



**HAL**  
open science

# Etudes structurales et fonctionnelles de la nucléoprotéine et de la polymérase du virus de la grippe en association avec leur transporteur nucléaire humain

Sébastien Boulo

► **To cite this version:**

Sébastien Boulo. Etudes structurales et fonctionnelles de la nucléoprotéine et de la polymérase du virus de la grippe en association avec leur transporteur nucléaire humain. Biochimie [q-bio.BM]. Université Claude Bernard - Lyon I, 2008. Français. NNT: . tel-00356602

**HAL Id: tel-00356602**

**<https://theses.hal.science/tel-00356602>**

Submitted on 27 Jan 2009

**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.

# THESE

Pour l'obtention du diplôme de Docteur de l'Université Claude Bernard - Lyon I

Discipline : Biochimie - Biologie Structurale

Présentée et soutenue publiquement par

**Sébastien BOULO**

le 11 décembre 2008

## **Etudes structurales et fonctionnelles de la nucléoprotéine et de la polymérase du virus de la grippe en association avec leur transporteur nucléaire humain**

JURY :

Dr Thierry GAUDE

Dr Nadia NAFFAKH

Dr Bernard DELMAS

Pr Bruno LINA

Dr Vincent LOTTEAU

Dr Florence BAUDIN

Président

Rapporteur

Rapporteur

Examineur

Directeur de thèse

Directeur de thèse

Travaux effectués à l'**UVHCI** - Grenoble (Unit of Virus Host Cell Interactions) UMR-5233  
et à l'**EMBL** - Heidelberg (European Molecular Biology Laboratory)

# **Remerciements**

*Mes premiers remerciements s'adressent à mes directeurs de thèse, le Dr **Florence Baudin** et le Dr **Vincent Lotteau**. A **Florence**, pour avoir su me guider dans mon travail de recherche, pour sa disponibilité, ses connaissances, sa pertinence et pour sa patience. A **Vincent**, pour m'avoir donné l'opportunité de réaliser cette thèse.*

*J'exprime tous mes remerciements au Pr **Rob Ruigrok**, directeur de l'UVHCI, pour m'avoir accueilli dans son Institut au sein du groupe de recherche sur les virus à ARN négatif.*

*Je suis reconnaissant envers le Dr **Nadia Naffakh** et le Dr **Bernard Delmas** qui ont accepté d'être les rapporteurs de ce travail. De même, je tiens à remercier le Pr **Bruno Lina** et le Dr **Thierry Gaude** pour avoir accepté de participer à mon jury de thèse.*

*Je remercie le Dr **Iain Mattaj**, directeur de l'EMBL-Heidelberg pour m'avoir accueilli dans son Institut ainsi que le Dr **Jan Ellenberg**, pour m'avoir reçu dans l'unité "gene expression".*

*Merci au Dr **Vladimir Rybin**, pour son expertise et ses conseils en biophysique, ainsi qu'à **Claire Batisse** pour son aide en microscopie électronique.*

*Je tiens également à remercier la meute de chacos de l'UVHCI avec qui j'ai partagé de grands moments d'actions tels que le WLRC, avec en première ligne, **Alexandre Dias**, **Antoine Maillard**, **Pierre-Alain Coly** (non non pas *E. coli*), **Julien Pérard**, **Euripedes Ribeiro** et **Julien Lupo**. Dans la même lignée, je remercie **Francine Gérard**, **Sophie Torres**, **Céline Fabry**, **Hatice Akarsu**, **Aurélien Albertini**, **Majida El Bakkouri** et **Lucy Freeman**.*

*Un grand merci à mon best friend **Romain Riboud** qui est toujours présent au delà du lieu et de la distance.*

*Pour avoir su me donner le choix et la possibilité de suivre de longues études, je remercie **mes parents**. Merci également de m'avoir donné la priorité sur le "Sommet Bruxellois" !*

*Enfin, je tiens à remercier **Sandra Bailly**, ma pure ps, ma compagne et ma complice pour m'avoir suivi dans cette aventure.*

# Table des matières

<b>I. INTRODUCTION .....</b>	<b>5</b>
<b>PARTIE A – LA GRIPPE, GENERALITES ET STRUCTURE .....</b>	<b>6</b>
1. <i>La grippe, la maladie</i> .....	7
1.1. Pathologie de la grippe .....	7
1.2. Prévention & traitement .....	8
1.3. Épidémiologie .....	9
2. <i>Les virus</i> .....	10
2.1. Les virus à ARN négatif .....	11
3. <i>Le virus de la grippe</i> .....	13
3.1. Les Orthomyxoviridae .....	13
3.2. Les différents sous-types .....	13
3.3. La dérive génétique .....	14
3.4. Le réassortiment génétique .....	16
3.5. La grippe aviaire .....	16
4. <i>Structure du virus de la grippe</i> .....	18
4.1. La particule virale .....	18
4.2. Le génome viral .....	20
5. <i>Les protéines du virus de la grippe</i> .....	22
6. <i>Le cycle de multiplication de virus de la grippe</i> .....	28
6.1. L'entrée du virus : attachement, endocytose, fusion et libération du génome .....	28
6.2. Migration des particules ribonucléoprotéiques virales (RNPv) dans le noyau cellulaire .....	30
6.3. La transcription et la réplication du génome viral .....	30
6.4. L'export nucléaire des RNPv .....	33
6.5. L'emballage des segments dans la particule virale .....	33
6.6. Bourgeonnement et relargage des particules virales .....	34
<b>PARTIE B – LE TRANSPORT NUCLEOCYTOPLASMIQUE .....</b>	<b>35</b>
1. <i>Introduction</i> .....	37
1.1. Le pore nucléaire .....	37
2. <i>Le transport nucléaire, import et export</i> .....	38
2.1. Distribution asymétrique de RanGTP et RanGDP entre le noyau et la cytoplasme .....	39
2.2. L'import nucléaire .....	41
2.3. Les importines .....	43
2.4. L'export nucléaire .....	45
3. <i>Le transport nucléocytoplasmique des protéines et RNP du virus de la grippe</i> .....	47
3.1. L'import nucléaire des RNPv .....	47
3.2. L'import nucléaire des protéines virales .....	49
3.3. L'export nucléaire de RNPv .....	50
<b>OBJECTIFS DE LA THESE .....</b>	<b>52</b>
<b>II. RESULTATS ET DISCUSSION .....</b>	<b>55</b>
<b>CHAPITRE 1 – LE COMPLEXE IMPORTINE ALPHA 5 / SOU-UNITE PB2 DE LA POLYMERASE .....</b>	<b>56</b>
1. <i>Introduction</i> .....	58
2. <i>Isolement et structure RMN du domaine soluble de PB2 (C-ter)</i> .....	60
2.1. Identification et isolement du domaine soluble .....	60
2.2. Spécificité de l'hôte .....	62
2.3. Structure RMN de DPDE en solution .....	63
3. <i>Le complexe Imp<math>\alpha</math>5-DPDE</i> .....	65
3.1. Affinité de PB2 pour l'importine $\alpha$ 5 humaine .....	65
3.2. Structure cristallographique du complexe entre PB2 (C-ter) et l'importine $\alpha$ 5 humaine .....	68
4. <i>Localisation in vivo de PB2</i> .....	75
<b>CHAPITRE 2 – POLYMERISATION DE LA NUCLEOPROTEINE ET ACTIVITE DE FIXATION A L'ARN .....</b>	<b>77</b>
1. <i>Introduction</i> .....	79
1.1. Structure de la NP du virus de la grippe .....	79
2. <i>Caractérisation des différentes formes de nucléoprotéines : NPv, NPrec et NPmono</i> .....	82
2.1. Purification de NPv, de NPrec et du mutant NPmono .....	82
2.2. Caractérisation du mutant NPmono .....	82
2.3. Analyse par microscopie électronique des différentes formes de NP .....	85
3. <i>Les nucléoprotéines des virus à ARN négatifs</i> .....	87
3.1. Etat d'oligomérisation des différentes NP de virus à ARN négatif .....	87

3.2.	Comparaison des structures des nucléoprotéines de virus à ARN négatif .....	89
4.	<i>Affinité de la nucléoprotéine pour l'ARN</i> .....	92
4.1.	Site de fixation à l'ARN .....	92
4.2.	<i>Filter Binding Assays</i> (FBA) .....	94
4.3.	Modélisation de la plateforme de fixation à l'ARN.....	96
<b>CHAPITRE 3 – LE COMPLEXE IMPORTINE ALPHA 5 / NUCLEOPROTEINE</b> .....	<b>99</b>	
1.	<i>Introduction</i> .....	100
2.	<i>Le complexe importine <math>\alpha 5</math> humaine / nucléoprotéine du virus de la grippe</i> .....	104
2.1.	Purification de la NPrec et de l'Imp $\alpha 5$ et formation du complexe NPrec/Imp $\alpha 5$ .....	104
2.2.	Protéolyse limitée du complexe NPrec/Imp $\alpha 5$ .....	107
3.	<i>Caractérisation du complexe NPmono-Imp<math>\alpha 5</math></i> .....	109
3.1.	Caractérisation de l'importine $\alpha 5$ humaine et de la NP monomérique (NPmono).....	109
3.2.	Formation et caractérisation du complexe Imp $\alpha 5$ / NPmono.....	112
3.3.	Affinité du complexe NP-Imp $\alpha 5$ pour l'ARN .....	116
4.	<i>Affinité de la NP pour l'importine <math>\alpha 5</math>, importance de l'acide aminé 319</i> .....	119
<b>III. MA CONCLUSION</b> .....	<b>122</b>	
<b>IV. MATERIEL ET METHODES</b> .....	<b>125</b>	
1.	<i>Expression et purification des protéines</i> .....	126
1.1.	Préparation de la nucléoprotéine virale (NPv).....	126
1.2.	Expression et purification de la nucléoprotéine recombinante et des différents mutants .....	126
1.3.	Expression et purification de la partie C-terminale de PB2 (DPDE : 678-759).....	128
1.4.	Expression et purification de l'importine $\alpha 5$ humaine (66-512).....	129
2.	<i>Les complexes protéine cellulaire - protéine virale</i> .....	130
2.1.	Le complexe importine $\alpha 5$ - DPDE .....	130
2.2.	Le complexe NPrec-imp $\alpha 5$ [66-512] .....	131
2.3.	Le complexe NPmono – Imp $\alpha 5$ .....	131
3.	<i>Ultracentrifugation analytique</i> .....	132
4.	<i>Microscopie électronique à transmission (MET) en coloration négative</i> .....	132
5.	<i>Affinité de la nucléoprotéine (NPv, NPrec et NPmono) ou du complexe NPrec-Imp<math>\alpha 5</math> pour l'ARN</i> .....	132
5.1.	Synthèse et marquage de l'ARN <i>panhandle</i> par transcription <i>in vitro</i> .....	132
5.2.	<i>Filter Binding Assays</i> (FBA) .....	133
6.	<i>Calorimétrie de titration isotherme (ITC - Isothermal Titration Calorimetry)</i> .....	133
<b>V. REFERENCES BIBLIOGRAPHIQUES</b> .....	<b>134</b>	
<b>ANNEXE</b> .....	<b>149</b>	

# Abréviations

AA : Acide aminé  
Å : Angström  
ADN : Acide désoxyribonucléique  
ARM : Armadillo repeat (motif répété de 3 hélices)  
ARN : Acide ribonucléique  
β-SH : β-mercaptoéthanol  
Da : Dalton  
DO : Densité Optique  
DTT : Dithiothreitol  
GEIG : Groupe d'Expertise et d'Information sur la Grippe  
IPTG : IsoPropyl-beta-D-ThioGalactopyranoside  
LB : Luria-Bertani Broth  
LMB : Leptomycin B  
MET : Microscopie Electronique à Transmission  
MM : Masse moléculaire  
MPD : 2 methyl-1,3 pentanediol  
OMS : Organisation Mondiale de la Santé  
NES : Nuclear Export Signal (signal d'export nucléaire)  
NLS : Nuclear Localisation Signal (signal de localisation nucléaire)  
NPC : Nuclear Pore Complex (le complexe du pore nucléaire)  
PDB : Protein Data Bank (base de données des structures cristallographiques de protéines)  
PMSF : Phenylmethylsulphonylfluoride  
Rmsd : Root mean square deviation  
rpm : rotation par minute  
SDS-PAGE : Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis  
SW : Swing  
TEV : Tobacco Etch Virus  
Tris : 2-Amino-2-hydroxymethyl-1,3-propanediol  
°C : degré Celsius  
h : heure  
min : minute  
3D : Trois dimensions

## Code à une et trois lettres des 20 acides aminés :

Alanine :	A	Ala	Méthionine :	M	Met
Cystéine :	C	Cys	Asparagine :	N	Asn
Acides aspartique :	D	Asp	Proline :	P	Pro
Acide glutamique :	E	Glu	Glutamine :	Q	Gln
Phénylalanine :	F	Phe	Arginine :	R	Arg
Glycine :	G	Gly	Sérine :	S	Ser
Histidine :	H	His	Thréonine :	T	Thr
Isoleucine :	I	Ile	Valine :	V	Val
Lysine :	K	Lys	Tryptophane :	W	Trp
Leucine :	L	Leu	Tyrosine :	Y	Tyr

# **I. INTRODUCTION**

## **Partie A**

—

### **La grippe, généralités et structure**



# Partie A - La grippe, généralités et structure

## 1. La grippe, la maladie

La grippe (ou *influenza*) est une maladie respiratoire infectieuse, aiguë et très contagieuse. Sa transmission est typiquement interhumaine par voie respiratoire, *via* des gouttelettes riches en virus provenant de la toux et des éternuements des sujets infectés. Sévissant vraisemblablement déjà dans l'antiquité, ce n'est qu'au XIV<sup>e</sup> siècle, à Florence, que le terme "influenza" est employé. Il évoque l'influence des astres et du froid et rappelle le caractère saisonnier de la maladie, qui laisse supposer l'influence de l'environnement extérieur sur l'homme. Le mot français *grippe* aurait une origine germanique, *Grippen* signifiant "saisir brusquement". Autrement dit : on n'attrape pas la grippe, c'est elle qui nous "agrippe". Le virus grippal infecte d'autres mammifères que l'homme pouvant être terrestres et marins. Cependant, le réservoir naturel reste confiné aux oiseaux.

### 1.1. Pathologie de la grippe

Cette maladie est causée par un virus et est caractérisée par l'apparition de symptômes tels que maux de tête, fièvre, frissons, douleurs musculaires, toux, difficulté respiratoire, affaiblissement de l'organisme, accompagné d'un sentiment de malaise général. Ces symptômes évoquent le diagnostic grippal par leur apparition soudaine, leur arrivée en période d'épidémie de grippe et leur disparition généralement après quelques jours d'évolution. Habituellement la fièvre diminue 2 à 3 jours suivant l'apparition des premiers symptômes, mais la toux et le malaise général peuvent persister durant 1 à 2 semaines. L'infection affecte et détruit partiellement l'épithélium des voies respiratoires, sa régénération pouvant prendre jusqu'à un mois pour être complète. Dans les cas les plus sérieux, la maladie se complique en infections bactériennes secondaires et en pneumonies virales parfois fatales (Tashiro *et al.*, 1987 ; Murphy & Webster, 1990). Les décès observés alors ne sont pas directement imputables à la grippe proprement dite mais à d'autres agents pathogènes (viraux ou bactériens) secondaires à l'infection grippale. L'épithélium respiratoire mis à mal par le développement du virus de la grippe devient une porte d'entrée pour les autres pathogènes. La grippe est particulièrement dangereuse pour les personnes appartenant à certains groupes dits à risque élevé. Ils comprennent les nouveaux nés et enfants en bas âge, les personnes âgées,

les femmes enceintes ou encore les personnes souffrant de maladies cardiovasculaires ou pulmonaires. Considérant le nombre de personnes atteintes chaque année dans le monde, la grippe constitue une des maladies infectieuses majeures.

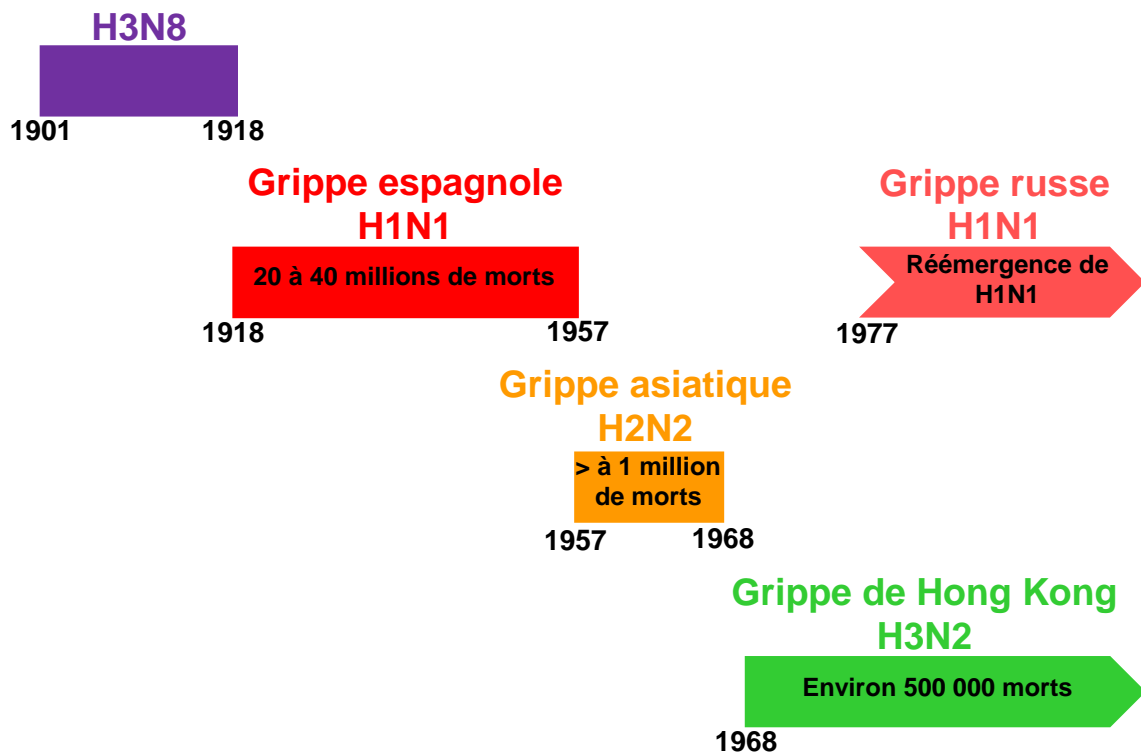
## 1.2. Prévention & traitement

Outre l'hygiène, la vaccination dans les pays où elle est accessible semble être la meilleure parade. Elle en diminue significativement le nombre d'hospitalisations ainsi que la mortalité. Selon l'[Organisation Mondiale de la Santé \(OMS\)](#), il semblerait que chez les personnes âgées, la vaccination diminuerait de 60 % la morbidité, et de 70 à 80 % la mortalité due à la grippe. Chez les adultes en bonne santé, le vaccin est très efficace (70 à 90 %) pour ce qui est de diminuer la morbidité, et il a été montré que la vaccination présentait des avantages substantiels pour cette tranche d'âge, tant du point de vue sanitaire qu'économique. La mise en quarantaine est un moyen efficace, mais difficile à mettre en œuvre. Dans certains pays, par exemple en France et en Belgique, les sujets à risque ont accès à une vaccination gratuite. Elle est recommandée aux professionnels de santé et aux personnes travaillant à des postes " socialement " exposés (enseignants, administration, transports en commun...). Le vaccin le plus commun est une suspension de particules virales inactivées ou fractionnées qui offre une protection contre trois souches virales. Il comporte toujours des particules de deux types de virus influenza A et d'un type de virus influenza B. Une vaccination annuelle est nécessaire en raison de la variabilité antigénique des virus grippaux. Les souches virales utilisées dans les vaccins sont réalisées en fonction des souches en circulation. Les moyens de lutte contre la grippe sont peu nombreux et d'efficacité moyenne. Les antiviraux qui existent sont utilisés avec beaucoup de précaution à cause d'importants effets secondaires pour certains d'entre eux. D'autre part, notons que des cas résistances à ces antiviraux ont été observés. La rimantadine, qui n'est désormais plus commercialisée en France, et l'amantadine ont une efficacité de l'ordre de 60 % si administrés à titre préventif ([Jefferson et al., 2006](#)). Ces deux molécules se fixent à une protéine virale (la protéine canal M2) impliquée dans la libération du matériel génétique du virus en phase précoce d'infection. Il existe également des inhibiteurs d'une autre protéine virale localisée en surface du virus, la neuraminidase (NA). Ces inhibiteurs sont l'ozeltamivir (Tamiflu®), commercialisé par Roche, et le zanamivir (Relenza®) par GlaxoSmithKline. Ces traitements, pris précocement, peuvent diminuer un peu l'importance des symptômes et la durée de l'affection. Le blocage de l'activité enzymatique de la NA empêche la libération des virus nouvellement produits de la cellule

hôte, et donc la diffusion de l'infection virale. Mais ils sont coûteux, ne sont pas remboursés par la sécurité sociale en France, en Belgique et en Suisse et doivent être pris dans les 48 heures après l'apparition des symptômes.

### 1.3. Épidémiologie

Banalisée dans l'esprit populaire comme synonyme de rhume ou de " coup de froid ", la grippe est une maladie autrement sérieuse et un problème majeur de santé publique à l'échelle planétaire. Sévissant sur un mode épidémique saisonnier, elle est responsable dans le monde d'une morbidité élevée et de la mort de plusieurs centaines de milliers de personnes chaque année.



**Figure 1. Historique des pandémies de grippe chez l'homme au XXe siècle.** Des données sérologiques suggèrent qu'au début du XXème siècle un virus de grippe de type A et de sous-type H3N8 circulait chez l'homme. L'apparition d'un virus d'origine aviaire, H1N1, en 1918 a provoqué la pandémie appelée "grippe espagnole" : 20 à 40 millions de morts selon les estimations. En 1957, un nouveau sous-type, H2N2, a provoqué la pandémie dite "asiatique" : plus d'1 million de morts. En 1968, le virus de sous-type H3N2 a évincé H2N2 causant la pandémie dite de "Hong Kong" : environ 500 000 morts. Depuis 1977, des virus de sous-types H1N1 et H3N2 co-circulent dans la population humaine.

Même si ces chiffres sont difficiles à évaluer, il est estimé que les épidémies annuelles entraînent entre 3 et 5 millions de cas graves et 250 000 à 500 000 décès par an dans le monde (source : OMS). En France, ce sont 1500 à 2000 personnes qui décèdent chaque année des suites de la grippe. Cette maladie présente également un impact économique annuel important puisqu'en France, selon l'intensité de l'épidémie, le coût en médecine ambulatoire (hors hospitalisation et perte de production) varie entre 230 et 840 millions d'euros (Source : Groupe d'Expertise et d'Information sur la Grippe (GEIG)). Indépendamment des épidémies, des pandémies grippales sont susceptibles de survenir lors de l'émergence de nouvelles souches virales. Depuis le début du XXe siècle, il y eu trois épidémies de grippe exceptionnelles par leur extension et leur gravité. Elles résultaient de l'émergence de nouveaux sous-types de virus vis-à-vis desquels la population humaine ne possédait pas d'anticorps. Ayant atteint la totalité du globe rapidement, touché des sujets jeunes et présenté un taux de mortalité très élevé, ces épidémies furent ainsi dénommées pandémies. La plus connue et la plus meurtrière à ce jour est celle de la "grippe espagnole" (virus influenza A, sous-type H1N1) qui sévit entre 1918 et 1919. Selon les estimations, elle serait responsable de 20 à 40 millions de morts. Ce surnom provient du fait que seule l'Espagne – non impliquée dans la première guerre mondiale – a pu, en 1918, publier librement les informations relatives à cette épidémie. Les journaux français parlaient donc de la "grippe espagnole" qui faisait des ravages en Espagne. Et ce, sans mentionner les cas français qui étaient tenus secrets pour ne pas faire savoir à l'ennemi que l'armée était affaiblie. Les pandémies suivantes furent moins dévastatrices. La "grippe asiatique" de 1957 (virus de type A, souche H2N2) responsable de plus d'1 million de morts et la "grippe de Hong Kong" de 1968 (virus de type A, souche H3N2) moins sévère mais néanmoins responsable de la mort d'environ 500 000 personnes (**figure 1**). Depuis son apparition en 1997 en Asie (Hong Kong), la souche de virus aviaire de type A et de sous-type H5N1 fait planer le risque d'une nouvelle pandémie grippale.

## 2. Les virus

Le rhume, la varicelle, la mononucléose infectieuse et la grippe sont des exemples de maladies humaines virales relativement courantes. Parmi les maladies infectieuses les plus meurtrières, beaucoup sont également causées par des virus : les fièvres hémorragiques (virus Ebola et Marburg), le sida (VIH), la rougeole, la rage, la pneumonie atypique (SRAS), la variole ou encore la grippe aviaire. Certains de ces virus, comme c'est le cas pour la rage, sont

connus depuis des siècles. En revanche, d'autres sont des virus émergents. Ils s'adaptent rapidement à de nouveaux hôtes et sont ainsi responsables d'épidémies à forte mortalité.

Tout d'abord, un virus est une entité biologique. Il a besoin d'une cellule hôte et de ses constituants pour se multiplier et donc pour exister. Il se caractérise ainsi par son incapacité à se multiplier par division. Cependant un virus possède un génome contenant l'information nécessaire à sa réplication. Cette information est stockée à l'intérieur d'une molécule d'acide nucléique, de l'ADN ou de l'ARN, qui peut être simple ou double brin. En fonction de cette caractéristique, les virus sont ainsi appelés virus à ADN ou virus à ARN. Contrairement aux virus à ADN qui ont accès à la machinerie cellulaire de réplication d'ADN, les virus à ARN sont obligés de posséder leur propre réplicase (polymérase virale) puisque la cellule ne possède pas d'ARN polymérase ARN-dépendante (Palese & Shaw, 2007).

### 2.1. Les virus à ARN négatif

Lorsqu'il est simple brin, l'ARN viral génomique peut être soit de polarité positive soit, comme c'est le cas de la grippe, de polarité négative. La polarité négative de l'ARN implique qu'il ne peut pas être traduit en protéines. Un brin d'ARN complémentaire doit d'abord être synthétisé par la polymérase virale. Cet ARN sera ainsi de polarité positive (ARN(+)) et équivalent à l'ARN messenger cellulaire. En d'autres termes, les virus à ARN négatif ont un génome dont le sens est complémentaire à celui de l'ARN messenger.

Le groupe des virus à ARN négatif (ARN(-)) comporte 7 familles de virus dont la plupart est responsable de graves pathologies chez l'homme (**tableau 1**). Certains de ces virus provoquent des maladies très impressionnantes telles que les fièvres hémorragiques. D'autres sont la cause de maladies beaucoup plus fréquentes comme la rougeole (en Afrique), la rage (en Asie et en Afrique) ou encore les oreillons et la grippe.

Les virus à ARN(-) diffèrent entre eux selon certains critères : la segmentation ou non de leur génome et le lieu de leurs étapes de transcription et de réplication dans la cellule. La taxonomie des virus est la suivante : ordre/famille/genre. Dans cette classification, on retrouve les virus à ARN(-) dans deux groupes : celui des virus dont le génome est non-segmenté, ordre des *Mononegavirales*, et celui des virus dont le génome est segmenté (**tableau 1**). Seules les familles des *Bornaviridae* (ordre des *Mononegavirales*) et des *Orthomyxoviridae* (génome segmenté), dont la grippe, ont une étape nucléaire durant leur cycle réplcatif, les autres virus à ARN(-) restent dans le cytoplasme. De façon générale, la plupart des virus à ARN humains restent dans le cytoplasme de la cellule hôte pour amplifier leur génome et

exprimer les protéines virales. En ce qui concerne le virus de la grippe, à ARN(-) segmenté, ses étapes de transcription et de réplication ont également lieu dans le noyau. Certains segments du génome viral codent pour plusieurs protéines, signifiant qu'une étape d'épissage des ARN messagers viraux est nécessaire. Le virus de la grippe utilise ainsi la machinerie d'épissage cellulaire qui se trouve dans le noyau.

**Tableau 1. Les virus à ARN négatif simple brin.**

<b>Ordre</b> <b>Famille</b> Genre	<b>Exemples de virus connus</b>	<b>Taille du génome (kb)</b>
<b>Mononegavirales</b>		
<b>Rhabdoviridae</b> Vesiculovirus Lyssavirus	Virus de la stomatite vésiculaire (VSV) Virus de la rage	<b>13-16</b>
<b>Filoviridae</b> Type Marburg Type Ebola	Virus de Marburg Virus Ebola	<b>13</b>
<b>Paramyxoviridae</b> Respirovirus Morbillivirus Rubulavirus	Virus de Sendai Virus de la rougeole Virus des oreillons	<b>16-20</b>
<b>Bornaviridae</b> Bornavirus	Borna disease virus	<b>9</b>
<b>Génomés segmentés</b>		
<b>Orthomyxoviridae</b> Influenza A Influenza B Influenza C Thogotovirus Isavirus	Virus de la grippe A Virus de la grippe B Virus de la grippe C Virus thogoto / Dhori Anémie infectieuse du saumon	<b>12-15 kb</b> dans : 8 segments 8 segments 7 segments 6/7 segments 8 segments
<b>Bunyaviridae</b> Bunyavirus Hantavirus Phlebovirus Nairovirus Tospovirus	Bunyamwera Virus de séoul Virus de la vallée du rift Virus Congo-Crimée Tomato Spotted Wilt Virus	<b>11-20 kb</b> dans 3 seg.
<b>Arenaviridae</b> Arenavirus	Virus de la fièvre de Lassa	<b>10-14 kb</b> dans 2 segments

### 3. Le virus de la grippe

#### 3.1. Les Orthomyxoviridae

Les virus grippaux, responsables d'infections respiratoires aiguës chez l'homme, appartiennent à la famille des Orthomyxoviridae (génomme à ARN segmenté simple brin de polarité négative). Cette famille comprend 5 genres : les influenza A, B et C ainsi que les genres Thogotovirus et Isavirus. Les gripes saisonnières que l'on connaît, correspondent aux genres A et B qui circulent en parallèle durant les épidémies. Le virus de la grippe de type A possède un large spectre d'hôtes. Il peut infecter les humains et d'autres espèces de mammifères terrestres (les porcs, les chevaux, les chiens) ou marins (les baleines, les phoques). Mais le réservoir naturel est avant tout celui des oiseaux sauvages aquatiques. Le virus de la grippe de type B, en revanche, est quasiment spécifique de l'espèce humaine même si on a pu le rencontrer chez le phoque ([Osterhaus et al., 2000](#)). A la suite d'un prélèvement effectué en 1947 ([Taylor, 1949](#)) un virus d'un troisième type a été isolé : la grippe de type C. Ce virus infecte également plusieurs espèces dont l'homme. Cependant, il est généralement admis que ces infections causent des syndromes grippaux peu sévères, de plus il n'est pas responsable d'épidémies annuelles chez l'homme. L'information génétique des virus grippaux de types A et B est constituée de 8 segments d'ARN viral, contre 7 pour les virus de type C.

#### 3.2. Les différents sous-types

On distingue les 3 types de virus grippaux (A, B et C) sur la base de l'antigénicité d'une protéine interne, la nucléoprotéine (NP). Parmi les virus de type A (dont certains sont hautement pathogènes pour l'homme) il existe différents sous-types antigéniques définis par la nature des deux glycoprotéines de surface : l'hémagglutinine (HA) et la neuraminidase (NA). Ainsi il se distingue 16 espèces de HA (H1 à H16) et 9 espèces de NA (N1 à N9) menant à différentes combinaisons possibles (ex : H1N1, H3N2...). Toutes ces espèces sont présentes chez les oiseaux aquatiques qui sembleraient constituer le réservoir de la diversité génétique des virus grippaux. En revanche chez l'homme ou chez d'autres mammifères (porc, cheval), seuls certains sous-types ont été documentés. Ainsi la grippe annuelle chez l'homme est provoquée par des virus à H1, H2, H3 pour l'HA et N1 ou N2 pour la NA et les souches de grippe circulant ces dernières décennies ont principalement H1N1 et H3N2 pour sous-types. Toutefois, les dernières épidémies de grippe aviaire en Asie ont montré la capacité de certains sous-types à infecter l'homme (H5, H7, H9). Les sous-types à H5 et H7 sont actuellement

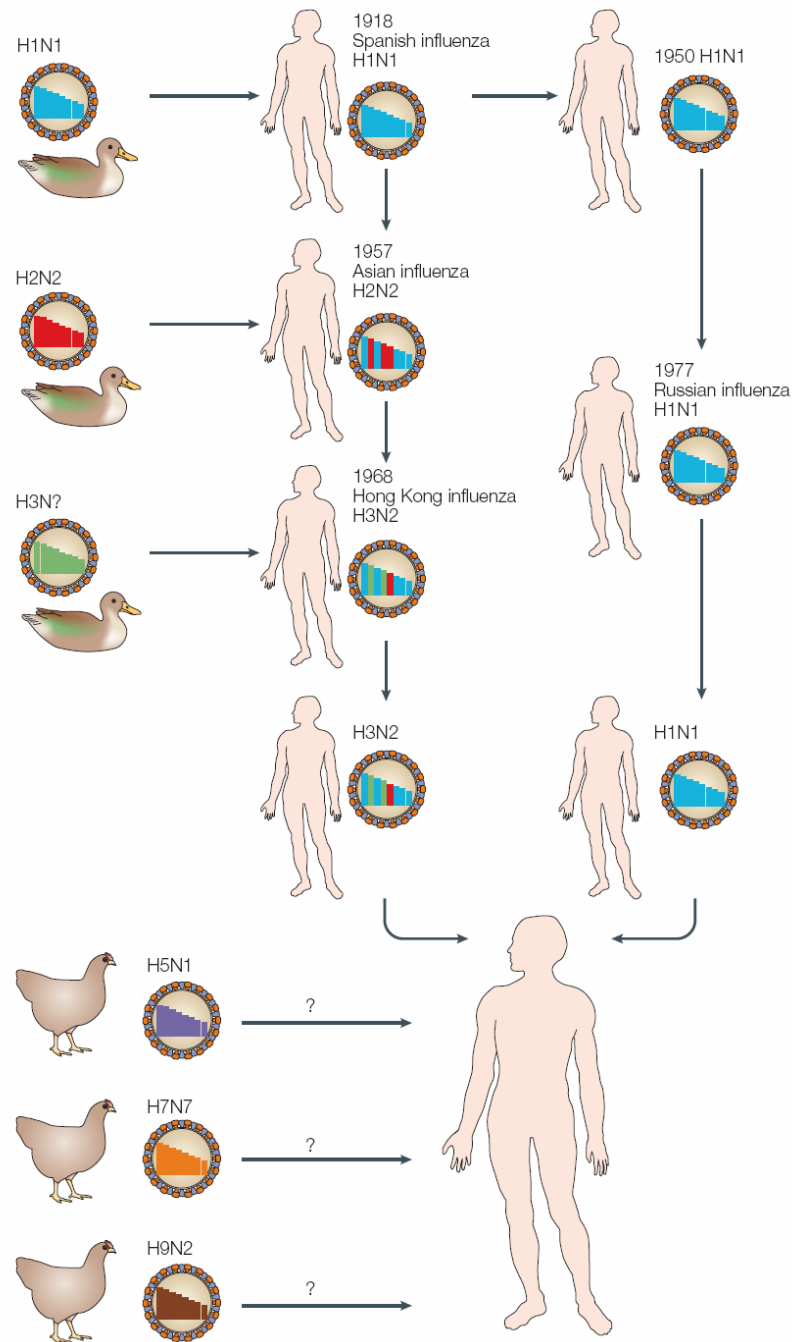
réputés hautement pathogènes chez les oiseaux. Parmi eux, le virus H5N1, se propage sous forme de vaste épidémie dans le monde aviaire et ne se transmet qu'occasionnellement à l'homme. La transmission interhumaine quant à elle n'a pas encore été établie.

La nomenclature actuelle des souches grippales inclut le type, le lieu, le numéro et l'année de l'isolement ainsi que la nature des antigènes de surface : A/Puerto Rico/8/34 (H1N1) par exemple. Pour les virus animaux, le lieu d'isolement est précédé par le nom de l'espèce animale d'origine : A/chicken/Shantou/423/2003 (H5N1) par exemple.

### 3.3. La dérive génétique

Ce phénomène classique est le résultat de l'accumulation de mutations lorsque le virus réplique son génome. Ces mutations sont dues à la faible fidélité ainsi qu'à l'absence d'activité correctrice de l'ARN polymérase-ARN dépendante. En effet, le taux d'erreur élevé de cette réplique est estimé à  $4 \cdot 10^{-3}$  et  $4,4 \cdot 10^{-3}$  substitutions/site/an pour la HA des virus A(H3N2) et A(H1N1) respectivement (Cox *et al.*, 1993). De nombreux variants de virus grippaux atténués ou non-viables sont ainsi produits, mais la pression de sélection exercée par le système immunitaire de l'hôte, notamment les anticorps neutralisants, ne conduit qu'à la viabilité de certains virus. L'évolution des gènes qui codent pour les protéines de surface (HA et NA) et ainsi beaucoup plus rapide que celle des protéines situées à l'intérieur du virus. Pour les virus grippaux humains 4 à 19% des substitutions nucléotidiques, se traduisant par des changements en acides aminés, se produisent pour les séquences codant les protéines PB2, NP, ou M1 alors qu'elle est de 43% pour la séquence de la HA (Daly *et al.*, 1998). Ainsi, au sein d'un même individu, la fréquence élevée d'erreurs dues à l'ARN polymérase durant le cycle répliatif, induit une population de virus grippaux qui n'est pas homogène. Cet ensemble forme un complexe appelé "quasi espèce" constitué de variants dont les génomes diffèrent les uns des autres par une ou plusieurs mutations.





**Figure 2. Pandémies passées et potentielles à venir du virus de la grippe (d'après Horimoto & Kawaoka, 2005, revue).** La pandémie de 1918 aurait été causée par un virus H1N1 aviaire transmis à l'homme. Celle de 1957 fut causée par un nouveau virus H2N2 issu du réassortiment du sous-type H1N1 adapté à l'homme avec un sous-type H2N2 aviaire. Ce nouveau sous-type nommé H2N2 possède les segments PB1, HA et NA d'origine aviaire (H2N2 aviaire) dans un environnement de génotype du virus de la grippe adapté à l'homme (H1N1). H1N1 disparaît alors de la circulation. La pandémie de 1968 (grippe de Hong Kong, sous-type H3N2) fut provoquée par la combinaison des segments PB1 et HA d'un sous-type H3N2? aviaire avec H2N2 humain. Les nombreux cas d'épizooties en Asie du Sud-Est de ces dernières années semblent se propager progressivement vers l'ouest, faisant craindre l'apparition de la première pandémie du 21ème siècle.

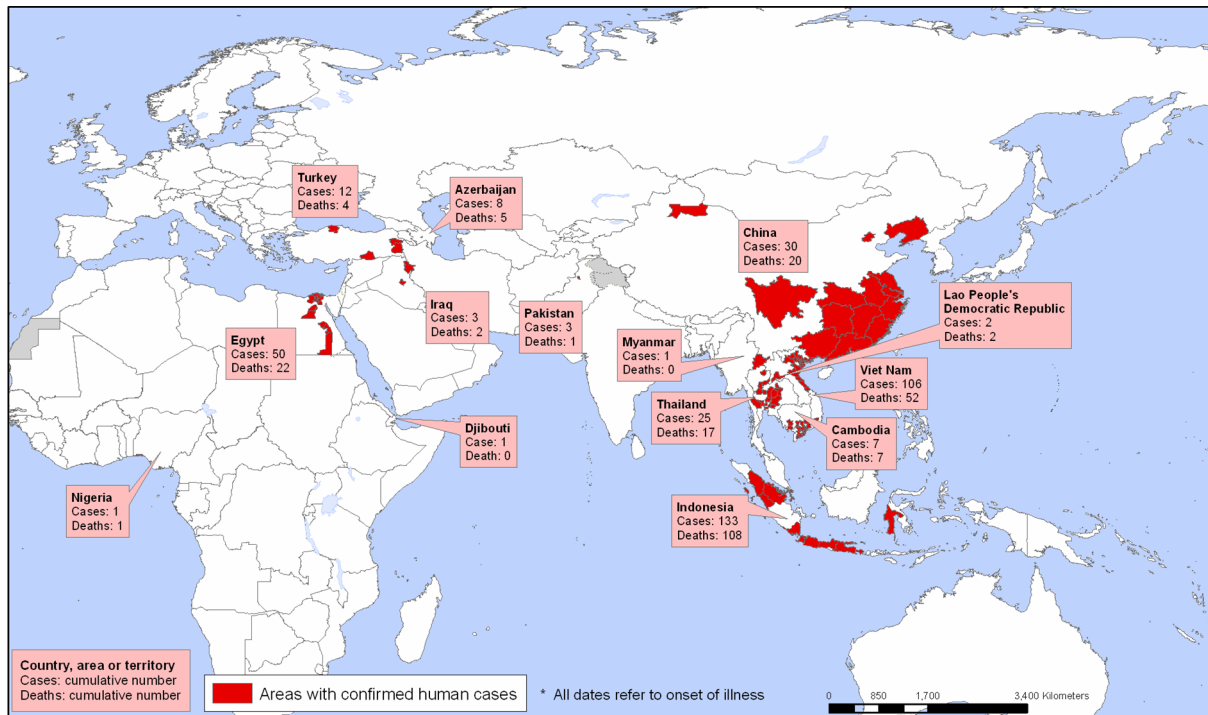
### 3.4. Le réassortiment génétique

La nature segmentée du génome des virus grippaux est à l'origine de ce phénomène. Il se produit à la suite d'une co-infection d'un même hôte par différents virus, lorsque les segments génomiques de ces virus sont redistribués aléatoirement au sein des virus produits. Ce processus semble se produire régulièrement entre virus humains (Lindstrom *et al.*, 1998 ; Obenauer *et al.*, 2006). Par ailleurs, des événements de réassortiments entre virus de type A humain et aviaire peuvent se produire. Lorsque de tels réassortiments surviennent, ils peuvent engendrer l'émergence de nouveaux sous-types capables d'une part, de passer la barrière de l'espèce, et d'autre part, être responsables de pandémies de grippe. Ainsi dans la population humaine, ce mécanisme serait à l'origine des pandémies de 1957 (H2N2, grippe asiatique) et de 1968 (H3N2). En 1957, une souche H2N2 d'origine aviaire aurait échangé ses segments génomique (PB1, HA et NA) avec la souche adaptée à l'humain H1N1. Dans la pandémie de 1968, à nouveau une souche d'origine aviaire H3N2 (sous-type de NA non connu) aurait incorporé des segments génomiques (PB1 et HA) dans le génome de la souche H2N2 adaptée à l'homme issue du virus de 1957 (Olsen, 2002 ; Horimoto & Kawaoka, 2005). Concernant la pandémie de 1918 (H1N1), son origine est moins claire. Un virus purement aviaire se serait soit directement adapté à l'homme, soit il serait passé par une phase d'adaptation chez un autre mammifère. Ainsi dérive et réassortiment génétiques conduisent à l'apparition de nouvelles souches potentiellement très virulentes pour l'homme et donc à de graves pandémies cycliques (**figure 2**).

### 3.5. La grippe aviaire

L'appellation "grippe aviaire" définit la pathologie de la maladie chez les oiseaux. Cependant, il est à noter que l'ensemble des différentes souches de grippe sévissant dans les populations animales, y compris l'homme, a pour point de départ un virus d'origine aviaire. Bien souvent banalisée comme H5N1, la grippe aviaire est également causée par d'autres sous-types. Depuis 1959, l'infection humaine par un virus grippal aviaire n'a été établie qu'à 10 reprises (source : OMS). Selon cette organisation, sur les centaines de sous-types de virus grippaux aviaires A, seul un certain nombre a provoqué des infections humaines : par exemple H5N1, H7N3, H7N7 et H9N2. En général, l'infection humaine par ces virus n'entraîne que des symptômes légers et une maladie bénigne, à une exception notable près : le virus H5N1 hautement pathogène. Il est de tous les virus grippaux, en circulation dans la population aviaire, le plus préoccupant pour la santé humaine principalement pour deux raisons.

Premièrement, il a provoqué le plus grand nombre de cas humains très graves et le plus grand nombre de décès avec un taux de mortalité proche de 60%. Il a franchi la barrière de l'espèce à plusieurs reprises. Tout d'abord à Hong Kong en 1997, 18 personnes furent infectées, dont 6 sont décédées, puis en 2003 à Hong Kong également (2 cas, dont 1 mortel) et lors des flambées qui ont commencé en décembre 2003. A l'heure actuelle sur la planète (Sept. 2008), c'est plus de 380 cas d'infections humaines dont 245 décès répertoriés par l'OMS (figure 3).



**Figure 3. Zones affectées par des cas humains confirmés de grippe aviaire (H5N1) depuis 2003 (en date de sept. 2008) (source : Organisation Mondiale de la Santé).**

La seconde raison, de loin la plus préoccupante, est le risque que le virus H5N1 puisse, s'il en a l'occasion, acquérir les caractéristiques nécessaires pour déclencher chez l'homme, une nouvelle pandémie de grippe. Le virus remplit toutes les conditions requises sauf une : la capacité de se transmettre efficacement d'une personne à l'autre. Le risque de pandémie vient donc du fait que ce virus de la grippe aviaire puisse améliorer sa transmissibilité par dérive ou réassortiment génétique. Le taux de mortalité chez les personnes déclarant les symptômes est nettement supérieur à 50%. Cependant, il n'est pas exclu que si le virus H5N1 devient un jour transmissible entre hommes, il ne perde de sa virulence.

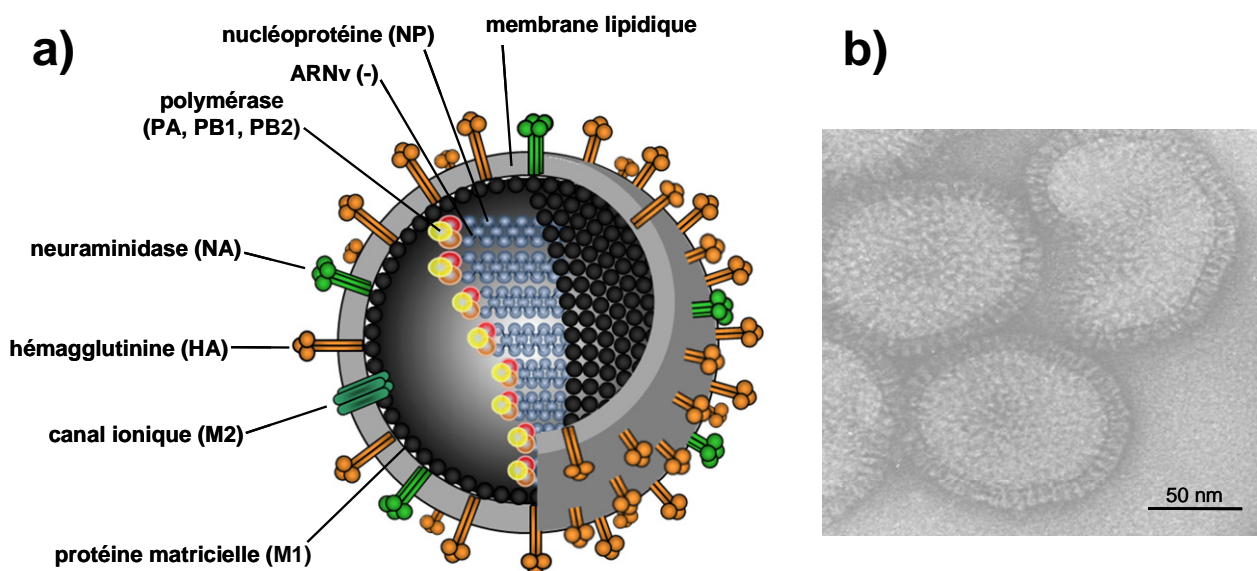
La transmission d'un virus, d'une espèce à l'autre, dépend tout d'abord des protéines de surface (revue, Neumann & Kawaoka, 2006). De récents travaux ont également montré

l'implication de la polymérase et de la nucléoprotéine dans l'adaptation d'un virus aviaire à un hôte mammifère, la souris ([Gabriel \*et al.\*, 2005](#)).

## 4. Structure du virus de la grippe

### 4.1. La particule virale

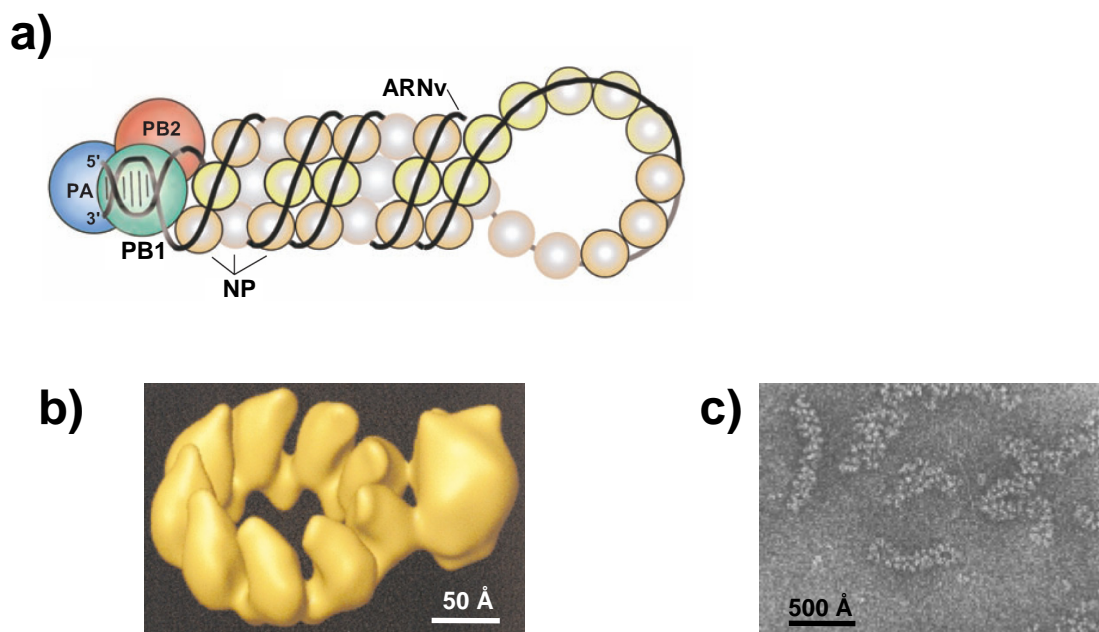
Le virus de la grippe se présente sous forme d'un virion sphérique d'environ 120 nm de diamètre. Il est délimité par une bicouche lipidique qui dérive de la membrane plasmique d'une cellule infectée. Insérées dans cette membrane se trouvent deux glycoprotéines de surface transmembranaires associées, en trimère pour l'hémagglutinine et en tétramère pour la neuraminidase (**figure 4a**). La HA, par ses fonctions d'attachement aux récepteurs (les acides sialiques) et de fusion, assure l'entrée du virus dans la cellule. La NA, par son activité de sialidase, assure la destruction des récepteurs et favorise ainsi la libération des particules virales néoformées. Egalement insérée dans la bicouche lipidique, mais en quantité minoritaire, la protéine M2 associée en homotétramère, joue le rôle de canal ionique. Elle intervient durant l'étape de libération du génome dans le cytoplasme de la cellule nouvellement infectée.



**Figure 4. Structure des virus grippaux.** a) Représentation schématique de la structure d'un virion. Les glycoprotéines HA et NA ainsi que le canal ionique M2 sont insérés dans la membrane lipidique de l'enveloppe. La protéine M1 tapisse la face interne de l'enveloppe. Les 8 segments d'ARNv sont associés à la NP et aux protéines du complexe de polymérase formant la particule ribonucléoprotéique (RNP). b) Image en microscopie électronique à coloration négative de virus grippaux ([cliché, par Rob Ruigrok](#)).

Les coupes de microscopie électronique montrent que la face interne de l'enveloppe virale est tapissée par une protéine, c'est la protéine matricielle M1 (**figure 4b**). Elle assure en partie la structure de la particule virale et est essentielle à son assemblage. Elle fait également le lien entre les constituants internes et externes, ainsi elle interagit d'une part avec les queues cytoplasmiques des glycoprotéines et d'autre part avec le génome (*Ye et al., 1999*).

A l'intérieur de la particule virale, les segments génomiques sont étroitement empaquetés (**figure 4**). Chacun d'eux est entièrement recouvert par la nucléoprotéine (NP) qui oligomérisse le long de ces derniers, l'ensemble formant une structure hélicoïdale d'environ 9 nm de diamètre. La stoechiométrie d'interaction entre la NP et l'ARNv est estimée à une NP pour 24 nucléotides d'ARNv (*Ortega et al., 2000*). Associée à cette structure, on trouve la polymérase virale, composée de trois sous-unités : PA, PB1 et PB2, fixée aux extrémités 5' et 3' de chacun des segments d'ARN (*Compans et al., 1972*). Cet ensemble ARN, nucléoprotéines et polymérase constitue la particule ribonucléoprotéique ou RNP (**figure 5**). Cette molécule représente à elle seule une unité active et indépendante vis à vis de la transcription et de la réplication du génome viral, au sein du noyau des cellules infectées, lorsque lui sont fournis les 4 différents nucléotides ainsi qu'une amorce.



