



**HAL**  
open science

# Interaction between the influenza replication machineries and the cellular nuclear import system

Amélie Donchet

► **To cite this version:**

Amélie Donchet. Interaction between the influenza replication machineries and the cellular nuclear import system. Structural Biology [q-bio.BM]. Université Grenoble Alpes [2020-..], 2021. English. NNT : 2021GRALV003 . tel-03246771

**HAL Id: tel-03246771**

**<https://theses.hal.science/tel-03246771v1>**

Submitted on 2 Jun 2021

**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.

## THÈSE

Pour obtenir le grade de

**DOCTEUR DE L'UNIVERSITE GRENOBLE ALPES**

Spécialité : **Biologie Structurale et Nanobiologie**

Arrêté ministériel : 25 mai 2016

Présentée par

**Amélie DONCHET**

Thèse dirigée par le Docteur Thibaut CREPIN

préparée au sein de l'Institut de Biologie Structurale (IBS)  
dans l'École Doctorale Chimie et Science du Vivant (EDCSV)

# Interaction between the influenza replication machineries and the nuclear import system

Thèse soutenue publiquement le 19 janvier 2021,  
devant le jury composé de :

**Docteur Sandie MUNIER**

MC, Université de Paris, Rapportrice

**Docteur Roland MARQUET**

DR, CNRS, Rapporteur

**Professeur Martin SCHWEMMLE**

PR, University of Freiburg, Examineur

**Professeur Winfried WEISSEHORN**

PR, Université Grenoble Alpes, Président du Jury

**Docteur Thibaut CREPIN**

DR, CNRS, Directeur de Thèse





## Remerciements

Je tiens à remercier tous ceux qui m'ont accompagnée lors de cette aventure intense, palpitante et riche en enseignements.

Je souhaite tout d'abord remercier Thibaut Crépin, mon directeur de thèse, pour sa confiance, ses conseils, et le précieux temps qu'il m'a toujours accordée. Merci de m'avoir guidé sur ce chemin ces quatre années. Merci d'avoir relevé ce défi et de m'avoir permis de devenir une vraie biochimiste !

Je veux également remercier Rob Ruigrok, pour tous les riches échanges que l'on a pu avoir, pour ses conseils, sa présence, sa bienveillance et son soutien.

Merci à Laura, Christopher et Alice, pour m'avoir accueillie, formée et présentée leurs projets. Merci de m'avoir mis le pied à l'étrier !

Un grand merci à tous les membres du laboratoire, pour tous les bons moments. Plus particulièrement, merci à Jean-Marie, Francine, Pascal, Emilie, Philippe, Nicolas, Caroline pour leur temps, pour toutes nos discussions, scientifiques et un peu moins, et pour leur aide très précieuse.

Je souhaite aussi adresser un merci spécial à Carla "Carliña" et Nicolas "Jojo" sans qui ces années au laboratoire auraient été bien différentes et bien plus pâles. Merci pour ces bouffées d'oxygène, pour les rires et les blagues, pour le soutien essentiel lorsque la science faisait de la résistance !

Merci à tous les petits piou-pious qui sont passés en stage au laboratoire (et particulièrement à Mai, Justine et Bastien pour leur aide) et aux grands piou-pious qui ont commencé leur thèse. Les prochains, c'est vous ! ;)

Je souhaite remercier également Antoine (parmi tant d'autres, le Atlas de mon P-body, qui défie littéralement les lois de la probabilité et retire cruellement le pont sous ses pieds !), Audrey (ma force tranquille dans le chaos). Il n'y a pas besoin de mot, je pense, car cela va au-delà de ça, pour te remercier Alice. Merci pour tout.

Merci aux copains de promo, aux thésards d'ici et d'ailleurs, à la team Glob'Alps et BIOTechno pour tous les bons instants.

Merci à mes parents qui m'ont permis d'arriver jusqu'ici de la meilleure des façons, et à toute la famille qui a toujours essayé de comprendre ce que je pouvais bien faire au labo, mais qui a adoré les photos de cristaux et de protéines fluo !

Merci d'avoir fait un bout de chemin avec moi, et quel chemin ! Thank you all!

**Amélie**

*A mes étoiles,*

*à mes petits héros.*



# Table of Contents



<b><u>List of abbreviations</u></b> .....	<b>9</b>
<b><u>List of figures</u></b> .....	<b>13</b>
<b><u>List of tables</u></b> .....	<b>17</b>
<b><u>Résumé détaillé</u></b> .....	<b>21</b>
<b><u>Chapter 1: Introduction</u></b> .....	<b>31</b>
<b>I. Influenza viruses</b> .....	<b>33</b>
A. Features of influenza viruses .....	33
i. Taxonomy .....	33
ii. Viral particle structure .....	34
iii. The ribonucleoproteins, the replicative units of the virus .....	37
iv. Genome organization, protein content and host tropism .....	37
v. Evolution: antigenic drift and antigenic shift .....	42
vi. Disease and epidemiology .....	43
B. Influenza virus cycle .....	45
C. Viral transcription and replication machinery .....	48
i. Nucleoprotein.....	48
ii. Heterotrimeric RNA-dependent RNA-polymerase .....	51
<b>II. Nucleocytoplasmic shuttling of proteins</b> .....	<b>56</b>
A. Nuclear transport pathways .....	57
B. The nuclear transport pathway based on Ran and karyopherins .....	58
i. The Ras-like small GTPase Ran .....	58
ii. The nuclear import and export: importins- $\alpha$ , importins- $\beta$ , and exportins .....	58
iii. The nuclear localization and export signals .....	60
C. Similarities and specificities of importins- $\alpha$ .....	61
D. <i>In vitro</i> and <i>in cellulo</i> interaction parameters of the complex between cargo and importins: affinity and rates .....	66
E. Nuclear import regulation through cellular machinery .....	67
i. Importin- $\alpha$ :importin- $\beta$ 1 sub-complex versus importins- $\beta$ as transport receptors .....	67
ii. Post-translational modifications .....	68
iii. Other regulatory mechanisms .....	69
<b>III. Nuclear import of the influenza replication machinery</b> .....	<b>70</b>
A. Nucleoprotein .....	70
B. Heterotrimeric RNA-dependent RNA polymerase .....	73
C. Ribonucleoproteins.....	74
<b><u>Objectives of the PhD thesis</u></b> .....	<b>77</b>
<b><u>Chapter 2: Structural analysis of the influenza D nucleoprotein and characterization of its nuclear localization signal</u></b> .....	<b>81</b>
<b><u>Chapter 3: Comparative analysis of the interaction involving each influenza nucleoprotein types and importins-<math>\alpha</math> involved in influenza nuclear import</u></b> .....	<b>99</b>
<b><u>Chapter 4: Structural and biochemical characterization of the Asiatic toad influenza NP</u></b> .....	<b>119</b>



I. Characterization of the interaction with importins- $\alpha$ .....	122
II. Characterization of the interaction with RNAs.....	125
III. Structural characterization of Toad/NP .....	128
IV. Discussion .....	131
<b><u>Conclusions</u></b> .....	<b>137</b>
<b><u>Annexes</u></b> .....	<b>141</b>
Annexe 1: Sequence identities between importins- $\alpha$ .....	142
Annexe 2: Conservation of NP <sub>TAIL</sub> within each influenza type.....	143
Annexe 3: Protein accession numbers .....	147
Annexe 4: Supplementary Data Chapter 2.....	148
Annexe 5: Supplementary Data Chapter 3.....	154
<b><u>Bibliography</u></b> .....	<b>162</b>

## List of abbreviations



ARM repeats: Armadillo repeats  
BLAST: Basic Local Alignment Search Tool  
BRD: Bovine Respiratory Disease  
CAS: Cellular Apoptosis Susceptibility  
CCP4: Collaborative Computational Project Number 4  
CCP4i: CCP4 Interface  
cNLS: conventional/classical Nuclear Localization Signal  
COOT: Crystallographic Object-Oriented Toolkit  
Crm1: Chromosome region maintenance 1 protein  
cRNA: complementary RNA  
CTD: C-terminal domain  
DNA: Deoxyribonucleic Acid  
EM: Electron Microscopy  
EMBL: European Molecular Biology Laboratory  
ESCRT: Endosomal Sorting Complex Required for Transport  
FA: Fluorescence Anisotropy  
FAM: Fluorescein  
FG-repeats: phenylalanine-glycine repeats  
GDP: Guanosine DiPhosphate  
GTP: Guanosine TriPhosphate  
HA: Hemagglutinin  
HEAT repeats: **H**untington, **E**longation factor 3, **P**R65/**A**, **T**OR  
HEF: Hemagglutinin-Esterase-Fusion  
HPAIs: Highly Pathogenic Avian Influenza viruses  
HSQC: Heteronuclear Single Quantum Correlation  
HTX: High Throughput Crystallization laboratory  
IAV: Influenza A Viruses  
IBB domain: Importin- $\beta$  Binding domain  
IBV: Influenza B Viruses  
ICTV: International Committee on Taxonomy of Viruses  
ICV: Influenza C Viruses  
IDV: Influenza D Viruses  
IPTG: isopropyl- $\beta$ -D-thiogalactopyranoside  
IRD: Influenza Research Database  
Isavirus: Infection salmon anemia virus  
ISBG: Integrated Structural Biology Grenoble  
 $K_d$ : Dissociation constant  
kDa: kilo Daltons  
LPAIs: Low-Pathogenic Avian Influenza viruses  
M1: Matrix protein 1  
M2: Matrix protein 2  
MAD: Multi-wavelength Anomalous Dispersion  
MDa: Mega Daltons  
mRNA: messenger RNA  
MW: Molecular Weight  
MxA: Human myxovirus resistance protein 1  
NA: Neuraminidase  
NCBI: National Center for Biotechnology Information  
ncNLS: non-conventional/non-classical Nuclear Localization Signal  
NEP: Nuclear Export Protein  
NES: Nuclear Export Signal  
Neu5Ac: N-Acetylneuraminic acid  
NLS: Nuclear Localization Signal

NMR: Nuclear Magnetic Resonance  
NP: Nucleoprotein  
NPC: Nuclear Pore Complex  
NS1: Non-Structural protein 1  
PA: Polymerase Acidic protein  
PB1: Polymerase Basic protein 1  
PB2: Polymerase Basic protein 2  
PEG: PolyEthylene Glycol  
pH: potential of Hydrogen  
pHi: isoelectric point  
PoII: DNA-dependent RNA polymerase II  
PTM: Post-Translational Modification  
PY NLS: Proline-Tyrosine rich Nuclear Localization Signal  
RanBP1: Ran-Binding Protein 1  
RanBP5: Ran-Binding Protein 5  
RanGAP: Ran GTPase-Activating Protein  
RCC1/RanGEF: Regulator of Chromosome Condensation 1 / Ran Guanine nucleotide Exchange Factor  
RdRp: RNA-dependent RNA polymerase  
RNA: Ribonucleic Acid  
RNP: Ribonucleoprotein  
SAXS: Small Angle X-rays Scattering  
SEC: Size Exclusion Chromatography  
SEC-MALLS: Size Exclusion Chromatography coupled with Multi Angle Laser Light Scattering  
SNP: Single Nucleotide Polymorphism  
SPR: Surface Plasmon Resonance  
STEM: Scanning Transmission Electron Microscopy  
SUMO: Small Ubiquitin-like Modifier  
TEV: Tobacco Etch Virus  
ToadV: Wuhan Asiatic Toad influenza Virus  
TEM: Transmission electron microscopy  
vRNA: viral RNA  
WHO: World Health Organization  
YFP: Yellow Fluorescent Protein

## List of figures



Figure 1: Viral particle structure and genome organization.....	35
Figure 2: Main and accessory proteins encoded by influenza A viral mRNAs. ....	38
Figure 3: Host tropism of influenza viruses. ....	39
Figure 4: Structure of influenza HA, HEF, NA, M2 and M1 proteins.....	41
Figure 5: Major historical outbreaks in the shared human and influenza history. ....	44
Figure 6: Influenza viral cycle.....	46
Figure 7: Biochemical features of influenza nucleoproteins.....	49
Figure 8: Influenza nucleoprotein structures and structural features. ....	50
Figure 9: Structure of Isavirus NP.....	51
Figure 10: Bat IAV, IBV, ICV and IDV RdRp structures. ....	52
Figure 11: Reconstituted model of influenza RdRp transcription.....	53
Figure 12: Replication of the influenza genome.....	54
Figure 13: Nuclear pore complex (NPC). ....	56
Figure 14: The nuclear transport pathway based on Ran.....	59
Figure 15: Evolution, diversification and global organization of importins- $\alpha$ .....	62
Figure 16: Structure of importins- $\alpha$ .....	63
Figure 17: Diversity of the roles played by importins- $\alpha$ in cells.....	65
Figure 18: Summary of the nuclear import of the influenza replication machinery.....	72
Figure 19: Toad/NP constructs.....	121
Figure 20: Profiles of the SEC purification step for each Toad/NP constructs.....	122
Figure 21: Interaction assays between Toad/NP constructs and human importin- $\alpha$ 7 by size exclusion chromatography. ....	123
Figure 22: Characterization of the interaction between Toad/NP <sub>TAIL</sub> -YFP and the human importin- $\alpha$ 7 by SPR. ....	124
Figure 23: Titration against FAM-labelled RNAs of Toad/NP <sub>FULL</sub> and Toad/NP <sub>CORE</sub> by fluorescence anisotropy at 150 mM NaCl. ....	126
Figure 24: <sup>1</sup> H- <sup>15</sup> N HSQC spectra of <sup>15</sup> N-labelled Toad/NP <sub>TAIL</sub> . ....	129
Figure 25: Crystals of Selenomethionine-labelled Toad/NP <sub>TAIL</sub> and Toad/NP <sub>TAIL</sub> solved structure.....	130
Figure 26: Crystals of Selenomethionine-labelled Toad/NP <sub>CORE</sub> and its crystallization condition.....	131
Figure 27: Structural features and comparison between Toad/NP and B/NP. ....	132
Toad/NP N-terminal structure is a result of this PhD work whereas B/NP modelling was published by <sup>177</sup> . ....	132
Figure 28: Diversity of newly identified NPs.....	135





## List of tables



<b>Table 1: Full taxonomy of the Orthomyxoviridae family members. ....</b>	<b>33</b>
<b>Table 2: Summary of the influenza main features depending on the virus type.....</b>	<b>36</b>
<b>Table 3: Sequence identities between influenza types.....</b>	<b>48</b>
<b>Table 4: Reference NLS sequences and consensus sequences for 6 classes of conventional and nonconventional basic NLSs interacting with importins-<math>\alpha</math>. ....</b>	<b>60</b>
<b>Table 5: NLS sequences of the influenza A replication machinery. ....</b>	<b>70</b>
<b>Table 6: Sequence identities between Toad/NP and the other influenza NPs.....</b>	<b>120</b>
<b>Table 7: Toad/NP<sub>TAIL</sub> sequences. ....</b>	<b>121</b>
<b>Table 8: Biacore steady state affinities between Toad/NP<sub>TAIL</sub> or B/NP<sub>TAIL</sub> and human <math>\Delta</math>IBB importins-<math>\alpha</math>.....</b>	<b>125</b>
<b>Table 9: Toad/NP:RNA apparent dissociation constants determined by fluorescence anisotropy.....</b>	<b>127</b>
<b>Table 10: Sequence identities between influenza NP types. ....</b>	<b>134</b>
<b>Table S1: Sequence identities between importins-<math>\alpha</math> paralogs of distinct species of interest. ....</b>	<b>142</b>
<b>Table S2: Consensus residues and frequencies of A/NP<sub>TAIL</sub>. ....</b>	<b>143</b>
<b>Table S3: Consensus residues and frequencies of B/NP<sub>TAIL</sub>. ....</b>	<b>144</b>
<b>Table S4: Consensus residues and frequencies of C/NP<sub>TAIL</sub>. ....</b>	<b>145</b>
<b>Table S5: Consensus residues and frequencies of D/NP<sub>TAIL</sub>. ....</b>	<b>146</b>
<b>Table S6: Protein accession numbers. ....</b>	<b>147</b>



## Résumé détaillé



Les virus grippaux font partie de la famille des *Orthomyxoviridae*. Cette famille comporte sept genres de virus, dont les virus de la grippe de type A, B, C et D (IAV, IBV, ICV et IDV). Cette classification est destinée à évoluer, car la recherche n'a fait qu'effleurer la surface de la diversité virale : une étude de 2018 a identifié 3 nouveaux membres de la famille des *Orthomyxoviridae*, notamment un virus de crapaud à Wuhan. Ces trois virus présentent des identités de séquence plus importantes avec IBV qu'avec les autres virus grippaux. Le virus de crapaud, étudié durant ce travail de thèse, sera appelé ToadV. Les virus grippaux possèdent une solide base commune en terme de structure et de fonctionnement, mais ils ont également de claires spécificités.

Les virus grippaux sont des virus enveloppés, à génome ARN segmenté, simple brin et de polarité négative. L'unité répliquative du virus, appelée ribonucléoprotéine (RNP), est composée du segment génomique, encapsidé par de multiples copies de la nucléoprotéine (NP) et associé par ses extrémités 3' et 5' à une ARN polymérase ARN dépendante (RdRp) virale hétérotrimérique. NP est la protéine la plus abondante des RNPs et elle est nécessaire aux activités de transcription et de réplication virales. NP est également impliquée dans la protection du génome viral, dans l'interaction avec les partenaires cellulaires et dans le trafic cellulaire des RNPs. NP est une protéine virale centrale, dont les fonctions clés sont conservées entre les NPs de IAV, IBV, ICV et IDV (A/NP, B/NP, C/NP et D/NP) mais avec des spécificités structurales et fonctionnelles dépendantes du type viral.

Les ions et les petites molécules (< 30-60 kDa) diffusent librement dans le noyau des cellules. Cependant, les protéines plus grosses nécessitent d'être importées par transport actif. La voie la plus étudiée est le système impliquant la petite GTPase Ran et les récepteurs de transport nucléaire appelés importines- $\alpha$  et - $\beta$ . Les importines- $\alpha$ , agissant comme des adaptateurs, interagissent avec un signal spécifique de la protéine à importer (cargo), appelé signal de localisation nucléaire (NLS), et avec l'importine- $\beta$ , *via* son domaine d'interaction avec l'importine- $\beta$  (domaine IBB). Ce domaine IBB, en absence de l'importine- $\beta$ , agit comme auto-inhibiteur en se repliant dans les deux sillons de l'importine- $\alpha$  où les NLSs peuvent se fixer. Certaines importines- $\beta$  n'ont pas besoin d'adaptateurs et peuvent directement interagir avec le NLS du cargo. Il existe deux types de NLS, les NLS classiques/conventionnels et les NLS non classiques/non conventionnels. Un NLS est généralement riche en Arginine



et en Lysine, et généralement composé d'un ou deux motifs basiques (NLS monopartite ou bipartite respectivement). Les importines sont très conservées au cours de l'évolution. Chez l'homme, il y a sept paralogues d'importines- $\alpha$  et au moins dix importines- $\beta$  responsables de l'import nucléaire. Cette diversité s'accompagne d'une spécificité de cargos, de comportement et d'expression dans les cellules et tissus : les importines ne sont pas interchangeableables.

Les virus grippaux font partie des rares virus à ARN à répliquer et transcrire leur génome dans le noyau de la cellule hôte. La machinerie répliquative ainsi que le génome ont donc besoin d'être importés dans le noyau, et pour cela détournent la machinerie des importines- $\alpha$  et - $\beta$ . PB1, PB2 et NP possèdent des NLS pour interagir avec les importines : PB2 et NP avec les importines- $\alpha$ , PB1 avec une importine- $\beta$  RanBP5, et PA est essentielle pour l'import efficace de PB1. L'import initial du génome dans le noyau sous forme de RNPs est supposé être médié par la NP, un autre des rôles essentiels de cette protéine virale dans le trafic viral.

Le NLS fonctionnel de A/NP et B/NP a été identifié sur l'extrémité intrinsèquement désordonnée N-terminale d'une longueur de 21 et 71 acides aminés respectivement (A/NP<sub>TAIL</sub> et B/NP<sub>TAIL</sub>). Le NLS de A/NP appartient aux NLS non conventionnels, interagit avec les importines- $\alpha$ 1, - $\alpha$ 3, - $\alpha$ 5 et - $\alpha$ 7, notamment avec une faible affinité pour l'importine- $\alpha$ 1 (constante de dissociation entre 1 et 5  $\mu$ M). De nombreux motifs basiques sont présents sur B/NP<sub>TAIL</sub> et les résidus 30 à 71 ont été identifiés comme participant à l'interaction avec l'importine- $\alpha$ 7, qui est caractérisée par une constante de dissociation de 844 nM. C/NP et D/NP, de leur côté, présentent une longue extrémité intrinsèquement désordonnée de 62 et 52 acides aminés respectivement, mais en C-terminale. C/NP<sub>TAIL</sub> et D/NP<sub>TAIL</sub> possèdent des motifs basiques qui ressemblent clairement à des NLS conventionnels bipartite. Des travaux publiés durant ce travail de thèse par une équipe concurrente ont montré que C/NP<sub>TAIL</sub> interagit avec les importines- $\alpha$ 1, - $\alpha$ 3, - $\alpha$ 4, - $\alpha$ 5 and - $\alpha$ 7, et son affinité pour l'importine- $\alpha$ 1 est de 50 nM.

Au début de ce travail de thèse, la caractérisation de la machinerie répliquative des virus de type D, et plus particulièrement de D/NP, était déjà initiée par l'équipe, en parallèle de ses travaux sur A/NP et B/NP. Ces dernières présentaient des affinités pour les importines- $\alpha$  trop faibles pour être considérées complètement fonctionnelles.

L'existence de ToadV et de sa NP n'était pas encore connue. Malgré le travail effectué sur les NPs grippales et leurs interactions avec les importines- $\alpha$ , de nombreuses questions restaient en suspens, d'autant plus que le mécanisme de fonctionnement d'une NP, ainsi que ses caractéristiques structurales et biochimiques, ne sont pas forcément transposables aux autres. Il a été décidé de caractériser dans un premier temps D/NP, structuralement ainsi que dans ses fonctions clés. Par la suite, dans l'optique de pouvoir comprendre comment les différents types grippaux accédaient au noyau et en quoi leurs stratégies convergeaient et/ou divergeaient, une stratégie expérimentale a été mise au point dans le but de pouvoir comparer aussi fiablement que possible les différentes NP<sub>TAILS</sub> et les différentes importines- $\alpha$  humaines. Enfin, avec la découverte inédite du virus grippal de crapaud, proche phylogénétiquement du virus de type B strictement humain, il a semblé judicieux d'initier sa caractérisation, aussi bien structurale que fonctionnelle, afin d'intégrer une vision évolutive aux fonctions cœur de la NP grippale et de mieux comprendre le fonctionnement de chaque NP individuellement.

Découvert en 2011 et classifié comme un nouveau genre grippal en 2013, le virus de la grippe de type D circule largement au niveau mondial chez les bovins. Peu de choses sont connues sur les protéines virales et son cycle.

A l'aide de la diffraction des rayons-X, de la microscopie électronique, du dichroïsme circulaire, de la chromatographie d'exclusion de taille couplée ou non à la diffusion de la lumière à plusieurs angles, de tests de stabilisation thermique, de l'anisotropie de fluorescence et de la résonance des plasmons de surface, la caractérisation de D/NP, à savoir ses activités d'interactions avec l'ARN et avec l'importine- $\alpha 7$ , son comportement oligomérique et sa structure, a été réalisée.

D/NP forme principalement des tétramères en solution, et la concentration en sel n'a pas d'impact sur cet état oligomérique. La structure a été résolue à une résolution de 2.4 Å, montrant un corps replié très conservé. D/NP se localise dans le noyau de cellules infectées et/ou transfectées et, contrairement à A/NP et B/NP, D/NP possède une extrémité flexible intrinsèquement désordonnée en C-terminale, avec un évident NLS bipartite. L'interaction entre l'importine- $\alpha 7$  et D/NP<sub>TAIL</sub> se caractérise par une affinité environ dix fois plus forte que celles déterminées pour A/NP<sub>TAIL</sub> et B/NP<sub>TAIL</sub>, de 100 nM, une affinité caractéristique des NLS fonctionnels. La délétion du motif basique

C-terminal du NLS conduit à une interaction défectueuse avec l'importine- $\alpha 7$  et une localisation nucléaire affectée, tandis que la délétion des deux motifs conduit à une abolition complète des deux mécanismes. Ces résultats ont fait l'objet d'une publication rattachée à ce travail de thèse.

Les fonctions clés de la nucléoprotéine grippale sont retrouvées avec D/NP : sa structure, son état d'oligomérisation, son NLS sur une extrémité désordonnée, sa localisation cellulaire. Cependant, chacun de ces mécanismes montrent aussi de claires spécificités, propre à D/NP.

Un large nombre d'études se sont intéressées à la caractérisation de l'interaction NP:importine- $\alpha$ , apportant de précieux renseignements au point de vue fonctionnel, moléculaire et atomique. Cependant, ces connaissances restaient lacunaires, autant du côté des importines- $\alpha$  que des NPs. Les seuls complexes étudiés dans le détail et impliquant A/NP et B/NP, A/NP<sub>TAIL</sub>:importine- $\alpha 1$  et B/NP<sub>TAIL</sub>:importine- $\alpha 7$ , montraient une faible affinité, probablement au-delà de la limite de fonctionnalité d'un NLS. Au contraire, C/NP<sub>TAIL</sub>:importine- $\alpha 1$  et D/NP<sub>TAIL</sub>:importine- $\alpha 7$  présentent des affinités consistantes avec une activité fonctionnelle *in cellulo*, respectivement 50 et 100 nM. Cependant, alors que toutes les NPs grippales ont commencé à être investiguées, il y avait toujours un manque criant d'intégration des différentes importines- $\alpha$ , alors même qu'il est bien connu désormais qu'il s'agit d'entités uniques, spécifiques, avec leurs propres régulations et partenaires.

Grâce à la chromatographie d'exclusion de taille couplée (ou non) à la diffusion de la lumière, à la résonance des plasmons de surfaces et à l'anisotropie de fluorescence, la première stratégie comparative complète, impliquant toutes les NP<sub>TAILS</sub> déjà décrites et toutes les importines- $\alpha$ , a pu être mise au point et a révélé des spécificités surprenantes. Les NP<sub>TAILS</sub> ont été fusionnées à la YFP, de manière à conserver un contexte globulaire sans contribution de l'oligomérisation propre à la NP. Toujours capables d'interagir avec les importines, la caractérisation cinétique a pu être réalisée. Les vitesses d'interaction impliquant d'une part, les importines- $\alpha 3$ , - $\alpha 5$  et - $\alpha 7$ , et d'autre part, A/NP<sub>TAIL</sub>, C/NP<sub>TAIL</sub> and D/NP<sub>TAIL</sub> se sont révélées extrêmement rapides, tant l'association que la dissociation. D/NP<sub>TAIL</sub> est capable d'interagir avec une forte affinité avec toutes ces importines- $\alpha$ , sans spécificité identifiable. A/NP<sub>TAIL</sub> et C/NP<sub>TAIL</sub> présentent des affinités plus variables, chacune interagissant mieux avec une

importine- $\alpha$  particulière, respectivement l'importine- $\alpha 7$  et l'importine- $\alpha 3$ . B/NP<sub>TAIL</sub> s'est montrée extrêmement surprenante: les vitesses d'association et de dissociation se sont révélées encore plus rapides, et les affinités pour toutes les importines- $\alpha$  testées se sont avérées très faibles.

L'importine- $\alpha 1$ , présentant une propension à dimériser en solution, n'a pas pu être testée par la stratégie mise en place. A l'aide d'un test de compétition et d'anisotropie de fluorescence, des valeurs d'affinités ont pu être obtenues et ont révélées une forte affinité pour D/NP<sub>TAIL</sub> et C/NP<sub>TAIL</sub>, une affinité basse pour A/NP<sub>TAIL</sub> et une affinité médiocre pour B/NP<sub>TAIL</sub>.

Ainsi, alors que des différences étaient attendues dans les caractéristiques de ces complexes, elles se sont avérées encore plus importantes, avec une contribution égale de chacun des partenaires à l'interaction. Ces résultats ont fait l'objet d'une seconde publication au cours de travail de thèse. Ils établissent une base solide pour continuer à disséquer et détailler cette interaction hôte:pathogène extrêmement complexe et essentielle.

En 2018, une étude métagénomique a identifié plus de 200 nouveaux virus à génome à ARN, et notamment 3 nouveaux virus appartenant à la famille des *Orthomyxoviridae*. Un de ces nouveaux virus infecte un crapaud et, de façon assez surprenante, est plus proche des virus de type B que des autres virus grippaux. Sa NP (Toad/NP), de la même façon, est phylogénétiquement plus proche de B/NP que des autres NPs grippales. Une prédiction structurale a mis en évidence une architecture potentiellement similaire à celle de B/NP, avec une longue extrémité N-terminale désordonnée (Toad/NP<sub>TAIL</sub>), un corps central replié de taille similaire (Toad/NP<sub>CORE</sub>) au autres NPs et une courte extrémité C-terminale désordonnée. Cette prédiction a aussi identifié des traits très particuliers de Toad/NP<sub>TAIL</sub>: elle pouvait potentiellement posséder un domaine replié. De plus, longue d'environ 130 acides aminés, cette extrémité N-terminale l'est près de deux fois plus que celle de B/NP et présente de nombreux motifs basiques. Il a donc été décidé de commencer à caractériser l'interaction de Toad/NP avec les importines- $\alpha$  via la stratégie SPR mise en place précédemment, ainsi que son interaction avec l'ARN et sa structure. La cristallographie, la microscopie électronique, la chromatographie d'exclusion de taille,

l'anisotropie de fluorescence et la résonance des plasmons de surface ont été utilisés à ces fins.

Toad/NP a tendance à former des tétramères en solution, et son état d'oligomérisation n'est pas impacté non plus par la concentration en NaCl, de façon similaire à D/NP. Toad/NP<sub>TAIL</sub> interagit avec les importines- $\alpha$  d'une façon très similaire à B/NP<sub>TAIL</sub>, avec des vitesses d'association et de dissociation très rapides, hors des limites de détection de l'appareil. Son affinité pour les importines- $\alpha 5$  et - $\alpha 7$  est très basse, hors des limites supposées d'un NLS fonctionnel, mais son affinité pour l'importine- $\alpha 3$  indique par contre qu'il pourrait s'agir d'une interaction fonctionnelle pour l'import nucléaire. Toad/NP<sub>CORE</sub> interagit avec l'ARN de façon similaire à A/NP et B/NP, avec coopérativité, sans spécificité de séquence et avec une potentielle saturation du site de fixation de l'ARN avec un ARN 12-mer. Toad/NP possède en revanche une spécificité intéressante : son extrémité N-terminale semble réguler négativement l'interaction avec l'ARN, peut-être *via* une région riche en glutamate. La structure de Toad/NP<sub>CORE</sub> n'a pas encore été résolue à ce jour mais le travail est encore en cours. Il est hautement probable que ce corps, comme le corps de toutes les autres NPs, soit structuralement très conservé. En revanche, des cristaux du domaine N-terminal ont permis de résoudre la structure d'un domaine replié en 3 longues hélices- $\alpha$  antiparallèles (acides aminés 4 à 68). Il s'agit du premier domaine replié identifié dans l'une des extrémités désordonnées des NPs des virus grippaux. Le rôle et la fonction de ces deux domaines (le domaine replié et le domaine désordonné) ne sont pas encore clairement identifiés et nécessite davantage de travaux.

La découverte et la caractérisation de Toad/NP a mis en évidence une forte conservation du fonctionnement de la NP malgré son tropisme vers un animal à sang froid et a également révélé des traits uniques de cette NP, comme ce domaine N-terminal très long, et partiellement replié.

En conclusion, ce travail de thèse a permis de disséquer les interactions entre les différents types de NP grippales et les importines- $\alpha$  de manière comparative, en évaluant la contribution de deux partenaires. Il s'agit d'une base importante et essentielle pour finalement comprendre cette interaction complexe et très régulée. Il a aussi permis de caractériser les NPs de deux nouveaux membres de la famille des virus de la grippe et mis en évidence à la fois la forte base conservée structurelle et

fonctionnelle des NPs grippales et des caractéristiques très spécifiques, propres à chaque type grippal et adaptés à son hôte.



# Chapter 1: Introduction





## I. Influenza viruses

The results detailed in this manuscript concern all influenza virus types. However, because influenza A is a major concern for the human health, the majority of the publications is based on influenza A viruses.

### A. Features of influenza viruses

#### i. *Taxonomy*

Influenza viruses, as viruses with a segmented negative-sense single-stranded RNA genome, belong to the *Riboviria* realm, the *Negarnaviricota* phylum, the *Articulavirales* order and are part of the *Orthomyxoviridae* family (**Table 1**). According to the last release of the International Committee on Taxonomy of Viruses in July 2019<sup>134</sup>, the *Orthomyxoviridae* family contains seven genera, all composed by one or two species. Amongst these seven genera, 55 *Orthomyxoviridae* viruses are currently unclassified.

---

**Table 1: Full taxonomy of the Orthomyxoviridae family members.**

---

<u>Taxonomy</u>	<u>Features</u>
<i>Riboviria</i> realm	RNA viruses
<i>Orthornavirae</i> kingdom	
<i>Negarnaviricota</i> phylum	Negative single-stranded RNA genome
<i>Polyploviricotina</i> sub-phylum	Cap-snatching RNA viruses
<i>Insthoviricetes</i> class	Contraction of <b>influenza</b> , <b>isavirus</b> and <b>thogotovirus</b>
<i>Articulavirales</i> order	Segmented genome
<i>Orthomyxoviridae</i> family	
<i>Alphainfluenzavirus</i> genus	
<i>Betainfluenzavirus</i> genus	
<i>Gammainfluenzavirus</i> genus	
<i>Deltainfluenzavirus</i> genus	
<i>Isavirus</i> genus	
<i>Quaranjavirus</i> genus	
<i>Thogotovirus</i> genus	

---

This classification is bound to evolve, as researchers have only scratched the surface of the viral diversity. The science enters the area of metagenomics and new studies are currently focusing on virus sampling in organisms that have been

overlooked before <sup>151,188</sup>. As an example, in 2018 a study identified 214 new RNA viruses found in reptiles, amphibians and fishes <sup>329</sup>. In this study, three new *Orthomyxoviridae* family members were identified: the Wuhan Asiatic toad influenza virus, the Wuhan spiny eel influenza virus and the Wenling hagfish influenza virus. None of these viruses has been included in the ICTV taxonomy yet. Unexpectedly, these viruses present sequence identities closer to the narrow-host range influenza B viruses rather than to the wide-host range influenza A viruses. It has been liberally chosen to use Toad influenza virus (ToadV) to refer to the Wuhan Asiatic toad influenza virus studied in this manuscript, due to a clear lack of insights on the virus cycle and specificities.

In the *Orthomyxoviridae* family, only influenza A viruses (IAV) are further classified into subtypes, based on the antigenicity of their two surface glycoproteins, the hemagglutinin (HA) and the neuraminidase (NA). Up to now, 18 different hemagglutinins (H1 - H18) and 11 neuraminidases (N1 - N11) subtypes have been identified. Influenza B viruses (IBV) present only one HA and NA subtype while influenza C and D viruses (ICV and IDV) possess only one surface glycoprotein, the hemagglutinin-esterase-fusion (HEF), combining both HA and NA activities <sup>117,120,335</sup>.

The current nomenclature <sup>393</sup> for influenza virus includes the influenza type (influenza A, B, C or D), the host species from which it has been isolated (except for human strains), the geographical area, the strain number, the year of isolation and, for influenza A viruses, the HA and NA subtypes (HxNx).

-A/WSN/1933(H1N1),

-B/Memphis/13/03,

-C/Ann-Arbor/1/50/ and

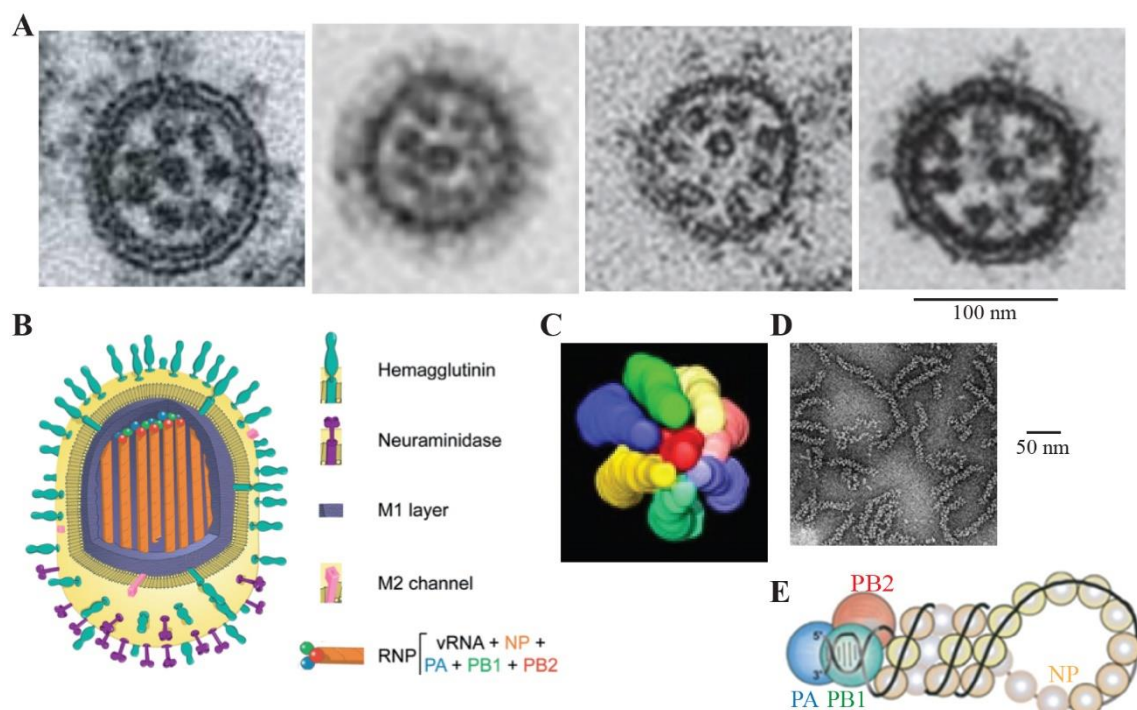
-D/bovine/France/2986/2012 influenza strains

were used for this PhD thesis work along with the newly discovered Wuhan Asiatic toad influenza virus (ToadV). Proteins will be labelled as X/Y, with X the influenza type (A, B, C, D or Toad) and Y the name of the protein.

## ii. *Viral particle structure*

Influenza viruses (**Fig. 1A and 1B**) are enveloped viruses with a lipid envelop deriving from the plasma membrane of the host cell, and so enriched in host lipids such

as cholesterol<sup>317</sup>. During the virion release, several cytoplasmic and membrane host proteins are integrated into the envelope, such as lipid raft-associated proteins and cytoskeleton proteins<sup>327</sup>. By electron microscopy (EM), envelop-inserted HA and NA can be distinguished by their specific features, respectively a smaller globular head and stalk and a high flat head (Fig. 1A and 1B)<sup>31,229,308,402</sup>. Approximately 350 trimeric HA and 50 tetrameric NA are inserted into a single viral particle envelop; this HA/NA ratio allows the functional receptor binding of HA<sup>115</sup>. Balance in HA and NA activities seems required for the fitness of the viral particles<sup>376</sup>. C/HEF and D/HEF are distributed hexagonally at the virion surface, with a distinct mushroom shape<sup>121,234</sup>.



**Figure 1: Viral particle structure and genome organization.**

(A) Electron microscopy (EM) pictures of IAV, IBV, ICV and IDV particles, from left to right<sup>234,235,250</sup>, (B) schematic representation of influenza A virus particle<sup>62</sup>, (C) modelization from STEM tomography of the packaging of the genome inside virions<sup>234</sup>, (D) EM picture of influenza ribonucleoproteins (RNPs) purified from virions<sup>310</sup>, (E) schematic representation of a RNP, with the viral genomic segment in black, the nucleoprotein in light yellow and the three subunits of the polymerase in blue, green and red<sup>231</sup>.

Influenza virions are pleomorphic<sup>43,115</sup>. All influenza viral particles present similar sizes and structures, with both spherical and filamentous phenotypes. The size of the viral particles is usually around 100-120 nm in diameter (Table 2)<sup>26,117,155,254</sup>. In addition, ICV filamentous particles can present an assembly in cord-like structures, up

to several micrometers in length <sup>247</sup>. IAV, IBV and IDV particles occasionally form such structures, seemingly without packaged RNPs <sup>42,200,234</sup>. The morphology of the viral particle seems to be cell- and/or strain-dependent and is modified by serial passages in culture. Viral particles isolated from clinical samples tend to form more filamentous structures while cell culture-adapted or eggs-adapted virus shed virions far more homogeneous and spherical, as a result of host adaptation <sup>42,159,300</sup>.

**Table 2: Summary of the influenza main features depending on the virus type.**

Sph, Fil and Cord respectively stand for Spherical, Filamentous and Cord-like. \* While reassortment can occur for IBV due to the two circulating lineages, this does not give rise to an antigenically distinct virus, only to a genetically distinct virus.

<u>Virus type</u>	IAV	IBV	ICV	IDV	ToadV
<u>Viral particle phenotypes</u>	Sph: 120 nm Fil: 120 nm Cord: several µm	Sph: 120 nm Fil: 150 nm Cord: several µm	Sph: 100 nm Fil: 100 nm Cord: several µm	Sph: 100 nm Fil: 100 nm Cord: several µm	?
<u>Genomic segments and main proteins</u>	Eight segments S1: PB2 S2: PB1 S3: PA S4: HA S5: NP S6: NA S7: M1/M2 S8: NS1/NEP	Eight segments S1: PB2 S2: PB1 S3: PA S4: HA S5: NP S6: NA S7: M1/M2 S8: NS1/NEP	Seven segments S1: PB2 S2: PB1 S3: P3 S4: HEF S5: NP S6: M1/M2 S7: NS1/NEP	Seven segments S1: PB2 S2: PB1 S3: P3 S4: HEF S5: NP S6: M1/M2 S7: NS1/NEP	Eight segments?
<u>Viable reassortment</u>	Yes, except some strain combinations	No*	No	No	?
<u>Reservoir</u>	Aquatic birds	Humans	Humans	Cattle	Toad?
<u>Tropism</u>	Wide spectrum Mammals and birds	Narrow spectrum Mainly humans	Narrow spectrum Humans and swines	Medium spectrum Mammals	?

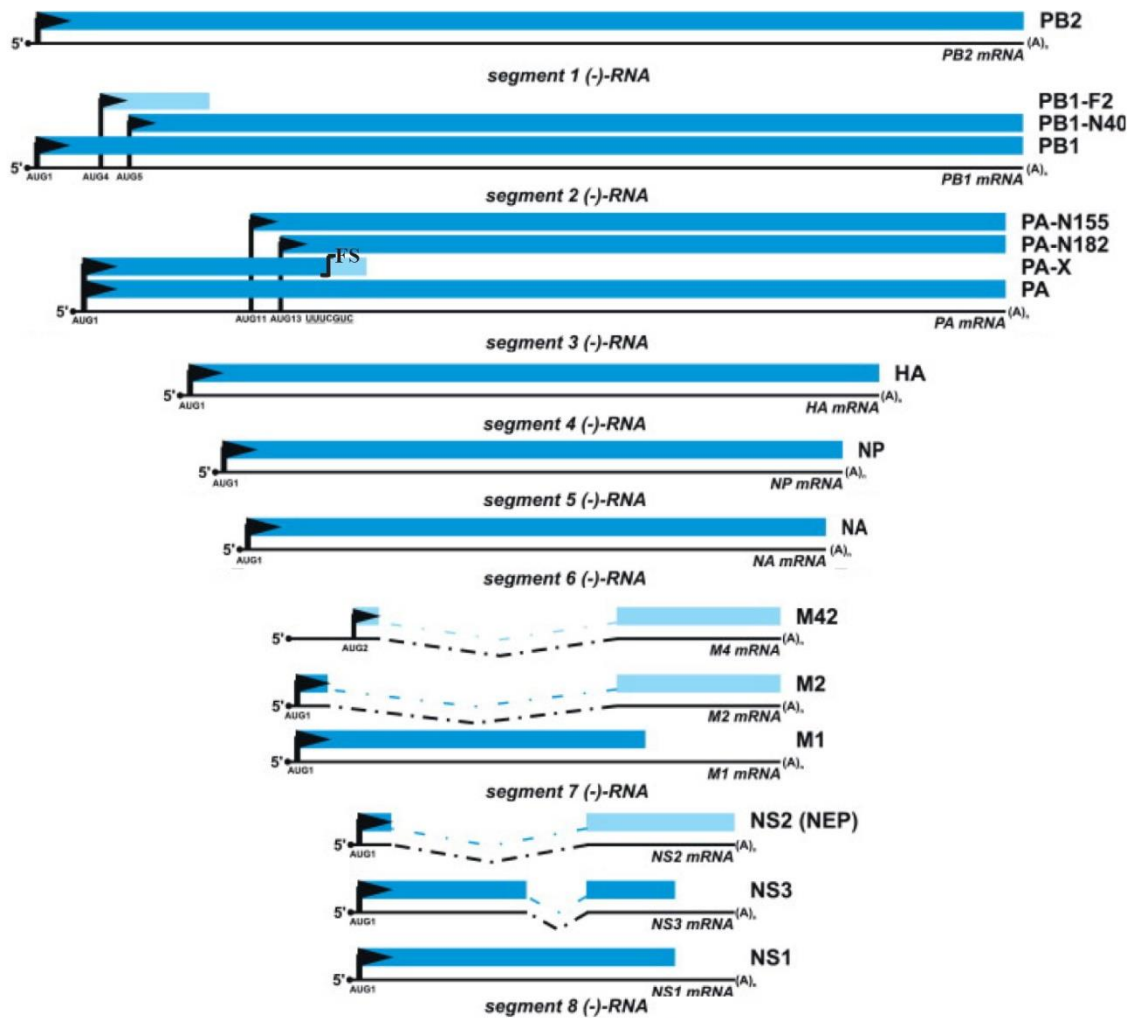
Using electron tomography and EM, each segment of the viral genome, can be visualized as a dense rod-like structure (**Fig. 1A and 1D**). Eight of these structures, called ribonucleoproteins (RNPs), are localized at the distal end of influenza virions (**Fig. 1A, 1B and 1C**) <sup>234,250</sup>.

### *iii. The ribonucleoproteins, the replicative units of the virus*

The functional replicative unit is composed by the viral RNA genome (vRNA) with viral proteins necessary and sufficient for an efficient replication and transcription. In influenza viruses, such a macromolecular complex is called an RNP (**Fig. 1**). One RNP is composed by a single genomic segment, encapsidated by multiple copies of the nucleoprotein (NP) and with one heterotrimeric RdRp (PA:PB1:PB2) associated to both 3' and 5' extremities (**Fig. 1E**)<sup>7,230</sup>. This structure is helicoidal, supercoiled and highly flexible (**Fig. 1D and 1E**)<sup>46,162,277</sup>. The dimensions of an RNP are respectively approximatively for 30-120 nm x 10-15 nm, depending on the encapsidated segment<sup>46,162,277,309</sup>. The oligomerization of RNA-bound NP gives RNPs their rod-like structure<sup>280,353</sup>. Influenza viruses possess 7 or 8 genomic segments and consequently 7 or 8 unique, distinct RNPs. Interestingly, even if ICV and IDV possess only seven viral genomic segments, it has been shown that all influenza viruses present a 7 + 1 packaging pattern of RNPs in virions (**Fig. 1C**)<sup>115,234</sup>. Whether ICV and IDV package an extra vRNA segment or one of the host RNA is still unclear<sup>249</sup>.

### *iv. Genome organization, protein content and host tropism*

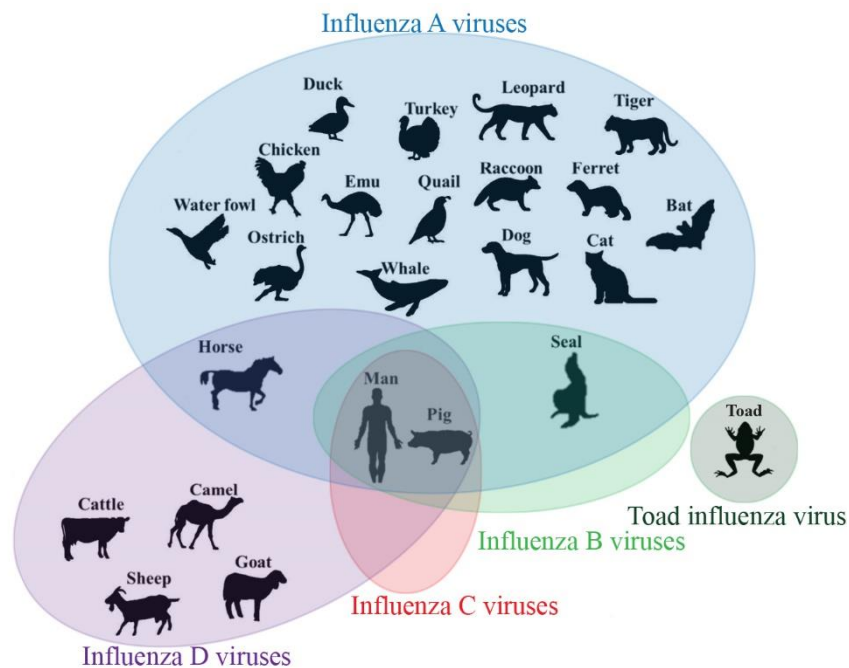
IAV and IBV possess eight viral RNA genomic segments whereas ICV and IDV genomes are composed by seven segments<sup>117,265,298</sup>. Each segment encodes for one or several proteins. Some segments are known to contain distinct open reading frames, encoding distinct proteins. Transcribed mRNAs can also be processed by the cellular splicing machinery, to generate other proteins (**Fig. 2**)<sup>67,90,233,408</sup>. The segments are numbered and named depending on their length and the main protein they encode (**Table 2, Fig. 2**). The difference in segment number is due to the difference in the viral glycoprotein composition: IAV and IBV possess HA and NA while ICV and IDV only possess HEF. While the full genome is not yet available for ToadV, two glycoproteins (HA and NA) are encoded by the genome, meaning that it seems to be made of eight segments, like both IAV and IBV. The details of the proteins is shown with the **Figure 2**.



**Figure 2: Main and accessory proteins encoded by influenza A viral mRNAs.**

Adapted from <sup>370</sup>. Black arrow and dotted lines represent respectively the mRNA translation start and spliced parts. Each blue shade corresponds to one of the 3 open reading frames. For segment 3 PA-X mRNA generation, one frameshift occurs (FS).

Beside the main proteins, several accessory type-dependent proteins have been identified. Up to now, seven non-essential proteins have been described for IAV, which seem to play many roles in the regulation, the virulence and/or the host adaptation processes (**Fig. 2**) <sup>370,409</sup>.



**Figure 3: Host tropism of influenza viruses.**

The known host range of each influenza virus is labelled as a coloured background, with blue for IAV, light green for IBV, red for ICV, purple for IDV and dark green for ToadV. Adapted from <sup>171</sup>.

Each influenza virus possesses its own specific host tropism, which relies partly on the behaviour and specificities of its proteins (**Fig. 3, Table 2**). Aquatic birds are the reservoir for IAV but they are able to infect a wide spectrum of animals. Some serological reports show evidence of infection in reptiles and amphibians or at least some sort of cross-reactivity, in addition to mammals, bats and birds <sup>36,369,387,415</sup>. IBV and ICV present a restricted host spectrum: IBV mainly infect humans, but some cases have been reported in seals, and ICV can infect swine in addition to humans <sup>255,260,417</sup>. IDV have been discovered in 2011, initially identified as an ICV subtype, but officially considered as a new influenza genus in 2013. IDV show a large spectrum of host organisms in mammals, from camelids to pigs and goats, but their reservoir is suspected to be the cattle <sup>68,116,117,285</sup>. They can replicate in ferrets, the animal model for human infection <sup>117</sup>. Humans, and more particularly people working closely to the cattle, have been shown to seroconvert by producing anti-influenza D antibodies <sup>117,363,391</sup>. Because of its recent discovery, ToadV has not been studied so far, without any further information in terms of reservoir and tropism.

Several factors restrict the host and the cell tropisms. The growth temperature is one of them. IDV have been found to efficiently replicate at both 33°C and 37°C, likely



helping in its broad host range <sup>117,124,391</sup>. IAV and IBV present an optimal growth temperature of 37-39°C and 35°C respectively, in agreement with the temperature of their respective reservoir hosts and cell environments <sup>200</sup>. ICV tend to grow better at lower temperatures (*i.e.* 33°C rather than 37°C) <sup>117,254</sup>.

In term of cell tropism, the main and most studied factor is the HA/HEF glycoprotein, which binds receptors modified with sialic acids (**Fig. 4A**). The A/HA activation seems to be done through a restricted number of proteases and its optimal temperature ranges from 33 to 39°C depending on the HA type and the targeted host. In contrast, B/HA demonstrates a large panel of activating proteases with an optimal temperature of 33°C, well adapted to human airways <sup>182</sup>. D/HEF and C/HEF are structurally well conserved and are both able to bind to a large spectrum of host tissues, but with a better D/HEF attachment to the cell surface than C/HEF. D/HEF presents a more open binding cavity, likely allowing more diverse receptors and/or partners interactions <sup>335</sup>. C/HEF is rather sensitive to the temperature, with an optimal close to 33°C <sup>348</sup>. Interestingly, IDV is the most resistant of all influenza viruses to an acidic or hot environment and its stability was shown to be mainly due to D/HEF <sup>416</sup>.

Several other viral proteins are involved at host:pathogen interfaces and are responsible for these differences in host and cell tropisms such as NS1, the polymerase subunits and NP.

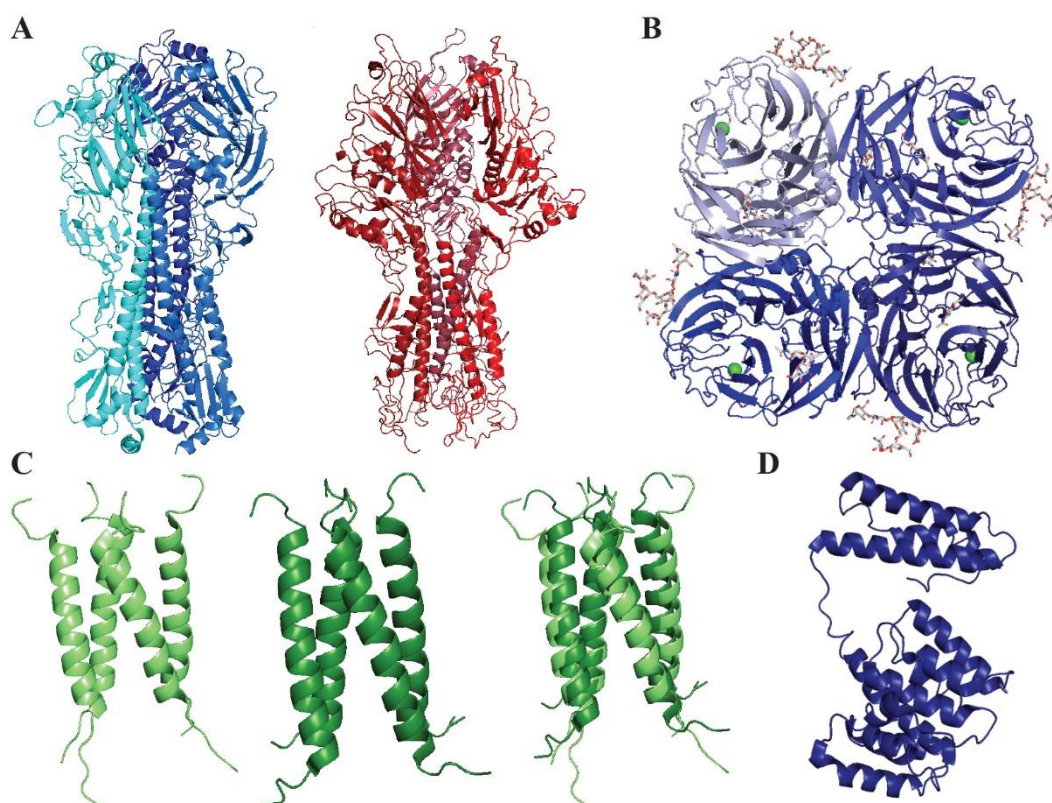
The subunits of the RdRp, and in particular PB2, are considered as determinants of the host tropism <sup>23,351</sup>. It is now well-known that the A/PB2 E627K substitution is an important marker of the mammal adaptation, for the replication at lower temperatures <sup>343,352</sup>. When residue 627 is in its avian form, the D701N substitution can compensate, allowing a successful replication, probably in its interaction with the cellular nuclear import machinery <sup>24,341</sup>. Other discreet substitutions can also compensate for a fully avian phenotype at these two positions, as seen in the case of the pandemic 2009 H1N1 strain <sup>214</sup>.

NP can be found at several host:pathogen interfaces. Several mutations (N319K, G102R, M105K or D375N) have been identified as positions involved in nuclear import <sup>93,94,291,292</sup>. Other mutations (G16D, L283P, F313Y or Q357K) allow the viral escape from the MxA restriction <sup>202</sup>.

Outside of a few specific mutations, the diversity of adaptive mutations suggests that adaptation to a new host is a largely synergic mechanism, with several distinct mutations likely required to overcome this barrier <sup>203,319</sup>.

NS1 is strongly involved in the host defences regulation, and more particularly against interferon signalling <sup>110</sup>. In H5N1 viruses, the residue 92 tends to be a glutamate or an aspartate, and could lead to a stronger resistance to antiviral responses and to higher virulence <sup>271,324</sup>. A/NS1, B/NS1, C/NS1 and D/NS1, while sharing the same antiviral activities, are not completely interchangeable for the replication and the host response inhibition purposes <sup>251</sup>.

While not directly involved in the host tropism, M1 and M2 seem to possess a very specific behaviour depending on their viral type.



**Figure 4: Structure of influenza HA, HEF, NA, M2 and M1 proteins.**

(A) Crystal structures of trimeric A/HA (blue, PDB 3LZG) and C/HEF (red, PDB 1FLC). (B) Crystal structure of A/NA (PDB 5HUK). (C) Crystal structures of the open channel of B/M2 (left), the close conformation (middle) and the superimposition (right) (PDB 6PVR and 6PVT). (D) Cryo-EM structure of the A/M1 matrix protein (PDB 7JM3).

The influenza M2 protein is an ion channel essential during the viral cycle and more particularly for the cytoplasmic release. A/M2 and B/M2 (**Fig. 4C**) share some features, such as a mostly proton channel role, histidines playing the part of proton sensor and activator and a similar transmembrane domain length. They also seem to present specificities, with a different pH of activation and a different direction of the proton flux<sup>179,198,420</sup>. In contrast, C/M2 and D/M2 are primarily chloride ion channel with a slight proton permeability, suggesting a different mechanism in the viral cycle<sup>21,125,158,227</sup>.

The highly flexible influenza M1 protein (**Fig. 4D**) is heavily involved in the structure of the viral particle and the budding process<sup>303</sup>. M1 is often found at the interface between viral and host partners, and it is known to interact with viral RNPs, several viral and cellular factors, and lipids.

#### *v. Evolution: antigenic drift and antigenic shift*

Influenza viruses are known for their fast continuous evolution, depending on two mechanisms: the antigenic drift and/or the antigenic shift.

Like the majority of RNA virus RdRp, the influenza RdRp does not possess a proofreading mechanism. Therefore, the influenza RdRp integrates about one error every genome replication, at a rate of about  $10^{-5}$ - $10^{-6}$  substitution/nucleotide site/cycle<sup>23,66,269,340</sup>. It has been shown previously that IAV present a higher mutation rate than IBV, with rates of about  $2 \cdot 10^{-3}$  and  $0.6 \cdot 10^{-3}$  substitutions/nucleotide site/year respectively<sup>248,390</sup>. Influenza viruses are under a strong selective pressure from both the host immune system and the host machineries they hijack. NA and HA (**Fig. 4A and 4B**) are in the front line, recognized swiftly by the immune system and the main target for host antibodies<sup>141,276</sup>. This requires a fine balance between escape mutations and/or mutations increasing the virus fitness and non-deleterious mutations to preserve the protein functions. This process is called the antigenic drift and occurs also in host adaptation in cell culture<sup>334</sup>.

The antigenic shift is a brutal event that is due to the direct jump from one host species to another and/or the reassortment. The latter is specific to the segmented nature of the viral genome. It occurs when two viruses co-infect the same cell and exchange one or several genomic segments during the packaging step. However, it is

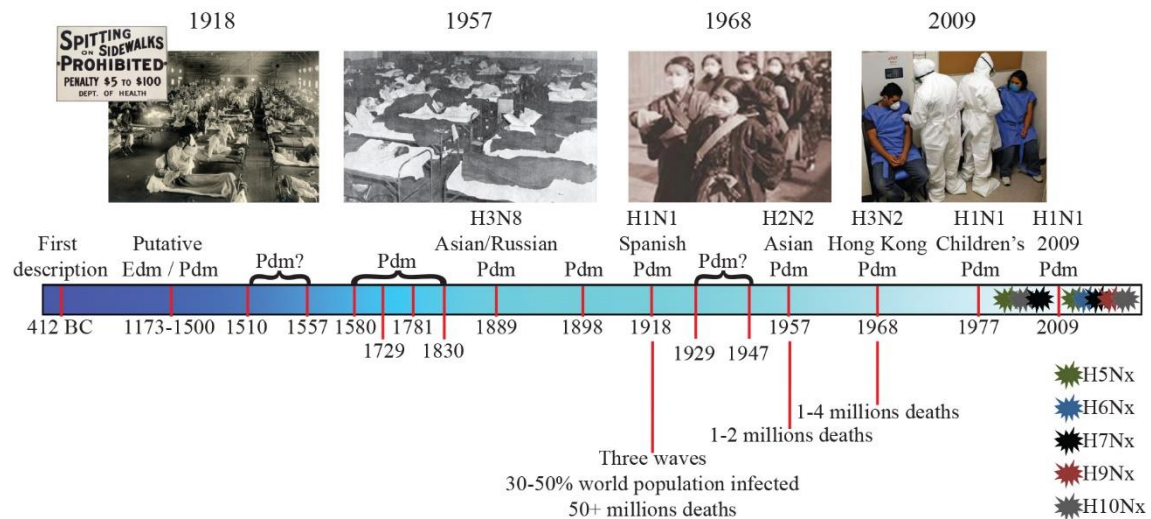
required for the two viruses to be of the same type and compatible<sup>187</sup>. Genomic segment exchange often leads to pandemic strains of IAV<sup>319,384</sup>. As an example, the pandemic 2009 H1N1 genome came from four distinct IAV parental strains<sup>225,253</sup>.

#### *vi. Disease and epidemiology*

Influenza viruses are mostly known due to the seasonal epidemic and sporadic pandemic they cause in humans. Seasonal influenza is an acute and contagious disease affecting the human respiratory tract with several non-specific symptoms, such as high fever, cough, asthenia, headaches and muscular aches. In most cases, no complication arises. However, severe or deadly outcomes can befall to high-risk people, including pregnant women, immunocompromised or people affected by a chronic disease, the elderly and new-borns. Upon influenza virus infection, they could develop severe respiratory tract infection, associated or not with coinfection, and extreme inflammatory reaction. The severity is intrinsically linked to the virus and the host immune response<sup>149</sup>. IAV and IBV are responsible for these yearly events, infecting around 10% of the world population, resulting in 3 to 5 million severe cases and 290 000 to 650 000 respiratory-related deaths each year<sup>138</sup>. The average annual economic burden of a seasonal outbreak has been recently estimated at \$11.2 billion in the USA or €78.3 million in Germany<sup>284,318</sup>.

Seasonal epidemics occur during fall and winter in temperate climates<sup>400</sup>. In the tropical and equatorial parts of the globe, influenza viruses circulate all around the year, with unpredictable epidemics, assuring virus constant exchange between the Northern and the Southern hemispheres<sup>312,374</sup>. The virus is mainly transmitted through infectious droplets emitted by infected patients, but it can also be spread through the contact with an infected surface or person<sup>213,252</sup>.

IAV and only IAV can cause pandemics, corresponding to worldwide sporadic occurrences: only IAV possess several antigenically distinct HA and NA proteins, leading to large-scale outbreaks for a lack of previous immunity.



**Figure 5: Major historical outbreaks in the shared human and influenza history.**

Putative and confirmed pandemics (Pdm) and epidemics (Edm) are illustrated as red bars. Sporadic outbreaks of deadly highly pathogenic avian influenza viruses are shown as spiky forms of distinct colours for distinct HA types. This schematic representation is not at scale.

