de 10^{-3} mM) est compensée par la forte densité de récepteurs cellulaires et d'hémagglutinines virales. La fixation aux récepteurs conduit à l'internalisation du virus par endocytose ②. Le faible pH de l'endosome (pH 5-6) induit un changement de conformation de HA (Cf. § II.3 l'hémagglutinine), conduisant à la fusion des membranes virale et endosomale ③. Lors de cette étape, le canal ionique M2 permet l'afflux de protons depuis l'endosome dans la particule virale. Cette acidification provoque aussi un changement conformationnel des protéines M1 qui a pour effet de rompre les interactions entre elles. Les RNP se retrouvent alors libres dans le cytoplasme ④ (Matlin et al., 1981).

Migration des particules ribonucléoprotéiques virales dans le noyau cellulaire

Les RNP sont ensuite transportées jusqu'au noyau ⑤ où ont lieu la réplication et la transcription (Harrison, 2008b; Gamblin and Skehel, 2010). Du fait de leur taille importante (environ 15 nm de large), les RNP ne peuvent pas diffuser à travers la membrane nucléaire (Compans et al., 1972; Martín-Benito et al., 2001). Le virus détourne alors la machinerie nucléocytoplasmique des cellules hôtes pour importer les RNP au noyau où se déroule la synthèse des ARNv (Boulo et al., 2007; Cros and Palese, 2003).

La transcription et la réplication du génome viral

A l'intérieur du noyau, les trois sous-unités de la polymérase, la NP et l'ARNv sont les molécules indispensables aux étapes de transcription ⑥ et de réplication ⑨. Durant ces deux processus, trois types d'ARN sont synthétisés : les ARN génomiques complémentaires (ARNc) de polarité positive, permettant la synthèse d'ARNv de polarité négative et les ARNm viraux directement synthétisés à partir de l'ARNv. La réplication correspond à la phase d'amplification du génome viral alors que la transcription correspond à l'étape de synthèse d'ARNm viraux (Cf. § IV.2). Une fois produits, les ARNm sont exportés ⑦ vers le cytoplasme pour être traduits en protéines. Les trois sous-unités de l'ARN-polymérase ainsi que la NP

néo-synthétisée sont importées ⑧ au noyau pour former de nouvelles RNP avec les ARNv issus de la réplication. Les autres protéines grippales néo-synthétisées sont directement adressées à la membrane plasmique où se déroule l'assemblage final pour former de nouveaux virions.

L'export nucléaire des RNPv

Une fois synthétisées, les RNP vont être transportées à l'extérieur du noyau, au niveau de la membrane plasmique ⑩ (Shapiro et al., 1987) où a lieu le processus de bourgeonnement cellulaire. L'export des RNP repose sur la machinerie d'export du noyau et la protéine M1 ainsi que NEP/NS2 sont impliquées dans cet export (Boulo et al., 2007; Cros and Palese, 2003; Elton et al., 2001; Neumann et al., 2000; O'Neill et al., 1998). M1 interagirait à la fois avec les RNPv néo-synthétisées et avec NEP qui recrute la machinerie d'export par son interaction avec Crm1 (protéine d'export nucléaire, *Chromosome Region Maintenance 1*) permettant l'export des RNPv vers le cytoplasme (Hutchinson and Fodor, 2013).

L'emballage des segments dans la particule virale

Un assemblage correct des 8 segments d'ARN génomique est essentiel pour qu'un virion soit infectieux. Depuis une dizaine d'années, de nombreuses études fournissent des arguments génétiques et biochimiques en faveur d'une incorporation sélective des RNPv dans la particule virale (Gerber et al., 2014; Hutchinson et al., 2010). Il a d'ailleurs été montré par hybridation in situ en fluorescence que la majorité des particules virales ne possèdent qu'une seule copie de chaque ARN (Chou et al., 2012). Cette incorporation sélective reposerait sur un mécanisme opérant à deux niveaux. D'une part, les extrémités non codantes U12 et U13 permettraient la discrimination des RNPv (Luytjes et al., 1989; Muramoto et al., 2006). D'autre part, les séquences non codantes spécifiques à chaque segment ainsi qu'une partie des régions codantes en 3' et 5' serviraient de signal d'incorporation sélective pour les RNPv (Goto et al., 2013). Il semblerait que certains ARNv jouent un rôle plus important que d'autres dans le phénomène d'incorporation et cela suggère qu'il s'agit donc d'un procédé organisé de manière hiérarchique (Gao et al., 2012; Hutchinson et al., 2008, 2009; Marsh et al., 2008). Bien que l'existence de segment d'incorporation spécifique ait été clairement démontrée, les bases moléculaires de ce processus restent encore floues. Il impliquerait la création d'un réseau supramoléculaire complexe maintenu par des interactions entre les RNPv (Hutchinson et al., 2010), au sein duquel il existerait des interactions directes (de type Watson-Crick) entre les segments

d'ARNv (Fujii et al., 2003). Plusieurs étude *in vitro,* (Fournier et al., 2012a, 2012b; Gavazzi et al., 2013a) montrent notamment que les segments d'ARNv du virus sont impliqués dans un unique réseau d'interaction au sein duquel chaque ARNv interagit avec au moins un autre ARNv. De plus une étude a montré que ces interactions, notamment celles existant entre les ARNv 2 et 8 sont essentielles au processus de réplication et d'incorporation sélective des ARNv dans les virions néoformés (Gavazzi et al., 2013b). Pour qu'un tel modèle d'incorporation soit possible, il est nécessaire que certaines régions de l'ARNv restent accessibles afin que le réseau d'interaction ARNv-ARNv puisse se faire. Au sein des virions, des études de microscopie (Arranz et al., 2012; Moeller et al., 2012) montrent que les RNPv adoptent une structure en double hélice avec deux brins antiparallèles (Cf. § IV.1). Au sein de ces structures l'ARN reste accessible car la liaison des NP aux phosphates des ARNv permet l'exposition des bases nucléotidiques au solvant. Le nombre de nucléotides fixés par une NP au sein des RNP reste toujours inconnu. Cependant une récente étude suggère que la NP de l'anémie infectieuse du saumon ne fixerait que 12 nucléotides d'ARN (Zheng et al., 2013) et non 20 - 24 comme décrit précédemment (Compans et al., 1972; Martín-Benito et al., 2001; Ortega et al., 2000) (Cf. § IV.3). La séquence ARN restante pourrait adopter une structure secondaire permettant la formation d'un complexe supramoléculaire de RNPv médié par des interactions ARN-ARN spécifiques. Les processus d'incorporation des segments d'ARNv semblent être en relation directe avec les processus de réassortiment et une meilleure compréhension de l'incorporation permettrait de mieux comprendre les réassortiments entre différentes souches virales.

Bourgeonnement et relargage des particules virales

Les RNP sont formées et adressées à la membrane plasmique. HA, NA et M2 sont aussi adressées aux sites apicaux de la membrane plasmique qui sont riches en radeaux lipidiques (Scheiffele et al., 1999; Nayak and Hui, 2004). M1 interagit avec toutes les queues cytoplasmiques de NA, HA et M2 et les ancres dans la nucléocapside (Rossman and Lamb, 2011). Le bourgeonnement ⑪ des particules néoformées dépend principalement de la neuraminidase (Cf. § II.3) qui clive la liaison entre l'acide sialique et l'hémagglutinine.

III. Evolution du virus

Le virus influenza infecte une très large variété d'hôtes au sein desquels il est à la fois capable de détourner la machinerie cellulaire à son avantage mais aussi de contrer de nombreux mécanismes de défense. Il peut aisément échapper au système immunitaire de l'hôte grâce à un processus d'adaptation impliquant l'accumulation de mutations au cours du temps. Il s'agit de la dérive génétique. De plus, le virus possède une capacité particulière à passer d'une espèce à une autre grâce à des réassortiments de segments d'ARNv entre virus lors de la co-infection d'une cellule par plusieurs virus. C'est la propriété de réassortiment génétique.

1. La dérive génétique

La dérive génétique correspond à l'accumulation de mutations au cours du temps. Ces évolutions du virus par mutations sont appelées « glissements ». C'est un procédé lent permettant un processus d'adaptation constant du virus afin de contrer l'immunité acquise de la cellule. Cette dérive est causée par les mutations ponctuelles survenant au sein du génome viral du fait de la faible fidélité de la RdRp ainsi qu'à son absence d'activité correctrice. Il est estimé que la polymérase fait environ 1 erreur sur 10 000 nucléotides au cours de la réplication (Boivin et al., 2010; Drake, 1993). Chacun des virions contient donc au moins 1 mutation (Stray and Air, 2001). Les virions viables possédant une mutation sur HA ou NA ne sont donc plus reconnus par les anticorps neutralisants et continuent les cycles d'infection en échappant facilement au système immunitaire (Bush et al., 1999; Plotkin and Dushoff, 2003). La dérive antigénique est un phénomène crucial pour le virus permettant ainsi de maintenir des souches épidémiques en circulation. Ces souches échappent facilement aux formulations des vaccins.

2. Le réassortiment génétique

C'est la nature segmentée du génome viral qui est à l'origine du réassortiment des brins d'ARNv entre virus. Ce phénomène, appelé « cassure », se produit à la suite d'une co-infection d'une même cellule par différents virus. Les différents segments génomiques de chaque virus sont alors redistribués aléatoirement au sein des nouveaux virions. Ces réassortiments sont observés au sein d'un même type viral mais jamais entre différents types. Sur le plan antigénique ces mécanismes peuvent conduire à des changements radicaux ayant pour conséquences de générer de nouveaux motifs antigéniques de HA ou NA. Ces motifs diffèrent de ceux présents dans les souches en circulation et peuvent

potentiellement augmenter l'infectivité virale du fait que la population est immunologiquement naïve. Ce mécanisme imprédictible peut donc conduire à de nouvelles pandémies (Schrauwen et al., 2014; Wright, P.F et al., 2013). Il est d'ailleurs impliqué dans les 4 dernières pandémies (figure 12). La pandémie de 1918 serait due à un échange de segments génomiques entre un virus aviaire et un virus humain. Lors de la pandémie de 1957, le virus H2N2 d'origine aviaire aurait échangé ses segments génomiques HA/PB1/NA avec la souche humaine H1N1. En 1968, une souche aviaire H3N? (NA inconnue) a échangé ses segments HA/PB1 avec H2N2 issue de 1957 (Schrauwen et al., 2014; Webster et al., 1992; Yoon et al., 2014). La pandémie H1N1 de 2009 a, quant à elle, été générée par de multiples réassortiments entre virus humain, aviaire et porcin (Munier et al., 2010). Ces réassortiments affectant les humains ne sont qu'une fraction des réassortiments existant dans la nature.

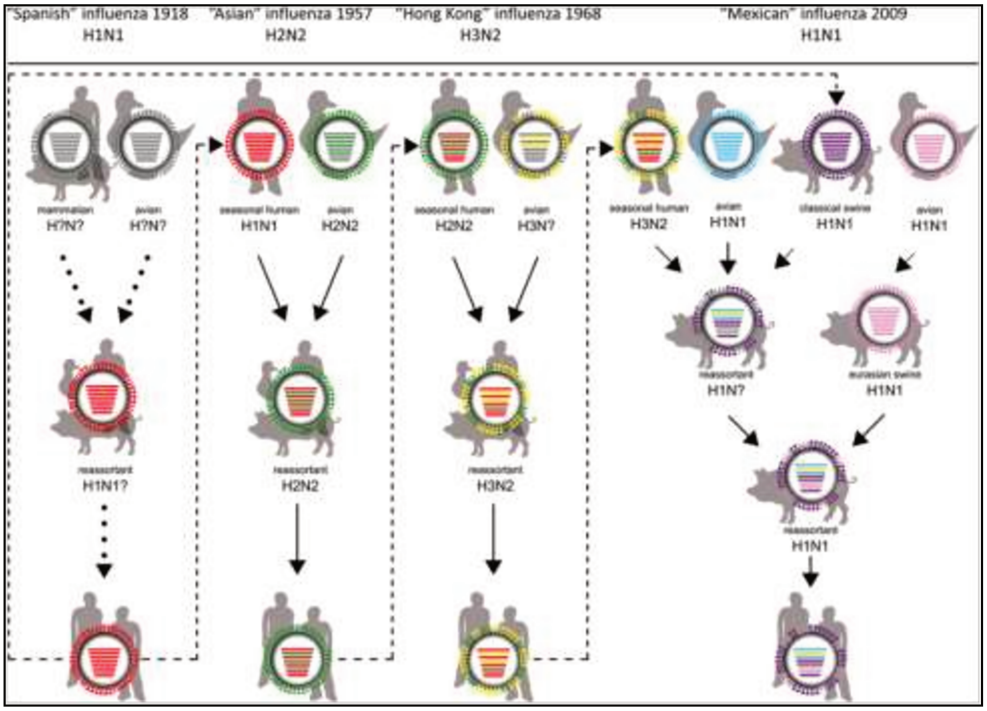


Figure 12 : Evènements de réassortiments génétiques ayant conduit à des pandémies de grippe A (Adaptée de Schrauwen et al., 2014).

IV. La ribonucléoprotéine, machinerie virale de synthèse de l'ARN

1. Structure de la ribonucléoprotéine

Chaque segment d'ARNv existe au sein du complexe de la RNP dans lequel l'ARN est recouvert de multiples copies de nucléoprotéine et a ses extrémités 5' et 3' liées à la polymérase (Murti et al., 1988). Au sein des virions purifiés, les RNP possèdent une structure flexible en forme de ruban fermé et surenroulé (figure 13A) (Compans et al., 1972; Heggeness et al., 1982; Pons et al., 1969). La polymérase est responsable de la fermeture du ruban car elle fixe à la fois les extrémités 3' et 5' conservées de chaque segment d'ARNv (Cf. § II.1) (Compans et al., 1972; Klumpp et al., 1997). Les RNP sont des structures d'une quinzaine de nanomètre de largeur (Ruigrok and Baudin, 1995) et possèdent une longueur variable comprise entre 30 et 110 nm en fonction de la taille du fragment d'ARNv recouvert (Compans et al., 1972; Heggeness et al., 1982; Pons et al., 1969). Leur forme varie en fonction des courbures présentes sur leur longueur (Gallagher et al., 2016). Bien que les RNP regroupent ARN, NP et polymérase, il est aussi possible d'observer des structures similaires sous certaines conditions dans lesquelles la NP purifiée à partir de virus est libre de tout ARN (figure 13C) (Ruigrok and Baudin, 1995). Ces observations laissent à penser que la forme des RNP serait donc majoritairement due aux propriétés d'homo-oligomérisation de la NP (Prokudina-Kantorovich and Semenova, 1996; Tarus et al., 2012). Récemment, une étude a montré que la flexibilité des RNP serait d'ailleurs due à la plasticité des interactions existantes entre nucléoprotéines. En effet, au sein même des RNP, des cours segments de NP pourraient adopter une conformation locale sous forme d'anneau provoquant alors une rupture dans leur structure en ruban (Gallagher et al., 2016). Cette plasticité est indispensable aux RNP pour accomplir leurs multiples fonctions telles que la synthèse d'ARN, l'import/l'export nucléaire ou encore l'assemblage du virus (Eisfeld et al., 2015).

Au sein des RNP, l'ARNv reste sensible aux RNAses (Duesberg, 1969) et peut facilement être déplacé par le polyvinylsulfate, un polymère chargé négativement (Goldstein and Pons, 1970). Cela supporte l'idée que l'ARNv est enroulé autour de la face externe de la NP avec les bases exposées au solvant pour qu'elles restent accessibles à la polymérase sans casser la structure de la RNP (Baudin et al., 1994).

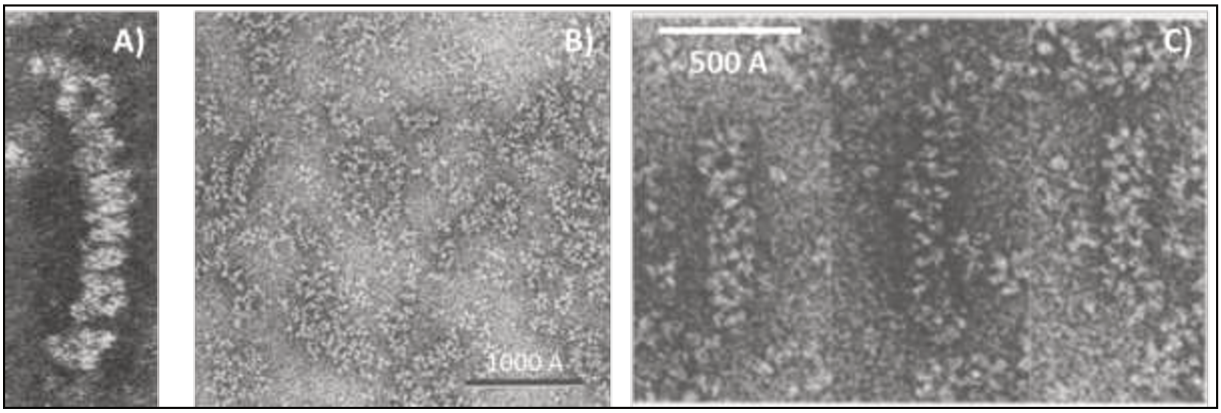


Figure 13 : La ribonucléoprotéine et la nucléoprotéine du virus de la grippe A. A) Ribonucléoprotéines purifiées à partir de virus. La boucle présente à l'extrémité supérieure de la RNP témoigne de la structure fermée de la particule (adaptée de Compans et al., 1972) B) Nucléoprotéines purifiées et libres d'ARN. Plusieurs espèces oligomériques, dont des structures similaires aux RNP sont présentes sur l'image. C) NP libre de tout ARN formant des structures similaires aux RNP (adaptée de Ruigrok and Baudin, 1995).

Il y a quelques années, une structure à basse résolution d'une RNP recombinante a été obtenue par coloration négative en microscopie électronique (Martín-Benito et al., 2001; Ortega et al., 2000). Des mini-RNP ont été reconstituées in vivo à partir des trois sous-unités de la polymérase, de NP et d'un court brin d'ARN. Cette étude montre qu'en fonction de la taille de l'ARN, les RNP peuvent adopter des structures circulaires ou elliptiques. Une structure tridimensionnelle (3D) de ces mini-RNP a été obtenue en cryo-microscopie électronique (figure 14A) (Coloma et al., 2009). Ces RNP recombinantes contenant une version tronquée de l'ARN sont capables de mener à bien le processus de transcription (Resa-Infante et al., 2011). La structure 3D montre que les mini-RNP contiennent 9 monomères de NP dont deux connectés aux sous-unités PB1 et PB2 de la polymérase. Une résolution de 12 Å a été obtenue pour la NP au sein de l'anneau et de 18 Å pour la NP adjacente à la polymérase. La structure cristallographique de la NP (Ye et al., 2006) peut être facilement placée dans le volume de chaque monomère, ce qui suggère que sa structure est bien préservée au sein des RNP (figure 14B) (Coloma et al., 2009). Cependant, la faible résolution ne permet pas de discerner l'ARN.

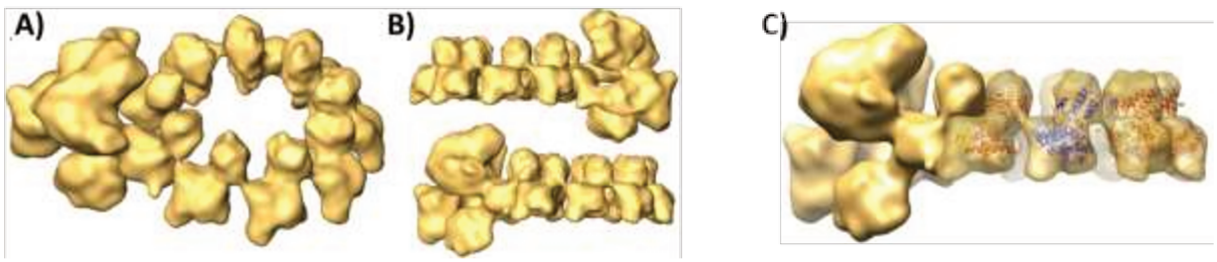


Figure 14 : Modèles 3D d'une mini-RNP recombinante. A) vue de perspective et B) vues de côté. 9 nucléoprotéines sont assemblées pour former une structure circulaire. Deux des NP sont en contact direct avec la polymérase. C) La structure cristallographique de la NP se positionne parfaitement à l'intérieur du modèle. Cela témoigne du fait que sa structure est bien préservée au sein des ribonucléoprotéines (adaptée de Coloma et al., 2009).

Des reconstructions de RNP recombinantes (Moeller et al., 2012) et directement issues de virions (figure 15) (Arranz et al., 2012) ont été obtenues par cryo-microscopie électronique. Ces reconstructions montrent que les RNP sont organisées sous la forme d'une double hélice en épingle à cheveux faite de deux brins de NP de polarité opposée (figure 15A). Arranz et al, définissent l'existence de deux sillons au niveau des RNP : le sillon majeur sans connexion entre les brins et le sillons mineur maintenu par des interactions entre deux NP appartenant chacune à un brin de l'épingle à cheveux (figure 15A et 15B) (Arranz et al., 2012; Ye et al., 2012). Au sein d'un même brin, les NP interagissent via la boucle d'oligomérisation (figure 15B). Quant à elle, la polymérase est située à l'extrémité ouverte de l'épingle à cheveux et interagit simultanément avec les extrémités 3' et 5' de l'ARNv (figure 15A). Au niveau de cette extrémité, la polymérase interagit aussi avec la NP, notamment via des interactions avec PB1 et PB2 (Albo et al., 1995; Biswas et al., 1998; Marklund et al., 2012; Turan et al., 2004). L'extrémité fermée des RNP adopte la structure d'une petite boucle formée par 3 à 5 NP (figure 15A). Ces deux études ont laissé des questions quant au sens de la double hélice mais récemment les travaux de Gallagher et al ont montré qu'il s'agirait plutôt d'une hélice droite (Gallagher et al., 2016). Des différences de structures entre les deux reconstructions laissent les questions ouvertes pour des analyses structurales de plus haute résolution.

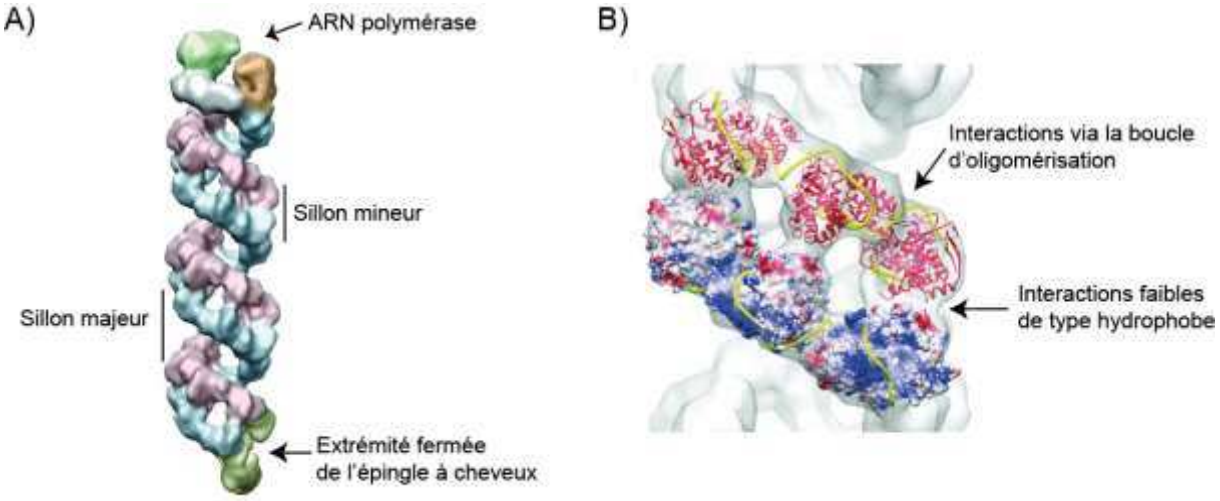


Figure 15: Structure de la ribonucléoprotéine du virus de la grippe A. A) Modèle 2D obtenu à partir des images de cryo-microscopie électronique des échantillons de ribonucléoprotéines. Les brins de polarité opposée sont représentés en bleu et rose. L'extrémité fermée de l'épingle à cheveux (en vert foncé) est composée de 3 à 5 NP. B) Modèle de la localisation de la matrice ARN (trait jaune) dans la structure hélicoïdale des RNP montrant un des brins dans une représentation ribbon et l'autre en potentiel énergétique de surface (adaptée de Arranz et al., 2012)

Au sein des virions les RNP sont organisées selon un modèle bien précis avec 7 RNP entourant une huitième (Noda et al., 2006, 2012). De plus les RNP formeraient un complexe supramoléculaire en interagissant les unes avec les autres via des segments d'ARN (Fournier

et al., 2012a, 2012b). Ces interactions semblent indispensables aux processus de formation des virions néo-synthétisés (Cf. § II.4.L'emballage des segments dans la particule virale).

2. Transcription versus réplication

Contrairement à la majorité des virus à ARN, les étapes de **transcription** mais aussi de **réplication** de l'information virale sont toutes les deux réalisées au sein du noyau. Ces deux processus sont réalisés par l'ARN-polymérase-ARN-dépendante qui fonctionne uniquement dans le cadre des RNP. La transcription permet l'obtention d'ARNm à partir des ARNv. Les ARNm vont ensuite dans le cytosol afin d'être traduits en protéine par la machinerie cellulaire. La réplication permet l'obtention d'un grand nombre de copies d'ARNv à partir d'ARNc (Hay et al., 1982). Le fait que la réplication ait lieu dans le noyau permet aux ARNv d'être encapsidés directement sous forme de RNP avant qu'ils ne rencontrent les protéines de types RIG-I qui provoqueraient une réponse immunitaire antivirale (Cf.§.II.4) (Rehwinkel et al., 2010).

Le processus de basculement de la polymérase entre transcription et réplication demeure encore mal connu aujourd'hui.

La transcription

La transcription permet la synthèse d'un brin d'ARNm, de polarité positive, codant pour les protéines virales. A la différence d'autres complexes polymérases, la RdRp de la grippe ne possède pas d'activité guanylyl-méthyltransférase requise pour la synthèse de la coiffe 7-méthyl-guanosine à l'extrémité 5' des ARNm. La présence de la coiffe en 5', tout comme celle de la queue poly-A en 3' sont indispensables pour la traduction des ARNm en protéines. L'évolution a donc poussé le virus de la grippe à générer des ARNm coiffés à travers un mécanisme de « vol de coiffe » (Bouloy et al., 1978; Plotch et al., 1979, 1981).

Le mécanisme de transcription fait appel aux trois sous-unités de la polymérase. Tout d'abord, PB2 va fixer la coiffe des pré-ARNm cellulaires, via son domaine de fixation à la coiffe (figure 16A). Ensuite, le domaine endonucléase de PA va cliver le pré-ARNm sur une longueur de 12-13 nucléotides (figure 16A). Si une guanine est présente dans la séquence, c'est à son extrémité 3' que PA va préférentiellement cliver le pré-ARNm (Datta et al., 2013). Ensuite, le domaine de fixation à la coiffe de PB2 tournerait librement et viendrait positionner la coiffe au niveau du site actif où se situe l'activité polymérase de PB1. Cela permettrait l'appariement des bases avec l'extrémité 3' de l'ARNv (figure 16B) (Pflug et al., 2014; Reich et al., 2014). PB1 est ensuite capable de réaliser la phase d'élongation en lisant l'ARNv de l'extrémité 3' à 5' (figure 16C). Une fois l'extrémité 5' atteinte, la polymérase

génère la queue poly-A par un mécanisme de bégaiement (figure 16D) (Tiley et al., 1994; Poon et al., 1999; González and Ortín, 1999). Des preuves expérimentales semblent indiquer que la RdRp recrute l'ARN polymérase II de l'hôte afin de favoriser l'exposition des pré-ARNm et donc de faciliter le processus de « vol de coiffe » (Engelhardt et al., 2005; Mayer et al., 2007; Rodriguez et al., 2007).

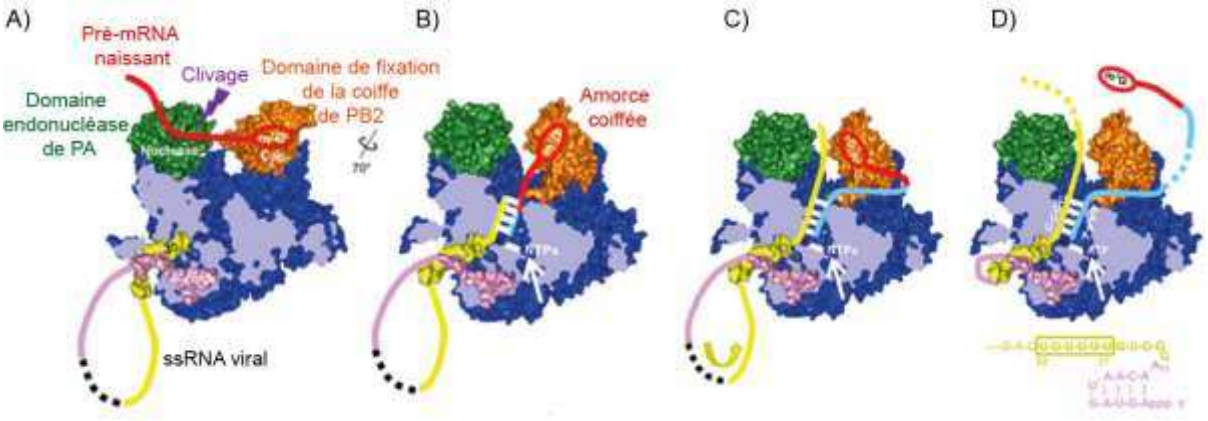


Figure 16: Mécanisme « du vol de coiffe » avec la structure de la polymérase de la grippe. A) La RdRp liée à l'ARNv capture un ARNm naissant avec le domaine de fixation à la coiffe de PB2 en orange et le clive sur une longueur de 12-13 nucléotides avec le domaine endonucléase de PA en vert. B) Le domaine de fixation à la coiffe effectue alors une rotation de 70° pour présenter l'ARNm clivé au niveau du site actif de la polymérase en bleu, et se lie avec une complémentarité partielle à l'extrémité 3' de l'ARNv. C) Lors de la phase d'élongation la matrice ARNv et le produit sont relargués de manière séparée. D) La queue poly A est générée via un mécanisme de bégaiement (adaptée de Reich et al., 2014).

La réplication

La réplication est un procédé menant à l'accumulation de nouveaux ARNv de polarité négative (-) (Hay et al., 1982). Ce procédé nécessite deux étapes. Tout d'abord, l'ARNv(-) est répliqué en ARNc de polarité positive (+). Ensuite, L'ARNc(+) sert de matrice pour générer de nouveaux brins d'ARNv(-)(Hay et al., 1977). Lors de ce processus, l'ARNc est lui aussi recouvert de NP nouvellement synthétisées et fixe la RdRp, générant ainsi des RNP(+) transitoires. Bien que la balance entre transcription et réplication ne soit pas complètement comprise, il est admis que la présence de nombreuses copies de NP et de la polymérase sont indispensables pour le processus de réplication. La NP serait notamment un des facteurs principaux de régulation, potentiellement par le biais d'interactions polymérase/NP (Biswas et al., 1998; Mena et al., 1999; Poole et al., 2004) mais aussi d'interactions avec la matrice ARNv à répliquer (Fodor et al., 1994b; Hsu et al., 1987a; Klumpp et al., 1997). D'autres facteurs régulent la balance réplication/transcription comme par exemple les facteurs cellulaires tels que la polymérase de type II (Engelhardt et al., 2005) ou l'hélicase réplicative MCM (Kawaguchi and Nagata, 2007). NEP est lui aussi un facteur de régulation négatif de la synthèse (Bullido et al., 2001; Robb et al., 2009). Dans tous les cas, plusieurs études

montrent que l'accumulation d'ARNv est un processus ayant lieu après une première phase de transcription (Wolstenholme et al., 1980; Thierry and Danos, 1982; Falcón et al., 2004; Hay et al., 1977; Huang et al., 1990) permettant d'accumuler les protéines nécessaires à la réplication.

Pour initier la réplication, la RdRp va synthétiser un dinucléotide (pppApG) à partir de la région promotrice hautement conservée de l'extrémité 3' de la matrice ARN (Deng et al., 2006; Zhang et al., 2010). Cependant il faut noter que ce mécanisme présente des différences en fonction de la matrice ARNv(-) ou ARNc(+) du fait de quelques différences structurales entre les deux types de matrice (Bae et al., 2001; Cheong et al., 1999; Park et al., 2003).

Pour la seconde étape de réplication (synthèse d'ARNv à partir d'ARNc) deux modèles ont été proposés. Le plus courant est celui de « réplication en trans » (figure 17) (Jorba et al., 2009; Resa-Infante et al., 2011) où l'extrémité 5' de la matrice ARNc est protégée par une polymérase résidente. Une autre RdRp non résidente est fixée à l'extrémité 3' de l'ARNc. Cette ARN polymérase non résidente initie la synthèse de l'ARNv et une seconde ARN polymérase non-résidente vient se fixer et protéger l'extrémité 5' de l'ARNv en cours de synthèse. Ceci induirait alors l'association avec NP et son oligomérisation autour de l'ARNv. Durant le processus, d'autres RdRp non-résidentes peuvent venir se fixer à l'extrémité 3' pour augmenter l'efficacité de la réplication. Ce modèle corrobore les données montrant que les deux extrémités sont nécessaires pour la réplication (Hagen et al., 1994; Olson et al., 2010; Vreede and Brownlee, 2007).

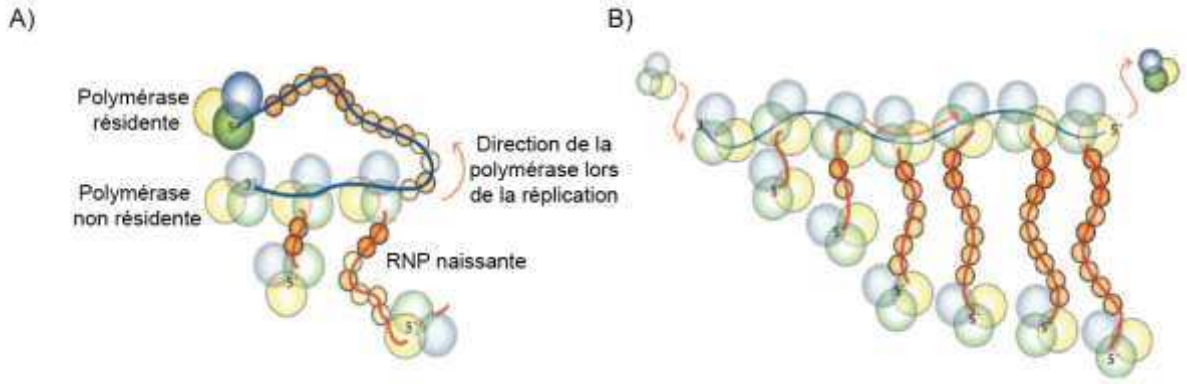


Figure 17: Modèle pour la synthèse des RNP. A) Phase d'initiation et d'élongation d'une polymérase non résidente présente à l'extrémité 3' de l'ARNv. Le brin naissant fixe une autre polymérase non résidente à son extrémité 5', de façon à la protéger d'une éventuelle dégradation B) Etape de terminaison où la polymérase résidente est forcée de se dissocier de l'extrémité 5' de l'ARNv (adaptée de Resa-Infante et al., 2011).

3. Interaction nucléoprotéine-ARN

La NP du virus de la grippe est une protéine riche en arginine et possède une charge nette positive à pH neutre qui reflète son activité de fixation à l'ARN. L'analyse de la séquence protéique ne montre pas de site spécifique de fixation à l'ARN. Ce dernier interagirait sûrement avec la NP au niveau du sillon basique présent à la jonction des domaines tête et corps (figure 18). La NP du virus H5N1 possède 16 résidus chargés au sein de ce sillon et 11 de ces 16 résidus sont conservés dans tous les types A, B, C et D. Cela suggère qu'il serait fortement impliqué dans la fixation à l'ARNv. Ces résidus sont retrouvés au niveau de deux clusters riches en arginine situés de part et d'autre du sillon basique ainsi que sur une boucle flexible, elle aussi située au niveau du sillon. La mutation des arginines de ces clusters et/ou la délétion de la boucle flexible entraine une perte importante d'affinité pour l'ARN (Ng et al., 2008). De plus, les différentes reconstructions obtenues en microscopie électronique (Cf. § IV.1) montrent que le sillon basique est exposé à la surface des RNP. D'après ces observations, Ng et al., ont émis l'hypothèse que l'ARN pourrait être capturé par la boucle flexible pour ensuite être libéré au niveau du sillon basique où il viendrait interagir via son squelette phosphodiester avec les arginines des deux clusters précédemment décrits (Elton et al., 1999).

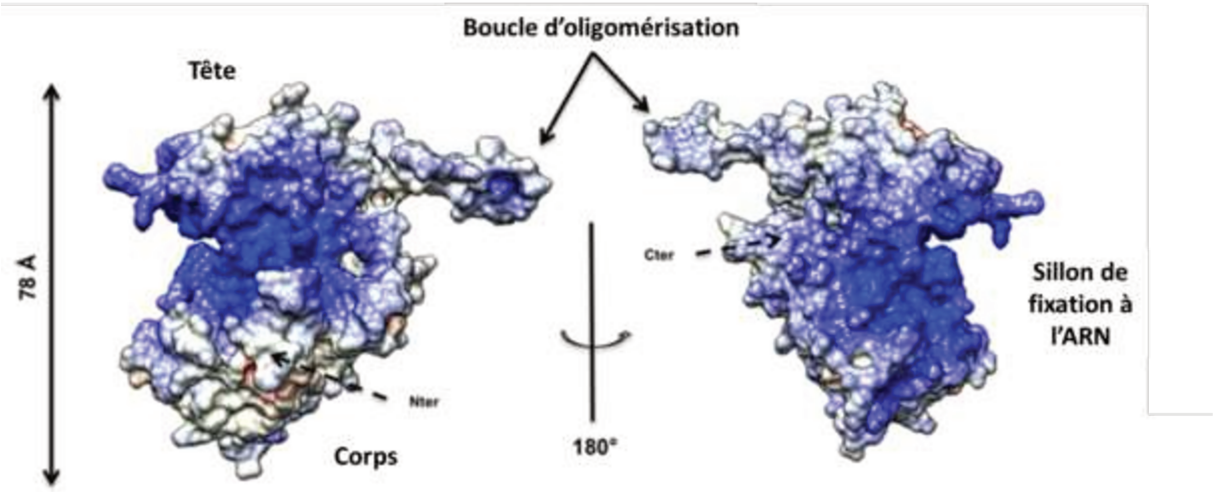


Figure 18 : Représentation de la surface électrostatique de la nucléoprotéine du virus de la grippe H1N1 (PDB ID 2IQH). La couleur bleue indique les régions chargées positivement (+10 kT/e) tandis que le rouge indique les régions chargées négativement (+10kT/e) (adaptée de la thèse de Christopher Swale, 2015).

Quant à lui, le mutant monomérique R416A ne possède affinité très faible pour l'ARN (≈ 2 x 10^{-6} M) (Boulo et al., 2011 ; Tarus et al., 2012). Cela est dû au réarrangement de certains résidus du fait de l'absence de liaison ionique avec le glutamate 339 d'un monomère voisin (Cf. II.3.La nucléoprotéine). La boucle vient se replier au niveau du site de fixation de l'ARN, empêchant alors la fixation de ce dernier (Chenavas et al., 2013a).

De nombreuses études montrent que la NP fixe l'ARN de manière coopérative (Yamanaka et al., 1990) avec une très forte affinité, (constante de dissociation de 20 nM) et sans spécificité de séquence (Baudin et al., 1994; Elton et al., 1999). Cependant le nombre de nucléotides fixés par nucléoprotéine reste encore incertain. La taille de l'ARNv est beaucoup plus importante que ce que peut accueillir le site proposé d'interaction et les modèles courants prédisent que la NP lie en moyenne 22 à 28 nt (Compans et al 1972, Ortega et al., 2000; Martín-Benito et al., 2001). Cela suggère que l'ARN est certainement enroulé autour de chaque monomère de NP (Jennings et al., 1983; Martín-Benito et al., 2001). Il est à noter que le nombre de nucléotides fixés par la NP semble différer de celui des autres virus à ARN négatif tel que les *rhabovirus* (Albertini et al., 2006), les *paramyxovirus* (Tawar et al., 2009a), les *Bunyavirus* (Reguera et al., 2013), ou encore les *Bornavirus* (Rudolph et al., 2003).

V. Cas des autres virus à ARN négatif

1. La ribonucléoprotéine des virus à ARN négatif

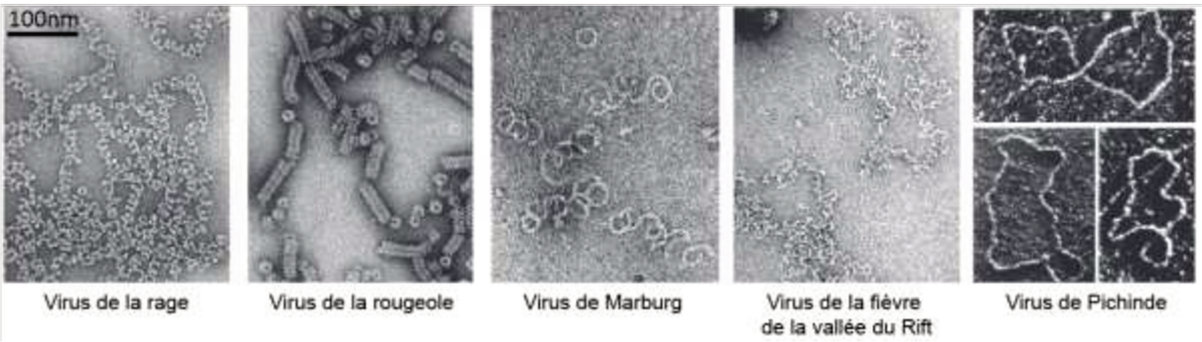


Figure 19 : Images de microscopie électronique de RNP de virus à ARN négatif. Les ribonucléoprotéines de la rage, de Pichinde et de la fièvre de la vallée du Rift ont été isolées à partir de virions tandis que les RNP de la rougeole et de Marburg sont recombinantes (adaptée de Ruigrok et al., 2011)

Une des caractéristiques des virus à ARN négatif est que leur génome est toujours trouvé assemblé à de multiples copies de nucléoprotéines (NP pour les virus à ARN segmenté et N pour les *Mononegavirales*) et au complexe de la polymérase (Reguera et al., 2014; Ruigrok et al., 2011). L'assemblage à la nucléoprotéine est concomitant à la réplication du génome de telle sorte qu'aucun ARNv de polarité positive ou négative n'est jamais retrouvé nu au sein de la cellule. Les nucléoprotéines sont multifonctionnelles : En recouvrant le génome elles le protègent de la dégradation, évitent la formation d'ARN double brin et compactent l'ARN au sein des RNP. Dans chacun des cas, NP/N est l'élément le plus abondant des RNP et c'est elle qui donne une forme hélicoïdale aux RNP. Chez les *Mononegavirales*, la RdRp dite protéine L nécessite la présence d'un cofacteur, la phosphoprotéine (P) pour se lier à l'ARNv (Curran et al., 1994)

Les RNP des *Mononegavirales* (figure 19, virus de la rage, de la rougeole et de Marburg) forment de longues structures hélicoïdales flexibles et assez régulières qui sont généralement linéaires (Gutsche et al., 2015). Lorsque les nucléocapsides sont intégrées dans les particules virales, notamment pour les *Rhabdovirus* et les *filovirus* (Tableau 1), elles forment alors des hélices plus rigides donnant aux virions leur forme globale (Desfosses et al., 2013; Ge et al., 2010; Schoehn et al., 2004; Tawar et al., 2009; Welsch et al., 2010).

A l'inverse, chez les virus à ARN négatif segmenté (figure 19, virus de la fièvre de la vallée du Rift, de Pichinde), les RNP adoptent une conformation hélicoïdale beaucoup plus flexible (Raymond et al., 2010). Dans le cas du virus de la grippe, comme décrit précédemment, les RNP sont même surenroulées du fait des interactions existantes entre les extrémités 5' et 3' de l'ARNv, interactions médiées par la RdRp (Hsu et al., 1987).

La différence de flexibilité entre les deux groupes de virus à ARN négatif peut s'expliquer par la structure de leur NP/N respective. Chez les *Mononegavirales,* les deux bras d'oligomérisation de la nucléoprotéine possèdent une flexibilité limitée de sorte que la N forme des contacts identiques avec tous ses voisins et génère donc une structure hélicoïdale régulière. A l'inverse, la boucle d'oligomérisation ou les bras amino et carboxy-terminaux (Cf. § V.2) qui médient l'interaction entre NP chez les virus à ARN segmenté forme des charnières flexibles qui permettent aux protéines de s'assembler dans des conformations différentes. Il faut tout de même noter que la flexibilité des RNP des virus à ARN segmenté reste relative et n'est pas identique pour tous les virus. Par exemple, des études par microscopie électronique montrent une différence nette entre les *Orthobunyavirus* et les *Orthomyxovirus*. La nature beaucoup plus régulière de la double hélice surenroulée des RNP du virus de la grippe (Arranz et al., 2012; Moeller et al., 2012) contraste avec la structure beaucoup plus flexible et moins hélicoïdale de *Bunyavirus* (Reguera et al., 2014).

2. Nucléoprotéines des virus à ARN négatif

Les NP/N sont les protéines les plus abondantes des RNP des virus à ARN négatif. Ce sont elles qui sont responsables de la structure hélicoïdale des RNP. De plus elles sont indispensables pour que les processus de transcription et de réplication s'effectuent en totalité (Ivanov et al., 2011; Reguera et al., 2014; Ruigrok et al., 2011, 2010). Les protéines NP/N des virus à ARN négatif possèdent toutes une forme de croissant formé de deux domaines. La plupart des N contiennent un domaine N-terminal et un domaine C-terminal (Albertini et al., 2006; Green et al., 2006; Hastie et al., 2011b; Raymond et al., 2010; Rudolph et al., 2003; Tawar et al., 2009) alors que les NP des *Orthomyxovirus* contiennent à la fois des régions amino et carboxy-terminales dans chacun des deux domaines (Ng et al., 2012; Ye et al., 2006; Zheng et al., 2013).

Une des propriétés essentielles des NP/N est leur capacité à oligomériser bien que les mécanismes d'oligomérisation diffèrent entre les virus à ARN segmenté ou non. Dans le cas des *Monenegavirales*, des interactions fortes impliquant les extensions N-terminale et/ou C-terminale, s'établissent entre les monomères de N (Cf. Structure des *Mononegavirales*) (Albertini et al., 2006; Green et al., 2006; Tawar et al., 2009). Cet arrangement est responsable de la structure hélicoïdale régulière des RNP des virus à ARN négatif non segmenté (Ge et al., 2010; Schoehn et al., 2004; Tawar et al., 2009). Dans le cas des virus à ARN segmenté, l'oligomérisation des NP se fait de manière beaucoup moins rigide. Comme décrit précédemment pour le virus de la grippe (Cf. § II.3.La nucléoprotéine), chez les *Orthomyxovirus,* l'oligomérisation se fait via une boucle flexible reliée au corps d'un premier monomère et qui vient se fixer sur un second monomère (Arranz et al., 2012; Chan et al.,

2010; Coloma et al., 2009; Ng et al., 2008, 2012; Ye et al., 2006). Cela permet entre autre aux NP de s'arranger sous forme de petits oligomères tels que des dimères (Zheng et al., 2013), des trimères (Ng et al., 2008; Ye et al., 2006) ou encore des tétramères (Ng et al., 2012). De plus, des structures plus complexes peuvent aussi être observées, notamment dans le cas de la grippe (Ruigrok and Baudin, 1995). Une telle oligomérisation est aussi observée chez les *Bunyavirus*, pour lesquels l'oligomérisation dépend de l'insertion d'un bras N-terminal d'un premier monomère dans un sillon hydrophobe d'un monomère voisin (Ferron et al., 2011).

La seconde caractéristique commune aux NP/N est leur capacité à fixer l'ARN. La plupart des nucléoprotéines des *Mononegavirales* et de quelques virus à ARN segmenté, ont été cristallisées en présence d'ARN. Dans chacun des cas, l'ARN se fixe au niveau d'un sillon chargé positivement et situé à l'interface entre les deux domaines protéiques des NP/N. Bien qu'il existe peut-être quelques exceptions (Osborne and Elliott, 2000), la fixation à l'ARN se fait de manière non-spécifique via des interactions avec le squelette phosphate et/ou des motifs sucrés bien que certains contacts sont directement établis avec les bases. A l'inverse de la NP du virus de la grippe (Cf. § IV.3), dans la plupart des cas les bases ne sont pas accessibles au solvant et l'ARN doit nécessairement se dissocier de la nucléoprotéine afin d'être utilisé comme matrice lors des étapes de transcription et de réplication. De même, la stœchiométrie NP/N-ARN n'est pas la même pour les *Orthomyxovirus* que pour les autres virus à ARN négatif. Alors que les *Mononegavirales*, les *Arenavirus* et les *Bunyavirus* fixent entre 7 et 11 nucléotides, le virus de la grippe fixerait entre 20 et 24 nucléotides (Compans et al., 1972 ; Martín-Benito et al., 2001; Ortega et al., 2000). Toutefois, des analyses biophysiques et des mesures d'affinité menées sur le virus de l'anémie infectieuse du saumon, un *Orthomyxovirus*, indique que la NP ne fixerait que 12 nucléotides suggérant qu'une fraction de l'ARN présent dans les RNP des *Orthomyxovirus* pourrait être associée de manière indirecte à un monomère de NP (Zheng et al., 2013).

3. Structure de la nucléoprotéine des *Mononegavirales*

De nombreuses structures de nucléoprotéines de *Mononegavirales* sont aujourd'hui disponibles (Reguera et al., 2014; Ruigrok et al., 2011). Exception faite du virus de Borna (Rudolph et al., 2003), les structures de ces nucléoprotéines ont toutes été obtenues en complexe avec l'ARN. Ces complexes N-ARN forment des anneaux contenant entre 10 et 11 protomères. Les premières structures N-ARN obtenues ont été celles des virus de la rage (Albertini et al., 2006) et du virus de la stomatite vésiculaire (Green et al., 2006) fixant chacune 9 bases de l'ARN. Par la suite la structure N-ARN du virus respiratoire syncytial montre que N fixe 7 nucléotides (Renner et al., 2016; Tawar et al., 2009). Plus récemment,

les structures de deux *paramyxovirus,* la rougeole (Gutsche et al., 2015) et le virus parainfluenza 5 (Alayyoubi et al., 2015) montrent que la N fixe 6 nucléotides.

Toutes ces N sont repliées en deux domaines, amino et carboxy-terminaux. L'extrémité amino-terminale est totalement désordonnée tandis qu'une hélice-α est retrouvée au niveau de l'extrémité carboxy-terminale (figure 20). Au sein des complexes N-RNA, les protomères établissent des contacts via leur domaine C-terminal. Ces contacts sont renforcés via un domaine d'échange additionnel constitué par l'extension N-terminale partant d'un protomère et allant se fixer au dos du protomère voisin. Pour les *Rhabdovirus*, l'extension C-terminale d'une nucléoprotéine établit elle aussi des contacts additionnels avec une N voisine, en allant se fixer sur une troisième nucléoprotéine. Cette fixation se fait du côté opposé au domaine d'échange de l'extension amino-terminale (Albertini et al., 2006). Il en résulte alors une structure N-ARN extrêmement stable. Au contraire, pour le virus respiratoire syncytial, l'extrémité carboxy-terminale est très flexible. Cette extrémité est homologue à celles du virus de la rougeole et du virus de Sendai qui sont intrinsèquement désordonnées et fixent la phosphoprotéine P, cofacteur de la polymérase L (Houben et al., 2007; Longhi et al., 2003).

Dans les anneaux N-ARN, la position de l'ARN diffère d'un virus à l'autre. En effet, l'ARN est situé à l'intérieur de l'anneau dans le cas des rhabdovirus alors qu'il est à l'extérieur dans le cas du virus respiratoire syncytial. Pour le virus de Borna, la mutation de certains résidus au niveau de ce sillon empêche la fixation de l'ARN suggérant que l'ARN vient aussi se fixer au niveau de ce site (Hock et al., 2010).

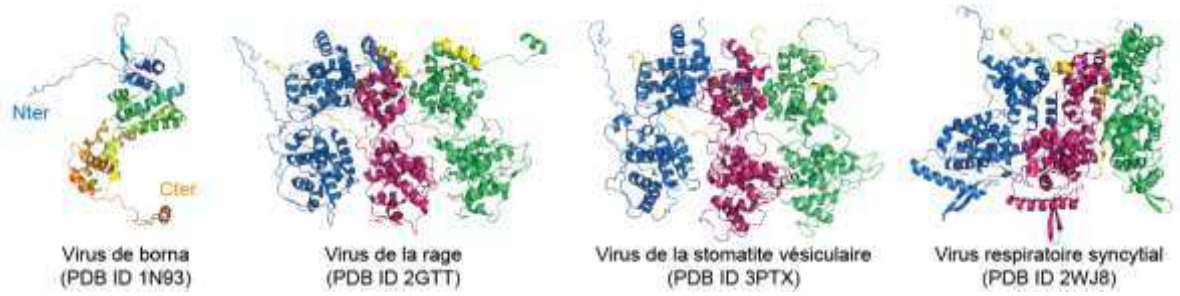


Figure 20 : Nucléoprotéines de *Mononegavirales*. Toutes les nucléoprotéines des *Mononegavirales* possèdent un repliement similaire. La représentation du monomère du virus de Borna (Rudolph et al., 2003) permet d'apprécier les deux domaines N et C-terminaux. Les autres virus (Albertini et al., 2006; Green et al., 2006; Tawar et al., 2009) permettent de visualiser les contacts (en jaune) établies par chaque protomère avec ses voisins.

In vivo, les NP/N des virus à ARN négatif fixent uniquement l'ARNv et l'ARNc. Chez les *Mononegavirales,* pour prévenir la fixation aux ARN cellulaires et préserver les N sous formes solubles, la phosphoprotéine agit comme une chaperonne en fixant la nucléoprotéine et en empêchant la fixation à l'ARN cellulaire (Mavrakis et al., 2006). Comme

la synthèse du génome viral et l'assemblage en complexe N-ARN sont des évènements couplés (Gubbay et al., 2001), il se pourrait que l'interaction entre la P et la nucléoprotéine sans ARN soit dissociée durant la réplication virale permettant ainsi à la N de fixer l'ARN naissant.

4. Structure des nucléoprotéines des virus à ARN segmenté

De nombreuses structures de NP de virus à ARN segmenté sont disponibles. C'est notamment le cas pour plusieurs *Orthomyxovirus, influenza* A (Chenavas et al., 2013; Ng et al., 2008; Ye et al., 2006, 2012), *influenza* B (Ng et al., 2012), le virus de l'anémie infectieuse du saumon (Zheng et al., 2013), pour le virus de la fièvre de la vallée du Rift appartenant au *Bunyavirus* (Raymond et al., 2010), pour le virus de La Crosse, un *Orthobunyavirus* (Reguera et al., 2013), et pour un Arenavirus, le virus de la fièvre de Lassa (Hastie et al., 2011a, 2011b; Qi et al., 2010). Contrairement aux *Mononegavirales* dont les N ont toutes des repliements très similaires, les NP des virus à ARN non segmenté possèdent des structures diverses.

Bien que ces NP présentent des structures très différentes, elles sont toutes repliées en hélice-α et pourvues d'une ou plusieurs extensions flexibles impliquées dans l'oligomérisation. Cette extension peut tout aussi bien être une boucle interne comme chez les *Orthomyxovirus*, des extrémités carboxy et amino-terminales comme pour les *Orthobunyavirus* (figure 21) ou encore des sous-domaines ou domaines flexibles comme pour le virus de Lassa.

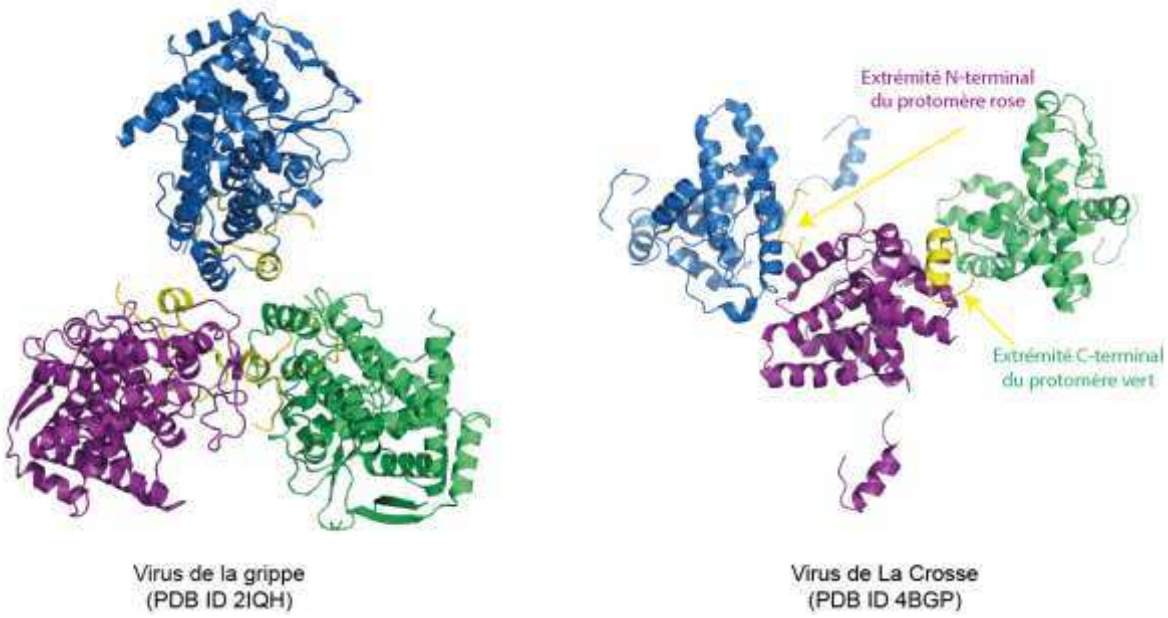


Figure 21 : Nucléoprotéines de virus à ARN négatif segmenté. Les représentations permettent de visualiser les contacts établis par chaque monomère avec ses voisins. Dans la cas du virus de la grippe (Ye et al., 2006) une boucle d'oligomérisation (représentée en jaune) sur chaque monomère établit un contact avec le monomère voisin. Dans le cas du virus de la Crosse (Reguera et al., 2013), les extrémités amino et carboxy-terminale (en jaune) établissent chacune une interaction avec un protomère adjacent.

Pour le virus de la vallée du rift tout comme le virus de Lassa, deux structures cristallographiques de nucléoprotéines en présence et en absence d'ARN ont été obtenues (Ferron et al., 2011; Hastie et al., 2011a; Qi et al., 2010; Raymond et al., 2010). Ces structures permettent de fournir les premiers indices sur la manière dont la nucléoprotéine reste monomérique *in vivo*, lorsqu'elle n'est pas fixée à l'ARNv. De même que les nucléoprotéines des *Mononegavirales* fixent la phosphoprotéine et se retrouvent alors dans une conformation ne permettant pas la fixation de l'ARN, les nucléoprotéines des virus de Lassa et de la fièvre de la vallée du Rift pourraient subir un changement conformationnel couplé au processus de réplication. Ce changement conformationnel permettrait alors la fixation à l'ARN.

Plusieurs structures cristallographiques de NP de virus à ARN négatif segmenté ont été obtenues en présence d'ARN. Les NP ayant été cristallisées en absence d'ARN sont généralement cristallisées sous forme oligomériques, démontrant la tendance inhérente des NP à oligomériser. Ainsi les NP des virus *influenza* A et B ont été respectivement cristallisées sous forme de trimères et de tétramères, les NP des Phlebovirus sous forme d'hexamères et celle du virus de Lassa sous forme trimériques. A ce jour, aucun rôle biologique n'a été attribué à ces formes oligomériques. Concernant les structures obtenues en présence d'ARN, elles présentent à la fois des similitudes et des différences dans la manière dont les NP fixent l'ARN. Ainsi il n'y a jamais de spécificité de séquence et le nombre de nucléotides fixés par la NP est variable. La NP du virus de Lassa fixe entre 6 et 7 nucléotides (Hastie et al., 2011a,

2011b), la NP du virus de la fièvre de la vallée du Rift en fixe 7 (Raymond et al., 2012) et les *Orthobunyavirus* en fixe 11 (Ariza et al., 2013; Dong et al., 2013; Niu et al., 2013; Reguera et al., 2013). Tout comme pour les *Mononegavirales*, l'ARN se fixe au niveau d'un sillon basique. Par exemple, chez les *Arenavirus*, l'interaction se fait entre le squelette phosphate de l'ARN et les lysines, les Arginines ainsi que le squelette de la chaîne polypeptidique. Chez les *Orthobunyavirus,* il y a des interactions clés établies avec le phosphate mais la plupart des interactions polaires se font avec les sucres et les bases de manière non-spécifique. Dans tous les cas, des liaisons hydrogènes sont établies avec le O2' du ribose, ce qui confère une spécificité de liaison à l'ARN plutôt qu'à l'ADN. L'interaction de l'ARN avec les NP de virus à ARN négatif segmenté indique qu'il faut au moins une dissociation partielle pour que l'ARN polymérase puisse agir.

VI. Transport nucléocytoplasmique de la nucléoprotéine du virus de la grippe

Dans les cellules eucaryotes, la réplication de l'ADN et la transcription en ARN sont deux processus spatialement séparés de la synthèse des protéines par la double membrane de l'enveloppe nucléaire. Les ARNm sont exportés dans le cytoplasme pour être traduits en protéine par les ribosomes. A l'inverse, les protéines nécessaires à la réplication et à la transcription telles que les histones, les polymérases ou encore les facteurs de transcription sont transférés au noyau depuis le compartiment cytosolique. Le transport de ses nombreuses molécules a lieu à travers le complexe du pore nucléaire (NPC).

1. Protéines impliquées dans le transport actif à travers le complexe du pore nucléaire

Les ions et les molécules de poids moléculaire inférieur à 40 kDa peuvent diffuser librement à travers le pore nucléaire. A l'inverse, les molécules dont la masse est supérieure à 40 kDa doivent être transportées de manière active à travers le NPC. Grâce au progrès de la microscopie électronique de nombreuses structures du NPC sont aujourd'hui disponibles (Akey and Radermacher, 1993; Alber et al., 2007; Beck et al., 2004, 2007; Hinshaw et al., 1992; Stoffler et al., 2003; Yang et al., 1998). Le NPC se présente sous la forme d'une structure cylindrique insérée dans la double membrane de l'enveloppe nucléaire (figure 22). Il possède un canal central de 10 nm de diamètre au repos et pouvant atteindre les 40 nm (Panté and Kann, 2002). Ce canal est prolongé par des fibrilles cytoplasmiques et nucléaires. Du côté nucléaire, les fibrilles sont toutes reliées à un anneau formant une structure dite en panier de basket. Le passage à travers ce pore s'effectue de manière bidirectionnelle (Nachury and Weis, 1999). Chez les mammifères, chaque pore est un complexe protéique d'environ 125 MDa. Il est composé d'une trentaine de protéines différentes. Ces protéines appelées nucléoporines (Nups) sont présentes en plusieurs exemplaires au sein du NPC (Schwartz, 2005).

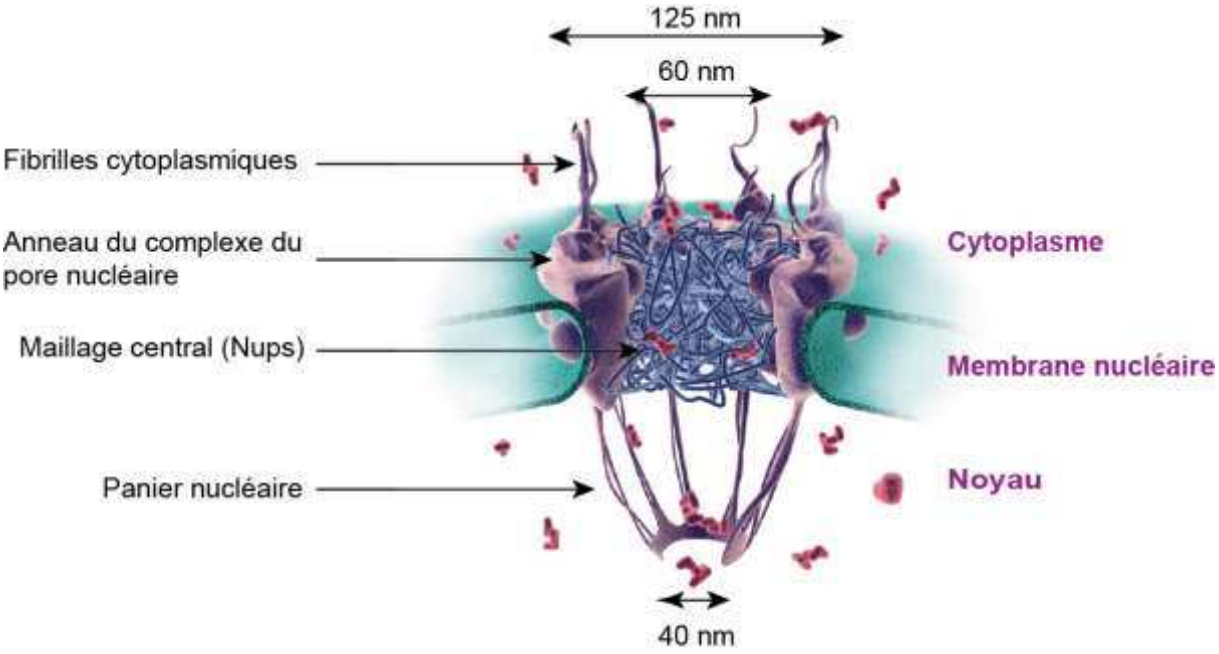


Figure 22 : Architecture du pore nucléaire. Les filaments cytoplasmiques sont arrangés autour du canal central. L'anneau distal du panier est connecté à l'anneau nucléaire par les filaments nucléaires (adaptée de Samir S.Patel : https://sites.google.com/site/sspatel/nuclearporecomplex2).

Les protéines à transporter à travers le NPC sont appelées protéines cargos. Les principales protéines impliquées dans le transport actif des protéines cargos sont les importines-β (aussi connues sous le nom de Karyophérine-β1) (Cingolani et al., 1999). Ces protéines fonctionnent de manière indépendante en se fixant directement au cargo ou bien, comme dans le cas de la nucléoprotéine du virus de la grippe, en s'associant à un adaptateur telles que les importine-α, qui lui, se fixe au cargo, (figure 23) (Goldfarb et al., 2004).

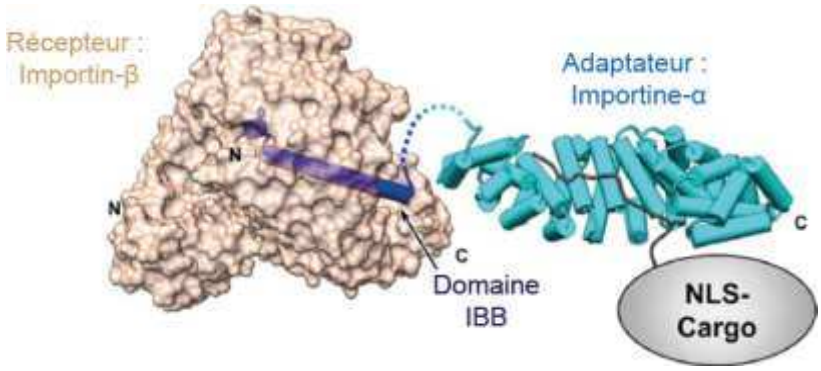


Figure 23 : Complexe d'import nucléaire classique. La surface d'une importine-β humaine est représentée (beige) liée au domaine IBB de l'importine-α (violet) (PDB ID 1QGK). L'importine-α1 de souris (cyan) est représentée en cartoon liée au domaine NLS du cargo (gris) (PDB ID 1EJY). Les extrémités amino et carboxy-terminales sont annotées N et C pour chaque protéine (adaptée de Pumroy and Cingolani, 2015)

L'importine-α

Les importines-α sont des protéines de la famille des Karyophérines. Chez les vertébrés, il existe 7 isoformes d'importines-α (1, 3, 4, 5, 6, 7 et 8). Les différentes isoformes d'importines-α humaines sont très bien conservées avec 26 % d'identité et 42 % de conservation dans leur séquence en acides-aminés (Pumroy and Cingolani, 2015). En fonction de l'hôte, les protéines grippales n'utilisent pas les mêmes isoformes d'importine pour leur import nucléaire. Pour une réplication efficace, la sous-unité PB2 et la NP des virus aviaires se lient à l'importine-α3 alors que les virus de mammifères montrent une spécificité pour l'isoforme-α7 (Gabriel et al., 2011). Au cours de ce travail de thèse, impliquant des protéines humaines et bovines, l'importine-α7 a été utilisée.

La reconnaissance du cargo par l'importine se fait via une séquence peptidique basique spécifique, KRxR, appelée signal de localisation nucléaire (NLS) (Kalderon et al., 1984; Lanford and Butel, 1984). Une NLS conventionnelle peut-être bipartite avec une NLS majeure et une NLS mineure : KRx(12-15)KRxR (Dingwall and Laskey, 1991). Il existe aussi des NLS dites non conventionnelles qui ne présentent pas de séquence aussi bien conservée comme c'est notamment le cas pour la NP du virus de la grippe (Cf. § VI.4).

L'importine-α est une protéine allongée, légèrement courbée et hélicoïdale. Les importines-α sont constituées d'une répétition de 10 motifs armadillo en tandem. Chaque armadillo est lui-même composé d'une succession de 3 hélices-α. Elles possèdent un domaine d'interaction de 40 acides-aminés avec l'importine-β, le domaine IBB (Importin-β binding domain) à leur extrémité amino-terminale. Au sein de ce domaine, la séquence KRR est reconnue par l'importine-β. Le domaine IBB est très similaire à la séquence NLS et possède un rôle d'auto-inhibition : en absence de cargo, il vient se fixer sur l'importine-α (Chook and Blobel, 2001). A l'inverse, lorsqu'une protéine cargo interagit avec l'importine-α, cela provoque la libération du domaine IBB qui vient fixer l'importine-β (Adam and Adam, 1994; Görlich et al., 1995; Imamoto et al., 1995; Moroianu et al., 1995). Ce processus constitue la voie classique d'import nucléaire (figure 23).

Malgré leur haut degré de similarité de séquence, les isoformes d'importine présentent quelques différences dans leur structure 3D et dans leurs propriétés biochimiques. Ces différences contribuent à leur spécificité *in vivo* pour le transport des cargos (Alvisi and Jans, 2015) et ces régulations commencent seulement à être comprises.

L'importine-β

L'importine-β est une protéine de 97 kDa composée de 19 répétitions du motif HEAT. Chacun de ces motifs est constitué de 2 hélices-α. C'est une protéine flexible dont la conformation varie en fonction de la molécule associée. L'importine-β peut en effet se lier à différents partenaires cellulaires. Elle reconnaît le domaine IBB de l'importine-α (figure 23) (Cingolani et al., 1999; Zachariae and Grubmüller, 2008), les motifs Phénylalanine-Glycine des Nups du NPC (Bayliss et al., 2000; Terry and Wente, 2007) ou encore RanGTP (Lee et al., 2005).

2. Transport actif des protéines cargos au noyau

Une fois le complexe hétérotrimèrique (importine-β/importine-α/cargo) formé, il passe à travers le NPC et vient délivrer le cargo à l'intérieur du noyau. Ce procédé se fait avec l'aide de petites molécules RanGTPases.

L'importine-β interagit par des liaisons hydrophobes avec les répétitions phénylalanines et glycines des nucléoporines du complexes du pore nucléaire (Otsuka et al., 2008). Une fois à l'intérieur du noyau, l'importine-β se lie à une petite protéine G (de la famille des GTPase) nommée RanGTP. Cette liaison entraîne un changement de conformation et la dissociation de l'hétérotrimère. Le cargo est relargué de l'importine-α par compétition du domaine IBB et avec l'aide de la nucléoporine Nup50 (Rexach and Blobel, 1995; Sun et al., 2008). L'importine-β complexée à RanGTP est directement renvoyée dans le cytoplasme pour y être recyclée. Quant à elle, l'importine-α lie la karyophérine-β CAS et RanGTP pour être exportée vers le cytoplasme (Kutay et al., 1997). Dans le cytoplasme, la protéine RanGAP (Ran GTPAse activating protein) stimule l'hydrolyse du GTP générant ainsi RanGDP, qui se dissocie alors de l'importine-α et la rend disponible pour un nouveau cycle d'import nucléaire (figure 24).

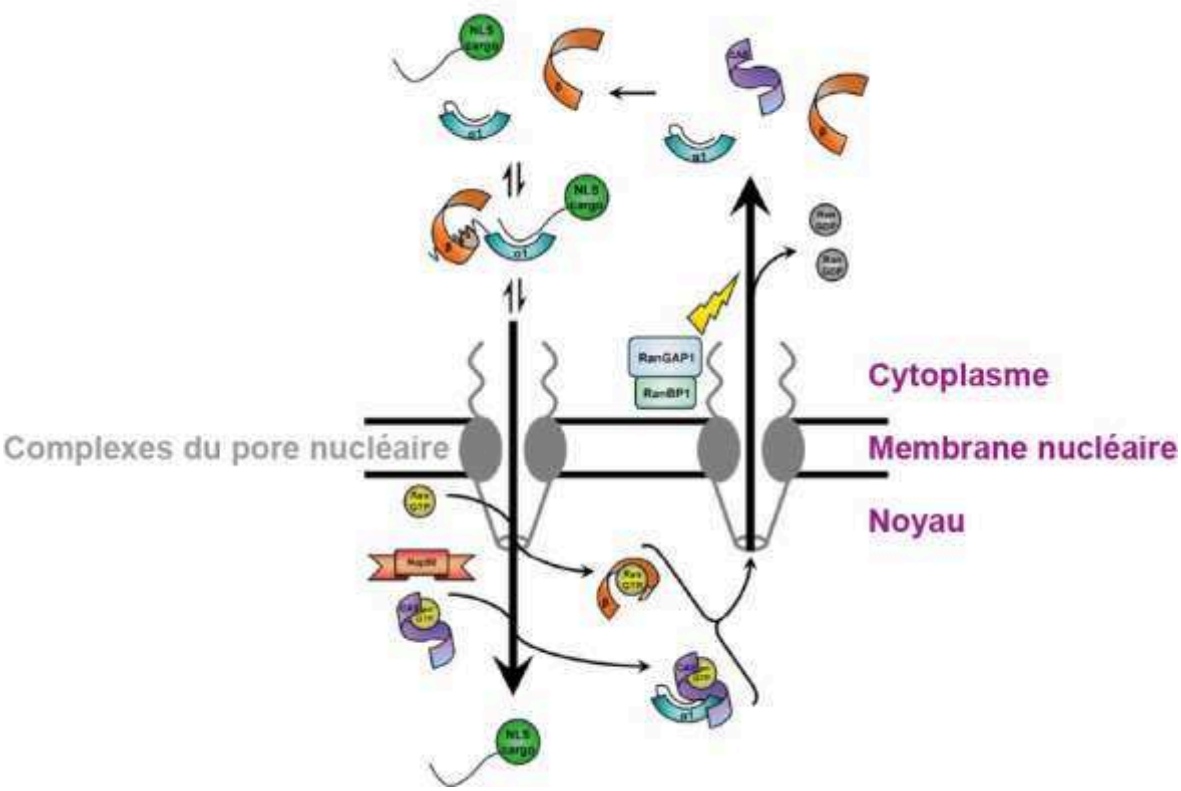


Figure 24 : Voie classique d'import nucléaire. L'importine-β, l'importine-α et le cargo s'assemblent dans le cytoplasme avant de passer à travers le pore nucléaire. La fixation de RanGTP à l'importine-β entraîne la libération du domaine IBB. Le facteur de recyclage CAS, les nucléoporines Nup50 et la fixation compétitive du domaine IBB contribuent tous trois à la libération du cargo de l'importine-α. Les importines libres sont alors recyclées dans le cytoplasme où elles sont relarguées via l'hydrolyse de RanGTP.

Ce mécanisme n'est pas le seul moyen d'import nucléaire actif mais il représente néanmoins le modèle le plus répandu dans la littérature, et est utilisé notamment par le génome du virus de la grippe.

3. Cas de la nucléoprotéine du virus de la grippe

Les étapes de réplication et de transcription du génome virale nécessitent le transport de la NP néo-synthétisée au noyau.

Dans le cas de la nucléoprotéine du virus de la grippe A, il a été montré que la protéine possède une NLS non conventionnelle située à son extrémité N-terminale au niveau des acides aminés 1 à 13 (Cros et al., 2005; Neumann et al., 1997; Wang et al., 1997). Il existerait aussi une NLS conventionnelle bipartite retrouvée au milieu de la séquence protéique au niveau des acides aminés 198 à 216 (Weber et al., 1998; Wu et al., 2007). Cependant il semblerait que cette NLS soit non-fonctionnelle (Cros et al., 2005). D'autres travaux, notamment ceux concernant la structure de la NP (Chen and Holmes, 2008; Ye et al., 2006)

ont montré que le motif bipartite ne possède pas les critères requis pour une NLS. Toutefois sa présence serait nécessaire à la réplication virale (Ozawa et al., 2007). De plus, des travaux récents ont permis d'obtenir une structure cristallographique de la NLS non conventionnelle en complexe avec l'importine-α1 de souris (Nakada et al., 2015) (figure 25). Les résultats montrent que la NLS de NP lie le sillon mineur du domaine de liaison de l'importine-α1. L'extrémité N-terminale de NP étant désordonnée et libre à la surface de la protéine (Ng et al., 2008), le domaine NLS est facilement accessible à l'importine. Cependant, contrairement à la majorité des protéines fixant l'importine avec une affinité assez forte, comprise entre 1 et 10 nM, la NP semble lier l'importine avec une faible affinité (K_D de l'ordre du µM) (Nakada et al., 2015).

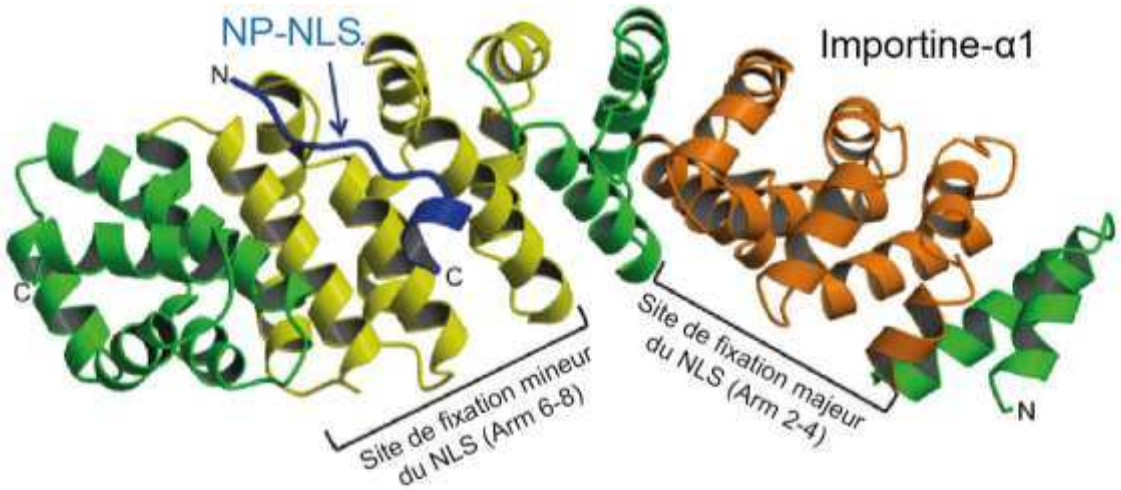


Figure 25 : NLS non-conventionnelle de NP fixée au niveau du sillon mineur de fixation de la NLS sur l'importine-α. Représentation Ribbon de la structure de l'importine-α avec ses répétitions de domaines armadillo (Arm). Les motifs se trouvant au niveau des sites de fixation majeurs et mineurs sont représentés respectivement en orange et jaune. La NLS non-conventionnelle de NP est représentée en bleu (adaptée de Nakada et al., 2015).