**Figure 5. Organisation structurale des RNP des virus grippaux de type A.** a) Schéma de la structure des RNP (*Naffakh et al., 2008 adapté d'après Portela & Digard, 2002*). Les sphères bleues représentent la NP associée à l'ARNv, en noir. Les extrémités 5' et 3' de l'ARNv s'associent pour former une courte région appariée où se fixe le complexe de la polymérase. b) Modèle 3D de la structure d'une mini RNP recombinante résolue à 23 Å, réalisé en microscopie électronique (*Martin-Benito et al., 2001*). c) Image en microscopie électronique à coloration négative de RNP intacts (*Ruigrok & Baudin, 1995*).

#### 4.2. Le génome viral

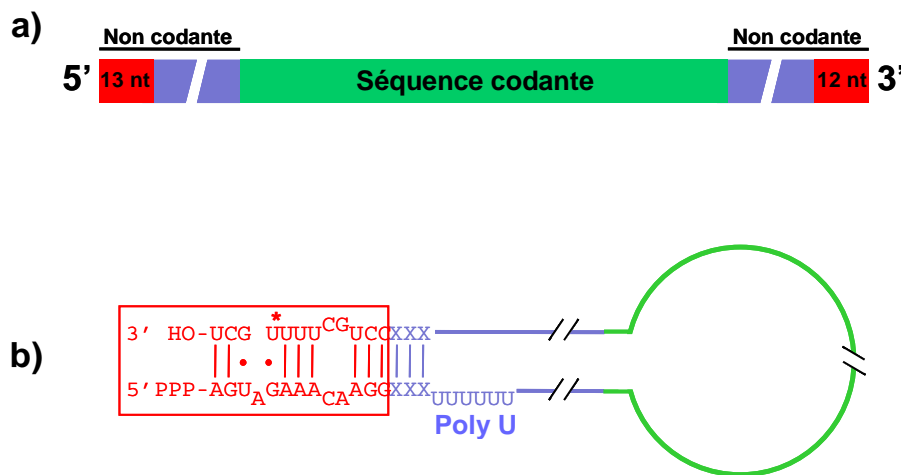
L'intérieur du virus contient le génome qui présente la particularité d'être segmenté. Il comprend 8 segments, de taille variable d'ARN viral (ARNv) monocaténaire et de polarité négative (ARN(-)), codant pour 11 protéines différentes (**tableau 2**). Certains segments codent pour une seule protéine virale comme c'est le cas du segment 1 (sous-unité PB2 de la polymérase), du segment 3 (sous-unité PA de la polymérase), du segment 4 (HA), du segment 5 (la nucléoprotéine NP) et du segment 6 (NA). Les autres segments codent pour deux protéines virales différentes.

Segment	Taille (nucléotides)	Protéine codée	Abréviation	Taille (acides aminés)	Poids moléculaire
1	2341 nt	Protéine basique 2	PB2	759	85,7 kDa
2	2341 nt	Protéine basique 1	PB1	757	86,5 kDa
		Protéine basique 1-F2	PB1-F2	87	10,5 kDa
3	2233 nt	Protéine acide	PA	716	84,2 kDa
4	1778 nt	Hémagglutinine	HA	566	61,5 kDa*
5	1565 nt	Nucléoprotéine	NP	498	56,1 kDa
6	1413 nt	Neuraminidase	NA	454	50,1 kDa*
7	1027 nt	Protéine matrice 1	M1	252	27,8 kDa
		Protéine matrice 2	M2	97	11,0 kDa
8	890 nt	Non structurale 1	NS1	230	26,8 kDa
		Nuclear export protein	NEP	121	14,2 kDa

**Tableau 2. Protéines codées par le génome du virus de la grippe de type A (A/PR/8/34).** La plupart des segments sont transcrits en un ARNm qui code pour seule protéine. Cependant l'ARNm transcrit à partir de PB1 présente un second cadre de lecture initiant la traduction de la protéine PB1-F2. Les ARNm transcrit à partir des segments 7 et 8 subissent un épissage alternatif de sorte qu'ils codent pour 2 protéines, M1 et M2 ou NS1 et NEP, respectivement. La masse moléculaire indiquée ne prend pas en compte l'effet de la glycosylation (d'après Lamb & Krug, 2001).

Ainsi le segment 2 code à la fois pour la sous-unité PB1 de la polymérase et pour la protéine PB1F2 découverte en 2001 (Chen *et al.*, 2001), le segment 7 code pour les protéines virales M1 et M2 (Lamb *et al.*, 1981), enfin le segment 8 code pour la protéine non structurale NS1 et la protéine d'export nucléaire NEP anciennement appelée NS2 (Lamb & Chopin, 1979 ; Inglis *et al.*, 1979). Alors que l'expression des protéines virales PB1 et PB1-F2 est basée sur une traduction alternative du segment 2 en fonction de deux cadres de lecture, celle des protéines virales M1/M2 et NS1/NS2 est liée à un épissage des transcrits codés par les segments 7 et 8 respectivement.

Sur chaque segment d'ARNv, les séquences codantes sont flanquées en 5' et en 3' par des séquences non codantes de longueurs variables en fonction du segment (**figure 6a**). Parmi les virus grippaux de type A, et ce, quelque soit le segment et quelle que soit l'origine du virus, les 13 premiers nucléotides de l'extrémité 5' et les 12 premiers de l'extrémité 3' des ARNv sont extrêmement conservés (Desselberger *et al.*, 1980). Ces séquences sont partiellement complémentaires et l'un des modèles de structure secondaire proposé, est celui de la structure nommée *panhandle*, faisant référence à une poignée de "poêle à frire" (Hsu *et al.*, 1987) (**Figure 6b**). Les extrémités conservées ainsi que les 3 nucléotides adjacents de chaque segment représentent les séquences minimales à la transcription et à la réplication, et constituent le promoteur de l'ARN viral (Fodor *et al.*, 1994 ; Luytjes *et al.*, 1989).



**Figure 6. Structure des ARN des virus grippaux de type A.** a) Représentation schématique d'un ARNv. En vert : la séquence codante. Les séquences non codantes en 5' et 3' de l'ARNv sont indiquées par les barres noires. En rouges : les nucléotides en 5' et 3' conservés, communs aux huit segments. En bleu : les séquences non codantes, non conservées et de longueur variable suivant le segment. b) Modèle de la structure secondaire des extrémités 5' et 3' des ARNv (d'après Hsu *et al.*, 1987). Les extrémités 5' et 3' s'associent pour former une structure nommée *panhandle* (poêle à frire). Encadrées en rouge sont indiquées les séquences communes aux 8 segments d'ARNv, à l'exception de l'uridine marquée par une astérisque, qui est remplacée quelques fois par une cytosine dans les segments 1, 2, 3 et 6.

## 5. Les protéines du virus de la grippe

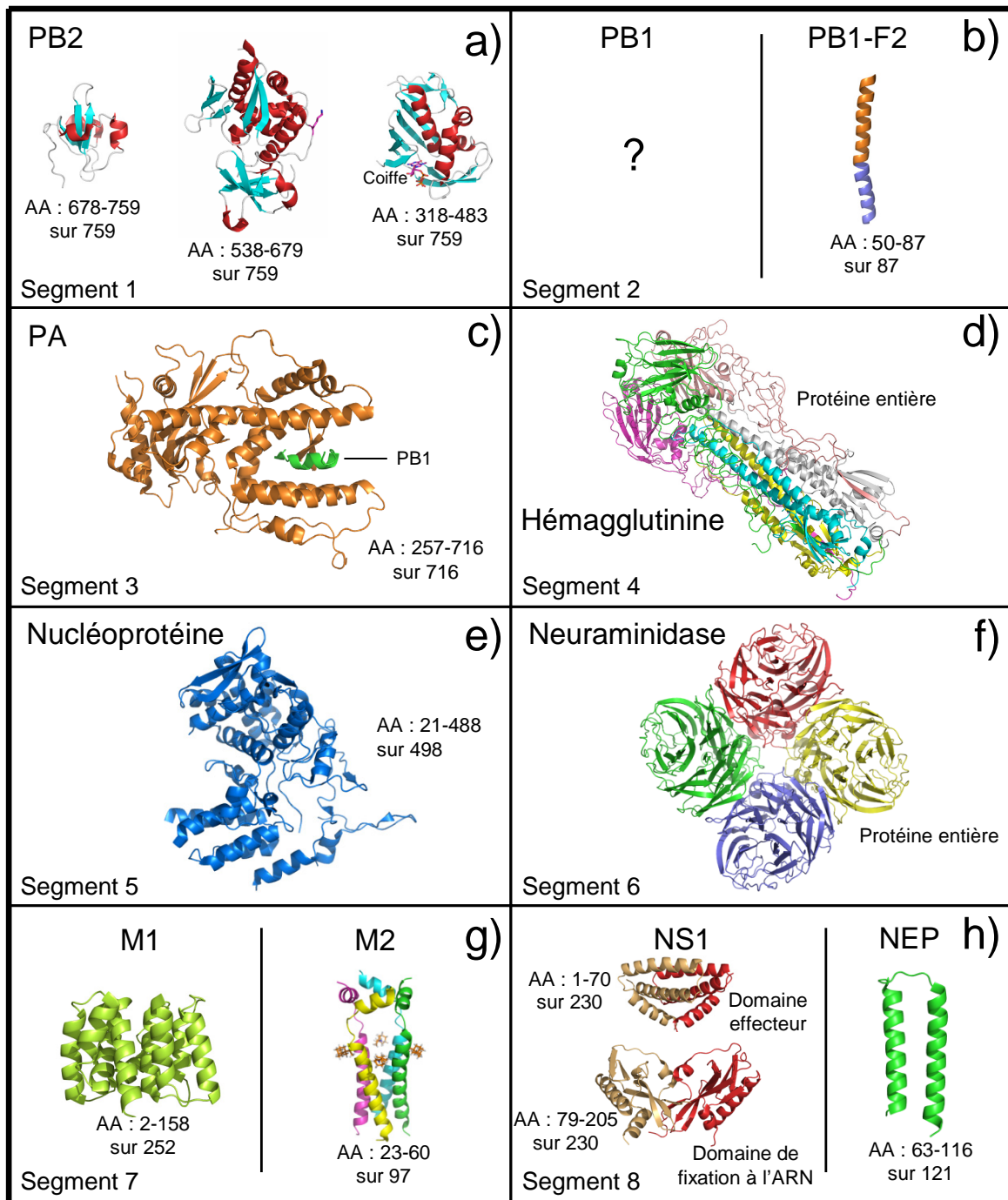
Selon les protéines, traduction et particulièrement maturation s'effectuent à différents endroits. Les protéines de surface : la HA, la NA et le canal ionique M2 (**figure 4**), enchâssées dans l'enveloppe virale provenant de la membrane plasmique cellulaire, sont fabriquées par les ribosomes liés au réticulum endoplasmique. Elles empruntent ensuite l'appareil de Golgi, où seules la NA et la HA seront glycosylées. Ces 3 protéines, qui sont exprimées tardivement durant le cycle viral, seront ensuite insérées dans la membrane plasmique.

Les autres protéines, celles qui sont dites internes ou celles qui ne sont pas présentes dans la structure du virus, sont synthétisées par des ribosomes libres dans le cytoplasme. Ce sont les protéines précoces NP, les trois sous-unités de la polymérase (PB1, PB2, PA), NS1 et les deux protéines tardives M1 et NEP.

**L'Hémagglutinine HA.** Présente à la surface des virions cette protéine est un homotrimer dont le poids moléculaire de chaque sous-unité est de 61,5 kDa (effet de glycosylation non pris en compte). La protéine est codée par le segment 4 du génome viral. Son nom provient du fait que le virus de la grippe peut conduire à l'agglutination des globules rouges. Outre son rôle dans l'entrée du virus dans la cellule hôte, la HA est l'antigène majeur du virus de la grippe. Tout au long de sa synthèse, elle va subir de nombreuses modifications structurales. Fabriquée dans le réticulum endoplasmique, la HA se présente tout d'abord sous la forme d'un précurseur (HA<sub>0</sub>) qui s'associe en trimère. Elle est ensuite clivée en deux sous-unités (HA<sub>1</sub> et HA<sub>2</sub>) à l'extrémité N-terminale de sa région appelée "peptide de fusion", ces deux domaines restant liés par un pont disulfure (Skehel & Waterfield, 1975). Lors de l'adoption de la conformation native de HA, il se forme d'autres ponts disulfures, de nombreuses glycosylations (Ward *et al.*, 1980 ; Naim & Roth, 1993) et une trimérisation de la protéine. C'est cette structure clivée, trimérique et contenant des modifications post traductionnelles qui est présente à la surface des virus infectieux.

La structure cristallographique de l'HA est connue depuis plus de 25 ans désormais (Wilson *et al.*, 1981). Depuis, la connaissance de sa structure dans différentes conformations (obtenues à différents pH) a permis de mieux comprendre certains points du fonctionnement du virus de la grippe. Notamment, les mécanismes d'interaction avec le récepteur cellulaire et l'entrée du virus dans la cellule grâce aux changements conformationnels de HA (**figure 7d**).





**Figure 7. Structures 3D des protéines des virus grippaux de type A. En raison du nombre de structures à ce jour résolues, seul un exemplaire a été choisi pour chaque protéine. a)** Structures cristallographiques des différents domaines de la sous-unité PB2 de la polymérase à 2 Å de résolution (Tarendeau *et al.*, 2007 et 2008 ; Guilligay *et al.*, 2008). **b)** Structure RMN de la protéine PB1-F2 (Bruns *et al.*, 2007), la structure de PB1 n'est pas encore disponible. **c)** Structure cristallographique de la sous-unité PA de la polymérase à 2,3 Å de résolution (Obayashi *et al.*, 2008). **d)** Structure cristallographique du trimère de l'hémagglutinine (HA) à 3 Å de résolution (Weis *et al.*, 1990). **e)** Structure cristallographique de la nucléoprotéine (NP) à 3,2 Å de résolution (Ye *et al.*, 2006). **f)** Structure cristallographique du tétramère de la neuraminidase (NA) à 2,2 Å de résolution (Varghese & Colman 1991). **g)** (*A gauche*) Structure cristallographique de la protéine matricielle M1 résolue à 2,3 Å (Arzt *et al.*, 2001) et (*à droite*) structure du canal ionique M2 complexé à la rimantadine obtenue par RMN (Schnell & Chou, 2008). **h)** (*A gauche*) Structures cristallographiques des dimères du domaine de fixation à l'ARN à 1,9 Å de résolution (Liu *et al.*, 1997) et du domaine effecteur à 2,1 Å (Bornholdt & Prasad, 2006) de la protéine non structurale (NS1) et (*à droite*), structure cristallographique du domaine C-terminal de la protéine d'export nucléaire (NEP) à 2,6 Å de résolution (Akarsu *et al.*, 2003).

**La neuraminidase NA.** Codée par le segment 6 du génome viral, cette protéine présente un poids moléculaire de 50 kDa (effet de glycosylation non pris en compte). Elle est également présente à la surface du virion sous la forme d'un tétramère (**figure 7f**). Sa structure cristallographique est connue depuis longtemps, aussi bien seule (Varghese *et al.*, 1983) qu'en complexe avec de l'acide sialique (Burmeister *et al.*, 1992) ou un analogue de l'acide sialique (White *et al.*, 1995). Comme la HA, elle est exprimée en phase tardive du cycle viral, au niveau du réticulum endoplasmique où elle subit des glycosylations, et est aussi un déterminant antigénique. La NA est impliquée dans la phase finale du cycle viral, lors du relargage des virions. En clivant les acides sialiques, elle évite que les virions ne restent collés à la surface cellulaire à cause des interactions HA/acide sialique. Cette protéine est donc impliquée dans la propagation de l'infection virale et est la cible d'analogues d'acides sialiques couramment utilisés contre le virus de la grippe.

**Le canal à proton M2.** C'est l'une des deux protéines codées par le segment 7 grâce à un épissage alternatif du transcrit viral (Lamb *et al.*, 1981). Elle présente un poids moléculaire de 11 kDa et forme un homotétramère. Une première partie de sa structure avait été résolue par RMN (Résonance Magnétique Nucléaire) quelques années auparavant (Nishimura *et al.*, 2002). Finalement la structure du tétramère a été entièrement résolue par RMN récemment (Schnell & Chou, 2008). La protéine M2 est une protéine canal à protons jouant un rôle primordial lors de l'acidification de l'intérieur du virus en phase précoce d'infection (**figure 7g**).

**La protéine matricielle M1.** Dès les premières études - en microscopie électronique notamment (Schulze, 1972) - portant sur la structure du virus de la grippe, la protéine M1 de 28 kDa apparaît comme un élément abondant de la particule virale. C'est la deuxième protéine, avec M2, codée par le segment 7 du génome (Lamb *et al.*, 1981). Les nombreuses expériences menées sur M1 montrent qu'elle possède plusieurs fonctions essentielles au virus durant les différentes étapes du cycle d'infection. Tout d'abord, M1 est une protéine structurale qui participe à l'architecture du virion. Elle tapisse toute la face interne de l'enveloppe virale (d'où son nom de protéine matricielle). Elle interagit d'une part avec la bicouche phospholipidique (Gregoriades, 1980 ; Ruigrok *et al.*, 2000 ; Baudin *et al.*, 2001) et d'autre part avec les segments génomiques viraux (Ye *et al.*, 1987, 1999 ; Watanabe *et al.*, 1996). Plus précisément, il a été démontré que M1 interagissait - sans être enchâssée - avec la membrane lipidique par des interactions électrostatiques (Ruigrok *et al.*, 2000), via son

domaine N-terminal (Baudin *et al.*, 2001). Quant au domaine C-terminal de M1, il semble être impliqué dans la liaison aux RNP (Ye *et al.*, 1987 ; Baudin *et al.*, 2001). M1 participerait de cette façon à la compaction des RNP dans le virion et à l'architecture de celui-ci.

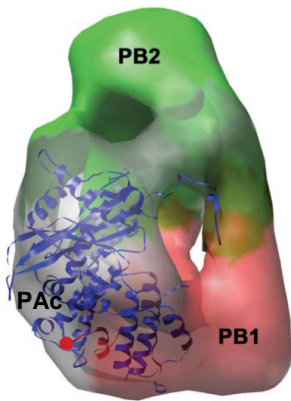
Exprimée en phase tardive du cycle viral, cette protéine joue également un rôle très important dans l'export nucléaire des RNP nouvellement synthétisées. Bien qu'ayant un poids moléculaire de 28 kDa, et donc pouvant traverser par simple diffusion passive, M1 possède un signal de localisation nucléaire (NLS) (Ye *et al.*, 1995), elle est ainsi détectée à la fois dans le cytoplasme et le noyau.

La structure cristallographique du domaine N-terminal de M1 a pu être résolue à deux reprises (**figure 7g**). Une première fois grâce une M1 purifiée à partir du virus et dont les cristaux obtenus ont permis de résoudre la structure à 2 Å de résolution (Sha & Luo, 1997). Une seconde fois à partir d'une M1 exprimée dans *E.coli* dont la structure a été résolue à 2.3 Å (Arzt *et al.*, 2001).

**La nucléoprotéine NP.** C'est une protéine de 56 kDa (498 acides aminés) codée par le segment 5. Cette protéine, riche en acides aminés basiques, présente une charge nette positive à pH neutre (point isoélectrique [pI] théorique de 9,3) et recouvre le génome. Ainsi, elle possède des domaines spécifiques pour l'interaction avec l'ARN (Elton *et al.*, 1999b) (**figure 5**). En effet, la NP basique, se lie *in vitro* à l'ARN simple brin avec une forte affinité mais sans spécificité de séquence (Baudin *et al.*, 1994). Cette liaison de la NP avec l'ARN se fait sur le squelette phosphate de l'ARN, les bases restant ainsi exposées vers le solvant (Baudin *et al.*, 1994 ; Klumpp *et al.*, 1997). De plus elle présente des domaines d'interaction NP-NP et s'oligomérisent spontanément en solution (Kingsbury & Webster, 1969 ; Pons *et al.*, 1969 ; Ruigrok & Baudin, 1995). La nucléoprotéine se fixe également à la sous-unité PB2 de la polymérase virale ainsi qu'à la protéine matricielle M1. Il a également été montré que la NP pouvait interagir avec des partenaires cellulaires appartenant à la famille des protéines d'import nucléaire, les importines. Dans cette optique, de nombreuses études menées sur les signaux de transport nucléaire (NLS) de la nucléoprotéine ont mis en évidence la présence de plusieurs de ces séquences sur la protéine (Wang *et al.*, 1997 ; Neumann *et al.*, 1997). Il semble donc probable que les fonctions de la NP dans le cycle viral, soient orchestrées par l'ensemble des interactions dans lesquelles la protéine est impliquée. La NP est depuis longtemps connue pour être le composant majeur des RNPv et sa fonction est d'être impliquée dans les processus de transcription et de réplication du génome (Huang *et al.*, 1990 ; Kimura *et al.*, 1992 ; de la Luna *et al.*, 1993). Très récemment, la NP (souche H1N1) a été cristallisée

sous forme d'un trimère dont la structure a été résolue à 3,2 Å de résolution (Ye *et al.*, 2006) (**figure 7e**) puis celle du virus H5N1 à également été résolue (Ng *et al.*, 2008). La structure cristallographique de la protéine a révélé la présence d'un sillon dans lequel se fixerait probablement l'ARN. La surface de ce sillon est largement composé d'acides aminés basiques, de type Lys et Arg, qui devraient intervenir en s'associant à l'ARN par interaction avec le squelette phosphodiester.

**Le complexe polymérase de la grippe.** La polymérase virale est composée de trois sous-unités : PB1 (86,5 kDa), PB2 (85,7 kDa) qui sont 2 protéines basiques et PA (84,2 kDa) qui est une protéine acide. Il a été montré que PB1 interagit avec PB2 et PA (Gonzalez *et al.*, 1996). Les trois sous-unités sont nécessaires à un cycle complet de transcription mais également de réplication durant lequel, les RNPv dont la polarité est négative doivent, pour être dupliquées, passer par un intermédiaire répliatif. Cet intermédiaire, nommé RNPc (RNP complémentaire), contient l'ARN de polarité positive (+). Cependant, chaque sous-unité possède une fonction propre. Brièvement, la sous-unité PB1 est codée par le segment 2, elle possède l'activité de polymérase proprement dite et une activité d'endonucléase. Cette seconde fonction lui permet de cliver les coiffes situées en 5' des ARNm cellulaires (Li *et al.*, 2001). PB2, codée par le segment 1, reconnaît les coiffes 5' m<sup>7</sup>Gppp (Blaas *et al.*, 1982 ; Shi *et al.*, 1996) et permet ainsi leur clivage par PB1 (Li *et al.*, 2001). La fonction de PA, codée par le segment 3, reste encore obscure. Elle semble, à la suite de certains travaux, jouer un rôle central à la fois dans la transcription et la réplication (Fodor *et al.*, 2002 ; Huarte *et al.*, 2003). Notamment, la région N-terminale de PA participerait à la stabilité de la protéine, et serait impliquée dans l'activité endonucléase ainsi que dans la fixation de la coiffe et du promoteur sur l'ARNv (Hara *et al.*, 2006). La NP et les trois sous-unités de la polymérase représentent les quatre protéines virales indispensables aux étapes de transcription et de réplication du génome dans le noyau cellulaire. Deux reconstructions sur la polymérase, effectuées en microscopie électronique à transmission (MET), ont été publiées par le même groupe. Tout d'abord, un modèle du complexe polymérase associé dans une RNPv recombinante a été obtenu à 23 Å de résolution (Area *et al.*, 2004). Puis, un modèle du trimère de la polymérase seule, a été déterminé à 26 Å de résolution (Torreira *et al.*, 2007) (**figure 8**). Ce deuxième modèle présente une structure semblable à celle obtenue précédemment mais laisse apparaître une conformation davantage ouverte, du fait de l'absence de la molécule d'ARN. Courant 2008, d'autres domaines de la polymérase ont également été résolus (**figure 7a et c**).



**Figure 8. Modèle 3D du complexe polymérase isolé.** Insertion de la sous-unité PA (partie C-terminale) (He *et al.*, 2008) dans le modèle 3D de la polymérase obtenue par Torreira *et al.*, 2007.

**La protéine non structurale NS1.** NS1 est l'une des deux protéines, avec NEP, codées par le plus petit segment du génome viral, le segment 8 (Lamb & Chopin, 1979 ; Inglis *et al.*, 1979). Elle est dénommée ainsi car elle est exprimée dans les cellules infectées mais n'est pas présente au sein de la particule virale. NS1 est constitué de 230 acides aminés, correspondant à un poids moléculaire de 27 kDa. Elle dispose d'un domaine de liaison à l'ARN localisé entre les acides aminés 19 et 38 ainsi que d'un domaine effecteur localisé entre les acides aminés 73 et 230. La protéine est retrouvée sous forme dimérique *in vivo* et *in vitro* (Wang *et al.*, 1999) et possède de nombreuses fonctions jouant un rôle capital au cours du cycle viral. Ainsi, l'une de ses fonctions majeures est d'être impliquée dans l'inhibition de la réponse antivirale. En effet, la présence de NS1 durant le cycle viral est essentiel pour inhiber la réponse aux interférons (IFN- $\alpha/\beta$ ) (Garcia-Sastre *et al.*, 1998 ; Kochs *et al.*, 2007). Il a également été montré que NS1 inhibe la maturation des ARNm cellulaires en se fixant à la plus petite sous unité (30 kDa) du complexe "Cleavage and polyadenylation specificity factor" (CPSF) (Nemeroff *et al.*, 1998), ainsi qu'à la poly (A) binding protein II (PABP II) (Chen *et al.*, 1999). Le contrôle qualité des ARNm mis en place au niveau du noyau entraîne la rétention nucléaire de ces pré-ARNm coiffés en 5' mais non polyadénylés. La protéine NS1 serait également impliquée dans l'augmentation de la traduction des ARNm viraux sans pour autant influencer celle des ARNm cellulaires (Enami *et al.*, 1994 ; de la Luna *et al.*, 1995) et de façon moins claire, dans la réponse à l'apoptose cellulaire (revue, Hale *et al.*, 2008).

La structure tridimensionnelle du domaine de liaison à l'ARN (partie N-terminale), a été résolue aux rayons X (Liu *et al.*, 1997) (figure 7h) et par RMN (Chien *et al.*, 1997). Et, très récemment la structure cristallographique aux rayons X du domaine effecteur de NS1 (partie C-terminale) a été résolue à 2,1 Å (figure 7h) par Bornholdt & Prasad (2006).

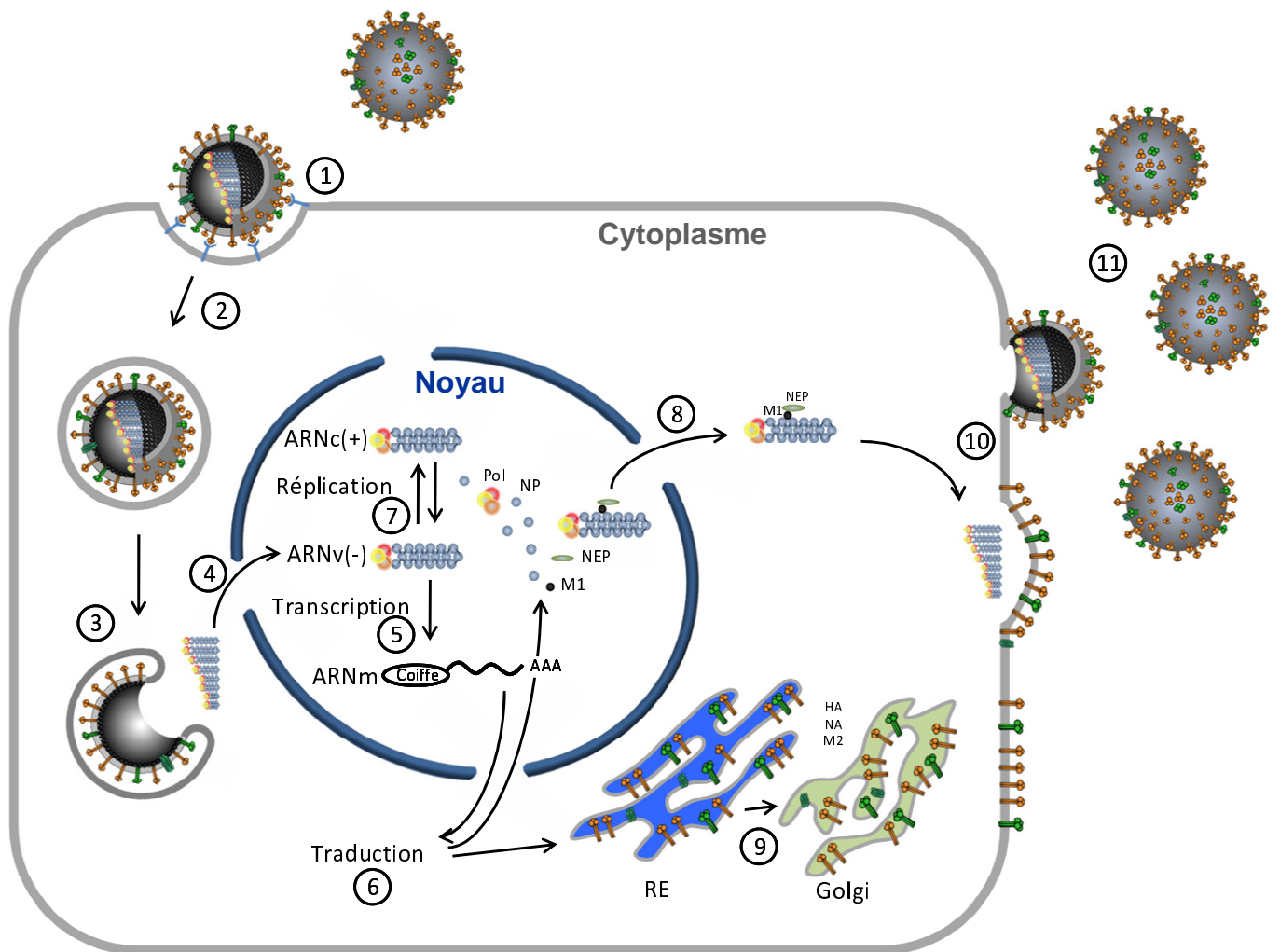
**La protéine d'export nucléaire NEP.** Cette protéine virale, codée par le segment 8, qui portait à l'origine le nom de protéine non structurale 2 (NS2), est en fait présente en plusieurs exemplaires (100-230) dans la particule virale (Richardson & Akkina, 1991 ; Yasuda *et al.*, 1993). C'est une petite protéine de 14 kDa. Comme M1, NEP est une protéine tardive qui n'a pas de localisation intracellulaire précise, elle est détectée dans le noyau et le cytoplasme. Le domaine C-terminal (acides aminés 59 à 116) a été cristallisé et sa structure résolue à 2,6 Å (Akarsu *et al.*, 2003) (**figure 7h**). La protéine NEP joue un rôle fondamental dans l'export nucléaire des RNP, d'où l'origine de son nom. Grâce à son motif d'export nucléaire (NES), elle se lie à l'exportine cellulaire CRM1 (*Chromosome Region Maintenance protein 1*). La protéine NEP interagit avec M1, et toutes deux, jouent un rôle fondamental dans l'export nucléaire des RNP virales (O'Neill *et al.*, 1998 ; Neumann *et al.*, 2000 ; Akarsu *et al.*, 2003).

**PB1-F2.** C'est la protéine du virus de la grippe la plus récemment découverte (Chen *et al.*, 2001). Sa traduction est initiée par un second cadre de lecture présent sur le gène codant pour PB1, donnant lieu à une protéine de 10,5 kDa. Localisée au niveau des membranes internes et externes de la mitochondrie, PB1-F2 semble être impliquée dans l'apoptose (Gibbs *et al.*, 2003). Récemment, sa région C-terminale a été résolue en RMN et montre une structure alpha hélicale (**figure 7b**) (Bruns *et al.*, 2007).

## 6. Le cycle de multiplication de virus de la grippe (Figure 9)

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

Le virus de la grippe a pour cible les cellules épithéliales du système respiratoire. Le cycle d'infection débute par la liaison spécifique de la particule virale à la cellule hôte. La reconnaissance se fait via la HA qui se lie aux récepteurs, les acides sialiques, fixés sur les glycoprotéines ou glycolipides cellulaires. La faible interaction (de l'ordre de  $10^{-3}$  M) est compensée par la densité de récepteurs cellulaires. Une fois accroché à la cellule, le virus pénètre dans la cellule hôte par endocytose à l'aide ou indépendamment des vésicules à clathrine (revue, Marsh & Helenius, 2006). Le virus se retrouve alors dans un endosome. Sous l'action des pompes à protons cellulaire, l'intérieur de ce compartiment devient acide. Cette diminution du pH va induire un changement de conformation de la HA, ce qui a pour effet de fusionner les membranes virale et endosomale.



**Figure 9. Schéma du cycle d'infection des virus grippaux de type A.** 1) Le virus s'attache à la cellule par fixation de l'hémagglutinine virale sur ses récepteurs cellulaires. 2) Entrée du virus par endocytose. 3) Suite à l'acidification de l'endosome et de l'intérieur du virus, les membranes du virus et de l'endosome fusionnent. 4) Transport des RNPv vers le noyau. 5) Transcription des ARNm viraux. 6) Synthèse des protéines virales. 7) Réplication du génome viral. 8) Export des RNPv nouvellement synthétisés vers le cytoplasme. 9) Insertion des protéines de l'enveloppe dans la membrane plasmique. 10) Assemblage et bourgeonnement des particules virales. 11) Libération des virions suite à l'action de la neuraminidase.

Par ailleurs, cette acidification dans l'endosome a pour effet d'activer le canal à proton, la protéine M2 (Holsinger *et al.*, 1994 ; Pinto *et al.*, 1992). L'intérieur de la particule virale s'acidifie à son tour, ce qui va dissocier les RNPv de la couche de protéines matricielle M1. La combinaison de ces deux événements, permet ainsi la libération des 8 segments génomiques dans le cytoplasme.

## 6.2. Migration des particules ribonucléoprotéiques virales (RNPv) dans le noyau cellulaire

Libérés dans le cytoplasme de la cellule hôte, les 8 RNPv migrent ensuite vers le noyau, siège de la transcription et de la réplication. Pour cela, l'ensemble des segments viraux va emprunter le pore nucléaire qui est l'unique voie de passage entre le noyau et le cytoplasme. Le transport vers le noyau fait intervenir des signaux d'adressages nucléaires (NLS), présents sur les RNPv (potentiellement la polymérase ou la NP) et reconnus par des protéines de transport cellulaire, les importines. Le complexe d'import ainsi formé s'associe au pore nucléaire et le recrutement du facteur Ran-GDP apporte l'énergie nécessaire au transport des RNPv dans le noyau. Ce point sera davantage développé au cours de la partie B.

## 6.3. La transcription et la réplication du génome viral

A l'intérieur du noyau, les trois sous-unités de la polymérase, la NP et l'ARNv sont les molécules nécessaires aux étapes de transcription et de réplication. Durant ces deux processus, trois types d'ARN sont synthétisés : les ARN messagers (ARNm) viraux, les ARN génomiques (ARNv) de polarité négative et les ARN génomiques complémentaires (ARNc) de polarité positive.

Dans un premier temps, sont transcrits les ARNm viraux précoces de PB1, PB2, PA, NP et NS1 (Hatada *et al.*, 1989). Ceci permet notamment la synthèse de nouveaux complexes polymérase. Il s'en suit la réplication des ARNv et la synthèse des ARNm viraux tardifs (HA, NA, M(1/2) et NEP).

**La transcription des ARNm viraux.** A l'extrémité 5' des ARNm viraux se trouve une coiffe ( $m^7Gppp$ ) ainsi que 10 à 13 nucléotides provenant d'ARNm cellulaires. A leur extrémité 3', ils sont polyadénylés grâce à une séquence riche en uridines située proche de l'extrémité 5' de l'ARNv matrice (Robertson *et al.*, 1981 ; Luo *et al.*, 1991).

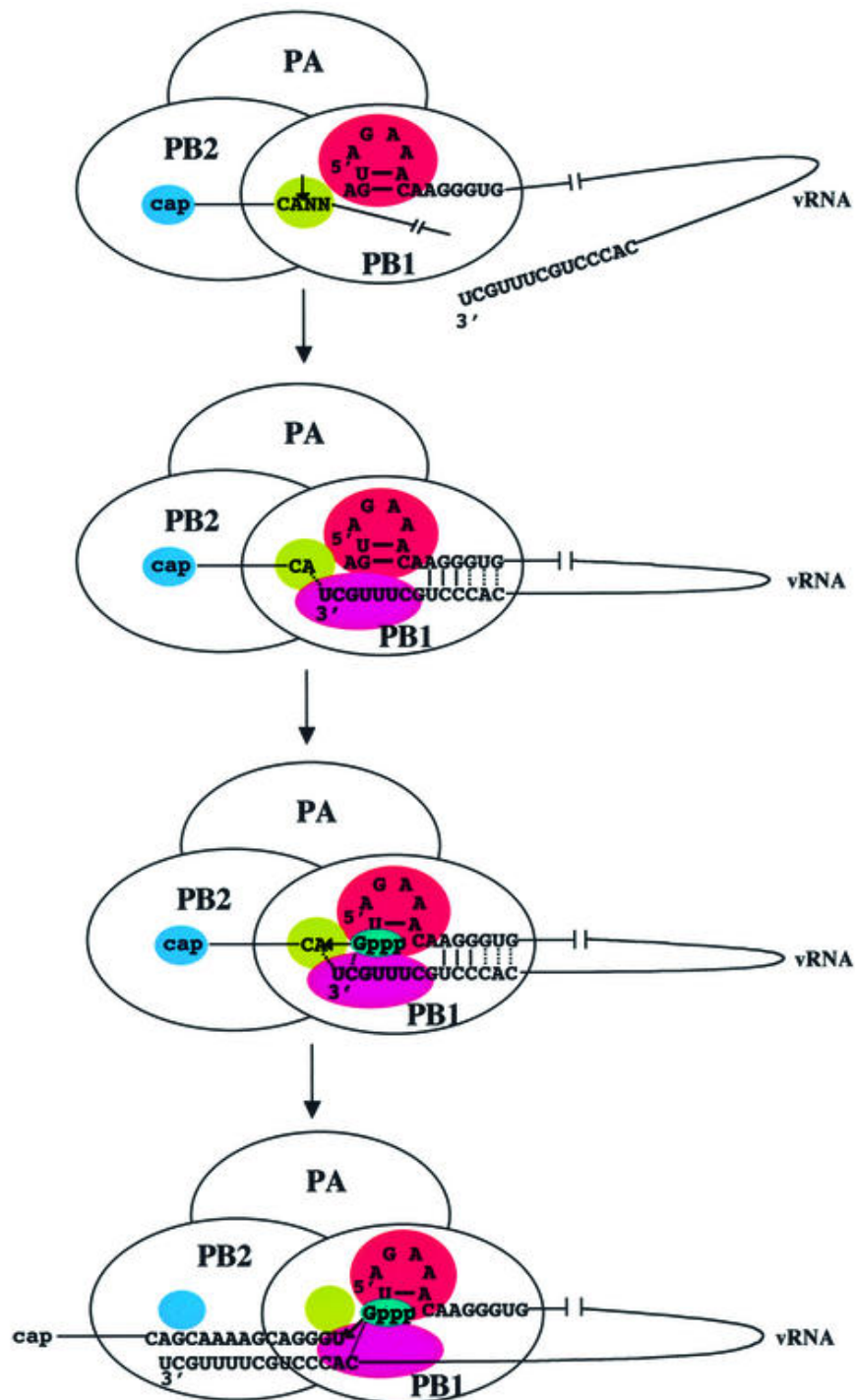
Ainsi, une particularité dans la transcription des ARNm du virus de la grippe, est qu'elle est initiée par un mécanisme nommé : "capture de coiffe" des ARNm cellulaires. Ce



processus a été initialement décrit (Bouloy *et al.*, 1978) puis les mécanismes progressivement élucidés (Krug, 1981 ; Plotch *et al.*, 1979 et 1981 ; Rao *et al.*, 2003) (**Figure 10**). L'interaction de PB1 avec le promoteur situé à l'extrémité 5' de l'ARNv induit un changement de conformation du complexe polymérase. Ceci active le site de capture de coiffe porté par PB2 ainsi que la fonction endonucléase de la sous-unité PB1. Les extrémités coiffées des ARN pré-messagers cellulaires se retrouvent liés à la protéine PB2 puis clivés par la protéine PB1 à une distance de 10 à 13 nucléotides de leur extrémité 5'. La coupure se fait au niveau des nucléotides CA. Cette séquence serait reconnue de façon stable par PB1 et aurait la capacité d'être associée à l'extrémité 3' de l'ARN viral (Rao *et al.*, 2003). La base A située à l'extrémité clivée de l'ARN pré-messager coiffé peut ainsi interagir avec la base U terminale de l'extrémité 3' de l'ARNv. Du fait de ces positionnements, les fragments d'ARN cellulaires coiffés serviront d'amorce à la synthèse des ARNm viraux. L'élongation est ensuite catalysée par la protéine PB1 qui porte l'activité d'ARN polymérase ARN dépendante, permettant ainsi la synthèse d'un ARNm viral qui n'est pas recouvert par la NP.

**La réplication des ARNv.** Ce processus débute par la production d'un intermédiaire, un ARN de polarité positive strictement complémentaire (ARNc) à la matrice d'ARNv. Les molécules d'ARNc serviront à leur tour de matrice pour la synthèse de nouvelles molécules d'ARNv. Le mécanisme de synthèse des ARN viraux et complémentaires est clairement différent de celui de la synthèse de l'ARNm puisque son initiation ne nécessite pas d'amorce. Tout comme l'ARNv, l'ARNc est recouvert par la NP et est associé au complexe polymérase pour former une RNPC.

Le passage de la transcription à la réplication reste inexplicé, cependant il semblerait que la NP dont la concentration augmente au cours du cycle viral, joue un rôle régulateur (revue, Portela & Digard, 2002). Ainsi, par le passé deux hypothèses ont été proposées pour expliquer le lien entre l'augmentation de la quantité de NP au cours du cycle et le passage de la transcription à la réplication. Selon la première hypothèse, la quantité croissante de NP induirait une interaction de la protéine avec la polymérase, modifiant ainsi la conformation du complexe (PA, PB1, PB2) (Biswas *et al.*, 1998 ; Mena *et al.*, 1999). Dans la deuxième, l'augmentation de la concentration en NP entraînerait une compétition entre NP et ARN polymérase pour la fixation du brin 5' de l'ARNv (Klumpp *et al.*, 1997). Dans un cas comme dans l'autre, ceci aurait pour conséquence d'altérer la synthèse d'ARNm viral au profit de l'initiation de la synthèse d'un ARN c ou v non coiffé.



**Figure 10. Modèle du mécanisme de capture de coiffe (d'après Rao *et al.*, 2003).** La fixation de l'extrémité 5' de l'ARNv sur la protéine PB1 (rouge) active le site de fixation de la coiffe (bleu) sur la protéine PB2 et initie l'activité endonucléase de PB1 (vert). L'extrémité coiffée de l'ARNm cellulaire est clivée au niveau d'une séquence CA. L'extrémité 3' de l'ARNv se fixe à la protéine PB1 (rose) et à l'extrémité 5' de l'ARNv. La base A située à l'extrémité clivée de l'ARNm coiffé interagit par complémentarité avec la base U terminale en 3' de l'ARNv et permettra à l'ARN cellulaire coiffé de servir d'amorce pour la synthèse d'ARN messager viral.

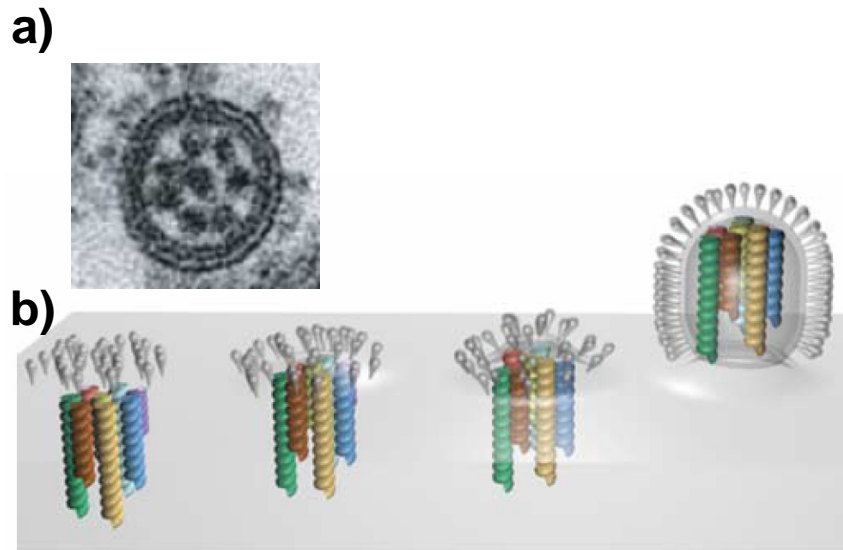
Cependant, ces modèles sont basés sur le fait que l'ARNc n'est pas synthétisé en début d'infection. Plus récemment, [Vreede et al., \(2004\)](#) ont observé simultanément la présence d'ARNm et d'ARNc viraux dès le début de l'infection, suggérant l'absence de "bascullement" entre transcription et réplication. Ainsi, un nouveau modèle a été proposé, dans lequel les ARNm viraux sont protégés par leur coiffe et leur queue poly A alors que les ARNc seraient dégradés en début de cycle pour être ensuite stabilisés par l'expression de la NP. Ceci permettrait ainsi la transition vers la réplication.

#### 6.4. L'export nucléaire des RNPv

Les ARNv nouvellement produits, recouverts par la NP et associés au complexe polymérase, sont exportés dans le cytoplasme vers la membrane plasmique, lieu de bourgeonnement des particules virales ([revue, Cros & Palese, 2003](#)). A l'intérieur du noyau, la protéine M1 fait le lien entre les RNPv néo-synthétisées ([Whittaker et al., 1996](#)) et la protéine d'export nucléaire NEP ([Yasuda et al., 1993](#) ; [Ward et al., 1995](#)). La sortie des RNPv se fait alors au travers du pore nucléaire, via le système d'export spécifique CRM1 (Chromatin Maintenance Protein 1) qui interagit avec la protéine NEP ([Elton et al., 2001](#) ; [Ma et al., 2001](#) ; [Watanabe et al., 2001](#) ; [Akarsu et al., 2003](#)). Les RNPv sont ensuite transportées vers la membrane cellulaire sur laquelle sont insérées les protéines HA, NA et M2 qui ont transité via le réticulum endoplasmique et l'appareil de Golgi. Ce point sera repris au cours de la partie B.

#### 6.5. L'emballage des segments dans la particule virale

Le mécanisme d'incorporation des 8 segments d'ARNv dans la particule virale n'est pas encore totalement compris. Alors que des travaux anciens proposaient un emballage aléatoire des huit segments ([Enami et al., 1991](#) ; [Bancroft & Parlsow, 2002](#)), de récents travaux indiquent la présence de séquences spécifiques sur chaque segment d'ARNv qui pourrait mener à l'incorporation d'un jeu de 8 segments par virion ([Fujii et al., 2003 et 2005](#) ; [Liang et al., 2005](#)). L'analyse de coupes de virus en MET, a montré que les RNPv sont organisées de façon caractéristique dans la particule virale en bourgeonnement ([Noda et al., 2006](#)). On peut ainsi observer sept segments de différentes longueurs qui entourent un huitième segment central (**Figure 11**). Il reste encore à élucider l'arrangement caractéristique entre les différents segments au sein de la particule virale. Une possibilité serait que l'ARNv contienne une séquence favorisant un recrutement des segments durant l'emballage, et qui leur servirait également de lien ([Noda et al., 2006](#)).



**Figure 11. Organisation des RNPv dans la particule virale en bourgeonnement (d'après Noda *et al.*, 2006).** a) Coupe transversale de virion en bourgeonnement observée en microscopie électronique à transmission. b) Modèle schématique de l'organisation des RNPv lors du bourgeonnement. On distingue 7 segments entourant un huitième.

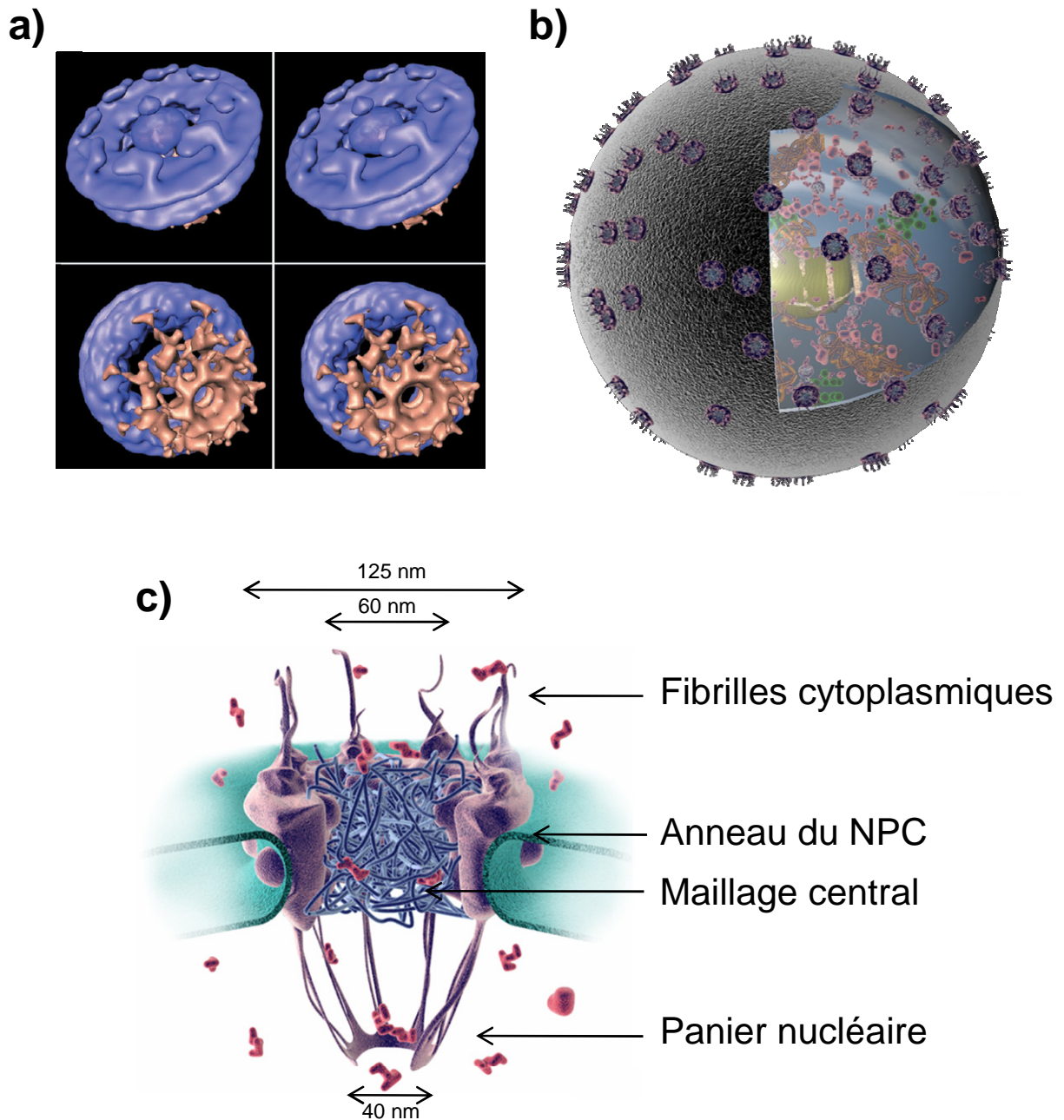
#### 6.6. Bourgeonnement et relargage des particules virales

La protéine M1 est essentiel pour connecter les queues cytoplasmiques des glycoprotéines NA et HA aux RNPv ainsi que pour le bourgeonnement des particules virales (revue, Nayak *et al.*, 2004). Le bourgeonnement des particules virales nouvellement formées dépend principalement de la neuraminidase. En effet, cette glycoprotéine de surface possède une activité de sialidase. Elle clive ainsi la liaison entre l'acide sialique et l'hémagglutinine détruisant ainsi le récepteur cellulaire de la HA. Sans cette étape, les virions nouvellement formés se recolleraient immédiatement aux récepteurs via l'hémagglutinine, ce qui empêcherait le virus de se propager. Cette fonction cruciale de la NA dans le cycle d'infection du virus de la grippe est la cible des antiviraux les plus utilisés actuellement dans le traitement des gripes A et B. C'est suite à la détermination de la structure cristalline de la NA (Varghese & Colman, 1991) que ces analogues (zanamivir et l'oseltamivir) de l'acide sialique ont été mis au point.

## **Partie B**

—

## **Le transport nucléocytoplasmique**



**Figure 12. Architecture du pore nucléaire (NPC).** **a)** Images en stéréo du NPC de *Dictyostelium discoideum* obtenues par tomographie. Les filaments cytoplasmiques, recourbés, sont arrangés autour du canal central. L'anneau distal du panier (représenté en rose) est connecté à l'anneau nucléaire par les filaments nucléaires (Beck *et al.*, 2004). **b)** Pores nucléaires incorporés dans l'enveloppe du noyau cellulaire. **c)** Représentation du NPC et de ses composants. La structure a une symétrie apparente d'axe 8 perpendiculaire au plan de la membrane nucléaire (Patel *et al.*, 2007). Images **b** et **c** : Samir S. Patel.