The human history is tightly intertwined with influenza viruses (**Fig. 5**): in 412 B.C. in Ancient Greece, Hippocrates described a disease strikingly similar to influenza and no less than 13 pandemics have been identified these last 400 years<sup>174,279,315</sup>. The worst pandemic in modern history was caused by a H1N1 influenza virus, submerging the world in three waves in 1918-1920. As World War I was raging across the world, involved belligerents did not report the pandemic. Spanish censorship-free newspapers did, and thus the pandemic was wrongly named “the Spanish flu”. The mortality rate was estimated at least 2.5-5%, far higher than any other influenza outbreak. The virus infected almost half of the world population, killing at least 50 million people, mainly young adults<sup>145,330,400</sup>. Beyond the new antigenic variant and the virus intrinsic virulence, other reasons can explain this catastrophe<sup>135</sup>. In 1918, the responsible agent was unknown and therapeutic, prophylactic and sanitary measures were lacking<sup>144,208,336</sup>. Antibiotics to treat severe secondary bacterial infections were not discovered yet, becoming widely available from 1939 onwards.

Since the 1920’s, four pandemics occurred, with a high death toll but far from the level of 1918-1919 (**Fig. 5**). Nowadays, experts are closely monitoring several influenza viruses, and more particularly zoonotic viruses of the H5Nx, H6Nx, H7Nx, H9Nx and H10Nx subtypes<sup>222,400</sup>. Sporadic transmission of these viruses to humans occurs regularly, without sustainable human-to-human transmission. However, the

mortality rate in these infections can be up to 50%, suggesting that a pandemic could result in a disaster <sup>147,323,394</sup>.

In addition to human infections, influenza viruses widely circulate in animals, more particularly in birds for IAV <sup>169,262</sup>. While wild aquatic birds tend to be asymptomatic or present little symptoms, this is not the case in poultry <sup>226</sup>. Low-pathogenic avian influenza viruses (LPAIs) and highly pathogenic avian influenza viruses (HPAIs) tend to spread quickly in bird populations. LPAIs tend to induce in worse cases mild symptoms, such as a decrease in egg production and slight respiratory issues; HPAIs, in contrast, usually result in a severe disease, killing domesticated birds in a few days, with a mortality rate that can be close to 100% <sup>262,347,392,401</sup>. The severity of the IAV infection depends on the strain and the host and ranges from little to no symptom to systemic infections and high mortality rates <sup>338</sup>.

Even if its epidemic and circulation patterns remain largely unknown, ICV infect humans and more particularly children, leading to a mild respiratory disease <sup>321</sup>. Swine are also susceptible but with no evident signs of the disease.

IDV reservoir seems to be the cattle. The clinical symptoms for IDV alone are typically weak but IDV coinfection with other pathogens seems to be involved in a more severe disease, called the bovine respiratory disease (BRD) <sup>80,216</sup>. Swine and various farm animals such as goat but also camel present an anti-IDV serology but the pathology has not been extensively studied. Swine and animal models like ferret do not present influenza-like symptoms and lesions <sup>79,117</sup>.

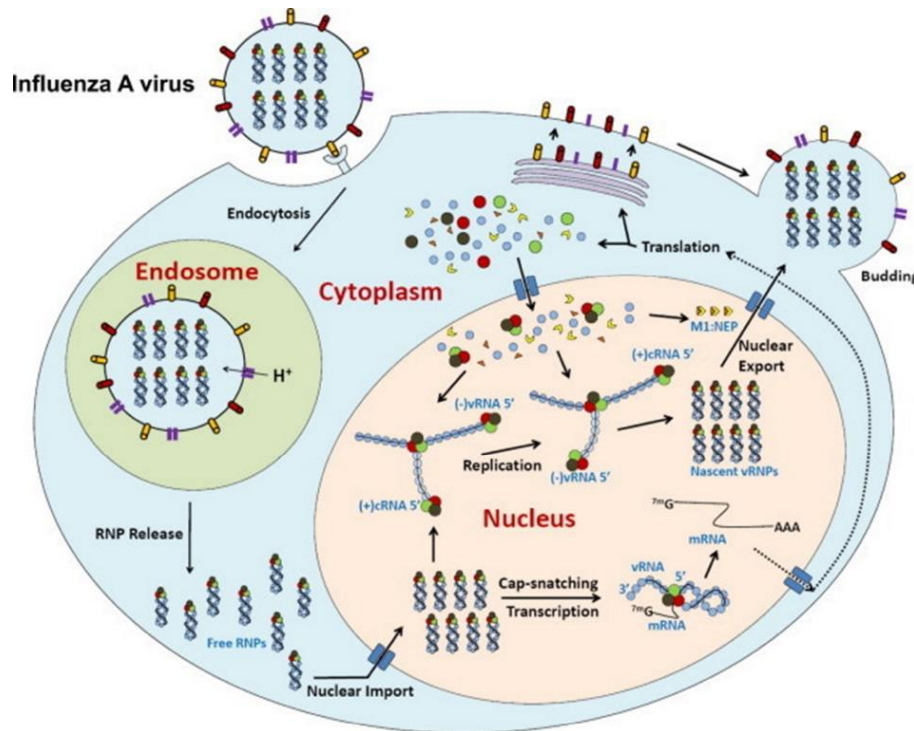
Regarding the newly identified influenza viruses infecting cold-blooded animals, it is not known if they can induce symptoms, infect other animals or cause any diseases.

## B. Influenza virus cycle

The first step of the viral cycle is the attachment to the host cell through HA/HEF (**Fig. 6 and 4A**). Both recognize sialic acid molecules, highly expressed and displayed on cell surfaces.

A/HA and B/HA recognize the N-Acetylneuraminic acid (Neu5Ac) sialic acid, found on glycolipids or glycoproteins <sup>344</sup>. Human A/HA bind preferentially to  $\alpha$ 2,6-linked

Neu5Ac, abundant on human epithelial cells in the upper respiratory tract <sup>47,375</sup>. In contrast, avian A/HA interact with  $\alpha$ 2,3-linked Neu5Ac, predominantly found in the digestive tract of aquatic bird and the human lower respiratory tract <sup>137</sup>. Both  $\alpha$ 2,3- and  $\alpha$ 2,6-linked sialic acids are highly expressed in the pig trachea <sup>136,375</sup>. ICV and IDV HEF bind to the N-Acetyl-9-O-acetylneuraminic acid <sup>301</sup>. The high HA/HEF density could compensate the weak binding affinity by a high avidity <sup>316,349</sup>.



**Figure 6: Influenza viral cycle.**

The viral proteins are represented with different forms and colours: NP, PA, PB1 and PB2 are represented as round forms in blue, brown, green and red respectively. M1 and NEP are triangular forms, in orange and yellow respectively. HA, NA and M2 are cylindrical forms, red, yellow and purple respectively. Adapted from <sup>423</sup>.

The viral entry in cells mainly occurs through the clathrin-dependent mechanism but clathrin-independent endocytosis and macropinocytosis have also been reported <sup>210,270,331,377</sup>. Influenza virions then go through the endosomal/lysosomal pathway. From early endosomes to lysosomes, the internal pH of the endosomal particles drops from 6.5 to 4.5. The ion channel M2, inserted into the virion envelop, leads to the internal acidification of the viral particle, resulting in the M1 dissociation from the RNPs (**Fig. 4C**) <sup>204</sup>. HA/HEF goes through a conformational change, triggering the viral

particle envelop fusion with the endosomal membrane and the RNPs release in the cytosol<sup>96,210,424</sup>.

The RNPs then require access to the nucleus to initiate the transcription/replication steps. Influenza viruses hijack the host nuclear import machinery, and in particular the importin- $\alpha/\beta$  pathway (cf. part. II)<sup>28,54</sup>. RNPs are rapidly translocated through the nuclear pore complex, and transcription is initiated by the viral polymerase<sup>205</sup>. Viral messenger RNAs (mRNAs) are directly transcribed from the viral genome. They are exported in the cytoplasm, where they will be translated by the host machinery<sup>54</sup>. Newly synthesized subunits of the RdRp, NP, M1, NS1 and NEP will be imported or will diffuse in the nucleus to contribute to the synthesis of viral RNAs<sup>28,57,93,381</sup>. Replication is the other important process in the nucleus. For this, negative-sense genomic RNAs are firstly used to generate positive-sense RNA (cRNA) which is used as a template for genomic negative-sense RNA synthesis. As the viral RNAs are never naked, the encapsidation by both NP and the RdRp occurs simultaneously to their synthesis. Once reconstituted, RNPs are exported in the cytosol by using the cellular nuclear export machinery, M1 and NEP. NEP interacts with the cellular Chromosome Region Maintenance 1 (Crm1) export protein and M1, which in turn interacts with RNPs<sup>2,37,72,126,240,267,385</sup>. RNPs are then addressed to the lipid rafts at the plasma membrane, where HA/HEF, NA and M2 are already localized to be packaged into virions<sup>132</sup>.

A functional viral particle requires the seven/eight individual genomic segments packaged in a single particle. While the molecular basis behind this step is still unclear, selective packaging of RNPs is the favoured hypothesis<sup>25,97,98,235</sup>. A viral RNA segment seems to present at its extremities packaging signals, which interact directly with other distinct RNP packaging signals<sup>88,89,197,228</sup>. In addition, parts of coding regions are suspected to enhance a correct packaging<sup>99,105</sup>. They bring together distinct segments, resulting in a close parallel alignment of distinct RNPs. The budding of infectious virions is a mechanism poorly understood. However, the budding seems to be initiated at lipid rafts by the presence of the RNP complexes, HA/HEF, M1 and NA, in an ESCRT-independent manner<sup>238,313,317,388</sup>. M2, localized at the border of lipid rafts, is responsible for the membrane scission<sup>204,302</sup>. NA then cleaves sialic acids from receptors associated with HA, making free the new virion.



C. Viral transcription and replication machineryi. *Nucleoprotein*

The nucleoprotein is the most abundant protein of the RNPs and also in the infected cells. NP is required for both viral transcription and replication activities and could act as a switch between the two processes <sup>15,325</sup>. NP is also involved in the viral RNA genome protection, in the interaction with cellular partners and in the cellular trafficking.

**Table 3: Sequence identities between influenza types.**

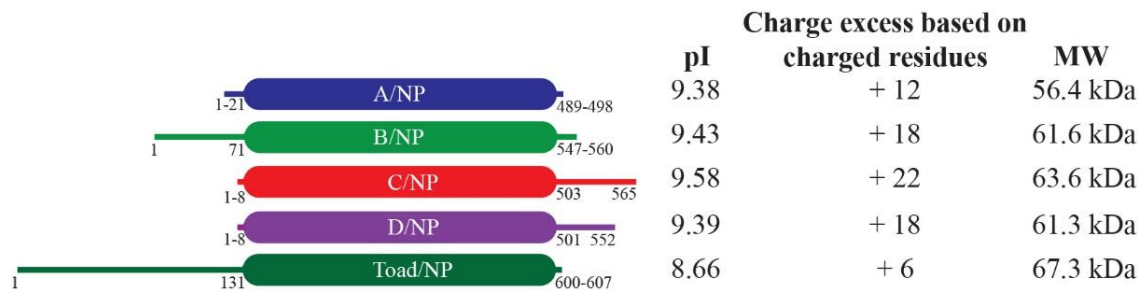
Sequences were recovered from the Influenza Research Database <sup>419</sup>. The strains used for alignment are those used in this work.

	A/NP	B/NP	C/NP	D/NP	Toad/NP
A/NP	100	37.2	21.8	21.1	22.5
B/NP		100	23.1	23.8	28.7
C/NP			100	37.5	22.1
D/NP				100	20.1
Toad/NP					100

Globally, sequence identities between all NP types range from 20 % to 38 %, with biochemical properties and structure predictions rather well conserved (**Fig. 7, Table 3**). A/NP, B/NP, C/NP, D/NP and Toad/NP molecular mass and protein chain length vary between 56 and 67 kDa, ranging from 498 to 607 amino acids.

NP is a highly basic protein with a basic groove, a feature to be linked to its RNA binding activity (**Fig. 8B**) <sup>242,243</sup>. NP does not seem to bind to RNA bases but rather to the phosphate skeleton <sup>13</sup>. It is still unclear how NP and RNA interact, as recent studies showed that NP was not distributed homogeneously on viral genomic segments with a preferential binding mode <sup>183,184,396</sup>. The monomeric form of NP is probably the active form in the interaction with RNAs, with the NP oligomerization triggering the formation of the RNP architecture; the high affinity binding of the trimeric A/NP to RNA could impair the oligomerization process required for the RNP formation <sup>178</sup>. A/NP and B/NP bind RNAs with nanomolar affinity, with cooperativity, but behave differently depending on the salt concentration and the RNA length <sup>13,70,71,178,407</sup>. The hypothetical length of vRNA bound by one monomer of NP in the context of the RNP is thought to be around

24 nucleotides<sup>46,207,259</sup>. Interestingly, NP-bound RNAs are still susceptible to RNases, consistently with the suspected binding pattern of RNAs, wrapped around NP<sup>143,207</sup>.



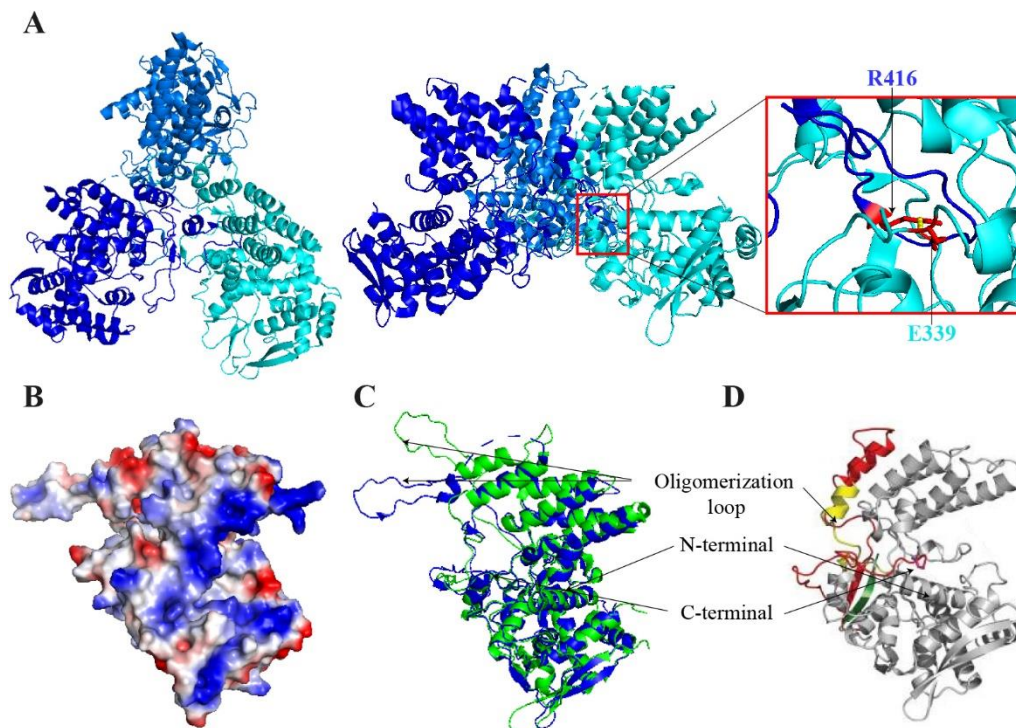
**Figure 7: Biochemical features of influenza nucleoproteins.**

Intrinsically disordered tails and structured domains are represented as thin lines and boxes respectively.

Influenza NPs are globular, tightly folded proteins. To date, several influenza NP structures have been solved, but none of them with RNA (**Fig. 8**). More particularly, two A/NP and one B/NP have been crystallized and solved as trimers and tetramers respectively, highlighting the NP tendency to form oligomers in solution and in absence of RNA. NP oligomerization is affected by the RNA, the salt concentration, its concentration, its phosphorylation state and the temperature<sup>39,178,353,365</sup>.

While A/NP and B/NP share only 37% of sequence identity, they are structurally well conserved (**Fig. 8C**). They are folded mainly in  $\alpha$ -helices. The two major differences come from **i**) the oligomerization loop orientation in the crystal structure, due to the oligomeric state (trimer *v.s.* tetramer) and **ii**) a longer N-terminal (N-ter) domain not visible in the structure and intrinsically disordered (i.e. 21 residues for A/NP against 71 for B/NP). The structure of the monomeric R416A mutant gives precious insights into the oligomerization mechanism (**Fig. 8D**). In the native structure, the Arg-416 is localized in a protruding loop (from residue 402 to 428) making the interaction interface with a second protomer. The Arg-416 makes an essential salt bridge with the side chain of the Glu-339. The R416A mutant loop, in contrast, was not extended and was folded into its own NP-binding groove (**Fig. 8A, 8C and 8D**). Putatively due to the oligomerization loop conformation, the monomeric mutant is drastically impaired in its interaction with RNA (micromolar affinity)<sup>27,40,353</sup>. The B/NP oligomerization

mechanism seems similar to the A/NP one, forming a conserved salt bridge between residues R472 and E395.

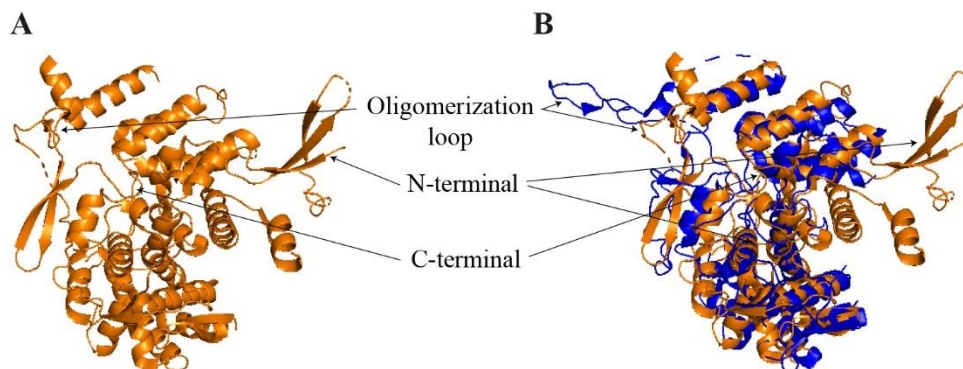


**Figure 8: Influenza nucleoprotein structures and structural features.**

(A) Crystal structure of trimeric influenza A/NP with the essential salt bond responsible for the oligomerization process, (B) Electrostatic potential of a protomer of the trimeric A/NP, with positive charges in blue and negative charges in red (C) Surimposition of A/NP (PDB 2IQH), coloured in blue; with B/NP (PDB 3TJ0), in green (D) Monomeric R416A mutant of A/NP, adapted from <sup>40</sup>.

In the *Orthomyxoviridae* family, another virus, the infectious salmon anemia virus (Isavirus) and its NP (ISA/NP) have been studied. ISA/NP has been crystallized as a dimer. It is structurally well conserved with influenza NPs, but it possesses a supplementary partly structured N-ter domain (**Fig. 9**) <sup>422</sup>. ISA/NP seems to be mainly dimeric in solution. Each monomer binds to 12 nucleotides long RNA molecules and 24 nucleotides long RNAs seem wrapped around ISA/NP, as it has also been suggested for influenza NPs. The activity or the functionality of the unique N-terminal domain of ISA/NP has not been elucidated yet. It is interesting to note that ISA/NP N-ter extremity (NP<sub>TAIL</sub>) is longer than those of A/NP and B/NP (respectively 111, 21 and 71 amino acids). Furthermore, A/NP<sub>TAIL</sub> and B/NP<sub>TAIL</sub> have a well-defined function, playing an essential role in the nuclear translocation (cf. part III.). Although, due its

size, B/NP<sub>TAIL</sub> could probably have a yet unravelled function.



**Figure 9: Structure of Isavirus NP.**

(A) Crystal structure of ISA virus NP (PDB 47EWC) (B) Surimposition of A/NP (PDB 2IQH), coloured in blue, with ISA/NP (PDB 47EWC), in orange.

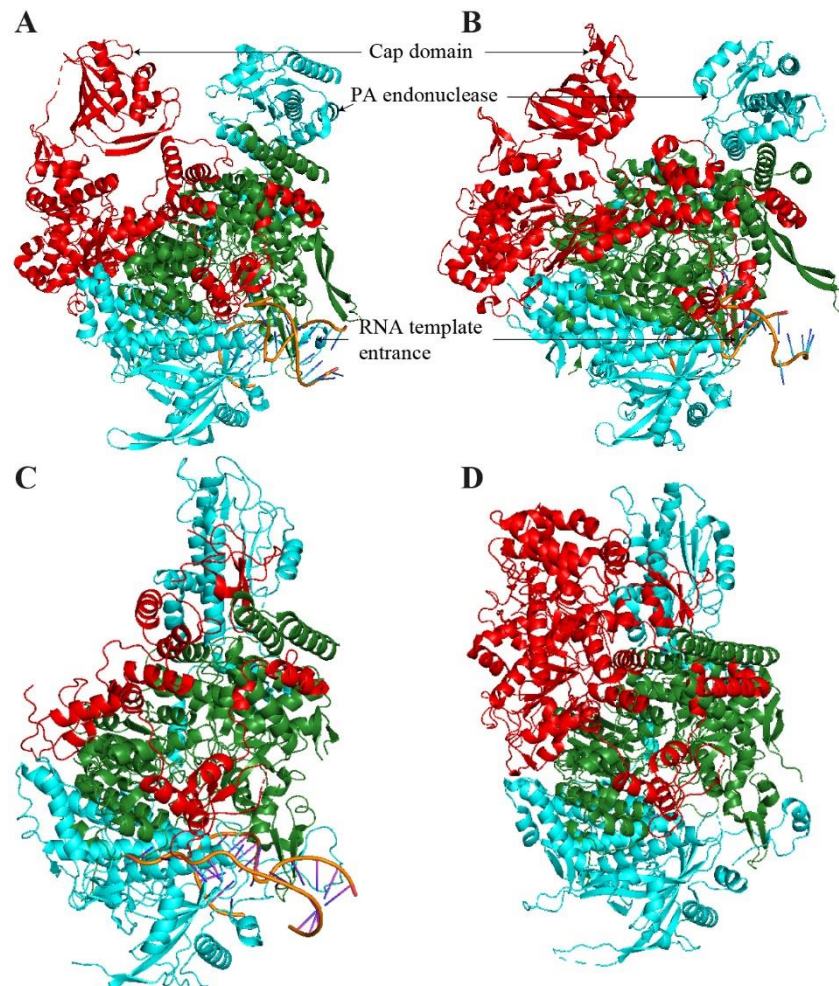
While no structure for C/NP or D/NP was solved at the beginning of this PhD thesis work (cf. Chapter 2), it was possible to make assumptions on their potential folding. The disorder prediction indicates that in contrast to A/NP and B/NP, C/NP and D/NP present only a few unstructured N-terminal residues and a long intrinsically disordered C-terminal (C-ter) region spanning about 50-60 residues (**Fig. 7**). This change from an N-ter NP<sub>TAIL</sub> to a C-ter one is puzzling. Identity between C/NP and D/NP sequences is 38%, similar to the one between A/NP and B/NP, suggesting that their folding could very well be alike.

## ii. *Heterotrimeric RNA-dependent RNA-polymerase*

The influenza RdRp is an intricate heterotrimeric complex of PA, PB1 and PB2, each of them being composed of 700 to 800 residues with a total molecular mass close to 270 kDa.

The influenza polymerase structure has eluded researchers for a long time, due to issues in producing it in a soluble amount. The issue was tackled by firstly solving individual domains of the complex, like the PB2-627 domain<sup>311,352</sup>. Subunit interfaces were initially roughly determined: PA<sub>C-ter</sub> interacting with PB1<sub>N-ter</sub>, which in turn interacts with PB2<sub>N-ter</sub> through its C-ter part. EM analysis showed that the polymerase is a tightly folded complex, suggesting supplementary inter-subunit interaction regions. These last

few years, major breakthroughs have been achieved using X-rays crystallography and cryo-EM allowing to solve structures of IAV, IBV, ICV and IDV polymerases (**Fig. 10**).

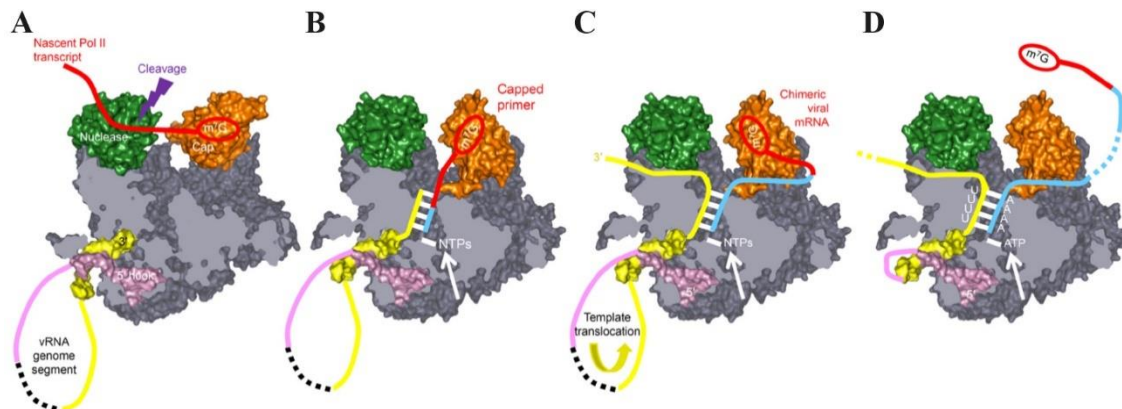


**Figure 10: Bat IAV, IBV, ICV and IDV RdRp structures.**

(A) Crystal structure of bat IAV RdRp (PDB 4WSB) and (B) crystal structure of IBV RdRp (PDB 4WSA). (C) Crystal structure of ICV RdRp (PDB 5D9A) with bound RNA (D) Cryo-EM structure of IDV Apo RdRp (PDB 6KUT). PB1, PB2 and PA are respectively in green, red and cyan. Well-defined domains such as PB2 cap domain and PA endonuclease domain have been identified when fully crystallized. RNA promoter is coloured in orange.

The first two influenza RdRp structures solved were a bat IAV RdRp and an IBV RdRp (**Fig. 10A and 10B**)<sup>273,289</sup>. These structures reveal further subunit interfaces, with each subunit tightly entangled with each other. They nevertheless possess an important intrinsic flexibility, as other structures were released with domains caught in alternative positions<sup>119,359</sup>. These two first macromolecular complexes are structurally well conserved, bare for the intrinsic flexibility of the cap-binding domain of PB2, which

presents itself in a different orientation. Both structures were solved with a bound promoter RNA. Structural insights have also been snatched from transcribing and replicating IAV, IBV, ICV and IDV polymerases, giving precious step-by-step descriptions of the polymerase activity (**Fig. 10C, 10D and 11**)<sup>75,119,168,272,380</sup>.



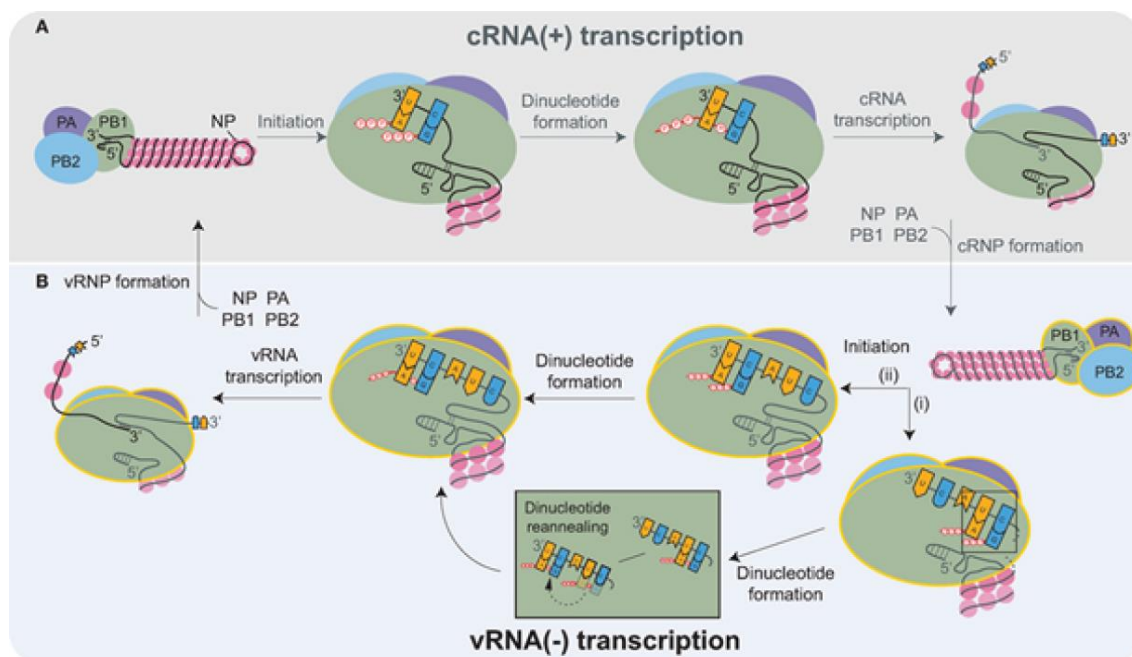
**Figure 11: Reconstituted model of influenza RdRp transcription.**

(A) PB2 performs its cap-snatching activity on a cellular mRNA and PA cleaves the mRNA before (B) PB2 rotates to present the cleaved capped primer to PB1 (C) that performs the elongation up to (D) the stuttered polyadenylation. Adapted from<sup>274</sup>.

Each subunit presents essential specific features and activities. PB1 is the core protein for transcription and replication, possessing the RNA polymerase activity. PB2 and PA present activities related to transcription and more particularly its initiation: as the influenza polymerase does not possess a guanylyl-methyltransferase activity for the mRNA capping and as viral mRNAs use the cap-requiring cellular machineries for transport and translation, a mechanism called ‘cap-snatching’ is used (**Fig. 11**)<sup>29,275</sup>. Briefly, PA is able to bind to the C-terminal domain (CTD) of the cellular DNA-dependent RNA polymerase II (Pol II)<sup>195,379</sup>. PB2 binds to the cap of host mRNAs, by snatching it from the transcribing cellular Pol II<sup>108</sup>. PB2 presents this capped RNA through a rotational change to the PA subunit. With its endonuclease activity, PA cleaves the host capped mRNA 10 to 13 nucleotides downstream<sup>58</sup>. These few nucleotides left are then used as primers by PB1 to synthesize the viral mRNA. The polyA tail is synthesized in a particular way by the RdRp, by stuttering over a patch of 5 to 7 uridines at the 5’ part of the vRNA<sup>84,278</sup>. The influenza RdRp has been shown to adopt two major distinct conformational states, *i.e.* transcriptionally active or inactive. It has been shown that the interaction between the influenza RdRp and the cellular

RNA polIII is required for the RdRp to switch from an inactive state to an active state<sup>379</sup>. The dissociation timing from the cellular polIII remains obscure.

In contrast to transcription, due to the negative sense of the viral genome, influenza polymerase realizes the replication in two steps (**Fig. 12**). Firstly, viral segments of positive sense are synthesized to serve as a template. This template then goes through another round of synthesis to perform the actual viral genome replication and generate negative sense genomic segments<sup>118</sup>. PB1 seems to be able to initiate and terminate the replication on its own, in a primer-independent manner, but making replication 300 times slower than transcription<sup>288</sup>. Interestingly, during the replication, several RdRp complexes can be found on the viral RNA, a feature absent during the transcription<sup>148,218</sup>.



**Figure 12: Replication of the influenza genome.**

The replication occurs in two steps: **(A)** the viral replication machinery transcribes cRNAs and promotes cRNPs assembly followed by **(B)** the transcription of vRNAs. Adapted from<sup>65</sup>.

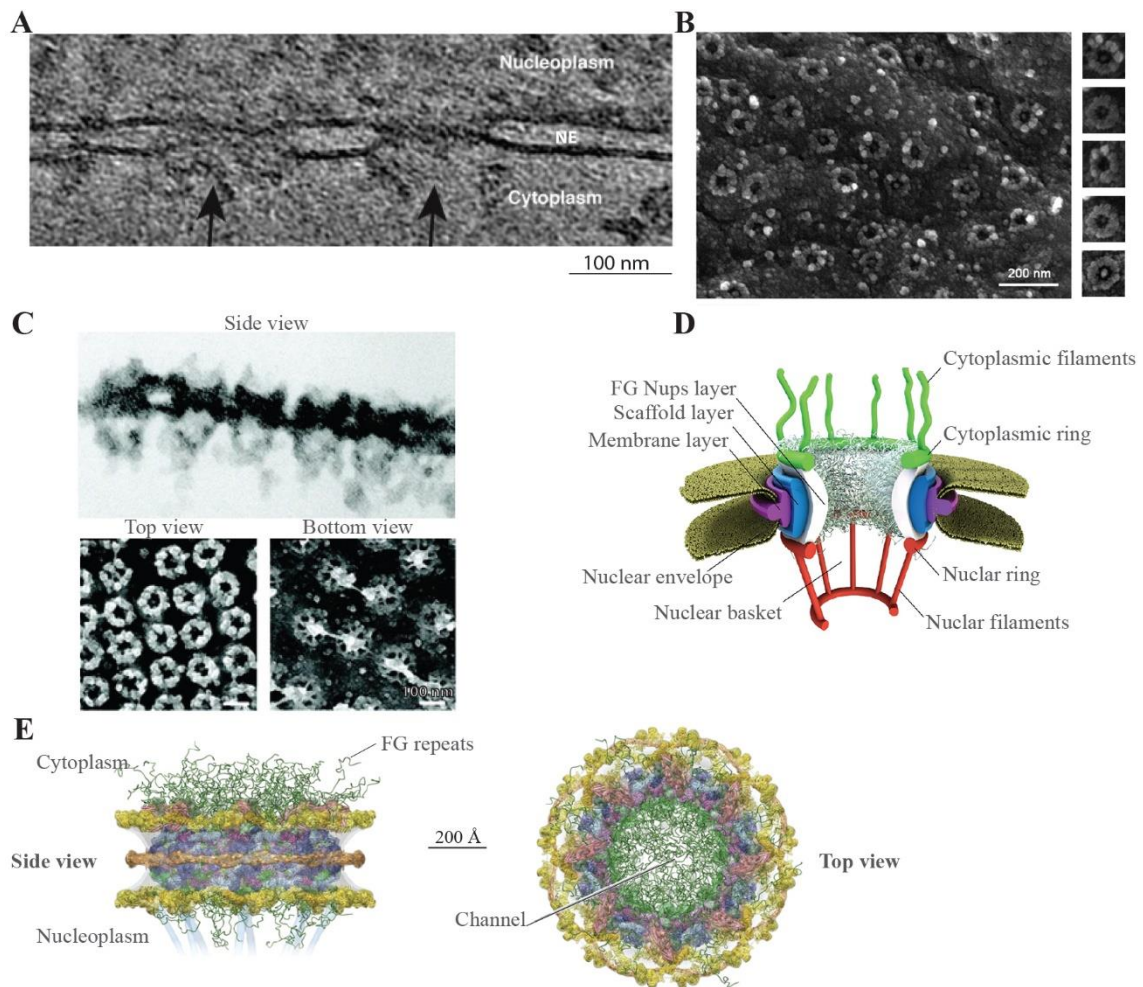
The balance between transcription and replication is not understood. However, the current model suggests that the switch from the former to the latter is not clearly cut but rather a progressive transition. The polymerase dimerization is thought to play an important role in the replication, and more particularly, in the synthesis of negative-sense RNAs from positive sense RNAs<sup>75,378</sup>. The NP:polymerase interaction is also a viral factor putatively playing a part in this process. Investigations of the role of several

cellular factors such as the cellular PolIII and cellular capped-mRNA have been carried out <sup>371</sup>. A second model emerged during the last decade, based on coexisting replication *in trans* and transcription *in cis*, where the newly synthesized RdRps would perform the replication while the RdRp imported with the viral genome as RNPs would do the transcription <sup>83,372</sup>. In spite of huge breakthroughs this decade, further work is required to dissect completely this particular transition.



## II. Nucleocytoplasmic shuttling of proteins

In cells, a huge number of different chemical and biological processes occur. Compartmentalization is a key feature to separate them, leading to a higher efficiency and allowing distinct specific functions to be carried out at the same time <sup>307</sup>.



**Figure 13: Nuclear pore complex (NPC).**

(A) Picture of tomographic slices of HeLa cell nuclear envelope with NPCs identified with black arrows. NE stands for nuclear envelope. Adapted from <sup>30</sup>. (B) Field emission scanning electron microscopy pictures of COS-1 cell nuclei. Adapted from <sup>326</sup>. (C) Transmission electron microscopy (TEM) pictures of the nuclear envelope with NPCs. Cytoplasmic filaments and nuclear baskets can be seen respectively in the top and bottom view (cellular/nuclear TEM pictures) as well as in the side view picture. Adapted from <sup>74</sup>. (D) Schematic view of the NPC. Adapted from <sup>10</sup>. (E) Structure of the NPC with modelled disordered FG-repeats. Adapted from <sup>160</sup>.

The physical separation, either membrane-based or protein-based, is commonly found in eukaryotes and to an increasing extent in prokaryotes<sup>52,59,92</sup>. One example is the nucleus, a compartment containing the largest part of the genetic information, housing the gene regulation, the RNA processing and the transcription. While the driving force behind the emergence of this particular organelle is still unclear, several hypotheses are currently debated, most of them agreeing to an efficiency and integrity advantage<sup>142,191,206,397</sup>. In contrast with other membrane-based organelles, in eukaryotes, the nucleus is separated from the cytoplasm by two membranes (**Fig. 13A**). A way to cross this physical barrier is through a large macromolecular complex acting as a channel, called the nuclear pore complex (NPC) (**Fig. 13**)<sup>56,78,81</sup>.

This complex, conserved across all eukaryotes, is made of several hundred copies of around thirty distinct nucleoporins and reaches a molecular weight of about 60 MDa in yeast, up to about 120 MDa in vertebrates (**Fig. 13E**)<sup>4,201,290</sup>. Nucleoporins localized in the NPC channel are enriched in disordered phenylalanine-glycine repeats (FG repeats), represent almost half of the molecular weight of the NPC and are essential for the restriction of the NPC permeability and for the active transport (**Fig. 13D and 13E**)<sup>306</sup>. The NPC shows an eight-fold symmetry, with a diameter and a height estimated at 100-150 nm and 30-80 nm respectively, from yeast to vertebrates (**Fig 13**)<sup>4,306,410</sup>.

It is of note that a single cell possesses between several hundred to thousands NPCs and the composition of these NPCs varies depending on the cell type or maybe even within one single cell<sup>95,194,212,258,287,320</sup>.

#### A. Nuclear transport pathways

The NPC allows ions, small molecules and proteins (usually <30-60 kDa) to diffuse freely through the 40 nm wide channel. Some features of this passive diffusion are still debated, such as the existence of a clearly cut size threshold of 40 kDa or a progressive threshold<sup>30,156,263,264,360</sup>. However, for bigger proteins, active transport by nuclear transport receptors is often required, although some proteins could directly interact with the FG-repeats localized in the channel, allowing a transporter-independent nuclear translocation<sup>50,129,172,173,266</sup>. The most studied transporter-dependent pathway is the system involving the small GTPase Ran and the importins-

$\alpha/\beta$  nuclear transport receptors<sup>1,219,241</sup>. The nuclear transport can also be mediated by the cytoskeleton, which allows a faster nuclear delivery<sup>32</sup>. Some molecules to be translocated (called cargo molecules) can be imported by alternative transport receptors, like lectins<sup>69</sup>. The cargo translocation can also be calcium-dependent, modulated by the calmodulin<sup>113</sup>.

Several pathways exist for the import of cargo molecules that seem to be able to use more than one system, likely in order to prevent the complete blockade of any essential metabolic and/or regulatory pathways if one transport system is impaired.

## B. The nuclear transport pathway based on Ran and karyopherins

### *i. The Ras-like small GTPase Ran*

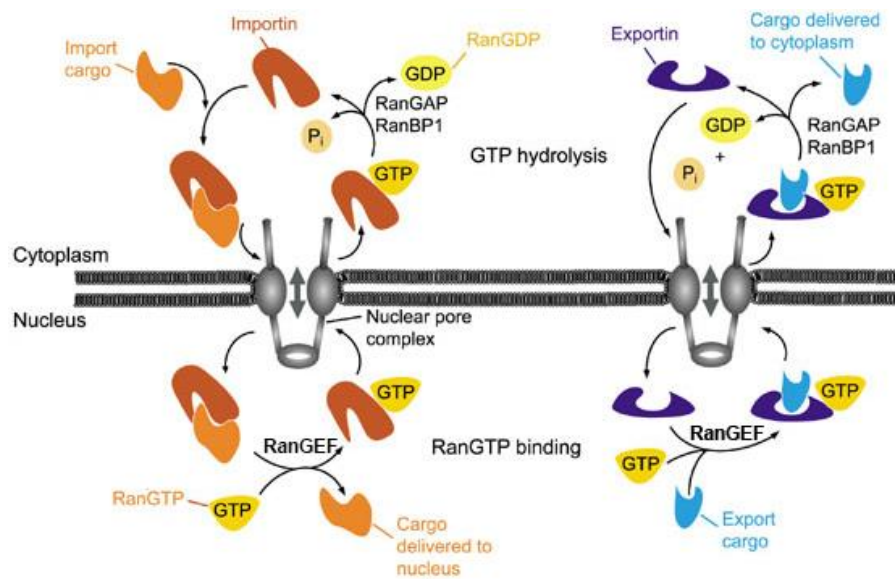
The keystone of the Ras-like small GTPase Ran system is based on the RanGDP/RanGTP gradient existing between the cytosol and the nucleus<sup>103,139</sup>.

Two nucleoside-bound forms of Ran coexist: the GDP-bound Ran form, mostly found in the cytosol and the GTP-bound form abundant in the nucleus. Like the majority of small GTPases, the GTP hydrolysis mediated by Ran is very slow. Two proteins, RanGAP and RanBP1 are found in the cytosol to accelerate the GTP hydrolysis<sup>22,211</sup>. Following the gradient, the GDP-bound Ran ends up in the nucleus. Another partner of Ran, the chromatin-associated RCC1 (also called RanGEF), is localized in the nucleus and induces nucleoside exchange from GDP to GTP<sup>22</sup>. Once more in the GTP-bound form, Ran is redirected towards the cytosol by the gradient.

These differential concentrations of GTP and GDP due to Ran partners lead to the gradient, the shuttling of both Ran forms and drive the bidirectionality of the transport.

### *ii. The nuclear import and export: importins- $\alpha$ , importins- $\beta$ , and exportins*

The Ran import system relies on nuclear transport receptors (**Fig. 14**). Import and export receptors, usually members of the karyopherin superfamily, use distinct nucleoside-bound forms of Ran, to make this system efficient<sup>49</sup>.



**Figure 14: The nuclear transport pathway based on Ran.**

The small GTPase Ran mediates two nuclear transport processes: the nuclear import via the importins- $\alpha/\beta$  and the nuclear export relying on the exportins. Adapted from <sup>342</sup>.

The nuclear export is mediated by proteins called exportins that interact with the GTP-bound form of Ran. These exportins recognize specifically their cargo through a specific signal called Nuclear Export Signal (NES), usually made of hydrophobic residues <sup>109,339</sup>. The most known and studied is the leucine-rich NES <sup>176</sup>.

The nuclear import is mediated by importins, and more particularly importins- $\alpha$  and importins- $\beta$ . Importins- $\beta$  require the GDP-bound form of Ran for the efficient import of its cargo, as RanGTP induces the cargo dissociation, and can directly bind the cargo for its transport <sup>102,295</sup>. In some cases, adapter proteins, namely importins- $\alpha$ , are required. They are able to interact with the cargo through a specific signal called a Nuclear Localization Signal (NLS) and with the importin- $\beta$  through an autoinhibitory importin- $\beta$  binding (IBB) domain that is otherwise folded into the importin- $\alpha$  NLS binding pockets <sup>101</sup>. Upon the hydrolysis of GTP and the availability of importins- $\alpha/\beta$  and the cargo, a likely cooperative, synergic mechanism allows the formation of a complex between the importin- $\beta$ , the importin- $\alpha$  and the cargo (or more directly between an importin- $\beta$  and a cargo with a specific NLS) <sup>193,373</sup>. As the NLS binding pockets of the importin- $\alpha$  are free from the IBB domain, bound by the importin- $\beta$ , the cargo can be loaded onto the dimeric transport complex and translocated into the nucleus. The transport through the NPC channel is mediated by transient interactions between the

importin- $\beta$  and the FG-repeats of nucleoporins, although the precise mechanism is still debated <sup>18,104</sup>. The release of the cargo in the nucleus relies on the conformational change of the importin- $\beta$  induced by a RanGDP-to-GTP exchange, along with the IBB domain autoinhibition and the nucleoporins contribution <sup>100,114,185,186</sup>.

### iii. The nuclear localization and export signals

NLSs can be split into two families: the conventional/classical (cNLS) and the non-conventional/classical NLS (ncNLS) <sup>123,167,181,406</sup>.

**Table 4: Reference NLS sequences and consensus sequences for 6 classes of conventional and nonconventional basic NLSs interacting with importins- $\alpha$ .**

Basic residues are represented in blue. The red X can be replaced by any residue. Z stands for any amino acids but Asp and Glu. \* Most optimal sequences of bipartite NLSs show a central linker region enriched with acidic residues and terminal regions enriched with Pro and basic residues. Adapted from <sup>167</sup>.

Reference NLS	Sequences
Monopartite SV40 Large T Antigen NLS	PKKKRKV
Bipartite nucleoplasmin NLS	KRPAATKKAGQAKKKK
NLS class	Consensus sequences
Class 1	1/ KR(K/R)R 2/ K(K/R)RK
Class 2	(P/R)XXKR(Z)(K/R)
Class 3	KRX(W/F/Y)XXAF
Class 4	(R/P)XXKR(K/R)(Z)
Class 5	LGKR(K/R)(W/F/Y)
Bipartite	1/ KRX <sub>10-12</sub> K(K/R)(K/R)* 2/ KRX <sub>10-12</sub> K(K/R)X(K/R)*

The most detailed NLSs belong to the classical NLS group and are usually made of small patches of basic residues, rich in arginine and lysine. This group presents two types of basic patches: monopartite and bipartite NLSs. The first NLSs identified in the monopartite and bipartite categories are respectively the Simian Virus (SV) 40 Large T antigen and the nucleoplasmin NLSs (**Table 4**) <sup>61,152,299,332</sup>. The discovery of these NLSs was the starting step to determine the consensus sequences, albeit loose ones (**Table 4**). Monopartite NLSs present a single basic patch separated in diverse classes and consensus sequences <sup>167</sup>. Bipartite NLSs present two highly basic patches separated by a linker of 10 to 12 residues with a consensus sequence usually KRX<sub>10-</sub>

$_{12}K(K/R)X(K/R)$  or  $KRX_{10-12}K(K/R)(K/R)$ . Longer linker sequences have also been described<sup>86,180</sup>.

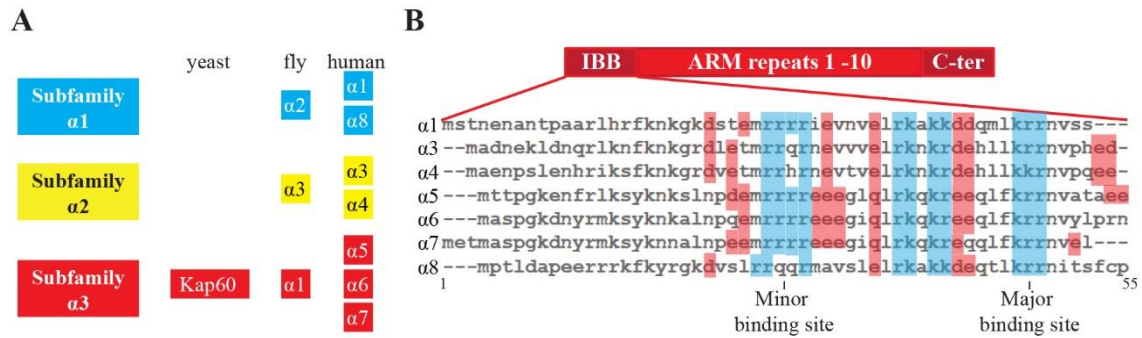
In addition, some NLSs do not show this basic feature but rather a hydrophobic pattern interacting with the importin- $\alpha$  using the same binding pockets as the basic NLS<sup>38,399</sup>. The classical NLSs require the importin- $\beta$ -bound importin- $\alpha$  adapter protein to be translocated, while non-classical NLSs tend to bypass these proteins and to interact directly with importins- $\beta$ . Such atypical cargos can present a proline-tyrosine rich sequence (PY NLS)<sup>366,406</sup>. PY NLSs are not as studied as basic NLSs but work is done towards their characterization. They tend to be longer, disordered and more variable. Identifying NLSs is a tricky task, also due to the fact that some NLSs are not linear but conformational<sup>6,161</sup>.

The dissections of nuclear localization signals have been greatly improved these last decades but the full story is not completely clear, with few other mechanisms and NLSs still investigated<sup>337</sup>. With several thousand cargos requiring nuclear translocation in a timely manner and new putative motifs identified recently, it is to be expected that some discreet patterns still elude discovery<sup>161</sup>.

### C. Similarities and specificities of importins- $\alpha$

Importins are highly conserved through evolution (cf. **Annexe 1**)<sup>209</sup>. Whereas the yeast possesses only one importin- $\alpha$  and 14 importins- $\beta$ , the human genome encodes as much as seven forms of importins- $\alpha$  and at least ten importins- $\beta$  involved in nuclear import<sup>41,256,342</sup>.

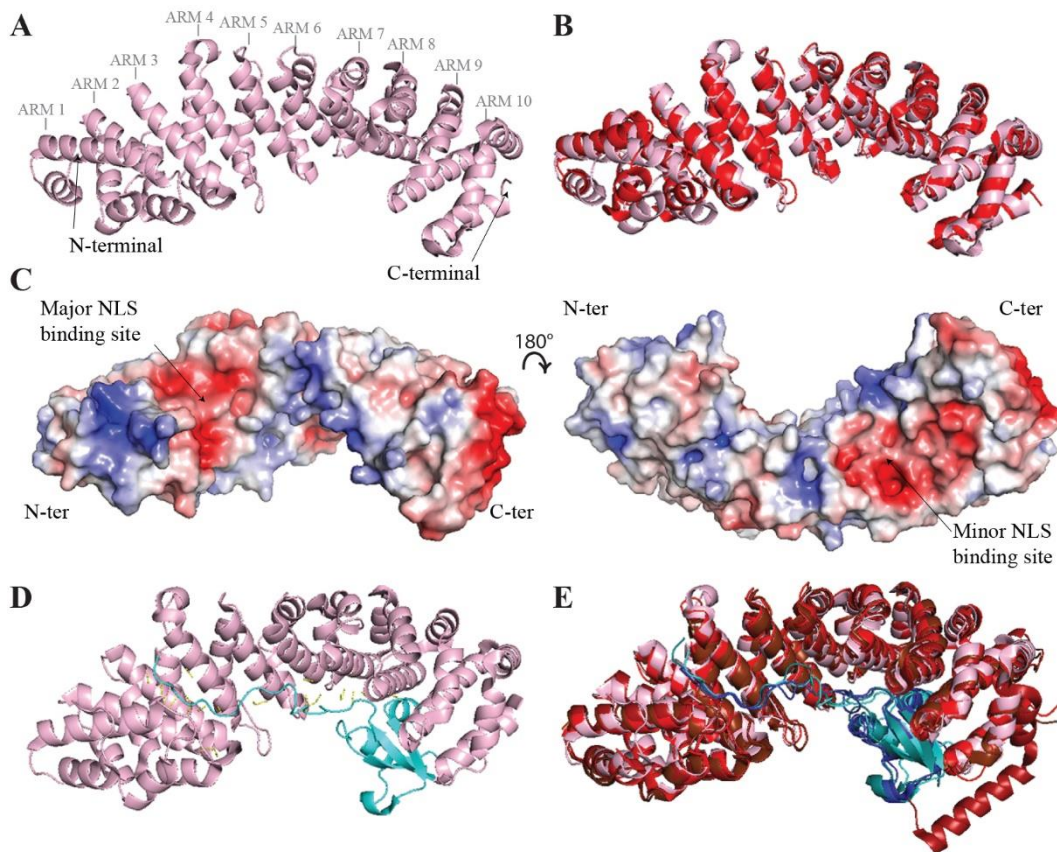
The seven human importins- $\alpha$  are split in three groups (**Fig. 15A**)<sup>282</sup>. The  $\alpha$ 1 clade includes the importin- $\alpha$ 1 and - $\alpha$ 8, which are the more divergent, with a sequence identity of 53 %. The  $\alpha$ 2 clade includes the importin- $\alpha$ 3 and the importin- $\alpha$ 4 that present a strong sequence identity (86 %). The last clade ( $\alpha$ 3) includes the importin- $\alpha$ 5, - $\alpha$ 6 and - $\alpha$ 7. These importins- $\alpha$  are well conserved, with sequence identities between 79 % and 84 %.



**Figure 15: Evolution, diversification and global organization of importins- $\alpha$ .**

(A) Diversity and subfamilies of importins- $\alpha$  in yeast, fly and human. Blue, yellow and red correspond respectively to the  $\alpha$ 1,  $\alpha$ 2 and  $\alpha$ 3 subfamilies. Inspired from <sup>282</sup>. (B) Schematic organization of importins- $\alpha$  and importin- $\beta$  binding (IBB) domain alignment. Basic and acidic charged residues of the NLS-like sequence are respectively in blue and red. Basic patches binding to NLS binding pockets are identified as 'minor binding site' and 'major binding site', accordingly to the auto-inhibited state. Sequences and accession numbers were recovered from NCBI/Uniprot databases for human (*Homo sapiens*) importins- $\alpha$  (cf. **Annexe 3**).

All importins- $\alpha$  present a common, conserved architecture (**Fig. 15B**) <sup>164</sup>. They are crescent-shaped, rather flat proteins, composed by one N-terminal IBB domain (with a NLS-like sequence), responsible for the auto-inhibition, the binding to the importin- $\beta$  and the cargo release in the nucleus <sup>44,114,193</sup>. The IBB domain is made of two basic patches separated by an acidic linker sequence of around 15 residues and flanked by acidic residues. Following this flexible domain, importins- $\alpha$  are composed by 10 armadillo (ARM) repeats, each one composed of repeats of hydrophobic residues forming three  $\alpha$ -helices. The C-terminal part is an acidic patch, responsible for the nuclear export of the cargo-free importin- $\alpha$  by the Cellular Apoptosis Susceptibility protein (CAS) <sup>175</sup>. The best-characterized importin- $\beta$ , the importin- $\beta$ 1, on the other hand, is structured by 19 HEAT repeats, with each repeat composed by two  $\alpha$ -helices. However, other importins- $\beta$  present a variable number of HEAT repeats: for example, RanBP5 is composed of 24 HEAT repeats <sup>345</sup>. While ARM and HEAT repeats are ubiquitous, closely related structures, the latter type of repeats confers a higher degree of flexibility, while ARM repeats result in a more rigid structure <sup>5,154,283</sup>.



**Figure 16: Structure of importins- $\alpha$ .**

(A) Crystal structure of  $\Delta$ IBB importin- $\alpha$ 1 in the free state (PDB 3WPT). (B) Structural comparison of  $\Delta$ IBB importin- $\alpha$ 1 (light pink) and  $\Delta$ IBB importin- $\alpha$ 3 (PDB 6BVZ) (red) in their cargo-free state. (C)  $\Delta$ IBB importin- $\alpha$ 1 represented as surface with its electrostatic potential and NLS binding pockets. (D) Crystal structure of  $\Delta$ IBB importin- $\alpha$ 1 in complex with influenza PB2 NLS (PDB 4UAF), bound to both NLS binding sites. (E) Structural comparison of complexes involving PB2 NLS with  $\Delta$ IBB importin- $\alpha$ 1 (cyan with light pink),  $\Delta$ IBB importin- $\alpha$ 3 (PDB 4UAE) (deep teal with red),  $\Delta$ IBB importin- $\alpha$ 5 (PDB 2JDQ) (blue with dark red) and  $\Delta$ IBB importin- $\alpha$ 7 (PDB 4UAD) (dark blue with chocolate).

Importins- $\alpha$  present two NLS-binding pockets. The major binding pocket, localized over the ARM repeats 2 to 4, is the interaction site for most of the monopartite NLSs with sites to accommodate five NLS residues. The minor binding pocket possesses four binding positions and is part of the ARM repeats 6 to 8. Bipartite NLSs tend to associate with both binding sites (**Fig. 16**)<sup>48</sup>. Interestingly, some NLSs tend to interact so strongly with importins- $\alpha$  that they can directly compete with the IBB domain and take its place, without any contribution of the importin- $\beta$ . Crystallographic structures of importins- $\alpha$  show few differences between paralogs and while they possess some sort of flexibility, it is not to the extent of the one of importins- $\beta$ , which goes through major structural changes (**Fig. 16B and 16E**)<sup>283</sup>.



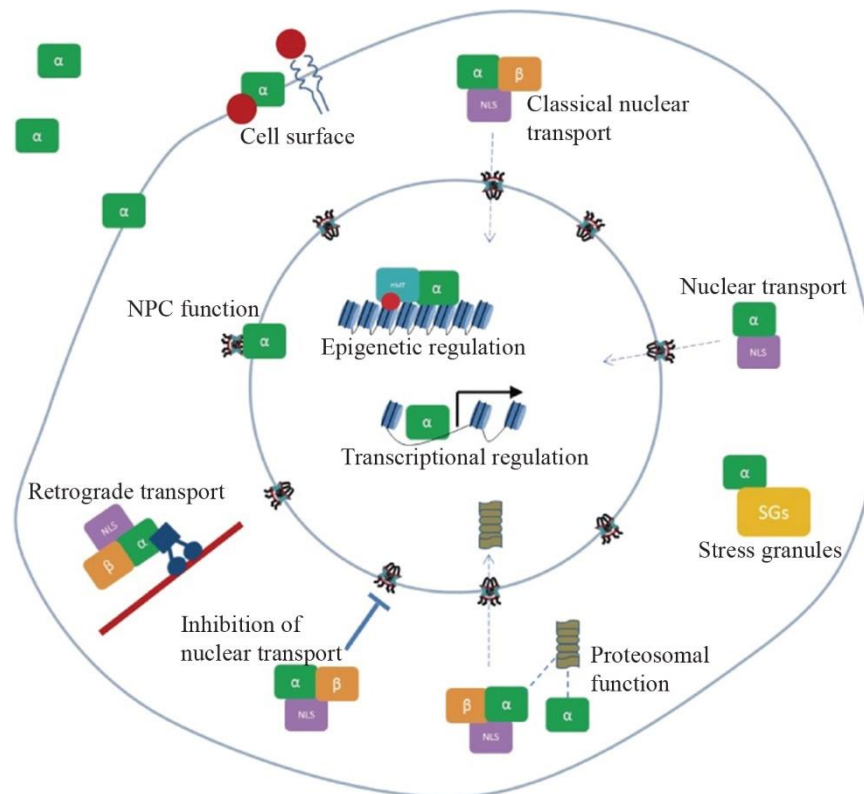
Despite these shared features, importins- $\alpha$  are not interchangeable proteins. While it has been shown that cargoes are able to interact with several paralogs of importins- $\alpha$ , they possess a preferential adapter protein<sup>91,286</sup>. Similarly, importins- $\alpha$  also show preferential binding to specific cargoes<sup>166,389</sup>.

An interesting specificity of importins- $\alpha$  comes from their biochemical properties. It has been shown that the importin- $\alpha 3$  is unusually flexible for an importin- $\alpha$ , but far from the flexibility of importins- $\beta$ <sup>283</sup>. It has been proposed that this peculiar importin- $\alpha$  was conserved through evolution for fast nuclear translocation, and/or for interacting with specific cargoes, with less common NLS or possessing a globular domain close by, requiring the adapter to curve and bend to bind it efficiently<sup>314</sup>. The importin- $\alpha 3$  seems less well adapted to interact with classical monopartite NLSs. In contrast, the importin- $\alpha 1$  appears rigid and tends to better interact with monopartite NLSs. The importin- $\alpha 7$  seems to be more flexible than the importin- $\alpha 1$  but not as much as the importin- $\alpha 3$ <sup>283</sup>. For other importins- $\alpha$ , such analysis has not been carried out yet.

Another specificity is the strength of the importin- $\alpha$  auto-inhibition. It has been found that the importin- $\alpha 1$  presents the strongest auto-inhibition while the importin- $\alpha 3$  auto-inhibition is the weakest; the importin- $\alpha 7$  auto-inhibition seems in between<sup>217,283</sup>. This is partly because subtle differences exist between paralogs, more particularly in the NLS sequence of the IBB domain, and in the flanking regions that are more or less acidic (**Fig. 15B**).

Importins- $\alpha$  are expressed in all cell types and tissues to some extent, the only exceptions being the importins- $\alpha 6$  and - $\alpha 8$ , which are more restricted respectively to the testis and to the ovaries. However, some tissues and cell types are enriched in one or several importin- $\alpha$  paralogs, and importins- $\alpha$  themselves are not homogeneously expressed. For example, based on the Human Protein Atlas, the importins- $\alpha 7$  is highly expressed in a large number of tissues while the importin- $\alpha 1$  is moderately expressed in most tissues, sometimes highly, lowly or not expressed at all<sup>357,367</sup>. This is particularly interesting when put in a more cellular context, with importins- $\alpha$  specificities towards cargoes. For example, the importin- $\alpha 3$  seems to be the most expressed importin- $\alpha$  in the mammalian respiratory tract, followed by the importin- $\alpha 7$ , while importin- $\alpha 1$  is lowly expressed in most respiratory tissues<sup>246,358</sup>. In contrast, the importin- $\alpha 1$  is highly expressed in most types of cancers and as most of the research

is based on immortalized and/or cancer cell lines, caution is required for the design and the interpretation of studies involving such an importin and such cells <sup>112</sup>.



**Figure 17: Diversity of the roles played by importins-α in cells.**

Besides their well-known role in nuclear transport, importins-α are involved in numerous other processes in cells, as regulatory proteins or in trafficking. Adapted from <sup>256</sup>.

In addition to their role in the nuclear import, importins play other, more discreet parts in the cell life (**Fig. 17**). The importin-α has been reported for sequestering a cargo in the cytoplasm, adding another layer for the regulation of nuclear processes <sup>87,322</sup>. Importins-α are expressed in a development-dependent, cell type-dependent manner: the regulation of the nuclear import through the importins-α differential expression is one of the switches used during the cell development, the differentiation and the aging <sup>153,165,354,364,413</sup>. It is of note that knockout or knockdown of specific importins-α often lead to issues in reproduction and development. Such a defect can be lethal <sup>220,305</sup>. Importins-α can also be found in an active form on the cell surface of a few cancer lines, where they putatively can bind to growth factors and transport them to the nucleus <sup>256</sup>. Importins play essential roles during the mitosis, for the regulation

of NLS-bearing proteins and for the spindle assembly<sup>107,395</sup>. They can also control the gene expression, by transporting transcription factors or by binding DNA during various stresses or in cancer cell lines<sup>19,412</sup>. Their role as cytoplasmic chaperones has been investigated, to prevent the aggregation of highly basic proteins, although only a few importins- $\beta$  have been identified yet<sup>140</sup>.

D. *In vitro* and *in cellulo* interaction parameters of the complex between cargo and importins: affinity and rates

A number of studies focused on the molecular basis for the interaction between importins and cargo proteins. While the *in cellulo* affinity, thermodynamics and rates of interactions eluded researchers for a long time, dissecting these parameters was more doable using *in vitro* characterization with purified partners. Most of the current knowledge comes from a large number of biochemical, biophysical and structural methods, which bring precious information about these interactions. *In vitro* characterizations highlighted that all partners in this complex nuclear import pathway (*i.e.* cargo proteins, importins, NPC proteins) have a surprisingly high affinity for each other. Interactions between cargo proteins and IBB-truncated importins, between importin- $\alpha$  and importin- $\beta$  or between importins and nucleoporins typically yield a dissociation constant in the nanomolar range<sup>18,35,76,192,356</sup>.  $\Delta$ IBB importins- $\alpha$  often show a nanomolar affinity for NLSs, with sometimes a low nanomolar affinity to bipartite NLSs or a low micromolar affinity. Some cargo proteins are even able to compete with the IBB domain. It has been shown previously that bipartite NLSs can displace the NLS-like sequence of the IBB domain in the absence of the importin- $\beta$ <sup>190,283</sup>. The same bipartite NLSs tend to interact strongly ( $K_d$  around 10 nM) with several importins- $\alpha$  *in vitro*, regardless of their cargo specificities. In mouse, the affinity between the importin- $\beta$  and the importin- $\alpha$ 2 was measured at around 15 nM while the affinity for the sole IBB domain was 7 nM. Pre-associated importin- $\beta$ :importin- $\alpha$ 2 complex binds to the SV40 Large T-Antigen and nucleoplasmin NLSs with a dissociation constant of 40 nM<sup>35</sup>. A functional NLS is expected to possess an affinity *in vitro* in the nanomolar range. Outside of this *in vitro* range, the cargo would be too tightly or too loosely bound to its transport receptor.