Afin de distinguer les NP à envoyer au noyau et les NP destinées à rester dans le cytoplasme, il est possible que la régulation se fasse via des phosphorylations sur des sérines/thréonines (Bui et al., 2002; Bullido et al., 2000). Il a notamment été montré que la phosphorylation des acide aminés sérines 9 et 3 et de la tyrosine 10 au cours de la phase tardive de l'infection conduit à une diminution de l'activité NLS (Bullido et al., 2000; Nakada et al., 2015; Zheng et al., 2015).

Dans le cas de la nucléoprotéine du virus de la grippe B, des études montrent que la NLS se situe au niveau de l'extrémité flexible N-terminale de la protéine comprenant 71 résidus (Liu et al., 2015; Sherry et al., 2014; Wanitchang et al., 2012). Au niveau de cette extrémité se trouve un motif KRXR en position 44-47 qui est hautement conservé entre les différentes souches de virus de type B et qui pourrait constituer un signal NLS (Wanitchang et al., 2013). De plus, la région comprenant les résidus 3-15 semble indispensable pour la localisation

nucléaire. Cette région ne présentant aucune des caractéristiques d'une séquence NLS pourrait être requise pour protéger l'extrémité N-terminale du clivage protéolytique (Wanitchang et al., 2013).

Concernant les virus des grippes C et D, aucune étude n'a été publiée à ce jour mais il semblerait que la NLS soit localisée au niveau de l'extrémité flexible C-terminale de ces protéines.

Objectifs du projet de Thèse

Le virus de la grippe reste aujourd'hui encore un problème majeur de santé public et est responsable d'un grand nombre de décès chaque année à travers le monde. Le premier virion de type A a été isolé en 1933 et bien que le virus ait été largement étudié depuis, le fonctionnement précis de la machinerie virale de transcription et de réplication ne reste que partiellement compris. Il est donc capital de poursuivre l'étude du virus de la grippe.

Ce projet de thèse se focalise sur l'étude des nucléoprotéines des virus de type A, B et D. Les virus de type A et B sont responsables d'épidémies chez l'homme et même de pandémies pour le type A tandis que le virus de type D est connu pour infecter principalement le bétail. L'étude de ces virus prend donc son sens dans la lutte contre les infections humaines dans le cas des types A et B et dans la protection des élevages destinés à l'industrie alimentaire dans le cas du type D. De plus, du fait de certaines similitudes entre les différents types de virus grippaux, il n'est pas impossible que le virus *influenza* D soit un jour transmis à l'homme.

La nucléoprotéine est une des protéines les plus abondantes du virus de la grippe et la protéine majoritaire des RNP. Au sein du complexe ribonucléoprotéique, elle recouvre entièrement l'ARN viral dont les extrémités 5' et 3' sont liées au complexe de l'ARN-polymérase-ARN-dépendante. C'est la nucléoprotéine qui régulerait la balance entre la réplication et la transcription lors du cycle viral (Skorko et al., 1991), en fonction notamment de son abondance et de sa disponibilité au sein du noyau des cellules infectées. Bien que de nombreuses études traitent de ses propriétés d'oligomérisation (Tarus et al., 2012; Ye et al., 2012) et de sa capacité à fixer l'ARN (Baudin et al., 1994; Martín-Benito et al., 2001; Yamanaka et al., 1990), son mécanisme de polymérisation autour de ce dernier reste encore inconnu. Au commencement de cette thèse, plusieurs structures cristallographiques de NP étaient déjà disponibles : 3 pour *influenza* A (Chenavas et al., 2013a; Ng et al., 2008; Ye et al., 2006) et 1 pour *influenza* B (Ng et al., 2012). Cependant, elles ont été toutes obtenues sans ARN bien que de nombreux groupes aient essayé d'obtenir une structure du complexe.

Au cours de ce travail de thèse, l'objectif principal était de déterminer la façon précise dont la nucléoprotéine fixe l'ARN afin de mieux comprendre le fonctionnement de la particule ribonucléoprotéique et notamment des processus de transcription et de réplication virales. Depuis de nombreuses années, plusieurs études biochimiques ont montré que l'ARN fixe la NP au niveau d'un sillon basique. Afin d'optimiser les chances de cristalliser un tel complexe, nous avons entrepris de travailler avec la NP du virus de type B dont la structure sans ARN avait été obtenue plus récemment (Ng et al., 2012). Nous nous sommes tout d'abord focalisés sur l'étude biochimique de la protéine afin de comprendre son comportement et plus particulièrement ses capacités d'oligomérisation en absence et en

présence d'un ARN de 8 nucléotides et en fonction de la concentration en sel (50, 150 et 300 mM NaCl). Malheureusement, malgré de nombreuses tentatives de cristallisation en présence de l'ARN nous n'avons pas réussi à obtenir de cristaux. Toutefois, le comportement biochimique assez différent de NP B par rapport à NP A nous a conduit à réaliser une étude comparative approfondie des deux nucléoprotéines en présence de différentes concentrations en sel et de plusieurs ARN de petites tailles allant de 5 à 24 nucléotides (article 1). Jusqu'à ce jour, de nombreuses études s'étaient intéressées à la fixation de l'ARN sur la NP mais aucune étude comparative de ce type n'avait été entreprise. Pour étudier et comparer le comportement de NP A et NP B, nous avons eu recours à la microscopie électronique et au SEC-MALLS afin de visualiser les différents états oligomériques des NP en fonction de la concentration en sel et de la taille de l'ARN ainsi qu'à l'anisotropie de fluorescence permettant de mesurer les affinités des NP pour les différents ARN. Les résultats obtenus montrent de nombreuses différences entre *influenza* A et B. Ces différences restent difficiles à expliquer clairement sans avoir la structure cristallographique du complexe NP-ARN. Cependant il a été possible d'établir que pour cristalliser la protéine en présence d'ARN un compromis entre concentration en sel et taille de l'ARN est nécessaire. Ainsi, à 150 mM NaCl, un ARN de 12 nucléotides semble être l'ARN idéal pour tenter de cristalliser le complexe. De plus nous avons pu formuler une nouvelle hypothèse concernant la manière dont l'ARN fixerait la nucléoprotéine.

Nous étant heurtés au problème de la cristallisation d'un complexe NP B-ARN, nous avons entrepris d'utiliser la nucléoprotéine de type D pour tenter de répondre à l'objectif initial étant de comprendre précisément la manière dont la NP fixe l'ARN. Aucune donnée n'ayant été précédemment publiée sur NP D, nous avons commencé par la caractérisation biochimique et structurale de la protéine (Article 2). Cette protéine possède un comportement assez différent de celui observé pour NP A et NP B, notamment en ce qui concerne les conditions physico-chimiques de purification. Après avoir réussi à isoler une forme tétramérique stable, nous avons obtenus des cristaux diffractant à 2.7 Å et la structure a pu être résolue par remplacement moléculaire à partir de NP B. Dans le cas de NP D l'extrémité C-terminale n'est pas visible sur la carte de densité électronique contrairement aux NP d'*influenza* A et B pour lesquelles l'extrémité N-terminale n'est pas visible.

En effet, alors que NP A et NP B présentent respectivement des extensions N-terminales de 21 et 71 résidus désordonnés, NP D possède une queue flexible de 51 acides aminés. Des analyses de séquences suggèrent aussi de l'existence d'une queue C-terminale pour NP C. Concernant NP A et NP B plusieurs études témoignaient de la présence d'un site NLS au niveau de cette queue flexible (Liu et al., 2015; Sherry et al., 2014; Wanitchang et al., 2012,

2013). Pour NP A une structure de la NLS en complexe avec l'importine-α1 de souris a même été obtenue (Nakada et al., 2015). Après l'obtention de la structure de NP D et quelques analyses de séquence suggérant de la présence potentielle d'un site NLS nous avons donc postulé que NP D fixait l'importine-α *via* son extrémité carboxy-terminale.

Afin de mettre en évidence l'interaction de la queue C-terminale avec l'importine, nous avons utilisé des mutants pour lesquels l'extrémité carboxy-terminale était tronquée afin de mettre en év. Enfin, afin d'étudier l'interaction avec l'ARN nous avons tenté de cristalliser la protéine en présence d'ARN. N'ayant obtenu que des cristaux dépourvus d'ARN, nous avons entrepris une étude similaire à celle réalisée avec NP A et NP B avec des ARN de petite taille faisant 5 à 24 nucléotides de long. En utilisant la forme totale et les formes tronquées nous avons pu mettre en évidence l'importance de l'extrémité C-terminale dans le processus de fixation à l'ARN.

Ainsi, ce travail de thèse, bien que n'ayant pas entièrement répondu à l'objectif initial de cristalliser la nucléoprotéine en présence d'ARN, a apporté de nombreuses informations biochimiques concernant cette interaction. De plus, il a permis de déterminer la structure de la nucléoprotéine du virus de type D et d'apporter une information essentielle quant à l'importance de l'extrémité C-terminale de NP D dans la fixation à l'ARN. Enfin il a permis de caractériser l'interaction de NP D en présence de l'importine.

Ce projet s'inscrit donc dans une dynamique globale visant à mieux comprendre le fonctionnement du virus de la grippe afin de pouvoir lutter aux mieux contre cette maladie qui fait toujours de très nombreuses victimes à travers le monde.

.

Chapitre 2 : Principes des méthodes expérimentales

De nombreuses techniques de biologie moléculaire, de biochimie et de biophysique ont été utilisées pour mener à bien ce travail de thèse. Toutes les procédures expérimentales sont décrites dans les parties matériels et méthodes des chapitres 3 à 5. Quant à ce chapitre, il s'efforce de décrire au mieux les principes des techniques utilisées.

I. Chromatographie d'exclusion de taille couplée au MALLS

(*MALLS: Multi Angle Laser Light Scattering*)

Les expériences de MALLS ont été réalisées sur la plateforme de biophysique du PSB de Grenoble.

Le système de diffusion de lumière à angles multiples (SEC-MALLS-RI: *Size Exclusion Chromatography Multi Angle Light Scattering-Refractive Index*) permet de déterminer la dispersité et la masse moléculaire absolue d'une molécule en se basant sur la façon dont elle diffracte la lumière.

Le système se compose d'une colonne de chromatographie d'exclusion de taille couplée à un détecteur de lumière à angles multiples et à un réfractomètre. Le détecteur permet de mesurer l'intensité de la lumière diffusée à différents angles tandis que le réfractomètre permet une mesure précise de la concentration en protéine. Cela est lié au fait que l'indice de réfraction de l'échantillon varie en fonction de la concentration protéique.

Lorsqu'une lumière polarisée traverse la cellule contenant des macromolécules en solution, l'excès de lumière diffusée par les molécules est directement proportionnel à la masse moléculaire (Oliva et al., 2004; Wyatt, 1991, 1993). La cellule est entourée de plusieurs détecteurs mesurant chacun l'intensité de lumière diffusée à différents angles. Les mesures d'intensité de la lumière diffusée à chaque angle et la concentration de l'échantillon mesurée au fur et à mesure de l'élution, grâce au réfractomètre, sont couplées pour déterminer la masse moléculaire qui est calculée à partir du graphique de Debye (ou Zimm). Ce dernier représente les paramètres de diffusion ($KC/R(\theta)$) en fonction de l'angle ($Sin2(\theta/2)$) (figure 26A).

Généralement, la courbe d'analyse finale générée par le logiciel de traitement des données représente l'excès d'indice de réfraction et la masse moléculaire en fonction du volume d'élution. L'analyse de cette courbe permet de déterminer la dispersité de l'échantillon en fonction de la stabilité de la masse moléculaire sous chaque pic (figure 26B).

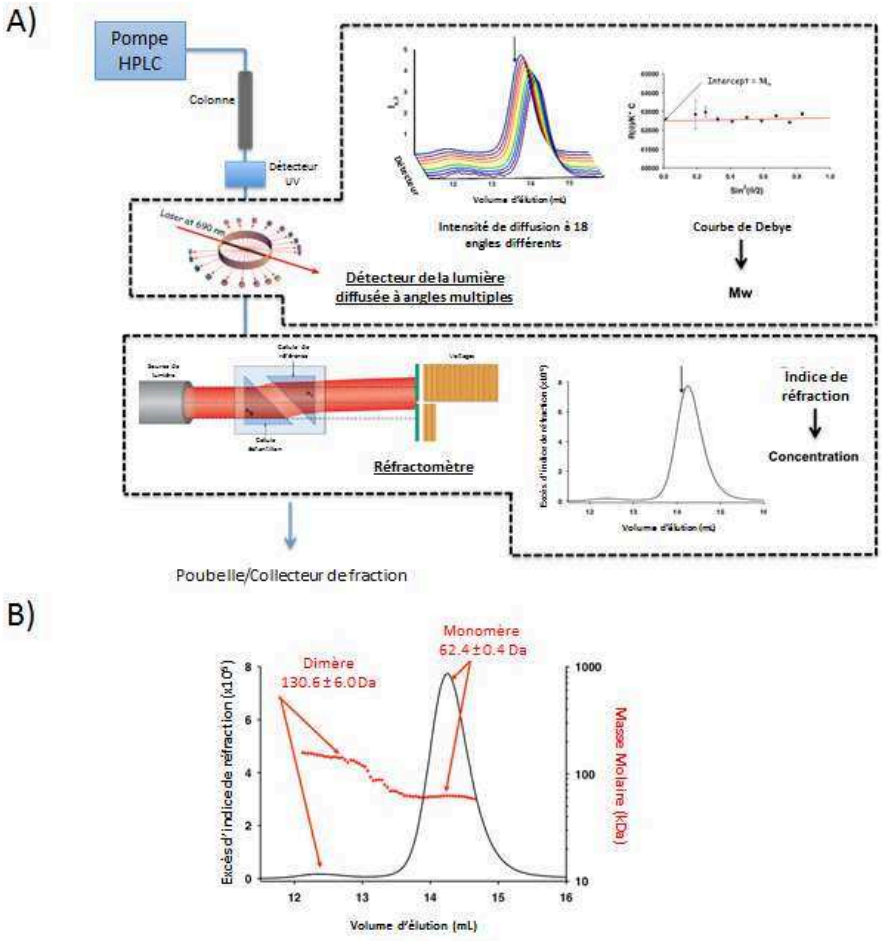


Figure 26 : Représentation schématique d'un système de chromatographie d'exclusion de taille couplé à un détecteur MALLS et à un réfractomètre. (A) Mesures réalisées par le détecteur MALLS et le réfractomètre. **(B)** Résultats d'une expérience de SEC-MALLS-RI après injection d'Albumine. La courbe noire représente l'excès d'indice de réfraction et les points rouges correspondent à la masse moléculaire estimée à chaque volume d'élution donné. (Image adaptée de Marc Jamin, IBS Grenoble).

.

II. Microscopie électronique à transmission

La microscopie électronique a été réalisée sur la plateforme de microscopie du PSB de Grenoble.

Le premier microscope électronique a été conçu en 1931 par Max Knoll et son étudiant Ernst Ruska. En 1986, Ernst Ruska reçut le prix Nobel de physique pour cette invention. Depuis, les progrès techniques ont sans cesse permis d'améliorer la qualité des images et la résolution des microscopes électroniques.

1. Principe de la microscopie électronique à transmission

Pour voir un objet de taille *d*, il faut un rayonnement dont la longueur d'onde λ, associée est telle que λ<*d*. Avec la lumière visible, la limite de résolution est donc d'environ 0,2 µm, ce qui est insuffisant pour observer des petits objets tels que des macromolécules biologiques. Dans ce cas, il faut recourir à l'utilisation d'une onde électromagnétique possédant une longueur d'onde plus faible que celle des photons. La solution est l'utilisation d'électrons dont longueur d'onde, λ est largement inférieure à l'angström (0,0251 Å à 200 KV). Cependant en raison de l'imperfection des lentilles électromagnétiques, les meilleurs microscopes dédiés à la biologie atteignent une résolution de l'ordre de l'angström ce qui est 50 000 fois plus que les microscopes optiques.

Le principe de fonctionnement d'un microscope électronique à transmission (MET) est très proche de celui d'un microscope optique. La différence principale étant l'utilisation des électrons à la place des photons. Un MET se compose principalement des éléments suivants :

- d'un système de pompes à vide,
- d'un canon à électrons composé d'une source d'électrons, d'un système de focalisation et d'un accélérateur d'électrons,
- d'une colonne optique électronique contenant les lentilles électromagnétiques et les diaphragmes,
- d'un porte échantillon (équipé d'un système de refroidissement à l'azote pour les cryo-microscopes),
- d'un détecteur d'électrons,
- éventuellement d'autres éléments électroniques de mesure.

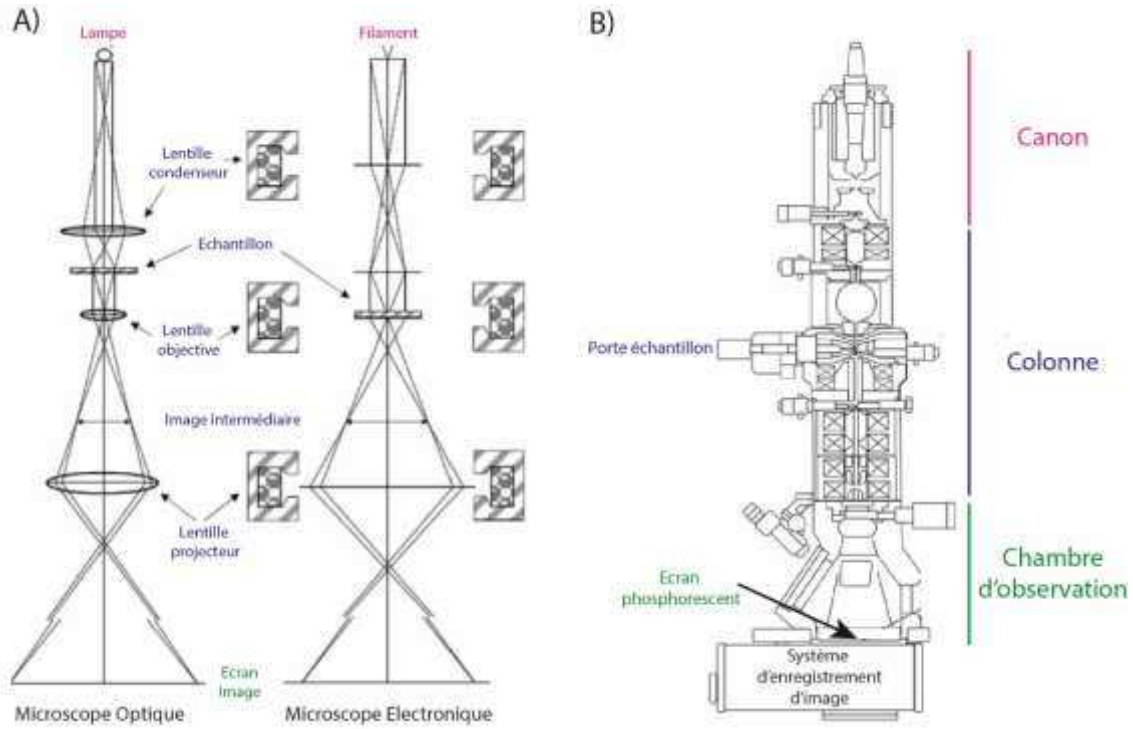


Figure 27 : Microscopie électronique à transmission. A) Représentation schématique comparant un système de microscopie optique et un système de microscopie électronique. **B)** Représentation schématique d'un microscope électronique à transmission. Pour **A)** et **B)**, les composants des trois principaux éléments du microscope, canon, colonne et chambre d'observation sont respectivement notés en rose, bleu et vert.

Au niveau du canon, la source d'électrons peut être un simple filament de tungstène, un cristal d'hexaborure de lanthane (respectivement chauffés entre 3000°C et 1500°C) ou encore une source à émission de champ. Cette dernière, plus récente, consiste à soumettre une pointe de tungstène (recouverte d'oxyde de zirconium) à une forte différence de potentiel électrique. Cela entraîne l'émission d'électrons délivrés en un faisceau de meilleure qualité, (plus brillant et plus cohérent que les deux premières sources). Les électrons sont ensuite accélérés par une forte tension d'accélération (de 60 à 400 kV) avant d'être propagés dans la colonne.

La colonne du MET ainsi que le canon et la chambre de projection sont gardés sous vide poussé en permanence car les électrons interagissent fortement avec la matière et peuvent donc aussi interagir avec les molécules d'air. Le faisceau d'électrons va d'abord être condensé par la lentille « condenseur » et se propage jusqu'à la lentille « objectif » où il va rencontrer l'échantillon. Les électrons du faisceau vont alors interagir avec ce dernier avant d'être projetés par la lentille « projecteur » sur un détecteur (généralement une caméra CCD) permettant une intégration informatique directe des images. Dans le cas d'un microscope optique, l'image vient directement se former sur la rétine de l'observateur.

C'est la rencontre des électrons avec l'échantillon qui va permettre de générer une image. L'épaisseur de l'échantillon étant faible, les électrons vont, soit passer à travers sans

interagir soit subir un phénomène de diffusion. La diffusion est inélastique dans le cas où un électron du faisceau percute un électron de l'échantillon. L'électron du faisceau perd de son énergie et sa trajectoire peut changer. La diffusion est élastique dans le cas où l'électron incident passe à proximité du noyau de l'atome et est attiré par la charge positive du noyau. La trajectoire de l'électron est alors largement déviée. Le contraste de l'image est en partie dû à la diffusion élastique des électrons par l'échantillon et notamment à l'élimination des électrons très déviés (par les parois du microscope ou par un diaphragme d'objectif qui permet d'éliminer des électrons plus faiblement déviés). Ceci entraîne la visualisation sur l'image de zones sombres (absence d'électrons) pour les parties de l'échantillon déviants fortement les électrons et de zones claires au niveau des zones déviant faiblement les électrons. Le marquage de l'échantillon par des sels de métaux lourds (technique de la coloration négative) exploite cette propriété pour augmenter le contraste dans l'image. Le contraste dépend aussi de la valeur de défocalisation de la lentille « objectif ».

2. Principe de la coloration négative

Les électrons constituent un rayonnement très destructeur pour les échantillons biologiques. De plus, la colonne d'un microscope électronique est soumise à un vide poussé. Il est donc nécessaire de préparer l'échantillon à observer en respectant certains principes afin d'en préserver sa structure.

Diverses méthodes de préparation des échantillons existent. Une des plus simples à mettre en œuvre et des plus répandues est celle de la coloration négative avec des sels de métaux lourds introduite par Hall en 1955 et décrite quelques années plus tard par Brenner et Horne (Brenner and Horne, 1959). Cette technique permet l'observation d'échantillons biologiques d'assez petite taille pouvant aller jusqu'à 50 kDa. Toutefois, elle ne permet pas d'obtenir une structure précise des particules observées.

La technique de la coloration négative consiste à fixer l'échantillon sur une grille par séchage d'une solution de sels de métaux lourds. Les métaux lourds possédant des noyaux très chargés, ils permettent de fortement dévier les électrons du faisceau et augmentent ainsi le contraste de l'échantillon.

Il existe plusieurs façons de procéder pour préparer les échantillons à observer. Dans ce manuscrit, c'est la technique de flottaison du carbone qui a été utilisée (figure 28). Une goutte (≈ 4 µL) de solution contenant l'échantillon à observer est déposée entre une surface de carbone et une couche de mica. Le carbone sert de support physique à l'échantillon et est obtenu par évaporation de graphite sur une lame de mica. Le tout est plongé dans une

solution contenant les sels de métaux lourds. Le carbone, sur lequel est attaché l'échantillon biologique, se détache du mica et les sels de métaux lourds viennent se déposer sur le carbone partout où il n'y a pas d'échantillon. Le carbone est alors récupéré sur une grille métallique (généralement en cuivre). L'excès de colorant est enlevé à l'aide de papier absorbant de type papier Whatman et la grille finit de sécher quelques minutes à l'air libre avant d'être visualisée ou stockée au sec dans une boîte prévue à cet effet. L'échantillon laisse une empreinte d'exclusion, sur le carbone, par rapport aux zones couvertes de métaux lourds et c'est cette empreinte qui va être observée. Le carbone est un matériau très fin (quelques nanomètres) et de faible absorbance, il laisse donc passer un maximum d'électrons contrairement aux métaux lourds : les électrons passent donc à travers la grille partout où il n'y a pas de métaux lourds, d'où le nom de coloration négative.

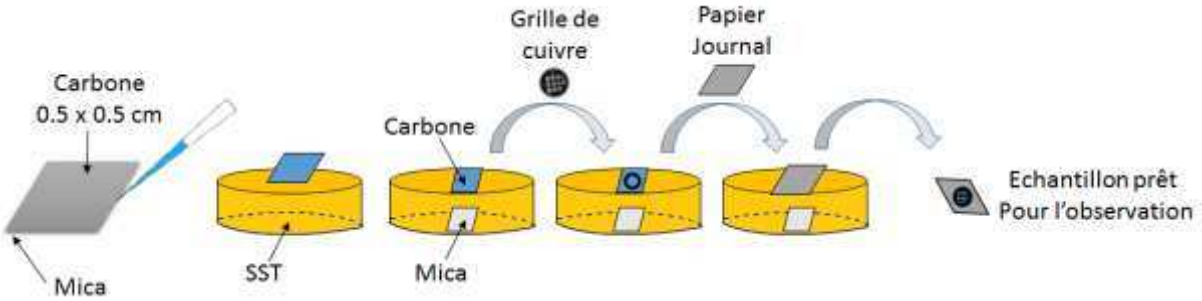


Figure 28 : Coloration négative par la technique de flottaison du carbone. L'échantillon est déposé entre le carbone et le mica. Le tout est plongé dans un bain de métaux lourds. Une grille de cuivre est déposée sur le carbone et récupérée à l'aide d'un papier journal.

Plusieurs solutions de colorant peuvent être utilisées comme le Silico tungstate de Sodium (SST), l'acétate d'Uranyle (UrAc) ou le Molybdate d'Ammonium (AmMb). Le choix du colorant dépend de l'échantillon à fixer ainsi que du pH de la solution de colorant. Le SST utilisé au cours de ce travail de thèse est un colorant de choix qui produit un grain assez fin donnant un contraste important quand on observe un échantillon de petite taille (50 kDa) et peu concentré. Cependant il est important de noter que l'adsorption sur une surface, la présence de colorants et le séchage vont entraîner des déformations importantes de l'échantillon.

.

III. L'anisotropie de fluorescence

L'anisotropie de fluorescence ou polarisation de fluorescence est une technique qui permet notamment de mesurer en temps réel les interactions entre les protéines et les séquences nucléotidiques (Heyduk and Lee, 1990).

La fluorescence est la capacité qu'ont les molécules à absorber un rayonnement lumineux et à restituer l'énergie sous forme de rayonnement lumineux de longueur d'onde plus élevée. La fluorophore absorbe un photon permettant à un électron de la molécule de passer d'un état fondamental à un état énergétique plus élevé. Le retour de l'électron à l'état fondamental se fait via l'émission d'un nouveau photon. Le temps de vie de l'état excité dépend du fluorophore et de l'environnement. Ce temps, de l'ordre de la nanoseconde, correspond au temps moyen que prends une molécule excité pour revenir à l'état fondamental.

L'anisotropie de fluorescence est basée sur l'excitation photo-sélective d'un fluorophore par une lumière polarisée linéairement : seule les fluorophores qui sont orientés de sorte que leur dipôle d'excitation soit colinéaire avec la lumière incidente polarisée vont être excités. Les fluorophores excités constituent un sous-ensemble de la population totale réémettant alors une lumière polarisée de façon cohérente.

Lorsqu'une solution de fluorophores est éclairée par une lumière polarisée incidente, les fluorophores excités se réorientent continuellement dans la solution en bougeant par diffusion rotationnelle durant un temps de vie de fluorescence. La polarisation de la lumière qu'il émet sera alors déterminée par sa nouvelle position. Si une population entière de fluorophores se réoriente aléatoirement durant le temps de vie de fluorescence, la lumière émise va devenir complètement dépolarisée puisque les positions de tous les dipôles seront aléatoires. La polarisation et l'intensité de la lumière émise dépendent donc de la diffusion rotationnelle pour une population donnée de fluorophores. C'est sur cette caractéristique que se fonde la mesure des interactions intermoléculaires par anisotropie de fluorescence.

L'anisotropie qui décrit la capacité des fluorophores pour maintenir le plan de la lumière polarisée peut être décrite par la formule suivante :

$$r = \frac{\mathrm{r0}}{1 + {}^{\tau}/_{\phi}}$$

où r représente l'anisotropie observée, r_0 l'anisotropie intrinsèque d'un fluorophore donné, τ le temps de vie de fluorescence et ϕ la constante de temps de corrélation de rotation. Pour une particule sphérique en solution, ϕ peut aussi s'exprimer selon la formule suivante :

$$\phi = \frac{\eta V}{RT}$$

où R représente la constante des gaz parfaits, T la température, η la constante de viscosité de la solution et V le volume moléculaire. Ainsi, lors d'une mesure d'anisotropie de fluorescence avec un ARN marqué, à température et viscosité constante, le seul facteur évoluant lors de la fixation de l'ARN à la molécule en solution est V. D'après la formule générale de l'anisotropie, l'anisotropie augmente lorsque la taille de la molécule augmente. Ainsi, dans le cas des interactions protéines/ARN, la mesure d'anisotropie de la lumière réémise se base sur le fait que les grosses molécules se réorientent plus lentement et fournissent donc des émissions de lumière plus polarisées. A l'inverse, les petites molécules se réorientent très rapidement et sont alors plus efficaces pour dépolariser la lumière. Ainsi, en solution, une petite molécule d'ARN marquée par un fluorophore va se réorienter plus rapidement que si elle est fixée à une protéine. L'augmentation de l'anisotropie de fluorescence due à une plus faible diffusion rotationnelle des complexes protéine/ARN par rapport à l'ARN libre est directement reliée à la fraction d'ARN fixée (figure 29).

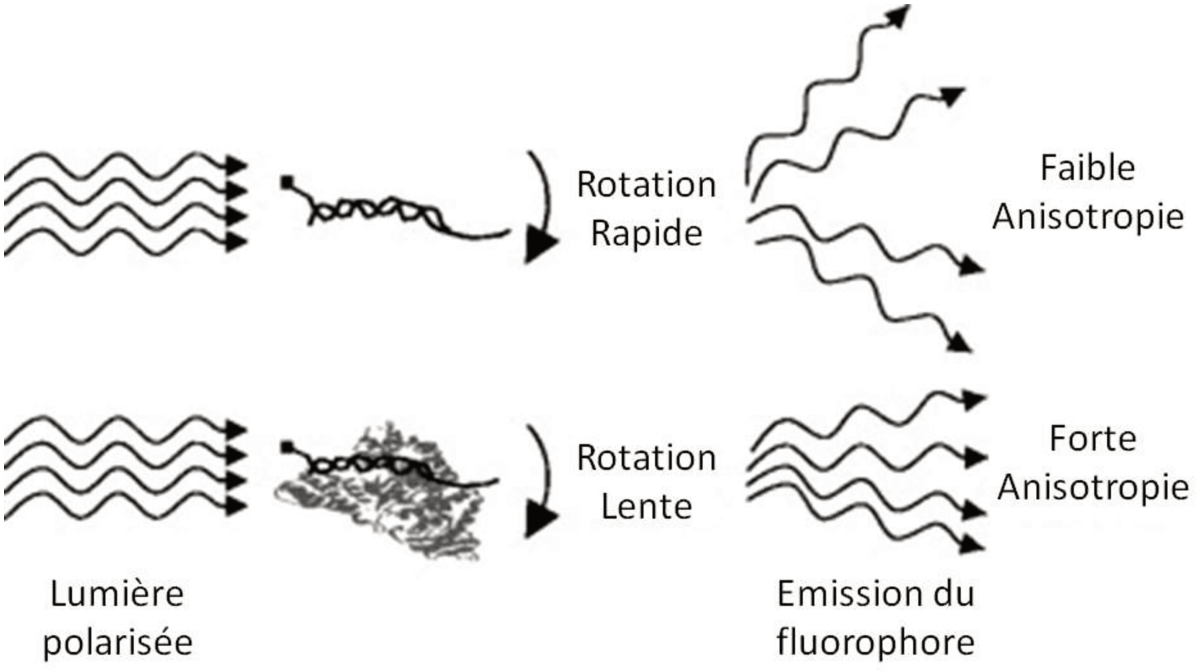


Figure 29 : Représentation schématique de l'anisotropie d'une molécule d'ADN fluorescente. L'anisotropie est accrue lors de la liaison d'un partenaire plus important qui ralentit la rotation de la molécule d'ADN. (Adaptée de LiCata and Wowor, 2008).

En pratique, une mesure d'anisotropie de fluorescence requiert l'utilisation d'un fluorimètre ayant une très forte sensibilité et possédant deux polariseurs. L'un positionné à

la longueur d'onde d'excitation et l'autre à la longueur d'onde d'émission. Pour chaque point de titration, deux mesures de fluorescence sont réalisées avec les polarisateurs alignés soit parallèlement (II) soit perpendiculairement (⊥) l'un par rapport à l'autre. Grâce à ces valeurs expérimentales, l'anisotropie peut être exprimée par l'équation suivante :

$$r = \frac{(I_{II} - I_{\perp})}{(I_{II} + 2I_{\perp})}$$

où I_{II} et $I_{\perp}$ sont les intensités de fluorescence respectivement parallèle et perpendiculaire au plan de la lumière polarisée. Les mesures d'affinités par anisotropie de fluorescence sont généralement réalisées en utilisant une concentration en ARN de l'ordre du nanomolaire.

IV. Cristallographie aux rayons-X

Tout comme le SAXS, la cristallographie aux rayons-X est une méthode basée sur la diffusion élastique des électrons. Cependant, à la différence du SAXS, avec cette méthode, l'information sur la manière dont sont orientées les molécules les unes par rapport aux autres dans l'échantillon est conservée. Cela conduit à l'obtention d'une carte de densité électronique qui permet de construire un modèle de la protéine à haute résolution.

Trois grandes étapes limitantes à l'obtention d'une structure cristallographique peuvent être décrites. La première est celle de la cristallogenèse qui consiste en l'obtention de cristaux et dépend principalement des propriétés biochimiques de la protéine (pureté, stabilité, flexibilité, …). La seconde est l'étape d'acquisition et d'intégration des données qui est limitée par la qualité/l'intensité de diffraction du cristal. Enfin, la troisième étape concerne le problème de la phase, cette dernière étant indispensable pour obtenir une carte de densité électronique. La phase peut être résolue par remplacement moléculaire en utilisant des protéines homologues et/ou par phasage expérimental. L'obtention d'un modèle protéique haute résolution est conditionnée par la réussite de ces trois étapes.

1. La cristallogenèse

La cristallogenèse consiste à induire la transition d'un état soluble à un état solide cristallin. L'état cristallin se définit par un haut degré d'agencement de l'organisation interne. Dans un cristal, les macromolécules sont assemblées de manière régulière, c'est-à-dire qu'un même motif est répété à l'identique un grand nombre de fois selon un réseau régulier tridimensionnel. La plus petite entité du réseau permettant de recomposer l'empilement correspond à la maille.

Quelque soit la protéine d'intérêt, quatre étapes sont requises pour la cristallogenèse : La pré-nucléation, la nucléation, la croissance et l'arrêt de la croissance cristalline. Ces étapes sont décrites au sein du diagramme de phase (figure 30).

Pour cristalliser des protéines, il existe différentes méthodes. La technique de diffusion de vapeur, relativement facile à mettre en œuvre et utilisée dans ce projet de thèse, est la plus courante. Dans une enceinte close, un puit contient une solution de précipants dont le volume dépasse largement celui de l'échantillon protéique. Les précipitants les plus utilisés sont généralement des sels et des précipitants organiques tels que le polyéthylène glycol ou l'isopropanol. Dans cette même enceinte, l'échantillon protéique est mélangé avec un peu

de solution de précipitants afin de former une goutte de quelques microlitres. Cette goutte est soit assise sur une surface adjacente, soit suspendue sur la partie supérieure du puit. Au cours du temps, l'eau et les solvants vont diffuser entre la goutte et le réservoir afin d'égaliser les concentrations de précipants entre la goutte et le puit. La protéine va alors progressivement se concentrer jusqu'à atteindre la zone de sursaturation dans laquelle la concentration de la protéine est supérieure à sa limite de solubilité. Il s'agit d'une zone frontière entre l'état soluble et l'état précipité à partir de laquelle la nucléation (point de départ de la cristallisation) peut intervenir ou alors la protéine précipite. Les zones de nucléation correspondent à des amas moléculaires à partir desquels la croissance cristalline peut se faire. De nombreux paramètres comme la température, le pH ou encore certains additifs présents au sein des solutions protéiques peuvent jouer un rôle important dans le processus de cristallisation qui reste un phénomène empirique.

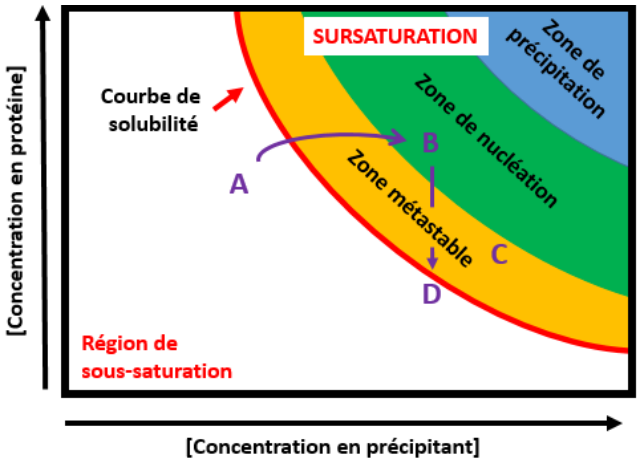


Figure 30 : Diagramme de phase de la cristallisation d'une protéine. Initialement, la protéine est sous forme soluble dans la région de sous-saturation (A). Lorsque la concentration en protéine augmente, elle dépasse la courbe de solubilité et atteint la région de sursaturation. Dans cette zone il peut potentiellement y avoir nucléation (B), puis croissance cristalline (C) au sein de la zone métastable. La croissance cristalline est stoppée (D) lorsque l'état d'équilibre est atteint.

Le cristal peut croître en quelques heures ou bien plusieurs semaines en fonction des protéines. Une fois la condition de cristallisation identifiée, un criblage plus fin est effectué afin d'obtenir des cristaux plus gros et le plus réguliers possible. Les cristaux sont ensuite pêchés et congelés dans de l'azote liquide pour être stockés avant exposition aux rayons-X. Généralement, des solutions de cryo-protectants sont utilisées pour garantir l'intégrité du cristal lors de la congélation.

2. Acquisition et intégration des données

Quand la matière est exposée à un faisceau incident de rayons-X, les électrons constituant la matière absorbent en partie les rayons-X et passent dans un état excités. Afin de revenir à leur état énergétique fondamental, ils réémettent des rayons-X dans toutes les directions de l'espace. Au sein d'un cristal, les protéines sont organisées périodiquement dans un ensemble de plans réticulaires sur lesquels le faisceau de rayons-X diffracte dans des directions privilégiées. Cela génère des interférences constructives produisant des tâches de diffraction (figure 31). Cette propriété est contrôlée par la loi de Bragg :

$$2d\ sin\theta = \eta\lambda$$

où d représente la distance entre deux plans réticulaires, λ la longueur d'onde du faisceau incident, η est un nombre entier positif appelé « ordre de diffraction » et θ l'angle entre le faisceau incident et le réseau réticulaire.

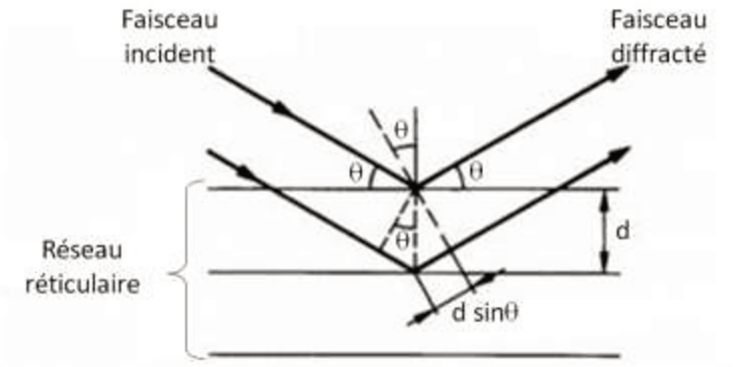


Figure 31 : Représentation schématique de la loi de Bragg. Le cristal constitue un ensemble de plans réticulaires sur lesquels le faisceau de rayons-X se réfléchit. D'après l'équation, 2dsinθ = ηλ, des interférences constructives sont obtenues uniquement si la différence de trajet dsinθ est un multiple de la longueur d'onde.

Un cliché de diffraction est constitué de tâches, ayant chacune des coordonnées précises et définissant le réseau réciproque. Chaque réflexion ou tâche de diffraction contient de l'information sur les atomes de la structure et chaque atome contribue à l'intensité des tâches des réflexions. La maille du réseau réciproque à des dimensions inversement proportionnelles à celle du réseau direct. A partir du spectre de diffraction il est possible de remonter à l'agencement des atomes dans le cristal grâce à l'image de la densité électronique de la molécule par la transformée de Fourier inverse. Cette image est obtenue en utilisant les phases ainsi que les amplitudes de chaque réflexion. Les amplitudes peuvent être tirées de l'intensité mesurée des réflexions mais aucun n'outil n'est disponible pour mesurer directement les phases des réflexions.

Une des premières étapes d'analyse informatique pendant la collecte de données est l'indexation. Cette étape peut être exécutée sur deux ou plusieurs images idéalement séparées par un angle de 90° et permet d'avoir une idée des dimensions de l'unité asymétrique du cristal à partir desquelles le groupe d'espace peut être déterminé. La détermination de groupe d'espace permet alors d'envisager une stratégie de collecte des données afin de limiter au maximum l'exposition du cristal aux rayons-X et donc les dommages dus aux radiations. Par exemple, si un cristal possède un groupe d'espace de symétrie P4, 90° d'oscillation seront nécessaires pour une collecte complète. Un angle d'oscillation assez faible pourra être utilisé pour la collecte. A l'inverse, pour un cristal ayant une symétrie P1, 360° d'oscillation seront nécessaires. L'angle d'oscillation utilisé devra donc être plus grand afin de ne pas trop irradier le cristal et de le préserver sur les 180°. Une fois le groupe d'espace établi, une collecte complète de données peut alors être entreprise pour mesurer toutes les réflexions dans l'espace réciproque. Généralement, il est considéré qu'il est nécessaire d'enregistrer au moins 98 % de réflexions pour avoir un jeu complet de données. Sur les lignes de lumière les plus modernes, il est possible d'enregistrer des oscillations avec un angle très faible (0.2°) ce qui génère un nombre très important d'images (> 1000 images) pour une collecte complète. Toutes ces images sont alors re-indexées avec des paramètres de maille affinés. Cela permet d'attribuer une coordonnée relative (l'indice de Miller, h, k, l) à toutes les tâches de réflexion enregistrées. L'indexation suivante correspond à l'intégration où toutes les images sont converties dans un fichier unique contenant toutes les informations sur chacune des réflexions : les indices h, k, l, son intensité, et l'erreur mesurée. Finalement les dernières étapes de préparation des données incluent la graduation et les étapes de fusion qui combinent les données et corrigent d'éventuelles contradictions survenues pendant la collecte (absorption due à la boucle, variations des paramètres sur la ligne de lumière, dommages dus aux radiations…).

3. Détermination de la phase

Pour déterminer les phases, plusieurs méthodes peuvent être utilisées. La première et la plus facile à mettre en œuvre est celle du remplacement moléculaire. Il s'agit d'une méthode purement informatique qui repose sur l'utilisation de structures précédemment résolues (Rossmann and Blow, 1962). En pratique, les programmes de remplacement moléculaires (Phaser ou Molrep) comparent les cartes de Patterson des données expérimentales avec les cartes du modèle, c'est-à-dire la protéine homologue. Une carte de Patterson est une carte vectorielle interatomique produite en élevant au carré toutes les amplitudes de facteur de structure et en mettant les phases au zéro. Une carte de Patterson est toujours centrosymétrique ainsi le remplacement moléculaire est divisé en deux étapes,

une recherche rotative (fonction d'auto-rotation) et une recherche de translation. Lors d'un remplacement moléculaire réussi, les deux étapes conduisent à une unique solution ou un nombre limité de solutions possibles. Le programme peut alors prévoir une solution de phasage en utilisant le modèle homologue. Un des principaux facteurs de réussite du remplacement moléculaire est l'homologie structurale (40 % au moins) entre le modèle et la structure réelle dans le cristal. Pour ce travail de thèse, le remplacement moléculaire est l'unique technique de phasage qui a été utilisée.

La seconde méthode est le phasage expérimental. Pour cela, différentes techniques peuvent être utilisées, et nécessitant donc l'ajout de différents composés dans le cristal natif. La première technique à avoir été utilisée historiquement est celle du Remplacement Isomorphe Simple (SIR, *Single Isomorphous Replacement)*. Cette technique nécessite le trempage des cristaux dans des solutions de métaux lourds. Les métaux lourds sont absorbés sans qu'il y ait de modifications dans les paramètres de maille du cristal. Pour une même longueur d'onde, les amplitudes des tâches de diffraction diffèrent grandement (de 15 à 25 % de différence) entre le cristal natif et le cristal trempé en raison de la contribution des atomes lourds. La position des atomes lourds au sein de l'unité asymétrique peut être déterminée par calculs informatiques en utilisant la fonction de Patterson. Ceci permet alors de fournir une solution de phasage pour tous les autres atomes de la structure. Cependant, le remplacement isomorphe simple peut conduire à une ambiguïté de la solution, du fait que deux solutions de phasage sont possibles par facteur de structure. Pour contrer ce problème, il est possible de combiner les jeux de données en utilisant un second ou plusieurs autres dérivés de métaux lourds pour faire du remplacement isomorphe multiple (MIR, *Multiple Isoporphous Remplacement)*.

Aujourd'hui, la méthode la plus utilisée pour la phasage expérimental est la Dispersion Anomale Simple (SAD, *Single Anomalous Dispersion)*. Cette méthode est analogue au SIR dans le sens où elle exige aussi l'introduction d'un composant supplémentaire dans le cristal natif (Hendrickson et al., 1990). Pour ce faire, la sélénométhionine est généralement utilisée en remplacement de la méthionine pendant l'expression de la protéine. Il est nécessaire que le nombre de méthionines présentes dans la séquence protéique soit au minimum de 1.5 % par rapport au nombre total de résidus. Le signal anomal est quantifié par des calculs informatiques et attribué aux atomes de sélénium présents dans la structure cristalline. Comme le SIR, le SAD peut produire des solutions de phasage ambiguës, qui peuvent être réduites en utilisant la Dispersion Anomale Multiple (MAD, *Multiple Anomalous Dispersion*), prenant en compte deux ensembles de données : une maximisant la dispersion anomale comme le SAD et une autre avec un signal de dispersion anomal faible. En raison des différences d'amplitude très faibles du signal anomal (1-5 %), une complétude élevée des données est donc nécessaire.

Les méthodes de phasage expérimental peuvent aussi être combinées pour produire un modèle plus précis. Les méthodes de SIRAS/MIRAS (*Single/Multiple Isomorphous Replacement with Anomalous Scattering)* tiennent notamment comptent des différences de diffraction comme source d'information de phase supplémentaire.

Une fois qu'une solution de phasage est trouvée, des programmes informatiques permettent de construire un modèle atomique de la protéine. Après cette première étape de construction, le modèle peut être affiné manuellement ainsi qu'informatiquement par d'autres programmes permettant de faire correspondre au mieux le modèle aux données expérimentales. Deux facteurs de corrélation majeurs sont principalement utilisés pour évaluer le modèle. Le *R-factor* qui calcule la différence entre les amplitudes calculées au dernier affinement et les amplitudes mesurées à l'étape d'affinement précédent. Pour évaluer le modèle final, le *free-R-factor* est utilisé afin de comparer les différences d'amplitudes entre les amplitudes affinées et 5 % des données initiales qui n'ont pas été soumises à l'affinement.

V. Test de stabilité thermique (TSA)

(TSA: Thermal Shift Assay)

Le test de stabilité thermique est utilisé pour étudier l'état de repliement d'une protéine dans des conditions données de température et de tampon. Cette méthode repose sur l'utilisation d'une sonde hydrophobe fluorescente, le Sypro dye qui ne fluoresce pas en solution aqueuse (Ericsson et al., 2006). Lors d'une expérience de TSA, sous l'effet de la chaleur la protéine se dénature et expose progressivement son cœur hydrophobe. La sonde vient se fixer aux patches hydrophobes de la protéine dépliée et émet alors un signal de fluorescence. Lorsque le signal a atteint son maximum, cela correspond à l'état totalement dénaturé de la protéine, la fluorescence va alors décroître du fait de l'agrégation de la protéine combinée à la dissociation du Sypro à des températures plus élevées (figure 32). Le Tm (température de fusion) d'une protéine donnée correspond à la température moyenne entre le point d'inflexion minimum et le point d'inflexion maximum de fluorescence. Cette technique a pour avantage d'être très facile et très rapide à mettre en œuvre. De plus elle permet d'accéder à des connaissances sur la stabilité thermique de la protéine en fonction du tampon, des additifs ou des ligands présents dans le milieu.

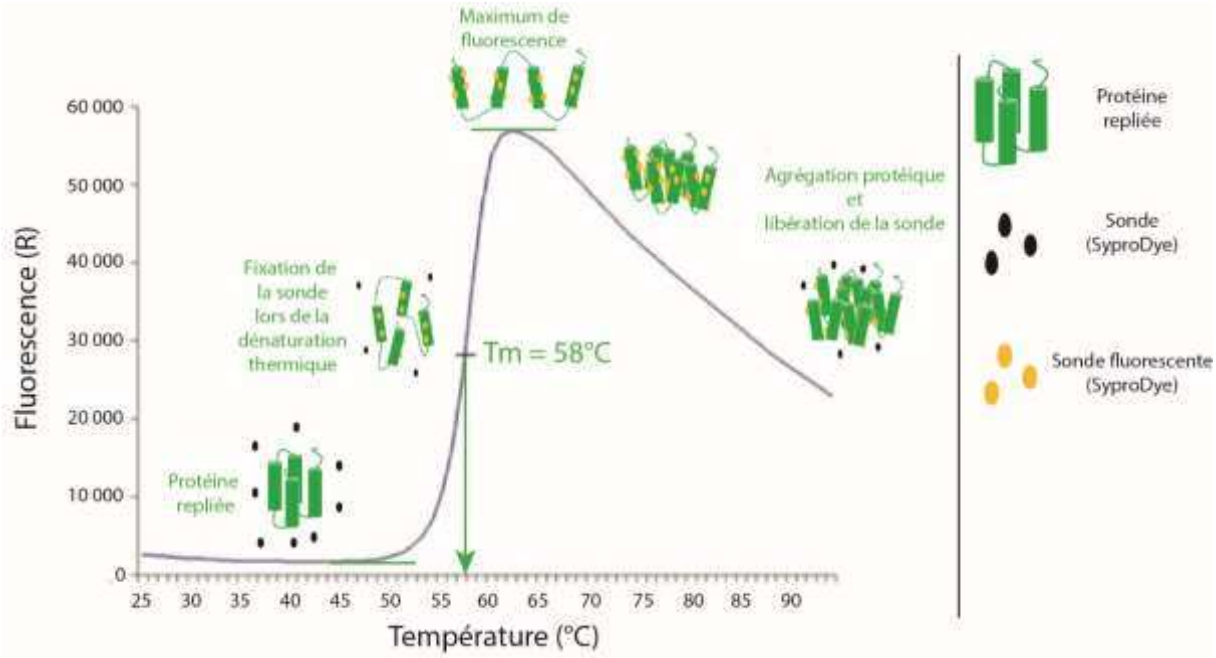


Figure 32 : Expérience de dénaturation thermique. Sous l'effet de la température, la protéine se dénature et expose ses séquences hydrophobes qui fixent la sonde. La sonde émet alors un signal de fluorescence quantifiable. Un fois le maximum de fluorescence atteint, la protéine, totalement dénaturée, s'agrège et la sonde est libérée dans le milieu.

VI. Dichroïsme circulaire

Le dichroïsme circulaire (CD) est une technique de spectroscopie permettant l'étude de l'effet de l'environnement sur la structure d'une molécule (tampon, température, ligand, …). Dans le cas des protéines, elle permet d'obtenir des informations sur leur structures secondaires (hélices-α, feuillets β et structures désordonnées).

Le CD repose sur la capacité d'une molécule chirale (acides nucléiques, sucres, protéines, …) à absorber différemment la lumière polarisée circulairement à droite et la lumière polarisée circulairement à gauche. En traversant l'échantillon biologique, ces ondes ne sont pas absorbées de la même façon. La différence d'absorption correspond à $\Delta A = A_G - A_D$ où A_G et A_D représentent respectivement l'absorbance de la lumière polarisée à gauche et à droite. Cela permet de définir le dichroïsme circulaire Δε :

$$\Delta\varepsilon = \varepsilon_G - \varepsilon_D$$

où ε_G et ε_G correspondent respectivement au coefficient d'extinction molaire du milieu en $mol^{-1}.L.cm^{-1}$ pour les deux directions de polarisation de la lumière. Or la loi de Beer-Lambert décrit que :

$$A = c\ x\ l\ x\ \varepsilon$$

où A est l'absorbance, c la concentration molaire en $mol.L^{-1}$ et l la longueur de la cuve en cm. D'après ces équations :

$$\Delta\varepsilon = \frac{\Delta A}{l\ x\ c}$$

Mais le dichroïsme circulaire est également accessible via une mesure de l'ellipticité molaire correspondant $[\theta] = 3300\ x\ \Delta\varepsilon$ en $degré.cm^2.dmol^{-1}$.

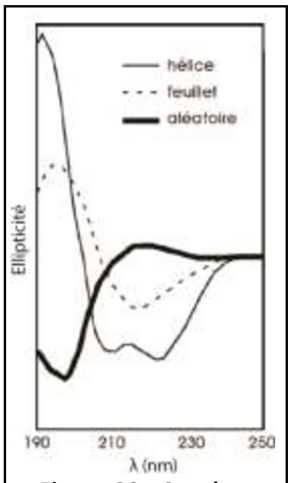


Figure 33 : Courbes caractéristiques obtenues en dichroïsme circulaire.

Pour effectuer une mesure de dichroïsme circulaire, l'échantillon est placé dans un faisceau dont la polarisation est alternativement gauche et droite. L'intensité lumineuse varie alors périodiquement et l'amplitude de ces variations permet de déterminer ΔA. Il est alors possible de tracer une courbe de l'ellipticité en fonction de la longueur d'onde. C'est à partir de l'allure de ces courbes que les informations sur la structure secondaire des protéines sont déduites (figure 33). Dans le cas des protéines, la mesure du CD est réalisée dans l'UV lointain, 180 – 220 nm, zone d'absorption de la liaison peptidique.

Chapitre 3 : Etude de la fixation de l'ARN à NP A et NP B

Viruses, 2016

Article 1: Binding of RNA by the nucleoproteins of influenza viruses A and B

Alice Labaronne, Christopher Swale, Alexandre Monod, Guy Schoehn, Thibaut Crépin et Rob W.H. Ruigrok

Note d'introduction au chapitre 3

Les ribonucléoprotéines du virus de la grippe représentent le cœur de la machinerie virale. Au sein des RNP, l'ARNv, entouré de multiples copies de nucléoprotéines, constitue le support des étapes de transcription et de réplication réalisées par l'ARN-polymérase-ARN-dépendante. Dans le but de comprendre le fonctionnement de cette machinerie, de nombreuses études fonctionnelles et structurales ont été entreprises depuis plus de 40 ans. De récents travaux ont permis de résoudre la structure de l'ARN-polymérase-ARN-dépendante en présence et en absence d'ARN (Pflug et al., 2014; Reich et al., 2014). Ces études ont permis d'apporter des précisions sur le fonctionnement de cette polymérase et plus particulièrement sur la façon dont elle interagit avec l'ARN. Depuis quelques années, plusieurs structures de nucléoprotéines ont été résolues pour les virus des grippes de type A et B (Chenavas et al., 2013; Ng et al., 2008, 2012; Ye et al., 2006). Toutefois, aucune de ces structures n'a été obtenue en présence d'ARN. Certaines études définissent qu'une vingtaine de bases serait fixées au niveau du sillon basique de la NP (Compans et al., 1972; Martín-Benito et al., 2001; Ortega et al., 2000) mais ce nombre reste sujet à controverse (Zheng et al., 2013). De nombreuses questions concernant la fixation et la polymérisation de la NP autour de l'ARNv dans le contexte des RNP restent donc encore en suspens.

De nombreux papiers traitent des différents états d'oligomérisation de la nucléoprotéine en fonction de la concentration en sel et de sa fixation à l'ARN (notamment pour NP A). De plus, de nombreux groupes ont mesuré l'affinité de NP pour l'ARN. Cependant ces études utilisent toutes des techniques de mesures différentes avec des NP recombinantes ou non et ayant été purifiées dans des conditions très variables de tampons et de températures.

Dans ce papier nous avons tenté de comprendre les processus d'oligomérisation de NP A et NP B dans des conditions très précises de tampons et en incrémentant progressivement la longueur des ARN utilisés. La microscopie électronique a permis de suivre l'état d'oligomérisation de NP A et NP B en fonction de différentes concentrations en sel (50, 150 et 300 mM NaCl) et de différentes longueurs d'ARN (ARN poly-UC de 5 à 24 nucléotides). Des expériences de SEC-MALLS réalisées dans des conditions similaires ont permis de confirmer les résultats obtenus en microscopie électronique. Enfin, nous avons utilisé la technique d'anisotropie de fluorescence pour mesurer l'affinité des ARN de différentes tailles pour NP A et NP B.

Nos résultats montrent de nombreuses différences entre NP A et NP B. Elles pourraient être en partie expliquées par une différence d'orientation de la boucle d'oligomérisation entre les deux protéines qui maintiendrait NP A dans une conformation plus fermée que NP B. Cependant, ces différences restent difficiles à expliquer du fait que nous n'ayons pas

encore la structure de NP en présence de l'ARN. Nos travaux nous ont toutefois permis de déterminer les conditions de tampons nécessaires et la taille idéale d'ARN pour tenter de cristalliser la NP en présence de ce dernier. Nous avons aussi mis en évidence un phénomène intéressant se produisant avec NP A dans des conditions précises de tampons et de longueur d'ARN. A 150 mM NaCl et en présence d'un ARN de 12 nucléotides marqué par une sonde fluorescente à son extrémité 3', nous avons observé des particules similaires aux RNP directement purifiées à partir de virions. Une étude cinétique en microscopie électronique nous a permis de montrer que l'ARN fixe la NP monomérique et ensuite les complexes oligomérisent pour former les *RNP-like.* Bien que nous ne puissions pas expliquer le mécanisme permettant l'apparition de ces espèces, ces dernières constituent néanmoins un outil très intéressant pour poursuivre l'étude de l'interaction entre la NP et l'ARN.

viruses

MDPI

Article

Binding of RNA by the Nucleoproteins of Influenza Viruses A and B

Alice Labaronne [1], Christopher Swale [1,2], Alexandre Monod [1], Guy Schoehn [1], Thibaut Crépin [1,*] and Rob W. H. Ruigrok [1]

1 Institut de Biologie Structurale (IBS), University Grenoble Alpes, CEA, CNRS, 38044 Grenoble, France; alice.tissot@ibs.fr (A.L.); cswale@embl.fr (C.S.); alexandre.monod@orange.fr (A.M.); guy.schoehn@ibs.fr (G.S.); rob.ruigrok@ibs.fr (R.W.H.R.)

2 EMBL Grenoble Outstation, 71 Avenue des Martyrs, BP181, F-38042 Grenoble CEDEX 9, France

* Correspondence: thibaut.crepin@ibs.fr; Tel.: +33-476-209-439

Academic Editors: Roland Marquet and Polly Roy

Received: 8 July 2016; Accepted: 29 August 2016; Published: 13 September 2016

Abstract: This paper describes a biochemical study for making complexes between the nucleoprotein of influenza viruses A and B (A/NP and B/NP) and small RNAs (polyUC RNAs from 5 to 24 nucleotides (nt)), starting from monomeric proteins. We used negative stain electron microscopy, size exclusion chromatography-multi-angle laser light scattering (SEC-MALLS) analysis, and fluorescence anisotropy measurements to show how the NP-RNA complexes evolve. Both proteins make small oligomers with 24-nt RNAs, trimers for A/NP, and dimers, tetramers, and larger complexes for B/NP. With shorter RNAs, the affinities of NP are all in the same range at 50 mM NaCl, showing that the RNAs bind on the same site. The affinity of B/NP for a 24-nt RNA does not change with salt. However, the affinity of A/NP for a 24-nt RNA is lower at 150 and 300 mM NaCl, suggesting that the RNA binds to another site, either on the same protomer or on a neighbour protomer. For our fluorescence anisotropy experiments, we used 6-fluorescein amidite (FAM)-labelled RNAs. By using a $(UC)_6$-FAM$^{3'}$ RNA with 150 mM NaCl, we observed an interesting phenomenon that gives macromolecular complexes similar to the ribonucleoprotein particles purified from the viruses.

Keywords: influenza virus; nucleoprotein; RNA; oligomerization; assembly; ribonucleoprotein

1. Introduction

Influenza belongs to the group of negative strand RNA viruses (NSVs). These viruses have a RNA genome that is coated with nucleoprotein (NP); the nucleoprotein–RNA complex (N–RNA) is the template for both transcription and replication by the viral polymerase. NP, viral RNA (vRNA), and polymerase constitute the ribonucleoparticle (RNP). Within the group of negative-strain RNA viruses, two genome organizations have been described: viruses with their entire genome made up of a single RNA molecule (non-segmented NSVs) and viruses with segmented genomes. The influenza viruses belong to the latter, with eight segments for types A and B.

To date, many X-ray structures of nucleoproteins bound to RNA have been published [1,2]. The first structures of the nucleoprotein–RNA complexes of non-segmented NSVs were these of vesicular stomatitis virus and rabies and showed that each nucleoprotein binds nine bases [3,4]. Thereafter, the structure of N–RNA of the respiratory syncytial virus showed seven bases bound to the nucleoprotein [5,6]. More recently, the structures from paramyxoviruses such as measles [7] and parainfluenza virus 5 [8] were solved. In both structures, six bases bind per nucleoprotein, confirming previous work based on biochemistry, electron microscopy (EM), and molecular virology [9–11]. The structures of the nucleoprotein–RNA complex from segmented RNA viruses are these of Lassa, binding six to seven bases [12,13], Rift Valley fever virus with seven bases [14], and Orthobunyaviruses

with 11 bases [15–19]. Several structures of other segmented NSVs' nucleoproteins have been obtained, including those of influenza, Ebola, and Borna viruses, but without RNA [20–27].

For influenza virus, two structures of the nucleoprotein of influenza A (A/NP) and influenza B (B/NP) have been obtained: A/NP trimers [25,27] and its R416A monomeric mutant [28], and B/NP tetramers [24]. Figure 1 shows the structures of the A/NP trimer and B/NP tetramer. The way that this protein makes oligomers is through an oligomerization loop that goes from one protomer into its neighbour. The loop makes an angle from their head/body core, thus making either trimers or tetramers (Figure 1). Negative staining electron microscopy of NP from influenza A/PR/8/34 showed monomers and trimers and many larger oligomers, but no dimers [29]. Within this paper, we will show that B/NP can make dimers, probably because the angle of the loop allows dimers in contrast to A/NP. No structure of nucleoprotein–RNA complex exists for influenza virus. Early work using biochemistry and electron microscopy on RNPs showed that each protomer of nucleoprotein binds about 20 bases [30–33]. Later, Ortega et al. [34] reconstituted RNPs from NP, viral RNA polymerase, and several vRNAs with different sizes. It was found that the RNPs with 48 bases per pair of NPs amplify best. Hutchinson et al. [35] used mass spectrometry on purified A/WSN/33 virions and found that NP binds 26 bases (13,588 bases for 530 copies of NP per virion). However, the work on the nucleoprotein of infectious salmon anaemia virus (ISAV), a fish orthomyxovirus [36], suggested that NP only binds 12 bases, although the sizes of the RNPs are the same when compared to influenza A and B viruses. These authors suggested that only 12 bases bind in the potential RNA-binding groove, whereas the other 12 bases bind in another way to the nucleoprotein.

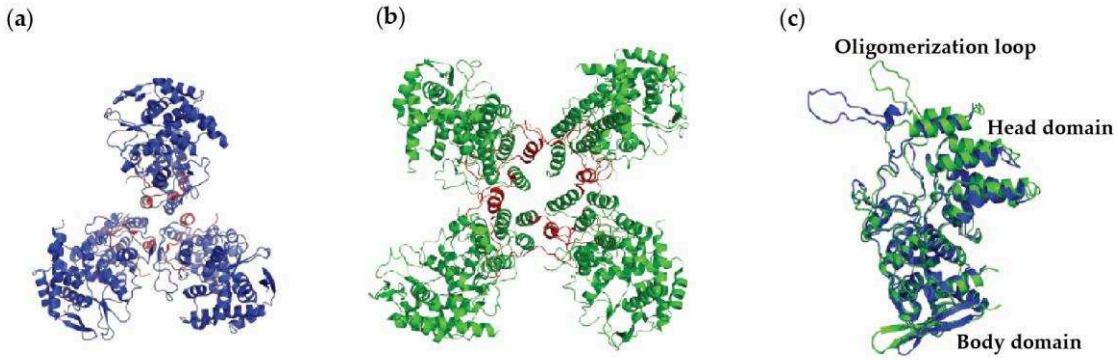


Figure 1. X-ray structures of influenza virus nucleoproteins. (**a**) Influenza A nucleoprotein (A/NP) (PDB ID: 2IQH) in blue forms a trimer, whereas (**b**) Influenza B nucleoprotein (B/NP) (PDB ID: 3TJ0) in green forms a tetramer. The oligomerization loops are coloured red on both (**a**) and (**b**). (**c**) Overlay of A/NP and B/NP. The overall folds of A/NP and B/NP are very similar with a head and a body domain. The root-mean-square deviation (r.m.s.d.) is 1.49 Å for 458 Cα. The exchange domains of the two structures point in different directions because they are bound to different neighbours.

Several groups have measured the affinities between the influenza virus A/NP and B/NP towards RNA with different methods (Table 1). Most studies used wild-type or mutant recombinant protein, either His-tagged or tagged with a maltose-binding protein (MBP) tag that was removed before the measurements. Baudin et al. [37] used NP from influenza A/PR/8/34 virus. For both NPs from A and B viruses, most of the apparent dissociation constants (K_d) were found to range from 4 to 70 nM, with experiments done at room temperature, 25 °C, and 37 °C. One test at 4 °C gave a 10-fold higher K_d (380 nM). The monomeric mutants (R416A and E339A) bind RNA with much higher K_d, between 1 and 10 μM [38]. The RNA-free recombinant or virus-purified NP displays a dynamic oligomeric behaviour, forming monomers, trimers/tetramers, and higher order oligomers with structures similar to real RNPs, depending on the protein concentration and salt content [29,39]. Tarus et al. [39] found that monomeric NP binds a 24-nucleotide (nt) RNA with a K_d of around 40 nM, whereas the trimer of NP binds with an affinity of 24 nM. They also found that trimers remain trimeric

when bound to RNA, whereas the monomer makes longer N–RNA complexes. This suggests that the trimers with RNA are blocked for making larger complexes.

In this paper, we show the effects of RNA binding on the oligomeric state of the NP from influenza viruses A and B, using size exclusion chromatography-multi-angle laser light scattering (SEC-MALLS), electron microscopy, and K_d measurements. We wanted to define a condition to obtain a homogeneous NP monomer bound to RNA for X-ray crystallization, but we did not find such condition. This paper describes the problems in making such a complex between NP and RNA. However, we have observed an interesting phenomenon between influenza A nucleoprotein and a 12-nucleotide RNA labelled with a fluorescein probe, which form very large protein–RNA complexes, similar to RNPs purified from the virus.

Table 1. Affinities of the nucleoproteins of influenza A and B for different RNAs. This table summarizes the results of affinity measurements from different studies. These studies used protein from different strains (A/WSN/1933; A/PR/8/34; A/HK/483/97(H5N1); B/Panama/45/90). Proteins are recombinant protein or directly purified from the virus. Recombinant proteins are MBP- or His-tagged. The MBP tag is cleaved after the first purification step. Different techniques were used to measure the affinities between nucleoprotein (NP) and RNA: filter binding assay (FBA), surface plasmon resonance (SPR), and polarization of fluorescence (Fluo-pol). Experiments were done at different temperatures, with different NaCl concentrations (from 100 to 300 mM) and with different RNAs (from eight to 890 nucleotides).

Virus	Protein	Method/Temp/[NaCl]	RNA	RNA Size (nt)	K_d (nM)	Reference
			Influenza A			
PR/8	virus	FBA/37 °C/100 mM	segment 8	890 nt	23	[37]
PR/8	virus	FBA/37 °C/100 mM	dsRNA	ds 20 nt	71	[37]
PR/8	(MBP)-NP [1]	FBA/100 mM (KCl)	panhandle	178 nt	16	[40]
PR/8	His-NP	FBA/4 °C/100 mM	panhandle	81 nt	380	[41]
WSN	His-NP [2]	SPR/25 °C/300 mM	5′biotin-RNA	24 nt	30	[39]
WSN	NP-His [2]	SPR/25 °C/300 mM	5′biotin-RNA	24 nt	41	[39]
WSN	His-NP	SPR/25 °C/300 mM	5′biotin-RNA	24 nt	14	[39]
WSN	NP-His	SPR/25 °C/300 mM	5′biotin-poly	8 nt	70	[39]
WSN	His-NP	SPR/25 °C/300 mM	5′biotin-RNA	40 nt	17	[39]
H5N1	(MBP)-NP [1]	SPR/25 °C/100 mM	2′O-methylated	24 nt	23	[25]
H5N1	(MBP)-NP [1]	SPR/25 °C	2′O-methylated	24 nt	16	[38]
WSN	(MBP)-NP [1]	Fluo-pol/RT/200 mM	5′fluo-RNA	20 nt	3.6	[42]
Influenza A mutants						
PR/8	E339A	FBA/4 °C/100 mM	panhandle	81 nt	1600	[41]
PR/8	R416A	FBA/4 °C/100 mM	panhandle	81 nt	2600	[41]
WSN	R416A	SPR/25 °C/300 mM	5′biotin-RNA	24 nt	10,000	[39]
H5N1	E339A	SPR/25 °C	2′O-methylated	24 nt	858	[38]
H5N1	R416A	SPR/25 °C	2′O-methylated	24 nt	975	[38]
WSN	Δ402-429	Fluo-pol/RT/200 mM	5′fluo-RNA	20 nt	1.7	[42]
Influenza B						
B	MBP-NP	SPR/25 °C/150 mM	2′O-methylated	24 nt	13	[24]

[1] The parentheses indicate that the MBP-Tag has been removed for the K_d measurements. [2] Monomeric nucleoprotein. Temp: temperature; PR/8: influenza A/PR/8/34; WSN: influenza A/WSN/1933; H5N1: A/HK/483/97(H5N1); B: influenza B/Panama/45/90; MBP: maltose-binding protein; NP: nucleoprotein.

2. Materials and Methods

2.1. Protein Expression and Purification of Influenza A and B NPs

The full-length NP genes of the H1N1 strain A/WSN/1933 and of the strain B/Memphis/13/03 were used. A/NP was cloned in pET22b (Novagen, Darmstadt, Germany) and B/NP in pETM11 (EMBL). *Escherichia coli* BL21 (DE3) cells were transformed with the resulting plasmids. Expression of the recombinant protein was induced by adding 0.3 mM isopropyl-β-D-thiogalactopyranoside (IPTG; Euromedex, Souffelweyersheim, France). Cultures were grown at 18 °C for 12 h and cells were

collected by centrifugation at 4000 *g*, for 15 min at 4 °C. Pellets were resuspended and sonicated in lysis buffer (50 mM Tris-HCl pH 7.5, 300 mM NaCl, 1 M NDSB201 (Sigma, Saint-Quentin Fallavier, France), 5 mM β-mercaptoethanol (β-ME; Roth, Lagny-sur-Marne, France) containing a cOmplete™ protease inhibitor cocktail (Roche, Meylan, France). Purifications were performed at room temperature. Proteins were purified by Ni^{2+} affinity chromatography (Ni-NTA, Qiagen; Les Ulis, France) followed by a heparin column (GE-Healthcare, Dutscher, Brumath, France). Heparin elution fractions were dialyzed against 20 mM Tris-HCl pH 7.5 at 50 mM, 150 mM or 300 mM NaCl, and 5 mM β-ME. The last purification step was a size-exclusion chromatography using a Superdex increase S200 10/300 GL column (GE-Healthcare). The proteins were eluted in 50, 150, or 300 mM NaCl according to the experimental needs. Peak fractions were concentrated using a 10 kDa concentrator (Amicon, Dutscher, Brumath, France). The protein concentrations were determined using the extinction coefficients at 280 nm for A/NP ε = 55,537 $M^{-1}\cdot cm^{-1}$ and for B/NP ε = 27,975 $M^{-1}\cdot cm^{-1}$.

2.2. Fluorescence Anisotropy Measurements

Fluorescence anisotropy assays were performed on a Clariostar (BMG Labtech, Champigny-sur-Marne, France) microplate reader, fitted with polarization filters to measure fluorescence anisotropy. The binding assays were done in 384-well plates at room temperature (60 µL reaction volume) in 20 mM Tris-HCl pH 7.5, 50, 150, or 300 mM NaCl and 5 mM β-ME. Concentrations ranging from 0 to 3 µM of A/NP or B/NP were titrated into 5 nM $^{5'}$phosphate RNA labelled in 3′ with 6-fluorescein amidite (FAM). Five to 24 nt polyUC RNA (Integrated DNA Technologies, Coralville, IA, USA) were used. After subtracting the polarization values obtained for RNA alone, the mean value of three independent experiments were fitted to the standard binding equation:

$$y = ((\text{Bmax} \times x\text{h}) / (Kd\text{h} + x\text{h})) \quad (1)$$

assuming a single binding site with Hill slope (h) using Prism (Version 7, GraphPad, La Jolla, CA, USA). A control titration with A/NP and fluorescein alone was performed in order to confirm that the protein does not bind directly to fluorescein (Figure S4a). Bmax is the maximum specific binding in the same units as y.

2.3. Electron Microscopy

Samples (concentrations close to 0.05 $mg\cdot mL^{-1}$) were applied between a carbon and a mica layer. The carbon was then floated on the top of a heavy atom salt drop (2% (w/v) sodium silicotungstate, pH 7.0). The carbon film was covered by a copper grid. Both were fished using a small piece of paper and air dried before insertion in the electron microscope [43,44]. Charge-coupled Device (CCD) frames were taken with a T12 microscope (FEI, Hillsboro, OR, USA) operating at 120 kV and a nominal magnification of 45,000 times. The $^{5'}$phosphate-polyUC-FAM$^{3'}$-NP complex was incubated overnight at room temperature using 100 µM of proteins in 20 mM Tris-HCl pH 7.5, 50, 150 or 300 mM NaCl and 5 mM β-ME and with a final ratio NP:RNA of 1:1. The dilutions for EM were performed with a buffer with the same salt concentration.

2.4. SEC-MALLS Experiments

SEC was performed with a Superdex increase S200 10/300 GL column (GE healthcare) equilibrated with 20 mM Tris-HCl, pH 7.5, 50, 150, or 300 mM NaCl, and 5 mM β-ME. Analytical runs were performed at 20 °C with a flow rate of 0.5 $mL\cdot min^{-1}$. Fifty microlitres of a sample between 3 and 5 $mg\cdot mL^{-1}$ were injected. $^{5'}$phosphate-polyUC-OH$^{3'}$-NP complexes with a ratio NP:RNA of 1:1 were incubated for 1 h at room temperature before injection. MALLS detection was performed with a DAWN-HELEOS II detector (Wyatt Technology, Toulouse, France) using a laser emitting at 690 nm and protein concentration was measured on-line with the use of differential refractive-index measurements, with an Optilab T-rEX detector (Wyatt Technology) and

a refractive-index increment (dn/dc) of 0.185 mL·g^{-1}. Weight-average molar masses (Mw) were calculated with ASTRA (Wyatt Technology) as previously described [45].

3. Results

The nucleoproteins of influenza viruses A and B were purified as monomers with only 5% of the protein forming larger oligomers (Figure 2b). The UV spectra of each sample showed no nucleic acid contamination. The B/NP is longer than A/NP because of a longer N-terminal tail (Figure 2a and Figure S1).

In order to assess the influence of both salt and RNA size on the binding of monomeric NP, a series of salt conditions (50, 150 and 300 mM NaCl) were tested against size-incremented polyUC RNA from 5 to 24 nt. The experiments were performed using SEC-MALLS (Figure 3 and Figure S2a,b), negative staining EM (Figures 4 and 5) and fluorescence anisotropy (Figure 6 and Figure S4). A $^{5'}$phosphate-polyUC-OH$^{3'}$ was used for the MALLS experiments and $^{5'}$phosphate-polyUC-FAM$^{3'}$ for the negative staining EM and for the fluorescence anisotropy experiments.

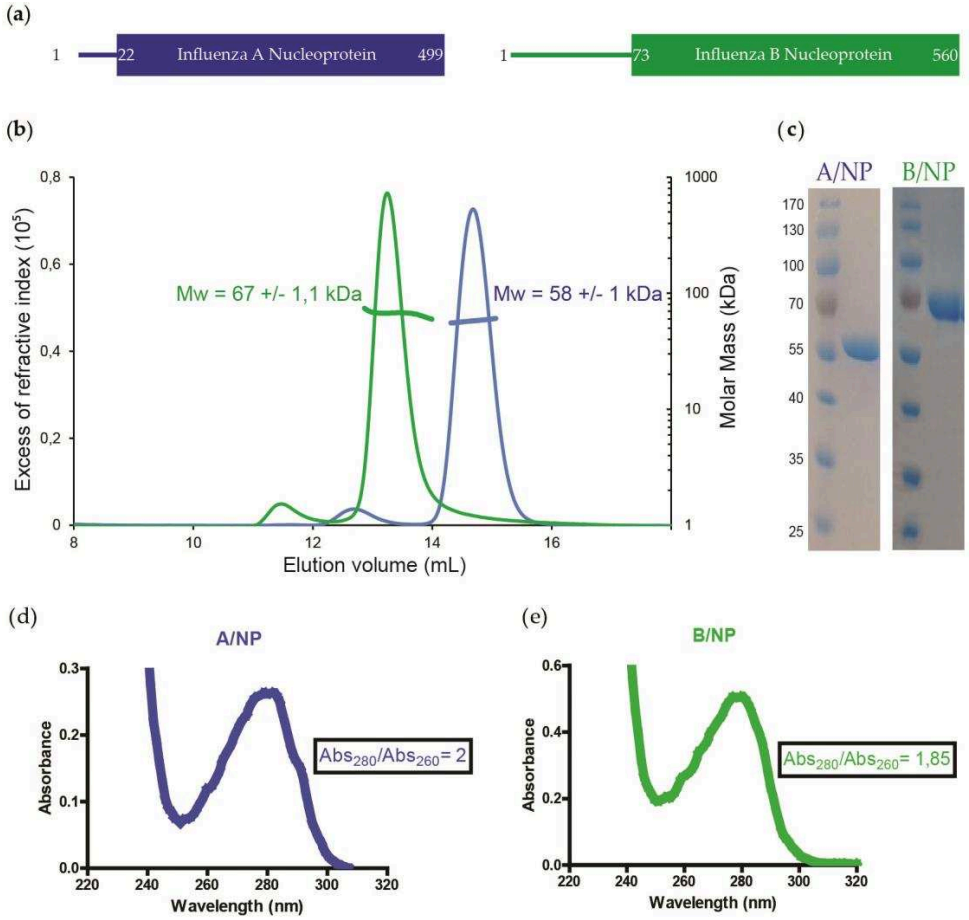


Figure 2. Homogeneous sample analysis of nucleoproteins of influenza A and B viruses. (**a**) Schematic representations of A/NP (blue) and B/NP (green). The A/NP presents a short 21 amino acid N-terminus, whereas B/NP has a longer N-tail of 72 amino acids; (**b**) Comparison of the size exclusion chromatography-multi-angle laser light scattering (SEC-MALLS) analysis of A/NP and B/NP at 50 mM NaCl. Experimental molecular weights, next to the peaks, are coherent with the expected masses of 58 and 65 kDa for A/NP and B/NP, respectively; (**c**) Coomassie blue-stained sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) gel of the purified A/NP and B/NP; Panels (**d**) and (**e**) are, respectively, the ultraviolet (UV) spectra of the A/NP and B/NP. Mw: molecular weight; Abs: absorbance.