## Partie B - Le transport nucléocytoplasmique

### 1. Introduction

A l'intérieur des cellules eucaryotes, la réplication de l'ADN et la transcription de l'ARN sont séparées spatialement de la synthèse de protéines par la double membrane du pore nucléaire. En conséquence, des centaines de macromolécules sont transférées entre le noyau et le cytoplasme. Les ARNm transcrits sont exportés du noyau pour rejoindre les ribosomes. A l'opposé, les protéines telles que les histones, les polymérase et les facteurs de transcription sont importées dans le noyau pour atteindre le matériel génétique. Le transport nucléocytoplasmique se fait à travers le pore nucléaire.

#### 1.1. Le pore nucléaire

Inscrit dans la double membrane de l'enveloppe nucléaire (**figure 12b**), le pore nucléaire représente l'unique voie de passage entre le noyau et le cytoplasme et permet un transport bidirectionnel (Feldherr *et al.*, 1984 ; Nachury & Weis, 1999). Il est imperméable à la plupart des macromolécules à l'exception notable des facteurs de transport nucléaire. Ces protéines peuvent ainsi se déplacer à travers le pore nucléaire en emportant avec elles d'autres macromolécules. Il existe plusieurs centaines de pores nucléaires par noyau et leur nombre varie suivant le type et l'état cellulaires. Chaque pore est en réalité un immense complexe protéique (NPC pour *nuclear pore complex*) d'environ 125 MDa chez les mammifères. Il est composé d'une trentaine de protéines présentes en plusieurs exemplaires, appelées nucléoporines (Nups) (Schwartz, 2005). L'organisation du NPC a été affinée tout au long des années grâce aux progrès de la microscopie électronique et plusieurs reconstructions en 3 dimensions (3D) sont aujourd'hui disponibles (Hinshaw *et al.*, 1992 ; Akey & Radermacher, 1993 ; Yang *et al.*, 1998 ; Stoffler *et al.*, 2003 ; Beck *et al.*, 2004 et 2007 ; Alber *et al.*, 2007). Le pore nucléaire apparaît comme une structure cylindrique avec une symétrie d'ordre 8 (**figure 12a**). Il est constitué d'un canal central de 10 nm de diamètre au repos, relié par 8 bras à 2 anneaux, l'un cytoplasmique et l'autre nucléaire. Des fibrilles décorent ces anneaux, formant une structure dite en "panier de basket" du côté nucléaire (**figure 12c**). Les ions et les molécules de poids moléculaire inférieur à 40 kDa peuvent diffuser librement au travers de ce pore nucléaire. En revanche, le transport de toute molécule de poids moléculaire supérieur à

40 kDa doit s'opérer de façon active à travers le canal central dont le diamètre peut s'élargir jusqu'à près de 40 nm (Panté & Kann, 2002 ; Alber *et al.*, 2007).

Les études portant sur la structure du NPC montrent qu'il s'agit d'une structure dynamique où les nucléoporines (Nups) peuvent se lier au cargo et l'escorter durant le transport (Beck *et al.*, 2004 ; Rabut *et al.*, 2004). Parmi les Nups, certaines d'entre elles disposent de nombreux motifs répétitifs d'acides aminés phénylalanine (F) et glycine (G) (Otsuka *et al.*, 2008), nommés motifs FG, reconnus par des facteurs de transport. Ces domaines FG ne sont pas structurés et sont composés d'enroulements aléatoires (Denning *et al.*, 2003). Ils forment ainsi un réseau de filaments qui empêchent le passage aux molécules intruses.

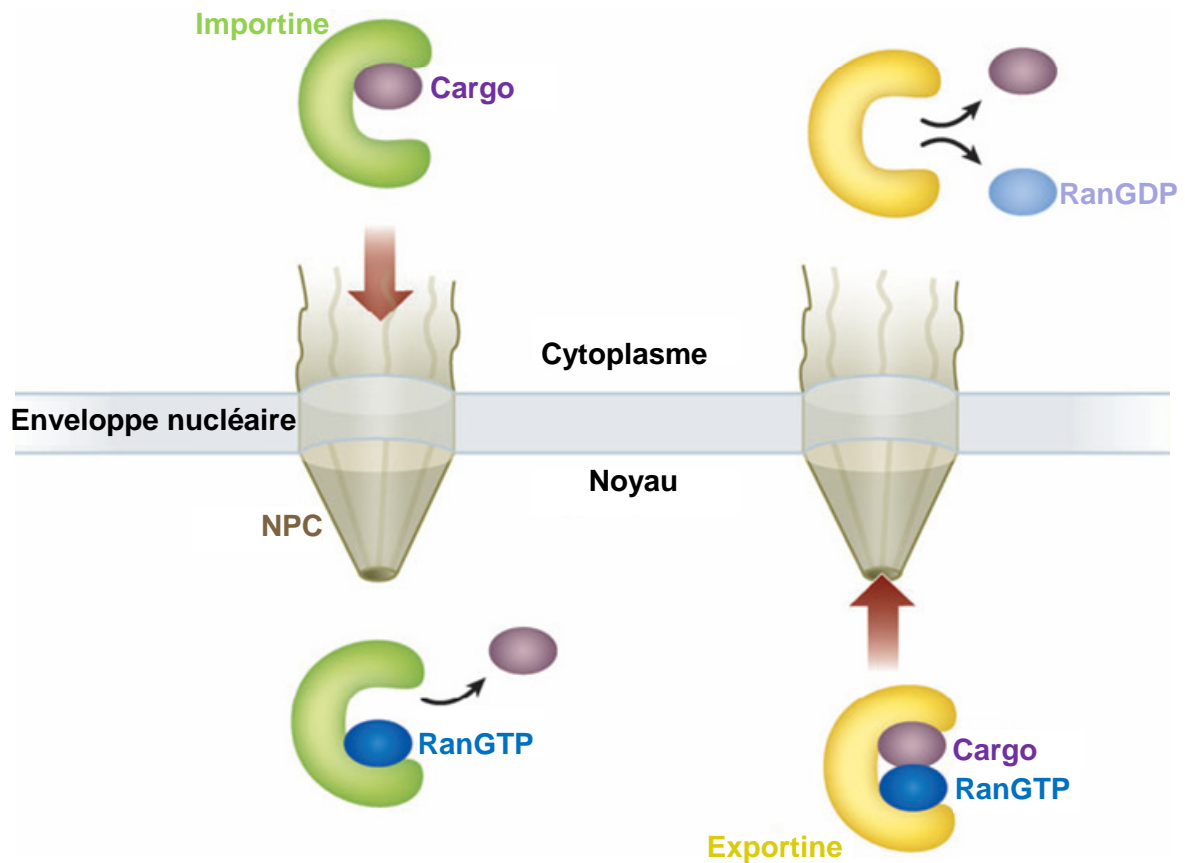
## 2. Le transport nucléaire, import et export

La majeure partie des facteurs de transport nucléocytoplasmiques, appartiennent à la famille des karyophérines  $\beta$ , également connues comme importines  $\beta$  suite à l'identification du premier facteur de transport (Adam & Adam, 1994 ; Chi *et al.*, 1995). Les karyophérines  $\beta$  impliquées dans l'import nucléaire sont généralement appelées importines, alors que celles qui sont impliquées dans l'export nucléaire sont nommées exportines.

Les importines s'associent à la protéine à transporter (le cargo), soit directement, soit indirectement via un adaptateur. Elles sont dirigées vers le NPC, transitent dans le NPC grâce aux motifs FG des nucléoporines (Bayliss *et al.*, 2000) et relâchent leur cargo dans le noyau (**figure 13**). A l'intérieur du noyau, c'est l'association d'une petite protéine G (de la famille des GTPase) nommée Ran qui, dans son état lié au GTP (RanGTP), dissocie le complexe en se liant aux importines et libère le cargo (Rexach & Blobel, 1995 ; Sun *et al.*, 2008).

L'export nucléaire est essentiellement un processus réciproque, dans lequel la reconnaissance du cargo se produit dans le noyau en présence de RanGTP. Le complexe d'export est ensuite dissocié dans le cytoplasme suite à l'hydrolyse de RanGTP (**figure 13**). Contrairement à RanGTP, Ran associée au GDP (RanGDP) n'est pas capable de se fixer aux karyophérines.



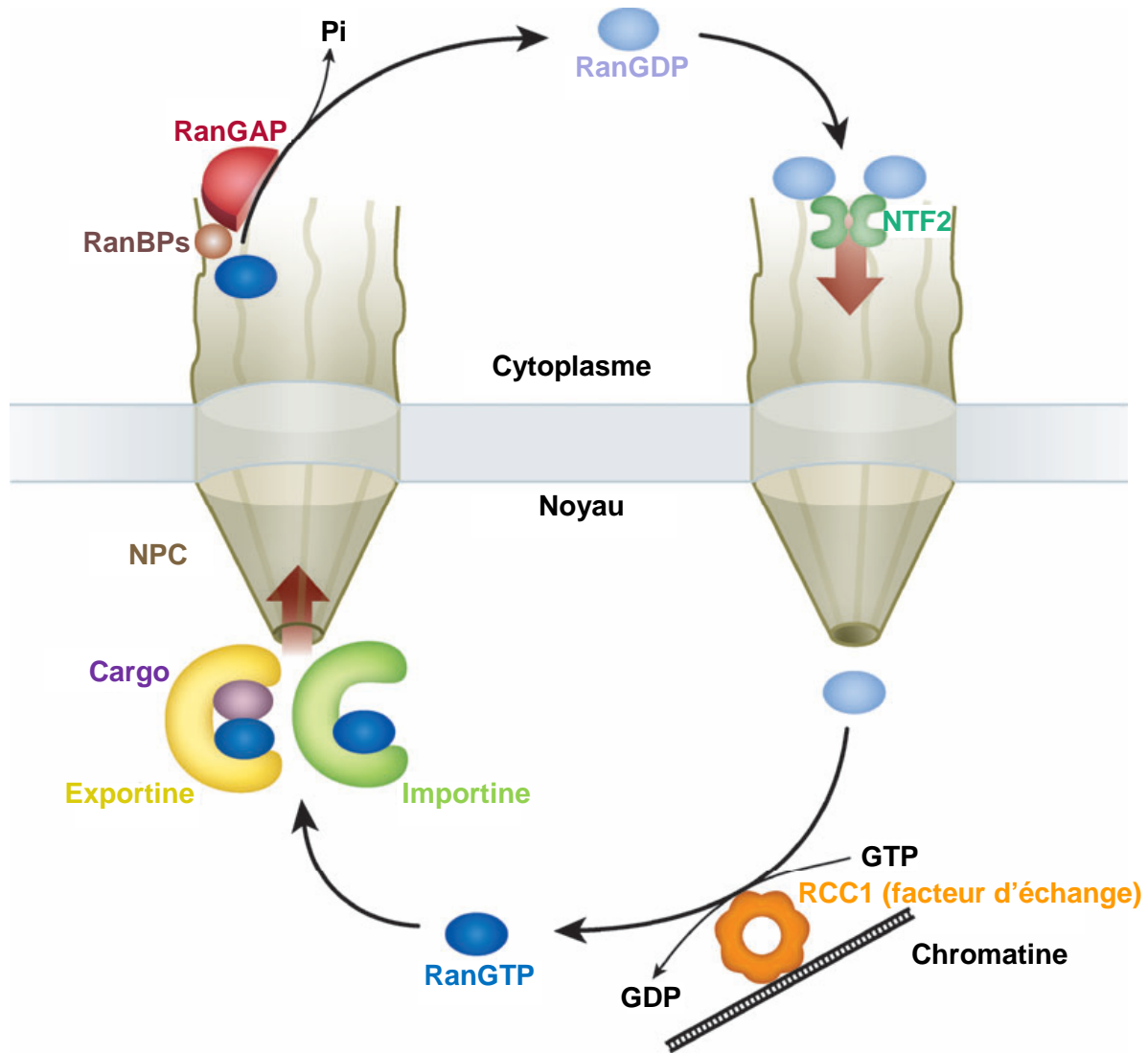


**Figure 13. Illustration schématique des processus d'import et d'export géré par les karyophérine  $\beta$  (revue, Cook *et al.*, 2007).** En haut à gauche, une importine se fixe au cargo dans le cytoplasme et le relâche une fois associée à RanGTP dans le noyau. En bas à droite, Une exportine se fixe à la fois au cargo et à RanGTP dans le noyau et les relâche suite à la conversion de RanGTP en RanGDP dans le cytoplasme.

### 2.1. Distribution asymétrique de RanGTP et RanGDP entre le noyau et la cytoplasme

Pour assurer la directionnalité du transport, la cellule eucaryote possède un système de régulation. Il est basé sur la distribution asymétrique, entre le noyau et le cytoplasme, de la protéine Ran (Ras-related Nuclear protein). Ran est une protéine G de 25kDa que l'on retrouve chez l'ensemble des vertébrés (Görlich *et al.*, 1996 ; Scheffzek *et al.*, 1995). Comme toutes les protéines G, Ran bascule d'un état lié au GTP à celui d'un état lié au GDP par hydrolyse et par échange de nucléotide.

La concentration de Ran-GTP est prédominante dans le noyau alors que celle de Ran-GDP est prédominante dans le cytoplasme. Cette distribution dépend de la localisation des protéines qui régulent la forme nucléotidique de Ran (**figure 14**).



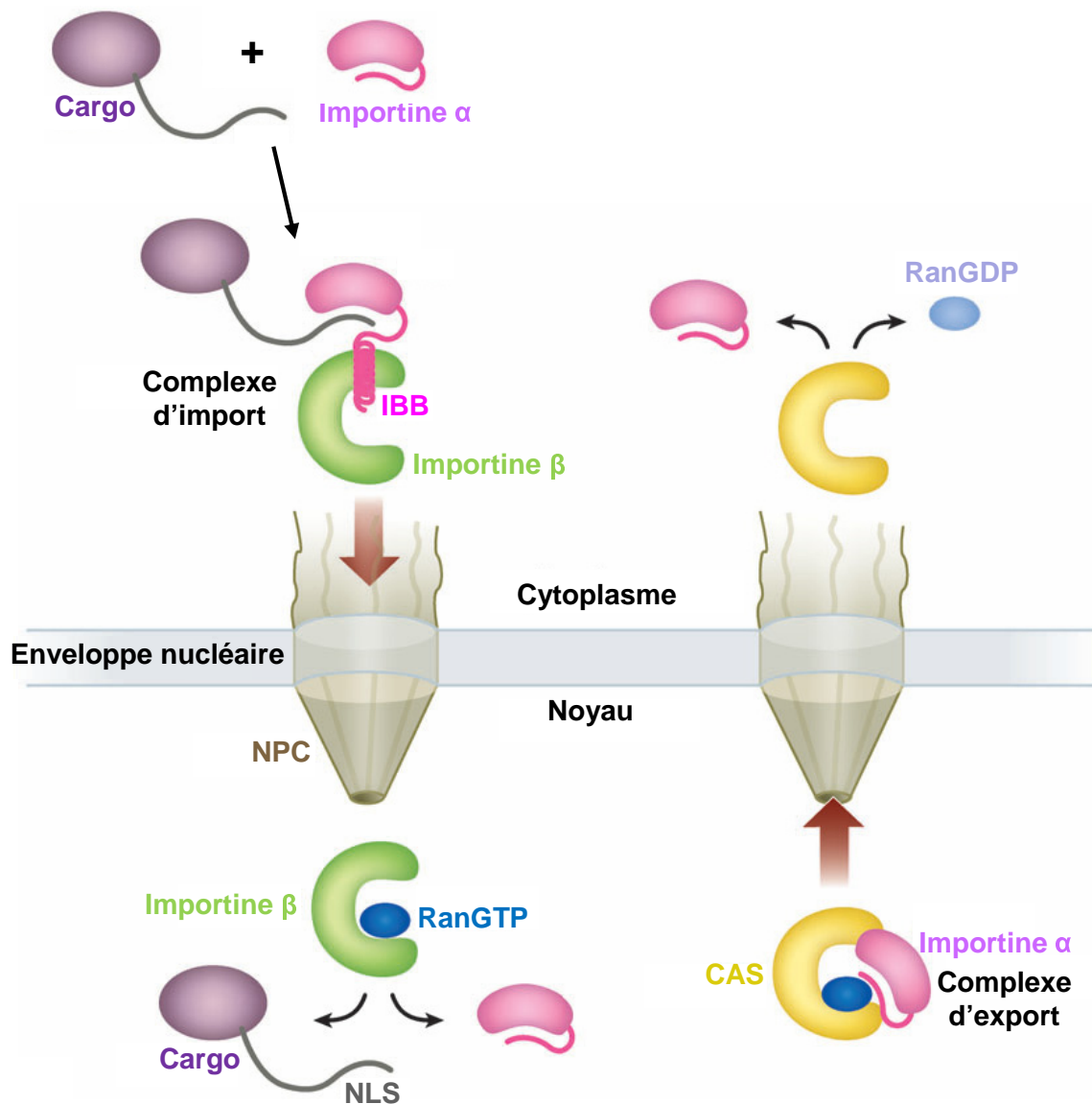
**Figure 14. Régulation asymétrique de l'état de la protéine Ran (revue, Cook *et al.*, 2007).** La haute concentration de RanGDP dans le cytoplasme est maintenue par RanGAP (*GTPase-activating protein*), qui est fixé aux fibres cytoplasmiques du NPC. RanGAP hydrolyse le GTP en GDP suite à son import associé aux karyophérines. RanGDP est recyclé vers le noyau par NTF2 (*nuclear transport factor*). La haute concentration de RanGTP dans le noyau est maintenue par le facteur RCC1 (*regulator of chromosome condensation 1*) fixé à la chromatine, qui est un facteur d'échange de nucléotides. Le gradient de concentration de chacune des 2 formes de la protéine Ran est ainsi maintenu.

Ainsi l'hydrolyse du GTP nécessite la propriété GTPase de RanGAP (*Ran GTPase-activating protein*), qui est présent dans le cytoplasme, accroché aux fibres du pore nucléaire (Mattaj & Englmeier, 1998 ; Kuersten *et al.*, 2001). A l'inverse, l'échange de GDP en GTP, sur la protéine Ran, est catalysé par le facteur d'échange de nucléotide RCC1 (*regulator of chromosome condensation 1*) qui est strictement localisé dans le noyau et majoritairement associé à la chromatine (**figure 14**) (Nemergut *et al.*, 2001 ; Hood & Clarke, 2007).

Des facteurs additionnels contribuent également au cycle de Ran. Ainsi, RanGAP nécessite la présence du co-facteur RanBP(1/2) (*Ran-binding proteins*) pour agir sur les complexes d'export qui atteignent le cytoplasme. Puis, suite à l'hydrolyse, RanGDP est recyclé vers le noyau grâce au facteur de transport NTF2 (*nuclear transport factor*). Cette distribution asymétrique des formes GDP et GTP de Ran entre le cytoplasme et le noyau orchestre ainsi la direction du transport.

## 2.2. L'import nucléaire

Les protéines ou complexes ARN-protéines qui suivent les voies d'import nucléaires sont identifiées grâce à des séquences d'adressage spécifiques. Dans le système d'import le plus courant et le plus étudié, ces séquences d'adressage sont appelées séquences de localisation nucléaire (NLS) et sont reconnues par les importines (Sitterlin, 2005). Les signaux d'import nucléaire les mieux caractérisés sont les NLS dits classiques. Ils sont définis par une forte concentration d'acides aminés basiques (lysine et/ou arginine) et peuvent être monopartite ou bipartite. Ainsi, le NLS monopartite consiste en une séquence linéaire de trois à cinq acides aminés basiques chargés positivement. Il fait référence au peptide NLS de l'antigène T de SV40 (NLS monopartite : -PKKKRKV-) identifié pour la première fois (Kalderon *et al.*, 1984 ; Lanford & Butel, 1984). En ce qui concerne le NLS bipartite, il est constitué de deux groupes de 2 à 3 acides aminés basiques chacun et séparés par environ 10 à 12 acides aminés (Hodel *et al.*, 2001). La référence de ce second motif correspond au NLS de la nucléoplasmine (NLS bipartite : -KR<sub>P</sub>AATK<sub>K</sub>AGQA<sub>K</sub>KKK-) (Dingwall *et al.*, 1988). Il existe d'autres séquences faisant office de NLS qui sont différentes de ces séquences consensus et que l'on appelle NLS non conventionnel.



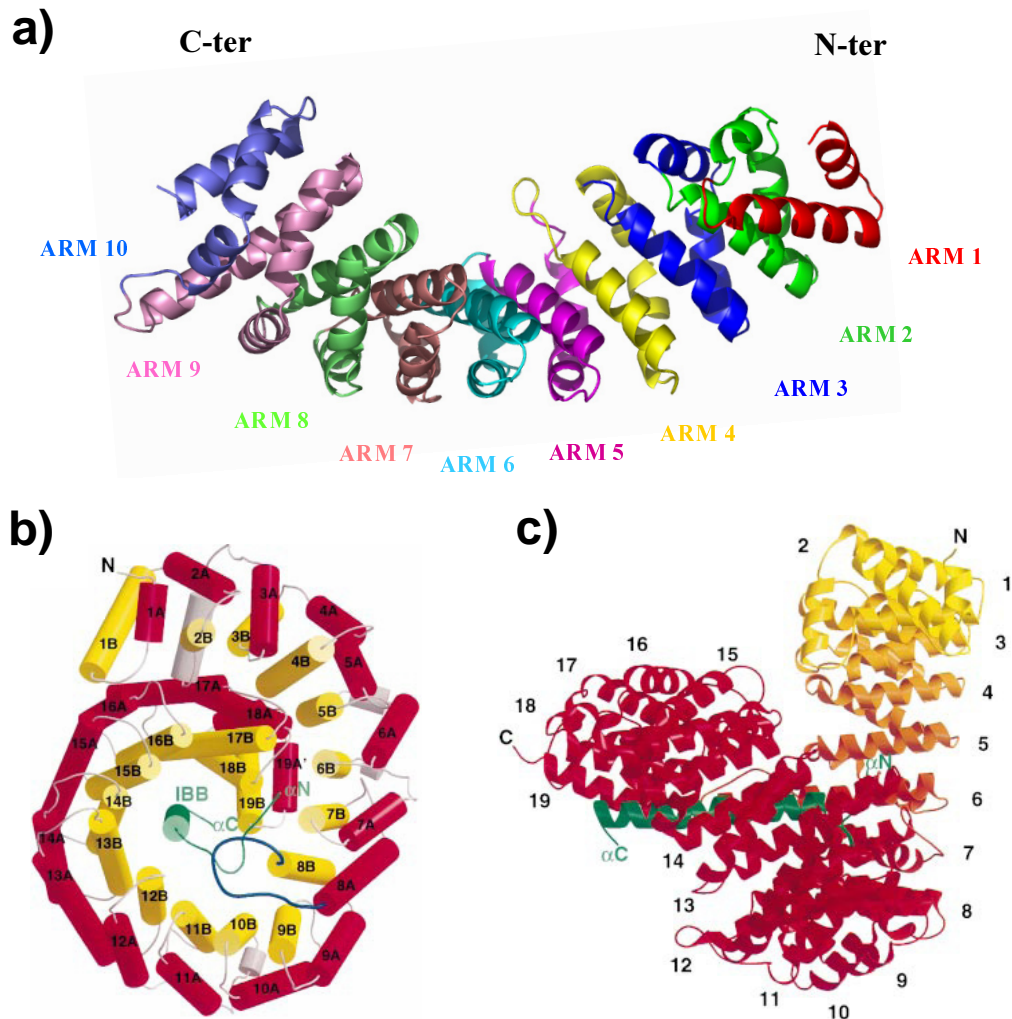
**Figure 15. Illustration schématique du cycle classique d'import nucléaire (revue, Cook *et al.*, 2007).** (En haut à gauche) Dans le cytoplasme, le motif NLS présent sur le cargo est reconnu par la molécule adaptatrice, l'importine  $\alpha$ . Le domaine IBB (*importin- $\beta$  binding domain*) de l'importine  $\alpha$  se fixe à l'importine  $\beta$ . Le complexe est ensuite transporté dans le noyau au travers du pore nucléaire (NPC) où il est dissocié suite à la formation du complexe RanGTP-importine  $\beta$ . En l'absence de l'importine  $\beta$ , le domaine IBB est auto-inhibiteur, puis qu'il se fixe au site de fixation du NLS. (En bas à droite) L'importine  $\alpha$  est recyclée vers le noyau par CAS en complexe avec RanGTP. Suite à l'hydrolyse de RanGTP en RanGDP, l'importine  $\alpha$  est relâchée dans le cytoplasme et est prête pour une nouvelle étape d'import.

La première voie de transport nucléaire des macromolécules mise en évidence, dite "classique", fait intervenir en réalité un complexe entre deux importines, l'importine  $\alpha$  et l'importine  $\beta$  (Adam & Adam, 1994 ; Moroianu *et al.*, 1995 ; Görlich *et al.*, 1995 ; Imamoto *et al.*, 1995 ; Radu *et al.*, 1995). L'importine  $\alpha$  possède deux domaines : un domaine structuré situé sur la partie C-terminale et un domaine flexible appelé IBB (pour *Importin  $\beta$  Binding domain*) en N-terminal. Ce domaine IBB, composé de 40 acides aminés dont la séquence KRR, est replié sur l'importine  $\alpha$  et interagit avec le domaine structuré en C-terminal (revue, Goldfarb *et al.*, 2004). Le cargo possédant une séquence NLS, se fixe au domaine structuré de l'importine  $\alpha$  et rejette ainsi le domaine IBB (**figure 15**). Celui-ci devient à son tour accessible à l'importine  $\beta$  et devient structuré dans le complexe importine  $\alpha/\beta$  (Cingolani *et al.*, 1999 ; Zachariae & Grubmüller, 2008). Ainsi, l'importin  $\alpha$  joue le rôle de molécule adaptatrice entre le cargo et l'importine  $\beta$  (**figure 15**) et le domaine IBB représente un inhibiteur compétitif régulant la fixation du cargo à l'importine  $\alpha$  (Fanara *et al.*, 2000 ; Catimel *et al.*, 2001).

Le complexe d'import ainsi formé se fixe au pore nucléaire grâce à l'importine  $\beta$  qui interagit avec les motifs FG situés sur les fibres du NPC (Bayliss *et al.*, 2000 ; Terry & Wentz, 2007), ce qui conduit au transport du cargo dans le noyau (**figure 15**). A l'intérieur du noyau le complexe est dissocié suite à la fixation RanGTP sur l'importine  $\beta$ . Les importines  $\alpha$  et  $\beta$  sont ensuite recyclées. L'importine  $\alpha$  retourne dans le cytoplasme via son transporteur spécifique nommé CAS lié à RanGTP (Kutay *et al.*, 1997), l'importine  $\beta$  quant à elle utilise directement la protéine RanGTP pour être recyclée. Tout ce mécanisme énoncé n'est pas le seul moyen d'import nucléaire actif mais il représente néanmoins le modèle le plus répandu dans la littérature, utilisé notamment par le génome du virus de la grippe.

### 2.3. Les importines

Aujourd'hui, six importines  $\alpha$  ont été identifiées :  $\alpha 1$ ,  $\alpha 3$ ,  $\alpha 4$ ,  $\alpha 5$ ,  $\alpha 6$  et  $\alpha 7$  (aucune n'ayant pris le nom  $\alpha 2$ ). Ces importines reconnaissent différents cargos mais il n'existe qu'une seule importine  $\beta$ . Suite à l'analyse de différentes lignées de cellules humaines, il apparaît que l'importine  $\alpha 1$  est la plus abondante des isoformes (Köhler *et al.*, 2002). Ainsi, l'importine  $\alpha$  représente la molécule adaptatrice entre le cargo et le système d'import nucléaire. Les importines  $\alpha$  sont des protéines de 50 kDa et des structures tridimensionnelles appartenant à différents organismes sont disponibles.



**Figure 16. Les protéines d'import nucléaire : les importines. a) Structure cristallographique de l'importine  $\alpha$  de levure ( $Kap\alpha 50$ ) (Conti *et al.*, 1998). La molécule contient 10 motifs ARM répétés en tandem. Excepté le premier motif, chaque ARM est composé de trois hélices  $\alpha$  (H1, H2, H3). L'axe superhélicoïdal de la molécule est vertical. b et c) Structure cristallographique de l'importine  $\beta$  humaine, en complexe avec le domaine IBB de l'importine  $\alpha$  humaine (Cingolani *et al.*, 1999). b) Vue suivant l'axe superhélicoïdal : les hélices A et B sont représentées en rouge et jaune respectivement, avec les résidus de connexion en gris. En vert, le domaine IBB de l'importine  $\alpha$ . c) Vue latérale : rotation de  $90^\circ$  suivant un axe vertical par rapport à b.**

Ainsi, les structures des importines  $\alpha$  de la levure *S. cerevisiae* (Kap $\alpha$ 50) et de la souris ont été résolues aussi bien seules qu'en complexe avec les peptides NLS classiques monopartite (NLS de l'antigène T de SV40) et bipartite (NLS de la nucléoplasmine) (Conti *et al.*, 1998 ; Kobe, 1999 ; Conti & Kuriyan, 2000 ; Fontes *et al.*, 2000). Les structures d'importines  $\alpha$  de la levure et de la souris présentent un haut degré de conservation, à la fois concernant la topologie de la protéine, mais aussi sur le mécanisme de reconnaissance avec le NLS. L'importine  $\alpha$  est une protéine allongée, légèrement courbée et hélicoïdale. Sa structure montre qu'elle est constituée de dix motifs répétés en tandem appelés "armadillo" (ARM) (**Figure 16a**). Chaque armadillo est lui-même composé d'une succession de trois hélices  $\alpha$ . La protéine possède deux sillons de liaison au NLS, ce qui suggère un mécanisme de reconnaissance des NLS bipartites mais n'exclut pas les NLS monopartites.

L'importine  $\beta$  est une protéine de 97 kDa qui peut se lier à différents partenaires cellulaires. Elle reconnaît ainsi le domaine IBB de différentes importines  $\alpha$  (Cingolani *et al.*, 1999 ; Zachariae & Grubmüller, 2008), les motifs FG du NPC (Bayliss *et al.*, 2000 ; Terry & Wente, 2007), ou encore la protéine RanGTP (Lee *et al.*, 2005). La structure de l'importine  $\beta$  en complexe avec le peptide IBB de l'importine  $\alpha$  a été résolue à 2,5 Å par cristallographie aux rayons X (Cingolani *et al.*, 1999) (**figure 16b et c**). Cette protéine, dont la forme est enroulée, est composée d'une succession de 19 motifs répétitifs, appelés motifs HEAT, constitués chacun par deux hélices  $\alpha$ . L'importine  $\beta$  est une molécule flexible dont la conformation varie en fonction de la protéine qui lui est associée. Ainsi, la liaison de RanGTP à l'importine  $\beta$  est une étape cruciale dans la libération du cargo dans le noyau. Cette liaison conduit à un changement de conformation de l'importine  $\beta$  (Vetter *et al.*, 1999 ; Weis, 2003 ; Lee *et al.*, 2005) et à la dissociation du complexe d'import nucléaire (Rexach & Blobel, 1995 ; Gilchrist *et al.*, 2002).

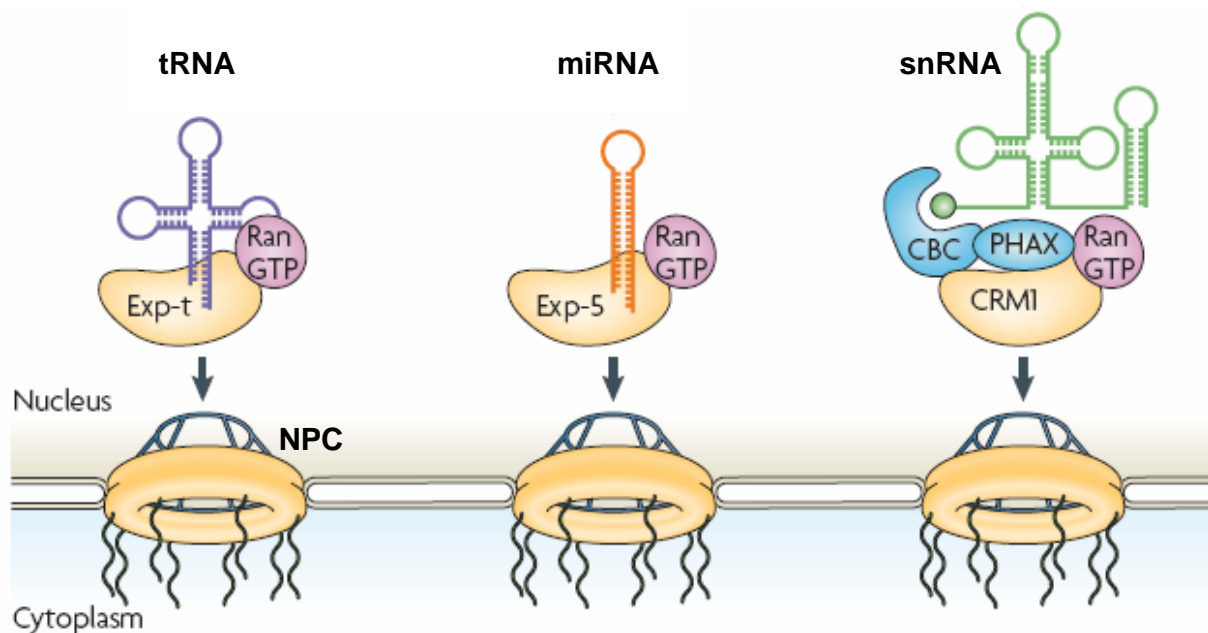
#### 2.4. L'export nucléaire

Les protéines ou complexes ARN-protéines qui suivent les voies d'export nucléaires sont identifiés grâce à des motifs appelés séquences d'export nucléaires (NES pour *nuclear export signal*). Ces séquences d'adressage sont reconnues par les protéines d'export nucléaire, les exportines (Kosugi *et al.*, 2008). Ce sont des facteurs solubles qui interagissent avec le pore nucléaire. Comme dans le cas des importines, il existe différentes exportines qui reconnaissent différents cargos (revue, Cook *et al.*, 2007).

Ainsi, la protéine CAS, déjà mentionnée en amont, est le récepteur spécifique de l'importine  $\alpha$  qui permet son recyclage (Kutay *et al.*, 1997) (**figure 15**). Elle utilise RanGTP pour une fixation coopérative du substrat. L'importine  $\alpha$  ne présente pas de signal d'export linéaire pour CAS, mais elle est reconnue au niveau de l'ARM 10.

L'exportine-t, également membre de la famille des karyophérines  $\beta$ , est impliquée dans l'export nucléaire des ARN de transfert (ARNt) (Kutay *et al.*, 1998 ; Arts *et al.*, 1998). Elle se fixe directement aux ARNt en présence de RanGTP (**figure 17**), en reconnaissant des motifs d'ARN secondaires et/ou tertiaires (Lipowsky *et al.*, 1999 ; Arts *et al.*, 1998).

Les micros ARN (miRNA), qui participent à la régulation des gènes, sont exportés du noyau par une autre karyophérine, l'exportine-5 (Exp-5). Ces pré-micro ARN ayant un recouvrement d'environ deux nucléotides en 3', sont spécifiquement reconnus par l'exportine-5 et transportés dans le cytoplasme (**figure 17**).



**Figure 17. Voies d'export d'ARN par les karyophérines  $\beta$  (revue, Köhler & Hurt, 2007).** Voies d'export des ARN de transfert (tRNA), des micros ARN (miRNA) et des petits ARN nucléaires (snRNA). Les récepteurs d'export (Exp-t, Exp-5 et CRM1) sont représentés en orange, les molécules adaptatrices (CBC et PHAX) sont représentées en bleu. Le transport est dépendant de RanGTP et s'effectue à travers le pore nucléaire (NPC). Le complexe d'export est ensuite dissocié suite à l'hydrolyse de RanGTP en RanGDP.



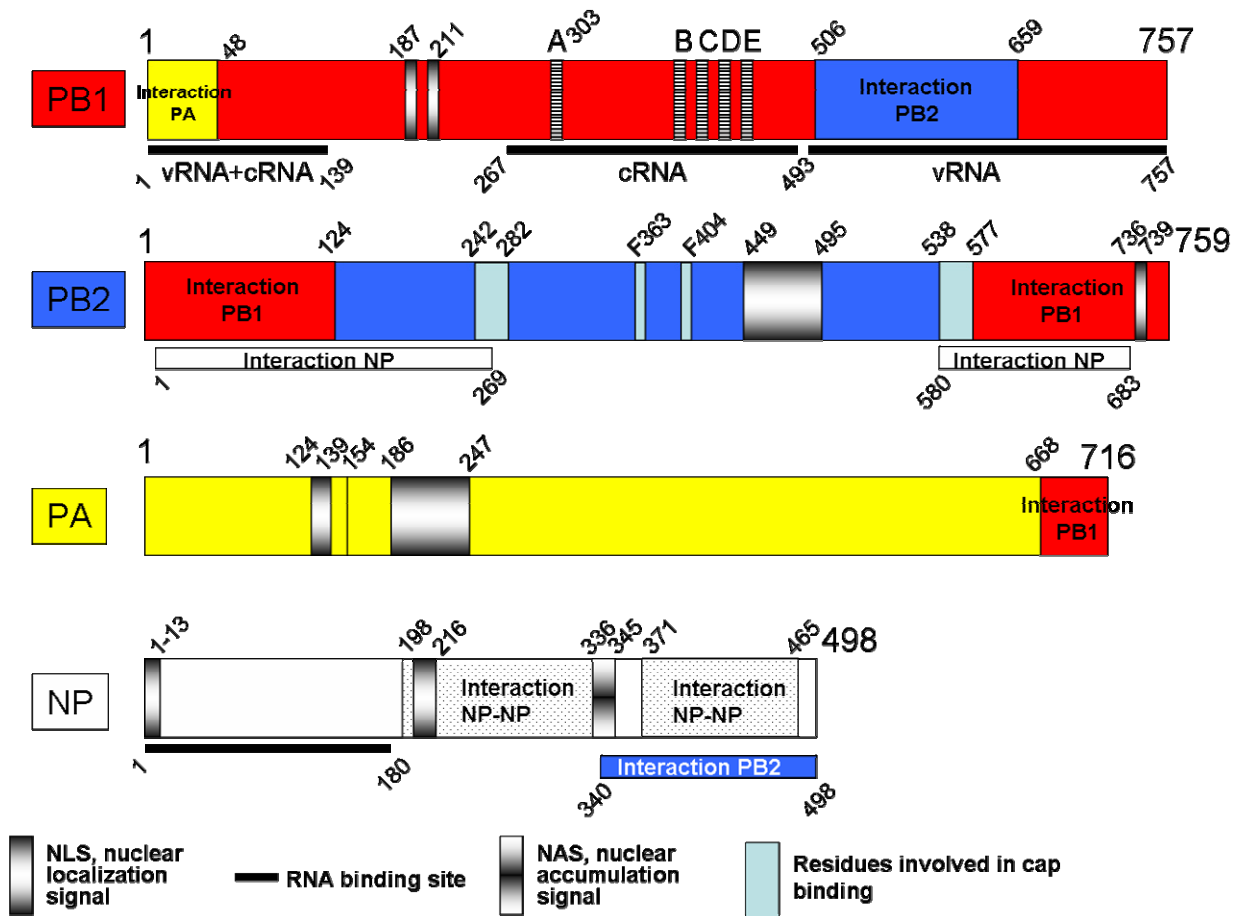
L'exportine CRM1 (pour *Chromosome Region Maintenance protein 1*), aussi appelée exportin 1 (Fornerod *et al.*, 1997 ; Ossareh-Nazari *et al.*, 1997 ; Weis, 2002), est l'un des transporteurs très utilisés pour les protéines. Le NES reconnu par CRM1 est généralement une courte séquence peptidique riche en leucines. La séquence consensus est LxxxLxxLxL, bien que certains autres résidus hydrophobes puissent remplacer les leucines. Elle appartient également à la famille des karyophérines  $\beta$  (Askjaer *et al.*, 1999 ; Petosa *et al.*, 2004) et représente le transporteur généralement utilisé pour l'export nucléaire de protéines et de certains ARN s'ils possèdent un adaptateur. En conséquence, CRM1 ne s'associe pas forcément au cargo, mais requiert parfois la présence de protéines adaptatrices (Schwartz, 2005). Par exemple, CRM1 est impliqué dans l'export des snRNA (petits ARN nucléaires impliqués dans la maturation d'ARNm primaires) et nécessite la présence du complexe CBC (*cap-binding protein*) ainsi que l'adaptateur PHAX qui, lui, contient le motif NES (**figure 17**) (Ohno *et al.*, 2000). L'agent antifongique issu de *Streptomyces*, la leptomycine B, est un inhibiteur spécifique et irréversible de CRM1 (Kudo *et al.*, 1999). Son utilisation permet de caractériser spécifiquement cette voie de transport.

Sur l'ensemble des exportines citées, toutes appartiennent à la famille des karyophérines  $\beta$  et sont dépendantes de RanGTP. Ainsi, la dissociation et le relargage du cargo se produisent dès l'entrée du complexe d'export, dans le cytoplasme, grâce à l'hydrolyse de RanGTP par RanGAP.

### 3. Le transport nucléocytoplasmique des protéines et RNP du virus de la grippe

#### 3.1. L'import nucléaire des RNPv

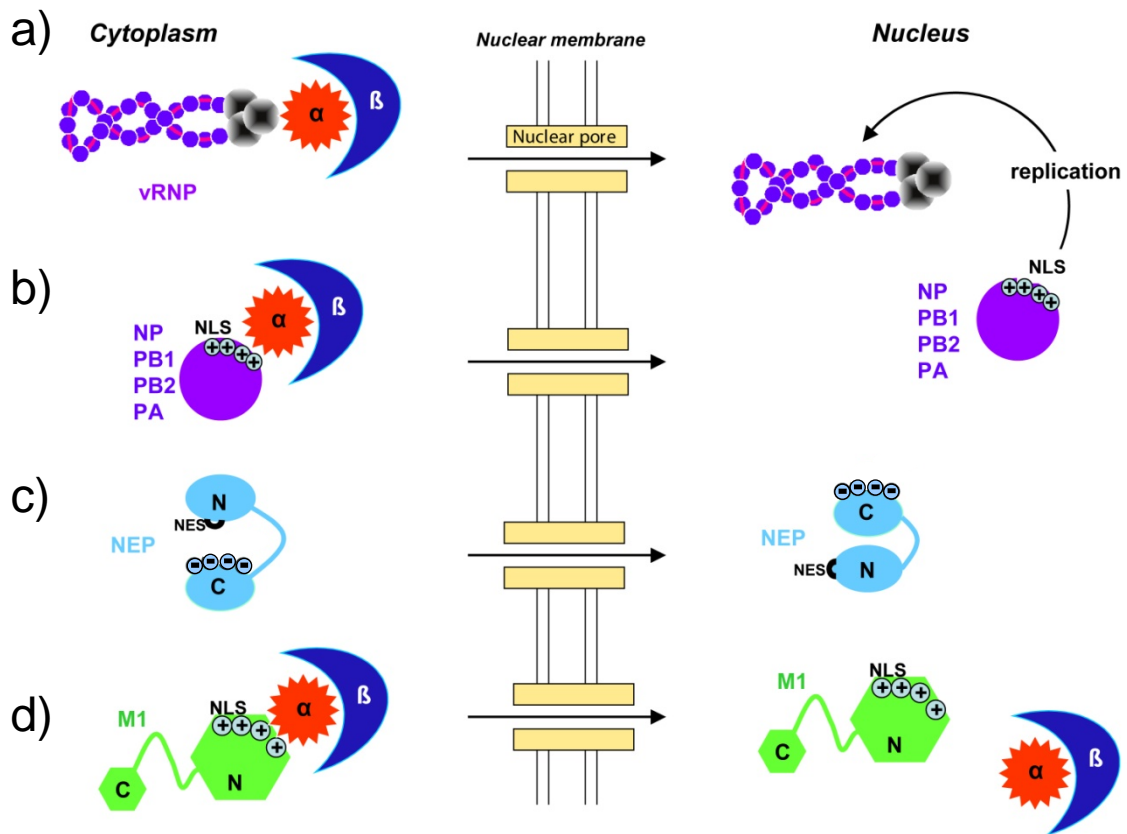
Les RNP ont un diamètre de 15 nm et une longueur variable de 50 à 100 nm en fonction du segment d'ARN viral encapsidé (Elton *et al.*, 1999b). Composée entre autre de nucléoprotéines et de la polymérase, la RNP présente plusieurs candidats potentiels portant les signaux d'import (**figure 18**). Cependant, il n'a toujours pas été établi si les RNPv entrent séparément ou groupées. Néanmoins, elles contiennent l'information nécessaire pour leur import nucléaire et utilisent le système d'import importine  $\alpha$  / importine  $\beta$  décrit précédemment (**figure 19a**). En effet, des RNPv injectées seules dans le cytoplasme de cellules de mammifère sont rapidement accumulées dans le noyau (Kemler *et al.*, 1994). Pour l'instant, le motif directement responsable du transport nucléaire des RNPv n'a pas été défini.



**Figure 18. Les différents domaines des protéines constitutives des RNPv (Boulo *et al.*, 2007).** Les trois sous-unités de la polymérase et la nucléoprotéine sont représentées avec leurs différents domaines NLS ainsi que ceux de leurs activités biologiques respectives. PB1 interagit avec PB2 et PA et elle possède deux signaux de localisation nucléaire (NLS) et trois sites de fixation pour l'ARN viral (ARNv, brin négatif) et l'ARN complémentaire (ARNc, brin positif, intermédiaire de la réplication). Les cinq barres rayées (A-E) indiquent les cinq motifs typiques des ARN polymérase-ARN-dépendantes (Poch *et al.*, 1989). PB2 interagit avec PB1 et avec la nucléoprotéine. La protéine possède plusieurs acides aminés impliqués dans la fixation de la coiffe des ARN messagers cellulaires, utilisés comme amorce pour la transcription et un NLS (aa 736-739) ainsi qu'un domaine dont la délétion entraîne une distribution cytoplasmique (449-495). Le résidu 627 est un déterminant de restriction d'hôte (Subbarao *et al.*, 1993). PA interagit seulement avec PB1 et possède également deux motifs NLS. La nucléoprotéine a des domaines d'autopolymérisation, un domaine d'interaction avec PB2 et un domaine d'interaction avec l'ARN viral. La molécule possède deux NLS et un domaine servant à l'accumulation nucléaire (NAS).

### 3.2. L'import nucléaire des protéines virales

Afin de participer à l'assemblage des RNPv nouvellement produites, la NP et les trois sous-unités de la polymérase doivent rejoindre le noyau. Ces protéines présentent toutes des séquences NLS et empruntent la voie d'import importine  $\alpha$  / importine  $\beta$  (**figure 19**). En plus des protéines nécessaires pour l'assemblage des nouvelles RNP, deux autres protéines virales, M1 et NEP, se retrouvent localisées dans le noyau. Concernant la protéine M1, elle possède un motif NLS classique (Ye *et al.*, 1995) (**figure 19c**). Quant à NEP, aucune séquence NLS n'a été montrée et il est possible qu'elle entre dans le noyau par diffusion sans intervention d'importines (**figure 19d**).



**Figure 19. Processus d'import des RNPv et des protéines nouvellement synthétisées vers le noyau (Boulo *et al.*, 2007).** Les transporteurs cellulaires sont représentés par les importines  $\alpha$  et  $\beta$ , en rouge et en bleu respectivement. Import **a)** des RNP, **b)** de la nucléoprotéine (NP) et des sous-unités de la polymérase (PB1, PB2 et PA), **c)** de la protéine d'export nucléaire (NEP) et **d)** de la protéine matricielle (M1).

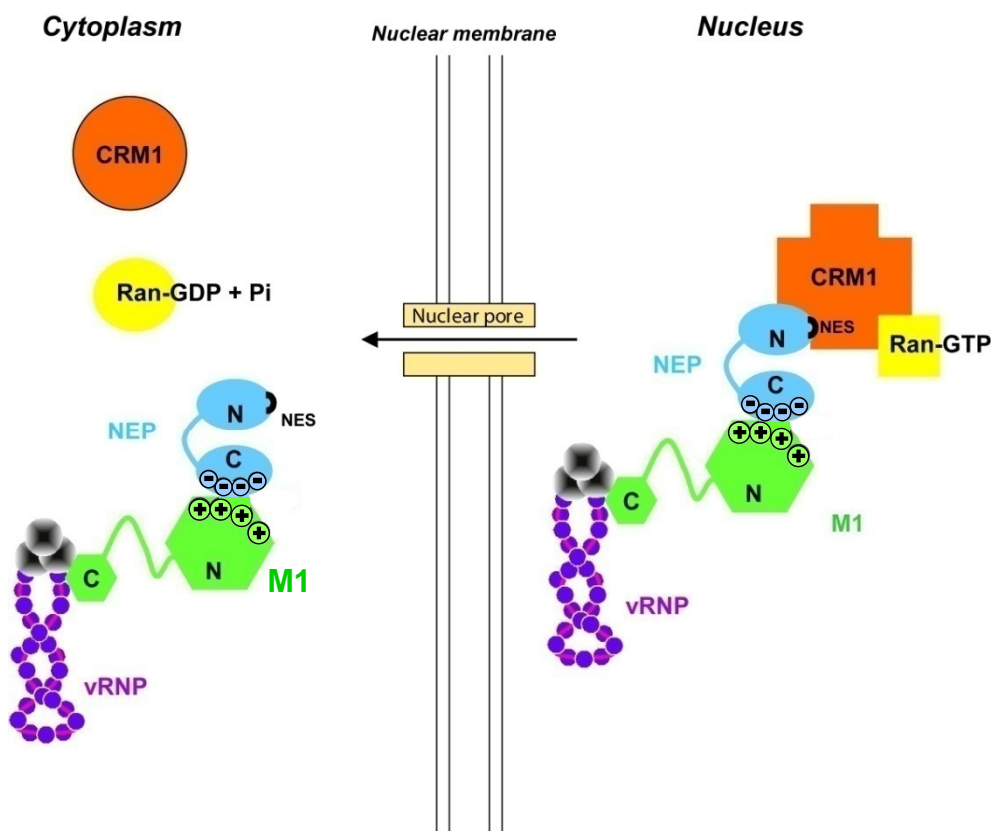
**La polymérase.** La sous-unité PB2 possède un motif NLS bipartite sur la région C-terminale (**Figure 18**). De plus, la délétion de la région comprenant les acides aminés 449 à 495 entraîne une distribution cytoplasmique périnucléaire de PB2, suggérant une étape de transit de la protéine avant son transport dans le noyau (Mukaigawa & Nayak, 1991). Quand la sous-unité PB1 est exprimée en l'absence d'autres protéines virales, elle s'accumule dans le noyau grâce au signal présent sur les résidus 187-RKRR-190 et 207-KRKQR-207 (Nath & Nayak, 1990). Néanmoins, pour être efficacement transportée dans le noyau, PB1 requiert la co-expression de la sous-unité PA (Fodor & Smith, 2004). La sous-unité PA possède deux régions impliquées dans son transport nucléaire : la région située entre les acides aminés 124 à 139 et la région située entre les acides aminés 186 à 247 (Nieto *et al.*, 1994). De plus, la mutation ponctuelle du résidu 154 abolit totalement le transport (Nieto *et al.*, 1994).

**La nucléoprotéine.** Différentes études ont montré que cette protéine présentait plusieurs signaux de transport. Ainsi, un motif NLS dit "non conventionnel" a été établi dans la région N-terminale (acides aminés 1 à 13) (Wang *et al.*, 1997 ; Neumann *et al.*, 1997) et un NLS classique bipartite a été cartographié entre les acides aminés 198-216 (Weber *et al.*, 1998) (**figure 18**). De récents travaux portant notamment sur la structure de la NP (Ye *et al.*, 2006), ont montré que le motif bipartite ne possédait pas les critères requis pour un NLS. Cependant, sa présence serait nécessaire à la réplication du virus (Ozawa *et al.*, 2007). Par ailleurs, des études utilisant des œufs de xénope ont mis en évidence la présence d'une séquence dite d'accumulation/rétention nucléaire (NAS/NRS). Cette séquence située entre les acides aminés 336 et 345 serait à l'origine de l'accumulation de la nucléoprotéine dans le noyau (Davey *et al.*, 1985). De plus, le transport de la nucléoprotéine est régulé par la phosphorylation de la sérine présente en position 3 de la nucléoprotéine (Neumann *et al.*, 1997 ; Arrese & Portela, 1996).

### 3.3. L'export nucléaire de RNPv

Sachant que le transport des RNPv est inhibé par la leptomycine B (Elton *et al.*, 2001 ; Watanabe *et al.*, 2001), il se pourrait que le transporteur CRM1 soit impliqué dans ce processus. Les protéines M1 et NEP sont également présentes dans le noyau. M1 est nécessaire pour déclencher le processus d'export des RNPv (Kemler *et al.*, 1994 ; Sakaguchi *et al.*, 2003). Et différentes études ont montré que la protéine NEP est également impliquée dans le processus d'export nucléaire des RNPv (Yasuda *et al.*, 1993 ; Neumann *et al.*, 2000).