As it has been shown that the affinity correlates well with the nuclear localization, the majority of studies realized on importin: cargo complexes rely on the dissociation constant obtained *in vitro* to assess the efficiency and functionality of the nuclear import *in cellulo*<sup>122</sup>. However, while this relationship is still true, it simplifies to an extreme the relationship between the transporter and its cargo. As in all interactions, the strength is an insightful parameter but the rates are equally important<sup>20,51</sup>. The rates of association and dissociation are the parameters determining the affinity, and they give precious insights on the half-life of a complex. This is what makes the difference between an efficient and an unproductive interaction: too slow or too fast an interaction leads either way to an issue in addressing the right partner to the right location on time, to carry out its fundamental function.

The major issue with dissociation constants determined *in vitro* for this kind of interaction comes from the fact that a nanomolar affinity correlates with a complex half-life of several minutes<sup>356</sup>. However, with the nucleocytoplasmic shuttling of importins being around 30 milliseconds, this raises questions about the results obtained *in vitro* in light of what can be observed *in cellulo*<sup>170,411</sup>. These last few years, studies have been tackling this issue. Several recent studies showed that the cell environment considerably affects the nuclear transport<sup>33,34,356,361</sup>. Looking at importins: nucleoporins or importins: cargos interactions, it has been shown *in vitro* that non-specific binding, even in the form of extremely diluted bacterial lysate, drastically reduced the affinity between partners. It has been estimated that the environment contribution to nuclear transport interactions is similar to a specific interaction in the micromolar range with a non-specific competition in the millimolar range<sup>33,361</sup>. These estimations are this time much more in agreement with the nuclear import rates, as the half-life of such a complex is in the millisecond range.

## E. Nuclear import regulation through cellular machinery

### *i. Importin- $\alpha$ :importin- $\beta$ sub-complex versus importins- $\beta$ as transport receptors*

Importin- $\beta$  transport receptors are perfectly able to bind and translocate cargo proteins by themselves, without importins- $\alpha$ . The direct interaction of importins- $\beta$  with

cargo leads to faster rates and a more efficient nuclear localization than when the importin- $\alpha$  is required <sup>297</sup>. Furthermore, the importin- $\alpha$ -mediated nuclear import is energy-consuming: while one nucleocytoplasmic cycle relying on importins- $\beta$  requires the hydrolysis of one GTP to GDP, if importin- $\alpha$  is required to play its part as an adapter, the shuttling cost increases to 2 GTP-to-GDP hydrolyses, for importin- $\alpha$  recycling. This raises the question of the conservation and the generation of new importins- $\alpha$  through evolution.

It has been shown that, despite these drawbacks of using importins- $\alpha$  in nuclear import, they have their own advantages: it adds a supplementary, larger layer of regulation and control and offers a broader, preferential choice for cargo loading and delivery depending on the cellular requirements and conditions <sup>296,297</sup>.

## *ii. Post-translational modifications*

As mentioned previously, the nuclear import is a multi-layered, highly regulated process. Another obvious and easy way to achieve this level of control is through post-translational modifications (PTMs) on transport receptors, cargo proteins or even NPC proteins.

Phosphorylation is the most described PTM and is heavily involved in transport regulation. It is well documented that the phosphorylation of cargos can improve or worsen the affinity for an importin and consequently the nuclear delivery <sup>236,355</sup>. In extreme cases, it can completely abolish the interaction or unmask a NLS. Methylation, acetylation, ubiquitination and SUMOylation have also been reported to regulate cargos:transporters assembly, similarly to phosphorylation <sup>17,281,355,382</sup>. For instance, acetylation of an importin- $\alpha$  seems to enhance its binding to importin- $\beta$  and Ran is regulated in almost all its functions by this PTM <sup>12,163</sup>. All of these PTMs are not simply dichotomic; they can be inhibiting or activating modifications, with the most diverse mechanisms. PTMs of the NPC proteins are equally diverse and extensive: phosphorylation, acetylation, ubiquitination, SUMOylation, O- and N-glycosylation have been reported <sup>16</sup>. They are involved in NPC permeability, assembly and disassembly during mitosis.

It is to be expected that other, less-studied PTMs are likely regulating the nuclear import, and further work is required to continue to shed light on this complex mechanism.

*iii. Other regulatory mechanisms*

Cytoskeleton networks and more particularly microtubules and associated transporters are able to regulate this specific nuclear trafficking pathway and are involved in viral nuclear trafficking<sup>221,304</sup>. Interestingly, unusual regulatory pathways are also involved in nucleocytoplasmic transport, such as lipids signalling or NLS/NES proteolysis<sup>14,77</sup>.

### III. Nuclear import of the influenza replication machinery

While influenza viruses are amongst the only RNA viruses performing replication and transcription in the nucleus, several RNA viruses require some sort of nuclear targeting, for replication and/or immune regulation purposes <sup>9</sup>. Therefore, an efficient interaction with the cellular nuclear import machinery is one of the keystones for the viral cycle and is determinant for virulence and fitness <sup>292</sup>. The influenza replication machinery requires nuclear import at two major points of the viral cycle: firstly in the RNPs context and secondly for the newly-synthesized replication machinery proteins. PB1, PB2 and NP interact with either importins- $\alpha$  or importins- $\beta$ , and PA is required for successful nuclear import of PB1. They present basic patches, a sign of potential NLSs (**Table 5**). Studies have long try to dissect this particular host:pathogen interaction, but many questions are still unanswered <sup>205,239,257</sup>.

**Table 5: NLS sequences of the influenza A replication machinery.**

Basic residues are represented in blue.

Protein	NLS type	Sequences
A/NP	ncNLS	NLS1: <sup>3</sup> T <b>KGT</b> KRSYEQM <sup>13</sup>
	Bipartite NLS	NLS2: <sup>198</sup> <b>KRGIND</b> RNFWR <b>GENGR</b> RTR <sup>216</sup>
A/PB1	Bipartite NLS?	<sup>187</sup> <b>RKRR</b> VRDNMT <b>KKM</b> V <b>TQ</b> R <b>TIG</b> <b>KRK</b> Q <b>R</b> <sup>211</sup>
A/PB2	Bipartite NLS	NLS1: <sup>450</sup> GIESIDNVMGMIGILPDMTPSTEMSM <b>R</b> GV <b>RISK</b> MGVDEYSSAE
		KIV <sup>495</sup> NLS2: <sup>736</sup> <b>KRKR</b> NSSILTDSQTAT <b>KRIR</b> MAIN <sup>759</sup>

#### A. Nucleoprotein

The nuclear import of NP and its interaction with importins- $\alpha$  (**Fig. 18**) are likely one of the most extensively studied of all influenza replication machinery proteins.

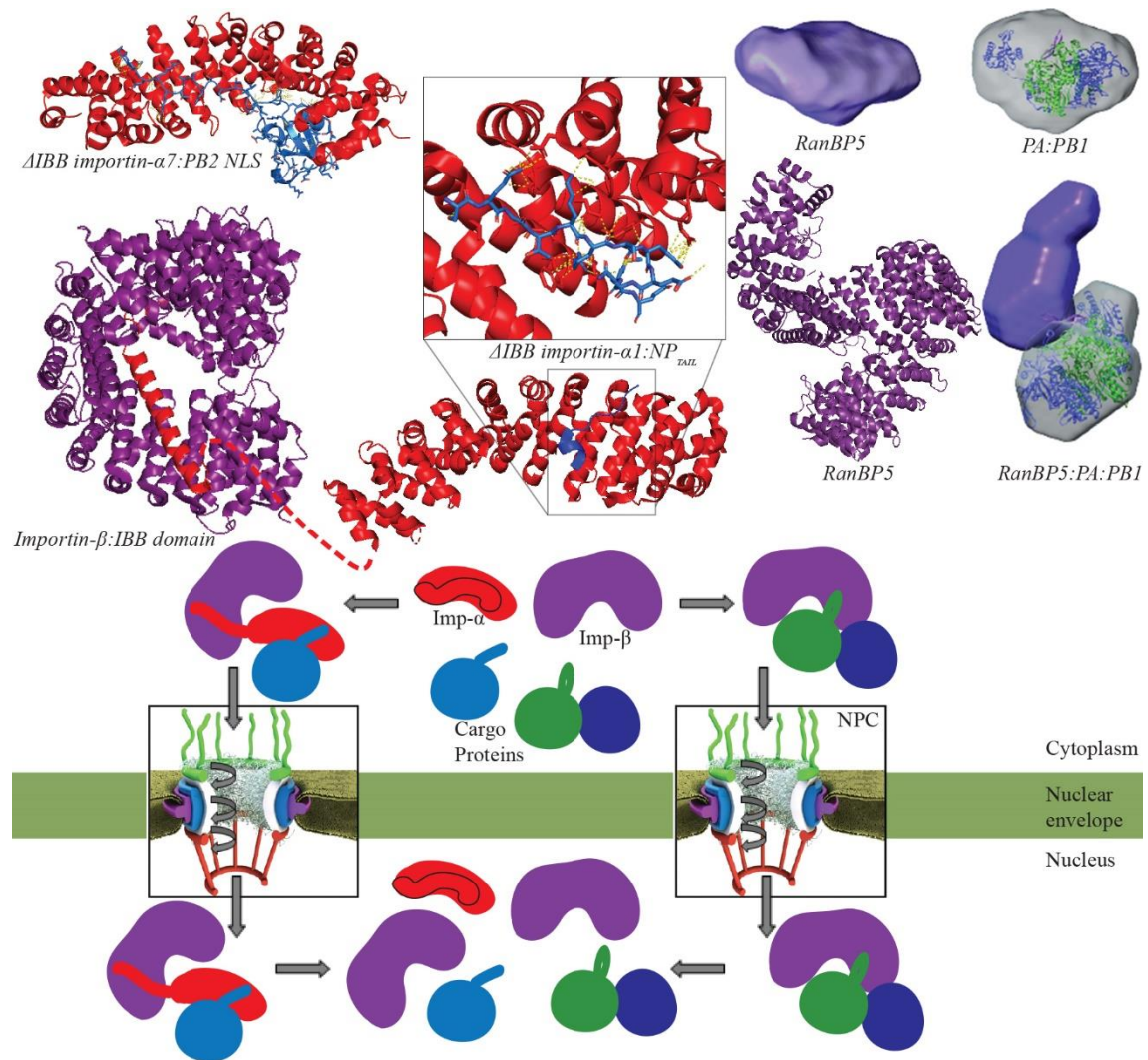
Two basic clusters have been identified in A/NP as putative NLSs. The first one (NLS1) is localized on the flexible N-ter NP<sub>TAIL</sub>, within the 20 first residues, is considered to be part of the category of non-conventional NLS and binds to the minor binding pocket of importins- $\alpha$  <sup>215,239,257,381</sup>. The second (NLS2), considered as a bipartite NLS, is localized in a loop, spanning residues 198 to 216 <sup>261,386</sup>. However its role is unclear:

NLS2 basic patches are structurally too close to constitute a platform for importin binding, NLS2 is not easily accessible and mutagenesis did not prevent nuclear import, only nucleolar localization<sup>53,243,414</sup>. A/NP interacts with several paralogs of importins- $\alpha$ , namely importin- $\alpha$ 1, - $\alpha$ 3, - $\alpha$ 5 and - $\alpha$ 7<sup>93,215,257,294,381</sup>. The crystal structure and characterization of the A/NP<sub>TAIL</sub> in complex with the mouse  $\Delta$ IBB importin- $\alpha$ 1 shows that the A/NP<sub>TAIL</sub> binds to the minor NLS binding pocket of importins- $\alpha$  with a weak affinity for the importin- $\alpha$ 1, from 1 to 5  $\mu$ M<sup>232,403</sup>. A/NP NLS2 showed an even lower affinity, with a dissociation constant of 72  $\mu$ M. The synergic behaviour of both NLSs has been suggested to explain the functionality of the nuclear import at such low affinities<sup>403</sup>. This last hypothesis is still debated. The sole NP<sub>TAIL</sub> is also suggested as the unique A/NP NLS: while the NLS on the A/NP<sub>TAIL</sub> presents a weak affinity for importin- $\alpha$ 1, it seems more plausible for it to be the A/NP NLS than the internal putative NLS sequence.

B/NP, C/NP and D/NP do not seem to present an internal basic NLS sequence. However, the N-terminal flexible tail of B/NP, while longer (71 amino acids), presents several basic residues across its sequence (**Fig. 7**). No classical NLS pattern could be predicted from the sequence alone, but works realized during this PhD thesis validated the B/NP<sub>TAIL</sub> NLS activity. B/NP interacts slightly more strongly with importin- $\alpha$ 7, with an affinity of 844 nM. Interestingly, a long stretch of residues is involved in the interaction, from residue 30 to 71<sup>177</sup>.

In the meantime, it has been shown that the C-terminal flexible tail of C/NP and D/NP (C/NP<sub>TAIL</sub> and D/NP<sub>TAIL</sub>) are also highly basic, possess what seems to be clearly bipartite NLSs, and are able to bind to importins- $\alpha$  (**Fig. 7**). The C/NP<sub>TAIL</sub> is able to interact with importin- $\alpha$ 1, - $\alpha$ 3, - $\alpha$ 4, - $\alpha$ 5 and - $\alpha$ 7 and its affinity for importin- $\alpha$ 1 is around 50 nM, a dissociation constant in complete agreement with previously characterized functional NLSs<sup>350</sup>. Point mutations of the second patch of basic residues (residues 546-549) impaired dramatically the nuclear import and the affinity for importin- $\alpha$ . Structural and nuclear import insights have been gained on D/NP during this PhD and will be addressed in the Chapter 2.





**Figure 18: Summary of the nuclear import of the influenza replication machinery.**

Schematic representation of nuclear import of influenza NP and RdRp subunits along with crystal structures of RanBP5 (PDB 6XTE),  $\Delta$ IBB importin- $\alpha$ 7:PB2 NLS (PDB 4UAD),  $\Delta$ IBB importin- $\alpha$ 1:NP NLS (PDB 4ZDU) and of importin- $\beta$ :IBB (PDB 1QGK) complexes. SAXS envelopes were adapted from <sup>346</sup>. Schematic representation of NPC was adapted from <sup>10</sup>.

The affinities determined so far of A/NP and B/NP for importins- $\alpha$  are too low to lead to an efficient nuclear transport. Another mechanism and/or partner is clearly still eluding discovery to explain the nuclear import functionality observed *in cellulo* during the infection course.

It is of note that most of the work to determine *in vitro* affinities between importins- $\alpha$  and NPs is realized using only peptides -whether the NP<sub>TAIL</sub> or the putative internal A/NP NLS peptide-, which is both an essential step and a bias. It is now known that for some cargos, both the NLS sequence and the 3-dimensional context are important.

NP is not different: other A/NP residues such as amino acids 102, 105 and 375 are able to mutate upon adaptation to  $\Delta$ importin- $\alpha$ 7 mice, to improve the binding to other importin paralogs, such as importin- $\alpha$ 3 in the case of residues 102 and 375 <sup>291</sup>. The interaction modelling showed that at least another non-NLS interface exists in the importin- $\alpha$ :NP interaction.

It is well known now that the A/NP nuclear import is regulated by phosphorylation: phosphorylation at residues Ser-9 or Tyr-10 inhibits the nuclear import while phosphorylation at the position Tyr-296 and Thr-188 leads to a nuclear retention <sup>189,421</sup>. While the B/NP phosphorylation function has not been studied yet, it has been shown that B/NP is also phosphorylated, at positions structurally conserved with A/NP <sup>130</sup>. The nuclear translocation and/or retention could be regulated by SUMOylation at Lys-7 and putatively at Lys-4 <sup>111</sup>.

Several cellular partners have also been identified as able to sequester NP in the cytoplasm, by preventing the interaction with importins <sup>60,196,418</sup>.

The nuclear import of influenza NPs is slowly unravelled, but a number of questions remain unanswered. Along with the highly divergent NLSs exhibited by distinct NP type, it is expected that what is found for one influenza NP will likely not be transposable to the other NPs (cf. **Annexe 2**). Work is still required to completely understand this essential process in the viral cycle.

## B. Heterotrimeric RNA-dependent RNA polymerase

The influenza RdRp is constituted by 3 subunits, each one requiring to be imported in the nucleus upon synthesis. The current scientific consensus is that PB2 is imported by importins- $\alpha$  while PA and PB1 are imported as a subcomplex by the importin- $\beta$  RanBP5 (**Fig. 17**) <sup>131</sup>.

PB2 is able to interact with several  $\Delta$ IBB importins- $\alpha$  (importin- $\alpha$ 1, - $\alpha$ 3, - $\alpha$ 5 and - $\alpha$ 7) with a strong affinity (1-10 nM), due to its bipartite NLS <sup>24</sup>. Two regions are involved in the nuclear localization: the essential bipartite motif spanning the residues 737-759 and a longer sequence between residues 449 and 495 <sup>224</sup>. The mutation of each residue from residue 736 to residue 739 (KRKR motif) results in a cytoplasmic localization, while the deletion of the 449-495 region leads to a perinuclear localization.

Interestingly, the bipartite NLS is localized in a folded C-terminal domain and unfold to bind efficiently to importins <sup>351</sup>. The NLS binds importins- $\alpha$  in an extended conformation.

Several mutations in PB2 are known to impact importin- $\alpha$  bindings. The E627K and D701N mutations are well-known in the bird-to-mammal adaptation, by improving the nuclear import and binding to the mammalian importin- $\alpha$ 1 <sup>93,293,341</sup>. The E627K mutation is also involved in the nuclear import, but more likely by a structural change, not affecting the binding affinity but the protein context, acting as a positive feature for the importin- $\alpha$ 1 and importin- $\alpha$ 7 specificity <sup>127</sup>. R702K and S714R mutations also play a part in the importin binding, increasing the association rate of the complex, thus increasing the binding affinity to importin- $\alpha$ 1 <sup>24</sup>.

The interaction between PA:PB1 and RanBP5 is less well understood. PA and PB1 both present NLSs but the nuclear import seems to be mediated by the PB1 NLS <sup>3,237,244,245,333</sup>. PB1 possesses a bipartite NLS on two protruding  $\beta$ -strands, spanning the residues 187-190 and 207-211, interacting with RanBP5 <sup>57,133</sup>. The absence of PA or the mutation of the PA linker in the PA:PB1 subcomplex leads to an impaired nuclear accumulation <sup>55,85</sup>. Getting insight in the binding affinity has not been achieved yet, but the RanBP5 binding interface has been recently determined: the PB1 NLS could bind as a secondary structure to RanBP5, through RanBP5 residues 362, 405, 486, 489, 490, 528 and 529 <sup>345</sup>.

Interestingly, the use of these distinct nuclear import pathways by PA:PB1 and PB2 could very well prevent the assembly of the polymerase in the cytoplasm, where the replication and the transcription would be unproductive <sup>57,128,346</sup>.

### C. Ribonucleoproteins

The nuclear import of ribonucleoproteins, and viral genomes in general, is an event still poorly understood. However, it has been shown that RNPs go through the NPC to gain access to the nucleus and use the importin- $\alpha$ - $\beta$  cellular nuclear import machinery <sup>157,205,257</sup>. Interactions between RNPs and the NPC, and the NPC crossing are not spontaneous events, with dissociation rates ranging from 1 s to 100 s. They are also regulated, for example with the binding prevented at later infection times by M1 <sup>11</sup>. The

exact mechanism behind the RNP nuclear import is still a mystery; it is well known that the NPC permeability is variable and modified during the cell life course but whether the 7-8 RNPs are transported together to dissociate in the nucleus or the individual RNPs are translocated after dissociation is unknown. One RNP is a long and large structure but it could be small enough to pass through the NPC if constrained and kept straight; such a feat could probably not occur if RNPs are translocated as a bundle. At late infection times, the nuclear export of RNPs is facilitated by the cellular caspase activation that enlarges NPCs; mechanical stress is known to induce breaks and enlargements in NPCs, so maybe a similar mechanism or something along the line of mechanical stress exists for the nuclear import <sup>82,223</sup>.

It is currently admitted that the RNP translocation is mediated by the NP<sub>TAIL</sub> NLS, which is readily accessible in the RNP context <sup>7,53,106,257,350,404,405</sup>. With the large number of NP coating the vRNA, it is estimated that a large number of NP<sub>TAILS</sub> are available to interact with importins- $\alpha$  <sup>362</sup>. Such a high NLS density and high avidity could explain the fast nuclear accumulation. However, it is still not known precisely how the avidity plays a part in the RNP nuclear import or if multiple importins- $\alpha$  mediate the nuclear translocation.



# Objectives of the PhD thesis

The nucleoprotein is a central protein of influenza viruses, essential for many processes, from transcription/replication to cellular trafficking, viral genome protection and viral escape. Yet, many of its features are still shrouded in mysteries.

The nuclear import of the viral genome is most likely mediated by NP that had to mediate its own trafficking by supposedly intrinsically disordered extremities. A/NP<sub>TAIL</sub> has been the most studied, and yet showed a weak binding to the cellular nuclear import factors, the importins- $\alpha$  ( $K_d$  1-5  $\mu$ M)<sup>93,215,232,239,257,294,381,403</sup>. Similarly, B/NP<sub>TAIL</sub> interacts with the importin- $\alpha$ 7 with an affinity in the same range (*i.e.* 844 nM) of the one of A/NP<sub>TAIL</sub><sup>177</sup>. In contrast, C/NP<sub>TAIL</sub> showed a functional binding to the importins- $\alpha$ , but the study was limited to a thorough characterization with only one importin- $\alpha$ , the importin- $\alpha$ 1 (48 nM)<sup>350</sup>. D/NP<sub>TAIL</sub> dissection was initiated before my arrival in the team, and finalized during this PhD work. All of these NP<sub>TAILS</sub> are extremely conserved within one influenza type, but highly divergent between types, suggesting that their behaviour with importins- $\alpha$  is a specific and an optimized feature (*cf.* **Annexe 2**). Beyond these differences in affinities and partners, another issue comes from the fact that distinct methods and techniques had been used to get these insights, making any comparison a difficult process. The metagenomics studies from these last few years showed that the understanding of the *Orthomyxoviridae* family and of the influenza viruses is in fact rather limited: the discovery of several new influenza-like viruses was an extraordinary opportunity to understand the differences, the specificities and the evolution processes at work.

The first objective of my work was to get functional data on the rates and affinities of the complexes between the D/NP<sub>TAIL</sub> and the importin- $\alpha$ 7 with a new strategy that was planned to be set-up for all types of NP:importins- $\alpha$  interactions. Once done, this strategy was then extended to the interaction between all the influenza NP<sub>TAILS</sub> and all the cellular partners previously identified, namely the importin- $\alpha$ 1, - $\alpha$ 3, - $\alpha$ 5 and  $\alpha$ 7, and maybe later on to the other proteins of the influenza replication machinery. This allows for the first time to compare and exactly characterize which importin- $\alpha$  could be a preferential partner of each NP<sub>TAIL</sub>, and to dissect more precisely the features, similarities and specificities of all these partners.

The second aim was to get structural and functional data on the NPs of two recently discovered influenza viruses: the influenza D virus and the Wuhan Asiatic toad

influenza virus. The influenza D virus is more closely related to the influenza C virus, infects a broad spectrum of mammals and its reservoir seems to be the cattle <sup>8</sup>. The Wuhan Asiatic toad influenza virus is more closely related to influenza B viruses and has been isolated from a toad <sup>329</sup>. Detailing firstly the nuclear import function, along with the other specific functions of the influenza NP, provide precious information on the role and behaviour of NPs, along with important insights in the evolutionary relationship within the influenza family.

The results obtained during this PhD project revealed how differently each influenza NP was behaving, more particularly in the interaction with the cellular import machinery despite their common structural and functional basis, and to what extent the importins- $\alpha$  were individual entities in this context, able to pick a specific, preferential cargo amongst others, in contrast with the conventional view on these adapters.





Chapter 2: Structural analysis of the influenza D nucleoprotein and characterization of its nuclear localization signal

In 2011, a new influenza virus affecting the cattle was discovered, while it was still believed to be spared by the influenza virus infections<sup>117</sup>. Initially thought to be a C-like influenza virus, further experiments lead to its classification as a new type of influenza virus in 2013, the influenza D virus (IDV)<sup>116</sup>. In the following years, most of the research on this new virus were focused on epidemiologic, pathologic and phylogenetic analyses. It was found to be distributed worldwide with a high seroprevalence in the cattle and able to infect a diversity of mammal species<sup>68,116,117,285</sup>. IDV does not seem to cause any severe disease on its own, only when associated with other pathogens<sup>68,216</sup>. Currently, two distinct lineages circulate<sup>8</sup>.

More recent studies are now focusing on the IDV proteins. The first IDV protein studied was the surface glycoprotein, HEF, followed by NS1<sup>251,335</sup>.

In this paper, we carried out a biochemical and structural analysis of this novel influenza NP, using X-rays crystallography, electron microscopy, circular dichroism, SEC, SEC-MALLS analysis, thermal shift assays, fluorescence anisotropy (FA) and surface plasmon resonance (SPR). Insights in the structure, the oligomerization, the RNA- and the importin- $\alpha$ - binding processes were obtained and highlighted both D/NP common features with the other influenza NPs and its own specificities.

D/NP forms tetramers in solution, and, in contrast with A/NP and B/NP, its oligomeric state could not be affected by modifying the salt concentration. The crystal structure of D/NP was solved at 2.4 Å resolution. A/NP, B/NP and D/NP structures are strikingly well-conserved, with a R.M.S.D. value between 1.6 and 2.1 Å. In contrast with A/NP and B/NP, which possess a flexible, NLS-bearing, intrinsically disordered N-terminal NP<sub>TAIL</sub>, D/NP shows such a disordered extension in C-terminal. D/NP<sub>TAIL</sub> is responsible for the binding to the importin- $\alpha$ 7 with about a 10-fold stronger affinity (100 nM), compared to A/NP. D/NP<sub>TAIL</sub> contains a clear bipartite nuclear localization signal with two basic patches (residues 514-516 and 531-535). The NP truncations of the C-terminal basic patch (residue 529-552), drastically impaired the *in vitro* importin- $\alpha$  binding as well as the *in cellulo* nuclear import while the deletion of both basic patches (residues 511-552), completely abolished the two processes.

This paper highlights some of the conserved features of influenza NPs, such as the overall conserved structure, the mechanism for the nuclear transport and the cellular

localization, along with more specific features, such as the C-terminal localization of its bipartite nuclear localization signal and its strong affinity for the importin- $\alpha 7$ .

---

**Article: The structure of the nucleoprotein of Influenza D shows that all Orthomyxoviridae nucleoproteins have a similar NPcore with or without a NPtail for nuclear transport**

Donchet A., Oliva J., Labaronne A., Tengo L., Miloudi M., Gerard F.C.A., Mas C., Schoehn G., Ruigrok R.W.H., Ducatez M. and Crépin T.  
Scientific Reports (2019)

---

# SCIENTIFIC REPORTS

OPEN

## The structure of the nucleoprotein of Influenza D shows that all *Orthomyxoviridae* nucleoproteins have a similar NP<sub>CORE</sub>, with or without a NP<sub>TAIL</sub> for nuclear transport

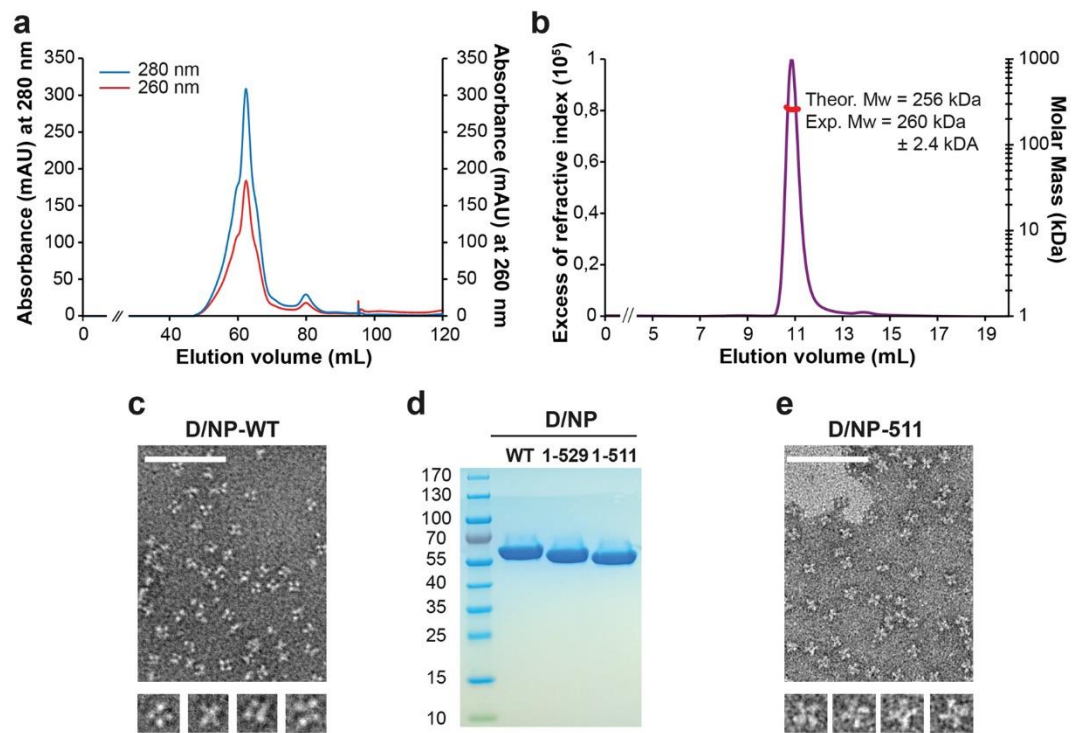
Amélie Donchet<sup>1</sup>, Justine Oliva<sup>2</sup>, Alice Labaronne<sup>1</sup>, Laura Tengo<sup>1</sup>, Myriam Miloudi<sup>1</sup>, Francine C.A. Gerard<sup>1</sup>, Caroline Mas<sup>3</sup>, Guy Schoehn<sup>1</sup>, Rob W.H. Ruigrok<sup>1</sup>, Mariette Ducatez<sup>2</sup> & Thibaut Crépin<sup>1</sup>

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 of infectious salmon anemia (ISA/NP) virus have been solved. Here we purified, characterized and solved the X-ray structure of the tetrameric D/NP at 2.4 Å resolution. The crystal structure of its core is 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<sub>TAIL</sub>). We show that D/NP<sub>TAIL</sub> harbors a bipartite NLS and designed C-terminal truncated mutants to demonstrate the role of D/NP<sub>TAIL</sub> for nuclear transport.

In 2011, a virus was isolated from a pig with Influenza-like symptoms in Oklahoma (USA). Electron microscopy showed features of an Influenza virus particle and real-time RT-PCR revealed that it was neither an Influenza A virus (IAV) nor an Influenza B virus (IBV). Next-generation sequencing analyses allowed the identification of RNA segments with around 50% identity to human Influenza C virus (ICV). Further serological analyses showed that antibodies against this new virus failed to cross-react with IAV, IBV or ICV (1). All these results suggested it was a new genus of the *Orthomyxoviridae*, temporarily named C/swine/Oklahoma/1334/2011 (C/swine/OK) and then Influenza D virus (IDV). IDV is widely distributed around the world; it was detected in North America<sup>1-4</sup>, Asia<sup>5-7</sup>, Europe<sup>8-10</sup> and Africa<sup>11</sup>. A serological study in Nebraska (USA) found a seroprevalence of 80% in cattle serum from 2003<sup>3</sup>.

The *Orthomyxoviridae* family includes different *genera*, Influenza A, B, C and D viruses, infectious salmon anemia (ISA) virus, Thogoto virus, Quaranjavirus and others. These viruses are segmented negative strand RNA viruses. Their genomes are made of a set of RNA segments coated with multiple copies of the nucleoprotein (NP) and associated to the viral heterotrimeric polymerase. 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 viral particle, 8 segments for IAV and IBV, 7 for IVC and IVD and 6 for Thogoto virus. These ribonucleoproteins (RNPs) are competent for both transcription and replication. 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<sup>12,13</sup> and one for the monomeric R416A mutant<sup>14</sup>. The X-ray structure of the tetrameric Influenza B nucleoprotein (B/NP) is also known<sup>15</sup>, whereas the structure of Isavirus NP (ISA/NP) was solved as a dimer<sup>16</sup>. The overall folds of A/NP and B/NP are very similar with a root-mean-square deviation (rmsd) of 1.6 Å.

<sup>1</sup>Institut de Biologie Structurale (IBS), Univ. Grenoble Alpes, CEA, CNRS, 38044, Grenoble, France. <sup>2</sup>IHAP, Université de Toulouse, INRA, ENVT, Toulouse, France. <sup>3</sup>Integrated Structural Biology Grenoble (ISBG) - UMS 3518 (CNRS-CEA-UJF-EMBL), 38044, Grenoble, France. Correspondence and requests for materials should be addressed to T.C. (email: [thibaut.crepin@ibs.fr](mailto:thibaut.crepin@ibs.fr))



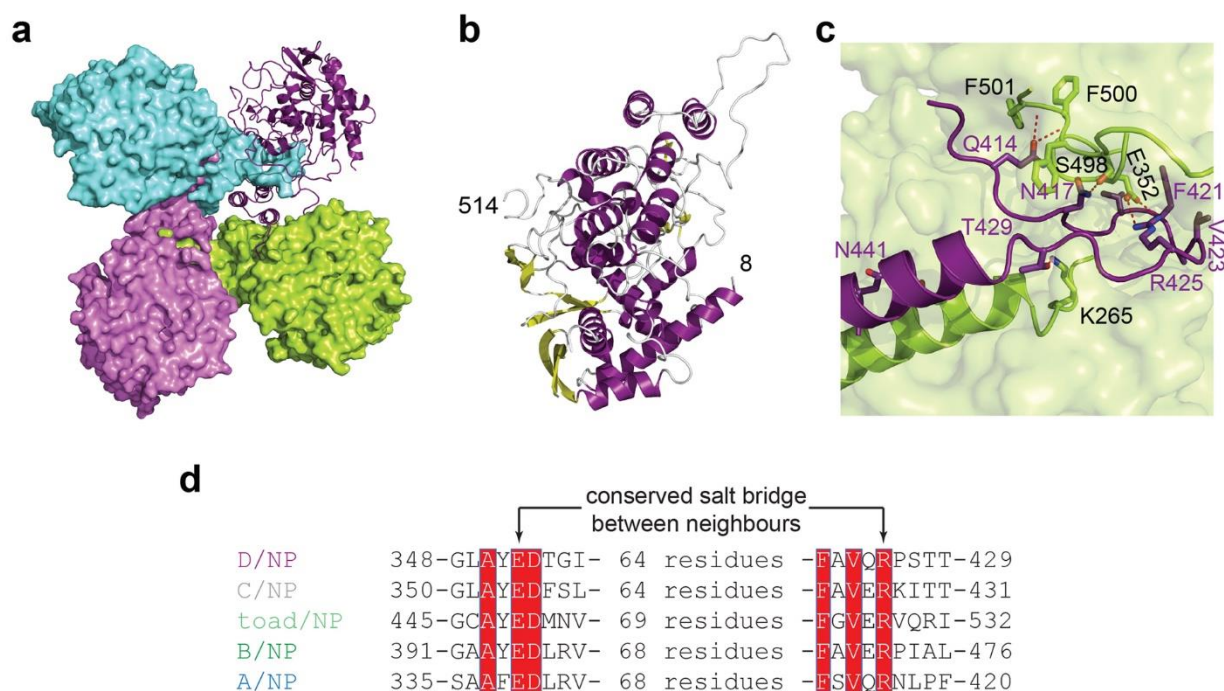
**Figure 1.** Purification and characterized of Influenza D nucleoprotein. **(a)** Size exclusion chromatography profile of wild-type D/NP. The sample was loaded on a Hiloal™ 16/600 S200 column equilibrated with the running buffer 20 mM Tris-HCl pH 7.5, 300 mM NaCl and 5 mM  $\beta$ -mercaptoethanol. **(b)** SEC-MALLS-RI analysis of D/NP. SEC was performed with a Superdex™ 200 increase 10/300 GL column equilibrated with 20 mM Tris-HCl pH 7.5, 150 mM NaCl and 5 mM  $\beta$ -ME. The panel shows the theoretical Mw and the measured Mw. **(c)** and **(e)** Electron microscopy images of the elution peak of D/NP and D/NP-511. Samples show different oligomeric states although most oligomers are tetramers. The scale bar corresponds to 100 nm. **(d)** 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).

Influenza viruses transcribe and replicate in the nucleus of the infected cells and the NPs and the polymerase subunits need 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- $\alpha$ , are located within the flexible N-terminal tail<sup>17–24</sup>. Recently, crystal structures of the two NLSs of A/NP bound to importin- $\alpha$  have been solved<sup>25,26</sup>.

In this paper, we characterized D/NP and solved the X-ray structure of its tetramer. The C-terminal D/NP<sub>TAIL</sub> harbouring a classical bipartite nuclear localization signal (NLS) was not visible in the structure. We designed two C-terminal truncated mutants (D/NP-511 and D/NP-529) to study the interaction of D/NP with importin- $\alpha$ . Our biochemical experiments demonstrate that D/NP<sub>TAIL</sub> is involved in the interaction with importins- $\alpha$  and immunofluorescence showed that the wild-type D/NP goes into the nucleus whereas the mutants stay in the cytoplasm.

## 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-tagged recombinant protein. D/NP was purified with a nickel affinity chromatography followed by a heparin column and a final gel filtration. Figure 1a shows a typical gel filtration elution profile using the absorbance signals at 260 and 280 nm (ratio 280/260 > 1.75). The protein could then be concentrated at 2 to 6 mg.mL<sup>-1</sup>. Polyacrylamide gels and SEC-MALLS-RI experiments have confirmed the homogeneity and the molecular weight of the recombinant tetrameric D/NP (Fig. 1b,d). By electron microscopy (negative staining), we showed that D/NP forms mainly tetramers in solution (Fig. 1c). Previously, it was shown that the oligomerization of recombinant A/NP and B/NP can be modulated by the NaCl concentration<sup>14,18,27,28</sup>. 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 dialysis at 150 mM NaCl, D/NP was eluted from gel filtration (with a 150 mM NaCl running buffer) in the same volume, meaning that a smooth reduction of the salt concentration did not change its oligomeric state. However, a decrease to 50 mM NaCl induced an irreversible and total precipitation of D/NP, even with a stepwise reduction at 150 mM NaCl. Therefore, the experiments on D/NP and its mutants were carried out at 300 mM NaCl.



**Figure 2.** Structure of Influenza D nucleoprotein. **(a)** Structure of the tetrameric D/NP, with 3 protomers shown in surface (respectively in green, pink and cyan) and the fourth in cartoon (deep purple). **(b)** Cartoon representation of one monomer of D/NP with the  $\alpha$ -helices in deep purple and  $\beta$ -strands in yellow. **(c)** Detail of the interactions between two protomers of the tetrameric D/NP as shown in **(a)**. The conserved R425 of one protomer (shown as deep purple cartoon) stabilized the position of the oligomerization loop at the surface of the neighbouring protomer, through its conserved E352. **(d)** Sequence alignment of the salt bridges of the oligomerization-loop of one protomer to the NP<sub>CORE</sub> of the neighbour protomers. For the sequences see Table 1.

	A/NP	B/NP	toad/NP	C/NP	D/NP	ISA/NP
A/NP		1.6 Å (383 C $\alpha$ )	—	—	2.1 Å (356 C $\alpha$ )	2.9 Å (340 C $\alpha$ )
B/NP	38%		—	—	1.9 Å (365 C $\alpha$ )	2.8 Å (340 C $\alpha$ )
toad/NP	25%	30%			—	—
C/NP	22%	25%	24%		—	—
D/NP	24%	25%	23%	38%		2.9 Å (330 C $\alpha$ )
ISA/NP	20%	18%	18%	16%	20%	
Tho/NP	18%	21%	20%	18%	17%	20%
WfB/NP	20%	19%	19%	20%	18%	17%

**Table 1.** Nucleoproteins of Orthomyxoviruses. A/NP, Influenza A virus nucleoprotein, strain A/WSN/1933(H1N1) (Uniprot accession number B4URF1, PDB entry 2IQH); B/NP, Influenza B virus nucleoprotein, strain B/Managua/4577.01/2008 (Uniprot accession number C4LQ26, PDB entry 3TJ0); Toad/NP, Wuhan asiatic toad Influenza virus nucleoprotein (GenBank accession number AVM87634); C/NP, Influenza C virus nucleoprotein, strain C/AnnArbor/1/1950 (Uniprot accession number Q6I7C0); D/NP, Influenza D virus nucleoprotein, strain D/bovine/France/2986/2012 (Uniprot accession number A0A0E3VZU8, PDB entry 5N2U); ISA/NP, infectious salmon anemia virus nucleoprotein, isolate salmon/Norway/810/9/99 (Uniprot accession number Q8V3T7, PDB entry 4EWC); Tho/NP, Thogoto virus nucleoprotein (Uniprot accession number A0A0B6VKB5); WfB/NP, Wellfleet Bay virus nucleoprotein (Uniprot accession number A0A0A1E9N5). Sequence identities were obtained using MUSCLE<sup>77</sup> and structure comparisons were calculated using PDBeFold<sup>78</sup>.

Like the nucleoproteins of A, B and ISA, D/NP bound single-stranded RNA. We measured the  $K_d$  of 14 nM with a fluorescence anisotropy assay with an RNA of 24 nucleotides of polyUC labeled in 3' with 6-fluorescein amidite (FAM)<sup>27</sup>. The  $K_d$  of 14 nM can be compared with the  $K_d$ s of 7 nM of the trimer of A/NP, 31 nM of the tetramer of B/NP<sup>29</sup> and 24 nM for the dimer of ISA/NP<sup>16</sup>.

**Structure of D/NP.** Full length D/NP was crystallized in sodium malonate as small fine needles that diffract X-rays up to 2.4 Å resolution (Supplementary Table 1). The structure was solved by molecular replacement using a starting model of the monomer of the A/NP R416A mutant<sup>14</sup>. The automatic search gave the position for three



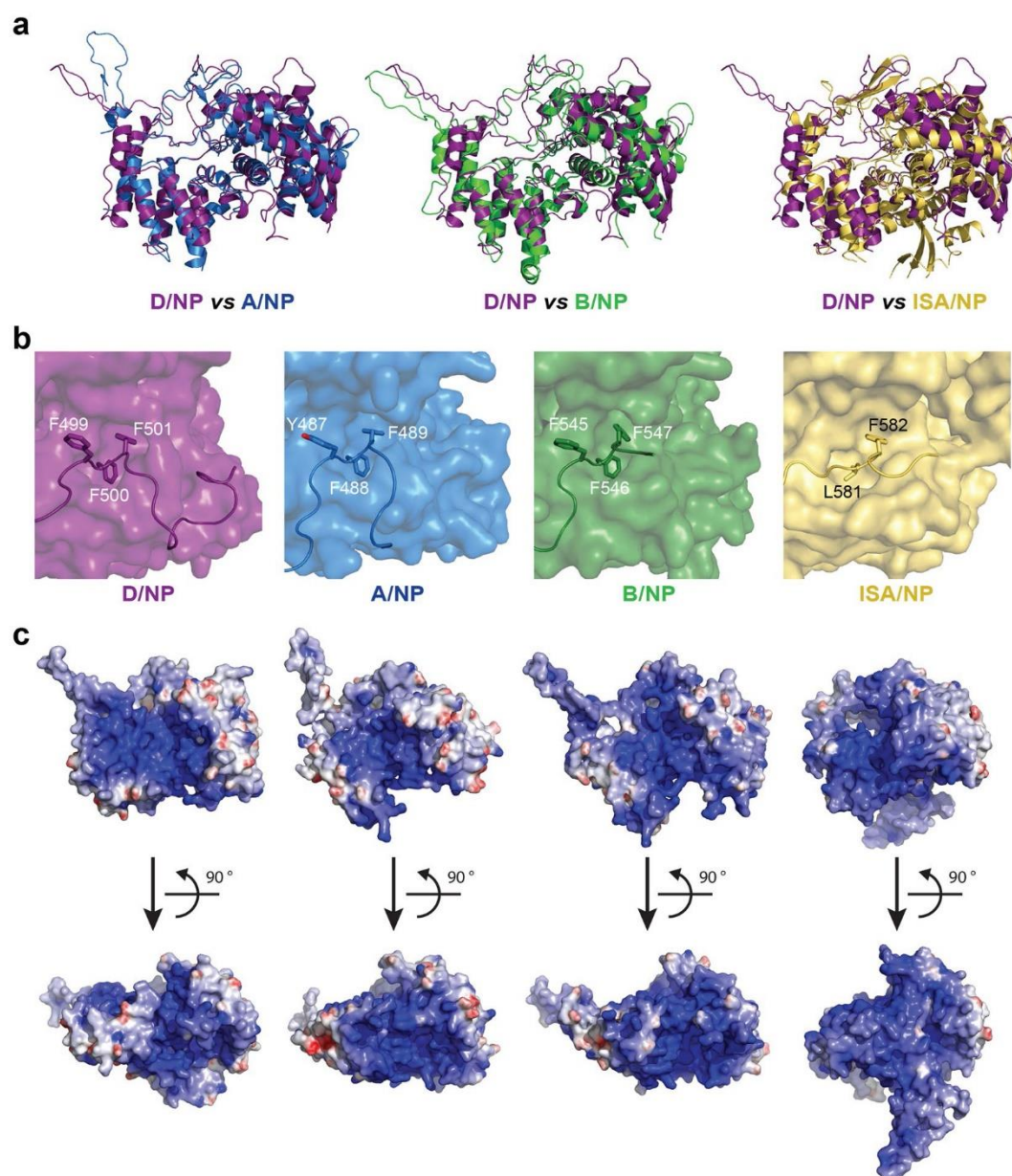
molecules and a fourth protomer was fitted manually after the analysis of the electron density. The molecular replacement using the tetrameric B/NP model did not give any acceptable solutions. The asymmetric unit contained four molecules arranged as a tetramer (Fig. 2a and refinement statistics in Supplementary Table 1), consistent with the observations made by electron microscopy. Based on the electron density, 87% of the model could be built, from residues 8 to 514 (Fig. 2b). Several internal loops were missing and the C-terminal 50 residues were disordered. The oligomerization-loop of one protomer plugged into a cavity of its neighbour (Fig. 2c). The N-terminal part of the loop started with a strictly conserved glutamate residue (Gln-414) that interacted with the backbone of the strictly conserved consecutive aromatic residues  $_{499}\text{FFF}_{501}$ . The loop was then stabilized by several kind of contacts; mainly hydrophobic (by the side chain of the strictly conserved Phe-421 and Val-423) and of several aromatic residues lying in the pocket of the neighbouring monomer but also the salt bridge between an arginine from one monomer (Arg-425) interacting with a glutamate of the neighbour (Glu-352; Fig. 2c,d). These two residues, also strictly conserved in all Influenza NPs (Fig. 2d and Supplementary Fig. 1), are well known for modulating their oligomeric state. Once mutated in alanine, the corresponding single mutants formed monomers, unable to self-assemble<sup>13,14,28,30–33</sup>. Phosphorylation state of A/NP has been shown to also be a key factor of NP/NP interfaces, especially for positions Ser-165 and Ser-407<sup>34,35</sup>. Sequence alignment showed that D/NP contains a threonine (Thr-161) and a serine (Ser-416) respectively, suggesting a conserved mechanism for the regulation of the oligomerization (Supplementary Fig. 1).

**Core structure of all nucleoproteins (NP<sub>CORE</sub>) with tails (NP<sub>TAIL</sub>).** The overall folds of A/NP, B/NP, ISA/NP and D/NP were very similar (Fig. 3a) with root-mean-square deviations (rmsd) between 1.6 Å (for 383 C $\alpha$  for the comparison A/NP with B/NP) and 2.9 Å (for 340 C $\alpha$  for the comparison D/NP with ISA/NP, Table 1). Including ISA/NP, a common architecture for all Influenza-like NPs can be defined based on the X-ray structures with the NP<sub>CORE</sub> starting with the first  $\alpha$ -helix for A/NP, B/NP and D/NP and finishing with three hydrophobic residues, anchored into the surface of the protein (Supplementary Fig. 1). Figure 3b shows the X-ray structures of A/NP, B/NP, D/NP and ISA/NP with a zoom on the C-terminal regions, where a superimposed patch made by three consecutive aromatic residues is found in Influenza NPs ( $_{487}\text{YFF}_{489}$  for A/NP,  $_{545}\text{FFF}_{547}$  for B/NP and  $_{499}\text{FFF}_{501}$  for D/NP). A similar patch is also present in the ISA/NP model ( $_{580}\text{GLF}_{582}$ ). The NP<sub>CORE</sub> contains the large and shallow positively charged surface which might bind RNA (Fig. 3c), without sequence specificity (see above)<sup>12,13,36,37</sup>. Twenty or 71 residues without any structure are found N-terminally before the cores of A/ and B/NP whereas for D/NP, the core starts with a very short version of an N-terminal tail (only 7 residues), which seems to be compensated by the presence of a carboxy-terminal NP<sub>TAIL</sub> of 51 residues (Supplementary Fig. 1). The length of the NP<sub>TAILs</sub> is variable and its location seems to be specific for each genus. Because D/NP differs from A/ and B/NP in its global organization, we decided to further analyse the C-terminal D/NP<sub>TAIL</sub>.

**D/NP<sub>TAIL</sub> behaves as an intrinsically disordered protein.** Based on the crystal structure and a disorder prediction (Fig. 4a), we designed a construct (from residues 505 to 552) for the expression in *E. coli* as an N-terminal His-tagged recombinant version. Considering 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 (Fig. 4a,b). D/NP<sub>TAIL</sub> was eluted from gel filtration as a 40 kDa-protein (data not shown) 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<sup>38,39</sup>, a SEC-MALLS-RI experiment (Fig. 4c) confirmed the monodispersity of the purified sample with a Mw of 8.2 kDa for D/NP<sub>TAIL</sub> (with the His-tag). A circular dichroism analysis confirmed that D/NP<sub>TAIL</sub> did not contain any significant stable secondary structure in solution (Fig. 4d), even in presence of TMAO (Supplementary Fig. 2), a chemical known to force unfolded proteins to fold to native-like species with significant functional activity<sup>40–42</sup>.

**D/NP<sub>TAIL</sub> interacts with importin- $\alpha$ .** The N-terminal tails of A/NP and B/NP are known to be involved in the nuclear import by interacting with importins- $\alpha$ <sup>17–19,21,23,25,43</sup>. The sequence analysis of D/NP showed no import signals in its N-terminal part but suggested the presence of a bipartite NLS within D/NP<sub>TAIL</sub> ( $_{514}\text{KRR-X}_{14}\text{-KKRGR}_{535}$ ; Fig. 4a). We thus tested the bipartite NLS with different constructs. First, we showed that the human importin- $\alpha 7$  is co-eluted with the His-tagged D/NP<sub>TAIL</sub> from gel filtration (Fig. 5a). The interaction was confirmed by thermal-shift experiments, where importin- $\alpha 7$  appeared more stable in presence of D/NP<sub>TAIL</sub> (Fig. 5b). We also confirmed the interaction between the two partners using the full-length D/NP. Using surface plasmon resonance, we measured a *K<sub>d</sub>* of 100 nM between immobilized D/NP<sub>TAIL</sub> and importin- $\alpha 7$  (Fig. 5c), ten-fold higher than the affinities measured for the N-terminally A/ and B/NP<sub>TAILs</sub><sup>18,25</sup>. By gel filtration, we showed that the D/NP:importin- $\alpha 7$  complex could be eluted as a single peak. A molecular weight of 474 kDa has been measured by SEC-MALLS-RI for the complex (Fig. 5d), corresponding to the interaction of four molecules of importins- $\alpha 7$  (Mw = 4 × 55 kDa) per tetramer of D/NP (Mw = 260 kDa). To confirm the role of D/NP<sub>TAIL</sub> in the interaction with importin- $\alpha$ , we 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 D/NP-529 contains only the first part (Fig. 4a). They both behave as the wild-type protein during the purification (Fig. 1d,e). Using pull-down experiments on nickel-affinity resin, we showed that D/NP-529 still slightly retained importin- $\alpha 7$  whereas the interaction between D/NP-511 and importin- $\alpha 7$  was totally abolished (Fig. 5e). This suggests that importin- $\alpha 7$  binds to the NP<sub>TAIL</sub> of D/NP.

**Nuclear transport of D/NP and its mutants.** First, human HEK 293T cells were infected with D/bovine/France/5920/2014 (*moi* of 5) and indirect immunofluorescence was used to localize D/NP within the cells. After 6 h, most of D/NP was observed in the nucleus (Fig. 6a). As a control, HEK 293T cells were infected with Influenza A/PR/8/34 and we found very similar results (data not shown). We then transfected HEK 293T

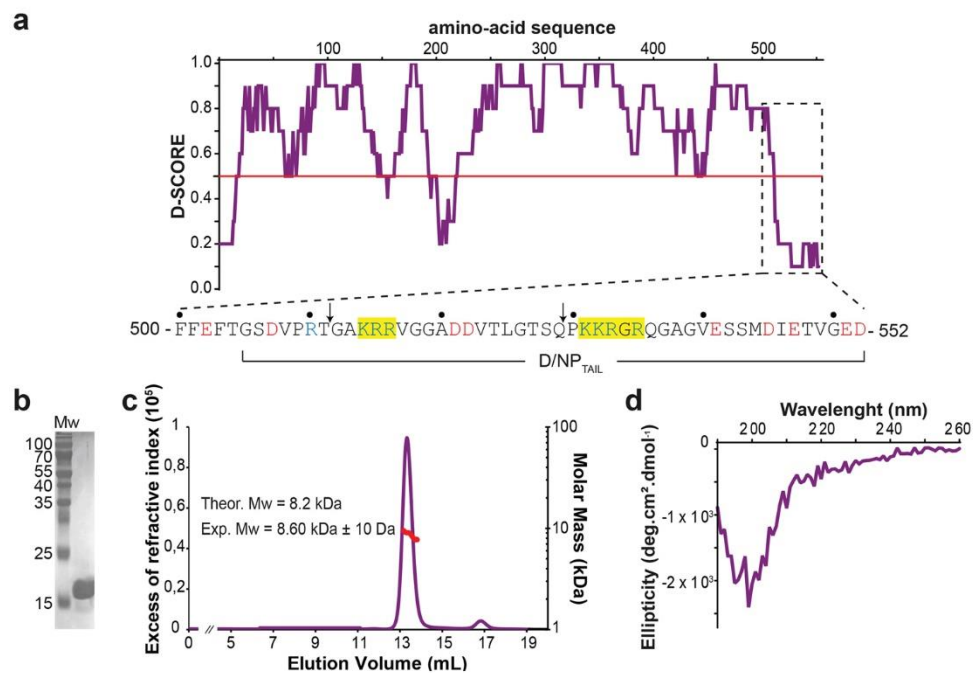


**Figure 3.** Comparison of D/NP with other segmented negative sRNA virus nucleoproteins. **(a)** The structure of one protomer of D/NP (deep purple) has been superimposed from left to right, with one protomer of A/NP (blue; PDB id: 2IQH), B/NP (forest; PDB id: 3TJ0) and ISA/NP (light orange; PDB id: 4EWC). The rmsd values are given in Table 1. **(b)** Anchoring of the C-terminus on NP<sub>CORE</sub> by a patch of 3 consecutive aromatic residues. The panel corresponds to a zoom of the superimposed structures shown in panel A with NP<sub>CORE</sub> represented in surface. **(c)** Electrostatic surface potentials of one protomer of D/NP, A/NP, B/NP and ISA/NP. The electrostatic surfaces were calculated from the crystal structures using DelPhi<sup>75</sup>. The potential scales range from  $-10.0$  kT/e (red) to  $10.0$  kT/e (blue).

cells with a plasmid containing wild-type D/NP and we found that D/NP can be localized in the cytoplasm but mainly in the nucleus (Fig. 6b). With this transfection strategy, D/NP-529 was located mainly in the cytoplasm with a little staining observed in the nucleus, whereas, D/NP-511 was located exclusively in the cytoplasm of the transfected cells (Fig. 6b), clearly showing that the cell used the bipartite NLS between residues 513 and 535 for the nuclear import of D/NP.

### Discussion

In 2011, a new virus has been isolated from pigs exhibiting Influenza-like illness. Subsequent studies identified an *Orthomyxovirus* with seven RNA segments sharing 50% overall amino acids identity with the human Influenza C virus<sup>2</sup>. 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

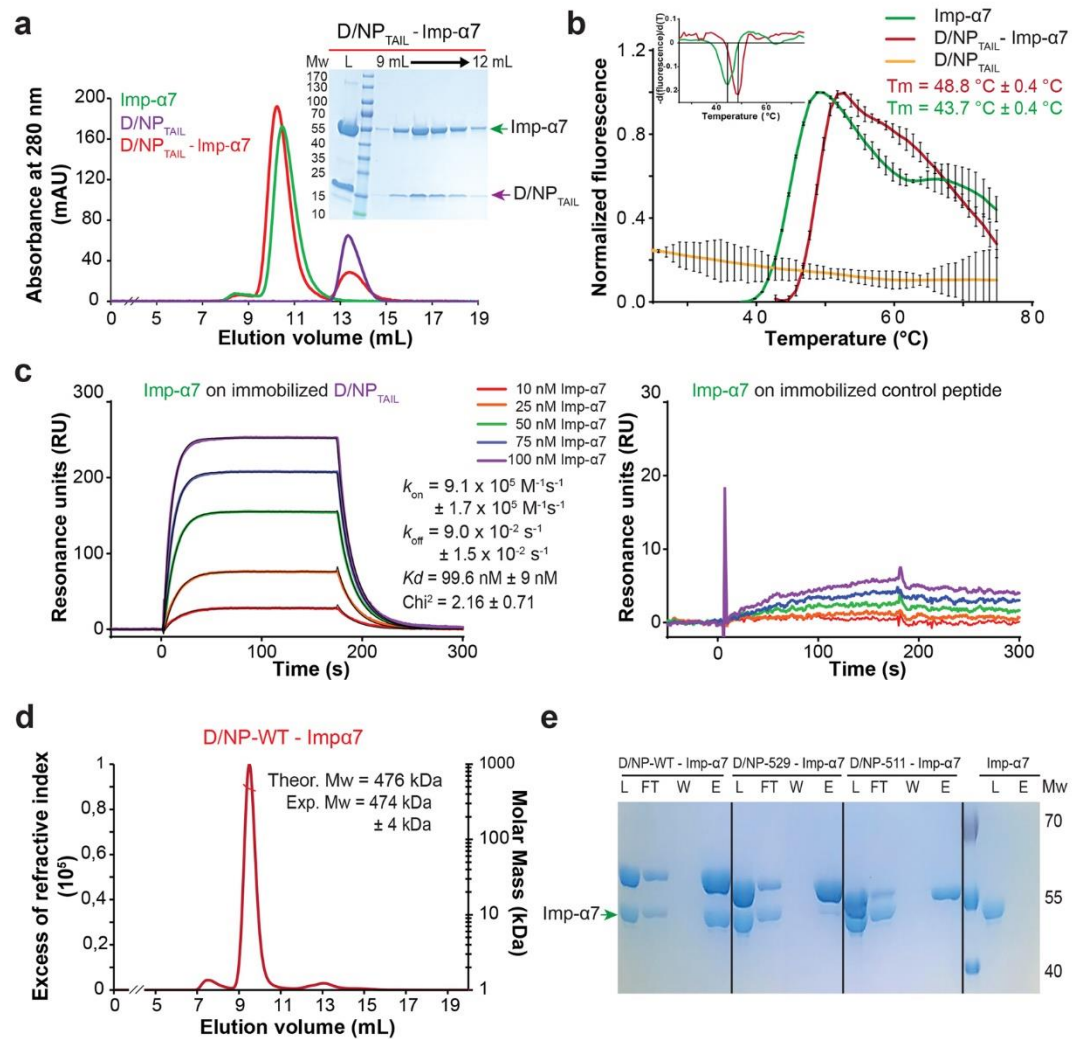


**Figure 4.** Biophysical characterization of D/NP<sub>TAIL</sub>. **(a)** D-score (score for disorder as a function of residue) of D/NP with a zoom (below the graph) on the last 50 residues. The prediction is based on 22 predictor web servers and the D-score was calculated by adding the values for each residue and dividing by the number of used algorithms. We arbitrarily defined a threshold level at 0.50; residues with a D-score < 0.50 were assigned as disordered<sup>38</sup>. The yellow boxes on the sequence are to highlight the putative NLS motifs. The arrows indicate the sequences were cut for making D/NP-529 and D/NP-511. **(b)** Coomassie blue-stained SDS-PAGE (Tris-Tricine; 15% polyacrylamide) of the purified D/NP<sub>TAIL</sub>. It migrates at a higher molecular weight (17 kDa approximately) than expected (8 kDa). **(c)** SEC-MALLS-RI analysis of D/NP<sub>TAIL</sub> loaded on S75 10/300 GL column. For this experiment, we have chosen to keep the His-tag encoded with the pETM11 plasmid, for an optimal detection of D/NP<sub>TAIL</sub> with UV. The experimental molecular weight is consistent with the expected mass. **(d)** Circular dichroism of D/NP<sub>TAIL</sub>. CD is a biophysical method based on the polarization of light, used for a fast determination of the secondary structures within the proteins in solution.  $\alpha$ -Helical proteins show negative bands at 222 nm and 208 nm and a positive band at 193 nm, proteins with well-defined antiparallel  $\beta$ -sheets have negative bands at 218 nm and positive bands at 195 nm and disordered proteins have very low ellipticity above 210 nm and negative bands near 195 nm.

circulate all around the world in many mammals, with cattle as a possible reservoir, and with serological data suggesting it may transmit to humans<sup>44</sup>. The recent structure of IDV Hemagglutinin-Esterase-Fusion glycoprotein (HEF) has described an open receptor-binding cavity capable of accommodating diverse extended glycan moieties that could be one reason for its broad cell tropism<sup>45</sup>.

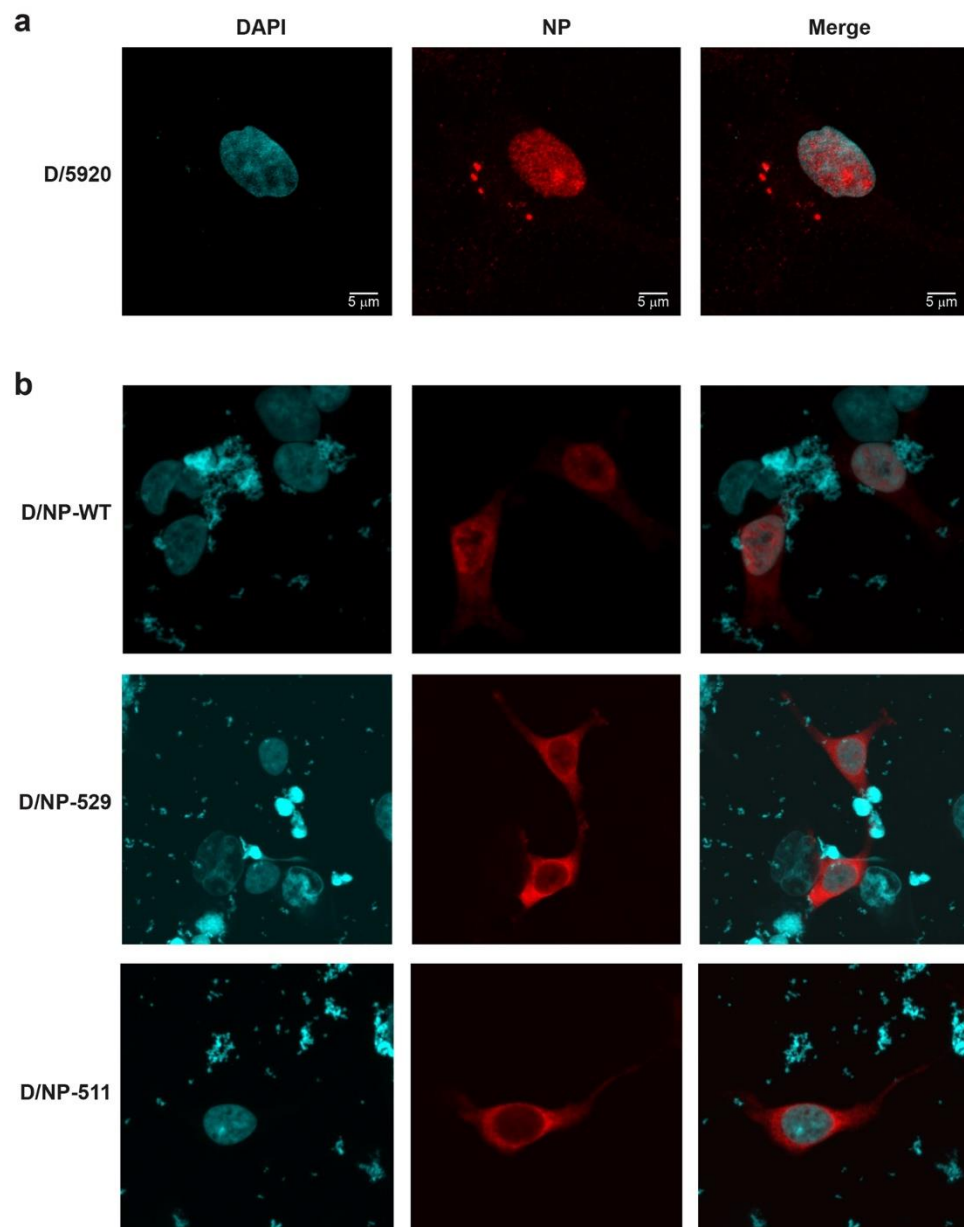
With this paper, we provide a detailed characterization of IDV nucleoprotein. We determined the structure of the tetramer of D/NP and we failed to find the biochemical conditions to stabilize the monomer. For A and B/NP we had found that the monomer stays in solution at low salt, 50 mM NaCl<sup>14,27,28</sup> but the D/NP protein makes aggregates in salt conditions lower than 150 mM. Post-translation changes like phosphorylation or ubiquitination may be able to stabilise the monomer, as it has been shown for A/NP<sup>14,34,46,47</sup>. The phosphorylation on A and B/NP are summarized in Supplementary Fig. 1.