3.1. SEC-MALLS

All the SEC-MALLS curves of A/NP and B/NP at 50, 150, and 300 mM NaCl are shown in Figure 3 (Figure S2a,b correspond to individual curves with the respective masses). For A/NP without RNA, the protein is mainly monomeric at 50 and 150 mM NaCl. When NP binds to RNA, the NP–RNA complexes form trimers rapidly. At 300 mM NaCl, there are only a few monomers and most of the protein makes trimers, with or without RNA. Only with a 24-nt RNA do the resulting complexes make larger structure entities. The behaviour of B/NP is quite different since it remains monomeric at 50 and 150 mM NaCl without RNA. With the RNA at 50 mM NaCl, dimers are obtained with 5- and 8-nt RNAs and then tetramers and much bigger entities with longer RNAs. At 150 mM NaCl, all complexes remain monomeric, except when a 24-nt RNA is added, which results in the formation of tetramers and larger complexes. At 300 mM the peak corresponding to the monomer becomes smaller with longer RNAs and we observe clear peaks corresponding to dimers, tetramers, pentamers, and several larger complexes, larger than those of A/NP with RNA, but not trimers.

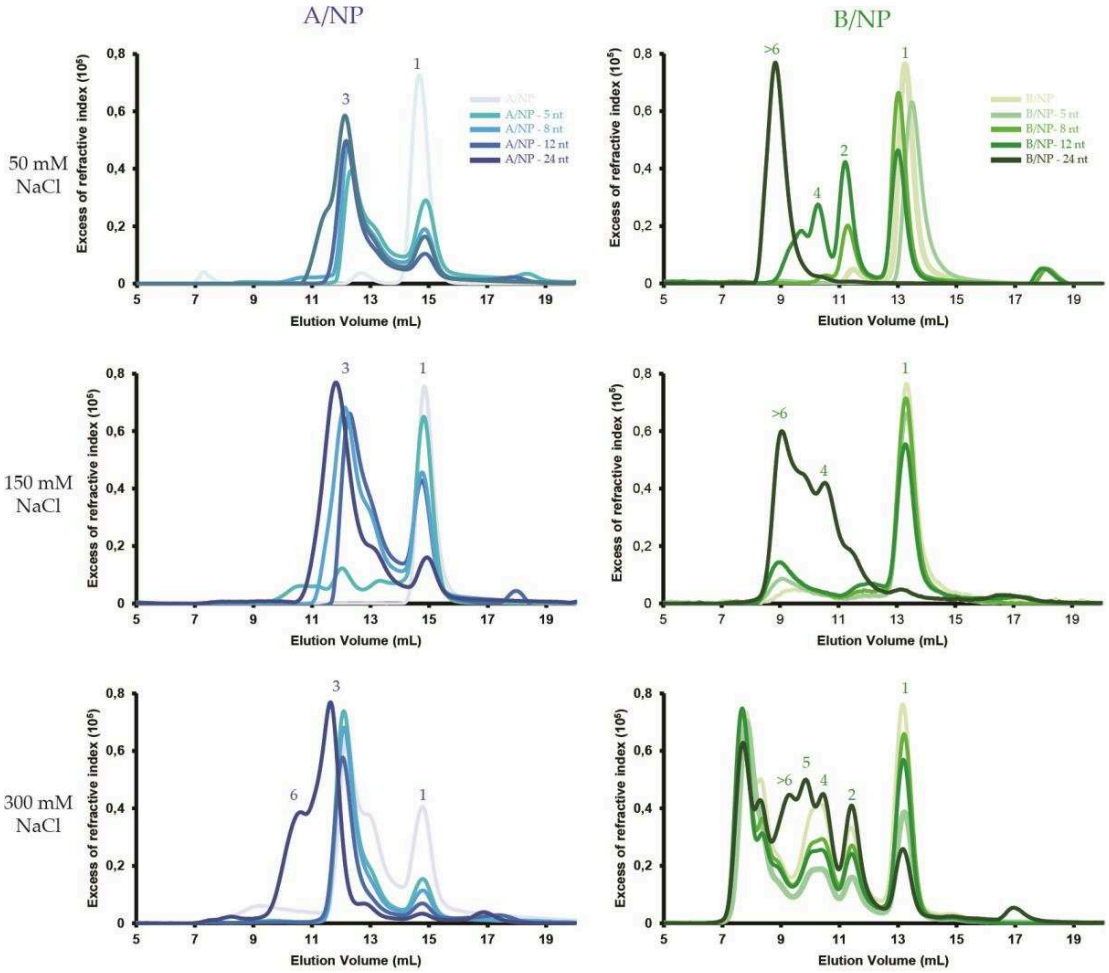


Figure 3. SEC-MALLS analysis of oligomeric states of the nucleoproteins of influenza A and B virus according to RNA length. SEC-MALLS analysis were performed on a Superdex 200 increase column equilibrated with 20 mM Tris-HCl pH 7.5 and different NaCl concentrations, from 50 to 300 mM. Protein is injected alone or in complex (ratio 1:1) with a short $^{5'}$phosphate-polyUC-OH$^{3'}$ RNA from five to 24 nucleotides. Complexes are incubated 1 h at room temperature before injection. For each buffer condition, the relative proportion of oligomers eluted from the column gradually increased upon increasing RNA length. Peaks are labelled 1 for monomer, 2 for dimer, 3 for trimer, 4 for tetramer, 5 for pentamer, and >6 for larger than hexamer.

3.2. Negative Staining EM

For these experiments different RNA molecules were used with a different time of incubation, overnight for EM instead of 1 h as for SEC-MALLS (Figure 4). For A/NP, at 50 mM NaCl, without RNA and with a 5-nt RNA, we observe monomers and with longer RNAs we see mainly trimers, even with 24-nt RNA. These data are coherent with SEC-MALLS results. At 150 and 300 mM NaCl, mainly trimers are observed in all cases, except when including the 12-nt RNA at 150 mM NaCl (this will be discussed later). For B/NP, at 50 mM only monomers are observed until the 12-nt RNA and with longer RNAs we see oligomers larger than with A/NP. At 150 and 300 mM, we observe monomers when RNAs of up to 6-nt are added, and with longer RNAs we see a mixture of dimers, tetramers, pentamers, and larger complexes making rings and "curly strings" [29].

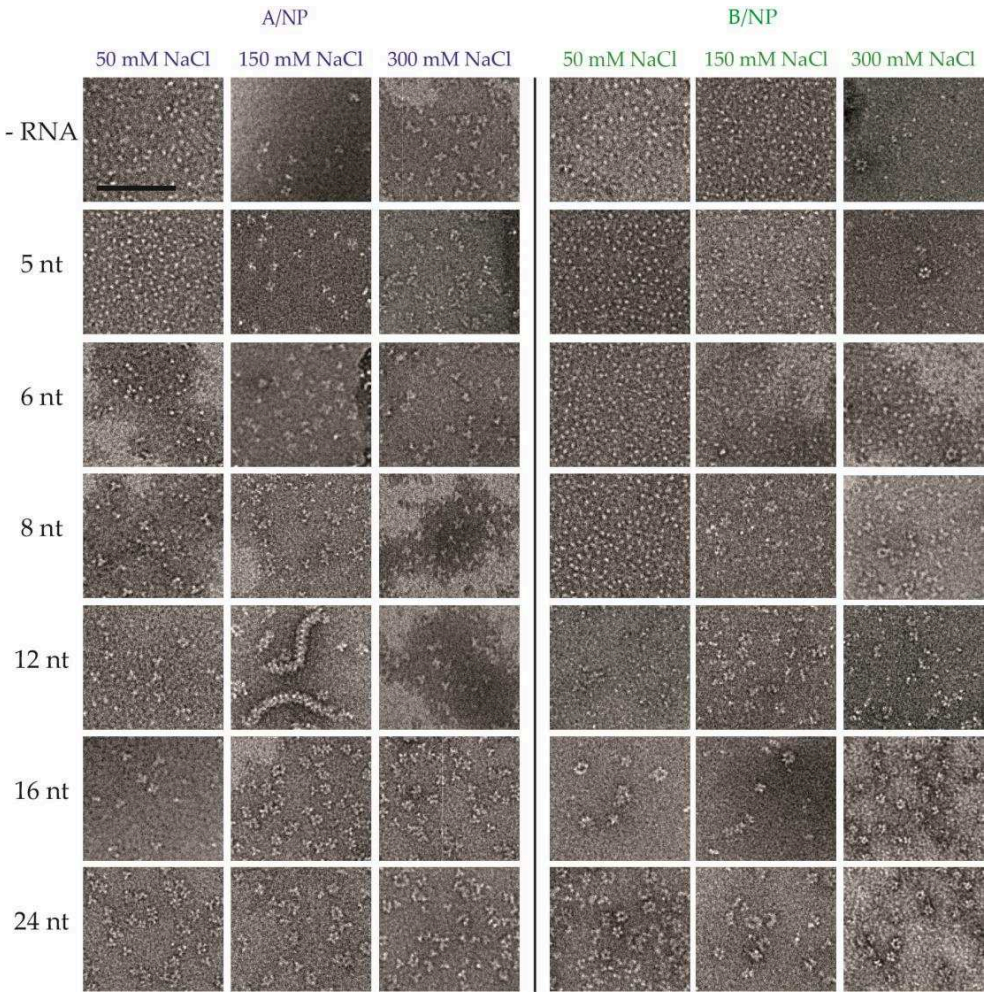


Figure 4. Electron microscopy of the nucleoproteins of influenza A and B viruses in the absence or presence of different RNA and different buffers. The A/NP or B/NP were incubated overnight at room temperature with or without $^{5'}$phosphate-polyUC-FAM$^{3'}$ in 20 mM Tris-HCl pH 7.5, 5 mM β-mercaptoethanol (β-ME), and different NaCl concentrations from 50 to 300 mM. Samples were negatively stained with sodium silicotungstate. Electron micrographs show different oligomeric states depending on the salt concentration and the RNA used. The scale bar corresponds to 100 nm.

In the case of A/NP, but not with B/NP, using a 12-nt $^{5'}$phosphate-$(UC)_6$-FAM$^{3'}$, we can observe oligomeric entities extremely different from all the other conditions. Figure 5 shows a time course of the sample by negative staining EM. Ten minutes after the preparation, the protein looks still monomeric and becomes trimeric after 30 min, then forming short rods. With time, the rods become longer and

after one night at room temperature most of the protein has formed long worm-like entities similar to RNPs with a few remaining trimers. However, with a 12-nt $^{5'}$phosphate-$(UC)_6$ the sample shows mainly trimers and a few larger complexes but no rods or worm-like structures. This means that the 6-fluorescein amidite at the 3′-end of the RNA provokes the reorganization of the NP–RNA complexes to these RNP-like entities. We found that the worm-like structures appear only at 150 mM NaCl, but not at 50 mM or 300 mM NaCl. This phenomenon is observed with fluorescein-labelled RNA ranging from 10 to 13 nt but not with 8 or 16 nt (Figure 4). Even if using 10, 11, and 13 nt RNAs gives similar results, the best homogeneous preparation is obtained with 12 nt.

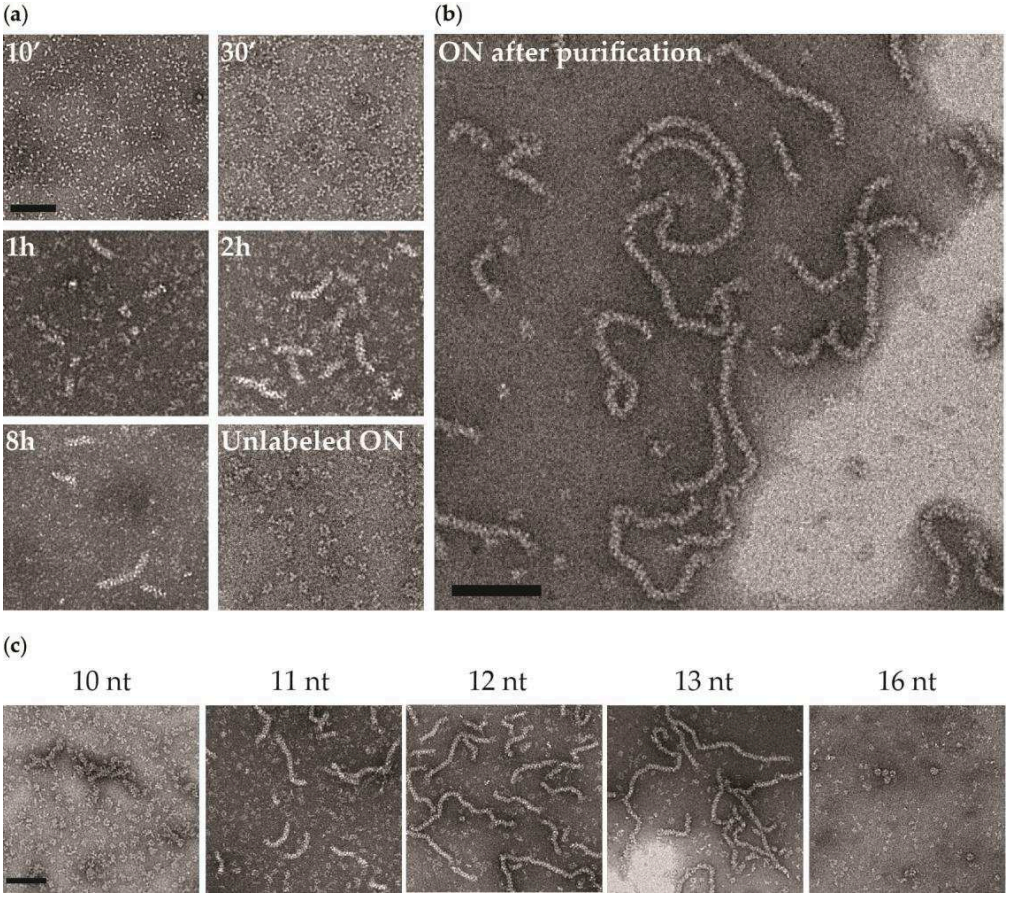


Figure 5. Ribonucleoprotein (RNP)-like species of influenza A virus nucleoprotein. Negative stain electron micrographs of A/NP with $^{5'}$phosphate-$(UC)_6$-FAM$^{3'}$. Complexes (ratio 1:1) were incubated at room temperature in 20 mM Tris-HCl pH 7.5, 150 mM NaCl, and 5 mM β-ME. Samples were negatively stained with sodium silicotungstate. (**a**) Kinetics of RNP-like formation with a 12-nucleotide RNA. Electron micrographs were taken at different times after incubation with several RNAs at room temperature. With an unlabelled RNA, no worm-like structure is observed, even overnight; (**b**) Worm-like structures after purification by centrifugation (10 min, 11,000× *g*, room temperature). The length of these species is random and they appear flexible. Similar to RNPs, these entities seem to be helical; (**c**) RNP-like species obtained with different RNA $^{5'}$phosphate-polyUC-FAM$^{3'}$ lengths. With a 10-nt RNA, only a few short worm-like structures are observed. With 11 nt the worm-like structures are longer. With 12- and 13-nt RNAs, the worms are very similar with fewer trimers/tetramers in the background. The scale bars correspond to 100 nm. ON: overnight.

3.3. Fluorescence Anisotropy Measurements

All the fits were good with R^2 values above 0.98 (Figure S4c). Affinities between NP and RNA were measured using fluorescence anisotropy after 10 min incubation. From then on, the signal did

not change. Figure 6 summarises the values of the K_d in histograms and all the affinities are detailed in Figure S4b,c. In all conditions, A/NP binds RNA with a higher affinity than B/NP. For A/NP at 50 mM NaCl, all the strongest affinities are obtained with 8- and 12-nt RNAs. For 150 and 300 mM NaCl the affinities for the smaller RNAs, between 5 and 12 nt RNAs are much lower than with 50 mM NaCl and at 300 mM NaCl; the 5-nt RNA does not bind the protein. However, the affinity with a 24-nt RNA is enhanced at higher salt. For B/NP we see the same trends as with A/NP, although the RNA with 24 nt RNA does not bind better at a higher salt concentration.

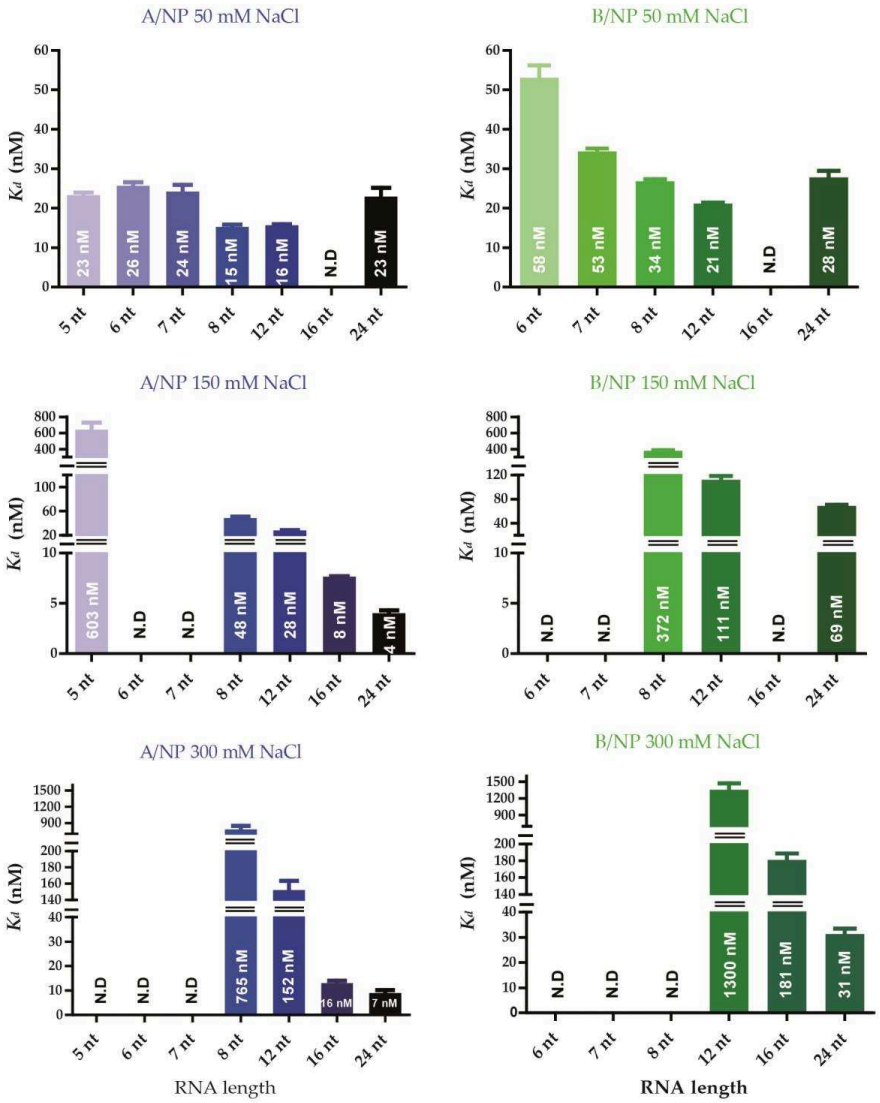


Figure 6. Affinities of both nucleoproteins of influenza A and B viruses for different RNAs. All the fluorescence anisotropy experiments were performed in triplicate in 20 mM Tris-HCl pH 7.5, 5 mM β-ME, and different NaCl concentrations (from 50 to 300 mM). The titration of NP was done against $^{5'}$phosphate-polyUC-FAM$^{3'}$ RNAs ranging from 5 to 24 nt. The mixes were incubated 5 min at room temperature. Curves are shown in Figure S4. N.D: non-determined.

4. Discussion

Using purified recombinant proteins without phosphorylation and with short RNAs, the experiments provide a rationale towards reconstructing viral-like nucleoprotein–RNA complexes. Nucleoproteins from influenza A and B were characterized and we found few biochemical differences between the two proteins.

Both proteins are monomeric at 50 and 150 mM NaCl. Both undergo a time-dependent oligomerization process with RNA, higher salt, and increasing protein concentration [25,27,29,39]. At 300 mM NaCl, both proteins are monomeric and form small oligomers. The experiment shown in Figure 5 and the results obtained by fluorescence anisotropy suggest that the RNA is rapidly bound on a monomeric protein and then, through a slower process, forms oligomers, at least in solution with 50 and 150 mM NaCl.

4.1. Nucleoprotein of A

At 50 mM NaCl we found that NP remains monomeric with small RNAs (up to 6 nt; Figure 7a). With 8 nt and longer RNA molecules, the monomers bind the RNA and then make trimers. At 150 mM NaCl, the monomers bind the RNA and then make trimers and larger complexes with 24-nt RNA. At 300 mM NaCl, part of the protein is already trimeric, so the binding of the RNA is to monomers and trimers and then all the complexes become trimers, making larger complexes with 24-nt RNA. The affinities at 50 mM NaCl are all similar, suggesting that the RNA is bound at the same site on the protein although 12-nt RNA binds slightly better than a 24-nt RNA (Figure 6). With more salt, shorter RNAs bind much less well but 24-nt RNA binds better than at 50 mM NaCl. At 300 mM NaCl, the trend is even stronger. Because the binding with a 24-nt RNA is stronger, this suggests that the RNA binding site is different for a 24-nt RNA at 150 and 300 mM NaCl, probably because under these conditions the RNA binds trimers and not monomers. Because the affinity of the RNA for the trimer is so strong, the trimers may have a problem disassembling and forming longer complexes.

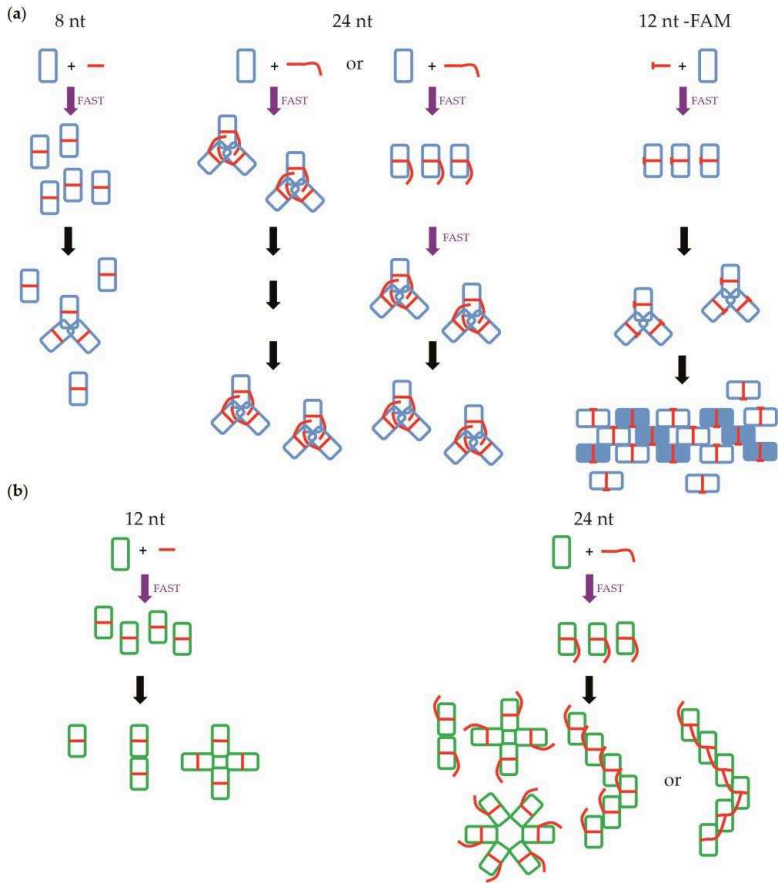


Figure 7. Schematic representation of the oligomerization of influenza A and B viruses' nucleoproteins depending on the size of the RNA molecules at 150 mM NaCl. Panels (**a**) and (**b**) are for A/NP and B/NP, respectively.

4.2. Nucleoprotein of B

For NP of influenza B, we observed similar results (Figure 7b). The only differences are that NP stays monomeric with 12-nt RNA at 50 mM NaCl. With 6-nt RNA at 150 mM and 300 mM NaCl, we still observed monomers. However, we did not obtain trimers but dimers and tetramers, as seen also in Figure 3 with SEC-MALLS experiments. Using a 24-nt RNA, the complexes are larger than those from A/NP. The affinities measurements have shown that the 24-nt RNA binds with a lower affinity than the 12-nt RNA. However, the affinity towards the 24-nt RNA remains comparable at all tested salt concentrations. This suggests that the initial binding of the 24-nt RNA is always on the monomer and not to another binding site.

Three structures of A/NP [25,27,28] and one of B/NP [24] were determined previously. In A/NP crystal structures, the protein is always organized as trimers, whereas it forms tetramers within crystals of B/NP. This difference can be explained by the orientation of the oligomerization loop, which is not the same, leading to different oligomeric forms. The trimeric A/NP seems to be in a more closed conformation than the B/NP tetramer, which appears to be in a more open conformation. Maybe the differences in the affinities observed for A/NP and B/NP for RNA result from the oligomer states. Moreover, the sequence alignments show that the C-terminus of B/NP contains two aspartic acids that are not present in A/NP. This C-terminus might bind in the basic RNA groove and then keep B/NP in a monomeric form with longer RNA than A/NP.

4.3. NP of Influenza A with $^{5'}$Phosphate-(UC)$_6$-FAM$^{3'}$

Many studies use "FAM" for fluorescence anisotropy to measure affinities between proteins and RNAs. Zheng and co-workers [36] have used the same method for the binding of NP of infectious salmon anaemia virus with a $^{5'}$fluorescein labelled-(UC)$_6$. Figure 5 shows that the RNA binds to the monomer and then the NP-RNA makes trimers, larger complexes, and flexible worm-like structures. These structures have the same diameter as the RNPs purified from influenza virus particles [29] and show fine lines, suggesting that these structures are helical. They are only obtained with fluorescein-labelled RNA. Without the fluorescein, NP assembles as trimers and does not generate these longer worm-like structures (Figure 5). With shorter RNAs, these structures are not observed either, probably because the NP-RNA remains monomeric or the reaction is too slow to be observed. With longer RNAs ($\geq$16 nt RNA), the same is observed, maybe because the NP-RNA makes trimers rapidly and stops the growth of these worm-like structures. There is also an effect of the salt concentration: at 50 mM NaCl, the complexes remain monomeric, whereas at 300 mM they make trimers. Figure 6 suggests that the 12-nt RNA binds to the monomer at 50 mM but the 24-nt RNA could be too long for a monomer. The structure of the monomeric A/NP-R416A [28] is different from the trimer [27]. One side of the monomer has the C-terminus bound to the RNA binding site, changing the charge of the RNA groove. Zheng et al. concluded that each monomer of NP of ISAV binds about 12 bases but the RNA in the RNPs binds additional parts of the proteins [36]. In our experiments, only the monomeric NP with a $^{5'}$phosphate-(UC)$_6$-FAM$^{3'}$ can form worm-like structures. However, it is still to be determined why the labelled RNA makes these interesting entities that are not observed with the unlabelled RNA. However, interestingly, we also show by MALLS that A/NP and B/NP, in complex with labelled RNA, have a tendency to form higher oligomeric species at 150 mM NaCl compared to RNA without FAM (Figure S3). Fluorescein may thus have a role in the oligomerization process even if it does not bind NP directly (Figure S4a).

5. Conclusions

In this paper, we tried to understand the oligomerization process of A/NP and B/NP with respect to the salt concentration and the RNA length. Our results show many differences between the two proteins but it remains difficult to explain these without crystal structures of NP with RNA. However, we show that to crystallise the protein in the presence of RNA, it is necessary to find

a compromise between the salt concentration and the RNA length. An unlabelled 12-nt RNA could be a good compromise between these two criteria because it binds strongly to the protein and is not too long to interfere with the crystallisation process. We also show that it is possible to reconstitute an RNP-like structure using recombinant NP in the absence of the RNA polymerase. We cannot yet explain why we can only have this species with a labelled RNA but it is a very interesting tool to start further investigations towards better understanding the interaction between NP and RNA.

Supplementary Materials: The following are available online at www.mdpi.com/1999-4915/8/9/247/s1, Figure S1: Sequence alignment of the nucleoproteins of influenza A and B viruses; Figure S2a: Oligomeric states of influenza A virus nucleoprotein according to RNA length by size exclusion chromatography-multi-angle laser light scattering (SEC-MALLS) analysis; Figure S2b: Oligomeric states of influenza B virus nucleoprotein according to RNA length by SEC-MALLS analysis; Figure S3: Oligomeric state of nucleoproteins of influenza A and B viruses using polyUC-FAM$^{3'}$ RNAs by SEC-MALLS analysis; Figure S4: Titration measurements against RNAs by fluorescence anisotropy. References [46,47] are cited in the supplementary materials.

Acknowledgments: We are very grateful to Daphna Fenel, Laura Tengo, Myriam Miloudi, Caroline Mas, Joanna Timmins, and Sigrid Milles for their help. We thank Darren J. Hart for a critical reading of the manuscript. Alice Labaronne was funded through the Labex GRAL (ANR-10-LABX-49-01), Christopher Swale by the Hofmann-La Roche pRED external collaboration program, and Alexandre Monod by the Région Rhône-Alpes (ARC 1 Santé). This work was supported by the French Agency for Research through the ANR RNAP-IAV (ANR-14-CE09-0017) and the European Commission project Flupharm (contract number 259751). This study used the platforms of the Grenoble Instruct centre (ISBG; UMS 3518 CNRS-CEA-UJF-EMBL) with support from FRISBI (ANR-10-INSB-05-02) and GRAL within the Grenoble Partnership for Structural Biology (PSB). The electron microscope facility was supported by the Rhône-Alpes Region, the Fondation pour la Recherche Medicale (FRM), and FEDER funds.

Author Contributions: A.L. and R.W.H.R. conceived and designed the experiments; A.L., C.S., and A.M. performed the experiments; A.L., G.S., T.C., and R.W.H.R. analysed the data; A.L., T.C., and R.W.H.R. wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Reguera, J.; Cusack, S.; Kolakofsky, D. Segmented negative strand RNA virus nucleoprotein structure. *Curr. Opin. Virol.* **2014**, *5*, 7–15.
2. Ruigrok, R.W.; Crepin, T.; Kolakofsky, D. Nucleoproteins and nucleocapsids of negative-strand RNA viruses. *Curr. Opin. Microbiol.* **2011**, *14*, 504–510. [CrossRef] [PubMed]
3. Albertini, A.A.; Wernimont, A.K.; Muziol, T.; Ravelli, R.B.; Clapier, C.R.; Schoehn, G.; Weissenhorn, W.; Ruigrok, R.W. Crystal structure of the rabies virus nucleoprotein-RNA complex. *Science* **2006**, *313*, 360–363. [CrossRef] [PubMed]
4. Green, T.J.; Zhang, X.; Wertz, G.W.; Luo, M. Structure of the vesicular stomatitis virus nucleoprotein-RNA complex. *Science* **2006**, *313*, 357–360. [CrossRef] [PubMed]
5. Renner, M.; Bertinelli, M.; Leyrat, C.; Paesen, G.C.; Saraiva de Oliveira, L.F.; Huiskonen, J.T.; Grimes, J.M. Nucleocapsid assembly in pneumoviruses is regulated by conformational switching of the N protein. *Elife* **2016**, *5*, e12627. [CrossRef] [PubMed]
6. Tawar, R.G.; Duquerroy, S.; Vonrhein, C.; Varela, P.F.; Damier-Piolle, L.; Castagne, N.; MacLellan, K.; Bedouelle, H.; Bricogne, G.; Bhella, D.; et al. Crystal structure of a nucleocapsid-like nucleoprotein-RNA complex of respiratory syncytial virus. *Science* **2009**, *326*, 1279–1283. [CrossRef] [PubMed]
7. Gutsche, I.; Desfosses, A.; Effantin, G.; Ling, W.L.; Haupt, M.; Ruigrok, R.W.; Sachse, C.; Schoehn, G. Structural virology. Near-atomic cryo-EM structure of the helical measles virus nucleocapsid. *Science* **2015**, *348*, 704–707. [PubMed]
8. Alayyoubi, M.; Leser, G.P.; Kors, C.A.; Lamb, R.A. Structure of the paramyxovirus parainfluenza virus 5 nucleoprotein-RNA complex. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E1792–E1799. [CrossRef] [PubMed]
9. Calain, P.; Roux, L. The rule of six, a basic feature for efficient replication of Sendai virus defective interfering RNA. *J. Virol.* **1993**, *67*, 4822–4830. [PubMed]
10. Egelman, E.H.; Wu, S.S.; Amrein, M.; Portner, A.; Murti, G. The Sendai virus nucleocapsid exists in at least four different helical states. *J. Virol.* **1989**, *63*, 2233–2243. [PubMed]

11. Kolakofsky, D.; Pelet, T.; Garcin, D.; Hausmann, S.; Curran, J.; Roux, L. Paramyxovirus RNA synthesis and the requirement for hexamer genome length: The rule of six revisited. *J. Virol.* **1998**, *72*, 891–899. [PubMed]
12. Hastie, K.M.; Kimberlin, C.R.; Zandonatti, M.A.; MacRae, I.J.; Saphire, E.O. Structure of the Lassa virus nucleoprotein reveals a dsRNA-specific 3′ to 5′ exonuclease activity essential for immune suppression. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 2396–2401. [CrossRef] [PubMed]
13. Hastie, K.M.; Liu, T.; Li, S.; King, L.B.; Ngo, N.; Zandonatti, M.A.; Woods, V.L., Jr.; de la Torre, J.C.; Saphire, E.O. Crystal structure of the Lassa virus nucleoprotein-RNA complex reveals a gating mechanism for RNA binding. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 19365–19370. [CrossRef] [PubMed]
14. Raymond, D.D.; Piper, M.E.; Gerrard, S.R.; Skiniotis, G.; Smith, J.L. Phleboviruses encapsidate their genomes by sequestering RNA bases. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 19208–19213. [CrossRef] [PubMed]
15. Ariza, A.; Tanner, S.J.; Walter, C.T.; Dent, K.C.; Shepherd, D.A.; Wu, W.; Matthews, S.V.; Hiscox, J.A.; Green, T.J.; Luo, M.; et al. Nucleocapsid protein structures from orthobunyaviruses reveal insight into ribonucleoprotein architecture and RNA polymerization. *Nucleic Acids Res.* **2013**, *41*, 5912–5926. [CrossRef] [PubMed]
16. Dong, H.; Li, P.; Bottcher, B.; Elliott, R.M.; Dong, C. Crystal structure of Schmallenberg orthobunyavirus nucleoprotein-RNA complex reveals a novel RNA sequestration mechanism. *RNA* **2013**, *19*, 1129–1136. [CrossRef] [PubMed]
17. Dong, H.; Li, P.; Elliott, R.M.; Dong, C. Structure of Schmallenberg orthobunyavirus nucleoprotein suggests a novel mechanism of genome encapsidation. *J. Virol.* **2013**, *87*, 5593–5601. [CrossRef] [PubMed]
18. Niu, F.; Shaw, N.; Wang, Y.E.; Jiao, L.; Ding, W.; Li, X.; Zhu, P.; Upur, H.; Ouyang, S.; Cheng, G.; Liu, Z.J. Structure of the Leanyer orthobunyavirus nucleoprotein-RNA complex reveals unique architecture for RNA encapsidation. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 9054–9059. [CrossRef] [PubMed]
19. Reguera, J.; Malet, H.; Weber, F.; Cusack, S. Structural basis for encapsidation of genomic RNA by La Crosse Orthobunyavirus nucleoprotein. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 7246–7251. [CrossRef] [PubMed]
20. Dong, S.; Yang, P.; Li, G.; Liu, B.; Wang, W.; Liu, X.; Xia, B.; Yang, C.; Lou, Z.; Guo, Y.; Rao, Z. Insight into the Ebola virus nucleocapsid assembly mechanism: Crystal structure of Ebola virus nucleoprotein core domain at 1.8 A resolution. *Protein Cell* **2015**, *6*, 351–362. [CrossRef] [PubMed]
21. Dziubanska, P.J.; Derewenda, U.; Ellena, J.F.; Engel, D.A.; Derewenda, Z.S. The structure of the C-terminal domain of the Zaire ebolavirus nucleoprotein. *Acta Crystallogr. D Biol. Crystallogr.* **2014**, *70*, 2420–2429. [CrossRef] [PubMed]
22. Kirchdoerfer, R.N.; Abelson, D.M.; Li, S.; Wood, M.R.; Saphire, E.O. Assembly of the Ebola Virus Nucleoprotein from a Chaperoned VP35 Complex. *Cell Rep.* **2015**, *12*, 140–149. [CrossRef] [PubMed]
23. Leung, D.W.; Borek, D.; Luthra, P.; Binning, J.M.; Anantpadma, M.; Liu, G.; Harvey, I.B.; Su, Z.; Endlich-Frazier, A.; Pan, J.; et al. An Intrinsically Disordered Peptide from Ebola Virus VP35 Controls Viral RNA Synthesis by Modulating Nucleoprotein-RNA Interactions. *Cell Rep.* **2015**, *11*, 376–389. [CrossRef] [PubMed]
24. Ng, A.K.; Lam, M.K.; Zhang, H.; Liu, J.; Au, S.W.; Chan, P.K.; Wang, J.; Shaw, P.C. Structural basis for RNA binding and homo-oligomer formation by influenza B virus nucleoprotein. *J. Virol.* **2012**, *86*, 6758–6767. [CrossRef] [PubMed]
25. Ng, A.K.; Zhang, H.; Tan, K.; Li, Z.; Liu, J.H.; Chan, P.K.; Li, S.M.; Chan, W.Y.; Au, S.W.; Joachimiak, A.; et al. Structure of the influenza virus A H5N1 nucleoprotein: Implications for RNA binding, oligomerization, and vaccine design. *FASEB J.* **2008**, *22*, 3638–3647. [CrossRef] [PubMed]
26. Rudolph, M.G.; Kraus, I.; Dickmanns, A.; Eickmann, M.; Garten, W.; Ficner, R. Crystal structure of the borna disease virus nucleoprotein. *Structure* **2003**, *11*, 1219–1226. [CrossRef] [PubMed]
27. Ye, Q.; Krug, R.M.; Tao, Y.J. The mechanism by which influenza A virus nucleoprotein forms oligomers and binds RNA. *Nature* **2006**, *444*, 1078–1082. [CrossRef] [PubMed]
28. Chenavas, S.; Estrozi, L.F.; Slama-Schwok, A.; Delmas, B.; Di Primo, C.; Baudin, F.; Li, X.; Crépin, T.; Ruigrok, R.W. Monomeric nucleoprotein of influenza A virus. *PLoS Pathog.* **2013**, *9*, e1003275. [CrossRef] [PubMed]
29. Ruigrok, R.W.; Baudin, F. Structure of influenza virus ribonucleoprotein particles. II. Purified RNA-free influenza virus ribonucleoprotein forms structures that are indistinguishable from the intact influenza virus ribonucleoprotein particles. *J. Gen. Virol.* **1995**, *76*, 1009–1014. [PubMed]

30. Compans, R.W.; Content, J.; Duesberg, P.H. Structure of the ribonucleoprotein of influenza virus. *J. Virol.* **1972**, *10*, 795–800. [PubMed]
31. Duesberg, P.H. Distinct subunits of the ribonucleoprotein of influenza virus. *J. Mol. Biol.* **1969**, *42*, 485–499. [CrossRef]
32. Kingsbury, D.W.; Webster, R.G. Some properties of influenza virus nucleocapsids. *J. Virol.* **1969**, *4*, 219–225. [PubMed]
33. Pons, M.W.; Schulze, I.T.; Hirst, G.K.; Hauser, R. Isolation and characterization of the ribonucleoprotein of influenza virus. *Virology* **1969**, *39*, 250–259. [CrossRef]
34. Ortega, J.; Martin-Benito, J.; Zurcher, T.; Valpuesta, J.M.; Carrascosa, J.L.; Ortin, J. Ultrastructural and functional analyses of recombinant influenza virus ribonucleoproteins suggest dimerization of nucleoprotein during virus amplification. *J. Virol.* **2000**, *74*, 156–163. [CrossRef] [PubMed]
35. Hutchinson, E.C.; Charles, P.D.; Hester, S.S.; Thomas, B.; Trudgian, D.; Martinez-Alonso, M.; Fodor, E. Conserved and host-specific features of influenza virion architecture. *Nat. Commun.* **2014**, *5*, 4816. [CrossRef] [PubMed]
36. Zheng, W.; Olson, J.; Vakharia, V.; Tao, Y.J. The crystal structure and RNA-binding of an orthomyxovirus nucleoprotein. *PLoS Pathog.* **2013**, *9*, e1003624. [CrossRef] [PubMed]
37. Baudin, F.; Bach, C.; Cusack, S.; Ruigrok, R.W. Structure of influenza virus RNP. I. Influenza virus nucleoprotein melts secondary structure in panhandle RNA and exposes the bases to the solvent. *EMBO J.* **1994**, *13*, 3158–3165. [PubMed]
38. Chan, W.H.; Ng, A.K.; Robb, N.C.; Lam, M.K.; Chan, P.K.; Au, S.W.; Wang, J.H.; Fodor, E.; Shaw, P.C. Functional analysis of the influenza virus H5N1 nucleoprotein tail loop reveals amino acids that are crucial for oligomerization and ribonucleoprotein activities. *J. Virol.* **2010**, *84*, 7337–7345. [CrossRef] [PubMed]
39. Tarus, B.; Bakowiez, O.; Chenavas, S.; Duchemin, L.; Estrozi, L.F.; Bourdieu, C.; Lejal, N.; Bernard, J.; Moudjou, M.; Chevalier, C.; et al. Oligomerization paths of the nucleoprotein of influenza A virus. *Biochimie* **2012**, *94*, 776–785. [CrossRef] [PubMed]
40. Elton, D.; Medcalf, L.; Bishop, K.; Harrison, D.; Digard, P. Identification of amino acid residues of influenza virus nucleoprotein essential for RNA binding. *J. Virol.* **1999**, *73*, 7357–7367. [PubMed]
41. Boulo, S.; Akarsu, H.; Lotteau, V.; Muller, C.W.; Ruigrok, R.W.; Baudin, F. Human importin alpha and RNA do not compete for binding to influenza A virus nucleoprotein. *Virology* **2011**, *409*, 84–90. [CrossRef] [PubMed]
42. Ye, Q.; Guu, T.S.; Mata, D.A.; Kuo, R.L.; Smith, B.; Krug, R.M.; Tao, Y.J. Biochemical and structural evidence in support of a coherent model for the formation of the double-helical influenza A virus ribonucleoprotein. *MBio* **2012**, *4*, e00467-12. [CrossRef] [PubMed]
43. Fender, P.; Moriscot, C.; Ruigrok, R.W.; Schoehn, G. Electron Microscopy of Viruses: Techniques to Prepare Viruses and Viral Proteins for Observation by Electron Microscopy. In *Reference Module in Biomedical Sciences*; Elsevier: Amsterdam, The Netherlands, 2014.
44. Valentine, R.C.; Wrigley, N.G.; Scrutton, M.C.; Irias, J.J.; Utter, M.F. Pyruvate carboxylase. 8. The subunit structure as examined by electron microscopy. *Biochemistry* **1966**, *5*, 3111–3116. [PubMed]
45. Gerard, F.C.; Ribeiro Ede, A., Jr.; Leyrat, C.; Ivanov, I.; Blondel, D.; Longhi, S.; Ruigrok, R.W.; Jamin, M. Modular organization of rabies virus phosphoprotein. *J. Mol. Biol.* **2009**, *388*, 978–996. [CrossRef] [PubMed]
46. Larkin, M.A.; Blackshields, G.; Brown, N.P.; Chenna, R.; McGettigan, P.A.; McWilliam, H.; Valentin, F.; Wallace, I.M.; Wilm, A.; Lopez, R.; et al. Clustal W and Clustal X version 2.0. *Bioinformatics* **2007**, *23*, 2947–2948. [CrossRef] [PubMed]
47. Robert, X.; Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **2014**, *42*, W320–W324. [CrossRef] [PubMed]

© 2016 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/).

Supplementary Materials: Binding of RNA by the Nucleoproteins of Influenza Viruses A and B

Alice Labaronne, Christopher Swale, Alexandre Monod, Guy Schoehn, Thibaut Crépin and Rob W. H. Ruigrok

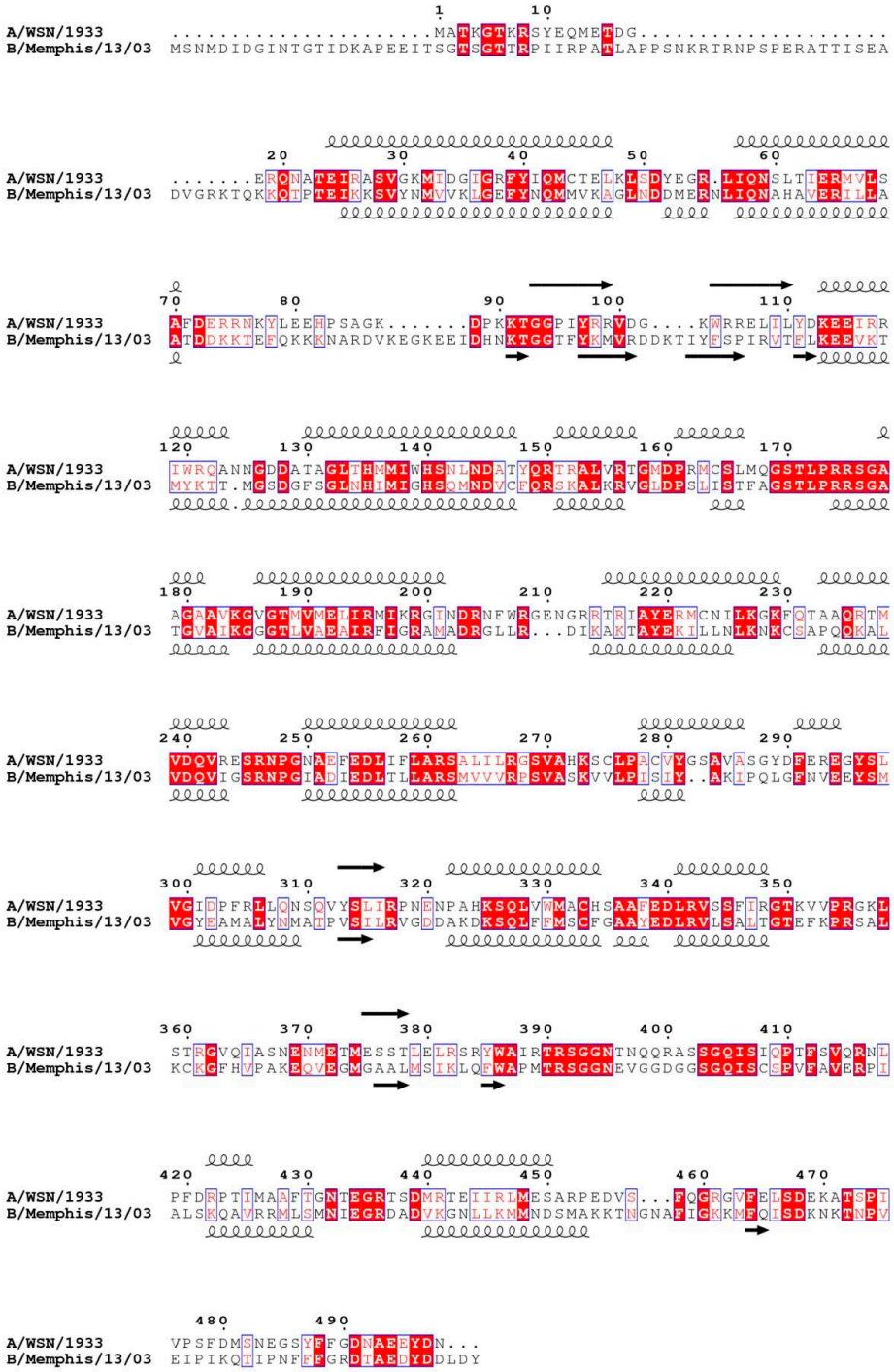


Figure S1. Sequence alignment of the nucleoproteins of influenza A and B viruses. The alignment has been made with the sequences of the two human influenza A and influenza B viruses used for this work: A/WSN/1933 (UniProtKB access number: Q1I2B5) and B/Memphis/13/03 (UniProtKB access number: Q5V8X3). The secondary sequences elements shown over and below the sequences alignment correspond to A/NP, A/WSN/1933 (PDB ID: 2IQH) and B/NP, B/Managua/4577.01/2008 (PDB ID: 3TJ0). The sequence identity between these two B/NP is 98.6%. The alignment has been done using Clustal W [46] and drawn with ESPript [47].

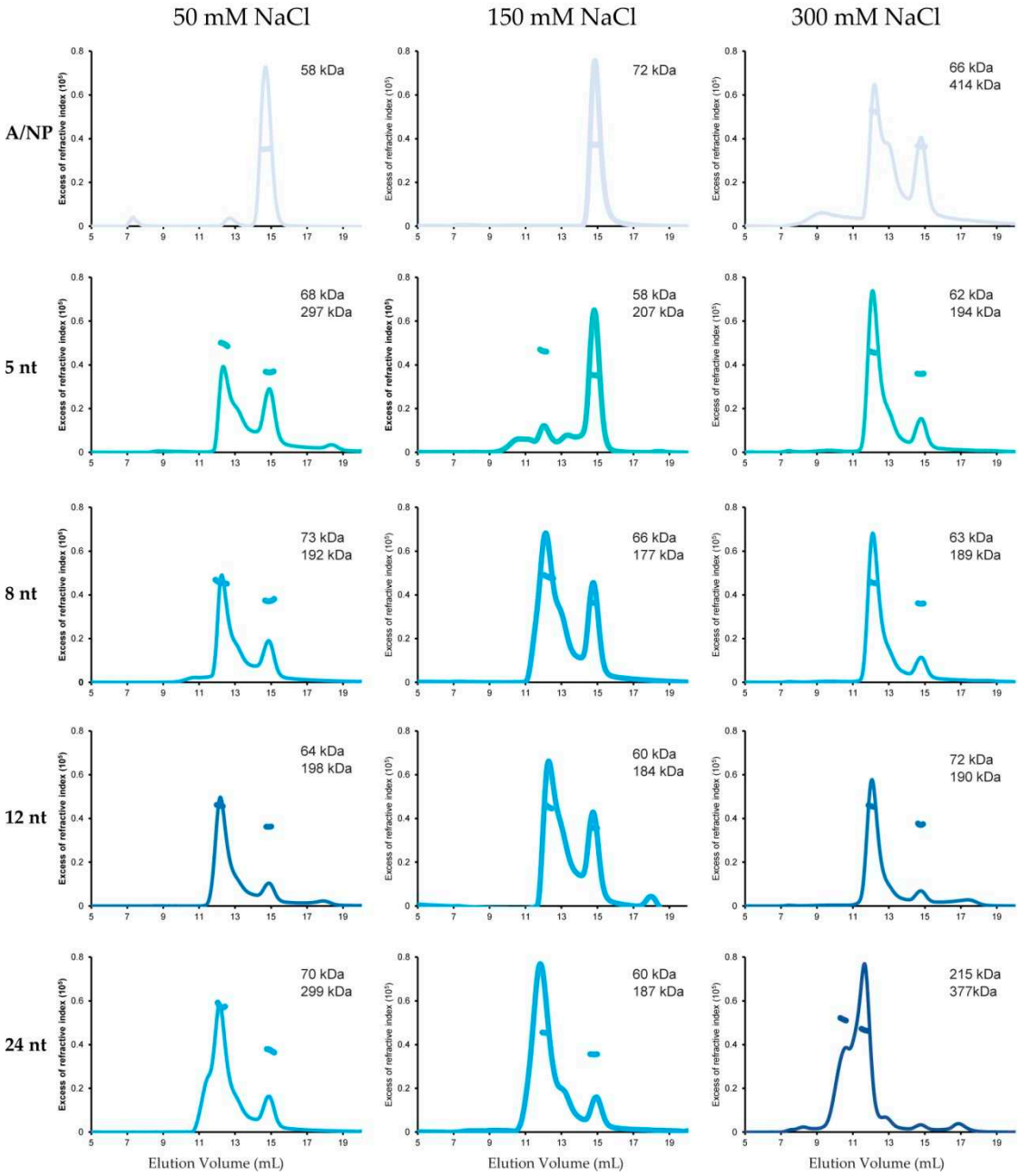


Figure S2a. Oligomeric states of influenza A virus nucleoprotein according to RNA length by size exclusion chromatography-multi-angle laser light scattering (SEC-MALLS) analysis. SEC-MALLS analysis were performed on a Superdex 200 increase 10/300 GL column equilibrated with 20 mM Tris-HCl pH 7.5 and different concentrations of NaCl (from 50 to 300 mM). Protein was injected alone or in complex (ratio 1:1) with short $^{5'}$phosphate-polyUC-OH$^{3'}$ RNAs (from 5 to 24 nt). Complexes were incubated 1 h at room temperature before injection. SEC-MALLS chromatograms are represented with the refractive index as a back trace and molecular weights are estimated at the left.

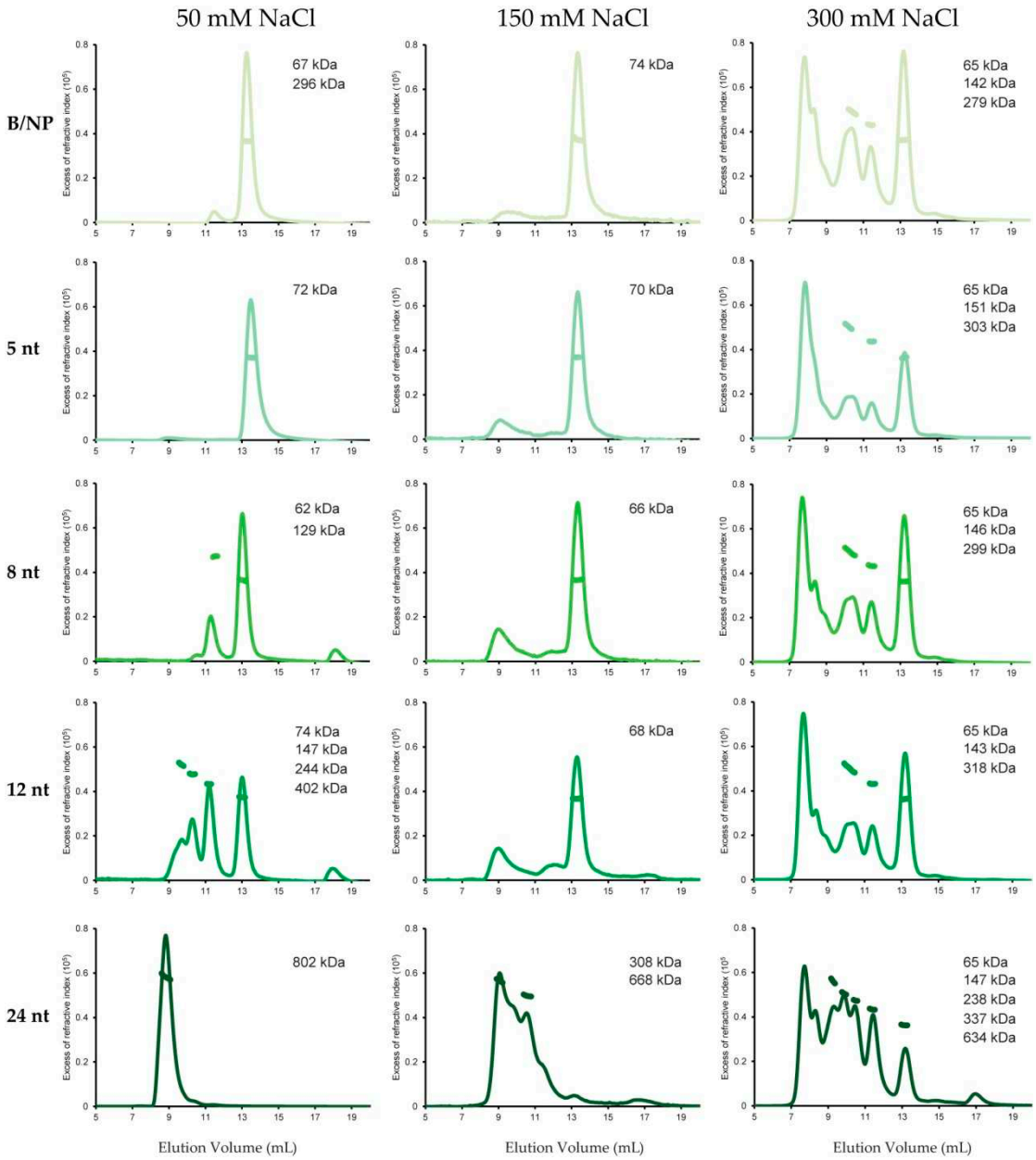


Figure S2b. Oligomeric states of influenza B virus nucleoprotein according to RNA length by SEC-MALLS analysis. SEC-MALLS analysis were performed on a Superdex 200 increase 10/300 GL column equilibrated with 20 mM Tris-HCl pH 7.5 and different concentrations of NaCl (from 50 to 300 mM). Protein was injected alone or in complex (ratio 1:1) with short $^{5'}$phosphate-polyUC-OH$^{3'}$ RNAs (from 5 to 24 nt). Complexes were incubated 1 h at room temperature before injection. SEC-MALLS chromatograms are represented with the refractive index as a back trace and molecular weights are estimated at the left.

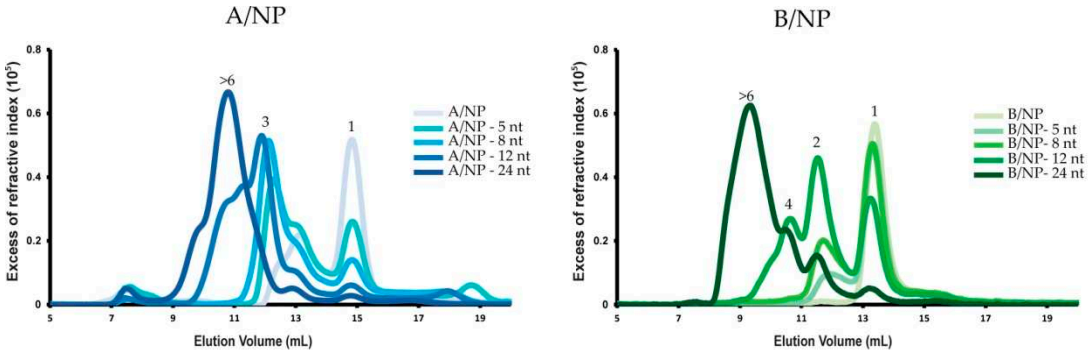
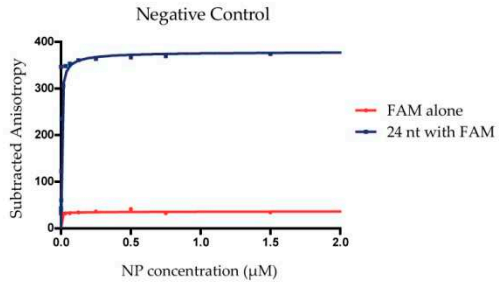
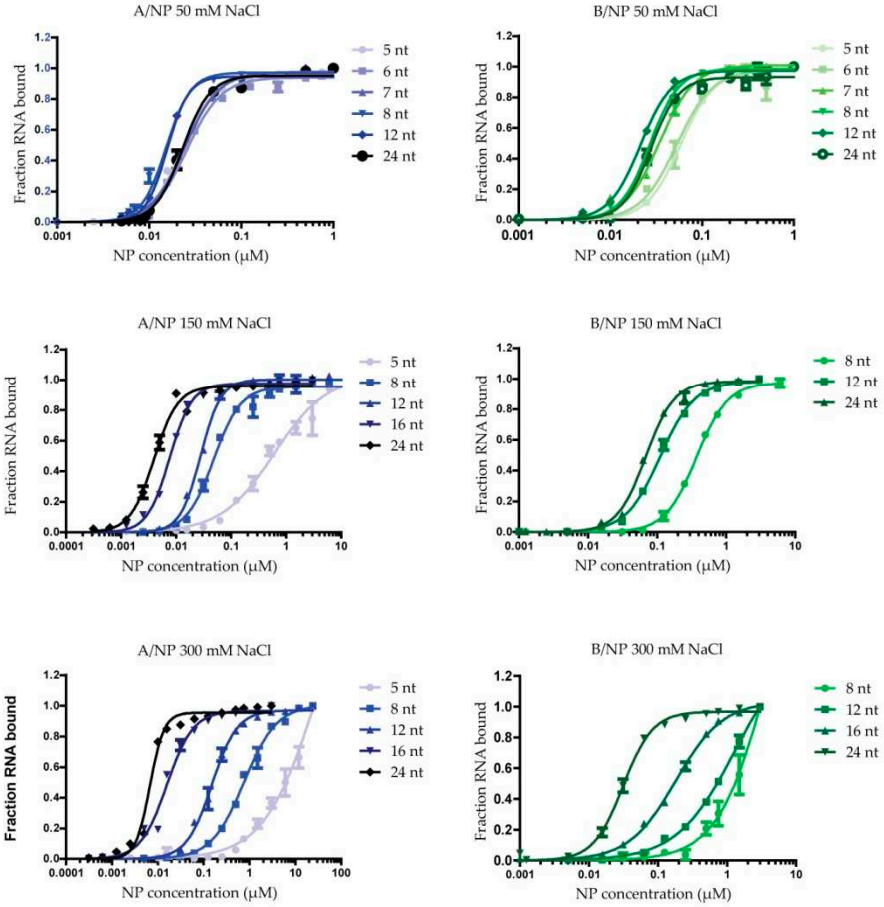


Figure S3. Oligomeric state of nucleoproteins of influenza A and B viruses using polyUC-FAM$^{3'}$ RNAs by SEC-MALLS analysis. SEC-MALLS analysis were performed on a Superdex 200 increase 10/300 GL column equilibrated with 20 mM Tris-HCl pH 7.5 at 150 mM NaCl. The protein was injected alone or in complex (ratio 1:1) with a single short $^{5'}$phosphate-polyUC-FAM$^{3'}$ RNA (5 to 24 nucleotides). The complexes were incubated 1 h at room temperature before injection. Peaks are labeled 1 for monomer, 2 for dimer, 3 for trimer, 4 for tetramer and 6 for hexamer.

(**a**)



(**b**)



(**c**)

	50 mM NaCl		150 mM NaCl		300 mM NaCl	
RNA	K_d **(nM)**	R^2	K_d **(nM)**	R^2	K_d **(nM)**	R^2
5 nt	23 ± 0.8	0.9952	603 ± 93	0.9736		
6 nt	26 ± 0.4	0.9923				
7 nt	24 ± 0.5	0.9943				
8 nt	15 ± 0.8	0.9852	48 ± 2	0.9890	765 ± 40	0.9925
12 nt	16 ± 0.8	0.9908	28 ± 0.8	0.9925	151 ± 8	0.9899
16 nt			8 ± 0.2	0.9969	16 ± 0.8	0.9874
24 nt	23 ± 0.6	0.9917	4 ± 0.2	0.9873	7 ± 0.2	0.9917

	50 mM NaCl		150 mM NaCl		300 mM NaCl	
RNA	K_d **(nM)**	R^2	K_d **(nM)**	R^2	K_d (nM)	R^2
5 nt	58 ± 0.8	0.9978				
6 nt	53 ± 3	0.9748				
7 nt	34 ± 0.8	0.9919				
8 nt	27 ± 0.4	0.9985	372 ± 8	0.9973		
12 nt	21 ± 0.5	0.9968	113 ± 3	0.9975	1306 ± 167	0.9953
16 nt					0.9857	0.9974
24 nt	28 ± 0.8	0.9857	69 ± 2	0.9963	31 ± 0.5	0.9939

Figure S4. Titration measurements against RNAs by fluorescence anisotropy. All the experiments have been made in triplicate in 20 mM Tris-HCl pH 7.5, 5 mM β-mercaptoethanol (β-ME) and different salt concentrations (from 50 to 300 mM NaCl). The mixes were incubated 5 min at room temperature; (**a**) Binding titration of A/NP at 50 mM NaCl using 6-fluorescein amidite (FAM) alone; (**b**) Binding titration of A/NP and B/NP at 50, 150 and 300 mM NaCl using different length of short RNAs (from 5 to 24 nt); (**c**) Table of the different values obtained from the titrations.

Chapitre 4 : Caractérisation biochimique et structurale de NP D

Journal of virology, article sous révision

Article 2: Structure of influenza D virus nucleoprotein and function of NP_{Tail}

Alice Labaronne, Myriam Miloudi, Laura Tengo, Caroline Mas, Guy Schoehn, Rob W.H. Ruigrok et Thibaut Crépin

Note d'introduction au chapitre 4

En 2011 un nouveau membre de la famille des *Orthomyxoviridae* a été isolé à partir d'un porc (Hause et al., 2013) . Ce virus, partageant 50 % d'identité de séquence avec le virus *influenza* C humain, a d'abord été apparenté à ce dernier. Finalement il a été classé en tant que nouveau membre de la famille des *Orthomyxoviridae* comptant désormais 7 genres : les virus *influenza* de type A, B, C et D, le virus de l'anémie infectieuse du saumon, le virus Thogoto ainsi que les Quaranfilivirus.

Les ribonucléoprotéines constituent l'unité fonctionnelle du virus de la grippe assurant à la fois la réplication et la transcription du virus au sein du noyau des cellules infectées. Chaque RNP est constituée d'un brin d'ARN négatif recouvert de multiples copies de nucléoprotéines. Les différents virus *influenza* se distinguent principalement par le nombre de RNP : les virus de types A et B en possèdent 8 alors que les virus de types C et D en possèdent seulement 7.

Plusieurs structures de nucléoprotéines d'*Orthomyxoviridae* sont disponibles (Chenavas et al., 2013; Ng et al., 2008, 2012a; Ye et al., 2006; Zheng et al., 2013). Bien que ces structures ont été obtenues en absence d'ARN, il a été montré depuis longtemps que le cœur des nucléoprotéines contient un large domaine basique impliqué dans la fixation de ce dernier (Baudin et al., 1994; Elton et al., 1999; Ng et al., 2008; Ye et al., 2006). Cependant le nombre de nucléotides fixant la protéine est toujours en débat. La comparaison des structures obtenues permet de mettre en évidence un repliement très similaire entre les NP des virus *influenza* A et B avec un rmsd de 1.6 Å sur les carbones α. Une différence plus marquée est observée avec la NP du virus ISA, présentant un domaine amino-terminal additionnel replié et une extension carboxy-terminale à l'inverse de NP A et NP B qui possèdent une extension amino-terminale désordonnée. Avec les différentes structures disponibles et à l'aide d'un alignement de séquences il est montré dans ce papier que NP A, NP B et NP Wfb se composent d'un corps prolongé par une queue flexible N-terminale tandis que NP C, NP D et NP ISA possèdent une queue C-terminale. La longueur de cette extension varie en fonction du virus.

La réplication et la transcription ayant lieu dans le noyau des cellules infectées, il est indispensable que les composants du virus interagissent avec le système de transport nucléaire. La NP fixe donc l'importine-α afin d'être acheminée jusqu'au noyau en passant par le complexe du pore nucléaire. Pour les virus *influenza* de type A et B, plusieurs études ont montré que le NLS fixé par l'importine-α se localise au niveau de la queue amino-terminale (Cros et al., 2005; Nakada et al., 2015; Ozawa et al., 2007; Sherry et al., 2014; Wang et al., 1997; Wanitchang et al., 2013; Wu et al., 2007).

Dans ce papier nous avons réalisé une étude biochimique et structurale de la nucléoprotéine du virus de la grippe de type D. Dans un premier temps, la cristallographie aux rayons-X nous a permis d'obtenir la structure de NP D. L'utilisation de la microscopie électronique, du dichroïsme circulaire, de la chromatographie d'exclusion de taille couplée ou non au MALLS, et du TSA nous ont permis de réaliser une caractérisation biophysique de la protéine et de deux mutants tronqués d'une partie plus ou moins longue de l'extrémité carboxy-terminale. Dans le but d'étudier la fixation de NP D à différents ARN de petite taille (entre 5 et 36 nucléotides) nous avons utilisé l'anisotropie de fluorescence.

Afin de purifier la protéine, différentes concentration en sels (de 50 mM à 300 mM NaCl) ont été testées. Contrairement à NP A et NP B, pouvant être purifiées à de faibles concentrations en sel (50 mM NaCl) (Chenavas et al., 2013a; Labaronne et al., 2016; Tarus et al., 2012) NP D n'a pu être purifiée sous forme stable qu'à 300 et 150 mM NaCl. Les résultats de MALLS et de microscopie électronique montrent que nous purifions une protéine sous forme tétramérique. Une structure du tétramère à 2.35 Å très similaire à NP A et NP B a été obtenue par cristallographie aux rayons-X. En se basant sur cette structure et les prédictions de désordre nous avons conçu deux mutants dépourvus d'une partie de l'extrémité C-terminale à partir des résidus 529 ou 511 ainsi qu'une construction permettant d'exprimer uniquement l'extrémité C-terminale (résidus 505-552). Des expériences de dichroïsme circulaire nous ont permis de confirmer l'absence de repliement de la queue C-terminale. Toutefois, grâce à la chromatographie d'exclusion de taille, nous avons pu mettre en évidence l'implication de cette queue C-terminale dans la fixation à l'importine-α. Finalement, l'anisotropie de fluorescence a permis de montrer l'importance de l'extrémité C-terminale dans la régulation de la fixation à l'ARN.

Ce papier fournit donc une caractérisation structurale et biochimique détaillée de la nucléoprotéine du virus de la grippe de type D. Nous mettons aussi en évidence l'implication de la queue carboxy-terminale dans le transport nucléaire via son interaction avec l'importine-α. De plus nous montrons ici pour la première fois l'importance de l'extrémité flexible de la NP dans la régulation négative de la fixation à l'ARN.

Structure of influenza D virus nucleoprotein and functions of NP_{TAIL}

Alice Labaronne[a], Myriam Miloudi[a], Laura Tengo[a], Caroline Mas[b], Guy Schoehn[a], Rob WH Ruigrok[a] & Thibaut Crépin[a]#

Institut de Biologie Structurale (IBS), Univ. Grenoble Alpes, CEA, CNRS, 38044 Grenoble, France[a]. Integrated Structural Biology Grenoble (ISBG) - UMS 3518 (CNRS-CEA-UJF-EMBL), 38044 Grenoble, France[b]

Running Head: RNA binding of IDV NP is regulated by NP_{TAIL}

#Address correspondence to Thibaut Crépin, thibaut.crepin@ibs.fr

Submitted to Journal of Virology, 2017/01/31

Abstract

This paper focuses on the nucleoprotein (NP) of the newly identified member of the *Orthomyxoviridae* family, influenza D virus. To date several X-ray structures of NP of influenza A (A/NP) and B (B/NP) viruses and the infectious salmon anemia (ISA/NP) virus were solved, highlighting a conserved overall core for the nucleoprotein. Here we purified, characterized and solved the X-Ray structure of the tetrameric D/NP at 2.35 Å resolution. The crystal structure of its core is globally similar to NP of other influenza viruses. However, unlike A/NP and B/NP which possess a flexible amino-terminal tail containing nuclear localization signals (NLS) for their nuclear import, D/NP possesses a carboxy terminal tail (D/NP_{TAIL}). We show that D/NP_{TAIL} harbors a bipartite NLS and designed two C-terminal truncated mutants (D/NP-511 and D/NP-529) to demonstrate the role of D/NP_{TAIL} in the interaction with human importin-α7. We also determine the affinity of D/NP for RNA using different RNAs from 5 to 32 nucleotides in length and we show that D/NP_{TAIL} controls the RNA binding of D/NP.

Importance

The RNAs of influenza viruses are covered by a viral protein; the nucleoprotein. The sequence and the structure of influenza D virus nucleoprotein was compared with the nucleoproteins of influenza A, B and C and 3 other nucleoproteins from animal Orthomyxoviruses. All of these proteins share a core starting with an α-helix and ending with 3 consecutive aromatic residues that plug a hydrophobic hole. Both influenza A and B nucleoproteins possess N-terminal disordered tails before the core that contain a NLS sequence. Influenza C and D viruses have C-terminal tails after the core. Here we show that the tail of influenza D nucleoprotein containing a bipartite NLS, regulates also the binding of the RNA to the protein.

Introduction

Influenza viruses are segmented negative strand RNA viruses. Their genome is made by the ribonucleoproteins (RNPs) corresponding to a set of each RNA segments coated with multiple copies of the nucleoprotein (NP) and associated to the viral heterotrimeric polymerase. RNPs are competent for both transcription and replication. The number of vRNA segments is specific to each type of influenza viruses and related to the number of glycoproteins at the surface of the vial particle. Influenza D virus, isolated for the first time in 2011 from swine, is a novel member of the *Orthomyxiviridae* family (15). Its genome is divided into 7 vRNA segments, similar to influenza C virus.

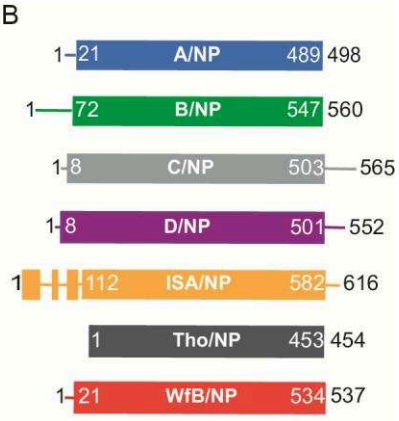
Orthomyxoviridae family includes 7 different *genera*, influenza A, B, C and D viruses, the infectious salmon anemia (ISA) virus, Thogoto virus, and Quaranfilivirus. To date, several X-ray structures of NP have been published, each of them without RNA. There are three structures of influenza A nucleoprotein (A/NP): Two with NP assembled as a trimer (29, 48) and one for the monomeric R416A mutant (5). The X-ray structure of the tetrameric influenza B nucleoprotein (B/NP) is also known (28), whereas the structure of isavirus NP (ISA/NP) was solved as a dimer (49). The overall folds of A/NP and B/NP are very similar with a root-mean-square deviation (rmsd) of 1.6 Å. They share less structure similarity with ISA/NP which presents an additional N-terminal folded domain (Fig. 1A). Figure 1B shows a schematic representation of the overall architecture of NP from Orthomyxoviruses. A/NP, B/NP and WfB/NP (NP of Wellfleet bay virus) are made by a core with an amino-terminal extension (NP_{TAIL}). The core contains the large and shallow positively charged surface which might bind RNA without sequence specificity (2, 9, 29, 48). For NP of influenza C (C/NP) and D (D/NP) viruses, their cores possess a short very version of the N-terminal tail (only 7 residues), which seems to be compensated by the presence of a carboxy-terminal NP_{TAIL}. This carboxy-

Figure-1

A

	A/NP	B/NP	C/NP	D/NP	ISA/NP
A/NP (2IQH)		1.6 Å (383 Cα)	-	2.1 Å (356 Cα)	2.9 Å (340 Cα)
B/NP (3TJ0)	37 %		-	1.9 Å (365 Cα)	2.8 Å (340 Cα)
C/NP	18 %	21 %		-	-
D/NP	17 %	21 %	38 %		2.9 Å (330 Cα)
ISA/NP (4EWC)	12 %	12 %	14 %	14 %	
Tho/NP	12 %	15 %	13 %	11 %	13 %
WfB/NP	12 %	12 %	14 %	13 %	12 %

B



terminal NP_{TAIL} is also present in ISA/NP in parallel of its additional N-terminal folded domain. The length of NP_{TAIL} is variable and seems to be specific for each genus.

It is known that influenza viruses replicate in the nucleus of the infected cells, meaning that all their replication machinery needs to interact with the nuclear transport system of the cell. For influenza A and B viruses, several studies have shown that the nuclear localization signals (NLS) of NP recognized by the cellular importins-α for their transport from the cytoplasm to the nucleus, are located within the flexible N-terminal tail (8, 21, 30, 34, 42, 43, 47) Recently a crystal structure between the non-classical NLS of A/NP and the importin-α1 has been solved (27).

In this paper, we characterized D/NP and solved the structure of its tetramer. The C-terminal D/NP_{TAIL} harbouring a classical bipartite nuclear localization signal (NLS) was not visible in the structure and we designed two C-terminal truncated mutants (D/NP-511 and D/NP-529) to study the interaction of D/NP with importin-α7. Our experiments clearly demonstrate that D/NP_{TAIL} is involved in the interaction with importins-α. Moreover we have determined affinity of D/NP for RNA using different RNAs from 5 to 24 nucleotides in length and demonstrated the implication of D/NP_{TAIL} in the RNA binding of D/NP, whereas NP_{TAIL} alone does not bind the RNA.

Materials and Methods

Molecular biology

The DNA coding sequence of D/bovine/France/2986/2012 NP was bought at GeneArt (ThermoFisher Scientific), optimized for bacterial expression. The different constructs have been PCR amplified and cloned following the supplier procedures (New England Biolabs). D/NP, D/NP-511 and D/NP-529 DNA coding sequences have been cloned in pETM13 (EMBL) to express C-terminal His-tagged proteins whereas D/NP_{TAIL} (505-D/NP-552) DNA coding sequence has been cloned in pETM11 (EMBL) to express an N-terminal His-tagged fragment. The DNA coding sequence of the human importin-α7 (KPNA6; Uniprot O60684) was cloned without its IBB domain (amino-acids 58 to 536) in pET9a (Novagen) to express an N-terminal His-tagged protein (4). Sequencing has been done by Eurofins.

Expression and purification

Escherichia coli BL21 RIL (DE3) cells (Life Technology) were transformed with the resulting plasmids. Cultures were induced 12 h by adding 0.3 mM isopropyl-β-D-thiogalactopyranoside (IPTG) at 18 °C and collected by centrifugation. The pellets were resuspended and sonicated in 50 mM Tris-HCl pH 7.5, 300 mM NaCl, 1 M NDSB-201 (Sigma), 2 mM β-mercaptoethanol (β-ME) and complete protease inhibitor cocktail (Roche) for D/NP constructs and 50 mM Tris-HCl pH 8, 500 mM NaCl, 1 mM β-ME and complete protease inhibitor cocktail for importin-α7. All purifications were performed at room temperature. All D/NP constructs were purified by Nickel affinity chromatography (Ni-NTA, Qiagen) followed, in the case of D/NP, D/NP-511 and D/NP-529, by a HiTrap™ Heparin HP column (GE-Healthcare) on a NGC system (BioRad). Elution fractions of D/NP, D/NP-511, D/NP-529 and D/NP_{TAIL} were dialyzed against 20 mM Tris-HCl pH 7.5 at 50 mM or 300 mM NaCl, 5 mM β-ME. The last purification step was a size-exclusion chromatography using a Hiload™ 16/600 S200 column (GE-Healthcare) for D/NP,

D/NP-511 and D/NP-529 or a S75 10/300 GL column (GE-Healthcare) for D/NP_{TAIL}. Importin-α7 was purified by Nickel affinity chromatography and elution fraction were dialyzed with TEV (1/100) against 20 mM Tris-HCl pH 7.5 at 150 mM, 5 mM β-ME and 20 mM imidazol. Finally a size-exclusion chromatography using Superdex^{TM} 200 increase 10/300 GL column (GE-Healthcare) was performed in 20 mM Tris-HCl pH 7.5 at 150 mM, 5 mM β-ME. Peak fractions were concentrated using a 10 kDa Amicon concentrator. Protein concentrations were determined using the extinction coefficients at 280 nm, ε = 44 537 $M^{-1}.cm^{-1}$ for D/NP, ε = 41 558 $M^{-1}.cm^{-1}$ for D/NP-529 and D/NP-511, ε = 2980 $M^{-1}.cm^{-1}$ for D/NP_{TAIL}, and ε = 46 785 $M^{-1}.cm^{-1}$ for importin-α7.