NEP possède un motif NES (O'Neill *et al.*, 1998) mis en évidence par des essais d'interaction directe avec CRM1 (Neumann *et al.*, 2000 ; Akarsu *et al.*, 2003). D'après des résultats *in vitro*, M1 et NEP qui sont deux molécules qui interagissent entre elles, constituent l'interface entre les vRNP et le système d'export cellulaire CRM1-RanGTP (Akarsu *et al.*, 2003). Ainsi, un modèle pourrait être proposé, dans lequel le complexe viral RNPv-M1-NEP est exporté du noyau en utilisant le système d'export cellulaire CRM1- RanGTP (**figure 20**). La protéine directement responsable de la liaison à CRM1 est NEP (Neumann *et al.*, 2000 ; Akarsu *et al.*, 2003). Lorsque NEP est cytoplasmique, son domaine N-terminal flexible contenant le NES serait positionné contre le domaine C-terminal (Akarsu *et al.*, 2003), rendant le NES inaccessible. Une fois NEP entrée dans le noyau, la liaison de son domaine C-terminal avec M1 pourrait exposer le motif NES du domaine N-terminal et rendre le motif NES accessible pour l'interaction avec CRM1.



**Figure 20. Modèle du processus d'export des vRNP en fin d'infection virale (Boulo *et al.*, 2007).** Seule la protéine NEP permet l'export de ses partenaires. Le transporteur cellulaire est représenté par CRM1 associé à Ran-GTP lors de la formation du complexe ternaire.

**- Objectifs de la thèse -**

## Objectifs de la thèse

Durant ces trois années de thèse, je me suis intéressé aux protéines impliquées dans le complexe de réplication et de transcription du virus de la grippe, c'est-à-dire à la polymérase et à la nucléoprotéine. Ces protéines représentent les molécules clés de ces deux processus. Peu d'informations sont aujourd'hui disponibles en raison des difficultés à les produire de façon soluble et en quantité suffisante pour réaliser des études structurales. De plus, l'intérêt de ces travaux réside dans le fait que ces protéines ont récemment été mises en évidence comme étant impliquées dans la spécificité de l'hôte et dans la transmission du virus inter-espèces.

La réplication du virus de la grippe passe par une étape nucléaire. C'est dans ce contexte que nous nous sommes intéressés aux interactions entre virus et cellule hôte. Au sein d'un laboratoire de virologie axé vers des techniques structurales, nous avons tenté de réaliser des complexes, protéine virale - protéine cellulaire, stables en vue d'une co-cristallisation. Ensuite, j'ai eu la chance de partir pour ma 3<sup>ème</sup> année de thèse à l'EMBL de Heidelberg dans l'unité *Gene Expression*. Je me suis ainsi orienté vers une étude moins cristallographique, pour compléter mes travaux en utilisant davantage d'expériences biochimiques et biophysiques.

**Etude de la polymérase.** Pour pallier aux difficultés d'expression et d'obtention des sous-unités de la polymérase, il a été possible, en collaboration avec une équipe de l'EMBL-Grenoble, d'isoler des clones exprimant des fragments solubles de la sous-unité PB2. Ainsi un polypeptide comprenant la région C-terminale (98 derniers acides aminés) a été mis en évidence. Cette région présente des acides aminés impliqués dans la spécificité de l'hôte, mais également une séquence de localisation nucléaire. Afin de déterminer le mécanisme d'import nucléaire, nous avons étudié son interaction avec son transporteur nucléaire, l'importine  $\alpha 5$  humaine.

**Etude de la nucléoprotéine.** La NP est l'autre protéine clé impliquée dans les processus de réplication et de transcription du génome viral. Afin de comprendre le mode de fonctionnement de cette protéine, nous avons étudié tout d'abord son interaction avec l'ARN. La NP polymérise le long des segments génomiques viraux, et de ce fait, nous avons tenté de faire un lien entre son état d'oligomérisation et son affinité pour l'ARN. La NP emprunte également la voie d'import nucléaire classique. Nous avons donc analysé son interaction avec le transporteur nucléaire, l'importine  $\alpha 5$  humaine. Enfin, parce que la nucléoprotéine s'avère

également être un déterminant de la spécificité de l'hôte, nous avons étudié la mutation N319 en lysine impliquée dans la conversion d'un virus aviaire à un virus adapté chez le mammifère. L'objectif de cette étude étant de voir si cet acide aminé présente une meilleure affinité pour son transporteur, résultant ainsi en une activité accrue de la polymérase.

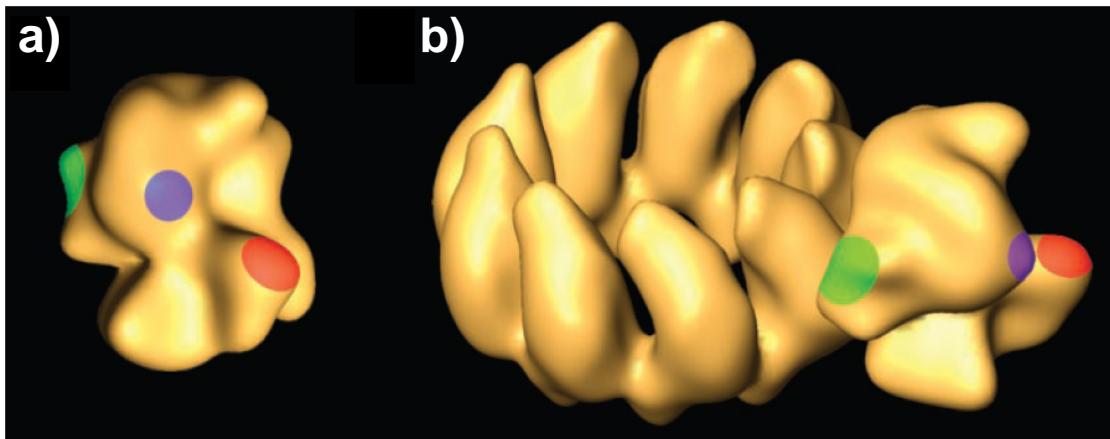


## **II. RESULTATS ET DISCUSSION**

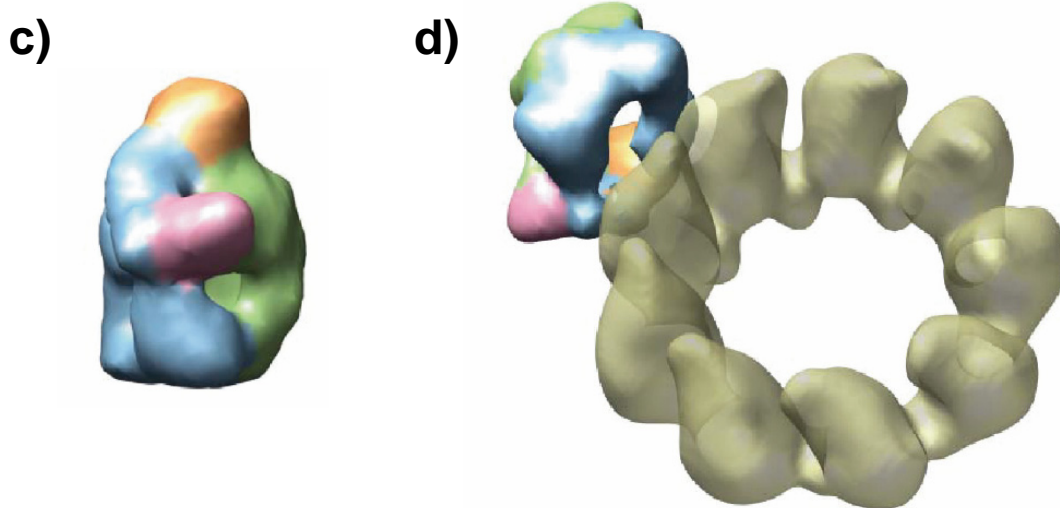
# Chapitre 1

—

## Le complexe d'import nucléaire : importine $\alpha 5$ humaine / sous-unité PB2 de la polymérase du virus de la grippe



*Area et al., 2004*



*Torreira et al., 2007*

**Figure 21. Modèles en 3 dimensions de la polymérase du virus de la grippe. a et b) Modèle 3D de la polymérase obtenu à partir de mini RNP (Area et al., 2004).** Les zones colorées indiquent la position de domaines spécifiques aux sous-unités : en rouge, le domaine N-terminal de PB2 ; en bleu, le domaine C-terminal de PA ; en vert, le domaine C-terminal de PB1. **a)** Modèle obtenu par reconstruction 3D à partir d'images de microscopie sur des RNP recombinantes. **b)** Position des sous-unités de la polymérase dans le contexte de la structure d'une RNP. **c et d) Modèle 3D du complexe polymérase soluble (Torreira et al., 2007).** **c)** Modèle obtenu par reconstruction 3D à partir d'images de microscopie de la polymérase soluble. **d)** Modèle chimérique entre la polymérase soluble (isolée) et la structure d'une RNP. Le modèle est représenté avec ses domaines structuraux de couleurs différentes. La densité de protéine située vers l'avant est colorée en vert, tandis que l'arrière est bleu. Le sommet est représenté en orange et la protubérance est colorée en rose.

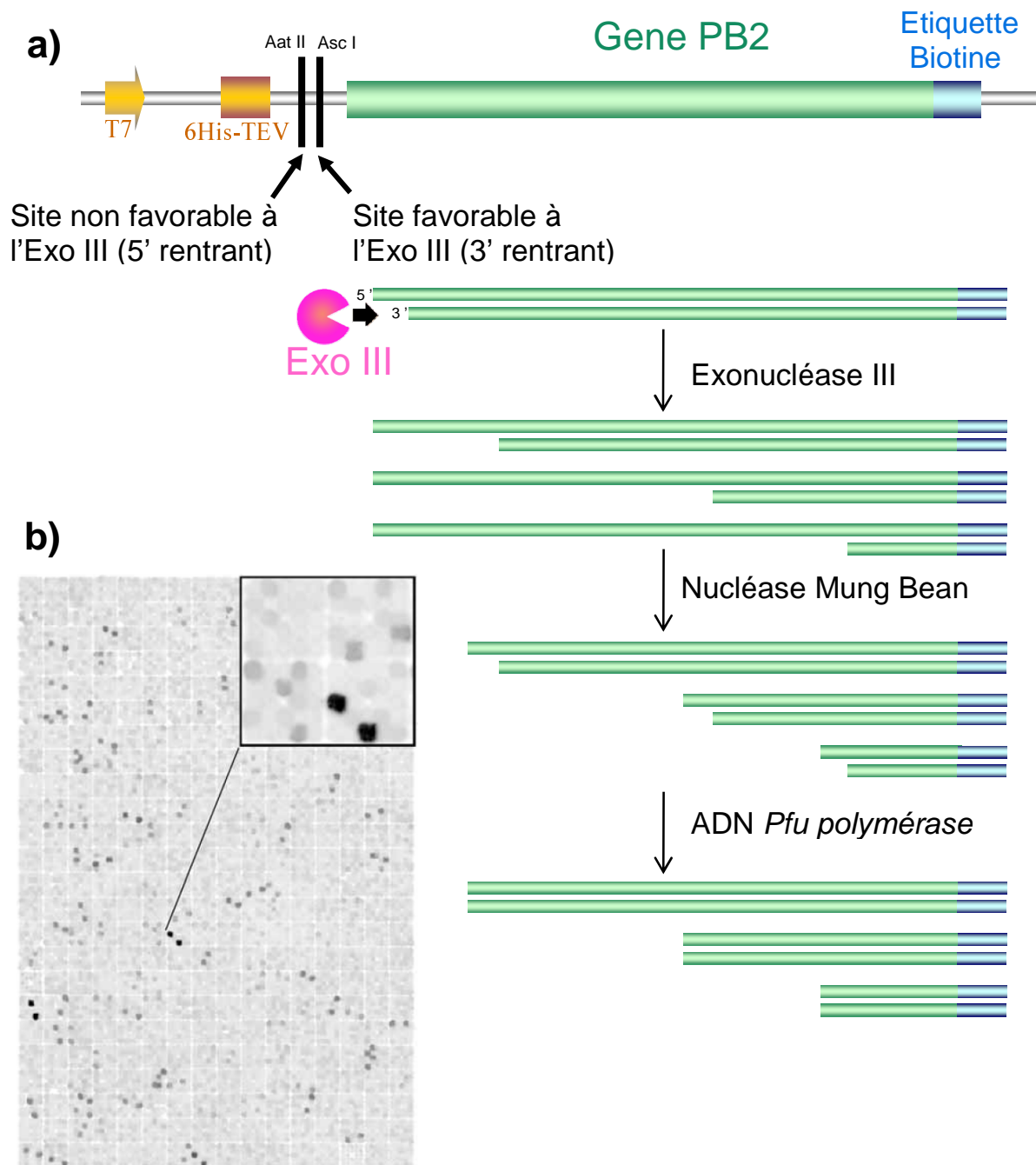
# Chapitre 1

## Le complexe d'import nucléaire : importine $\alpha 5$ humaine / sous-unité PB2 de la polymérase du virus de la grippe

### 1. Introduction

La préoccupation actuelle concernant l'adaptation à l'homme de la souche de la grippe aviaire, H5N1, nécessite de comprendre ce qui détermine la virulence ainsi que la transmissibilité du virus (Russell & Webster, 2005). Des études récentes portant sur des mutations ponctuelles des sous-unités de la polymérase et de la nucléoprotéine ont attiré l'attention. Elles accompagnent l'adaptation des virus aviaires en virus mammifères (Li *et al.*, 2005 ; Salomon *et al.*, 2006), comme ce fut le cas de la souche pandémique de 1918 (Taubenberger *et al.*, 2005). Cependant, les implications fonctionnelles de ces mutations – à savoir, si elles affectent l'activité de la polymérase, l'assemblage ou la stabilité des interactions avec les partenaires de l'hôte – sont mal connues. Ceci est dû en partie au manque d'informations structurales détaillées sur la polymérase.

Des modèles, obtenus en microscopie électronique, de la polymérase sont néanmoins disponibles. Il s'agit tout d'abord d'une reconstruction résolue à 23 Å par microscopie électronique en coloration négative des trois sous-unités présentes dans une mini RNP (Area *et al.*, 2004). Grâce à l'utilisation d'anticorps dirigés contre PB2 ou PA et de l'étiquette TAP sur PB1, ce modèle présente la polymérase dans sa globalité ainsi que la localisation de chacune des sous-unités du complexe (**figure 21a et b**). Plus récemment, un modèle du trimère de la polymérase seule a été déterminé à 26 Å de résolution, également en microscopie électronique, par le même groupe (Torreira *et al.*, 2007) (**figure 21c et d**). La comparaison des deux modèles montre une conformation plus ouverte pour la polymérase "soluble", c'est-à-dire isolée. Cette différence de conformation peut être attribuée soit à l'absence de l'ARN, soit à l'absence des nucléoprotéines. Cependant les détails atomiques des trois sous-unités ne sont pas encore accessibles et seule une structure à résolution atomique pourrait apporter des informations nécessaires à l'élaboration de cibles thérapeutiques.



**Figure 22. Identification du domaine soluble de la partie C-terminale de PB2.** **a)** Principe de la méthode utilisée pour tronquer le gène de la sous unité PB2 de la polymérase. Après ouverture du plasmide par des enzymes de restriction, l'exonuclease III se charge de dégrader le brin d'ADN associé à l'extrémité 3' rentrante du gène de PB2. La Mung Bean nucléase dégrade ensuite le simple brin. Enfin, une étape de polissage est réalisée avec l'ADN *Pfu* polymérase afin d'obtenir des bouts francs. **b)** Criblage de l'expression de protéines sur plus de 26 000 constructions aléatoires du gène PB2. Le vecteur utilisé permet de produire la protéine fusionnée à une étiquette biotine. L'expression de protéines solubles autorise alors un marquage par la biotine. La détection est faite à l'aide de streptavidine couplée à un fluorophore (Tarendeau *et al.*, 2007).

Exprimées dans *Escherichia coli*, les sous-unités de la polymérase présentent un caractère totalement insoluble. Dans les cellules d'insectes ou de mammifères, leur expression est très faible, il est donc difficile d'en obtenir en quantité suffisante pour réaliser des études structurales. De plus, la connaissance biochimique ou bioinformatique sur la structure des domaines de la polymérase est limitée.

En collaboration avec l'EMBL-Grenoble, une nouvelle approche appelée ESPRIT (Expression of Soluble Proteins by Random Incremental Truncation) a été utilisée pour mettre en évidence des domaines solubles de la polymérase. Par cette méthode, il a été possible d'isoler des clones exprimant des fragments solubles de la polymérase à partir d'une banque d'ADN tronqués, soit de l'extrémité N-terminale, soit de l'extrémité C-terminale.

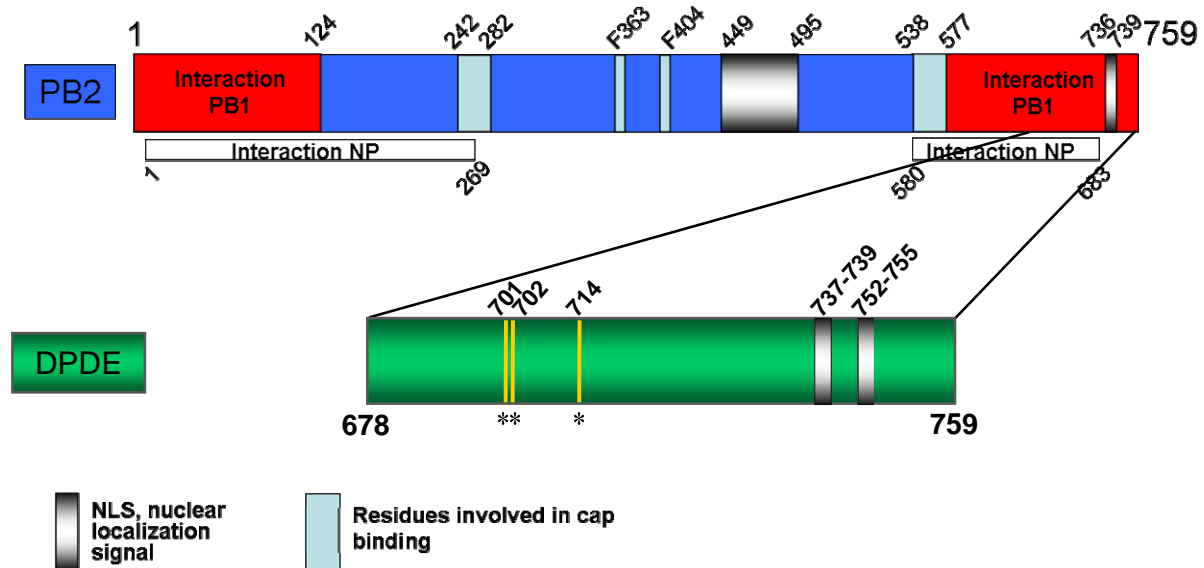
Ainsi un polypeptide comprenant les acides aminés 678 à 759 de PB2 a été mis en évidence. Ce domaine C-terminal de la sous-unité PB2 de la polymérase n'a pas pu cristalliser, cependant la structure a pu être résolue par RMN en collaboration avec le groupe de RMN à l'IBS (Institut de Biologie Structurale) de Grenoble. En parallèle, nous avons tenté d'associer ce domaine à l'importine  $\alpha 5$  humaine et nous avons réussi à isoler un complexe dont la structure a été résolue par cristallographie aux rayons X. L'ensemble de ces résultats a fait l'objet d'une publication (Tarendeau *et al.*, 2007).

## 2. Isolement et structure RMN du domaine soluble de PB2 (C-ter)

### 2.1. Identification et isolement du domaine soluble

Le gène de PB2 employé dans ces travaux provient de la souche A/Victoria/3/75 (H3N2) qui dérive du virus responsable de la grippe de Hong Kong en 1968 (de la Luna *et al.*, 1989). La stratégie utilisée pour mettre à jour un domaine soluble de PB2 repose sur la création d'une librairie de gènes tronqués. La mise en place de cette librairie passe par des techniques de biologie moléculaire à haut débit (plateforme à haut débit, EMBL-Grenoble) et notamment par l'utilisation de l'exonucléase III qui intervient dans la troncation du gène d'intérêt (Ostermeier & Lutz, 2003). Le vecteur, dans lequel le gène de la sous-unité PB2 de la polymérase a été cloné, est ouvert par des enzymes de restriction de manière à ce que seule une extrémité 3' rentrante se retrouve du côté du gène viral. L'exonucléase III dégrade alors le brin d'ADN associé à cette extrémité 3' rentrante puis la Mung Bean nucléase élimine l'ADN simple brin complémentaire non hydrolysé (**figure 22a**). Une étape de polissage, réalisée par une ADN polymérase permet d'obtenir des bouts francs. Le vecteur est ensuite

refermé sur lui-même afin qu'il puisse être utilisé dans une transformation. En jouant ainsi sur la cinétique de dégradation par l'exonucléase III, une librairie de troncation du gène de PB2 est obtenue (**figure 22a**). Pour à la fois distinguer les fragments de PB2 en phase et ceux qui sont solubles, le vecteur utilisé permet l'expression d'une protéine fusionnée avec une étiquette biotine (**figure 22a**). Ainsi, si la protéine exprimée est correctement repliée et soluble, la biotine ligase BirA, exprimée par la bactérie ([Chapman-Smith & Cronan, 1999](#)), va pouvoir lier la biotine à la protéine fusion. En utilisant l'avidité de la streptavidine pour la biotine, il devient aisé de révéler les clones exprimant une protéine soluble marquée par la biotine, en couplant la streptavidine à un fluorophore. Ceci est visualisé par des points de très forte intensité sur la **figure 22b**. En revanche, si la protéine exprimée est mal repliée, dégradée ou séquestrée dans des corps d'inclusion, le marquage par la biotine n'aura pas lieu et les clones correspondant ne seront pas révélés. Ainsi un polypeptide de 81 acides aminés (environ 10 kDa) correspondant à la partie C-terminale de PB2 (acides aminés 678 – 759) a été mis en évidence. Sa séquence en N-terminal débute par les acides aminés DPDE, nous avons ainsi nommé ce polypeptide (**figure 23**).



**Figure 23. Domaines de la sous-unité PB2 de la polymérase du virus de la grippe et fragment soluble de la partie C-terminale : DPDE (acides aminés : 678 à 759).** PB2 possède deux NLS, un motif interne (acides aminés 449-495) et un motif situé dans la région C-terminale (acides aminés 736-739). Sont indiqués sur DPDE le motif NLS en C-terminal et par une astérisque, les acides aminés impliqués dans la transmissibilité du virus inter-espèces.

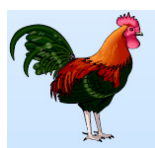
## 2.2. Spécificité de l'hôte

De récents travaux, portant sur la transmissibilité du virus de la grippe, ont montré que des mutations étaient responsables de la conversion d'un virus aviaire, SC35, à un virus adapté et létale chez le mammifère (la souris), SC35M (Gabriel *et al.*, 2005). Ainsi, le virus nommé SC35 dérive de la souche d'origine aviaire A/SEAL/Massachusetts/1/80 (H7N7) par passage multiple sur des cellules embryonnaires de poulet devient 100% mortel chez le poulet. Le virus SC35 est ensuite passé à plusieurs reprises sur des cellules de poumons de souris, générant ainsi le variant adapté à la souris, SC35M. Ainsi, après séquençage du génome des deux souches virales, SC35 et SC35M diffèrent par la présence de neuf mutations uniquement, dont six principalement sur la polymérase (PB2, PB1 et PA), mais aussi une sur la nucléoprotéine, une sur l'hémagglutinine et une sur la neuraminidase (**figure 24**).

Concernant la sous-unité PB2, la substitution D701N contribuerait considérablement à rendre pathogènes des souches aviaires H5N1 (Li *et al.*, 2005) et H7N7 (Gabriel *et al.*, 2005) chez les souris. De plus, cette substitution a pu être observée dans des cas fatals, ou non, d'infections humaines au Vietnam (de Jong *et al.*, 2006). La substitution de la sérine 714 en arginine, augmente l'activité de la polymérase sur des RNP reconstituées dans les cellules 293T de mammifère. De plus, cette mutation en commun avec la substitution D701N, augmente davantage cette activité de polymérase. Ainsi, les mutants comportant l'Asn701 et l'Arg714 présentent une meilleure activité de polymérase (Gabriel *et al.*, 2005).

Il est très intéressant de voir que sur ces neuf acides aminés impliqués dans la transmissibilité du virus inter-espèce, deux sont présents sur le fragment soluble DPDE (**figures 23 et 24**). De plus, l'acide aminé 702, également présent sur DPDE, est aussi un déterminant majeur de la spécificité d'hôte. C'est généralement une arginine dans des isolats humains (incluant le virus pandémique de 1918, H1N1) et une lysine dans les souches aviaires (Taubenberger *et al.*, 2005).





Souche SC35 (H7N7), Létale chez le poulet  
Faiblement pathogène chez la souris



Souche SC35M (H7N7), Létale chez la souris  
Virulence atténuée chez le poulet

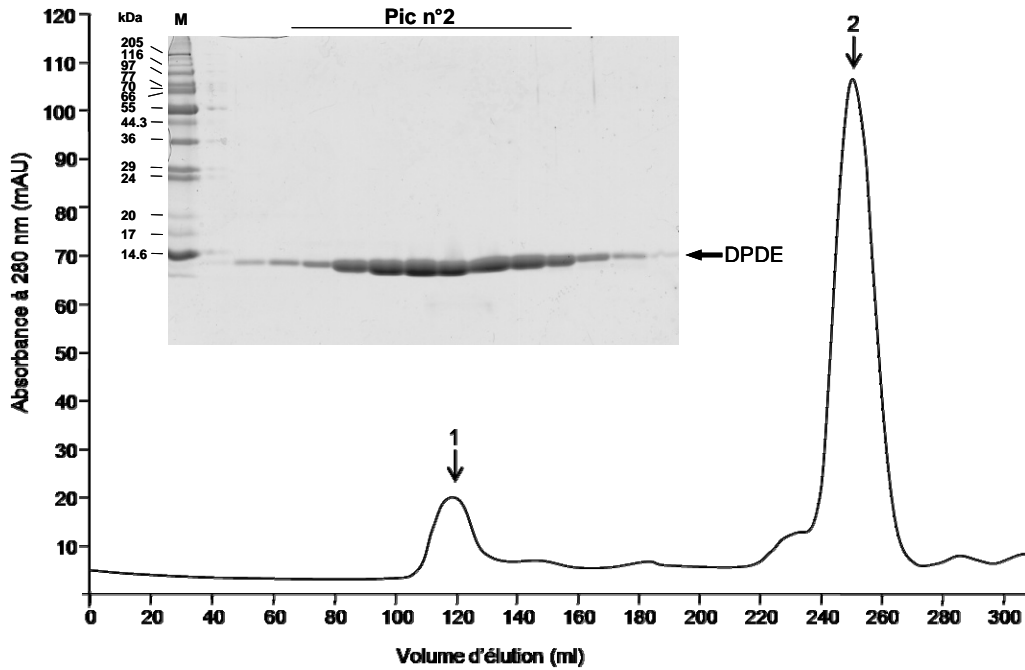
	PB1		PB2			PA	HA	NP	NA
	13	678	333	701	714	615	340	319	328
<b>SC35</b>	L eu	Ser	Thr	Asp	Ser	Lys	Gly	Asn	Ala
<b>SC35M</b>	Pro	Asn	Ile	Asn	Arg	Asn	Arg	Lys	Ser

**Figure 24. Adaptation d'un virus aviaire à un hôte mammifère.** Le virus aviaire de souche SC35 devient adapté à la souris (SC35M) par passages successifs sur des cellules de poumons de souris. Après séquençage, il en résulte une différence de 9 mutations entre les gènes de SC35 et de SC35M. Parmi elles, deux sont présentes sur le fragment soluble DPDE. [Adapté de Gabriel et al., 2005.](#)

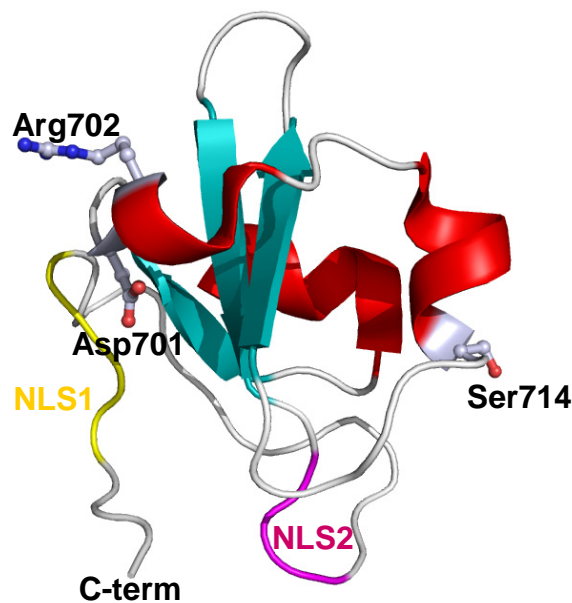
### 2.3. Structure RMN de DPDE en solution

La mise en évidence de ce domaine soluble de la protéine permet une production et une purification aisée de plusieurs dizaines de milligrammes pour seulement un litre de culture cellulaire. Ainsi à la suite d'une étape de purification par chromatographie d'exclusion, DPDE est éluée de façon très pure (**figure 25**). Les premiers essais de cristallogénèse réalisés sur ce domaine C-terminal de PB2 n'ont pas abouti à des cristaux. Cependant, la structure en résonance magnétique nucléaire (RMN) de ce domaine a été résolue en collaboration, avec l'équipe du Dr J.P. Simorre à l'IBS (**figure 26**). Elle révèle un domaine compact  $\alpha$ - $\beta$  dont le cœur hydrophobe est constitué de 3 feuillets  $\beta$  antiparallèles, sur lequel est positionnée une courte hélice  $\alpha$ .

De plus, sont présents sur ce domaine C-terminal de PB2, trois des acides aminés impliqués dans la spécificité de l'hôte. Ainsi, l'aspartate 701, l'arginine 702 et la sérine 714 sont visibles et exposés vers le solvant, cependant, il n'est pas exclu que cette exposition diffère sur la protéine entière.



**Figure 25. Purification de DPDE.** Profil d'élution de la partie C-terminale de PB2 (DPDE : acides aminés 678 à 759) par chromatographie d'exclusion de taille sur une colonne Superdex 200 HiLoad (GE Healthcare). Les flèches indiquent (1) le volume mort et (2) le volume d'élution de DPDE (*MM* 10 *kDa*). Analyse par électrophorèse sur gel de polyacrylamide (SDS-PAGE 12% Laemmli) du pic n°2 correspondant DPDE. M = Marqueur moléculaire.



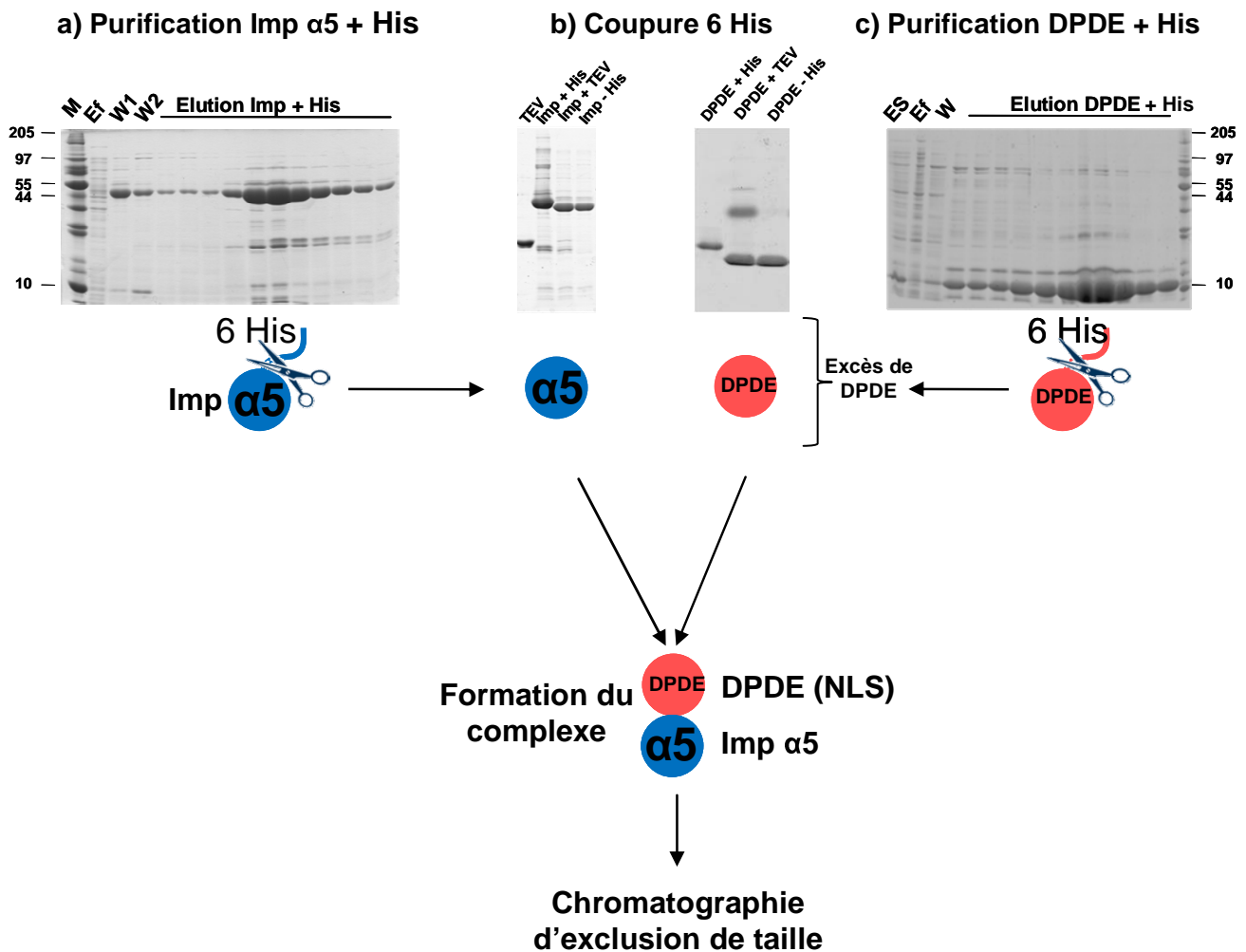
**Figure 26. Structure RMN de la partie C-terminale de PB2 (DPDE : acides aminés 678 à 759) en solution (Tarendeau *et al.*, 2007).** Sont indiquées les régions basiques correspondant au site mineur (violet) et site majeur (jaune) du NLS bipartite. Sont également représentés les acides aminés Asp701, Arg702 et Ser714 qui sont impliqués dans la transmission inter-espèces.

### 3. Le complexe Imp $\alpha$ 5-DPDE

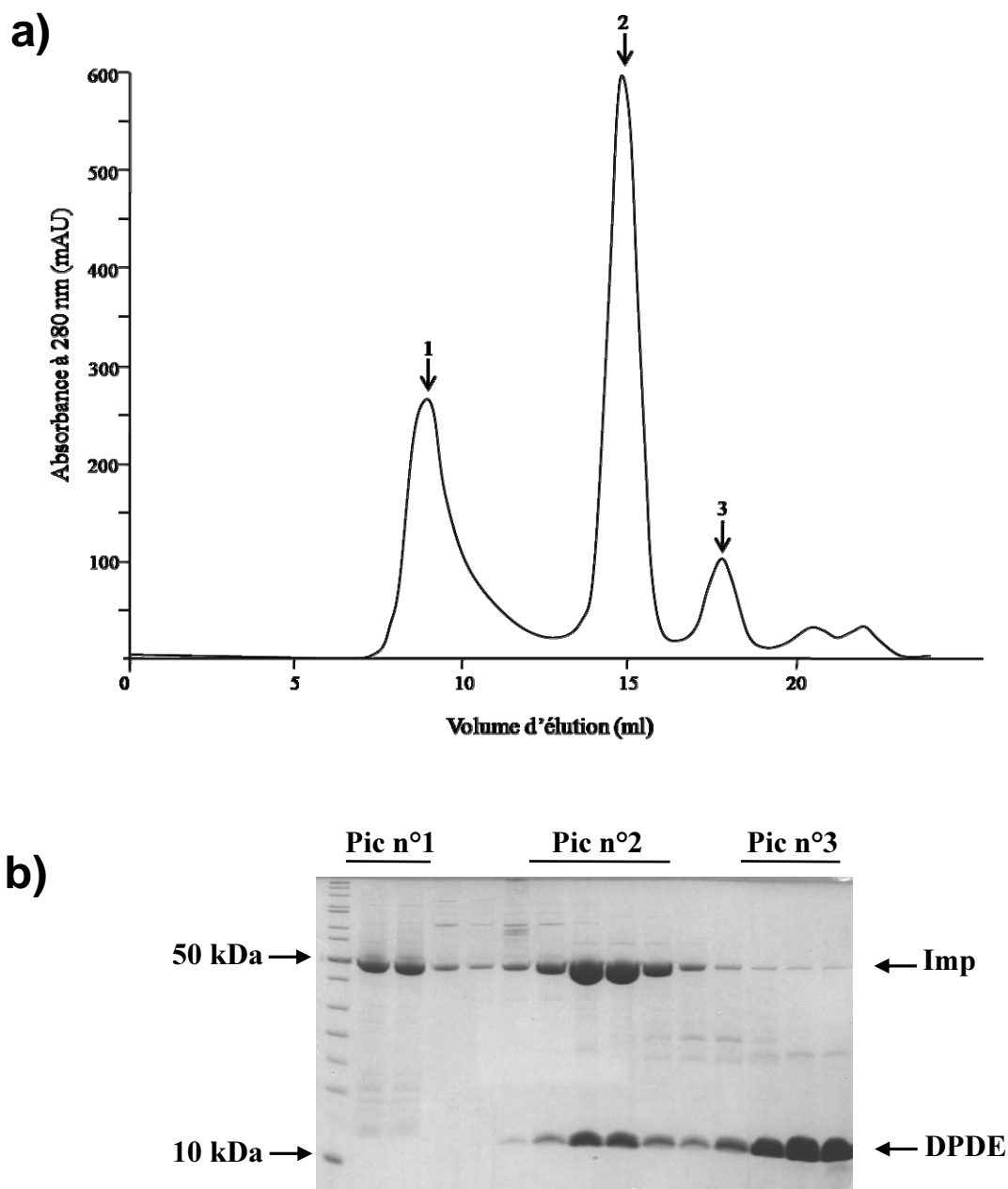
#### 3.1. Affinité de PB2 pour l'importine $\alpha$ 5 humaine

La localisation nucléaire de PB2 est censée dépendre d'un motif interne (acides aminés 449-495) et d'un motif NLS monopartite, 736-KRKR-739 (Mukaigawa & Nayak, 1991), présent sur la partie C-terminale de PB2 (**figure 23**). Du fait de cette séquence NLS présente sur DPDE nous avons voulu tester son interaction avec son facteur d'import nucléaire, l'importine  $\alpha$ 5. Nous avons cloné l'importine  $\alpha$ 5 en délétant les régions flexibles qui pourraient par la suite gêner le processus de cristallogénèse. Ainsi, notre gène de l'importine  $\alpha$ 5 ne présente que la séquence codant pour le domaine structuré de la protéine (acides aminés 66-512). Les séquences codant pour le domaine IBB (en N-ter) et l'extrémité C-terminale (26 acides aminés) ont été délétées. Nous avons tout d'abord, exprimé indépendamment DPDE ainsi que l'importine  $\alpha$ 5 humaine (Imp  $\alpha$ 5), toutes deux avec une étiquette polyhistidine située à l'extrémité N-terminale. Les deux protéines sont ainsi purifiées par chromatographie d'affinité (sur résine de cobalte) (**figure 27a et c**). Avant de former le complexe entre ces deux molécules, les étiquettes polyhistidines des deux protéines sont enlevées grâce à l'action de la protéase TEV Tobacco Etch Virus (TEV) (**figure 27b**). Cette protéase possède l'avantage d'être très spécifique pour son site de coupure (ENLYFQG). Afin de former le complexe, DPDE en excès est mise en présence de l'Imp  $\alpha$ 5 (3:1 ; DPDE:Imp  $\alpha$ 5) (**figure 27**). Le mélange ainsi formé est ensuite purifié par chromatographie d'exclusion de taille. La colonne utilisée consiste en une résine qui sépare les molécules en fonction de leur taille et de leur forme, c'est-à-dire selon leur masse moléculaire apparente. La taille des pores de la résine est connue et donne la limite d'exclusion des molécules. La Superdex 200 (S200) (Pharmacia/Amersham) sépare des molécules dont le poids est compris entre 10 et 600 kDa. Après cette étape de purification par chromatographie d'exclusion du mélange entre DPDE et l'importine  $\alpha$ 5, nous obtenons 3 pics (**figure 28a**). Le premier correspond aux agrégats présents, le deuxième, élué à 14,9 ml, correspond au complexe attendu et le troisième pic, à 17,9 ml représente l'excès de DPDE (**figure 28a**). Selon la courbe standard (non présentée), un volume de 14,9 ml correspondrait à une molécule globulaire d'environ 70 kDa, donc légèrement supérieure à la masse de l'importine  $\alpha$ 5 et de DPDE associées (environ 60 kDa). Ce décalage peut être expliqué par la forme allongée de l'importine  $\alpha$ 5 complexée à DPDE. L'analyse du gel dénaturant (SDS-PAGE) confirme la présence des deux protéines dans les fractions issues du pic n°2 ainsi que l'excédant de DPDE non complexée (**figure 28b**). Il n'est

pas évident de conclure quant à la stœchiométrie du complexe mais il est tout à fait probable que celle-ci soit de 1:1. Nous avons ainsi montré que la partie C-terminale de PB2 interagit avec l'importine  $\alpha 5$  et forme un complexe stable *in vitro*.



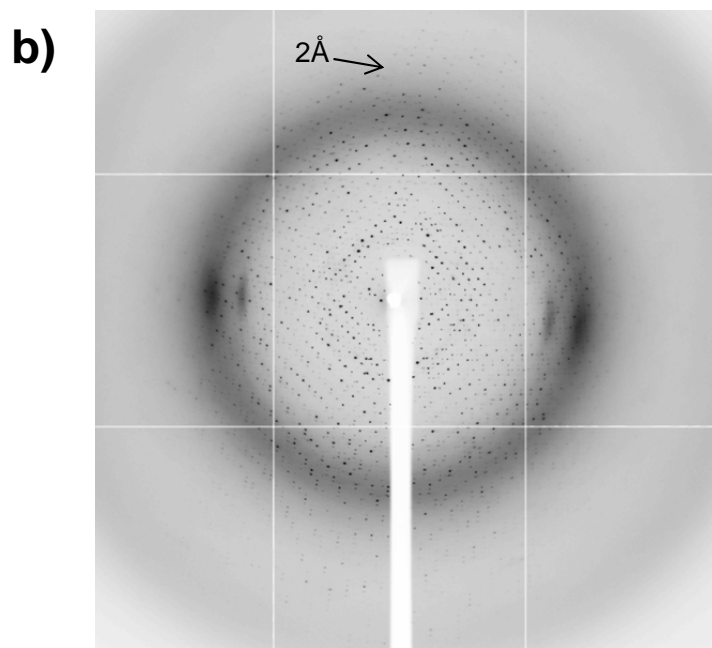
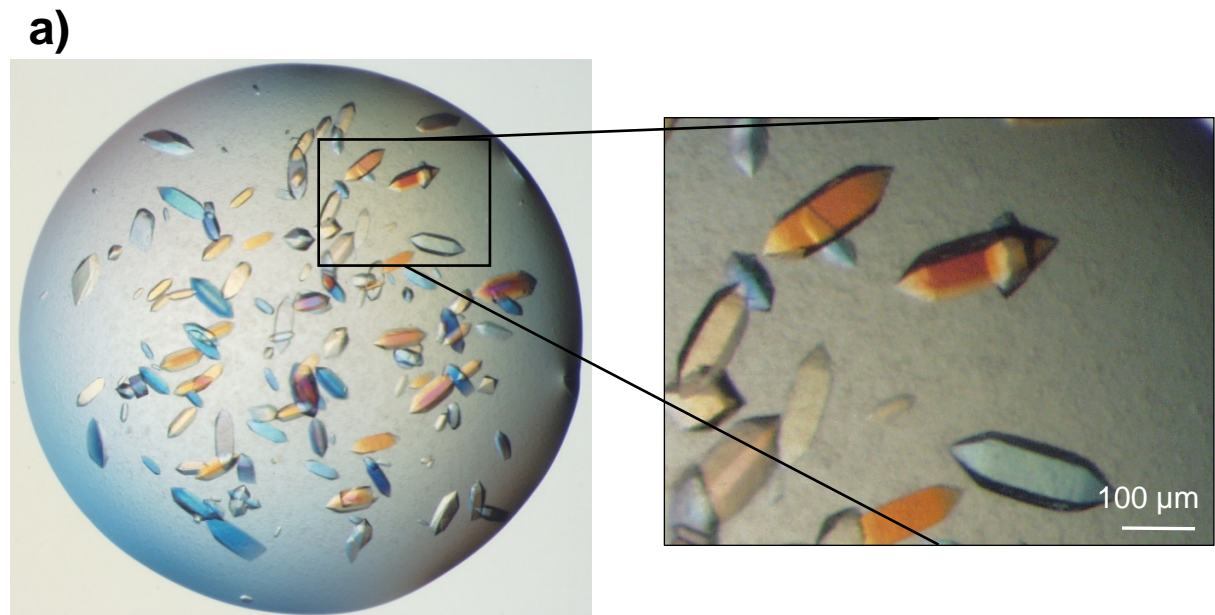
**Figure 27. Stratégie de formation du complexe entre la partie C-terminale de la sous-unité PB2 de la polymérase (DPDE : acides aminés 678 à 759) et l'importine  $\alpha 5$  humaine.** Les étiquettes polyhistidines de l'imp  $\alpha 5$  et DPDE sont coupées grâce à l'action de la proéase TEV. La protéase TEV et les étiquettes polyhistidines sont éliminées par chromatographie par chromatographie d'affinité. Enfin, le complexe est formé en ajoutant la DPDE trois fois en excès à l'Imp $\alpha 5$ . M : Marquer moléculaire, ES : Extrait soluble, Ef : Effluent, W : lavage.



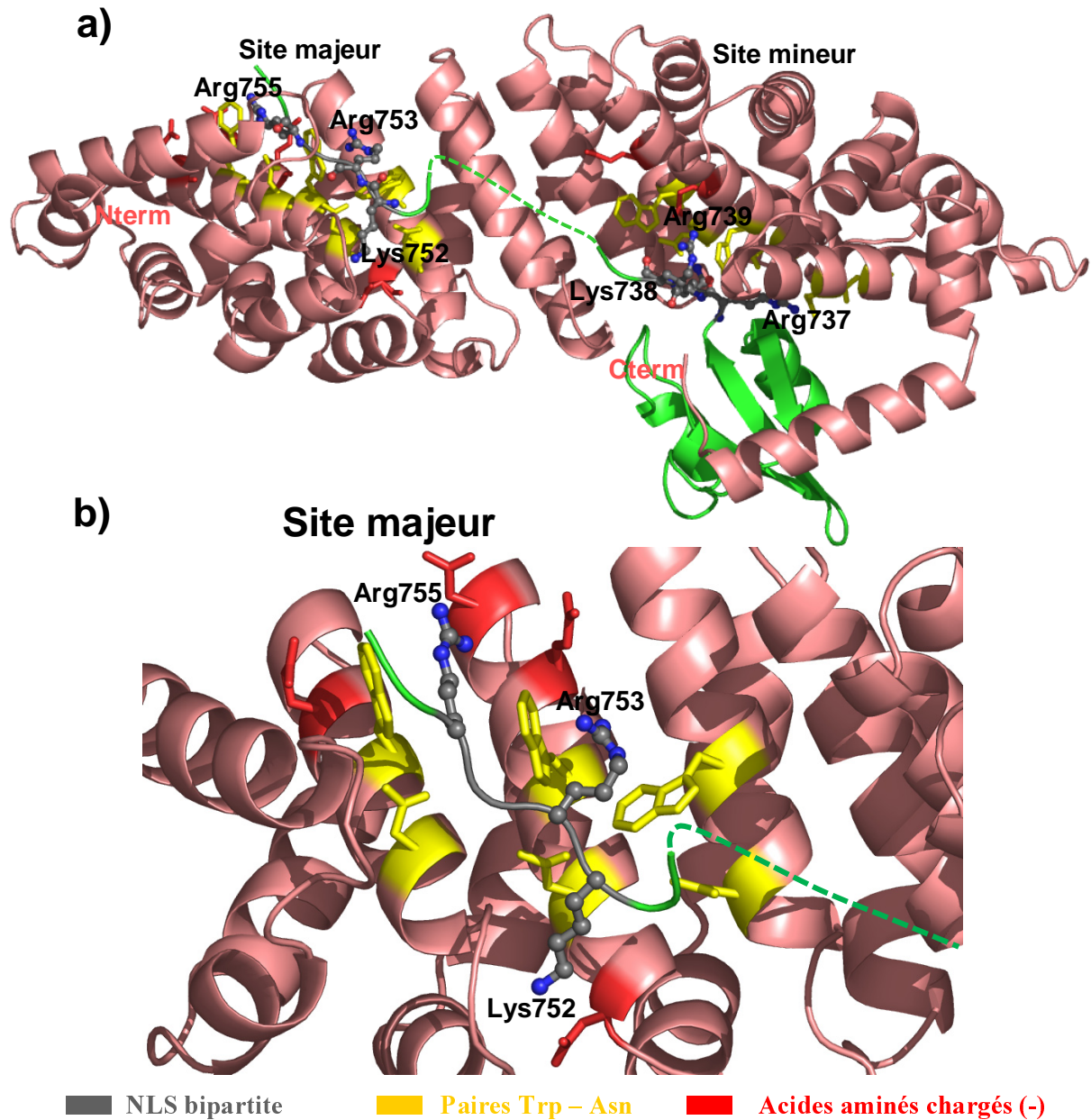
**Figure 28. Interaction du domaine C-terminal de la sous unité PB2 de la polymérase (DPDE) avec l'importine  $\alpha 5$  humaine.** La partie C-terminale de PB2 (DPDE : acides aminés 678-759) et l'importine  $\alpha 5$  (Imp  $\alpha 5$  : acides aminés 66-512) sont rassemblées selon un rapport DPDE / Imp  $\alpha 5$  de 5/1 (mole/mole). Le complexe est ensuite purifié par chromatographie d'exclusion de taille. **a)** Profil d'élution du complexe sur colonne Superdex 200 (GE Healthcare). Les flèches indiquent le volume mort à 8,9 ml (1), le complexe DPDE / Imp  $\alpha 5$  à 14,9 ml (2) et l'excès de DPDE à 17,9 ml (3). **b)** Analyse par électrophorèse sur gel de polyacrylamide (SDS-PAGE 12% Tris-Tricine) des pics issus du profil a.

### 3.2. Structure cristallographique du complexe entre le domaine C-terminal de PB2 et l'importine $\alpha 5$ humaine

Nous disposons maintenant d'un complexe stable et de qualité. Afin de le cristalliser, les fractions correspondant au complexe, issues de la purification par chromatographie d'exclusion (**figure 28a pic n°2**), sont rassemblées puis concentrées à 20 mg/ml. La plateforme de cristallisation à haut débit Tecan développée par l'EMBL-Grenoble permet de réaliser des nanogouttes de 0.2  $\mu$ l (0.1  $\mu$ l de précipitant + 0.1  $\mu$ l de solution protéique). Nous avons ainsi obtenu des cristaux dans le tampon : 20 mM  $\text{CaCl}_2$ , 0.1 M AcNa pH 4.6, 16% MPD (2-methyl-2, 4-pentandiol). Ces cristaux peuvent être reproduits manuellement (**figure 29a**). Ces cristaux dont les arrêtes les plus imposantes atteignent environ une centaine de micromètres ont été cryo-protégés puis diffractés à l'ESRF (European Synchrotron Radiation Facility). Un jeu de données a pu être collecté sur la ligne ID23-1 et la structure du complexe a été résolue à 2.2 Å (**figure 30**). Le groupe d'espace déterminé est  $P2_12_12_1$  et deux complexes sont présents à l'intérieur de l'unité asymétrique. Le remplacement moléculaire a été employé pour résoudre la structure en utilisant le modèle PDB (Protein Data Bank) de l'importine  $\alpha$  de souris (1Q1S). La structure du complexe résolue, le mécanisme d'interaction entre DPDE et l'importine  $\alpha 5$  nous apparaît. La reconnaissance se fait bien par la région C-terminale de DPDE, cependant le motif NLS observé n'est pas monopartite comme on aurait pu s'y attendre. Il est bipartite (737-RKR<sub>12</sub>KRIR-755) (**en gris sur la figure 30a**) et écarte ainsi le présumé NLS monopartite (736-KRKR-739) décrit par [Mukaigawa & Nayak, \(1991\)](#). Trois acides aminés basiques à la fois sur la première portion (NLS1 : 737-RKR-739) et sur la deuxième portion (NLS2 : 752-KRIR-755) du NLS bipartite interagissent avec deux régions distinctes, situées à l'intérieur du sillon de l'importine  $\alpha 5$ . Comme c'est également le cas des importines de levure et de souris, l'importine  $\alpha 5$  humaine présente deux sites de fixation pour les NLS. Le plus grand de deux sites nommé site majeur, est présent dans la région N-terminale de l'importine (ARM 2, 3, 4). Le deuxième site qui est plus petit, est nommé site mineur et est situé au niveau des ARM 7 et 8 (région C-terminale). Cette forte interaction de la région basique du NLS avec des acides aminés chargés négativement, de types acide aspartique et acide glutamique (**en rouge sur la figure 30b**), est stabilisée par la présence de liaisons hydrogènes additionnelles notamment entre les acides aminés de types tryptophane et asparagine (**en jaune sur la figure 30b**) et le NLS de DPDE. Probablement dû à leur flexibilité, les 12 acides aminés faisant le lien dans le NLS bipartite ne sont pas visibles dans la densité électronique.