The comparison of the structures of the nucleoproteins of Influenza A, B, D and of infectious salmon anemia virus (ISAV) shows that of all these proteins share a common structural core (NP<sub>CORE</sub>) that start with the first  $\alpha$ -helix in the structure of A, B and D, up to three consecutive aromatic residues (Fig. 3 and Supplementary Fig. 1). Looking at the sequences of all known *Orthomyxoviruses* nucleoproteins shows that such hydrophobic patch is present in C/NP (501 FFF<sub>503</sub>) and could be present in Tho/NP (451 YLF<sub>453</sub>) and Wfb/NP (Wellfleet bay virus nucleoprotein; 532 VIY<sub>534</sub>). The only structure that differs is the ISAV/NP that has a folded domain upstream of the first  $\alpha$ -helix of the NP<sub>CORE</sub> (Fig. 7). One could also see that Tho/NP is constituted only by the NP<sub>CORE</sub> without any other appended sequences. NPs of IAV and IBV have an N-terminal NP<sub>TAIL</sub> whereas the NPs of ICV and IDV have a C-terminal NP<sub>TAIL</sub>. The NP<sub>TAIL</sub> of A is only 20 residues whereas the tails of the other Influenza viruses NPs are much longer; 71 residues for IBV, 62 for ICV, 51 for IDV. Recently, new large-scale methods for finding RNA viruses in vertebrates other than mammals and birds, identified up to 240 “new” viruses<sup>48</sup>. These authors found new Influenza viruses close to IBV, a Wuhan spiny eel Influenza virus and a Wuhan Asiatic toad Influenza virus. The NP for the toad virus (toad/NP) is very close to B/NP (identity of 30%, Table 1) with a very long N-terminal tail of 126 residues (Fig. 7 and Supplementary Fig. 1). Interestingly, toad/NP possess an alanine residue (Ala-278) aligned with A/NP Ser-165 and D/NP Thr-161 and a phenylalanine (Phe-519) aligned with A/NP Ser-407 and D/NP Ser-416.



**Figure 5.** Interaction of D/NP and D/NP<sub>TAIL</sub> with importin- $\alpha$ 7. **(a)** Size exclusion chromatography profile of a mixture between human importin- $\alpha$ 7 and D/NP<sub>TAIL</sub>. The mixture (molar ratio 1 importin- $\alpha$ 7 for 2 D/NP<sub>TAIL</sub>) was incubated 1 hour at room temperature and then loaded on a Superdex<sup>TM</sup> 75 10/300GL column equilibrated with the running buffer 20 mM Tris-HCl pH 7.5, 250 mM NaCl, 5 mM  $\beta$ -mercaptoethanol. **(b)** Thermal stability assay of importin- $\alpha$ 7 in absence (green) or in presence (red) of D/NP<sub>TAIL</sub> using ThermoFluor<sup>76</sup>. In presence of D/NP<sub>TAIL</sub>, the melting temperature of importin- $\alpha$ 7 is 5 °C higher. D/NP<sub>TAIL</sub> alone using ThermoFluor did not give any denaturation signal (yellow curve). The upper insert corresponds to the derivative of the fluorescence signal for a precise measure of the melting temperature. **(c)** Affinity of importin- $\alpha$ 7 for D/NP<sub>TAIL</sub> by measured by surface plasmon resonance (SPR). Biotinylated D/NP<sub>TAIL</sub> (left) and control peptide (right) were captured on a streptavidin-coated sensor chip surface before injections of several importin- $\alpha$ 7 concentrations (10 nM in red, 25 nM in orange, 50 nM in green, 75 nM in blue and 100 nM in purple). The sensorgrams of the interaction between D/NP<sub>TAIL</sub> and importin- $\alpha$ 7 were fitted under a Langmuir 1:1 binding model with mass-transfer (black line). **(d)** SEC-MALLS analysis of D/NP in complex with importin- $\alpha$ 7. The mixture (molar ratio 1 D/NP for 1.2 importin- $\alpha$ 7) was incubated 1 hour at room temperature and then loaded on a Superdex<sup>TM</sup> 200 increase 10/300 GL. The experimental molecular weight is consistent with the expected mass of four importin- $\alpha$ 7 bound per D/NP tetramer. **(e)** Pull-down assays of human importin- $\alpha$ 7 by D/NP and the two C-terminal truncated mutants (D/NP-529 and D/NP-511). The his-tags are on D/NP. The mixtures (molar ratio 1 D/NP for 1.2 importin- $\alpha$ 7) were incubated 1 hour and the experiment was done as described in panel (a). The figure shows the coomassie blue-stained SDS-PAGE (12% polyacrylamide) with the Load, FlowThrough, Wash and the second fractions (E2).

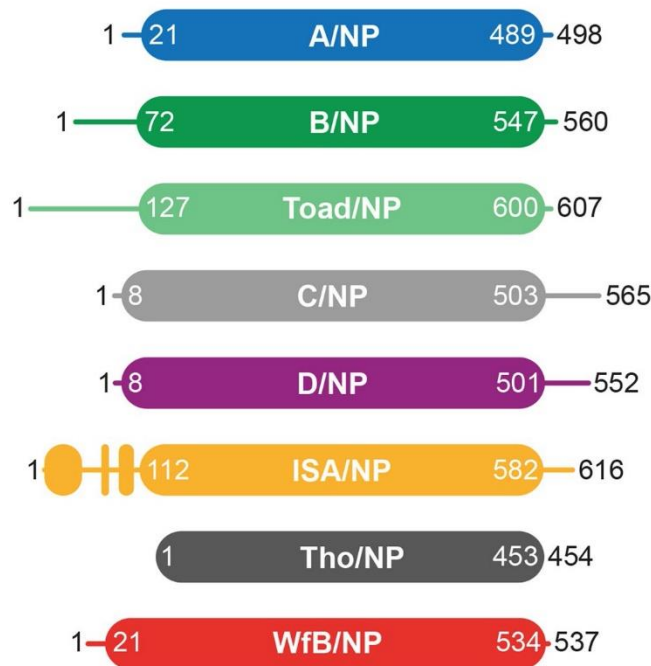
We demonstrated that the C-terminal D/NP<sub>TAIL</sub> presents all the characteristics of an intrinsically disordered protein and that is involved in the nuclear import through its interaction with cellular importins- $\alpha$ . This role is not a big surprise since the presence of such NLS motifs was identified within the flexible NP<sub>TAILS</sub> of viruses A and B located at the N-terminus of these proteins<sup>17,18,21,23,25,43,49</sup>. Tho/NP has been shown to be imported in the nucleus<sup>50</sup> but this nucleoprotein does not possess any additional intrinsically disordered protein-like peptide with an NLS. Recently, a second NLS on A/NP (strain A/X-31) at the rim of the RNA binding channel was shown. A crystal structure shows that this internal NLS (<sub>198</sub>Lys-Arg-INDRNFWRGENC-Arg-Arg-T-Arg<sub>216</sub>) binds to mouse



**Figure 6.** Nuclear transport of D/NP in HEK 293T cells. Microscopy pictures of HEK 293T cells with DAPI, NP and merged DAPI + NP staining (left, middle and right panels, respectively). **(a)** D/NP after infection of HEK 293T cells at 6 hours post infection with a *moi* of 5. **(b)** Cells were transfected with wt D/NP, D/NP-529 and D/NP-511. After 24 h cells were fixed, permeabilized and analyzed by indirect immunofluorescence. NP localization was observed using an in house rabbit hyperimmune NP-IDV serum and anti-rabbit IgG labeled Rhodamine RX (in red). Cells were mounted with DAPI-Vectashield and observed with a Leica Zeiss 710 (magnification: x63). In red: NP protein; in blue: nucleus. The bar in (A) represents 5 μm.

importin- $\alpha$ <sup>126</sup>. Likewise, for the NLSs of Tho/NP and ISA/NP, it has been suggested that the rims on the RNA binding surfaces can be used as an NLS for nuclear import for viral replication<sup>16,50,51</sup>.

Finally, we have shown that the NP<sub>TAILs</sub> of these nucleoproteins are intrinsically disordered proteins. The preservation of such intrinsically disordered protein-like peptide with a conserved function highlights its importance. The name for N<sub>TAIL</sub> was first used for the C-terminal disordered tail of measles virus nucleoproteins<sup>52</sup>, which is about 125 residues long. This tail is disordered but can adopt a short  $\alpha$ -helical structure (residue D484-A502) to promote the interaction with the phosphoproteins of these viruses which bind themselves to the corresponding RNA polymerase<sup>53,54</sup>. For these viruses, the tails keep the nucleocapsids in a flexible conformation. When cleaved off with trypsin, the helical nucleocapsids become rigid<sup>55,56</sup>. The tails do not change the affinity of the nucleoprotein for RNA but change the kinetics for the assembly of the nucleocapsids<sup>57</sup>. The RNPs of Influenza A and B viruses are also very flexible<sup>58-61</sup> which makes it difficult to handle for high resolution structures. The EM structure of the measles nucleocapsid with a resolution of 3.6 Å<sup>62</sup> was performed on the nucleocapsids without



**Figure 7.** Schemas for nucleoproteins of *Orthomyxoviruses*. Schematic representation of the nucleoprotein based on the amino acid sequence identity and structure analysis of the protein from representative members of the *Orthomyxoviridae* family. The schema respects the size of the proteins. The protein accession numbers are the same as Table 1. The flexible tails are represented with simple lines whereas the folded parts and the cores are represented with filled boxes.

the C-terminal  $N_{TAIL}$ . If we could generate Influenza RNPs with NP without  $NP_{TAIL}$  to rigidify the RNPs, it would make it easier solving a high resolution structure to observe the RNA on the nucleoproteins.

## Methods

**Molecular biology for expression of nucleoprotein of Influenza D.** 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 were cloned in pETM13 (EMBL) to express C-terminal His-tagged proteins whereas D/ $NP_{TAIL}$  (505-D/NP-552) DNA coding sequence was cloned in pETM11 (EMBL) to express an N-terminal His-tagged fragment. The DNA coding sequence of the human importin- $\alpha 7$  (KPNA6; Uniprot accession number O60684) was cloned without its IBB domain (amino-acids 58 to 536) in pET9a (Novagen) to express an N-terminal His-tagged protein<sup>63</sup>. Sequencing was performed by Eurofins.

**Expression and purification of proteins.** *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- $\beta$ -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  $\beta$ -mercaptoethanol ( $\beta$ -ME) and complete protease inhibitor cocktail (Roche) for D/NP constructs and 50 mM Tris-HCl pH 8, 500 mM NaCl, 1 mM  $\beta$ -ME and complete protease inhibitor cocktail for importin- $\alpha 7$ . 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<sup>TM</sup> 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  $\beta$ -ME. The last purification step was a size-exclusion chromatography using a Hiloal<sup>TM</sup> 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- $\alpha 7$  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 NaCl, 5 mM  $\beta$ -ME and 20 mM imidazol. Finally, a size-exclusion chromatography using Superdex<sup>TM</sup> 200 increase 10/300 GL column (GE-Healthcare) was performed in 20 mM Tris-HCl pH 7.5 at 150 mM NaCl, 5 mM  $\beta$ -ME. Peak fractions were concentrated using a 10 kDa Amicon concentrator. Protein concentrations were determined using the extinction coefficients at 280 nm,  $\epsilon = 44\,537\text{ M}^{-1}\cdot\text{cm}^{-1}$  for D/NP,  $\epsilon = 41\,558\text{ M}^{-1}\cdot\text{cm}^{-1}$  for D/NP-529 and D/NP-511,  $\epsilon = 2980\text{ M}^{-1}\cdot\text{cm}^{-1}$  for D/ $NP_{TAIL}$ , and  $\epsilon = 46\,785\text{ M}^{-1}\cdot\text{cm}^{-1}$  for importin- $\alpha 7$ .

**SEC-MALLS-RI analysis.** Multi-angle laser light scattering (MALLS) coupled with size exclusion chromatography (SEC) and refractometry (RI) is a method for measuring the absolute molecular mass of a particle in solution that is independent of its dimensions and shape<sup>64</sup>. SEC was performed with a column (Superdex<sup>TM</sup> 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  $\beta$ -ME. Analytical runs were performed at 20 °C with a flow rate of 0.5 mL·min<sup>-1</sup>. 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<sup>-1</sup>. Weight-average molar masses (Mw) were calculated with ASTRA (Wyatt Technology) as previously described<sup>38</sup>.

**Electron microscopy.** Samples (concentrations around 0.05 mg·mL<sup>-1</sup>) 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<sup>65</sup>. Charge-coupled Device (CCD) frames were taken with a FEI T12 microscope operating at 120 kV and a nominal magnification of 45 000 times. The dilutions for EM were performed with the size-exclusion buffer right before preparing the grid.

**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 of 1.5 mg·mL<sup>-1</sup>. The crystals were directly flash-frozen without any cryoprotectant. Data were collected at the ESRF (beam-line ID30A-3) and processed with the XDS package<sup>66</sup>. 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)<sup>67-71</sup>. The final refinement was done using BUSTER<sup>72</sup>. The coordinates have been deposited in the Protein Data Bank under PDB ID code 5N2U. The protein structure figures were drawn using PyMOL<sup>73</sup>.

**Thermal shift assays.** Thermal shift assays were performed following the established protocols of thermofluor experiments<sup>74</sup> using a fluorescent probe (SYPRO Orange). SYPRO Orange dye binds non-specifically to hydrophobic surfaces and water strongly quenches its fluorescence. When a protein unfolds with the increase of the temperature, the exposed hydrophobic surfaces bind the dye, resulting in an increase of the fluorescence by excluding water. Samples were diluted at 0.25 mg·mL<sup>-1</sup> in 50 mM Tris-HCl pH 7.5, 300 mM NaCl and 5 mM  $\beta$ -ME and 5X SYPRO Orange dye (Invitrogen) in a final volume of 40  $\mu$ L. 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 75 °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 derivative curves.

**Circular dichroism.** Circular dichroism (CD) spectroscopy is an efficient tool for rapid determination of the secondary structure and folding properties of a protein by measuring the absorption difference between left and right circularly polarized light. 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<sub>TAIL</sub> at 6.4  $\mu$ M in presence or in absence of trimethylamine N-oxide (TMAO). TMAO is a stabilizing chemical agent known to promote protein folding. It can be used with CD spectroscopy to assess putative folding of unfolded peptides and proteins. In addition to D/NP<sub>TAIL</sub> without TMAO, final concentrations of 0.5, 1, 2 and 3 M TMAO were used in phosphate buffer (10 mM, pH 7.2) to validate the unfolded nature of D/NP<sub>TAIL</sub>. Due to the high absorbance of TMAO around 200 nm, spectra were recorded up to the point before the HT rose above 700 V (190 nm without TMAO, 205 nm with 0.5 M TMAO, 207 nm with 1 M TMAO, 210 nm with 2 M TMAO and 212 nm with 3 M TMAO). Spectra were collected at 20 °C with fifteen runs using a cuvette with a path-length of 1 mm. After blank subtraction, the CD signal (in mdeg) was converted to mean molar residue ellipticity (in deg·cm<sup>2</sup>·dmol<sup>-1</sup>).

**Surface plasmon resonance (SPR) measurements and analysis.** The D/NP<sub>TAIL</sub> sequence was fused to the sequence encoding a biotinylation motif (NGSGGGLNDIFEA-QKIEWHE) and cloned in pETM11 (EMBL). *E. coli* BL21 RIL (DE3) were transformed and grown up to an OD of 0.6–0.8 in a medium supplemented with biotin (12.5  $\mu$ g·mL<sup>-1</sup>) in order to biotinylate the peptide *in vivo* during the expression (12 h, 18 °C, 0.3 mM IPTG). Purification was performed at room temperature in two steps, a Nickel affinity chromatography, followed by TEV cleavage and size exclusion chromatography (S75 10/300 GL, GE-Healthcare). A control biotinylated peptide (LEEMK-KGHLERECMEETCSYEEAREVFEDSEKTFNEFWNK-biotin) was also used as negative control.

SPR experiments were carried out on a Biacore 3000 (GE Healthcare). Streptavidin was first immobilized on a CM5 sensor chip (GE Healthcare) surface through amine coupling at a flow rate of 5  $\mu$ L·min<sup>-1</sup>. The biotinylated NP<sub>TAIL</sub> (10  $\mu$ g·mL<sup>-1</sup>) or the control biotinylated peptide (10  $\mu$ g·mL<sup>-1</sup>) were diluted in surfactant-supplemented HBS-N running buffer (10 mM HEPES pH 7.4, 150 mM NaCl, 0.05% Tween 20) (GE Healthcare) and the attachment was carried out through biotin:streptavidin interaction, at a flow rate of 5  $\mu$ L·min<sup>-1</sup> up to 400–500 RU. A flow-cell containing immobilized streptavidin only was used as reference while the flow-cell with the streptavidin:biotinylated NP<sub>TAIL</sub> complex was used as active flow-cell. For kinetic measurements, the importin- $\alpha$ 7 analyte was serially diluted in surfactant-supplemented HBS-N running buffer to final concentrations of 10, 25, 50, 75 and 100 nM. Each concentration was injected in triplicate in both the reference and active flow-cells. Analyte injection and following buffer injection times were set at 180 s and 150 s respectively at a flow-rate of 15  $\mu$ L·min<sup>-1</sup>. As the analyte spontaneously and completely dissociated, no regeneration was required and complete dissociation was achieved in buffer for 300 s. All sensorgrams were reference-subtracted and injection points were aligned to be processed with the analysis. Kinetic data were analysed using the Biacore 3000 Evaluation software (GE Healthcare) under a Langmuir 1:1 binding model with mass-transfer. Chi<sup>2</sup> values for the sensorgram fits used to

determine the kinetic parameters were kept below 3, due to the obtained high signal, and T-value for rate parameters above at least 30. The residuals for the fitting were kept between 1 and -1.

**Nuclear import of WT and NP mutants in HEK 293T cells.** *Plasmids.* The D/NP, D/NP-511 and D/NP-529 genes were amplified by PCR, cloned into the multiple cloning site of pSC-A-amp/kan vector (Strataclone blunt kit, Agilent) and expressed into competent bacterial cells, according to the manufacturer's instructions. The NP genes in pSC-A-amp/kan vector were cloned into KpnI and BamHI restriction sites of an eukaryotic expression vector, pCDNA3.1 (+), driven by a CMV promoter. The constructs sequences were confirmed by Sanger sequencing (GATC Biotech platform, Germany).

*Rabbit antibodies.* To generate rabbit hyperimmune serum anti-D/NP, three 7 to 8 weeks old New Zealand white rabbits (Pôle Experimental Cunicole de Toulouse, Castanet-Tolosan) were immunized at day 0 and day 35 with 100–200 µg D/NP in incomplete Freund adjuvant. Rabbits were anesthetized and bled two weeks post boost (50 mL blood collected intra-cardiac before humane euthanasia).

*Infection.* Human embryonic kidney cells (HEK 293T; ATCC) were infected with D/bovine/France/5920/2014 at a multiplicity of infection of 5. Briefly, cells were washed with PBS, the inoculum was added and the cells were incubated 1 hour at 37 °C + 5% CO<sub>2</sub>. After incubation, infection media (OPTI-MEM) was added and the cells were incubated until the fixation.

*Transfection.* HEK 293T cells were grown on glass coverslips in Dulbecco's modified Eagle's medium (DMEM) complemented with 10% fetal bovine serum (FBS) and penicillin-streptomycin (Dominique DUTSCHER SAS) and incubated at 37 °C + 5% CO<sub>2</sub> for 24 hours. Two hours before transfection, DMEM media was removed and replaced by OPTI-MEM media (Dominique DUTSCHER SAS). Cells were transfected with 2.5 µg of the respective plasmids using Mirus-TransIT reagent (Mirus) following the manufacturers' instructions. Cells were then incubated for 24 h at 37 °C + 5% CO<sub>2</sub>.

*Immunofluorescence.* Cells were pre-fixed using 4% paraformaldehyde (Bio-Rad) for 10 minutes at room temperature and then, fixed and permeabilized using ice cold ethanol:acetone for 10 minutes at room temperature. After a blocking step using PBS complemented with 5% horse serum, cells were incubated with a rabbit hyperimmune D/NP serum at a 1:100 dilution for 2 hours at room temperature. After three washing steps with PBS-triton 0.05%, cells were incubated with anti-rabbit IgG labeled Rhodamine RX (Jackson Immuno; Research, 711.296.152) at 1:200 dilution for 1 hour in the dark and at room temperature. Then, cells were washed with PBS-triton 0.05% and coverslips were mounted on glass slides using Vectashield-DAPI mounting medium (Vector Laboratories) and analyzed with a Leica Zeiss 710 at the Cell Imagery Platform in Purpan (Toulouse).

**Ethics Statement.** Experimentations were conducted in accordance with European and French legislations on Laboratory Animal Care and Use (French Decree 2001-464 and European Directive CEE86/609) and the animal protocol was approved by the Ethics Committee "Sciences et santé animale" number 115.

## References

1. Ferguson, L. *et al.* Influenza D virus infection in Mississippi beef cattle. *Virology* **486**, 28–34, <https://doi.org/10.1016/j.virol.2015.08.030> (2015).
2. Hause, B. M. *et al.* Isolation of a novel swine influenza virus from Oklahoma in 2011 which is distantly related to human influenza C viruses. *PLoS Pathog* **9**, e1003176, <https://doi.org/10.1371/journal.ppat.1003176> (2013).
3. Luo, J. *et al.* Serological evidence for high prevalence of Influenza D Viruses in Cattle, Nebraska, United States, 2003–2004. *Virology* **501**, 88–91, <https://doi.org/10.1016/j.virol.2016.11.004> (2017).
4. Mitra, N., Cernicchiaro, N., Torres, S., Li, F. & Hause, B. M. Metagenomic characterization of the virome associated with bovine respiratory disease in feedlot cattle identified novel viruses and suggests an etiologic role for influenza D virus. *J Gen Virol* **97**, 1771–1784, <https://doi.org/10.1099/jgv.0.000492> (2016).
5. Horimoto, T. *et al.* Nationwide Distribution of Bovine Influenza D Virus Infection in Japan. *PLoS One* **11**, e0163828, <https://doi.org/10.1371/journal.pone.0163828> (2016).
6. Murakami, S. *et al.* Influenza D Virus Infection in Herd of Cattle, Japan. *Emerg Infect Dis* **22**, 1517–1519, <https://doi.org/10.3201/eid2208.160362> (2016).
7. Zhai, S. L. *et al.* Influenza D Virus in Animal Species in Guangdong Province, Southern China. *Emerg Infect Dis* **23**, 1392–1396, <https://doi.org/10.3201/eid2308.170059> (2017).
8. Chiapponi, C. *et al.* Detection of Influenza D Virus among Swine and Cattle, Italy. *Emerg Infect Dis* **22**, 352–354, <https://doi.org/10.3201/eid2202.151439> (2016).
9. Ducatez, M. F., Pelletier, C. & Meyer, G. Influenza D virus in cattle, France, 2011–2014. *Emerg Infect Dis* **21**, 368–371, <https://doi.org/10.3201/eid2102.141449> (2015).
10. Foni, E. *et al.* Influenza D in Italy: towards a better understanding of an emerging viral infection in swine. *Sci Rep* **7**, 11660, <https://doi.org/10.1038/s41598-017-12012-3> (2017).
11. Salem, E. *et al.* Serologic Evidence for Influenza C and D Virus among Ruminants and Camelids, Africa, 1991–2015. *Emerg Infect Dis* **23**, 1556–1559, <https://doi.org/10.3201/eid2309.170342> (2017).
12. Ng, A. K. *et al.* Structure of the influenza virus A H5N1 nucleoprotein: implications for RNA binding, oligomerization, and vaccine design. *FASEB J* **22**, 3638–3647, <https://doi.org/10.1096/fj.08-112110> (2008).
13. Ye, Q., Krug, R. M. & Tao, Y. J. The mechanism by which influenza A virus nucleoprotein forms oligomers and binds RNA. *Nature* **444**, 1078–1082, <https://doi.org/10.1038/nature05379> (2006).
14. Chenavas, S. *et al.* Monomeric nucleoprotein of influenza A virus. *PLoS Pathog* **9**, e1003275, <https://doi.org/10.1371/journal.ppat.1003275> (2013).
15. Ng, A. K. *et al.* Structural basis for RNA binding and homo-oligomer formation by influenza B virus nucleoprotein. *J Virol* **86**, 6758–6767, <https://doi.org/10.1128/JVI.00073-12> (2012).



16. Zheng, W., Olson, J., Vakharia, V. & Tao, Y. J. The crystal structure and RNA-binding of an orthomyxovirus nucleoprotein. *PLoS Pathog* **9**, e1003624, <https://doi.org/10.1371/journal.ppat.1003624> (2013).
17. Cros, J. F., Garcia-Sastre, A. & Palese, P. An unconventional NLS is critical for the nuclear import of the influenza A virus nucleoprotein and ribonucleoprotein. *Traffic* **6**, 205–213, <https://doi.org/10.1111/j.1600-0854.2005.00263.x> (2005).
18. Labaronne, A. *et al.* Structural analysis of the complex between influenza B nucleoprotein and human importin- $\alpha$ . *Sci Rep* **7**, 17164, <https://doi.org/10.1038/s41598-017-17458-z> (2017).
19. Liu, M. *et al.* The Functional Study of the N-Terminal Region of Influenza B Virus Nucleoprotein. *PLoS One* **10**, e0137802, <https://doi.org/10.1371/journal.pone.0137802> (2015).
20. Ozawa, M. *et al.* Contributions of two nuclear localization signals of influenza A virus nucleoprotein to viral replication. *J Virol* **81**, 30–41, <https://doi.org/10.1128/JVI.01434-06> (2007).
21. Sherry, L., Smith, M., Davidson, S. & Jackson, D. 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, <https://doi.org/10.1128/JVI.01542-14> (2014).
22. Wang, P., Palese, P. & O'Neill, R. E. 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 (1997).
23. Wanitchang, A., Narkpuk, J. & Jongkaewwattana, A. Nuclear import of influenza B virus nucleoprotein: involvement of an N-terminal nuclear localization signal and a cleavage-protection motif. *Virology* **443**, 59–68, <https://doi.org/10.1016/j.virol.2013.04.025> (2013).
24. Wu, W. W., Sun, Y. H. & Pante, N. Nuclear import of influenza A viral ribonucleoprotein complexes is mediated by two nuclear localization sequences on viral nucleoprotein. *Virology* **44**, 49, <https://doi.org/10.1186/1743-422X-4-49> (2007).
25. Nakada, R., Hirano, H. & Matsuura, Y. Structure of importin- $\alpha$  bound to a non-classical nuclear localization signal of the influenza A virus nucleoprotein. *Sci Rep* **5**, 15055, <https://doi.org/10.1038/srep15055> (2015).
26. Wu, W. *et al.* Synergy of two low-affinity NLSs determines the high avidity of influenza A virus nucleoprotein NP for human importin  $\alpha$  isoforms. *Sci Rep* **7**, 11381, <https://doi.org/10.1038/s41598-017-11018-1> (2017).
27. Labaronne, A. *et al.* Binding of RNA by the Nucleoproteins of Influenza Viruses A and B. *Viruses* **8**, <https://doi.org/10.3390/v8090247> (2016).
28. Tarus, B. *et al.* Oligomerization paths of the nucleoprotein of influenza A virus. *Biochimie* **94**, 776–785, <https://doi.org/10.1016/j.biochi.2011.11.009> (2012).
29. Leung, D. W. *et al.* An Intrinsically Disordered Peptide from Ebola Virus VP35 Controls Viral RNA Synthesis by Modulating Nucleoprotein-RNA Interactions. *Cell Rep* **11**, 376–389, <https://doi.org/10.1016/j.celrep.2015.03.034> (2015).
30. Chan, W. H. *et al.* Functional analysis of the influenza virus H5N1 nucleoprotein tail loop reveals amino acids that are crucial for oligomerization and ribonucleoprotein activities. *J Virol* **84**, 7337–7345, <https://doi.org/10.1128/JVI.02474-09> (2010).
31. Elton, D., Medcalf, E., Bishop, K. & Digard, P. Oligomerization of the influenza virus nucleoprotein: identification of positive and negative sequence elements. *Virology* **260**, 190–200, <https://doi.org/10.1006/viro.1999.9818> (1999).
32. Shen, Y. F. *et al.* E339...R416 salt bridge of nucleoprotein as a feasible target for influenza virus inhibitors. *Proc Natl Acad Sci USA* **108**, 16515–16520, <https://doi.org/10.1073/pnas.1113107108> (2011).
33. Turrell, L., Lyall, J. W., Tiley, L. S., Fodor, E. & Vreede, F. T. The role and assembly mechanism of nucleoprotein in influenza A virus ribonucleoprotein complexes. *Nat Commun* **4**, 1591, <https://doi.org/10.1038/ncomms2589> (2013).
34. Mondal, A., Potts, G. K., Dawson, A. R., Coon, J. J. & Mehle, A. Phosphorylation at the homotypic interface regulates nucleoprotein oligomerization and assembly of the influenza virus replication machinery. *PLoS Pathog* **11**, e1004826, <https://doi.org/10.1371/journal.ppat.1004826> (2015).
35. Mondal, A. *et al.* Influenza virus recruits host protein kinase C to control assembly and activity of its replication machinery. *Elife* **6**, <https://doi.org/10.7554/eLife.26910> (2017).
36. 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).
37. 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* **73**, 7357–7367 (1999).
38. Gerard, F. C. *et al.* Modular organization of rabies virus phosphoprotein. *J Mol Biol* **388**, 978–996, <https://doi.org/10.1016/j.jmb.2009.03.061> (2009).
39. Uversky, V. N. Size-exclusion chromatography in structural analysis of intrinsically disordered proteins. *Methods Mol Biol* **896**, 179–194, [https://doi.org/10.1007/978-1-4614-3704-8\\_11](https://doi.org/10.1007/978-1-4614-3704-8_11) (2012).
40. Baskakov, I. & Bolen, D. W. Forcing thermodynamically unfolded proteins to fold. *J Biol Chem* **273**, 4831–4834 (1998).
41. Baskakov, I., Wang, A. & Bolen, D. W. Trimethylamine-N-oxide counteracts urea effects on rabbit muscle lactate dehydrogenase function: a test of the counteraction hypothesis. *Biophys J* **74**, 2666–2673, [https://doi.org/10.1016/S0006-3495\(98\)77972-X](https://doi.org/10.1016/S0006-3495(98)77972-X) (1998).
42. Baskakov, I. V. *et al.* Trimethylamine N-oxide-induced cooperative folding of an intrinsically unfolded transcription-activating fragment of human glucocorticoid receptor. *J Biol Chem* **274**, 10693–10696 (1999).
43. Stevens, M. P. & Barclay, W. S. The N-terminal extension of the influenza B virus nucleoprotein is not required for nuclear accumulation or the expression and replication of a model RNA. *J Virol* **72**, 5307–5312 (1998).
44. White, S. K., Ma, W., McDaniel, C. J., Gray, G. C. & Lednicky, J. A. Serologic evidence of exposure to influenza D virus among persons with occupational contact with cattle. *J Clin Virol* **81**, 31–33, <https://doi.org/10.1016/j.jcv.2016.05.017> (2016).
45. Song, H. *et al.* 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, <https://doi.org/10.1371/journal.ppat.1005411> (2016).
46. Liao, T. L., Wu, C. Y., Su, W. C., Jeng, K. S. & Lai, M. M. Ubiquitination and deubiquitination of NP protein regulates influenza A virus RNA replication. *EMBO J* **29**, 3879–3890, <https://doi.org/10.1038/emboj.2010.250> (2010).
47. Turrell, L., Hutchinson, E. C., Vreede, F. T. & Fodor, E. Regulation of influenza A virus nucleoprotein oligomerization by phosphorylation. *J Virol* **89**, 1452–1455, <https://doi.org/10.1128/JVI.02332-14> (2015).
48. Shi, M. *et al.* The evolutionary history of vertebrate RNA viruses. *Nature* **556**, 197–202, <https://doi.org/10.1038/s41586-018-0012-7> (2018).
49. Hause, B. M. *et al.* Characterization of a novel influenza virus in cattle and Swine: proposal for a new genus in the Orthomyxoviridae family. *MBio* **5**, e00031–00014, <https://doi.org/10.1128/mBio.00031-14> (2014).
50. 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, <https://doi.org/10.1006/viro.1998.9329> (1998).
51. Aspehaug, V. *et al.* Infectious salmon anemia virus (ISAV) genomic segment 3 encodes the viral nucleoprotein (NP), an RNA-binding protein with two monopartite nuclear localization signals (NLS). *Virus Res* **106**, 51–60, <https://doi.org/10.1016/j.virusres.2004.06.001> (2004).
52. Longhi, S. *et al.* 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, <https://doi.org/10.1074/jbc.M300518200> (2003).
53. Bloyet, L. M. *et al.* Modulation of Re-initiation of Measles Virus Transcription at Intergenic Regions by PXD to NTAIL Binding Strength. *PLoS Pathog* **12**, e1006058, <https://doi.org/10.1371/journal.ppat.1006058> (2016).

54. Communie, G. *et al.* Atomic resolution description of the interaction between the nucleoprotein and phosphoprotein of Hendra virus. *PLoS Pathog* **9**, e1003631, <https://doi.org/10.1371/journal.ppat.1003631> (2013).
55. Jenson, M. R. *et al.* Intrinsic disorder in measles virus nucleocapsids. *Proc Natl Acad Sci USA* **108**, 9839–9844, <https://doi.org/10.1073/pnas.1103270108> (2011).
56. Schoehn, G. *et al.* The 12 Å structure of trypsin-treated measles virus N-RNA. *J Mol Biol* **339**, 301–312, <https://doi.org/10.1016/j.jmb.2004.03.073> (2004).
57. Milles, S. *et al.* Self-Assembly of Measles Virus Nucleocapsid-like Particles: Kinetics and RNA Sequence Dependence. *Angew Chem Int Ed Engl* **55**, 9356–9360, <https://doi.org/10.1002/anie.201602619> (2016).
58. Arranz, R. *et al.* The structure of native influenza virion ribonucleoproteins. *Science* **338**, 1634–1637, <https://doi.org/10.1126/science.1228172> (2012).
59. Klumpp, K., Ruigrok, R. W. & Baudin, F. Roles of the influenza virus polymerase and nucleoprotein in forming a functional RNP structure. *EMBO J* **16**, 1248–1257, <https://doi.org/10.1093/emboj/16.6.1248> (1997).
60. Moeller, A., Kirchdoerfer, R. N., Potter, C. S., Carragher, B. & Wilson, I. A. Organization of the influenza virus replication machinery. *Science* **338**, 1631–1634, <https://doi.org/10.1126/science.1227270> (2012).
61. Gallagher, J. R., Torian, U., McCraw, D. M. & Harris, A. K. Structural studies of influenza virus RNPs by electron microscopy indicate molecular contortions within NP supra-structures. *J Struct Biol*, <https://doi.org/10.1016/j.jsb.2016.12.007> (2016).
62. Gutsche, I. *et al.* Structural virology. Near-atomic cryo-EM structure of the helical measles virus nucleocapsid. *Science* **348**, 704–707, <https://doi.org/10.1126/science.aaa5137> (2015).
63. Boivin, S. & Hart, D. J. Interaction of the influenza A virus polymerase PB2 C-terminal region with importin alpha isoforms provides insights into host adaptation and polymerase assembly. *J Biol Chem* **286**, 10439–10448, <https://doi.org/10.1074/jbc.M110.182964> (2011).
64. Wyatt, P. J. Submicrometer Particle Sizing by Multiangle Light Scattering following Fractionation. *J Colloid Interface Sci* **197**, 9–20 (1998).
65. Gaudin, Y., Ruigrok, R. W., Tuffereau, C., Knossow, M. & Flamand, A. Rabies virus glycoprotein is a trimer. *Virology* **187**, 627–632 (1992).
66. Kabsch, W. X. *Acta Crystallogr D Biol Crystallogr* **66**, 125–132, <https://doi.org/10.1107/S0907444909047337> (2010).
67. Cohen, S. X. *et al.* ARP/wARP and molecular replacement: the next generation. *Acta Crystallogr D Biol Crystallogr* **64**, 49–60, <https://doi.org/10.1107/S0907444907047580> (2008).
68. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486–501, <https://doi.org/10.1107/S0907444910007493> (2010).
69. McCoy, A. J. *et al.* Phaser crystallographic software. *J Appl Crystallogr* **40**, 658–674, <https://doi.org/10.1107/S0021889807021206> (2007).
70. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**, 240–255, <https://doi.org/10.1107/S0907444996012255> (1997).
71. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67**, 235–242, <https://doi.org/10.1107/S0907444910045749> (2011).
72. Smart, O. S. *et al.* Exploiting structure similarity in refinement: automated NCS and target-structure restraints in BUSTER. *Acta Crystallogr D Biol Crystallogr* **68**, 368–380, <https://doi.org/10.1107/S0907444911056058> (2012).
73. Schrodinger, LLC. *The PyMOL Molecular Graphics System, Version 1.8* (2015).
74. Pantoliano, M. W. *et al.* High-density miniaturized thermal shift assays as a general strategy for drug discovery. *J Biomol Screen* **6**, 429–440, <https://doi.org/10.1089/108705701753364922> (2001).
75. Rocchia, W. *et al.* Rapid grid-based construction of the molecular surface and the use of induced surface charge to calculate reaction field energies: applications to the molecular systems and geometric objects. *J Comput Chem* **23**, 128–137, <https://doi.org/10.1002/jcc.1161> (2002).
76. Ericsson, U. B., Hallberg, B. M., Detitta, G. T., Dekker, N. & Nordlund, P. Thermofluor-based high-throughput stability optimization of proteins for structural studies. *Anal Biochem* **357**, 289–298, <https://doi.org/10.1016/j.ab.2006.07.027> (2006).
77. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**, 1792–1797, <https://doi.org/10.1093/nar/gkh340> (2004).
78. Krissinel, E. & Henrick, K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr* **60**, 2256–2268, <https://doi.org/10.1107/S0907444904026460> (2004).

## Acknowledgements

We are highly grateful to Jean-Marie Bourhis, Daphna Fenel and Emmanuelle Neumann for their help. We thank Darren J Hart for the plasmid to express the human importin- $\alpha$ 7. We thank Martin Blackledge, Anny Slama-Schwok and Bernard Delmas for discussion. We thank Stéphane Bertagnoli for his help with the rabbit immunizations. We thank the staff of the ESRF-EMBL Joint Structural Biology Group for access to ESRF beamlines. AD, AL and MM were funded through the Labex GRAL (ANR-10-LABX-49-01), LT by the ANR RNAP-IAV (ANR-14-CE09-0017) and AD and JO are supported by a PhD scholarship of the French Ministry of Higher Education and Research. This work was supported by the Fondation de la Recherche Médicale (FRM; Equipe DEQ. 20170336754) and used the electron microscopy and the biophysical platforms of the Grenoble Instruct-ERIC Center (ISBG; UMS 3518 CNRS-CEA-UGA-EMBL) with support from FRISBI (ANR-10-INSB-05-02) and GRAL, within the Grenoble Partnership for Structural Biology (PSB), as well as the Imagery platform of Centre de Physiopathologie de Toulouse Purpan. The IBS Electron Microscope facility is supported by the Rhône-Alpes Region, the Fonds Feder, the FRM and GIS-IBISA.

## Author Contributions

T.C. and M.D. conceived the experiments. A.D., J.O., A.L., L.T., M.M., F.C.A.G., C.M., G.S., M.D. and T.C. performed the experiments. A.D., J.O., A.L., F.C.A.G., G.S., R.W.H.R., M.D. and T.C. analysed the data. R.W.H.R., M.D. and T.C. wrote the paper.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-018-37306-y>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019

Chapter 3: Comparative analysis of the interaction involving each influenza nucleoprotein types and importins- $\alpha$  involved in influenza nuclear import

The nuclear import of the influenza replication machinery is a critical step of the viral cycle and this host:pathogen interaction is one of the tropism determinants. The initial nuclear import of the ribonucleoproteins is believed to be mediated through the NP<sub>TAIL</sub> NLS, while the newly synthesized NP and RNA polymerase subunits possess their own NLS for nuclear import purposes<sup>28,54,57,283,346</sup>. The nuclear translocation is a complex, highly regulated system. Small ions and molecules can passively diffuse in the nucleus, but bigger proteins such as NP and the viral RNA polymerase require an active transport in the nucleus, and more particularly by hijacking the Ran/importin- $\alpha$ /importin- $\beta$  system<sup>139,156,211,263,266,306</sup>.

A large number of studies have tried to describe and characterize the interaction between the cellular nuclear import machinery and the influenza proteins. A/NP has been the most characterized in complex with importins- $\alpha$ , and has been shown to interact with importin- $\alpha$ 1, - $\alpha$ 3, - $\alpha$ 5 and - $\alpha$ 7<sup>27,232,261,381,403</sup>. While precious insights were gained, the affinities obtained so far is in the low micromolar, outside of the *in vitro* range of a functional NLS<sup>360</sup>. Similar results have been obtained for B/NP<sup>177,328,383,404</sup>. In contrast C/NP and D/NP present *in vitro* affinities more consistent with a nuclear import efficiency *in cellulo*, but these studies did not thoroughly investigate all the nuclear import partners<sup>350</sup>. Importins- $\alpha$  are the direct partners of influenza NPs in the nuclear shuttling, but they were, so far, not studied extensively and comparatively. It is now well-known that importins- $\alpha$  are distinct entities, with their own regulatory behaviours and binding preferences<sup>94,283,360</sup>.

In this paper, we performed a complete comparative analysis of the interactions between IAV, IBV, ICV, IDV NP<sub>TAILS</sub> and the four human importins- $\alpha$  identified as partners. Using SEC, SPR, SEC-MALLS analysis and FA, we obtained affinities for all these complexes and unravelled striking specificities.

Using YFP-fused NP<sub>TAILS</sub> to avoid the NP oligomerization and avidity contribution while keeping a globular context, we showed that these constructs were still able to interact with the importins *in vitro*. SPR revealed interactions with fast kinetics for A/NP<sub>TAIL</sub>, C/NP<sub>TAIL</sub> and D/NP<sub>TAIL</sub>. D/NP<sub>TAIL</sub> was able to interact with strong affinities with the importins- $\alpha$ 3, - $\alpha$ 5 and - $\alpha$ 7. Both A/NP<sub>TAIL</sub> and C/NP<sub>TAIL</sub> interact with varying affinities with these importins- $\alpha$ , with seemingly one preferential partner. B/NP<sub>TAIL</sub>, on the other hand, showed a completely different behaviour, with even faster association and

dissociation rates and a low affinity for all importins. The contribution of previously identified key residues of B/NP<sub>TAIL</sub> and D/NP<sub>TAIL</sub> was analyzed by alanine scanning and resulted in a drastic decrease in affinity. The importin- $\alpha$ 1, showing homodimerization behaviour in solution, could not be used in the surface plasmon resonance experiment. Instead, a SEC-MALLS competition assay and fluorescence anisotropy measurements were used and revealed a strong affinity of importin- $\alpha$ 1 for D/NP<sub>TAIL</sub> and C/NP<sub>TAIL</sub>, medium-to-low affinity for A/NP<sub>TAIL</sub> and poor affinity for B/NP<sub>TAIL</sub>.

This paper provides insights on the specificity of these crucial interactions, for which both partners contributes equally. It also provides a strong basis for further studies aiming at dissecting this particular, essential host:pathogen interaction.

---





**Article: Differential behaviors and preferential binding for influenza nucleoproteins and importins- $\alpha$**

Donchet A., Stermann E., Gerard F.C.A., Ruigrok R.W.H. and Crépin T  
Viruses (2020)

---

Article

# Differential Behaviours and Preferential Bindings of Influenza Nucleoproteins on Importins- $\alpha$

Amélie Donchet, Emilie Vassal-Stermann , Francine C. A. Gérard <sup>†</sup> , Rob W. H. Ruigrok  and Thibaut Crépin <sup>\*</sup> 

Centre National de la Recherche Scientifique (CNRS), Institut de Biologie Structurale (IBS), University Grenoble Alpes, Commissariat à l’Energie Atomique et aux Energies Alternatives (CEA), 38044 Grenoble, France; amelie.donchet@ibs.fr (A.D.); emilie.stermann@ibs.fr (E.V.-S.); francine.baraggia@grenoble-inp.fr (F.C.A.G.); rob.ruigrok@ibs.fr (R.W.H.R.)

\* Correspondence: thibaut.crepin@ibs.fr; Tel.: +33-(0)476-209-439

† Current address: Laboratoire des Matériaux et du Génie Physique (LMGP), University Grenoble Alpes, CNRS, Institut Polytechnique de Grenoble (Grenoble INP), F-38000, Grenoble, France.

Received: 30 May 2020; Accepted: 27 July 2020; Published: 30 July 2020



**Abstract:** Influenza viruses are negative single-stranded RNA viruses with nuclear transcription and replication. They enter the nucleus by using the cellular importin- $\alpha$ / $\beta$  nuclear import machinery. Influenza nucleoproteins from influenza A, B, C and D viruses possess a nuclear localization signal (NLS) localized on an intrinsically disordered extremity (NP<sub>TAIL</sub>). In this paper, using size exclusion chromatography (SEC), SEC-multi-angle laser light scattering (SEC-MALLS) analysis, surface plasmon resonance (SPR) and fluorescence anisotropy, we provide the first comparative study designed to dissect the interaction between the four NP<sub>TAIL</sub>s and four importins- $\alpha$  identified as partners. All interactions between NP<sub>TAIL</sub>s and importins- $\alpha$  have high association and dissociation rates and present a distinct and specific behaviour. D/NP<sub>TAIL</sub> interacts strongly with all importins- $\alpha$  while B/NP<sub>TAIL</sub> shows weak affinity for importins- $\alpha$ . A/NP<sub>TAIL</sub> and C/NP<sub>TAIL</sub> present preferential importin- $\alpha$  partners. Mutations in B/NP<sub>TAIL</sub> and D/NP<sub>TAIL</sub> show a loss of importin- $\alpha$  binding, confirming key NLS residues. Taken together, our results provide essential highlights of this complex translocation mechanism.

**Keywords:** influenza nucleoprotein; nuclear transport; importin- $\alpha$ ; nuclear localization signal; influenza-host interaction; surface plasmon resonance; fluorescence anisotropy

## 1. Introduction

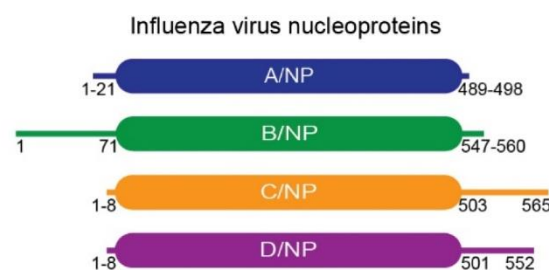
Influenza viruses are part of the *Orthomyxoviridae* family and form the genus *Influenzavirus*. Up to now, the International Committee officially recognizes four types (respectively types A, B, C and D) in its Taxonomy of Viruses (ICTV), but because of actual large-scale meta-transcriptomic approaches new specimens are regularly discovered, making this classification a constantly evolving entity. While influenza viruses were thought to be uniquely warm-blooded organism pathogens, a study has recently uncovered their existence in fish and amphibians in China [1]. A first criterion of distinction between the actual four types of *Influenzavirus* is based on their tropism. Influenza type A viruses (IAV) are viruses from water fowl but some of the viruses can also infect a range of hosts from chickens to swine and humans; type B (IBV) almost exclusively infect humans; type C (ICV) can infect both human and swine whereas type D (IDV) has a broad host tropism but limited to mammals. Influenza viruses are negative-sense single-stranded RNA viruses with segmented genomes. The RNA segments are packaged in ribonucleoproteins (RNPs), a complex architecture made by a single RNA molecule encapsidated by multiple copies of the nucleoprotein (NP) and the heterotrimeric RNA-dependent



RNA polymerase (RdRp) bound to the strictly conserved 3' and 5' end [2,3]. The number of RNPs varies depending on the type: IAV and IBV genomes consist of eight RNA segments, i.e., there are eight different RNPs, but only seven for ICV and IDV viruses. This is related to the inherent organisation of the viral spikes, respectively the haemagglutinin (HA) plus the neuraminidase (NA) for A and B or a single hemagglutinin-esterase-fusion (HEF) polypeptide with both concatenated activities for C and D.

Unlike most RNA viruses, influenza viruses transcribe and replicate their genome in the nucleus of infected cells. In general, the access to the nuclear compartment is through the highly regulated nuclear pore complex (NPC) [4–7]. Several pathways are involved in cytoplasmic-nuclear trafficking. Small molecules (< 30–60 kDa) are able to diffuse passively through the NPC but bigger molecules often require an active mechanism to be imported [8–10]. The main active transport is based on the importins- $\alpha/\beta$  (coupled to the small GTPase Ran) pathway. Importins- $\alpha$  and - $\beta$  can both interact with the cargos but only the last ones mediate directly the translocation, and importin- $\alpha$ :cargo complexes must first interact with importin- $\beta$  [11,12]. Cargos are recognized by importins- $\alpha/\beta$  through nuclear localization signals (NLSs), usually made of basic or hydrophobic motifs [13,14]. Additionally to the interaction with the NLS motifs, recent karyopherins- $\beta$ /cargo-protein structures have demonstrated that these complexes are dependent on multiple interactions, mediated through intricate 3D interfaces which involve residues well beyond the NLS peptide stretch [15–17].

The human genome encodes for seven importins- $\alpha$  and at least twenty importins- $\beta$  [9,18,19], meaning that the importins- $\alpha/\beta$  pathway is a difficult problem. Nuclear import of influenza virus proteins has been of a major interest for a long time [20,21]. Considering the neo-synthesized components of the replication machinery, two subunits (PA and PB1) of the heterotrimeric RdRp are imported as a sub-complex by the importin- $\beta$  RanBP5 [17,22–24] whereas the third subunit (PB2) and NP interact with importins- $\alpha$  [25–28]. Because of its role in the RNPs formation, NP is critical for influenza replication but also shown to be the major contributor to the nuclear transport of the RNPs [29]. The NLS of NPs are located in intrinsically disordered tails of the proteins, either in  $N_{\text{terminal}}$  for A/ and B/NP or in  $C_{\text{terminal}}$  for C/ and D/NP (Figure 1). A minor putative NLS has been located in a loop of the core of A/NP, between residues 198 to 216, but its role remains unclear [30,31]. A/NP is known to interact with several importin- $\alpha$  isoforms, including the universal - $\alpha$ 1, - $\alpha$ 3, - $\alpha$ 5 and possibly - $\alpha$ 7 [32–36]. In recent years, our group has worked on the mechanisms of the interaction between two different NPs and the human nuclear trafficking machinery [37,38], in particular the tails containing the NLS of influenza B (B/NP<sub>TAIL</sub>) and D (D/NP<sub>TAIL</sub>) NPs. We showed that the NLS of B/NP<sub>TAIL</sub> is highly extended from residues 30 to 71, around a single basic Lysine-Arginine (44-KR-45) patch, with a low affinity of 844 nM for the human importin- $\alpha$ 7. In the case of the  $C_{\text{terminal}}$  D/NP<sub>TAIL</sub>, the NLS was composed of two basic KR patches, respectively 514-KR-515 and 532-KR-533. Both contribute to the nuclear trafficking and we estimated that the affinity of D/NP<sub>TAIL</sub> for the human importin- $\alpha$ 7 was at 100 nM by surface plasmon resonance (SPR). The literature also contains several studies detailing the relation between A/NP and C/NP and importins- $\alpha$  [21,29,31,33,39–43] (Table 1). However, recent work highlighted that each importin- $\alpha$  is unique, with its own features and preferential partners [28,35,44,45].



**Figure 1.** *Influenzavirus* nucleoproteins. The schematic representations of *Influenzavirus* nucleoproteins are based on this structural analysis [37]. The schema respects the size of the proteins. The flexible tails with the NLSs are represented with simple lines whereas the cores are represented with filled boxes.

**Table 1.** Binding of influenza nucleoproteins or their peptides to importins- $\alpha$  according to the literature. Abbreviations: ITC, isothermal titration calorimetry; SPBA, solid phase binding assay; MT, microscale thermophoresis; SPR, surface plasmon resonance.

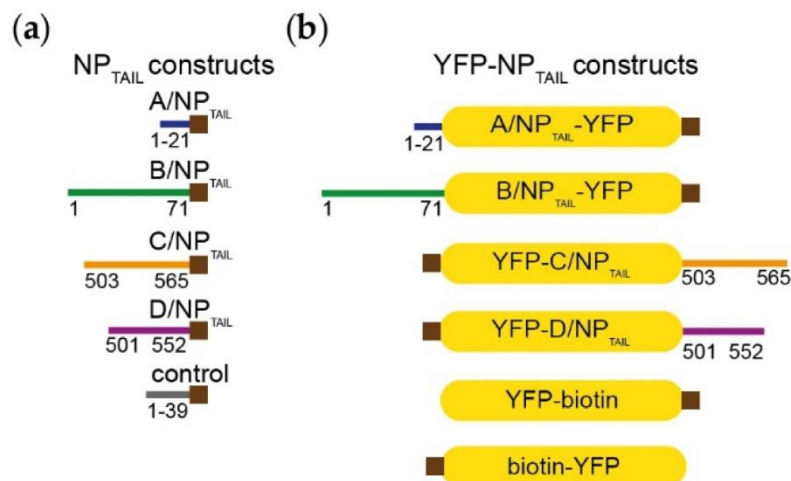
	Construct	Importin- $\alpha$	Method	$K_d$	Ref
A/NP	R416A	$\alpha 5$	ITC	26 nM	[41]
A/NP	2–15	$\alpha 1$	SPBA	1.7 $\mu$ M	[42]
A/NP	1–13	$\alpha 1$	ITC	5 $\mu$ M	[31]
A/NP	198–216	$\alpha 1$	ITC	72 $\mu$ M	[31]
B/NP	1–70	$\alpha 7$	ITC	844 nM	[38]
C/NP	506–565	$\alpha 1$	MT	48 nM	[43]
D/NP	505–552	$\alpha 7$	SPR	100 nM	[37]

Because we were confronted with a wide range of values (from 26 nM to 72  $\mu$ M) obtained using different techniques and a plethora of different partner constructs, we wanted to establish a systematic comparative study of the *in vitro* interaction between the flexible extremities (NP<sub>TAILs</sub>) of all types of influenza NPs and the different human importin- $\alpha$  isoforms described in the literature. Our strategy was design, with a specific focus upon the short NLS sequences, and built with the objective of avoiding the inherent oligomerization of NPs, while guaranteeing the architecture of the tails in a unique globular environment. Using size exclusion chromatography experiments, we show that all NP<sub>TAILs</sub> interact with importins- $\alpha$  in solution and then used surface plasmon resonance measurements to determine the affinity and binding kinetics of these interactions. In the current paper, we show that D/NP<sub>TAIL</sub> interacts with all importins- $\alpha$  isoforms with strong affinity. Both A/NP<sub>TAIL</sub> and C/NP<sub>TAIL</sub> interact with varying affinities with the corresponding importins- $\alpha$ . On the other hand, B/NP<sub>TAIL</sub> shows a completely different behaviour, with even faster association and dissociation rates and a low affinity with all the human importins- $\alpha$  we measured.

## 2. Materials and Methods

### 2.1. Molecular Biology and Constructs

Human importin- $\alpha 1$  (residues 69 to 529), - $\alpha 3$  (residues 64 to 521) and - $\alpha 5$  (residues 66 to 538) DNA coding sequences were bought at GENEART GmbH (Thermo Fischer Scientific, Regensburg, Germany) and subcloned into pETM11 bacterial expression vector (EMBL) without their importin- $\beta$  (IBB) domains. Human importin- $\alpha 7$  (residues 58 to 536) cloning was described in [26]. NLS-bearing NP<sub>TAILs</sub> from strains A/WSN/1933 (residues 1 to 21), B/Memphis/13/03 (residues 1 to 71), C/Ann-Arbor/1/50 (residues 508 to 565) and D/bovine/France/2986/2012 (residues 505 to 552) were ordered from GeneART, fused to a biotinylation sequence in C<sub>terminal</sub> (Figure 2, and details of the sequences are provided in Table S1). The NP<sub>TAILs</sub> were fused to the yellow fluorescent protein (YFP) either in C<sub>terminal</sub> (C/NP<sub>TAIL</sub> and D/NP<sub>TAIL</sub>) or in N<sub>terminal</sub> (A/NP<sub>TAIL</sub>, B/NP<sub>TAIL</sub>) according to their native localization; at the other YFP extremity, the biotinylation sequence was added (Figure 2). NP<sub>TAILs</sub>-YFP constructs were subcloned into pETM11 bacterial expression vector. All constructs were expressed as N-terminally His-tagged proteins. Sequencing was performed by Eurofins Genomics (Köln, Germany). The control peptide was a synthetic peptide (GENPEP, St-Jean-de-Védas, France) corresponding to a GLA domain-derived polypeptide of the human coagulation factor X [46].



**Figure 2.** Constructs. (a) *Influenzavirus* nucleoproteins tails were expressed and purified independently fused with a biotinylation sequence (brown squares) in their C<sub>terminal</sub> end. The details of the sequences are provided in Table S1. (b) The corresponding tails were fused to the yellow fluorescent protein (YFP) in order to integrate then in the context of a globular protein. The yellow filled boxes represent the YFP and the brown squares indicate the position of the biotinylation sequence.

## 2.2. Expression and Purification

*Escherichia coli* BL21 RIL (DE3) cells (Life Technologies, Thermo Fischer Scientific, Courtaboeuf, France) were used for protein expression. Cultures were supplemented with biotin (12.5  $\mu\text{g}\cdot\text{mL}^{-1}$ , Sigma, Saint-Quentin-Fallavier, France) to biotinylate the proteins *in cellulo*, induced for 12 h by adding 0.3 mM isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG; Euromedex, Souffelweyersheim, France) at 18 °C and collected by centrifugation. Pellets were resuspended in 50 mM Tris-HCl pH 7.5, 300 mM NaCl, 2 mM  $\beta$ -mercaptoethanol ( $\beta$ -ME; Roth, Lagny-sur-Marne, France) and cComplete™, ethylenediaminetetraacetic acid (EDTA)-free protease inhibitor cocktail (Roche, Meylan, France) before sonication. All purifications included a nickel affinity chromatography (resin Ni-NTA; Qiagen, Les Ulis, France), a Tobacco Etch Virus (TEV) protease cleavage, a second nickel affinity chromatography and a size-exclusion chromatography (SEC). SEC was performed on a NGC system (Bio-Rad, Marnes-La-Coquette, France) with a S75 10/300 GL column (GE-Healthcare, Dutscher, Brumath, France). The final buffer composition for all proteins was 10 mM Hepes pH 7.5, 150 mM NaCl. NP<sub>TAILs</sub>-YFP protein concentrations were determined by measuring the absorbance at 280 nm and using the theoretical extinction coefficients ( $\epsilon$ ) 28,942  $\text{M}^{-1}\cdot\text{cm}^{-1}$  for A/NP<sub>TAIL</sub>-YFP and C/NP<sub>TAIL</sub>-YFP, of 27,452  $\text{M}^{-1}\cdot\text{cm}^{-1}$  for B/NP<sub>TAIL</sub>-YFP, D/NP<sub>TAIL</sub>-YFP and YFP control.  $\epsilon$  at 280 nm used for  $\Delta$ IBB importins- $\alpha$  protein concentrations were 50,607  $\text{M}^{-1}\cdot\text{cm}^{-1}$  for importin- $\alpha$ 1, 42,190  $\text{M}^{-1}\cdot\text{cm}^{-1}$  for importin- $\alpha$ 3, 52,285  $\text{M}^{-1}\cdot\text{cm}^{-1}$  for importin- $\alpha$ 5 and 46,785  $\text{M}^{-1}\cdot\text{cm}^{-1}$  for importin- $\alpha$ 7. The concentrations were then obtained using the Beer–Lambert law.

## 2.3. Interaction Assays by Size Exclusion Chromatography

All size exclusion chromatography (SEC) experiments were performed in the same buffer (10 mM Hepes pH 7.4, 150 mM NaCl) using a Superdex™ 200 increase 10/300GL column (GE-Healthcare) for YFP-NP<sub>TAIL</sub> fusion proteins and a Superdex™ 75 10/300GL column (GE-Healthcare) for all the other NP<sub>TAILs</sub>. Samples were diluted to be used at 30  $\mu\text{M}$  for YFP-NP<sub>TAIL</sub> fusion proteins, 60  $\mu\text{M}$  for NP<sub>TAILs</sub> and 25  $\mu\text{M}$  importin- $\alpha$ 7 in 300  $\mu\text{L}$ . They were incubated for one hour at room temperature before injection on a NGC system (Bio-Rad).

#### 2.4. SEC-MALLS-RI Analysis

Size exclusion chromatography (SEC) followed by multi-angle laser light scattering (MALLS) and refractometry (RI) analysis allow the determination of the molecular mass of a protein or a complex in solution that is independent of its dimensions and shape [47]. SEC was performed with a column (Superdex™ 200 increase 10/300 GL or Superdex 75 10/300 GL) equilibrated with 10 mM Hepes pH 7.5, 150 mM NaCl. Analytical runs were performed at 20 °C with a flow rate of 0.5 mL·min<sup>-1</sup>. Multi-angle laser light scattering (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<sup>-1</sup>. Data were analyzed and weight-averaged molar masses were calculated using the ASTRA software (Wyatt Technology Corp., Santa Barbara, CA, USA).

#### 2.5. Surface Plasmon Resonance

Surface plasmon resonance (SPR) experiments were carried out at 25 °C on a Biacore T200 (GE Healthcare Life Sciences Europe GmbH, Velizy-Villacoublay, France). Avidin (50 µg·mL<sup>-1</sup> in 50 mM acetate buffer pH 4.5; Sigma) was first immobilized on a Serie S CM5 sensor chip (GE-Healthcare) surface through amine coupling chemistry according to the manufacturer instructions, until a coupling level of about 17,000 resonance units (RUs) was reached. The biotinylated NP<sub>TAIL</sub>-YFP (1–10 µg·mL<sup>-1</sup>) or biotinylated YFP control (1–10 µg·mL<sup>-1</sup>) were then captured in surfactant-supplemented HBS-N running buffer (10 mM Hepes pH 7.4, 150 mM NaCl, 0.05% Tween 20; GE Healthcare) at a flow rate of 30 µL·min<sup>-1</sup> until a coupling level of 100–200 RU was obtained. A flow-cell with immobilized avidin only was used as negative control while the flow-cell with the avidin: biotinylated NP<sub>TAIL</sub>-YFP complex was used as active flow-cell. For kinetic measurements, importins-α were serially diluted in surfactant-supplemented HBS-N running buffer. Each concentration was injected in triplicate on both the reference and active flow-cells. Analyte injection and following buffer injection times were set between 55–180 s and 55–200 s respectively, depending on the examined partners, at a flow-rate of 30 µL·min<sup>-1</sup>. Regeneration was required and achieved by pulse injection of 5 M NaCl. All sensorgrams were reference-subtracted. Data were analyzed with Steady State Analysis using the Biacore T200 Evaluation software (GE Healthcare) under a Langmuir 1:1 binding model, assuming that the binding is equivalent and independent for all binding site. Chi<sup>2</sup> values for the fitting were kept below 1 for most of the analysis with a maximum of 2.5, and T-value means for the dissociation constant above 10.

#### 2.6. Fluorescence Anisotropy

Fluorescence anisotropy experiments were performed on a Clariostar microplate reader (BMG Labtech, Champigny-sur-Marne, France) set and dedicated to fluorescence anisotropy measurements, with excitation and emission wavelengths fixed at 482 and 530 nm, respectively. Importin-α1 was serially diluted in buffer (10 mM Hepes pH 7.4, 150 mM NaCl) and mixed in dilution series with the YFP-fused NP<sub>TAILS</sub>. YFP construct final concentration was set at 25 nM. Measurements were done in 384-well plates at room temperature. Blank subtraction was done using the condition with YFP-fused NP<sub>TAILS</sub> alone. Data were normalized using the equation:

$$B = \frac{(A_{exp} - A_{min})}{(A_{max} - A_{min})} \quad (1)$$

where  $B$  corresponds to the YFP-NP<sub>TAIL</sub> bound fraction (value between 0 and 1),  $A_{exp}$  is the anisotropy value observed for one given importin-α1 concentration, and  $A_{min}$  and  $A_{max}$  stand for the fluorescence anisotropy values of the free YFP-NP<sub>TAIL</sub> and bound YFP-NP<sub>TAIL</sub>, respectively. Titrations were then fitted to the GraphPad Prism model “Single binding site with Hill slope (h)”.