SEC-MALLS-RI analysis

SEC was performed with a column (Superdex^{TM} 200 increase 10/300 GL or Superdex 75 10/300 GL) equilibrated with 20 mM Tris-HCl pH 7.5, 150 mM NaCl and 5 mM β-ME. Analytical runs were performed at 20 °C with a flow rate of 0.5 $mL.min^{-1}$. MALLS detection was performed with a DAWN-HELEOS II detector (Wyatt Technology) using a laser emitting at 690 nm and protein concentration was measured on-line with the use of differential refractive-index measurements, with an Optilab T-rEX detector (Wyatt Technology) and a refractive-index increment, dn/dc, of 0.185 $mL.g^{-1}$. Weight-average molar masses (Mw) were calculated with ASTRA (Wyatt Technology) as previously described (13).

Electron microscopy

Samples (concentrations around 0.05 $mg.mL^{-1}$) were applied between a carbon and a mica layer. The carbon was then floated on the top of a 2% (w/v) sodium silicotungstate, pH 7.0 solution. The carbon film was covered by a copper grid. Both were fished using a small piece of journal paper and air dried before insertion in the electron microscope (12). Charge-coupled Device (CCD) frames were taken with a FEI T12 microscope operating at 120 kV and

Depuis de nombreuses années, la communauté scientifique étudie le fonctionnement des virus à ARN négatif afin de mieux lutter contre eux. Dans le cas du virus de la grippe, de nombreux efforts sont faits pour mieux comprendre le fonctionnement de la machinerie de réplication et de transcription virale constituée par le complexe de la ribonucléoprotéine. Depuis une quinzaine d'années, plusieurs structures basses résolutions de RNP sont disponibles (Arranz et al., 2012; Coloma et al., 2009; Martín-Benito et al., 2001b; Moeller et al., 2012; Ortega et al., 2000a) ainsi que de nombreuses structures de nucléoprotéines (Chenavas et al., 2013a; Ng et al., 2008, 2012, Ye et al., 2006, 2012). Récemment des structures de la polymérase en complexe avec l'ARN ont permis de mieux appréhender les mécanismes réplicatif et transcriptionnel du virus (Pflug et al., 2014; Reich et al., 2014).

Structure de la nucléoprotéine du virus de la grippe

Au cours de ce travail de thèse, la structure de la nucléoprotéine du virus de la grippe D a été résolue par cristallographie aux rayons X. La structure de NP D a été confrontée aux structures de NP précédemment résolues et bien que biochimiquement différentes, il apparait clairement que les NP des *Orthomyxoviridae* et notamment des virus grippaux ont des structures très similaires : NP A et NP B possèdent un rmsd de 1.6 Å sur leur carbone-α tandis que NP D possède respectivement un rmsd de 2.1 Å et 1.9 Å avec NP A et NP B. La comparaison des structures des nucléoprotéines des virus de la grippe A, B et D montrent que les protéines possèdent toutes un corps s'étendant de la première hélice présente dans la structure jusqu'à une séquence de trois résidus aromatiques venant recouvrir une zone hydrophobe de la protéine (Article 2). D'un point de vue structural, la différence majeure entre les NP du virus de la grippe concerne leurs extrémités amino ou carboxy-terminales qui sont intrinsèquement désordonnées. NP A possède une courte queue amino-terminale (21 résidus) alors que NP B en possède une plus longue (72 résidus). NP C et NP D possèdent quant à elles une extension carboxy-terminale d'une cinquantaine de résidus. La position variable de ces extensions n'est pas encore comprise mais cela pourraient avoir un lien avec le type d'hôte infecté et les protéines cellulaires spécifiques de l'hôte interagissant avec la NP.

Bien que nous n'ayons pas pu obtenir de cristaux de NP D en présence d'ARN, la structure de NP D nous a fourni des informations intéressantes pour nous permettre d'étudier l'interaction de cette protéine avec l'importine-α.

a nominal magnification of 45 000 times. The dilutions for EM were performed with the size-exclusion buffer right before preparing the grid.

Fluorescence anisotropy measurements

Fluorescence anisotropy assays were performed on a Clariostar (BMG Labtech) microplate reader, fitted with polarization filters to measure fluorescence anisotropy. The binding assays were done in 384-well plates at room temperature (60 µL reaction volume) in 20 mM Tris-HCl pH 7.5, 150 or 300 mM NaCl and 5 mM β-ME. 0 to 3 µM of D/NP were titrated in 5 nM $^{5'}$Phosphate-RNA labelled in 3' with fluorescein amidite (FAM; Integrated DNA Technologies). 5 to 36 nucleotides polyUC RNA (Integrated DNA Technologies) were used. After substracting the polarization values obtained for RNA alone, the mean value of three independent experiments were fitted to the standard binding equation y = ((Bmax x xh)/(Kdh = xh)). All fluorescence anisotropy experiments were done in triplicate.

Thermal shift assays

Thermofluor experiments were performed following the established protocols (31). Sample were diluted at 0.25 mg.mL^{-1} in 50 mM Tris-HCl pH 7.5, 300 mM NaCl and 5 mM β-ME and 5X SYPRO orange dye (Invitrogen) in a final volume of 40 µL. The thermal stability was measured using a real time PCR machine (Mx3005P Q-PCR, Stratagene). The dye was excited at 488 nm and the emission light was recorded at 585 nm while the temperature was increased by increments of 1 °C per minute from 25 to 95 °C. The relative fluorescence emission was then plotted against its corresponding temperature to produce the thermal shift profile curve. The melting temperatures were estimated from the curve as the midpoint temperature between the baseline and maximum value of the curve.

Circular dichroism

A JASCO J-810 CD spectropolarimeter equipped with a temperature-controller (Peltier system) was used to record the far-UV CD spectrum of D/NP_{TAIL} at 20 µM. The spectrum was collected at 20°C with ten accumulations using a cuvette with a path-length of 1 mm. After blank signal subtraction, the CD signal (in mdeg) was converted to mean molar residue ellipticity (in $deg.cm^2\ dmol^{-1}$).

Crystallisation and structure determination

D/NP was crystallised by vapor diffusion using the sitting drop method. The crystals were obtained in 2.4 M sodium malonate pH 4 with a protein concentration at 1.5 $mg.mL^{-1}$. The crystals were directly flash-frozen without any cryoprotectant. Data were collected at the ESRF (beamline ID30A-3) and processed with the XDS package (18). The structure was solved by molecular replacement using the influenza A virus R416A monomeric nucleoprotein structure (PDB ID code 3ZDP) without the oligomerization loop (from residues 391 to 439) as a model. Model building and refinement were performed using CCP4i suite program for crystallography (PHASER, ARP/wARP, REFMAC5, COOT) (6, 10, 23, 26, 46). The final refinement was done using BUSTER (36). The coordinates have been deposited in the Protein Data Bank under PDB ID code XXX. The protein structure figures were drawn using PyMOL (33).

Results

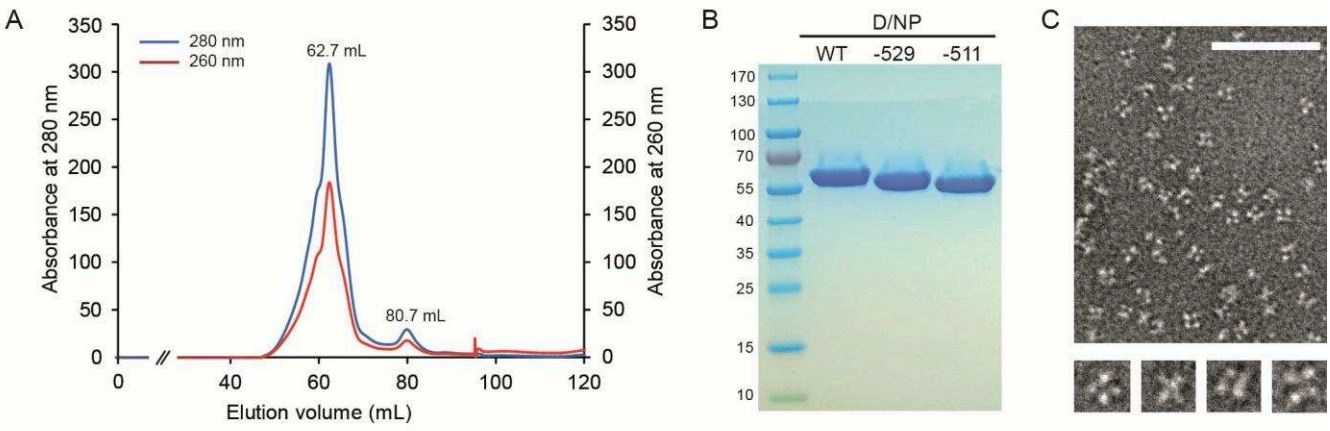
Recombinant D/NP forms a tetramer in solution

The DNA coding sequence of D/NP was cloned in a bacterial expression plasmid and over-expressed in *Escherichia coli* as a C-terminal His-tag recombinant protein. D/NP was purified in three steps; a Ni^{2+} resin followed by a heparin column and a final gel filtration. Figure 2A shows a typical gel filtration elution profile using the absorbance signals at 260 and 280 nm. The protein can be then concentrated between 2 and 6 $mg.mL^{-1}$. The sample was pure (Fig. 2B) without nucleic acid contamination (ratio 280/260 > 1.75). By electron microscopy (negative staining), we show that D/NP forms mainly tetramers in solution (Fig. 2C). A SEC-MALLS-RI experiment has confirmed the homogeneity and the molecular weight of the recombinant tetrameric D/NP (supplementary Fig. 1). It has been previously shown that the oligomerization of recombinant A/NP and B/NP can be modulated by the NaCl concentration of the purification buffers (5, 20, 39). Starting from purified and stable oligomeric samples (trimers for A/NP and tetramers for B/NP), monomeric proteins can be obtained by decreasing stepwise the NaCl concentration. After purification at 300 mM NaCl followed by its dialysis at 150 mM NaCl, D/NP is eluted from gel filtration (with a 150 mM NaCl running buffer) at the same volume, meaning that a smooth reduction of the salt concentration does not change its oligomeric state. However, a more drastic decrease induces an irreversible and total precipitation of D/NP, even with a stepwise reduction at 150 mM NaCl. For the rest of the experiments on D/NP and its mutants, all the samples were purified at 300 mM NaCl.

D/NP structure

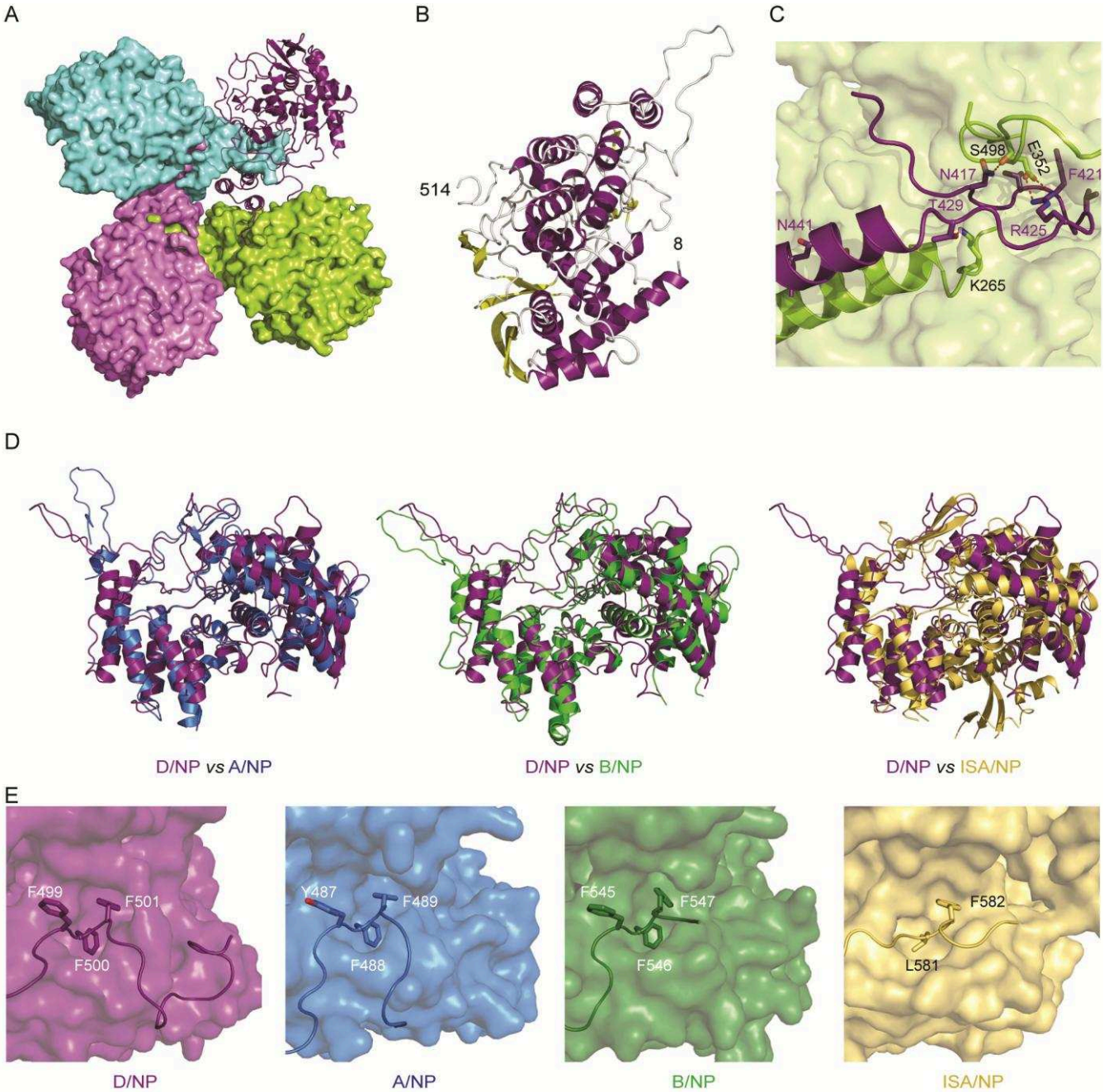
Full length D/NP was crystallized in sodium malonate as small fine needles that diffract up to 2.35 Å resolution (Table 1). The structure was solved by molecular replacement using a starting model made by one monomer of the A/NP R416A mutant (5), in which the parts

Figure-2


A
280 nm
260 nm
62.7 mL
80.7 mL
Absorbance at 280 nm
Absorbance at 260 nm
Elution volume (mL)
B
D/NP
WT
-529
-511
170
130
100
70
55
40
35
25
15
10
C

197 involved in the oligomerization are removed. The automatic search has provided the position
198 for three molecules and a forth one was then fitted manually after the analysis of the electron
199 density. Retrospectively, the molecular replacement using the tetrameric B/NP model did not
200 provide any acceptable solution. The asymmetric unit contains four molecules arranged as a
201 tetramer (Fig. 3A and refinement statistics in Table 1), coherent with the observations made
202 by electron microscopy. Based on the electron density, 87 % of the model could be built, from
203 residues 8 to 514 (Fig. 3B). Several internal loops are missing and the NP_{TAIL} was disordered.
204 The D/NP monomer was compared to the already known structures (Fig. 3D with the statistics
205 included in Fig. 1A). The closest model is B/NP (r.m.s.d of 1.9 Å for 365 C_α aligned), the
206 highest deviation between the two models comes from the part of the protein involved in the
207 oligomerization process; the oligomerization-loop of one monomer plugged into a cavity of its
208 neighbour. The loop is then stabilized by several kind of contacts, mainly hydrophobic
209 because of several aromatic residues lying in the pocket of the neighbouring monomer but
210 also a salt bridge between an arginine from one monomer (Arg-425) interacting with a
211 glutamate of the neighbour (Glu-352; Fig. 3C). These two residues, strictly conserved in all
212 influenza NPs, are well known for modulating their oligomeric state. Once mutated in alanine,
213 the corresponding single mutants form monomers, unable to self-assemble (5, 9, 39, 40, 48).
214 Including ISA/NP, a common architecture for all influenza-like NPs can be defined based on
215 the X-ray structures. It explains the limits we have included on the figure 1A, with the NP_{CORE}
216 starting with an α-helix and finishing with an unfolded part, anchored into the surface of the
217 protein. Figure 3E shows the X-ray structures of A/NP, B/NP, D/NP and ISA/NP with a zoom
218 on the C-termini, where a superimposed patch made by 3 consecutive aromatic residues is
219 found in influenza NPs ($_{487}$YFF$_{489}$ for A/NP, $_{545}$FFF$_{547}$ for B/NP and $_{499}$FFF$_{501}$ for D/NP). A
220 similar patch is also presents in the ISA/NP model ($_{580}$GLF$_{582}$) even if the upstream unfolded

Figure-3


A
B
514
8
C
S498
E352
N417
F421
T429
N441
R425
K265
D
D/NP vs A/NP
D/NP vs B/NP
D/NP vs ISA/NP
E
F499
F501
F500
Y487
F489
F488
F545
F547
F546
F582
L581
D/NP
A/NP
B/NP
ISA/NP

part is not as much superimposed. The sequences analysis shows such patch is present in C/NP ($_{501}$FFF$_{503}$) and could be present in Tho/NP ($_{451}$YLF$_{453}$) and Wfb/NP ($_{532}$VIY$_{534}$). Because D/NP differs from A/NP and B/NP in its global architecture, we have decided to further analyze the C-terminal extension: D/NP$_{TAIL}$.

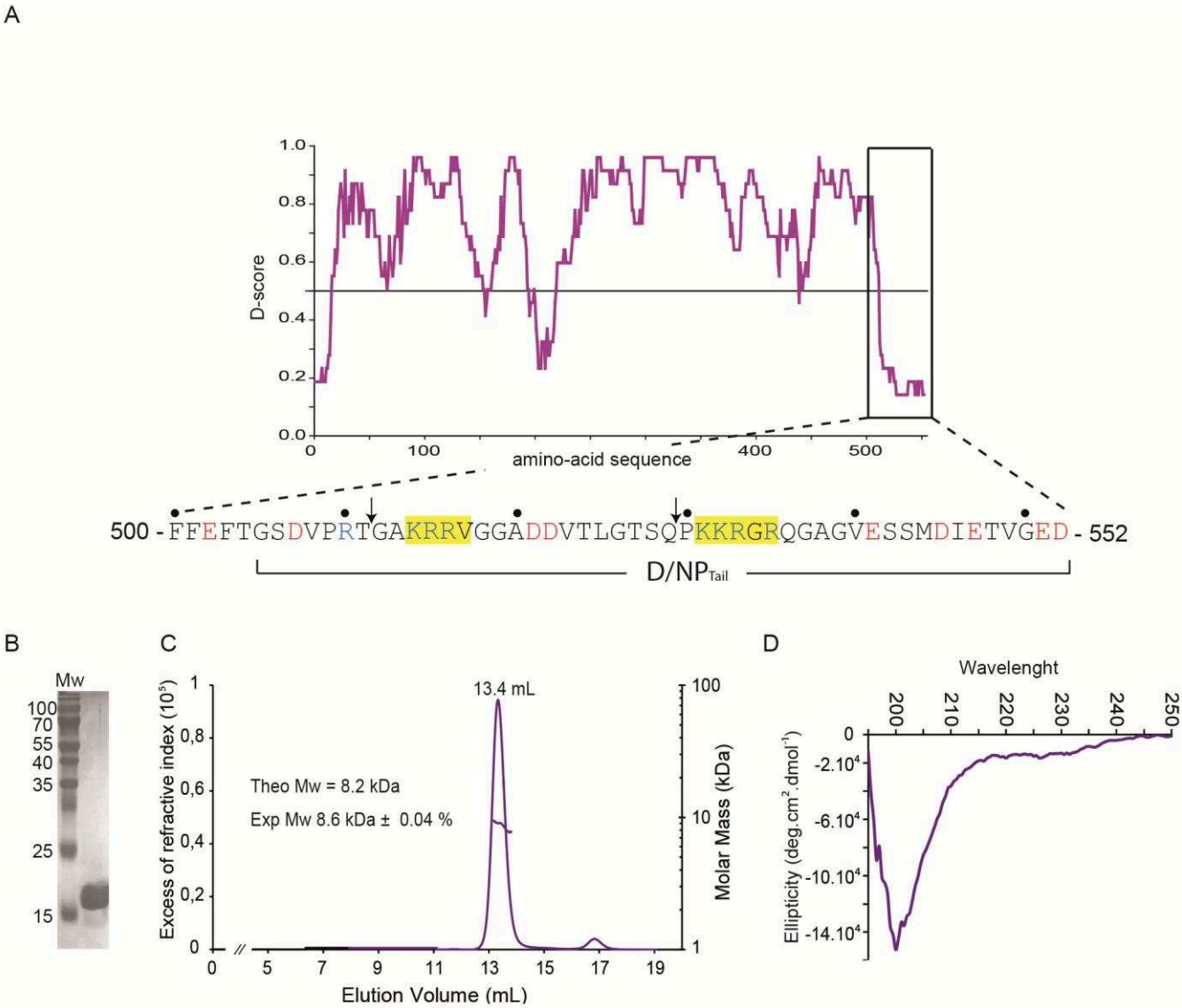
D/NP$_{TAIL}$ behaves as an intrinsically disordered protein

Based on the crystal structure and the disorder prediction (Fig. 4A), we have designed a construct (from residues 505 to 552) for the expression in *E. coli* as a N-terminal His-tag recombinant version. Based on the structure, even if we consider that the core of D/NP ends after residue Phe-501, we have chosen to start the construct at Gly-505, in order to avoid the hydrophobicity of Phe-503. D/NP$_{TAIL}$ was eluted from gel filtration as a 40 kDa-protein but with a normal migration pattern on SDS-PAGE (Fig. 4B). As intrinsically disordered proteins are known to be eluted from gel filtration with an aberrant volume (13, 41). A SEC-MALLS-RI experiment (Fig. 4C) confirmed the monodispersity of the purified sample with a Mw of 8.2 kDa for D/NP$_{TAIL}$ (with the His-tag). The CD analysis confirms that D/NP$_{TAIL}$ does not contain any significant stable secondary structure in solution (Fig. 4D).

D/NP$_{TAIL}$ interacts with importin-α

The N-terminal tails of A/NP and B/NP are known to be involved in their nuclear import by interacting with importins-α (8, 21, 27, 34, 38, 43). The sequence analysis of D/NP shows no import signals in its N-terminal part but suggests the presence of a bipartite NLS within D/NP$_{TAIL}$ ($_{514}$KRR-X$_{14}$-KKRGR$_{545}$; Fig. 4A). We have tested the bipartite NLS with different constructs. First, we could show that the human importin-α7 is co-eluted with the His-tagged D/NP$_{TAIL}$ from a nickel-affinity resin (Fig. 5A). We then have confirmed the interaction by thermal-shift experiments where importin-α7 appears to be more stable in presence of

Figure-4


A
D-score
amino-acid sequence
500 - FFEFTGSDVPRTGAKRRVGGADDVTLGTSQPKKRGRQGAGVESSMDIETVGED - 552
D/NPTail
B
Mw
100
70
55
40
35
25
15
C
13.4 mL
Theo Mw = 8.2 kDa
Exp Mw 8.6 kDa ± 0.04 %
Excess of refractive index (10⁵)
Molar Mass (kDa)
Elution Volume (mL)
D
Wavelenght
Ellipticity (deg.cm².dmol⁻¹)

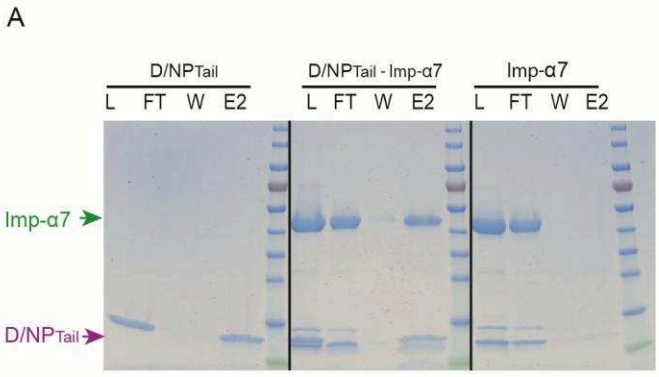
D/NP_{TAIL} (Fig. 5B), and could purified the D/NP_{TAIL}:importin-α7 complex by gel filtration on a S75 column (sup Fig. 2). We have confirmed the interaction between the two partners using the full-length D/NP. By gel filtration, we have shown that the D/NP:importin-α7 complex can be eluted as a single peak. A molecular weight of 474 kDa has been measured by MALLS for the complex (Fig. 5C) corresponding to the interaction of 4 importins-α7 (Mw = 55 kDa) per tetramer of D/NP (Mw = 260 kDa). To confirm the role of D/NP_{TAIL} in the interaction with importin-α, we have designed two C-terminal truncated constructs of D/NP, D/NP-511 and D/NP-529, ending respectively at residues 511 and 529. D/NP-511 is deprived of the whole bipartite NLS whereas contains D/NP-529 the first part. They both behave as the wild-type protein during the purification. Using pull-down experiments on nickel-affinity resin, we show that D/NP-529 still slightly retains importin-α7 whereas the interaction between D/NP-511 and importin-α7 is totally abolished (Fig.5D).

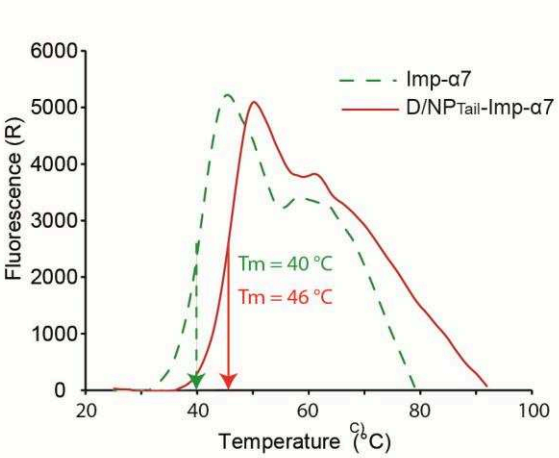
D/NP_{TAIL} regulates the binding of RNA

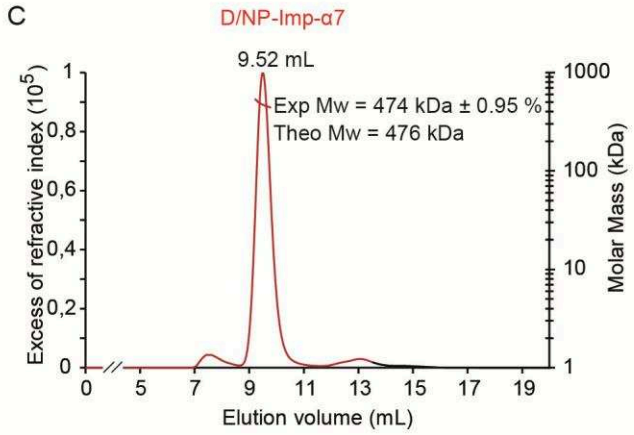
By using $^{5'}$phosphate-$(UC)_x$-FAM$^{3'}$ in fluorescence anisotropy experiments, we have measured the affinities of D/NP for RNA molecules, ranging from 5 to 36 nucleotides in length (Fig. 6 and Table 2). We show that for small RNA molecules, there is a clear effect of the NaCl concentration on the affinity, with higher affinities obtained at “low” salt concentration. For longer RNA molecules (> 16 nt), the affinities are stronger, with no significant effect of the ionic strength neither of the length of RNA. The tipping point seems to be around 12 nucleotides, similar to what has been observed for A/NP and B/NP (20) but also for ISA/NP (49).

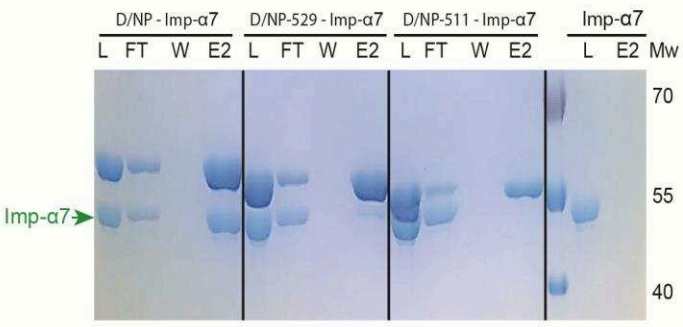
By focusing on the 10-16 nt fluorescent RNA molecules, we have performed the same experiment using the D/NP-511 and D/NP-529 truncated constructs. We have chosen the 300 mM NaCl condition to get a better tuning of the results. For the 10 nt and 16 nt RNA

Figure-5


A
D/NPTail
D/NPTail - Imp-α7
Imp-α7
L FT W E2
L FT W E2
L FT W E2
Imp-α7
D/NPTail

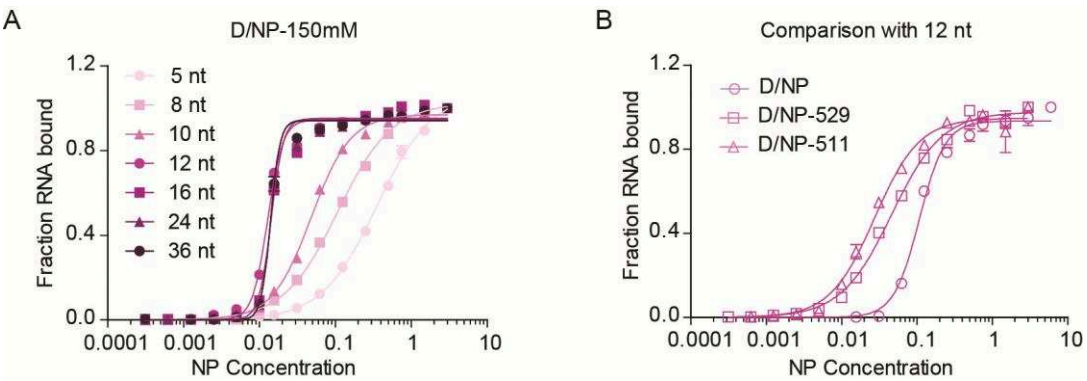

B
Imp-α7
D/NPTail-Imp-α7
Fluorescence (R)
Tm = 40 °C
Tm = 46 °C
Temperature (°C)


C
D/NP-Imp-α7
9.52 mL
Exp Mw = 474 kDa ± 0.95 %
Theo Mw = 476 kDa
Excess of refractive index (10^5)
Molar Mass (kDa)
Elution volume (mL)


D
D/NP - Imp-α7
D/NP-529 - Imp-α7
D/NP-511 - Imp-α7
Imp-α7
L FT W E2
L FT W E2
L FT W E2
L E2
Mw
Imp-α7

molecules, there is no significant difference in the affinities between the two mutants and the *wt* protein. However, using the 12 nt fluorescent RNA, we observed that the binding of D/NP-511 and -529 is more efficient than for D/NP (Fig. 6B and Table 2), suggesting that D/NP_{TAIL} controls the binding of RNA onto the core of the nucleoprotein. We have also confirmed that there is no detectable change in the anisotropy when mixing D/NP_{TAIL} and the corresponding RNAs.

Figure-6


A
D/NP-150mM
Fraction RNA bound
1.2
0.8
0.4
0.0
5 nt
8 nt
10 nt
12 nt
16 nt
24 nt
36 nt
0.0001
0.001
0.01
0.1
1
10
NP Concentration
B
Comparison with 12 nt
Fraction RNA bound
D/NP
D/NP-529
D/NP-511
NP Concentration

Discussion

In 2011, a new virus has been isolated from pigs exhibiting influenza-like illness. Later, cattle has been found to be the primary reservoir and subsequent electron microscopic, biochemical and genetic studies identified an Orthomyxovirus with seven RNA segments sharing 50 % overall amino acids identity with the human influenza C virus (15). First considered as a subtype of the influenza C virus, it has since been officially named influenza D virus (IDV) by the International Committee of Taxonomy of Viruses in 2016. Since its first isolation, IDV appears to be spread all around the world in many mammals, with serological evidence for its transmission to humans (45). The recent structure of IDV Hemagglutinin-Esterase-Fusion glycoprotein (HEF) has described a rather open receptor-binding cavity capable of accommodating diverse extended glycan moieties that could be one reason for its broad cell tropism (37).

With this paper, we provide a detailed characterization of IDV nucleoprotein. The comparison of the nucleoproteins of influenza A, B, D and of infectious salmon anemia virus (ISAV) structures shows that of these proteins share a core (NP_{CORE}) that start with the first helix in the structure, up to three aromatic residues that plug a hydrophobic hole in the protein. The only structure that differs is the ISAV/NP that possesses a small folded domain before the first helix of the NP_{CORE}. Looking to the sequences of all known Orthomyxoviruses nucleoproteins, one can observe that the NP of Thogoto virus has only the NP_{CORE} without any else. The NPs of influenza A and B viruses have an N-terminal NP_{TAIL} whereas the NPs of influenza C, D and ISAV have a C-terminal NP_{TAIL}. The NP_{TAIL} of A is only 20 residues whereas the tails of the other influenza viruses are much longer; 71 residues for B, 62 for C, 51 for D. We demonstrate that D/NP_{TAIL} presents all the characteristics of an intrinsically disordered protein (IDP) and that is involved in the nuclear import through its interaction with cellular importins-α.

This role is not a big surprise as it had already been shown the presence of the NLS motifs within the flexible NP_{TAILs} of viruses A and B (8, 21, 27, 34, 38, 43). Whereas Tho/NP has been shown to be imported in the nucleus (44), this NP does not a IDP-like additional peptide with an NLS.

Like influenza A and B viruses nucleoproteins, D/NP binds single strand RNA. We found that the 12-16 nt long RNAs bind with a better affinity than smaller RNAs. For A/NP and B/NP we found that a 12-mer RNA binds strongly, as if there is strong binding site for 12 nt, and that longer RNAs do not bind better. For the ISA/NP similar findings were found (49). Using a 12-mer RNA and 300 mM salt, we found that the D/NP binds with a K_d of 108 nM whereas the shorter proteins D/NP-529 and -511 bind the same RNA with K_ds of 46 and 28, close to the K_d obtained when using longer RNAs. This suggests that D/NP_{TAIL} could overlap with the RNA binding site. It is known that IDPs contain many charged residues, more than structured parts of proteins (7). D/NP_{TAIL} has 9 negative charges for 8 positive (7 in the NLSs), suggesting that the charges alone will not influence the RNA binding. We find indeed that D/NP_{TAIL} alone doesn't bind RNA, suggesting that the presence of the D/NP_{TAIL} can control the kinetics of the RNA binding.

The preservation of such IDP-like peptide with a conserved function highlights the amplitude of its importance. The name for N_{TAIL} was first used for the C-terminal disordered tails of Sendai and measles viruses nucleoproteins (22), which are about 125 residues long. These tails are disordered but they can adopt a short helix structure (residue D484-A502) for the binding to the phosphoproteins of these viruses that themselves bind to the corresponding RNA polymerase (3). For these viruses, the tails keep the nucleocapsids in a flexible conformation. When cleaved off with trypsin, the helical nucleocapsids become rigid (17, 32). However, the tails do not change the affinity of N for the RNA but the assembly of the capsids with a slower kinetics (24). The RNPs of influenza A and B viruses are very flexible (1, 11, 19,

25) which makes it difficult to handle for high resolution structures. The EM structure of the measles nucleocapsid with a resolution of 3.6 Å (14) was done on the nucleocapsids without the N_{TAIL}. Maybe it would be interesting to make influenza RNPs with NP without NP_{TAIL} for making rigid RNPs that would be easier for making a structure for seeing the RNA on the nucleoprotein.

Acknowledgements

We are highly grateful to Jean-Marie Bourhis, Daphna Fenel and Joanna Timmins for their help. We thank Darren J Hart for the plasmid to express the human importin-α7. We thank Mariette Ducatez and Anny Slama-Schwok for discussion. We thank the staff of the ESRF-EMBL Joint Structural Biology Group for access to ESRF beamlines. AL and MM were funded through the Labex GRAL (ANR-10-LABX-49-01). This work was supported by the French agency for Research through the ANR RNAP-IAV (ANR-14-CE09-0017). This study used the platforms of the Grenoble Instruct centre (ISBG; UMS 3518 CNRS-CEA-UJF-EMBL) with support from FRISBI (ANR-10-INSB-05-02) and GRAL, within the Grenoble Partnership for Structural Biology (PSB).

Author contributions

TC conceived the experiments. AL, MM, LT, CM and TC performed the experiments. AL, GS, RWHR and TC analysed the data. AL and TC wrote the first version of the paper. AL, RWHR, and TC wrote the final version of the paper

Competing interest statement

The authors declare no competing financial interests.

Figures legends

FIG. 1. Nucleoproteins of Orthomyxoviruses. (A) Pairwise root-mean-square deviation (rmsd) of Cα positions and sequence identity between NP of the members of the *Orthomyxoviridae* family. The sequence identity and the rmsd values have been calculated using respectively Clustal Omega (35) and DaLi (16). The identity between Tho/NP and WfB/NP is 16 %. (B) Schematic representation of the nucleoprotein based on the sequence and structure analysis of the protein from representative members of the *Orthomyxoviridae* family. The schema respects the size of the proteins. A/NP (influenza A virus, strain A/WSN/1933(H1N1); Uniprot id: B4URF1), B/NP (influenza B virus, strain B/Memphis/13/2003; Uniprot id: Q5V913), C/NP (influenza C virus, strain C/AnnArbor/1/1950; Uniprot id: Q6I7C0), D/NP (influenza D virus, strain D/bovine/France/2986/2012; Uniprot id: A0A0E3VZU8), ISA/NP (infectious salmon anemia virus, isolate salmon/Norway/810/9/99; Uniprot id: Q8V3T7), Tho/NP (Thogoto virus; Uniprot id: A0A0B6VKB5) and WfB/NP (Wellfleet Bay virus; Uniprot id: A0A0A1E9N5). The flexible tails are represented with simple lines whereas the folded parts and the cores are represented with filled boxes.

FIG. 2. Purification of influenza D nucleoprotein. (A) Size exclusion chromatography profile of wild-type D/NP. The sample was loaded on a Hiload^{TM} 16/600 S200 column equilibrated with the running buffer 20 mM Tris-HCl pH 7.5, 300 mM NaCl and 5 mM β-mercaptoethanol. (B) Coomassie blue-stained SDS-PAGE (4-20 % gradient polyacrylamide) showing the purified wild-type D/NP and the two C-terminal truncated mutants (D/NP-529 and D/NP-511). (C) Electron microscopy images of the elution peak of D/NP. Samples show different oligomeric states although most oligomers are tetramers.

529 FIG. 3. Structure of influenza D nucleoprotein. (A) Structure of the tetrameric D/NP, with 3
530 monomers shown in surface (respectively in green, pink and cyan) and the fourth in cartoon
531 (deep purple). (B) Ribbon diagram of one monomer of D/NP with the α-helixes in deep purple
532 and β-strands in yellow. (C) Detail of the interactions between two monomers of the
533 tetrameric D/NP as shown in (A). The conserved R425 of one monomer (shown as deep
534 purple cartoon) stabilized the position of the oligomerization loop at the surface of the
535 neighbouring protomer, through its conserved E352. (D) Comparison of D/NP with other
536 nucleoproteins negative sRNA viruses. The structure of one monomer of D/NP (deep purple)
537 has been superimposed from left to right, with one monomer of A/NP (blue; PDB id: 2IQH),
538 B/BP (forest; PDB id: 3TJ0) and ISA/NP (light orange; PDB id: 4EWC). The rmsd values are
539 in Figure 1A. (E) Anchoring of the C-terminus on NP_{CORE} by a patch of 3 consecutive aromatic
540 residues. The panel corresponds to a zoom of the superimposed structures shown in panel C
541 with NP_{CORE} represented in surface.

542 FIG. 4. Biophysical characterization of D/NP_{TAIL}. (A) D-score (score for disorder as a function
543 of residue) of D/NP with a zoom (below the graph) on the last 50 residues. The prediction is
544 based on 22 predictor web servers. The disordered regions have a D-score < 0.50 (13). The
545 yellow boxes correspond to the putative NLS motifs. (B) Coomassie blue-stained SDS-PAGE
546 (Tris-Tricine; 15% polyacrylamide) of the purified D/NP_{TAIL}. It migrates at a higher molecular
547 weight (17 kDa approximately) than expected (8 kDa). (C) SEC-MALLS analysis of D/NP_{TAIL}
548 loaded on S75 10/300 GL column. For this experiment, we have chosen to keep the His-tag
549 encoded with the pETM11 plasmid, for an optimal detection of D/NP_{TAIL}. The experimental
550 molecular weight is coherent with the expected mass. (D) Circular dichroïsm of D/NP_{TAIL}.

551 FIG. 5. Interaction with importin-α7. (A) Pull-down assays of human importin-α7 by His-
552 tagged D/NP_{TAIL}. The mixtures (200 μL; molar ratio 1 D/NP_{TAIL} for 1 importin-α7) were
553 incubated 1 hour at room temperature and then loaded (L) on 200 μL of Ni-NTA resin. The
554 flow throughs (FT) were recovered by gravity and the resins were washed with 50 volumes of
555 the equilibration buffer. The final 200 μL of the washing step were recovered (W). The
556 proteins were then eluted with the buffer 20 mM Tris-HCl pH 7.5, 250 mM NaCl, 800 mM
557 imidazol and 5 mM β-mercaptoethanol (100 μL fractions). The figure shows the coomassie
558 blue-stained SDS-PAGE (15 % polyacrylamide) with the L, FT, W and the second fractions of
559 the pull-down (E2). The importin-α7 used for this experiment presents a small degradation.
560 (B) Thermal stability assay of importin-α7 in absence (green) or in presence (purple) of
561 D/NP_{TAIL} using Thermofluor. In presence of D/NP_{TAIL}, the melting temperature of importin-α7
562 is 6°C higher. (C) SEC-MALLS analysis of D/NP in complex with importin-α7. The mixture (50
563 μL; molar ratio 1 D/NP monomer for 1.2 importin-α7) was incubated 1 hour at room
564 temperature and then loaded on a Superdex^{TM} 200 increase 10/300 GL. The experimental
565 molecular weight (below the peak) is coherent with the expected mass of four importins-α7
566 bound per D/NP tetramer. (D) Pull-down assays of human importin-α7 by D/NP and the two
567 C-terminal truncated mutants (D/NP-529 and D/NP-511). The his-tags are on D/NP. The
568 mixtures (200 μL; molar ratio 1 D/NP monomer for 1.2 importin-α7) were incubated 1 hour
569 and the experience was done as described in panel (A). The figure shows the coomassie
570 blue-stained SDS-PAGE (12 % polyacrylamide) with the L, FT, W and the second fractions
571 (E2).

572 FIG. 6. RNA interaction. (A) Titrations of NP were done against $^{5'}$phosphate-polyUC-FAM$^{3'}$
573 ranging from 8 to 36 nucleotides. The mixtures were incubated 5 min at room temperature.

574 RNA binding titration curves of D/NP at 300 mM. (B) Comparison of the titration curves
575 obtained for the wild-type D/NP and the two truncated mutants (D/NP-529 and D/NP-511) in
576 presence $^{5'}$phosphate-poly(UC)$_6$-FAM$^{3'}$. The K_d obtained for these experiments are listed in
577 Table 2.
578

579

Table 1: Crystallographic data collection and refinement statistics. Values in parentheses are for highest resolution shell.

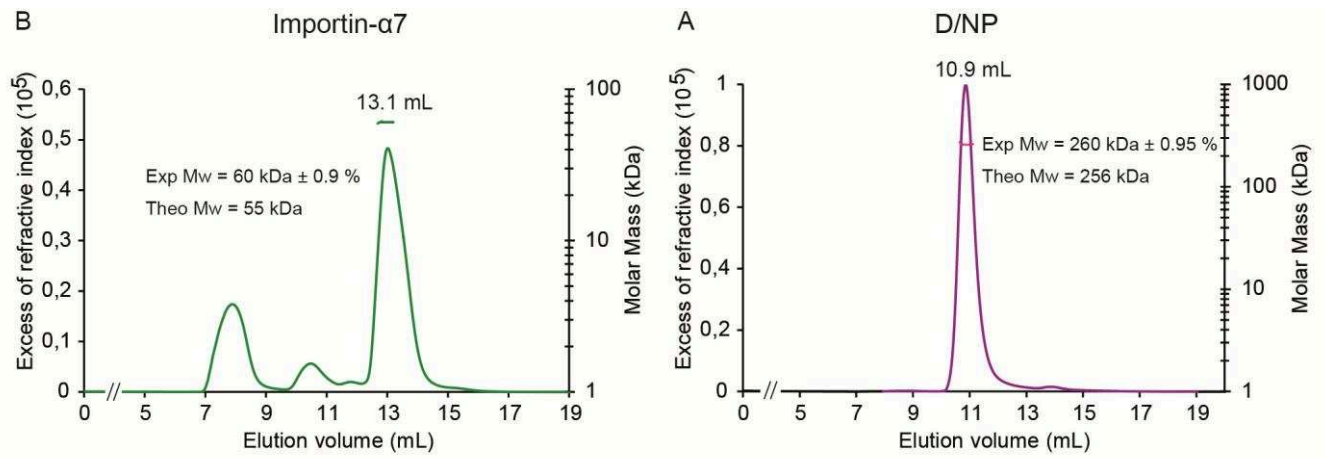
	D/NP
Data collection	
X-ray source	ID30A3
Wavelength (Å)	0.9677
Space group	P1
a (Å)	75.16
b (Å)	85.16
c (Å)	103.38
α (º)	91.203
β (º)	101.940
γ (º)	101.017
Resolution range (Å)	40 - 2.35 (2.48 - 2.35)
Completeness (%)	91.4 (93.3)
Rsym (I) [(a)] (%)	9.8 (90.0)
I/σI	7.36 (0.98)
Total reflections	204 313 (31 120)
Unique reflections	93 545 (14 226)
Multiplicity	2.2 (2.2)
Refinement statistics	
molecules/AU	4
R-factor (%)	20.4
R_{free} (%)	26.3
R.m.s deviations from ideal values	
Bond length (Å)	0.009
Bond angle (°)	1.19

580

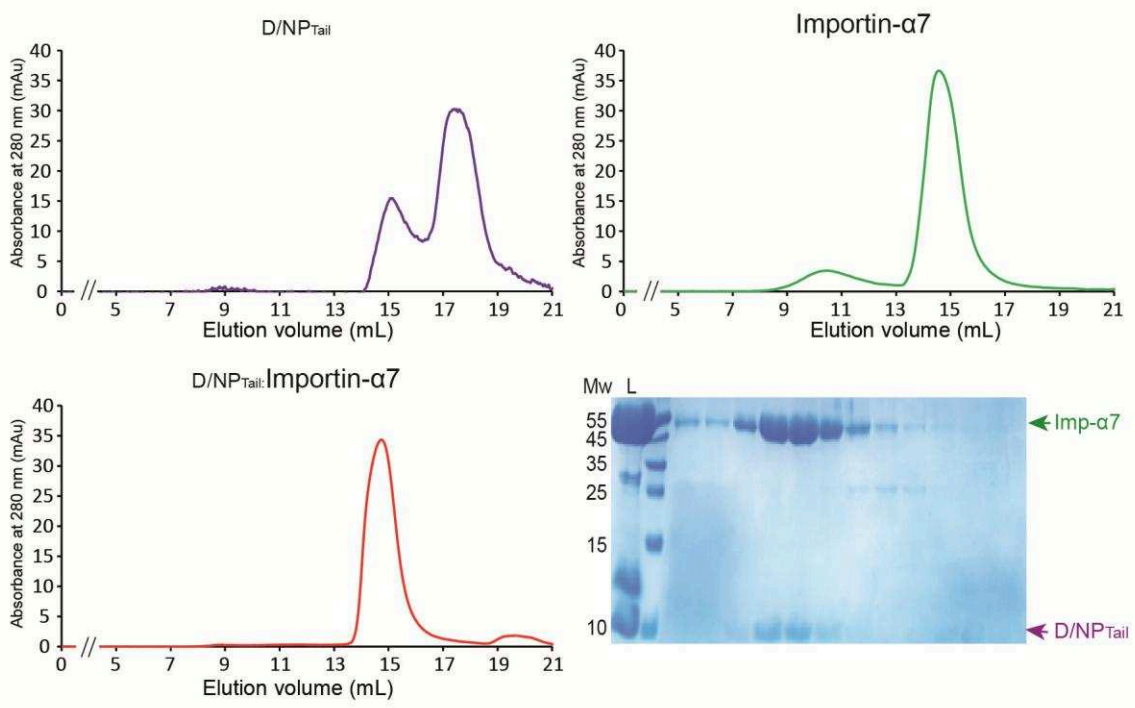
Table 2: K_d values of the D/NP:RNA interactions obtained from the titration curves. The fluorescence anisotropy measurements have been performed at 150 and 300 mM NaCl using $^{5'}$phosphate-polyUC-FAM$^{3'}$.

	D/NP				D/NP-529		D/NP-511	
[NaCl]	150 mM		300 mM		300 mM		300 mM	
RNA	K_d (nM)	R^2	K_d (nM)	R^2	K_d (nM)	R^2	K_d (nM)	R^2
5 nt	352 ± 14	0.997	nd	nd	nd	nd	nd	nd
8 nt	108 ± 3.3	0.998	1064 ± 75	0.998	nd	nd	nd	nd
10 nt	48 ± 1.2	0.997	229 ± 9	0.997	221 ± 7.4	0.998	260 ± 19	0.997
12 nt	13 ± 0.3	0.993	108 ± 4.2	0.989	46.1 ± 1.7	0.996	28 ± 1.4	0.991
16 nt	14 ± 0.3	0.988	14 ± 0.3	0.990	11 ± 0.3	0.993	13 ± 0.3	0.987
24 nt	14 ± 0.3	0.991	10 ± 0.2	0.995	nd	nd	nd	nd
32 nt	14 ± 0.2	0.994	16 ± 0.4	0.990	nd	nd	nd	nd

supFigure-1


B
Importin-α7
13.1 mL
Exp Mw = 60 kDa ± 0.9 %
Theo Mw = 55 kDa
Excess of refractive index (10^5)
Molar Mass (kDa)
Elution volume (mL)
A
D/NP
10.9 mL
Exp Mw = 260 kDa ± 0.95 %
Theo Mw = 256 kDa
Excess of refractive index (10^5)
Molar Mass (kDa)
Elution volume (mL)

supFigure-2


D/NPTail
Absorbance at 280 nm (mAu)
Elution volume (mL)
Importin-α7
Absorbance at 280 nm (mAu)
Elution volume (mL)
D/NPTail:Importin-α7
Absorbance at 280 nm (mAu)
Elution volume (mL)
Mw L
55
45
35
25
15
10
Imp-α7
D/NPTail

References

1. **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:**1634-1637.
2. **Baudin, F., C. Bach, S. Cusack, and R. W. Ruigrok.** 1994. Structure of influenza virus RNP. I. Influenza virus nucleoprotein melts secondary structure in panhandle RNA and exposes the bases to the solvent. The EMBO journal **13:**3158-3165.
3. **Bloyet, L. M., J. Brunel, M. Dosnon, V. Hamon, J. Erales, A. Gruet, C. Lazert, C. Bignon, P. Roche, S. Longhi, and D. Gerlier.** 2016. Modulation of Re-initiation of Measles Virus Transcription at Intergenic Regions by PXD to NTAIL Binding Strength. PLoS pathogens **12:**e1006058.
4. **Boivin, S., and D. J. Hart.** 2011. Interaction of the influenza A virus polymerase PB2 C-terminal region with importin alpha isoforms provides insights into host adaptation and polymerase assembly. The Journal of biological chemistry **286:**10439-10448.
5. **Chenavas, S., L. F. Estrozi, A. Slama-Schwok, B. Delmas, C. Di Primo, F. Baudin, X. Li, T. Crepin, and R. W. Ruigrok.** 2013. Monomeric nucleoprotein of influenza A virus. PLoS pathogens **9:**e1003275.
6. **Cohen, S. X., M. Ben Jelloul, F. Long, A. Vagin, P. Knipscheer, J. Lebbink, T. K. Sixma, V. S. Lamzin, G. N. Murshudov, and A. Perrakis.** 2008. ARP/wARP and molecular replacement: the next generation. Acta crystallographica. Section D, Biological crystallography **64:**49-60.
7. **Communie, G., R. W. Ruigrok, M. R. Jensen, and M. Blackledge.** 2014. Intrinsically disordered proteins implicated in paramyxoviral replication machinery. Current opinion in virology **5:**72-81.

8. **Cros, J. F., A. Garcia-Sastre, and P. Palese.** 2005. An unconventional NLS is critical for the nuclear import of the influenza A virus nucleoprotein and ribonucleoprotein. Traffic **6:**205-213.
9. **Elton, D., L. Medcalf, K. Bishop, D. Harrison, and P. Digard.** 1999. Identification of amino acid residues of influenza virus nucleoprotein essential for RNA binding. Journal of virology **73:**7357-7367.
10. **Emsley, P., B. Lohkamp, W. G. Scott, and K. Cowtan.** 2010. Features and development of Coot. Acta crystallographica. Section D, Biological crystallography **66:**486-501.
11. **Gallagher, J. R., U. Torian, D. M. McCraw, and A. K. Harris.** 2016. Structural studies of influenza virus RNPs by electron microscopy indicate molecular contortions within NP supra-structures. Journal of structural biology.
12. **Gaudin, Y., R. W. Ruigrok, C. Tuffereau, M. Knossow, and A. Flamand.** 1992. Rabies virus glycoprotein is a trimer. Virology **187:**627-632.
13. **Gerard, F. C., A. Ribeiro Ede, Jr., C. Leyrat, I. Ivanov, D. Blondel, S. Longhi, R. W. Ruigrok, and M. Jamin.** 2009. Modular organization of rabies virus phosphoprotein. Journal of molecular biology **388:**978-996.
14. **Gutsche, I., A. Desfosses, G. Effantin, W. L. Ling, M. Haupt, R. W. Ruigrok, C. Sachse, and G. Schoehn.** 2015. Structural virology. Near-atomic cryo-EM structure of the helical measles virus nucleocapsid. Science **348:**704-707.
15. **Hause, B. M., M. Ducatez, E. A. Collin, Z. Ran, R. Liu, Z. Sheng, A. Armien, B. Kaplan, S. Chakravarty, A. D. Hoppe, R. J. Webby, R. R. Simonson, and F. Li.** 2013. Isolation of a novel swine influenza virus from Oklahoma in 2011 which is distantly related to human influenza C viruses. PLoS pathogens **9:**e1003176.

16. **Holm, L., and P. Rosenstrom.** 2010. Dali server: conservation mapping in 3D. Nucleic acids research **38:**W545-549.
17. **Jensen, M. R., G. Communie, E. A. Ribeiro, Jr., N. Martinez, A. Desfosses, L. Salmon, L. Mollica, F. Gabel, M. Jamin, S. Longhi, R. W. Ruigrok, and M. Blackledge.** 2011. Intrinsic disorder in measles virus nucleocapsids. Proceedings of the National Academy of Sciences of the United States of America **108:**9839-9844.
18. **Kabsch, W.** 2010. Xds. Acta crystallographica. Section D, Biological crystallography **66:**125-132.
19. **Klumpp, K., R. W. Ruigrok, and F. Baudin.** 1997. Roles of the influenza virus polymerase and nucleoprotein in forming a functional RNP structure. The EMBO journal **16:**1248-1257.
20. **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**.
21. **Liu, J., Y. Cong, R. Yin, C. Ding, S. Yu, X. Liu, C. Wang, and Z. Ding.** 2014. The deletion of an extra six nucleotides in the 5' -untranslated region of the nucleoprotein gene of Newcastle disease virus NA-1 decreases virulence. BMC veterinary research **10:**964.
22. **Longhi, S., V. Receveur-Brechot, D. Karlin, K. Johansson, H. Darbon, D. Bhella, R. Yeo, S. Finet, and B. Canard.** 2003. The C-terminal domain of the measles virus nucleoprotein is intrinsically disordered and folds upon binding to the C-terminal moiety of the phosphoprotein. The Journal of biological chemistry **278:**18638-18648.
23. **McCoy, A. J., R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, and R. J. Read.** 2007. Phaser crystallographic software. Journal of applied crystallography **40:**658-674.

24. **Milles, S., M. R. Jensen, G. Communie, D. Maurin, G. Schoehn, R. W. Ruigrok, and M. Blackledge.** 2016. Self-Assembly of Measles Virus Nucleocapsid-like Particles: Kinetics and RNA Sequence Dependence. Angewandte Chemie **55:**9356-9360.
25. **Moeller, A., R. N. Kirchdoerfer, C. S. Potter, B. Carragher, and I. A. Wilson.** 2012. Organization of the influenza virus replication machinery. Science **338:**1631-1634.
26. **Murshudov, G. N., A. A. Vagin, and E. J. Dodson.** 1997. Refinement of macromolecular structures by the maximum-likelihood method. Acta crystallographica. Section D, Biological crystallography **53:**240-255.
27. **Nakada, R., H. Hirano, and Y. Matsuura.** 2015. Structure of importin-alpha bound to a non-classical nuclear localization signal of the influenza A virus nucleoprotein. Scientific reports **5:**15055.
28. **Ng, A. K., M. K. Lam, H. Zhang, J. Liu, S. W. Au, P. K. Chan, J. Wang, and P. C. Shaw.** 2012. Structural basis for RNA binding and homo-oligomer formation by influenza B virus nucleoprotein. Journal of virology **86:**6758-6767.
29. **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 journal : official publication of the Federation of American Societies for Experimental Biology **22:**3638-3647.
30. **Ozawa, M., K. Fujii, Y. Muramoto, S. Yamada, S. Yamayoshi, A. Takada, H. Goto, T. Horimoto, and Y. Kawaoka.** 2007. Contributions of two nuclear localization signals of influenza A virus nucleoprotein to viral replication. Journal of virology **81:**30-41.
31. **Pantoliano, M. W., E. C. Petrella, J. D. Kwasnoski, V. S. Lobanov, J. Myslik, E. Graf, T. Carver, E. Asel, B. A. Springer, P. Lane, and F. R. Salemme.** 2001. High-density

446 miniaturized thermal shift assays as a general strategy for drug discovery. Journal of
447 biomolecular screening **6:**429-440.
448 32. **Schoehn, G., M. Mavrakis, A. Albertini, R. Wade, A. Hoenger, and R. W. Ruigrok.**
449 2004. The 12 A structure of trypsin-treated measles virus N-RNA. Journal of molecular
450 biology **339:**301-312.
451 33. **Schrodinger, LLC.** 2015. The PyMOL Molecular Graphics System, Version 1.8.
452 34. **Sherry, L., M. Smith, S. Davidson, and D. Jackson.** 2014. The N terminus of the
453 influenza B virus nucleoprotein is essential for virus viability, nuclear localization, and
454 optimal transcription and replication of the viral genome. Journal of virology **88:**12326-
455 12338.
456 35. **Sievers, F., A. Wilm, D. Dineen, T. J. Gibson, K. Karplus, W. Li, R. Lopez, H.**
457 **McWilliam, M. Remmert, J. Soding, J. D. Thompson, and D. G. Higgins.** 2011. Fast,
458 scalable generation of high-quality protein multiple sequence alignments using Clustal
459 Omega. Molecular systems biology **7:**539.
460 36. **Smart, O. S., T. O. Womack, C. Flensburg, P. Keller, W. Paciorek, A. Sharff, C.**
461 **Vonrhein, and G. Bricogne.** 2012. Exploiting structure similarity in refinement:
462 automated NCS and target-structure restraints in BUSTER. Acta crystallographica.
463 Section D, Biological crystallography **68:**368-380.
464 37. **Song, H., J. Qi, Z. Khedri, S. Diaz, H. Yu, X. Chen, A. Varki, Y. Shi, and G. F. Gao.**
465 2016. An Open Receptor-Binding Cavity of Hemagglutinin-Esterase-Fusion Glycoprotein
466 from Newly-Identified Influenza D Virus: Basis for Its Broad Cell Tropism. PLoS
467 pathogens **12:**e1005411.
468 38. **Stevens, M. P., and W. S. Barclay.** 1998. The N-terminal extension of the influenza B
469 virus nucleoprotein is not required for nuclear accumulation or the expression and
470 replication of a model RNA. Journal of virology **72:**5307-5312.

39. **Tarus, B., O. Bakowiez, S. Chenavas, L. Duchemin, L. F. Estrozi, C. Bourdieu, N. Lejal, J. Bernard, M. Moudjou, C. Chevalier, B. Delmas, R. W. Ruigrok, C. Di Primo, and A. Slama-Schwok.** 2012. Oligomerization paths of the nucleoprotein of influenza A virus. Biochimie **94:**776-785.

40. **Turrell, L., J. W. Lyall, L. S. Tiley, E. Fodor, and F. T. Vreede.** 2013. The role and assembly mechanism of nucleoprotein in influenza A virus ribonucleoprotein complexes. Nature communications **4:**1591.

41. **Uversky, V. N.** 2012. Size-exclusion chromatography in structural analysis of intrinsically disordered proteins. Methods in molecular biology **896:**179-194.

42. **Wang, P., P. Palese, and R. E. O'Neill.** 1997. The NPI-1/NPI-3 (karyopherin alpha) binding site on the influenza a virus nucleoprotein NP is a nonconventional nuclear localization signal. Journal of virology **71:**1850-1856.

43. **Wanitchang, A., J. Narkpuk, and A. Jongkaewwattana.** 2013. Nuclear import of influenza B virus nucleoprotein: involvement of an N-terminal nuclear localization signal and a cleavage-protection motif. Virology **443:**59-68.

44. **Weber, F., G. Kochs, S. Gruber, and O. Haller.** 1998. A classical bipartite nuclear localization signal on Thogoto and influenza A virus nucleoproteins. Virology **250:**9-18.

45. **White, S. K., W. Ma, C. J. McDaniel, G. C. Gray, and J. A. Lednicky.** 2016. Serologic evidence of exposure to influenza D virus among persons with occupational contact with cattle. Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology **81:**31-33.

46. **Winn, M. D., C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin, and K. S. Wilson.**

2011. Overview of the CCP4 suite and current developments. Acta crystallographica. Section D, Biological crystallography **67:**235-242.

47. **Wu, W. W., Y. H. Sun, and N. Pante.** 2007. Nuclear import of influenza A viral ribonucleoprotein complexes is mediated by two nuclear localization sequences on viral nucleoprotein. Virology journal **4:**49.
48. **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:**1078-1082.
49. **Zheng, W., J. Olson, V. Vakharia, and Y. J. Tao.** 2013. The crystal structure and RNA-binding of an orthomyxovirus nucleoprotein. PLoS pathogens **9:**e1003624.

CONCLUSION

Interaction de la nucléoprotéine du virus de la grippe avec l'importine-α

Pour qu'il y'ait une réplication efficace *in* vivo, il a été montré que la NP fixe différents types d'importine-α en fonction de l'hôte (Gabriel et al., 2011). Dans le cas d'un virus aviaire, NP A fixe l'importine-α3 alors que dans le cas d'un virus humain, elle fixe l'importine-α7. Cette propriété a été déterminante dans le choix de l'importine utilisée au cours de cette thèse pour étudier l'interaction de cette protéine avec la nucléoprotéine. De manière concomitante au travail présenté ici, une structure de la NLS de NP A en complexe avec l'importine a été obtenue mais avec l'importine-α1 de souris (Nakada et al., 2015). Nous avons donc souhaité étudier l'interaction en présence d'une importine humaine.

Contrairement à NP A qui fixe l'importine via une NLS non conventionnelle située au niveau des 14 premiers résidus de la protéine (Boulo et al., 2011; Cros et al., 2005), les NLS de NP C et NP D semblaient logiquement situées sur leur queue C-terminale. Au cours de ce travail de thèse, nous avons déterminé plus précisément le site de la NLS sur NP D. Pour étudier cette interaction nous avons délibérément choisi d'utiliser l'importine-α7 en considérant la forte similitude (50 % d'identité) existant entre ce virus et le virus humain de type C (Hause et al., 2013) rendant donc fort possible la transmission du virus à l'homme (White et al., 2016) notamment grâce aux propriétés d'adaptation de la protéine membranaire HEF (Song et al., 2016). Grâce à plusieurs expériences de chromatographie d'exclusion de taille et de *pull-down* nous avons montré que NP D possède une NLS bipartite $_{514}$KRR-X$_{14}$-KKRGR$_{545}$. Cette NLS située sur l'extrémité carboxy-terminale interagit avec l'importine-α7. Il est a noté qu'un autre point de régulation essentiel non abordé au cours de ce travail de thèse, se situe au niveau de la phosphorylation de NP. Cette phosphorylation pourrait empêcher la fixation de l'importine sur la nucléoprotéine (Hutchinson and Fodor, 2013; Zheng et al., 2013).

Interaction de la nucléoprotéine avec l'ARN

Dans le but de parfaire la compréhension du fonctionnement de la machinerie virale, une attention particulière a été portée sur l'interaction entre la nucléoprotéine et l'ARN au cours de ce travail de thèse. Contrairement à la majorité des virus appartenant à l'ordre des *Mononegavirales* et de quelques virus à ARN segmenté (Reguera et al., 2014; Ruigrok et al., 2011) aucune structure de nucléoprotéine du virus *influenza* n'a été obtenue en présence d'ARN. Toutefois, de nombreuses études montrent que la NP du virus de la grippe fixerait entre 20 et 24 nucléotides sans spécificité de séquence (Compans et al., 1972; Martín-Benito et al., 2001; Ortega et al., 2000). Ce nombre est bien plus élevé que la plupart des autres virus à ARN négatif fixant en moyenne entre 6 et 11 bases (Reguera et al., 2014; Ruigrok et

al., 2011). Au cours de cette thèse, les expériences d'anisotropie de fluorescence réalisées avec les NP des virus de type A, B et D montrent que la NP possède une très forte affinité pour un ARN poly-UC de 12 nucléotides (Articles 1 et 2).

Pour NP A, les structures *RNP-like* clairement visibles en microscopie électronique et obtenues uniquement en présence d'un ARN de 12 nucléotides appuient fortement cette hypothèse (Article 1). Ce nombre de nucléotides est en accord avec celui obtenu pour la nucléoprotéine du virus de l'anémie infectieuse du saumon, appartenant lui aussi à la famille des *Orthomyxoviridae* (Zheng et al., 2013). Toutefois cette information remettrait en cause de nombreux travaux réalisés au cours des 50 dernières années. Afin de mieux interpréter les différences observées entre ces études et la nôtre nous pouvons proposer un modèle dans lequel la NP fixerait avec une forte affinité uniquement 12 bases d'un ARN tandis que les 12 autres bases seraient fixées avec une affinité beaucoup plus faible. Ces 12 bases plus ou moins libres pourraient alors s'organiser de manière à venir fixer une RNP voisine permettant la mise en place du complexe supramoléculaire nécessaire à la formation des virions néo-synthétisés (Fournier et al., 2012a, 2012b) (Figure 34).

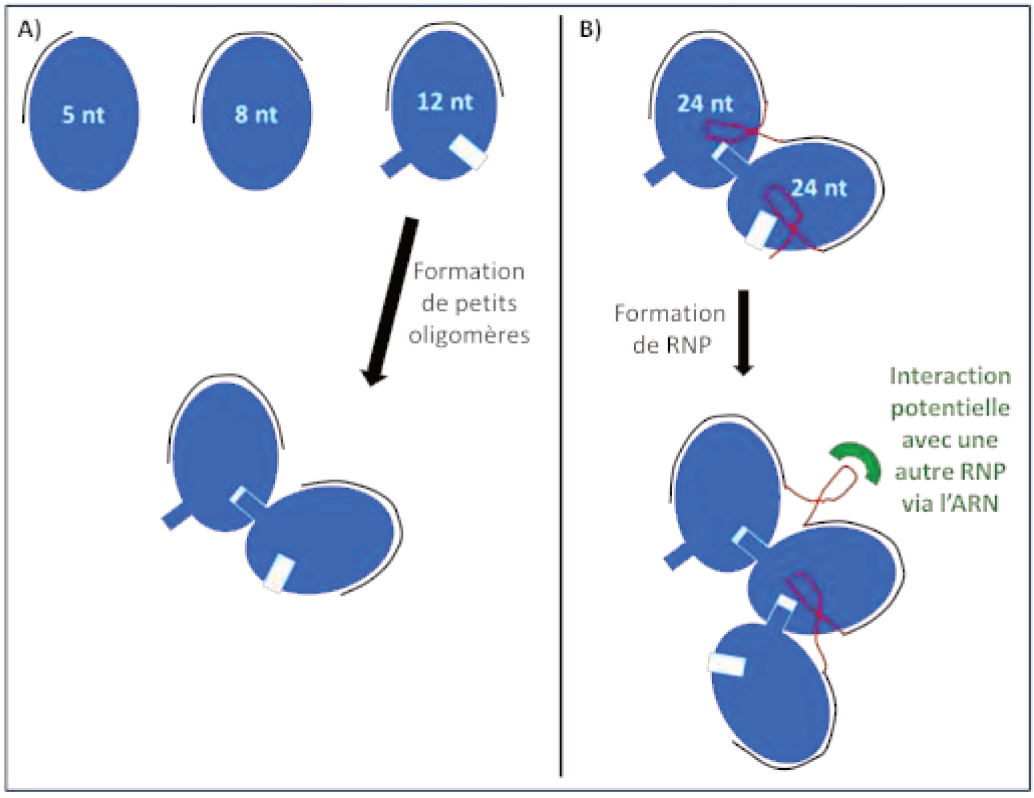


Figure 34 : Représentation schématique concernant le mécanisme hypothétique de fixation de la nucléoprotéine d'influenza A à l'ARN. A) NP A est capable de lier de petits ARN avec une affinité croissante en fonction de la taille. La fixation à l'ARN entraine la formation de petits oligomères, notamment des trimères et des tétramères. B) Chaque nucléoprotéine serait en fait capable de lier 24 nucléotides de l'ARNv. Sur ces 24 bases seulement 12 seraient liées à NP A avec une très forte affinité tandis que les 12 autres bases pourraient être impliquées dans le recrutement des RNP voisines notamment lors des processus de formation des nouveaux virions.

Concernant NP D, nous avons aussi montré par des expériences d'anisotropie de fluorescence que l'extrémité C-terminale régule de manière négative la fixation à l'ARN. En effet en présence de la queue C-terminale l'affinité de la protéine pour l'ARN diminue de manière significative.

Ainsi, ce travail de thèse, bien que n'ayant pas répondu parfaitement à l'objectif initial étant d'obtenir une structure de NP en présence de l'ARN afin de mieux comprendre le fonctionnement de la machinerie virale a toutefois permis d'apporter de nombreuses informations biochimiques concernant cette interaction. De plus, il a permis de déterminer la structure de la nucléoprotéine du virus de type D, de caractériser son interaction avec l'importine-α7 et d'apporter une information essentielle quant à l'importance de l'extrémité C-terminale de NP D dans la fixation à l'ARN. Ces informations ont permis de formuler de nouvelles hypothèses quant au fonctionnement du virus de la grippe et permettre d'inscrire ce projet de thèse dans une dynamique globale de lutte contre ce virus.

Listes des abréviations

ARN : Acide RiboNucléique

ARNc : Acide RiboNucléique complétaire

ARNm: Acide RiboNucléique Messager

ARNv: Acide RiboNucléique Viral

CD: Dichroïsme Circulaire *(Circular Dichroïsm)*

ESRF : *European Synchrotron Radiation Facility*

FAM : Fluoréscéine

HA : Hémagglutinine

HEF : Hémagglutinine-Estérase de Fusion

IPTG: Isopropyl β-D-1-thiogalactopyranoside

ITC : Titration Calorimètrique Isotherme (*Isothermal Titration Calorimetry)*

MAD : Dispersion Anomale Multiple *(Multiple Anomalous Dispersion)*

MET : Microscopie Electronique à Transmission

MIR : Remplacement Isopmorphe Multiple *(Multiple Isomorphous Remplacement)*

NA : Neuraminidase

NLS : Signal de Localisation Nucléaire *(Nuclear Localisation Signal)*

NP : Nucléoprotéine

NPC : Complexe du Pore Nucléaire

NS1 : *Non Structural Protein 1*

NS2/NEP : *Non Structural Protein 2*

ORF : Phase Ouverte de Lecture

P : Phosphoprotéine

PA : Protéine Acide

PB1/PB2 : Protéine Basique 1/2

RdRp : ARN-polymérase-ARN-dépendante

RE: Réticulum Endoplasmique

RMN : Résonance Magnétique Nucléaire

RNP : Ribonucléoprotéine

SAD: Dispersion Anomale Simple *(Single Anomalous Dispersion)*

SIR: Remplacement Isomorphe Simple *(Single Isomorphous Remplacement)*

Tm : Température de fusion *(Melting Temperature)*

TSA: Test de Stabilité Thermique *(Thermal Shift Assay)*

UTR: Untranslated Region

Bibliographie

Adam, E.J., and Adam, S.A. (1994). Identification of cytosolic factors required for nuclear location sequence-mediated binding to the nuclear envelope. J. Cell Biol. *125*, 547–555.

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

Akey, C.W., and Radermacher, M. (1993). Architecture of the Xenopus nuclear pore complex revealed by three-dimensional cryo-electron microscopy. J. Cell Biol. *122*, 1–19.

Alamgir, A.S., Matsuzaki, Y., Hongo, S., Tsuchiya, E., Sugawara, K., Muraki, Y., and Nakamura, K. (2000). Phylogenetic analysis of influenza C virus nonstructural (NS) protein genes and identification of the NS2 protein. J. Gen. Virol. *81*, 1933–1940.

Alayyoubi, M., Leser, G.P., Kors, C.A., and Lamb, R.A. (2015). Structure of the paramyxovirus parainfluenza virus 5 nucleoprotein-RNA complex. Proc. Natl. Acad. Sci. U. S. A. *112*, E1792-1799.

Alber, F., Dokudovskaya, S., Veenhoff, L.M., Zhang, W., Kipper, J., Devos, D., Suprapto, A., Karni-Schmidt, O., Williams, R., Chait, B.T., et al. (2007). The molecular architecture of the nuclear pore complex. Nature *450*, 695–701.

Albertini, A.A.V., Wernimont, A.K., Muziol, T., Ravelli, R.B.G., Clapier, C.R., Schoehn, G., Weissenhorn, W., and Ruigrok, R.W.H. (2006). Crystal structure of the rabies virus nucleoprotein-RNA complex. Science *313*, 360–363.

Albo, C., Valencia, A., and Portela, A. (1995). Identification of an RNA binding region within the N-terminal third of the influenza A virus nucleoprotein. J. Virol. *69*, 3799–3806.

Alvisi, G., and Jans, D.A. (2015). Basis of cargo recognition by importin αs: the power of structure. Struct. Lond. Engl. 1993 *23*, 251–252.

Apostolov, K., and Flewett, T.H. (1969). Further observations on the structure of influenza viruses A and C. J. Gen. Virol. *4*, 365–370.

Ariza, A., Tanner, S.J., Walter, C.T., Dent, K.C., Shepherd, D.A., Wu, W., Matthews, S.V., Hiscox, J.A., Green, T.J., Luo, M., et al. (2013). Nucleocapsid protein structures from orthobunyaviruses reveal insight into ribonucleoprotein architecture and RNA polymerization. Nucleic Acids Res. *41*, 5912–5926.

Arranz, R., Coloma, R., Chichón, F.J., Conesa, J.J., Carrascosa, J.L., Valpuesta, J.M., Ortín, J., and Martín-Benito, J. (2012). The structure of native influenza virion ribonucleoproteins. Science *338*, 1634–1637.

Arzt, S., Baudin, F., Barge, A., Timmins, P., Burmeister, W.P., and Ruigrok, R.W. (2001). Combined results from solution studies on intact influenza virus M1 protein and from a new crystal form of its N-terminal domain show that M1 is an elongated monomer. Virology *279*, 439–446.

Babu, Y.S., Chand, P., Bantia, S., Kotian, P., Dehghani, A., El-Kattan, Y., Lin, T.H., Hutchison, T.L., Elliott, A.J., Parker, C.D., et al. (2000). BCX-1812 (RWJ-270201): discovery of a novel, highly potent, orally active, and selective influenza neuraminidase inhibitor through structure-based drug design. J. Med. Chem. *43*, 3482–3486.

Bae, S.H., Cheong, H.K., Lee, J.H., Cheong, C., Kainosho, M., and Choi, B.S. (2001). Structural features of an influenza virus promoter and their implications for viral RNA synthesis. Proc. Natl. Acad. Sci. U. S. A. *98*, 10602–10607.

Baltimore, D. (1971). Expression of animal virus genomes. Bacteriol. Rev. *35*, 235–241.

Bancroft, C.T., and Parslow, T.G. (2002). Evidence for segment-nonspecific packaging of the influenza a virus genome. J. Virol. *76*, 7133–7139.

Baquero, E., Albertini, A.A., Vachette, P., Lepault, J., Bressanelli, S., and Gaudin, Y. (2013). Intermediate conformations during viral fusion glycoprotein structural transition. Curr. Opin. Virol. *3*, 143–150.

Baudin, F., Bach, C., Cusack, S., and Ruigrok, R.W. (1994). Structure of influenza virus RNP. I. Influenza virus nucleoprotein melts secondary structure in panhandle RNA and exposes the bases to the solvent. EMBO J. *13*, 3158–3165.

Baudin, F., Petit, I., Weissenhorn, W., and Ruigrok, R.W.H. (2001). In Vitro Dissection of the Membrane and RNP Binding Activities of Influenza Virus M1 Protein. Virology *281*, 102–108.

Bayliss, R., Littlewood, T., and Stewart, M. (2000). Structural basis for the interaction between FxFG nucleoporin repeats and importin-beta in nuclear trafficking. Cell *102*, 99–108.

Beck, M., Förster, F., Ecke, M., Plitzko, J.M., Melchior, F., Gerisch, G., Baumeister, W., and Medalia, O. (2004). Nuclear pore complex structure and dynamics revealed by cryoelectron tomography. Science *306*, 1387–1390.

Beck, M., Lucić, V., Förster, F., Baumeister, W., and Medalia, O. (2007). Snapshots of nuclear pore complexes in action captured by cryo-electron tomography. Nature *449*, 611–615.

Bedford, T., Cobey, S., Beerli, P., and Pascual, M. (2010). Global migration dynamics underlie evolution and persistence of human influenza A (H3N2). PLoS Pathog. *6*, e1000918.

Belshe, R.B., Burk, B., Newman, F., Cerruti, R.L., and Sim, I.S. (1989). Resistance of influenza A virus to amantadine and rimantadine: results of one decade of surveillance. J. Infect. Dis. *159*, 430–435.

Betakova, T., Nermut, M.V., and Hay, A.J. (1996). The NB protein is an integral component of the membrane of influenza B virus. J. Gen. Virol. *77 (Pt 11)*, 2689–2694.

Beveridge, W.I. (1977). The start of pandemics: site, season and spread. Dev. Biol. Stand. *39*, 443–444.

Biswas, S.K., Boutz, P.L., and Nayak, D.P. (1998). Influenza virus nucleoprotein interacts with influenza virus polymerase proteins. J. Virol. *72*, 5493–5501.

Blaas, D., Patzelt, E., and Kuechler, E. (1982). Identification of the cap binding protein of influenza virus. Nucleic Acids Res. *10*, 4803–4812.

Blok, J., Air, G.M., Laver, W.G., Ward, C.W., Lilley, G.G., Woods, E.F., Roxburgh, C.M., and Inglis, A.S. (1982). Studies on the size, chemical composition, and partial sequence of the neuraminidase (NA) from type A influenza viruses show that the N-terminal region of the NA is not processed and serves to anchor the NA in the viral membrane. Virology *119*, 109–121.