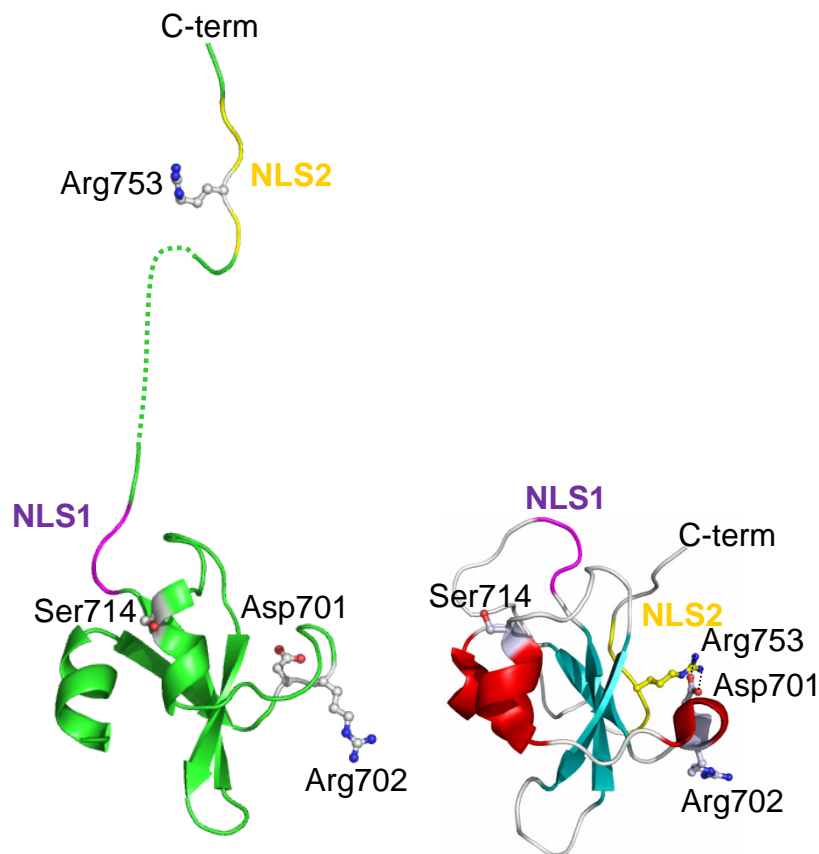
**Figure 29. a) Cristaux d'importine  $\alpha 5$  humaine complexée au domaine C-terminal de PB2.** La concentration du complexe dans la goutte est de  $10 \text{ mg.ml}^{-1}$  et les cristaux ont poussé par diffusion de vapeur dans la condition suivante :  $20 \text{ mM CaCl}_2$ ,  $100 \text{ mM AcNa}$  pH 4.6, avec comme agent précipitant 16% MPD (2-méthyl-2, 4-pentandiol). **b) Spectre de diffraction** du complexe à  $2.2 \text{ \AA}$  de résolution collecté sur la ligne ID23-1 de l'installation européenne de rayonnement synchrotron (ESRF-Grenoble).



**Figure 30. Structure cristallographique aux rayons X, à 2.2 Å de résolution, de l'importine  $\alpha 5$  humaine complexée avec la partie C-terminale de la sous unité PB2 de la polymérase du virus de la grippe (Tarendeau *et al.*, 2007). a) La représentation ruban montre la partie C-terminale de PB2, DPDE (en vert), associée à l'importine  $\alpha 5$  (rose) constituée d'une répétition de 10 armadillos, chacun étant composé d'une succession de 3 hélices  $\alpha$ . Le NLS bipartite (RKR<sub>12</sub>KRIR, gris) en C-ter de DPDE se fixe à l'intérieur du sillon de l'importine  $\alpha 5$ . b) Grossissement de la région de fixation des acides aminés 752-KRIR-755 du NLS à l'importine  $\alpha 5$  (site majeur). Les acides aminés constitutifs du NLS présentent une forte interaction avec les acides aminés chargés négativement (en rouge) de type acide aspartique et glutamique. L'ensemble est stabilisé par la présence de liaison hydrogène entre les acides aminés de type tryptophanes, asparagines (en jaune) avec le NLS de DPDE.**



Si l'on compare maintenant les structures RMN et cristallographique de DPDE, nous voyons que la protéine libre en solution est très similaire à sa structure cristalline dans son état lié à l'importine  $\alpha 5$  (la déviation r.m.s. des C $\alpha$  est de 1.87 Å pour 41 acides aminés alignés) (**figure 31**). Cependant la structure aux rayons X montre davantage de structures secondaires. La comparaison révèle comment la région C-terminale se délie et permet ainsi la fixation du NLS bipartite (**en violet et en jaune sur la figure 31**) sur son facteur d'import. La structure RMN, de DPDE en solution, montre que le groupement carboxyl de l'Asp701 forme un pont salin avec l'Arg753 appartenant au motif NLS bipartite. De ce fait, le domaine C-terminal est ramené vers le corps de la protéine. La fixation à l'importine  $\alpha 5$  nécessite que cette interaction soit cassée comme le montre la structure de DPDE lorsqu'elle est impliquée dans le complexe (**figure 31 à gauche**). Notons que la mutation D701N romprait le pont salin avec l'Arg753 du NLS et pourrait ainsi modifier l'énergie de fixation à l'importine  $\alpha 5$ .

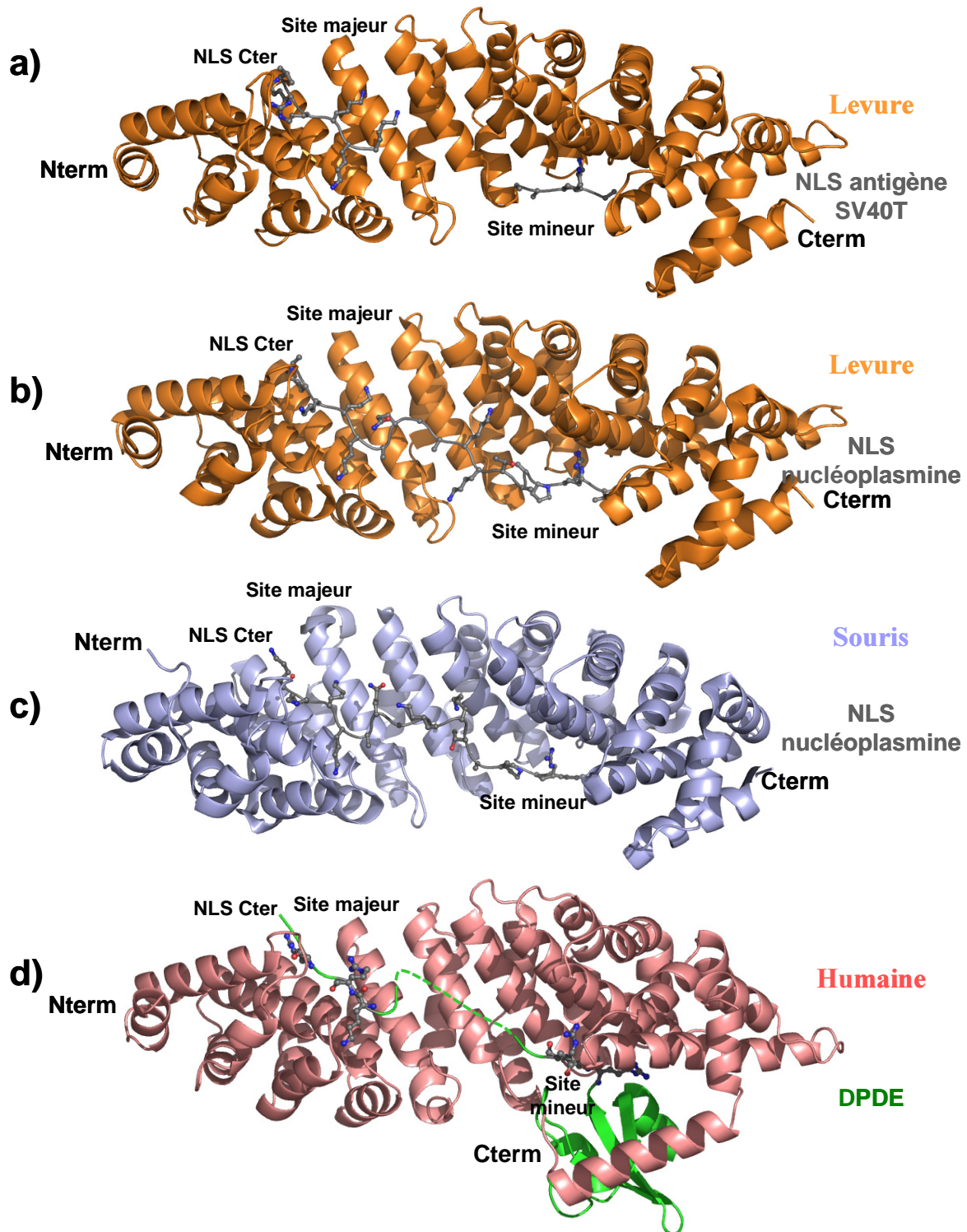


**Figure 31. Comparaison de la structure aux rayons X de DPDE en complexe avec l'imp  $\alpha 5$  (à gauche) ou dans son état libre en solution (à droite) obtenue par RMN (Tarendeau *et al.*, 2007). Les acides aminés constitutifs du NLS (en violet et jaune) sont repliés sur la molécule lorsqu'elle est dans son état libre en solution. Sont également indiqués les acides aminés Asp701, Arg702 et Ser714 qui sont impliqués dans la transmission du virus inter-espèces ainsi que le pont salin entre l'Arg753 et l'Asp701.**

Différentes séquences de NLS (peptides de 5 à 12 acides aminés), monopartite ou bipartite ont déjà été utilisées (en *soaking* ou en co-cristallisation) pour décrire les mécanismes d'interaction. L'importance des acides aminés basiques dans le transport nucléaire a été établie pour la première fois lors de la caractérisation du NLS de l'antigène T de SV40 (*Virus Simien 40*) (Kalderon *et al.*, 1984 ; Lanford & Butel, 1984). Les déterminants spécifiques de reconnaissance du NLS monopartite ont été élucidés quand la structure de l'importine de levure en complexe avec le peptide NLS de l'antigène T de SV40 a été résolue (Conti *et al.*, 1998). Ce peptide, 126-PKKKRKV-132, se fixe à deux sites à l'intérieur du sillon de l'importine de levure (**figure 32a**) et à l'époque suggérait un modèle pour la reconnaissance du NLS bipartite.

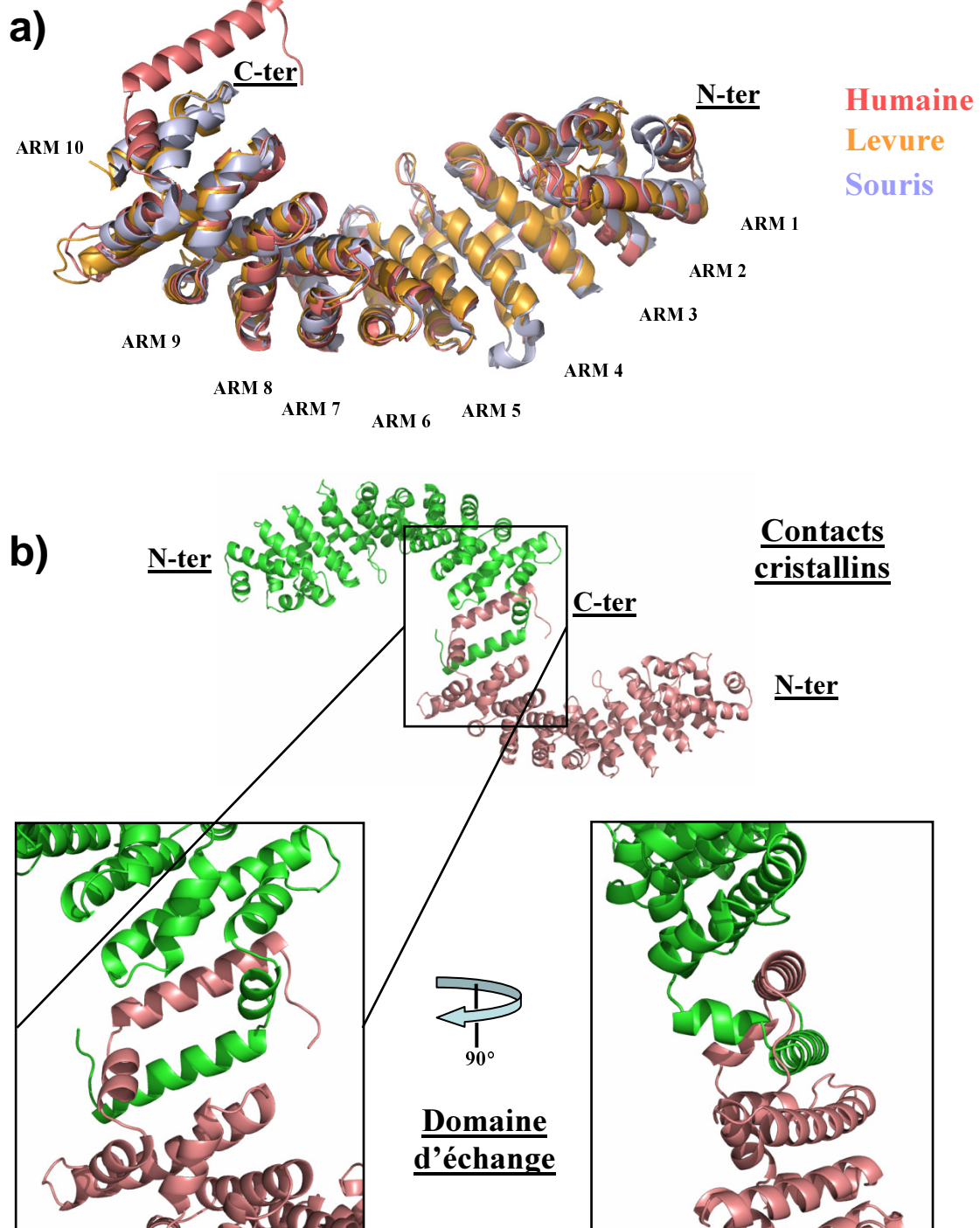
La nucléoplasmine impliquée dans l'assemblage de la chromatine dans des œufs de xénope est également une protéine de choix pour étudier l'import nucléaire. En effet, cette protéine présente également une séquence NLS mais de taille supérieure à celle du NLS de l'antigène SV40 T (Dingwall *et al.*, 1988). Plusieurs fois résolu en complexe avec l'importine  $\alpha$  de levure (Conti & Kuriyan, 2000) ou encore avec celle de la souris (Fontes *et al.*, 2000 & 2003), ce peptide présente un mode d'interaction bipartite, similaire à celui de DPDE (**figure 32b, c et d**). Le NLS de nucléoplasmine, 155-KR<sub>PAATKKAGQA</sub>KKKK-170, se fixe simultanément aux deux sites de fixation situés sur l'importine  $\alpha$  (**figure 32b et c**). Comme pour l'interaction avec DPDE, le site mineur de l'importine, se fixe spécifiquement aux acides aminés situés en N-terminal du NLS (RKR pour DPDE et KR pour le peptide NLS de nucléoplasmine). Et le groupement d'acides aminés basiques, en C-terminal des NLS, se fixe au site majeur (**figure 32b, c et d**). Les séquences, du NLS de DPDE ou de la nucléoplasmine, faisant le lien entre les deux régions d'acides aminés basiques, sont moins ordonnées et engendrent peu de liaisons avec l'importine. Cette structure d'importine  $\alpha 5$  présentée ici, est la première de la famille importine  $\alpha$  humaine. De plus, c'est également la première structure de complexe d'une importine avec un domaine replié, si l'on se réfère à d'autres structures d'importines contenant simplement un peptide NLS comme substrat.

La structure de l'importine  $\alpha 5$  comprend une répétition de 10 armadillos et est similaire à celles des importines de la levure et de la souris (**figure 33a**). Cependant, nous observons un arrangement différent notable au niveau de la dernière hélice  $\alpha$  située en C-terminal. La disposition périodique et ordonnée des molécules constitue la structure cristalline. Suite à l'analyse des contacts entre les différentes molécules du cristal, nous avons observé le phénomène d'échange de domaines ("domain swapping").



**Figure 32. Reconnaissance de séquences de localisation nucléaire (NLS) par différentes importines  $\alpha$ .** **a)** Structure de l'importine  $\alpha$  de levure avec le peptide NLS de l'antigène de SV40 fixé à deux sites (Conti *et al.*, 1998). **b)** Structure du complexe importine  $\alpha$  de la levure / peptide NLS bipartite de nucléocytoplasmine (Conti & Kuriyan, 2000). **c)** Structure du complexe importine  $\alpha$  de souris / peptide NLS bipartite de nucléocytoplasmine (Fontes *et al.*, 2003). **d)** Structure du complexe importine  $\alpha 5$  humaine / domaine C-terminal de PB2 du virus de la grippe (DPDE) (Tarendeau *et al.*, 2007).

Superposition de  
différentes importines



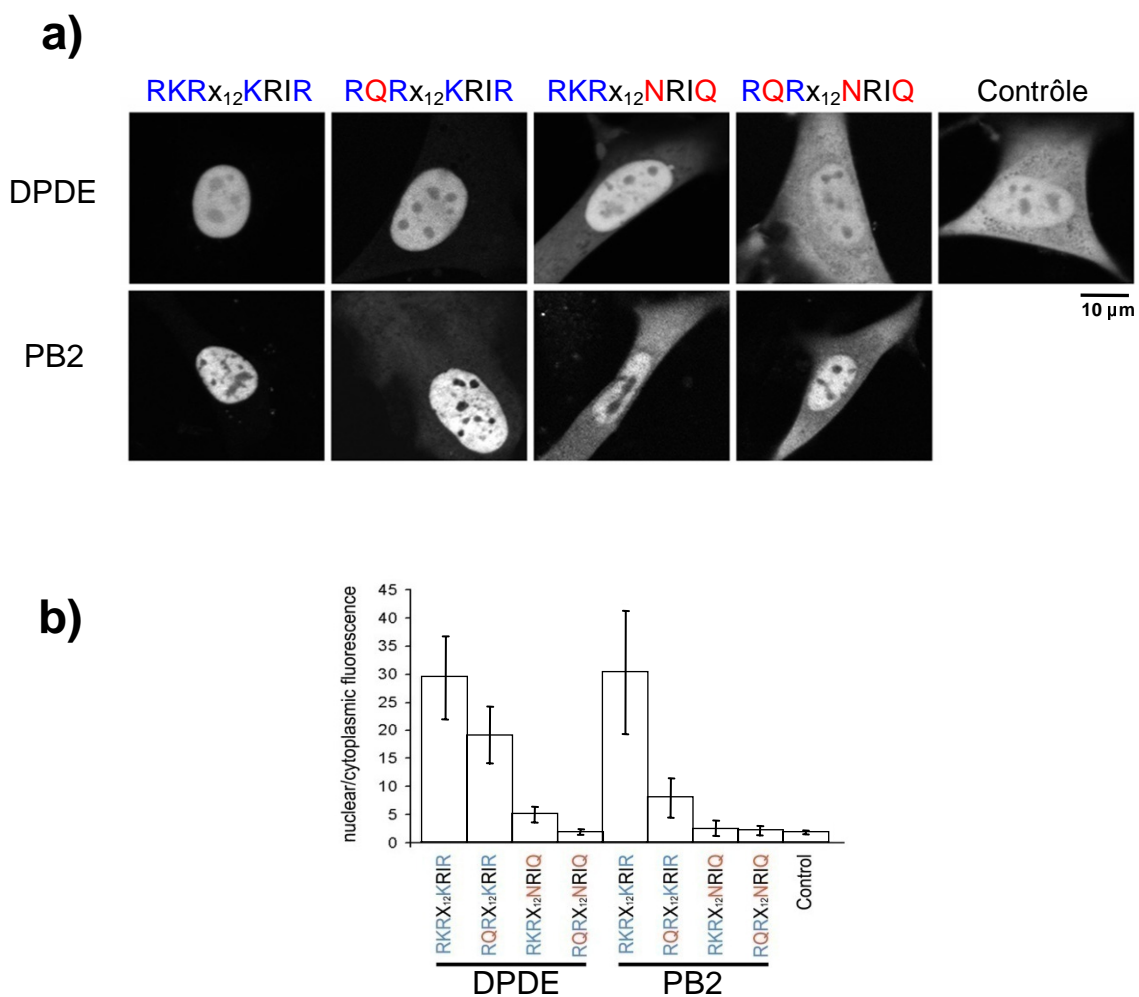
**Figure 33. L'importine  $\alpha 5$  humaine.** **a)** Superposition des structures d'importines de la souris et de la levure avec l'importine  $\alpha 5$  humaine. **b)** Contacts cristallins entre 2 molécules d'importine  $\alpha 5$  humaine (haut) et visualisation du domaine d'échange (swapping domain) entre les hélices  $\alpha$  C-terminales de 2 molécules d'importine  $\alpha 5$  (bas).

C'est un mécanisme où deux molécules voire plus, vont former des dimères ou des oligomères grâce à l'échange d'éléments de structures identiques (Liu & Eisenberg, 2002). Si le monomère et le dimère d'une molécule existent sous forme stable, dans laquelle le dimère adopte une conformation de *domain swapping* et que le monomère adopte une conformation proche, alors cette molécule peut être considérée comme un bon exemple de *domain swapping*. Certaines protéines dimérisent selon un mécanisme apparent à celui de *domain swapping* sans pour autant qu'un monomère de conformation proche ne soit connu. Si ces protéines ont des homologues connus pour être de proches monomères, alors ces oligomères sont considérés comme "quasidomain swapped". Ainsi, pour l'importine  $\alpha 5$  humaine, l'hélice  $\alpha$  C-terminale de la molécule rose est insérée dans la molécule verte avec un changement de conformation, et vice et versa pour l'hélice  $\alpha$  C-terminale de la molécule verte (**figure 33b**). De plus, les homologues de l'importine de levure et de souris étant connus, le terme *quasidomain swapped* peut être utilisé pour l'importine  $\alpha 5$  humaine. Plus d'une cinquantaine de protéines ont déjà été rapportées dans la bibliographie pour répondre à ce mécanisme de *domain swapping*. Ces protéines sont souvent impliquées dans différents phénomènes biologiques. Par exemple, ce mécanisme a été proposé pour la formation de fibre amyloïde à partir du prion humain (Cohen & Prusiner, 1998). Cependant, sur l'ensemble des protéines répondant à ce mécanisme, toutes n'ont pas une pertinence biologique et le phénomène pourrait aussi être dû à un artefact de cristallisation. Les importines de la levure et de la souris présentent un arrangement quasi identique. De ce fait, il semble probable que la conformation de l'importine  $\alpha 5$  humaine soit également le fruit d'un artefact de cristallisation.

#### 4. Localisation in vivo de PB2

Afin de mettre davantage en évidence ce mécanisme révisé d'import nucléaire de PB2, des études *in vivo* de localisation cellulaire ont été menées en collaboration avec l'équipe du Dr J. Ellenberg (EMBL-Heidelberg). Ces expériences ont été réalisées à l'aide de la protéine de fusion DPDE-GFP (green fluorescent protein) qui est retrouvée quasi exclusivement dans le noyau ( $29 \pm 8$  fois plus), alors que le contrôle, GFP seule, reste équilibré entre le noyau et le cytoplasme (**figure 34a et b**). Ainsi le NLS bipartite de la protéine de fusion DPDE-GFP a été modifié (**figure 34a et b**). La mutation de la première partie du NLS (NLS1), K738Q, diminue de façon modérée mais significative l'import nucléaire. Ce qui est en accord avec les études réalisées sur la protéine entière (Mukaigawa & Nayak, 1991). La double mutation de la

deuxième partie du NLS (NLS2), K752N et R755Q, diminue clairement l'import nucléaire. La combinaison des mutations du NLS1 et du NLS2 abolie l'import nucléaire dont le taux de transport se retrouve comparable à celui de l'import de la GFP seule. Les résultats obtenus avec la protéine PB2 entière sont similaires (**figure 34a et b**). Ainsi, la séquence C-terminale de PB2, 737-RKR<sub>12</sub>KRIR-755, est bien confirmée pour porter la fonction de NLS bipartite classique pour l'import nucléaire. Bien que ce NLS soit fonctionnel sur le domaine C-terminal isolé et sur PB2 entière, il ne peut être exclu la possibilité que d'autres séquences soient impliquées dans la faible localisation nucléaire observée chez PB2 entière lorsque le NLS1 et le NLS2 sont mutés.

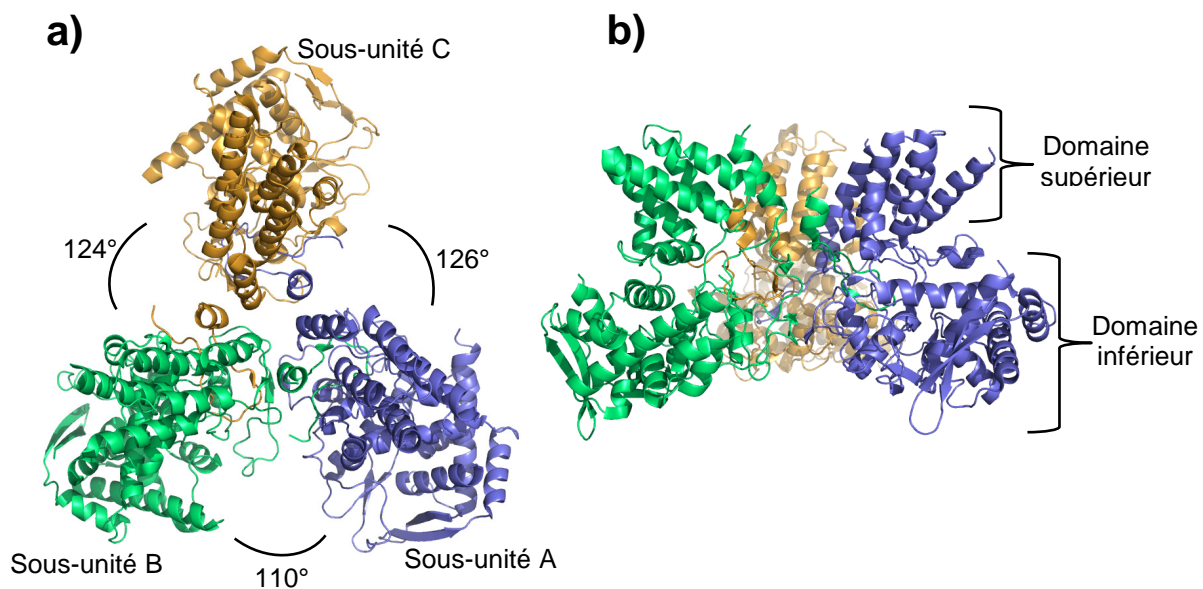


**Figure 34. L'import nucléaire de la partie C-terminale de PB2 et de la protéine entière est dirigé par un NLS bipartite (Tarendeau *et al.*, 2007). a)** Localisation subcellulaire de la partie C-terminale de PB2 (DPDE) ou de la protéine entière (PB2), fusionnée à la GFP dans des cellules NIH 3T3. Les constructions possèdent soit le NLS sauvage (737-RKR<sub>12</sub>KRIR-755) soit des points de mutations représentés en rouge. Les images montrent des sections confocales de 2 μm de la fluorescence de la GFP. La barre représente 10 μm. **b)** Rapport de la concentration entre le noyau et le cytoplasme de la protéine fusionnée à la GFP.

## **Chapitre 2**

—

### **Polymérisation de la nucléoprotéine et activité de fixation à l'ARN**



**Figure 35. Structure cristallographique de la nucléoprotéine du virus de la grippe souche A/WSN/33 (code pdb : 2IQH) (Ye *et al.*, 2006). a) Le trimère de la nucléoprotéine est observé selon l'axe 3 de symétrie non cristallographique, avec pour chaque sous-unité une couleur différente. Les angles de rotations reliant les 3 sous-unités sont indiqués. b) Vue de coté du trimère de la nucléoprotéine avec les couleurs identiques à a).**



## Chapitre 2

### Polymérisation de la nucléoprotéine et activité de fixation à l'ARN

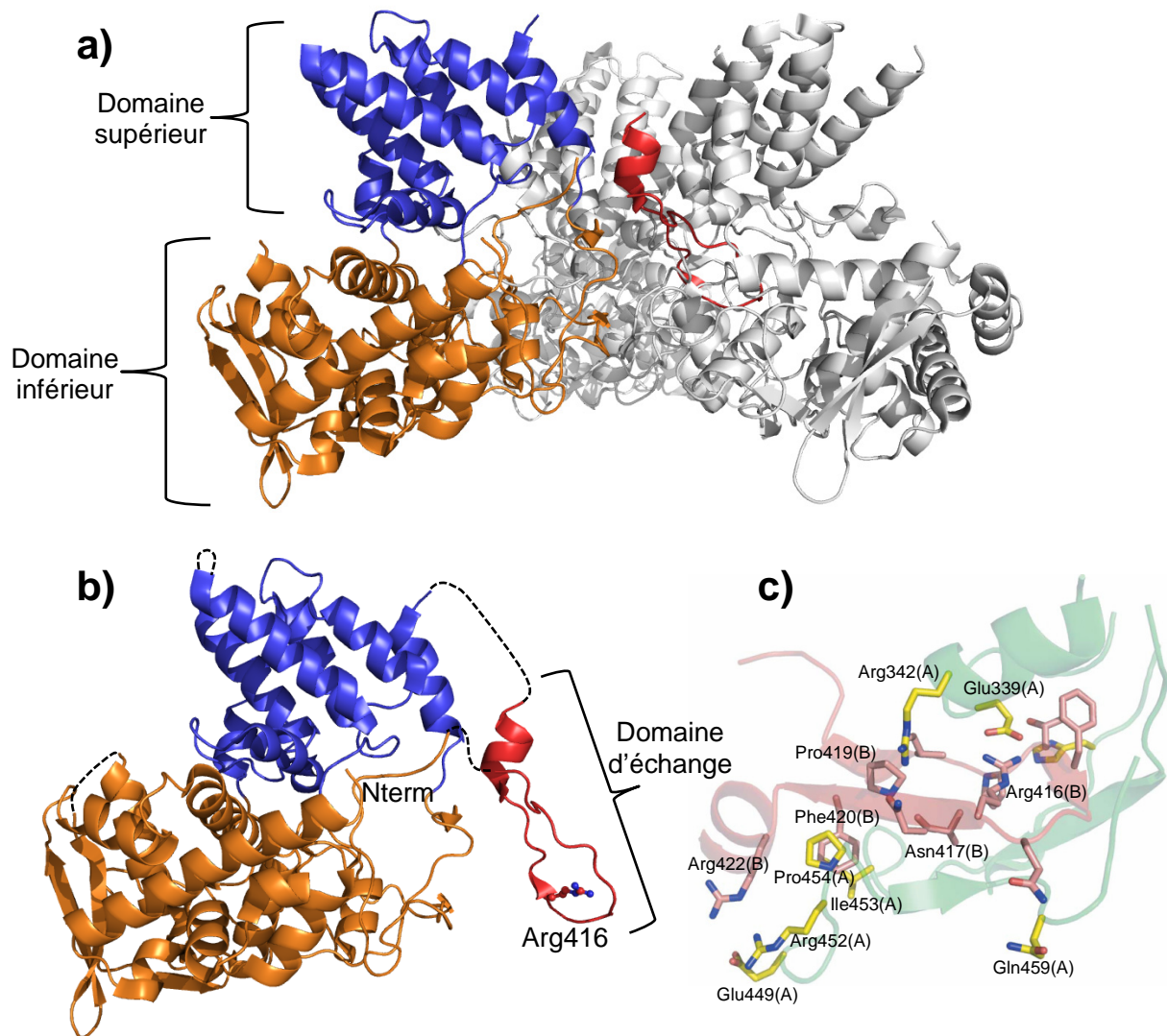
#### 1. Introduction

La nucléoprotéine du virus de la grippe est une protéine de 56 kDa, codée par le segment n°5 du génome viral. Cette protéine de 498 acides aminés est le principal constituant des RNPv. La séquence de cette protéine est très conservée dans les souches de virus grippaux de type A, B et C (Shu *et al.*, 1993). Les fonctions biologiques de la NP sont multiples. Tout d'abord, avec les trois sous-unités de la polymérase virale, la NP est indispensable aux processus de transcription et de réplication du virus (Huang *et al.*, 1990 ; Kimura *et al.*, 1992 ; de la Luna *et al.*, 1993). De plus, c'est une protéine chargée positivement qui est riche en acides aminés basiques de type arginine, glycine et sérine. Ainsi, elle présente des domaines spécifiques pour interagir avec l'ARN (Elton *et al.*, 1999a et 1999b). En effet, elle le recouvre entièrement à raison de 1 NP pour 24 bases (Ortega *et al.*, 2000) et protège ainsi le génome viral des nucléases cellulaires. Caractérisée dans le passé par notre laboratoire, la liaison de la NP à l'ARN révéla que la protéine se fixe sur le squelette de l'acide nucléique sans spécificité de séquence et expose les bases vers le solvant (Baudin *et al.*, 1994). De ce fait, l'ARN polymérase virale pourrait transcrire l'ARNv sans pour autant que la NP ne s'en dissocie. Par le passé, deux études ont permis d'identifier des séquences minimales, situées sur le premier tiers de la nucléoprotéine, capables d'interagir avec l'ARN (Albo *et al.*, 1995 ; Kobayashi *et al.*, 1994). Cependant, ces fragments tronqués de la nucléoprotéine se fixent à l'ARN avec une affinité moins élevée que la protéine entière, indiquant la nécessité d'autres régions pour une activité de fixation complète (Albo *et al.*, 1995). Plus récemment des travaux ont contredit Albo *et al.*, (1995) et Kobayashi *et al.*, (1994), au profit d'une distribution d'acides aminés impliqués dans la fixation à l'ARN, situés tout au long de la séquence de la NP (Elton *et al.*, 1999b).

#### 1.1. Structure de la NP du virus de la grippe

Récemment, la structure de la nucléoprotéine résolue à 3,2 Å a été publiée (**figure 35**) (Ye *et al.*, 2006). Le fait que cette protéine présente cette propension à polymériser rend

difficiles les études structurales et particulièrement les essais de cristallogénèse. Néanmoins c'est sous forme d'un trimère que [Ye et al., \(2006\)](#) ont réussi à cristalliser la nucléoprotéine du virus de la grippe (**figure 35a**). Bien que cristallisée sans ARN, la structure dévoile un sillon à l'intérieur duquel pourrait probablement se loger l'acide nucléique (**figure 35b**). Cette structure est majoritairement composée d'hélices  $\alpha$  et chaque monomère est incurvé selon une forme de "mâchoire". L'ensemble de la structure peut être divisé en deux domaines, l'un inférieur, l'autre supérieur (**figure 35b**). Ces deux parties sont composées par des régions polypeptidiques non contiguës : le domaine supérieur comprend les acides aminés 150-272 et 438-452, le domaine inférieur consiste en trois segments composés des acides aminés 21-149, 273-396 et 453-489. Ainsi, ce repliement et cette topologie de la NP du virus de la grippe sont différents de ceux que l'on peut observer chez des nucléoprotéines appartenant à d'autres virus à ARN négatifs, où les deux domaines sont formés par des séquences polypeptidiques colinéaires. En effet, la structure cristalline de la NP du bornavirus a déjà été rapportée ([Rudolph et al., 2003](#)). Et plus récemment, ont été publiées les structures de la NP de la rage ([Albertini et al., 2006](#)) ainsi que celle du virus de la stomatite vésiculaire (VSV) ([Green et al., 2006](#)). Les travaux de [Ye et al., \(2006\)](#), sur la structure de la NP du virus de la grippe, ont mis en évidence la présence d'un domaine d'échange (acides aminés 402-428), située au dos de la protéine (**figure 36**). La conformation de ce domaine est relativement étendue et interagit avec une molécule voisine du trimère où il est inséré. Différents types d'interaction se manifestent entre le domaine d'échange et la molécule voisine (**figure 36b et c**). Elles sont de types intermoléculaires entre feuillets  $\beta$  hydrophobes. De plus, nous pouvons remarquer la présence de ponts salins, notamment entre un acide glutamique (Glu339) et une arginine (Arg416). Les mutations R416A, E339A ou encore la délétion des acides aminés 402-428 (domaine d'échange) évaluées par chromatographie d'exclusion de taille, ont permis d'obtenir des mutants de la NP de la grippe sous forme monomérique ([Ye et al., 2006](#)). Nous avons à notre tour réalisé ces différentes mutations sur la nucléoprotéine du virus de la grippe (souche A/PR/8/34) et parmi ces trois mutants, le R416A fut le candidat retenu pour sa stabilité et son taux d'expression. Pour une plus grande clarté, nous nommerons ce mutant R416A, NPmono.



**Figure 36. Interactions entre sous-unités générées par le domaine d'échange (d'après Ye *et al.*, 2006).** Trimère (a) et sous-unité B (b) du trimère de la nucléoprotéine. La tête est représentée en bleu, le corps en orange et le domaine d'échange en rouge. L'arginine 416 est indiquée. c) Interaction entre les acides aminés de la sous unité A en vert et B en rouge. Sont représentés les acides aminés impliqués dans les interactions importantes.

Afin de caractériser la NP en terme d'oligomérisation mais aussi d'activité de fixation à l'ARN, nous avons réalisé nos expériences à la fois sur la NP virale (NPv), la NP sauvage recombinante (NPrec) et sur le mutant R416A (NPmono). Il est à noter que cette mutation ne touche pas la plateforme de fixation à l'ARN puisqu'elle se situe au dos de la protéine (**figure 36b**). Nous avons ainsi travaillé sur la caractérisation de ce mutant R416A, que nous avons comparé à la protéine sauvage ainsi qu'à la NP directement extraite à partir du virus. Enfin par des expériences de *filter binding assay* (FBA), nous avons tenté de faire un lien entre l'état d'oligomérisation de la nucléoprotéine et son affinité pour l'ARN. Les trois échantillons de NP correspondent toutes à la souche de virus A/PR/8/34.

## 2. Caractérisation des différentes formes de nucléoprotéines : NP virale (NPv), NP recombinante (NPrec) et NPmono

### 2.1. Purification de NPv, de NPrec et du mutant NPmono

Nous avons employé deux modes de purification distincts. Le premier pour purifier la nucléoprotéine directement à partir du virus et le second, pour les nucléoprotéines recombinantes exprimées chez la bactérie *E. coli*. Concernant la nucléoprotéine virale (NPv), elle est purifiée à partir d'une suspension de virus A/PR/8/34 produits sur des embryons de poulets (Baudin *et al.*, 1994). Cette préparation de NPv nécessite tout d'abord une étape de purification des RNPv. Brièvement, les virus sont ouverts par l'action de détergents (lyzolécithine et triton X-100) et les RNPv récupérées après avoir été déposées puis centrifugées sur un gradient de glycérol. L'étape suivante consiste à séparer la NPv de l'ARN et de la polymérase virale. C'est à nouveau une étape d'ultracentrifugation qui va nous permettre de récupérer la NPv, mais cette fois-ci sur un gradient de glycérol contenant du chlorure de césium (CICs). Ce composé salin décroche l'ARN de la NPv qui est ensuite récupérée dans le gradient de CICs (**figure 37a et b**).

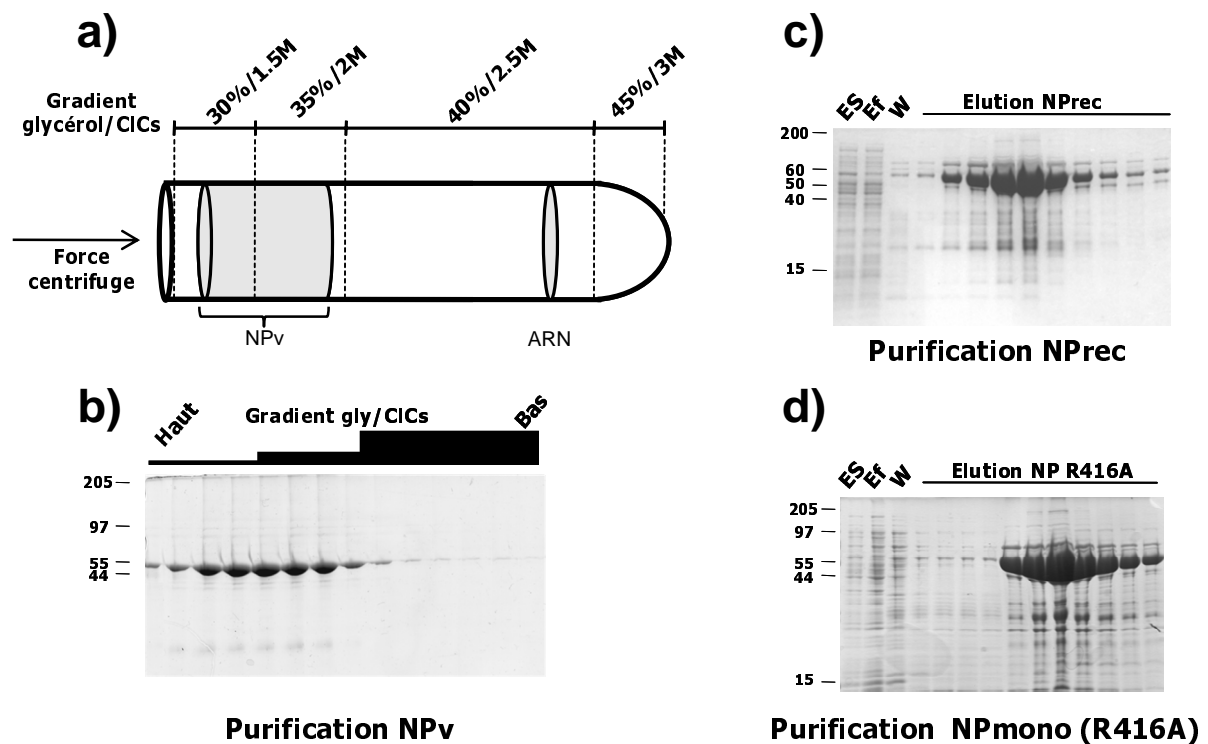
Le gène de la nucléoprotéine recombinante (NPrec) présente la même séquence que celle de la NPv et la mutation R416A (NPmono), consistant à substituer une arginine par une alanine en position 416, est obtenu à partir du gène NPrec. NPrec et NPmono sont exprimés chez la bactérie, puis purifiés par chromatographie d'affinité grâce à la présence d'une étiquette polyhistidine située à l'extrémité N-terminale des gènes (**figure 37c et d**).

### 2.2. Caractérisation du mutant NPmono

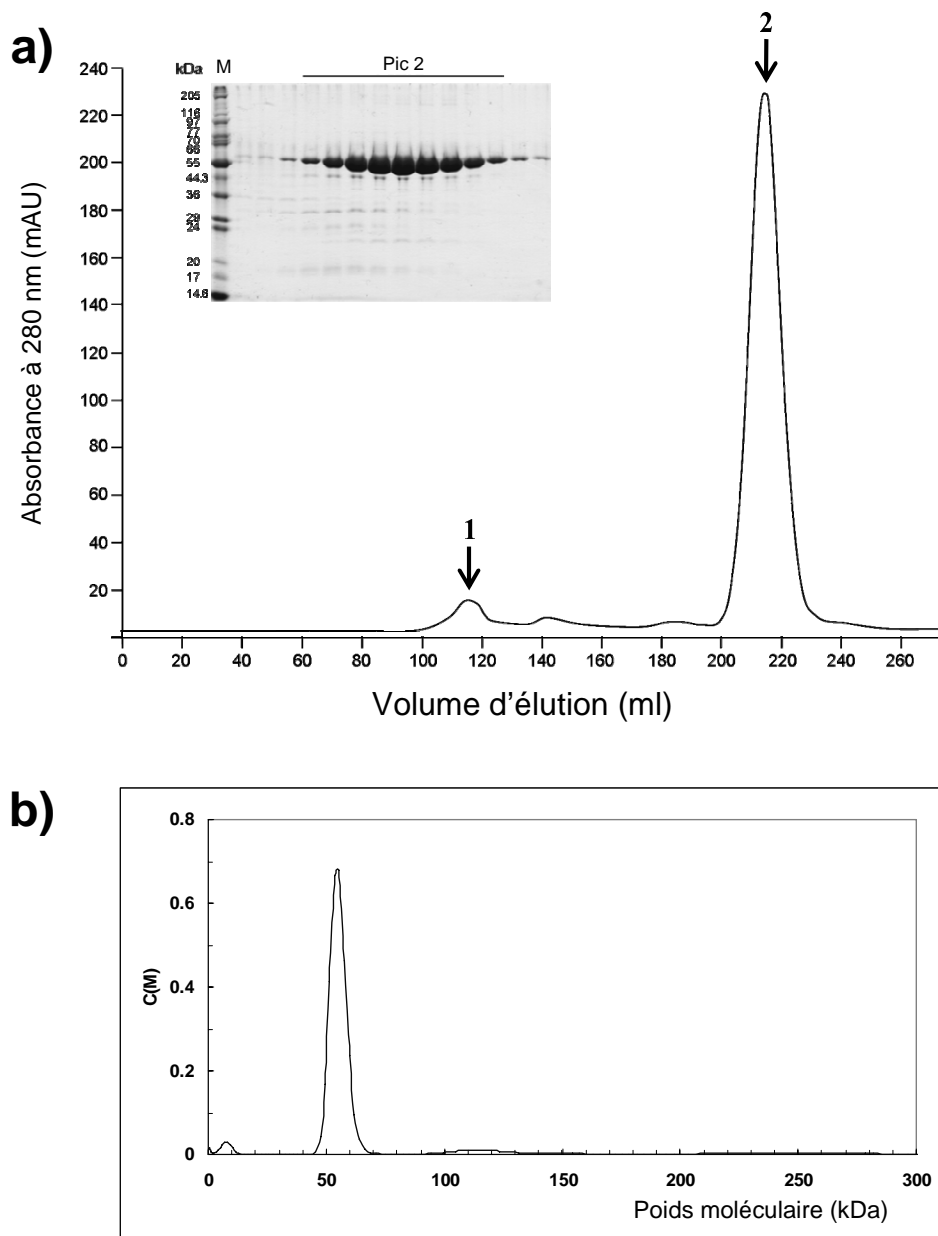
Alors que la nucléoprotéine recombinante, NPrec, comme la NP virale (NPv) s'oligomérisent pour former des chapelets de tailles variables, la NPmono présente un caractère monomérique très homogène. En effet, suite à une chromatographie d'exclusion de taille (sur Superdex 200 HiLoad), la NPmono est éluée en un pic majoritaire à un volume de 215 ml (**figure 38a**). Selon les courbes de calibration standard, ce volume correspond à une protéine d'environ 55 kDa. L'analyse par électrophorèse sur gel dénaturant du pic n° 2 révèle une protéine très pure (**figure 38a**). Nous pouvons cependant noter la présence d'un léger pic élué à 115 ml dans le volume mort de la colonne et correspondant à quelques agrégats (**figure 38a, pic 1**). Nous avons ensuite réalisé des expériences d'ultracentrifugation analytique (AUC) sur

NPmono (**figure 38b**). Cette méthode permet d'étudier les macromolécules en solution lorsqu'elles sont soumises à une force de gravitation. Le comportement hydrodynamique d'une macromolécule et son coefficient de sédimentation sont liés au volume, à la forme, à la densité et à la masse de la molécule.

Ainsi cette méthode permet de déterminer directement la masse d'une protéine. S'il s'agit d'un oligomère ou d'un complexe, le nombre de sous-unités ou la stœchiométrie peuvent ainsi être déterminés. De plus, des informations sur l'homogénéité de l'échantillon peuvent être obtenues. La NPmono est ainsi sédimentée pour un poids moléculaire de 55 kDa, ce qui indique que la protéine en solution se comporte bel et bien comme un monomère.



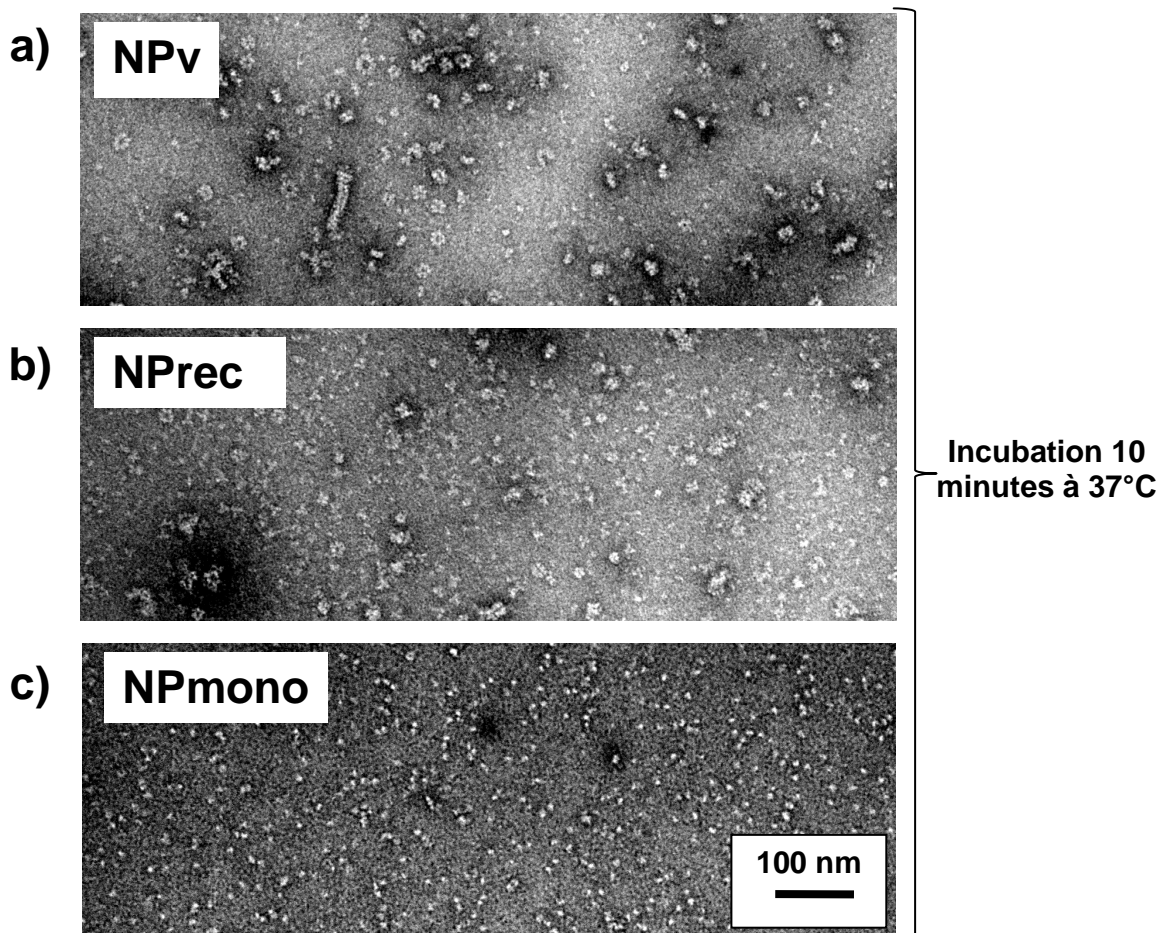
**Figure 37. Purification des différentes formes de nucléoprotéines.** a) Principe de purification de la NP virale (NPv) sur gradient de glycérol et de CICs après ultracentrifugation et b) analyse par SDS-PAGE 12% des fractions issues de l'ultracentrifugation. Purification c) de la NP recombinante (Nprec) et d) du mutant R416A (NPmono) par chromatographie d'affinité analysés par SDS-PAGE 12%. ES : extrait soluble, Ef : effluent, W : lavage.



**Figure 38. Caractérisation du mutant R416A de la nucléoprotéine (NPmono) (sans His).** **a)** Profil d'éluion de NP R416A (NPmono) du virus de la grippe (A/PR/8/34) par chromatographie d'exclusion de taille sur une colonne Superdex 200 HiLoad (GE Healthcare). Les flèches indiquent (1) le volume mort et (2) le volume d'éluion de la NP (MM 56 kDa). Et analyse par électrophorèse sur gel de polyacrylamide (SDS-PAGE 12% Laemmli) du pic n°2 correspondant au mutant R416A de la NP. M = Marqueur moléculaire. **b)** Résultats d'ultracentrifugation analytique pour le mutant NP R416A (NPmono) du virus de la grippe (A/PR/8/34). La distribution de coefficient de sédimentation provient du profil de vitesse de sédimentation analysé à l'aide du programme SEDIFT. La concentration en protéine était de 0,8 mg/ml.