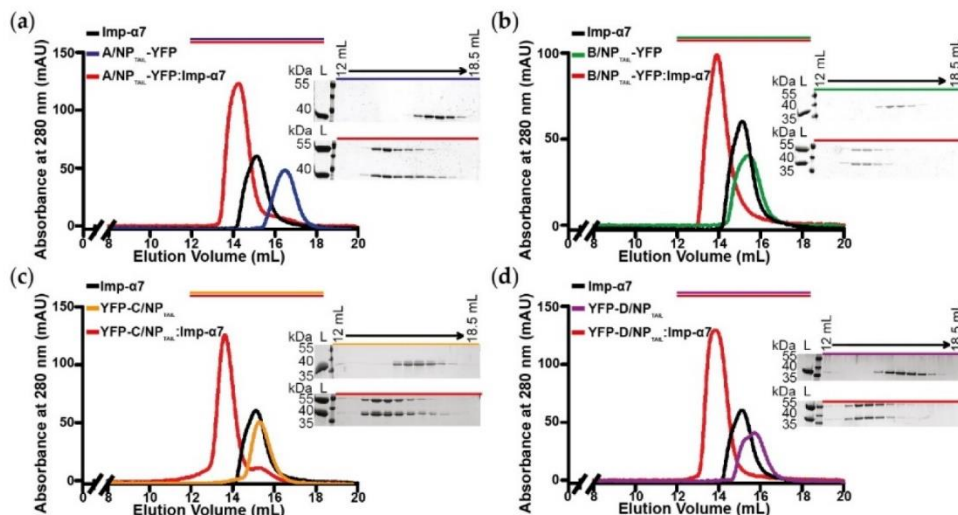
### 3. Results

#### 3.1. Biochemical and Biophysical Analysis of the Interaction between NP<sub>TAILS</sub> and Importin- $\alpha$ 7

In all experiments with NP and its partners, measurements are difficult because all NPs tend to form oligomers depending on the buffer concentrations [41,48], and therefore, we used only the disordered NP<sub>TAILS</sub>. We designed our study with a specific focus upon these short sequences that do not address the effect of domains beyond the NP<sub>TAILS</sub>. In order to decipher the specificities of the different NP<sub>TAILS</sub> for importins- $\alpha$ , we first use the corresponding peptides, in size exclusion chromatography (SEC) experiments combined with surface plasmon resonance (SPR) measurements to determine the affinity and binding kinetics. The first method was used to confirm the ability of the tails to interact with its partner in solution, whereas the second method was used for the determination of the kinetic parameters of the interactions. In SPR, one partner (called the ligand) is immobilized onto a biosensor surface. The interacting partner (called the analyte) diluted in a buffer is then continuously flowed across the biosensor surface, where it binds to the ligand. The binding is measured as a change in resonance units (RUs) on the biosensor surface. Measuring the increase in binding over time for a given analyte set of concentrations gives the association rate ( $k_{on}$ ) up to the point where the system is at equilibrium (i.e., as many association events are observed as dissociation events). Ceasing flow and changing to buffer alone then allows the analyte to wash off the ligand. Measuring the decrease in bound partner over time gives the dissociation rate ( $k_{off}$ ). The equilibrium dissociation constant ( $K_d$ ) is calculated from the kinetic association and dissociation rates ( $k_{off}/k_{on}$ ). When both association and dissociation rates are too fast, it is possible to use the equilibrium part of the curves to extract the dissociation constant, without relying on the rates. For the SPR experiments, we chose to immobilize the tails on the biosensor surface by using the importins- $\alpha$  as the analytes. The immobilization of the tails was done by fusing a biotinylation sequence in C<sub>terminal</sub> (Figure 2a), with a capture strategy based on a biotin: avidin interaction. The corresponding constructs were recombinantly produced in bacteria (with 12.5  $\mu\text{g}\cdot\text{mL}^{-1}$  biotin in the growth media) and purified. We first tested their ability to interact with importin- $\alpha$ 7 by SEC experiments. They all formed a stable complex in solution with importin- $\alpha$ 7 (Figure S1). However, for SPR measurements, the kinetic parameters could not all be established because of technical issues. In particular, for B/NP<sub>TAIL</sub>, the inability to stabilize the SPR signal because of a putative mass transfer phenomenon made it impossible to exploit the raw data. In consequence, we adapted our strategy.

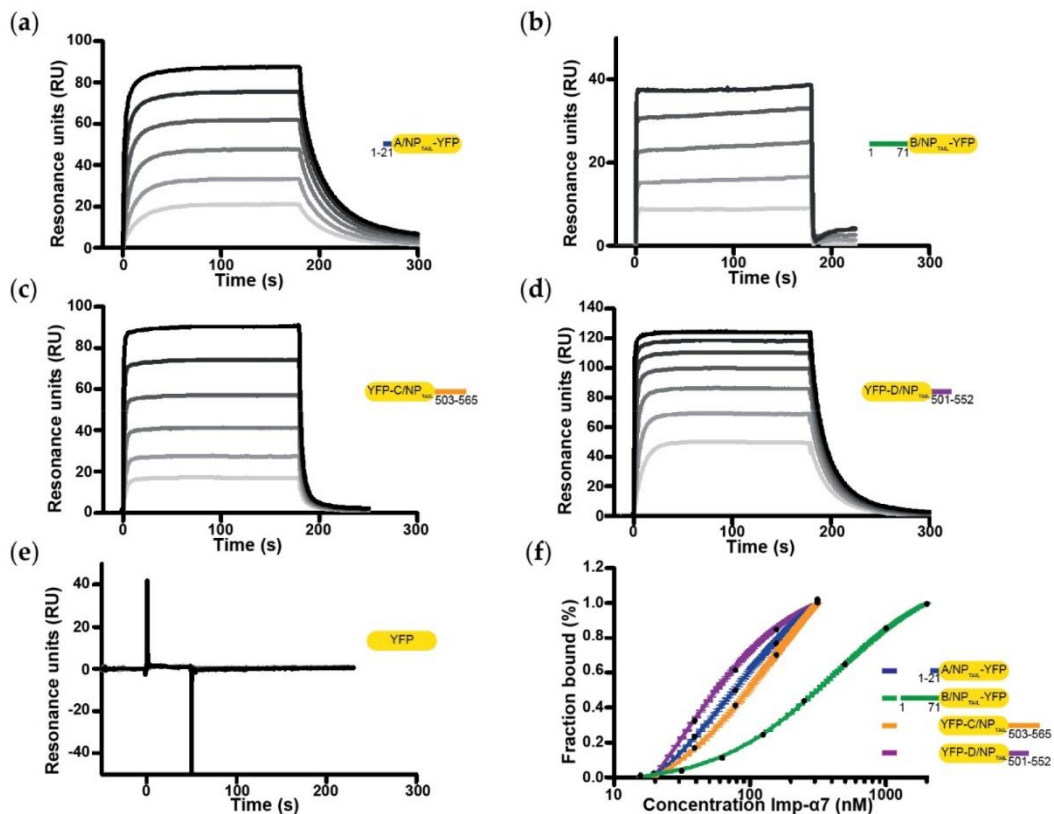
#### 3.2. Biochemical and Biophysical Analysis of the Interaction between NP<sub>TAILS</sub> Fused to Yellow Fluorescent Protein (YFP) and Importin- $\alpha$ 7

We decided to increase the size of the tails by fusing them to a known protein, unrelated to the influenza virus life cycle. A larger ligand should allow the experiment to be performed with lower ligand densities, resulting in lower mass transport rates. We chose the yellow fluorescent protein because of its small size and compact folding, with an unchanged capture strategy on the biosensor surface. A YFP construct was fused in N<sub>terminal</sub> for C/NP<sub>TAIL</sub> and D/NP<sub>TAIL</sub> and in C<sub>terminal</sub> for A/NP<sub>TAIL</sub> and B/NP<sub>TAIL</sub>, to respect their position in the wild-type proteins (Figure 2b). The recombinant proteins were produced, purified (Figure S2) and used in SEC experiments to test their ability to interact with importin- $\alpha$ 7. We confirmed that adding the YFP has no visible effect on the interaction between the NP<sub>TAILS</sub> and the human import factor in solution. For each tail, a peak corresponding to the complex was observed prior those of each separate samples (Figure 3). On the contrary, when the control YFP was mixed with importin- $\alpha$ 7, a double peak corresponding to the sum of each sample injected separately was observed (Figure S3).



**Figure 3.** Analysis of the complexes between YFP- $\text{NP}_{\text{TAILs}}$  and human importin- $\alpha 7$  in solution. Each YFP- $\text{NP}_{\text{TAIL}}$  fusion protein ( $30 \mu\text{M}$ ) was injected on a Superdex<sup>TM</sup> 200 increase 10/300GL column, alone or in the presence of importin- $\alpha 7$  ( $25 \mu\text{M}$ ). The figure shows the superimposition of the SEC profiles obtained for (a) A/ $\text{NP}_{\text{TAIL}}$ -YFP, (b) B/ $\text{NP}_{\text{TAIL}}$ -YFP, (c) YFP-C/ $\text{NP}_{\text{TAIL}}$  and (d) YFP-D/ $\text{NP}_{\text{TAIL}}$ . The colors correspond to the code used on Figure 2b with all the complexes in red and the importin- $\alpha 7$  alone in black. For each panel, the 12% SDS-PAGE corresponding to the elution of the YFP- $\text{NP}_{\text{TAIL}}$  fusion protein alone (top) and its complex with importin- $\alpha 7$  (bottom) are shown. The control experiment is shown on Figure S3.

Using the corresponding samples, we measured the kinetic parameters of the different interactions by SPR (Figure 4; Tables 2 and 3 respectively for affinities ( $K_d$ ) and kinetic parameters values ( $k_{on}$  and  $k_{off}$ )). Gradient concentration of importin- $\alpha 7$  were injected into the SPR flowing channels at a speed of  $30/\mu\text{L}\cdot\text{min}^{-1}$  for 50–180/s to reach equilibrium. As shown in Figure 4, a binding signal was generated immediately after injection of importin- $\alpha 7$  which dissociated from the sensor chip very quickly, suggesting that importin- $\alpha 7$  bound to the immobilized  $\text{NP}_{\text{TAILs}}$  in a quick “in and out” manner. Equilibrium analysis was then chosen to derive affinity binding constants ( $K_d$ ) for the interaction between the immobilized  $\text{NP}_{\text{TAILs}}$  and the importin- $\alpha 7$  in solution. Our reference was the data obtained for the interaction between D/ $\text{NP}_{\text{TAIL}}$  and importin- $\alpha 7$  previously measured using the same method and the same strategy [37] but with a previous generation instrument. Using a Biacore 3000, we had measured a  $K_d$  of  $99 \text{ nM} \pm 9 \text{ nM}$  for this interaction and  $29 \text{ nM} \pm 1 \text{ nM}$  using a new generation Biacore T200. This time, we were able to fully exploit the SPR data obtained with all  $\text{NP}_{\text{TAILs}}$ . In particular, we showed that the architecture of the NLS (mono- or bipartite NLS) has no importance for importin- $\alpha 7$ ; the affinity for the non-conventional monopartite A/ $\text{NP}_{\text{TAIL}}$  was better than for the bipartite C/ $\text{NP}_{\text{TAIL}}$ , with  $K_d$  of  $73 \pm 4 \text{ nM}$  and  $146 \pm 8 \text{ nM}$  respectively, but weaker than D/ $\text{NP}_{\text{TAIL}}$ . B/ $\text{NP}_{\text{TAIL}}$  appeared to be the lowest value with a  $K_d$  of  $405 \pm 5 \text{ nM}$ , but comparable to the value obtained using isothermal titration calorimetry [38]. Despite the nanomolar affinities, all of these interactions displayed a highly dynamic behaviour. Indeed, kinetics parameters show extremely high  $on$  rates, suggesting that very little energy is needed to form the interaction, but also very high  $off$  rates. Both association and dissociation rates were close to the upper detection limits of the apparatus, respectively around  $10^{-6} \text{ M}^{-1}\cdot\text{s}^{-1}$  and  $0.1 \text{ s}^{-1}$  and even above in the case of the B/ $\text{NP}_{\text{TAIL}}$  (Table 3).



**Figure 4.** Affinity between YFP-NP<sub>TAILs</sub> and human importin- $\alpha$ 7 by surface plasmon resonance (SPR). The figure shows the sensorgrams of the interaction between importin- $\alpha$ 7 and (a) A/NP<sub>TAIL</sub>-YFP, (b) B/NP<sub>TAIL</sub>-YFP, (c) YFP-C/NP<sub>TAIL</sub>, (d) YFP-D/NP<sub>TAIL</sub> and (e) the YFP control. For each panel, the gradation of grey represents the different concentrations of importin- $\alpha$ 7 used for each titration, from the lowest (light-grey) to the highest (dark-grey). The concentrations range from 19.5 nM to 625 nM for (a) and (c), from 125 nM to 2  $\mu$ M for (b), from 19.5 nM to 1.25  $\mu$ M for (d) and from 80.5 nM to 10.3  $\mu$ M for (e). (f) Steady state analysis was performed on Biacore evaluation software, the resulting fitting normalized and plotted on a log<sub>10</sub> scale.

**Table 2.** Affinity between wild-type influenza nucleoprotein tails and human importins- $\alpha$  obtained by SPR. The apparent dissociation constants ( $K_d$ ) are expressed in nM; n.i. means ‘no interaction’.

	Imp- $\alpha$ 3	Imp- $\alpha$ 5	Imp- $\alpha$ 7
A/NP <sub>TAIL</sub>	301 $\pm$ 21	128 $\pm$ 20	73 $\pm$ 4
B/NP <sub>TAIL</sub>	308 $\pm$ 106	770 $\pm$ 24	405 $\pm$ 5
C/NP <sub>TAIL</sub>	59 $\pm$ 9	287 $\pm$ 31	146 $\pm$ 8
D/NP <sub>TAIL</sub>	22 $\pm$ 15	21 $\pm$ 5	29 $\pm$ 1
YFP control	n.i.	n.i.	n.i.

**Table 3.** Kinetic parameters of the interaction between wild-type influenza nucleoprotein tails and human importins- $\alpha$  obtained by SPR. The association rate constants ( $k_{on}$ ) are expressed in M<sup>-1</sup>·s<sup>-1</sup> and the dissociation rate constants ( $k_{off}$ ) values are expressed in s<sup>-1</sup>; n.i. means ‘no interaction’.

	Imp- $\alpha$ 3		Imp- $\alpha$ 5		Imp- $\alpha$ 7	
	$k_{on}$ ( $\times 10^6$ )	$k_{off}$ ( $\times 10^{-1}$ )	$k_{on}$ ( $\times 10^6$ )	$k_{off}$ ( $\times 10^{-1}$ )	$k_{on}$ ( $\times 10^6$ )	$k_{off}$ ( $\times 10^{-1}$ )
A/NP <sub>TAIL</sub>	1.31 $\pm$ 0.03	2.81 $\pm$ 0.04	1.0 $\pm$ 0.6	1.6 $\pm$ 0.1	1.6 $\pm$ 0.1	0.7 $\pm$ 0.1
B/NP <sub>TAIL</sub>	>10	>10	>10	>10	>10	>10
C/NP <sub>TAIL</sub>	2.4 $\pm$ 1.1	1.3 $\pm$ 0.4	6.1 $\pm$ 3.6	2.6 $\pm$ 0.4	2.6 $\pm$ 0.4	2.5 $\pm$ 0.1
D/NP <sub>TAIL</sub>	2.8 $\pm$ 0.2	0.07 $\pm$ 0.01	1.1 $\pm$ 0.1	5.0 $\pm$ 0.2	5.0 $\pm$ 0.2	1.04 $\pm$ 0.03
YFP control	n.i.	n.i.	n.i.	n.i.	n.i.	n.i.

### 3.3. Systematic Analysis of the Interactions between NP<sub>TAILs</sub> and Importins- $\alpha$

Once the strategy for measuring the kinetics of all NP<sub>TAILs</sub> for the importin- $\alpha$ 7 was established, we decided to expand it to other members of the importins- $\alpha$  family. The human genome encodes seven isoforms of importin- $\alpha$ , divided into three subfamilies known as  $\alpha$ 1,  $\alpha$ 2 and  $\alpha$ 3: the  $\alpha$ 1 subfamily contains importin- $\alpha$ 1 and - $\alpha$ 8; the  $\alpha$ 2 subfamily contains importin- $\alpha$ 3 and - $\alpha$ 4; and the  $\alpha$ 3 subfamily contains importin- $\alpha$ 5, - $\alpha$ 6 and - $\alpha$ 7 [18,49]. We chose to work with one representative isoform of each subfamily (importin- $\alpha$ 1 and importin- $\alpha$ 3, respectively for the  $\alpha$ 1 and  $\alpha$ 2 subfamilies) plus a second isoform for the  $\alpha$ 3 subfamily (i.e., importin- $\alpha$ 5). Furthermore, all these isoforms are described in the literature regarding the influenza virus infection. We produced and purified the recombinant isoforms without their importin- $\beta$  binding (IBB) domain. The case of importin- $\alpha$ 1 is shown below. Using SEC-MALLS-RI experiments, we first confirmed that both  $\Delta$ IBB importins- $\alpha$ 3 and - $\alpha$ 5 are monomeric in solution (Figure S4), before we used them in our SPR experiments to determine their respective kinetics (Tables 2 and 3). Importin- $\alpha$ 5 shows a similar interaction profile as importin- $\alpha$ 7, with  $K_d$  D/NP<sub>TAIL</sub> < A/NP<sub>TAIL</sub> < C/NP<sub>TAIL</sub> < B/NP<sub>TAIL</sub> but with a higher amplitude in the values (from 21 to 770 nM). On the contrary, the architecture of the NLS seems essential for recognition by importin- $\alpha$ 3. The two bipartite NLS tails (i.e., C/NP<sub>TAIL</sub> and D/NP<sub>TAIL</sub>) have a higher affinity than the two monopartite (i.e., A/NP<sub>TAIL</sub> and B/NP<sub>TAIL</sub>) which both show identical values. Based on these results, we further deciphered the interactions by introducing point mutations in the positive charge patches of the NLSs which we had analyzed in the past (i.e., B/NP<sub>TAIL</sub> and D/NP<sub>TAIL</sub>) [37,38]. We show that for bipartite NLS, the second basic patch could be stronger in relation to the stability of the cargo: importin- $\alpha$  than the first patch (Table 4). Based on the affinities of each single patch mutants, D/NP<sub>TAIL</sub> seems to present a cooperative behaviour in importin- $\alpha$  binding. On the other hand, importin- $\alpha$ 7 seems to be more prompt in compensating for mutations in the first patch whereas importin- $\alpha$ 3 looks more permissive to mutations in the second. The results obtained with B/NP<sub>TAIL</sub> mutant confirm our nuclear magnetic resonance (NMR) observations on the major contribution of the whole tail, from residues 30 to 71, in the interaction with importins- $\alpha$ , and not only the regions surrounding the KR motif [38]. Regarding the B/NP<sub>TAIL</sub> mut1-YFP, even if we were not able to quantify precisely the equilibrium dissociation constants from the raw data, probably because importin- $\alpha$  require a longer injection period to saturate all the ligand binding sites in this condition, the sensorgrams show clear concentration dependent-responses in contrast to the YFP-D/NP<sub>TAIL</sub> mut3 construct (Figure S5).

**Table 4.** Effects of mutations in B/NP<sub>TAIL</sub> and D/NP<sub>TAIL</sub> on the affinities for the human importins- $\alpha$ . The  $K_d$  obtained by SPR are expressed in nM; n.i. means ‘no interaction’.

	Imp- $\alpha$ 3	Imp- $\alpha$ 5	Imp- $\alpha$ 7
B/NP <sub>TAIL</sub>	308 ± 106	770 ± 24	405 ± 5
B/NP <sub>TAIL</sub> mut1	>5000	>5000	>5000
D/NP <sub>TAIL</sub>	22 ± 15	21 ± 5	29 ± 1
D/NP <sub>TAIL</sub> mut1	922 ± 77	1150 ± 154	658 ± 208
D/NP <sub>TAIL</sub> mut2	2457 ± 184	>5000	>5000
D/NP <sub>TAIL</sub> mut3	n.i.	n.i.	n.i.

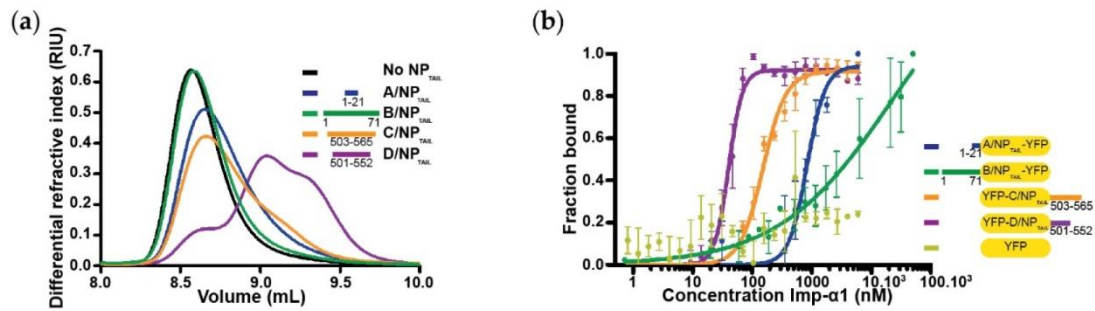
  

### 3.4. The Case of the Importin- $\alpha$ 1

The SEC-MALLS-RI experiments showed a different behaviour of importin- $\alpha$ 1 in solution (Figure S4). Whereas the experimental molecular mass for  $\Delta$ IBB importins- $\alpha$ 3, - $\alpha$ 5 and - $\alpha$ 7 were similar to the theoretical values, the results obtained with  $\Delta$ IBB importin- $\alpha$ 1 suggest a propensity



for dimerization of this construct. The corresponding sample was eluted precociously from the S200 column with an apparent molecular mass of 88 kDa rather than the theoretical 50 kDa (Figure S4a). Dimerization was reported previously [50] with a monomer–dimer  $K_d$  estimated to be 8  $\mu\text{M}$  and the equilibrium could be shifted to the monomeric form by adding NLS-mimicking peptides. In a second SEC-MALLS-RI set of experiments but using a Superdex 75 column, we showed that with the different NP<sub>TAIL</sub> peptides the peaks corresponding to the importin- $\alpha$ 1 were shifted to the smaller molecular mass (Figure 5a). The largest effect is observed with D/NP<sub>TAIL</sub> suggesting that a quasi-quantitative estimation of the interaction could be extracted from these results, with apparent  $K_d$  D/NP<sub>TAIL</sub> < C/NP<sub>TAIL</sub> < A/NP<sub>TAIL</sub> < B/NP<sub>TAIL</sub>.



**Figure 5.** Binding of influenza nucleoproteins tails to importin- $\alpha$ 1. (a) Dissociation of importin- $\alpha$ 1 dimers by competition with NP<sub>TAIL</sub> using a SEC-MALLS-RI experiments. Human importin- $\alpha$ 1 (45  $\mu\text{M}$ ) alone or incubated with each NP<sub>TAIL</sub> (45  $\mu\text{M}$ ) was injected on a Superdex<sup>TM</sup> 75 increase 10/300 G. Successful competition resulted in a delayed elution volume compared to the importin- $\alpha$ 1 alone. Three distinct species can be identified from importin- $\alpha$ 1:D/NP<sub>TAIL</sub> injection and two species from importin- $\alpha$ 1:C/NP<sub>TAIL</sub> one. (b) YFP-NP<sub>TAIL</sub>s: importin- $\alpha$ 1 affinities by fluorescence anisotropy. Titration with each NP<sub>TAIL</sub> (25 nM) was carried out by fluorescence anisotropy with importin- $\alpha$ 1 concentrations ranging from 0.7 nM to 50  $\mu\text{M}$  for B/NP<sub>TAIL</sub>-YFP or from 0.8 nM to 6  $\mu\text{M}$  for the other constructs. The titrations were normalized and fitted with GraphPad analysis model “specific binding with Hill slope”. The raw data are shown in Figure S6.

When we tried to measure the corresponding kinetics by SPR, we could not obtain a stable equilibrium, probably due to the competition between the monomer-dimer transition of importin- $\alpha$ 1 and its interaction with the immobilized tails. This experimental artifact is inherent to the nature of the analyte, which despite the purification steps may dimerize in time and significantly affect affinity measurements by SPR. We then found an alternative method; with the YFP-NP<sub>TAIL</sub>s we could measure the fluorescence anisotropy. Fluorescence anisotropy is a measure of light emitted by a fluorophore, unequal along different axes of polarization. This measure depends on rotational behaviour of this fluorescent molecule in solution. Depending on the fluorophore environment, fluorescence anisotropy will vary from a low value because of its high rotational rate when the molecule is free, to a high value in case of its interaction with a partner. This feature allows the determination of an apparent binding constant when the fluorescent molecule is kept at a constant concentration upon titration of its partner. By titration of our  $\Delta\text{IBB}$  importin- $\alpha$ 1 construct in presence of the different YFP-NP<sub>TAIL</sub>s (Figure 5b), the corresponding apparent  $K_d$  was determined (Table 5). The curve of the B/NP<sub>TAIL</sub> could not be estimated because the  $K_d$  of the B/NP<sub>TAIL</sub> for importin- $\alpha$ 1 was too close to the  $K_d$  of the dimerization of importin- $\alpha$ 1.

**Table 5.** Affinity between wild-type influenza nucleoprotein tails and human importin- $\alpha$ 1 obtained by fluorescence anisotropy. The apparent  $K_d$  are expressed in nM; n.d. and n.i. mean ‘not determined’ and ‘no interaction’, respectively.

	A/NP <sub>TAIL</sub>	B/NP <sub>TAIL</sub>	C/NP <sub>TAIL</sub>	D/NP <sub>TAIL</sub>	YFP Control
imp- $\alpha$ 1	621 $\pm$ 89	n.d.	149 $\pm$ 1	41 $\pm$ 12	n.i.

#### 4. Discussion

The interaction between viral proteins and cellular factors is of growing interest for understanding the evolution mechanisms of viruses, and RNA viruses in particular. These are major potential therapeutic targets, as these interfaces are less prone to mutations. It has been decades since the interactions between importin- $\alpha$  and the influenza virus nucleoprotein have been studied but numerous features and mechanisms are still unknown [32,33,35,51]. Importins- $\alpha$  are essential actors in cells and were diversified and highly conserved through evolution (Table S2). Each importin- $\alpha$  shows a high sequence identity in higher eukaryotes with a specific expression pattern and behaviour but also preferential bindings to cargo and functions.

In this paper, we provide a comparative characterization of the interactions between four importin- $\alpha$  isoforms with the tails from A/NP, B/NP, C/NP and D/NP. The affinities are between the nanomolar and the low micromolar range, with association and dissociation rates surprisingly high. Two different behaviours were identified: (i) A/NP<sub>TAIL</sub>, C/NP<sub>TAIL</sub> and D/NP<sub>TAIL</sub> present a dynamic complex formation and release with high-affinity preferential binding and (ii) B/NP<sub>TAIL</sub> shows even faster association and dissociation rates with very low affinity.

The A/NP<sub>TAIL</sub> was isolated from a human strain of influenza A virus and interacted preferentially with importin- $\alpha$ 7. This is consistent with the importins- $\alpha$  expression pattern in the human respiratory track [52] as well as of previously reported work [31]. It is of note that PB2 isolated from human strains also preferentially bind to importin- $\alpha$ 7, while avian PB2s tend to interact with importin- $\alpha$ 3 [28]. The poor binding of A/NP<sub>TAIL</sub> to importin- $\alpha$ 3 is consistent with the fact that importin- $\alpha$ 3 tends not to interact well with monopartite NLS; while A/NP<sub>TAIL</sub> is not a monopartite NLS *per se*, it possesses only a short NLS motif. In contrast, C/NP<sub>TAIL</sub> presents a higher affinity for importin- $\alpha$ 3. Importin- $\alpha$ 3 has a low auto-inhibition and a unique intrinsic flexibility, acting as an adapter for unusual and complex NLS sequences or NLS close to globular domains [53,54]. Such NLSs could not associate with a less flexible importin- $\alpha$  due to steric hindrance. The NLS sequence of C/NP<sub>TAIL</sub> is similar to a bipartite NLS, with two basic patches, but separated by a long linker sequence (more than 20 residues). Such a sequence could require the flexibility of importin- $\alpha$ 3 to achieve optimal binding. D/NP<sub>TAIL</sub> is a textbook case of bipartite NLS, which binds tightly to importins- $\alpha$ , using both the minor and major binding pockets [55–57]. D/NP<sub>TAIL</sub> strongly interacts with all tested importins with a high affinity (20–41 nM), in the same range as other bipartite NLSs ( $K_d \approx 10$  nM) [28,58]. In other studies, it has been shown that the *in vitro* affinities using purified proteins and peptides were in the nanomolar range whereas the interaction between NLSs and importins- $\alpha$  *in cellulo* is in the micromolar range. Addition of cytosol in *in vitro* assays led to the destabilization of importin: NLS complexes, even in the case of bacterial lysate, highlighting the contribution of non-specific interaction and competition in this cellular process [44,59]. Importins cycles between both compartments have been estimated to be fast, around 30 ms; nanomolar affinity leads to a residence time of several seconds or up to several minutes, which is far too long to allow optimal release in the nucleus, while micromolar affinity put it in the millisecond range [59].

The sensorgram shapes of C/NP<sub>TAIL</sub> and D/NP<sub>TAIL</sub> along with the affinities resulting from D/NP<sub>TAIL</sub> mutations clearly suggest that the interaction follows a 1:1 binding model with cooperativity. At the other end of the spectrum, B/NP<sub>TAIL</sub> behaviour towards importins- $\alpha$  is still puzzling. It is possible that the B/NP<sub>TAIL</sub> is not properly processed in our settings. It is well known that A/NP and B/NP nuclear import are regulated through post-translational modifications (PTM), in particular

through phosphorylation [60,61]. A discrete PTM could be an activating signal for nuclear import regulation [62–66]. The absence of the NP<sub>CORE</sub> in our experiment could also be responsible for the lack of functionality of B/NP<sub>TAIL</sub>: three-dimensional context could matter for importin specificity, with putative additional contacts and constrains, and for NLS docking [18,28,53,67]. Another hypothesis comes from the nuclear import receptor involved and the complex to be imported. While importin- $\alpha/\beta$  complex is largely believed to be the most common nuclear import pathway, this is still debated. Other adaptor proteins besides importin- $\alpha$  are known to interact with cargo and importin- $\beta$  [68,69]. Co-import of neo-synthesized B/NP with another cargo is also a possibility, taking in account that RNP nuclear import functionality putatively relies on avidity rather than affinity [70,71]. Therefore, it is not be excluded that B/NP uses another receptor/complex to gain access to the nucleus. Further investigations into other nuclear import receptors could shed light on successful B/NP nuclear translocation. One interesting feature of B/NP is its long, intrinsically disordered N<sub>terminal</sub> tail. Such a tail could be involved in the formation of liquid phase separation and help with protein recruitment [72,73]. If B/NP exhibits such behaviour, it could be interesting to dissect it further as nucleoporins in the nuclear pore complex channel tend themselves to be involved in phase separation [74–76].

Our work focuses on human importins- $\alpha$ , as three out of the four NP<sub>TAILS</sub> were derived from human viruses. However, due to the high conservation of these importins through evolution (Table S2), it is likely that extrapolation to other mammal importins can be achieved to some degree. Further work would be required to validate this completely.

While a lot of work is still required to fully understand the mechanistic and specificity of the interactions between influenza nucleoproteins and importins- $\alpha$ , we provide here this first comparative study on the interaction of NP<sub>TAILS</sub> with importins- $\alpha$ 1, - $\alpha$ 3, - $\alpha$ 5 and - $\alpha$ 7. NP<sub>CORE</sub> involvement, NP oligomerization, cellular environment, IBB domain contribution, and post-translational modification of both partners are other factors which need to be investigated in the future, in the hope of fully dissecting this host-pathogen interaction and trying to target and disrupt this essential hijacking mechanism.

## 5. Conclusions

While the interactions between influenza NPs and importins- $\alpha$  have been extensively studied, this has mainly been done through the prism of A/NP, with a large number of highly different methods and sometimes without the essential focus required on importins- $\alpha$ . Whereas each method and study brought useful characterization insights, it is a complex task to compare affinity from widely different assays, with significant changes in conditions, partners, settings and analysis. Here, we provide the first comparative study to dissect several interactions between influenza NPs and importins- $\alpha$  using the same strategy and under the same conditions. We showed that A/NP<sub>TAIL</sub>, C/NP<sub>TAIL</sub> and D/NP<sub>TAIL</sub> interact strongly with one or several importins- $\alpha$ . On the contrary, the human specific B/NP<sub>TAIL</sub> interacts poorly with selected importins- $\alpha$ , raising new questions about this essential interaction in the influenza virus cycle.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/1999-4915/12/8/834/s1>, Figure S1: analysis of the complexes between NP<sub>TAILS</sub> and human importin- $\alpha$ 7 in solution, Figure S2: SEC-MALLS-RI analysis of the different YFP-fused tails used in that work, Figure S3: absence of interaction between YFP and human importin- $\alpha$ 7 in solution, Figure S4: SEC-MALLS-RI analysis of the different importins- $\alpha$  used in that work, Figure S5: interaction between human importin- $\alpha$ 7 and YFP-NP<sub>TAILS</sub> mutants, Figure S6: interaction between human importin- $\alpha$ 1 and YFP-NP<sub>TAILS</sub>, Table S1: sequences, Table S2: sequence identity between the importins- $\alpha$  used in the studies.

**Author Contributions:** T.C. conceived and designed the experiments; A.D., E.V.-S. and F.C.A.G. performed the experiments; A.D., E.V.-S., F.C.A.G. and T.C. analysed the data; A.D. and T.C. wrote the original draft; A.D., E.V.-S., F.C.A.G., R.W.H.R. and T.C. edited the final manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the Agence Nationale pour la Recherche, grant number ANR-14-CE09-0017.

**Acknowledgments:** We thank Justine Sensini and Louis Saliceti for their contribution in the samples preparation. We are highly grateful to Jean-Marie Bourhis, Pascal Fender, Corinne Mercier, Carlo Petosa and Antoine Royant for discussions. We thank Darren J Hart for the plasmid to express the human importin- $\alpha$ 7. This work used the

platforms of the Grenoble Instruct-ERIC center (ISBG; UMS 3518 CNRS-CEA-UGA-EMBL) within the Grenoble Partnership for Structural Biology (PSB), supported by FRISBI (ANR-10-INBS-05-02) and GRAL, financed by the University Grenoble Alpes - Ecoles Universitaires de Recherche CBH-EUR-GS (ANR-17-EURE-0003).

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

## References

- Shi, M.; Lin, X.D.; Chen, X.; Tian, J.H.; Chen, L.J.; Li, K.; Wang, W.; Eden, J.S.; Shen, J.J.; Liu, L.; et al. The evolutionary history of vertebrate RNA viruses. *Nature* **2018**, *556*, 197–202. [[CrossRef](#)] [[PubMed](#)]
- Klump, K.; Ruigrok, R.W.; Baudin, F. Roles of the Influenza virus polymerase and nucleoprotein in forming a functional RNP structure. *EMBO J.* **1997**, *16*, 1248–1257. [[CrossRef](#)] [[PubMed](#)]
- Coloma, R.; Arranz, R.; de la Rosa-Trevin, J.M.; Sorzano, C.O.S.; Munier, S.; Carlero, D.; Naffakh, N.; Ortin, J.; Martin-Benito, J. Structural insights into Influenza A virus ribonucleoproteins reveal a processive helical track as transcription mechanism. *Nat. Microbiol.* **2020**, *5*, 727–734. [[CrossRef](#)] [[PubMed](#)]
- Feldherr, C.M.; Kallenbach, E.; Schultz, N. Movement of a karyophilic protein through the nuclear pores of oocytes. *J. Cell Biol.* **1984**, *99*, 2216–2222. [[CrossRef](#)] [[PubMed](#)]
- Davis, L.I. The nuclear pore complex. *Annu. Rev. Biochem.* **1995**, *64*, 865–896. [[CrossRef](#)]
- Mans, B.J.; Anantharaman, V.; Aravind, L.; Koonin, E.V. Comparative genomics, evolution and origins of the nuclear envelope and nuclear pore complex. *Cell Cycle* **2004**, *3*, 1612–1637. [[CrossRef](#)]
- Lin, D.H.; Hoelz, A. The structure of the nuclear pore complex (an update). *Annu. Rev. Biochem.* **2019**, *88*, 725–783. [[CrossRef](#)]
- Bonner, W.M. Protein migration into nuclei. II. Frog oocyte nuclei accumulate a class of microinjected oocyte nuclear proteins and exclude a class of microinjected oocyte cytoplasmic proteins. *J. Cell Biol.* **1975**, *64*, 431–437. [[CrossRef](#)]
- Chook, Y.M.; Suel, K.E. Nuclear import by karyopherin-betas: Recognition and inhibition. *Biochim. Biophys. Acta* **2011**, *1813*, 1593–1606. [[CrossRef](#)]
- Timney, B.L.; Raveh, B.; Mironska, R.; Trivedi, J.M.; Kim, S.J.; Russel, D.; Wenthe, S.R.; Sali, A.; Rout, M.P. Simple rules for passive diffusion through the nuclear pore complex. *J. Cell Biol.* **2016**, *215*, 57–76. [[CrossRef](#)]
- Gorlich, D.; Vogel, F.; Mills, A.D.; Hartmann, E.; Laskey, R.A. Distinct functions for the two importin subunits in nuclear protein import. *Nature* **1995**, *377*, 246–248. [[CrossRef](#)] [[PubMed](#)]
- Chook, Y.M.; Blobel, G. Karyopherins and nuclear import. *Curr. Opin. Struct. Biol.* **2001**, *11*, 703–715. [[CrossRef](#)]
- Lange, A.; Mills, R.E.; Lange, C.J.; Stewart, M.; Devine, S.E.; Corbett, A.H. Classical nuclear localization signals: Definition, function, and interaction with importin alpha. *J. Biol. Chem.* **2007**, *282*, 5101–5105. [[CrossRef](#)] [[PubMed](#)]
- Lee, B.J.; Cansizoglu, A.E.; Suel, K.E.; Louis, T.H.; Zhang, Z.; Chook, Y.M. Rules for nuclear localization sequence recognition by karyopherin beta 2. *Cell* **2006**, *126*, 543–558. [[CrossRef](#)] [[PubMed](#)]
- Bono, F.; Cook, A.G.; Grunwald, M.; Ebert, J.; Conti, E. Nuclear import mechanism of the EJC component Mago-Y14 revealed by structural studies of importin 13. *Mol. Cell* **2010**, *37*, 211–222. [[CrossRef](#)]
- Padavannil, A.; Sarkar, P.; Kim, S.J.; Cagatay, T.; Jiou, J.; Brautigam, C.A.; Tomchick, D.R.; Sali, A.; D’Arcy, S.; Chook, Y.M. Importin-9 wraps around the H2A-H2B core to act as nuclear importer and histone chaperone. *eLife* **2019**, *8*. [[CrossRef](#)] [[PubMed](#)]
- Swale, C.; Da Costa, B.; Sedano, L.; Garzoni, F.; McCarthy, A.A.; Berger, I.; Bieniossek, C.; Ruigrok, R.W.H.; Delmas, B.; Crepin, T. X-ray structure of the human karyopherin RanBP5, an essential factor for Influenza polymerase nuclear trafficking. *J. Mol. Biol.* **2020**, *432*, 3353–3359. [[CrossRef](#)]
- Pumroy, R.A.; Cingolani, G. Diversification of importin-alpha isoforms in cellular trafficking and disease states. *Biochem. J.* **2015**, *466*, 13–28. [[CrossRef](#)]
- Oka, M.; Yoneda, Y. Importin alpha: Functions as a nuclear transport factor and beyond. *Proc. Jpn. Acad. Ser. B Phys. Biol. Sci.* **2018**, *94*, 259–274. [[CrossRef](#)]
- Martin, K.; Helenius, A. Transport of incoming Influenza virus nucleocapsids into the nucleus. *J. Virol.* **1991**, *65*, 232–244. [[CrossRef](#)]
- Neumann, G.; Castrucci, M.R.; Kawaoka, Y. Nuclear import and export of Influenza virus nucleoprotein. *J. Virol.* **1997**, *71*, 9690–9700. [[CrossRef](#)] [[PubMed](#)]

22. Deng, T.; Engelhardt, O.G.; Thomas, B.; Akoulitchev, A.V.; Brownlee, G.G.; Fodor, E. Role of ran binding protein 5 in nuclear import and assembly of the Influenza virus RNA polymerase complex. *J. Virol.* **2006**, *80*, 11911–11919. [[CrossRef](#)] [[PubMed](#)]
23. 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.* **2011**, *92*, 1859–1869. [[CrossRef](#)] [[PubMed](#)]
24. Swale, C.; Monod, A.; Tengo, L.; Labaronne, A.; Garzoni, F.; Bourhis, J.M.; Cusack, S.; Schoehn, G.; Berger, I.; Ruigrok, R.W.; et al. Structural characterization of recombinant IAV polymerase reveals a stable complex between viral PA-PB1 heterodimer and host RanBP5. *Sci. Rep.* **2016**, *6*, 24727. [[CrossRef](#)] [[PubMed](#)]
25. Tarendeau, F.; Boudet, J.; Guilligay, D.; Mas, P.; Bougault, C.; Boulo, S.; Baudin, F.; Ruigrok, R.W.H.; Daigle, N.; Ellenberg, J.; et al. Structure and nuclear import function of the C-terminal domain of Influenza virus polymerase PB2 subunit. *Nat. Struct. Mol. Biol.* **2007**, *14*, 229–233. [[CrossRef](#)] [[PubMed](#)]
26. Boivin, S.; Hart, D.J. Interaction of the Influenza A virus polymerase PB2 C-terminal region with importin alpha isoforms provides insights into host adaptation and polymerase assembly. *J. Biol. Chem.* **2011**, *286*, 10439–10448. [[CrossRef](#)] [[PubMed](#)]
27. Hudjetz, B.; Gabriel, G. Human-like PB2 627K Influenza virus polymerase activity is regulated by importin-alpha1 and -alpha7. *PLoS Pathog.* **2012**, *8*, e1002488. [[CrossRef](#)]
28. Pumroy, R.A.; Ke, S.; Hart, D.J.; Zachariae, U.; Cingolani, G. Molecular determinants for nuclear import of Influenza A PB2 by importin alpha isoforms 3 and 7. *Structure* **2015**, *23*, 374–384. [[CrossRef](#)]
29. Cros, J.F.; García-Sastre, A.; Palese, P. An unconventional NLS is critical for the nuclear import of the Influenza A virus nucleoprotein and ribonucleoprotein. *Traffic* **2005**, *6*, 205–213. [[CrossRef](#)]
30. Weber, F.; Kochs, G.; Gruber, S.; Haller, O. A classical bipartite nuclear localization signal on Thogoto and Influenza A virus nucleoproteins. *Virology* **1998**, *250*, 9–18. [[CrossRef](#)]
31. Wu, W.; Sankhala, R.S.; Florio, T.J.; Zhou, L.; Nguyen, N.L.T.; Lokareddy, R.K.; Cingolani, G.; Pante, N. Synergy of two low-affinity NLSs determines the high avidity of Influenza A virus nucleoprotein NP for human importin alpha isoforms. *Sci. Rep.* **2017**, *7*, 11381. [[CrossRef](#)] [[PubMed](#)]
32. O'Neill, R.E.; Palese, P. NPI-1, the human homolog of SRP-1, interacts with Influenza virus nucleoprotein. *Virology* **1995**, *206*, 116–125. [[CrossRef](#)]
33. Wang, P.; Palese, P.; O'Neill, R.E. The NPI-1/NPI-3 (karyopherin alpha) binding site on the Influenza A virus nucleoprotein NP is a nonconventional nuclear localization signal. *J. Virol.* **1997**, *71*, 1850–1856. [[CrossRef](#)] [[PubMed](#)]
34. Melen, K.; Fagerlund, R.; Franke, J.; Kohler, M.; Kinnunen, L.; Julkunen, I. Importin alpha nuclear localization signal binding sites for STAT1, STAT2, and Influenza A virus nucleoprotein. *J. Biol. Chem.* **2003**, *278*, 28193–28200. [[CrossRef](#)] [[PubMed](#)]
35. Gabriel, G.; Klingel, K.; Otte, A.; Thiele, S.; Hudjetz, B.; Arman-Kalcek, G.; Sauter, M.; Schmidt, T.; Rother, F.; Baumgarte, S.; et al. Differential use of importin-alpha isoforms governs cell tropism and host adaptation of Influenza virus. *Nat. Commun.* **2011**, *2*, 156. [[CrossRef](#)]
36. Sasaki, Y.; Hagiwara, K.; Kakisaka, M.; Yamada, K.; Murakami, T.; Aida, Y. Importin alpha3/Qip1 is involved in multiplication of mutant Influenza virus with alanine mutation at amino acid 9 independently of nuclear transport function. *PLoS ONE* **2013**, *8*, e55765. [[CrossRef](#)]
37. Donchet, A.; Oliva, J.; Labaronne, A.; Tengo, L.; Miloudi, M.; Gerard, F.C.A.; Mas, C.; Schoehn, G.; Ruigrok, R.W.H.; Ducatez, M.; et al. The structure of the nucleoprotein of Influenza D shows that all *Orthomyxoviridae* nucleoproteins have a similar NP<sub>CORE</sub>, with or without a NP<sub>TAIL</sub> for nuclear transport. *Sci. Rep.* **2019**, *9*, 600. [[CrossRef](#)]
38. Labaronne, A.; Milles, S.; Donchet, A.; Jensen, M.R.; Blackledge, M.; Bourhis, J.M.; Ruigrok, R.W.H.; Crepin, T. Structural analysis of the complex between Influenza B nucleoprotein and human importin-alpha. *Sci. Rep.* **2017**, *7*, 17164. [[CrossRef](#)]
39. Ozawa, M.; Fujii, K.; Muramoto, Y.; Yamada, S.; Yamayoshi, S.; Takada, A.; Goto, H.; Horimoto, T.; Kawaoka, Y. Contributions of two nuclear localization signals of Influenza A virus nucleoprotein to viral replication. *J. Virol.* **2007**, *81*, 30–41. [[CrossRef](#)]
40. 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](#)]

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. Nakada, R.; Hirano, H.; Matsuura, Y. Structure of importin-alpha bound to a non-classical nuclear localization signal of the Influenza A virus nucleoprotein. *Sci. Rep.* **2015**, *5*, 15055. [[CrossRef](#)] [[PubMed](#)]
43. Tang, Y.S.; Lo, C.Y.; Mok, C.K.; Chan, P.K.; Shaw, P.C. The extended C-terminal region of Influenza C virus nucleoprotein is important for nuclear import and ribonucleoprotein activity. *J. Virol.* **2019**, *93*, e02048-18. [[CrossRef](#)] [[PubMed](#)]
44. Timney, B.L.; Tetenbaum-Novatt, J.; Agate, D.S.; Williams, R.; Zhang, W.; Chait, B.T.; Rout, M.P. Simple kinetic relationships and nonspecific competition govern nuclear import rates *in vivo*. *J. Cell Biol.* **2006**, *175*, 579–593. [[CrossRef](#)] [[PubMed](#)]
45. Mackmull, M.T.; Klaus, B.; Heinze, I.; Chokkalingam, M.; Beyer, A.; Russell, R.B.; Ori, A.; Beck, M. Landscape of nuclear transport receptor cargo specificity. *Mol. Syst. Biol.* **2017**, *13*, 962. [[CrossRef](#)]
46. Sumarheni, S.; Hong, S.S.; Josserand, V.; Coll, J.L.; Boulanger, P.; Schoehn, G.; Fender, P. Human full-length coagulation factor X and a GLA domain-derived 40-mer polypeptide bind to different regions of the adenovirus serotype 5 hexon capsomer. *Hum. Gene Ther.* **2014**, *25*, 339–349. [[CrossRef](#)]
47. Wyatt, P.J. Submicrometer particle sizing by multiangle light scattering following fractionation. *J. Colloid Interface Sci.* **1998**, *197*, 9–20. [[CrossRef](#)]
48. Labaronne, A.; Swale, C.; Monod, A.; Schoehn, G.; Crepin, T.; Ruigrok, R.W. Binding of RNA by the nucleoproteins of Influenza viruses A and B. *Viruses* **2016**, *8*, 247. [[CrossRef](#)]
49. Mason, D.A.; Stage, D.E.; Goldfarb, D.S. Evolution of the metazoan-specific importin alpha gene family. *J. Mol. Evol.* **2009**, *68*, 351–365. [[CrossRef](#)]
50. Miyatake, H.; Sanjoh, A.; Unzai, S.; Matsuda, G.; Tatsumi, Y.; Miyamoto, Y.; Dohmae, N.; Aida, Y. Crystal structure of human importin-alpha1 (Rch1), revealing a potential autoinhibition mode involving homodimerization. *PLoS ONE* **2015**, *10*, e0115995. [[CrossRef](#)]
51. Resa-Infante, P.; Paterson, D.; Bonet, J.; Otte, A.; Oliva, B.; Fodor, E.; Gabriel, G. Targeting importin-alpha7 as a therapeutic approach against pandemic Influenza viruses. *J. Virol.* **2015**, *89*, 9010–9020. [[CrossRef](#)] [[PubMed](#)]
52. Ninpan, K.; Suptawiwat, O.; Boonarkart, C.; Phuangphung, P.; Sathirareuangchai, S.; Uprasertkul, M.; Auewarakul, P. Expression of importin-alpha isoforms in human nasal mucosa: Implication for adaptation of avian Influenza A viruses to human host. *Virol. J.* **2016**, *13*, 90. [[CrossRef](#)] [[PubMed](#)]
53. Sankhala, R.S.; Lokareddy, R.K.; Begum, S.; Pumroy, R.A.; Gillilan, R.E.; Cingolani, G. Three-dimensional context rather than NLS amino acid sequence determines importin alpha subtype specificity for RCC1. *Nat. Commun.* **2017**, *8*, 979. [[CrossRef](#)] [[PubMed](#)]
54. Smith, K.M.; Tsimbalyuk, S.; Edwards, M.R.; Cross, E.M.; Batra, J.; Soares da Costa, T.P.; Aragao, D.; Basler, C.F.; Forwood, J.K. Structural basis for importin alpha 3 specificity of W proteins in Hendra and Nipah viruses. *Nat. Commun.* **2018**, *9*, 3703. [[CrossRef](#)] [[PubMed](#)]
55. Robbins, J.; Dilworth, S.M.; Laskey, R.A.; Dingwall, C. Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: Identification of a class of bipartite nuclear targeting sequence. *Cell* **1991**, *64*, 615–623. [[CrossRef](#)]
56. Fontes, M.R.; Teh, T.; Kobe, B. Structural basis of recognition of monopartite and bipartite nuclear localization sequences by mammalian importin-alpha. *J. Mol. Biol.* **2000**, *297*, 1183–1194. [[CrossRef](#)]
57. Hodel, M.R.; Corbett, A.H.; Hodel, A.E. Dissection of a nuclear localization signal. *J. Biol. Chem.* **2001**, *276*, 1317–1325. [[CrossRef](#)]
58. Lokareddy, R.K.; Hapsari, R.A.; van Rheenen, M.; Pumroy, R.A.; Bhardwaj, A.; Steen, A.; Veenhoff, L.M.; Cingolani, G. Distinctive properties of the nuclear localization signals of inner nuclear membrane proteins Heh1 and Heh2. *Structure* **2015**, *23*, 1305–1316. [[CrossRef](#)]
59. Cardarelli, F.; Bizzarri, R.; Serresi, M.; Albertazzi, L.; Beltram, F. Probing nuclear localization signal-importin alpha binding equilibria in living cells. *J. Biol. Chem.* **2009**, *284*, 36638–36646. [[CrossRef](#)]
60. Hutchinson, E.C.; Denham, E.M.; Thomas, B.; Trudgian, D.C.; Hester, S.S.; Ridlova, G.; York, A.; Turrell, L.; Fodor, E. Mapping the phosphoproteome of Influenza a and B viruses by mass spectrometry. *PLoS Pathog.* **2012**, *8*, e1002993. [[CrossRef](#)]

61. Zheng, W.; Li, J.; Wang, S.; Cao, S.; Jiang, J.; Chen, C.; Ding, C.; Qin, C.; Ye, X.; Gao, G.F.; et al. Phosphorylation controls the nuclear-cytoplasmic shuttling of Influenza A virus nucleoprotein. *J. Virol.* **2015**, *89*, 5822–5834. [[CrossRef](#)] [[PubMed](#)]
62. Terry, L.J.; Shows, E.B.; Wentz, S.R. Crossing the nuclear envelope: Hierarchical regulation of nucleocytoplasmic transport. *Science* **2007**, *318*, 1412–1416. [[CrossRef](#)] [[PubMed](#)]
63. Bedford, M.T.; Clarke, S.G. Protein arginine methylation in mammals: Who, what, and why. *Mol. Cell* **2009**, *33*, 1–13. [[CrossRef](#)] [[PubMed](#)]
64. Nardozi, J.; Wenta, N.; Yasuhara, N.; Vinkemeier, U.; Cingolani, G. Molecular basis for the recognition of phosphorylated STAT1 by importin alpha5. *J. Mol. Biol.* **2010**, *402*, 83–100. [[CrossRef](#)] [[PubMed](#)]
65. Wang, Y.E.; Pernet, O.; Lee, B. Regulation of the nucleocytoplasmic trafficking of viral and cellular proteins by ubiquitin and small ubiquitin-related modifiers. *Biol. Cell.* **2012**, *104*, 121–138. [[CrossRef](#)] [[PubMed](#)]
66. Ptak, C.; Wozniak, R.W. SUMO and nucleocytoplasmic transport. *Adv. Exp. Med. Biol.* **2017**, *963*, 111–126.
67. Friedrich, B.; Quensel, C.; Sommer, T.; Hartmann, E.; Kohler, M. Nuclear localization signal and protein context both mediate importin alpha specificity of nuclear import substrates. *Mol. Cell. Biol.* **2006**, *26*, 8697–8709. [[CrossRef](#)]
68. Jullien, D.; Gorlich, D.; Laemmli, U.K.; Adachi, Y. Nuclear import of RPA in *Xenopus* egg extracts requires a novel protein XRIPalpha but not importin alpha. *EMBO J.* **1999**, *18*, 4348–4358. [[CrossRef](#)]
69. Narayanan, U.; Ospina, J.K.; Frey, M.R.; Hebert, M.D.; Matera, A.G. SMN, the spinal muscular atrophy protein, forms a pre-import snRNP complex with snurportin1 and importin beta. *Hum. Mol. Genet.* **2002**, *11*, 1785–1795. [[CrossRef](#)]
70. Hodges, J.L.; Leslie, J.H.; Mosammamarast, N.; Guo, Y.; Shabanowitz, J.; Hunt, D.F.; Pemberton, L.F. Nuclear import of TFIIB is mediated by Kap114p, a karyopherin with multiple cargo-binding domains. *Mol. Biol. Cell* **2005**, *16*, 3200–3210. [[CrossRef](#)]
71. Tome-Amat, J.; Ramos, I.; Amanor, F.; Fernandez-Sesma, A.; Ashour, J. Influenza A virus utilizes low-affinity, high-avidity interactions with the nuclear import machinery to ensure infection and immune evasion. *J. Virol.* **2019**, *93*, e01046-18. [[CrossRef](#)] [[PubMed](#)]
72. Schuster, B.S.; Reed, E.H.; Parthasarathy, R.; Jahnke, C.N.; Caldwell, R.M.; Bermudez, J.G.; Ramage, H.; Good, M.C.; Hammer, D.A. Controllable protein phase separation and modular recruitment to form responsive membraneless organelles. *Nat. Commun.* **2018**, *9*, 2985. [[CrossRef](#)] [[PubMed](#)]
73. Guseva, S.; Milles, S.; Jensen, M.R.; Salvi, N.; Kleman, J.P.; Maurin, D.; Ruigrok, R.W.H.; Blackledge, M. Measles virus nucleo- and phosphoproteins form liquid-like phase-separated compartments that promote nucleocapsid assembly. *Sci. Adv.* **2020**, *6*, eaaz7095. [[CrossRef](#)] [[PubMed](#)]
74. Milles, S.; Mercadante, D.; Aramburu, I.V.; Jensen, M.R.; Banterle, N.; Koehler, C.; Tyagi, S.; Clarke, J.; Shammass, S.L.; Blackledge, M.; et al. Plasticity of an ultrafast interaction between nucleoporins and nuclear transport receptors. *Cell* **2015**, *163*, 734–745. [[CrossRef](#)] [[PubMed](#)]
75. Zilman, A. Aggregation, phase separation and spatial morphologies of the assemblies of FG nucleoporins. *J. Mol. Biol.* **2018**, *430*, 4730–4740. [[CrossRef](#)] [[PubMed](#)]
76. Dormann, D. FG-nucleoporins caught in the act of liquid-liquid phase separation. *J. Cell Biol.* **2020**, *219*, e201910211. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Chapter 4: Structural and biochemical  
characterization of the Asiatic toad  
influenza NP



As a newly discovered virus in a 2018 metagenomic study, no further information are known on the Asiatic toad influenza virus (ToadV), except for the partial genome sequences released <sup>329</sup>. Six genomic segments have been recovered so far: the three segments encoding the RdRp subunits (PA, PB1 and PB2), and the HA, NA and NP segments. With both HA and NA encoded by the genome, this virus seems structurally/genetically closer to IAV and IBV. Furthermore, the sequence alignment of ToadV nucleoprotein sequence with all other influenza nucleoproteins shows a higher percent identity with B/NP than A/NP (**Table 6**), which is surprising, considering the huge difference in term of host spectrum between these two well-known influenza viruses. This discovery is particularly interesting: the evolutionary relationship between influenza viruses is still poorly understood, and the differences between the four types of influenza viruses are at the same time extensive and enigmatic.

**Table 6: Sequence identities between Toad/NP and the other influenza NPs.**

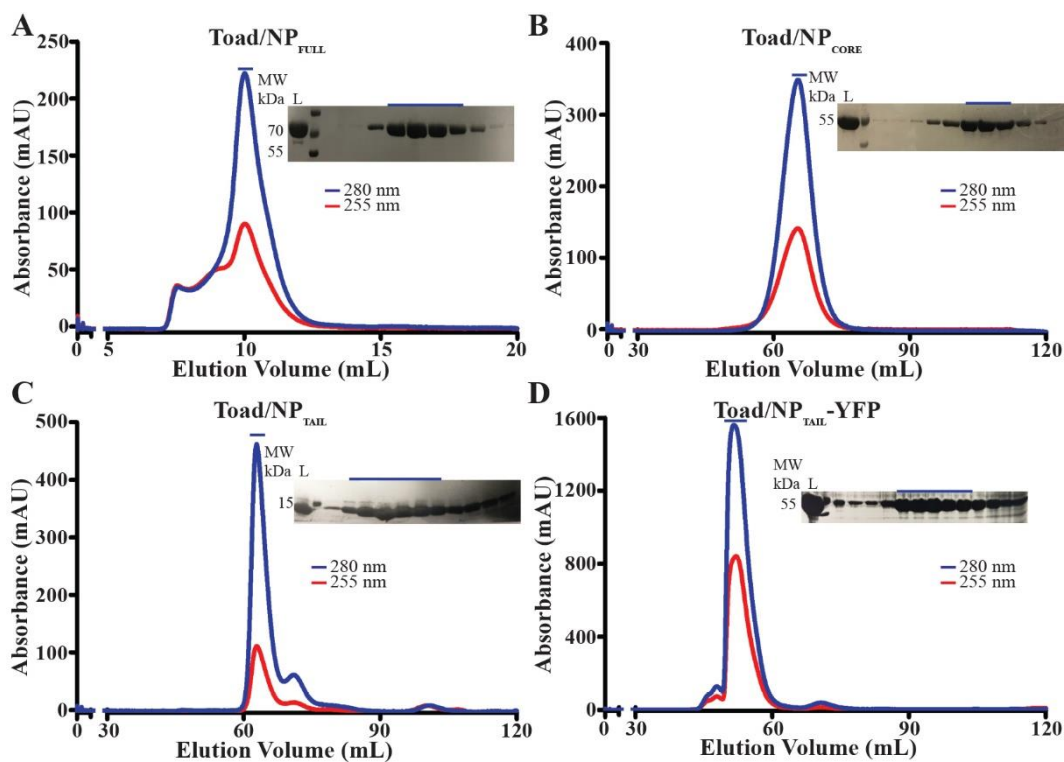
The sequences were obtained from the Influenza Research Database <sup>419</sup>. The strains used for the alignment are those used in this PhD work.

	Toad/NP
A/NP	22.5 %
B/NP	28.7 %
C/NP	22.1 %
D/NP	20.1 %

Based on the characteristics established since the publication of D/NP structure <sup>63</sup>, the structure prediction of Toad/NP (**Figure 19**) revealed that, similarly to B/NP, it possesses a long N-terminal tail ( $\approx$  130 amino acids) with a folded core and a short C-ter disordered tail (7 residues). Toad/NP N-terminal tail was almost twice as long as B/NP<sub>TAIL</sub> ( $\approx$  130 residues against 71), with a similar pattern of scattered basic residues (**Table 7**). Interestingly, Toad/NP<sub>TAIL</sub> could putatively possess a folded domain, as found in ISA/NP, despite their evolutionary distance. In spite of a credible alignment of Toad/NP<sub>CORE</sub> and B/NP<sub>CORE</sub>, the actual N-ter edge of Toad/NP<sub>CORE</sub> could not be strongly pinpointed: the B/NP Proline 72 is aligned with the Toad/NP Glutamine 132 and two Prolines at plausible positions could be found close by (P116 and P123).



affected by the salt concentration used (150 mM and 300 mM) and a 50 mM NaCl buffer lead to a degraded elution, maybe through aggregation. The NP<sub>TAIL</sub> constructs were purified with a nickel affinity chromatography followed by a size exclusion chromatography with a likewise TEV protease cleavage. Toad/NP<sub>FULL</sub> and Toad/NP<sub>CORE</sub> final buffer was 20 mM Tris-HCl pH 7.5, 150 or 300 mM NaCl, 5 mM  $\beta$ -mercaptoethanol while all Toad/NP<sub>TAIL</sub> constructs were kept in 10 mM Hepes pH 7.5, 150 mM NaCl. The ratio 280/260 was always close to two, indicative of the absence of a nucleic acid contamination (**Fig. 20**).



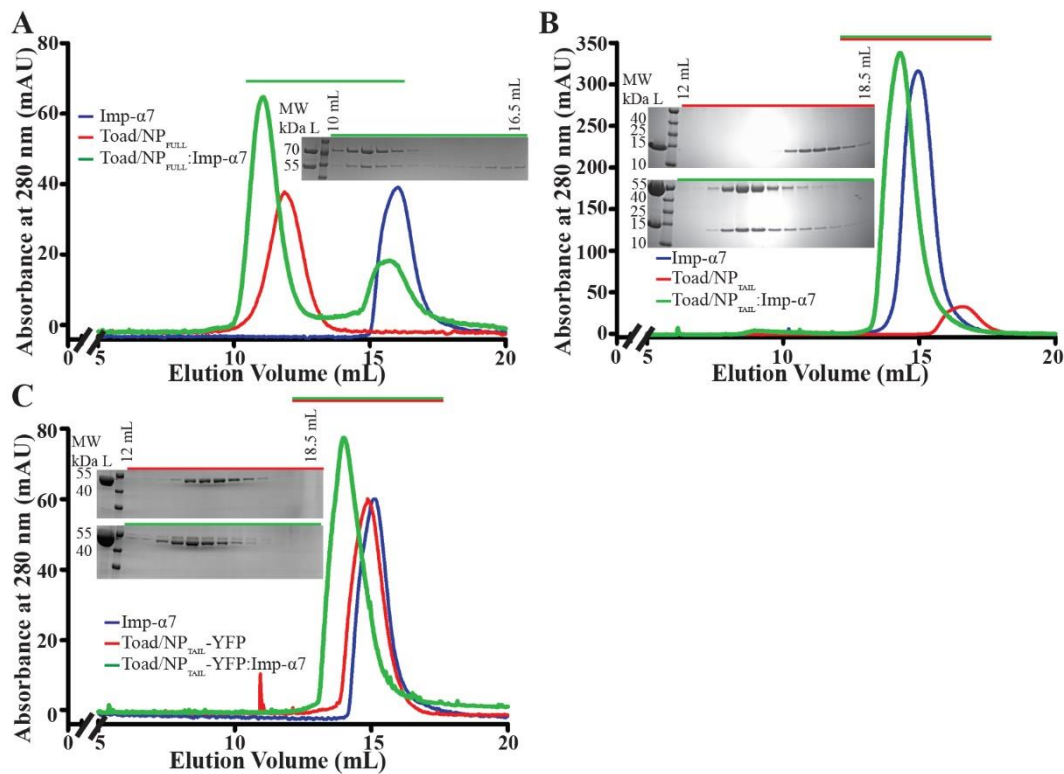
**Figure 20: Profiles of the SEC purification step for each Toad/NP constructs.**

Gel filtration profiles and the corresponding SDS-PAGE of (A) Toad/NP<sub>FULL</sub> on a Superdex<sup>TM</sup> 200 10/300 GL column (B) Toad/NP<sub>CORE</sub> on a HiLoad<sup>TM</sup> 16/600 Superdex<sup>TM</sup> 200 column, (C) Toad/NP<sub>TAIL</sub> and (D) Toad/NP<sub>TAIL</sub>-YFP on a HiLoad<sup>TM</sup> 16/600 Superdex<sup>TM</sup> 75 column. L and MW stand for Load and Molecular Weight respectively.