Boivin, S., Cusack, S., Ruigrok, R.W.H., and Hart, D.J. (2010). Influenza A virus polymerase: structural insights into replication and host adaptation mechanisms. J. Biol. Chem. *285*, 28411–28417.

Booy, F.P., Ruigrok, R.W., and van Bruggen, E.F. (1985). Electron microscopy of influenza virus. A comparison of negatively stained and ice-embedded particles. J. Mol. Biol. *184*, 667–676.

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

Boulo, S., Akarsu, H., Ruigrok, R.W.H., and Baudin, F. (2007). Nuclear traffic of influenza virus proteins and ribonucleoprotein complexes. Virus Res. *124*, 12–21.

Boulo, S., Akarsu, H., Lotteau, V., Müller, C.W., Ruigrok, R.W.H., and Baudin, F. (2011). Human importin alpha and RNA do not compete for binding to influenza A virus nucleoprotein. Virology *409*, 84–90.

Bouloy, M., Plotch, S.J., and Krug, R.M. (1978). Globin mRNAs are primers for the transcription of influenza viral RNA in vitro. Proc. Natl. Acad. Sci. U. S. A. *75*, 4886–4890.

Brenner, S., and Horne, R.W. (1959). A negative staining method for high resolution electron microscopy of viruses. Biochim. Biophys. Acta *34*, 103–110.

Broor, S., Krishnan, A., Roy, D.S., Dhakad, S., Kaushik, S., Mir, M.A., Singh, Y., Moen, A., Chadha, M., Mishra, A.C., et al. (2012). Dynamic patterns of circulating seasonal and pandemic A(H1N1)pdm09 influenza viruses from 2007-2010 in and around Delhi, India. PloS One *7*, e29129.

Brundage, J.F. (2006). Interactions between influenza and bacterial respiratory pathogens: implications for pandemic preparedness. Lancet Infect. Dis. *6*, 303–312.

Bui, M., Myers, J.E., and Whittaker, G.R. (2002). Nucleo-cytoplasmic localization of influenza virus nucleoprotein depends on cell density and phosphorylation. Virus Res. *84*, 37–44.

Bulletin of the World Health Organization (1980). A revision of the system of nomenclature for influenza viruses: a WHO Memorandum. Bull. World Health Organ. *58*, 585–591.

Bullido, R., Gómez-Puertas, P., Albo, C., and Portela, A. (2000). Several protein regions contribute to determine the nuclear and cytoplasmic localization of the influenza A virus nucleoprotein. J. Gen. Virol. *81*, 135–142.

Bullido, R., Gómez-Puertas, P., Saiz, M.J., and Portela, A. (2001). Influenza A virus NEP (NS2 protein) downregulates RNA synthesis of model template RNAs. J. Virol. *75*, 4912–4917.

Bush, R.M., Bender, C.A., Subbarao, K., Cox, N.J., and Fitch, W.M. (1999). Predicting the evolution of human influenza A. Science *286*, 1921–1925.

Carrillo, B., Choi, J.-M., Bornholdt, Z.A., Sankaran, B., Rice, A.P., and Prasad, B.V.V. (2014). The influenza A virus protein NS1 displays structural polymorphism. J. Virol. *88*, 4113–4122.

CDC (2016). Influenza (flu) including seasonal, avian, swine, pandemic, and other.

Chan, J., Holmes, A., and Rabadan, R. (2010). Network analysis of global influenza spread. PLoS Comput. Biol. *6*, e1001005.

Chan, J.F.-W., To, K.K.-W., Tse, H., Jin, D.-Y., and Yuen, K.-Y. (2013). Interspecies transmission and emergence of novel viruses: lessons from bats and birds. Trends Microbiol. *21*, 544–555.

Chen, B.J., and Lamb, R.A. (2008). Mechanisms for enveloped virus budding: Can some viruses do without an ESCRT? Virology *372*, 221–232.

Chen, R., and Holmes, E.C. (2008). The evolutionary dynamics of human influenza B virus. J. Mol. Evol. *66*, 655–663.

Chen, W., Calvo, P.A., Malide, D., Gibbs, J., Schubert, U., Bacik, I., Basta, S., O'Neill, R., Schickli, J., Palese, P., et al. (2001). A novel influenza A virus mitochondrial protein that induces cell death. Nat. Med. *7*, 1306–1312.

Chenavas, S., Estrozi, L.F., Slama-Schwok, A., Delmas, B., Di Primo, C., Baudin, F., Li, X., Crépin, T., and Ruigrok, R.W.H. (2013a). Monomeric nucleoprotein of influenza A virus. PLoS Pathog. *9*, e1003275.

Chenavas, S., Crépin, T., Delmas, B., Ruigrok, R.W.H., and Slama-Schwok, A. (2013b). Influenza virus nucleoprotein: structure, RNA binding, oligomerization and antiviral drug target. Future Microbiol. *8*, 1537–1545.

Cheng, A., Wong, S.M., and Yuan, Y.A. (2009). Structural basis for dsRNA recognition by NS1 protein of influenza A virus. Cell Res. *19*, 187–195.

Cheong, H.K., Cheong, C., Lee, Y.S., Seong, B.L., and Choi, B.S. (1999). Structure of influenza virus panhandle RNA studied by NMR spectroscopy and molecular modeling. Nucleic Acids Res. *27*, 1392–1397.

Chien, Y.-W., Klugman, K.P., and Morens, D.M. (2009). Bacterial pathogens and death during the 1918 influenza pandemic. N. Engl. J. Med. *361*, 2582–2583.

Chook, Y.M., and Blobel, G. (2001). Karyopherins and nuclear import. Curr. Opin. Struct. Biol. *11*, 703–715.

Chou, Y., Vafabakhsh, R., Doğanay, S., Gao, Q., Ha, T., and Palese, P. (2012). One influenza virus particle packages eight unique viral RNAs as shown by FISH analysis. Proc. Natl. Acad. Sci. *109*, 9101–9106.

Chua, K.B., Bellini, W.J., Rota, P.A., Harcourt, B.H., Tamin, A., Lam, S.K., Ksiazek, T.G., Rollin, P.E., Zaki, S.R., Shieh, W., et al. (2000). Nipah virus: a recently emergent deadly paramyxovirus. Science *288*, 1432–1435.

Cingolani, G., Petosa, C., Weis, K., and Müller, C.W. (1999). Structure of importin-beta bound to the IBB domain of importin-alpha. Nature *399*, 221–229.

Claas, E.C., Osterhaus, A.D., van Beek, R., De Jong, J.C., Rimmelzwaan, G.F., Senne, D.A., Krauss, S., Shortridge, K.F., and Webster, R.G. (1998a). Human influenza A H5N1 virus related to a highly pathogenic avian influenza virus. Lancet Lond. Engl. *351*, 472–477.

Claas, E.C., de Jong, J.C., van Beek, R., Rimmelzwaan, G.F., and Osterhaus, A.D. (1998b). Human influenza virus A/HongKong/156/97 (H5N1) infection. Vaccine *16*, 977–978.

Claudia Dominguer (2010). Claudia Domínguez Ilustración.

Clodfelter, M. (2002). Warfare and Armed Conflicts: A Statistical Reference to Casualty and Other Figures, 1500-2000 (McFarland).

Cockburn, W.C., Delon, P.J., and Ferreira, W. (1969). Origin and progress of the 1968-69 Hong Kong influenza epidemic. Bull. World Health Organ. *41*, 345–348.

Collin, E.A., Sheng, Z., Lang, Y., Ma, W., Hause, B.M., and Li, F. (2014). Cocirculation of Two Distinct Genetic and Antigenic Lineages of Proposed Influenza D Virus in Cattle. J. Virol. *89*, 1036–1042.

Colman, P.M., Varghese, J.N., and Laver, W.G. (1983). Structure of the catalytic and antigenic sites in influenza virus neuraminidase. Nature *303*, 41–44.

Coloma, R., Valpuesta, J.M., Arranz, R., Carrascosa, J.L., Ortín, J., and Martín-Benito, J. (2009). The Structure of a Biologically Active Influenza Virus Ribonucleoprotein Complex. PLoS Pathog. *5*.

Compans, R.W., and Choppin, P.W. (1975). Reproduction of Myxoviruses. In Comprehensive Virology, H. Fraenkel-Conrat, and R.R. Wagner, eds. (Springer US), pp. 179–252.

Compans, R.W., Content, J., and Duesberg, P.H. (1972). Structure of the ribonucleoprotein of influenza virus. J. Virol. *10*, 795–800.

Connor, R.J., Kawaoka, Y., Webster, R.G., and Paulson, J.C. (1994). Receptor specificity in human, avian, and equine H2 and H3 influenza virus isolates. Virology *205*, 17–23.

Cros, J.F., and Palese, P. (2003). Trafficking of viral genomic RNA into and out of the nucleus: influenza, Thogoto and Borna disease viruses. Virus Res. *95*, 3–12.

Cros, J.F., García-Sastre, A., and Palese, P. (2005). An unconventional NLS is critical for the nuclear import of the influenza A virus nucleoprotein and ribonucleoprotein. Traffic Cph. Den. *6*, 205–213.

Curran, J., Pelet, T., and Kolakofsky, D. (1994). An acidic activation-like domain of the Sendai virus P protein is required for RNA synthesis and encapsidation. Virology *202*, 875–884.

Danieli, T., Pelletier, S.L., Henis, Y.I., and White, J.M. (1996). Membrane fusion mediated by the influenza virus hemagglutinin requires the concerted action of at least three hemagglutinin trimers. J. Cell Biol. *133*, 559–569.

Datta, K., Wolkerstorfer, A., Szolar, O.H.J., Cusack, S., and Klumpp, K. (2013). Characterization of PA-N terminal domain of Influenza A polymerase reveals sequence specific RNA cleavage. Nucleic Acids Res. *41*, 8289–8299.

Davies, W.L., Grunert, R.R., Haff, R.F., Mcgahen, J.W., Neumayer, E.M., Paulshock, M., Watts, J.C., Wood, T.R., Hermann, E.C., and Hoffmann, C.E. (1964). ANTIVIRAL ACTIVITY OF 1-ADAMANTANAMINE (AMANTADINE). Science *144*, 862–863.

Deng, T., Vreede, F.T., and Brownlee, G.G. (2006). Different de novo initiation strategies are used by influenza virus RNA polymerase on its cRNA and viral RNA promoters during viral RNA replication. J. Virol. *80*, 2337–2348.

Desfosses, A., Jr, E.A.R., Schoehn, G., Blondel, D., Guilligay, D., Jamin, M., Ruigrok, R.W.H., and Gutsche, I. (2013). Self-organization of the vesicular stomatitis virus nucleocapsid into a bullet shape. Nat. Commun. *4*, 1429.

Desselberger, U., Racaniello, V.R., Zazra, J.J., and Palese, P. (1980). The 3’ and 5’-terminal sequences of influenza A, B and C virus RNA segments are highly conserved and show partial inverted complementarity. Gene *8*, 315–328.

Digard, P., Blok, V.C., and Inglis, S.C. (1989). Complex formation between influenza virus polymerase proteins expressed in Xenopus oocytes. Virology *171*, 162–169.

Dingwall, C., and Laskey, R.A. (1991). Nuclear targeting sequences--a consensus? Trends Biochem. Sci. *16*, 478–481.

Dolin, R., Reichman, R.C., Madore, H.P., Maynard, R., Linton, P.N., and Webber-Jones, J. (1982). A controlled trial of amantadine and rimantadine in the prophylaxis of influenza A infection. N. Engl. J. Med. *307*, 580–584.

Donald, H.B., and Isaacs, A. (1954). Counts of influenza virus particles. J. Gen. Microbiol. *10*, 457–464.

Dong, H., Li, P., Böttcher, B., Elliott, R.M., and Dong, C. (2013). Crystal structure of Schmallenberg orthobunyavirus nucleoprotein-RNA complex reveals a novel RNA sequestration mechanism. RNA N. Y. N *19*, 1129–1136.

Drake, J.W. (1993). Rates of spontaneous mutation among RNA viruses. Proc. Natl. Acad. Sci. U. S. A. *90*, 4171–4175.

Ducatez, M.F., Pelletier, C., and Meyer, G. (2015). Influenza D virus in cattle, France, 2011-2014. Emerg. Infect. Dis. *21*, 368–371.

Duesberg, P.H. (1969). Distinct subunits of the ribonucleoprotein of influenza virus. J. Mol. Biol. *42*, 485–499.

Dutkowski, R., Thakrar, B., Froehlich, E., Suter, P., Oo, C., and Ward, P. (2003). Safety and pharmacology of oseltamivir in clinical use. Drug Saf. *26*, 787–801.

Eisfeld, A.J., Neumann, G., and Kawaoka, Y. (2015). At the centre: influenza A virus ribonucleoproteins. Nat. Rev. Microbiol. *13*, 28–41.

Elton, D., Medcalf, E., Bishop, K., and Digard, P. (1999). Oligomerization of the influenza virus nucleoprotein: identification of positive and negative sequence elements. Virology *260*, 190–200.

Elton, D., Simpson-Holley, M., Archer, K., Medcalf, L., Hallam, R., McCauley, J., and Digard, P. (2001). Interaction of the influenza virus nucleoprotein with the cellular CRM1-mediated nuclear export pathway. J. Virol. *75*, 408–419.

Enami, K., Sato, T.A., Nakada, S., and Enami, M. (1994). Influenza virus NS1 protein stimulates translation of the M1 protein. J. Virol. *68*, 1432–1437.

Enami, M., Sharma, G., Benham, C., and Palese, P. (1991). An influenza virus containing nine different RNA segments. Virology *185*, 291–298.

Engelhardt, O.G., Smith, M., and Fodor, E. (2005). Association of the influenza A virus RNA-dependent RNA polymerase with cellular RNA polymerase II. J. Virol. *79*, 5812–5818.

Ericsson, U.B., Hallberg, B.M., DeTitta, G.T., Dekker, N., and Nordlund, P. (2006). Thermofluor-based high-throughput stability optimization of proteins for structural studies. Anal. Biochem. *357*, 289–298.

Falcón, A.M., Marión, R.M., Zürcher, T., Gómez, P., Portela, A., Nieto, A., and Ortín, J. (2004). Defective RNA replication and late gene expression in temperature-sensitive influenza viruses expressing deleted forms of the NS1 protein. J. Virol. *78*, 3880–3888.

Ferron, F., Li, Z., Danek, E.I., Luo, D., Wong, Y., Coutard, B., Lantez, V., Charrel, R., Canard, B., Walz, T., et al. (2011). The hexamer structure of Rift Valley fever virus nucleoprotein suggests a mechanism for its assembly into ribonucleoprotein complexes. PLoS Pathog. *7*, e1002030.

Fodor, E., Pritlove, D.C., and Brownlee, G.G. (1994). The influenza virus panhandle is involved in the initiation of transcription. J. Virol. *68*, 4092–4096.

Fodor, E., Crow, M., Mingay, L.J., Deng, T., Sharps, J., Fechter, P., and Brownlee, G.G. (2002). A single amino acid mutation in the PA subunit of the influenza virus RNA polymerase inhibits endonucleolytic cleavage of capped RNAs. J. Virol. *76*, 8989–9001.

Fournier, E., Moules, V., Essere, B., Paillart, J.-C., Sirbat, J.-D., Isel, C., Cavalier, A., Rolland, J.-P., Thomas, D., Lina, B., et al. (2012). A supramolecular assembly formed by influenza A virus genomic RNA segments. Nucleic Acids Res. *40*, 2197–2209.

Francis, T. (1940). A NEW TYPE OF VIRUS FROM EPIDEMIC INFLUENZA. Science *92*, 405–408.

Fujii, Y., Goto, H., Watanabe, T., Yoshida, T., and Kawaoka, Y. (2003). Selective incorporation of influenza virus RNA segments into virions. Proc. Natl. Acad. Sci. U. S. A. *100*, 2002–2007.

Fujiyoshi, Y., Kume, N.P., Sakata, K., and Sato, S.B. (1994). Fine structure of influenza A virus observed by electron cryo-microscopy. EMBO J. *13*, 318–326.

Gabriel, G., Klingel, K., Otte, A., Thiele, S., Hudjetz, B., Arman-Kalcek, G., Sauter, M., Shmidt, T., Rother, F., Baumgarte, S., et al. (2011). Differential use of importin-α isoforms governs cell tropism and host adaptation of influenza virus. Nat. Commun. *2*, 156.

Gallagher, J.R., Torian, U., McCraw, D.M., and Harris, A.K. (2016). Structural studies of influenza virus RNPs by electron microscopy indicate molecular contortions within NP supra-structures. J. Struct. Biol.

Gamblin, S.J., and Skehel, J.J. (2010). Influenza hemagglutinin and neuraminidase membrane glycoproteins. J. Biol. Chem. *285*, 28403–28409.

Gao, Q., Lowen, A.C., Wang, T.T., and Palese, P. (2010). A nine-segment influenza a virus carrying subtype H1 and H3 hemagglutinins. J. Virol. *84*, 8062–8071.

Gao, Q., Chou, Y.-Y., Doğanay, S., Vafabakhsh, R., Ha, T., and Palese, P. (2012). The influenza A virus PB2, PA, NP, and M segments play a pivotal role during genome packaging. J. Virol. *86*, 7043–7051.

Gastaminza, P., Perales, B., Falcón, A.M., and Ortín, J. (2003). Mutations in the N-terminal region of influenza virus PB2 protein affect virus RNA replication but not transcription. J. Virol. *77*, 5098–5108.

Gavazzi, C., Isel, C., Fournier, E., Moules, V., Cavalier, A., Thomas, D., Lina, B., and Marquet, R. (2013). An in vitro network of intermolecular interactions between viral RNA segments of an avian H5N2 influenza A virus: comparison with a human H3N2 virus. Nucleic Acids Res. *41*, 1241–1254.

Ge, P., Tsao, J., Schein, S., Green, T.J., Luo, M., and Zhou, Z.H. (2010). Cryo-EM model of the bullet-shaped vesicular stomatitis virus. Science *327*, 689–693.

GEIG, Groupe d'Expertise et d'Information sur la Grippe Accueil.

Gerber, M., Isel, C., Moules, V., and Marquet, R. (2014). Selective packaging of the influenza A genome and consequences for genetic reassortment. Trends Microbiol. *22*, 446–455.

Goldfarb, D.S., Corbett, A.H., Mason, D.A., Harreman, M.T., and Adam, S.A. (2004). Importin alpha: a multipurpose nuclear-transport receptor. Trends Cell Biol. *14*, 505–514.

Goldstein, E.A., and Pons, M.W. (1970). The effect of polyvinylsulfate on the ribonucleoprotein of influenza virus. Virology *41*, 382–384.

González, S., and Ortín, J. (1999). Distinct regions of influenza virus PB1 polymerase subunit recognize vRNA and cRNA templates. EMBO J. *18*, 3767–3775.

Gorai, T., Goto, H., Noda, T., Watanabe, T., Kozuka-Hata, H., Oyama, M., Takano, R., Neumann, G., Watanabe, S., and Kawaoka, Y. (2012). F1Fo-ATPase, F-type proton-translocating ATPase, at the plasma membrane is critical for efficient influenza virus budding. Proc. Natl. Acad. Sci. U. S. A. *109*, 4615–4620.

Görlich, D., Vogel, F., Mills, A.D., Hartmann, E., and Laskey, R.A. (1995). Distinct functions for the two importin subunits in nuclear protein import. Nature *377*, 246–248.

Goto, H., Muramoto, Y., Noda, T., and Kawaoka, Y. (2013). The genome-packaging signal of the influenza A virus genome comprises a genome incorporation signal and a genome-bundling signal. J. Virol. *87*, 11316–11322.

Green, T.J., Zhang, X., Wertz, G.W., and Luo, M. (2006). Structure of the vesicular stomatitis virus nucleoprotein-RNA complex. Science *313*, 357–360.

Gubbay, O., Curran, J., and Kolakofsky, D. (2001). Sendai virus genome synthesis and assembly are coupled: a possible mechanism to promote viral RNA polymerase processivity. J. Gen. Virol. *82*, 2895–2903.

Guilligay, D., Chenavas, S., and Cusack, S. (2010). L'Arn polymérase du virus de la grippe : de la recherche fondamentale à l'élaboration de nouveaux antiviraux.

Gultyaev, A.P., Fouchier, R.A.M., and Olsthoorn, R.C.L. (2010). Influenza virus RNA structure: unique and common features. Int. Rev. Immunol. *29*, 533–556.

Guo, Y.J., Jin, F.G., Wang, P., Wang, M., and Zhu, J.M. (1983). Isolation of influenza C virus from pigs and experimental infection of pigs with influenza C virus. J. Gen. Virol. *64 (Pt 1)*, 177–182.

Gutsche, I., Desfosses, A., Effantin, G., Ling, W.L., Haupt, M., Ruigrok, R.W.H., Sachse, C., and Schoehn, G. (2015). Structural virology. Near-atomic cryo-EM structure of the helical measles virus nucleocapsid. Science *348*, 704–707.

Hagen, M., Chung, T.D., Butcher, J.A., and Krystal, M. (1994). Recombinant influenza virus polymerase: requirement of both 5' and 3' viral ends for endonuclease activity. J. Virol. *68*, 1509–1515.

Hale, B.G., Barclay, W.S., Randall, R.E., and Russell, R.J. (2008). Structure of an avian influenza A virus NS1 protein effector domain. Virology *378*, 1–5.

Hale, B.G., Kerry, P.S., Jackson, D., Precious, B.L., Gray, A., Killip, M.J., Randall, R.E., and Russell, R.J. (2010a). Structural insights into phosphoinositide 3-kinase activation by the influenza A virus NS1 protein. Proc. Natl. Acad. Sci. U. S. A. *107*, 1954–1959.

Hale, B.G., Albrecht, R.A., and García-Sastre, A. (2010b). Innate immune evasion strategies of influenza viruses. Future Microbiol. *5*, 23–41.

Haller, O., Staeheli, P., and Kochs, G. (2009). Protective role of interferon-induced Mx GTPases against influenza viruses. Rev. Sci. Tech. Int. Off. Epizoot. *28*, 219–231.

Harris, A., Forouhar, F., Qiu, S., Sha, B., and Luo, M. (2001). The crystal structure of the influenza matrix protein M1 at neutral pH: M1-M1 protein interfaces can rotate in the oligomeric structures of M1. Int. Congr. Ser. *1219*, 405–410.

Harrison, S.C. (2008). Viral membrane fusion. Nat. Struct. Mol. Biol. *15*, 690–698.

Hastie, K.M., Liu, T., Li, S., King, L.B., Ngo, N., Zandonatti, M.A., Woods, V.L., de la Torre, J.C., and Saphire, E.O. (2011a). Crystal structure of the Lassa virus nucleoprotein-RNA complex reveals a gating mechanism for RNA binding. Proc. Natl. Acad. Sci. U. S. A. *108*, 19365–19370.

Hastie, K.M., Kimberlin, C.R., Zandonatti, M.A., MacRae, I.J., and Saphire, E.O. (2011b). Structure of the Lassa virus nucleoprotein reveals a dsRNA-specific 3’ to 5’ exonuclease activity essential for immune suppression. Proc. Natl. Acad. Sci. U. S. A. *108*, 2396–2401.

Hause, B.M., Ducatez, M., Collin, E.A., Ran, Z., Liu, R., Sheng, Z., Armien, A., Kaplan, B., Chakravarty, S., Hoppe, A.D., et al. (2013). Isolation of a novel swine influenza virus from Oklahoma in 2011 which is distantly related to human influenza C viruses. PLoS Pathog. *9*, e1003176.

Hause, B.M., Collin, E.A., Liu, R., Huang, B., Sheng, Z., Lu, W., Wang, D., Nelson, E.A., and Li, F. (2014). Characterization of a Novel Influenza Virus in Cattle and Swine: Proposal for a New Genus in the Orthomyxoviridae Family. MBio *5*.

Hay, A.J., Lomniczi, B., Bellamy, A.R., and Skehel, J.J. (1977). Transcription of the influenza virus genome. Virology *83*, 337–355.

Hay, A.J., Skehel, J.J., and McCauley, J. (1982). Characterization of influenza virus RNA complete transcripts. Virology *116*, 517–522.

Hayden, F.G., and Hay, A.J. (1992). Emergence and transmission of influenza A viruses resistant to amantadine and rimantadine. Curr. Top. Microbiol. Immunol. *176*, 119–130.

Heggeness, M.H., Smith, P.R., Ulmanen, I., Krug, R.M., and Choppin, P.W. (1982). Studies on the helical nucleocapsid of influenza virus. Virology *118*, 466–470.

Helenius, A. (1992). Unpacking the incoming influenza virus. Cell *69*, 577–578.

Hendrickson, W.A., Horton, J.R., and LeMaster, D.M. (1990). Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. EMBO J. *9*, 1665–1672.

Henkel, M., Mitzner, D., Henklein, P., Meyer-Almes, F.-J., Moroni, A., Difrancesco, M.L., Henkes, L.M., Kreim, M., Kast, S.M., Schubert, U., et al. (2010). The proapoptotic influenza A virus protein PB1-F2 forms a nonselective ion channel. PloS One *5*, e11112.

Herfst, S., Imai, M., Kawaoka, Y., and Fouchier, R. a. M. (2014). Avian influenza virus transmission to mammals. Curr. Top. Microbiol. Immunol. *385*, 137–155.

Herrler, G., Dürkop, I., Becht, H., and Klenk, H.D. (1988). The glycoprotein of influenza C virus is the haemagglutinin, esterase and fusion factor. J. Gen. Virol. *69 (Pt 4)*, 839–846.

Heyduk, T., and Lee, J.C. (1990). Application of fluorescence energy transfer and polarization to monitor Escherichia coli cAMP receptor protein and lac promoter interaction. Proc. Natl. Acad. Sci. *87*, 1744–1748.

Hinshaw, J.E., Carragher, B.O., and Milligan, R.A. (1992). Architecture and design of the nuclear pore complex. Cell *69*, 1133–1141.

Hock, M., Kraus, I., Schoehn, G., Jamin, M., Andrei-Selmer, C., Garten, W., and Weissenhorn, W. (2010). RNA induced polymerization of the Borna disease virus nucleoprotein. Virology *397*, 64–72.

Hoehling AA (1961). The Great Epidemic. Boston: Little, Brown & Co.; 1961. . In The Great Epidemic, (Boston: Lillte Brown and Company), p.

Holsinger, L.J., and Lamb, R.A. (1991). Influenza virus M2 integral membrane protein is a homotetramer stabilized by formation of disulfide bonds. Virology *183*, 32–43.

Honda, A., Mizumoto, K., and Ishihama, A. (2002). Minimum molecular architectures for transcription and replication of the influenza virus. Proc. Natl. Acad. Sci. U. S. A. *99*, 13166–13171.

Houben, K., Marion, D., Tarbouriech, N., Ruigrok, R.W.H., and Blanchard, L. (2007). Interaction of the C-terminal domains of sendai virus N and P proteins: comparison of polymerase-nucleocapsid interactions within the paramyxovirus family. J. Virol. *81*, 6807–6816.

Hsu, M.T., Parvin, J.D., Gupta, S., Krystal, M., and Palese, P. (1987). Genomic RNAs of influenza viruses are held in a circular conformation in virions and in infected cells by a terminal panhandle. Proc. Natl. Acad. Sci. U. S. A. *84*, 8140–8144.

Huang, T.S., Palese, P., and Krystal, M. (1990). Determination of influenza virus proteins required for genome replication. J. Virol. *64*, 5669–5673.

Hutchinson, E.C., and Fodor, E. (2013). Transport of the influenza virus genome from nucleus to nucleus. Viruses *5*, 2424–2446.

Hutchinson, E.C., Curran, M.D., Read, E.K., Gog, J.R., and Digard, P. (2008). Mutational analysis of cis-acting RNA signals in segment 7 of influenza A virus. J. Virol. *82*, 11869–11879.

Hutchinson, E.C., Wise, H.M., Kudryavtseva, K., Curran, M.D., and Digard, P. (2009). Characterisation of influenza A viruses with mutations in segment 5 packaging signals. Vaccine *27*, 6270–6275.

Hutchinson, E.C., von Kirchbach, J.C., Gog, J.R., and Digard, P. (2010). Genome packaging in influenza A virus. J. Gen. Virol. *91*, 313–328.

Imamoto, N., Shimamoto, T., Takao, T., Tachibana, T., Kose, S., Matsubae, M., Sekimoto, T., Shimonishi, Y., and Yoneda, Y. (1995). In vivo evidence for involvement of a 58 kDa component of nuclear pore-targeting complex in nuclear protein import. EMBO J. *14*, 3617–3626.

Ivanov, I., Yabukarski, F., Ruigrok, R.W.H., and Jamin, M. (2011). Structural insights into the rhabdovirus transcription/replication complex. Virus Res. *162*, 126–137.

Jagger, B.W., Wise, H.M., Kash, J.C., Walters, K.-A., Wills, N.M., Xiao, Y.-L., Dunfee, R.L., Schwartzman, L.M., Ozinsky, A., Bell, G.L., et al. (2012). An overlapping protein-coding region in influenza A virus segment 3 modulates the host response. Science *337*, 199–204.

Jennings, P.A., Finch, J.T., Winter, G., and Robertson, J.S. (1983). Does the higher order structure of the influenza virus ribonucleoprotein guide sequence rearrangements in influenza viral RNA? Cell *34*, 619–627.

Jivraj, N., and Butler, A. (2013). The 1918-19 influenza pandemic revisited. J. R. Coll. Physicians Edinb. *43*, 347–352.

Johnson, N.P.A.S., and Mueller, J. (2002). Updating the accounts: global mortality of the 1918-1920 “Spanish” influenza pandemic. Bull. Hist. Med. *76*, 105–115.

de Jong, J.C., Claas, E.C., Osterhaus, A.D., Webster, R.G., and Lim, W.L. (1997). A pandemic warning? Nature *389*, 554.

Jorba, N., Coloma, R., and Ortín, J. (2009). Genetic trans-complementation establishes a new model for influenza virus RNA transcription and replication. PLoS Pathog. *5*, e1000462.

Junge, C. (2011). Morbidity: A personal response. Nature *480*, S14–S15.

Kalderon, D., Roberts, B.L., Richardson, W.D., and Smith, A.E. (1984). A short amino acid sequence able to specify nuclear location. Cell *39*, 499–509.

Kaplan, B.S., and Webby, R.J. (2013). The avian and mammalian host range of highly pathogenic avian H5N1 influenza. Virus Res. *178*, 3–11.

Katz, G., Benkarroum, Y., Wei, H., Rice, W.J., Bucher, D., Alimova, A., Katz, A., Klukowska, J., Herman, G.T., and Gottlieb, P. (2014). Morphology of influenza B/Lee/40 determined by cryo-electron microscopy. PloS One *9*, e88288.

Kawaguchi, A., and Nagata, K. (2007). De novo replication of the influenza virus RNA genome is regulated by DNA replicative helicase, MCM. EMBO J. *26*, 4566–4575.

Kawai, T., and Akira, S. (2011). Toll-like receptors and their crosstalk with other innate receptors in infection and immunity. Immunity *34*, 637–650.

Kerry, P.S., Ayllon, J., Taylor, M.A., Hass, C., Lewis, A., García-Sastre, A., Randall, R.E., Hale, B.G., and Russell, R.J. (2011). A transient homotypic interaction model for the influenza A virus NS1 protein effector domain. PloS One *6*, e17946.

Kilbourne, E.D., and Murphy, J.S. (1960). Genetic studies of influenza viruses. I. Viral morphology and growth capacity as exchangeable genetic traits. Rapid in ovo adaptation of early passage Asian strain isolates by combination with PR8. J. Exp. Med. *111*, 387–406.

Kimura, N., Nishida, M., Nagata, K., Ishihama, A., Oda, K., and Nakada, S. (1992). Transcription of a recombinant influenza virus RNA in cells that can express the influenza virus RNA polymerase and nucleoprotein genes. J. Gen. Virol. *73 (Pt 6)*, 1321–1328.

Kiso, M., Takahashi, K., Sakai-Tagawa, Y., Shinya, K., Sakabe, S., Le, Q.M., Ozawa, M., Furuta, Y., and Kawaoka, Y. (2010). T-705 (favipiravir) activity against lethal H5N1 influenza A viruses. Proc. Natl. Acad. Sci. U. S. A. *107*, 882–887.

Kistner, O., Müller, H., Becht, H., and Scholtissek, C. (1985). Phosphopeptide fingerprints of nucleoproteins of various influenza A virus strains grown in different host cells. J. Gen. Virol. *66 (Pt 3)*, 465–472.

Kistner, O., Müller, K., and Scholtissek, C. (1989). Differential phosphorylation of the nucleoprotein of influenza A viruses. J. Gen. Virol. *70 (Pt 9)*, 2421–2431.

Klumpp, K., Ruigrok, R.W., and Baudin, F. (1997). Roles of the influenza virus polymerase and nucleoprotein in forming a functional RNP structure. EMBO J. *16*, 1248–1257.

Koyama, K., Takahashi, M., Oitate, M., Nakai, N., Takakusa, H., Miura, S., and Okazaki, O. (2009). CS-8958, a prodrug of the novel neuraminidase inhibitor R-125489, demonstrates a favorable long-retention profile in the mouse respiratory tract. Antimicrob. Agents Chemother. *53*, 4845–4851.

Krug, R.M., Ueda, M., and Palese, P. (1975). Temperature-sensitive mutants of influenza WSN virus defective in virus-specific RNA synthesis. J. Virol. *16*, 790–796.

Kutay, U., Bischoff, F.R., Kostka, S., Kraft, R., and Görlich, D. (1997). Export of importin alpha from the nucleus is mediated by a specific nuclear transport factor. Cell *90*, 1061–1071.

Labaronne, A., Swale, C., Monod, A., Schoehn, G., Crépin, T., and Ruigrok, R.W.H. (2016). Binding of RNA by the Nucleoproteins of Influenza Viruses A and B. Viruses *8*.

Lamb, R.A., Zebedee, S.L., and Richardson, C.D. (1985). Influenza virus M2 protein is an integral membrane protein expressed on the infected-cell surface. Cell *40*, 627–633.

Lanford, R.E., and Butel, J.S. (1984). Construction and characterization of an SV40 mutant defective in nuclear transport of T antigen. Cell *37*, 801–813.

Lee, S.J., Matsuura, Y., Liu, S.M., and Stewart, M. (2005). Structural basis for nuclear import complex dissociation by RanGTP. Nature *435*, 693–696.

Lewis, P.A., and Shope, R.E. (1931). SWINE INFLUENZA : II. A HEMOPHILIC BACILLUS FROM THE RESPIRATORY TRACT OF INFECTED SWINE. J. Exp. Med. *54*, 361–371.

LiCata, V.J., and Wowor, A.J. (2008). Applications of fluorescence anisotropy to the study of protein-DNA interactions. Methods Cell Biol. *84*, 243–262.

Lin, Y.P., Xiong, X., Wharton, S.A., Martin, S.R., Coombs, P.J., Vachieri, S.G., Christodoulou, E., Walker, P.A., Liu, J., Skehel, J.J., et al. (2012). Evolution of the receptor binding properties of the influenza A(H3N2) hemagglutinin. Proc. Natl. Acad. Sci. U. S. A. *109*, 21474–21479.

Liu, J., Lynch, P.A., Chien, C.Y., Montelione, G.T., Krug, R.M., and Berman, H.M. (1997). Crystal structure of the unique RNA-binding domain of the influenza virus NS1 protein. Nat. Struct. Biol. *4*, 896–899.

Liu, M., Lam, M.K.-H., Zhang, Q., Elderfield, R., Barclay, W.S., and Shaw, P.-C. (2015). The Functional Study of the N-Terminal Region of Influenza B Virus Nucleoprotein. PloS One *10*, e0137802.

Longhi, S., Receveur-Bréchot, V., Karlin, D., Johansson, K., Darbon, H., Bhella, D., Yeo, R., Finet, S., and Canard, B. (2003). The C-terminal Domain of the Measles Virus Nucleoprotein Is Intrinsically Disordered and Folds upon Binding to the C-terminal Moiety of the Phosphoprotein. J. Biol. Chem. *278*, 18638–18648.

de la Luna, S., Martín, J., Portela, A., and Ortín, J. (1993). Influenza virus naked RNA can be expressed upon transfection into cells co-expressing the three subunits of the polymerase and the nucleoprotein from simian virus 40 recombinant viruses. J. Gen. Virol. *74 (Pt 3)*, 535–539.

de la Luna, S., Fortes, P., Beloso, A., and Ortín, J. (1995). Influenza virus NS1 protein enhances the rate of translation initiation of viral mRNAs. J. Virol. *69*, 2427–2433.

Luytjes, W., Krystal, M., Enami, M., Parvin, J.D., and Palese, P. (1989). Amplification, expression, and packaging of foreign gene by influenza virus. Cell *59*, 1107–1113.

Mahy, B.W.J. (1983). Mutants of Influenza Virus. In Genetics of Influenza Viruses, P.D.P. Palese, and P.D.D.W. Kingsbury, eds. (Springer Vienna), pp. 192–254.

Manuguerra, J.C., and Hannoun, C. (1992). Natural infection of dogs by influenza C virus. Res. Virol. *143*, 199–204.

Mänz, B., Brunotte, L., Reuther, P., and Schwemmle, M. (2012). Adaptive mutations in NEP compensate for defective H5N1 RNA replication in cultured human cells. Nat. Commun. *3*, 802.

Marklund, J.K., Ye, Q., Dong, J., Tao, Y.J., and Krug, R.M. (2012). Sequence in the influenza A virus nucleoprotein required for viral polymerase binding and RNA synthesis. J. Virol. *86*, 7292–7297.

Marsh, G.A., Rabadán, R., Levine, A.J., and Palese, P. (2008). Highly conserved regions of influenza a virus polymerase gene segments are critical for efficient viral RNA packaging. J. Virol. *82*, 2295–2304.

Martín-Benito, J., Area, E., Ortega, J., Llorca, O., Valpuesta, J.M., Carrascosa, J.L., and Ortín, J. (2001). Three-dimensional reconstruction of a recombinant influenza virus ribonucleoprotein particle. EMBO Rep. *2*, 313–317.

Matlin, K.S., Reggio, H., Helenius, A., and Simons, K. (1981). Infectious entry pathway of influenza virus in a canine kidney cell line. J. Cell Biol. *91*, 601–613.

Mavrakis, M., Méhouas, S., Réal, E., Iseni, F., Blondel, D., Tordo, N., and Ruigrok, R.W.H. (2006). Rabies virus chaperone: identification of the phosphoprotein peptide that keeps nucleoprotein soluble and free from non-specific RNA. Virology *349*, 422–429.

Mayer, D., Molawi, K., Martínez-Sobrido, L., Ghanem, A., Thomas, S., Baginsky, S., Grossmann, J., García-Sastre, A., and Schwemmle, M. (2007). Identification of cellular interaction partners of the influenza virus ribonucleoprotein complex and polymerase complex using proteomic-based approaches. J. Proteome Res. *6*, 672–682.

McCown, M.F., and Pekosz, A. (2006). Distinct domains of the influenza a virus M2 protein cytoplasmic tail mediate binding to the M1 protein and facilitate infectious virus production. J. Virol. *80*, 8178–8189.

McDevitt, J., Rudnick, S., First, M., and Spengler, J. (2010). Role of absolute humidity in the inactivation of influenza viruses on stainless steel surfaces at elevated temperatures. Appl. Environ. Microbiol. *76*, 3943–3947.

Mena, I., Jambrina, E., Albo, C., Perales, B., Ortín, J., Arrese, M., Vallejo, D., and Portela, A. (1999). Mutational analysis of influenza A virus nucleoprotein: identification of mutations that affect RNA replication. J. Virol. *73*, 1186–1194.

Moeller, A., Kirchdoerfer, R.N., Potter, C.S., Carragher, B., and Wilson, I.A. (2012). Organization of the influenza virus replication machinery. Science *338*, 1631–1634.

Mondal, A., Potts, G.K., Dawson, A.R., Coon, J.J., and Mehle, A. (2015). Phosphorylation at the homotypic interface regulates nucleoprotein oligomerization and assembly of the influenza virus replication machinery. PLoS Pathog. *11*, e1004826.

Monod, A., Swale, C., Tarus, B., Tissot, A., Delmas, B., Ruigrok, R.W., Crépin, T., and Slama-Schwok, A. (2015). Learning from structure-based drug design and new antivirals targeting the ribonucleoprotein complex for the treatment of influenza. Expert Opin. Drug Discov. *10*, 345–371.

Morens, D.M., Taubenberger, J.K., Folkers, G.K., and Fauci, A.S. (2010). Pandemic influenza's 500th anniversary. Clin. Infect. Dis. Off. Publ. Infect. Dis. Soc. Am. *51*, 1442–1444.

Moriuchi, H., Katsushima, N., Nishimura, H., Nakamura, K., and Numazaki, Y. (1991). Community-acquired influenza C virus infection in children. J. Pediatr. *118*, 235–238.

Moroianu, J., Blobel, G., and Radu, A. (1995). Previously identified protein of uncertain function is karyopherin alpha and together with karyopherin beta docks import substrate at nuclear pore complexes. Proc. Natl. Acad. Sci. U. S. A. *92*, 2008–2011.

Munier, S., Moisy, D., Marc, D., and Naffakh, N. (2010). [Interspecies transmission, adaptation to humans and pathogenicity of animal influenza viruses]. Pathol. Biol. (Paris) *58*, e59-68.

Muraki, Y., Washioka, H., Sugawara, K., Matsuzaki, Y., Takashita, E., and Hongo, S. (2004). Identification of an amino acid residue on influenza C virus M1 protein responsible for formation of the cord-like structures of the virus. J. Gen. Virol. *85*, 1885–1893.

Muramoto, Y., Takada, A., Fujii, K., Noda, T., Iwatsuki-Horimoto, K., Watanabe, S., Horimoto, T., Kida, H., and Kawaoka, Y. (2006). Hierarchy among viral RNA (vRNA) segments in their role in vRNA incorporation into influenza A virions. J. Virol. *80*, 2318–2325.

Murti, K.G., Webster, R.G., and Jones, I.M. (1988). Localization of RNA polymerases on influenza viral ribonucleoproteins by immunogold labeling. Virology *164*, 562–566.

Nachury, M.V., and Weis, K. (1999). The direction of transport through the nuclear pore can be inverted. Proc. Natl. Acad. Sci. U. S. A. *96*, 9622–9627.

Naffakh, N., Tomoiu, A., Rameix-Welti, M.-A., and van der Werf, S. (2008). Host restriction of avian influenza viruses at the level of the ribonucleoproteins. Annu. Rev. Microbiol. *62*, 403–424.

Nakada, R., Hirano, H., and Matsuura, Y. (2015). Structure of importin-α bound to a non-classical nuclear localization signal of the influenza A virus nucleoprotein. Sci. Rep. *5*, 15055.

Nakada, S., Creager, R.S., Krystal, M., Aaronson, R.P., and Palese, P. (1984). Influenza C virus hemagglutinin: comparison with influenza A and B virus hemagglutinins. J. Virol. *50*, 118–124.

Nakada, S., Graves, P.N., Desselberger, U., Creager, R.S., Krystal, M., and Palese, P. (1985). Influenza C virus RNA 7 codes for a nonstructural protein. J. Virol. *56*, 221–226.

Nayak, D.P., and Hui, E.K.W. (2004). The role of lipid microdomains in virus biology. Subcell. Biochem. *37*, 443–491.

Nayak, D.P., Balogun, R.A., Yamada, H., Zhou, Z.H., and Barman, S. (2009). Influenza virus morphogenesis and budding. Virus Res. *143*, 147–161.

Nemeroff, M.E., Qian, X.Y., and Krug, R.M. (1995). The influenza virus NS1 protein forms multimers in vitro and in vivo. Virology *212*, 422–428.

Nemeroff, M.E., Barabino, S.M., Li, Y., Keller, W., and Krug, R.M. (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*, 991–1000.

Neumann, G., Castrucci, M.R., and Kawaoka, Y. (1997). Nuclear import and export of influenza virus nucleoprotein. J. Virol. *71*, 9690–9700.

Neumann, G., Hughes, M.T., and Kawaoka, Y. (2000). Influenza A virus NS2 protein mediates vRNP nuclear export through NES-independent interaction with hCRM1. EMBO J. *19*, 6751–6758.

Ng, A.K.-L., Zhang, H., Tan, K., Li, Z., Liu, J., Chan, P.K.-S., Li, S.-M., Chan, W.-Y., Au, S.W.-N., Joachimiak, A., et al. (2008). Structure of the influenza virus A H5N1 nucleoprotein: implications for RNA binding, oligomerization, and vaccine design. FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol. *22*, 3638–3647.

Ng, A.K.-L., Lam, M.K.-H., Zhang, H., Liu, J., Au, S.W.-N., Chan, P.K.-S., Wang, J., and Shaw, P.-C. (2012). Structural basis for RNA binding and homo-oligomer formation by influenza B virus nucleoprotein. J. Virol. *86*, 6758–6767.

Nishimura, H., Hara, M., Sugawara, K., Kitame, F., Takiguchi, K., Umetsu, Y., Tonosaki, A., and Nakamura, K. (1990). Characterization of the cord-like structures emerging from the surface of influenza C virus-infected cells. Virology *179*, 179–188.

Niu, F., Shaw, N., Wang, Y.E., Jiao, L., Ding, W., Li, X., Zhu, P., Upur, H., Ouyang, S., Cheng, G., et al. (2013). Structure of the Leanyer orthobunyavirus nucleoprotein-RNA complex reveals unique architecture for RNA encapsidation. Proc. Natl. Acad. Sci. U. S. A. *110*, 9054–9059.

Noda, T., Sagara, H., Yen, A., Takada, A., Kida, H., Cheng, R.H., and Kawaoka, Y. (2006). Architecture of ribonucleoprotein complexes in influenza A virus particles. Nature *439*, 490–492.

Noda, T., Sugita, Y., Aoyama, K., Hirase, A., Kawakami, E., Miyazawa, A., Sagara, H., and Kawaoka, Y. (2012). Three-dimensional analysis of ribonucleoprotein complexes in influenza A virus. Nat. Commun. *3*, 639.

Normile, D. (2006). Avian influenza. Human transmission but no pandemic in Indonesia. Science *312*, 1855.

Noti, J.D., Blachere, F.M., McMillen, C.M., Lindsley, W.G., Kashon, M.L., Slaughter, D.R., and Beezhold, D.H. (2013). High humidity leads to loss of infectious influenza virus from simulated coughs. PloS One *8*, e57485.

O'Callaghan, R.J., Gohd, R.S., and Labat, D.D. (1980). Human antibody to influenza C virus: its age-related distribution and distinction from receptor analogs. Infect. Immun. *30*, 500–505.

Odagiri, T., Hong, J., and Ohara, Y. (1999). The BM2 protein of influenza B virus is synthesized in the late phase of infection and incorporated into virions as a subviral component. J. Gen. Virol. *80 (Pt 10)*, 2573–2581.

Oestereich, L., Lüdtke, A., Wurr, S., Rieger, T., Muñoz-Fontela, C., and Günther, S. (2014). Successful treatment of advanced Ebola virus infection with T-705 (favipiravir) in a small animal model. Antiviral Res. *105*, 17–21.

Ohishi, K., Ninomiya, A., Kida, H., Park, C.-H., Maruyama, T., Arai, T., Katsumata, E., Tobayama, T., Boltunov, A.N., Khuraskin, L.S., et al. (2002). Serological evidence of transmission of human influenza A and B viruses to Caspian seals (Phoca caspica). Microbiol. Immunol. *46*, 639–644.

Ohwada, K., Kitame, F., Sugawara, K., Nishimura, H., Homma, M., and Nakamura, K. (1987). Distribution of the antibody to influenza C virus in dogs and pigs in Yamagata Prefecture, Japan. Microbiol. Immunol. *31*, 1173–1180.

Oliva, A., Llabrés, M., and Fariña, J.B. (2004). Applications of multi-angle laser light-scattering detection in the analysis of peptides and proteins. Curr. Drug Discov. Technol. *1*, 229–242.

Olson, A.C., Rosenblum, E., and Kuchta, R.D. (2010). Regulation of influenza RNA polymerase activity and the switch between replication and transcription by the concentrations of the vRNA 5’ end, the cap source, and the polymerase. Biochemistry (Mosc.) *49*, 10208–10215.

OMS, Organisation mondiale de la santé, 2017. www.who.

O’Neill, R.E., Talon, J., and Palese, P. (1998). The influenza virus NEP (NS2 protein) mediates the nuclear export of viral ribonucleoproteins. EMBO J. *17*, 288–296.

Ortega, J., Martín-Benito, J., Zürcher, T., Valpuesta, J.M., Carrascosa, J.L., and Ortín, J. (2000). Ultrastructural and functional analyses of recombinant influenza virus ribonucleoproteins suggest dimerization of nucleoprotein during virus amplification. J. Virol. *74*, 156–163.

Osborne, J.C., and Elliott, R.M. (2000). RNA binding properties of bunyamwera virus nucleocapsid protein and selective binding to an element in the 5’ terminus of the negative-sense S segment. J. Virol. *74*, 9946–9952.

Osterhaus, A.D., Rimmelzwaan, G.F., Martina, B.E., Bestebroer, T.M., and Fouchier, R.A. (2000). Influenza B virus in seals. Science *288*, 1051–1053.

Otsuka, S., Iwasaka, S., Yoneda, Y., Takeyasu, K., and Yoshimura, S.H. (2008). Individual binding pockets of importin-beta for FG-nucleoporins have different binding properties and different sensitivities to RanGTP. Proc. Natl. Acad. Sci. U. S. A. *105*, 16101–16106.

Ozawa, M., Fujii, K., Muramoto, Y., Yamada, S., Yamayoshi, S., Takada, A., Goto, H., Horimoto, T., and Kawaoka, Y. (2007). Contributions of Two Nuclear Localization Signals of Influenza A Virus Nucleoprotein to Viral Replication. J. Virol. *81*, 30–41.

Palese, P., and Compans, R.W. (1976). Inhibition of influenza virus replication in tissue culture by 2-deoxy-2,3-dehydro-N-trifluoroacetylneuraminic acid (FANA): mechanism of action. J. Gen. Virol. *33*, 159–163.

Palese, P., and Schulman, J.L. (1976). Mapping of the influenza virus genome: identification of the hemagglutinin and the neuraminidase genes. Proc. Natl. Acad. Sci. U. S. A. *73*, 2142–2146.

Palese, P., Tobita, K., Ueda, M., and Compans, R.W. (1974). Characterization of temperature sensitive influenza virus mutants defective in neuraminidase. Virology *61*, 397–410.

Panté, N., and Kann, M. (2002). Nuclear Pore Complex Is Able to Transport Macromolecules with Diameters of ~39 nm. Mol. Biol. Cell *13*, 425–434.

Park, C.-J., Bae, S.-H., Lee, M.-K., Varani, G., and Choi, B.-S. (2003). Solution structure of the influenza A virus cRNA promoter: implications for differential recognition of viral promoter structures by RNA-dependent RNA polymerase. Nucleic Acids Res. *31*, 2824–2832.

Paterson, D., and Fodor, E. (2012). Emerging roles for the influenza A virus nuclear export protein (NEP). PLoS Pathog. *8*, e1003019.

Patterson, K.D. (1985). Pandemic and epidemic influenza, 1830-1848. Soc. Sci. Med. 1982 *21*, 571–580.

Patterson, K.D., and Pyle, G.F. (1991). The geography and mortality of the 1918 influenza pandemic. Bull. Hist. Med. *65*, 4–21.

Peiris, J.S.M., Yu, W.C., Leung, C.W., Cheung, C.Y., Ng, W.F., Nicholls, J.M., Ng, T.K., Chan, K.H., Lai, S.T., Lim, W.L., et al. (2004). Re-emergence of fatal human influenza A subtype H5N1 disease. Lancet Lond. Engl. *363*, 617–619.

Pekosz, A., and Lamb, R.A. (1997). The CM2 protein of influenza C virus is an oligomeric integral membrane glycoprotein structurally analogous to influenza A virus M2 and influenza B virus NB proteins. Virology *237*, 439–451.

Pflug, A., Guilligay, D., Reich, S., and Cusack, S. (2014). Structure of influenza A polymerase bound to the viral RNA promoter. Nature *516*, 355–360.

Pielak, R.M., and Chou, J.J. (2011). Influenza M2 proton channels. Biochim. Biophys. Acta BBA - Biomembr. *1808*, 522–529.

Plotch, S.J., Bouloy, M., and Krug, R.M. (1979). Transfer of 5'-terminal cap of globin mRNA to influenza viral complementary RNA during transcription in vitro. Proc. Natl. Acad. Sci. U. S. A. *76*, 1618–1622.

Plotch, S.J., Bouloy, M., Ulmanen, I., and Krug, R.M. (1981). A unique cap(m7GpppXm)-dependent influenza virion endonuclease cleaves capped RNAs to generate the primers that initiate viral RNA transcription. Cell *23*, 847–858.

Plotkin, J.B., and Dushoff, J. (2003). Codon bias and frequency-dependent selection on the hemagglutinin epitopes of influenza A virus. Proc. Natl. Acad. Sci. U. S. A. *100*, 7152–7157.

Pons, M.W., Schulze, I.T., Hirst, G.K., and Hauser, R. (1969). Isolation and characterization of the ribonucleoprotein of influenza virus. Virology *39*, 250–259.

Poole, E., Elton, D., Medcalf, L., and Digard, P. (2004). Functional domains of the influenza A virus PB2 protein: identification of NP- and PB1-binding sites. Virology *321*, 120–133.

Poon, L.L., Pritlove, D.C., Fodor, E., and Brownlee, G.G. (1999). Direct evidence that the poly(A) tail of influenza A virus mRNA is synthesized by reiterative copying of a U track in the virion RNA template. J. Virol. *73*, 3473–3476.

Prokudina-Kantorovich, E.N., and Semenova, N.P. (1996). Intracellular oligomerization of influenza virus nucleoprotein. Virology *223*, 51–56.

Pumroy, R.A., and Cingolani, G. (2015). Diversification of importin-α isoforms in cellular trafficking and disease states. Biochem. J. *466*, 13–28.

Qi, X., Lan, S., Wang, W., Schelde, L.M., Dong, H., Wallat, G.D., Ly, H., Liang, Y., and Dong, C. (2010). Cap binding and immune evasion revealed by Lassa nucleoprotein structure. Nature *468*, 779–783.

Quast, M., Sreenivasan, C., Sexton, G., Nedland, H., Singrey, A., Fawcett, L., Miller, G., Lauer, D., Voss, S., Pollock, S., et al. (2015). Serological evidence for the presence of influenza D virus in small ruminants. Vet. Microbiol. *180*, 281–285.

Racaniello, V.R., and Palese, P. (1979). Influenza B virus genome: assignment of viral polypeptides to RNA segments. J. Virol. *29*, 361–373.

Randall, R.E., and Goodbourn, S. (2008). Interferons and viruses: an interplay between induction, signalling, antiviral responses and virus countermeasures. J. Gen. Virol. *89*, 1–47.

Raymond, D.D., Piper, M.E., Gerrard, S.R., and Smith, J.L. (2010). Structure of the Rift Valley fever virus nucleocapsid protein reveals another architecture for RNA encapsidation. Proc. Natl. Acad. Sci. U. S. A. *107*, 11769–11774.

Raymond, D.D., Piper, M.E., Gerrard, S.R., Skiniotis, G., and Smith, J.L. (2012). Phleboviruses encapsidate their genomes by sequestering RNA bases. Proc. Natl. Acad. Sci. U. S. A. *109*, 19208–19213.

Regan, J.F., Liang, Y., and Parslow, T.G. (2006). Defective assembly of influenza A virus due to a mutation in the polymerase subunit PA. J. Virol. *80*, 252–261.

Reguera, J., Malet, H., Weber, F., and Cusack, S. (2013). Structural basis for encapsidation of genomic RNA by La Crosse Orthobunyavirus nucleoprotein. Proc. Natl. Acad. Sci. U. S. A. *110*, 7246–7251..

Reguera, J., Cusack, S., and Kolakofsky, D. (2014). Segmented negative strand RNA virus nucleoprotein structure. Curr. Opin. Virol. *5*, 7–15.

Rehwinkel, J., Tan, C.P., Goubau, D., Schulz, O., Pichlmair, A., Bier, K., Robb, N., Vreede, F., Barclay, W., Fodor, E., et al. (2010). RIG-I detects viral genomic RNA during negative-strand RNA virus infection. Cell *140*, 397–408.

Reich, S., Guilligay, D., Pflug, A., Malet, H., Berger, I., Crépin, T., Hart, D., Lunardi, T., Nanao, M., Ruigrok, R.W.H., et al. (2014). Structural insight into cap-snatching and RNA synthesis by influenza polymerase. Nature *516*, 361–366.

Reperant, L.A., Rimmelzwaan, G.F., and Kuiken, T. (2009). Avian influenza viruses in mammals. Rev. Sci. Tech. Int. Off. Epizoot. *28*, 137–159.

Resa-Infante, P., Jorba, N., Coloma, R., and Ortín, J. (2011). The influenza RNA synthesis machine. RNA Biol. *8*, 207–215.

Rexach, M., and Blobel, G. (1995). Protein import into nuclei: association and dissociation reactions involving transport substrate, transport factors, and nucleoporins. Cell *83*, 683–692.

Richardson, J.C., and Akkina, R.K. (1991). NS2 protein of influenza virus is found in purified virus and phosphorylated in infected cells. Arch. Virol. *116*, 69–80.

Ritchey, M.B., Palese, P., and Schulman, J.L. (1976). Mapping of the influenza virus genome. III. Identification of genes coding for nucleoprotein, membrane protein, and nonstructural protein. J. Virol. *20*, 307–313.

Robb, N.C., Smith, M., Vreede, F.T., and Fodor, E. (2009). NS2/NEP protein regulates transcription and replication of the influenza virus RNA genome. J. Gen. Virol. *90*, 1398–1407.

Rodriguez, A., Pérez-González, A., and Nieto, A. (2007). Influenza virus infection causes specific degradation of the largest subunit of cellular RNA polymerase II. J. Virol. *81*, 5315–5324.

Rossman, J.S., and Lamb, R.A. (2011). Influenza Virus Assembly and Budding. Virology *411*, 229–236.

Rossman, J.S., Jing, X., Leser, G.P., Balannik, V., Pinto, L.H., and Lamb, R.A. (2010). Influenza virus m2 ion channel protein is necessary for filamentous virion formation. J. Virol. *84*, 5078–5088.

Rossmann, M.G., and Blow, D.M. (1962). The detection of sub-units within the crystallographic asymmetric unit. Acta Crystallogr. *15*, 24–31.

Rudolph, M.G., Kraus, I., Dickmanns, A., Eickmann, M., Garten, W., and Ficner, R. (2003). Crystal structure of the borna disease virus nucleoprotein. Struct. Lond. Engl. 1993 *11*, 1219–1226.

Ruigrok, R.W., and Baudin, F. (1995). Structure of influenza virus ribonucleoprotein particles. II. Purified RNA-free influenza virus ribonucleoprotein forms structures that are indistinguishable from the intact influenza virus ribonucleoprotein particles. J. Gen. Virol. *76 (Pt 4)*, 1009–1014.

Ruigrok, R.W., Calder, L.J., and Wharton, S.A. (1989). Electron microscopy of the influenza virus submembranal structure. Virology *173*, 311–316.

Ruigrok, R.W., Barge, A., Durrer, P., Brunner, J., Ma, K., and Whittaker, G.R. (2000). Membrane interaction of influenza virus M1 protein. Virology *267*, 289–298.

Ruigrok, R.W., Crépin, T., and Kolakofsky, D. (2011). Nucleoproteins and nucleocapsids of negative-strand RNA viruses. Curr. Opin. Microbiol. *14*, 504–510.

Ruigrok, R.W.H., Crépin, T., Hart, D.J., and Cusack, S. (2010). Towards an atomic resolution understanding of the influenza virus replication machinery. Curr. Opin. Struct. Biol. *20*, 104–113.

Russell, C.A., Jones, T.C., Barr, I.G., Cox, N.J., Garten, R.J., Gregory, V., Gust, I.D., Hampson, A.W., Hay, A.J., Hurt, A.C., et al. (2008). The Global Circulation of Seasonal Influenza A (H3N2) Viruses. Science *320*, 340–346.

Samson, M., Pizzorno, A., Abed, Y., and Boivin, G. (2013). Influenza virus resistance to neuraminidase inhibitors. Antiviral Res. *98*, 174–185.

Santé publique France (2017). https://invs.santepubliquefrance.fr.

Scheiffele, P., Rietveld, A., Wilk, T., and Simons, K. (1999). Influenza viruses select ordered lipid domains during budding from the plasma membrane. J. Biol. Chem. *274*, 2038–2044.

Schmitt, A.P., and Lamb, R.A. (2005). Influenza virus assembly and budding at the viral budozone. Adv. Virus Res. *64*, 383–416.

Schnell, J.R., and Chou, J.J. (2008). Structure and mechanism of the M2 proton channel of influenza A virus. Nature *451*, 591–595.

Schoehn, G., Mavrakis, M., Albertini, A., Wade, R., Hoenger, A., and Ruigrok, R.W.H. (2004). The 12 A structure of trypsin-treated measles virus N-RNA. J. Mol. Biol. *339*, 301–312.

Schrauwen, E.J.A., de Graaf, M., Herfst, S., Rimmelzwaan, G.F., Osterhaus, A.D.M.E., and Fouchier, R. a. M. (2014). Determinants of virulence of influenza A virus. Eur. J. Clin. Microbiol. Infect. Dis. Off. Publ. Eur. Soc. Clin. Microbiol. *33*, 479–490.

Schulze, I.T. (1972). The structure of influenza virus. II. A model based on the morphology and composition of subviral particles. Virology *47*, 181–196.

Schwartz, T.U. (2005). Modularity within the architecture of the nuclear pore complex. Curr. Opin. Struct. Biol. *15*, 221–226.

Selman, M., Dankar, S.K., Forbes, N.E., Jia, J.-J., and Brown, E.G. (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*, e42.

Sha, B., and Luo, M. (1997). Structure of a bifunctional membrane-RNA binding protein, influenza virus matrix protein M1. Nat. Struct. Biol. *4*, 239–244.

Shapiro, G.I., Gurney, T., and Krug, R.M. (1987). Influenza virus gene expression: control mechanisms at early and late times of infection and nuclear-cytoplasmic transport of virus-specific RNAs. J. Virol. *61*, 764–773.

Sherry, L., Smith, M., Davidson, S., and Jackson, D. (2014). The N Terminus of the Influenza B Virus Nucleoprotein Is Essential for Virus Viability, Nuclear Localization, and Optimal Transcription and Replication of the Viral Genome. J. Virol. *88*, 12326–12338.

Shope, R.E. (1931). SWINE INFLUENZA : I. EXPERIMENTAL TRANSMISSION AND PATHOLOGY. J. Exp. Med. *54*, 349–359.

Shortridge, K.F. (1999). Poultry and the influenza H5N1 outbreak in Hong Kong, 1997: abridged chronology and virus isolation. Vaccine *17 Suppl 1*, S26-29.

Shortridge, K.F., Zhou, N.N., Guan, Y., Gao, P., Ito, T., Kawaoka, Y., Kodihalli, S., Krauss, S., Markwell, D., Murti, K.G., et al. (1998). Characterization of avian H5N1 influenza viruses from poultry in Hong Kong. Virology *252*, 331–342.

Simonsen, L., Clarke, M.J., Schonberger, L.B., Arden, N.H., Cox, N.J., and Fukuda, K. (1998). Pandemic versus epidemic influenza mortality: a pattern of changing age distribution. J. Infect. Dis. *178*, 53–60.

Sims, L.D., Domenech, J., Benigno, C., Kahn, S., Kamata, A., Lubroth, J., Martin, V., and Roeder, P. (2005). Origin and evolution of highly pathogenic H5N1 avian influenza in Asia. Vet. Rec. *157*, 159–164.

Skorko, R., Summers, D.F., and Galarza, J.M. (1991). Influenza A virus in vitro transcription: roles of NS1 and NP proteins in regulating RNA synthesis. Virology *180*, 668–677.

Smith, D.J., Lapedes, A.S., de Jong, J.C., Bestebroer, T.M., Rimmelzwaan, G.F., Osterhaus, A.D.M.E., and Fouchier, R.A.M. (2004). Mapping the antigenic and genetic evolution of influenza virus. Science *305*, 371–376.

Song, L. (2014). It is Unlikely That Influenza Viruses Will Cause a Pandemic Again Like What Happened in 1918 and 1919. Front. Public Health *2*, 39.

Song, H., Qi, J., Khedri, Z., Diaz, S., Yu, H., Chen, X., Varki, A., Shi, Y., and Gao, G.F. (2016). An Open Receptor-Binding Cavity of Hemagglutinin-Esterase-Fusion Glycoprotein from Newly-Identified Influenza D Virus: Basis for Its Broad Cell Tropism. PLoS Pathog. *12*, e1005411.

Stoffler, D., Feja, B., Fahrenkrog, B., Walz, J., Typke, D., and Aebi, U. (2003). Cryo-electron tomography provides novel insights into nuclear pore architecture: implications for nucleocytoplasmic transport. J. Mol. Biol. *328*, 119–130.

Stray, S.J., and Air, G.M. (2001). Apoptosis by influenza viruses correlates with efficiency of viral mRNA synthesis. Virus Res. *77*, 3–17.

Subbarao, K., Klimov, A., Katz, J., Regnery, H., Lim, W., Hall, H., Perdue, M., Swayne, D., Bender, C., Huang, J., et al. (1998). Characterization of an avian influenza A (H5N1) virus isolated from a child with a fatal respiratory illness. Science *279*, 393–396.

Sugrue, R.J., and Hay, A.J. (1991). Structural characteristics of the M2 protein of influenza A viruses: evidence that it forms a tetrameric channel. Virology *180*, 617–624.

Sun, C., Yang, W., Tu, L.-C., and Musser, S.M. (2008). Single-molecule measurements of importin alpha/cargo complex dissociation at the nuclear pore. Proc. Natl. Acad. Sci. U. S. A. *105*, 8613–8618.

Tarus, B., Bakowiez, O., Chenavas, S., Duchemin, L., Estrozi, L.F., Bourdieu, C., Lejal, N., Bernard, J., Moudjou, M., Chevalier, C., et al. (2012). Oligomerization paths of the nucleoprotein of influenza A virus. Biochimie *94*, 776–785.

Taubenberger, J.K., and Morens, D.M. (2006). 1918 Influenza: the mother of all pandemics. Emerg. Infect. Dis. *12*, 15–22.

Tauber, S., Ligertwood, Y., Quigg-Nicol, M., Dutia, B.M., and Elliott, R.M. (2012). Behaviour of influenza A viruses differentially expressing segment 2 gene products in vitro and in vivo. J. Gen. Virol. *93*, 840–849.

Tawar, R.G., Duquerroy, S., Vonrhein, C., Varela, P.F., Damier-Piolle, L., Castagné, N., MacLellan, K., Bedouelle, H., Bricogne, G., Bhella, D., et al. (2009a). Crystal structure of a nucleocapsid-like nucleoprotein-RNA complex of respiratory syncytial virus. Science *326*, 1279–1283.

Tawar, R.G., Duquerroy, S., Vonrhein, C., Varela, P.F., Damier-Piolle, L., Castagné, N., MacLellan, K., Bedouelle, H., Bricogne, G., Bhella, D., et al. (2009b). Crystal structure of a nucleocapsid-like nucleoprotein-RNA complex of respiratory syncytial virus. Science *326*, 1279–1283.

Taylor, R.M. (1949). Studies on survival of influenza virus between epidemics and antigenic variants of the virus. Am. J. Public Health Nations Health *39*, 171–178.

Teissier, E., Zandomeneghi, G., Loquet, A., Lavillette, D., Lavergne, J.-P., Montserret, R., Cosset, F.-L., Böckmann, A., Meier, B.H., Penin, F., et al. (2011). Mechanism of inhibition of enveloped virus membrane fusion by the antiviral drug arbidol. PloS One *6*, e15874.

Terry, L.J., and Wente, S.R. (2007). Nuclear mRNA export requires specific FG nucleoporins for translocation through the nuclear pore complex. J. Cell Biol. *178*, 1121–1132.

Thierry, F., and Danos, O. (1982). Use of specific single stranded DNA probes cloned in M13 to study the RNA synthesis of four temperature-sensitive mutants of HK/68 influenza virus. Nucleic Acids Res. *10*, 2925–2938.

Tiley, L.S., Hagen, M., Matthews, J.T., and Krystal, M. (1994). Sequence-specific binding of the influenza virus RNA polymerase to sequences located at the 5’ ends of the viral RNAs. J. Virol. *68*, 5108–5116.

Tong, S., Zhu, X., Li, Y., Shi, M., Zhang, J., Bourgeois, M., Yang, H., Chen, X., Recuenco, S., Gomez, J., et al. (2013). New world bats harbor diverse influenza A viruses. PLoS Pathog. *9*, e1003657.

Turan, K., Mibayashi, M., Sugiyama, K., Saito, S., Numajiri, A., and Nagata, K. (2004). Nuclear MxA proteins form a complex with influenza virus NP and inhibit the transcription of the engineered influenza virus genome. Nucleic Acids Res. *32*, 643–652.

Ulmanen, I., Broni, B.A., and Krug, R.M. (1981). Role of two of the influenza virus core P proteins in recognizing cap 1 structures (m7GpppNm) on RNAs and in initiating viral RNA transcription. Proc. Natl. Acad. Sci. U. S. A. *78*, 7355–7359.

Ungchusak, K., Auewarakul, P., Dowell, S.F., Kitphati, R., Auwanit, W., Puthavathana, P., Uiprasertkul, M., Boonnak, K., Pittayawonganon, C., Cox, N.J., et al. (2005). Probable person-to-person transmission of avian influenza A (H5N1). N. Engl. J. Med. *352*, 333–340.

Vahlenkamp, T.W., and Harder, T.C. (2006). Influenza virus infections in mammals. Berl. Münch. Tierärztl. Wochenschr. *119*, 123–131.

Viboud, C., Alonso, W.J., and Simonsen, L. (2006). Influenza in tropical regions. PLoS Med. *3*, e89.

Viswanathan, K., Chandrasekaran, A., Srinivasan, A., Raman, R., Sasisekharan, V., and Sasisekharan, R. (2010). Glycans as receptors for influenza pathogenesis. Glycoconj. J. *27*, 561–570.

Vreede, F.T., and Brownlee, G.G. (2007). Influenza virion-derived viral ribonucleoproteins synthesize both mRNA and cRNA in vitro. J. Virol. *81*, 2196–2204.

Wan, Z., Ye, J., Xu, L., Shao, H., Jin, W., Qian, K., Wan, H., and Qin, A. (2014). Antigenic mapping of the hemagglutinin of an H9N2 avian influenza virus reveals novel critical amino acid positions in antigenic sites. J. Virol. *88*, 3898–3901.

Wang, P., Palese, P., and O'Neill, R.E. (1997). The NPI-1/NPI-3 (karyopherin alpha) binding site on the influenza a virus nucleoprotein NP is a nonconventional nuclear localization signal. J. Virol. *71*, 1850–1856.

Wanitchang, A., Narkpuk, J., Jaru-ampornpan, P., Jengarn, J., and Jongkaewwattana, A. (2012). Inhibition of influenza A virus replication by influenza B virus nucleoprotein: an insight into interference between influenza A and B viruses. Virology *432*, 194–203.

Wanitchang, A., Narkpuk, J., and Jongkaewwattana, A. (2013). Nuclear import of influenza B virus nucleoprotein: involvement of an N-terminal nuclear localization signal and a cleavage-protection motif. Virology *443*, 59–68.

Weber, F., Kochs, G., Gruber, S., and Haller, O. (1998). A Classical Bipartite Nuclear Localization Signal on Thogoto and Influenza A Virus Nucleoproteins. Virology *250*, 9–18.

Webster, R.G., Bean, W.J., Gorman, O.T., Chambers, T.M., and Kawaoka, Y. (1992). Evolution and ecology of influenza A viruses. Microbiol. Rev. *56*, 152–179.

Welsch, S., Kolesnikova, L., Krähling, V., Riches, J.D., Becker, S., and Briggs, J.A.G. (2010). Electron tomography reveals the steps in filovirus budding. PLoS Pathog. *6*, e1000875.

Weycker, D., Edelsberg, J., Elizabeth Halloran, M., Longini Jr., I.M., Nizam, A., Ciuryla, V., and Oster, G. (2005). Population-wide benefits of routine vaccination of children against influenza. Vaccine *23*, 1284–1293.

White, S.K., Ma, W., McDaniel, C.J., Gray, G.C., and Lednicky, J.A. (2016). Serologic evidence of exposure to influenza D virus among persons with occupational contact with cattle. J. Clin. Virol. *81*, 31–33.

Wise, H.M., Foeglein, A., Sun, J., Dalton, R.M., Patel, S., Howard, W., Anderson, E.C., Barclay, W.S., and Digard, P. (2009). A complicated message: Identification of a novel PB1-related protein translated from influenza A virus segment 2 mRNA. J. Virol. *83*, 8021–8031.

Wolstenholme, A.J., Barrett, T., Nichol, S.T., and Mahy, B.W. (1980). Influenza virus-specific RNA and protein syntheses in cells infected with temperature-sensitive mutants defective in the genome segment encoding nonstructural proteins. J. Virol. *35*, 1–7.

Wright, P.F, Neumann, G, and Kawaoka, Y (2013). Chapter 41, Orthomyxoviruses D.M. Knipe, and P. Howley, eds. (Lippincott Williams & Wilkins). In Fields Virology 6th Edition, (Knipe, and P.Howley, eds. (Lippincott Williams & Wilkins)), pp. 1186–1243.

Wu, W.W.H., Sun, Y.-H.B., and Panté, N. (2007). Nuclear import of influenza A viral ribonucleoprotein complexes is mediated by two nuclear localization sequences on viral nucleoprotein. Virol. J. *4*, 49.

Wyatt, P.J. (1991). Combined differential light scattering with various liquid chromatography separation techniques. Biochem. Soc. Trans. *19*, 485.

Wyatt, P.J. (1993). Light scattering and the absolute characterization of macromolecules. Anal. Chim. Acta *272*, 1–40.

Yamanaka, K., Ishihama, A., and Nagata, K. (1990). Reconstitution of influenza virus RNA-nucleoprotein complexes structurally resembling native viral ribonucleoprotein cores. J. Biol. Chem. *265*, 11151–11155.

Yamashita, M., Krystal, M., and Palese, P. (1989). Comparison of the three large polymerase proteins of influenza A, B, and C viruses. Virology *171*, 458–466.

Yang, Q., Rout, M.P., and Akey, C.W. (1998). Three-dimensional architecture of the isolated yeast nuclear pore complex: functional and evolutionary implications. Mol. Cell *1*, 223–234.

Yasuda, J., Nakada, S., Kato, A., Toyoda, T., and Ishihama, A. (1993). Molecular Assembly of Influenza Virus: Association of the NS2 Protein with Virion Matrix. Virology *196*, 249–255.

Ye, Q., Krug, R.M., and Tao, Y.J. (2006). The mechanism by which influenza A virus nucleoprotein forms oligomers and binds RNA. Nature *444*, 1078–1082.

Ye, Q., Guu, T.S.Y., Mata, D.A., Kuo, R.-L., Smith, B., Krug, R.M., and Tao, Y.J. (2012). Biochemical and Structural Evidence in Support of a Coherent Model for the Formation of the Double-Helical Influenza A Virus Ribonucleoprotein. MBio *4*.

Yen, H.-L., Hoffmann, E., Taylor, G., Scholtissek, C., Monto, A.S., Webster, R.G., and Govorkova, E.A. (2006). Importance of neuraminidase active-site residues to the neuraminidase inhibitor resistance of influenza viruses. J. Virol. *80*, 8787–8795.

Yoneyama, M., Kikuchi, M., Natsukawa, T., Shinobu, N., Imaizumi, T., Miyagishi, M., Taira, K., Akira, S., and Fujita, T. (2004). The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses. Nat. Immunol. *5*, 730–737.

Yoon, S.-W., Webby, R.J., and Webster, R.G. (2014). Evolution and ecology of influenza A viruses. Curr. Top. Microbiol. Immunol. *385*, 359–375.

Zachariae, U., and Grubmüller, H. (2008). Importin-beta: structural and dynamic determinants of a molecular spring. Struct. Lond. Engl. 1993 *16*, 906–915.

Zebedee, S.L., and Lamb, R.A. (1988). Influenza A virus M2 protein: monoclonal antibody restriction of virus growth and detection of M2 in virions. J. Virol. *62*, 2762–2772.

Zhang, S., Wang, J., Wang, Q., and Toyoda, T. (2010). Internal initiation of influenza virus replication of viral RNA and complementary RNA in vitro. J. Biol. Chem. *285*, 41194–41201.

Zheng, W., Olson, J., Vakharia, V., and Tao, Y.J. (2013). The crystal structure and RNA-binding of an orthomyxovirus nucleoprotein. PLoS Pathog. *9*, e1003624.

Zheng, W., Li, J., Wang, S., Cao, S., Jiang, J., Chen, C., Ding, C., Qin, C., Ye, X., Gao, G.F., et al. (2015). Phosphorylation controls the nuclear-cytoplasmic shuttling of influenza A virus nucleoprotein. J. Virol. *89*, 5822–5834.

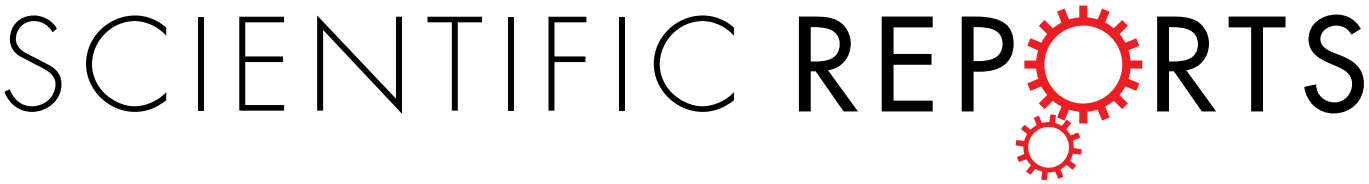
Zhu, X., McBride, R., Nycholat, C.M., Yu, W., Paulson, J.C., and Wilson, I.A. (2012). Influenza virus neuraminidases with reduced enzymatic activity that avidly bind sialic Acid receptors. J. Virol. *86*, 13371–13383.

Publications

Scientif Reports, 2016

Structural characterization of recombinant IAV polymerase reveals a stable complex between viral PA-PB1 heterodimer and host RanBP5

Christopher Swale, Alexandre Monod, Laura Tengo, Alice Labaronne, Frédéric Garzoni, Jean-Marie Bourhis, Stephen Cusack, Guy Schoehn, Imre Berger, Rob WH Ruigrok et Thibaut Crépin



OPEN

Structural characterization of recombinant IAV polymerase reveals a stable complex between viral PA-PB1 heterodimer and host RanBP5

Received: 17 November 2015
Accepted: 04 April 2016
Published: 20 April 2016

Christopher Swale[1,2,3,*], Alexandre Monod[1,2,*], Laura Tengo[1,2,3], Alice Labaronne[1,2,3], Frédéric Garzoni[1,2,4], Jean-Marie Bourhis[1,2,3], Stephen Cusack[1,2,4], Guy Schoehn[1,2,3], Imre Berger[1,2,4,5], Rob WH Ruigrok[1,2,3] & Thibaut Crépin[1,2,3]

The genome of influenza A virus (IAV) comprises eight RNA segments (vRNA) which are transcribed and replicated by the heterotrimeric IAV RNA-dependent RNA-polymerase (RdRp). RdRp consists of three subunits (PA, PB1 and PB2) and binds both the highly conserved 3′- and 5′-ends of the vRNA segment. The IAV RdRp is an important antiviral target, but its structural mechanism has remained largely elusive to date. By applying a polyprotein strategy, we produced RdRp complexes and define a minimal human IAV RdRp core complex. We show that PA-PB1 forms a stable heterodimeric submodule that can strongly interact with 5′-vRNA. In contrast, 3′-vRNA recognition critically depends on the PB2 N-terminal domain. Moreover, we demonstrate that PA-PB1 forms a stable and stoichiometric complex with host nuclear import factor RanBP5 that can be modelled using SAXS and we show that the PA-PB1-RanPB5 complex is no longer capable of 5′-vRNA binding. Our results provide further evidence for a step-wise assembly of IAV structural components, regulated by nuclear transport mechanisms and host factor binding.