### 2.3. Analyse par microscopie électronique des différentes formes de NP

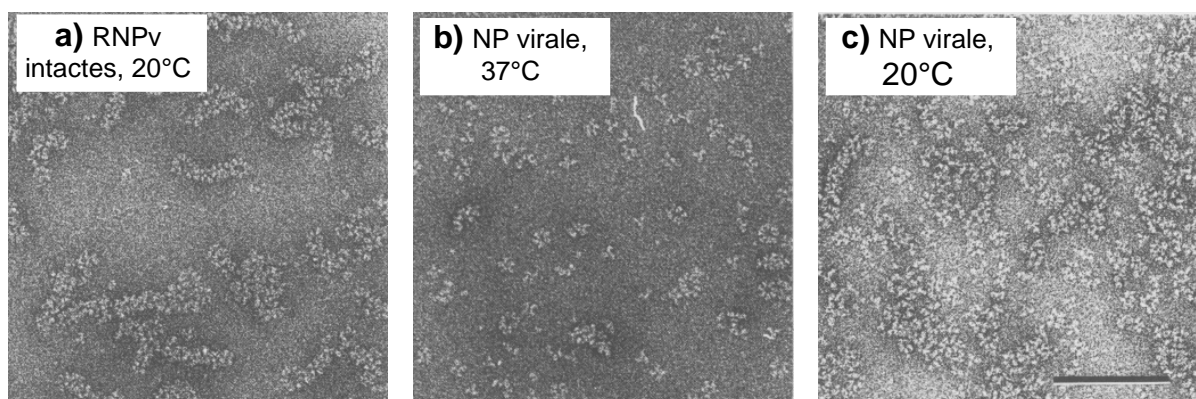
Afin de comparer les trois formes de nucléoprotéine (NPv, NPrec et NPmono) entre elles, les différents échantillons furent préparés pour être analysés en microscopie électronique à coloration négative. Après une étape d'incubation de 10 minutes à 37°C pour modérer l'état d'agrégation, les échantillons sont colorés par de l'acétate d'uranyle. Nous pouvons voir que la NP virale est hautement polymérisée et qu'elle forme des anneaux ainsi que des polymères d'ordres supérieurs semblables aux RNPv (**figure 39a**). Nous pouvons également noter qu'il y a relativement peu de petits oligomères de types dimères, trimères, etc.



**Figure 39. Images de microscopie électronique en coloration négative de la nucléoprotéine du virus de la grippe (A/PR/8/34). a) Nucléoprotéine virale (NPv) purifiée à partir du virus, b) NP recombinante (NPrec) et c) mutant NP R416A. La coloration négative est réalisée avec 2% d'acétate d'uranyle et les images sont prises avec le microscope électronique Morgagni FEI.**

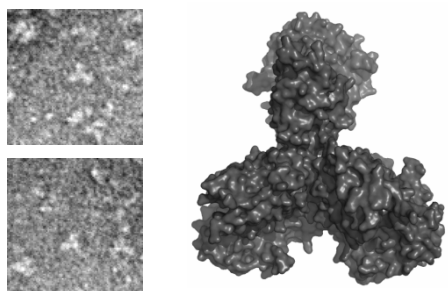
Largement étudiée dans le passé par notre équipe, en microscopie électronique à coloration négative, la nucléoprotéine purifiée à partir du virus, polymérise pour former des structures identiques aux RNPv natives (Ruigrok & Baudin 1995). Dans ces travaux, publiés en 1995, les RNPv et la NP, purifiés directement à partir du virus, furent incubés, tout d'abord à 37°C, pour éliminer l'état d'agrégation, puis à nouveau à différentes températures. Ainsi après 10 minutes à 20°C, les RNPv montrèrent une morphologie caractéristique pseudo hélicoïdale avec la présence de boucles aux extrémités (**figure 40a**). Concernant la nucléoprotéine, la même préparation de NP fut utilisée mais à des concentrations et températures d'incubation différentes. Alors que les molécules de NP, incubées à 37°C (**figure 40b**), s'associent en trimères, tétramères et en petits anneaux, l'auto-polymérisation observée est bien plus importante lorsque l'échantillon est incubé à 20°C (**figure 40c**). La comparaison des clichés (**figure 40a et b**) montre ainsi une morphologie similaire de la NP virale à celle des RNPv natives.

Concernant la NPrec, de nos expériences actuelles, elle se présente également sous forme d'oligomères semblables à ceux observés pour la NPv mais de taille inférieure (**figure 39b**). Nous pouvons ici observer la présence d'oligomères de petite taille tels que des pentamères, des tétramères ou encore des trimères dont l'agencement est fortement similaire à celui de la structure cristallographique du trimère de NP (**figure 41**). Enfin, le mutant R416A, nous montre la présence exclusive de particules de petite taille correspondant à la protéine sous forme monomérique (**figure 39c**) comme le suggèrent les expériences de sédimentation ainsi que les résultats de *Ye et al.*, (2006).



**Figure 40. Images de microscopie électronique en coloration négative de RNP intactes, et de la nucléoprotéine purifiée à partir du virus de la grippe (A/PR/8/34) (Ruigrok et Baudin 1995).** La suspension de RNP a été incubée à 20°C pour **a**), la préparation de NP a été incubée à 37°C pour **b**) et 20°C pour **c**) avant la préparation pour la microscopie électronique. La barre représente 100 nm.



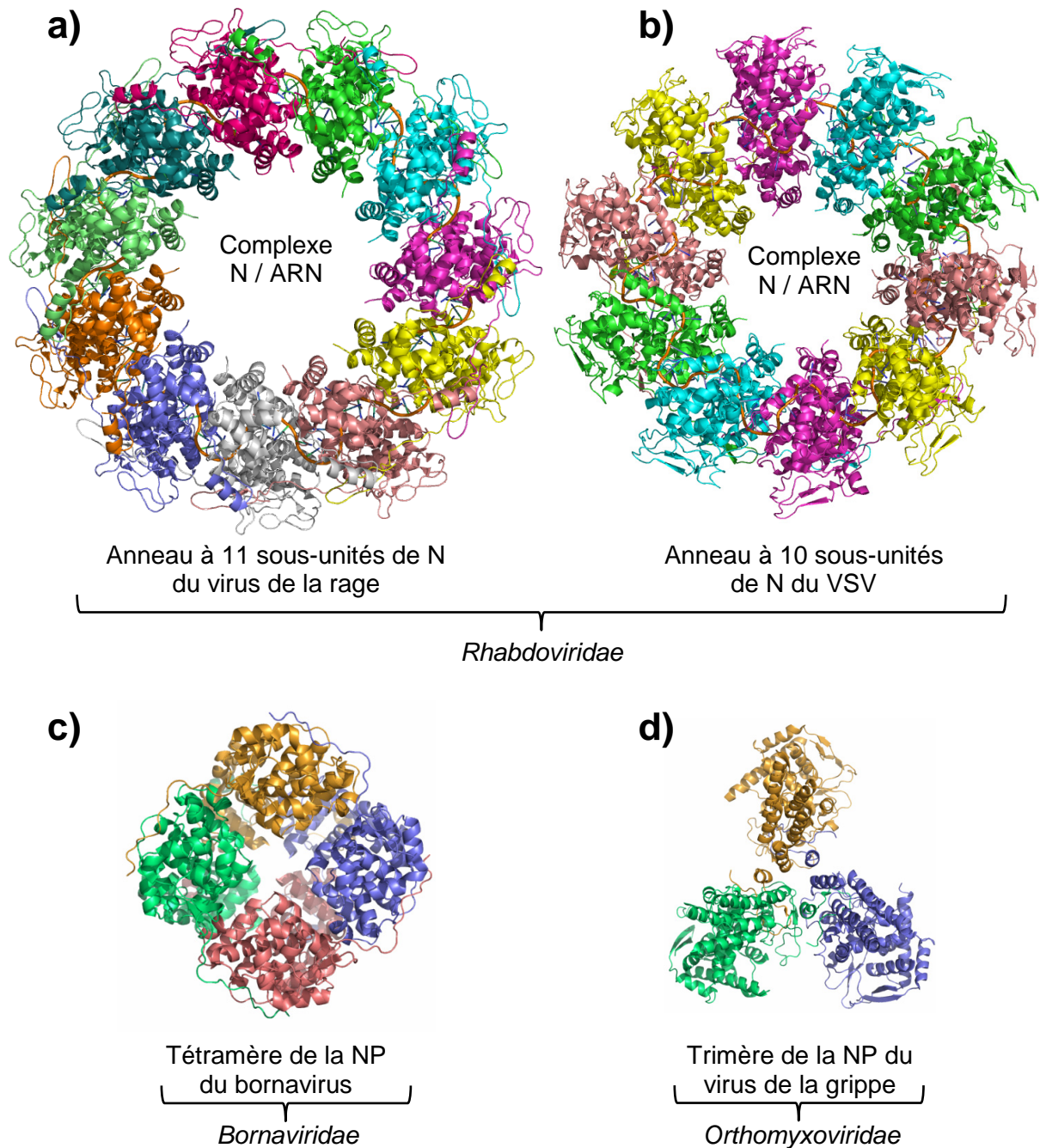


**Figure 41. Comparaison du trimère de la nucléoprotéine observé en microscopie électronique à coloration négative avec sa structure cristallographique (représentée en mode surface).**

### 3. Les nucléoprotéines des virus à ARN négatifs

#### 3.1. Etat d'oligomérisation des différentes NP de virus à ARN négatif

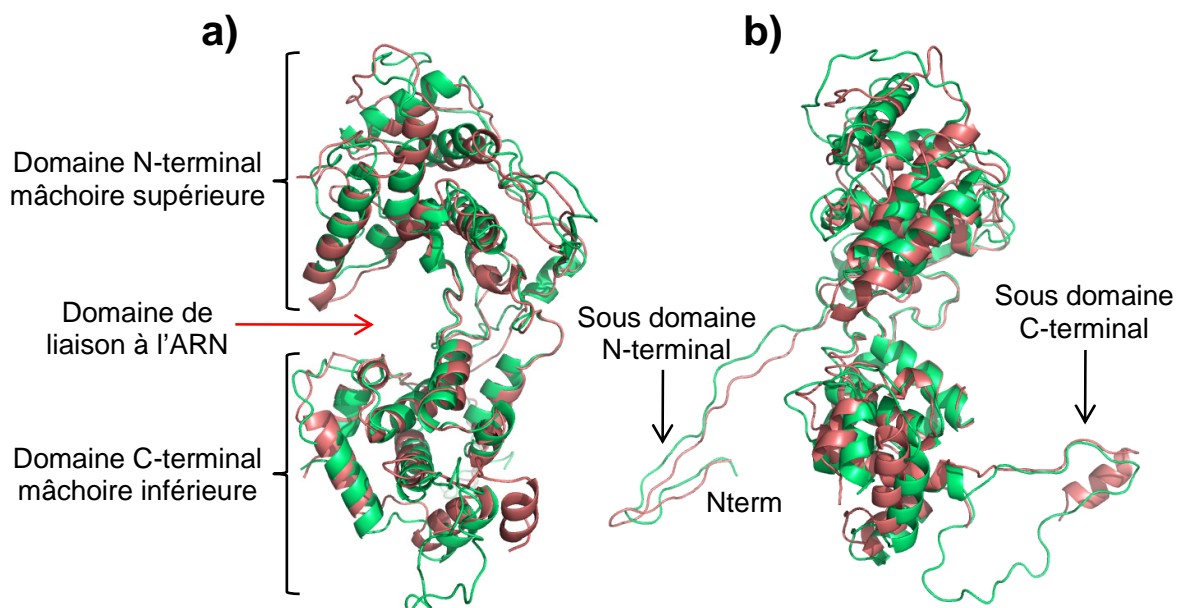
Afin de comprendre davantage la structure et l'état d'oligomérisation de la NP du virus de la grippe, il est intéressant de la replacer dans le contexte des nucléocapsides et des nucléoprotéines d'autres virus à ARN négatif. Le bornavirus (famille des *Bornaviridae*), comme le virus de la rage et le VSV, appartient à l'ordre des *Mononégavirales*. La structure de sa nucléoprotéine a été résolue à 1,76 Å et à l'intérieur des cristaux les sous-unités de NP sont arrangées en homo-tétramères (**figure 42c**) (Rudolph *et al.*, 2003). Contrairement à d'autres *Mononégavirales*, la NP de ce virus, a pu être cristallisée sans être associée à un ARN. En effet, concernant le virus de la rage et de son proche voisin le virus de la stomatite vésiculaire (VSV), tous deux membres de la famille des *Rhabdoviridae*, leurs nucléoprotéines (N) ont pu être cristallisées uniquement en présence d'ARN. Les structures cristallographiques des nucléoprotéines de ces deux virus ont été résolues et publiées en même temps. La nucléoprotéine de la rage (N) a été résolue à 3.5 Å, à l'intérieur d'un anneau constitué de 11 sous-unités de nucléoprotéines et d'un ARN de 99 bases (**figure 42a**) (Albertini *et al.*, 2006). La structure de la nucléoprotéine de VSV (N) résolue à 2.9 Å, représente également un anneau, mais cette fois-ci, composé de 10 sous-unités de nucléoprotéines associées à un ARN de 90 bases (**figure 42b**) (Green *et al.*, 2006). Nous pouvons noter que les structures actuelles des nucléoprotéines de 3 familles de virus différents, y compris celui de la grippe, dérivent toujours de polymères : soit en trimère, tétramère ou anneau circulaire (**figure 42**).



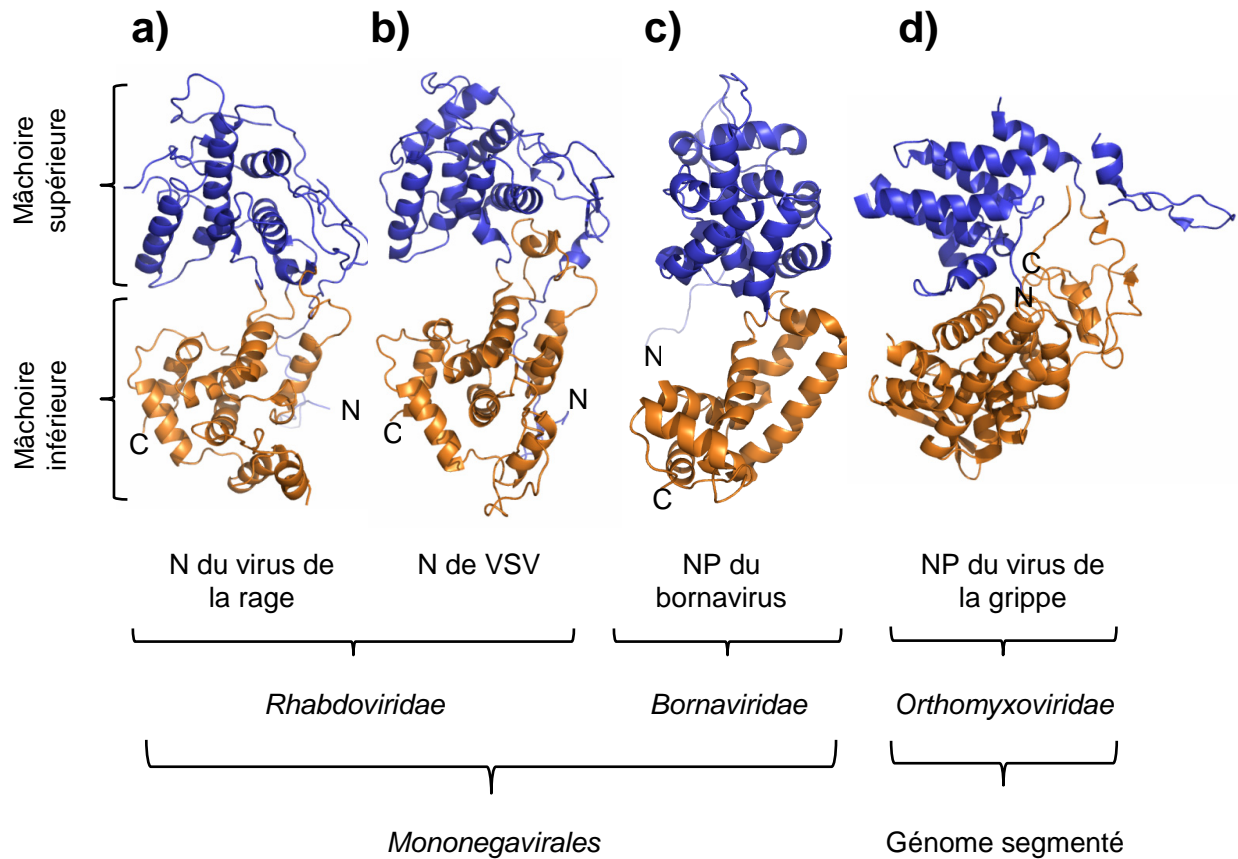
**Figure 42. Cristallisation et structures cristallographiques de nucléoprotéines de plusieurs familles de virus à ARN négatif.** Structures cristallographiques **a)** des anneaux à 11 sous-unités de la nucléoprotéine du virus de la rage en complexe avec l'ARN (code pdn : 2GTT) (Albertini *et al.*, 2006), **b)** des anneaux à 10 sous-unités de nucléoprotéine du virus de la stomatite vésiculaire (VSV) en complexe avec l'ARN (code pdb : 2GIC) (Green *et al.*, 2006), **c)** du tétramère de la NP du bornavirus (code pdb : 1N93) (Rudolph *et al.*, 2003) et **d)** du trimère de la NP du virus de la grippe (code pdb : 2IQH) (Ye *et al.*, 2006).

### 3.2. Comparaison des structures des nucléoprotéines de virus à ARN négatif

La NP du bornavirus comporte 370 acides aminés, les N du virus de la rage et du VSV comportent respectivement 450 et 422 acides aminés et la NP du virus de la grippe comporte 498 acides aminés. Il n'y a pas, ou très peu d'homologie de séquence entre ces différentes nucléoprotéines : près de 20% entre les deux N de rhabdovirus et moins de 6% pour toutes les autres. Cependant, en comparant les structures de la N de la rage et de VSV, à l'aide du programme FATCAT (Ye & Godzik, 2003), nous observons qu'elles sont très semblables et se superposent (rmsd de 2.98 Å sur les C $\alpha$ ) (**figure 43**). Elles peuvent être divisées en deux domaines principaux, que sont les mâchoires supérieure et inférieure, qui séquestrent l'ARN. Dans les deux cas, le sillon dans lequel se loge l'ARN est situé au même endroit sur les deux nucléoprotéines. Enfin, deux sous-domaines flexibles en N et C terminaux (**figure 43b**) interagissent avec des NP voisines (**figure 42a et b**).



**Figure 43. Superposition des structures des nucléoprotéines du virus de la rage et de VSV.** La N du virus de la rage est colorée en rose (code pdb : 2GTT), la N de VSV est colorée en vert (code pdb : 2GIC). Rmsd : 2.98 Å. La superposition est effectuée à l'aide du programme FATCAT (Ye & Godzik, 2003). **a)** Vue de côté, la flèche rouge indique le domaine de liaison à l'ARN. **b)** Vue de face.

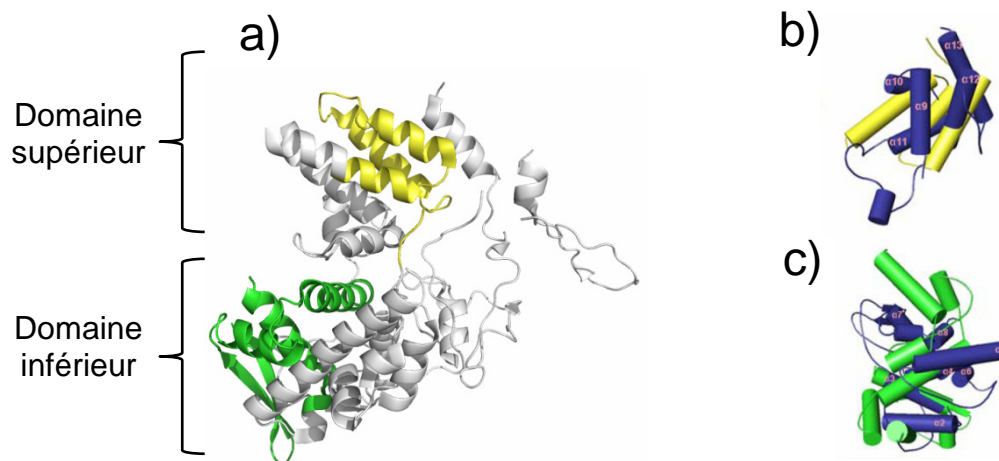


**Figure 44. Comparaison des structures cristallographiques de nucléoprotéine de plusieurs familles de virus à ARN négatif.** Structure cristallographique de la nucléoprotéine **a)** du virus de la rage cristallisée sous forme d'anneaux (11-mer) avec de l'ARN (ARN non montré) (Albertini *et al.*, 2006), **b)** du virus de la stomatite vésiculaire (VSV) cristallisée sous forme d'anneaux (10-mer) avec de l'ARN (ARN non montré) (Green *et al.*, 2006), **c)** du bornavirus cristallisée en tétramère sans ARN (Rudolph *et al.*, 2003) et **d)** du virus de la grippe cristallisée en trimère sans ARN (Ye *et al.*, 2006).

Concernant la nucléoprotéine du bornavirus, sa structure se superpose approximativement aussi bien avec celle de la N de la rage, qu'avec celle du VSV (Luo *et al.*, 2007). En effet l'organisation des éléments de structure secondaire est quelque peu différente. Néanmoins, cette structure présente des similitudes et est composée principalement d'hélices  $\alpha$  et peut être divisée en deux domaines, N et C-terminaux distincts (**figure 44c**). De plus, toujours sur cette même figure, il est possible de distinguer un sillon qui pourrait être impliqué dans la liaison à l'ARN. Cette protéine présente, elle aussi, deux sous-domaines (non distincts sur la figure vue de côté) qui émergent et qui pourraient être impliqués dans la formation de complexe NP-NP lorsque la protéine est liée à l'ARN. Enfin, il est possible de distinguer sur la **figure 44**, la similitude qu'il existe entre les différentes nucléoprotéines de certains virus à ARN négatif. Ceci s'observe particulièrement au niveau de l'organisation des domaines inférieur et supérieur. Cependant, bien que la NP de la grippe possède elle aussi

deux domaines distincts comme les autres nucléoprotéines, celle-ci présente néanmoins une morphologie structurale différente (**figure 44d**). Bien que ces quatre protéines possèdent une silhouette globalement semblable, nous pouvons davantage observer une morphologie de "mâchoire" sur les nucléoprotéines de *Mononegavirales* (**figure 44a, b et c**). De plus, le sillon de la NP de la grippe est plus ouvert et s'apparente davantage à une plateforme (**figure 44d**).

Dans les travaux de [Luo et al., \(2007\)](#), les auteurs ont cherché à comparer les différentes structures de nucléoprotéines de virus à ARN négatif. Ils ont observé que la région comprenant les acides aminés 21-271 de la NP du virus de la grippe présentait une certaine similarité de structure avec le domaine de liaison de à l'ARN de la N de VSV. Cette région 21-271 de la NP de la grippe est établie sur deux domaines, l'un sur la partie inférieure de la protéine (acides aminés 21-149), l'autre sur la partie supérieure (acides aminés 150-271) (**figure 45**). Cependant ces deux domaines de la NP de la grippe présentent une orientation très différente aux deux domaines similaires présents sur la structure de VSV. En conséquence, chacun des domaines sur les deux protéines peuvent uniquement être superposés séparément (**figure 45b et c**). Le reste de la protéine (acides aminés 272-489) ne présente aucune homologie, que ce soit avec la N du virus de la rage ou de VSV ([Luo et al., 2007](#)). Il n'est donc pas impossible que l'orientation des deux domaines de la NP du virus de la grippe change lorsque la protéine se retrouve liée à l'ARN.



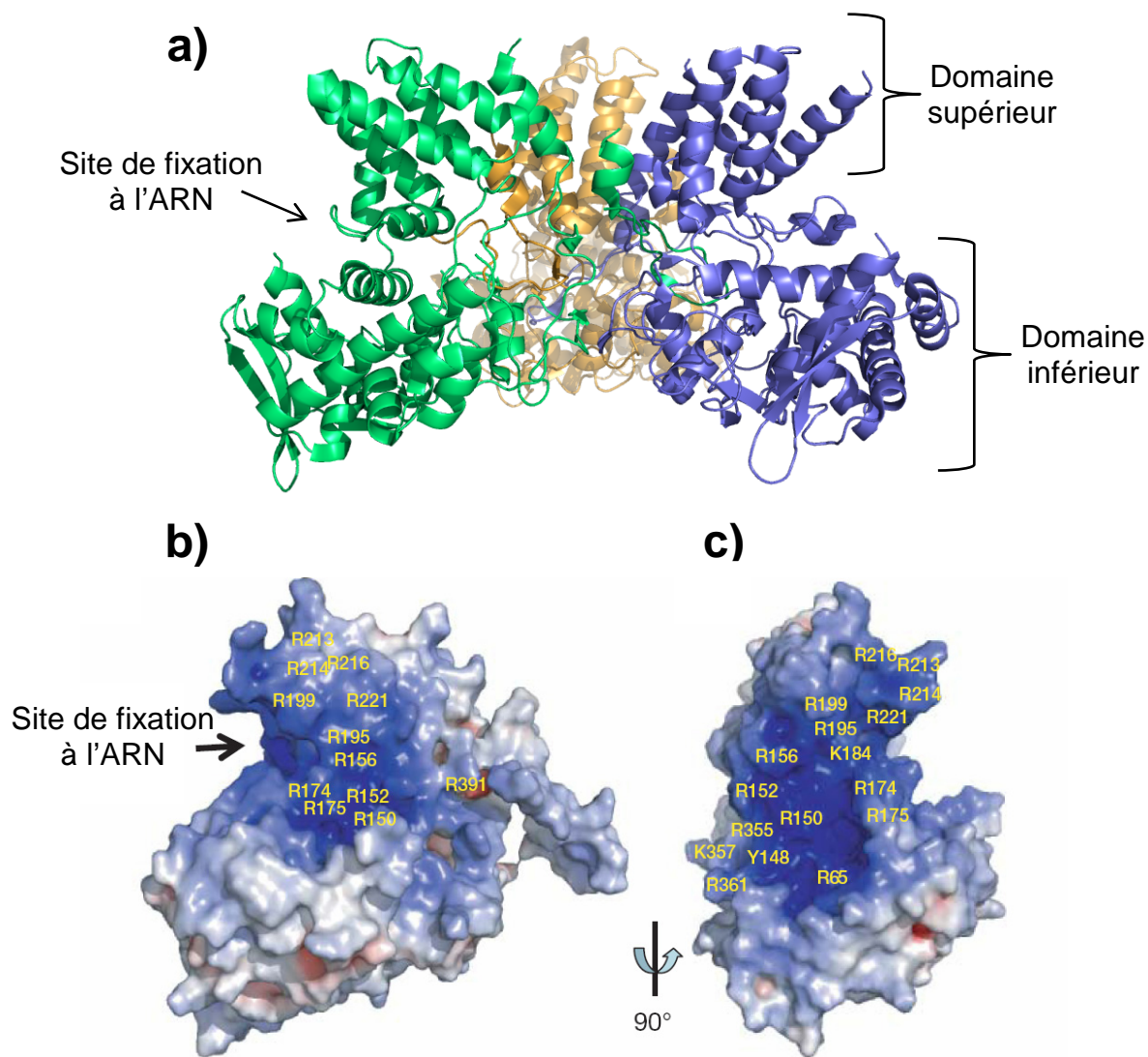
**Figure 45. Comparaison des structures des nucléoprotéines du virus de la grippe et du VSV. a)** Structure cristallographique du monomère de la NP du virus de la grippe avec représentées en jaune et vert les régions similaires à certains domaines de la N de VSV. **b)** Superposition d'une partie du domaine supérieur de la NP du virus de la grippe (en jaune) avec le domaine C-terminal de la N de VSV (en bleu). **c)** Superposition d'une partie du domaine inférieur de la NP du virus de la grippe (en vert) avec le domaine N-terminal de la N de VSV (en bleu), adapté de [Luo et al., 2007](#).

Comme discuté précédemment, les différentes nucléoprotéines polymérisent le long du génome viral. Nous avons vu que pour la N de la rage et de VSV, que les sous-domaines N et C-terminaux émergents faisaient le lien entre les N voisines et contribuaient à la stabilité du complexe. En revanche, l'arrangement du trimère de NP du virus de la grippe est tellement différent qu'il est impossible de le comparer à l'état d'oligomérisation de la N de la rage ou de VSV et qu'il est difficile de prévoir l'arrangement de la NP le long de l'ARN viral.

#### 4. Affinité de la nucléoprotéine pour l'ARN

##### 4.1. Site de fixation à l'ARN

Bien que la structure cristalline de la nucléoprotéine du virus de la grippe résolue par [Ye et al., \(2006\)](#) a été obtenue en absence d'ARN, elle révèle néanmoins la localisation probable de son site de fixation à l'ARN. En effet, nous pouvons observer la présence d'un sillon situé entre les domaines inférieur et supérieur de la protéine qui de plus, est exposé vers l'extérieur de l'oligomère (**figure 46a**). Si l'on regarde maintenant la distribution du potentiel électrostatique de la NP, nous voyons que la surface de ce sillon est constituée d'un important nombre d'acides aminés basiques, comprenant R65, R150, R152, R156, R174, R175, K184, R195, R199, R213, R214, R216, R221, R355, K357, R361 et R391 (**figure 46b et c**). Sur des complexes NP-ARN reconstitués, l'ARN se fixe à la NP grâce à son squelette ribose-phosphate ([Baudin et al., 1994](#)). Les acides aminés mentionnés plus haut, sont de bons candidats pour interagir directement avec l'ARN. Un acide aminé aromatique, la tyrosine en position 148, est situé au fond du sillon et serait susceptible d'interagir avec une base nucléotidique (**figure 46c**). Onze des seize acides aminés basiques observés à la surface du supposé site d'interaction à l'ARN, sont conservés parmi les virus de grippe A, B et C (souches A/WSN/33, B/Ann Arbor/1/66, et C/California/78, respectivement) confortant ainsi leur rôle dans la fixation à l'ARN. La tyrosine 148 est également conservée.

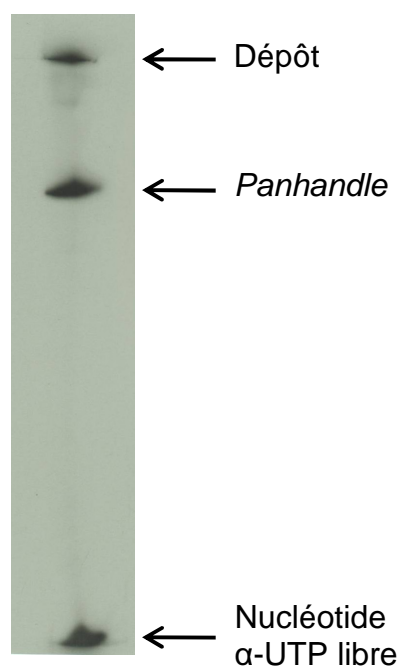


**Figure 46. Mise en évidence du site probable de fixation à l'ARN (Ye *et al.*, 2006). a)** vue de côté du trimère de la NP. **b et c)** Distribution du potentiel électrostatique sur la NP. Les potentiels électrostatiques, positif et négatif, sont représentés en bleu et rouge respectivement. Sont indiqués, les acides aminés basiques situés à l'intérieur du sillon. Une rotation de 90° est effectuée entre les images **a** et **b**.

#### 4.2. Filter Binding Assays (FBA)

Afin de tester l'activité de fixation de la NP à l'ARN en fonction de son état d'oligomérisation, nous avons mesuré l'affinité de nos trois échantillons (NPv, NPrec et NPmono) pour un mini ARN viral de type *panhandle*. L'ARN que nous avons utilisé est long de 81 nt. Décrit par [Baudin et al. \(1994\)](#), il comprend les extrémités 3' et 5' conservées chez le virus de la grippe de type A, connectées par une séquence dérivée du gène de la NP. Le terme *panhandle* a été employé pour décrire une structure en forme de "poêle à frire" comprenant un ARN circulaire dont les extrémités 3' et 5' s'associent pour former une structure en double brin ([Hsu et al., 1987](#)).

Nous avons déterminé l'affinité des différentes préparations de NP pour l'ARN *panhandle* par des expériences de *filter binding assays* (FBA). Cette technique est largement utilisée pour mesurer les constantes d'affinité entre deux molécules *in vitro*. Afin de déterminer cette constante il faut tout d'abord trouver un moyen de mesurer la quantité de complexe formé. Ceci peut être réalisé en marquant l'une des espèces avec une étiquette fluorescente, ou dans notre cas, une étiquette radioactive. Ainsi, à la suite d'une étape de transcription *in vitro*, l'ARN *panhandle* est marqué radioactivement grâce à l'incorporation d'uridine dont le phosphate  $\alpha$  est marqué par l'isotope  $P^{32}$ . L'ARN radioactif produit est ensuite purifié par électrophorèse sur gel (**figure 47**).

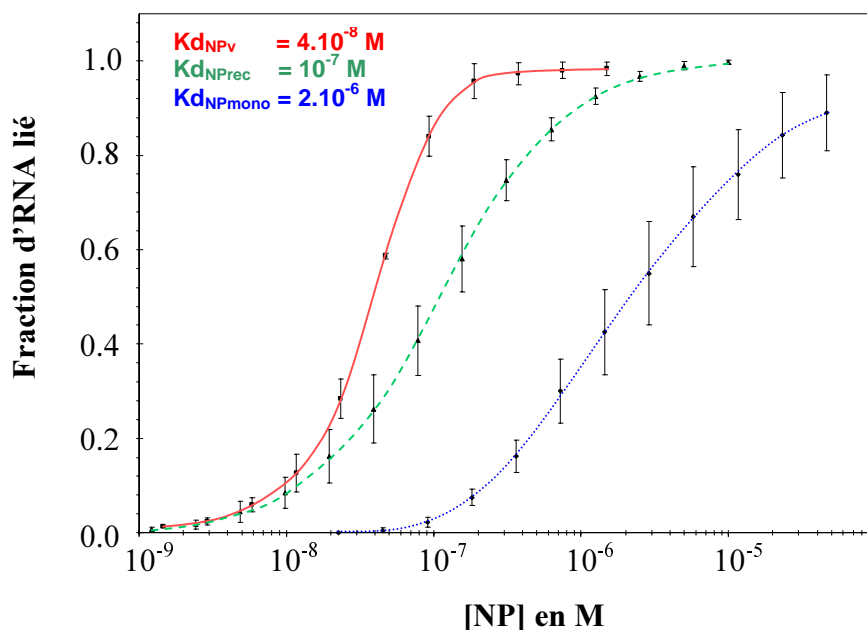


**Figure 47. Purification de l'ARN *panhandle* suite à la transcription radioactive.** Autoradiographie du gel polyacrylamide 12%, urée 8 M, après 1 min 30 sec d'exposition.



L'expérience de FBA proprement dite nécessite l'utilisation d'un filtre de nitrocellulose. Cette membrane est idéale pour immobiliser les protéines qui restent accrochées. Quant à l'ARN, qui est hydrophile, il n'est pas en mesure de s'accrocher au filtre de nitrocellulose alors qu'associé à une protéine il sera retenu. La proportion exacte d'ARN retenu sur le filtre, donc fixé à la protéine, est quantifiée par mesure de la radioactivité présente sur le filtre grâce à un compteur à scintillation.

Nous avons mesuré les affinités avec une faible quantité d'ARN radio-marqué à concentration constante, mélangée à des concentrations croissantes de NP (de  $10^{-9}$  M à  $10^{-4}$  M). La solution contenant le mélange est ensuite filtrée à travers la membrane de nitrocellulose. La constante de dissociation ( $K_d$ ), pour un équilibre entre deux molécules, peut être définie par la concentration en protéine nécessaire pour être à 50% de rétention. La quantité d'ARN utilisée étant extrêmement faible, (0.3 pmol), on peut considérer que la fraction d'ARN lié au plateau est représentatif d'une fixation complète. Ainsi, la concentration de NP à 50% de rétention, correspond à la constante de dissociation, le  $K_d$ . Les résultats sont présentés dans la **figure 48**. Nous pouvons observer que la NP virale se fixe à l'ARN *panhandle* avec un  $K_d$  de 40 nM ce qui est comparable à la valeur de 37 nM publiée dans [Baudin et al., \(1994\)](#). La NP recombinante se fixe à l'ARN *panhandle* avec une affinité 2 fois inférieure ( $K_d = 100$  nM). Le mutant monomérique R416A (NPmono) quant à lui se fixe très faiblement à l'ARN ( $K_d = 2$   $\mu$ M). Il existe une relation inverse entre l'état de polymérisation et l'affinité à l'ARN pour les trois échantillons de NP. Les courbes représentées, sont de formes sigmoïdales ce qui indique que la fixation à l'ARN est coopérative, comme montré précédemment par [Yamanaka et al., \(1990\)](#). De plus, si l'on observe la pente de chacune des trois courbes, nous remarquons que celle de NPv est plus importante que celle de NPrec qui est elle-même plus importante que celle de NPmono (**figure 48**). Ceci suggère que l'effet coopératif est moins important pour les protéines dont l'état d'oligomérisation est plus faible, notamment pour le mutant NPmono. Cependant, une alternative à l'effet coopératif de la NP pourrait être envisageable pour expliquer la différence entre ces courbes, celle-ci pourrait être due à l'état de polymérisation de la protéine avant sa fixation à l'ARN. En effet, lorsque la NP est sous forme de polymère, elle pourrait plus rapidement et avec une plus grande affinité se fixer à l'ARN. Une question se pose néanmoins. Les séquences en acides aminés de la NPv et de la NPrec sont identiques. Quelle pourrait être l'origine de cette différence d'état de polymérisation et d'affinité observée entre ces deux protéines ? Bien que ces séquences soient identiques, leurs modes d'expression sont cependant différents. En effet la NP recombinante est produite par la bactérie qui n'apporte aucune modification post-traductionnelle.

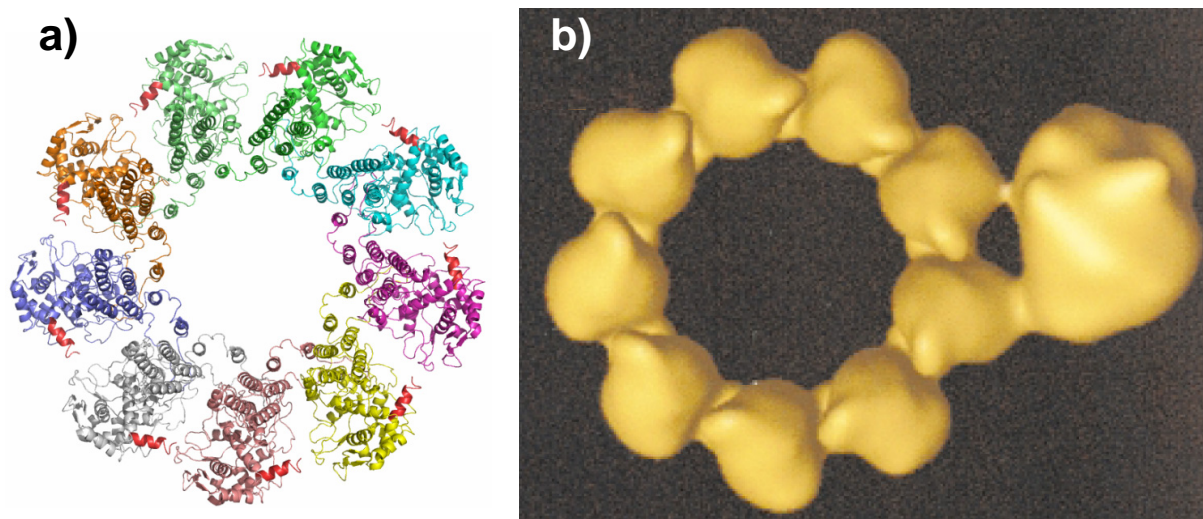


**Figure 48. Activité de fixation à l'ARN par la nucléoprotéine du virus de la grippe (souche A/PR/8/34).** Courbes de titration réalisées à partir de "filter binding assays" (FBA) sur un ARN de 81 bases (*panhandle*) avec la NPv (—), la NPrec (---) et le mutant NP R416A (.....). Le coefficient de dissociation apparent (Kd) correspond à la concentration de la protéine à 50 % de rétention de l'ARN.

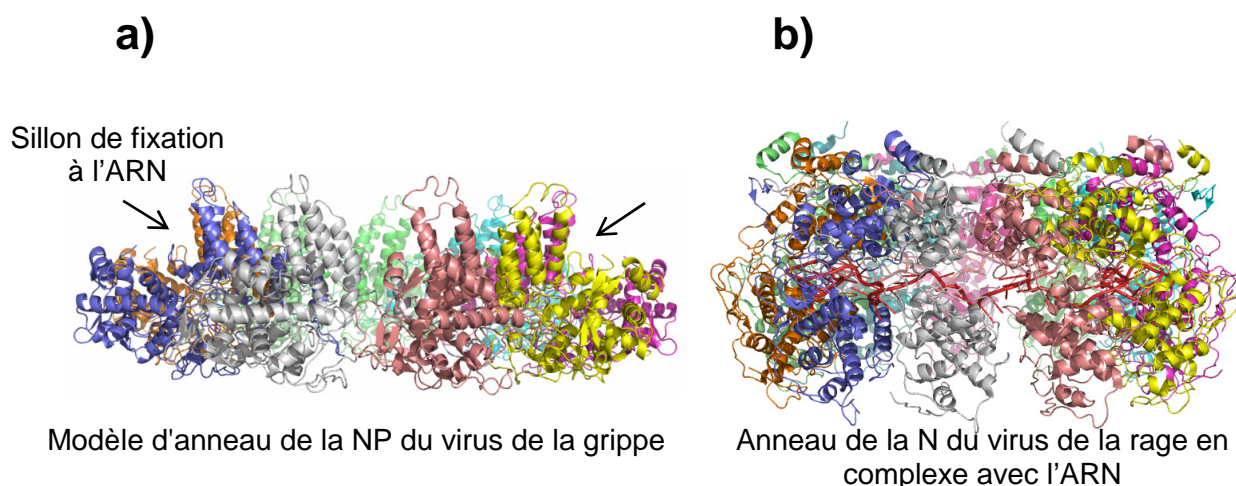
En revanche, la NP virale est produite grâce à l'utilisation de la machinerie cellulaire, ce qui lui permet d'acquérir des modifications post traductionnelles. Dans ce contexte, deux régions de phosphorylation ont été mises en évidence : la première, comprend les 88 premiers acides aminés N-terminaux, et la deuxième les 196 derniers acides aminés C-terminaux. Le site de phosphorylation principal consiste en la sérine en position 3 (Arrese & Portela, 1996), toutefois la sérine 473 pourrait également être phosphorylée (Kistner *et al.*, 1989). Le fait que la NPv puisse être phosphorylée pourrait induire un changement de conformation de la protéine et être à l'origine de la différence d'oligomérisation de la NP, et pourrait ainsi expliquer les différences observées entre NPv et NPrec dans les expériences de FBA.

#### 4.3. Modélisation de la plateforme de fixation à l'ARN

Elton *et al.*, (1999b) ont réalisé des expériences de pontage covalent sur différents mutants, dans le but de déterminer les acides aminés impliqués dans la fixation de la NP à l'ARN. Ils concluent que le mutant R416A ne se fixait pas à l'ARN. Nous avons cependant montré que même si l'affinité de ce mutant pour l'ARN est faible, la fixation est effective.



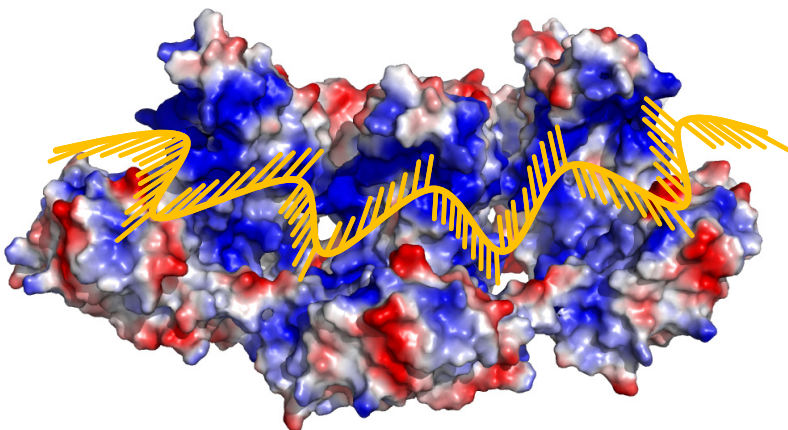
**Figure 49. a) Modèle d'un anneau formé par 9 sous-unités de nucléoprotéine du virus de la grippe (Ng *et al.*, 2008), basé sur b) la précédente structure du mini RNPv déterminée par microscopie électronique (Martin-Benito *et al.*, 2001). La région N-terminale de la NP est représentée en rouge.**



**Figure 50. Différents modèles d'encapsulation. a)** Le sillon de fixation à l'ARN est exposé vers l'extérieur dans le modèle de l'anneau de NP du virus de la grippe (Ng *et al.*, 2008), alors que **b)** l'ARN est séquestré à l'intérieur dans le modèle de l'anneau de la NP du virus de la rage. L'ARN est représenté en rouge (Albertini *et al.*, 2006).

Ceci était à prévoir suite à l'analyse de la structure cristallographique de la NP puisque la mutation R416A ne modifie pas la plateforme de fixation à l'ARN. Très récemment, la structure de la NP du virus de la grippe de type H5N1 a été déterminée (Ng *et al.*, 2008). Par des expériences de résonance plasmonique de surface sur différents mutants, les auteurs ont défini une plateforme de fixation à l'ARN sur la NP de H5N1. De plus, basé à la fois sur le modèle du mini RNPv en microscopie électronique (Martin-Benito *et al.*, 2001 ; Area *et al.*, 2004) et sur les structures des anneaux de la NP du virus de la rage (Albertini *et al.*, 2006) et de VSV (Green *et al.*, 2006), ils ont eux aussi proposé un modèle de la structure de la NP grippale sous forme d'un anneau (figure 49). Le modèle qu'ils défendent est un anneau de 9 sous-unités de NP, avec un site de fixation à l'ARN qui se trouve exposé vers l'extérieur de l'anneau. Cependant, le mode d'encapsidation de l'ARN proposé est nettement différent de celui des RNPv du virus de la rage et de VSV. En effet, dans le modèle sous forme d'anneau de la NP du virus de la grippe, le sillon de fixation à l'ARN est exposé vers l'extérieur. En revanche dans la structure en anneau de la NP du virus de la rage par exemple, l'ARN se trouve séquestré à l'intérieur (figure 50).

Les résultats que nous apportons ici, montrent le fait que la NP sous forme de polymères se fixe davantage à l'ARN que la NP monomérique. Ainsi, à l'aide de la structure de la NP nous pouvons modéliser un sillon continu issu de la polymérisation de la protéine (figure 51). Avec l'ARN fixé dans le sillon de la NP sous forme de polymère, la distance entre deux sites de fixation à l'ARN voisins serait de 70 Å (distance calculée entre deux NP voisines du trimère, Ye *et al.*, 2006). Ceci coïnciderait avec la stœchiométrie d'une nucléoprotéine pour 24 bases, comme cela fût précédemment déterminé par Ortega *et al.*, (2000). Bien entendu, la structure de la NP en complexe avec de l'ARN reste le moyen idéal non seulement de déterminer le site exact de fixation à l'ARN mais également l'arrangement structural de la nucléoprotéine lorsqu'elle le séquestre.



**Figure 51. Modèle schématique de la fixation de l'ARN à un oligomère de nucléoprotéine du virus de la grippe.** L'ARN représenté en orange est inséré dans le sillon basique issu de l'oligomérisation de 3 NP. Les potentiels électrostatiques, positif et négatif, sont représentés en bleu et rouge respectivement.

## **Chapitre 3**

—

### **Le complexe d'import nucléaire : importine $\alpha 5$ humaine / nucléoprotéine du virus de la grippe**

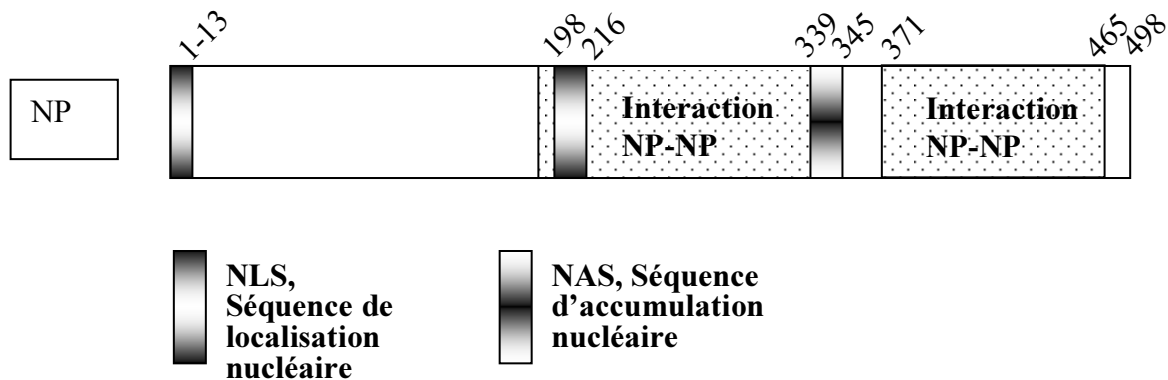
## Chapitre 3

### Le complexe d'import nucléaire : importine $\alpha 5$ humaine / nucléoprotéine du virus de la grippe

#### 1. Introduction

La transcription et la réplication du virus de la grippe s'effectuent à l'intérieur du noyau de la cellule infectée. Durant l'étape de réplication proprement dite les RNPv, contenant un ARN de polarité négative, passent par un intermédiaire complémentaire contenant l'ARN de polarité positive (RNPc), avant d'être à nouveau répliquées en RNPv. Suite à leur expression dans le cytoplasme, les protéines précoces que sont les trois sous-unités de la polymérase et notamment la nucléoprotéine, doivent rejoindre le noyau pour participer à la fois à la synthèse des RNPc et des RNPv. Ces protéines exprimées dans le cytoplasme nécessitent d'être transportées activement à l'intérieur du noyau. Cette étape fait ainsi appel à des signaux d'adressage spécifiques que sont les NLS (signal de localisation nucléaire), situés sur les protéines à transporter.

Les nombreuses études menées sur les signaux de transport de la nucléoprotéine ont mis en évidence la présence de plusieurs de ces séquences sur la protéine. Ainsi, un NLS non conventionnel a été cartographié dans la région N-terminale (acides aminés 1-13) (Wang *et al.*, 1997 ; Neumann *et al.*, 1997) puis un NLS classique bipartite entre les acides aminés 198 à 216 (Weber *et al.*, 1998) (**figure 52**).

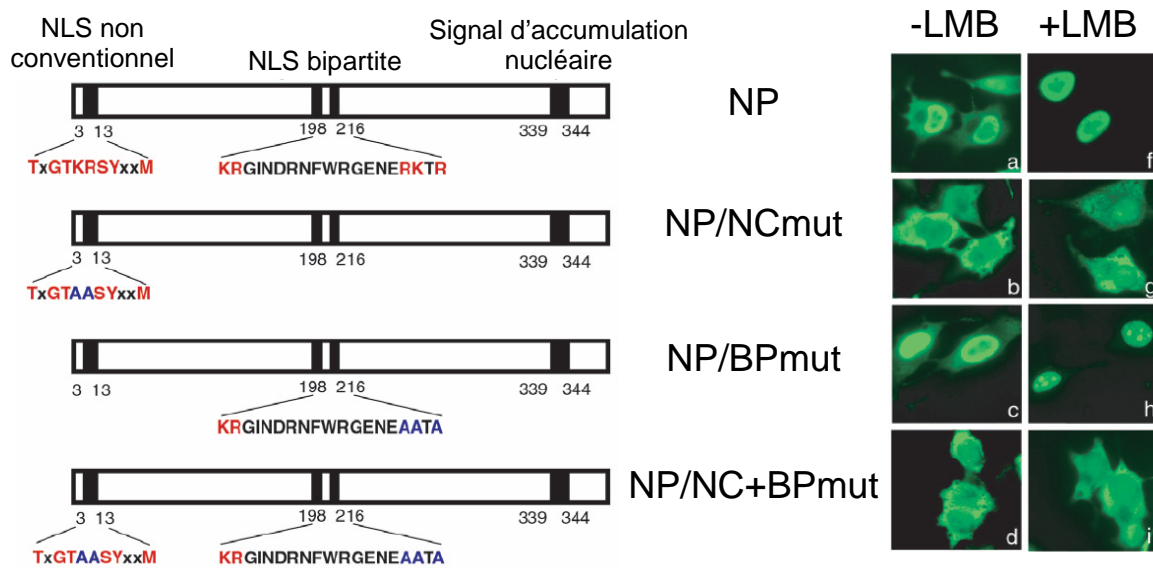


**Figure 52. Signaux d'adressage nucléaire de la nucléoprotéine du virus de la grippe.** La NP possède deux NLS et un domaine servant à l'accumulation nucléaire.