## I. Characterization of the interaction with importins- $\alpha$

It was suspected that, similarly to other NPs, Toad/NP should access to the nucleus to participate to the transcription and the replication processes, and probably has a crucial role in the initial import of the RNPs in the nucleus. Therefore, as Toad/NP<sub>TAIL</sub>

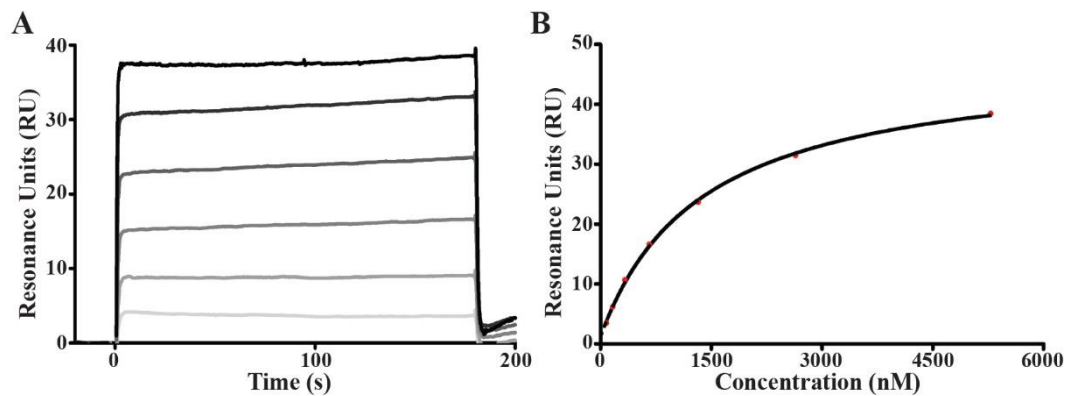
presented several basic patches, the first aim was to determine if it was responsible for an interaction with importins- $\alpha$ , similarly to other influenza NPs, but knowing that the importins- $\alpha$  we had used in initial experiments were human paralogs, potentially different despite their high sequence identities from homologous found in batrachians (**Annexe 1**).



**Figure 21: Interaction assays between Toad/NP constructs and human importin- $\alpha$ 7 by size exclusion chromatography.**

Interaction profiles between the importin- $\alpha$ 7 and (A) Toad/NP<sub>FULL</sub> (30 μM), (B) Toad/NP<sub>TAIL</sub> (70 μM), (C) Toad/NP<sub>TAIL</sub>-YFP (30 μM). The importin- $\alpha$ 7 was used at a concentration of (A) (C) 25 μM or (B) 70 μM. For each panel, the 4-20% SDS-PAGE of the complex elution, with or without the elution of the Toad/NP construct alone is associated. L and MW stand for Load and Molecular Weight respectively.

An interaction assay by gel filtration was achieved as previously<sup>63,64,177</sup>. Briefly, after a 1h incubation, three samples were successively injected on a SEC column (Superdex™ 200 increase 10/300 GL): the NP constructs alone, the importin- $\alpha$ 7 alone and a mix NP:importin- $\alpha$ 7 (**Fig. 21**). Toad/NP<sub>FULL</sub> and Toad/NP<sub>TAIL</sub> (fused or not with the YFP) were shown to be eluted in a complex with the importin- $\alpha$ 7. The experiment was done with Toad/NP<sub>CORE</sub> and in that case, it did not seem to interact with the importin- $\alpha$ 7. The experiment should be redone in the future, for validation.



**Figure 22: Characterization of the interaction between Toad/NP<sub>TAIL</sub>-YFP and the human importin- $\alpha$ 7 by SPR.**

(A) Sensorgrams and (B) associated steady state fit determined on the Biacore evaluation software of the interaction between Toad/NP<sub>TAIL</sub>-YFP and the importin- $\alpha$ 7. Each shade of gray represents one concentration, ranging from 83 nM to 5282 nM (light gray to dark gray).

Toad/NP<sub>TAIL</sub> fused to YFP was then used for a Biacore characterization of the interaction with the importins- $\alpha$ 3, - $\alpha$ 5 and - $\alpha$ 7. The attachment of Toad/NP<sub>TAIL</sub>-YFP on the sensor chip surface was done by the biotin in C-terminal, through the avidin:biotin interaction, up to 200 resonance units. For the kinetics, the importin- $\alpha$ 7 was injected at several concentrations (from 83 nM to up to 5  $\mu$ M), in a serially diluted manner (**Fig. 22**). The first interesting feature came from the rates of association and dissociation. Similarly to B/NP<sub>TAIL</sub> and in clear contrast with A/NP<sub>TAIL</sub>, C/NP<sub>TAIL</sub> and D/NP<sub>TAIL</sub>, they were extremely high, outside of the range of detection of the apparatus used (cf. **Chapter 3**). The second feature concerned the affinities. Toad/NP<sub>TAIL</sub> showed a low affinity for the importin- $\alpha$ 5 and - $\alpha$ 7 but a high affinity for the importin- $\alpha$ 3 (**Fig. 22** and **Table 8**), similar to the affinities obtained for other NP<sub>TAILS</sub><sup>64</sup>. The trend in affinity between the two related NP (*i.e.* B/NP and Toad/NP) was well conserved, with a wider range for Toad/NP<sub>TAIL</sub>. While these low affinities and this trend are still puzzling, the affinity between Toad/NP<sub>TAIL</sub> and the importin- $\alpha$ 3 seemed to suggest a putative preference for this partner.

**Table 8: Biacore steady state affinities between Toad/NP<sub>TAIL</sub> or B/NP<sub>TAIL</sub> and human  $\Delta$ IBB importins- $\alpha$ .**

$K_d$  values are expressed in nM.  $\chi^2$  values represent the goodness of the fit, the lower the better. Values for B/NP<sub>TAIL</sub> come from the Chapter 3.

	Imp $\alpha$ 3		Imp $\alpha$ 5		Imp $\alpha$ 7	
	$K_d$	$\chi^2$	$K_d$	$\chi^2$	$K_d$	$\chi^2$
Toad/NP <sub>TAIL</sub> -YFP	155.8 $\pm$ 23.1	0.02	1217 $\pm$ 63	0.7	1024 $\pm$ 341	0.14
B/NP <sub>TAIL</sub> -YFP	308 $\pm$ 106	0.5	770 $\pm$ 24	0.003	405 $\pm$ 5	0.1

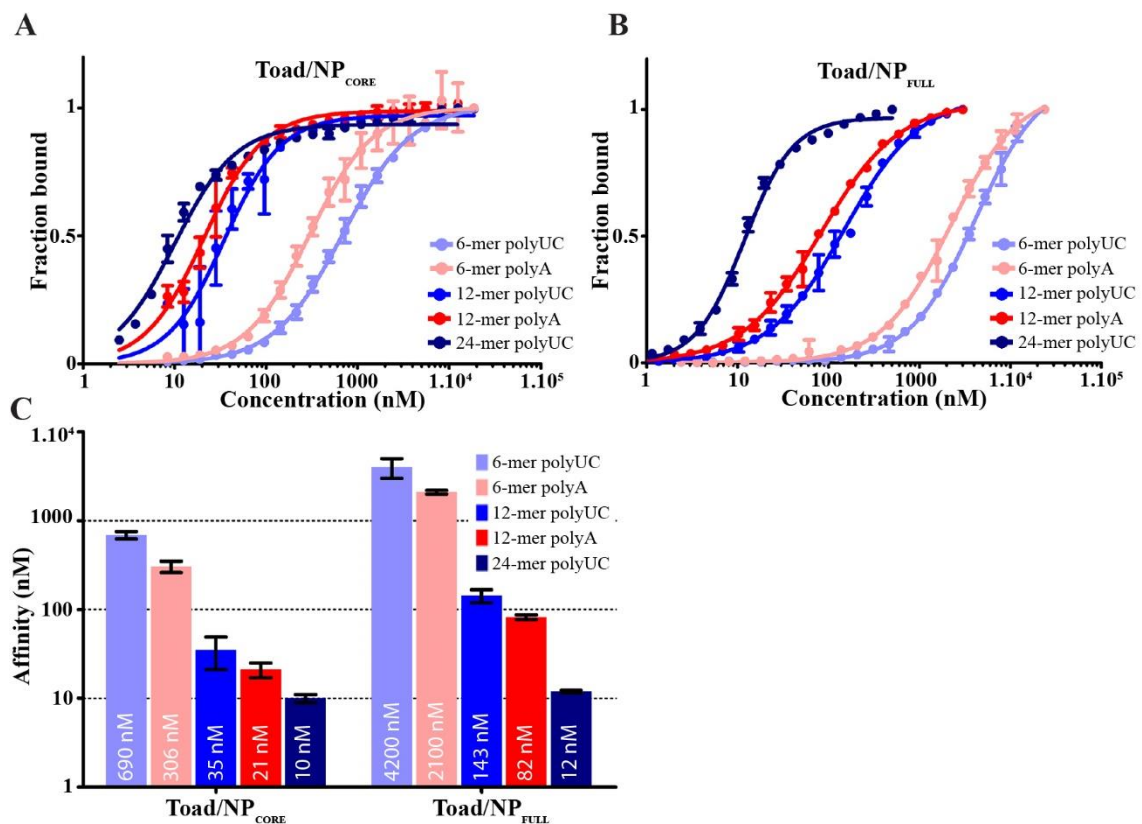
Further work is although required to determine more clearly the residues involved in this interaction, the *in cellulo* kinetics of the nuclear localization and whether other interfaces exist.

## II. Characterization of the interaction with RNAs

One of the main features of influenza nucleoproteins is their essential role in the viral/complementary RNA bindings, for protection and replication/transcription purposes. A/NP and B/NP RNA binding activities have been extensively studied, highlighting a nanomolar affinity with an important impact of the salt concentration<sup>13,27,71,178,242,353</sup>. It is actually not known exactly how many nucleotides a single NP binds, but supposedly ranging from 12 to 24 nucleotides. ISA/NP, a non-influenza NP belonging to the *Orthomyxoviridae* family, is known to bind to 20-mer, 24-mer and 28-mer RNAs with nanomolar affinities, a monomer binding to about 12 nucleotides<sup>422</sup>. In order to decipher its RNA binding activity and to dissect its specificities/similarities regarding others NPs, the affinities of both full length Toad/NP<sub>FULL</sub> and Toad/NP<sub>CORE</sub> for several RNAs were measured by fluorescence anisotropy.

Briefly, fluorescence anisotropy relies on the fact that a fluorescent molecule in solution is in rotation and therefore, upon excitation, polarizes the light. This movement depends on several parameters, and more particularly on the size of the molecule. The bigger the molecule is, the slower it diffuses and the more it polarizes the light. In our experiments, FAM-labelled RNAs (5 nM) were mixed with serially diluted Toad/NP<sub>FULL</sub> and/or Toad/NP<sub>CORE</sub>. The corresponding complex, far bigger in size, polarizes the light. This feature allows the determination of an apparent dissociation constant by titrating the RNAs from a fully free state to a fully bound state.

Measurements were done at 150 mM and 300 mM NaCl. In these salt concentrations, the constructs seem to be mainly tetrameric. Three sizes of RNAs were used: 6-mer, 12-mer and 24-mer RNA labelled in 3' with fluorescein (FAM), with either a polyA or a polyUC sequence. A short RNA (6-mer) is able to bind to the RNA-binding groove without saturating it and usually clearly highlights an impaired binding ability for any conditions or constructs. A 12-mer RNA could represent the size fitting the NP RNA binding groove<sup>178,422</sup>. Finally, the 24-mer RNA is the standard binding size used in influenza NP:RNA interactions.



**Figure 23: Titration against FAM-labelled RNAs of Toad/NP<sub>FULL</sub> and Toad/NP<sub>CORE</sub> by fluorescence anisotropy at 150 mM NaCl.**

Titration measurements against FAM-labelled RNAs with (A) Toad/NP<sub>CORE</sub> or (B) Toad/NP<sub>FULL</sub>. (C) Affinities of the full length Toad/NP or its core for RNAs at 150 mM NaCl by fluorescence anisotropy. The fitting was done using the GraphPad Prism 5 equation "One site – Specific Binding with Hill slope".

Toad/NP was serially diluted and mixed with 5 nM of FAM-labelled RNA in a 384-well microplate. The measurements and analyses were done as previously detailed (cf. Chapter 3), on a CLARIOstar microplate reader, with excitation and emission wavelengths fixed at 482 and 530 nm respectively, after a 10-min incubation at room

temperature (Fig. 23, Table 9). After blank subtraction and normalization, the fitting were represented using GraphPad Prism 5, with the equation “One site – specific binding with Hill slope”. The Hill slope coefficient was consistent with the cooperative features of the binding in all conditions: the longer the RNA, the higher the factor, although this was more marked for Toad/NP<sub>CORE</sub> than for Toad/NP<sub>FULL</sub>.

**Table 9: Toad/NP:RNA apparent dissociation constants determined by fluorescence anisotropy.**

NaCl Concentration	150 mM NaCl		300 mM NaCl
	Toad/NP <sub>FULL</sub>	Toad/NP <sub>CORE</sub>	Toad/NP <sub>CORE</sub>
<b>6-mer polyA</b>	2.1 $\mu\text{M} \pm 0.1 \mu\text{M}$	306 nM $\pm 44$ nM	7.8 $\mu\text{M} \pm 1.5 \mu\text{M}$
<b>6-mer polyUC</b>	4.2 $\mu\text{M} \pm 0.6 \mu\text{M}$	690 nM $\pm 64$ nM	12.8 $\mu\text{M} \pm 1.6 \mu\text{M}$
<b>12-mer polyA</b>	82.2 nM $\pm 4.9$ nM	21 nM $\pm 4$ nM	751 nM $\pm 13$ nM
<b>12-mer polyUC</b>	143 nM $\pm 24$ nM	35 nM $\pm 14$ nM	1.2 $\mu\text{M} \pm 0.1 \mu\text{M}$
<b>24-mer polyUC</b>	11.9 nM $\pm 0.4$ nM	10 nM $\pm 1$ nM	106 nM $\pm 7$ nM

At 150 mM NaCl, a concentration close to the intracellular KCl concentration, Toad/NP<sub>FULL</sub> and Toad/NP<sub>CORE</sub> did not show any strong preference for a polyA or a polyUC sequence, with only a 2-fold difference in term of affinity for polyA RNA molecules. Both Toad/NP<sub>FULL</sub> and Toad/NP<sub>CORE</sub> interacted with a low nanomolar affinity (10 nM) with a 24-mer polyUC RNA, saturating the binding sites. Toad/NP<sub>CORE</sub> showed similar affinities for 24-mer polyUC, 12-mer polyA and 12-mer polyUC, suggesting that the binding pattern was identical and putatively saturated the binding groove. These results are in agreement with the hypothesis that 12 nucleotides are strongly bound to one NP monomer, while the others are more loosely interacting, maybe responsible for the recruitment or interaction with other NPs. In contrast, the affinity of Toad/NP<sub>FULL</sub> for both 12-mer RNA molecules is lower (7- and 12-fold decrease in affinity). A sharp decrease in affinity was observed for the two constructs when the 6-mer RNAs were used. This pattern of decreasing affinities for decreasing RNA sizes is similar to those observed for A/NP and B/NP, and expected for cooperative binding with a binding groove that may be able to accommodate about 12 nucleotides<sup>178</sup>.

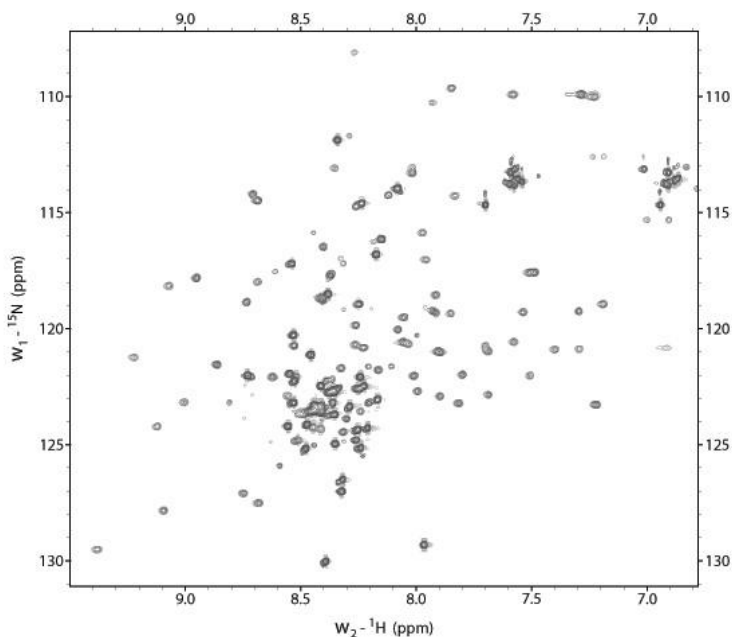
The most striking feature of these titrations came from the the comparison between Toad/NP<sub>FULL</sub> and Toad/NP<sub>CORE</sub>. Interestingly, when Toad/NP was deprived of its long N-terminal tail, the affinities are 4- to 7-fold higher than when the full-length protein was used, suggesting a putative regulatory role, direct or indirect, of the tail in the RNA-



binding activity. The RNA titrations with Toad/NP<sub>TAIL</sub> were done, to verify any RNA-binding activity, but the tail did not seem to bind RNAs specifically, with affinities determined for both 24-mer and 12-mer RNAs above 230  $\mu$ M. At 300 mM NaCl, Toad/NP<sub>CORE</sub> was clearly impaired in its global RNA binding activity in comparison to 150 mM NaCl, with more than a 10-fold decrease towards all RNAs. However, it still showed this pattern of decreasing affinity for shorter RNAs, with again a 2-fold difference between polyA and polyUC. This characterization should be continued, firstly with measurements with Toad/NP<sub>FULL</sub> at 300 mM NaCl. It could also be interesting to go into details of the oligomerization state/process of both Toad/NP, with and without RNAs, to better understand what the role of Toad/NP<sub>TAIL</sub> is and how Toad/NP interacts with RNAs and package the genome into RNPs.

### III. Structural characterization of Toad/NP

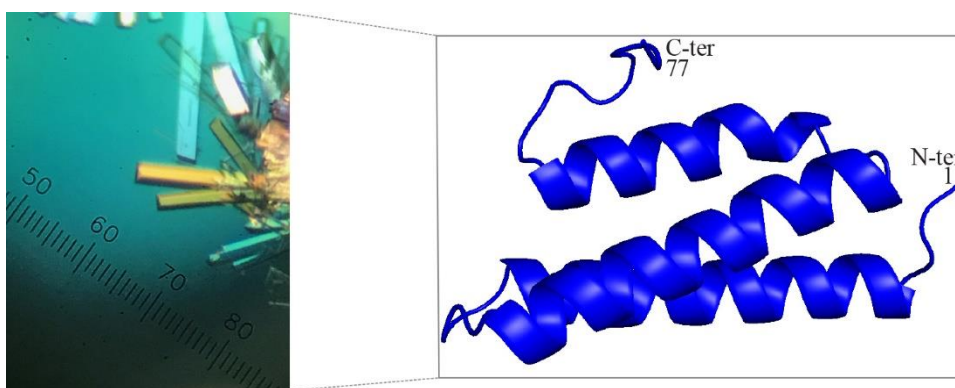
The structural analysis of this new influenza NP was conducted in parallel of the biochemical/biophysics experiments. In order to gain evidences on the propensity of Toad/NP<sub>TAIL</sub> to fold in secondary structure elements, a measurement of the <sup>15</sup>N-labelled Toad/NP<sub>TAIL</sub> by NMR was done on the ISBG NMR platform (**Fig. 24**). The spectra showed that, while Toad/NP<sub>TAIL</sub> contained an intrinsically disordered region (narrow dispersion of the peaks in the middle of the spectra,  $\omega_{2-1}^1\text{H} = 8 - 8.6$  ppm) but also, a part of the protein looked folded (proton chemical shifts present in the other parts of the spectra).



**Figure 24:**  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of  $^{15}\text{N}$ -labelled Toad/NP<sub>TAIL</sub>.

Therefore, based on the secondary structure predictions, a new NP<sub>TAIL</sub> construct, consisting in the first 98 N-ter residues, was engineered, produced and purified similarly to the full length NP<sub>TAIL</sub>. The sample was submitted to crystallization on the HTX platform (EMBL) and an initial crystallization condition using ammonium sulfate was identified and reproduced manually. The crystallization was carried out with 1:1 protein:crystallization solution ratio. The protein crystallized in 3 M ammonium sulfate in combination with various buffers. Crystals obtained in 0.1 M Tris-HCl pH 7.25, 3 M ammonium sulfate were analysed on PROXIMA-1 (synchrotron SOLEIL) and a native dataset was collected at 2.2 Å resolution. However, no structural homolog could be found to solve the structure by molecular replacement. Selenomethionine-labelled Toad/NP<sub>TAIL</sub> crystals were then produced (crystals grown in 0.1 M Tris pH 7.5-7.75, 2.85-2.95 M ammonium sulfate). Diffraction data were collected on PROXIMA-1 with the multiple-wavelength anomalous dispersion (MAD) technique and processed with XDS<sup>150</sup>. The structure was solved by Pierre Legrand (PROXIMA-1; SOLEIL) using the CCP4i programs for crystallography<sup>398</sup>. The final model was built using COOT<sup>73</sup> and refined at 2.2 Å resolution ( $R_{\text{fact}} = 20\%$ ;  $R_{\text{free}} = 24\%$ ) using REFMAC5 of the CCP4 suite (**Fig. 25**).

The asymmetric unit contained six molecules, with little difference in term of folding, flexibility and number of residues within the electron density. Residue 1 up to residue 77 could be assigned in the electron density map and residues 4 to 68 were folded in three distinct antiparallel  $\alpha$ -helices (**Fig. 25**). This domain (residues 4-68) was compact, highly charged (pHi of 9.44 with 34% of charged residues (Arg, Lys, Glu and Asp)) and stabilized by many electrostatic interactions between side chains. The C-terminal part of the tail (residues 78 to 98) was not visible in the electron density maps, supposedly intrinsically disordered. Interestingly, residues 80 to 98 were composed for half of the sequence in glutamate residues.

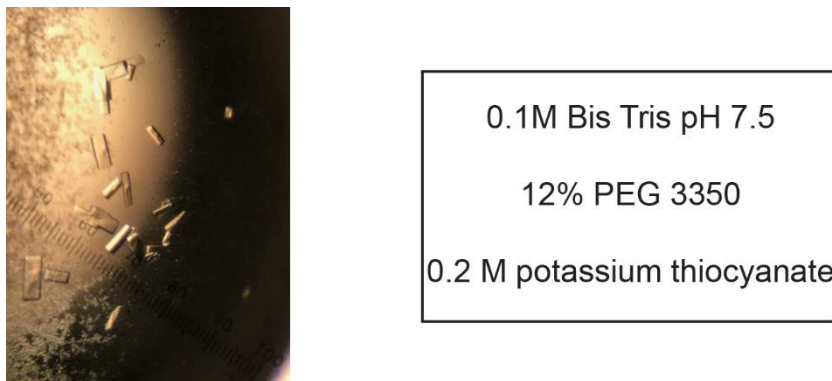


**Figure 25: Crystals of Selenomethionine-labelled Toad/NP<sub>TAIL</sub> and Toad/NP<sub>TAIL</sub> solved structure.**

One small graduation represents 10  $\mu\text{m}$ .

Toad/NP<sub>FULL</sub> were also crystallized using the HTX platform. Crystals were obtained in 0.1 M Bis Tris pH 7.5, 18-20 % PEG 3350, 0.2-0.25 M potassium thiocyanate. However, the packing seemed faulty: a five seconds-long exposure time was required to obtain a visible diffraction pattern at 8  $\text{\AA}$  resolution on a single image. The space group was P4 with cell parameters  $a = b = 256.4 \text{ \AA}$ ,  $c = 215.3 \text{ \AA}$  and  $\alpha = \beta = \gamma = 90^\circ$ . As the long, flexible N-terminal tail could negatively affect the crystal packing, Toad/NP<sub>CORE</sub> (residues 124 to 607) was crystallized to avoid such an issue. Toad/NP<sub>CORE</sub> crystals were obtained in 0.1 M Bis-Tris pH 7.5, 14% PEG 3350, 0.2-0.25 M potassium thiocyanate. These crystals diffracted at 3.5  $\text{\AA}$  (space group P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> with the following cell parameters  $a = 46.5 \text{ \AA}$ ,  $b = 56.8 \text{ \AA}$ ,  $c = 357.4 \text{ \AA}$  and  $\alpha = \beta = \gamma = 90^\circ$ , suggesting two molecules in the asymmetric unit. The molecular replacement did not

offer any satisfying solution, thus crystals of selenomethionine-labelled Toad/NP<sub>CORE</sub> were reproduced (0.1 M Bis-Tris pH 7.5, 12 % PEG 3350, 0.2 M potassium thiocyanate) for future synchrotron experiments (**Fig. 26**).



**Figure 26: Crystals of Selenomethionine-labelled Toad/NP<sub>CORE</sub> and its crystallization condition.**

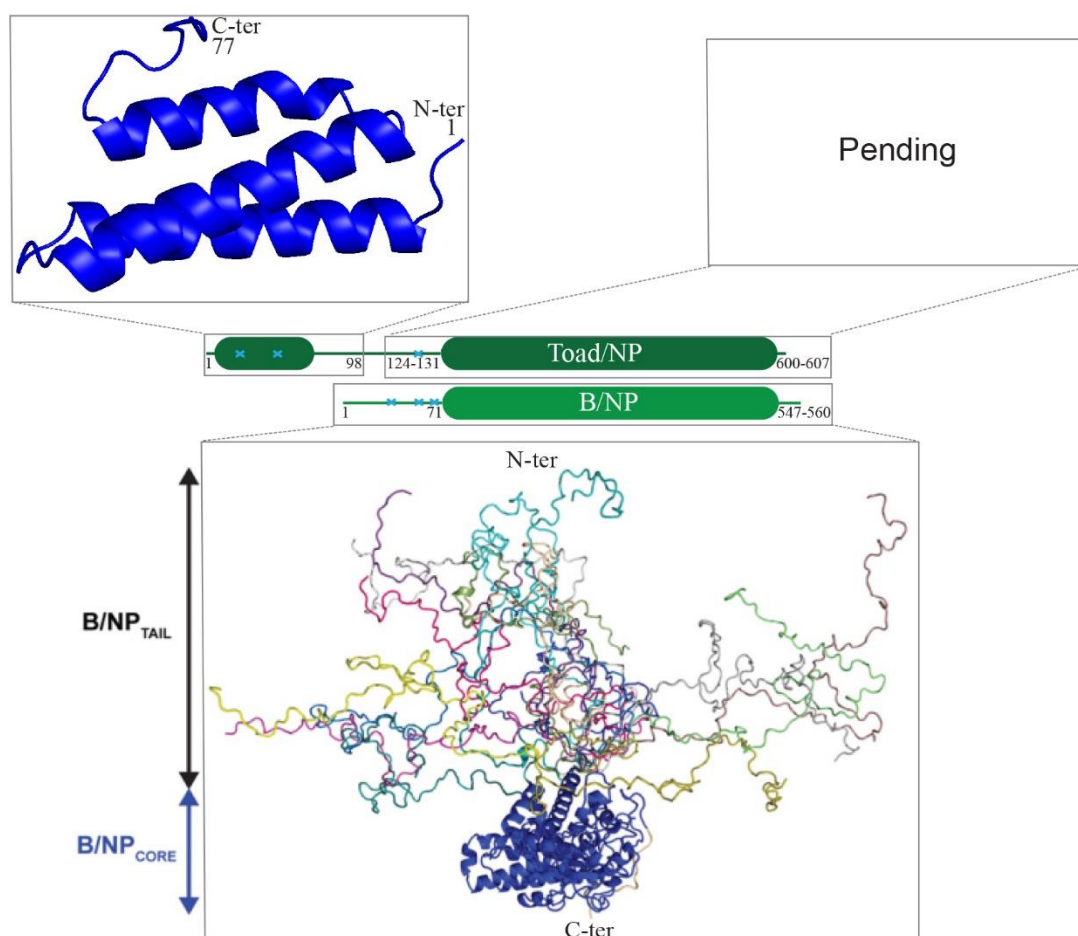
Two small graduations represent 10  $\mu\text{m}$ .

#### IV. Discussion

ToadV has shown its phylogenetic proximity with IBV, its NP sharing a part of B/NP features but has also its own, very specific behaviour and structures.

Toad/NP seems to have a tendency to form tetramers in solution. The salt concentration in the buffer (150 mM or 300 mM NaCl) does not have a clear effect on the oligomerization state, similarly to what was observed for D/NP and in contrast to B/NP<sup>63,178</sup>. Toad/NP possesses a long N-terminal tail, with several scattered basic residues, a pattern highly similar to the B/NP<sub>TAIL</sub> NLS. A long stretch of residues (amino acids 30 to 71) of B/NP<sub>TAIL</sub> is involved in the importins- $\alpha$  binding; it could be interesting to know if a similar behaviour is also found in Toad/NP<sub>TAIL</sub><sup>177</sup>. Furthermore, the exact NLS motif of the tail is not known yet and further investigation are required to highlight the molecular basis of this particular interaction. Toad/NP<sub>TAIL</sub> interactions with importins- $\alpha$  follow the same pattern as B/NP<sub>TAIL</sub>, with high association and dissociation rates (outside of the detection range of the apparatus) and low affinities for the importin- $\alpha$ 5 and importin- $\alpha$ 7 ( $K_d$  of 1217 nM and 1024 nM respectively). Interestingly, Toad/NP<sub>TAIL</sub> affinities are even weaker for these importins- $\alpha$  than B/NP<sub>TAIL</sub> affinities, in the micromolar range and putatively outside of the functional range. In contrast, it

interacts more strongly with the importin- $\alpha 3$  than B/NP<sub>TAIL</sub>, with an affinity (156 nM) suggesting that they could form a functional complex. However it is of note that human importins- $\alpha$  were used in this work. While the importin- $\alpha 3$ , - $\alpha 5$ , and - $\alpha 7$  from human and frog share a high sequence identity (94%, 94% and 90% respectively, **Annexe 1**), it is a possibility that these importins, and this nuclear import pathway as a whole, possess some species-specificities. The extrapolation, while possible, should be cautious.



**Figure 27: Structural features and comparison between Toad/NP and B/NP. Toad/NP N-terminal structure is a result of this PhD work whereas B/NP modelling was published by <sup>177</sup>.**

Intrinsically disordered parts are represented as lines while folded parts are represented as rounded squares. Blue crosses represent potential NLS.

Toad/NP interacts with RNAs with a pattern similar to other influenza NPs, seemingly without sequence specificity, with cooperativity, with usually a nanomolar

affinity<sup>13,70,71,178,407</sup>. One point was particularly interesting: while Toad/NP behaves similarly to B/NP, the Toad/NP<sub>TAIL</sub> seems to play a unique role in RNA binding regulation, maybe through its Glutamate-rich region between residues 80 and 98. Interestingly, a negatively charged region is also found in C-terminal of influenza NPs and ISA/NP with several glutamate and aspartate, suspected to regulate the RNA-binding activity. Only ISA/NP shows such a Glutamate-rich pattern, which has been shown to impact RNA binding, but this motif is localized in the last 16 C-terminal residues. As Toad/NP does not show such an acidic C-terminal patch, it is possible to speculate about a change in localization of this particular region. As the Toad/NP C-terminal disordered tail is the shortest (8 residues) and its N-terminal one the longest, it is possible that both N-ter and C-ter tail functions could be found on the sole N-terminal Toad/NP<sub>TAIL</sub>.

While the structure of Toad/NP<sub>CORE</sub> is not known yet, the structure in the N-terminal tail was solved. It is neatly packed in three long antiparallel  $\alpha$ -helices (14 - 20 residues), which do not seem too flexible, as no alternate conformation was found in the six molecules from the asymmetric unit. Regarding influenza NPs, this is the first domain found in a disordered N/C-terminal tail. B/NP, the Toad/NP closest relative, has a long N-terminal extremity that is entirely disordered. The contrast is striking between this folded domain within Toad/NP<sub>TAIL</sub> and modelled B/NP<sub>TAIL</sub>, extensively flexible and occupying a large volume (**Fig. 27**). This structural insight is particularly interesting for two points: firstly, the basic patches detected within the full length NP<sub>TAIL</sub> are involved in importins- $\alpha$  binding and secondly, Glu-rich regions and low-complexity-regions are well known for their role as molecular chaperones or in protein:protein, protein:nucleic acid interactions<sup>45,368</sup>. Both the folded N-terminal domain and the disordered segment require additional work to detail their roles.

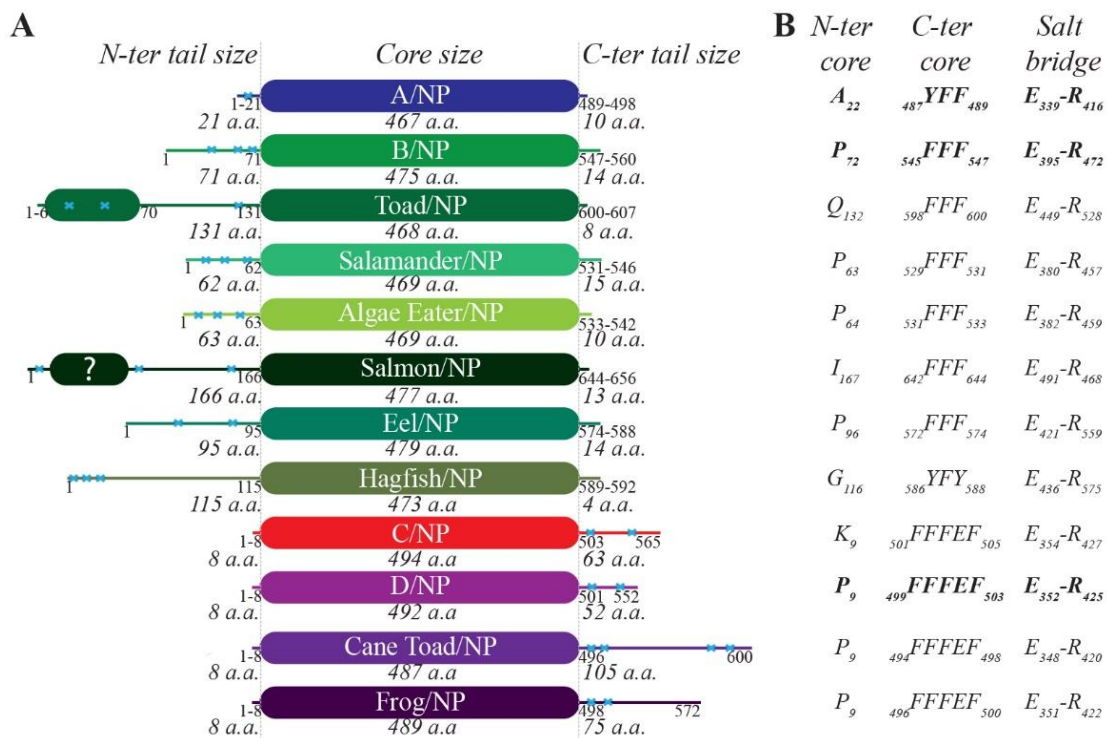
Outside of this specific domain, while Toad/NP<sub>CORE</sub> structure is still pending, it is highly plausible that it will be highly similar to the one of other NPs: their size is closely matched, their sequence relatively conserved and all *Orthomyxoviridae* NPs have so far revealed a structurally conserved core (**Fig. 27**).

**Table 10: Sequence identities between influenza NP types.**

Sequences were recovered from the Influenza Research Database <sup>419</sup>. The strains used for alignment are those used in this work. Alignment was realized using Clustal Omega <sup>199</sup>. Sala, AIEa and Salm stands for Salamander, Algae Eater and Salmon respectively.

	A NP	B NP	Toad NP	Sala NP	AIEa NP	Salm NP	Eel NP	Hagfish NP	C NP	D NP	Cane Toad NP	Frog NP
A NP	100	37.2	22.5	37.9	41.2	38.0	38.2	21.0	21.8	21.1	21.4	22.2
B NP		100	28.7	47.6	41.5	43.0	53.5	23.2	23.1	23.8	25.9	23.5
Toad NP			100	27.4	25.2	30.4	26.2	21.2	22.1	20.1	23.6	23.8
Sala NP				100	38.0	39.7	46.3	21.5	22.4	26.2	23.4	25.4
AIEa NP					100	37.8	39.1	22.2	21.9	21.1	23.6	22.3
Salm NP						100	41.2	19.7	22.8	22.4	24.7	24.1
Eel NP							100	21.6	23.7	23.8	23.4	22.2
Hagfish NP								100	21.4	20.1	20.8	22.3
C NP									100	37.5	38.3	38.7
D NP										100	47.3	51.2
Cane Toad NP											100	51.1
Frog NP												100

In September 2020, when the characterization of Toad/NP was still underway, a new study identified 5 new influenza-like viruses from extremely diverse organisms <sup>268</sup>. Three of them turned out to be more closely related to IBV (one salamander-, one algae eater- and one salmon-infecting virus), and two, for the first time, were identified to be more closely related to IDV (infecting either a cane toad or a frog). Preliminary predictive structural analyses can be done with these new NPs sequences, based on the work on Toad/NP. The more surprising point was the high sequence identities found between influenza NPs and influenza-like NPs. B/NP shared between 42% and 54% sequence identities with Salamander/NP, Algae Eater/NP, Salmon/NP and Eel/NP (**Table 10**). Toad/NP was more divergent, probably branching out earlier, while still relatively close to B/NP. Similarly, D/NP and D-like/NPs shared high sequence identities (between 47 and 51%).



**Figure 28: Diversity of newly identified NPs.**

Structural predictions were performed using PsiPred<sup>146</sup>. Intrinsically disordered parts are represented as lines while folded parts are represented as rounded squares. Blue crosses represent potential NLS. In (B), data coming from the crystal structures are in bold, in contrast with the predictive analyses.

All B-like/NPs seem to possess a structural organization similar to B/NP: a long intrinsically disordered N-terminal tail, a folded core, and a short C-ter tail (**Fig. 27A**). The N-terminal intrinsically disordered tail ( $NP_{TAIL}$ ) was highly variable in size. Interestingly, all  $NP_{TAILS}$  possessed basic patches that could be NLSs, and one basic patch, while not conserved in sequence, was always found within the last 20-30 residues of  $NP_{TAILS}$ . A similar result came from the comparison of the organization of D/NP and D-like/NPs (**Fig. 27A**). Other hallmark features seemed to be conserved (**Fig. 27B**). The salt bridge essential for the oligomerization process (salt bridge between Glu-339 and Arg-416 in A/NP), the aromatic motif in C-terminal of the folded core and to some extent the N-terminal residue of the core could be identified for B-like and D-like NPs.

Now, considering Toad/NP and Cane Toad/NP, one intriguing feature appeared: they both were infecting a toad and their structural organization was mirrored: a NLS-bearing N-terminal  $NP_{TAIL}$  of 127 residues or a NLS-bearing C-ter  $NP_{TAIL}$  of 105 residues



respectively, a similarly sized core and the opposite terminal extremity made of 8 amino acids. Furthermore, a part of the Cane Toad/NP C-terminal tail is predicted potentially folded in one or two  $\alpha$ -helices.

These predictive analyses confirm the judicious choice of characterizing Toad/NP, as it seems to have a unique position in the growing influenza family. This opens up many new avenues to probe at the evolutionary history and functionality of influenza viruses.

# Conclusions

A large number of studies tried to dissect the IAV, IBV and ICV replication machineries function through the years, along with its essential interactions with the host cell machinery. Host: pathogen interactions are one of the more crucial interfaces, in term of function, host specificity, adaptation and therapeutic targets. The interaction of the influenza replication machinery with the cellular nuclear import system is one of them, yet little is known despite the number of publications on this interaction. The discoveries in 2011 of the influenza D virus and, during the course of this PhD, of the Asiatic toad-infecting influenza virus in 2018 were game-changing news, as the influenza world was by the past limited to three genera since the discovery of the influenza C virus in 1947, more than half a century ago. It is well known that our understanding of viruses, viral diversity and viral evolution is largely biased, but these two viruses managed to be surprising. IDV is more closely related to one of the influenza family member with a strongly restricted host spectrum (ICV) instead of the other genus with a broader host spectrum (IAV). Similarly, ToadV was phylogenetically closer to IBV than any other influenza genera.

During this PhD thesis, with a focus on the importin- $\alpha$ :NP<sub>TAIL</sub> interaction, structural and functional insights have been gained on influenza NPs. Firstly, the structural and functional basis of D/NP initiated by Alice Labaronne, a previous PhD candidate in the lab, was finalised. In spite of its low sequence identity with A/NP and B/NP, the folded core of D/NP is structurally well conserved, with R.M.S.D. values of 1.9-2.1 Å. It presents the key conserved features of influenza NP. Firstly, the oligomerization mechanism is conserved, relying on a conserved salt bridge between a glutamate (*i.e.* E352, similar to A/NP E339) and an arginine (*i.e.* R425, similar to A/NP R416), along with other stabilizing contacts, more particularly hydrophobic contacts. The conserved hydrophobic aromatic motif localized at the C-terminal edge of the folded core (487YFF<sub>489</sub> and 545FFF<sub>647</sub> in A/NP and B/NP respectively) is also present in D/NP (499FFF<sub>501</sub>). Like B/NP, the N-terminal edge of the core is a Proline, at position 9. In contrast with A/NP and B/NP, only the 8 first residues are intrinsically disordered at the N-terminal extremity, while the last 52 amino acids are intrinsically disordered. This C-terminal NP<sub>TAIL</sub> shows a clear bipartite NLS, responsible for the nuclear localization and the interaction with importins- $\alpha$ , with a strong affinity (K<sub>d</sub> = 100 nM) for the importin- $\alpha$ 7<sup>63</sup>.

The huge differences between the tail localization, the NLS sequences, the importins- $\alpha$  studied and the affinities determined so far, along with the diversity of methods and techniques lead to the characterization and the comparison of all importin- $\alpha$ :NP<sub>TAIL</sub> complexes within the same settings, in order to get a broader, more precise view of this interaction. As the A/NP<sub>TAIL</sub>, B/NP<sub>TAIL</sub> and C/NP<sub>TAIL</sub> sequences came from human strains, it has been chosen to use human importins- $\alpha$  as partners; these proteins being highly conserved in higher eukaryotes, it should be possible to extrapolate the results to some extent. A SPR strategy has been set up to work on representatives of all NP types, so insights in the affinities and rates of these interactions were gained. In agreement with the cellular process, both the association and the dissociation rates are fast for such a strong *in vitro* interaction. Interestingly, each NP<sub>TAIL</sub> type showed a specific, preferential binding, with a clear contribution of the importin- $\alpha$  paralog. D/NP<sub>TAIL</sub>, as a bipartite NLS, interacted strongly with all importins- $\alpha$  without clear specificity, with a dissociation constant about 20-40 nM. C/NP<sub>TAIL</sub>, while also a bipartite NLS, was a bit more versatile in its interactions with the importins, with a  $K_d$  ranging from 60 to 290 nM, importin- $\alpha$ 3 being the best candidate for the interaction. A/NP<sub>TAIL</sub>, with its unclassical monopartite NLS, presented an affinity range similar to C/NP<sub>TAIL</sub> ( $K_d$  70-620 nM). The best “match” was this time the importin- $\alpha$ 7:A/NP<sub>TAIL</sub> one. These results are particularly interesting and consistent, as the importin- $\alpha$ 3 and - $\alpha$ 7 seem to be the importins- $\alpha$  most expressed in the mammalian respiratory tract. B/NP<sub>TAIL</sub> interaction features were the most puzzling fact, with faster rates and low affinity to all partners tested (no better than 310 nM), raising questions about its true importin partner, its state of post-translational modifications and its translocation process. These results provide a strong basis for the interaction between importins- $\alpha$  and influenza NPs, and highlight specificities from both NPs and importins not foreseen before. The major experienced drawback was the dimerization tendency of the importin- $\alpha$ 1, requiring another strategy to get insights on its behaviour towards NP<sub>TAILS</sub>. These results are essential in the process of understand how this complex, tightly regulated interaction and translocation process works <sup>64</sup>.

In the meantime, structural and functional investigations of Toad/NP were initiated, due to its surprising evolutionary proximity to B/NP and its tropism towards a cold-blooded animal. Integrated to our SPR strategy, the long N-terminal Toad/NP<sub>TAIL</sub> shows a strikingly similar behaviour to B/NP<sub>TAIL</sub>, with fast rates and low affinities. However, the

affinities range, depending on the importins- $\alpha$ , is more important, and the affinity of the Toad/NP<sub>TAIL</sub>:importin- $\alpha$ 3 complex suggests a potential preferential interaction ( $K_d$  160 nM). Several basic patches are identified within this tail, either in the intrinsically disordered part or in the  $\alpha$ -helices. It is not known yet which basic patch is responsible for the interaction. Regarding Toad/NP<sub>CORE</sub> interaction with RNAs, its behaviour matches the A/NP and B/NP ones determined by the team at 150 mM NaCl: it seems that a 12-mer RNA is sufficient to saturate the RNA-binding groove of the NP. The most particular feature comes from Toad/NP<sub>TAIL</sub>: the removal of the long tail positively affects the binding activity, suggesting a regulatory role, putatively carried out by the Glutamate-rich sequence found in the region 80-98. While the C-terminal extremity of influenza NPs has been shown to negatively regulate the RNA-binding activity, it is the first time another potential mechanism has been found at work. In term of structural organization and in contrast with other influenza NPs, Toad/NP<sub>TAIL</sub> showed a folded domain within the N-terminal tail, packed in three long  $\alpha$ -helices. While ISA/NP in the *Orthomyxoviridae* family also present a folded domain in N-terminal, this latter is organized around 3  $\beta$ -strands and 2  $\alpha$ -helices. So far, the function(s) of these domains is(are) not known. Structural insights on Toad/NP<sub>CORE</sub> has not been gained yet, but are expected in the future, as crystals of selenomethionine NP<sub>CORE</sub> has been obtained.

Globally, while this PhD work did not apply the SPR strategy to the polymerase subunits as was initially envisaged, it allowed the successful characterization of all NP<sub>TAILS</sub>:importins- $\alpha$ , which is a crucial interaction both at the NP and the RNP levels. Furthermore, the characterization of the newly discovered D/NP and Toad/NP revealed interesting features and specificities, highlighting how differently influenza NPs could behave and be organized, while still retaining a strong common structural and functional basis.

# Annexes

## Annexe 1: Sequence identities between importins- $\alpha$

**Table S1: Sequence identities between importins- $\alpha$  paralogs of distinct species of interest.**

Sequences and accession numbers were recovered from NCBI/Uniprot databases for human (*Homo sapiens*) importins- $\alpha$ . Each BLAST was done separately, using one human importin sequence against reference protein sequences from *Sus scrofa* (swine), *Bos Taurus* (bovine), *Gallus gallus* (chicken), *Anas platyrhynchos* (duck) and *Xenopus tropicalis* (frog). As no duck importin- $\alpha$ 7 was identified in the NCBI database, the corresponding sequence was recovered from Uniprot database and aligned with a BLAST against reference protein sequences from *Homo sapiens* (bold). 99+ corresponds to a sequence identity between 99.5 and 100 %. No importin- $\alpha$ 6 was found for *Xenopus tropicalis*. Importins- $\alpha$  from the same subfamily are highlighted in the same colour.

	Subfamily Member	Human						
		$\alpha$ 1		$\alpha$ 2		$\alpha$ 3		
		Imp- $\alpha$ 1	Imp- $\alpha$ 8	Imp- $\alpha$ 3	Imp- $\alpha$ 4	Imp- $\alpha$ 5	Imp- $\alpha$ 6	Imp- $\alpha$ 7
Human	Imp- $\alpha$ 1	100						
	Imp- $\alpha$ 8	55	100					
	Imp- $\alpha$ 3	51	45	100				
	Imp- $\alpha$ 4	50	46	86	100			
	Imp- $\alpha$ 5	46	42	47	48	100		
	Imp- $\alpha$ 6	48	42	48	49	80	100	
	Imp- $\alpha$ 7	48	42	49	48	82	84	100
Swine	Imp- $\alpha$ 1	97	54	51	49	45	48	47
	Imp- $\alpha$ 8	55	75	46	48	41	42	41
	Imp- $\alpha$ 3	52	45	99	86	47	48	48
	Imp- $\alpha$ 4	50	46	86	99	49	49	49
	Imp- $\alpha$ 5	46	42	46	48	99	80	81
	Imp- $\alpha$ 6	49	43	48	49	79	97	84
	Imp- $\alpha$ 7	48	42	48	48	81	84	99
Bovine	Imp- $\alpha$ 1	95	54	52	51	46	48	47
	Imp- $\alpha$ 8	55	75	46	46	42	42	41
	Imp- $\alpha$ 3	51	45	99+	86	47	48	48
	Imp- $\alpha$ 4	50	46	86	99	49	49	49
	Imp- $\alpha$ 5	46	42	46	48	99	80	81
	Imp- $\alpha$ 6	48	42	48	49	80	98	84
	Imp- $\alpha$ 7	48	42	48	49	81	84	99
Chicken	Imp- $\alpha$ 1	83	55	51	51	46	47	47
	Imp- $\alpha$ 8	60	61	49	50	43	45	44
	Imp- $\alpha$ 3	51	45	99	86	47	48	48
	Imp- $\alpha$ 4	50	46	86	98	49	49	49
	Imp- $\alpha$ 5	46	42	46	49	96	79	81
	Imp- $\alpha$ 6	49	44	48	49	81	94	84
	Imp- $\alpha$ 7	48	43	48	49	81	84	94
Duck	Imp- $\alpha$ 1	84	22	51	51	46	47	46
	Imp- $\alpha$ 8	60	60	49	48	42	44	43
	Imp- $\alpha$ 3	51	45	98	86	48	48	48
	Imp- $\alpha$ 4	51	46	86	99	49	49	48
	Imp- $\alpha$ 5	46	42	46	48	96	79	81
	Imp- $\alpha$ 6	48	44	47	49	81	94	84
	Imp- $\alpha$ 7	<b>48</b>	<b>42</b>	<b>49</b>	<b>49</b>	<b>82</b>	<b>85</b>	<b>95</b>
Frog	Imp- $\alpha$ 1	77	53	52	51	46	47	47
	Imp- $\alpha$ 8	62	60	49	49	43	44	44
	Imp- $\alpha$ 3	51	45	94	84	46	47	47
	Imp- $\alpha$ 4	50	45	85	95	48	50	49
	Imp- $\alpha$ 5	46	43	46	48	94	79	81
	Imp- $\alpha$ 7	49	44	48	51	79	83	90

## Annexe 2: Conservation of NP<sub>TAIL</sub> within each influenza type

**Table S2: Consensus residues and frequencies of A/NP<sub>TAIL</sub>.**

The sequences were recovered from the Influenza Research Database and processed with the Sequence Variation (SNP) tool. The frequency was determined manually.

Position	Consensus (57,774 seq)	Frequency	Polymorphism (> 10%)
1	Met	100 %	
2	Ala	99.9 %	
3	Ser	99.4 %	
4	Gln	99.9 %	
5	Gly	99.97 %	
6	Thr	99.97 %	
7	Lys	99.97 %	
8	Arg	99.99 %	
9	Ser	99.3 %	
10	Tyr	99.7 %	
11	Del	99.997 %	
12	Glu	99.98 %	
13	Gln	99.9 %	
14	Met	99.99 %	
15	Glu	99.9 %	
16	Thr	99.9 %	
17	Gly	66%	Asp (33 %)
18	Gly	99.9%	
19	Glu	72%	Asp (28 %)
20	Arg	99.95%	
21	Gln	99.9%	



**Table S3: Consensus residues and frequencies of B/NP<sub>TAIL</sub>.**

The sequences were recovered from the Influenza Research Database and processed with the Sequence Variation (SNP) tool. The frequency was determined manually.

Position	Consensus (9,209 seq)	Frequency	Polymorphism (> 10%)	Position	Consensus (9,209 seq)	Frequency	Polymorphism (> 10%)	Position	Consensus (9,209 seq)	Frequency	Polymorphism (> 10%)	
1	Met	100 %		25	Thr	98 %		49	Pro	100 %		
2	Ser	100 %		26	Ser	99.6 %		50	Ser	100 %		
3	Asn	99.99 %		27	Gly	99.97 %		51	Pro	99.8 %		
4	Met	100 %		28	Thr	99 %		52	Glu	99.9 %		
5	Asp	99.99 %		29	Thr	94.9 %		53	Arg	99.5 %		
6	Ile	99.96 %		30	Arg	99.97 %		54	Ala	99.6 %		
7	Asp	100 %		31	Pro	99.95 %		55	Thr	99.8 %		
8	Gly	99.5 %		32	Ile	99.9 %		56	Thr	99.8 %		
9	Ile	69 %		Met (30 %)	33	Ile		99.98 %	57	Ser		96 %
10	Asn	99.7 %			34	Arg		99.8 %	58	Ser		96 %
11	Thr	99.5 %	35		Pro	99.99 %	59	Glu	99.9 %			
12	Gly	99.9 %	36		Ala	99.98 %	60	Asp	92 %			
13	Thr	99.7 %	37		Thr	99.95 %	61	Asp	99.5 %			
14	Ile	99.7 %	38		Leu	100 %	62	Val	90 %			
15	Asp	99.8 %	39		Ala	99.9 %	63	Gly	98.95 %			
16	Lys	99.9 %	40		Pro	99.97 %	64	Arg	99 %			
17	Thr	94.7 %	41		Pro	99.95 %	65	Lys	99 %			
18	Pro	98.9 %	42		Ser	99.6 %	66	Thr	72 %			
19	Glu	99.95 %	43		Asn	99.9 %	67	Gln	99.8 %			
20	Glu	99.98 %	44	Lys	99.98 %	68	Lys	99.7 %				
21	Ile	99 %	45	Arg	100 %	69	Lys	99.95 %				
22	Thr	99 %	46	Thr	99.7 %	70	Gln	100 %				
23	Ser	99.7 %	47	Arg	99.99 %	71	Thr	99.99 %				
24	Gly	99.9 %	48	Asn	99.9 %							

**Table S4: Consensus residues and frequencies of C/NP<sub>TAIL</sub>.**

The sequences were recovered from the Influenza Research Database and processed with the Sequence Variation (SNP) tool. The frequency was determined manually.

Position	Consensus (50 seq)	Frequency	Polymorphism (> 10%)	Position	Consensus (50 seq)	Frequency	Polymorphism (> 10%)
503	Phe	100%		535	Gly	100%	
504	Glu	100%		536	Met	100%	
505	Phe	100%		537	Asp	100%	
506	Asp	100%		538	Glu	100%	
507	Pro	100%		539	Asn	100%	
508	Asp	100%		540	Ser	88%	Pro (12%)
509	Tyr	100%		541	Glu	100%	
510	Asn	100%		542	Ile	98%	
511	Pro	100%		543	Gly	100%	
512	Ile	100%		544	Gln	100%	
513	Arg	100%		545	Ala	100%	
514	Val	94%		546	Lys	100%	
515	Lys	100%		547	Lys	100%	
516	Arg	100%		548	Met	100%	
517	Pro	100%		549	Lys	100%	
518	Lys	100%		550	Pro	100%	
519	Lys	98%		551	Leu	98%	
520	Pro	100%		552	Asp	98%	
521	Ile	98%		553	Gln	98%	
522	Ala	94%		554	Leu	96%	
523	Lys	100%		555	Ala	98%	
524	Arg	100%		556	Ser	98%	
525	Asn	98%		557	Thr	100%	
526	Ser	100%		558	Ser	100%	
527	Asn	100%		559	Ser	98%	
528	Ile	100%		560	Asn	100%	
529	Ser	100%		561	Ile	100%	
530	Arg	100%		562	Pro	100%	
531	Leu	100%		563	Gly	100%	
532	Glu	100%		564	Glu	94%	
533	Glu	100%		565	Asn	100%	
534	Glu	100%					

**Table S5: Consensus residues and frequencies of D/NP<sub>TAIL</sub>.**

The sequences were recovered from the Influenza Reseach Database and processed with the Sequence Variation (SNP) tool. The frequency was determined manually.

Position	Consensus (34 seq)	Frequency	Position	Consensus (34 seq)	Frequency
501	Phe	100 %	527	Thr	100 %
502	Glu	100 %	528	Ser	100 %
503	Phe	100 %	529	Gln	94 %
504	Thr	100 %	530	Pro	100 %
505	Gly	100 %	531	Lys	100 %
506	Ser	100 %	532	Lys	100 %
507	Asp	100 %	533	Arg	100 %
508	Val	100 %	534	Gly	100 %
509	Pro	100 %	535	Arg	100 %
510	Arg	100 %	536	Gln	100 %
511	Thr	100 %	537	Gly	100 %
512	Gly	100 %	538	Ala	100 %
513	Ala	100 %	539	Gly	100 %
514	Lys	100 %	540	Ala	94 %
515	Arg	100 %	541	Glu	100 %
516	Arg	100 %	542	Ser	100 %
517	Val	94 %	543	Ser	100 %
518	Gly	100 %	544	Met	100 %
519	Gly	91 %	545	Asp	100 %
520	Ala	100 %	546	Ile	100 %
521	Asp	100 %	547	Glu	100 %
522	Asp	97 %	548	Thr	100 %
523	Val	91 %	549	Val	100 %
524	Thr	100 %	550	Gly	91 %
525	Pro	94 %	551	Glu	100 %
526	Gly	100 %	552	Asp	97 %

### Annexe 3: Protein accession numbers

**Table S6: Protein accession numbers.**

		NCBI	Uniprot
Human ( <i>Homo sapiens</i> )	Imp-α1	NP_001307540	P52292
	Imp-α8	NP_001139187.1	A9QM74
	Imp-α3	NP_002259.1	O00629
	Imp-α4	NP_002258.2	O00505
	Imp-α5	NP_002255.3	P52294
	Imp-α6	NP_001353233.1	O15131
	Imp-α7	NP_036448.1	O60684
Swine ( <i>Sus scrofa</i> )	Imp-α1	NP_001156876.1	
	Imp-α8	NP_001156883.1	
	Imp-α3	XP_020925394.1	
	Imp-α4	NP_001177159.1	
	Imp-α5	NP_001156877.1	
	Imp-α6	XP_005654449.1	
	Imp-α7	XP_020951465.1	
Bovine ( <i>Bos Taurus</i> )	Imp-α1	NP_001029621.1	
	Imp-α8	NP_001157419.1	
	Imp-α3	NP_001152788.1	
	Imp-α4	NP_001179702.3	
	Imp-α5	NP_001075202.1	
	Imp-α6	XP_002690143.2	
	Imp-α7	NP_001069422.1	
Chicken ( <i>Gallus gallus</i> )	Imp-α1	NP_001006209.1	
	Imp-α8	XP_414795.3	
	Imp-α3	NP_001007964.1	
	Imp-α4	NP_001180504.1	
	Imp-α5	NP_001025945.1	
	Imp-α6	XP_419770.3	
	Imp-α7	NP_001012859.2	
Duck ( <i>Anas platyrhynchos</i> )	Imp-α1	XP_005012369.2	
	Imp-α8	XP_012955615.2	
	Imp-α3	XP_027320524.1	
	Imp-α4	XP_027327221.1	
	Imp-α5	XP_027305514.1	
	Imp-α6	XP_027310463.1	
	Imp-α7		U3IY73
Frog ( <i>Xenopus tropicalis</i> )	Imp-α1	NP_001008155.1	
	Imp-α8	NP_001135557.1	
	Imp-α3	XP_002933281.2	
	Imp-α4	NP_001016060.1	
	Imp-α5	NP_989192.1	
	Imp-α7	NP_001008018.1	
Influenza viruses	A/NP	5TJW6_A	
	B/NP	AAU94830	
	C/NP	YP_089656	
	D/NP	CEE50063	
	Toad/NP	AVM87634	
	Eel/NP	AVM87623	
	Hagfish/NP	AVM87638	
	Salamander/NP	QOE76812	
	Salmon/NP	QOE76828	
	Algae Eater/NP	QOE76804	
Cane Toad/NP	QOE76821		
Frog/NP	QOE76796		

---

## **Annexe 4: Supplementary Data Chapter 2**

**The structure of the nucleoprotein of Influenza D shows that all *Orthomyxoviridae* nucleoproteins have a similar NP<sub>CORE</sub>, with or without a NP<sub>TAIL</sub> for nuclear transport**

Amélie DONCHET<sup>1</sup>, Justine OLIVA<sup>2</sup>, Alice LABARONNE<sup>1</sup>, Laura TENGO<sup>1</sup>, Myriam MILOUDI<sup>1</sup>, Francine C. A. GERARD<sup>1</sup>, Caroline MAS<sup>3</sup>, Guy SCHOEHN<sup>1</sup>, Rob W. H. RUIGROK<sup>1</sup>, Mariette DUCATEZ<sup>2</sup> and Thibaut CREPIN<sup>1\*</sup>

<sup>1</sup> Institut de Biologie Structurale (IBS), Univ. Grenoble Alpes, CEA, CNRS, 38044 Grenoble, France

<sup>2</sup> IHAP, Université de Toulouse, INRA, ENVT, Toulouse, France

<sup>3</sup> Integrated Structural Biology Grenoble (ISBG) - UMS 3518 (CNRS-CEA-UJF-EMBL), 38044 Grenoble, France

**Supplementary information**

---

**Supplementary Figure 1**

**Sequence alignment of NP from representative Influenza strains.** The sequences used are the same as in Table 1. The figure has been drawn using ESPript <sup>1</sup>. The upper and lower secondary structures correspond to the X-ray structures of D/NP (PDB entry 5N2U) and A/NP (PDB entry 2IQH), respectively. The stars indicate the phosphorylation sites identified for A/NP (closed symbol) and B/NP (open symbol) by Hutchinson *et al.* (red)<sup>2</sup>, Li *et al.* (green)<sup>3</sup> and Monda *et al.* (blue)<sup>4</sup>.

**Supplementary Figure 2**

**Effect of trimethylamine N-oxide (TMAO) on the folding of D/NP<sub>TAIL</sub>.** Far UV-circular dichroism spectra of either D/NP<sub>TAIL</sub> alone (purple) or in presence of TMAO (0.5 M in yellow, 1M in red, 2M in blue and 3 M in green) were collected. Because of the high background induced by the presence of TMAO around 200 nm, spectra were cut when the voltage (right bottom insert) became too high (> 700 V).

## References

- 1 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).
- 2 Hutchinson, E. C. *et al.* Mapping the phosphoproteome of influenza A and B viruses by mass spectrometry. *PLoS Pathog* **8**, e1002993, doi:10.1371/journal.ppat.1002993 (2012).
- 3 Li, Y. *et al.* Phosphorylation and dephosphorylation of threonine 188 in nucleoprotein is crucial for the replication of influenza A virus. *Virology* **520**, 30-38, doi:10.1016/j.virol.2018.05.002 (2018).
- 4 Mondal, A. *et al.* Influenza virus recruits host protein kinase C to control assembly and activity of its replication machinery. *Elife* **6**, doi:10.7554/eLife.26910 (2017).
- 5 Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**, 12-21, doi:10.1107/S0907444909042073 (2010).