The influenza virus is a negative sense RNA virus with a segmented genome belonging to the *Orthomyxoviridae* family. The viral RNA (vRNA) is divided into 8 segments that encode for a total of 10 core proteins and a few additional accessory proteins. Every vRNA segment is coated by numerous nucleoprotein (NP) units and by one RNA-dependent RNA-polymerase (RdRp) complex composed of PA (716 aa for influenza A virus (IAV) strains), PB1 (756 aa for IAV strains) and PB2 (757aa for IAV strains). The RdRp binds to the 5′- and 3′- terminal extremities of the vRNA which acts as a promoter region[1–3]. The macromolecular assembly between the vRNA, NP and the RdRp forms the ribonucleoprotein (RNP). The RdRp performs both replication and transcription of the vRNA genome in the nucleus of the cell. As such, the RdRp can produce either cRNA/vRNA through *de novo* replication or viral mRNA through "cap snatching" transcription[4]. The factors regulating the switch from a transcriptive to a replicative RdRp remain to be characterised.

Recently, the structures of influenza B[5], bat influenza A[6] and influenza C[7] were published, providing tremendous insight into the complex architecture of the RdRp. All three subunits are tightly intertwined and form intricate quaternary structures at the vRNA promoter binding sites. Current models of RdRp assembly suggest that PB2 is imported via the importin-α pathway[8,9] whereas the PA-PB1 heterodimer is imported through direct interaction with RanBP5[10–12].

[1]Université Grenoble Alpes, Unit of Virus Host Cell Interactions, UMI 3265 UJF-EMBL-CNRS, 71 avenue des Martyrs, CS 90181, F-38042 Grenoble Cedex 9, France. [2]CNRS, Unit of Virus Host Cell Interactions, UMI 3265 UJF-EMBL-CNRS, 71 avenue des Martyrs, CS 90181, F-38042 Grenoble Cedex 9, France. [3]Institut de Biologie Structurale (IBS), Univ. Grenoble Alpes, CEA, CNRS, 38044 Grenoble, France. [4]EMBL Grenoble, 71 avenue des Martyrs, CS 90181, F-38042 Grenoble Cedex 9, France. [5]The School of Biochemistry, University of Bristol, Clifton BS8 1TD, United Kingdom. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to T.C. (email: thibaut.crepin@ibs.fr)

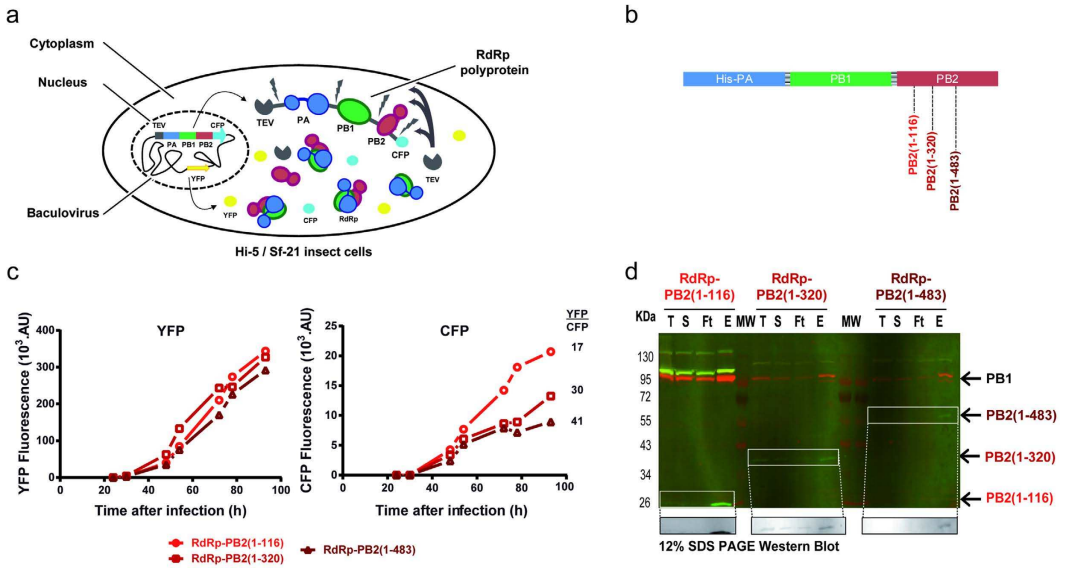


Figure 1. Truncated human-IAV RdRp polyprotein expression. (**a**) Logic of the TEV-PA-PB1-PB2-CFP polyprotein expression in a schematic view. During expression, TEV will process the polyprotein ensuring a stoichiometric assembly of PA, PB1 and PB2. YFP and CFP are produced during the process and monitor respectively baculovirus proliferation and polyprotein translation. (**b**) Truncated RdRp constructs where PB2 is incrementally extended until residues 116, 320 or 483. (**c**) YFP (left) and CFP (right) fluorescence kinetics measured during Hi-5 insect cells culture of truncated RdRp constructs. YFP ($\lambda_{exi} = 488$ nm, $\lambda_{emi} = 525$ nm) and CFP ($\lambda_{exi} = 430$ nm, $\lambda_{emi} = 480$ nm) measurements were performed on cellular extracts prepared by sonicating 1×10^6 cells in PBS (500 µL) follow by centrifugation. Fluorescence intensities are plotted against time after infection. (**d**) Small scale nickel resin purification analysis by western blot. Purifications were performed on the 50 mL of Hi-5 insect cells cultures used for the YFP and CFP fluorescence kinetics (**c**). Deposits feature total lysate after freeze/thaw (T), supernatant after centrifugation (S), resin flow through (Ft) and the primary elution fraction (E). After migration on a 12% SDS-PAGE, proteins were transferred on PVDF membrane. Primary antibodies targeting human-IAV PB2 (rabbit IgG) and human-IAV PB1 (mouse IgG) have been used. Revelation was performed with secondary goat antibodies coupled with Alexa Fluor 532 ($\lambda_{exi} = 632$ nm, $\lambda_{emi} = 647$ nm) and Alexa Fluor 633 ($\lambda_{exi} = 531$ nm, $\lambda_{emi} = 554$ nm) targeting mouse and rabbit H + L domains respectively, using a Typhoon Trio imaging system (GE Healthcare). After integration of the raw data, PB1 and PB2 revelation are visible in red and green respectively. Black and white signal of PB2 is also shown (bottom) to highlight the PB2 truncations. The upper bands appearing in green/yellow, correspond to unprocessed polyproteins.

Structural studies of RdRp from the human infecting influenza A (human-IAV) strains in contrast, have been limited to date, partly due to difficulties to produce recombinant H3N2 or H5N1 polymerases. Nonetheless, a cryo-EM reconstruction was recently obtained of a truncated influenza A/H5N1 polymerase[13]. In the present article, we identify PB2 as the main bottleneck hampering complete recombinant polymerase expression in our insect cells expression system. Moreover, we present the characterisation of several constructs of human-IAV RdRp, including the biologically relevant PA-PB1 heterodimer in solution. The titration measurements against the 5′- and 3′-vRNA promoters show a strong sub-nanomolar affinity of the PA-PB1 heterodimer towards the 5′-vRNA whereas the specific binding of the 3′-vRNA requires the presence of PB2. By applying our co-expression strategy, we show that the previously proposed PA-PB1-RanBP5 import complex can be purified, which we characterise using small angle X-ray scattering (SAXS). Biochemical analysis of this complex reveals a role for RanBP5 in hindering 5′-vRNA binding. Taken together, these data provide evidence for a cellular RdRp assembly pathway following a sequential and conditional process of assembly.

Results

PB2 is a key limiting factor for recombinant expression of the heterotrimer. Polyproteins are naturally used by certain viruses to structure their proteome[14,15]. Recombinant polyproteins mimicking viral polyproteins have recently emerged as a powerful tool to express protein complexes for functional characterisation as well as structural determination (reviewed in[16]). Using this strategy[17], soluble and active heterotrimeric RNA-dependent RNA-polymerases of influenza B virus (IBV) and bat influenza A (bat-IAV) virus could be produced, crystallized and the structure determined at high resolution[5,6]. We have applied a similar strategy to the RdRp of two human-IAV strains, A/Victoria/3/1975(H3N2) and the highly pathogenic A/Viet-Nam/1203/2004(H5N1). The three genes of each heterotrimeric complex were combined with Tobacco Etch Virus (TEV) protease and Cyan Fluorescent Protein (CFP) in a single large open reading frame (ORF). Each gene was separated by a DNA sequence encoding for a peptide segment comprising a short serine/glycine linker and a TEV protease cleavage site (Fig. 1a and supplementary Figure 1). The constructs were optimized for recombinant expression in insect

	Construct number	PA	PB1	PB2	RanBP5	YFP/CFP	Duplicate in H5N1	Solubility
heterodimers	1	1–716	1–686	–	–	6	✓	yes
	2	197–716	1–686	–	–	4	✓	yes
	3	197–716	1–660	–	–	4	✓	yes
	4	197–716	1–560	–	–	4	✓	no
	5	210–716	1–686	–	–	6	✓	no
	6	222–716	1–686	–	–	6	✓	no
	7	231–716	1–686	–	–	6	✓	no
	8	240–716	1–686	–	–	6	✓	no
	9	250–716	1–686	–	–	6	✓	no
	10	263–716	1–686	–	–	6	✓	no
	11	197–263	16–686	–	–	n.d.*	✓	no
heterotrimers	12	1–716	1–757	1–759	–	50	✓	n.d.
	13	1–716	1–757	1–36	–	15	✓	yes
	14	197–716	1–757	1–116	–	7	✓	yes
	15	1–716	1–757	1–116	–	17	✓	yes
	16	1–716	1–757	1–250	–	28	✓	n.d.
	17	1–716	1–757	1–320	–	30	✓	n.d.
	18	1–716	1–757	1–483	–	41	✓	n.d.
	19	1–716	1–757	1–116 – MBP	–	17	✓	yes
	20	1–716	1–757	1–116 – 320–483	–	20	✓	yes
	21	1–716	1–686	–	1–1115	12	–	yes

Table 1. Details of human-IAV RdRp polyproteins. The expressions have been made on both A/Victoria/3/1975(H3N2) and A/Viet-Nam/1203/2004(H5N1) strains. Identical results were obtained with the 2 strains. The table provides the data obtained with the A/Victoria/3/1975(H3N2) strains. The ratio YFP/CFP were calculated with the 2 maxima of the fluorescence spectra measured on the same sample (YFP: $\lambda_{exi} = 488$ nm, $\lambda_{emi} = 525$ nm ; CFP: $\lambda_{exi} = 430$ nm, $\lambda_{emi} = 480$ nm) corresponding to 1×10^6 infected cells. This ratio is proportional to the level of expression of the corresponding constructs. The table contains also data on the solubility of each construct in the classical purification buffer (i.e.: 50 mM Tris-HCl pH 8.5, 300 mM NaCl, 2 mM β-mercaptoethanol, 2–10% glycerol). *This construct has been cloned in pFastBac-HTB without TEV nor CFP. The only fluorescent reporter protein was the YFP. Its signal was in accordance with the YFP values obtained for the other constructs.

cells using the MultiBac system[18,19]. During expression, the TEV protease cleaves the polyprotein co-translationally, resulting in a stoichiometric assembly of the RdRp that can then be selectively purified using an affinity nickel resin. The production of the polyprotein is monitored using the fluorescence of two reporter proteins: CFP encoded within the polyprotein reports directly on recombinant protein yield, while YFP, integrated in the baculovirus genome, monitors virus performance[19]. The ratio of YFP/CFP in our system is a highly useful criterion to determine recombinant polyprotein translation efficacy of the different constructs.

Table 1 summarizes the results obtained with all constructs tested. Several attempts were made to express the H3N2 polymerase heterotrimer (construct number 12, Table 1). Although YFP and CFP can be detected and quantified, CFP values were extremely low compared to YFP with a YFP/CFP ratio of 50, the highest observed for any polyprotein constructs that we expressed. No clear bands were observed on western blot using specific PA, PB1 and PB2 antibodies. Similar results were obtained with the H5N1 RdRp polyprotein, indicating failure of these recombinant expression experiments for these full-length IAV heterotrimeric polymerases. As a comparison, using the same strategy, the YFP/CFP ratio for the expression of IBV or Bat-IAV heterotrimer was close to 20, similar to the ratios we obtained for the expressions of constructs 13 and 15 (i.e. without most of PB2).

In order to understand the difference between the expression of the human-IAV and IBV polymerases, we designed new constructs with a C-terminally truncated PB2 based on available structural information. These were PB2(1–36)[20], PB2(1–116)[21] and based on the cap-binding domain[22], PB2(322–483). Expression and purification experiments were undertaken on constructs 15, 17 and 18, all of which all have full-length PA, PB1 and truncations of PB2 extending to residues 116, 320 and 483 respectively (Fig. 1b). As the PB2 extension increases beyond residue 116, we observed a loss of expression of the polyprotein, thus identifying a critical region in PB2 that limits yield. This can first of all be observed through the monitoring of YFP/CFP reporter genes during expression (Fig. 1c). YFP fluorescence values reach a similar plateau for all constructs whereas CFP values are reduced significantly for constructs 17 and 18. Moreover, after nickel resin pull-down on lysates from an equal amount of cells, we observed by western blot (Fig. 1d) decreased intensity bands of not only PB2 but also PB1, indicative of an overall expression loss. Polyprotein size increase alone could not explain the reduced expression because other control constructs (19 and 20) with additional protein sequences beyond PB2 residue 116 conserve lower YFP/CFP ratios. Furthermore, replacing the PB2(1–116) by another large protein such as RanBP5 (residues 1–1115) generates a construct with a lower YFP/CFP ratio than construct 15. The analysis of the amino

acids sequences of PB2 does not provide more indication on the real nature of this phenomenon (supplementary Figure 2).

To determine whether this "loss-of-expression" phenomenon results from our polyprotein expression strategy, which combines PA-PB1 and PB2 in one ORF, we generated a MultiBac baculovirus co-expressing a PA-PB1 fusion and PB2 from an independent expression cassette[23]. This strategy also resulted in high YFP/CFP ratios ($\geq$35), even though the CFP gene was only fused to PA-PB1. These results identify PB2, especially when extended beyond the residue 116, as the bottleneck for expression of complete human-IAV RdRp.

Dissecting the heterotrimer. In order to identify a minimal active core of the human-IAV polymerase, we have systematically dissected the heterotrimeric complex for further characterisation using the polyprotein strategy. The expression of PA-PB1 increases significantly when PB2 is totally removed (compare constructs 1 and 13/15; Table 1). The same observation is made after removing the endonuclease domain (i.e. PA-Nter; compare constructs 14 and 13/15). After purification, PA-PB1 forms homogeneous monomeric and stable particles (Fig. 2b) whereas the complexes with PB2 (constructs 15 and 16) give rise to dimers in solution (Fig. 2a). The dimers are stable enough to withstand salt concentrations up to 2 M NaCl during purification, but the SEC elution peaks are broad suggesting that the oligomerization process is quite dynamic. By adding vRNA promoter-like molecules (e.g. the IAV panhandle)[24], we found that the dimers dissociated into monomeric RNA:protein complexes. We were able to show that the truncated PA-PB1 (construct 1) and the PA-PB1-PB2 (construct 14) exhibit expected polymerase activities. The endonuclease activity is similar to that of the isolated PA-Nter domain (supplementary Figure 3), can be inhibited by the same point mutation (i.e. PA-E80A) or compound (i.e. DPBA) and shows comparable dependency on manganese[25,26]. The constructs are also functional in RNA synthesis. Upon addition of a mixture of IAV panhandle, NTPs and [α-^{32}P]-UTP, both truncated PA-PB1 and PA-PB1-PB2 are able to generate an 80-nucleotide long product plus other minor products in absence or in presence of ApG (supplementary Figure 4).

We have also shown that further N-terminally truncating PA and/or C-terminally truncating PB1 has no impact on the expression of the corresponding polyproteins (constructs 1 to 11). The heterodimer PA-PB1 can be purified with or without the N-terminal domain of PA (i.e. PA-Nter; compare constructs 1 and 2). However, as soon as the PA-hinge (residues 200–260) is removed, PB1 becomes insoluble. This implies that the PA-hinge linking the Nter- and Cter-domains of PA is crucial for the stability of PB1. However, it is not sufficient in itself since the C-terminal region of PA also interacts with the N-terminal part of PB1[27,28] and cannot be deleted either (compare constructs 2 and 5 to 11). PB1 can be shortened on its C-terminus but when purifying the construct 4 (i.e. PB1 until amino acid 560), only PA-Cter remains soluble. Thus, the minimal construct required to obtain soluble PB1 is construct 3 with PA from 197 to 716 and PB1 from 1 to 660. All the PA-PB1 constructs without the N-terminal domain of PA form dimers in solution (supplementary Figure 5) and adding vRNA-like molecules has no incidence on the dimerization.

Over all the constructs tested, crystals have only been obtained using construct 2 (PA from 197 to 716 and PB1 from 1 to 686). Crystals grew in a few days in low PEG content solutions. They diffract poorly, with diffraction limited to typically $\approx$ 10 Å.

PA-PB1 dimer exhibits high-affinity interaction with 5′-vRNA but requires PB2 for binding 3′-vRNA. All influenza RNA segments have the same organization, a central coding region flanked by 2 un-translated regions containing the highly conserved and complementary 5′- and 3′-ends[29–31]. The viral polymerase specifically interacts with both the 5′- and 3′-ends, as recently visualized in crystal structures[5,6], and uses them as a promoter[1–3]. During their biochemical characterisation, we have seen significant effects of vRNA-like RNA molecules on the stability and/or the oligomeric state of most of the constructs. The initial RNA molecule used was the 80-nt panhandle[24] from which we made the 5′-vRNAp corresponding to the 5′-end (5′-AGUAGAAACAAGGGUA) and it's 3′ equivalent called 3′-vRNAp (AUACCCUGCUUUUGCU-3′). By thermal shift assay experiments, a 10 °C stabilisation was observed when 5′-vRNAp was added to all the constructs, whereas when 3′-vRNAp was added, the effect was less significant and dependent on the construct. Similar data had already been published[32].

We used fluorescence anisotropy to measure the interaction between our truncated polymerase constructs and the conserved vRNA ends by following the fluorescence polarization increase of a fluorescently labelled RNA when it binds the polymerase. For this purpose, we used 5′-vRNAp and 3′-vRNAp, both labelled with fluorescein amidite (FAM) at the opposite extremity of the putative interaction, i.e. at the 3′-end for the 5′-vRNAp and *vice versa* (Fig. 3 and Table 2). At 300 mM NaCl, the Kd for the 5′-vRNAp are similar (sub-nanomolar) for both constructs 1 and 14, whereas for the 3′-vRNAp, a Kd of 36 nM for the construct with PB2 is obtained (Fig. 3a) and no binding was observed for the construct without PB2 (Fig. 3c). As the estimated Kd for the 5′-end was in a range below the working concentration and beyond the sensitivity limit of the fluorescence anisotropy measurement, another method was required to determine the 5′-vRNAp affinity. Filter binding assays (FBA) were then therefore used for precise Kd determination (Fig. 3b). With the corresponding radioactive probes, we found values of 0.2–0.4 nM for the 5′-end. For the 3′-end, the Kd was 100-fold higher and critically dependent on the presence of PB2. Salt concentration is also an important parameter. We show different binding characteristics at 150 and 300 mM NaCl of the PA-PB1 towards different RNAs. The Kd value of PA-PB1 for the 5′-end remains sub-nanomolar in the range of 150 to 300 mM NaCl. In contrast, at 150 mM NaCl, 3′-vRNAp binds to PA-PB1 with the same affinity than for a poly-UC (Fig. 3c), and so is non-specific. All relevant titrations were performed at 300 mM NaCl to prevent non-specific affinity measurements. Recently published affinity data[33] working at 500 mM NaCl likewise consistent with very high affinity of the 5′-vRNAp towards IAV polymerase but this study failed to measure 3′-vRNAp binding within the nanomolar range, indicating that 3′-vRNAp binding is also tightly influenced by salt concentration.

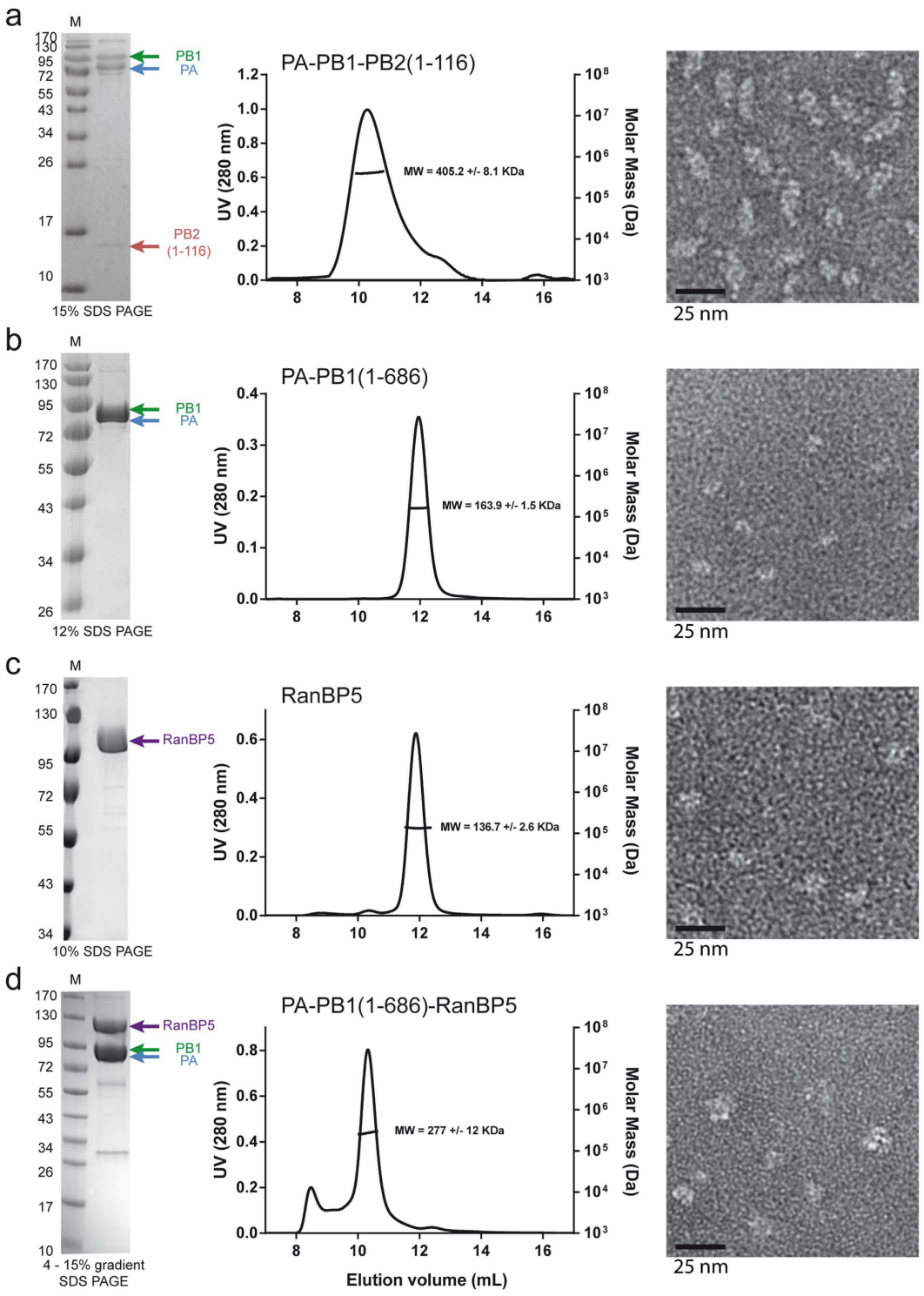


Figure 2. Homogeneous sample analysis of recombinant polymerases, RanBP5 and polymerase-RanBP5. Purified product analysis are horizontally grouped for (**a**) PA-PB1-PB2(1-116), (**b**) PA-PB1(1-686), (**c**) RanBP5 and (**d**) PA-PB1(1-686)-RanBP5. On the left are the Coomassie blue colored SDS PAGE gel of the purified sample with corresponding subunits bands indicated by colored arrows (PA in blue, PB1 in green and RanBP5 in purple). In the middle are the SEC-MALLS chromatograms with the UV signal as a backtrace and molecular weight estimate below the peak. Estimated average molecular weight for each sample is also detailed. SEC-MALLS-RI runs of PA-PB1(1-686), RanBP5 and PA-PB1(1-686)-RanBP5 were performed in the same buffer (50 mM Tris-HCl pH 8.0 and 150 mM NaCl) whereas the SEC-MALLS-RI run of PA-PB1-PB2(1-116) was performed with 50 mM Tris-HCl pH 8.0, 300 mM NaCl and 10% glycerol. (right) corresponding electron microscopy images.

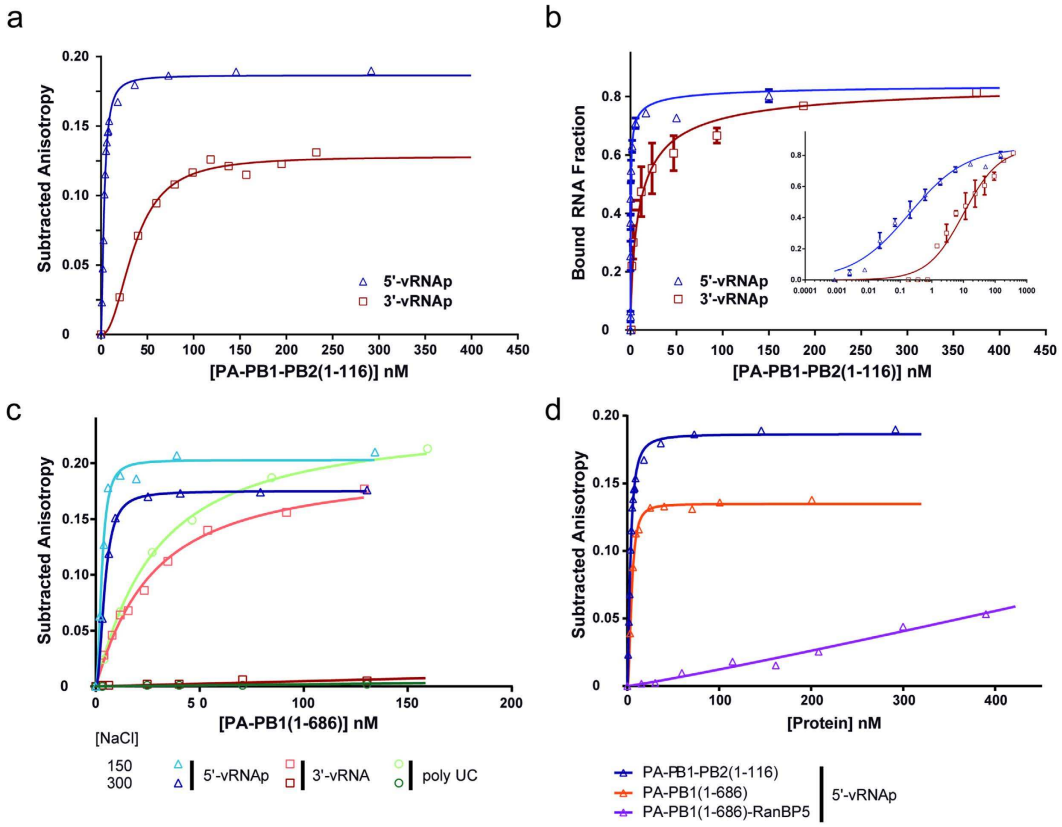


Figure 3. vRNA binding and specificity. (**a**) Binding titration of the truncated trimer PA-PB1-PB2(1-116) towards the 5′-vRNAp (blue triangle) and 3′-vRNAp (red square) sequences using fluorescence anisotropy at 300 mM NaCl. (**b**) Binding titration performed by filter binding assay against P^{32} labelled 5′-vRNAp (blue triangle) and 3′-vRNAp (red square) using 300 mM NaCl. Bound RNA fraction is plotted as a function of polymerase concentration. (**c**) Binding titration of the truncated dimer PA-PB1(1-686) performed at 150 and 300 mM NaCl against the 5-′vRNAp (dark and light blue triangles), 3′-vRNAp (orange and dark red squares) and polyUC RNA (light and dark green circles) by fluorescence anisotropy. (**d**) Binding titration of different polymerases and polymerase-RanBP5 constructs against the 5′-vRNAp at 300 mM NaCl by fluorescence anisotropy. PA-PB1-PB2(116) and PA-PB1(1-686) are depicted by blue and orange triangles respectively, PA-PB1(1-686)-RanBP5 is depicted with purple triangles. For all anisotropy titrations (**a**,**c**,**d**) subtracted anisotropy is plotted as a function of protein concentration.

Construct	NaCl (mM)	RNA	Kd (nM)	std error (nM)	R^2
construct **14**:PA-PB1-PB2(1-116)		5′-vRNAp	0.83 (0.19)	0.08 (0.02)	0.99 (0.97)
	300	3′-vRNAp	36 (17)	2 (2)	0.99 (0.95)
		polyUC	≥1000	n.d.	n.d.
construct **1**:PA-PB1(1-686)		5′-vRNAp	0.87 (0.38)	0.1 (0.05)	0.99 (0.99)
	300	3′-vRNAp	≥1000	n.d.	n.d.
		polyUC	≥1000	n.d.	n.d.
		5′-vRNAp	0.53	0.12	0.99
	150	3′-vRNAp	22.1	1.0	0.99
		polyUC	22.8	0.7	0.99
construct **21**:PA-PB1(1-686) RanBP5	300	5′-vRNAp	≥1000	n.d.	n.d.
NP	300	polyUC	29	1.2	0.94

Table 2. Titration measurements against vRNA promoters. Values in parentheses correspond to the values obtained by filter-binding assay experiments. All the experiments have been made in triplicate.

RanBP5 interacts tightly with PA-PB1(1-686) when co-expressed. During the viral cycle, the assembly of influenza RdRp follows a multi-factorial pathway involving many host partners proteins. After transcription in the nucleus, viral mRNAs are exported to the cytoplasm to be translated by the cellular machinery. The components of the replication machinery (i.e. PA, PB1, PB2 and NP) then must be imported into the nucleus.

	construct 1:PA-PB1(1-686)	construct 21:PA-PB1 (1-686)-RanBP5	construct 22 :RanBP5
Data collection parameters			
Instrument	ESRF - BM29		
Beam size at sample (μm)	700 × 700		
Wavelength (Å)	0.9919		
q range ($Å^{-1}$)	0.25–50		
Detector	Pilatus 1 M		
Detector distance (m)	2.867		
Exposure (s per image)	1		
Column	S200inc 5/150 GL		
Flow rate ($mL.min^{-1}$)	0.5	0.4	0.5
Injected sample concentrations ($mg.mL^{-1}$)	3.4	4.3	8.5
Injection volume (μL)	50		
Temperature (K)	293		
Structural parameters			
R_g (Å) [from $P(r)$]	37.9 ± 0.1	52.2 ± 0.1	39.4 ± 0.2
R_g (Å) [from Guinier]	36.2 ± 0.4	51.8 ± 0.5	38.8 ± 0.8
D_{max} (Å)	128	181	136
Porod volume estimate ($Å^3$)	254 630	577 070	215 890
Molecular-mass determination			
Molecular mass M_r (Da) [from Rambo]	146 493	323 311	143 827
Calculated M_r (Da) from sequence	165 915	291 983	125 892
Software employed			
Primary data reduction	PRIMUS		
Data processing	PRIMUS		
Ab initio analysis	DAMMIF		
Validation and averaging	DAMAVER & DAMMIN/MONSA		
Computation of model intensities	CRYSOL		
3D graphics representations	CHIMERA		

Table 3. SAXS data-collection and scattering-derived parameters.

Whereas PB2 and NP use importins-α[9,34–37], PA and PB1 are conjointly imported as a preformed heterodimeric submodule by the importin-β RanBP5[10,12,38]. Having confirmed the existence of a stable PA-PB1 complex we sought to demonstrate its interaction with RanPB5. The human *IPO5* gene which encodes for RanBP5, was expressed in insect cells. After purification, RanBP5 forms a homogeneous monomeric sample (Fig. 2c) that has been used in mixing experiments with freshly purified PA-PB1(1-686) followed by SEC-MALLS-RI experiments. The two samples were mixed in stoichiometric ratio and incubated before size exclusion chromatography (SEC). Both RanBP5 and PA-PB1(1-686) were eluted in the same single peak, but MALLS-RI indicated a molecular weight corresponding of a mixture rather than a stable ternary complex (supplementary Figure 6). We then attempted to produce a PA-PB1(1-686)-RanBP5 heterotrimer by the self-processing polyprotein strategy. A long ORF encoding for the trimeric complex was created by inserting the *IPO5* gene downstream of the PB1(1-686) coding sequence (Table 1, construct 21). Remarkably, from this polyprotein construct, PA-PB1 co-purifies with RanBP5, forming a homogeneous and stable heterotrimer (Fig. 2d).

Online SEC-SAXS was used for the characterisation of PA-PB1(1-686), RanBP5 and PA-PB1(1-686)-RanBP5, respectively in solution. Using the Vc determination method[39] on the diffusion data of PA-PB1(1-686) shows a Mw estimate of 146 kDa (Table 3 and supplementary Figure 7). The calculated Mw is 166 kDa. The Guinier transform measures a hydrodynamic radius (Rg) of 36.2 Å and GNOM produces the pair distribution function fit with a Dmax of 128 Å. Further analysis of the SAXS curve (Fig. 4a) shows a visually adequate correlation with the CRYSOL curve which uses the bat-IAV polymerase structure (PDB id: 4WSB) as a model. To evaluate the statistical similarity between experimental intensities and those computed from a model has been derived. The high chi^2 values are observed for this dataset, due to a noisy SAXS curve. A slight deviation of the fit is observed above a q range of 0.8 nm^{-1}, suggesting conformational differences between the scattering curve and the crystal structure coordinates. *Ab-initio* modelling was performed using 15 DAMMIF models which were averaged by DAMAVER. Averaged model correlation with the diffusion curve was then checked using the damstart file as a starting envelope for DAMMIN. The polymerase DAMAVER envelope (Fig. 4a) appears as a pear shaped structure in which the homologous model can be fitted. Additional envelope volume is visible close to the endonuclease domain, implying that it adopts multiple conformations in solution.

Using the same methodology, we determine RanBP5 to be a monomer in solution with a measured Mw of 144 kDa close to the theoretical Mw of 126 kDa (Table 3 and supplementary Figure 7). RanBP5 displays an Rg of 38.8 Å and a Dmax of 136 Å. With a Mw lower than that of PA-PB1(1-686), RanBP5 displays a larger Dmax and Rg indicating an extended structure. We have looked over the *Protein Data Base* (PDB) and found yeast

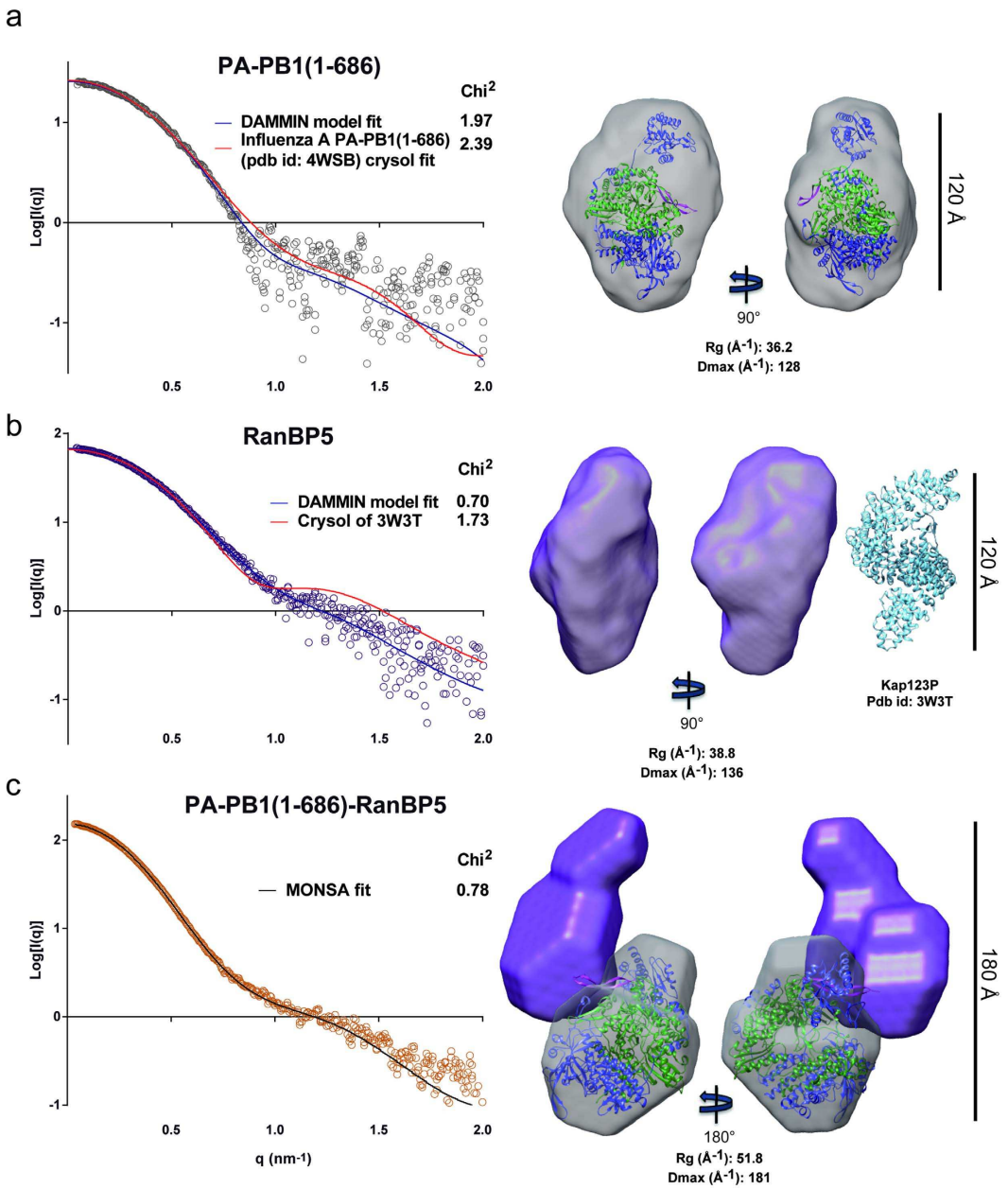


Figure 4. Online SAXS analysis of PA-PB1, RanBP5 and PA-PB1-RanBP5 complex in solution. Results are shown for (**a**) PA-PB1(1-686), (**b**) RanBP5 and (**c**) PA-PB1-RanBP5. On the left are the experimental Ln[I(q)] curves with the *ab-initio* DAMMIN curve fit (blue) using the Damstart (DAMAVER) as an initial constraint and the CRYSOL fit (red) of the closest homologous X-ray structure. The MONSA fit is also shown for PA-PB1-RanBP5 in black. The Chi^2, to evaluate the statistical similarity between experimental intensities and those computed from a model, are also shown. On the right, the homologous PDB structure of PA-PB1(1-686) from Bat-IAV polymerase is depicted in cartoon and fitted in both the DAMAVER and the MONSA generated envelops of PA-PB1(1-686) which are depicted in grey. PA is coloured in blue, PB1 in green and the NLS of PB1 in magenta. The DAMAVER and MONSA envelop of RanBp5 is coloured in purple. As a comparison, the atomic structure of Kap123P (pdb is: 3W3T) is depicted in cartoon with a sky blue colour alongside the DAMAVER envelop of RanBP5. All envelopes and models are adjusted at the same size scale.

karyopherin Kap121[40] to be the closest homologue available for RanBP5 (30% sequence identity over 1027 residues). CRYSOL curve fitting using the apo form of Kap121 (PDB id: 3W3T) against the diffusion curve reveals important structural and conformational differences between the two structures (Fig. 4b). However, envelope modelling confirms an elongated structure of RanBP5 in solution, comparable in size with yeast Kap121 in the crystal structure.

Analysis of PA-PB1(1-686)-RanBP5 describes a significantly larger complex in solution with an Rg of 51.8 Å and a Dmax of 181 Å. The Mw estimate of 323 KDa is close to the expected Mw of 292 KDa, confirming the

presence of a stoichiometric 1:1:1 PA-PB1(1-686)-RanBP5 complex in solution. MONSA builds *ab initio* shapes using the scattering curves for complexes together with those of their individual components[41]. MONSA was used to perform the *ab-initio* modelling of the complex PA-PB1(1-686)-RanBP5 (supplementary Figure 8). It separately uses the PA-PB1(1-686) and RanBP5 diffusion curves (Fig. 4a,b) in combination with that of PA-PB1(1-686)-RanBP5 to propose a consensus dual envelope (Fig. 4c). Model fitting was performed against the three experimental curves and matches all 3 curves with similar quality. With this *ab-initio* modelling we also find a comparable envelope for PA-PB1(1-686) like the one produced with DAMMIF/DAMAVER. RanBP5 on the other hand displays a different envelope suggesting conformational differences between the unbound and bound forms of RanBP5. The model suggests a close interaction between the two proteins, RanBP5 (violet envelope) interacts with PA-PB1(1-686) (grey envelope), suggesting molecular contacts well beyond those of the proposed PB1 nuclear localization signal (NLS) containing domain, which the crystal structure shows is a mobile and solvent exposed β-ribbon[6]. This model further supports the hypothesis that RanBP5 may play a chaperone function for PA-PB1 prior to assembly with PB2 in the nucleus to complete the polymerase.

RanBP5 regulates the binding of the 5′-vRNAp. The PA-PB1 complex not only exists as a stable submodule but can moreover recognise the 5′-vRNAp with a high affinity and acts as a replicase and endonuclease on its own, without PB2 present. The current model predicts the NLS of PA-PB1 to be present within PB1 (residues 187–211)[12], at the end of a beta-strand located parallel to the vRNAp binding sites, indicating a potential link between the binding to RanBP5 and/or the binding to vRNAp.

Subsequently to measuring the sub-nanomolar Kd of PA-PB1(1-686) and PA-PB1-PB2(1-116) to the 5′-vRNAp, an identical experiment was performed with the PA-PB1(1-686)-RanBP5 complex (Fig. 3d). The association curve clearly shows a complete loss of binding affinity towards the 5′-vRNAp when RanBP5 is present, both at 300 mM and 150 mM NaCl. Using SEC-MALLS-RI experiments, we have shown that an excess of 5′-vRNAp does not induce any complex dissociation of the PA-PB1(1-686)-RanBP5 trimer (supplementary Figure 9). Taken together, these experiments suggest that RanBP5 binding obscures the 5′-vRNAp binding site, thus blocking specific 5′-vRNAp binding.

Discussion

Expression of full-length IBV and bat-IAV polymerases has been successfully achieved using a polyprotein approach, resulting in atomic resolution structures by X-ray crystallography[5,6]. In marked contrast and unexpectedly, efficient expression of active full-length human or avian IAV polymerase at yields high enough for structural studies has not been reported to date, irrespective of the method used. Our studies identify PB2, and notably residues 120 to 250 as responsible for this poor expression. Interestingly, attempts by a different research group to reconstitute a full IAV polymerase (e.g. full IAV polymerase expression/purification for structural studies) likewise failed beyond the N-terminal PB1 interacting domain of PB2[13]. Furthermore, the recent bat-IAV, IBV and ICV RdRp crystal structures have shown that PB2 can adopt a plethora of conformations, demonstrating how dynamic PB2 is within the heterotrimer[5–7,42,43]. The reason why PB2 hinders polyprotein expression is not clear at the moment. We speculate that the extended fold of the PB2 region 1–250 is unstable unless bound to one side of PB1[5,6] and thus may require chaperones to maintain its stability prior to assembly onto PA-PB1. Of note, Hsp90 has been shown to stimulate influenza virus RNA synthesis[44] and the nuclear import of the subunits[35]. Moreover, Hsp90 inhibitors reduce influenza virus replication in cell culture[45]. We have shown previously that provision of co-factors in *trans* can ameliorate the quality and quantity of target proteins expressed in the baculovirus/insect cell system[46,47]. Provision of Hsp90 and the other chaperones, within the polyprotein or in *trans* on the baculovirus backbone, may provide a powerful handle to overcome the bottleneck of human-AIV polymerase production. Furthermore, although the cellular context between viral infection and the expression of recombinant proteins is different, it cannot be excluded that the polyprotein interferes with the cellular innate immunity. IAV RdRp, and PB2 in particular, has been shown to regulate host anti-viral response through the binding to IPS-1 and the inhibition of interferon β production[48].

PA-PB1 can be produced as a stable submodule and forms a discrete heterodimer in solution. Dimerization of PA-PB1 heterodimers can appear when either the endonuclease is removed or the N-terminal of PB2 is added which may indicate that the superstructure formation could be artificially induced by truncations of the full-length polymerase. PA-PB1 can specifically bind to the 5′-vRNAp with sub-nanomolar affinity in solution. In contrast, 3′-vRNAp binding requires the presence of PB2 and is within the low nanomolar range, indicating a much stronger preference of the polymerase towards the 5′-vRNAp. These results are consistent with the high-resolution crystal structures in which the 5′-vRNAp binding site is located between the PA and PB1 subunits[5,6]. The 3′-vRNA, on the other hand, interacts with all three subunits suggesting a sequential vRNAp binding mechanism.

Our results corroborate key steps of influenza polymerase assembly[10,38,49]. Purification of stoichiometric complexes using the polyprotein strategy confirms that PA-PB1 can exist as a preformed and stable submodule, independently of additional host factors, with PA acting as a putative "chaperone" for PB1. Our data also shows that PA-PB1 is functional in RNA synthesis, which has interesting physiological implications. Moreover, we demonstrate here that the PA-PB1(1-686) heterodimer can assemble with RanBP5 into a stable, heterotrimeric complex that can be purified *in vitro*. This likely corresponds to the import complex for these two subunits, although other chaperones such as HSP90 may be co-imported[35]. In contrast to PA-PB1 with RanBP5, we failed to reconstitute a binary PB2-importin-α complex using the polyprotein strategy even though functional and structural data show that PB2 nuclear import depends on the importin-α pathway, through a direct interaction between the PB2 C-terminal domain that contains the NLS[8,9]. This is probably partly due to an inherent instability of isolated IAV PB2 in insect cells but could also reflect the lack of sufficient chaperones, co-chaperones or other unidentified factors that stabilise PB2. Furthermore, cytoplasmic expression of PB2 appears to destabilise

many cellular processes[50,51]. We also still do not know the reason for the difference in the import pathway usage between the PA-PB1 submodule and PB2. The lack of high expression levels of fully assembled PA-PB1-PB2 heterotrimers unfortunately currently limits our efforts to elucidate the structure of the complete human IAV polymerase at atomic resolution. Nevertheless, docking of bat-IAV and IBV crystal structures into the PA-PB1(1-686)-RanBP5 SAXS envelope suggests that assembly of PB2 onto the PA-PB1-RanBP5 complex might be possible, based on steric considerations. Indeed RanBP5 could even facilitate trimer assembly perhaps concomitantly with Ran-dependent dissociation from PA-PB1. In contrast, RanBP5 profoundly affects the ability of PA-PB1 to bind vRNA. We show that whereas the 5′-vRNA is specifically recognized by PA-PB1 with high affinity, this is abolished in the ternary complex with RanBP5. This again suggests that RanBP5 disassociation may be coupled to assembly of the polymerase with the promoter RNA. In any case, only when PB2 associates is the polymerase able to associate with the 3′-vRNA end, consistent with the crystal structures[5,6]. But surprisingly, the affinity of the heterotrimer for 3′-vRNA is in the same order of magnitude than the affinity of NP for RNA substrates (Table 2), revealing a potential competition between the complete polymerase heterotrimer and the nucleoprotein.

We demonstrate here for the first time that the polyprotein expression strategy can be used not only to co-express subunits that form a stable complex, but also to successfully reconstitute interactions with downstream partner proteins *in vivo* in the expression host for direct purification and functional characterisation. Functional studies on IAV RdRp reveal that vRNAp promoter binding is completely abolished when PA-PB1(1-686) is bound to RanBP5, suggesting a possible role for RanBP5 in maintaining PA-PB1 in an inactive form during nuclear import. This discovery holds a promise for influenza research as it provides a novel point of therapeutic intervention for targeting and inhibiting influenza virus during assembly of its active components. We anticipate that compounds which block polymerase/RanBP5 binding or release may prove to be highly efficient, broad spectrum antiviral inhibitors in the treatment of influenza, in particular for patient groups where classical vaccination strategies fail.

Methods

Molecular biology. To avoid expression of alternative proteins, all the genes used for this work are synthetic genes, codon optimised for insect cells expression. The DNA coding sequences of A/Victoria/3/1975(H3N2), the highly pathogenic A/Viet-Nam/1203/2004(H5N1) polymerase subunits and human *IPO5* have been ordered to GeneArt (ThermoFisher Scientific), optimized for the expression in insect cells. Cloning has been achieved following the supplier procedures (New England Biolabs). The pPBAC plasmid was used for the polyprotein constructs[16,18]. Few constructs have been expressed using pFastBac-HTB (Life Technologies). vRNA-like molecules (i.e. IAV panhandle)[24] have been produced and purified using classical *in vitro* transcription protocols[52,53].

Expression and purification. Large scale suspension cultures expressing polymerase fusion constructs were prepared using High Five insect cells grown in Express Five media (Life Technologies) at 0.5×10^6 cells/mL infected at 0.2% (V/V) with the baculovirus mother solution. Cultures were maintained at $0.5–1 \times 10^6$ cells/mL until proliferation arrest (24–48 h after infection). Following the proliferation arrest YFP and CFP measurements were performed every 12 h until a fluorescence plateau was reached (72–96 h after infection). Cultures were then spun down at 800 g for 10 min and cell pellets were stored at −80 °C.

Cell pellets were resuspended in 50 mL of lysis buffer (50 mM Tris-HCl pH 8.5, 300 mM NaCl and 2 mM β-mercaptoethanol) per 500×10^6 cells in the presence of EDTA-free anti-protease cocktail (complete from Roche). Lysis was performed with two cycles of freezing (−180 °C)/thawing (26 °C) after which 10% of glycerol were added to the lysate before centrifugation (45 min, 40 000 g, 4 °C). After retrieval of the clarified lysate, 30 mM of imidazole pH 8.0 were added before loading on Ni-NTa superpose resin (Quiagen). After flowing the lysate through the resin, three wash steps with 10 column volumes (CV) of buffer A (lysis buffer with 30 mM imidazole), 10 CV of wash buffer B (50 mM Tris-HCl pH 8.5, 1 M NaCl, 10% glycerol and 2 mM β-mercaptoethanol) and 10 CV of wash buffer A were performed. Elution of the bound complex was performed with 300 mM imidazole. Elution fractions containing polymerase were then pooled together and directly injected on a 5 mL Hitrap heparin resin (GE healthcare) which had previously been equilibrated with 5 CV of buffer A. After binding to the resin, a 5 CV wash was performed with buffer A before eluting with a 40 mL salt gradient on a FPLC system. The elution peak corresponding to stoichiometric polymerase assemblies was then pooled and injected on an S200 (GE healthcare) size exclusion chromatography equilibrated in 50 mM Tris-HCl pH 8.5, 300 mM NaCl, 5 mM β-mercaptoethanol. Peak fractions were pooled and concentrated to the desired concentration using a 100 kDa concentrator. Once concentrated the protein prep could be stored by addition of 20% glycerol and flash freezing in liquid nitrogen.

NP was purified using the protocol described in[54,55].

SEC-MALLS-RI analysis. All MALLS runs were performed using a S200 increase SEC column (10/300 GL, GE Healthcare). Sample injection and buffer flow was controlled by a Hitachi L2130 pump, following the SEC column was a L-2400 UV detector (Hitachi), Optilab T-rEX refractometer (Wyatt technologies) and a DAWN HELEOS-II multi angle light scattering detector (Wyatt technologies). Prior to injection, columns and systems were equilibrated in 5 to 10 column volumes of running buffer. 50 μL injections were performed using protein samples concentrated at a minimum of 2 $mg.mL^{-1}$, constant flow rate of 0.5 $mL.min^{-1}$ was used. Accurate MALLS mass prediction was performed with the Astra software (Wyatt Technologies). Curves were represented with Graphpad (Prism).

Electron microscopy. Samples (0.1 $mg.mL^{-1}$) were applied to the clean side of carbon on mica. The carbon was then floated on sodium silicotungstate (2% v/w) or uranyl acetate (2% v/w) and a grid placed on the top of

it. After air-drying, the samples were observed in a T12 FEI electron microscope and images were taken using an Orius SC1000 CCD camera (Gtan Inc., Pleasanton, CA).

SAXS analysis. All datasets were collected on BM29 (ESRF). The initial trials using direct sample measurement showed a slight concentration dependency of the Guinier estimated Rg. Online SEC-SAXS was therefore used for all experiments in order to limit the contribution of partially aggregated, oligomeric or dissociated protein to the experimental curve. The experimental setup consists of a High Pressure Liquid Chromatography (HPLC) system connected to an analytical S200 increase column (5/150 GL, GE Healthcare) followed downstream by the SAXS sample capillary. SAXS measurements were performed every second with a Pilatus 1 M detector at distance of 2.87 m allowing a q range of 0.03 to 4.5 nm with a wavelength of 0.01 nm.

Following data collection, experimental curves were subtracted and analysed using Primus (ATSAS) or Scatter (Bioisis) programs suits[56]. To verify the molecular mass, the Rambo and Tainer method was used[39,57]. Rg predictions using the Guinier extrapolation were plotted against the elution volume to select the most monodisperse part of the protein elution peak. SAXS datasets within this zone were then scaled and averaged to produce one unique I(q) curve. GNOM[58] was then used to produce a P(r) function which was subsequently used by DAMMIF[59] to generate 15 *ab-initio* models. DAMAVER[60] was then used to generate an average model. Consistency of the DAMAVER averaging with the original experimental curve was assessed by using DAMMIN[41] and using the damstart.pdb envelope as a starting model. Homologue PDB structure comparison was assed using Crysol[61]. Multiple diffusion curve *ab-initio* modelling was performed using Bunch[62]. Homologue structure fitting within the DAMAVER envelope was performed with Chimera[63] and curve representations using Graphpad (Prism).

Fluorescence anisotropy assay. Equilibrium RNA binding experiments were performed with the following vRNA set synthetized by Integrated DNA Technologies (IDT): 5′-vRNAp (5′-pAGUAGAAACAAGG GUA-FAM3′), 3′-vRNAp (5′FAM-AUACCCUGCUUUUGCU-3′) and polyUC (5′-pUCUCUCUCUCU CUCUCUC-FAM3′). 5′-vRNA sequences were synthetized with a 5′-mono phosphate and labeled with 6-FAM in 3′OH. 3′ vRNA were labeled with FAM in 5′. Polymerase constructs was titrated into fluorescently labeled RNA using the buffer (50 mM HEPES pH 7.5, 5 mM β-mercaptoethanol plus 150 mM or 300 mM NaCl) at room temperature. Initial reactions were performed with 2–4 nM of RNA with an initial volume of 600 µL to limit the dilution effect of protein addition (<5%). Anisotropy measurements were undertaken using a PTI fluorometer equipped with automated polarizers. The excitation and emission wavelengths were 494 and 521 nm, respectively. 50 anisotropy measurements with a 1 second integration time were performed per titration point.

Filter Binding Assay. Equilibrium RNA binding experiments were undertaken with the same vRNA set synthetized by IDT. vRNA were labeled with^{32} P in 5′ using the T4 polynucleotide kinase (New England Biolabs). Double filter binding was used following the previously established protocol[64]. Protein concentration was titrated against constant concentrations of vRNA ($\leq$0,01 nM) within a final volume of 200 µL using the standard protein buffer (50 mM HEPES pH 7.5, 300 mM NaCl, 5 mM β-mercaptoethanol). 180 µL were then filtered on two membranes using a 96 well Whatman Minifold dot-blot apparatus. The first membrane was Protran BA 85 membrane (Whatman) to retain the polymerase and the second one is a nylon Hybond-N + membrane (Amersham Bioscience) to retain unbound nucleic acid. Both membranes were pre-incubated 1 h at room temperature within the protein buffer prior to use. After blotting, another 180 µL of protein buffer were run through the membranes to wash any unbound RNA before letting the vacuum dry the dot-blots for 30 seconds. Both membranes were then dismounted from the dot-blot apparatus and exposed to a storage phosphor screen (BAS storage phosphor screen, GE Healthcare) overnight. Revelation of the phosphor screen was then performed with a Typhoon Trio imaging system (GE Healthcare). Dot blot phosphorescence intensity was integrated using the Image J software. A derived bound ratio was then calculated using the following formula (1):

$$Br = \frac{Ibound}{Ibound + Ifree} \tag{1}$$

Where I_{bound} is the integrated intensity of protein/RNA complex which is retained by the protran membrane and I_{free} is the unbound RNA retained by the nylon Hybond-N + membrane.

Binding assay data analysis. RNA binding curves were plotted with the subtracted anisotropy or Bound RNA ratio as a function of the protein concentration. Sigmoidal binding curves were fitted to the data using GraphPad (Prism) with the following two equations:

(1) For the anisotropy titrations, [vRNA] ≈ Kd. Assuming a 1/1 stoichiometry the following equation (2) was used to estimate the Kd

$$\Delta A = \left\{ \frac{(R_T + E_T + K_d) - [(R_T + E_T + K_d)^2 - 4 \times R_T \times E_T]^{1/2}}{2 \times R_T} \right\} \times \Delta A_T \tag{2}$$

Where ΔA is the change in subtracted anisotropy, ΔA_T the total change of subtracted anisotropy. R_T is the total vRNA concentration, E_T is the total polymerase concentration at each given point and Kd is the dissociation constant.

(2) In the case of filter binding assay, a single site isotherm was used as we consider that [vRNA] ≪ Kd. This equation (3) therefor assumes an effective equality between the free polymerase concentration and the total polymerase concentration.

$$\Delta Br = \left\{ \frac{\Delta Br_T(E_T/K_d)}{(1 + E_T/K_d)} \right\} \quad (3)$$

Where ΔBr is the change in binding ratio, ΔBr_T is the total change in binding ratio. E_T is the total polymerase concentration at every point in titration and Kd is the dissociation constant.

Standard error calculations for the apparent Kd within a given condition were calculated through triplicate data measurement.

References

1. Cianci, C., Tiley, L. & Krystal, M. Differential activation of the influenza virus polymerase via template RNA binding. *J Virol* **69,** 3995–3999 (1995).
2. Hagen, M., Chung, T. D., Butcher, J. A. & Krystal, M. Recombinant influenza virus polymerase: requirement of both 5′ and 3′ viral ends for endonuclease activity. *J Virol* **68,** 1509–1515 (1994).
3. Tiley, L. S., Hagen, M., Matthews, J. T. & Krystal, M. Sequence-specific binding of the influenza virus RNA polymerase to sequences located at the 5′ ends of the viral RNAs. *J Virol* **68,** 5108–5116 (1994).
4. Plotch, S. J., Bouloy, M., Ulmanen, I. & Krug, R. M. A unique cap(m7GpppXm)-dependent influenza virion endonuclease cleaves capped RNAs to generate the primers that initiate viral RNA transcription. *Cell* **23,** 847–858 (1981).
5. Reich, S. *et al.* Structural insight into cap-snatching and RNA synthesis by influenza polymerase. *Nature* **516,** 361–366, doi: 10.1038/nature14009 (2014).
6. Pflug, A., Guilligay, D., Reich, S. & Cusack, S. Structure of influenza A polymerase bound to the viral RNA promoter. *Nature* **516,** 355–360, doi: 10.1038/nature14008 (2014).
7. Hengrung, N. *et al.* Crystal structure of the RNA-dependent RNA polymerase from influenza C virus. *Nature* **527,** 114–117, doi: 10.1038/nature15525 (2015).
8. Mukaigawa, J. & Nayak, D. P. Two signals mediate nuclear localization of influenza virus (A/WSN/33) polymerase basic protein 2. *J Virol* **65,** 245–253 (1991).
9. Tarendeau, F. *et al.* Structure and nuclear import function of the C-terminal domain of influenza virus polymerase PB2 subunit. *Nat Struct Mol Biol* **14,** 229–233 (2007).
10. Deng, T. *et al.* Role of ran binding protein 5 in nuclear import and assembly of the influenza virus RNA polymerase complex. *J Virol* **80,** 11911–11919 (2006).
11. Huet, S. *et al.* Nuclear import and assembly of influenza A virus RNA polymerase studied in live cells by fluorescence cross-correlation spectroscopy. *J Virol* **84,** 1254–1264 (2010).
12. Hutchinson, E. C., Orr, O. E., Man Liu, S., Engelhardt, O. G. & Fodor, E. Characterization of the interaction between the influenza A virus polymerase subunit PB1 and the host nuclear import factor Ran-binding protein 5. *J Gen Virol* **92,** 1859–1869 (2011).
13. Chang, S. *et al.* Cryo-EM structure of influenza virus RNA polymerase complex at 4.3 A resolution. *Mol Cell* **57,** 925–935, doi: 10.1016/j.molcel.2014.12.031 (2015).
14. Al-Tawfiq, J. A., Zumla, A. & Memish, Z. A. Coronaviruses: severe acute respiratory syndrome coronavirus and Middle East respiratory syndrome coronavirus in travelers. *Current opinion in infectious diseases* **27,** 411–417, doi: 10.1097/QCO.0000000000000089 (2014).
15. Lucas, S. & Nelson, A. M. HIV and the spectrum of human disease. *The Journal of pathology* **235,** 229–241, doi: 10.1002/path.4449 (2015).
16. Crepin, T. *et al.* Polyproteins in structural biology. *Curr Opin Struct Biol* **32,** 139–146, doi: 10.1016/j.sbi.2015.04.007 (2015).
17. Nie, Y., Bellon-Echeverria, I., Trowitzsch, S., Bieniossek, C. & Berger, I. Multiprotein complex production in insect cells by using polyproteins. *Methods in molecular biology* **1091,** 131–141, doi: 10.1007/978-1-62703-691-7_8 (2014).
18. Bieniossek, C., Imasaki, T., Takagi, Y. & Berger, I. MultiBac: expanding the research toolbox for multiprotein complexes. *Trends Biochem Sci* **37,** 49–57, doi: 10.1016/j.tibs.2011.10.005 (2012).
19. Trowitzsch, S., Bieniossek, C., Nie, Y., Garzoni, F. & Berger, I. New baculovirus expression tools for recombinant protein complex production. *J Struct Biol* **172,** 45–54 (2010).
20. Sugiyama, K. *et al.* Structural insight into the essential PB1-PB2 subunit contact of the influenza virus RNA polymerase. *Embo J* **28,** 1803–1811 (2009).
21. An, Y., Meresse, P., Mas, P. J. & Hart, D. J. CoESPRIT: a library-based construct screening method for identification and expression of soluble protein complexes. *PLoS One* **6,** e16261, doi: 10.1371/journal.pone.0016261 (2011).
22. Guilligay, D. *et al.* The structural basis for cap binding by influenza virus polymerase subunit PB2. *Nat Struct Mol Biol* **15,** 500–506 (2008).
23. Bieniossek, C., Richmond, T. J. & Berger, I. MultiBac: multigene baculovirus-based eukaryotic protein complex production. *Current protocols in protein science / editorial board, John E. Coligan …*[*et al.*] Chapter 5, Unit 5 **20,** doi: 10.1002/0471140864.ps0520s51 (2008).
24. Baudin, F., Bach, C., Cusack, S. & Ruigrok, R. W. Structure of influenza virus RNP. I. Influenza virus nucleoprotein melts secondary structure in panhandle RNA and exposes the bases to the solvent. *Embo J* **13,** 3158–3165 (1994).
25. Crepin, T. *et al.* Mutational and metal binding analysis of the endonuclease domain of the influenza virus polymerase PA subunit. *J Virol* **84,** 9096–9104 (2010).
26. Dias, A. *et al.* The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit. *Nature* **458,** 914–918 (2009).
27. He, X. *et al.* Crystal structure of the polymerase PA(C)-PB1(N) complex from an avian influenza H5N1 virus. *Nature* **454,** 1123–1126 (2008).
28. Obayashi, E. *et al.* The structural basis for an essential subunit interaction in influenza virus RNA polymerase. *Nature* **454,** 1127–1131 (2008).
29. Desselberger, U., Racaniello, V. R., Zazra, J. J. & Palese, P. The 3′ and 5′-terminal sequences of influenza A, B and C virus RNA segments are highly conserved and show partial inverted complementarity. *Gene* **8,** 315–328 (1980).
30. Robertson, J. S. 5′ and 3′ terminal nucleotide sequences of the RNA genome segments of influenza virus. *Nucleic Acids Res* **6,** 3745–3757 (1979).
31. Skehel, J. J. & Hay, A. J. Nucleotide sequences at the 5′ termini of influenza virus RNAs and their transcripts. *Nucleic Acids Res* **5,** 1207–1219 (1978).
32. Brownlee, G. G. & Sharps, J. L. The RNA polymerase of influenza a virus is stabilized by interaction with its viral RNA promoter. *J Virol* **76,** 7103–7113 (2002).

33. Tomescu, A. I., Robb, N. C., Hengrung, N., Fodor, E. & Kapanidis, A. N. Single-molecule FRET reveals a corkscrew RNA structure for the polymerase-bound influenza virus promoter. *Proc Natl Acad Sci USA* **111,** E3335–3342, doi: 10.1073/pnas.1406056111 (2014).
34. Melen, K. *et al.* Importin alpha nuclear localization signal binding sites for STAT1, STAT2, and influenza A virus nucleoprotein. *J Biol Chem* **278,** 28193–28200 (2003).
35. Naito, T., Momose, F., Kawaguchi, A. & Nagata, K. Involvement of Hsp90 in assembly and nuclear import of influenza virus RNA polymerase subunits. *J Virol* **81,** 1339–1349, doi: 10.1128/JVI.01917-06 (2007).
36. Portela, A. & Digard, P. The influenza virus nucleoprotein: a multifunctional RNA-binding protein pivotal to virus replication. *J Gen Virol* **83,** 723–734 (2002).
37. Weber, F., Kochs, G., Gruber, S. & Haller, O. A classical bipartite nuclear localization signal on Thogoto and influenza A virus nucleoproteins. *Virology* **250,** 9–18, doi: 10.1006/viro.1998.9329 (1998).
38. Fodor, E. & Smith, M. The PA subunit is required for efficient nuclear accumulation of the PB1 subunit of the influenza A virus RNA polymerase complex. *J Virol* **78,** 9144–9153, doi: 10.1128/JVI.78.17.9144-9153.2004 (2004).
39. Rambo, R. P. & Tainer, J. A. Accurate assessment of mass, models and resolution by small-angle scattering. *Nature* **496,** 477–481, doi: 10.1038/nature12070 (2013).
40. Kobayashi, J. & Matsuura, Y. Structural basis for cell-cycle-dependent nuclear import mediated by the karyopherin Kap121p. *J Mol Biol* **425,** 1852–1868, doi: 10.1016/j.jmb.2013.02.035 (2013).
41. Svergun, D. I. Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys. J.* **76,** 2879–2886 (1999).
42. Delaforge, E. *et al.* Large-Scale Conformational Dynamics Control H5N1 Influenza Polymerase PB2 Binding to Importin alpha. *Journal of the American Chemical Society* **137,** 15122–15134, doi: 10.1021/jacs.5b07765 (2015).
43. Thierry, E. *et al.* Influenza Polymerase Can Adopt an Alternative Configuration Involving a Radical Repacking of PB2 Domains. *Mol Cell* **61,** 125–137, doi: 10.1016/j.molcel.2015.11.016 (2016).
44. Momose, F. *et al.* Identification of Hsp90 as a stimulatory host factor involved in influenza virus RNA synthesis. *J Biol Chem* **277,** 45306–45314, doi: 10.1074/jbc.M206822200 (2002).
45. Chase, G. *et al.* Hsp90 inhibitors reduce influenza virus replication in cell culture. *Virology* **377,** 431–439, doi: 10.1016/j.virol.2008.04.040 (2008).
46. Fitzgerald, D. J. *et al.* Multiprotein expression strategy for structural biology of eukaryotic complexes. *Structure* **15,** 275–279, doi: 10.1016/j.str.2007.01.016 (2007).
47. Palmberger, D. & Rendic, D. SweetBac: Applying MultiBac Technology Towards Flexible Modification of Insect Cell Glycosylation. *Methods in molecular biology* **1321,** 153–169, doi: 10.1007/978-1-4939-2760-9_11 (2015).
48. Iwai, A. *et al.* Influenza A virus polymerase inhibits type I interferon induction by binding to interferon beta promoter stimulator 1. *J Biol Chem* **285,** 32064–32074, doi: 10.1074/jbc.M110.112458 (2010).
49. Deng, T., Sharps, J., Fodor, E. & Brownlee, G. G. *In vitro* assembly of PB2 with a PB1-PA dimer supports a new model of assembly of influenza A virus polymerase subunits into a functional trimeric complex. *J Virol* **79,** 8669–8674 (2005).
50. Graef, K. M. *et al.* The PB2 subunit of the influenza virus RNA polymerase affects virulence by interacting with the mitochondrial antiviral signaling protein and inhibiting expression of beta interferon. *J Virol* **84,** 8433–8445 (2010).
51. Patel, D., Schultz, L. W. & Umland, T. C. Influenza A polymerase subunit PB2 possesses overlapping binding sites for polymerase subunit PB1 and human MAVS proteins. *Virus Res* **172,** 75–80, doi: 10.1016/j.virusres.2012.12.003 (2013).
52. Degut, C., Monod, A., Brachet, F., Crepin, T. & Tisne, C. *In Vitro*/*In Vivo* Production of tRNA for X-Ray Studies. *Methods in molecular biology* **1320,** 37–57, doi: 10.1007/978-1-4939-2763-0_4 (2016).
53. Price, S. R., Ito, N., Oubridge, C., Avis, J. M. & Nagai, K. Crystallization of RNA-protein complexes. I. Methods for the large-scale preparation of RNA suitable for crystallographic studies. *J Mol Biol* **249,** 398–408 (1995).
54. Boulo, S. *et al.* Human importin alpha and RNA do not compete for binding to influenza A virus nucleoprotein. *Virology* **409,** 84–90 (2011).
55. Chenavas, S. *et al.* Monomeric nucleoprotein of influenza A virus. *PLoS Pathog* **9,** e1003275, doi: 10.1371/journal.ppat.1003275 (2013).
56. Petoukhov, M. V., Konarev, P. V., Kikhney, A. G. & Svergun, D. I. ATSAS 2.1 – towards automated and web-supported small-angle scattering data analysis. *J. Appl. Crystallogr.* **40,** s223–s228 (2007).
57. Rambo, R. P. & Tainer, J. A. Super-resolution in solution X-ray scattering and its applications to structural systems biology. *Annual review of biophysics* **42,** 415–441, doi: 10.1146/annurev-biophys-083012-130301 (2013).
58. Semenyuk, A. V. & Svergun, D. I. GNOM – a program package for small-angle scattering data processing. *J. Appl. Crystallogr.* **24,** 537–540 (1991).
59. Franke, D. & Svergun, D. I. DAMMIF , a program for rapid ab-initio shape determination in small-angle scattering. *J. Appl. Crystallogr.* **42,** 342–346 (2009).
60. Volkov, V. V. & Svergun, D. I. Uniqueness of ab initio shape determination in small-angle scattering. *J. Appl. Crystallogr.* **36,** 860–864 (2003).
61. Svergun, D., Barberato, C. & Koch, M. H. J. CRYSOL – a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *J. Appl. Crystallogr.* **28,** 768–773 (1995).
62. Petoukhov, M. V. & Svergun, D. I. Global rigid body modeling of macromolecular complexes against small-angle scattering data. *Biophysical journal* **89,** 1237–1250, doi: 10.1529/biophysj.105.064154 (2005).
63. Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25,** 1605–1612, doi: 10.1002/jcc.20084 (2004).
64. Wong, I. & Lohman, T. M. A double-filter method for nitrocellulose-filter binding: application to protein-nucleic acid interactions. *Proc Natl Acad Sci USA* **90,** 5428–5432 (1993).

Acknowledgements

We are very grateful to Daphna Fenel, Francine Gérard, Raphaelle Germi, Laurence Grossi, Darren J Hart, Marc Jamin, Thomas Lunardi, Hélène Malet, Stefan Reich and Adam Round and for their help. We thank Kausiki Datta, Klaus Klumpp and Christoph Bieniossek for discussion. We thank Juan Ortin for the gift of the RdRp antibodies. CS was funded by the Hofmann-La Roche pRED external collaboration program, AM by the Région Rhône-Alpes (ARC 1 Santé) and AL through the Labex GRAL. This work was supported by the French agency for Research through the ANR Flunucleovir (ANR-2010-Blanc-1307-1301), the ANR RNAP-IAV (ANR-14-CE09-0017) and the European Commission projects Flupharm (contract number 259751) and ComplexINC (contact number 279039). This study used the platforms of the Grenoble Instruct centre (ISBG; UMS 3518 CNRS-CEA-UJF-EMBL) with support from FRISBI (ANR-10-INSB-05-02) and GRAL (ANR-10-LABX-49-01) within the Grenoble Partnership for Structural Biology (PSB).

Author Contributions

T.C., I.B., S.C. and R.W.H.R. conceived the experiments. C.S., A.M., L.T., F.G., A.L. and T.C. performed the biochemistry. A.L. and G.S. performed EM. C.S., J.M.B. and T.C. performed SAXS experiments. C.S., A.M., S.C. and T.C. performed X-ray crystallography. C.S., A.M., A.L., R.W.H.R. and T.C. analysed the Data. C.S., G.S., J.M.B. and T.C. prepared the figures. C.S., A.M., S.C., G.S., J.M.B., I.B., R.W.H.R. and T.C. wrote the Paper.

Additional Information

Supplementary information accompanies this paper at http://www.nature.com/srep

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Swale, C. *et al.* Structural characterization of recombinant IAV polymerase reveals a stable complex between viral PA-PB1 heterodimer and host RanBP5. *Sci. Rep.* **6**, 24727; doi: 10.1038/srep24727 (2016).

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/

Structural characterization of recombinant IAV polymerase reveals a stable complex between viral PA-PB1 heterodimer and host RanBP5

Christopher SWALE[1,2,3$], Alexandre MONOD[1,2$], Laura TENGO[1,2,3], Alice LABARONNE[1,2,3], Frédéric GARZONI[1,2,4], Jean-Marie BOURHIS[1,2,3], Stephen CUSACK[1,2,4], Guy SCHOEHN[1,2,3], Imre BERGER[1,2,4,5], Rob WH RUIGROK[1,2,3] & Thibaut CREPIN[1,2,3*]

Supplementary information

Supplementary Figure 1

Sequences of the main polyprotein constructs. The figure details the open reading frames of constructs 12, 15 and 21. The accession numbers for the protein sequences of the TEV protease, the A/Victoria/3/1975(H3N2) and A/Viet-Nam/1203/2004(H5N1) RdRP subunits, human RanBP5 isoform 3 and the cyan fluorescent protein are provided in the right part of the schema. The amino acid sequences for the TEV-cleavage site and the polyhistidine affinity tag are coloured respectively in green and red. All the accession numbers correspond to UniProtKB access numbers, except the sequence for the TEV protease (NCBI accession number).

Supplementary Figure 2

Sequences alignment of PB2. The alignment has been made for the sequences of the 2 human-IAV strains used for this work plus the 2 strains for which the structures have been recently solved. bat-IAV corresponds to A/little yellow-shouldered bat/Guatemala/060/2010(H17N10) (UniProtKB access number: H6QM90), H3N2-IAV to A/Victoria/3/1975(H3N2) (H9XIJ5), H5N1-IAV to the highly pathogenic A/Viet-Nam/1203/2004(H5N1) (Q6DNN3) and IBV to B/Memphis/13/2003 (Q5V8X3). The secondary sequences elements shown over and below the sequences alignment correspond to bat-IAV (PDB access numbers 4WSB) and IBV (4WSA) respectively. The alignment has been done using Clustal X2 [1] and drawn with ESPript [2].

Supplementary Figure 3

Endonuclease activity. An unstructured 52 nt poly-UC RNA (10 µM) [3] was incubated with the wild-type (wt) or E80A mutant of PA-PB1-PB2(1-116) (12 µM) or the N-terminal domain of PA (PA-Nter) (12 µM) for 2 h at 37°C in a final volume of 30 µl. Reactions were done in the presence and/or in the absence of 0.5 mM $MnCl_2$ in the reaction buffer 50 mM Tris-HCl pH 8,5, 25 mM KCl, 2,5 mM NaCl, 10

mM β-mercaptoethanol as previously described [3,4]. Reactions were inhibited by addition of 2 mM 2,4-dioxo-4-phenylbutanoic acid (DPBA) in the same conditions. As controls, the RNA was incubated with 0.5 mM $MnCl_2$ or with 2 mM DPBA or alone (ctrl). Reactions were stopped by adding 20 mM EGTA. The reactions were then loaded on an 8 M urea polyacrylamide gel. After migration, the gel was revealed by methylene blue staining.

Supplementary Figure 4

***De novo* synthesis activity assay.** (**a**) PA-PB1(1-686) and (**b**) PA-PB1-PB2(1-116) polymerases were incubated at 37°C with the 80-mer panhandle RNA, NTPs and radiolabelled [αP^{32}] UTP. Reactions were stopped at 30, 60, 90 and 120 minutes with the addition of EGTA before loading on a 15 % urea PAGE. Revelation was performed using a phosphorus screen and Typhoon scanner (GE Healthcare).

Supplementary Figure 5

Effect of PA-Nter depletion on PA-PB1 oligomerzation. Individual SEC-MALLS chromatography runs of (**a**) PA-PB1(1-686) and (**b**) PA(196-716)-PB1(1-686) have been performed in the same buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 5 mM β-mercaptoethanol) using a S200increase (10/300 GL, GE Healthcare) column. UV absorbance at 280 mm is plotted against the elution volume. MALLS-RI molecular mass estimates are plotted below the curve.

Supplementary Figure 6

RanBP5 and PA-PB1(1-686) does not form a stable complex when mixed. PA-PB1(1-686) and RanBP5 were separately purified before being mixed in equimolar proportions. The mixture was incubated 1h at 20°C before analysis. Individual SEC-MALLS chromatography runs of (**a**) PA-PB1(1-686), (**b**) RanBP5 and (**c**) PA-PB1(1-686) + RanBP5 were performed in the same buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 5 mM β-mercaptoethanol) using a S200increase (10/300 GL, GE Healthcare) column. UV

absorbance at 280 mm is plotted against the elution volume. MALLS-RI molecular mass estimates are plotted below the curve.

Supplementary Figure 7

Rg determination by Guinier extrapolation. Guinier plots were calculated on low q regions and are linear for (**a**) PA-PB1(1-686), (**b**) RanBP5 and (**c**) PA-PB1(1-686)-RanBP5.

Supplementary Figure 8

MONSA multi component fitting. MONSA builds a triple consensus *ab initio* model by simultaneously fitting the LnI(q) curve of (**1**)PA-PB1(1-686), (**2**)RanBP5 and (**3**)PA-PB1(1-686)-RanBP5.

Supplementary Figure 9

The 5'-vRNAp does not dissociate or bind to the PA-PB1(1-686)-RanBP5 complex. Prior to injection on SEC-MALLS, the RdRp-RanBP5 complex was incubated with a 3 fold molar concentration of 5'-vRNAp for 1h at 20°C. SEC-MALLS chromatograms are shown for both RdRp-RanBP5 wihtout (**a**) and with (**b**) 5'-vRNAp. Both analyses were performed in the same buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 5 mM β-mercaptoethanol).

References

1 Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947-2948, doi:10.1093/bioinformatics/btm404 (2007).