Le terme "classique" employé pour définir les NLS, fait référence aux deux séquences prototypes de localisation nucléaire. L'une étant le peptide NLS de l'antigène T de SV40 (NLS monopartite : -PKKKRKV-) (Kalderon *et al.*, 1984 ; Lanford & Butel, 1984) et l'autre des ces séquences correspond au NLS de la nucléoplasmine (NLS bipartite : -KR<sub>PAATKKAGQA</sub>KKKK-) (Dingwall *et al.*, 1988). Ainsi sur la NP, la région comprenant les acides aminés 198 à 216 (198-KR<sub>GINDRNFWRGENGRK<sub>T</sub>R</sub>-216) pourrait être assimilable à un NLS bipartite. En revanche, la région N-terminale de la protéine allant des acides aminés 1 à 13 (1-MASQG**T**KRSYEQM-13) ne correspond ni à l'une ou l'autre de ces séquences prototypes et est ainsi nommée NLS non conventionnel. De plus, un autre point important de régulation du transport de la NP est lié à la phosphorylation de la molécule. La NP peut être phosphorylée et parmi les différentes régions susceptibles de l'être, il a été montré que la sérine 3, situées dans le NLS N-terminal, était le site majeur pour accepter un phosphate (Arrese & Portela, 1996). De plus, l'inhibition de la phosphorylation (par des inhibiteurs de la protéine kinase C) conduit à l'accumulation nucléaire de la NP (Neumann *et al.*, 1997 ; Bui *et al.*, 2002). Ceci suggère une importance de la phosphorylation et/ou de certaines protéines cellulaires dans le transport de la NP, mais le mécanisme mis en jeu n'est toujours pas clairement défini. D'autre part, des études utilisant des œufs de xénope ont montré la présence d'une séquence dite d'accumulation/rétention nucléaire (NAS/NRS). Cette séquence située entre les acides aminés 339 et 345 serait à l'origine de l'accumulation de nucléoprotéine dans le noyau (Davey *et al.*, 1985). Il apparaît ainsi étrange de trouver autant de signaux sur une protéine.

Selon les travaux de Cros *et al.*, (2005), le NLS non conventionnel situé en N-terminal jouerait un rôle prédominant et serait crucial pour l'import de la nucléoprotéine libre. Ainsi, afin de déterminer l'importance relative de chaque NLS, ces auteurs ont fabriqué différentes constructions dans lesquelles les différents NLS furent inactivés en substituant certains acides aminés par des alanines (**figure 53 à gauche**). Pour réaliser ses expériences, ils ont transfecté les différents plasmides dans des cellules HeLa, puis ils ont localisé la NP par immunofluorescence. La NP sauvage (NP) est principalement localisée dans le noyau (**figure 53, panneau a**). La mutation du NLS non conventionnel (NP/NCmut) conduit à une forte réduction de l'accumulation de la NP dans le noyau (**figure 53, panneau b**), laissant une faible coloration au niveau du noyau. La mutation du NLS bipartite (NP/BPmut) ne semble pas affecter la localisation de la NP (**figure 53, panneau c**). Enfin, l'inactivation simultanée des deux NLS (NP/NCmut+BPmut) présente le même résultat que le mutant NP/NCmut (**comparaison entre les panneaux b et d de la figure 53**). Ceci suggère que le NLS non conventionnel est le signal utilisé par la NP lorsqu'elle est exprimée en absence d'autres

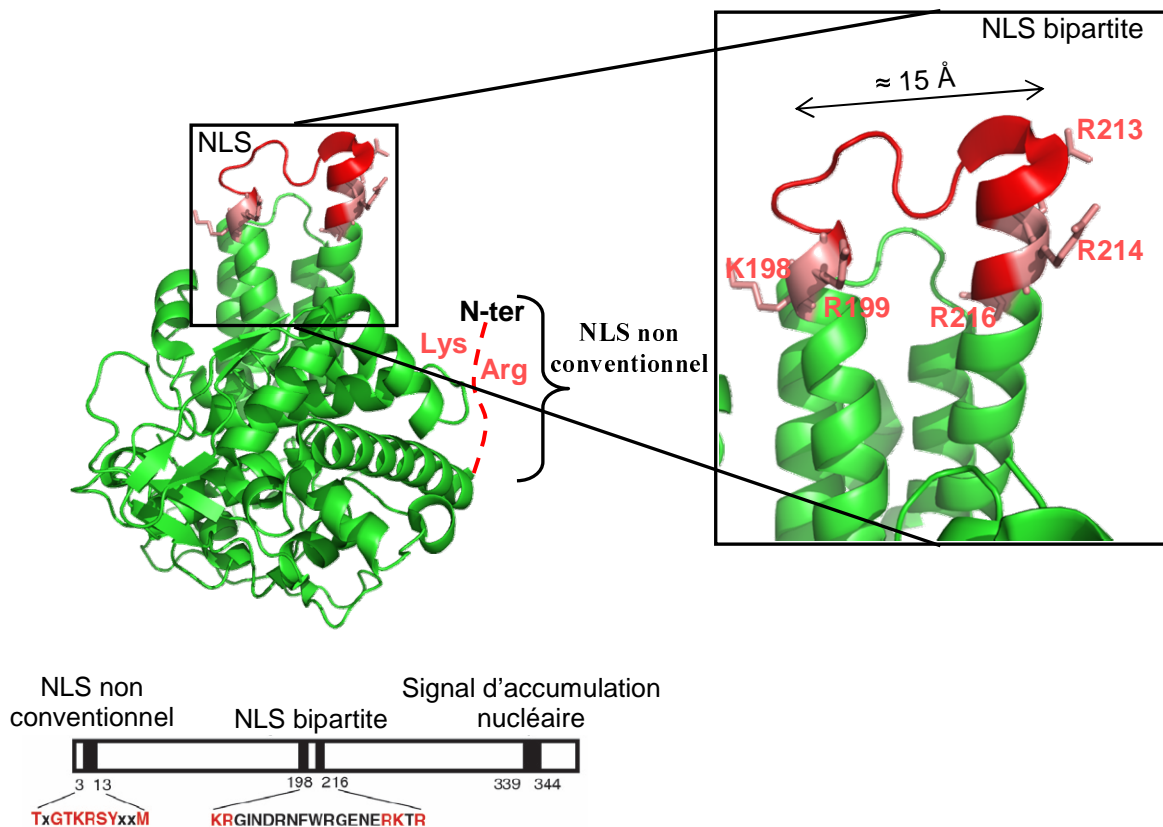
protéines virales ou en absence d'ARNv. Nous pouvons remarquer dans ces expériences que l'utilisation de la leptomycine B (LMB), inhibiteur connu du système d'export de Crm1, réduit la redistribution de la nucléoprotéine qui pourrait posséder un signal d'export nucléaire (**panneaux f, g, h et i figure 53**). Toujours dans le même article, des études similaires réalisées sur des cellules de foie de rat perméabilisées par la digitonine, renforce le fait que le NLS non conventionnel soit crucial pour l'import de la NP, et de plus, qu'il l'est également pour l'import de RNPv reconstituées (Cros *et al.*, 2005).



**Figure 53. Le NLS non conventionnel est crucial pour l'import nucléaire de la NP (d'après Cros *et al.*, 2005). À gauche, représentation schématique de la NP et des mutants de la NP du virus de la grippe. Les zones noires indiquent les signaux de localisation décrits : le NLS non conventionnel (Wang *et al.*, 1997), le NLS bipartite (Weber *et al.*, 1998) et le signal d'accumulation nucléaire (Davey *et al.*, 1985). Les acides aminés cruciaux pour l'activité des NLS sont représentés en rouge. Plusieurs de ces acides aminés ont été substitués par des alanines (en bleu) pour inactiver le NLS. A droite, localisation cellulaire de la NP et des mutants énoncés en a). Des cellules HeLa ont été transfectées avec un plasmide exprimant la NP ou les différents mutants puis fixées après 30 heures. La leptomycine B (LMB) a été ajoutée 12 heures suite à la transfection (panneaux de droite). La NP et les différents mutants ont été détectés par immunofluorescence grâce à l'utilisation de l'anticorps monoclonal anti-NP.**



Dans la structure cristallographique de la nucléoprotéine (Ye *et al.*, 2006), seul le NLS bipartite est visible, les vingt premiers acides aminés N-terminaux, portant le NLS non conventionnel étant trop flexibles, ils n'apparaissent pas dans la structure (**figure 54**). Dans la structure tridimensionnelle, le NLS bipartite dont la séquence est située au centre du polypeptide est exposé à la surface de la protéine, cependant, la distance séparant les motifs 198-**KR**-199 et 213-**RKTR**-216 est environ de 15 Å (**figure 54**) (Ye *et al.*, 2006). Cette longueur est nettement inférieure à la distance de séparation minimale de 28 Å nécessaire pour la fixation efficace d'un NLS bipartite (Fontes *et al.*, 2000). De ce fait, le signal bipartite ne semble pas approprié structuralement. Il est possible que la présence de ces acides aminés basiques soit fortuite dans le contexte d'un NLS, mais importante pour le site de fixation à l'ARN.



**Figure 54. Position des NLS sur la structure de la nucléoprotéine du virus de la grippe.** En rouge sont représentés et indiqués, les acides aminés constitutifs du NLS non conventionnel et du NLS bipartite.

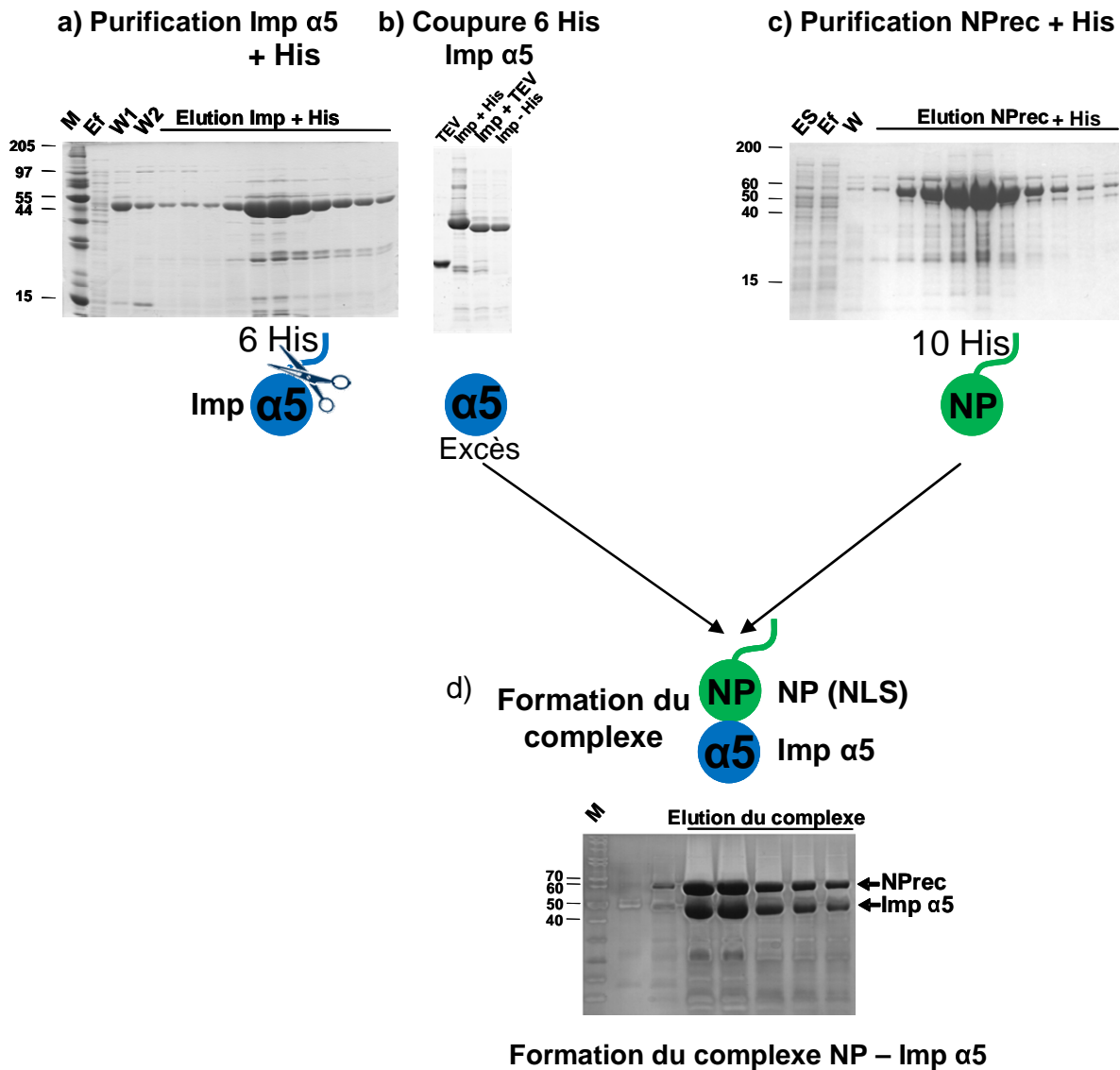
Différents travaux ont déjà montré que la nucléoprotéine isolée interagit *in vivo* (O'Neill *et al.*, 1995 ; Wang *et al.*, 1997) et *in vitro* (Melen *et al.*, 2003) avec plusieurs importines  $\alpha$  humaines ( $\alpha 1$ ,  $\alpha 3$  et  $\alpha 5$ ). Seule en solution, la NP a tendance à former des oligomères qui vont des dimères et trimères à de plus larges structures. Toutes ces espèces sont difficilement séparables et exploitables pour des études structurales. Notre objectif est de l'associer à l'importine  $\alpha 5$  humaine (acides aminés 66-512) pour la rendre monomérique et ainsi réaliser des essais de cristallogénèse. Bien que les structures à la fois de la nucléoprotéine et de l'importine  $\alpha 5$  humaine soient déjà résolues, l'obtention de cette structure nous permettrait de définir le site de fixation entre les deux molécules et par conséquent leur mécanisme d'interaction.

## 2. Le complexe importine $\alpha 5$ humaine / nucléoprotéine du virus de la grippe

### 2.1. Purification de la NPrec et de l'Imp $\alpha 5$ et formation du complexe NPrec/Imp $\alpha 5$

La stratégie employée pour former et purifier le complexe consiste à conserver une étiquette polyhistidine sur l'une des deux protéines. Ainsi, le complexe formé sera isolé par chromatographie d'affinité grâce à l'étiquette polyhistidine résiduelle. Du fait de la propension de la NP à polymériser, nous avons utilisé un excès d'importine  $\alpha 5$  pour déplacer l'équilibre polymères NP-NP au profit du dimère NP-Imp $\alpha 5$ . Et, pour retenir par affinité uniquement le complexe sans l'excès d'importine  $\alpha 5$ , l'étiquette polyhistidine sera portée par la protéine minoritaire, la NP.

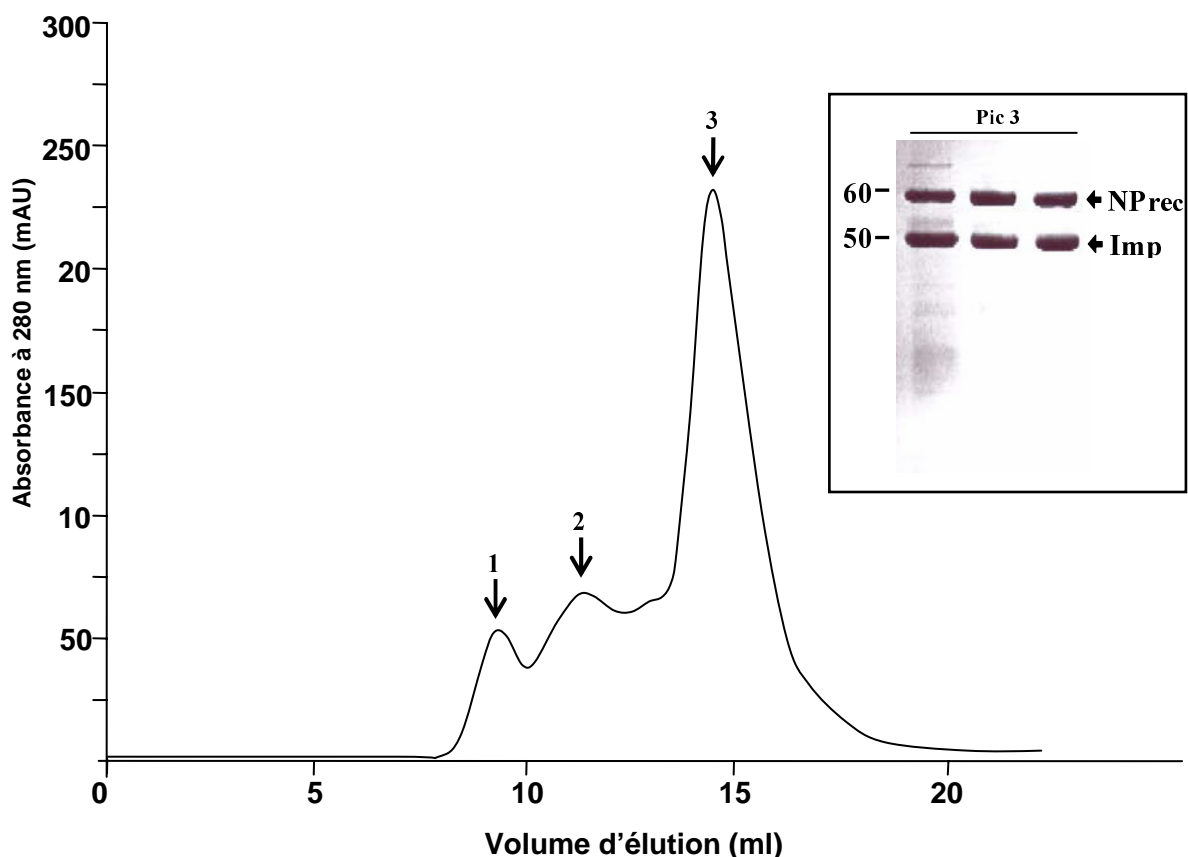
Nous avons exprimé indépendamment la nucléoprotéine recombinante (NPrec) ainsi que l'importine  $\alpha 5$  humaine (imp  $\alpha 5$ ), toutes deux munies d'une étiquette polyhistidine située à l'extrémité N-terminale. Les deux protéines sont ainsi purifiées par chromatographie d'affinité (**figure 55a et c**). L'étiquette polyhistidine de l'Importine  $\alpha 5$  est alors éliminée grâce à l'action de la protéase du Tobacco Etch Virus (TEV) (**figure 55b**). Afin de former le complexe, la nucléoprotéine est mise en présence de l'Imp  $\alpha 5$  en excès (3:1 ; Imp  $\alpha 5$ :NP). La NP est retenue par affinité grâce à sa séquence polyhistidine, et retient l'Imp $\alpha 5$  avec elle. Dans chacune des fractions de la **figure 55d**, sont présentes les 2 protéines, indiquant que le complexe a bien été formé. L'intensité des bandes sur gel de polyacrylamide nous permet de supposer que la stoechiométrie est de 1:1.



**Figure 55. Stratégie générale de formation du complexe entre la nucléoprotéine du virus de la grippe et l'importine  $\alpha 5$  humaine.** **a et c)** L'imp  $\alpha 5$  et la NP sont purifiées par chromatographie d'affinité via leur étiquette polyhistidine. **b)** Uniquement l'étiquette polyhistidine de l'imp  $\alpha 5$  est enlevée par l'action de la TEV. **d)** Le complexe est formé en ajoutant l'imp  $\alpha 5$  trois fois en excès par rapport à la NP puis isolé à nouveau par chromatographie d'affinité par le biais de l'étiquette polyhistidine de la NP. M = Marqueur moléculaire, Ef = effluent, W = lavage.

Nous réalisons ensuite une chromatographie d'exclusion de taille qui permet d'éliminer toutes les différentes formes d'oligomères de nucléoprotéines résiduelles. Ainsi les oligomères de taille supérieure sont élués dans le volume mort de la colonne, soit environ 9 ml, et dans le pic n°2 (**figure 56, pic 1 et 2**). Le pic majoritaire élué à 14,5 ml correspond à l'élué du complexe NPreC / Imp  $\alpha$ 5. D'après la courbe d'étalonnage de la colonne S200 analytique (non présentée), un tel volume correspond bien à la taille attendue du complexe : 58 kDa pour la NP recombinante et 50 kDa pour l'importine  $\alpha$  recombinante, soit environ 108 kDa au total. L'analyse du pic n°3 sur gel dénaturant montre que ce pic est constitué des protéines Imp  $\alpha$ 5 et NP et la stoechiométrie semble être de 1:1 (**figure 56**).

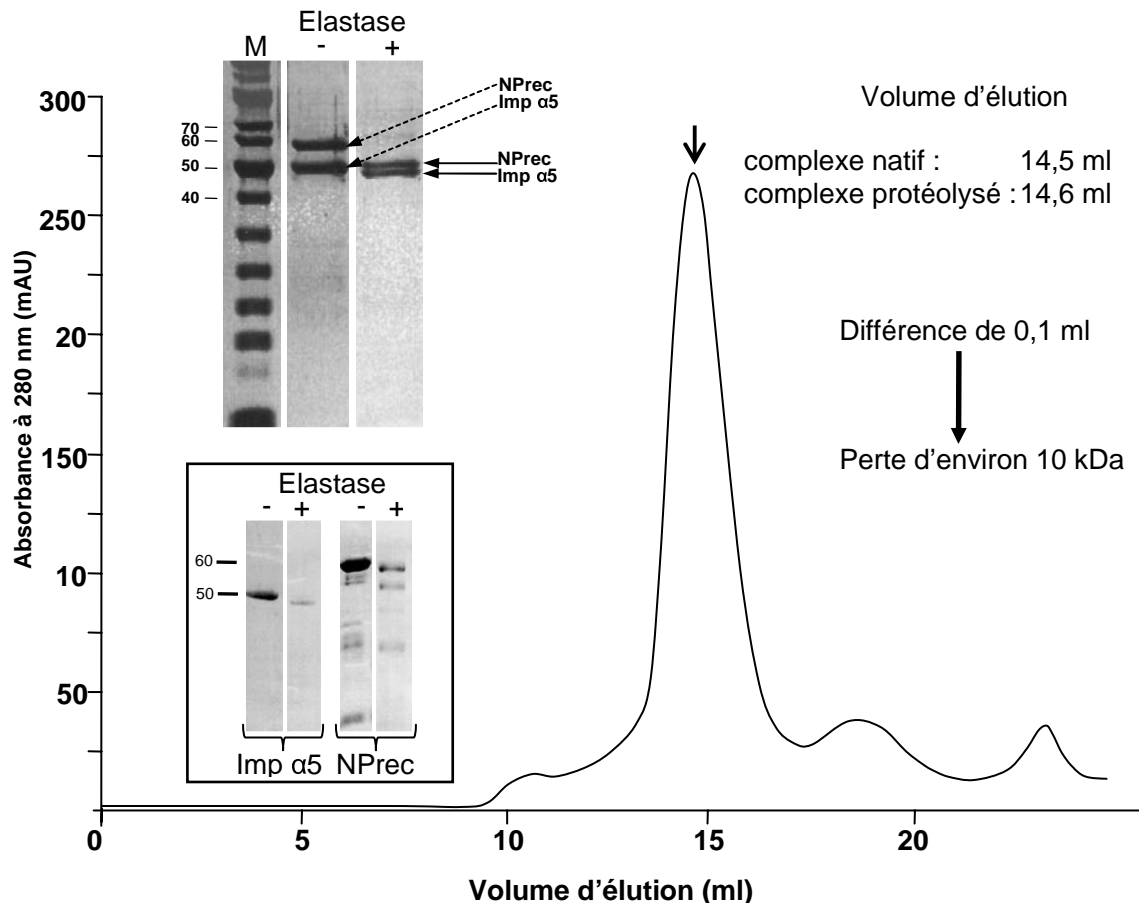
Ainsi, grâce à cette expérience nous avons réussi à éliminer l'état d'oligomérisation de la NP en l'associant à l'importine  $\alpha$ 5 pour former un complexe hétérodimère 1:1.



**Figure 56. Formation d'un hétérodimère entre la nucléoprotéine recombinante (NPreC) et l'importin  $\alpha$ 5 (Imp  $\alpha$ 5).** Profil d'élué du complexe Imp  $\alpha$ 5 humaine / NPreC du virus de la grippe (A/PR/8/34) par chromatographie d'exclusion de taille sur une colonne Superdex 200 analytique (GE Healthcare). Les flèches indiquent (1) le volume mort élué à 9 ml, (2) les polymères de NP élués à 11,3 ml et (3) le volume d'élué du complexe importine  $\alpha$ 5 / nucléoprotéine à 14,5 ml (*MM 108 kDa*). Et analyse par électrophorèse sur gel de polyacrylamide (SDS-PAGE 12% Tris-Tricine) du pic n° 3 correspondant au complexe.

## 2.2. Protéolyse limitée du complexe NPrec/Imp $\alpha 5$

Après concentration du complexe jusqu'à  $10 \text{ mg.ml}^{-1}$ , nous avons effectué des essais de cristallogène sur le complexe qui malheureusement n'ont pas abouti à la formation de cristaux. Nous avons alors, décidé d'effectuer une protéolyse limitée afin d'éliminer les parties flexibles qui souvent gênent le processus de cristallisation. A cette fin, le complexe NPrec/imp  $\alpha 5$  purifié sur Superdex 200 est protéolysé par l'élastase puis à nouveau purifié sur S200 afin d'éliminer les fragments de protéolyse ainsi que l'élastase (protéine de 26 kDa). Le volume d'éluion du complexe protéolysé ne varie pas de façon significative par rapport à celui du complexe natif (14.5 et 14.6 ml respectivement), ce qui indique que le complexe a relativement peu changé de poids moléculaire après la protéolyse (**figure 57**).



**Figure 57. Profil d'éluion du complexe NPrec/Imp $\alpha 5$  après protéolyse limitée par l'élastase.** Le profil d'éluion est obtenu par chromatographie d'exclusion de taille sur une colonne Superdex 200 analytique (GE Healthcare). Le pic majoritaire est indiqué par la flèche et correspond à l'éluion du complexe à 14,6 ml. Gel du haut, analyse du complexe, avant et après la protéolyse. Gel du bas, contrôle des protéolyses individuelles de l'Imp $\alpha 5$  et de la NP. Gels dénaturants SDS, 12% polyacrylamide. M = Marqueur moléculaire.

Le produit de la digestion est analysé par gel dénaturant (**gel du haut, figure 57**). Nous avons vu que lorsque que nous réalisons la protéolyse de chacune des deux protéines individuellement, il se produisait une dégradation importante de chacune d'elles (**gel du bas figure 57**). Ainsi, l'association des ces deux protéines favorise la formation d'un complexe stable dont seules les extrémités sont accessible à la protéolyse. Nous avons analysé par spectrométrie de masse le complexe protéolysé. Il s'avère que seule la masse d'une des deux protéines a pu être identifiée à 48 503 Da, la seconde protéine n'a pu être détectée par l'appareil. L'identification des bandes protéolysées par l'élastase a été effectuée par séquençage N-terminal. La bande à 48,5 kDa, correspond à l'importine  $\alpha 5$  protéolysée (50 kDa avant protéolyse). L'analyse du séquençage en N-terminal, de la protéine protéolysée, nous a révélé que la séquence -NMEMAPGG- était en accord avec celle de l'importine  $\alpha 5$  à partir de l'acide aminé 14 de notre gène. La différence d'environ 1,5 kDa correspond bien à une perte de 13 acides aminés, nous indiquant que l'importine  $\alpha 5$  a été coupée en N-terminal.

Concernant la bande à environ 51 kDa environ, le séquençage n'a pas permis d'identifier les premiers acides aminés N-terminaux. Ceci est probablement dû à une protection située à l'extrémité N-terminale de la protéine qui empêche ainsi la dégradation d'Edman de se réaliser. Cependant par élimination, cette bande est supposée correspondre à la NP protéolysée (58 kDa avant protéolyse) et la perte des 7 kDa a probablement dû se produire dans la région C-terminale. La NP et l'importine  $\alpha 5$  en complexe ont perdu quelques kDa, le poids moléculaire du complexe passe de 108 kDa à environ 99 kDa. Le léger décalage observé entre les profils d'élution du complexe intact et complexe protéolysé n'est que de 0.1 ml. Ce léger retard reflète donc la perte d'environ 10 kDa du complexe initial. Nous avons alors réalisé de nouveaux essais de cristallogénèse et avons obtenu de façon reproductible des cristaux qui n'ont cependant pas diffracté suite à leur exposition sous le faisceau de rayons X. Il arrive que l'arrangement des molécules à l'intérieur du cristal ne soit pas suffisamment régulier, entraînant ainsi la formation d'un cristal mais sans diffraction.

Une autre possibilité pourrait expliquer l'absence de cristallisation. Des expériences de mutagenèse dirigée sur l'importine  $\alpha 5$  humaine ont montré l'importance des ARM 7, 8 et 9 pour la reconnaissance avec la NP ([Melen \*et al.\*, 2003](#)). Ceci nous indique que le NLS de la NP se fixe au site mineur de l'importine  $\alpha 5$  situé sur la région C-terminale. Suite aux structures d'importines  $\alpha$  de la levure avec différents motifs NLS, nous savons que les signaux monopartites peuvent se fixer au site majeur de l'importine  $\alpha$  (en N-ter) et au site mineur (en C-ter). Concernant les signaux bipartites, le plus grand des deux clusters se fixe au site majeur de l'importine  $\alpha$  et le plus petit cluster, au site mineur ([Conti \*et al.\*, 1998](#)). Bien qu'il ne soit

pas conventionnel, le NLS de la NP n'est constitué que d'un seul cluster d'acides aminés basiques, comme l'est un NLS monopartite. De ce fait, on pourrait envisager que la NP ne soit pas uniquement fixée au site mineur comme cela est décrit par [Melen \*et al.\*, \(2003\)](#), mais qu'une faible quantité puisse se fixer au site majeur. Ceci pourrait engendrer un complexe en solution non homogène empêchant la cristallisation.

### 3. Caractérisation du complexe NPmono-Imp $\alpha$ 5

En décembre 2006, la structure publiée de la nucléoprotéine du virus de la grippe, à montré que les acides aminés R416, E339 et la région située entre les acides aminés 402 à 428 étaient impliqués dans l'oligomérisation de la protéine ([Ye \*et al.\*, 2006](#)). Grâce à la mutation R416A de la nucléoprotéine nous disposons ainsi de la NP sous une forme monomérique, NPmono (**décrite dans le chapitre 2**).

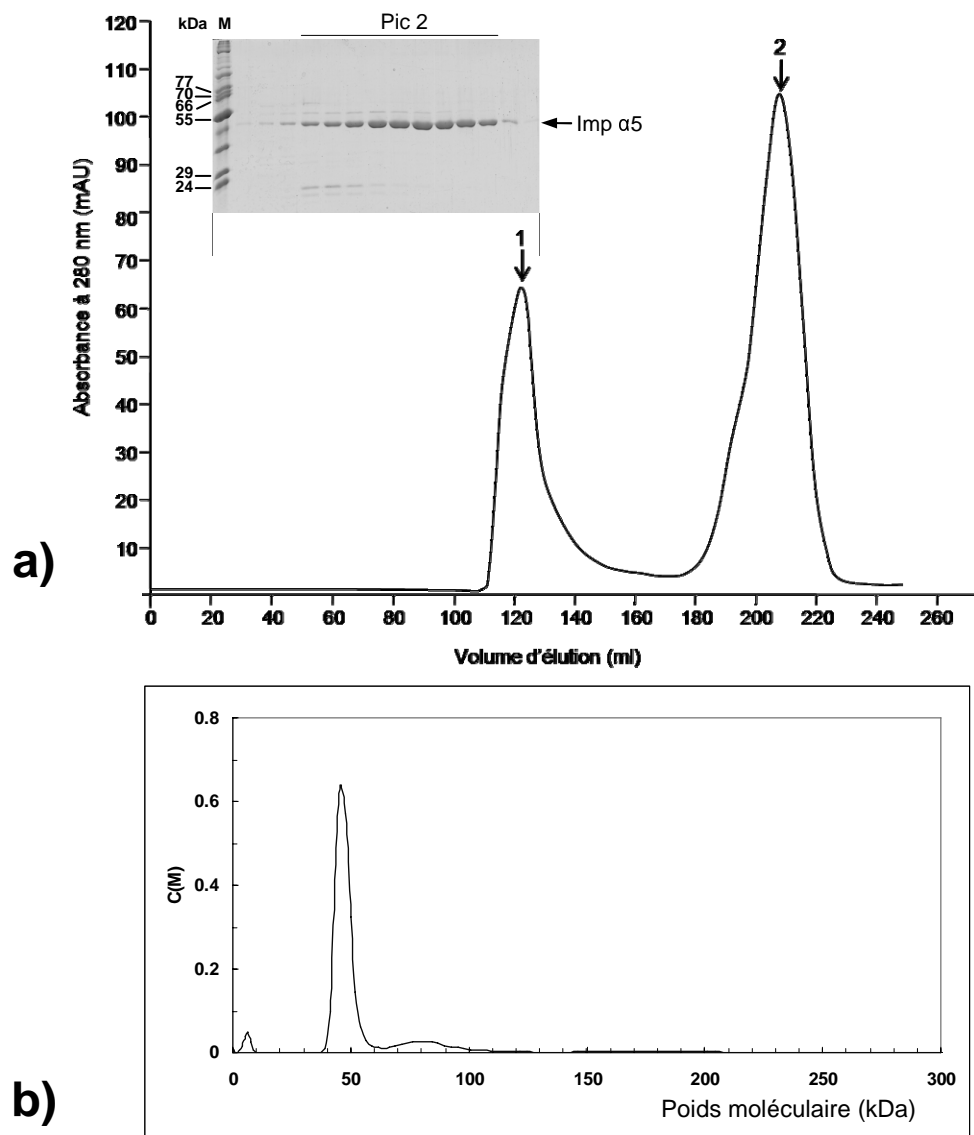
#### 3.1. Caractérisation de l'importine $\alpha$ 5 humaine et de la NP monomérique (NPmono)

Dans nos mains, à la suite d'une chromatographie d'exclusion de taille sur une colonne Superdex 200 HiLoad, l'importine  $\alpha$ 5 humaine est éluée selon deux pics (**figure 58a**). Le premier élué à 120 ml, représente la protéine agrégée et le second qui est le pic majoritaire est élué à 205 ml. Un tel volume correspond, selon la calibration de la colonne, à une molécule d'environ 70 kDa et non de 50 kDa comme nous pourrions nous y attendre. Les protéines de forme allongée comme l'importine  $\alpha$ 5, dont nous avons résolu la structure, présentent souvent une masse moléculaire apparente supérieure à celle attendue. Suite à l'analyse sur gel dénaturant (**figure 58a**) nous pouvons observer, au niveau des premières fractions d'éluion, deux contaminants aux environs de 23 kDa co-élués avec l'importine et certainement responsables de la légère épaulement observée sur le profil d'éluion de la protéine.

Afin de conclure sur l'état d'oligomérisation de l'importine  $\alpha$ 5 humaine, nous avons ensuite réalisé des expériences d'ultracentrifugation analytique (AUC pour analytical ultracentrifugation) (**figure 58b**). Cette méthode, décrite dans le chapitre 2, nous permet de déterminer directement la masse de la protéine et s'il s'agit d'un oligomère ou d'un complexe, le nombre de sous-unités impliquées et sa stœchiométrie. L'importine  $\alpha$ 5 est ainsi sédimentée pour un poids moléculaire près de 50 kDa, ce qui nous indique que la protéine en solution se comporte bel et bien comme un monomère.

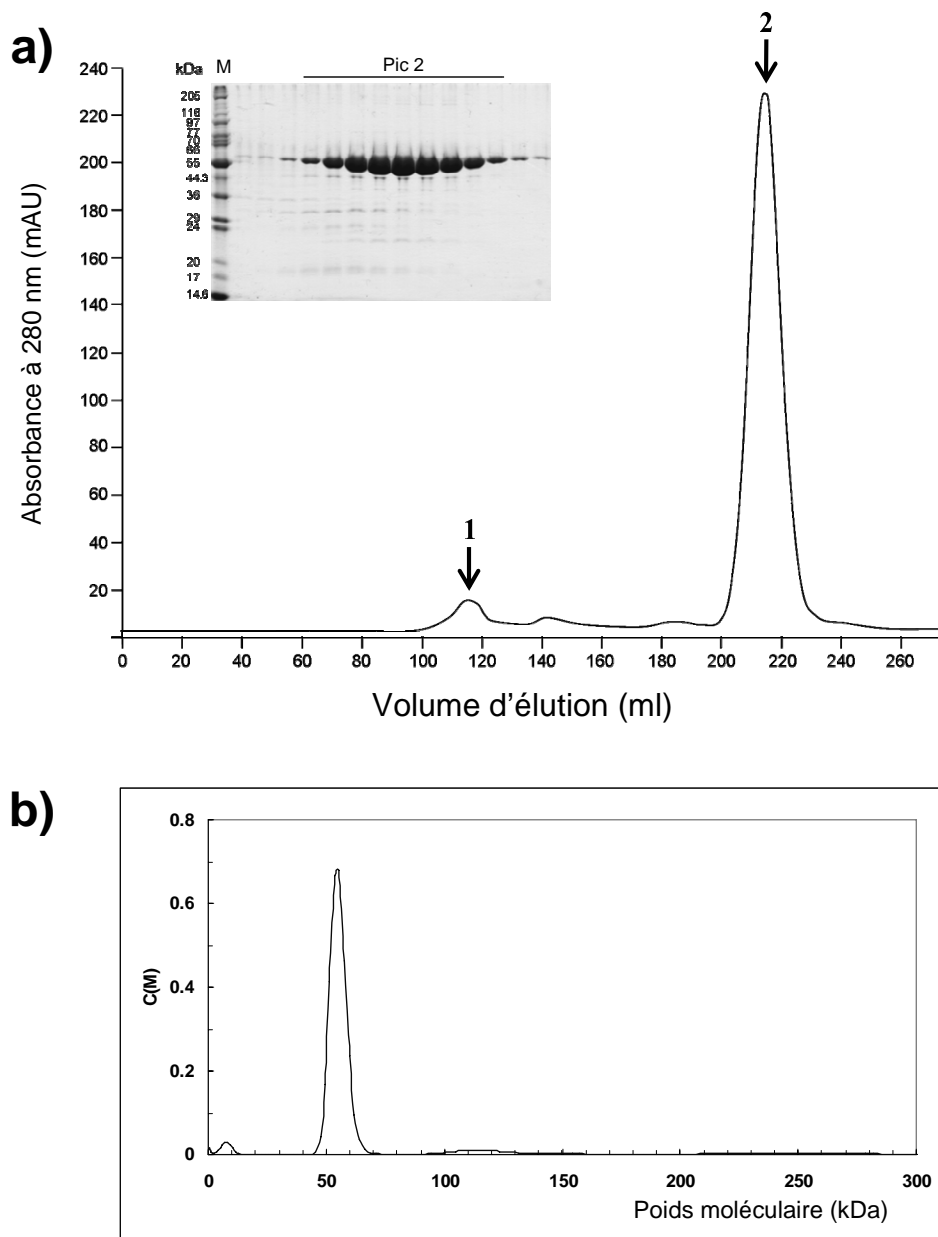
Décrit dans le **chapitre 2**, l'état d'oligomérisation de la nucléoprotéine du virus de la grippe peut être affecté par certaines mutations. Ainsi, le mutant R416A de la nucléoprotéine (NPmono), a été caractérisé et les résultats analysés dans le **chapitre 2**. Nous avons vu par chromatographie d'exclusion de taille et par AUC que le mutant R416A était monomérique avec comme masse moléculaire 56 kDa (**figure 59a et b**).

Une nouvelle caractérisation de notre complexe NP-Imp $\alpha$ 5 est maintenant possible, mêlant deux protéines dont les caractéristiques individuelles sont celles de deux protéines monomériques.



**Figure 58. Caractérisation de l'importine  $\alpha$ 5 humaine (sans His).** Profil d'éluion de l'importine  $\alpha$ 5 humaine (Imp  $\alpha$ 5) par chromatographie d'exclusion de taille sur une colonne Superdex 200 HiLoad (GE Healthcare). Les flèches indiquent (1) le volume mort et (2) le volume d'éluion de l'importine  $\alpha$ 5 (*MM* 50 kDa). Et analyse par électrophorèse sur gel de polyacrylamide (SDS-PAGE 12% Laemmli) du pic n°2 correspondant à l'importine  $\alpha$ 5. M = Marqueur moléculaire. **b)** Résultats d'ultracentrifugation analytique pour l'importine  $\alpha$ 5 humaine. La distribution de coefficient de sédimentation provient du profil de vitesse de sédimentation analysé à l'aide du programme SEDIFT. La concentration en protéine était de 0,8 mg/ml.



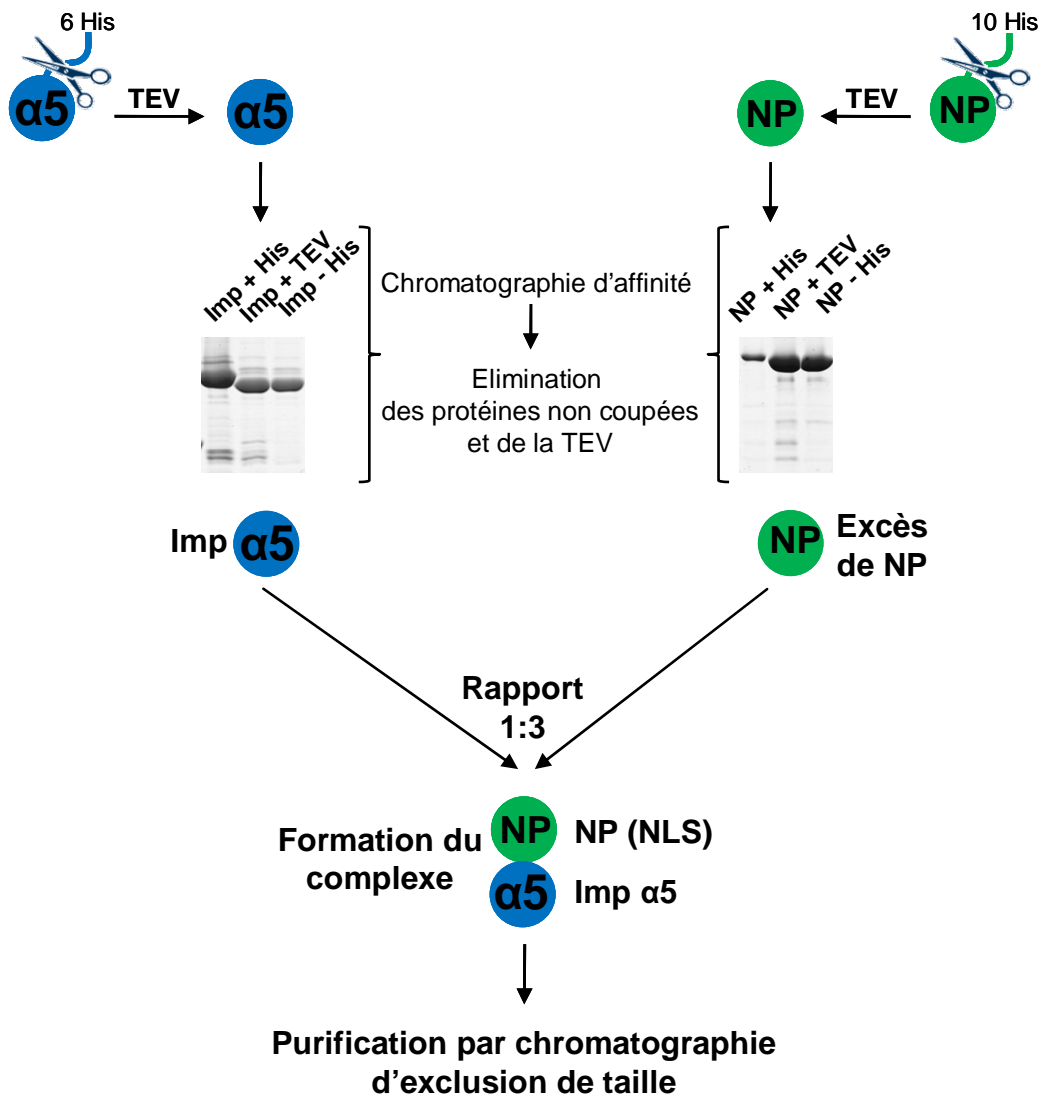


**Figure 59. Caractérisation du mutant R416A de la nucléoprotéine (NP mono) (sans His).** **a)** Profil d'élution de la NP R416A (NPmono) du virus de la grippe (A/PR/8/34) par chromatographie d'exclusion de taille sur une colonne Superdex 200 HiLoad (GE Healthcare). Les flèches indiquent (1) le volume mort et (2) le volume d'élution de la NP (*MM* 56 kDa). Et analyse par électrophorèse sur gel de polyacrylamide (SDS-PAGE 12% Laemmli) du pic n°2 correspondant au mutant R416A de la NP. M = Marqueur moléculaire. **b)** Résultats d'ultracentrifugation analytique pour le mutant NP R416A. La distribution de coefficient de sédimentation provient du profil de vitesse de sédimentation analysé à l'aide du programme SEDIFT. La concentration en protéine était de 0,8 mg/ml.

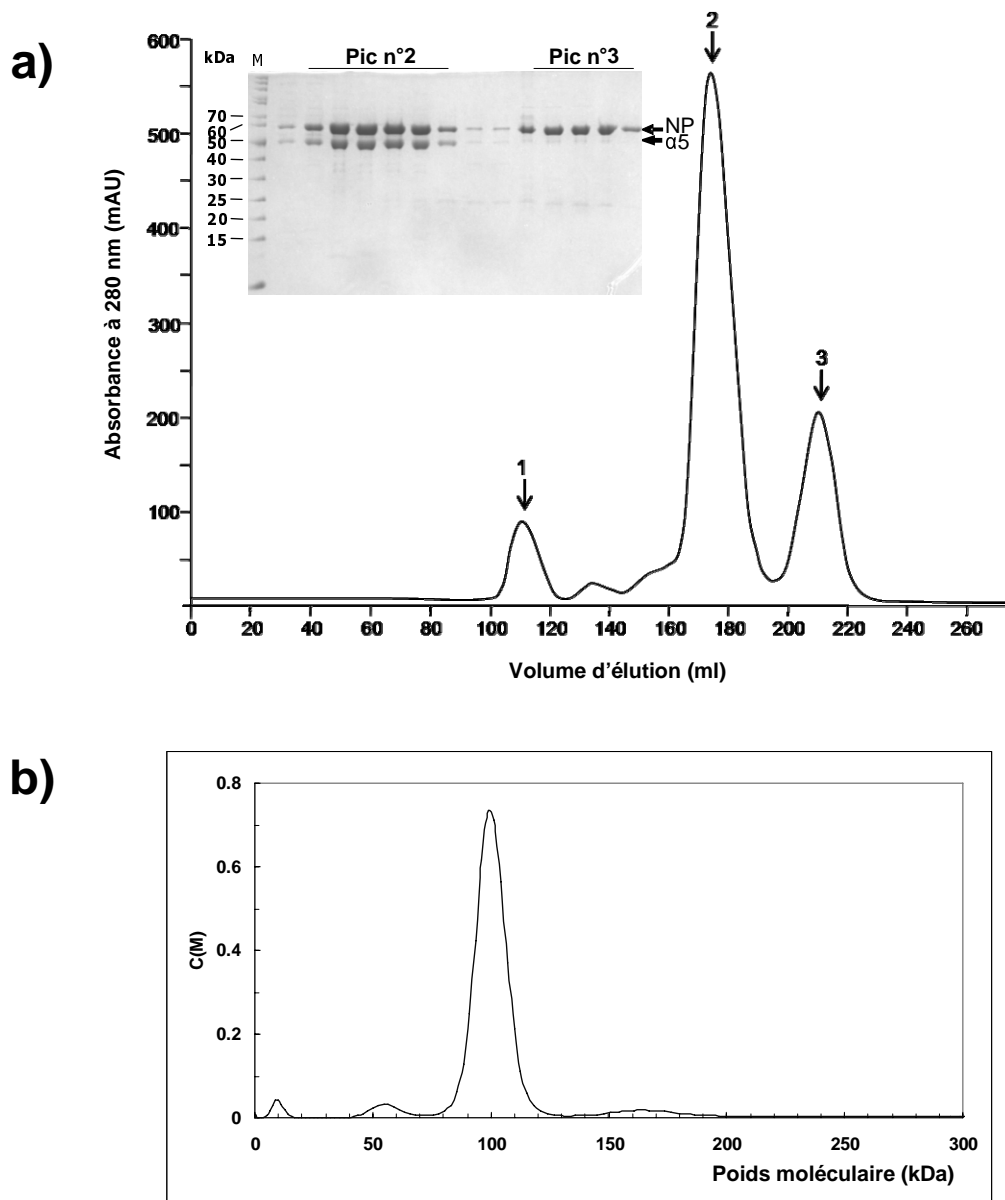
### 3.2. Formation et caractérisation du complexe Imp $\alpha$ 5 / NPmono

La formation du nouveau complexe (NPmono-Imp $\alpha$ 5) passe par la mise en place d'une nouvelle stratégie (**figure 60**). Tout d'abord, les étiquettes polyhistidines des deux protéines sont enlevées dès le départ grâce à l'action de la protéase TEV. Contrairement à la formation du complexe NPrec-Imp $\alpha$ 5 où l'importine était excédentaire, dans la formation du complexe NPmono-Imp $\alpha$ 5, c'est la NPmono qui est trois fois plus abondante. En effet, elle ne présente plus l'inconvénient de polymériser et les qualités d'expression obtenues pour la NPmono sont supérieures à celles de l'Imp $\alpha$ 5. Après injection sur une colonne de filtration sur gel, (superdex 200 HiLoad), le mélange est réparti selon trois pics (**figure 61a**). Les agrégats sont éliminés dans le premier pic à 115 ml d'élution, le complexe est élué dans le deuxième à 174 ml et le troisième pic élué à 215 ml représente l'excédent de NPmono, c'est-à-dire la protéine non complexée et donc monomérique. Analysé par gel dénaturant (**figure 61a**), le complexe ainsi formé est très pur. Nous avons ensuite réalisé une expérience d'ultracentrifugation analytique (AUC) sur ce complexe et les résultats révèlent un poids moléculaire autour de 100 kDa (**figure 61b**).

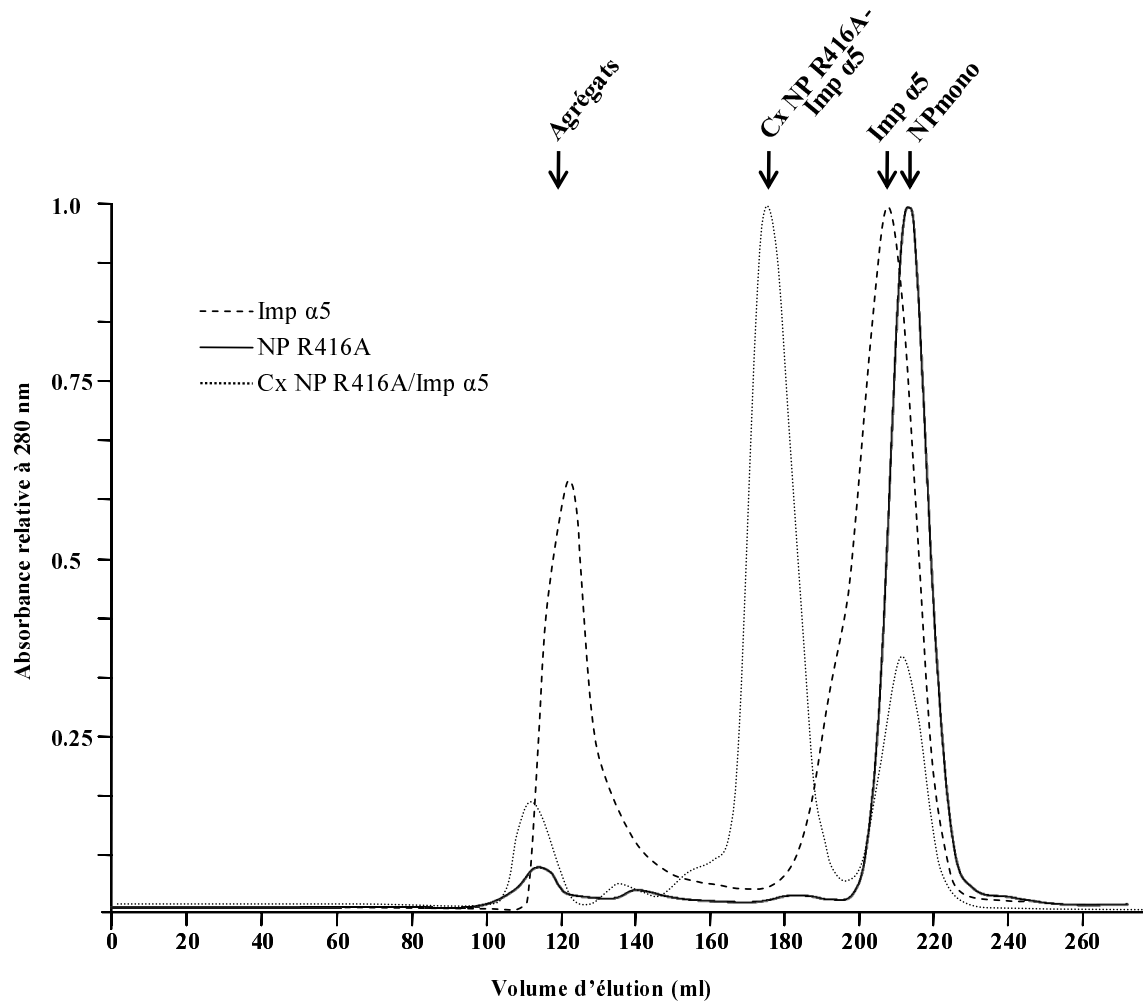
Le bilan effectué sur l'ensemble des expériences de filtration sur gel nous montre comment les pics correspondants à la NPmono (—) et l'Imp $\alpha$ 5 (---) (215 ml et 205 ml respectivement) sont décalés vers un pic à 174 ml lorsque le complexe est formé (····) (**figure 62**). Nous avons donc été en mesure de produire un complexe issu de l'association de deux protéines monomériques et dont la stœchiométrie est de 1:1. Même après avoir obtenu un matériel biologique d'aussi bonne qualité, nous n'avons pas été en mesure de produire des cristaux. Il est possible que la grande flexibilité des protéines complexées au niveau du site d'interaction empêche la formation de cristaux.