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

Crystal Parameters	D/NP
<b>Data collection</b>	
Beamline	ID30A3 - ESRF
Wavelength (Å)	0.9677
Space group	P1
Cell Dimensions a, b, c (Å)	75.17, 85.17, 103.39
$\alpha, \beta, \gamma$ (°)	91.20, 101.94, 101.02
Molecules per asymmetric unit	4
Wilson B-factor (Å <sup>2</sup> )	51.4
Resolution (Å)	40 - 2.4 (2.48 - 2.4)
Observed Reflections	191 888 (18 695)
Unique Reflections	87 824 (8 505)
Completeness (%)	91.3 (94.1)
$R_{\text{sym}}(I)^{(a)}$ (%)	9.3 (84.7)
$\langle I/\sigma(I) \rangle$	7.78 (1.05)
<b>Model quality indicators</b>	
$R_{\text{cryst}}$ (%)	20.2 (21.9)
$R_{\text{free}}$ (%)	25.2 (23.8)
Non-hydrogen atoms	14 646
Water atoms	235
Average B-factors (Å <sup>2</sup> )	55.7
B-factors per chains (Å <sup>2</sup> )	59.8, 54.3, 52.8, 56.3
B-factors water (Å <sup>2</sup> )	49.5
Ramachandran favoured / allowed (%)	94.0 / 5.1
rms deviations, bonds (Å) / angles (°)	0.009 / 1.15
Clashscore <sup>(b)</sup>	4.31
MolProbity score <sup>(b)</sup>	2.28

<sup>(a)</sup>  $R_{\text{sym}} = \frac{\sum_{hkl} \sum_j |I_{hkl,j} - \langle I_{hkl} \rangle|}{\sum_{hkl} \sum_j I_{hkl,j}}$

<sup>(b)</sup> calculated using MolProbity <sup>5</sup>



D/NP .....  
 C/NP .....  
 toad/NP MATQEQA...  
 B/NP .....  
 A/NP .....  
 MSNMDIDGINTGTIDKAPE

D/NP ..... MDSTKAC...  
 C/NP ..... MSDRRQNRKT...  
 toad/NP EERDENMSEED...  
 B/NP EITSGTSGTTRP...  
 A/NP ..... MATKGT...  
 \* \* \*

D/NP VSE...  
 C/NP LNG...  
 toad/NP LATA...  
 B/NP LND...  
 A/NP LSY...  
 \* \* \*

D/NP VEAL...  
 C/NP IVQ...  
 toad/NP LMD...  
 B/NP VAE...  
 A/NP LRR...  
 \* \* A/NP Ser-165 \* \*

D/NP VAS...  
 C/NP IAD...  
 toad/NP ASRA...  
 B/NP LARS...  
 A/NP VME...  
 \* \* \*

D/NP MLD...  
 C/NP IAE...  
 toad/NP IAQC...  
 B/NP LARS...  
 A/NP LARS...  
 \* \* \*

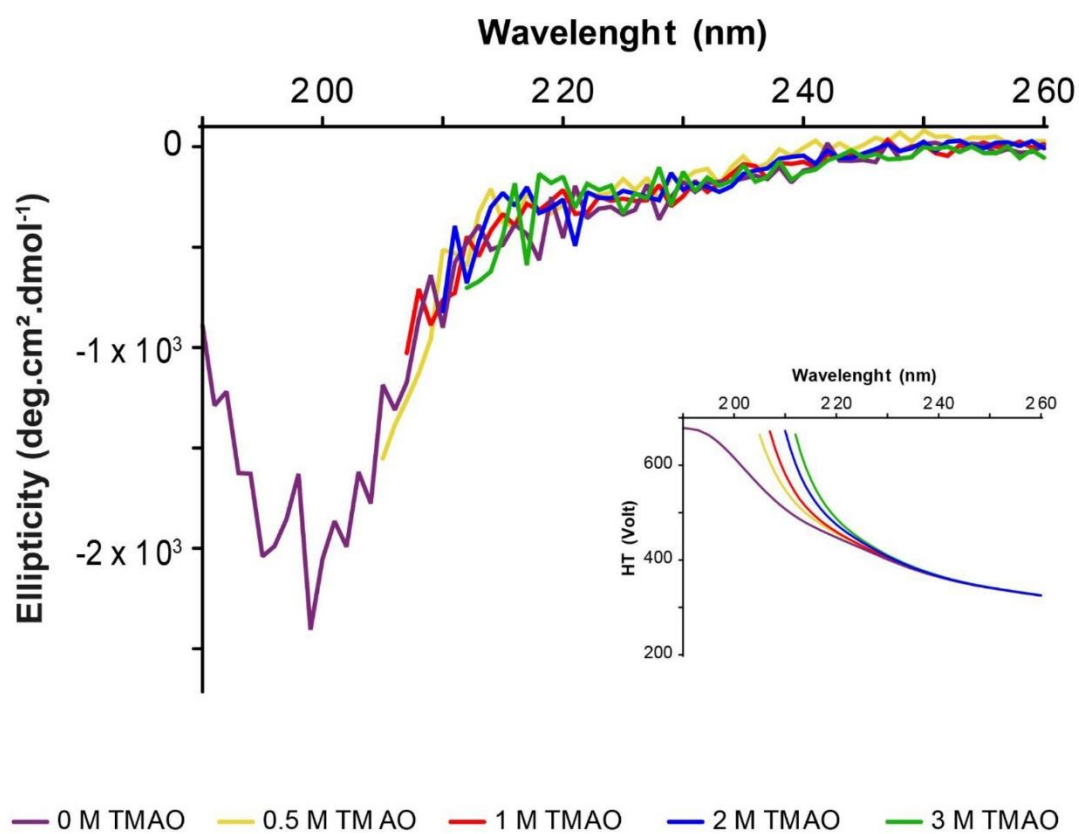
D/NP NDK...  
 C/NP EDK...  
 toad/NP NDK...  
 B/NP DKRS...  
 A/NP AHS...  
 \* \* \*

exchange domain

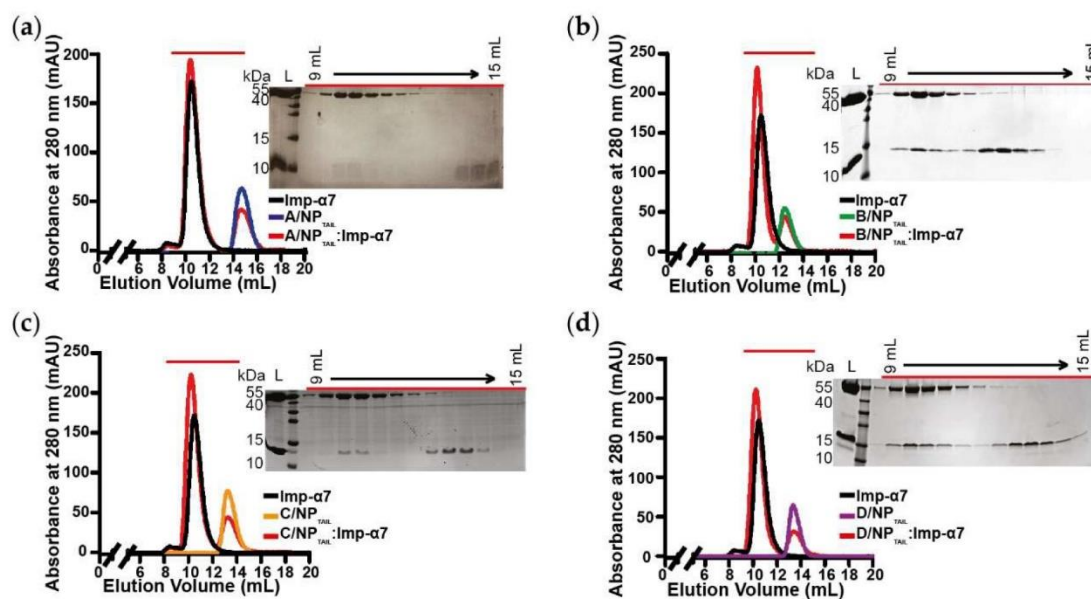
D/NP TGES...  
 C/NP SREG...  
 toad/NP DERG...  
 B/NP DGG...  
 A/NP RASS...  
 \* \* A/NP Ser-407 \* \*

D/NP NPK...  
 C/NP NPN...  
 toad/NP RPP...  
 B/NP NPVE...  
 A/NP SPV...  
 \* \* \*

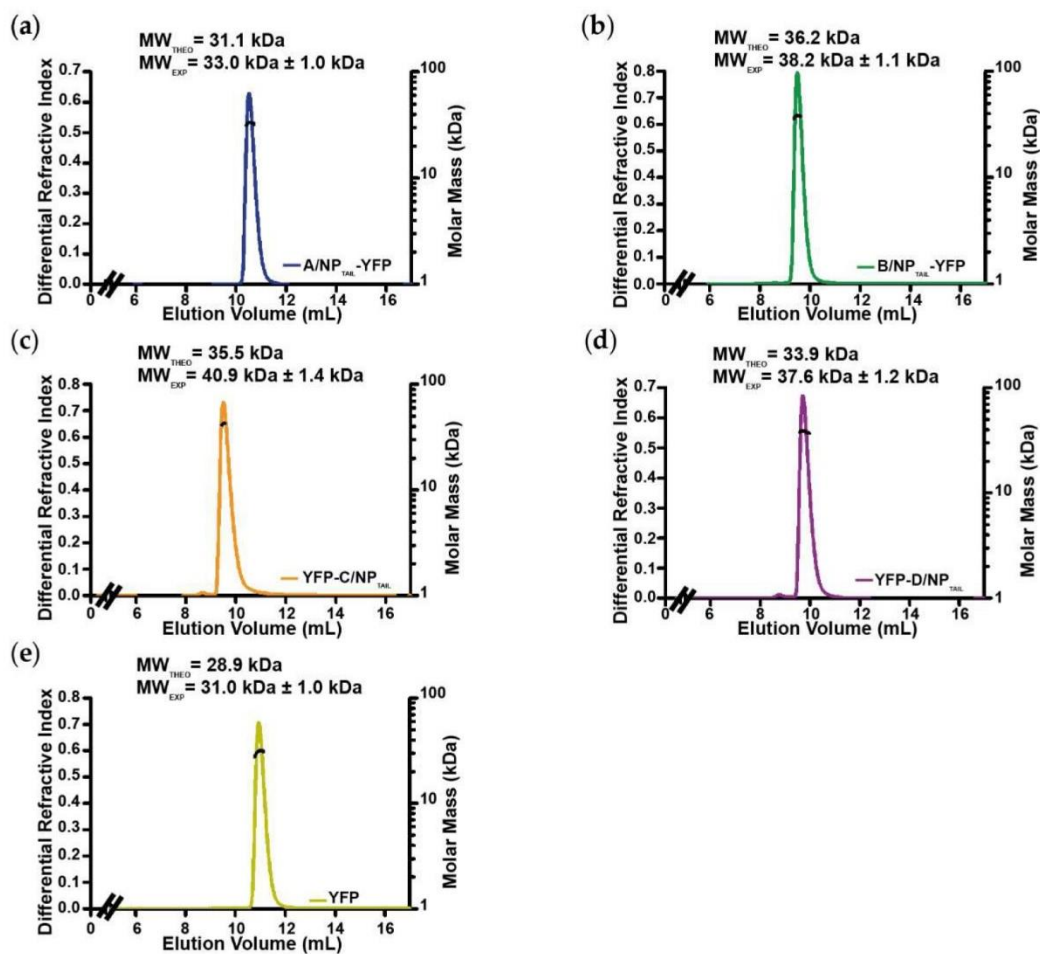
Supplementary Figure 2



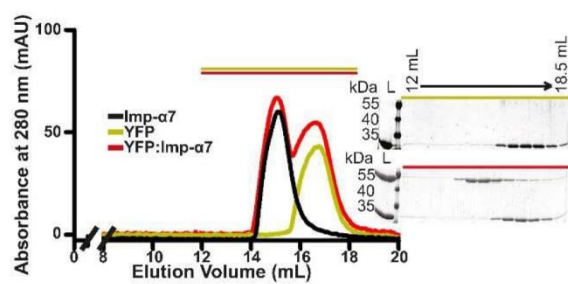
## Annexe 5: Supplementary Data Chapter 3



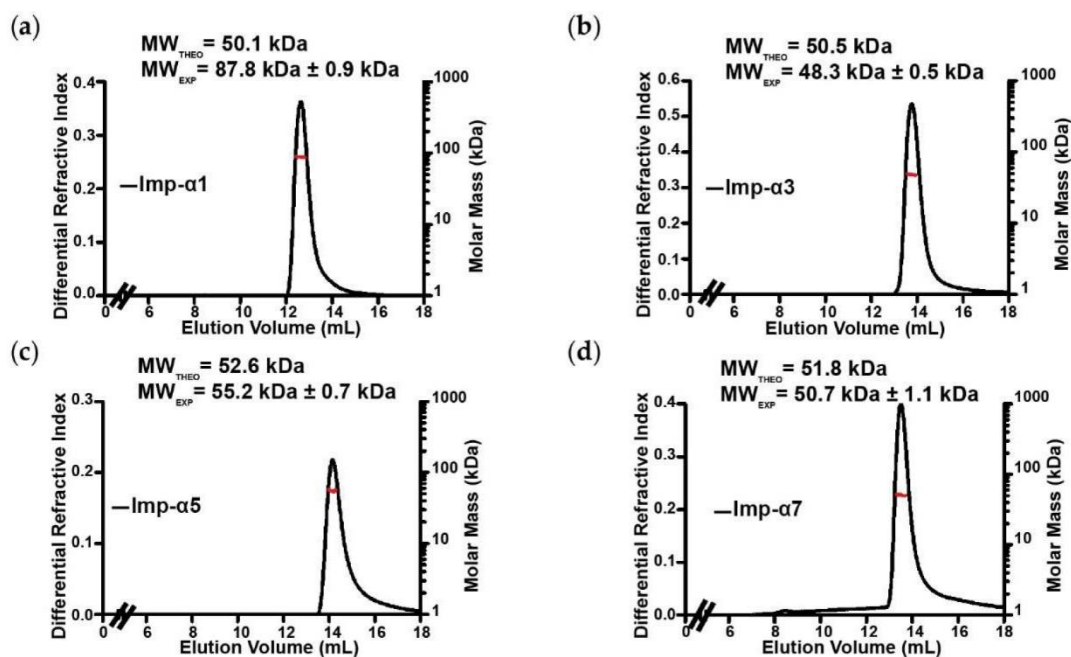
**Figure S1.** Analysis of the complexes between NP<sub>TAIL</sub> and human importin- $\alpha$ 7 in solution. Each NP<sub>TAIL</sub> (60  $\mu$ M) was injected on a Superdex<sup>TM</sup> 75 10/300GL column, alone or in presence of importin- $\alpha$ 7 (25  $\mu$ M). The figure shows the superimposition of the size exclusion chromatography profiles obtained for (a) A/NP<sub>TAIL</sub>, (b) B/NP<sub>TAIL</sub>, (c) C/NP<sub>TAIL</sub> and (d) D/NP<sub>TAIL</sub>. The colors correspond to the code used on Figure 1, with the complexes in red and the importin- $\alpha$ 7 alone in black. For each panel, the 18 % SDS-PAGE corresponding to the elution of the complex is shown.



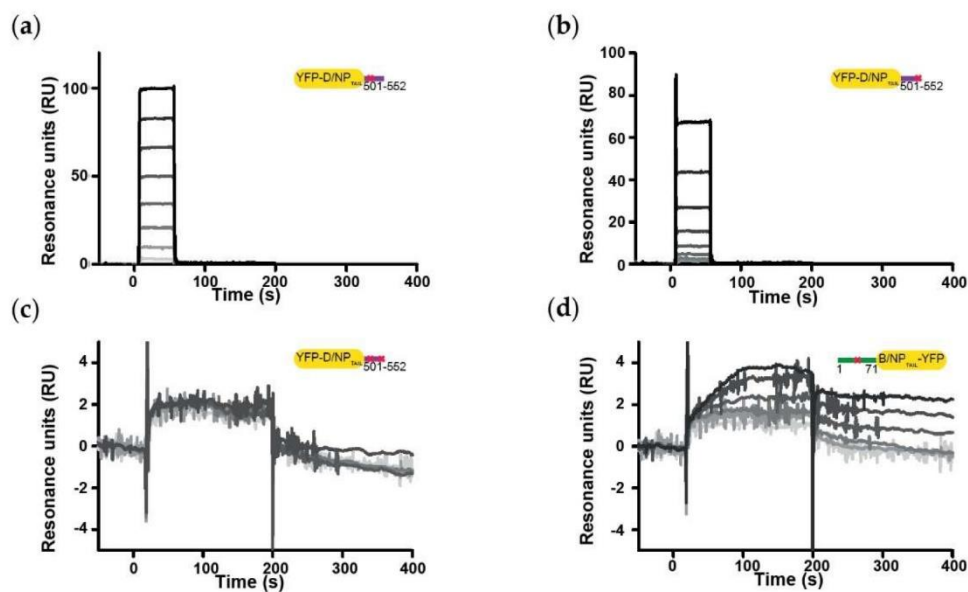
**Figure S2.** SEC-MALLS-RI analysis of the different YFP-fused tails used in that work. All the samples were injected on a Superdex<sup>TM</sup> 75 10/300GL column. Each panel indicates the theoretical (top) and experimental (bottom) molecular weights of (a) A/NP<sub>TAIL</sub>-YFP, (b) B/NP<sub>TAIL</sub>-YFP, (c) YFP-C/NP<sub>TAIL</sub>, (d) YFP-D/NP<sub>TAIL</sub> and (e) the control YFP. The black lines indicate the molecular weight as estimated below the peak.



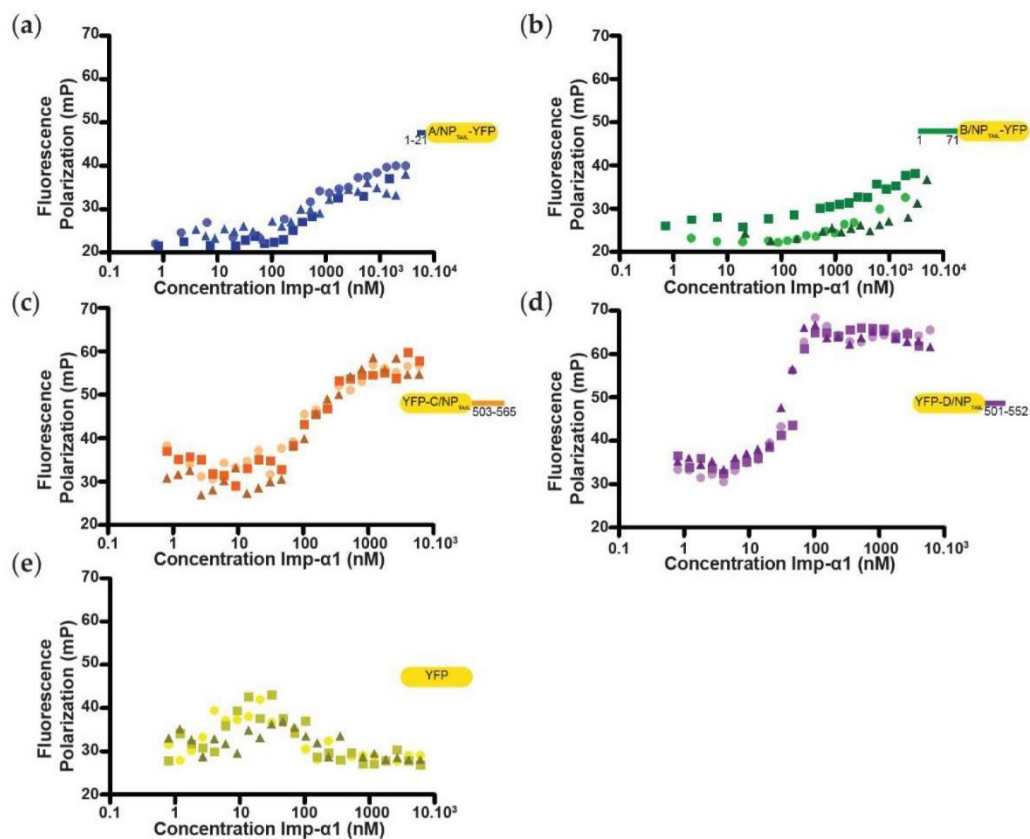
**Figure S3.** Absence of interaction between YFP and human importin- $\alpha$ 7 in solution. The YFP (30  $\mu$ M) was injected on a Superdex™ 200 increase 10/300GL column, alone or in presence of importin- $\alpha$ 7 (25  $\mu$ M). The 12 % SDS-PAGE corresponding to the elution of the YFP alone (top) and the sample containing both YFP and importin- $\alpha$ 7 (bottom) are shown.



**Figure S4.** SEC-MALLS-RI analysis of the different importins- $\alpha$  used in that work. All the samples were injected on a Superdex™ 200 increase 10/300GL column. Each panel indicates the theoretical (top) and experimental (bottom) molecular weights of (a) importin- $\alpha$ 1, (b) importin- $\alpha$ 3, (c) importin- $\alpha$ 5 and (d) importin- $\alpha$ 7. The red lines indicate the molecular weight as estimated below the peak.



**Figure S5.** Interaction between human importin- $\alpha$ 7 and YFP-NP<sub>TAILS</sub> mutants. The figure shows the sensorgrams of the interaction between importin- $\alpha$ 7 and (a) YFP-D/NP<sub>TAILS</sub> mut1, (b) YFP-D/NP<sub>TAILS</sub> mut2, (c) YFP-D/NP<sub>TAILS</sub> mut3, and (d) YFP-B/NP<sub>TAILS</sub> mut1. For each panel, the gradation of grey represents the different concentrations of importin- $\alpha$ 7 used for the titrations, from the lowest (light-grey) to the highest (dark-grey). The concentrations range from 80 nM to 10.3  $\mu$ M for (a) and (c) and from 41 nM to 4  $\mu$ M for (b) and (d). For a more accurate comparison, (c) and (d) sensorgrams were recovered from the same kinetic assay.



**Figure S6.** Interaction between human importin- $\alpha 1$  and YFP-NP<sub>TAILS</sub>. The figure shows the fluorescence polarization raw data of the interaction between importin- $\alpha 1$  and (a) A/NP<sub>TAIL</sub>-YFP, (b) B/NP<sub>TAIL</sub>-YFP, (c) YFP-C/NP<sub>TAIL</sub>, (d) YFP-D/NP<sub>TAIL</sub> and (e) YFP control. For each panel, one shade of colour with one specific symbol represent one of the three replicate.



**Table S1:** Sequences. The table details the sequences of the nucleoproteins tails and biotinylation motif used in this work. The basic residues of the tails are shown in blue and the corresponding mutations in red. Biotinylated lysine are in green.

---

<b>A/NP<sub>TAIL</sub></b>	
WT	1MATKGT <b>K</b> RSYEQMETDGERQN <sub>21</sub>
<b>B/NP<sub>TAIL</sub></b>	
WT	1MSNMDIDGINTGTIDKAP <b>E</b> ITS <b>G</b> TS <b>G</b> TRPIIRPATLAPPSN <b>K</b> RTRN <b>P</b> SPERATTISEADV <b>G</b> R <b>K</b> TQ <b>K</b> QT <sub>71</sub>
B/NP <sub>TAIL</sub> . mut1	1MSNMDIDGINTGTIDKAP <b>E</b> ITS <b>G</b> TS <b>G</b> TRPIIRPATLAPPSN <b>A</b> ATRN <b>P</b> SPERATTISEADV <b>G</b> R <b>K</b> TQ <b>K</b> QT <sub>71</sub>
<b>C/NP<sub>TAIL</sub></b>	
WT	503FEFDPDYNPIRV <b>K</b> R <b>P</b> K <b>K</b> PIA <b>K</b> RNSNISRLEEE <b>G</b> MDENSEIGQA <b>K</b> K <b>M</b> KPLDQLTSTSSNIPGKN <sub>565</sub>
<b>D/NP<sub>TAIL</sub></b>	
WT	501FEFTGSDVPRTGAKRRVGGADDVTLGTSQP <b>K</b> K <b>R</b> GRQGAGVESSMDIETVGED <sub>552</sub>
D/NP <sub>TAIL</sub> . mut1	501FEFTGSDVPRTGAA <b>A</b> RVGGADDVTLGTSQP <b>K</b> K <b>R</b> GRQGAGVESSMDIETVGED <sub>552</sub>
D/NP <sub>TAIL</sub> . mut2	501FEFTGSDVPRTGAKRRVGGADDVTLGTSQP <b>A</b> A <b>A</b> GAQGAGVESSMDIETVGED <sub>552</sub>
D/NP <sub>TAIL</sub> . mut3	501FEFTGSDVPRTGAA <b>A</b> RVGGADDVTLGTSQP <b>A</b> A <b>A</b> GAQGAGVESSMDIETVGED <sub>552</sub>
<b>Synthetic peptide</b>	
control	1LEEM <b>K</b> K <b>G</b> HLE <b>R</b> ECMEETCSYEEA <b>R</b> EVFEDSEKTNEFWN <b>K</b> <sub>39</sub>
<b>Biotinylation sequences</b>	
N-terminal	DIFEAQ <b>K</b> IEWHEGGNGSGGGLN
C-terminal	NGSGGGLNDIFEAQ <b>K</b> IEWHE

---

**Table S2:** Sequence identity between the importins- $\alpha$  used in the studies. Sequences and accession numbers were recovered from NCBI/Uniprot databases for human (*Homo sapiens*) importin- $\alpha$ 1 (NP\_001307540 / P52292), importin- $\alpha$ 3 (NP\_002259.1 / O00629), importin- $\alpha$ 5 (NP\_002255.3 / P52294) and importin- $\alpha$ 7 (NP\_036448.1 / O60684). Each BLAST was done separately, using one human importin sequence against reference protein sequences from *Sus scrofa* (swine), *Bos Taurus* (bovine), *Gallus gallus* (chicken) and *Anas platyrhynchos* (duck). As no duck importin- $\alpha$ 7 was identified in the NCBI database, the corresponding sequence was recovered from Uniprot database (Uniprot accession number U3IY73) and aligned with each human importin alpha to get a sequence identity by global alignment (bold), 99+ corresponds to a sequence identity between 99.5 and 100 %. Importins- $\alpha$  from the same subfamily are highlighted in the same colour.

	Subfamily Member	Human			
		sub- $\alpha$ 1	sub- $\alpha$ 2	sub- $\alpha$ 3	
		Imp- $\alpha$ 1	Imp- $\alpha$ 3	Imp- $\alpha$ 5	Imp- $\alpha$ 7
Human	Imp- $\alpha$ 1	100	51	46	48
	Imp- $\alpha$ 3		100	47	49
	Imp- $\alpha$ 5			100	82
	Imp- $\alpha$ 7				100
Swine	Imp- $\alpha$ 1	97	50	45	47
	Imp- $\alpha$ 3	52	99+	47	48
	Imp- $\alpha$ 5	46	46	99	81
	Imp- $\alpha$ 7	48	48	81	99
Bovine	Imp- $\alpha$ 1	95	82	46	47
	Imp- $\alpha$ 3	51	99+	47	48
	Imp- $\alpha$ 5	46	46	99	81
	Imp- $\alpha$ 7	48	48	81	99
Chicken	Imp- $\alpha$ 1	83	51	46	47
	Imp- $\alpha$ 3	51	99	47	48
	Imp- $\alpha$ 5	46	46	96	81
	Imp- $\alpha$ 7	48	48	81	94
Duck	Imp- $\alpha$ 1	84	51	46	46
	Imp- $\alpha$ 3	51	98	47	48
	Imp- $\alpha$ 5	46	46	96	81
	Imp- $\alpha$ 7	<b>46</b>	<b>48</b>	<b>82</b>	<b>94</b>

# Bibliography

1. Adam EJ, Adam SA. Identification of cytosolic factors required for nuclear location sequence-mediated binding to the nuclear envelope. *J Cell Biol.* 1994;125(3):547-555. doi:10.1083/jcb.125.3.547
2. Akarsu H, Burmeister WP, Petosa C, *et al.* Crystal structure of the M1 protein-binding domain of the influenza A virus nuclear export protein (NEP/NS2). *EMBO J.* 2003;22(18):4646-4655. doi:10.1093/emboj/cdg449
3. Akkina RK, Chambers TM, Londo DR, Nayak DP. Intracellular localization of the viral polymerase proteins in cells infected with influenza virus and cells expressing PB1 protein from cloned cDNA. *J Virol.* 1987;61(7):2217-2224. doi:10.1128/JVI.61.7.2217-2224.1987
4. Alber F, Dokudovskaya S, Veenhoff LM, *et al.* The molecular architecture of the nuclear pore complex. *Nature.* 2007;450(7170):695-701. doi:10.1038/nature06405
5. Andrade MA, Petosa C, O'Donoghue SI, Müller CW, Bork P. Comparison of ARM and HEAT protein repeats. *J Mol Biol.* 2001;309(1):1-18. doi:10.1006/jmbi.2001.4624
6. Armstrong EH, Goswami D, Griffin PR, Noy N, Ortlund EA. Structural basis for ligand regulation of the fatty acid-binding protein 5, peroxisome proliferator-activated receptor  $\beta/\delta$  (FABP5-PPAR $\beta/\delta$ ) signaling pathway. *J Biol Chem.* 2014;289(21):14941-14954. doi:10.1074/jbc.M113.514646
7. Arranz R, Coloma R, Chichón FJ, *et al.* The structure of native influenza virion ribonucleoproteins. *Science.* 2012;338(6114):1634-1637. doi:10.1126/science.1228172
8. Asha K, Kumar B. Emerging Influenza D Virus Threat: What We Know so Far! *J Clin Med.* 2019;8(2). doi:10.3390/jcm8020192
9. Audsley MD, Jans DA, Moseley GW. Roles of nuclear trafficking in infection by cytoplasmic negative-strand RNA viruses: paramyxoviruses and beyond. *J Gen Virol.* 2016;97(10):2463-2481. doi:10.1099/jgv.0.000575
10. Azimi M, Mofrad MRK. Higher Nucleoporin-Importin $\beta$  Affinity at the Nuclear Basket Increases Nucleocytoplasmic Import. *PLOS ONE.* 2013;8(11):e81741. doi:10.1371/journal.pone.0081741
11. Babcock HP, Chen C, Zhuang X. Using Single-Particle Tracking to Study Nuclear Trafficking of Viral Genes. *Biophys J.* 2004;87(4):2749-2758. doi:10.1529/biophysj.104.042234
12. Bannister AJ, Miska EA, Görlich D, Kouzarides T. Acetylation of importin-alpha nuclear import factors by CBP/p300. *Curr Biol CB.* 2000;10(8):467-470. doi:10.1016/s0960-9822(00)00445-0
13. Baudin F, Bach C, Cusack S, Ruigrok RW. 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(13):3158-3165.
14. Bauer NC, Doetsch PW, Corbett AH. Mechanisms Regulating Protein Localization. *Traffic.* 2015;16(10):1039-1061. doi:10.1111/tra.12310
15. Beaton AR, Krug RM. Transcription antitermination during influenza viral template RNA synthesis requires the nucleocapsid protein and the absence of a 5' capped end. *Proc Natl Acad Sci U S A.* 1986;83(17):6282-6286.
16. Beck M, Glavy JS. Toward understanding the structure of the vertebrate nuclear pore complex. *Nucleus.* 2014;5(2):119-123. doi:10.4161/nucl.28739
17. Bedford MT, Clarke SG. Protein Arginine Methylation in Mammals: Who, What, and Why. *Mol Cell.* 2009;33(1):1-13. doi:10.1016/j.molcel.2008.12.013
18. Ben-Efraim I, Gerace L. Gradient of increasing affinity of importin beta for nucleoporins along the pathway of nuclear import. *J Cell Biol.* 2001;152(2):411-417. doi:10.1083/jcb.152.2.411

19. Ben-Yaakov K, Dagan SY, Segal-Ruder Y, *et al.* Axonal transcription factors signal retrogradely in lesioned peripheral nerve. *EMBO J.* 2012;31(6):1350-1363. doi:10.1038/emboj.2011.494
20. Bernetti M, Cavalli A, Mollica L. Protein–ligand (un)binding kinetics as a new paradigm for drug discovery at the crossroad between experiments and modelling. *MedChemComm.* 2017;8(3):534-550. doi:10.1039/C6MD00581K
21. Betáková T, Kollerová E. pH modulating activity of ion channels of influenza A, B, and C viruses. *Acta Virol.* 2006;50(3):187-193.
22. Bischoff FR, Ponstingl H. Catalysis of guanine nucleotide exchange on Ran by the mitotic regulator RCC1. *Nature.* 1991;354(6348):80-82. doi:10.1038/354080a0
23. Boivin S, Cusack S, Ruigrok RWH, Hart DJ. Influenza A virus polymerase: structural insights into replication and host adaptation mechanisms. *J Biol Chem.* 2010;285(37):28411-28417. doi:10.1074/jbc.R110.117531
24. Boivin S, Hart DJ. Interaction of the influenza A virus polymerase PB2 C-terminal region with importin alpha isoforms provides insights into host adaptation and polymerase assembly. *J Biol Chem.* 2011;286(12):10439-10448. doi:10.1074/jbc.M110.182964
25. Bolte H, Rosu ME, Hagelauer E, García-Sastre A, Schwemmler M. Packaging of the Influenza Virus Genome Is Governed by a Plastic Network of RNA- and Nucleoprotein-Mediated Interactions. *J Virol.* 2019;93(4). doi:10.1128/JVI.01861-18
26. Booy FP, Ruigrok RW, van Bruggen EF. Electron microscopy of influenza virus. A comparison of negatively stained and ice-embedded particles. *J Mol Biol.* 1985;184(4):667-676. doi:10.1016/0022-2836(85)90312-2
27. Boulo S, Akarsu H, Lotteau V, Müller CW, Ruigrok RWH, Baudin F. Human importin alpha and RNA do not compete for binding to influenza A virus nucleoprotein. *Virology.* 2011;409(1):84-90. doi:10.1016/j.virol.2010.10.001
28. Boulo S, Akarsu H, Ruigrok RWH, Baudin F. Nuclear traffic of influenza virus proteins and ribonucleoprotein complexes. *Virus Res.* 2007;124(1-2):12-21. doi:10.1016/j.virusres.2006.09.013
29. Bouloy M, Plotch SJ, Krug RM. Globin mRNAs are primers for the transcription of influenza viral RNA *in vitro*. *Proc Natl Acad Sci U S A.* 1978;75(10):4886-4890.
30. Bui KH, von Appen A, DiGiulio AL, *et al.* Integrated structural analysis of the human nuclear pore complex scaffold. *Cell.* 2013;155(6):1233-1243. doi:10.1016/j.cell.2013.10.055
31. Calder LJ, Wasilewski S, Berriman JA, Rosenthal PB. Structural organization of a filamentous influenza A virus. *Proc Natl Acad Sci U S A.* 2010;107(23):10685-10690. doi:10.1073/pnas.1002123107
32. Campbell EM, Hope TJ. Role of the cytoskeleton in nuclear import. *Adv Drug Deliv Rev.* 2003;55(6):761-771. doi:10.1016/S0169-409X(03)00049-8
33. Cardarelli F, Bizzarri R, Serresi M, Albertazzi L, Beltram F. Probing Nuclear Localization Signal-Importin  $\alpha$  Binding Equilibria in Living Cells. *J Biol Chem.* 2009;284(52):36638-36646. doi:10.1074/jbc.M109.036699
34. Cardarelli F, Tosti L, Serresi M, Beltram F, Bizzarri R. Fluorescent Recovery after Photobleaching (FRAP) Analysis of Nuclear Export Rates Identifies Intrinsic Features of Nucleocytoplasmic Transport. *J Biol Chem.* 2012;287(8):5554-5561. doi:10.1074/jbc.M111.304899
35. Catimel B, Teh T, Fontes MR, *et al.* Biophysical characterization of interactions involving importin-alpha during nuclear import. *J Biol Chem.* 2001;276(36):34189-34198. doi:10.1074/jbc.M103531200

36. Chan JF-W, To KK-W, Tse H, Jin D-Y, Yuen K-Y. Interspecies transmission and emergence of novel viruses: lessons from bats and birds. *Trends Microbiol.* 2013;21(10):544-555. doi:10.1016/j.tim.2013.05.005
37. Chase GP, Rameix-Welti M-A, Zvirbliene A, *et al.* Influenza Virus Ribonucleoprotein Complexes Gain Preferential Access to Cellular Export Machinery through Chromatin Targeting. *PLoS Pathog.* 2011;7(9). doi:10.1371/journal.ppat.1002187
38. Chen M-H, Ben-Efraim I, Mitrousis G, Walker-Kopp N, Sims PJ, Cingolani G. Phospholipid scramblase 1 contains a nonclassical nuclear localization signal with unique binding site in importin alpha. *J Biol Chem.* 2005;280(11):10599-10606. doi:10.1074/jbc.M413194200
39. Chenavas S, Crépin T, Delmas B, Ruigrok RW, Slama-Schwok A. Influenza virus nucleoprotein: structure, RNA binding, oligomerization and antiviral drug target. *Future Microbiol.* 2013;8(12):1537-1545. doi:10.2217/fmb.13.128
40. Chenavas S, Estrozi LF, Slama-Schwok A, *et al.* Monomeric nucleoprotein of influenza A virus. *PLoS Pathog.* 2013;9(3):e1003275. doi:10.1371/journal.ppat.1003275
41. Chook YM, Süel KE. Nuclear import by Karyopherin- $\beta$ s: recognition and inhibition. *Biochim Biophys Acta.* 2011;1813(9):1593-1606. doi:10.1016/j.bbamcr.2010.10.014
42. Choppin PW, Murphy JS, Tamm I. Studies of two kinds of virus particles which comprise influenza A2 virus strains. III. Morphological characteristics: independence to morphological and functional traits. *J Exp Med.* 1960;112:945-952. doi:10.1084/jem.112.5.945
43. Chu CM, Dawson IM, Elford WJ. Filamentous forms associated with newly isolated influenza virus. *Lancet Lond Engl.* 1949;1(6554):602. doi:10.1016/s0140-6736(49)91699-2
44. Cingolani G, Petosa C, Weis K, Müller CW. Structure of importin- $\beta$  bound to the IBB domain of importin- $\alpha$ . *Nature.* 1999;399(6733):221-229. doi:10.1038/20367
45. Coletta A, Pinney JW, Solís DYW, Marsh J, Pettifer SR, Attwood TK. Low-complexity regions within protein sequences have position-dependent roles. *BMC Syst Biol.* 2010;4:43. doi:10.1186/1752-0509-4-43
46. Compans RW, Content J, Duesberg PH. Structure of the Ribonucleoprotein of Influenza Virus. *J Virol.* 1972;10(4):795-800.
47. Connor RJ, Kawaoka Y, Webster RG, Paulson JC. Receptor specificity in human, avian, and equine H2 and H3 influenza virus isolates. *Virology.* 1994;205(1):17-23. doi:10.1006/viro.1994.1615
48. Conti E, Uy M, Leighton L, Blobel G, Kuriyan J. Crystallographic Analysis of the Recognition of a Nuclear Localization Signal by the Nuclear Import Factor Karyopherin  $\alpha$ . *Cell.* 1998;94(2):193-204. doi:10.1016/S0092-8674(00)81419-1
49. Cook A, Bono F, Jinek M, Conti E. Structural biology of nucleocytoplasmic transport. *Annu Rev Biochem.* 2007;76:647-671. doi:10.1146/annurev.biochem.76.052705.161529
50. Cook AG, Conti E. Nuclear export complexes in the frame. *Curr Opin Struct Biol.* 2010;20(2):247-252. doi:10.1016/j.sbi.2010.01.012
51. Corzo J. Time, the forgotten dimension of ligand binding teaching. *Biochem Mol Biol Educ Bimon Publ Int Union Biochem Mol Biol.* 2006;34(6):413-416. doi:10.1002/bmb.2006.494034062678
52. Cremer T, Cremer M, Cremer C. The 4D Nucleome: Genome Compartmentalization in an Evolutionary Context. *Biochem Mosc.* 2018;83(4):313-325. doi:10.1134/S000629791804003X
53. Cros JF, García-Sastre A, Palese P. An Unconventional NLS is Critical for the Nuclear Import of the Influenza A Virus Nucleoprotein and Ribonucleoprotein. *Traffic.* 2005;6(3):205-213. doi:10.1111/j.1600-0854.2005.00263.x

54. Cros JF, Palese P. Trafficking of viral genomic RNA into and out of the nucleus: influenza, Thogoto and Borna disease viruses. *Virus Res.* 2003;95(1-2):3-12. doi:10.1016/s0168-1702(03)00159-x
55. Da Costa B, Sausset A, Munier S, *et al.* Temperature-Sensitive Mutants in the Influenza A Virus RNA Polymerase: Alterations in the PA Linker Reduce Nuclear Targeting of the PB1-PA Dimer and Result in Viral Attenuation. *J Virol.* 2015;89(12):6376-6390. doi:10.1128/JVI.00589-15
56. Davis LI. The nuclear pore complex. *Annu Rev Biochem.* 1995;64:865-896. doi:10.1146/annurev.bi.64.070195.004245
57. Deng T, Engelhardt OG, Thomas B, Akoulitchev AV, Brownlee GG, Fodor E. Role of Ran Binding Protein 5 in Nuclear Import and Assembly of the Influenza Virus RNA Polymerase Complex. *J Virol.* 2006;80(24):11911-11919. doi:10.1128/JVI.01565-06
58. Dias A, Bouvier D, Crépin T, *et al.* The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit. *Nature.* 2009;458(7240):914-918. doi:10.1038/nature07745
59. Diekmann Y, Pereira-Leal JB. Evolution of intracellular compartmentalization. *Biochem J.* 2013;449(2):319-331. doi:10.1042/BJ20120957
60. Digard P, Elton D, Bishop K, Medcalf E, Weeds A, Pope B. Modulation of Nuclear Localization of the Influenza Virus Nucleoprotein through Interaction with Actin Filaments. *J Virol.* 1999;73(3):2222-2231.
61. Dingwall C, Robbins J, Dilworth SM, Roberts B, Richardson WD. The nucleoplasmin nuclear location sequence is larger and more complex than that of SV-40 large T antigen. *J Cell Biol.* 1988;107(3):841-849. doi:10.1083/jcb.107.3.841
62. Domínguez C. *Claudia Domínguez Ilustración.*; 2010.
63. Donchet A, Oliva J, Labaronne A, *et al.* The structure of the nucleoprotein of Influenza D shows that all Orthomyxoviridae nucleoproteins have a similar NPCORE, with or without a NPTAIL for nuclear transport. *Sci Rep.* 2019;9(1):600. doi:10.1038/s41598-018-37306-y
64. Donchet A, Vassal-Stermann E, Gérard FCA, Ruigrok RWH, Crépin T. Differential Behaviours and Preferential Bindings of Influenza Nucleoproteins on Importins- $\alpha$ . *Viruses.* 2020;12(8). doi:10.3390/v12080834
65. Dou D, Revol R, Östbye H, Wang H, Daniels R. Influenza A Virus Cell Entry, Replication, Virion Assembly and Movement. *Front Immunol.* 2018;9. doi:10.3389/fimmu.2018.01581
66. Drake JW. Rates of spontaneous mutation among RNA viruses. *Proc Natl Acad Sci U S A.* 1993;90(9):4171-4175. doi:10.1073/pnas.90.9.4171
67. Dubois J, Terrier O, Rosa-Calatrava M. Influenza Viruses and mRNA Splicing: Doing More with Less. *mBio.* 2014;5(3). doi:10.1128/mBio.00070-14
68. Ducatez MF, Pelletier C, Meyer G. Influenza D virus in cattle, France, 2011-2014. *Emerg Infect Dis.* 2015;21(2):368-371. doi:10.3201/eid2102.141449
69. Duverger E, Carpentier V, Roche AC, Monsigny M. Sugar-dependent nuclear import of glycoconjugates from the cytosol. *Exp Cell Res.* 1993;207(1):197-201. doi:10.1006/excr.1993.1181
70. Elton D, Medcalf E, Bishop K, Digard P. Oligomerization of the influenza virus nucleoprotein: identification of positive and negative sequence elements. *Virology.* 1999;260(1):190-200. doi:10.1006/viro.1999.9818
71. 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(9):7357-7367.

72. Elton D, Simpson-Holley M, Archer K, *et al.* Interaction of the Influenza Virus Nucleoprotein with the Cellular CRM1-Mediated Nuclear Export Pathway. *J Virol.* 2001;75(1):408-419. doi:10.1128/JVI.75.1.408-419.2001
73. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr.* 2010;66(Pt 4):486-501. doi:10.1107/S0907444910007493
74. Fahrenkrog B, Aebi U. The nuclear pore complex: nucleocytoplasmic transport and beyond. *Nat Rev Mol Cell Biol.* 2003;4(10):757-766. doi:10.1038/nrm1230
75. Fan H, Walker AP, Carrique L, *et al.* Influenza A virus RNA polymerase structures provide insights into viral genome replication. *Nature.* 2019;573(7773):287-290. doi:10.1038/s41586-019-1530-7
76. Fanara P, Hodel MR, Corbett AH, Hodel AE. Quantitative analysis of nuclear localization signal (NLS)-importin alpha interaction through fluorescence depolarization. Evidence for auto-inhibitory regulation of NLS binding. *J Biol Chem.* 2000;275(28):21218-21223. doi:10.1074/jbc.M002217200
77. Faustino RS, Cheung P, Richard MN, *et al.* Ceramide regulation of nuclear protein import. *J Lipid Res.* 2008;49(3):654-662. doi:10.1194/jlr.M700464-JLR200
78. Feldherr CM, Kallenbach E, Schultz N. Movement of a karyophilic protein through the nuclear pores of oocytes. *J Cell Biol.* 1984;99(6):2216-2222. doi:10.1083/jcb.99.6.2216
79. Ferguson L, Luo K, Olivier AK, *et al.* Influenza D Virus Infection in Feral Swine Populations, United States. *Emerg Infect Dis.* 2018;24(6):1020-1028. doi:10.3201/eid2406.172102
80. Ferguson L, Olivier AK, Genova S, *et al.* Pathogenesis of Influenza D Virus in Cattle. *J Virol.* 2016;90(12):5636-5642. doi:10.1128/JVI.03122-15
81. Field MC, Rout MP. Pore timing: the evolutionary origins of the nucleus and nuclear pore complex. *F1000Research.* 2019;8. doi:10.12688/f1000research.16402.1
82. Flatt JW, Greber UF. Misdelivery at the Nuclear Pore Complex—Stopping a Virus Dead in Its Tracks. *Cells.* 2015;4(3):277-296. doi:10.3390/cells4030277
83. Fodor E. The RNA polymerase of influenza a virus: mechanisms of viral transcription and replication. *Acta Virol.* 2013;57(2):113-122. doi:10.4149/av\_2013\_02\_113
84. Fodor E, Pritlove DC, Brownlee GG. The influenza virus panhandle is involved in the initiation of transcription. *J Virol.* 1994;68(6):4092-4096.
85. 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.* 2004;78(17):9144-9153. doi:10.1128/JVI.78.17.9144-9153.2004
86. Fontes MRM, Teh T, Jans D, Brinkworth RI, Kobe B. Structural basis for the specificity of bipartite nuclear localization sequence binding by importin-alpha. *J Biol Chem.* 2003;278(30):27981-27987. doi:10.1074/jbc.M303275200
87. Forwood JK, Jans DA. Nuclear import pathway of the telomere elongation suppressor TRF1: inhibition by importin alpha. *Biochemistry.* 2002;41(30):9333-9340. doi:10.1021/bi025548s
88. Fournier E, Moules V, Essere B, *et al.* A supramolecular assembly formed by influenza A virus genomic RNA segments. *Nucleic Acids Res.* 2012;40(5):2197-2209. doi:10.1093/nar/gkr985
89. Fournier E, Moules V, Essere B, *et al.* Interaction network linking the human H3N2 influenza A virus genomic RNA segments. *Vaccine.* 2012;30(51):7359-7367. doi:10.1016/j.vaccine.2012.09.079



90. Fournier G, Chiang C, Munier S, *et al.* Recruitment of RED-SMU1 Complex by Influenza A Virus RNA Polymerase to Control Viral mRNA Splicing. *PLoS Pathog.* 2014;10(6). doi:10.1371/journal.ppat.1004164
91. Friedrich B, Quensel C, Sommer T, Hartmann E, Köhler M. Nuclear localization signal and protein context both mediate importin alpha specificity of nuclear import substrates. *Mol Cell Biol.* 2006;26(23):8697-8709. doi:10.1128/MCB.00708-06
92. Gabaldón T, Pittis AA. Origin and evolution of metabolic sub-cellular compartmentalization in eukaryotes. *Biochimie.* 2015;119:262-268. doi:10.1016/j.biochi.2015.03.021
93. Gabriel G, Herwig A, Klenk H-D. Interaction of Polymerase Subunit PB2 and NP with Importin  $\alpha$ 1 Is a Determinant of Host Range of Influenza A Virus. *PLoS Pathog.* 2008;4(2). doi:10.1371/journal.ppat.0040011
94. Gabriel G, Klingel K, Otte A, *et al.* Differential use of importin- $\alpha$  isoforms governs cell tropism and host adaptation of influenza virus. *Nat Commun.* 2011;2:156. doi:10.1038/ncomms1158
95. Galy V, Gadad O, Fromont-Racine M, Romano A, Jacquier A, Nehrbass U. Nuclear Retention of Unspliced mRNAs in Yeast Is Mediated by Perinuclear Mlp1. *Cell.* 2004;116(1):63-73. doi:10.1016/S0092-8674(03)01026-2
96. Gamblin SJ, Skehel JJ. Influenza Hemagglutinin and Neuraminidase Membrane Glycoproteins. *J Biol Chem.* 2010;285(37):28403-28409. doi:10.1074/jbc.R110.129809
97. Gerber M, Isel C, Moules V, Marquet R. Selective packaging of the influenza A genome and consequences for genetic reassortment. *Trends Microbiol.* 2014;22(8):446-455. doi:10.1016/j.tim.2014.04.001
98. Giese S, Bolte H, Schwemmle M. The Feat of Packaging Eight Unique Genome Segments. *Viruses.* 2016;8(6). doi:10.3390/v8060165
99. Gilbertson B, Zheng T, Gerber M, *et al.* Influenza NA and PB1 Gene Segments Interact during the Formation of Viral Progeny: Localization of the Binding Region within the PB1 Gene. *Viruses.* 2016;8(8). doi:10.3390/v8080238
100. Gilchrist D, Mykytka B, Rexach M. Accelerating the rate of disassembly of karyopherin.cargo complexes. *J Biol Chem.* 2002;277(20):18161-18172. doi:10.1074/jbc.M112306200
101. Görlich D, Kostka S, Kraft R, *et al.* Two different subunits of importin cooperate to recognize nuclear localization signals and bind them to the nuclear envelope. *Curr Biol CB.* 1995;5(4):383-392. doi:10.1016/s0960-9822(95)00079-0
102. Görlich D, Kutay U. Transport between the cell nucleus and the cytoplasm. *Annu Rev Cell Dev Biol.* 1999;15:607-660. doi:10.1146/annurev.cellbio.15.1.607
103. Görlich D, Panté N, Kutay U, Aebi U, Bischoff FR. Identification of different roles for RanGDP and RanGTP in nuclear protein import. *EMBO J.* 1996;15(20):5584-5594.
104. Görlich D, Vogel F, Mills AD, Hartmann E, Laskey RA. Distinct functions for the two importin subunits in nuclear protein import. *Nature.* 1995;377(6546):246-248. doi:10.1038/377246a0
105. Goto H, Muramoto Y, Noda T, Kawaoka Y. The genome-packaging signal of the influenza A virus genome comprises a genome incorporation signal and a genome-bundling signal. *J Virol.* 2013;87(21):11316-11322. doi:10.1128/JVI.01301-13
106. Götz V, Magar L, Dornfeld D, *et al.* Influenza A viruses escape from MxA restriction at the expense of efficient nuclear vRNP import. *Sci Rep.* 2016;6. doi:10.1038/srep23138
107. Gruss OJ, Carazo-Salas RE, Schatz CA, *et al.* Ran induces spindle assembly by reversing the inhibitory effect of importin alpha on TPX2 activity. *Cell.* 2001;104(1):83-93. doi:10.1016/s0092-8674(01)00193-3

108. 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):500-506. doi:10.1038/nsmb.1421
109. Güttler T, Görlich D. Ran-dependent nuclear export mediators: a structural perspective. *EMBO J.* 2011;30(17):3457-3474. doi:10.1038/emboj.2011.287
110. Hale BG, Randall RE, Ortín J, Jackson D. The multifunctional NS1 protein of influenza A viruses. *J Gen Virol.* 2008;89(Pt 10):2359-2376. doi:10.1099/vir.0.2008/004606-0
111. Han Q, Chang C, Li L, *et al.* Sumoylation of Influenza A Virus Nucleoprotein Is Essential for Intracellular Trafficking and Virus Growth. *J Virol.* 2014;88(16):9379-9390. doi:10.1128/JVI.00509-14
112. Han Y, Wang X. The emerging roles of KPNA2 in cancer. *Life Sci.* 2020;241:117140. doi:10.1016/j.lfs.2019.117140
113. Hanover JA, Love DC, Prinz WA. Calmodulin-driven Nuclear Entry: Trigger for Sex Determination and Terminal Differentiation. *J Biol Chem.* 2009;284(19):12593-12597. doi:10.1074/jbc.R800076200
114. Harreman MT, Hodel MR, Fanara P, Hodel AE, Corbett AH. The auto-inhibitory function of importin alpha is essential *in vivo*. *J Biol Chem.* 2003;278(8):5854-5863. doi:10.1074/jbc.M210951200
115. Harris A, Cardone G, Winkler DC, *et al.* Influenza virus pleiomorphy characterized by cryoelectron tomography. *Proc Natl Acad Sci U S A.* 2006;103(50):19123-19127. doi:10.1073/pnas.0607614103
116. Hause BM, Collin EA, Liu R, *et al.* Characterization of a novel influenza virus in cattle and Swine: proposal for a new genus in the Orthomyxoviridae family. *mBio.* 2014;5(2):e00031-00014. doi:10.1128/mBio.00031-14
117. Hause BM, Ducatez M, Collin EA, *et al.* Isolation of a novel swine influenza virus from Oklahoma in 2011 which is distantly related to human influenza C viruses. *PLoS Pathog.* 2013;9(2):e1003176. doi:10.1371/journal.ppat.1003176
118. Hay AJ, Lomniczi B, Bellamy AR, Skehel JJ. Transcription of the influenza virus genome. *Virology.* 1977;83(2):337-355. doi:10.1016/0042-6822(77)90179-9
119. Hengrung N, El Omari K, Serna Martin I, *et al.* Crystal structure of the RNA-dependent RNA polymerase from influenza C virus. *Nature.* 2015;527(7576):114-117. doi:10.1038/nature15525
120. Herrler G, Dürkop I, Becht H, Klenk HD. The glycoprotein of influenza C virus is the haemagglutinin, esterase and fusion factor. *J Gen Virol.* 1988;69 ( Pt 4):839-846. doi:10.1099/0022-1317-69-4-839
121. Hewat EA, Cusack S, Ruigrok RW, Verwey C. Low resolution structure of the influenza C glycoprotein determined by electron microscopy. *J Mol Biol.* 1984;175(2):175-193. doi:10.1016/0022-2836(84)90473-x
122. Hodel AE, Harreman MT, Pulliam KF, *et al.* Nuclear Localization Signal Receptor Affinity Correlates with *in vivo* Localization in *Saccharomyces cerevisiae*. *J Biol Chem.* 2006;281(33):23545-23556. doi:10.1074/jbc.M601718200
123. Hodel MR, Corbett AH, Hodel AE. Dissection of a nuclear localization signal. *J Biol Chem.* 2001;276(2):1317-1325. doi:10.1074/jbc.M008522200
124. Holwerda M, Kelly J, Laloli L, *et al.* Determining the Replication Kinetics and Cellular Tropism of Influenza D Virus on Primary Well-Differentiated Human Airway Epithelial Cells. *Viruses.* 2019;11(4). doi:10.3390/v11040377
125. Hongo S, Ishii K, Mori K, *et al.* Detection of ion channel activity in *Xenopus laevis* oocytes expressing Influenza C virus CM2 protein. *Arch Virol.* 2004;149(1):35-50. doi:10.1007/s00705-003-0209-3

126. Huang S, Chen J, Chen Q, *et al.* A Second CRM1-Dependent Nuclear Export Signal in the Influenza A Virus NS2 Protein Contributes to the Nuclear Export of Viral Ribonucleoproteins. *J Virol.* 2013;87(2):767-778. doi:10.1128/JVI.06519-11
127. Hudjetz B, Gabriel G. Human-like PB2 627K Influenza Virus Polymerase Activity Is Regulated by Importin- $\alpha$ 1 and - $\alpha$ 7. *PLoS Pathog.* 2012;8(1). doi:10.1371/journal.ppat.1002488
128. Huet S, Avilov SV, Ferbitz L, Daigle N, Cusack S, Ellenberg J. Nuclear Import and Assembly of Influenza A Virus RNA Polymerase Studied in Live Cells by Fluorescence Cross-Correlation Spectroscopy. *J Virol.* 2010;84(3):1254-1264. doi:10.1128/JVI.01533-09
129. Hülsmann BB, Labokha AA, Görlich D. The permeability of reconstituted nuclear pores provides direct evidence for the selective phase model. *Cell.* 2012;150(4):738-751. doi:10.1016/j.cell.2012.07.019
130. Hutchinson EC, Denham EM, Thomas B, *et al.* Mapping the Phosphoproteome of Influenza A and B Viruses by Mass Spectrometry. *PLOS Pathog.* 2012;8(11):e1002993. doi:10.1371/journal.ppat.1002993
131. Hutchinson EC, Fodor E. Nuclear import of the influenza A virus transcriptional machinery. *Vaccine.* 2012;30(51):7353-7358. doi:10.1016/j.vaccine.2012.04.085
132. Hutchinson EC, Fodor E. Transport of the Influenza Virus Genome from Nucleus to Nucleus. *Viruses.* 2013;5(10):2424-2446. doi:10.3390/v5102424
133. Hutchinson EC, Orr OE, Man Liu S, Engelhardt OG, 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.* 2011;92(8):1859-1869. doi:10.1099/vir.0.032813-0
134. ICTV, release 2019. International Committee on Taxonomy of Viruses (ICTV). Published March 2020. Accessed March 18, 2020. [https://talk.ictvonline.org/taxonomy/p/taxonomy\\_releases](https://talk.ictvonline.org/taxonomy/p/taxonomy_releases)
135. Institute of Medicine (US) Forum on Microbial Threats, Knobler SL, Mack A, Mahmoud A, Lemon SM. *The Story of Influenza*. National Academies Press (US); 2005. Accessed November 13, 2020. <https://www.ncbi.nlm.nih.gov/books/NBK22148/>
136. Ito T, Suzuki Y, Mitnaul L, Vines A, Kida H, Kawaoka Y. Receptor specificity of influenza A viruses correlates with the agglutination of erythrocytes from different animal species. *Virology.* 1997;227(2):493-499. doi:10.1006/viro.1996.8323
137. Ito T, Suzuki Y, Suzuki T, *et al.* Recognition of N-glycolylneuraminic acid linked to galactose by the  $\alpha$ 2,3 linkage is associated with intestinal replication of influenza A virus in ducks. *J Virol.* 2000;74(19):9300-9305. doi:10.1128/jvi.74.19.9300-9305.2000
138. Iuliano AD, Roguski KM, Chang HH, *et al.* Estimates of global seasonal influenza-associated respiratory mortality: a modelling study. *Lancet Lond Engl.* 2018;391(10127):1285-1300. doi:10.1016/S0140-6736(17)33293-2
139. Izaurralde E, Kutay U, von Kobbe C, Mattaj IW, Görlich D. The asymmetric distribution of the constituents of the Ran system is essential for transport into and out of the nucleus. *EMBO J.* 1997;16(21):6535-6547. doi:10.1093/emboj/16.21.6535
140. Jäkel S, Mingot J-M, Schwarzmaier P, Hartmann E, Görlich D. Importins fulfil a dual function as nuclear import receptors and cytoplasmic chaperones for exposed basic domains. *EMBO J.* 2002;21(3):377-386. doi:10.1093/emboj/21.3.377
141. Jang J, Bae S-E. Comparative Co-Evolution Analysis Between the HA and NA Genes of Influenza A Virus. *Virol Res Treat.* 2018;9. doi:10.1177/1178122X18788328
142. Jékely G. Origin of the nucleus and Ran-dependent transport to safeguard ribosome biogenesis in a chimeric cell. *Biol Direct.* 2008;3:31. doi:10.1186/1745-6150-3-31

143. Jennings PA, Finch JT, Winter G, Robertson JS. Does the higher order structure of the influenza virus ribonucleoprotein guide sequence rearrangements in influenza viral RNA? *Cell*. 1983;34(2):619-627. doi:10.1016/0092-8674(83)90394-x
144. Jivraj N, Butler A. The 1918-19 influenza pandemic revisited. *J R Coll Physicians Edinb*. 2013;43(4):347-352. doi:10.4997/JRCPE.2013.405
145. Johnson NPAS, Mueller J. Updating the accounts: global mortality of the 1918-1920 "Spanish" influenza pandemic. *Bull Hist Med*. 2002;76(1):105-115. doi:10.1353/bhm.2002.0022
146. Jones DT. Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol*. 1999;292(2):195-202. doi:10.1006/jmbi.1999.3091
147. de Jong MD, Simmons CP, Thanh TT, *et al*. Fatal outcome of human influenza A (H5N1) is associated with high viral load and hypercytokinemia. *Nat Med*. 2006;12(10):1203-1207. doi:10.1038/nm1477
148. Jorba N, Coloma R, Ortín J. Genetic trans-Complementation Establishes a New Model for Influenza Virus RNA Transcription and Replication. *PLOS Pathog*. 2009;5(5):e1000462. doi:10.1371/journal.ppat.1000462
149. Junge C. Morbidity: A personal response. *Nature*. 2011;480(7376):S14-15. doi:10.1038/480S14a
150. Kabsch W. Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr D Biol Crystallogr*. 2010;66(Pt 2):133-144. doi:10.1107/S0907444909047374
151. Käfer S, Paraskevopoulou S, Zirkel F, *et al*. Re-assessing the diversity of negative strand RNA viruses in insects. *PLoS Pathog*. 2019;15(12):e1008224. doi:10.1371/journal.ppat.1008224
152. Kalderon D, Roberts BL, Richardson WD, Smith AE. A short amino acid sequence able to specify nuclear location. *Cell*. 1984;39(3 Pt 2):499-509. doi:10.1016/0092-8674(84)90457-4
153. Kamei Y, Yuba S, Nakayama T, Yoneda Y. Three distinct classes of the alpha-subunit of the nuclear pore-targeting complex (importin-alpha) are differentially expressed in adult mouse tissues. *J Histochem Cytochem Off J Histochem Soc*. 1999;47(3):363-372. doi:10.1177/002215549904700310
154. Kappel C, Zachariae U, Dölker N, Grubmüller H. An Unusual Hydrophobic Core Confers Extreme Flexibility to HEAT Repeat Proteins. *Biophys J*. 2010;99(5):1596-1603. doi:10.1016/j.bpj.2010.06.032
155. Katz G, Benkarroum Y, Wei H, *et al*. Morphology of influenza B/Lee/40 determined by cryo-electron microscopy. *PloS One*. 2014;9(2):e88288. doi:10.1371/journal.pone.0088288
156. Keminer O, Peters R. Permeability of single nuclear pores. *Biophys J*. 1999;77(1):217-228. doi:10.1016/S0006-3495(99)76883-9
157. Kemler I, Whittaker G, Helenius A. Nuclear import of microinjected influenza virus ribonucleoproteins. *Virology*. 1994;202(2):1028-1033. doi:10.1006/viro.1994.1432
158. Kesinger E, Liu J, Jensen A, Chia CP, Demers A, Moriyama H. Influenza D virus M2 protein exhibits ion channel activity in *Xenopus laevis* oocytes. *PloS One*. 2018;13(6):e0199227. doi:10.1371/journal.pone.0199227
159. Kilbourne ED, Murphy JS. 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*. 1960;111:387-406. doi:10.1084/jem.111.3.387
160. Kim SJ, Fernandez-Martinez J, Nudelman I, *et al*. Integrative structure and functional anatomy of a nuclear pore complex. *Nature*. 2018;555(7697):475-482. doi:10.1038/nature26003