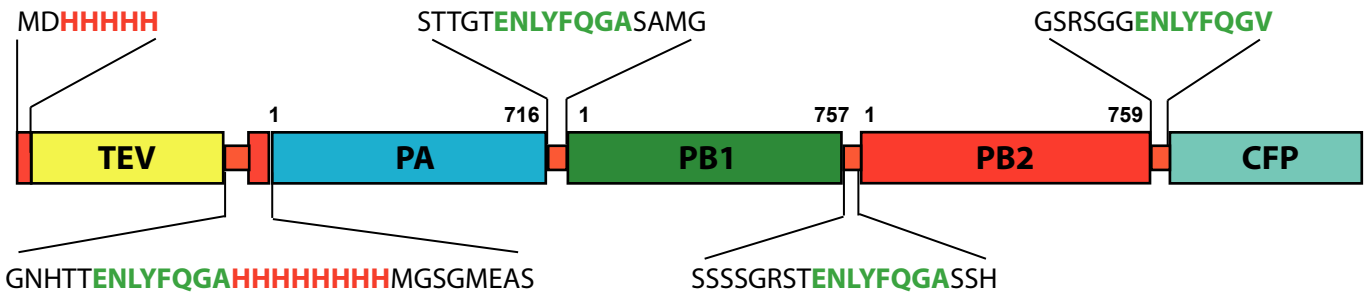
2 Robert, X. & Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* **42**, W320-324, doi:10.1093/nar/gku316 (2014).

3 Dias, A. *et al.* The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit. *Nature* **458**, 914-918 (2009).

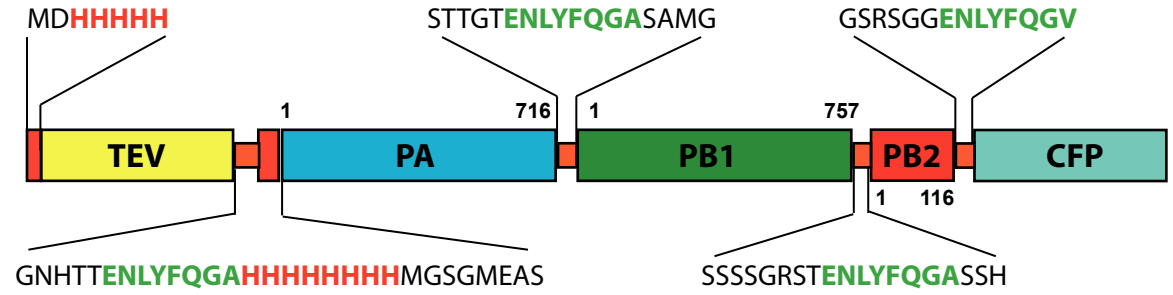
4 Crepin, T. *et al.* Mutational and metal binding analysis of the endonuclease domain of the influenza virus polymerase PA subunit. *J Virol* **84**, 9096-9104 (2010).

Supplementary Figure 1

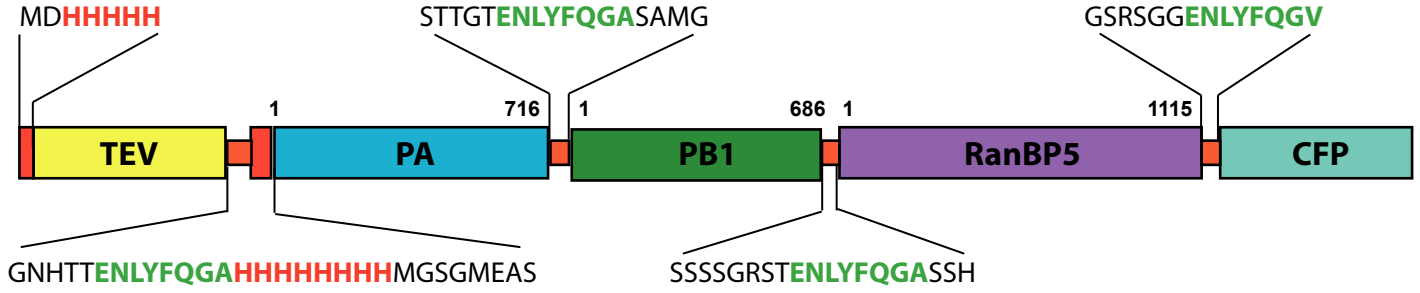
a Construct 12: PA-PB1-PB2



b Construct 15: PA-PB1-PB2(1-116)



c Construct 21: PA-PB1(1-686)-RanBP5

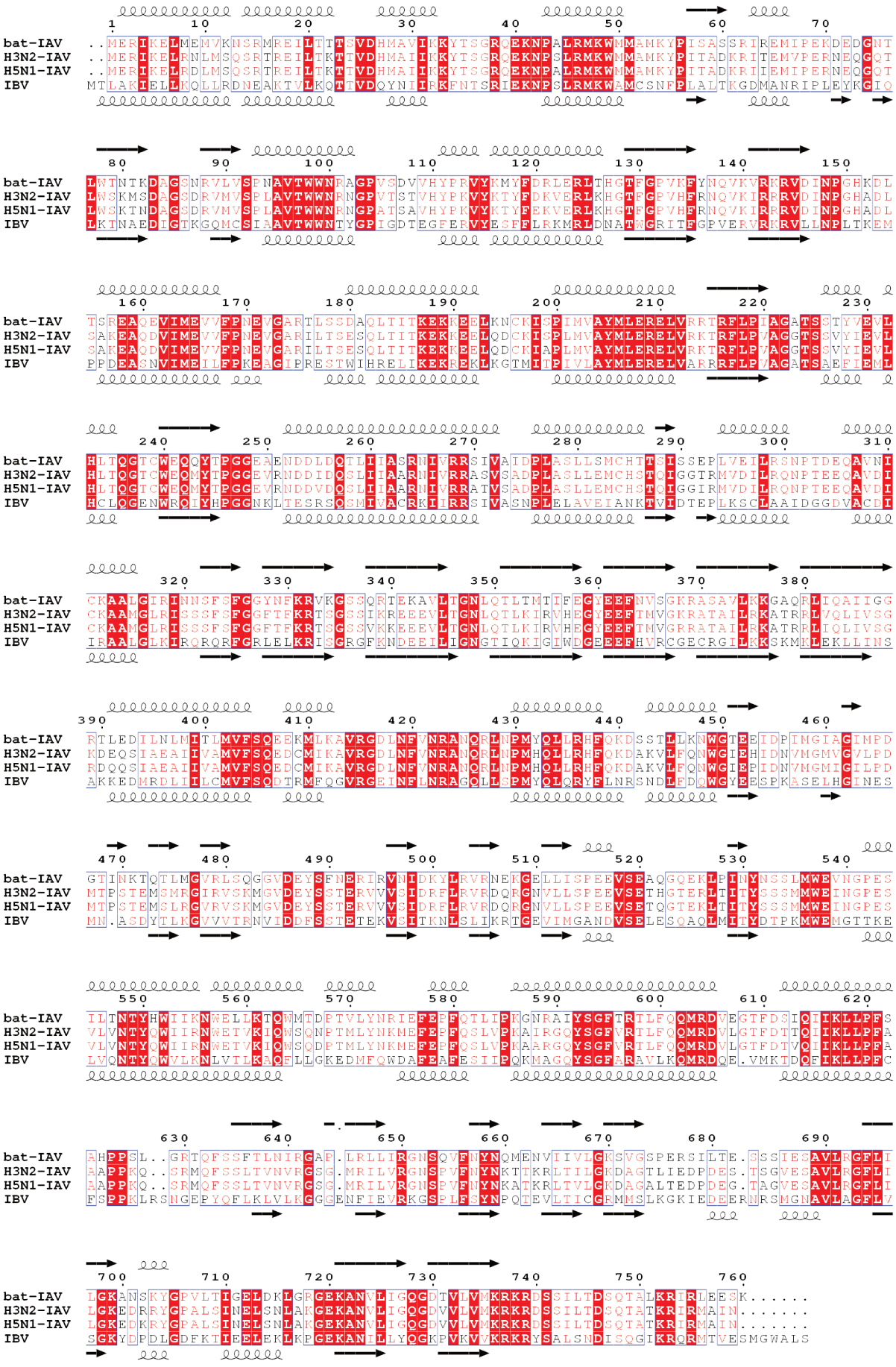


Protein	accession number
TEV	NP_734212.1
PA/H3N2	H9XIJ2
PA/H5N1	Q5EP34
PB1/H3N2	P31341
PB1/H5N1	Q5EP37
PB2/H3N2	H9XIJ5
PB2/H5N1	Q6DNN3
CFP	A0A059PIU2

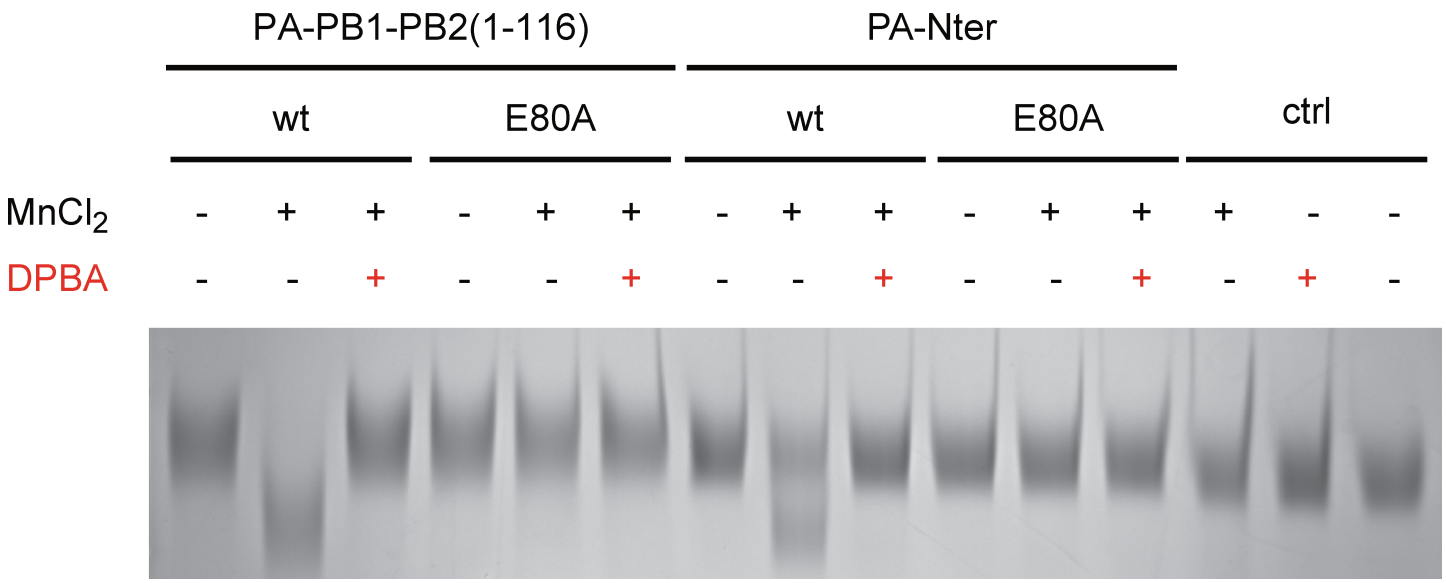
TEV	NP_734212.1
PA/H3N2	H9XIJ2
PA/H5N1	Q5EP34
PB1/H3N2	P31341
PB1/H5N1	Q5EP37
PB2/H3N2	H9XIJ5
PB2/H5N1	Q6DNN3
CFP	A0A059PIU2

TEV	NP_734212.1
PA/H3N2	H9XIJ2
PB1/H3N2	P31341
RanBP5	CAA70103_1
CFP	A0A059PIU2

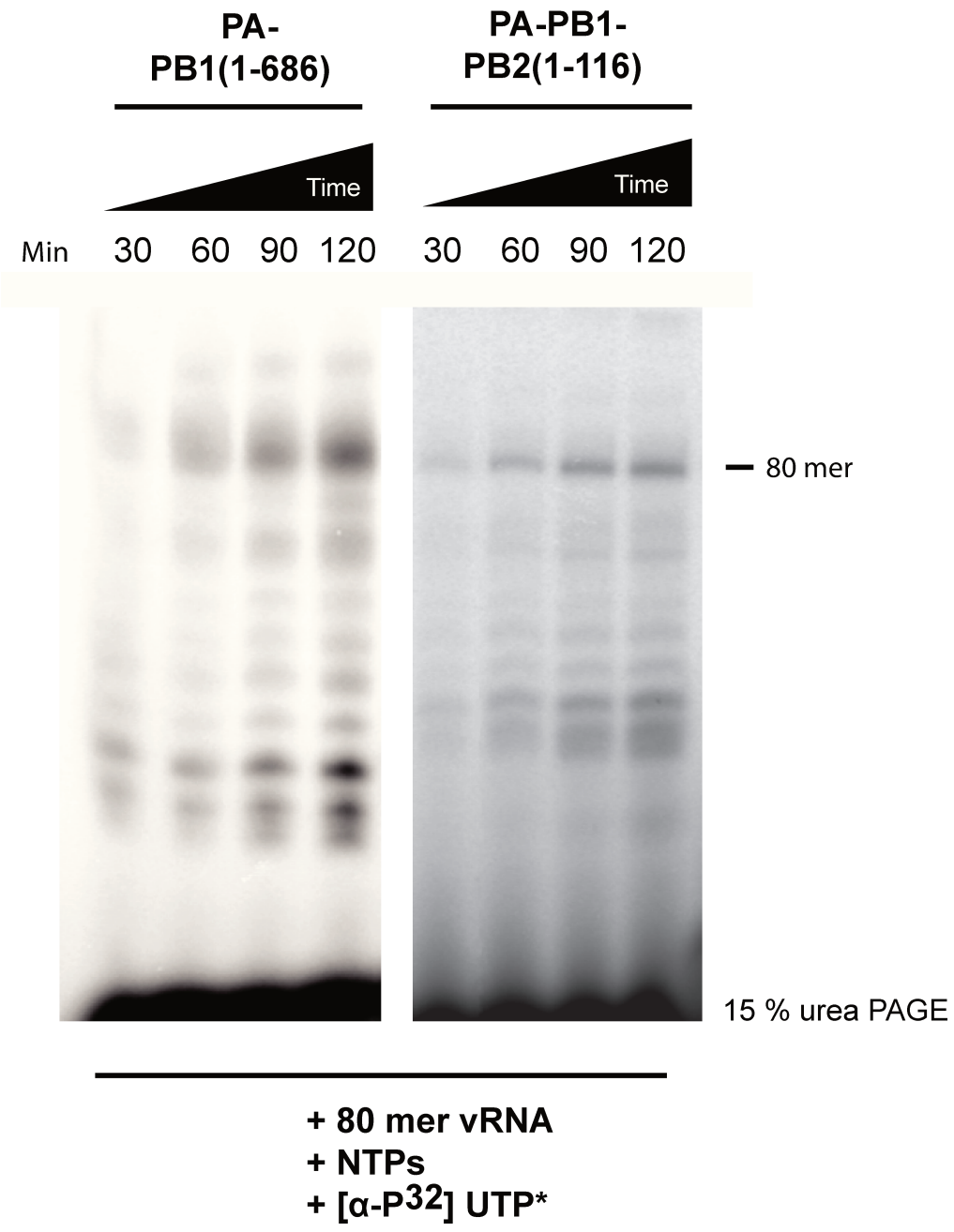
Supplementary Figure 2



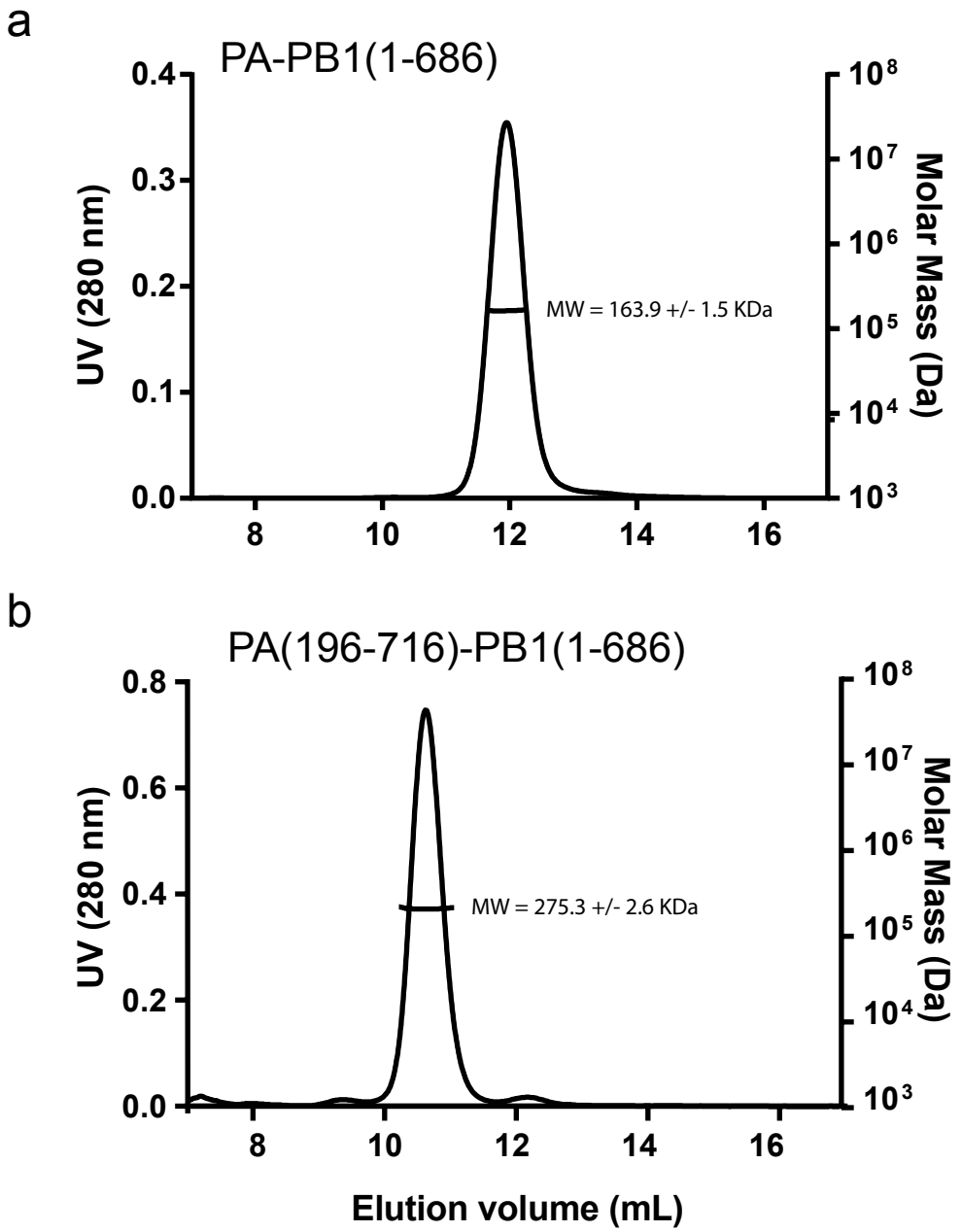
Supplementary Figure 3



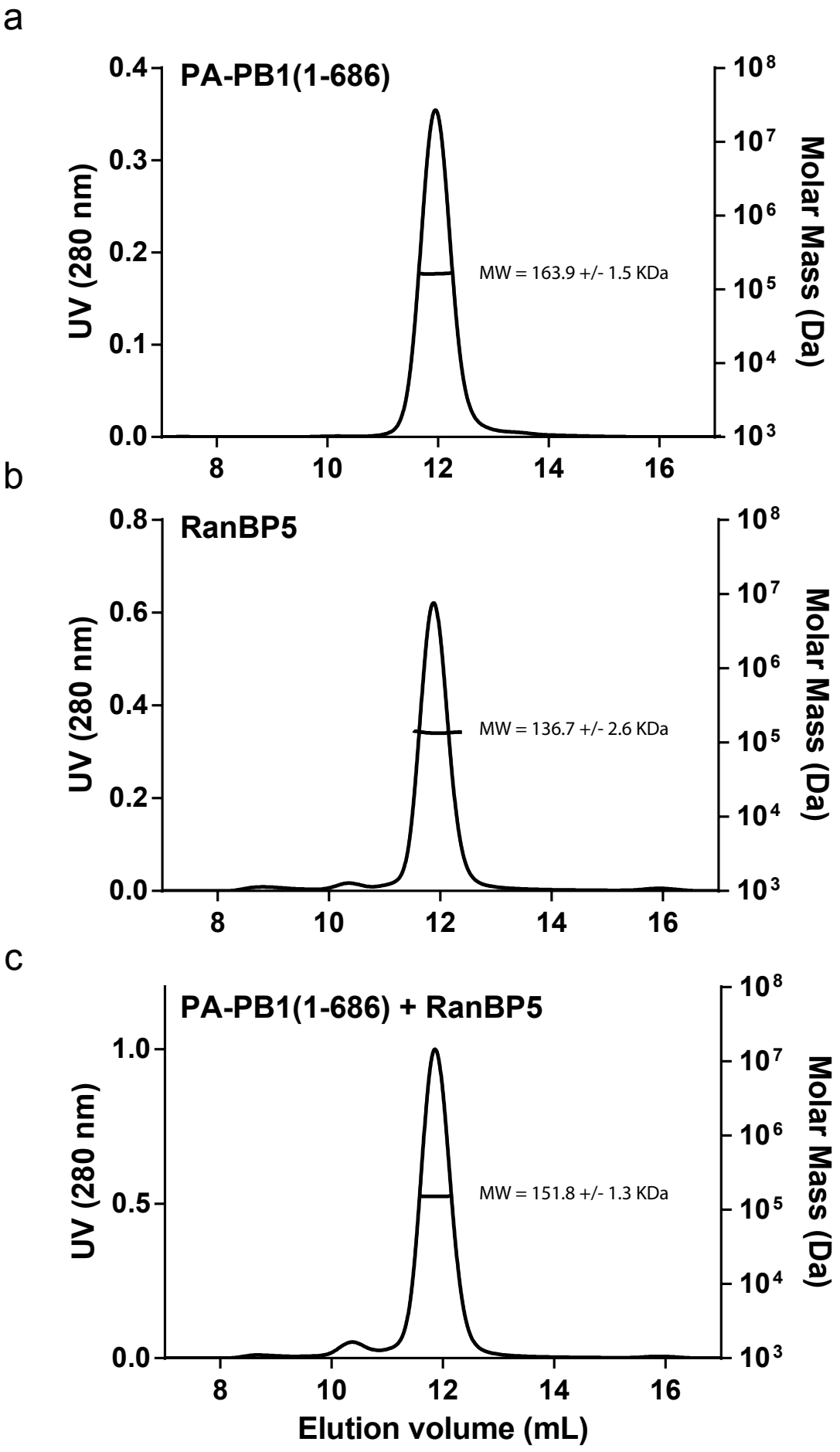
Supplementary Figure 4



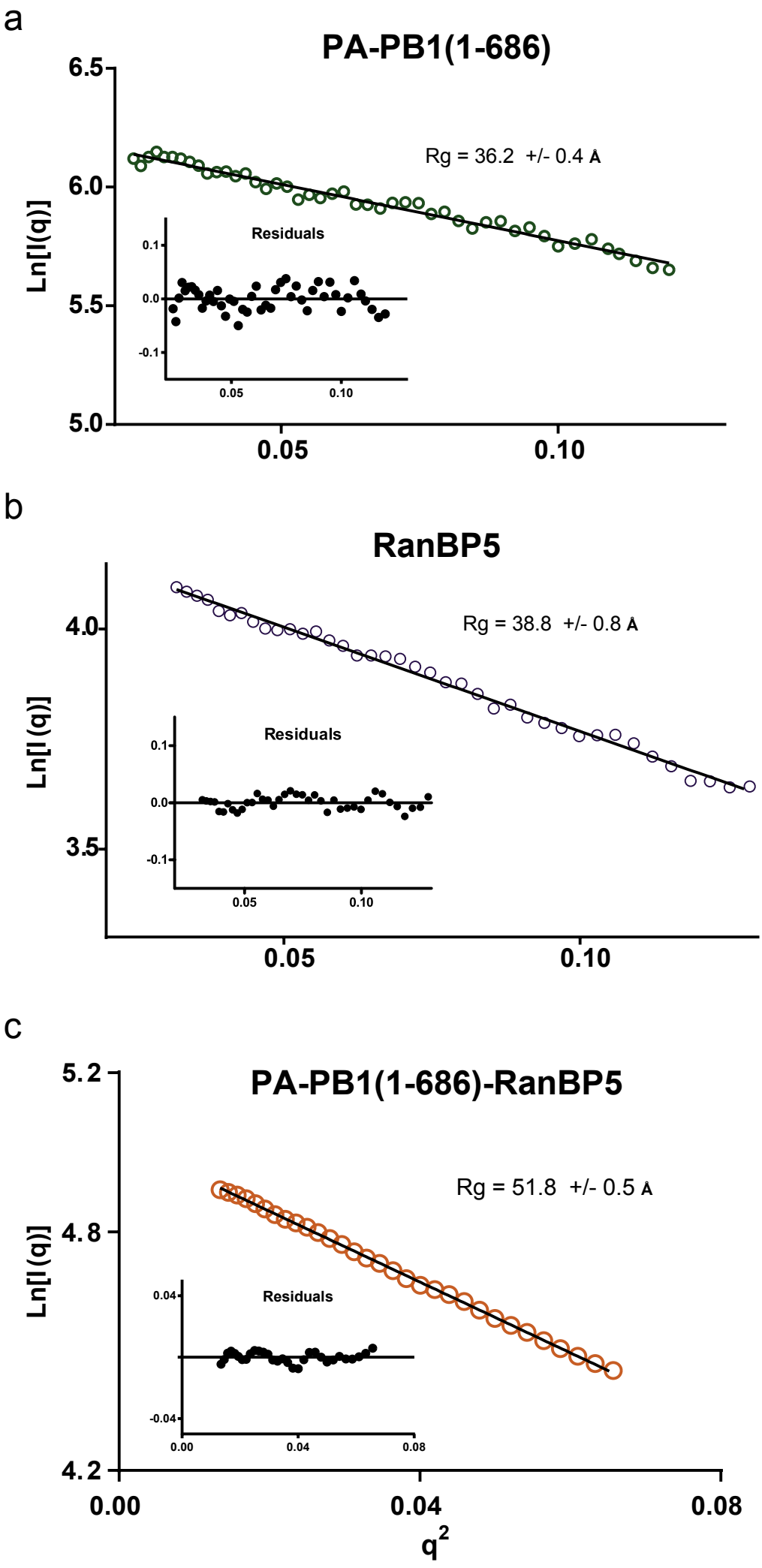
Supplementary Figure 5



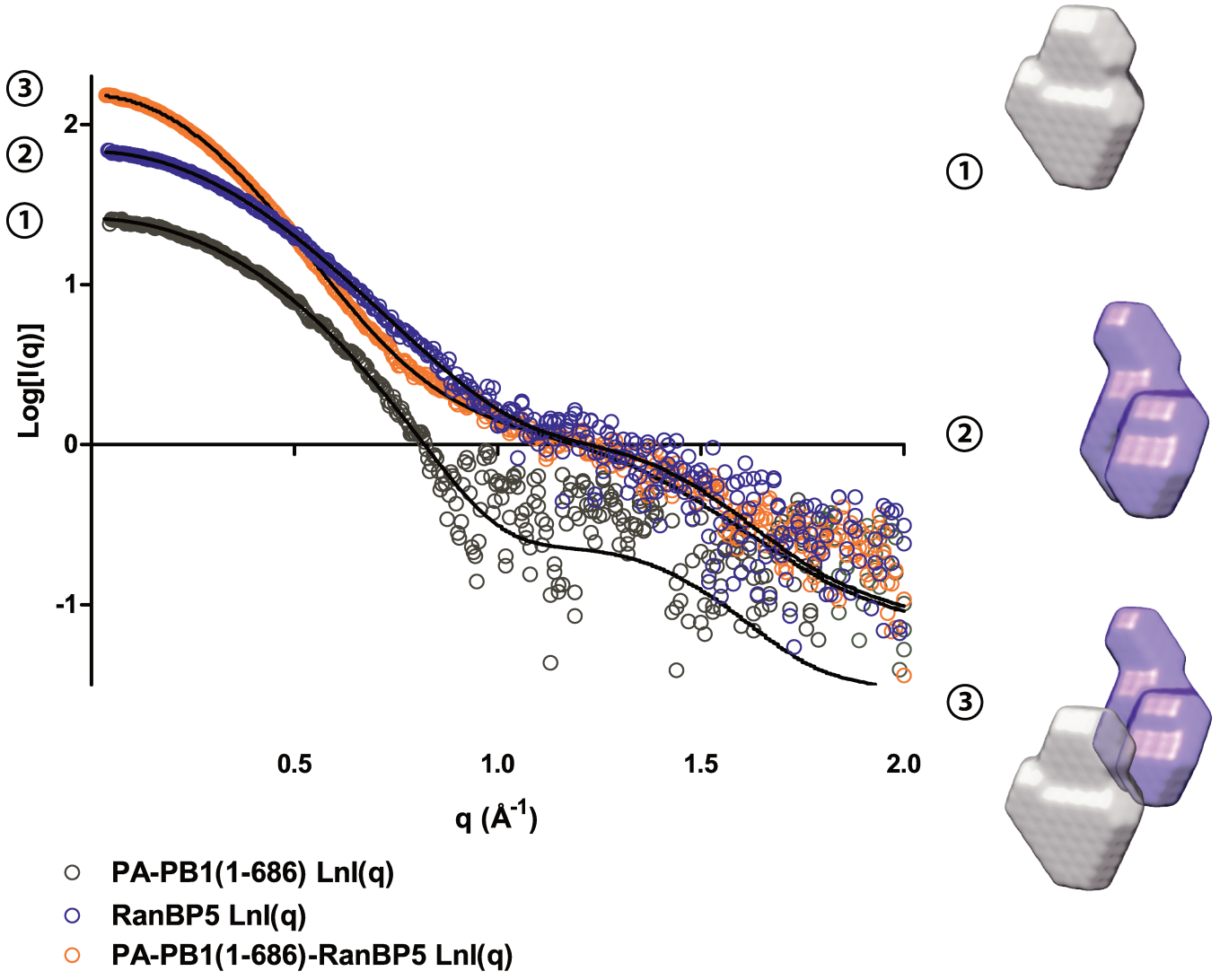
Supplementary Figure 6



Supplementary figure 7

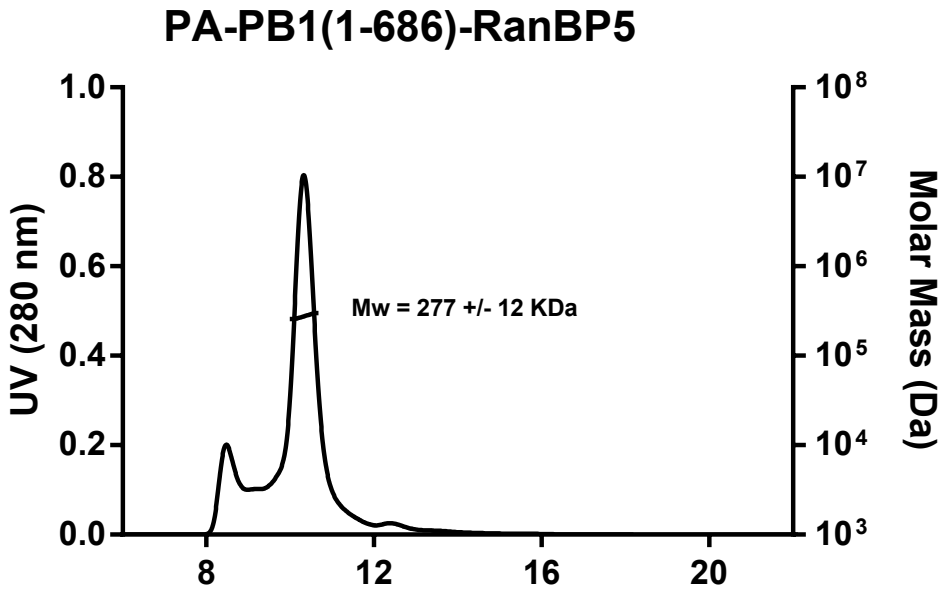


Supplementary figure 8

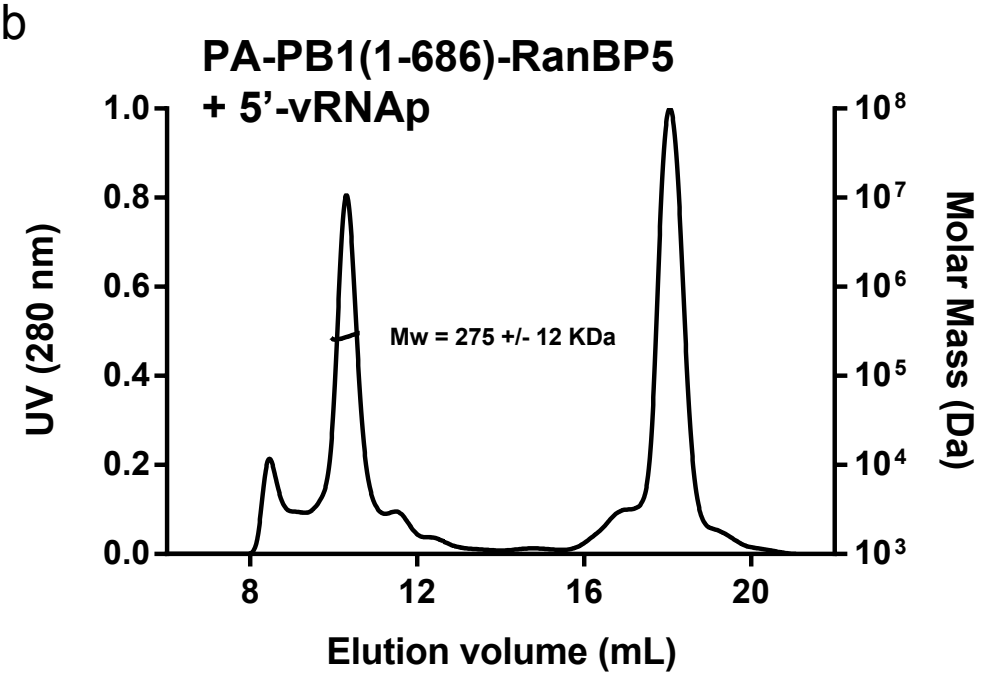


Supplementary Figure 9

a



b



Expert Opinion, 2015

Learning from structure-based drug design and new antivirals targeting the ribonucleoprotein complex for the treatment of influenza

Alexandre Monod, Chritopher Sxale, Bogdan Tarus, Alice Tissot, Bernard Delmas, Rob Ruigrok WH Ruigrok, Thibaut Crépin et Anny Slama-Schwok

EXPERT OPINION

1. Introduction
2. Novel antiviral candidates targeting external viral proteins
3. Structural basis for small-molecules inhibition of individual proteins of the RNP
4. Targeting the non-structural protein 1
5. Conclusion
6. Expert opinion

Expert Opin. Drug Discov. Downloaded from informahealthcare.com by INSERM on 04/10/15
For personal use only.

Learning from structure-based drug design and new antivirals targeting the ribonucleoprotein complex for the treatment of influenza

Alexandre Monod, Christopher Swale, Bogdan Tarus, Alice Tissot, Bernard Delmas, Rob WH Ruigrok, Thibaut Crépin & Anny Slama-Schwok†

†*Virologie et Immunologie Moléculaires, INRA UR, Jouy en Josas, France*

Introduction: Influenza viruses are a threat to human health. There are presently only two methods for treating influenza: vaccines, which require yearly updates, and two classes of antivirals that suffer with the problem of resistance by current human influenza viruses; this is especially the case with amantadine and rimantadine. Consequently, there is an urgent need for the development of new antivirals with new mechanisms of action.

Areas covered: In this review, the authors focus on viral protein domains, their associated activity and their inhibition by small molecules defined by a structure-based design with a special emphasis on the ribonucleoprotein complex and its inhibitors. Several new classes of antiviral candidates targeting viral replication through individual domains of the polymerase and the nucleoprotein (NP) have been developed through structure-based design.

Expert opinion: To date, the antivirals targeting neuraminidase are by far the most developed and potent. Antiviral candidates targeting the NP and polymerase domains are in the pipeline but their pharmacokinetics needs further studies. The recently published structures of the polymerase expand the possibilities for development of new antivirals. Combination therapies targeting conserved viral targets and new cellular proteins or exploiting drug promiscuity hold promises to fight against the emergence of resistance.

Keywords: antivirals, cap-binding domain, endonuclease domain, neuraminidase and matrix 2 inhibitors, nucleoprotein inhibitors, replication and transcription, resistance, ribonucleoprotein complex

Expert Opin. Drug Discov. (2015) ***10****(4):345-371*

1. Introduction

Influenza viruses are important pathogens of humans and animals. In humans, they cause seasonal epidemics, resulting in about 3 – 5 million yearly cases of severe illness and 500,000 deaths [1]. Highly transmissible influenza A viruses can lead to pandemics as it occurred in 1918, 1957, 1968 and 2009. Influenza viruses have a segmented RNA genome, and exchange of segments between different viruses can lead to viruses with new biological characteristics. In particular, when a human virus mixes its genomic segments with those of an animal virus, a new virus may emerge with novel hemagglutinin (HA) and neuraminidase (NA). Such an occasional event called antigenic shift is associated with pandemics, as previously observed in 1957 and 1968, with the emergence of H2N2 and H3N2 lineages in the human

informa healthcare

All rights reserved: reproduction in whole or in part not permitted
RIGHTSLINK

Article highlights.

- There is a need for new antivirals to protect the populations against resistant viral strains to approved antivirals targeting neuraminidase and matrix 2 (M2).
- To overcome resistance, new antiviral candidates targeting the conserved stem region of hemagglutinin and inhibiting its pH-induced structural change are being developed. New structure-based inhibitors of NA and active inhibitors against resistant strains of M2 were reported.
- Within the ribonucleoprotein, inhibitors of the nucleoprotein, modulating its oligomerization and its RNA binding, were developed. Two functional domains of the polymerase were targeted by new antiviral candidates: the PA endonuclease and the PB2 cap-binding domain. PB1–PA interactions were disrupted by novel inhibitors.
- Non-structural protein 1 is a conserved protein that is targeted both at the level of its RNA-binding ability and probably at its effector domain through its interface with cellular factors.
- The new structures of the polymerase hold a lot of possibility for developing new antivirals. Antiviral combinations may probably be the necessary solution for blocking resistance to antivirals.

This box summarizes key points contained in the article.

Expert Opin. Drug Discov. Downloaded from informahealthcare.com by INSERM on 04/10/15
For personal use only.

population, respectively. Antigenic drift results from gradual accumulation of mutations in the surface glycoproteins and leads to annual flu epidemics.

Due to the high mutation rate of the virus, influenza vaccines that promote antibody production against the HA confer protection for no more than a few years. Every year, the World Health Organization estimates which strains of the virus are expected to circulate in the next year, allowing production of vaccines that will provide a strong immunity against the emerging strains. The vaccine is reformulated once a year before the winter periods in the northern and southern hemispheres.

Influenza viruses belong to the Orthomyxoviridae family. Wild aquatic birds constitute the natural host reservoir for a large variety of influenza A viruses, but some influenza viruses infect humans, pigs and other mammals. The type A viruses are the most virulent among the three influenza types. The influenza A viruses can be divided into different serotypes, according to the antibody response against their HAs (H1 – H18) and the NA (N1 to N10), but only H1 – H3 and N1 and N2 are commonly found in humans (H2 was present only from 1957 to 1968). Influenza B almost exclusively infects humans and thus pandemics of influenza B do not occur. Influenza C viruses infect humans, dogs and pigs. Less common than other types, it only causes mild disease in children.

The influenza A genome contains eight RNA segments, each segment encoding one to two main proteins. Whereas three segments encode the HA, the NA and the matrix 2 (M2) proteins, the five others encode the replicative machinery and proteins controlling the host response and the traffic of the genome in the cell compartments. Entry of the particles into the cell is mediated by endocytosis, and fusion between the viral and the endosomal membranes occurs after endosome acidification. Using microtubule and actin-associated motors, viral ribonucleoproteins (RNPs) migrate toward the nucleus to ensure the primary transcription (**Figure 1**). The genomic segments are of negative sense, meaning that they are of opposite sense than mRNA, and transcription is the first activity of the virus after entry in the host cell. Transcription and replication are ensured by the viral RNA-dependent RNA-polymerase (RdRp), a multiprotein complex composed of three subunits, PA, PB1 and PB2 which interacts with both nucleoprotein (NP) and the viral genome [2,3]. While replication and transcription occur in the nucleus of the cell, several viral proteins (NEP and M1) are associated with the export of the RNP complexes from the nucleus toward the viral budding site in rafts at the apical side of the plasma membrane. The non-structural protein 1 (NS1) has multiple roles during the virus cycle, the best recognized role being its ability to downregulate the IFN induction in the infected cell.

What are the targets for available antivirals and their role in the viral cycle? Current antiviral drugs used against influenza are inhibitors of detachment of the viral particle from the host cell, NA inhibitors (oseltamivir and zanamivir) and M2 channel protein inhibitors (adamantane derivatives) [4,5]. Their efficiency remains questionable: American H3N2 circulating strains in 2005 were generally resistant to amantadine and rimantadine.

In this review, focus is made on viral protein domains, their associated activity and their inhibition by small molecules defined by a structure-based design with a special emphasis on the RNP complex and its inhibitors. Influenza antivirals were previously actively reviewed [1,2,4-7]. Nucleic acid-based and vaccine approaches are out of the scope of this review. We will first detail the external proteins HA, NA and M2 and their new inhibitors.

2. Novel antiviral candidates targeting external viral proteins

2.1 Hemagglutinin

The HA is a surface glycoprotein of influenza virus required for viral particle entry. HA is a homotrimeric protein formed by a globular head positioned over a stem domain. Each HA monomer contains two disulfide-linked polypeptides HA1 and HA2 that are derived from the proteolytic cleavage of a precursor by host protease. HA binds to sialic acid receptors of the host cell (**Figure 1** and **Table 1**). Seasonal vaccines provide neutralizing antibodies targeting the highly variable region of HA. Carbohydrate-binding agents that recognize glycosylation sites on HA prevent virus adsorption to the cell [8]. Besides its activity as binding to sialic acid receptors,

RIGHTSLINK

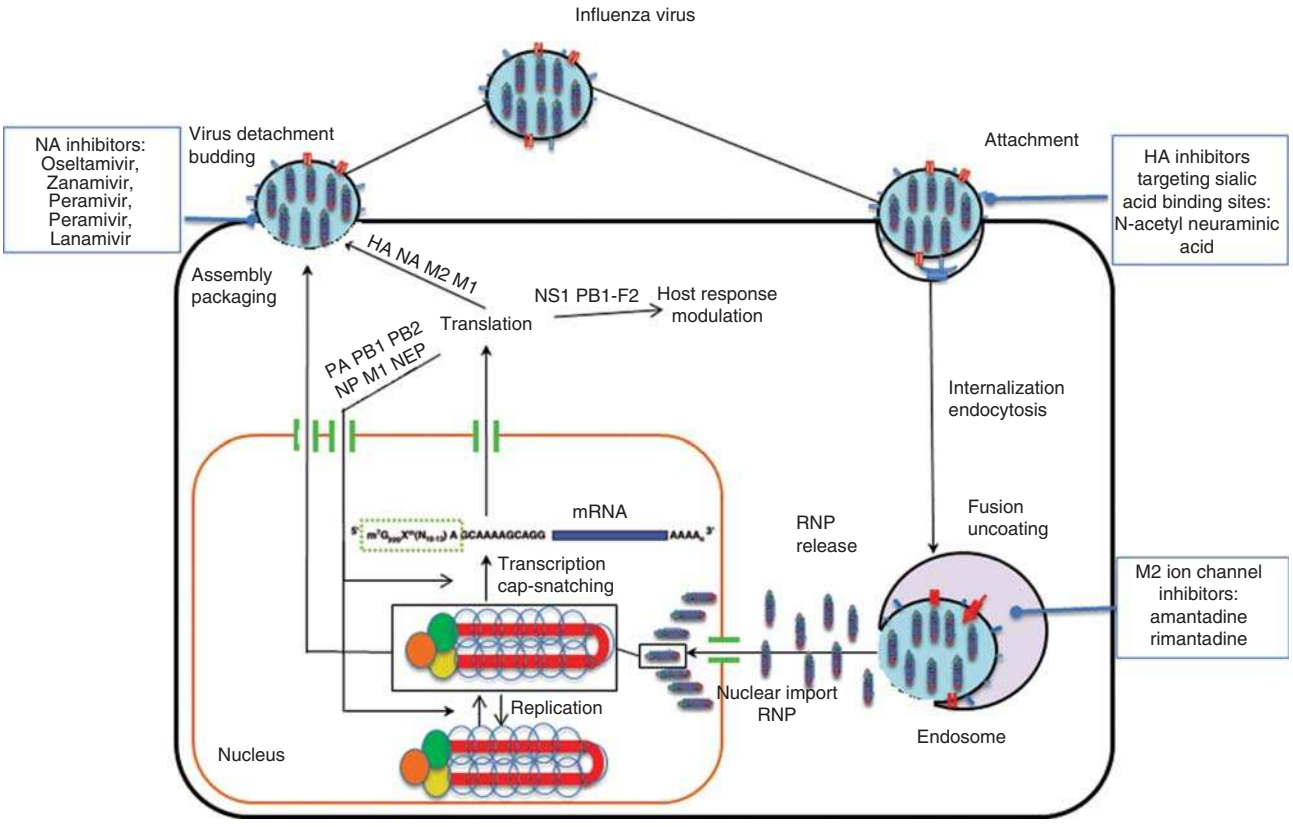


Figure 1. Illustration showing influenza virus' lifecycle: influenza viral particle attaches to the host cell (represented by a black rectangle) by binding to sialic acid receptors mediated by the hemagglutinin Internalization of the particle is mediated by endocytosis. Following internalization, endosomal acidification initiates conformational changes in HA, driving fusion of virus and cell membranes. The rupture of the membrane of the endosome releases the viral RNP complexes of each of the eight segments of genomic RNA, whereas acidification of the virion interior by the M2 proton channel drives the dissociation of M1 from the viral genome. Nuclear import of the RNP mediated by the cellular transport machinery allows transcription to begin in the nucleus involving a specific cap-snatching mechanism and the formation of a capped messenger RNA. The mRNAs are exported to the cytoplasm and translated into proteins by cellular ribosomes. Newly translated proteins are transported to the nucleus (PA, PB1, PB2, NP, M1, NEP). Formation of new progeny requires replication to take place in the nucleus by the RNP followed by translation. Progeny viral RNP are then exported to the cytoplasm with the assistance of M1 and NEP and assembly of novel viral particles containing HA, NA, M2 and M1 and their budding at the cell surface that requires the NA to detach. The figures also shows the presently available antivirals targeting M2, NA or HA, the latter being targeted by vaccines. HA: Hemagglutinin; M1: Matrix protein 1; M2: Matrix 2; NA: Neuraminidase; NEP: Nuclear export protein; RNP: Ribonucleoprotein.

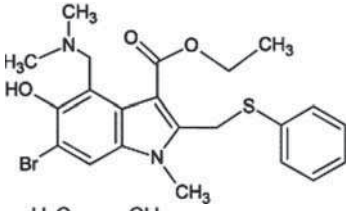
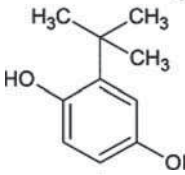
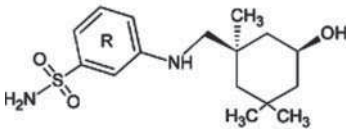
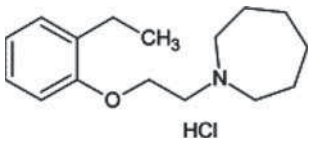
HA, together with the M2 proton channel, participates in virus entry by endocytosis through fusion of virus and cell membranes and release of the viral RNP complexes from the acidified endocytic particle. This activity is inhibited by a variety of small molecules that prevent the fusogenic activity by blocking the irreversible conformational change of HA induced by acidic pH. The drug arbidol is representative of such a class of antivirals which inhibits HA-mediated membrane fusion by increasing the stability of the HA to low pH [9,10]. The advantage of arbidol also resides in its broad spectrum activity not only against influenza A and B but also against other viruses such as respiratory syncytial virus (RSV), parainfluenza virus, Coxsackievirus, rhinovirus, hepatitis B virus and hepatitis C virus. This broad spectrum activity may also be enhanced by arbidol interactions with additional targets as phospholipids of the cell membrane. Binding of small molecules as *t*-butyl benzene-1,4 diol at the interface between two HA monomers likely inhibit HA by avoiding structural changes required for membrane fusion [11]. Structural optimization of salicylamide-based HA inhibitor led to the design of benzenesulfonamide derivatives that stabilized the pre-fusion of HA structure [12]. Inhibition of conserved sites in the stem region of HA holds promise for broad spectrum protective effects [13]. Small-molecule inhibitors derived from aminoalkyl phenol or sulfonamide scaffolds targeting pockets in the stem region were recently reported [14].

Expert Opin. Drug Discov. Downloaded from informahealthcare.com by INSERM on 04/10/15
For personal use only.

RIGHTSLINK

Expert Opin. Drug Discov. Downloaded from informahealthcare.com by INSERM on 04/10/15
For personal use only.

Table 1. External proteins of influenza virus and their inhibitors binding sites*.

Protein	Function	Structure	Binding site/ PDB	Inhibitor/status	IC_{50} or Kd strain/CC_{50}	Ref.
HA	Surface glycoprotein associated with membrane fusion		Interface between 2 HA in trimer	Arbidol/ approved in Russia and China	IC_{50} = 7 – 21 μM H2N2, H7N7 $CC_{50} \geq 120\mu M$	Leneva *et al.* (2009) [9,10]
			3EYK	Inhibitor of membrane fusion- *t*-butyl benzene-1,4 diol	nd	Russel *et al.* (2008) [11]
			HA2 in HA trimer	Benzenesulfonamide derivatives‡/preclinical	IC_{50} = 86 – 210 nM H1N1 CC_{50} = 71 – > 100 μM Kd = 200 nM	Tang *et al.* (2011) [12]
		–	Stem region 3R2X	Protein		
			Stem region	MBX2329§	IC_{50} = 0.3 – 5 μM (H1N1-resistant, H5N1); no activity against H3N2	Fleishman *et al.* (2011) [13] Basu *et al.* (2014) [14]

A = (2E,5S,9R,10S)-10-(acetylamino)-2-imino-4-oxo-9-(pentan-3-yloxy)-1-thia-3-azaspiro[4.5]dec-6-ene-7-carboxylic acid.

B = (5R,9R,10S)-10-(acetylamino)-2-amino-4-oxo-9-(pentan-3-yloxy)-1-thia-3-azaspiro[4.5]deca-2,6-diene-7-carboxylic acid.

C = (3S,4R,5R)-4-(acetylamino)-3-amino-5-(pentan-3-yloxy)cyclohex-1-ene-1-carboxylic acid.

D = (3S,5S,7S)-N-{[5-(thiophen-2-yl)-1,2-oxazol-3-yl]methyl}tricyclo[3.3.1.1 ~ 3,7 ~]decan-1-aminium.

*Only part of the inhibitor structures that are found in the PDB and those involving a structure-based design combined with experimental validation are summarized here.

‡The derivatives exhibit good water solubility and good stability in human liver microsomes (6 – 8 ml/min/kg) and no inhibition of major CYP450 isoforms.

§Both compounds exhibit strong synergy with oseltamivir.

¶These NA inhibitors currently approved in the US are only FDA-approved for uncomplicated influenza infections [72]. Peramivir is in clinical trials in US. p09N1: pandemic H1N1 2009, N5: H12N5.

#Laninamivir exists as an octanoate prodrug that is slowly processed in adult patients; the prodrug binds as oseltamivir and its efficacy is not improved as compared to oseltamivir against resistant strain. N5: H12N5.

**AV580: (3R,4R,5S)-5-[(diaminomethylene)amino]-3-(1-ethylpropoxy)-4-[(fluoroacetyl)amino] cyclohex-1-ene-1-carboxylic acid R = CF_3 W = 4,5-dihydro-1H-imidaz-2-yl. AV5080 has good solubility and metabolic properties and a favorable pharmacokinetics profile in animals.

‡‡The pharmacokinetics properties of the 2,3-difluorosialic acid (DFSA) derivative FaxGuDFSA were comparable to those of zanamivir, a 100% efficacy at 10 mg/kg/day with no side effects in infected mice (H3N2) [21].

§§S31N-resistant strain arising from wsn in Ref. 3.

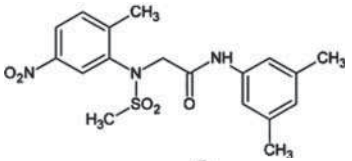
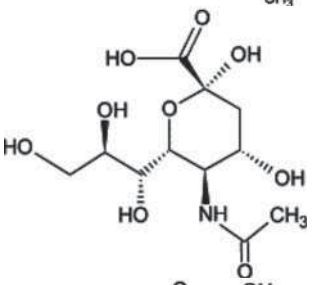
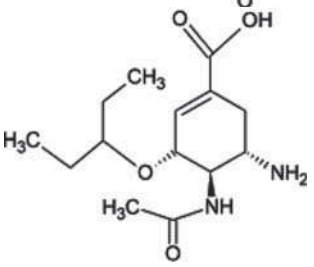
¶¶The full chemical names of the compounds represented in the third column are listed above.

HA: Hemagglutinin; M2: Matrix 2; NA: Neuraminidase; nd: Not determined; PDB: Protein Data Bank; WSN: Influenza A/WSN/33 H1N1.

RIGHTSLINK

Expert Opin. Drug Discov. Downloaded from informahealthcare.com by INSERM on 04/10/15
For personal use only.

Table 1. External proteins of influenza virus and their inhibitors binding sites* (continued).

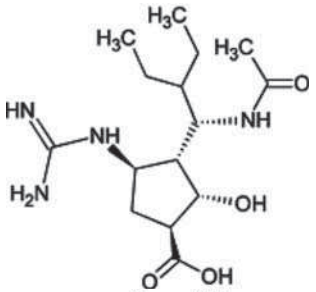
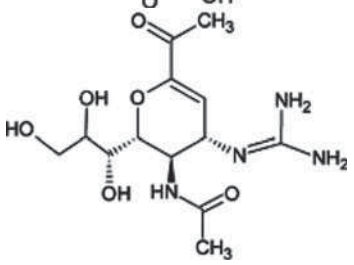
Protein	Function	Structure	Binding site/ PDB	Inhibitor/status	IC_{50} or Kd strain/CC_{50}	Ref.
				MBX2346/preclinical	CC_{50} > 100 μM	
NA	Surface glycoprotein releasing progeny viral particles			Sialic acid		
			Active site '150 cavity' 2HT7, 2HTU	Oseltamivir¶, /approved	IC_{50} = 0.8 – 0.5 nM N5 p09N1; CC_{50} > 35 μM	Russell *et al.* (2006) [16]

A = (2E,5S,9R,10S)-10-(acetylamino)-2-imino-4-oxo-9-(pentan-3-yloxy)-1-thia-3-azaspiro[4.5]dec-6-ene-7-carboxylic acid.
B = (5R,9R,10S)-10-(acetylamino)-2-amino-4-oxo-9-(pentan-3-yloxy)-1-thia-3-azaspiro[4.5]deca-2,6-diene-7-carboxylic acid.
C = (3S,4R,5R)-4-(acetylamino)-3-amino-5-(pentan-3-yloxy)cyclohex-1-ene-1-carboxylic acid.
D = (3S,5S,7S)-N-{[5-(thiophen-2-yl)-1,2-oxazol-3-yl]methyl}tricyclo[3.3.1.1 ~ 3,7 ~]decan-1-aminium.
*Only part of the inhibitor structures that are found in the PDB and those involving a structure-based design combined with experimental validation are summarized here.
‡The derivatives exhibit good water solubility and good stability in human liver microsomes (6 – 8 ml/min/kg) and no inhibition of major CYP450 isoforms.
§Both compounds exhibit strong synergy with oseltamivir.
¶These NA inhibitors currently approved in the US are only FDA-approved for uncomplicated influenza infections [72]. Peramivir is in clinical trials in US. p09N1: pandemic H1N1 2009, N5: H12N5.
#Laninamivir exists as an octanoate prodrug that is slowly processed in adult patients; the prodrug binds as oseltamivir and its efficacy is not improved as compared to oseltamivir against resistant strain. N5: H12N5.
**AV580: (3R,4R,5S)-5-[(diaminomethylene)amino]-3-(1-ethylpropoxy)-4-[(fluoroacetyl)amino] cyclohex-1-ene-1-carboxylic acid R = CF_3 W = 4,5-dihydro-1H-imidaz-2-yl. AV5080 has good solubility and metabolic properties and a favorable pharmacokinetics profile in animals.
‡‡The pharmacokinetics properties of the 2,3-difluorosialic acid (DFSA) derivative FaxGuDFSA were comparable to those of zanamivir, a 100% efficacy at 10 mg/kg/day with no side effects in infected mice (H3N2) [21].
§§S31N-resistant strain arising from wsn in Ref. 3.
¶¶The full chemical names of the compounds represented in the third column are listed above.
HA: Hemagglutinin; M2: Matrix 2; NA: Neuraminidase; nd: Not determined; PDB: Protein Data Bank; WSN: Influenza A/WSN/33 H1N1.

RIGHTSLINK

Expert Opin. Drug Discov. Downloaded from informahealthcare.com by INSERM on 04/10/15
For personal use only.

Table 1. External proteins of influenza virus and their inhibitors binding sites* (continued).

Protein	Function	Structure	Binding site/ PDB	Inhibitor/status	IC_{50} or Kd strain/CC_{50}	Ref.
			active site 2HTU	Peramivir/approved Japan	IC_{50} = 3.4 nM	
			Active site 3CKZ	Zanamivir/approved	IC_{50} = 0.6 – 1.1 nM N5 p09N1; CC_{50} > 30 μM	

A = (2E,5S,9R,10S)-10-(acetylamino)-2-imino-4-oxo-9-(pentan-3-yloxy)-1-thia-3-azaspiro[4.5]dec-6-ene-7-carboxylic acid.

B = (5R,9R,10S)-10-(acetylamino)-2-amino-4-oxo-9-(pentan-3-yloxy)-1-thia-3-azaspiro[4.5]deca-2,6-diene-7-carboxylic acid.

C = (3S,4R,5R)-4-(acetylamino)-3-amino-5-(pentan-3-yloxy)cyclohex-1-ene-1-carboxylic acid.

D = (3S,5S,7S)-N-{[5-(thiophen-2-yl)-1,2-oxazol-3-yl]methyl}tricyclo[3.3.1.1 ~ 3,7 ~]decan-1-aminium.

*Only part of the inhibitor structures that are found in the PDB and those involving a structure-based design combined with experimental validation are summarized here.

‡The derivatives exhibit good water solubility and good stability in human liver microsomes (6 – 8 ml/min/kg) and no inhibition of major CYP450 isoforms.

§Both compounds exhibit strong synergy with oseltamivir.

¶These NA inhibitors currently approved in the US are only FDA-approved for uncomplicated influenza infections [72]. Peramivir is in clinical trials in US. p09N1: pandemic H1N1 2009, N5: H12N5.

#Laninamivir exists as an octanoate prodrug that is slowly processed in adult patients; the prodrug binds as oseltamivir and its efficacy is not improved as compared to oseltamivir against resistant strain. N5: H12N5.

**AV580: (3R,4R,5S)-5-[(diaminomethylene)amino]-3-(1-ethylpropoxy)-4-[(fluoroacetyl)amino] cyclohex-1-ene-1-carboxylic acid R = CF_3 W = 4,5-dihydro-1H-imidaz-2-yl. AV5080 has good solubility and metabolic properties and a favorable pharmacokinetics profile in animals.

‡‡The pharmacokinetics properties of the 2,3-difluorosialic acid (DFSA) derivative FaxGuDFSA were comparable to those of zanamivir, a 100% efficacy at 10 mg/kg/day with no side effects in infected mice (H3N2) [21].

§§S31N-resistant strain arising from wsn in Ref. 3.

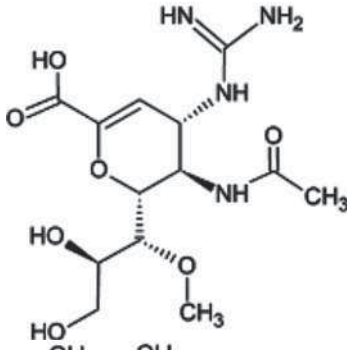
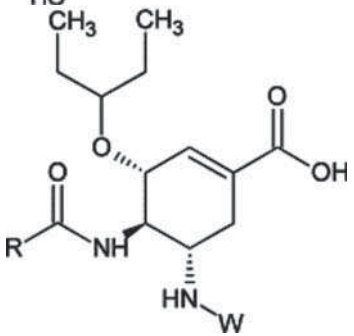
¶¶The full chemical names of the compounds represented in the third column are listed above.

HA: Hemagglutinin; M2: Matrix 2; NA: Neuraminidase; nd: Not determined; PDB: Protein Data Bank; WSN: Influenza A/WSN/33 H1N1.

RIGHTSLINK

Expert Opin. Drug Discov. Downloaded from informahealthcare.com by INSERM on 04/10/15
For personal use only.

Table 1. External proteins of influenza virus and their inhibitors binding sites* (continued).

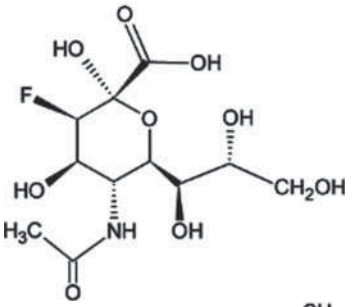
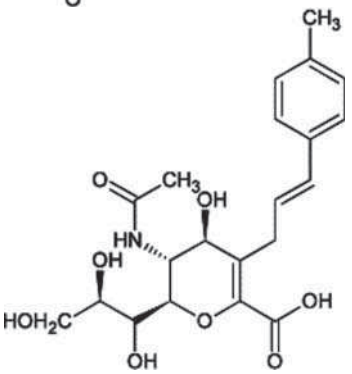
Protein	Function	Structure	Binding site/ PDB	Inhibitor/status	IC_{50} or Kd strain/CC_{50}	Ref.
			Active site 3TI8	Laninamivir[#]/approved in Japan	IC_{50} = 0.9 –1.8 nM N5 p09N1	Vavricka (2011) [18]
			Active site	Tight inhibitor AV5080**/preclinical	IC_{50} = 0.03 – 0.06 nM (H5N1 p09 N1); CC_{50} > 690μM IC_{50} = 0.7 – 1 μM p09N1 H5N1	Ivachtchenko *et al.* (2014) [19]

A = (2E,5S,9R,10S)-10-(acetylamino)-2-imino-4-oxo-9-(pentan-3-yloxy)-1-thia-3-azaspiro[4.5]dec-6-ene-7-carboxylic acid.
B = (5R,9R,10S)-10-(acetylamino)-2-amino-4-oxo-9-(pentan-3-yloxy)-1-thia-3-azaspiro[4.5]deca-2,6-diene-7-carboxylic acid.
C = (3S,4R,5R)-4-(acetylamino)-3-amino-5-(pentan-3-yloxy)cyclohex-1-ene-1-carboxylic acid.
D = (3S,5S,7S)-N-{[5-(thiophen-2-yl)-1,2-oxazol-3-yl]methyl}tricyclo[3.3.1.1 ~ 3,7 ~]decan-1-aminium.
*Only part of the inhibitor structures that are found in the PDB and those involving a structure-based design combined with experimental validation are summarized here.
[‡]The derivatives exhibit good water solubility and good stability in human liver microsomes (6 – 8 ml/min/kg) and no inhibition of major CYP450 isoforms.
[§]Both compounds exhibit strong synergy with oseltamivir.
[¶]These NA inhibitors currently approved in the US are only FDA-approved for uncomplicated influenza infections [72]. Peramivir is in clinical trials in US. p09N1: pandemic H1N1 2009, N5: H12N5.
[#]Laninamivir exists as an octanoate prodrug that is slowly processed in adult patients; the prodrug binds as oseltamivir and its efficacy is not improved as compared to oseltamivir against resistant strain. N5: H12N5.
**AV580: (3R,4R,5S)-5-[(diaminomethylene)amino]-3-(1-ethylpropoxy)-4-[(fluoroacetyl)amino] cyclohex-1-ene-1-carboxylic acid R = CF_3 W = 4,5-dihydro-1H-imidaz-2-yl. AV5080 has good solubility and metabolic properties and a favorable pharmacokinetics profile in animals.
[‡‡]The pharmacokinetics properties of the 2,3-difluorosialic acid (DFSA) derivative FaxGuDFSA were comparable to those of zanamivir, a 100% efficacy at 10 mg/kg/day with no side effects in infected mice (H3N2) [21].
[§§]S31N-resistant strain arising from wsn in Ref. 3.
[¶¶]The full chemical names of the compounds represented in the third column are listed above.
HA: Hemagglutinin; M2: Matrix 2; NA: Neuraminidase; nd: Not determined; PDB: Protein Data Bank; WSN: Influenza A/WSN/33 H1N1.

RIGHTSLINK

Expert Opin. Drug Discov. Downloaded from informahealthcare.com by INSERM on 04/10/15
For personal use only.

Table 1. External proteins of influenza virus and their inhibitors binding sites* (continued).

Protein	Function	Structure	Binding site/ PDB	Inhibitor/status	IC_{50} or Kd strain/CC_{50}	Ref.
			4H52	Covalent inhibitor‡‡ /preclinical	IC_{50} = 1.7 – 7.7 μM; p09N1 H5N1	Vavricka *et al.* (2013) [20] Kim *et al.* (2013) [21]
			3O9K	C3-subsituted sialic acid derivatives/preclinical	K_i = 0.46 nM H1N1, H3N2	Rudrawar *et al.* (2010) [22]

A = (2E,5S,9R,10S)-10-(acetylamino)-2-imino-4-oxo-9-(pentan-3-yloxy)-1-thia-3-azaspiro[4.5]dec-6-ene-7-carboxylic acid.

B = (5R,9R,10S)-10-(acetylamino)-2-amino-4-oxo-9-(pentan-3-yloxy)-1-thia-3-azaspiro[4.5]deca-2,6-diene-7-carboxylic acid.

C = (3S,4R,5R)-4-(acetylamino)-3-amino-5-(pentan-3-yloxy)cyclohex-1-ene-1-carboxylic acid.

D = (3S,5S,7S)-N-{[5-(thiophen-2-yl)-1,2-oxazol-3-yl]methyl}tricyclo[3.3.1.1 ~ 3,7 ~]decan-1-aminium.

*Only part of the inhibitor structures that are found in the PDB and those involving a structure-based design combined with experimental validation are summarized here.

‡The derivatives exhibit good water solubility and good stability in human liver microsomes (6 – 8 ml/min/kg) and no inhibition of major CYP450 isoforms.

§Both compounds exhibit strong synergy with oseltamivir.

¶These NA inhibitors currently approved in the US are only FDA-approved for uncomplicated influenza infections [72]. Peramivir is in clinical trials in US. p09N1: pandemic H1N1 2009, N5: H12N5.

#Laninamivir exists as an octanoate prodrug that is slowly processed in adult patients; the prodrug binds as oseltamivir and its efficacy is not improved as compared to oseltamivir against resistant strain. N5: H12N5.

**AV580: (3R,4R,5S)-5-[(diaminomethylene)amino]-3-(1-ethylpropoxy)-4-[(fluoroacetyl)amino] cyclohex-1-ene-1-carboxylic acid R = CF_3 W = 4,5-dihydro-1H-imidaz-2-yl. AV5080 has good solubility and metabolic properties and a favorable pharmacokinetics profile in animals.

‡‡The pharmacokinetics properties of the 2,3-difluorosialic acid (DFSA) derivative FaxGuDFSA were comparable to those of zanamivir, a 100% efficacy at 10 mg/kg/day with no side effects in infected mice (H3N2) [21].

§§S31N-resistant strain arising from wsn in Ref. 3.

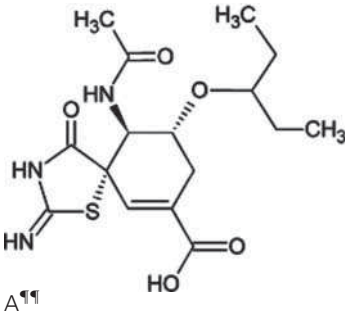
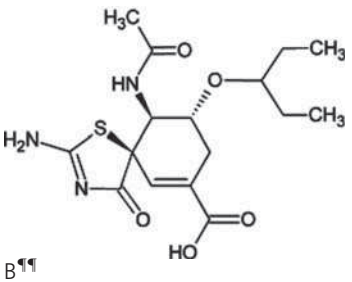
¶¶The full chemical names of the compounds represented in the third column are listed above.

HA: Hemagglutinin; M2: Matrix 2; NA: Neuraminidase; nd: Not determined; PDB: Protein Data Bank; WSN: Influenza A/WSN/33 H1N1.

RIGHTSLINK

Expert Opin. Drug Discov. Downloaded from informahealthcare.com by INSERM on 04/10/15
For personal use only.

Table 1. External proteins of influenza virus and their inhibitors binding sites* (continued).

Protein	Function	Structure	Binding site/ PDB	Inhibitor/status	IC_{50} or Kd strain/CC_{50}	Ref.
		 A¶¶	4MJV	Spirolactam derivatives (preclinical)	K_i = 1.5 nM, H1N1	Mohan *et al.* (2014) [23]
		 B¶¶	4MJU			

A = (2E,5S,9R,10S)-10-(acetylamino)-2-imino-4-oxo-9-(pentan-3-yloxy)-1-thia-3-azaspiro[4.5]dec-6-ene-7-carboxylic acid.

B = (5R,9R,10S)-10-(acetylamino)-2-amino-4-oxo-9-(pentan-3-yloxy)-1-thia-3-azaspiro[4.5]deca-2,6-diene-7-carboxylic acid.

C = (3S,4R,5R)-4-(acetylamino)-3-amino-5-(pentan-3-yloxy)cyclohex-1-ene-1-carboxylic acid.

D = (3S,5S,7S)-N-{[5-(thiophen-2-yl)-1,2-oxazol-3-yl]methyl}tricyclo[3.3.1.1 ~ 3,7 ~]decan-1-aminium.

*Only part of the inhibitor structures that are found in the PDB and those involving a structure-based design combined with experimental validation are summarized here.

‡The derivatives exhibit good water solubility and good stability in human liver microsomes (6 – 8 ml/min/kg) and no inhibition of major CYP450 isoforms.

§Both compounds exhibit strong synergy with oseltamivir.

¶These NA inhibitors currently approved in the US are only FDA-approved for uncomplicated influenza infections [72]. Peramivir is in clinical trials in US. p09N1: pandemic H1N1 2009, N5: H12N5.

#Laninamivir exists as an octanoate prodrug that is slowly processed in adult patients; the prodrug binds as oseltamivir and its efficacy is not improved as compared to oseltamivir against resistant strain. N5: H12N5.

**AV580: (3R,4R,5S)-5-[(diaminomethylene)amino]-3-(1-ethylpropoxy)-4-[(fluoroacetyl)amino] cyclohex-1-ene-1-carboxylic acid R = CF_3 W = 4,5-dihydro-1H-imidaz-2-yl. AV5080 has good solubility and metabolic properties and a favorable pharmacokinetics profile in animals.

‡‡The pharmacokinetics properties of the 2,3-difluorosialic acid (DFSA) derivative FaxGuDFSA were comparable to those of zanamivir, a 100% efficacy at 10 mg/kg/day with no side effects in infected mice (H3N2) [21].

§§S31N-resistant strain arising from wsn in Ref. 3.

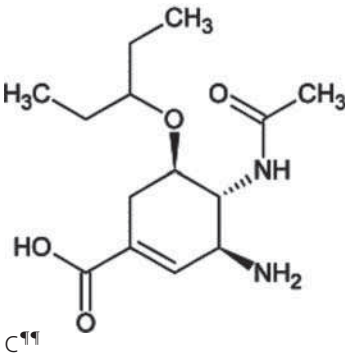
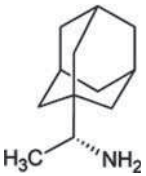
¶¶The full chemical names of the compounds represented in the third column are listed above.

HA: Hemagglutinin; M2: Matrix 2; NA: Neuraminidase; nd: Not determined; PDB: Protein Data Bank; WSN: Influenza A/WSN/33 H1N1.

RIGHTSLINK

Expert Opin. Drug Discov. Downloaded from informahealthcare.com by INSERM on 04/10/15
For personal use only.

Table 1. External proteins of influenza virus and their inhibitors binding sites* (continued).

Protein	Function	Structure	Binding site/ PDB	Inhibitor/status	IC_{50} or Kd strain/CC_{50}	Ref.
			4KS1	Triazole-substituted sialic acid derivatives /preclinical		Kerry *et al.* (2013) [24]
M2	Proton channel	C¶¶ 	Pore site 2RLF	Rimantadine/approved	0.1 μM wsn, 200 μM S31N[§§]	Schnell and Chou (2008) [26]

A = (2E,5S,9R,10S)-10-(acetylamino)-2-imino-4-oxo-9-(pentan-3-yloxy)-1-thia-3-azaspiro[4.5]dec-6-ene-7-carboxylic acid.

B = (5R,9R,10S)-10-(acetylamino)-2-amino-4-oxo-9-(pentan-3-yloxy)-1-thia-3-azaspiro[4.5]deca-2,6-diene-7-carboxylic acid.

C = (3S,4R,5R)-4-(acetylamino)-3-amino-5-(pentan-3-yloxy)cyclohex-1-ene-1-carboxylic acid.

D = (3S,5S,7S)-N-{[5-(thiophen-2-yl)-1,2-oxazol-3-yl]methyl}tricyclo[3.3.1.1 ~ 3,7 ~]decan-1-aminium.

*Only part of the inhibitor structures that are found in the PDB and those involving a structure-based design combined with experimental validation are summarized here.

[‡]The derivatives exhibit good water solubility and good stability in human liver microsomes (6 – 8 ml/min/kg) and no inhibition of major CYP450 isoforms.

[§]Both compounds exhibit strong synergy with oseltamivir.

[¶]These NA inhibitors currently approved in the US are only FDA-approved for uncomplicated influenza infections [72]. Peramivir is in clinical trials in US. p09N1: pandemic H1N1 2009, N5: H12N5.

[#]Laninamivir exists as an octanoate prodrug that is slowly processed in adult patients; the prodrug binds as oseltamivir and its efficacy is not improved as compared to oseltamivir against resistant strain. N5: H12N5.

**AV580: (3R,4R,5S)-5-[(diaminomethylene)amino]-3-(1-ethylpropoxy)-4-[(fluoroacetyl)amino] cyclohex-1-ene-1-carboxylic acid R = CF_3 W = 4,5-dihydro-1H-imidaz-2-yl. AV5080 has good solubility and metabolic properties and a favorable pharmacokinetics profile in animals.

[‡‡]The pharmacokinetics properties of the 2,3-difluorosialic acid (DFSA) derivative FaxGuDFSA were comparable to those of zanamivir, a 100% efficacy at 10 mg/kg/day with no side effects in infected mice (H3N2) [21].

[§§]S31N-resistant strain arising from wsn in Ref. 3.

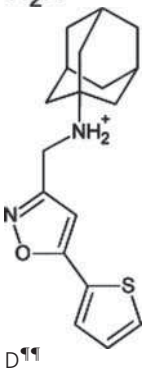
[¶¶]The full chemical names of the compounds represented in the third column are listed above.

HA: Hemagglutinin; M2: Matrix 2; NA: Neuraminidase; nd: Not determined; PDB: Protein Data Bank; WSN: Influenza A/WSN/33 H1N1.

RIGHTSLINK

Expert Opin. Drug Discov. Downloaded from informahealthcare.com by INSERM on 04/10/15
For personal use only.

Table 1. External proteins of influenza virus and their inhibitors binding sites* (continued).

Protein	Function	Structure	Binding site/ PDB	Inhibitor/status	IC_{50} or Kd strain/CC_{50}	Ref.
			Pore site 2KQT	Amantadine/approved	100 μM wsn, > 500 μM S31N	Cady *et al.* (2010) [27]
		 D¶¶	Pore site 2LY0	M2-WJ332/preclinical	26 μM S31N	Wang *et al.* (2013) [3,29]

A = (2E,5S,9R,10S)-10-(acetylamino)-2-imino-4-oxo-9-(pentan-3-yloxy)-1-thia-3-azaspiro[4.5]dec-6-ene-7-carboxylic acid.

B = (5R,9R,10S)-10-(acetylamino)-2-amino-4-oxo-9-(pentan-3-yloxy)-1-thia-3-azaspiro[4.5]deca-2,6-diene-7-carboxylic acid.

C = (3S,4R,5R)-4-(acetylamino)-3-amino-5-(pentan-3-yloxy)cyclohex-1-ene-1-carboxylic acid.

D = (3S,5S,7S)-N-{[5-(thiophen-2-yl)-1,2-oxazol-3-yl]methyl}tricyclo[3.3.1.1 ~ 3,7 ~]decan-1-aminium.

*Only part of the inhibitor structures that are found in the PDB and those involving a structure-based design combined with experimental validation are summarized here.

‡The derivatives exhibit good water solubility and good stability in human liver microsomes (6 – 8 ml/min/kg) and no inhibition of major CYP450 isoforms.

§Both compounds exhibit strong synergy with oseltamivir.

¶These NA inhibitors currently approved in the US are only FDA-approved for uncomplicated influenza infections [72]. Peramivir is in clinical trials in US. p09N1: pandemic H1N1 2009, N5: H12N5.

#Laninamivir exists as an octanoate prodrug that is slowly processed in adult patients; the prodrug binds as oseltamivir and its efficacy is not improved as compared to oseltamivir against resistant strain. N5: H12N5.

**AV580: (3R,4R,5S)-5-[(diaminomethylene)amino]-3-(1-ethylpropoxy)-4-[(fluoroacetyl)amino] cyclohex-1-ene-1-carboxylic acid R = CF_3 W = 4,5-dihydro-1H-imidaz-2-yl. AV5080 has good solubility and metabolic properties and a favorable pharmacokinetics profile in animals.

‡‡The pharmacokinetics properties of the 2,3-difluorosialic acid (DFSA) derivative FaxGuDFSA were comparable to those of zanamivir, a 100% efficacy at 10 mg/kg/day with no side effects in infected mice (H3N2) [21].

§§S31N-resistant strain arising from wsn in Ref. 3.

¶¶The full chemical names of the compounds represented in the third column are listed above.

HA: Hemagglutinin; M2: Matrix 2; NA: Neuraminidase; nd: Not determined; PDB: Protein Data Bank; WSN: Influenza A/WSN/33 H1N1.

RIGHTSLINK

Expert Opin. Drug Discov. Downloaded from informahealthcare.com by INSERM on 04/10/15
For personal use only.

2.2 Neuraminidase

The NA is a second glycoprotein expressed at the surface of influenza virus which releases the newly formed viral particle by cleaving the terminal sialic acid from HA receptors on cell membranes (**Figure 1** and **Table 1**). NA is a tetrameric enzyme. NA forms two phylogenetic groups with distinctive structural features in the active site, including a protein loop (150-loop) in an open conformation in group 1 (N1, N4, N5 and N8), whereas in group 2 NAs (N2, N3, N6, N7 and N9) this loop is in a closed conformation. The currently used antivirals are oseltamivir (Tamiflu) and zanamivir (Relenza) that target the active site of NA and represent successful early structure-based designed compounds (**Table 1**) [15,16]. Zanamivir has a poor solubility and is administered as aerosol, whereas Tamiflu is formulated as a prodrug. Although zanamivir is well tolerated and has few side effects, the route of administration of this drug can represent a problem especially for children and elderly patients who might not be able to inhale zanamivir suitably [3]. To circumvent these issues, an intravenous formulation of zanamivir has been developed and is currently under Phase III clinical trial. Peramivir is a cyclopentane compound, approved in Japan for use in adult and pediatric patients with influenza A virus and influenza B virus infection [17]; in addition, it is currently undergoing clinical trials in the US and in other countries. This compound is administered only as an intravenous formulation, due to its very low bioavailability. Laninamivir targets more specifically class 2 NA and is potent, in particular, against the 2009 pandemic strain [18]. The octanoate prodrug of laninamivir has binding and inhibitory properties that make it like Tamiflu (**Table 1**). Many mutations of NA have been identified in patients, of which some confer resistance to treatment [1]. In 2009, the swine flu pandemic (p09N1) was sensitive to oseltamivir but resistance develops in this strain as observed earlier in the H274Y H1N1-resistant strains.