161. Kim YH, Han M-E, Oh S-O. The molecular mechanism for nuclear transport and its application. *Anat Cell Biol.* 2017;50(2):77-85. doi:10.5115/acb.2017.50.2.77
162. Klumpp K, Ruigrok RW, Baudin F. Roles of the influenza virus polymerase and nucleoprotein in forming a functional RNP structure. *EMBO J.* 1997;16(6):1248-1257. doi:10.1093/emboj/16.6.1248
163. Knyphausen P, Kuhlmann N, de Boor S, Lammers M. Lysine-acetylation as a fundamental regulator of Ran function: Implications for signaling of proteins of the Ras-superfamily. *Small GTPases.* 2015;6(4):189-195. doi:10.1080/21541248.2015.1103399
164. Kobe B. Autoinhibition by an internal nuclear localization signal revealed by the crystal structure of mammalian importin alpha. *Nat Struct Biol.* 1999;6(4):388-397. doi:10.1038/7625
165. Köhler M, Fiebeler A, Hartwig M, et al. Differential expression of classical nuclear transport factors during cellular proliferation and differentiation. *Cell Physiol Biochem Int J Exp Cell Physiol Biochem Pharmacol.* 2002;12(5-6):335-344. doi:10.1159/000067903
166. Köhler M, Speck C, Christiansen M, et al. Evidence for distinct substrate specificities of importin alpha family members in nuclear protein import. *Mol Cell Biol.* 1999;19(11):7782-7791. doi:10.1128/mcb.19.11.7782
167. Kosugi S, Hasebe M, Matsumura N, et al. Six Classes of Nuclear Localization Signals Specific to Different Binding Grooves of Importin  $\alpha$ . *J Biol Chem.* 2009;284(1):478-485. doi:10.1074/jbc.M807017200
168. Kouba T, Drncová P, Cusack S. Structural snapshots of actively transcribing influenza polymerase. *Nat Struct Mol Biol.* 2019;26(6):460-470. doi:10.1038/s41594-019-0232-z
169. Krammer F, Smith GJD, Fouchier RAM, et al. Influenza. *Nat Rev Dis Primer.* 2018;4(1). doi:10.1038/s41572-018-0002-y
170. Kubitscheck U, Grünwald D, Hoekstra A, et al. Nuclear transport of single molecules: dwell times at the nuclear pore complex. *J Cell Biol.* 2005;168(2):233-243. doi:10.1083/jcb.200411005
171. Kuchipudi SV, Nissly RH. Novel Flu Viruses in Bats and Cattle: "Pushing the Envelope" of Influenza Infection. *Vet Sci.* 2018;5(3). doi:10.3390/vetsci5030071
172. Kuersten S, Ohno M, Mattaj IW. Nucleocytoplasmic transport: Ran, beta and beyond. *Trends Cell Biol.* 2001;11(12):497-503. doi:10.1016/s0962-8924(01)02144-4
173. Kumeta M, Yamaguchi H, Yoshimura SH, Takeyasu K. Karyopherin-independent spontaneous transport of amphiphilic proteins through the nuclear pore. *J Cell Sci.* 2012;125(Pt 21):4979-4984. doi:10.1242/jcs.109520
174. Kuszewski K, Brydak L. The epidemiology and history of influenza. *Biomed Pharmacother Biomedecine Pharmacother.* 2000;54(4):188-195. doi:10.1016/S0753-3322(00)89025-3
175. Kutay U, Bischoff FR, Kostka S, Kraft R, Görlich D. Export of Importin  $\alpha$  from the Nucleus Is Mediated by a Specific Nuclear Transport Factor. *Cell.* 1997;90(6):1061-1071. doi:10.1016/S0092-8674(00)80372-4
176. Kutay U, Güttinger S. Leucine-rich nuclear-export signals: born to be weak. *Trends Cell Biol.* 2005;15(3):121-124. doi:10.1016/j.tcb.2005.01.005
177. Labaronne A, Milles S, Donchet A, et al. Structural analysis of the complex between influenza B nucleoprotein and human importin- $\alpha$ . *Sci Rep.* 2017;7(1):17164. doi:10.1038/s41598-017-17458-z
178. Labaronne A, Swale C, Monod A, Schoehn G, Crépin T, Ruigrok RWH. Binding of RNA by the Nucleoproteins of Influenza Viruses A and B. *Viruses.* 2016;8(9). doi:10.3390/v8090247

179. Lamb RA. The Structure, Function, and Pathobiology of the Influenza A and B Virus Ion Channels. *Cold Spring Harb Perspect Med*. Published online January 27, 2020. doi:10.1101/cshperspect.a038505
180. Lange A, McLane LM, Mills RE, Devine SE, Corbett AH. Expanding the Definition of the Classical Bipartite Nuclear Localization Signal. *Traffic Cph Den*. 2010;11(3):311-323. doi:10.1111/j.1600-0854.2009.01028.x
181. Lange A, Mills RE, Lange CJ, Stewart M, Devine SE, Corbett AH. Classical Nuclear Localization Signals: Definition, Function, and Interaction with Importin  $\alpha$ . *J Biol Chem*. 2007;282(8):5101-5105. doi:10.1074/jbc.R600026200
182. Laporte M, Stevaert A, Raeymaekers V, *et al*. Hemagglutinin Cleavability, Acid Stability, and Temperature Dependence Optimize Influenza B Virus for Replication in Human Airways. *J Virol*. 2019;94(1). doi:10.1128/JVI.01430-19
183. Le Sage V, Nanni AV, Bhagwat AR, *et al*. Non-Uniform and Non-Random Binding of Nucleoprotein to Influenza A and B Viral RNA. *Viruses*. 2018;10(10). doi:10.3390/v10100522
184. Lee N, Le Sage V, Nanni AV, Snyder DJ, Cooper VS, Lakdawala SS. Genome-wide analysis of influenza viral RNA and nucleoprotein association. *Nucleic Acids Res*. 2017;45(15):8968-8977. doi:10.1093/nar/gkx584
185. Lee SJ, Imamoto N, Sakai H, *et al*. The adoption of a twisted structure of importin-beta is essential for the protein-protein interaction required for nuclear transport. *J Mol Biol*. 2000;302(1):251-264. doi:10.1006/jmbi.2000.4055
186. Lee SJ, Matsuura Y, Liu SM, Stewart M. Structural basis for nuclear import complex dissociation by RanGTP. *Nature*. 2005;435(7042):693-696. doi:10.1038/nature03578
187. Li C, Hatta M, Watanabe S, Neumann G, Kawaoka Y. Compatibility among polymerase subunit proteins is a restricting factor in reassortment between equine H7N7 and human H3N2 influenza viruses. *J Virol*. 2008;82(23):11880-11888. doi:10.1128/JVI.01445-08
188. Li C-X, Shi M, Tian J-H, *et al*. Unprecedented genomic diversity of RNA viruses in arthropods reveals the ancestry of negative-sense RNA viruses. *eLife*. 2015;4. doi:10.7554/eLife.05378
189. Li Y, Sun L, Zheng W, *et al*. Phosphorylation and dephosphorylation of threonine 188 in nucleoprotein is crucial for the replication of influenza A virus. *Virology*. 2018;520:30-38. doi:10.1016/j.virol.2018.05.002
190. Lokareddy RK, Hapsari RA, van Rheenen M, *et al*. Distinctive Properties of the Nuclear Localization Signals of Inner Nuclear Membrane Proteins Heh1 and Heh2. *Structure*. 2015;23(7):1305-1316. doi:10.1016/j.str.2015.04.017
191. López-García P, Moreira D. Selective forces for the origin of the eukaryotic nucleus. *BioEssays News Rev Mol Cell Dev Biol*. 2006;28(5):525-533. doi:10.1002/bies.20413
192. Lott K, Bhardwaj A, Mitrousis G, Pante N, Cingolani G. The importin beta binding domain modulates the avidity of importin beta for the nuclear pore complex. *J Biol Chem*. 2010;285(18):13769-13780. doi:10.1074/jbc.M109.095760
193. Lott K, Cingolani G. The importin  $\beta$  binding domain as a master regulator of nucleocytoplasmic transport. *Biochim Biophys Acta BBA - Mol Cell Res*. 2011;1813(9):1578-1592. doi:10.1016/j.bbamcr.2010.10.012
194. Lowe AR, Tang JH, Yassif J, *et al*. Importin- $\beta$  modulates the permeability of the nuclear pore complex in a Ran-dependent manner. *eLife*. 2015;4. doi:10.7554/eLife.04052
195. Lukarska M, Fournier G, Pflug A, *et al*. Structural basis of an essential interaction between influenza polymerase and Pol II CTD. *Nature*. 2017;541(7635):117-121. doi:10.1038/nature20594

196. Luo W, Zhang J, Liang L, *et al.* Phospholipid scramblase 1 interacts with influenza A virus NP, impairing its nuclear import and thereby suppressing virus replication. *PLoS Pathog.* 2018;14(1). doi:10.1371/journal.ppat.1006851
197. Luytjes W, Krystal M, Enami M, Parvin JD, Palese P. Amplification, expression, and packaging of foreign gene by influenza virus. *Cell.* 1989;59(6):1107-1113. doi:10.1016/0092-8674(89)90766-6
198. Ma C, Wang J. Functional studies reveal the similarities and differences between AM2 and BM2 proton channels from influenza viruses. *Biochim Biophys Acta Biomembr.* 2018;1860(2):272-280. doi:10.1016/j.bbamem.2017.10.026
199. Madeira F, Park Y mi, Lee J, *et al.* The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* 2019;47(W1):W636-W641. doi:10.1093/nar/gkz268
200. von Magnus P. The influenza virus: its morphology, immunology, and kinetics of multiplication. *Bull World Health Organ.* 1953;8(5-6):647-660.
201. Mans BJ, Anantharaman V, Aravind L, Koonin EV. Comparative genomics, evolution and origins of the nuclear envelope and nuclear pore complex. *Cell Cycle Georget Tex.* 2004;3(12):1612-1637. doi:10.4161/cc.3.12.1345
202. Mänz B, Dornfeld D, Götz V, *et al.* Pandemic influenza A viruses escape from restriction by human MxA through adaptive mutations in the nucleoprotein. *PLoS Pathog.* 2013;9(3):e1003279. doi:10.1371/journal.ppat.1003279
203. Mänz B, Schwemmle M, Brunotte L. Adaptation of Avian Influenza A Virus Polymerase in Mammals To Overcome the Host Species Barrier. *J Virol.* 2013;87(13):7200-7209. doi:10.1128/JVI.00980-13
204. Manzoor R, Igarashi M, Takada A. Influenza A Virus M2 Protein: Roles from Ingress to Egress. *Int J Mol Sci.* 2017;18(12). doi:10.3390/ijms18122649
205. Martin K, Helenius A. Transport of incoming influenza virus nucleocapsids into the nucleus. *J Virol.* 1991;65(1):232-244.
206. Martin W, Koonin EV. Introns and the origin of nucleus-cytosol compartmentalization. *Nature.* 2006;440(7080):41-45. doi:10.1038/nature04531
207. Martín-Benito J, Area E, Ortega J, *et al.* Three-dimensional reconstruction of a recombinant influenza virus ribonucleoprotein particle. *EMBO Rep.* 2001;2(4):313-317. doi:10.1093/embo-reports/kve063
208. Martini M, Gazzaniga V, Bragazzi NL, Barberis I. The Spanish Influenza Pandemic: a lesson from history 100 years after 1918. *J Prev Med Hyg.* 2019;60(1):E64-E67. doi:10.15167/2421-4248/jpmh2019.60.1.1205
209. Mason DA, Stage DE, Goldfarb DS. Evolution of the metazoan-specific importin alpha gene family. *J Mol Evol.* 2009;68(4):351-365. doi:10.1007/s00239-009-9215-8
210. Matlin KS, Reggio H, Helenius A, Simons K. Infectious entry pathway of influenza virus in a canine kidney cell line. *J Cell Biol.* 1981;91(3 Pt 1):601-613. doi:10.1083/jcb.91.3.601
211. Mattaj IW, Englmeier L. Nucleocytoplasmic transport: the soluble phase. *Annu Rev Biochem.* 1998;67:265-306. doi:10.1146/annurev.biochem.67.1.265
212. Maul GG, Deaven L. Quantitative determination of nuclear pore complexes in cycling cells with differing DNA content. *J Cell Biol.* 1977;73(3):748-760. doi:10.1083/jcb.73.3.748
213. McDevitt J, Rudnick S, First M, Spengler J. Role of Absolute Humidity in the Inactivation of Influenza Viruses on Stainless Steel Surfaces at Elevated Temperatures. *Appl Environ Microbiol.* 2010;76(12):3943-3947. doi:10.1128/AEM.02674-09
214. Mehle A, Doudna JA. Adaptive strategies of the influenza virus polymerase for replication in humans. *Proc Natl Acad Sci U S A.* 2009;106(50):21312-21316. doi:10.1073/pnas.0911915106

215. Melén K, Fagerlund R, Franke J, Köhler M, Kinnunen L, Julkunen I. Importin  $\alpha$  Nuclear Localization Signal Binding Sites for STAT1, STAT2, and Influenza A Virus Nucleoprotein. *J Biol Chem*. 2003;278(30):28193-28200. doi:10.1074/jbc.M303571200
216. Mitra N, Cernicchiaro N, Torres S, Li F, Hause BM. Metagenomic characterization of the virome associated with bovine respiratory disease in feedlot cattle identified novel viruses and suggests an etiologic role for influenza D virus. *J Gen Virol*. 2016;97(8):1771-1784. doi:10.1099/jgv.0.000492
217. Miyamoto Y, Yamada K, Yoneda Y. Importin  $\alpha$ : a key molecule in nuclear transport and non-transport functions. *J Biochem (Tokyo)*. 2016;160(2):69-75. doi:10.1093/jb/mvw036
218. Moeller A, Kirchdoerfer RN, Potter CS, Carragher B, Wilson IA. Organization of the Influenza Virus Replication Machinery. *Science*. 2012;338(6114):1631-1634. doi:10.1126/science.1227270
219. Moore MS, Blobel G. The GTP-binding protein Ran/TC4 is required for protein import into the nucleus. *Nature*. 1993;365(6447):661-663. doi:10.1038/365661a0
220. Moriyama T, Nagai M, Oka M, Ikawa M, Okabe M, Yoneda Y. Targeted disruption of one of the importin  $\alpha$  family members leads to female functional incompetence in delivery. *FEBS J*. 2011;278(9):1561-1572. doi:10.1111/j.1742-4658.2011.08079.x
221. Moseley GW, Roth DM, DeJesus MA, et al. Dynein light chain association sequences can facilitate nuclear protein import. *Mol Biol Cell*. 2007;18(8):3204-3213. doi:10.1091/mbc.e07-01-0030
222. Mostafa A, Abdelwhab EM, Mettenleiter TC, Pleschka S. Zoonotic Potential of Influenza A Viruses: A Comprehensive Overview. *Viruses*. 2018;10(9). doi:10.3390/v10090497
223. Mühlbauer D, Dzieciolowski J, Hardt M, et al. Influenza Virus-Induced Caspase-Dependent Enlargement of Nuclear Pores Promotes Nuclear Export of Viral Ribonucleoprotein Complexes. *J Virol*. 2015;89(11):6009-6021. doi:10.1128/JVI.03531-14
224. Mukaigawa J, Nayak DP. Two signals mediate nuclear localization of influenza virus (A/WSN/33) polymerase basic protein 2. *J Virol*. 1991;65(1):245-253.
225. Munier S, Moisy D, Marc D, Naffakh N. [Interspecies transmission, adaptation to humans and pathogenicity of animal influenza viruses]. *Pathol Biol (Paris)*. 2010;58(2):e59-68. doi:10.1016/j.patbio.2010.01.012
226. Munster VJ, Wallensten A, Baas C, et al. Mallards and Highly Pathogenic Avian Influenza Ancestral Viruses, Northern Europe. *Emerg Infect Dis*. 2005;11(10):1545-1551. doi:10.3201/eid1110.050546
227. Muraki Y, Okuwa T, Himeda T, Hongo S, Ohara Y. Effect of cysteine mutations in the extracellular domain of CM2 on the influenza C virus replication. *PloS One*. 2013;8(4):e60510. doi:10.1371/journal.pone.0060510
228. Muramoto Y, Takada A, Fujii K, et al. Hierarchy among viral RNA (vRNA) segments in their role in vRNA incorporation into influenza A virions. *J Virol*. 2006;80(5):2318-2325. doi:10.1128/JVI.80.5.2318-2325.2006
229. Murti KG, Brown PS, Bean WJ, Webster RG. Composition of the helical internal components of influenza virus as revealed by immunogold labeling/electron microscopy. *Virology*. 1992;186(1):294-299. doi:10.1016/0042-6822(92)90084-3
230. Murti KG, Webster RG, Jones IM. Localization of RNA polymerases on influenza viral ribonucleoproteins by immunogold labeling. *Virology*. 1988;164(2):562-566. doi:10.1016/0042-6822(88)90574-0



231. Naffakh N, Tomoiu A, Rameix-Welti M-A, van der Werf S. Host Restriction of Avian Influenza Viruses at the Level of the Ribonucleoproteins. *Annu Rev Microbiol.* 2008;62(1):403-424. doi:10.1146/annurev.micro.62.081307.162746
232. Nakada R, Hirano H, Matsuura Y. Structure of importin- $\alpha$  bound to a non-classical nuclear localization signal of the influenza A virus nucleoprotein. *Sci Rep.* 2015;5(1):1-9. doi:10.1038/srep15055
233. Nakada S, Graves PN, Palese P. The influenza C virus NS gene: evidence for a spliced mRNA and a second NS gene product (NS2 protein). *Virus Res.* 1986;4(3):263-273. doi:10.1016/0168-1702(86)90005-5
234. Nakatsu S, Murakami S, Shindo K, *et al.* Influenza C and D Viruses Package Eight Organized Ribonucleoprotein Complexes. *J Virol.* 2018;92(6). doi:10.1128/JVI.02084-17
235. Nakatsu S, Sagara H, Sakai-Tagawa Y, Sugaya N, Noda T, Kawaoka Y. Complete and Incomplete Genome Packaging of Influenza A and B Viruses. *mBio.* 2016;7(5). doi:10.1128/mBio.01248-16
236. Nardozzi JD, Lott K, Cingolani G. Phosphorylation meets nuclear import: a review. *Cell Commun Signal.* 2010;8(1):32. doi:10.1186/1478-811X-8-32
237. Nath ST, Nayak DP. Function of two discrete regions is required for nuclear localization of polymerase basic protein 1 of A/WSN/33 influenza virus (H1 N1). *Mol Cell Biol.* 1990;10(8):4139-4145. doi:10.1128/mcb.10.8.4139
238. Nayak DP, Hui EKW. The role of lipid microdomains in virus biology. *Subcell Biochem.* 2004;37:443-491. doi:10.1007/978-1-4757-5806-1\_14
239. Neumann G, Castrucci MR, Kawaoka Y. Nuclear import and export of influenza virus nucleoprotein. *J Virol.* 1997;71(12):9690-9700.
240. Neumann G, Hughes MT, Kawaoka Y. Influenza A virus NS2 protein mediates vRNP nuclear export through NES-independent interaction with hCRM1. *EMBO J.* 2000;19(24):6751-6758. doi:10.1093/emboj/19.24.6751
241. Newmeyer DD, Forbes DJ. Nuclear import can be separated into distinct steps *in vitro*: nuclear pore binding and translocation. *Cell.* 1988;52(5):641-653. doi:10.1016/0092-8674(88)90402-3
242. Ng AK-L, Lam MK-H, Zhang H, *et al.* Structural Basis for RNA Binding and Homo-Oligomer Formation by Influenza B Virus Nucleoprotein. *J Virol.* 2012;86(12):6758-6767. doi:10.1128/JVI.00073-12
243. Ng AK-L, 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):3638-3647. doi:10.1096/fj.08-112110
244. Nieto A, de la Luna S, Bárcena J, *et al.* Nuclear transport of influenza virus polymerase PA protein. *Virus Res.* 1992;24(1):65-75. doi:10.1016/0168-1702(92)90031-4
245. Nieto A, de la Luna S, Bárcena J, Portela A, Ortín J. Complex structure of the nuclear translocation signal of influenza virus polymerase PA subunit. *J Gen Virol.* 1994;75 ( Pt 1):29-36. doi:10.1099/0022-1317-75-1-29
246. Ninpan K, Suptawiwat O, Boonarkart C, *et al.* Expression of importin- $\alpha$  isoforms in human nasal mucosa: implication for adaptation of avian influenza A viruses to human host. *Virol J.* 2016;13:90. doi:10.1186/s12985-016-0546-y
247. Nishimura H, Hara M, Sugawara K, *et al.* Characterization of the cord-like structures emerging from the surface of influenza C virus-infected cells. *Virology.* 1990;179(1):179-188. doi:10.1016/0042-6822(90)90287-2
248. Nobusawa E, Sato K. Comparison of the mutation rates of human influenza A and B viruses. *J Virol.* 2006;80(7):3675-3678. doi:10.1128/JVI.80.7.3675-3678.2006

249. Noda T, Murakami S, Nakatsu S, *et al.* Importance of the 1+7 configuration of ribonucleoprotein complexes for influenza A virus genome packaging. *Nat Commun.* 2018;9. doi:10.1038/s41467-017-02517-w
250. Noda T, Sagara H, Yen A, *et al.* Architecture of ribonucleoprotein complexes in influenza A virus particles. *Nature.* 2006;439(7075):490-492. doi:10.1038/nature04378
251. Nogales A, Aydillo T, Ávila-Pérez G, *et al.* Functional Characterization and Direct Comparison of Influenza A, B, C, and D NS1 Proteins *in vitro* and *in vivo*. *Front Microbiol.* 2019;10. doi:10.3389/fmicb.2019.02862
252. Noti JD, Blachere FM, McMillen CM, *et al.* High Humidity Leads to Loss of Infectious Influenza Virus from Simulated Coughs. *PLoS ONE.* 2013;8(2). doi:10.1371/journal.pone.0057485
253. Novel Swine-Origin Influenza A (H1N1) Virus Investigation Team, Dawood FS, Jain S, *et al.* Emergence of a novel swine-origin influenza A (H1N1) virus in humans. *N Engl J Med.* 2009;360(25):2605-2615. doi:10.1056/NEJMoa0903810
254. O'Callaghan RJ, Loughlin M, Labat DD, Howe C. Properties of influenza C virus grown in cell culture. *J Virol.* 1977;24(3):875-882.
255. Ohwada K, Kitame F, Sugawara K, Nishimura H, Homma M, Nakamura K. Distribution of the Antibody to Influenza C Virus in Dogs and Pigs in Yamagata Prefecture, Japan. *Microbiol Immunol.* 1987;31(12):1173-1180. doi:10.1111/j.1348-0421.1987.tb01351.x
256. Oka M, Yoneda Y. Importin  $\alpha$ : functions as a nuclear transport factor and beyond. *Proc Jpn Acad Ser B Phys Biol Sci.* 2018;94(7):259-274. doi:10.2183/pjab.94.018
257. O'Neill RE, Jaskunas R, Blobel G, Palese P, Moroianu J. Nuclear Import of Influenza Virus RNA Can Be Mediated by Viral Nucleoprotein and Transport Factors Required for Protein Import. *J Biol Chem.* 1995;270(39):22701-22704. doi:10.1074/jbc.270.39.22701
258. Ori A, Banterle N, Iskar M, *et al.* Cell type-specific nuclear pores: a case in point for context-dependent stoichiometry of molecular machines. *Mol Syst Biol.* 2013;9:648. doi:10.1038/msb.2013.4
259. Ortega J, Martín-Benito J, Zürcher T, Valpuesta JM, Carrascosa JL, Ortín J. Ultrastructural and functional analyses of recombinant influenza virus ribonucleoproteins suggest dimerization of nucleoprotein during virus amplification. *J Virol.* 2000;74(1):156-163. doi:10.1128/jvi.74.1.156-163.2000
260. Osterhaus ADME, Rimmelzwaan GF, Martina BEE, Bestebroer TM, Fouchier R a. M. Influenza B Virus in Seals. *Science.* 2000;288(5468):1051-1053. doi:10.1126/science.288.5468.1051
261. Ozawa M, Fujii K, Muramoto Y, *et al.* Contributions of Two Nuclear Localization Signals of Influenza A Virus Nucleoprotein to Viral Replication. *J Virol.* 2007;81(1):30-41. doi:10.1128/JVI.01434-06
262. Ozawa M, Kawaoka Y. Crosstalk between animal and human influenza viruses. *Annu Rev Anim Biosci.* 2013;1:21-42. doi:10.1146/annurev-animal-031412-103733
263. Paine PL, Feldherr CM. Nucleocytoplasmic exchange of macromolecules. *Exp Cell Res.* 1972;74(1):81-98. doi:10.1016/0014-4827(72)90483-1
264. Paine PL, Moore LC, Horowitz SB. Nuclear envelope permeability. *Nature.* 1975;254(5496):109-114. doi:10.1038/254109a0
265. Palese P, Schulman JL. Mapping of the influenza virus genome: identification of the hemagglutinin and the neuraminidase genes. *Proc Natl Acad Sci U S A.* 1976;73(6):2142-2146. doi:10.1073/pnas.73.6.2142
266. Panté N, Kann M. Nuclear pore complex is able to transport macromolecules with diameters of about 39 nm. *Mol Biol Cell.* 2002;13(2):425-434. doi:10.1091/mbc.01-06-0308

267. Paragas J, Talon J, O'Neill RE, Anderson DK, García-Sastre A, Palese P. Influenza B and C Virus NEP (NS2) Proteins Possess Nuclear Export Activities. *J Virol.* 2001;75(16):7375-7383. doi:10.1128/JVI.75.16.7375-7383.2001
268. Parry R, Wille M, Turnbull OMH, Geoghegan JL, Holmes EC. Divergent Influenza-Like Viruses of Amphibians and Fish Support an Ancient Evolutionary Association. *Viruses.* 2020;12(9). doi:10.3390/v12091042
269. Parvin JD, Moscona A, Pan WT, Leider JM, Palese P. Measurement of the mutation rates of animal viruses: influenza A virus and poliovirus type 1. *J Virol.* 1986;59(2):377-383.
270. Patterson S, Oxford JS, Dourmashkin RR. Studies on the mechanism of influenza virus entry into cells. *J Gen Virol.* 1979;43(1):223-229. doi:10.1099/0022-1317-43-1-223
271. Peiris JSM, Yu WC, Leung CW, et al. Re-emergence of fatal human influenza A subtype H5N1 disease. *Lancet Lond Engl.* 2004;363(9409):617-619. doi:10.1016/S0140-6736(04)15595-5
272. Peng Q, Liu Y, Peng R, et al. Structural insight into RNA synthesis by influenza D polymerase. *Nat Microbiol.* 2019;4(10):1750-1759. doi:10.1038/s41564-019-0487-5
273. Pflug A, Guilligay D, Reich S, Cusack S. Structure of influenza A polymerase bound to the viral RNA promoter. *Nature.* 2014;516(7531):355-360. doi:10.1038/nature14008
274. Pflug A, Lukarska M, Resa-Infante P, Reich S, Cusack S. Structural insights into RNA synthesis by the influenza virus transcription-replication machine. *Virus Res.* 2017;234:103-117. doi:10.1016/j.virusres.2017.01.013
275. Plotch SJ, Bouloy M, Ulmanen I, Krug RM. A unique cap(m7GpppXm)-dependent influenza virion endonuclease cleaves capped RNAs to generate the primers that initiate viral RNA transcription. *Cell.* 1981;23(3):847-858. doi:10.1016/0092-8674(81)90449-9
276. Plotkin JB, Dushoff J. Codon bias and frequency-dependent selection on the hemagglutinin epitopes of influenza A virus. *Proc Natl Acad Sci U S A.* 2003;100(12):7152-7157. doi:10.1073/pnas.1132114100
277. Pons MW, Schulze IT, Hirst GK, Hauser R. Isolation and characterization of the ribonucleoprotein of influenza virus. *Virology.* 1969;39(2):250-259. doi:10.1016/0042-6822(69)90045-2
278. Poon LL, Pritlove DC, Fodor E, Brownlee GG. 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.* 1999;73(4):3473-3476.
279. Potter CW. A history of influenza. *J Appl Microbiol.* 2001;91(4):572-579. doi:10.1046/j.1365-2672.2001.01492.x
280. Prokudina-Kantorovich EN, Semenova NP. Intracellular oligomerization of influenza virus nucleoprotein. *Virology.* 1996;223(1):51-56. doi:10.1006/viro.1996.0454
281. Ptak C, Wozniak RW. SUMO and Nucleocytoplasmic Transport. *Adv Exp Med Biol.* 2017;963:111-126. doi:10.1007/978-3-319-50044-7\_7
282. Pumroy RA, Cingolani G. Diversification of importin- $\alpha$  isoforms in cellular trafficking and disease states. *Biochem J.* 2015;466(1):13-28. doi:10.1042/BJ20141186
283. Pumroy RA, Ke S, Hart DJ, Zachariae U, Cingolani G. Molecular determinants for nuclear import of influenza A PB2 by importin  $\alpha$  isoforms 3 and 7. *Struct Lond Engl* 1993. 2015;23(2):374-384. doi:10.1016/j.str.2014.11.015
284. Putri WCWS, Muscatello DJ, Stockwell MS, Newall AT. Economic burden of seasonal influenza in the United States. *Vaccine.* 2018;36(27):3960-3966. doi:10.1016/j.vaccine.2018.05.057

285. Quast M, Sreenivasan C, Sexton G, *et al.* Serological evidence for the presence of influenza D virus in small ruminants. *Vet Microbiol.* 2015;180(3-4):281-285. doi:10.1016/j.vetmic.2015.09.005
286. Quensel C, Friedrich B, Sommer T, Hartmann E, Kohler M. *In vivo* analysis of importin alpha proteins reveals cellular proliferation inhibition and substrate specificity. *Mol Cell Biol.* 2004;24(23):10246-10255. doi:10.1128/MCB.24.23.10246-10255.2004
287. Raices M, D'Angelo MA. Nuclear pore complex composition: a new regulator of tissue-specific and developmental functions. *Nat Rev Mol Cell Biol.* 2012;13(11):687-699. doi:10.1038/nrm3461
288. Reich S, Guilligay D, Cusack S. An *in vitro* fluorescence based study of initiation of RNA synthesis by influenza B polymerase. *Nucleic Acids Res.* 2017;45(6):3353-3368. doi:10.1093/nar/gkx043
289. Reich S, Guilligay D, Pflug A, *et al.* Structural insight into cap-snatching and RNA synthesis by influenza polymerase. *Nature.* 2014;516(7531):361-366. doi:10.1038/nature14009
290. Reichelt R, Holzenburg A, Buhle EL, Jarnik M, Engel A, Aebi U. Correlation between structure and mass distribution of the nuclear pore complex and of distinct pore complex components. *J Cell Biol.* 1990;110(4):883-894. doi:10.1083/jcb.110.4.883
291. Resa-Infante P, Bonet J, Thiele S, *et al.* Alternative interaction sites in the influenza A virus nucleoprotein mediate viral escape from the importin- $\alpha$ 7 mediated nuclear import pathway. *FEBS J.* 2019;286(17):3374-3388. doi:10.1111/febs.14868
292. Resa-Infante P, Gabriel G. The nuclear import machinery is a determinant of influenza virus host adaptation. *BioEssays News Rev Mol Cell Dev Biol.* 2013;35(1):23-27. doi:10.1002/bies.201200138
293. Resa-Infante P, Jorba N, Zamarrero N, Fernández Y, Juárez S, Ortín J. The Host-Dependent Interaction of  $\alpha$ -Importins with Influenza PB2 Polymerase Subunit Is Required for Virus RNA Replication. *PLoS ONE.* 2008;3(12). doi:10.1371/journal.pone.0003904
294. Resa-Infante P, Thieme R, Ernst T, *et al.* Importin- $\alpha$ 7 is required for enhanced influenza A virus replication in the alveolar epithelium and severe lung damage in mice. *J Virol.* 2014;88(14):8166-8179. doi:10.1128/JVI.00270-14
295. Rexach M, Blobel G. Protein import into nuclei: association and dissociation reactions involving transport substrate, transport factors, and nucleoporins. *Cell.* 1995;83(5):683-692. doi:10.1016/0092-8674(95)90181-7
296. Riddick G, Macara IG. A systems analysis of importin- $\alpha$ - $\beta$  mediated nuclear protein import. *J Cell Biol.* 2005;168(7):1027-1038. doi:10.1083/jcb.200409024
297. Riddick G, Macara IG. The adapter importin- $\alpha$  provides flexible control of nuclear import at the expense of efficiency. *Mol Syst Biol.* 2007;3:118. doi:10.1038/msb4100160
298. Ritchey MB, Palese P, Schulman JL. Mapping of the influenza virus genome. III. Identification of genes coding for nucleoprotein, membrane protein, and nonstructural protein. *J Virol.* 1976;20(1):307-313.
299. Robbins J, Dilworth SM, Laskey RA, Dingwall C. Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: identification of a class of bipartite nuclear targeting sequence. *Cell.* 1991;64(3):615-623. doi:10.1016/0092-8674(91)90245-t
300. Roberts PC, Compans RW. Host cell dependence of viral morphology. *Proc Natl Acad Sci U S A.* 1998;95(10):5746-5751. doi:10.1073/pnas.95.10.5746
301. Rogers GN, Herrler G, Paulson JC, Klenk HD. Influenza C virus uses 9-O-acetyl-N-acetylneuraminic acid as a high affinity receptor determinant for attachment to cells. *J Biol Chem.* 1986;261(13):5947-5951.

- 
302. Rossman JS, Jing X, Leser GP, Lamb RA. Influenza virus M2 protein mediates ESCRT-independent membrane scission. *Cell*. 2010;142(6):902-913. doi:10.1016/j.cell.2010.08.029
303. Rossman JS, Lamb RA. Influenza Virus Assembly and Budding. *Virology*. 2011;411(2):229-236. doi:10.1016/j.virol.2010.12.003
304. Roth DM, Moseley GW, Pouton CW, Jans DA. Mechanism of microtubule-facilitated "fast track" nuclear import. *J Biol Chem*. 2011;286(16):14335-14351. doi:10.1074/jbc.M110.210302
305. Rother F, Schmidt T, Popova E, et al. Importin  $\alpha 7$  is essential for zygotic genome activation and early mouse development. *PLoS One*. 2011;6(3):e18310. doi:10.1371/journal.pone.0018310
306. Rout MP, Aitchison JD, Suprpto A, Hjertaas K, Zhao Y, Chait BT. The yeast nuclear pore complex: composition, architecture, and transport mechanism. *J Cell Biol*. 2000;148(4):635-651. doi:10.1083/jcb.148.4.635
307. Rout MP, Field MC. The Evolution of Organellar Coat Complexes and Organization of the Eukaryotic Cell. *Annu Rev Biochem*. 2017;86:637-657. doi:10.1146/annurev-biochem-061516-044643
308. Ruigrok R w., Wrigley N g., Calder L j., et al. Electron microscopy of the low pH structure of influenza virus haemagglutinin. *EMBO J*. 1986;5(1):41-49. doi:10.1002/j.1460-2075.1986.tb04175.x
309. 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):1009-1014. doi:10.1099/0022-1317-76-4-1009
310. Ruigrok RW, Crépin T, Kolakofsky D. Nucleoproteins and nucleocapsids of negative-strand RNA viruses. *Curr Opin Microbiol*. 2011;14(4):504-510. doi:10.1016
311. Ruigrok RWH, Crépin T, Hart DJ, Cusack S. Towards an atomic resolution understanding of the influenza virus replication machinery. *Curr Opin Struct Biol*. 2010;20(1):104-113. doi:10.1016/j.sbi.2009.12.007
312. Russell CA, Jones TC, Barr IG, et al. The global circulation of seasonal influenza A (H3N2) viruses. *Science*. 2008;320(5874):340-346. doi:10.1126/science.1154137
313. Saletti D, Radzimanowski J, Effantin G, et al. The Matrix protein M1 from influenza C virus induces tubular membrane invaginations in an *in vitro* cell membrane model. *Sci Rep*. 2017;7. doi:10.1038/srep40801
314. Sankhala RS, Lokareddy RK, Begum S, Pumroy RA, Gillilan RE, Cingolani G. Three-dimensional context rather than NLS amino acid sequence determines importin  $\alpha$  subtype specificity for RCC1. *Nat Commun*. 2017;8(1):1-15. doi:10.1038/s41467-017-01057-7
315. Saunders-Hastings PR, Krewski D. Reviewing the History of Pandemic Influenza: Understanding Patterns of Emergence and Transmission. *Pathogens*. 2016;5(4). doi:10.3390/pathogens5040066
316. Sauter NK, Bednarski MD, Wurzburg BA, et al. Hemagglutinins from two influenza virus variants bind to sialic acid derivatives with millimolar dissociation constants: a 500-MHz proton nuclear magnetic resonance study. *Biochemistry*. 1989;28(21):8388-8396. doi:10.1021/bi00447a018
317. Scheiffele P, Rietveld A, Wilk T, Simons K. Influenza viruses select ordered lipid domains during budding from the plasma membrane. *J Biol Chem*. 1999;274(4):2038-2044. doi:10.1074/jbc.274.4.2038

318. Scholz S, Damm O, Schneider U, Ultsch B, Wichmann O, Greiner W. Epidemiology and cost of seasonal influenza in Germany - a claims data analysis. *BMC Public Health*. 2019;19(1):1090. doi:10.1186/s12889-019-7458-x
319. Schrauwen EJA, de Graaf M, Herfst S, Rimmelzwaan GF, Osterhaus ADME, Fouchier R a. M. Determinants of virulence of influenza A virus. *Eur J Clin Microbiol Infect Dis Off Publ Eur Soc Clin Microbiol*. 2014;33(4):479-490. doi:10.1007/s10096-013-1984-8
320. Schwartz TU. Modularity within the architecture of the nuclear pore complex. *Curr Opin Struct Biol*. 2005;15(2):221-226. doi:10.1016/j.sbi.2005.03.003
321. Sederdahl BK, Williams JV. Epidemiology and Clinical Characteristics of Influenza C Virus. *Viruses*. 2020;12(1). doi:10.3390/v12010089
322. Sekimoto T, Miyamoto Y, Arai S, Yoneda Y. Importin alpha protein acts as a negative regulator for Snail protein nuclear import. *J Biol Chem*. 2011;286(17):15126-15131. doi:10.1074/jbc.M110.213579
323. Sendor AB, Weerasuriya D, Sapra A. Avian Influenza. In: *StatPearls*. StatPearls Publishing; 2020. Accessed March 30, 2020. <http://www.ncbi.nlm.nih.gov/books/NBK553072/>
324. Seo SH, Hoffmann E, Webster RG. Lethal H5N1 influenza viruses escape host anti-viral cytokine responses. *Nat Med*. 2002;8(9):950-954. doi:10.1038/nm757
325. Shapiro GI, Krug RM. Influenza virus RNA replication *in vitro*: synthesis of viral template RNAs and virion RNAs in the absence of an added primer. *J Virol*. 1988;62(7):2285-2290. doi:10.1128/JVI.62.7.2285-2290.1988
326. Shaulov L, Harel A. Improved Visualization of Vertebrate Nuclear Pore Complexes by Field Emission Scanning Electron Microscopy. *Structure*. 2012;20(3):407-413. doi:10.1016/j.str.2012.01.022
327. Shaw ML, Stone KL, Colangelo CM, Gulcicek EE, Palese P. Cellular proteins in influenza virus particles. *PLoS Pathog*. 2008;4(6):e1000085. doi:10.1371/journal.ppat.1000085
328. Sherry L, Smith M, Davidson S, Jackson D. 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*. 2014;88(21):12326-12338. doi:10.1128/JVI.01542-14
329. Shi M, Lin X-D, Chen X, *et al*. The evolutionary history of vertebrate RNA viruses. *Nature*. 2018;556(7700):197-202. doi:10.1038/s41586-018-0012-7
330. Short KR, Kedzierska K, van de Sandt CE. Back to the Future: Lessons Learned From the 1918 Influenza Pandemic. *Front Cell Infect Microbiol*. 2018;8. doi:10.3389/fcimb.2018.00343
331. Sieczkarski SB, Whittaker GR. Influenza virus can enter and infect cells in the absence of clathrin-mediated endocytosis. *J Virol*. 2002;76(20):10455-10464. doi:10.1128/jvi.76.20.10455-10464.2002
332. Smith AE, Kalderon D, Roberts BL, *et al*. The nuclear location signal. *Proc R Soc Lond B Biol Sci*. 1985;226(1242):43-58. doi:10.1098/rspb.1985.0078
333. Smith GL, Levin JZ, Palese P, Moss B. Synthesis and cellular location of the ten influenza polypeptides individually expressed by recombinant vaccinia viruses. *Virology*. 1987;160(2):336-345. doi:10.1016/0042-6822(87)90004-3
334. Smith W, Westwood MA, Westwood JCN, Belyavin G. Spontaneous mutation of influenza virus A during routine egg passage. *Br J Exp Pathol*. 1951;32(5):422-432.
335. Song H, Qi J, Khedri Z, *et al*. An Open Receptor-Binding Cavity of Hemagglutinin-Esterase-Fusion Glycoprotein from Newly-Identified Influenza D Virus: Basis for Its Broad Cell Tropism. *PLoS Pathog*. 2016;12(1):e1005411. doi:10.1371/journal.ppat.1005411

336. Song L. It is Unlikely That Influenza Viruses Will Cause a Pandemic Again Like What Happened in 1918 and 1919. *Front Public Health*. 2014;2. doi:10.3389/fpubh.2014.00039
337. Soniat M, Chook YM. Nuclear localization signals for four distinct karyopherin- $\beta$  nuclear import systems. *Biochem J*. 2015;468(3):353-362. doi:10.1042/BJ20150368
338. Spickler A. Highly Pathogenic Avian Influenza. Published online 2016. [https://www.oie.int/fileadmin/Home/eng/Animal\\_Health\\_in\\_the\\_World/docs/pdf/Disease\\_cards/HPAI.pdf](https://www.oie.int/fileadmin/Home/eng/Animal_Health_in_the_World/docs/pdf/Disease_cards/HPAI.pdf)
339. Stade K, Ford CS, Guthrie C, Weis K. Exportin 1 (Crm1p) is an essential nuclear export factor. *Cell*. 1997;90(6):1041-1050. doi:10.1016/s0092-8674(00)80370-0
340. Stech J, Xiong X, Scholtissek C, Webster RG. Independence of evolutionary and mutational rates after transmission of avian influenza viruses to swine. *J Virol*. 1999;73(3):1878-1884.
341. Steel J, Lowen AC, Mubareka S, Palese P. Transmission of influenza virus in a mammalian host is increased by PB2 amino acids 627K or 627E/701N. *PLoS Pathog*. 2009;5(1):e1000252. doi:10.1371/journal.ppat.1000252
342. Ström A-C, Weis K. Importin-beta-like nuclear transport receptors. *Genome Biol*. 2001;2(6):reviews3008.1-reviews3008.9.
343. Subbarao EK, London W, Murphy BR. A single amino acid in the PB2 gene of influenza A virus is a determinant of host range. *J Virol*. 1993;67(4):1761-1764.
344. Suzuki Y. Sialobiology of Influenza: Molecular Mechanism of Host Range Variation of Influenza Viruses. *Biol Pharm Bull*. 2005;28(3):399-408. doi:10.1248/bpb.28.399
345. Swale C, Da Costa B, Sedano L, et al. X-ray Structure of the Human Karyopherin RanBP5, an Essential Factor for Influenza Polymerase Nuclear Trafficking. *J Mol Biol*. 2020;432(10):3353-3359. doi:10.1016/j.jmb.2020.03.021
346. Swale C, Monod A, Tengo L, et al. Structural characterization of recombinant IAV polymerase reveals a stable complex between viral PA-PB1 heterodimer and host RanBP5. *Sci Rep*. 2016;6. doi:10.1038/srep24727
347. Swayne DE, Suarez DL. Highly pathogenic avian influenza. *Rev Sci Tech Int Off Epizoot*. 2000;19(2):463-482. doi:10.20506/rst.19.2.1230
348. Takashita E, Muraki Y, Sugawara K, et al. Intrinsic temperature sensitivity of influenza C virus hemagglutinin-esterase-fusion protein. *J Virol*. 2012;86(23):13108-13111. doi:10.1128/JVI.01925-12
349. Takemoto DK, Skehel JJ, Wiley DC. A Surface Plasmon Resonance Assay for the Binding of Influenza Virus Hemagglutinin to Its Sialic Acid Receptor. *Virology*. 1996;217(2):452-458. doi:10.1006/viro.1996.0139
350. Tang Y-S, Lo C-Y, Mok CK-P, Chan PK-S, Shaw P-C. The Extended C-Terminal Region of Influenza C Virus Nucleoprotein Is Important for Nuclear Import and Ribonucleoprotein Activity. *J Virol*. 2019;93(9). doi:10.1128/JVI.02048-18
351. Tarendeau F, Boudet J, Guilligay D, et al. Structure and nuclear import function of the C-terminal domain of influenza virus polymerase PB2 subunit. *Nat Struct Mol Biol*. 2007;14(3):229-233. doi:10.1038/nsmb1212
352. Tarendeau F, Crepin T, Guilligay D, Ruigrok RWH, Cusack S, Hart DJ. 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). doi:10.1371/journal.ppat.1000136
353. Tarus B, Bakowicz O, Chenavas S, et al. Oligomerization paths of the nucleoprotein of influenza A virus. *Biochimie*. 2012;94(3):776-785. doi:10.1016/j.biochi.2011.11.009

354. Tejomurtula J, Lee K-B, Tripurani SK, Smith GW, Yao J. Role of importin alpha8, a new member of the importin alpha family of nuclear transport proteins, in early embryonic development in cattle. *Biol Reprod.* 2009;81(2):333-342. doi:10.1095/biolreprod.109.077396
355. Terry LJ, Shows EB, Wentz SR. Crossing the Nuclear Envelope: Hierarchical Regulation of Nucleocytoplasmic Transport. *Science.* 2007;318(5855):1412-1416. doi:10.1126/science.1142204
356. Tetenbaum-Novatt J, Hough LE, Mironska R, McKenney AS, Rout MP. Nucleocytoplasmic Transport: A Role for Nonspecific Competition in Karyopherin-Nucleoporin Interactions. *Mol Cell Proteomics MCP.* 2012;11(5):31-46. doi:10.1074/mcp.M111.013656
357. The Human Protein Atlas. Accessed September 8, 2020. <https://www.proteinatlas.org/>
358. Thiele S, Stanelle-Bertram S, Beck S, *et al.* Cellular Importin- $\alpha$ 3 Expression Dynamics in the Lung Regulate Antiviral Response Pathways against Influenza A Virus Infection. *Cell Rep.* 2020;31(3):107549. doi:10.1016/j.celrep.2020.107549
359. Thierry E, Guilligay D, Kosinski J, *et al.* Influenza Polymerase Can Adopt an Alternative Configuration Involving a Radical Repacking of PB2 Domains. *Mol Cell.* 2016;61(1):125-137. doi:10.1016/j.molcel.2015.11.016
360. Timney BL, Raveh B, Mironska R, *et al.* Simple rules for passive diffusion through the nuclear pore complex. *J Cell Biol.* 2016;215(1):57-76. doi:10.1083/jcb.201601004
361. Timney BL, Tetenbaum-Novatt J, Agate DS, *et al.* Simple kinetic relationships and nonspecific competition govern nuclear import rates *in vivo*. *J Cell Biol.* 2006;175(4):579-593. doi:10.1083/jcb.200608141
362. Tome-Amat J, Ramos I, Amanor F, Fernández-Sesma A, Ashour J. Influenza A Virus Utilizes Low-Affinity, High-Avidity Interactions with the Nuclear Import Machinery To Ensure Infection and Immune Evasion. *J Virol.* 2019;93(1). doi:10.1128/JVI.01046-18
363. Trombetta CM, Marchi S, Manini I, *et al.* Influenza D Virus: Serological Evidence in the Italian Population from 2005 to 2017. *Viruses.* 2020;12(1):30. doi:10.3390/v12010030
364. Tsuji L, Takumi T, Imamoto N, Yoneda Y. Identification of novel homologues of mouse importin alpha, the alpha subunit of the nuclear pore-targeting complex, and their tissue-specific expression. *FEBS Lett.* 1997;416(1):30-34. doi:10.1016/s0014-5793(97)01092-2
365. Turrell L, Hutchinson EC, Vreede FT, Fodor E. Regulation of Influenza A Virus Nucleoprotein Oligomerization by Phosphorylation. *J Virol.* 2014;89(2):1452-1455. doi:10.1128/JVI.02332-14
366. Twyffels L, Gueydan C, Kruys V. Transportin-1 and Transportin-2: Protein nuclear import and beyond. *FEBS Lett.* 2014;588(10):1857-1868. doi:10.1016/j.febslet.2014.04.023
367. Uhlén M, Fagerberg L, Hallström BM, *et al.* Tissue-based map of the human proteome. *Science.* 2015;347(6220). doi:10.1126/science.1260419
368. Uversky VN. The alphabet of intrinsic disorder. *Intrinsically Disord Proteins.* 2013;1(1):e24684. doi:10.4161/idp.24684
369. Vahlenkamp TW, Harder TC. Influenza virus infections in mammals. *Berl Munch Tierarztl Wochenschr.* 2006;119(3-4):123-131.
370. Vasin AV, Temkina OA, Egorov VV, Klotchenko SA, Plotnikova MA, Kiselev OI. Molecular mechanisms enhancing the proteome of influenza A viruses: an overview of recently discovered proteins. *Virus Res.* 2014;185:53-63. doi:10.1016/j.virusres.2014.03.015
371. te Velthuis AJW, Fodor E. Influenza virus RNA polymerase: insights into the mechanisms of viral RNA synthesis. *Nat Rev Microbiol.* 2016;14(8):479-493. doi:10.1038/nrmicro.2016.87



372. te Velthuis AJW, Turell L, Vreede FT, Fodor E. Uncoupling of Influenza A Virus Transcription and Replication through Mutation of the Unpaired Adenosine in the Viral RNA Promoter. *J Virol.* 2013;87(18):10381-10384. doi:10.1128/JVI.00636-13
373. Vetter IR, Arndt A, Kutay U, Görlich D, Wittinghofer A. Structural view of the Ran-Importin beta interaction at 2.3 Å resolution. *Cell.* 1999;97(5):635-646. doi:10.1016/s0092-8674(00)80774-6
374. Viboud C, Alonso WJ, Simonsen L. Influenza in Tropical Regions. *PLoS Med.* 2006;3(4). doi:10.1371/journal.pmed.0030089
375. Viswanathan K, Chandrasekaran A, Srinivasan A, Raman R, Sasisekharan V, Sasisekharan R. Glycans as receptors for influenza pathogenesis. *Glycoconj J.* 2010;27(6):561-570. doi:10.1007/s10719-010-9303-4
376. de Vries E, Du W, Guo H, de Haan CAM. Influenza A Virus Hemagglutinin-Neuraminidase-Receptor Balance: Preserving Virus Motility. *Trends Microbiol.* 2020;28(1):57-67. doi:10.1016/j.tim.2019.08.010
377. de Vries E, de Vries RP, Wienholts MJ, et al. Influenza A virus entry into cells lacking sialylated N-glycans. *Proc Natl Acad Sci U S A.* 2012;109(19):7457-7462. doi:10.1073/pnas.1200987109
378. Walker A, Fan H, Carrique L, et al. Dimerisation of the influenza virus RNA polymerase during viral genome replication. *Access Microbiol.* 2019;1(1A):292. doi:10.1099/acmi.ac2019.po0145
379. Walker AP, Fodor E. Interplay between Influenza Virus and the Host RNA Polymerase II Transcriptional Machinery. *Trends Microbiol.* 2019;27(5):398-407. doi:10.1016/j.tim.2018.12.013
380. Wandzik JM, Kouba T, Karuppasamy M, et al. A Structure-Based Model for the Complete Transcription Cycle of Influenza Polymerase. *Cell.* 2020;181(4):877-893.e21. doi:10.1016/j.cell.2020.03.061
381. Wang P, Palese P, O'Neill RE. The NPI-1/NPI-3 (karyopherin alpha) binding site on the influenza A virus nucleoprotein NP is a nonconventional nuclear localization signal. *J Virol.* 1997;71(3):1850-1856.
382. Wang YE, Pernet O, Lee B. Regulation of the nucleocytoplasmic trafficking of viral and cellular proteins by ubiquitin and small ubiquitin-related modifiers. *Biol Cell Auspices Eur Cell Biol Organ.* 2012;104(3):121-138. doi:10.1111/boc.201100105
383. Wanitchang A, Narkpuk J, Jongkaewwattana A. Nuclear import of influenza B virus nucleoprotein: involvement of an N-terminal nuclear localization signal and a cleavage-protection motif. *Virology.* 2013;443(1):59-68. doi:10.1016/j.virol.2013.04.025
384. Ward MJ, Lycett SJ, Avila D, Bollback JP, Leigh Brown AJ. Evolutionary interactions between haemagglutinin and neuraminidase in avian influenza. *BMC Evol Biol.* 2013;13:222. doi:10.1186/1471-2148-13-222
385. Watanabe K, Shimizu T, Noda S, Tsukahara F, Maru Y, Kobayashi N. Nuclear export of the influenza virus ribonucleoprotein complex: Interaction of Hsc70 with viral proteins M1 and NS2. *FEBS Open Bio.* 2014;4:683-688. doi:10.1016/j.fob.2014.07.004
386. Weber F, Kochs G, Gruber S, Haller O. A Classical Bipartite Nuclear Localization Signal on Thogoto and Influenza A Virus Nucleoproteins. *Virology.* 1998;250(1):9-18. doi:10.1006/viro.1998.9329
387. Webster RG, Bean WJ, Gorman OT, Chambers TM, Kawaoka Y. Evolution and ecology of influenza A viruses. *Microbiol Rev.* 1992;56(1):152-179.
388. Weissenhorn W, Poudévigne E, Effantin G, Bassereau P. How to get out: ssRNA enveloped viruses and membrane fission. *Curr Opin Virol.* 2013;3(2):159-167. doi:10.1016/j.coviro.2013.03.011

389. Welch K, Franke J, Köhler M, Macara IG. RanBP3 contains an unusual nuclear localization signal that is imported preferentially by importin- $\alpha$ 3. *Mol Cell Biol*. 1999;19(12):8400-8411. doi:10.1128/mcb.19.12.8400
390. Westgeest KB, Russell CA, Lin X, *et al*. Genomewide Analysis of Reassortment and Evolution of Human Influenza A(H3N2) Viruses Circulating between 1968 and 2011. *J Virol*. 2014;88(5):2844-2857. doi:10.1128/JVI.02163-13
391. White SK, Ma W, McDaniel CJ, Gray GC, Lednicky JA. Serologic evidence of exposure to influenza D virus among persons with occupational contact with cattle. *J Clin Virol Off Publ Pan Am Soc Clin Virol*. 2016;81:31-33. doi:10.1016/j.jcv.2016.05.017
392. WHO. Avian influenza : assessing the pandemic threat. Published online 2005. [https://apps.who.int/iris/bitstream/handle/10665/68985/WHO\\_CDS\\_2005.29.pdf;jsessionid=7753E3CCFA49748D5A97FAB681392DC1?sequence=1](https://apps.who.int/iris/bitstream/handle/10665/68985/WHO_CDS_2005.29.pdf;jsessionid=7753E3CCFA49748D5A97FAB681392DC1?sequence=1)
393. WHO memorandum. A revision of the system of nomenclature for influenza viruses: a WHO memorandum. *Bull World Health Organ*. 1980;58(4):585-591.
394. WHO/GIP. Cumulative number of confirmed human cases for avian influenza A(H5N1) reported to WHO, 2003-2020. Published online 2020.
395. Wilbur JD, Heald R. Mitotic spindle scaling during *Xenopus* development by kif2a and importin  $\alpha$ . *eLife*. 2013;2:e00290. doi:10.7554/eLife.00290
396. Williams GD, Townsend D, Wylie KM, *et al*. Nucleotide resolution mapping of influenza A virus nucleoprotein-RNA interactions reveals RNA features required for replication. *Nat Commun*. 2018;9. doi:10.1038/s41467-018-02886-w
397. Wilson KL, Dawson SC. Functional evolution of nuclear structure. *J Cell Biol*. 2011;195(2):171-181. doi:10.1083/jcb.201103171
398. Winn MD, Ballard CC, Cowtan KD, *et al*. Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr*. 2011;67(Pt 4):235-242. doi:10.1107/S0907444910045749
399. Wolff T, Unterstab G, Heins G, Richt JA, Kann M. Characterization of an unusual importin alpha binding motif in the borna disease virus p10 protein that directs nuclear import. *J Biol Chem*. 2002;277(14):12151-12157. doi:10.1074/jbc.M109103200
400. World Health Organization. World Health Organization. Accessed March 30, 2020. <https://www.who.int>
401. World Organisation for Animal Health. Accessed June 5, 2020. <https://www.oie.int/en/animal-health-in-the-world/update-on-avian-influenza/2020/>
402. Wrigley NG. Electron microscopy of influenza virus. *Br Med Bull*. 1979;35(1):35-38. doi:10.1093/oxfordjournals.bmb.a071539
403. Wu W, Sankhala RS, Florio TJ, *et al*. Synergy of two low-affinity NLSs determines the high avidity of influenza A virus nucleoprotein NP for human importin  $\alpha$  isoforms. *Sci Rep*. 2017;7(1):11381. doi:10.1038/s41598-017-11018-1
404. Wu WW, Sun Y-HB, Panté N. Nuclear import of influenza A viral ribonucleoprotein complexes is mediated by two nuclear localization sequences on viral nucleoprotein. *Virology*. 2007;4:49. doi:10.1186/1743-422X-4-49
405. Wu WWH, Weaver LL, Panté N. Ultrastructural analysis of the nuclear localization sequences on influenza A ribonucleoprotein complexes. *J Mol Biol*. 2007;374(4):910-916. doi:10.1016/j.jmb.2007.10.022
406. Xu D, Farmer A, Chook YM. Recognition of nuclear targeting signals by Karyopherin- $\beta$  proteins. *Curr Opin Struct Biol*. 2010;20(6):782-790. doi:10.1016/j.sbi.2010.09.008

407. Yamanaka K, Ishihama A, Nagata K. Reconstitution of influenza virus RNA-nucleoprotein complexes structurally resembling native viral ribonucleoprotein cores. *J Biol Chem*. 1990;265(19):11151-11155.
408. Yamashita M, Krystal M, Palese P. Evidence that the matrix protein of influenza C virus is coded for by a spliced mRNA. *J Virol*. 1988;62(9):3348-3355.
409. Yamayoshi S, Watanabe M, Goto H, Kawaoka Y. Identification of a Novel Viral Protein Expressed From the PB2 Segment of Influenza A Virus. *Journal of virology*. doi:10.1128/JVI.02175-15
410. Yang Q, Rout MP, Akey CW. Three-dimensional architecture of the isolated yeast nuclear pore complex: functional and evolutionary implications. *Mol Cell*. 1998;1(2):223-234. doi:10.1016/s1097-2765(00)80023-4
411. Yang W, Gelles J, Musser SM. Imaging of single-molecule translocation through nuclear pore complexes. *Proc Natl Acad Sci U S A*. 2004;101(35):12887-12892. doi:10.1073/pnas.0403675101
412. Yasuda Y, Miyamoto Y, Yamashiro T, *et al*. Nuclear retention of importin  $\alpha$  coordinates cell fate through changes in gene expression. *EMBO J*. 2012;31(1):83-94. doi:10.1038/emboj.2011.360
413. Yasuhara N, Shibasaki N, Tanaka S, *et al*. Triggering neural differentiation of ES cells by subtype switching of importin- $\alpha$ . *Nat Cell Biol*. 2007;9(1):72-79. doi:10.1038/ncb1521
414. Ye Q, Krug RM, Tao YJ. The mechanism by which influenza A virus nucleoprotein forms oligomers and binds RNA. *Nature*. 2006;444(7122):1078-1082. doi:10.1038/nature05379
415. Yoon S-W, Webby RJ, Webster RG. Evolution and ecology of influenza A viruses. *Curr Top Microbiol Immunol*. 2014;385:359-375. doi:10.1007/82\_2014\_396
416. Yu J, Hika B, Liu R, *et al*. The Hemagglutinin-Esterase Fusion Glycoprotein Is a Primary Determinant of the Exceptional Thermal and Acid Stability of Influenza D Virus. *mSphere*. 2017;2(4). doi:10.1128/mSphere.00254-17
417. Yuanji (Kuo Yuanchi) G, Fenggen J, Ping W, Min W, Jiming (Chu Chinming) Z. Isolation of Influenza C Virus from Pigs and Experimental Infection of Pigs with Influenza C Virus. *J Gen Virol*. 1983;64(1):177-182. doi:10.1099/0022-1317-64-1-177
418. Zhang J, Huang F, Tan L, *et al*. Host Protein Moloney Leukemia Virus 10 (MOV10) Acts as a Restriction Factor of Influenza A Virus by Inhibiting the Nuclear Import of the Viral Nucleoprotein. *J Virol*. 2016;90(8):3966-3980. doi:10.1128/JVI.03137-15
419. Zhang Y, Aebermann BD, Anderson TK, *et al*. Influenza Research Database: An integrated bioinformatics resource for influenza virus research. *Nucleic Acids Res*. 2017;45(D1):D466-D474. doi:10.1093/nar/gkw857
420. Zhang Y, Zhang H, Zheng Q. A unique activation-promotion mechanism of the influenza B M2 proton channel uncovered by multiscale simulations. *Phys Chem Chem Phys PCCP*. 2019;21(6):2984-2991. doi:10.1039/c9cp00130a
421. Zheng W, Li J, Wang S, *et al*. Phosphorylation Controls the Nuclear-Cytoplasmic Shuttling of Influenza A Virus Nucleoprotein. *J Virol*. 2015;89(11):5822-5834. doi:10.1128/JVI.00015-15
422. Zheng W, Olson J, Vakharia V, Tao YJ. The crystal structure and RNA-binding of an orthomyxovirus nucleoprotein. *PLoS Pathog*. 2013;9(9):e1003624. doi:10.1371/journal.ppat.1003624
423. Zheng W, Tao YJ. Structure and assembly of the influenza A virus ribonucleoprotein complex. *FEBS Lett*. 2013;587(8):1206-1214. doi:10.1016/j.febslet.2013.02.048

424. Zhirnov OP, Grigoriev VB. Disassembly of influenza C viruses, distinct from that of influenza A and B viruses requires neutral-alkaline pH. *Virology*. 1994;200(1):284-291. doi:10.1006/viro.1994.1188

## Abstract

In 2014, American and Chinese studies have identified the influenza D virus (IDV). This new member of the Orthomyxoviridae family shares only 50% sequence identity with the already well characterized influenza C virus and less than 30% sequence identity with influenza A and B viruses. IDV has been detected in pigs and cattle, and has been shown to replicate in the ferret, the main animal model for influenza virus studies in humans, suggesting a possible transmission. In 2018, a new study identified the first influenza viruses infecting cold-blooded animals. A phylogenetic comparison highlighted that these viruses, and more particularly a toad-infecting virus (ToadV), were closer to influenza B viruses than influenza A viruses. This relationship is surprising, taking into account the influenza B restricted host spectrum in comparison to the influenza A one.

After the work on influenza A and B viruses and in line with it, the team collaborates now with Mariette Ducatez (INRA - ENV Toulouse) to detail the IDV replication, but also the replication machinery of ToadV. The replication machinery comprises various proteins, including the RNA polymerase (RdRp) and the nucleoprotein (NP). The RdRp is composed of three subunits (PA, PB1 and PB2), binds both the highly conserved 3'- and 5'-ends of the vRNA segment which is covered by multiple copies of NP. This macromolecular complex replicates the viral genome in the nucleus of infected cells and makes many contacts with cellular partners for its assembly. We have gained evidences that IDV NP uses the importins- $\alpha$  system, similar to influenza A and B nucleoproteins. By using advanced techniques of modern biology and physics (X-ray, electron microscopy, surface plasmon resonance, fluorescence anisotropy, ...), the aims of this project are 1/ to provide structural and functional data on IDV and ToadV nucleoproteins and also on their interactions with host partners, and 2/ to highlight both the common and specific features of all types of influenza NPs, giving precious insights on the evolution processes occurring within this growing family of viruses.

## Résumé

En 2014, des études chinoises et américaines ont identifié le virus de la grippe de type D (IDV). Ce nouveau membre de la famille des Orthomyxoviridae partage seulement 50% d'identité de séquence avec le déjà bien caractérisé virus de type C et partage moins de 30% d'identité de séquence avec les virus de grippaux de type A et B. IDV a été détecté chez les cochons et les bovins et il a été montré qu'il était capable de se répliquer chez le furet, le principal modèle animal pour l'étude du virus de la grippe chez l'homme, suggérant une possible transmission. En 2018, une nouvelle étude a identifié les premiers virus grippaux infectant des animaux à sang froid. Une comparaison phylogénétique a mis en lumière que ces virus, notamment un virus isolé chez un crapaud (ToadV), sont plus proches des virus de type B que des virus de type A. Cette relation évolutive est surprenante, considérant le tropisme restreint du virus de type B comparé à celui des virus de type A. En droite ligne de ses travaux sur les virus grippaux de type A et B, l'équipe collabore avec Mariette Ducatez (INRA - ENV Toulouse) pour détailler la réplication d'IDV, mais également la machinerie répliquative de ToadV. La machinerie de réplication comprend de nombreuses protéines, incluant l'ARN polymérase (RdRp) et la nucléoprotéine (NP). La RdRp, composée de 3 sous-unités (PA, PB1 et PB2), lie les extrémités hautement conservées 3'- et 5'- du segment d'ARN viral qui est encapsidé par de multiples copies de NP. Ce complexe macromoléculaire réplique le génome viral dans le noyau des cellules infectées et interagit avec de nombreux partenaires cellulaires pour son assemblage. Nous avons obtenu des preuves que la NP d'IDV utilise le système d'import nucléaire basé sur les importines  $\alpha$ , de façon similaire aux NP des virus de type A et B. En utilisant des techniques avancées de biologie et de physique (rayons X, microscopie électronique, résonance des plasmons de surface, anisotropie de fluorescence,...), les buts de ce projet sont 1/ d'obtenir des données structurales et fonctionnelles sur les nucléoprotéines d'IDV et de ToadV ainsi que sur leurs interactions avec les partenaires cellulaires, et 2/ de mettre en lumière aussi bien les traits communs que les spécificités de tous les types de NP grippales, ce qui permettra également de comprendre les processus évolutifs en action au sein de cette grandissante famille de virus.