Based on crystallographic structures, novel anti-NA compounds have been developed, some of them being very tight binders as AV5080, whereas others act as 'suicide inhibitors' forming a covalent bond with NA active site [19-21]. Exploitation of the cavity formed by the open-loop conformation (the 150-cavity) led to the development of novel sialic acid derivatives, such as oseltamivir derivatives modified with a spirolactam ring or carrying various side-chains [22-24]. In the C4 position, the OH group of sialic acid was replaced by a guanidinium in peramivir, zanamivir and laninamivir and by an amine in oseltamivir and AV580. The amide group in the C5 position and carboxylic acid in the second position are shared by most compounds and sialic acid, the latter being recognized by an arginine (R292). Conjugates of polymers (biodegradable polyglutamine) with multiple copies of zanamivir interfered more efficiently than zanamivir with intracellular trafficking of the endocytosed viruses and the subsequent virus–endosome fusion [25]. Their advantage resides in a higher antiviral activity by 1000- to 10,000-folds and *in vivo* longer-lasting effects than the monomeric drug.

2.3 Matrix 2

M2 is a homotetramer composed of four parallel transmembrane α-helices positioned across the viral lipid envelope, forming the ion channel where proton conduction takes place. In response to low pH in the endosome, M2 creates a proton flux from the endosome to the virion core (**Figure 1**). M2 is also responsible for the delayed acidification of the Golgi before viral assembly to insure correct structures of the other viral proteins. The N-terminal end of the channel is gated by the V27 residue, and the channel width (pore) is controlled by four additional pore-facing residues such as A30, S31 (N31 in resistant strains), G34, H37 and W41 that constitute the functional core of the channel (**Figure 2**). H37 is the pH sensor, whereas W41 acts as a C-terminal gate and is adjacent to amphipathic helices stabilizing the channel. The above pore-controlling residues constitute the high-affinity binding site for amantadine and rimantadine that inhibit proton conduction by pore occlusion [26-28]. Adamantanes were the first anti-influenza drugs approved, although presenting CNS side effects. Resistance to adamantanes as adamantine is the main problem of most present circulating or threatening influenza virus, including avian H5N1. The hydrophobic adamantane moiety interacts with the channel gate V27, while their positively charged group interacts with H37. The channel adapts to saturated polycyclic compounds of various sizes [29] and amantadine substituted with aromatic side chain as thienyl group in M2WJ332 [3] were potent inhibitors of drug-resistant S31N mutant of M2 (**Figure 2** and **Table 1**).

Resistance to antiviral treatment addressed above led to the targeting of the RNP complex, the viral minimal machinery for all transcription and replication [2,30]. Because these proteins are usually less prone to mutations, and some of these proteins as the NP are unique to virus and have no equivalent in the host, targeting the RNP by small molecules could tend to an ultimate goal in antiviral treatment: i) to be broad spectrum and effective against various strains and ii) to generate minimal escape from treatment and if possible be effective on existing resistant strains.

3. Structural basis for small-molecules inhibition of individual proteins of the RNP

3.1 Nucleoprotein

NP is a highly conserved protein that binds to and protects the viral genomic or anti-genomic RNA [30-32]. NP has a structural chaperone-like role in the assembly of the RNP; however, much of the structural information on how RNA binds to NP is still incomplete, and the assembly of NP-RNA oligomers into RNP mainly relies on exciting cryo-electron microscopy (EM) studies at 20 Å resolution [33-35]. The reconstruction showed a double helical stem region of NP oligomers on viral single-stranded RNA in which NP

RIGHTSLINK

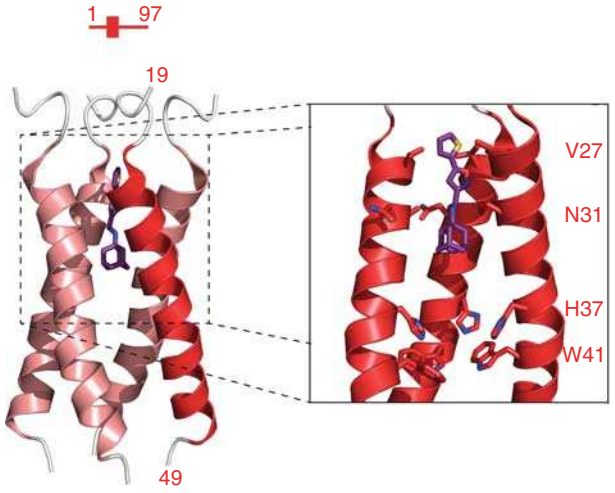


Figure 2. Illustration showing part of the M2 proton channel constituted by the assembly of four parallel transmembrane α-helices across the viral coat. The core of the channel involves V27 as the gate to the channel, and important residues N31, H37 and W41 are represented. The channel lumen is blocked by the inhibitor M2-WJ332 that is active against the resistant S31N protein [3,29].

mediated the assembly of the two strands. Interactions of NP with the polymerase subunits through other types of interactions were also observed, consistent with NP–PB2 interactions characterized by other methods [36-38]. How NP discriminates these viral RNAs (viral RNA of negative polarity [vRNA] and complementary RNA intermediate of positive polarity [cRNA]) from viral mRNA and cellular RNAs is still unknown [39]. NP is hypothesized to contribute to switching between transcription and replication that take place in the nucleus of infected cells; the exact role of NP in these processes is an active field of research [40-42]. NP binds single-stranded viral RNA as oligomers that are essential for RNP complex activity on long templates [43]. NP self-association is mediated by a flexible tail-loop that protrudes into a pocket in adjacent subunit on trimer formation; a dimeric interface was also defined [44-46]. The tail loop position (residues 402 – 428 of NP) differs within the influenza A H1N1 and H5N1 NP trimer and influenza B tetramer [47], thus suggesting the that tail loop may support different oligomers [48,49]. NP oligomerization is a dynamic process and NP can also adopt a monomeric structure [30], which is stabilized by phosphorylation and possibly through the interaction with other (cellular) proteins as importin α5 [30,50-53]. NP structure is characterized by a head and body domain that are separated by a highly basic groove involved in RNA-binding. **Figure 3** and **Table 2** recapitulate the present antiviral candidates targeting NP for which structure is known or estimated by molecular dynamic simulations and confirmed by experimental evidence. The ability of NP to self-associate and form NP-RNA oligomers required for RNP function was targeted by two opposite approaches (**Figure 3A**). In the first one, nucleozin stabilized the interface between NP monomers inducing the formation of nonfunctional higher order oligomers that modified the RNP trafficking to the nucleus (**Figure 3A**) [54-56]. Nucleozin was the first molecule identified by high-throughput screening using the nuclear transport of NP as readout. The nucleozin-binding sites were identified by structural and modeling studies and were confirmed by escape mutations to nucleozin analog treatment at the interface between two NP molecules in a hydrophobic pocket formed by Y289, N309 and Y52 and Y313 (**Figure 3C**) [57]. Nucleozin and its derivatives bear the arylpiperazine scaffold associated with a relatively low solubility and yet their pharmacokinetics was not reported. These compounds are presently the most potent compounds targeting NP. Shen *et al.* used an opposite approach that avoided oligomer formation by blocking the insertion of the oligomerization loop targeting the E339-R416 salt bridge with peptides or compound 3 bearing a thiazole carboxamide motif (**Figure 3A**, center and **Table 2**) [58]. Competing with NP binding to RNA was another approach that identified naproxen that bound in a small hydrophobic site formed by Y148, F489 and R361 in the RNA-binding groove of NP based on molecular modeling and confirmed by surface plasmon resonance (**Figure 3A**, right and **3B**) [59]. Naproxen combined an antiviral activity with an anti-inflammatory effect; being also a generic drug inhibiting the inducible COX, naproxen is an approved drug (anti-COX activity) and as such has favorable solubility and stability properties. Naproxen and its designed derivatives were shown to stabilize monomeric NP by modifying the cavity in which the oligomerization loop inserts, close to their binding site [60].

Within the context of the RNP, the RdRp, composed by the three subunits PA, PB1 and PB2, mediates replication and transcription of the segmented viral RNA. The transcription process involves a unique cap-snatching mechanism in which the RdRp binds to host cellular pre-mRNAs via their 5′-cap structure and then cleaves them to generate capped oligonucleotides, which are used as primers to initiate the transcription of viral mRNAs. Owing to the cap, viral mRNAs are translated into proteins by the cell machinery. The cap-snatching mechanism is an essential and conserved process in all members of the Orthomyxoviridae family, including influenza A, B and C viruses. Because the host cell has no analogous activity, cap-snatching inhibitors could be specific against all influenza subtypes and strains. Cap-snatching processes can be inhibited by targeting either the PB2 cap-binding or PA endonuclease domains. Influenza polymerase domains such as PA endonuclease [61,62], PA C-terminal domain associated with PB1 N-terminal peptide [63,64], PB1 C-terminal associated with PB2 N-terminal [65], PB2 Cap-binding domain [66] as well large parts of PB2 C-terminal [67] were solved and these structures provided the basis for structure-based design of specific inhibitors as detailed below. During the course of writing this review, the structures of the

Expert Opin. Drug Discov. Downloaded from informahealthcare.com by INSERM on 04/10/15
For personal use only.

RIGHTSLINK

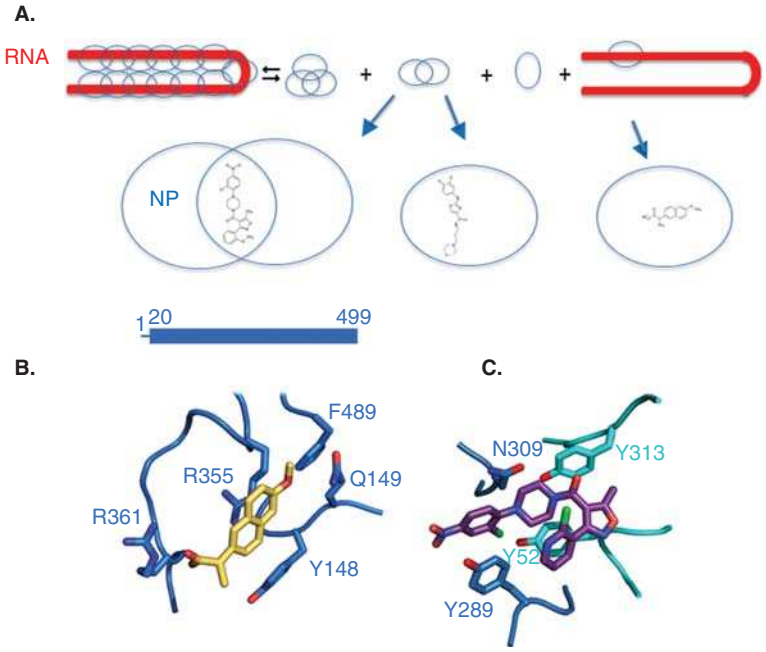


Figure 3. Illustration showing inhibitors of the nucleoprotein NP: two opposite strategies have been developed for modulating NP oligomerization loop as explained by the cartoon shown in panel (A). NP is found in various assemblies with and without single-stranded RNA that may interconvert depending on the conditions. Attempts to decrease NP self-association through binding to the oligomerization [58] (panel A, center) or to impede RNA binding to the binding groove of monomeric NP (panel (A), right and yellow naproxen in panel **(B)** [59,60] is presented. The nucleozin derivative, shown in purple, stabilizes the interface between two adjacent NP monomers and favors nonfunctional oligomers of NP (panel (A), left and panel **(C)** the residues of the two NP are shown in cyan and blue) [55,57].

whole polymerase of influenza A and B were published [68,69]. **Figure 4** shows the overall structure of Influenza A polymerase; the previously solved domains of PA endonuclease (green) and PB2 Cap-binding (orange) with bound inhibitors are highlighted.

3.2 PB2 cap-binding

Transcription by the cap-snatching mechanism requires that host-cell pre-mRNAs are bound via their 5′-cap structure to the PB2 subunit. Thus, the PB2 cap-binding site is a potential target for new drugs to inhibit viral replication. The X-ray structure of the cap-binding site with bound 7-methylguanosine triphosphate (m^7GTP) has been obtained [66] and revealed that the m^7GTP is sandwiched between two aromatic residues (Phe404 and His357), forming strong π–stacking interactions. The guanine base is also recognized by hydrogen bonds to key PB2 residues (Lys376 and Glu361) and several residues are involved in interactions with phosphate groups. Docking studies using the structure of the PB2 cap-binding domain suggested that different 7-methylguanine derivatives are good candidates [70]. Crystal structures of cap-binding domain–inhibitor complexes have been obtained with three active m^7GTP derivatives and confirmed a unique binding mode of the m^7GTP scaffold in the PB2 recognition site for these compounds and the m^7GTP [70]. More recently, a series of azaindole derivatives targeting PB2 has been identified [71]. The X-ray crystal structure of a pyrimidine azaindole derivative denoted VX787 (described in **Table 2**) bound to PB2 [71] confirmed that VX787 occupies the m^7GTP-binding site where the azaindole ring makes interactions with the key residues, Lys376 and Glu361, as well as stacking of the three aromatic rings of VX787 between the side chains of His357, Phe323 Phe404 and Phe325 (**Figure 4**, orange inset). Orally bioavailable, this compound shows a strong potency for a large number of influenza A strains, including pandemic 2009 H1N1 and avian H5N1 flu strains in a mouse influenza model. VX787 is the one of the two compounds targeting the polymerase actually under clinical trial. The other compound is T705, a ribonucleotide analog (6-fluoro-3-hydroxy-2-pyrazinecarboxamide) also known as favipiravir that inhibits many viral RdRps, thus being effective against a variety of viruses [72].

3.3 PA endonuclease

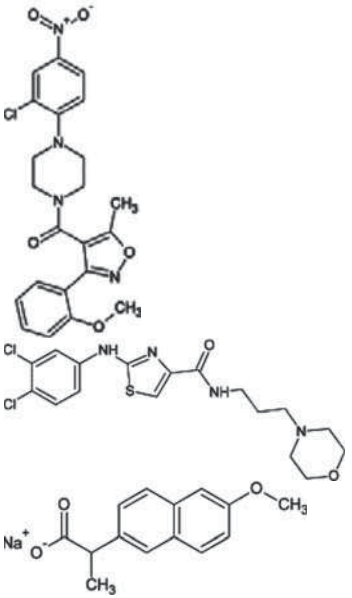
The endonuclease activity necessary for the cleavage of the host cellular pre-mRNAs is carried out by the N-terminal domain of PA subunit (Nter-PA, ~ 25 kDa, residues 1 – 196). The structure of Nter-PA has been solved in several crystal forms with (or without) various ligands [61,62,73-79]. Two years after the first structure of the PA catalytic domain

Expert Opin. Drug Discov. Downloaded from informahealthcare.com by INSERM on 04/10/15
For personal use only.

RIGHTSLINK

Expert Opin. Drug Discov. Downloaded from informahealthcare.com by INSERM on 04/10/15
For personal use only.

Table 2. Internal proteins involved in RNP function and in antagonizing host antiviral response against influenza A virus and their inhibitor binding sites.

Protein	Function	Structure*	Binding site/PDB	Inhibitor*/ Status‡	IC_{50} or Kd strain/ CC_{50}	Ref.
NP	RNA-binding		Hydrophobic pocket	Nucleozine	IC_{50} = 69 nM (WSN) 160 nM (H3N2) 330 nM (H5N1); CC_{50} > 250 μM	Kao *et al.* (2010) [55]
			3RO5	Nucleozine derivatives	IC_{50} = 520 nM (H1N1 2009), 2 μM (H3N2) 690 nM (H5N1); CC_{50} > 100 μM	Gerritz *et al.* (2011) [57]
			Oligomerization loop	Compound 3	IC_{50} = 2.7 μM (WSN); CC_{50} = 35.5 μM	Shen *et al.* (2011) [58]
			RNA-binding groove	Naproxen and derivatives	IC_{50} = 16 μM (WSN) 12 μM (H3N2), CC_{50} = 1605 μM	Lejal *et al.* (2013); Tarus *et al.* (2014) [59,60]

Adapted in part from Ref. [17] with permission of Future Medicine Ltd.

*The full chemical names of the compounds represented in the third column are listed below.

‡All compounds are in preclinical stage except VX787.

Nucleozine derivative: [4-(2-chloro-4-nitrophenyl)piperazin-1-yl] [3-(2-methoxyphenyl)-5-methyl-1,2-oxazol-4-yl]methanone.

Compound 3: 2-((3,4-dichlorophenyl)amino)-N-(3-morpholinopropyl)thiazole-5-carboxamide.

VX787: (2S,3S)-3-[[5-fluoranyl-2-(5-fluoranyl-1H-pyrrolo[2,3-b]pyridin-3-yl)pyrimidin-4-yl]amino] bicyclo[2.2.2]octane-2-carboxylic acid.

Cap analogue: 9-N-(3-Carboxy-4-hydrophenyl) ketomethyl-7-N-methylguanine.

Hydroxypyrimidine derivative: 2-(4-(1H-tetrazol-5-yl)phenyl)-5-hydroxypyrimidin-4(3H)-one.

Hydroxyquinoline derivative: 7-(4-fluorophenyl)-3-hydroxyquinolin-2(1H)-one.

Hydroxypyridine derivative: 5-(4-Fluorophenyl)-3-hydroxy-6-[4-(1H-1,2,3,4-tetrazol-5-yl)phenyl]-1,2-Dihydropyridin-2-one.

Phenol derivative: 4-(1H-imidazol-1-yl)phenol.

Diketo derivative: (2Z)-4-[(3S)-1-(benzylsulfonyl)-3-(4-chlorobenzyl)piperidin-3-yl]-2-hydroxy-4-oxobut-2-enoic acid.

Polyphenol EGCG: epigallocatechin gallate.

Quinoline derivative: 2-hydroxyisoquinoline-1,3(2H,4H)-dione.

DPBA: 2-4-Dioxo-4-phenylbutanoic acid.

Pyrimidine derivative: 2-[3-(acetylamino)phenyl]-5-hydroxy-6-oxo-1,6-dihydropyrimidine-4-carboxylic acid.

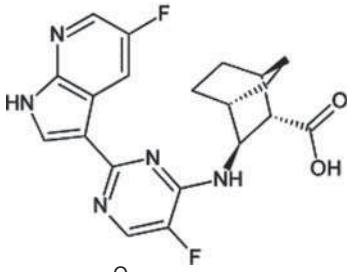
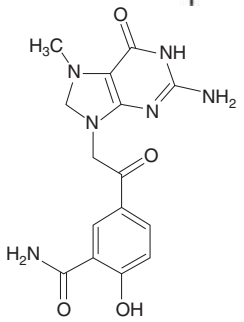
Tintori 2014: ethyl 4-[[2-[3-cyano-4-(3-ethoxy-4-phenylmethoxyphenyl)-6-phenylpyridin-2-yl]sulfanylacetyl]amino]benzoate.

Quinoxaline derivative: 2,3-Bis(2-furyl)quinoxaline-6-{1-[(2-furyl)methyl]aminocarbonyl}.Nd: Not determined; NP: Nucleoprotein; NS1: Non-structural protein 1; PDB: Protein Data Bank; WSN: Influenza A/ WSN/33 H1N1.

RIGHTSLINK

Expert Opin. Drug Discov. Downloaded from informahealthcare.com by INSERM on 04/10/15
For personal use only.

Table 2. Internal proteins involved in RNP function and in antagonizing host antiviral response against influenza A virus and their inhibitor binding sites (continued).

Protein	Function	Structure*	Binding site/PDB	Inhibitor*/ Status‡	IC_{50} or Kd strain/ CC_{50}	Ref.
PB2	Cap-binding		4P1U	Azaindole derivative (VX787)/ clinical trial	EC_{50} = 0.0026 μM (H1N1)	Clark *et al.* (2014) [71]
			Cap-binding site	Anthracene derivative	Nd	
			4NCE 4NCM	Pyrimidine derivative	Nd	
			4CB5	7-methylguanine derivatives	4CB5: IC_{50} = 7.5 μM (H3N2)	Pautus *et al.* (2013) [70]
			4CB4,4CB6, 4CB7			

Adapted in part from Ref. [17] with permission of Future Medicine Ltd.

*The full chemical names of the compounds represented in the third column are listed below.

‡All compounds are in preclinical stage except VX787.

Nucleozine derivative: [4-(2-chloro-4-nitrophenyl)piperazin-1-yl] [3-(2-methoxyphenyl)-5-methyl-1,2-oxazol-4-yl]methanone.

Compound 3: 2-((3,4-dichlorophenyl)amino)-N-(3-morpholinopropyl)thiazole-5-carboxamide.

VX787: (2S,3S)-3-[[5-fluoranyl-2-(5-fluoranyl-1H-pyrrolo[2,3-b]pyridin-3-yl)pyrimidin-4-yl]amino] bicyclo[2.2.2]octane-2-carboxylic acid.

Cap analogue: 9-N-(3-Carboxy-4-hydrophenyl) ketomethyl-7-N-methylguanine.

Hydroxypyrimidine derivative: 2-(4-(1H-tetrazol-5-yl)phenyl)-5-hydroxypyrimidin-4(3H)-one.

Hydroxyquinoline derivative: 7-(4-fluorophenyl)-3-hydroxyquinolin-2(1H)-one.

Hydroxypyridine derivative: 5-(4-Fluorophenyl)-3-hydroxy-6-[4-(1H-1,2,3,4-tetrazol-5-yl)phenyl]-1,2-Dihydropyridin-2-one.

Phenol derivative: 4-(1H-imidazol-1-yl)phenol.

Diketo derivative: (2Z)-4-[(3S)-1-(benzylsulfonyl)-3-(4-chlorobenzyl)piperidin-3-yl]-2-hydroxy-4-oxobut-2-enoic acid.

Polyphenol EGCG: epigallocatechin gallate.

Quinoline derivative: 2-hydroxyisoquinoline-1,3(2H,4H)-dione.

DPBA: 2-4-Dioxo-4-phenylbutanoic acid.

Pyrimidine derivative: 2-[3-(acetylamino)phenyl]-5-hydroxy-6-oxo-1,6-dihydropyrimidine-4-carboxylic acid.

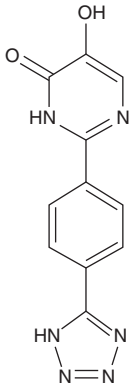
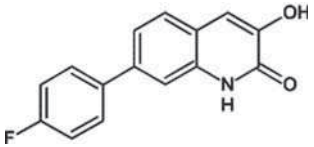
Tintori 2014: ethyl 4-[[2-[3-cyano-4-(3-ethoxy-4-phenylmethoxyphenyl)-6-phenylpyridin-2-yl]sulfanylacetyl]amino]benzoate.

Quinoxaline derivative: 2,3-Bis(2-furyl)quinoxaline-6-{1-[(2-furyl)methyl]aminocarbonyl}.Nd: Not determined; NP: Nucleoprotein; NS1: Non-structural protein 1; PDB: Protein Data Bank; WSN: Influenza A/ WSN/33 H1N1.

RIGHTSLINK

Expert Opin. Drug Discov. Downloaded from informahealthcare.com by INSERM on 04/10/15
For personal use only.

Table 2. Internal proteins involved in RNP function and in antagonizing host antiviral response against influenza A virus and their inhibitor binding sites (continued).

Protein	Function	Structure*	Binding site/PDB	Inhibitor*/ Status‡	IC_{50} or Kd strain/ CC_{50}	Ref.
PA	Endonuclease		Endonuclease active site 4W9S	Hydroxypyrimidine derivative	IC_{50} = 0.15 μM (endo. inhibition; H3N2)	Sagong *et al.* (2014) [77]
			Endonuclease active site 4KIL	Hydroxyquinoline derivative	IC_{50} = 0.5 μM (endo. inhibition; H3N2)	Sagong *et al.* (2013) [78]

Adapted in part from Ref. [17] with permission of Future Medicine Ltd.

*The full chemical names of the compounds represented in the third column are listed below.

‡All compounds are in preclinical stage except VX787.

Nucleozine derivative: [4-(2-chloro-4-nitrophenyl)piperazin-1-yl] [3-(2-methoxyphenyl)-5-methyl-1,2-oxazol-4-yl]methanone.

Compound 3: 2-((3,4-dichlorophenyl)amino)-N-(3-morpholinopropyl)thiazole-5-carboxamide.

VX787: (2S,3S)-3-[[5-fluoranyl-2-(5-fluoranyl-1H-pyrrolo[2,3-b]pyridin-3-yl)pyrimidin-4-yl]amino] bicyclo[2.2.2]octane-2-carboxylic acid.

Cap analogue: 9-N-(3-Carboxy-4-hydrophenyl) ketomethyl-7-N-methylguanine.

Hydroxypyrimidine derivative: 2-(4-(1H-tetrazol-5-yl)phenyl)-5-hydroxypyrimidin-4(3H)-one.

Hydroxyquinoline derivative: 7-(4-fluorophenyl)-3-hydroxyquinolin-2(1H)-one.

Hydroxypyridine derivative: 5-(4-Fluorophenyl)-3-hydroxy-6-[4-(1H-1,2,3,4-tetrazol-5-yl)phenyl]-1,2-Dihydropyridin-2-one.

Phenol derivative: 4-(1H-imidazol-1-yl)phenol.

Diketo derivative: (2Z)-4-[(3S)-1-(benzylsulfonyl)-3-(4-chlorobenzyl)piperidin-3-yl]-2-hydroxy-4-oxobut-2-enoic acid.

Polyphenol EGCG: epigallocatechin gallate.

Quinoline derivative: 2-hydroxyisoquinoline-1,3(2H,4H)-dione.

DPBA: 2-4-Dioxo-4-phenylbutanoic acid.

Pyrimidine derivative: 2-[3-(acetylamino)phenyl]-5-hydroxy-6-oxo-1,6-dihydropyrimidine-4-carboxylic acid.

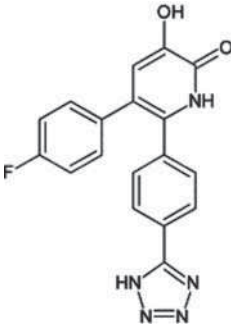
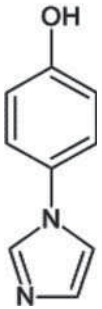
Tintori 2014: ethyl 4-[[2-[3-cyano-4-(3-ethoxy-4-phenylmethoxyphenyl)-6-phenylpyridin-2-yl]sulfanylacetyl]amino]benzoate.

Quinoxaline derivative: 2,3-Bis(2-furyl)quinoxaline-6-{1-[(2-furyl)methyl]aminocarbonyl}.Nd: Not determined; NP: Nucleoprotein; NS1: Non-structural protein 1; PDB: Protein Data Bank; WSN: Influenza A/ WSN/33 H1N1.

RIGHTSLINK

Expert Opin. Drug Discov. Downloaded from informahealthcare.com by INSERM on 04/10/15
For personal use only.

Table 2. Internal proteins involved in RNP function and in antagonizing host antiviral response against influenza A virus and their inhibitor binding sites (continued).

Protein	Function	Structure*	Binding site/PDB	Inhibitor*/ Status‡	IC_{50} or Kd strain/ CC_{50}	Ref.
			Endonuclease active site 4M5U	Hydroxypyridine derivative	IC_{50} = 0.023 μM (endo. inhibition; H1N1)	Parhi *et al.* (2013) [76]
			Endonuclease active site 4LN7, 4MK1, 4MK2, 4MK5, 4M5O, 4M4Q	Pyridine derivatives	IC_{50} = 0.041 μM (endo. inhibition; H1N1)	Bauman *et al.* (2013) [73]
			Endonuclease active site 4M5R	Phenol derivative	IC_{50} = 1000 μM (endo. inhibition; H1N1)	Kowalinski *et al.* (2012) [74]

Adapted in part from Ref. [17] with permission of Future Medicine Ltd.

*The full chemical names of the compounds represented in the third column are listed below.

‡All compounds are in preclinical stage except VX787.

Nucleozine derivative: [4-(2-chloro-4-nitrophenyl)piperazin-1-yl] [3-(2-methoxyphenyl)-5-methyl-1,2-oxazol-4-yl]methanone.

Compound 3: 2-((3,4-dichlorophenyl)amino)-N-(3-morpholinopropyl)thiazole-5-carboxamide.

VX787: (2S,3S)-3-[[5-fluoranyl-2-(5-fluoranyl-1H-pyrrolo[2,3-b]pyridin-3-yl)pyrimidin-4-yl]amino] bicyclo[2.2.2]octane-2-carboxylic acid.

Cap analogue: 9-N-(3-Carboxy-4-hydrophenyl) ketomethyl-7-N-methylguanine.

Hydroxypyrimidine derivative: 2-(4-(1H-tetrazol-5-yl)phenyl)-5-hydroxypyrimidin-4(3H)-one.

Hydroxyquinoline derivative: 7-(4-fluorophenyl)-3-hydroxyquinolin-2(1H)-one.

Hydroxypyridine derivative: 5-(4-Fluorophenyl)-3-hydroxy-6-[4-(1H-1,2,3,4-tetrazol-5-yl)phenyl]-1,2-Dihydropyridin-2-one.

Phenol derivative: 4-(1H-imidazol-1-yl)phenol.

Diketo derivative: (2Z)-4-[(3S)-1-(benzylsulfonyl)-3-(4-chlorobenzyl)piperidin-3-yl]-2-hydroxy-4-oxobut-2-enoic acid.

Polyphenol EGCG: epigallocatechin gallate.

Quinoline derivative: 2-hydroxyisoquinoline-1,3(2H,4H)-dione.

DPBA: 2-4-Dioxo-4-phenylbutanoic acid.

Pyrimidine derivative: 2-[3-(acetylamino)phenyl]-5-hydroxy-6-oxo-1,6-dihydropyrimidine-4-carboxylic acid.

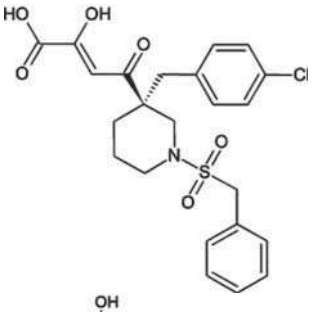
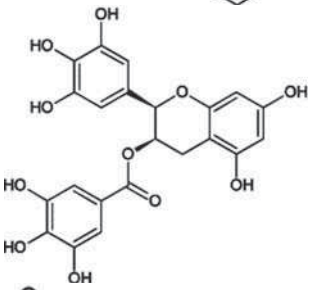
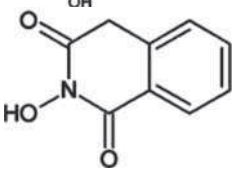
Tintori 2014: ethyl 4-[[2-[3-cyano-4-(3-ethoxy-4-phenylmethoxyphenyl)-6-phenylpyridin-2-yl]sulfanylacetyl]amino]benzoate.

Quinoxaline derivative: 2,3-Bis(2-furyl)quinoxaline-6-{1-[(2-furyl)methyl]aminocarbonyl}.Nd: Not determined; NP: Nucleoprotein; NS1: Non-structural protein 1; PDB: Protein Data Bank; WSN: Influenza A/ WSN/33 H1N1.

RIGHTSLINK

Expert Opin. Drug Discov. Downloaded from informahealthcare.com by INSERM on 04/10/15
For personal use only.

Table 2. Internal proteins involved in RNP function and in antagonizing host antiviral response against influenza A virus and their inhibitor binding sites (continued).

Protein	Function	Structure*	Binding site/PDB	Inhibitor*/ Status‡	IC_{50} or Kd strain/ CC_{50}	Ref.
			Endonuclease active site 4AWK 4AVG, 4AWG	Diketo derivatives	IC_{50} = 0.13 μM (endo. inhibition); IC_{50} = 20 μM (antiviral inhibition; H3N2); CC_{50} > 100 μM	
			Endonuclease active site 4AWF	DPBA	IC_{50} = 2.7 μM (endo. inhibition); IC_{50} = inactive (antiviral inhibition; H3N2); CC_{50} > 200 μM	
			4AWM	EGCG from green tea	IC_{50} = 1.9 μM (endo. inhibition; H3N2); IC_{50} = 1.1 μM (antiviral inhibition); CC_{50} > 54 μM	
			4E5F	Hydroxyquinoline derivative	IC_{50} = 15 μM (endo. Inhibition; H1N1);	DuBois *et al.* (2012) [75]
			4E5G	DPBA	IC_{50} = 20 μM (endo. inhibition; H1N1);	
			4E5H	Benzylpiperidine derivative	IC_{50} = 0.43 μM (endo. inhibition; H1N1);	

Adapted in part from Ref. [17] with permission of Future Medicine Ltd.

*The full chemical names of the compounds represented in the third column are listed below.

‡All compounds are in preclinical stage except VX787.

Nucleozine derivative: [4-(2-chloro-4-nitrophenyl)piperazin-1-yl] [3-(2-methoxyphenyl)-5-methyl-1,2-oxazol-4-yl]methanone.

Compound 3: 2-((3,4-dichlorophenyl)amino)-N-(3-morpholinopropyl)thiazole-5-carboxamide.

VX787: (2S,3S)-3-[[5-fluoranyl-2-(5-fluoranyl-1H-pyrrolo[2,3-b]pyridin-3-yl)pyrimidin-4-yl]amino] bicyclo[2.2.2]octane-2-carboxylic acid.

Cap analogue: 9-N-(3-Carboxy-4-hydrophenyl) ketomethyl-7-N-methylguanine.

Hydroxypyrimidine derivative: 2-(4-(1H-tetrazol-5-yl)phenyl)-5-hydroxypyrimidin-4(3H)-one.

Hydroxyquinoline derivative: 7-(4-fluorophenyl)-3-hydroxyquinolin-2(1H)-one.

Hydroxypyridine derivative: 5-(4-Fluorophenyl)-3-hydroxy-6-[4-(1H-1,2,3,4-tetrazol-5-yl)phenyl]-1,2-Dihydropyridin-2-one.

Phenol derivative: 4-(1H-imidazol-1-yl)phenol.

Diketo derivative: (2Z)-4-[(3S)-1-(benzylsulfonyl)-3-(4-chlorobenzyl)piperidin-3-yl]-2-hydroxy-4-oxobut-2-enoic acid.

Polyphenol EGCG: epigallocatechin gallate.

Quinoline derivative: 2-hydroxyisoquinoline-1,3(2H,4H)-dione.

DPBA: 2-4-Dioxo-4-phenylbutanoic acid.

Pyrimidine derivative: 2-[3-(acetylamino)phenyl]-5-hydroxy-6-oxo-1,6-dihydropyrimidine-4-carboxylic acid.

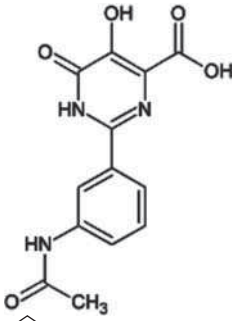
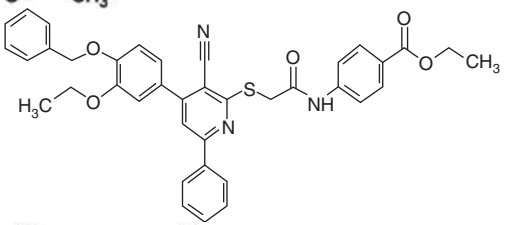
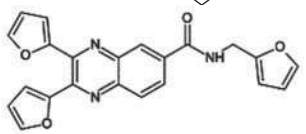
Tintori 2014: ethyl 4-[[2-[3-cyano-4-(3-ethoxy-4-phenylmethoxyphenyl)-6-phenylpyridin-2-yl]sulfanylacetyl]amino]benzoate.

Quinoxaline derivative: 2,3-Bis(2-furyl)quinoxaline-6-{1-[(2-furyl)methyl]aminocarbonyl}.Nd: Not determined; NP: Nucleoprotein; NS1: Non-structural protein 1; PDB: Protein Data Bank; WSN: Influenza A/ WSN/33 H1N1.

RIGHTSLINK

Expert Opin. Drug Discov. Downloaded from informahealthcare.com by INSERM on 04/10/15
For personal use only.

Table 2. Internal proteins involved in RNP function and in antagonizing host antiviral response against influenza A virus and their inhibitor binding sites (continued).

Protein	Function	Structure*	Binding site/PDB	Inhibitor*/ Status‡	IC_{50} or Kd strain/ CC_{50}	Ref.
			4E5J 4E5I, 4E5L	Pyrimidine derivatives	IC_{50} = 5 μM (endo. inhibition); IC_{50} > 50 μM (antiviral inhibition); CC_{50} > 60 μM (H1N1)	
PA–PB1 peptide				Benzoic acid (phenol) derivative Benzofurazan derivatives	30 < IC_{50} < 180 μM (H1N1) IC_{50} = 45 μM (endo. inhibition); IC_{50} = 1 μM (anti-viral inhib.); CC_{50} = 20 μM (WSN,H1N1)	Tintori *et al.* (2014) [84] Pagano *et al.* (2014) [83]
NS1	Antagonist of the cellular antiviral IFN response		RNA-binding site RNA-binding site	Quinoxaline derivatives Polyphenol E	 IC_{50} = 0.29 – 9 μM CC_{50} > 60 μM	You *et al.* (2011) [95] Cho *et al.* (2012) [94]

Adapted in part from Ref. [17] with permission of Future Medicine Ltd.

*The full chemical names of the compounds represented in the third column are listed below.

‡All compounds are in preclinical stage except VX787.

Nucleozine derivative: [4-(2-chloro-4-nitrophenyl)piperazin-1-yl] [3-(2-methoxyphenyl)-5-methyl-1,2-oxazol-4-yl]methanone.

Compound 3: 2-((3,4-dichlorophenyl)amino)-N-(3-morpholinopropyl)thiazole-5-carboxamide.

VX787: (2S,3S)-3-[[5-fluoranyl-2-(5-fluoranyl-1H-pyrrolo[2,3-b]pyridin-3-yl)pyrimidin-4-yl]amino] bicyclo[2.2.2]octane-2-carboxylic acid.

Cap analogue: 9-N-(3-Carboxy-4-hydrophenyl) ketomethyl-7-N-methylguanine.

Hydroxypyrimidine derivative: 2-(4-(1H-tetrazol-5-yl)phenyl)-5-hydroxypyrimidin-4(3H)-one.

Hydroxyquinoline derivative: 7-(4-fluorophenyl)-3-hydroxyquinolin-2(1H)-one.

Hydroxypyridine derivative: 5-(4-Fluorophenyl)-3-hydroxy-6-[4-(1H-1,2,3,4-tetrazol-5-yl)phenyl]-1,2-Dihydropyridin-2-one.

Phenol derivative: 4-(1H-imidazol-1-yl)phenol.

Diketo derivative: (2Z)-4-[(3S)-1-(benzylsulfonyl)-3-(4-chlorobenzyl)piperidin-3-yl]-2-hydroxy-4-oxobut-2-enoic acid.

Polyphenol EGCG: epigallocatechin gallate.

Quinoline derivative: 2-hydroxyisoquinoline-1,3(2H,4H)-dione.

DPBA: 2-4-Dioxo-4-phenylbutanoic acid.

Pyrimidine derivative: 2-[3-(acetylamino)phenyl]-5-hydroxy-6-oxo-1,6-dihydropyrimidine-4-carboxylic acid.

Tintori 2014: ethyl 4-[[2-[3-cyano-4-(3-ethoxy-4-phenylmethoxyphenyl)-6-phenylpyridin-2-yl]sulfanylacetyl]amino]benzoate.

Quinoxaline derivative: 2,3-Bis(2-furyl)quinoxaline-6-{1-[(2-furyl)methyl]aminocarbonyl}.Nd: Not determined; NP: Nucleoprotein; NS1: Non-structural protein 1; PDB: Protein Data Bank; WSN: Influenza A/ WSN/33 H1N1.

RIGHTSLINK

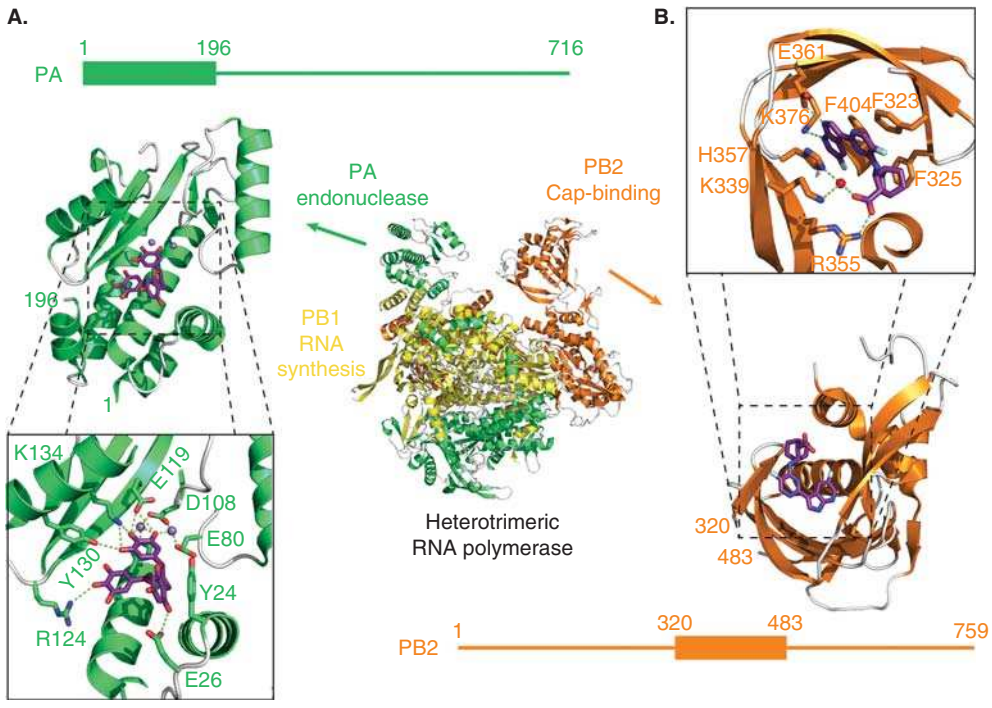


Figure 4. Illustration showing the binding of endonuclease and cap-binding inhibitors. The recent full heterotrimeric RNA-polymerase X-ray structure is shown as a reference ([68] PDB code: 4WSB), PA is represented in green, PB1 in yellow and PB2 in orange. (A) Bound green tea catechin to PA endonuclease active site is shown. Divalent cations (Mn^{2+} ions, purple spheres) and key active site residues that interact with the compound or are close to it are represented. Putative hydrogen bonds (< 3.2 Å) and additional possible interactions (< 3.6 Å) are shown as green dotted lines. Green inset: green tea catechin coordinates the two manganese ions and the active site residues are engaged in the interaction ([74] PDB code: 4AWM). (B) Bound VX787 to PB2 cap-binding domain is shown. Orange inset: the azaindole ring system is shown interacting with the key protein residues K376 and E361 (green dotted lines). The three aromatic rings are stacked between the side chains of H357, F404, F323 and F325. The carboxylic group of the ligand makes two water-mediated interactions (red spheres and green dotted lines) with the nitrogen of H357 and K339 ([71] PDB code: 4P1U).

was solved [61,62], two laboratories described high-resolution X-ray cocrystal structures of PA endonuclease domain with a series of specific compounds known or predicted to block the endonuclease activity [74,75]. This domain presents a high conservation among influenza A, B and C subtypes and identical residues are found in particular in the metal-binding residues His41, Glu80, Asp108, Glu119 and the catalytic residue Lys134. The cocrystal structure of Nter-PA with the well-known inhibitor 2,4 dioxo-4-phenylbutanoic acid (**Table 2**) [80] have been obtained by the two groups. Furthermore, whereas Dubois *et al.* showed crystal structures with different compounds like pyrimidine and hydroxyquinoline derivatives (**Table 2**), Kowalinski *et al.* characterized crystal structures with diketo compounds and a green tea catechin (**Figure 4**). All these inhibitors coordinate the two-metal endonuclease active site and engage the active site residues (**Figure 4**, green inset). The structures showed a structure–activity relationship of several known influenza inhibitors, opening the way for drug design optimization. In particular, residues away from the two-metal center may increase potency and new binding pockets bearing aromatic residues may deserve further investigation. Additionally, the large water molecule network found in the structures may be targeted by inhibitors for an entropic gain while keeping the metal center recognition moiety. Recently, hydroxypyridine derivative has been identified as a bimetal chelating ligand at the endonuclease active site through an X-ray crystallographic screening campaign of fragments library [73]. Thus, two new series of endonuclease inhibitors hydroxypyridine and hydroxyquoline derivatives have been reported [76,78] as well as their aza analogs [77].

3.4 Targeting PA-peptide PB1 interface

Efficient assembly of the influenza virus heterotrimeric RdRp is critical for virus replication and pathogenicity. Therefore, interfering with the assembly of the RdRp complex could be a way to fight against influenza infection. A large library of drugs, in particular benzofurazan derivatives, was found to act by inhibiting the viral RdRp complex through disruption of the complex formed by PA and PB1 subunits [81-84]. Since the association of these subunits is essential for replication [85-88] and the sequence of this domain is highly conserved [89,90],

Expert Opin. Drug Discov. Downloaded from informahealthcare.com by INSERM on 04/10/15
For personal use only.

RIGHTSLINK

this interaction represents an attractive target for inhibitors. Due to the crystallographic structures of the PA–peptide PB1 complex [63,64] a short peptide derived from the C terminus of PB1 has been shown to disrupt the interaction between the C-terminal part of PB1 and the N-terminal part of PB2 and to inhibit viral polymerase activity and replication [91]. Docking studies were performed to explore the structural features responsible for the biological activity of benzofurazan derivatives (**Table 2**) [83,84]. Docking suggested that their most reliable binding mode involves a л–stacking interactions with PA Trp706 of the benzofurazan scaffold, a hydrogen bond with Gln408 and electrostatic interactions with PA Lys643.

Last, most of the compounds targeting the RNP are still in a preclinical stage in terms of therapeutics for humans and many of drug-related aspects still require adequate studies. Likewise, no drug targeting NS1 as described below has been approved and the following are mostly preclinical studies.

4. Targeting the non-structural protein 1

NS1 is a highly conserved multifunctional protein. NS1 protects the virus against the antiviral responses of the host mediated by IFN-α/β pathways in a way that seems to be strain-specific [92]. NS1 antagonizes IFN by a number of ways, as binding to double-stranded (ds) viral RNA, thus protecting it from cellular factors, by blocking retinoic acid inducible gene–I and NF-kB activation. NS1 is active as a dimer and is divided into two functional domains: an N-terminal domain that binds dsRNA (RNA-binding domain [RBD]) and a C-terminal effector domain that binds multiple cellular proteins. The structures of the isolated RBD alone and with RNA bound provided the structural basis for small molecule inhibition of the RNA-protein association or destabilization of the homodimer interface required for dsRNA binding [93]. In particular, a deep cavity was identified in the conserved tracks of basic and hydrophilic residues complementary to the phosphate backbone of A-form dsRNA of NS1 as a site for small molecule binding as a polyphenol and quinoxaline derivatives (**Table 2**) [94,95]. The recent structure of the C-terminal domain of NS1 in complex with a chromatin-associated factor may open a new way to inhibit NS1 through associated cellular factor–NS1 interface(s) [96].

5. Conclusion

Many of the antiviral candidates described in this review still require preclinical and clinical studies to assess their safety, toxicity and efficacy profiles required for the drug to be used on humans. The goal will be to treat in priority more susceptible populations as immunocompromised, young infants and elderly people in the search of antivirals directed against influenza A virus and probably adapt to their specific immune conditions. This goal is, nevertheless, complicated to reach, owing to the difficulties of enrolling patients in appropriate clinical studies, complicated by comorbidities and placebo-controlled conditions which may be ethically unacceptable. FDA approval is thus challenging in these conditions. Future perspectives for new classes of antivirals targeting the RNP complex are very promising but these await pharmacokinetics and drug-related studies to be performed.

6. Expert opinion

6.1 Key findings and weaknesses in the research done so far

Monotherapy with antivirals holds a higher probability of resistance to treatment, in particular to those antivirals targeting external proteins with intrinsic relatively high mutation rate as NA, M2 and the variable region of HA.

To date, the antiviral candidates targeting NA are by far the most developed and for some of them extremely potent. The structure of the sialic acids derivatives in complex with the enzyme active site revealed an empty negatively charged pocket in the region of the C4 position on the sugar ring. The C4-OH was substituted by an amino or guanidinium group (oseltamivir, zanamivir, laninamivir) or by a bulky hydrophobic substituent (AV580 [19], A, B and C; **Table 1**). Substitution at other position(s) may help improve solubility and potency [23,24].

Novel antiviral candidates that target the RNP hold a great potential, both in terms of target conservation and of uniqueness/specificity compared to viral entry inhibitors. In terms of specificity, targeting the PB2 cap-binding domain of influenza virus may not differ so much from targeting cellular factors involved in cap-binding in the host cell. To date, inhibitors of the capping machinery are used to reverse chemoresistance since deregulation of cap-dependent translation is associated with cancer initiation [97]. Nevertheless, quite potent inhibitors of PB2 cap-binding were recently developed with few reported interactions with cellular partners [70,71]. The endonuclease domain represents a specific target with an excellent druggable site [73-78,83,84]. Moreover, the PA–PB1 interaction represents an attractive target for inhibitors since the association of PA and PB1 subunits is essential for replication [85-88], and the sequence of these domains is highly conserved [89,90]. The NP is also an excellent target for antivirals that was exploited through modulation of its self-association and its RNA-binding function [31,44,45,53,57,59,60,98].

6.2 Potential of this research and exciting area of research

The potential for disruption of specific protein–protein interactions within the RNP holds a lot of promises for the design of new antivirals. The field benefits from a breakthrough with the exciting crystal structures of influenza A and B polymerase with bound viral RNA promoter [68,69] that now opens brand new perspectives for the design of new antivirals perturbing

Expert Opin. Drug Discov. Downloaded from informahealthcare.com by INSERM on 04/10/15
For personal use only.

RIGHTSLINK

the polymerase assembly. Additionally, RNA-binding sites on the polymerase could represent newly defined targets probably through a better mechanistic understanding of the conserved RNA 3′-end recognition.

6.3 Antiviral combinations as a promising area of research

Multiple combinations as successfully exemplified by the tri-therapies developed against HIV may apply to severe influenza cases [1]. At present, combination therapies using dual NA antivirals did not reduce resistance. In contrast, preclinical studies showed additive effects of NA inhibitors with antiviral drugs with different mechanisms of action [1]. Future drug combinations targeting the viral RNP may involve compounds that inhibit more than one interaction type (protein–protein, protein–RNA). Alternatively, antiviral combination that target both a viral protein and a cellular factor involved in the host response to Influenza virus, holds promises for improved results for the host, human or animal [1,2]. The targeted host protein could be a factor required for transcription/replication or factors involved in chaperoning or transport of the RNP or its constituents. Moreover, combination therapies may reduce doses, thereby reducing the side effects and importantly, resistance against cellular factor are not observed.

Exploiting drug promiscuity (drug binding to more than one target) may be an advantage in severe cases of influenza [1]. A 'two-in-one' combination was recently suggested for naproxen, an approved anti-inflammatory drug and also inhibits influenza A NP [59].

6.4 How can it facilitate the discovery of new compounds or reduce attrition rates?

Improving our understanding of the structural biology of the virus and the underlying mechanisms of the targeted proteins in the viral lifecycle would be extremely important in the near future. As for any drug development, overcoming toxicity is an issue requiring careful and relevant assays. Improving solubility and cellular targeting to the proper cellular compartment may decrease side effects.

6.5 Alternative technologies/approaches for understanding structural and mechanistic issues on influenza

Mechanistic assembly and function of multiprotein complexes may now benefit from high-resolution cryo-EM detectors. It is clear that function of the polymerase involves coordinated conformational changes and appropriate tools monitoring these movements in real time as time-resolved fluorescence, force measurements using single molecule techniques will flourish and greatly add to our understanding on influenza virus.

Declaration of interest

The authors are funded by the Agence Nationale de la Recherche, FLUNUCLEOVIR ANR 2010 Blan-130701 and RNAP-IAV ANR Blanc -14-CE09-0017. This work was granted access to the HPC resources of IDRIS made by Grand Equipment National de Calcul Intensif under the grant 2012-076378 attributed. This work used the platforms of the Grenoble Instruct Center (Integrated Structural Biology Grenoble; UMS3518 CNRS-CEA-UJF-EMBL) with support from The French Infrastructure for Integrated Structural Biology (ANR-10-INSB-05-02) and The Alliance Grenobloise pour la Biologie Structurale et Cellulaire Intégrées (ANR-10-LABX-49-01) within the Grenoble Partnership for Structural Biology. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

Bibliography

Papers of special note have been highlighted as either of interest (•) or of considerable interest (••) to readers.

1. Dunning J, Baillie JK, Cao B, et al. Antiviral combinations for severe influenza. Lancet Infect Dis 2014;14(12):1259-70
• **Review of preclinical and clinical studies involving combination antiviral therapy for severe influenza.**

2. Du J, Cross TA, Zhou HX. Recent progress in structure-based anti-influenza drug design. Drug Discov Today 2012;17(19-20):1111-20
• **Nice review of structure-based antiviral candidates targeting influenza proteins.**

3. Wang J, Wu Y, Ma C, et al. Structure and inhibition of the drug-resistant S31N mutant of the M2 ion channel of influenza A virus. Proc Natl Acad Sci USA 2013;110(4):1315-20
• **Structure of a potent inhibitor of matrix 2 (M2) of amantadine-resistant influenza.**

4. Das K. Antivirals targeting influenza a virus. J Med Chem 2012;55:6263-77
• **Review of structural domains of the polymerase and non-structural protein 1 (NS1) as antiviral targets and their inhibitors.**

5. Das K, Aramini JM, Ma LC, et al. Structures of influenza A proteins and insights into antiviral drug targets. Nat Struct Mol Biol 2010;17(5):530-8
• **Review of structural domains of the polymerase and NS1 as antiviral targets and their inhibitors.**

6. Loregian A, Mercorelli B, Nannetti G, et al. Antiviral strategies against influenza virus: towards new therapeutic approaches. Cell Mol Life Sci 2014;71(19):3659-83
• **Interesting overview of novel strategies targeting the virus and/or the host cell for counteracting influenza virus infection.**

7. Muller KH, Kakkola L, Nagaraj AS, et al. Emerging cellular targets for influenza antiviral agents. Trends Pharmacol Sci 2012;33(2):89-99

Expert Opin. Drug Discov. Downloaded from informahealthcare.com by INSERM on 04/10/15
For personal use only.

RIGHTSLINK

Expert Opin. Drug Discov. Downloaded from informahealthcare.com by INSERM on 04/10/15
For personal use only.

8. Smee DF, Bailey KW, Wong MH, et al. Treatment of influenza A (H1N1) virus infections in mice and ferrets with cyanovirin-N. Antiviral Res 2008;80(3):266-71

9. Brancato V, Peduto A, Wharton S, et al. Design of inhibitors of influenza virus membrane fusion: synthesis, structure-activity relationship and in vitro antiviral activity of a novel indole series. Antiviral Res 2013;99(2):125-35

10. Leneva IA, Russell RJ, Boriskin YS, et al. Characteristics of arbidol-resistant mutants of influenza virus: implications for the mechanism of anti-influenza action of arbidol. Antiviral Res 2009;81(2):132-40

11. Russell RJ, Kerry PS, Stevens DJ, et al. Structure of influenza hemagglutinin in complex with an inhibitor of membrane fusion. Proc Natl Acad Sci USA 2008;105(46):17736-41

12. Tang G, Lin X, Qiu Z, et al. Design and synthesis of benzenesulfonamide derivatives as potent anti-influenza hemagglutinin inhibitors. ACS Med Chem Lett 2011;2(8):603-7

13. Fleishman SJ, Whitehead TA, Ekiert DC, et al. Computational design of proteins targeting the conserved stem region of influenza hemagglutinin. Science 2011;332(6031):816-21

14. Basu A, Antanasijevic A, Wang M, et al. New small molecule entry inhibitors targeting hemagglutinin-mediated influenza a virus fusion. J Virol 2014;88(3):1447-60

15. Kim CU, Lew W, Williams MA, et al. Influenza neuraminidase inhibitors possessing a novel hydrophobic interaction in the enzyme active site: design, synthesis, and structural analysis of carbocyclic sialic acid analogues with potent anti-influenza activity. J Am Chem Soc 1997;119(4):681-90

16. Russell RJ, Haire LF, Stevens DJ, et al. The structure of H5N1 avian influenza neuraminidase suggests new opportunities for drug design. Nature 2006;443(7107):45-9

17. Spanakis N, Pitiriga V, Gennimata V, Taskris A. A review of neuraminidase inhibitor susceptibility in influenza strains. Expert Rev Anti Infect Ther 2014;12:1325-36

18. Vavricka CJ, Li Q, Wu Y, et al. Structural and functional analysis of laninamivir and its octanoate prodrug reveals group specific mechanisms for influenza NA inhibition. PLoS Pathog 2011;7(10):e1002249

19. Ivachtchenko AV, Ivanenkov YA, Mitkin OD, et al. Novel oral anti-influenza drug candidate AV5080. J Antimicrob Chemother 2014;69(7):1892-902

20. Vavricka CJ, Liu Y, Kiyota H, et al. Influenza neuraminidase operates via a nucleophilic mechanism and can be targeted by covalent inhibitors. Nat Commun 2013;4:1491

21. Kim JH, Resende R, Wennekes T, et al. Mechanism-based covalent neuraminidase inhibitors with broad-spectrum influenza antiviral activity. Science 2013;340(6128):71-5

• **Novel sialic acid derivatives active against pandemic H1N1 and avian H5N1 influenza that target the active site.**

22. Rudrawar S, Dyason JC, Rameix-Welti MA, et al. Novel sialic acid derivatives lock open the 150-loop of an influenza A virus group-1 sialidase. Nat Commun 2010;1:113

• **Novel sialic acid derivatives active against pandemic H1N1 and avian H5N1 influenza that target the active site.**

23. Mohan S, Kerry PS, Bance N, et al. Serendipitous discovery of a potent influenza virus a neuraminidase inhibitor. Angew Chem Int Ed Engl 2014;53(4):1076-80

24. Kerry PS, Mohan S, Russell RJ, et al. Structural basis for a class of nanomolar influenza A neuraminidase inhibitors. Sci Rep 2013;3:2871

25. Lee CM, Weight AK, Haldar J, et al. Polymer-attached zanamivir inhibits synergistically both early and late stages of influenza virus infection. Proc Natl Acad Sci USA 2012;109(50):20385-90

26. Schnell JR, Chou JJ. Structure and mechanism of the M2 proton channel of influenza A virus. Nature 2008;451(7178):591-5

27. Cady SD, Schmidt-Rohr K, Wang J, et al. Structure of the amantadine binding site of influenza M2 proton channels in lipid bilayers. Nature 2010;463(7281):689-92

28. Gu RX, Liu LA, Wei DQ. Structural and energetic analysis of drug inhibition of the influenza A M2 proton channel. Trends Pharmacol Sci 2013;34(10):571-80

29. Duque MD, Ma C, Torres E, et al. Exploring the size limit of templates for inhibitors of the M2 ion channel of influenza A virus. J Med Chem 2011;54(8):2646-57

30. Chenavas S, Estrozi LF, Delmas B, et al. Monomeric nucleoprotein of Influenza A virus. Plos Pathogens 2013;9(3):e1003275

• **Structure of monomeric nucleoprotein mutant R416A showing that the oligomerization loop folds in the RNA binding groove.**

31. Cianci C, Gerritz SW, Deminie C, et al. Influenza nucleoprotein: promising target for antiviral chemotherapy. Antivir Chem Chemother 2012;23(3):77-91

32. Davis AM, Chabolla BJ, Newcomb LL. Emerging antiviral resistant strains of influenza A and the potential therapeutic targets within the viral ribonucleoprotein (vRNP) complex. Virol J 2014;11:167

33. Arranz R, Coloma R, Chichon FJ, et al. The structure of native influenza virion ribonucleoproteins. Science 2012;338(6114):1634-7

•• **Cryo-microscopic studies of the ribonucleoprotein (RNP) and discovery of the nucleoprotein (NP)-RNA double-stranded structure.**

34. Moeller A, Kirchdoerfer RN, Potter CS, et al. Organization of the influenza virus replication machinery. Science 2012;338(6114):1631-4

•• **Cryo-microscopic studies of the RNP and an electrom microscopic (EM) model of the native influenza RNP.**

35. Coloma R, Valpuesta JM, Arranz R, et al. The structure of a biologically active influenza virus ribonucleoprotein complex. PLoS Pathog 2009;5(6):e1000491

•• **High-resolution cryo-EM structure of a recombinant RNP containing an RNA molecule of 248 nucleotides, 9 NP subunits and a polymerase complex. The atomic structures of NP are also given.**

36. Ng AK, Chan WH, Choi ST, et al. Influenza Polymerase Activity Correlates with the Strength of Interaction between Nucleoprotein and PB2 through the Host-Specific Residue K/E627. PLoS ONE 2012;7(5):e36415

RIGHTSLINK

Expert Opin. Drug Discov. Downloaded from informahealthcare.com by INSERM on 04/10/15
For personal use only.

37. Marklund JK, Ye Q, Dong J, et al. Sequence in the influenza A virus nucleoprotein required for viral polymerase binding and RNA synthesis. J Virol 2012;86(13):7292-7

38. Gui X, Li R, Zhang X, et al. An important amino acid in nucleoprotein contributes to influenza A virus replication by interacting with polymerase PB2. Virology 2014;464-465:11-20

39. Gavazzi C, Yver M, Isel C, et al. A functional sequence-specific interaction between influenza A virus genomic RNA segments. Proc Natl Acad Sci USA 2013;110(41):16604-9

40. Turrell L, Lyall JW, Tiley LS, et al. The role and assembly mechanism of nucleoprotein in influenza A virus ribonucleoprotein complexes. Nat Commun 2013;4:1591

41. Portela A, Digard P. The influenza virus nucleoprotein: a multifunctional RNA-binding protein pivotal to virus replication. J Gen Virol 2002;83(Pt 4):723-34

42. Newcomb LL, Kuo RL, Ye Q, et al. Interaction of the influenza a virus nucleocapsid protein with the viral RNA polymerase potentiates unprimed viral RNA replication. J Virol 2009;83(1):29-36

43. Ruigrok RW, Baudin F. Structure of influenza virus ribonucleoprotein particles. II. Purified RNA-free influenza virus ribonucleoprotein forms structures that are indistinguishable from the intact influenza virus ribonucleoprotein particles. J Gen Virol 1995;76(Pt 4):1009-14

44. Ng AK, Zhang H, Tan K, et al. Structure of the influenza virus A H5N1 nucleoprotein: implications for RNA binding, oligomerization, and vaccine design. FASEB J 2008;22(10):3638-47

- **Atomic structure of H5N1 nucleoprotein and mutational analysis of the RNA-binding site.**

45. Ye Q, Krug RM, Tao YJ. The mechanism by which influenza A virus nucleoprotein forms oligomers and binds RNA. Nature 2006;444(7122):1078-82

- **Atomic structure of H1N1 influenza virus NP showing a domain exchange polymerization domain and a putative RNA-binding surface.**

46. Ye Q, Guu TS, Mata DA, et al. Biochemical and structural evidence in support of a coherent model for the formation of the double-helical influenza A virus ribonucleoprotein. MBio 2012;4(1):e00467-12

47. Ng AK, Lam MK, Zhang H, et al. Structural basis for RNA binding and homo-oligomer formation by influenza B virus nucleoprotein. J Virol 2012;86(12):6758-67

48. Ruigrok RW, Crepin T, Kolakofsky D. Nucleoproteins and nucleocapsids of negative-strand RNA viruses. Curr Opin Microbiol 2011;14(4):504-10

49. Reguera J, Cusack S, Kolakofsky D. Segmented negative strand RNA virus nucleoprotein structure. Curr Opin Virol 2014;5:7-15

50. Reguera J, Malet H, Weber F, et al. Structural basis for encapsidation of genomic RNA by La Crosse Orthobunyavirus nucleoprotein. Proc Natl Acad Sci USA 2013;110(18):7246-51

51. Tarus B, Bakowiez O, Chenavas S, et al. Oligomerization paths of the nucleoprotein of influenza A virus. Biochimie 2012;94:776-85

52. Tarus B, Chevalier C, Richard CA, et al. Molecular dynamics studies of the nucleoprotein of influenza a virus: role of the protein flexibility in RNA binding. PLoS ONE 2012;7(1):e30038

53. Boulo SA, Lotteau H, Mueller V, et al. Human importin alpha and RNA do not compete for binding to Influenza A nucleoprotein. Virology 2011;409(1):84-90

54. Amorim MJ, Kao RY, Digard P. Nucleozin targets cytoplasmic trafficking of viral ribonucleoprotein-Rab11 complexes in influenza A virus infection. J Virol 2013;87(8):4694-703

55. Kao RY, Yang D, Lau LS, et al. Identification of influenza A nucleoprotein as an antiviral target. Nat Biotechnol 2010;28(6):600-5

- **Identification of nucleozin by high-throughput screening, first inhibitor of NP, enhancing non-functional oligomerization of NP and impeding its transport to the nucleus.**

56. Cheng H, Wan J, Lin MI, et al. Design, synthesis, and in vitro biological evaluation of 1H-1,2,3-triazole-4-carboxamide derivatives as new anti-influenza A agents targeting virus nucleoprotein. J Med Chem 2012;55(5):2144-53

57. Gerritz SW, Cianci C, Kim S, et al. Inhibition of influenza virus replication via small molecules that induce the formation of higher-order nucleoprotein oligomers. Proc Natl Acad Sci USA 2011;108(37):15366-71

- **Structure of a nucleozin derivative stabilizing the interface between two NP subunits.**

58. Shen YF, Chen YH, Chu SY, et al. E339 R416 salt bridge of nucleoprotein as a feasible target for influenza virus inhibitors. Proc Natl Acad Sci USA 2011;108(40):16515-20

- **Targeting the oligomerization loop of NP by new compounds to decrease its oligomerization.**

59. Lejal N, Tarus B, Bouguyon E, et al. Structure-Based Discovery of the Novel Antiviral Properties of Naproxen against the Nucleoprotein of Influenza A Virus. Antimicrob Agents Chemother 2013;57(5):2231-42

- **Competing with RNA binding to the NP by naproxen, a known inhibitor of COX2.**

60. Tarus B, Bertrand H, Zedda G, et al. Structure-based design of novel naproxen derivatives targeting monomeric nucleoprotein of Influenza A virus. J Biomol Struct Dyn 2014; [Epub ahead of print

61. Dias A, Bouvier D, Crepin T, et al. The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit. Nature 2009;458(7240):914-18

- **Atomic structure of the endonuclease domain of the PA subunit containing two bound manganese ions and biochemical demonstration that the isolated domain has intrinsic divalent cation-dependent nuclease activity.**

62. Yuan P, Bartlam M, Lou Z, et al. Crystal structure of an avian influenza polymerase PA(N) reveals an endonuclease active site. Nature 2009;458(7240):909-13

- **Atomic structure of the endonuclease domain in the PA subunit containing a single-bound magnesium ion. The biological function of this domain was confirmed by a mutational analysis using recombinant polymerase showing that mutants of the metal-ligand**

RIGHTSLINK

Expert Opin. Drug Discov. Downloaded from informahealthcare.com by INSERM on 04/10/15
For personal use only.

residues were active in replication but inactive in cap-dependent transcription.

63. He X, Zhou J, Bartlam M, et al. Crystal structure of the polymerase PA(C)-PB1 (N) complex from an avian influenza H5N1 virus. Nature 2008;454(7208):1123-6
- **Atomic structure of the C-terminal domain of the PA subunit in complex with a short N-terminal peptide from the PB1 subunit.**

64. Obayashi E, Yoshida H, Kawai F, et al. The structural basis for an essential subunit interaction in influenza virus RNA polymerase. Nature 2008;454(7208):1127-31
- **Atomic structure of the C-terminal domain of the PA subunit in complex with a short N-terminal peptide from the PB1 subunit.**

65. Sugiyama K, Obayashi E, Kawaguchi A, et al. Structural insight into the essential PB1-PB2 subunit contact of the influenza virus RNA polymerase. EMBO J 2009;28(12):1803-11

66. Guilligay D, Tarendeau F, Resa-Infante P, et al. The structural basis for cap binding by influenza virus polymerase subunit PB2. Nat Struct Mol Biol 2008;15(5):500-6
- **Atomic structure of the PB2 cap-binding domain in complex with 7-methylguanosine triphosphate as well as binding studies and analysis using recombinant polymerase showing that mutations of residues involved in cap-binding leads to the inactivation of cap-dependent transcription but not cap-independent transcription or replication. The paper also describes the ESPRIT method for screening soluble internal fragments of proteins.**

67. Tarendeau F, Crepin T, Guilligay D, et al. Host determinant residue lysine 627 lies on the surface of a discrete, folded domain of influenza virus polymerase PB2 subunit. PLoS Pathog 2008;4(8):e1000136
- **Atomic structure of the C-terminal double domain of the PB2 subunit rich in amino acids that are mutated when avian viruses adapt to mammals. The most prominent of these is residue 627 that is a lysine in viruses that infect the human respiratory tract. The human virus domain contains an extensive positively charged surface that is disrupted by a glutamate substitution.**

68. Pflug A, Guiligay D, Reich S, et al. Structure of influenza A polymerase bound to the viral RNA promoter. Nature 2014;516(7531):355-60.
- **Atomic structure of the C-terminal double domain of the PB2 subunit rich in amino acids that are mutated when avian viruses adapt to mammals. The most prominent of these is residue 627 that is a lysine in viruses that infect the human respiratory tract. The human virus domain contains an extensive positively charged surface that is disrupted by a glutamate substitution.**

69. Reich S, Guilligay D, Malet H, et al. Structural insight into cap-snatching and RNA synthesis by influenza polymerase. Nature 2014;516(7531):361-6.
- **Structures of influenza B polymerase and evidence of movements of PB2 cap-binding domains. Conformational changes required to initiate polymerase function are suggested.**

70. Pautus S, Sehr P, Lewis J, et al. New 7-methylguanine derivatives targeting the influenza polymerase PB2 cap-binding domain. J Med Chem 2013;56(21):8915-30

71. Clark MP, Ledeboer MW, Davies I, et al. Discovery of a novel, first-in-class, orally bioavailable azaindole inhibitor (VX-787) of influenza PB2. J Med Chem 2014;57(15):6668-78

72. Wathen MW, Barro M, Bright RA. Antiviral in seasonal and pandemic influenza-future perspectives. Influenza Other Respir Viruses 2012;7(Suppl 1):76-80

73. Bauman JD, Patel D, Baker SF, et al. Crystallographic fragment screening and structure-based optimization yields a new class of influenza endonuclease inhibitors. ACS Chem Biol 2013;8:2501-8

74. Kowalinski E, Zubieta C, Wolkerstorfer A, et al. Structural analysis of specific metal chelating inhibitor binding to the endonuclease domain of influenza pH1N1 (2009) polymerase. PLoS Pathog 2012;8(8):e1002831
- **Structure of PA endonuclease domain with a series of specific inhibitors, including four diketo compounds and a green tea catechin.**

75. DuBois RM, Slavish PJ, Baughman BM, et al. Structural and biochemical basis for development of influenza virus inhibitors targeting the PA endonuclease. PLoS Pathog 2012;8(8):e1002830
- **Inhibitors of PA endonuclease.**

76. Parhi AK, Xiang A, Bauman JD, et al. Phenyl substituted 3-hydroxypyridin-2 (1H)-ones: inhibitors of influenza A endonuclease. Bioorg Med Chem 2013;21(21):6435-46

77. Sagong HY, Bauman JD, Patel D, et al. Phenyl Substituted 4-Hydroxypyridazin-3 (2H)-ones and 5-Hydroxypyrimidin-4 (3H)-ones: inhibitors of Influenza A Endonuclease. J Med Chem 2014;57(19):8086-98

78. Sagong HY, Parhi A, Bauman JD, et al. 3-Hydroxyquinolin-2(1H)-ones As Inhibitors of Influenza A Endonuclease. ACS Med Chem Lett 2013;4(6):547-50

79. Tefsen B, Lu G, Zhu Y, et al. The N-terminal domain of PA from bat-derived influenza-like virus H17N10 has endonuclease activity. J Virol 2014;88(4):1935-41

80. Tomassini J, Selnick H, Davies ME, et al. Inhibition of cap (m7GpppXm)-dependent endonuclease of influenza virus by 4-substituted 2,4-dioxobutanoic acid compounds. Antimicrob Agents Chemother 1994;38(12):2827-37

81. Muratore G, Mercorelli B, Goracci L, et al. Human cytomegalovirus inhibitor AL18 also possesses activity against influenza A and B viruses. Antimicrob Agents Chemother 2012;56(11):6009-13

82. Muratore G, Goracci L, Mercorelli B, et al. Small molecule inhibitors of influenza A and B viruses that act by disrupting subunit interactions of the viral polymerase. Proc Natl Acad Sci USA 2012;109(16):6247-52

83. Pagano M, Castagnolo D, Bernardini M, et al. The fight against the influenza A virus H1N1: synthesis, molecular modeling, and biological evaluation of benzofurazan derivatives as viral RNA polymerase inhibitors. ChemMedChem 2014;9(1):129-50

84. Tintori C, Laurenzana I, Fallacara AL, et al. High-throughput docking for the identification of new influenza A virus polymerase inhibitors targeting the PA-PB1 protein-protein interaction.

RIGHTSLINK

Bioorg Med Chem Lett 2014;24(1):280-2

- **Novel inhibitors targeting the PA–PB1 protein–protein interaction.**

85. Brownlee GG, Sharps JL. The RNA polymerase of influenza a virus is stabilized by interaction with its viral RNA promoter. J Virol 2002;76(14):7103-13
86. Fodor E, Crow M, Mingay LJ, et al. A single amino acid mutation in the PA subunit of the influenza virus RNA polymerase inhibits endonucleolytic cleavage of capped RNAs. J Virol 2002;76(18):8989-9001
87. Perales B, Ortin J. The influenza A virus PB2 polymerase subunit is required for the replication of viral RNA. J Virol 1997;71(2):1381-5
88. Perez DR, Donis RO. Functional analysis of PA binding by influenza a virus PB1: effects on polymerase activity and viral infectivity. J Virol 2001;75(17):8127-36
89. Ghanem A, Mayer D, Chase G, et al. Peptide-mediated interference with influenza A virus polymerase. J Virol 2007;81(14):7801-4
90. Wunderlich K, Mayer D, Ranadheera C, et al. Identification of a PA-binding peptide with inhibitory activity against influenza A and B virus replication. PLoS ONE 2009;4(10):e7517
91. Li C, Ba Q, Wu A, et al. A peptide derived from the C-terminus of PB1 inhibits influenza virus replication by interfering with viral polymerase assembly. FEBS J 2013;280(4):1139-49
92. Engel DA. The influenza virus NS1 protein as a therapeutic target. Antiviral Res 2013;99(3):409-16
93. Yin C, Khan JA, Swapna GV, et al. 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 2007;282(28):20584-92
94. Cho EJ, Xia S, Ma LC, et al. Identification of influenza virus inhibitors targeting NS1A utilizing fluorescence polarization-based high-throughput assay. J Biomol Screen 2012;17(4):448-59
95. You L, Cho EJ, Leavitt J, et al. Synthesis and evaluation of quinoxaline derivatives as potential influenza NS1A protein inhibitors. Bioorg Med Chem Lett 2011;21(10):3007-11
96. Qin S, Liu Y, Tempel W, et al. Structural basis for histone mimicry and hijacking of host proteins by influenza virus protein NS1. Nat Commun 2014;5:3952
97. Cencic R, Hall DR, Robert F, et al. Reversing chemoresistance by small molecule inhibition of the translation initiation complex eIF4F. Proc Natl Acad Sci USA 2011;108(3):1046-51
98. Chenavas S, Crepin T, Delmas B, et al. Influenza virus nucleoprotein: structure, RNA binding, oligomerization and antiviral drug target. Future Microbiol 2013;8:1537-45

- **Structure of monomeric nucleoprotein (NP)-mutant R416A showing that the oligomerization loop folds in the RNA-binding groove.**

Affiliation

Alexandre Monod[1], Christopher Swale[1], Bogdan Tarus[2], Alice Tissot[1], Bernard Delmas[2], Rob WH Ruigrok[1], Thibaut Crépin[1] & Anny Slama-Schwok[†2]
[†]Author for correspondence
[1]University of Grenoble Alpes-EMBL-CNRS, Unit for Virus Host-Cell Interactions, 71 avenue des Martyrs, 38042 Grenoble, France
[2]Virologie et Immunologie Moléculaires, INRA UR 892, 78350, Jouy en Josas, France
E-mail: Anny.Schwok@gmail.com

Expert Opin. Drug Discov. Downloaded from informahealthcare.com by INSERM on 04/10/15
For personal use only.

RIGHTSLINK

RESUME

Caractérisation structurale et biochimique de la Nucléoprotéine des virus grippaux de type A, B et D

Le virus de la grippe est un virus à ARN négatif appartenant à la famille des *Orthomyxoviridae* qui se compose de 7 membres dont les virus *influenza* A, B, C et D. Le génome viral comprend 7 à 8 particules ribonucléoprotéiques (RNP) au sein desquelles l'ARN viral (ARNv) est recouvert de multiples copies de nucléoprotéines (NP) et est associé à l'ARN polymérase virale *via* ses extrémités 3' et 5'. Au cours de ce travail de thèse, nous nous sommes tout d'abord focalisés sur l'étude biochimique de NP A et NP B et avons pu mettre en évidence des comportements différents en ce qui concerne leurs propriétés d'oligomérisation en présence ou en absence d'ARN et en fonction de la concentration en sel. Pour la première fois nous avons pu observer une structure similaire aux RNP mais reconstituée uniquement à partir de NP A et d'un ARN de 12 nucléotides. Nous avons pu formuler l'hypothèse que 12 nucléotides de l'ARN serait fixés à la NP avec une forte affinité tandis que le reste de l'ARN fixerait la NP avec une affinité beaucoup plus faible. En parallèle nous avons résolu la structure cristallographique de la nucléoprotéine de la grippe de type D et réaliser la caractérisation de son interaction avec l'importine-α7 humaine. Enfin nous avons étudié la fixation de l'ARN sur NP D et mis en évidence l'importance de l'extrémité C-terminale dans le processus de fixation à l'ARN. Ces informations ont permis de formuler de nouvelles hypothèses quant au fonctionnement du virus de la grippe et permettre d'inscrire ce projet de thèse dans une dynamique globale de lutte contre ce virus.

ABSTRACT

Structural and biochemical characterization of Nucleoprotein of influenza A, B and D viruses

Influenza virus is a negative RNA virus belongs to the *Orthomyxoviridae* family which consists of 7 members including influenza viruses A, B, C and D. The viral genome comprises 7 to 8 ribonucleoprotein particles (RNP) in which the viral RNA (vRNA) is coated with multiple copies of nucleoproteins (NP) and is associated with the viral RNA polymerase by its 3 'and 5' ends. In this thesis, we first focused on the biochemical study of NP A and NP B and we demonstrate that there are different behaviors with regard to their oligomerization properties in the presence or absence of RNA and as a function of the salt concentration. For the first time we were able to observe a structure very similar to RNP which was reconstituted only from NP A and a 12 nucleotide RNA. Thus, we formulate the hypothesis that 12 nucleotides of the RNA would bind NP with a very strong affinity while the rest of the RNA would bind NP with a lower affinity. In parallel, we solved the crystallographic structure of the nucleoprotein of influenza D virus and we characterized its interaction with human importin-α7. Finally, we studied the binding of RNA on NP D and we demonstrated the importance of the C-terminal end in the RNA binding process. This thesis project made it possible to formulate new hypotheses concerning the functioning of the influenza virus and to include this thesis project in a global dynamic of combating the influenza virus