**Figure 60. Stratégie de formation du complexe entre la nucléoprotéine monomérique du virus de la grippe et l'importine  $\alpha 5$  humaine.** Les étiquettes polyhistidines de l'imp  $\alpha 5$  et la NPmono sont coupées grâce à l'action de la protéase TEV. La protéase TEV et les protéines non coupées sont éliminées par chromatographie d'affinité. Enfin, le complexe est formé en ajoutant la NPmono trois fois en excès à l'Imp $\alpha 5$ .



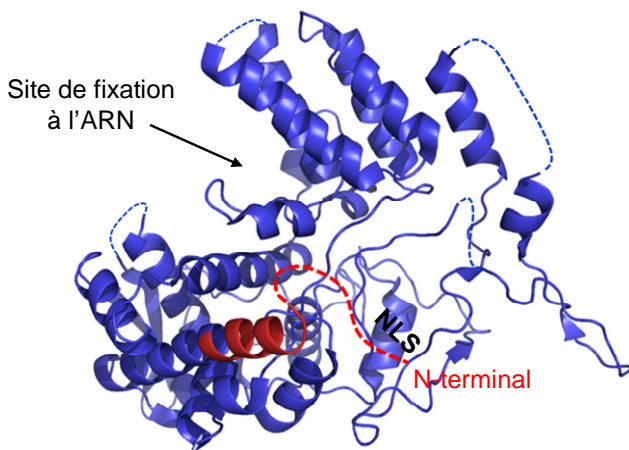
**Figure 61. Formation du complexe importine 5 humaine / nucléoprotéine du virus de la grippe.** Profil d'éluion du complexe Imp  $\alpha 5$  humaine / NP R416A du virus de la grippe (A/PR/8/34) par chromatographie d'exclusion de taille sur une colonne Superdex 200 HiLoad (GE Healthcare). Les flèches indiquent (1) le volume mort, (2) le volume d'éluion du complexe importine  $\alpha 5$  / nucléoprotéine (*MM* 106 kDa) et (3) l'excès de nucléoprotéine. Et analyse par électrophorèse sur gel de polyacrylamide (SDS-PAGE 12% Tris-Tricine) des pics n° 2 et 3 correspondant au complexe et à l'excès de NP respectivement. **b)** Résultats d'ultracentrifugation analytique pour le complexe Imp  $\alpha 5$  / NPmono. La distribution de coefficient de sédimentation provient du profil de vitesse de sédimentation analysé à l'aide du programme SEDIFT. La concentration en protéine était de 0,8 mg/ml.



**Figure 62. Formation du complexe NPmono-Imp $\alpha$ 5.** Profil d'élution de l'importine  $\alpha$ 5 humaine (Imp  $\alpha$ 5) (---), du mutant NP R416A (NPmono) du virus de la grippe (A/PR/8/34) (—) et du complexe Imp  $\alpha$ 5 / NPmono (····) par chromatographie d'exclusion de taille sur une colonne Superdex 200 HiLoad (GE Healthcare). Les différents pics sont indiqués.

### 3.3. Affinité du complexe NP-Imp $\alpha$ 5 pour l'ARN

D'un point de vue structural il serait intéressant de savoir si le complexe, une fois formé, peut se fixer à l'ARN. La structure de la NP, résolue par [Ye et al., \(2006\)](#), montre entre les deux domaines (supérieur et inférieur) une plateforme ouverte chargée positivement (**figure 63**). Celle-ci représente certainement le site de fixation à l'ARN. Il est dommage de constater que la structure cristallographique ne commence qu'à partir de l'acide aminé 21, laissant le NLS non conventionnel indéterminé. Cependant, lorsque l'on regarde la première hélice N-terminale visible, on peut modéliser la région NLS (**en pointillés rouges sur la figure 63**). Même si cette région n'est pas définie avec certitude, elle est néanmoins exposée vers la surface de la molécule, au niveau du domaine inférieur de la NP.

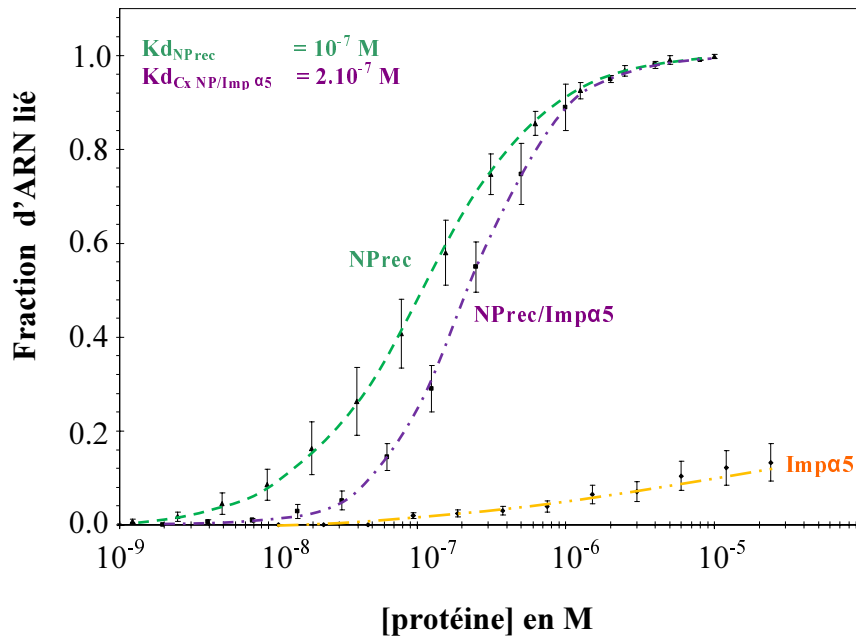


**Figure 63. Structure cristallographique de la NP du virus de la grippe, souche A/WSN/33 (code pdb : 2IQH) (Ye et al., 2006).** Le sillon potentiel de fixation à l'ARN est indiqué par la flèche et l'extrémité N-terminale est représentée en rouge. Les régions non visibles dans la structure sont représentées en pointillés.

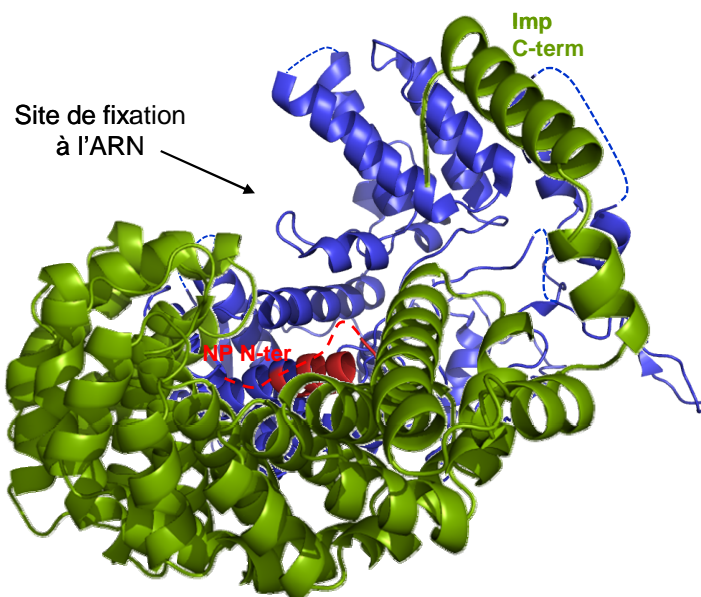
Il nous semblait intéressant de voir, lorsque l'importine  $\alpha$ 5 est fixée au NLS de la NP, si la plateforme était encore accessible pour l'ARN. Pour répondre à cette question nous avons regardé la différence d'affinité de la nucléoprotéine pour l'ARN, lorsque celle-ci est seule ou liée à l'importine  $\alpha$ 5. Afin d'étudier ces interactions, nous avons déterminé l'affinité de la NPrec seule, du complexe NPrec-Imp $\alpha$ 5 et de l'Imp $\alpha$ 5 seule pour l'ARN *panhandle*, grâce à des expériences de *filter binding assays* (FBA). Les résultats sont présentés dans la **figure 64**. La NPrec, comme cela a déjà été montré dans le chapitre 2, se fixe à l'ARN avec une affinité de 100 nM (**courbe ---- figure 64**). Lorsque la NPrec est associée à l'Imp $\alpha$ 5 nous pouvons déterminer un Kd de 200 nM (**courbe -.-.- figure 64**). Nous pouvons voir en contrôle, qu'aucune activité de liaison à l'ARN n'a été détectée pour l'Imp $\alpha$ 5 seule (**courbe ..— figure 64**).

Il faut prendre avec précaution ces résultats. Ici nous comparons la NPrec (qui forme de petits oligomères) avec la NPrec associée à l'importine, donc dans un état où l'oligomérisation de la NP est éliminé. Or, dans le chapitre 2 nous avons déterminé que la NPmono (mutation R416A) présentait une affinité pour l'ARN de 2  $\mu$ M, c'est-à-dire 10 fois moins bonne que celle de NPrec associée à l'importine  $\alpha 5$ . Il est parfois assez difficile de déterminer avec exactitude les concentrations de complexes protéiques. En effet, les acides aminés responsables de l'absorbance (tryptophanes et tyrosines) d'une protéine peuvent être masqués suite à la formation du complexe et modifier ainsi la valeur de la concentration. D'autre part, le fait que la NPrec dans le complexe puisse encore se lier à l'ARN pourrait éventuellement nuancer l'effet coopératif de la NP (mentionné dans le chapitre 2) au profit d'une implication de l'arginine 416. Il serait envisageable que la mutation de cet acide aminé en alanine apporte un changement de conformation de la NP modifiant ainsi l'ouverture du site de fixation à l'ARN et diminuant de ce fait son interaction. Les structures des nucléoprotéines de la rage ou de VSV en complexe avec l'ARN montrent une "mâchoire" davantage fermée. Il serait possible que l'acide aminé 416 soit impliqué au niveau de la charnière de la NP et que sa mutation en alanine produise peut-être une fermeture de la mâchoire. Quoiqu'il en soit, nos résultats montrent que la nucléoprotéine associée à l'importine  $\alpha 5$  est capable de se fixer à l'ARN. Des expériences complémentaires pourraient départager ces hypothèses.

Ainsi, le fait que le complexe puisse se lier à l'ARN apporte une réponse davantage structurale que biologique. En effet, si l'ARN est capable de s'associer au complexe d'import, alors l'importine  $\alpha 5$  doit être positionnée de façon à laisser accessible le sillon de fixation à l'ARN, de ce fait, l'Imp $\alpha 5$  serait située plutôt au niveau de la partie inférieure de la NP (**figure 65**). Bien évidemment ceci n'est qu'un modèle et seule la structure du complexe d'import, NP-Imp $\alpha 5$ , permettra de conclure définitivement sur la géométrie de ces deux molécules assemblées ainsi que sur de leur site d'interaction.



**Figure 64. Activité de fixation à l'ARN par le complexe nucléoprotéine / importine  $\alpha5$ .** Courbes de titration réalisées à partir de "filter binding assays" (FBA) sur l'ARN de 81 bases (panhandle) avec la NPreC (---), le complexe NPreC/Imp $\alpha5$  (-.-.-) et l'importine  $\alpha5$  seule (.....). Le coefficient de dissociation apparent ( $Kd$ ) correspond à la concentration de la protéine à 50 % de rétention de l'ARN.



**Figure 65. Hypothétique modèle du complexe NP-Imp $\alpha5$ .** L'importine  $\alpha5$  en vert est associée au niveau de la région inférieure de la nucléoprotéine en bleue. L'interaction implique le signal NLS non conventionnel, modélisé en pointillés rouges sur la NP.



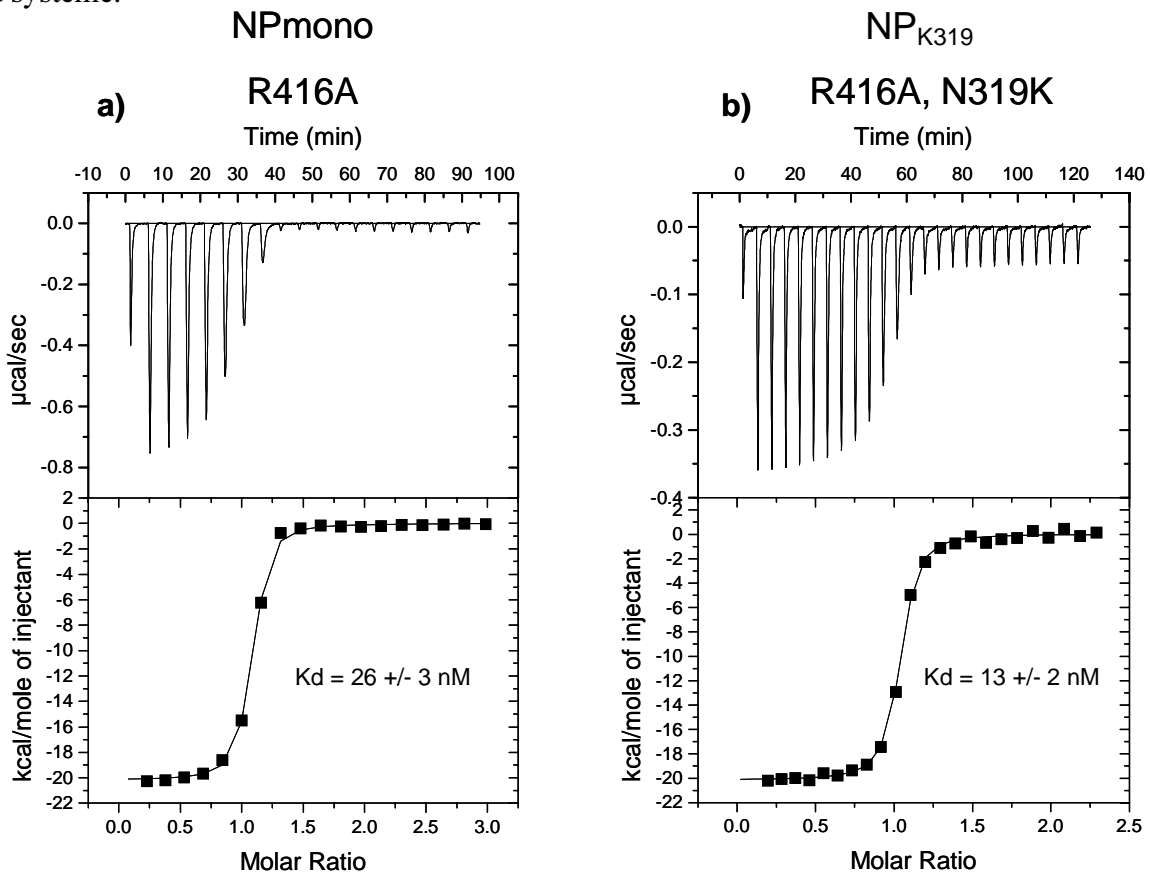
#### 4. Affinité de la NP pour l'importine $\alpha 5$ , importance de l'acide aminé 319

Sur l'ensemble des différentes importines  $\alpha$  humaines qui existent, chacune interagit avec différents partenaires cellulaires ou extracellulaires. Les travaux de [Melen \*et al.\*, \(2003\)](#) basés sur différentes importines  $\alpha$  humaines apportent des résultats intéressants. Ils ont montré que la NP était capable de se fixer aux importines  $\alpha 1$ ,  $\alpha 3$  et  $\alpha 5$ . Cependant, l'interaction avec l'importine  $\alpha 3$  était moins bonne qu'avec les importines  $\alpha 1$  et  $\alpha 5$ . Les importines sont donc capables d'interagir avec différentes classes de protéines à transporter et présentent des affinités variables pour chacune d'elles.

Les travaux de [Gabriel \*et al.\*, \(2005\)](#), déjà mentionnés au cours du **chapitre 1**, faisaient état de neuf mutations conduisant un virus de type aviaire à s'adapter au mammifère (la souris). Parmi celles-ci, une s'avère être également située sur la nucléoprotéine, il s'agit de la mutation N319K. Toujours dans les travaux de [Gabriel \*et al.\*, \(2005\)](#), les auteurs ont noté que cette mutation augmentait l'activité de la polymérase dans les cellules de mammifères mais pas dans les cellules aviaires. Plus récemment, la même équipe a identifié l'importine  $\alpha 1$  humaine, en tant que facteur impliqué dans la spécificité de l'hôte ([Gabriel \*et al.\*, 2008](#)). Ainsi, ils ont montré que la mutation N319K améliorait la fixation de la NP à l'importine  $\alpha 1$  dans des cellules de mammifères. De plus, ils ont pu observer que cette mutation augmentait le transport nucléaire de la NP dans des cellules de mammifères sans pour autant l'augmenter dans des cellules aviaires. Sur l'ensemble des séquences publiées de la NP, la lysine en position 319 n'apparaît que 19 fois sur 1 101 séquences, toutes les autres contiennent une asparagine. Sur ces 19 NP, 14 proviennent d'isolats d'humains et de chevaux, et 5 sont hautement pathogènes chez le poulet et leur virulence est proche de celle du sous-type H5N1 (A/HK/485/97) ([Gabriel \*et al.\*, 2005](#)).

Dans nos recherches, nous nous sommes intéressés à l'impact que pourrait produire cette mutation N319K sur l'interaction entre la NP et l'importine  $\alpha 5$  humaine. Concernant la nucléoprotéine, nous disposons du gène de la NP du virus la grippe de souche A/PR/8/34 et de sous-type H1N1, adaptée à l'homme, sur laquelle la mutation N319K de la nucléoprotéine n'est pas présente. Afin de déterminer l'affinité entre ces deux molécules, nous avons employé la méthode de calorimétrie de titration isotherme (*Isothermal Titration Calorimetry*), ITC. Le principe de l'ITC est le suivant : une macromolécule située dans la cellule de mesure d'un calorimètre isotherme est progressivement saturée, à température constante, par l'injection d'aliqotes d'un ligand à partir d'une seringue. A chaque ajout de ligand correspond un

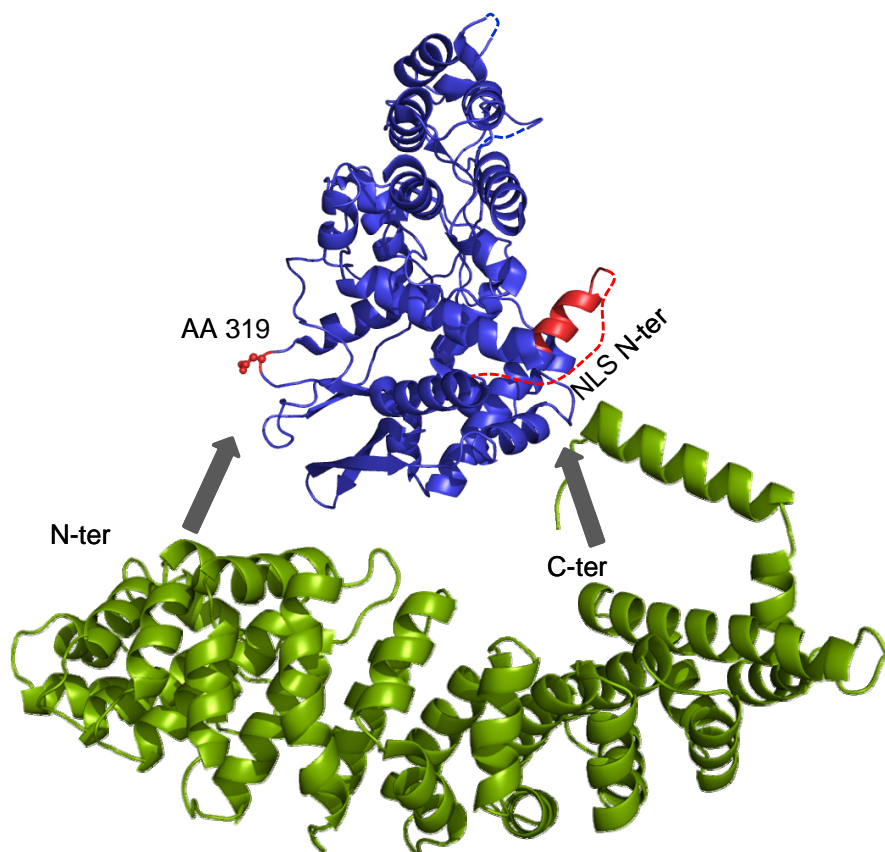
échange thermique signant l'interaction macromolécule-ligand, qui est un signal proportionnel à la quantité de complexe formé. Les quantités de chaleur mesurées permettent d'obtenir l'isotherme de liaison, et les paramètres d'interaction. L'ITC permet donc de mettre en évidence une interaction et de déterminer la constante d'affinité, ainsi que la stœchiométrie du système.



**Figure 66. Détermination de l'affinité de la nucléoprotéine du virus de la grippe pour l'importine  $\alpha 5$  humaine par ITC.** a) Titration du mutant R416A, NPmono (126  $\mu\text{M}$ ) dans une solution d'importine  $\alpha 5$  (11  $\mu\text{M}$ ) et b) titration du double mutant R416A, N319K, NP<sub>K319</sub> (53  $\mu\text{M}$ ) dans une solution d'importine  $\alpha 5$  (6,6  $\mu\text{M}$ ). Les expériences ont été réalisées dans le tampon 40 mM Tris-HCl, pH 7.5, 200 mM NaCl et 1 mM  $\beta$ -mercaptoethanol à 25 °C.

L'état d'oligomérisation de la nucléoprotéine présente un obstacle pour réaliser cette expérience, c'est donc à nouveau la nucléoprotéine monomérique (NPmono) qui a été utilisée. Ainsi, afin de tester l'implication de cette mutation dans l'activité de fixation à l'importine  $\alpha 5$ , nous avons remplacé l'asparagine située en position 319 de notre NP monomérique (NPmono) par une lysine. Pour une plus grande clarté de lecture nous avons nommé NP<sub>K319</sub> ce double mutant R416A, N319K. La NPmono ou la NP<sub>K319</sub> ont donc été titrées dans une solution d'importine  $\alpha 5$  humaine et l'interaction entre les partenaires a été mesurée par ITC (**figure 66**). La NPmono se fixe à l'importine  $\alpha 5$  avec une constante d'affinité de 26 nM (**figure 66a**). La

NP<sub>K319</sub> quant à elle, se fixe à l'importine  $\alpha 5$  avec une constante d'affinité de 13 nM, ce qui est deux fois plus fort que celle reliant la NP<sub>mono</sub> à l'Imp $\alpha 5$  (**figure 66b**). Ainsi, la mutation N319K de la nucléoprotéine améliore l'affinité de la NP pour l'importine  $\alpha 5$ . Pour expliquer ce résultat, nous pouvons à nouveau nous pencher sur la structure de la nucléoprotéine. Le domaine N-terminal comportant le NLS n'est pas situé à proximité de l'acide aminé 319. Cependant la forme allongée de l'importine  $5\alpha$  pourrait permettre la fixation au NLS et renforcer l'interaction grâce à une liaison supplémentaire réalisée avec l'acide aminé 319 lorsque celui-ci serait une lysine (**figure 67**). Cette différence de facteur 2, observée entre les affinités des complexes NP<sub>mono</sub>-Imp $\alpha 5$  et NP<sub>K319</sub>-Imp $\alpha 5$  est significative et pourrait éventuellement affecter l'efficacité de transport d'une nucléoprotéine portant la mutation N319K. Bien qu'elle soit retrouvée parmi les mutations adaptatrices, entre SC35 et SC35M, la mutation N319K n'est pas à elle seule responsable de la virulence du virus chez la souris. Cependant, l'augmentation de l'activité de la polymérase pourrait être due à une amélioration du transport nucléaire de la nucléoprotéine.



**Figure 67. Hypothétique mode d'interaction du complexe NP-Importine  $\alpha 5$ .** L'importine  $\alpha 5$  est représentée en vert, la nucléoprotéine en bleue. Les flèches indiquent les zones de contacts éventuelles entre les deux molécules. Le NLS non conventionnel sur la NP est modélisé en pointillés rouges et l'acide aminé en position 319 est indiqué.

## **III. Ma conclusion**

## -Ma conclusion-

Comment suis-je arrivé jusqu'ici ? C'est la question que je me pose en terminant ma thèse. Cela venait-il du temps où enfant je m'amusais à attraper des mouches et autres insectes que je congelais dans des bacs à glaçons avec l'espoir de les faire revivre ? "*Ah, encore avec tes expériences !*", s'écriaient mes parents quelque peu dégoutés, à qui bien sûr, je recommandais de ne rien jeter. Ou bien, est-ce parce que chaque semaine en allant au collège à Villars-de-Lans, je voyais se construire le synchrotron où l'on m'expliquait qu'en ces lieux, des chercheurs feraient, eux aussi, des expériences ? A cette époque, j'avais un peu plus d'une dizaine d'années et bien sûr aucune idée de ce qui pouvait réellement se tramer à l'intérieur du site de l'ESRF. Ce mystérieux anneau aurait-il éveillé ma curiosité ?

Au cours de mon cursus universitaire, j'ai eu l'opportunité de visiter tous les instituts de biologie qui encerclent le synchrotron. A cette époque, l'UVHCI n'existait pas encore. J'étais dans ma deuxième année d'université et à ce moment là, j'ai su deux choses. Premièrement, je voulais allier la structure et la biologie. Deuxièmement, les locaux de l'EMBL me paraissaient parfaits pour atteindre ce but. Quelques années plus tard, durant mon Master, je me retrouvais enfin en ces lieux pour y réaliser....des expériences. A la suite de rencontres déterminantes, notamment avec le Dr Florence Baudin, j'ai fait mes premières armes sur le virus de la grippe et sur le transport nucléaire. J'ai ainsi découvert le monde fascinant des virus et du piratage des fonctions cellulaires.

Enfin, j'ai eu l'opportunité de partir en Allemagne pour ma dernière année de thèse. Là, pas de synchrotron, mais des laboratoires situés au cœur d'une forêt. Et pourtant, ici aussi, comme à Grenoble, des chercheurs de l'Europe entière, s'activent, discutent et s'interrogent sur une partie des énigmes de notre univers.

Toutes ces rencontres, ces expériences, ces changements m'ont permis chaque jour de considérer mon sujet de thèse sous un angle différent et d'y puiser de nouvelles idées. Au final, même si tout n'a pas été forcément facile à vivre, l'enjeu en valait le coup et toutes ces années furent riches d'actions et d'émotions.

Aujourd'hui, en regardant l'ensemble de mon travail de thèse, je sais bien qu'il reste encore beaucoup de choses à découvrir. Parce que la menace d'une nouvelle pandémie de grippe se profile, un réseau de cinq équipes de chercheurs européens s'est allié pour étudier plus en détail la transcription et la réplication du virus de la grippe. Je suis fier d'appartenir à

l'une d'elles et j'ai pu, par mon travail, contribuer entre autre à la compréhension du mécanisme d'import de l'une des trois sous-unités de la polymérase.

Cependant, de nombreuses questions restent encore en suspens, en particulier sur l'identification de partenaires cellulaires qui interagissent avec les protéines virales, ou sur la résolution de nouveaux domaines de la polymérase, sur la détermination de la virulence des virus aviaires chez l'humain, etc.

Pour ma part, je me sens prêt à visiter de nouveaux labos et de nouveaux pays, à faire de nouvelles rencontres et suis impatient d'apporter mes connaissances acquises à un nouveau sujet de recherche.

# **IV. MATERIEL ET METHODES**

## 1. Expression et purification des protéines

### 1.1. Préparation de la nucléoprotéine virale (NPv)

Le virus de la grippe de souche **A/PR/8/34** est produit sur embryons de poulet (Pasteur-Mérieux, Marcy L'Etoile, France). Le virus est d'abord lavé. Pour cela, 10 ml de virus (à environ 3-4 mg/ml) sont dilués dans 35 ml finaux de PBS, puis centrifugés 1 h à 25000 rpm (rotor SW27) à 4°C. Le culot est ensuite re-suspendu dans 2 ml de PBS et conservé à 4°C sur la nuit. L'ouverture du virus lavé se fait par incubation, pendant 10 min à 30°C, dans le tampon Tris-HCl (pH 8) 50 mM, dithiothreitol (DTT) 1,5 mM, KCl 100 mM, MgCl<sub>2</sub> 5 mM, glycérol 5 %, Triton X-100 (1%), lyzolécythine (10 mg/ml) comme cela fût décrit par [Honda \*et al.\* \(1987\)](#). Il faut compter 5 ml de tampon d'ouverture pour 10 ml de suspension de virus lavé. Environ 2 ml de virus ouvert sont ensuite déposés sur un gradient linéaire de glycérol 10-30%, dans le tampon Tris-HCl (pH 8) 50 mM, NaCl 100 mM, DTT 1 mM, et centrifugés 16 h à 25000 rpm à 4°C (rotor SW27). Des fractions de 2 ml sont récoltées à partir du bas du gradient. Puis les fractions riches en RNPv sont réunies et dialysées et concentrées dans le tampon Tris-HCl (pH 8) 20 mM, β-SH 1 mM, KCl 150 mM, MgCl<sub>2</sub> 10 mM à 4°C. La suspension de RNPv est ensuite déposée sur un gradient discontinu de ClCs/glycérol (30%/1.5 M, 35%/2 M, 40%/2.5 M, 45%/3 M) dans le tampon Tris-HCl (pH 8) 50 mM, NaCl 100 mM, DTT 1 mM et centrifugée 20 h à 36000 rpm à 4°C (rotor SW41) comme cela fût décrit par [Honda \*et al.\* \(1987\)](#). Des fractions de 1 ml sont récoltées à partir du haut du gradient et un aliquot de chaque fraction est chauffé 3 min à 95°C et analysé sur SDS-PAGE, Laemmli 12 %. Les fractions de NPv sont collectées puis dialysées contre le tampon Tris-HCl (pH 7,5), NaCl 100 mM, MgCl<sub>2</sub> 2 mM, DTT 2 mM et glycérol 30%.

### 1.2. Expression et purification de la nucléoprotéine recombinante et des différents mutants

Le gène de la nucléoprotéine du virus de la grippe de souche A/PR/8/34 (NPrec pour recombinant), qui code pour une protéine de 498 acides aminés, est exprimé en utilisant le plasmide recombinant pET-16b (Novagen). Ce plasmide contient le gène entier avec une étiquette polyhistidine située en N-terminal. Afin d'enlever l'étiquette polyhistidine, nous avons changé le site de clivage Factor X en site TEV. La mutation R416A sur la NP (NPmono) a été construite en remplaçant le codon AGA (codant pour l'arginine 416) par le codon GCA (codant pour une alanine). La mutation N319K sur la NP (NP<sub>K319</sub>) a été



construite en remplaçant le codon AAT (codant pour l'asparagine 319) par le codon AAG (codant pour une lysine). Ces deux mutants ont été construits en utilisant une méthode de PCR inverse (Quickchange, Stratagene).

**Expression.** L'expression des différentes NP se fait dans la bactérie *Escherichia coli* de souche BL21 (DE3) pLysS (Stratagene). La cellule contient en plus le plasmide pLysS codant pour le T7 lysozyme, inhibiteur naturel de la T7 ARN polymérase (elle-même codée par DE3 dans le chromosome bactérien et sous contrôle du promoteur *lac*) et donc de l'expression du gène d'intérêt. En présence d'IPTG, l'inhibition par le répresseur *LacI* sur le gène d'intérêt et sur celui de la T7 ARN polymérase est levée et la protéine est surexprimée. Les cultures se font sous agitation (190 rpm) à 37°C dans du milieu LB contenant de l'ampicilline 50 µg/ml + chloramphénicol 34 µg/ml. L'induction se fait à une DO<sub>600nm</sub> de 0.8-1 avec de l'IPTG 0.5 mM final, et la culture se poursuit pendant 4-5 heures à 22°C. Les bactéries sont ensuite centrifugées à 7000 rpm pendant 10 minutes à 4°C et les culots sont congelés à -80°C.

**Lyse bactérienne.** Trente grammes de cellules sont repris dans 100 ml de tampon de lyse Tris-HCl (pH 7,9) 20 mM, NaCl 500 mM, imidazole 10 mM, β-SH 1 mM, glycérol 10%, auxquels sont ajoutés 1 pastille inhibitrice de protéases Complete (EDTA-free) et du lysozyme. Après ouverture des cellules par sonication, les débris cellulaires sont éliminés par centrifugation à 20000 rpm pendant 1 h à 4°C. Le surnageant est filtré sur une membrane à pores de 0.8 µm de diamètre (Nalgen), puis déposé sur une colonne d'affinité contenant une résine chargée en Co<sup>2+</sup> (Sigma).

**Purification par chromatographie d'affinité.** La colonne est constituée de 5 ml tassés de résine (Sigma) équilibrée dans 20 ml de tampon lyse. L'extrait soluble est appliqué sur la colonne, puis la colonne est lavée avec 15 ml de tampon de lyse, puis 20 ml de tampon de lavage Tris-HCl (pH 7,9) 20 mM, NaCl 1 M, imidazole 10 mM, β-SH 1 mM, glycérol 10%. L'élution des deux protéines se fait avec 12 ml de tampon d'élution Tris-HCl (pH 7,9) 20 mM, NaCl 500 mM, imidazole 500 mM, β-SH 1 mM, glycérol 10%. Des fractions d'élution de 1 ml sont récoltées et 10 µl d'aliquote sont déposés sur SDS-PAGE 12% Laemmli. Les fractions riches en NP sont rassemblées puis dialysées contre le tampon Tris-HCl (pH 7,5) 20 mM, NaCl 200 mM, glycérol 10%, et β-SH 1 mM. Quand cela est nécessaire, c'est-à-dire pour NP<sub>prec</sub> et NP<sub>K319</sub>, l'étiquette polyhistidine est enlevée. Ceci est réalisé en incubant la protéine, une nuit à température ambiante, avec la protéase du Tobacco Etch Virus (TEV) dans un rapport protéase/protéine de 1/100. Le produit de la digestion est ensuite déposé sur 800 µl tassés de résine Ni<sup>2+</sup> équilibrée dans le tampon Tris (pH 7,5) 20

mM, NaCl 200 mM,  $\beta$ -SH 1 mM, imidazole 10 mM, glycérol 10%. L'effluent comprenant la nucléoprotéine coupée est récolté. La protéine coupée est de plus purifiée et analysée par chromatographie d'exclusion de taille sur une colonne Superdex 200 HiLoad 16/60 (GE Healthcare) équilibrée dans Tris-HCl (pH 7,5) 40 mM, NaCl 200 mM et  $\beta$ -SH 1 mM.

### 1.3. Expression et purification de la partie C-terminale de PB2 (DPDE : 678-759)

Le gène de la partie C-terminale de PB2 (DPDE : acides aminés 678-759) du virus de la grippe de souche A/Victoria/3/75, a été déterminé par troncation unidirectionnelle du gène comme cela est décrit dans [Tarendeau et al. \(2007\)](#). Ce gène qui code pour une protéine de 98 acides aminés, est exprimé en utilisant le plasmide recombinant pETM-11. Ce plasmide dérive du pET24d (Novagen) par introduction du site de clivage par la protéase TEV entre l'étiquette polyhistidine et le site de clonage multiple (MCS).

**Expression.** L'expression de DPDE se fait dans la bactérie *Escherichia coli* de souche BL21 RIL (Stratagene). Les cultures se font sous agitation (190 rpm) à 37°C dans du milieu LB contenant de la kanamycine 50  $\mu$ g/ml. L'induction se fait à une  $DO_{600nm}$  de 0.9-1 avec de l'IPTG 0.25 mM final, et la culture se poursuit pendant toute la nuit à 25°C. Les bactéries sont ensuite centrifugées à 7000 rpm pendant 10 minutes à 4°C et les culots sont congelés à -80°C.

**Lyse bactérienne.** Dix grammes de cellules sont repris dans 50 ml de tampon de lyse Tris-HCl (pH 7,5) 40 mM, NaCl 300 mM, imidazole 10 mM, glycérol 10%, auxquels est ajoutée 1 pastille inhibitrice de protéases Complete (EDTA-free). Après ouverture des cellules par sonication, les débris cellulaires sont éliminés par centrifugation à 20000 rpm pendant 1 h à 4°C. Le surnageant est filtré sur une membrane à pores de 0.8  $\mu$ m de diamètre (Nalgen), puis déposé sur une colonne d'affinité contenant une résine chargée en  $Co^{2+}$  (Sigma).

**Purification par chromatographie d'affinité.** La colonne est constituée de 4 ml tassés de résine (Sigma) équilibrée dans 20 ml de tampon lyse. L'extrait soluble est appliqué sur la colonne, puis la colonne est lavée avec 12 ml de tampon de lyse, puis 12 ml de tampon de lavage Tris-HCl (pH 7,5) 40 mM, NaCl 1 M, imidazole 15 mM, glycérol 10%. L'élution des deux protéines se fait avec 12 ml de tampon d'élution Tris-HCl (pH 7,5) 40 mM, NaCl 200 mM, imidazole 400 mM, glycérol 10%. Des fractions d'élution de 1 ml sont récoltées et 10  $\mu$ l d'aliquote sont déposés sur SDS-PAGE 12% Laemmli. Les fractions riches en DPDE sont rassemblées puis dialysées contre le tampon Tris-HCl (pH 7,5) 40 mM, NaCl 200 mM, glycérol 10%. L'étiquette polyhistidine de DPDE est ensuite enlevée. Ceci est réalisé en incubant la protéine, une nuit à température ambiante, avec la protéase du Tobacco Etch Virus

(TEV) dans un rapport protéase/protéine de 1/100. Le produit de la digestion est ensuite déposé sur 1 ml tassé de résine Ni<sup>2+</sup> équilibrée dans le tampon Tris (pH 7,5) 40 mM, NaCl 200 mM, imidazole 10 mM, glycerol 10%. L'effluent comprenant DPDE coupée est récolté. La protéine coupée est à nouveau purifiée par chromatographie d'exclusion de taille sur une colonne Superdex 200 HiLoad 16/60 (GE Healthcare) équilibrée dans Tris-HCl (pH 7,5) 40 mM, NaCl 200 mM.

#### 1.4. Expression et purification de l'importine $\alpha 5$ humaine (66-512)

Le gène de l'importine  $\alpha 5$  humaine (Imp $\alpha 5$ ), qui code pour une protéine de 473 acides aminés, est exprimé en utilisant le plasmide recombinant pProEX HTb (Invitrogen). La séquence codant pour les 66 premiers acides aminés ainsi que les 20 derniers ont été délétés, la protéine utilisée correspond ainsi à la séquence [66-512]. La protéine exprimée présente une étiquette polyhistidine en N-terminal fournie par le vecteur ainsi qu'un site de clivage pour la protéase TEV.

**Expression.** L'expression de l'Imp $\alpha 5$  se fait dans la bactérie *Escherichia coli* de souche BL21 (DE3) (Stratagene). Les cultures se font sous agitation (190 rpm) à 37°C dans du milieu LB contenant de l'ampicilline à 50  $\mu$ g/ml. L'induction se fait à une DO<sub>600nm</sub> de 0.8 avec de l'IPTG 0.5 mM final, et la culture se poursuit pendant 4-5 heures à 22°C. Les bactéries sont ensuite centrifugées à 7000 rpm pendant 10 minutes à 4°C et les culots sont congelés à -80°C.

**Lyse bactérienne.** Le culot bactérien (environ 40 g pour 12 L), est repris dans 200 ml de tampon de lyse Tris-HCl (pH 8,45) 20 mM, NaCl 100 mM, imidazole 10 mM, glycérol 10%, auxquels est ajoutée 1 pastille inhibitrice de protéases Complete (EDTA-free). Après ouverture des cellules par sonication, les débris cellulaires sont éliminés par centrifugation à 20000 rpm pendant 1 h à 4°C. Le surnageant est filtré sur une membrane à pores de 0.8  $\mu$ m de diamètre (Nalgen), puis déposé sur une colonne d'affinité contenant une résine chargée en Co<sup>2+</sup> (Sigma).

**Purification par chromatographie d'affinité.** La colonne est constituée de 5 ml tassés de résine (Sigma) équilibrée dans 20 ml de tampon lyse. L'extrait soluble est appliqué sur la colonne, puis la colonne est lavée avec 15 ml de tampon de lyse, puis 20 ml de tampon de lavage Tris-HCl (pH 8,45) 20 mM, NaCl 500 mM, imidazole 10 mM,  $\beta$ -SH 1 mM, glycérol 10%. L'élution des deux protéines se fait avec 12 ml de tampon d'élution Tris-HCl (pH 8,45) 20 mM, NaCl 200 mM, imidazole 250 mM,  $\beta$ -SH 1 mM, glycérol 10%. Des

fractions d'éluion de 1 ml sont récoltées et 10  $\mu$ l d'aliquote sont déposés sur SDS-PAGE 12% Laemmli. Les fractions riches en NP sont rassemblées puis dialysées contre le tampon Tris-HCl (pH 7,5) 20 mM, NaCl 200 mM, glycérol 10%, et  $\beta$ -SH 1 mM. L'étiquette polyhistidine est enlevée grâce à l'action sur la nuit de la protéase TEV (1/100) à température ambiante. Le produit de la digestion est ensuite déposé sur 800  $\mu$ l tassés de résine Ni<sup>2+</sup> équilibrée dans le tampon Tris (pH 7,5) 20 mM, NaCl 200 mM,  $\beta$ -SH 1 mM, imidazole 10 mM, glycérol 10%. L'effluent comprenant la protéine coupée est récolté. La protéine coupée est de plus purifiée et analysée par chromatographie d'exclusion de taille sur une colonne Superdex 200 HiLoad 16/60 (GE Healthcare) équilibrée dans le tampon Tris-HCl (pH 7,5) 40 mM, NaCl 200 mM et  $\beta$ -SH 1 mM.

## 2. Les complexes protéine cellulaire - protéine virale

### 2.1. Le complexe importine $\alpha$ 5 - DPDE

**Formation du complexe Imp $\alpha$ 5-DPDE.** A ce stade, les deux protéines sont démunies de leurs étiquettes polyhistidines. Afin de constituer le complexe, l'Imp $\alpha$ 5 [66-512] est mise en présence de la partie C-terminale de PB2 (DPDE : 678-759) en excès selon le rapport 1 Imp $\alpha$ 5 / 5 DPDE (mole / mole). Le complexe formé est ensuite purifié par chromatographie d'exclusion de taille sur une colonne Superdex 200 équilibrée dans le tampon Tris-HCl (pH 7,5) 30 mM, NaCl 150 mM,  $\beta$ -SH 5 mM puis analysé sur SDS-PAGE 12%.

**Cristallogénèse.** Le pic correspondant au complexe Imp $\alpha$ 5-DPDE élué à 14,9 ml est récupéré puis concentré jusqu'à l'obtention d'un échantillon à 20 mg.ml<sup>-1</sup>. Des nanogouttes de 0.2  $\mu$ l (0.1  $\mu$ l de précipitant + 0.1  $\mu$ l de solution protéique) sont réalisées par l'utilisation de la plateforme de cristallisation à haut débit Tecan (EMBL-Grenoble). Les cristaux de l'importine  $\alpha$ 5 complexée à la partie C-terminale de PB2 ont poussé par diffusion de vapeur dans la condition suivante : 20 mM CaCl<sub>2</sub>, 100mM AcNa (pH 4.6), avec comme agent précipitant 16% MPD (2-methyl-2, 4-pentandiol). Les cristaux ont été reproduits manuellement (1  $\mu$ l de précipitant + 1  $\mu$ l de solution protéique), puis un jeu de données complet à 2.2 Å de résolution a été mesuré sur la ligne ID23-1 de l'installation européenne de rayonnement synchrotron (ESRF) de Grenoble. Deux complexes sont présents dans l'unité asymétrique. La structure du complexe a été résolue par remplacement moléculaire en utilisant le modèle PDB (Protein Data Bank) de l'importine  $\alpha$  de souris (1Q1S). Deux complexes sont présents dans l'unité

asymétrique. Pour de plus amples détails concernant la résolution de la structure, se référer à la publication [Tarendeau \*et al.\* \(2007\)](#).

## 2.2. Le complexe NPrec-imp $\alpha$ 5 [66-512]

**Formation du complexe.** La NPrec éluée de la colonne de  $\text{Co}^{2+}$  et dialysée contre le tampon Tris-HCl (pH 7,5) 20 mM, NaCl 200 mM, glycérol 10%,  $\beta$ -SH 1 mM est directement mélangée à l'Imp $\alpha$ 5 sans étiquette polyhistidine, dans le même tampon. L'Imp $\alpha$ 5 est mise en excès selon le rapport 3 Imp $\alpha$ 5 / 1 NPrec (mole / mole). Le complexe est ensuite déposé sur une colonne contenant 2 ml de résine  $\text{Ni}^{2+}$  équilibrée dans le tampon mentionné précédemment + imidazole 10 mM. L'excès d'Imp $\alpha$ 5 non complexée à NPrec est ainsi éliminé. La résine de  $\text{Ni}^{2+}$  est lavée avec 30 ml de Tris-HCl (pH 7,5) 20 mM, NaCl 200 mM, imidazole 20 mM,  $\beta$ -SH 1 mM. Le complexe est élué avec 20 ml de tampon d'éluion Tris-HCl (pH 7,5) 20 mM, NaCl 200 mM, imidazole 200 mM,  $\beta$ -SH 1 mM. L'éluion est ensuite analysée par SDS-PAGE 12%. Les fractions d'intérêt sont ensuite passées sur une colonne Superdex 200 équilibrée dans le tampon Tris-HCl (pH 7,5) 30 mM, NaCl 150 mM,  $\beta$ -SH 1 mM. Le profil d'éluion est à nouveau analysé par SDS-PAGE 12%. Le pic correspondant au complexe, élué à 14,5 ml est récupéré puis soumis à une protéolyse à l'élastase (rapport protéase / substrat de 1/80ème p/p) pendant 20 min à température ambiante. La protéolyse est arrêtée avec du PMSF 1mM final et le complexe protéolysé est purifié sur Superdex 200 dans le même tampon. Le complexe élué à 14.6 ml est concentré jusqu'à environ 5 mg/ml pour réaliser des essais de cristallogénèse à l'aide de la plateforme de cristallisation à haut débit Tecan (EMBL-Grenoble).

## 2.3. Le complexe NPmono – Imp $\alpha$ 5

**Formation du complexe.** Contrairement à la formation du complexe NPrec-Imp $\alpha$ 5 où seule l'étiquette polyhistidine de l'importine  $\alpha$ 5 était enlevée, dans le cas de la formation du complexe NPmono-Imp $\alpha$ 5, les étiquettes des deux protéines ont été clivées. Afin de constituer le complexe, l'Imp $\alpha$ 5 [66-512] est mise en présence de la NPmono (mutant R416A) en excès selon le rapport 1 Imp $\alpha$ 5 / 3 NPmono (mole / mole). Le complexe formé est ensuite purifié par chromatographie d'exclusion de taille sur une colonne Superdex 200 équilibrée dans le tampon Tris-HCl (pH 7,5) 40 mM, NaCl 200 mM,  $\beta$ -SH 1 mM puis analysé par SDS-PAGE 12%. Des essais de cristallogénèse ont également été réalisés.

### 3. Ultracentrifugation analytique

Les expériences de sédimentation de vitesse ont été réalisées à 20°C. La vitesse de rotation était de 46 000 rpm dans une centrifugeuse Optima XL-A (Beckman) ajustée avec le rotor AN-60. Les profils de sédimentation de vitesse ont été collectés en mesurant le signal d'absorbance à 280 nm dans le tampon Tris-HCl (pH 7,5) 40 mM, NaCl 200 mM,  $\beta$ -SH 1 mM. Le coefficient de sédimentation et les distributions de poids moléculaire furent analysés à l'aide du logiciel Sedfit (Schuck P., 2000). Les corrections concernant la densité et la viscosité du tampon ont été faites conformément aux données publiées par Laue *et al.* (1992).

### 4. Microscopie électronique à transmission (MET) en coloration négative

**Préparation des échantillons.** Les différents échantillons de nucléoprotéines sont incubés à 37°C pendant 10 minutes. Sont introduits par capillarité, 3  $\mu$ l de NPv, de NPrec ou de NPmono purifiée, à l'interface d'une feuille de carbone-mica. L'échantillon s'adsorbe alors sur le carbone. L'échantillon fixé sur le carbone est ensuite déposé à la surface d'une solution de colorant d'acétate d'uranyle à 2% (p/v). L'ensemble carbone/échantillon est pêché à l'aide d'une grille de cuivre/rhodium, séché puis introduit dans le microscope.

**Acquisition des images.** Les images ont été prises à l'aide d'un microscope Morgagni FEI sous une tension d'accélération de 100 kV et un grossissement nominal de 71 000X. Les différents échantillons de nucléoprotéines furent incubés à 37°C pendant 10 minutes.

### 5. Affinité de la nucléoprotéine (NPv, NPrec et NPmono) ou du complexe NPrec-Imp $\alpha$ 5 pour l'ARN

#### 5.1. Synthèse et marquage de l'ARN *panhandle* par transcription *in vitro*

Le plasmide codant pour l'ARN de 81 bases est décrit par Baudin *et al.* (1994), il comprend les extrémités 3' et 5' conservées chez le virus de la grippe de type A, connectées par une séquence dérivée du gène de la NP. L'ARN est transcrit *in vitro* à partir du plasmide devenu linéaire grâce à l'enzyme de restriction Eam1104I (Ksp632I). L'ADN linéaire est purifié par extraction au phénol/chloroforme et précipité à l'éthanol. Ce plasmide est sous le

contrôle du promoteur T3 et la réaction de transcription est réalisée par la T3 ARN polymérase dans le tampon standard et en utilisant comme isotope radioactif, l'UTP-[ $\alpha$ -<sup>32</sup>P], comme cela est décrit par [Baudin \*et al.\* \(1994\)](#). L'ARN marqué est ensuite purifié par électrophorèse sur gel de polyacrylamide, puis excisé et enfin précipité à l'éthanol à -20°C.

## 5.2. Filter Binding Assays (FBA)

Avant les expériences de FBA, l'ARN radioactif dans l'eau est chauffé à 95°C pendant une minute, et refroidi jusqu'à température ambiante. Les différents échantillons de NP (NPv, NPrec et NPmono) et le complexe NPrec-Imp $\alpha$ 5 sont dilués en série (de 10<sup>-4</sup> M à 10<sup>-9</sup> M) dans un volume final de 180  $\mu$ l suivi par l'addition de 20  $\mu$ l d'ARN marqué. Une incubation de 1 h à 4°C est réalisée dans le tampon de fixation, Tris-HCl (pH 7,5) 50 mM, NaCl 100 mM, MgCl<sub>2</sub> 10 mM, DTT 2,5 mM. La mixture est ensuite filtrée à travers un filtre de nitrocellulose pré-imbibé (Whatman pores de 0,45  $\mu$ m, 25 mm de diamètre) comme cela est décrit dans [Baudin \*et al.\* \(1994\)](#). Seul l'ARN complexé à l'échantillon protéique est retenu sur le filtre et est détecté par la méthode Cerenkov dans un compteur à scintillation. Comme la quantité d'ARN radioactif utilisée est très faible (0,03 pmol), la concentration en protéine à 50% de rétention correspond à la constante de dissociation apparente, le Kd.

## 6. Calorimétrie de titration isotherme (ITC - Isothermal Titration Calorimetry)

Les expériences d'ITC sont réalisées avec le calorimètre VP-ITC Microcal (Microcal, Northhampton, MA, USA). Toutes les protéines (NPmono, NP<sub>K319</sub> et Imp $\alpha$ 5) sont abondamment dialysées contre le tampon Tris-HCl (pH 7,5) 40 mM, NaCl 200 mM,  $\beta$ -SH 1 mM, avant chaque titration. Les expériences ont été réalisées à 25°C. Une titration typique consiste en plusieurs injections d'aliquotes de 12  $\mu$ l de NP (NPmono ou NP<sub>K319</sub>) à 50-70  $\mu$ M à partir d'une seringue, dans une solution d'Imp $\alpha$ 5 à 5-7  $\mu$ M située dans la cellule de mesure du calorimètre. Un intervalle de temps, de 5 minutes, entre chaque injection est respecté afin de s'assurer que le pic de titration retourne à la ligne de base. Enfin, les données d'ITC sont analysées grâce au programme Origin version 5, fourni par le fabricant.

# **V. REFERENCES BIBLIOGRAPHIQUES**



## Bibliographie

- Adam, E. J. & Adam, S. A. (1994).** Identification of cytosolic factors required for nuclear location sequence-mediated binding to the nuclear envelope. *The Journal of cell biology* **125**, 547-555.
- Akarsu, H., Burmeister, W. P., Petosa, C., Petit, I., Muller, C. W., Ruigrok, R. W. & Baudin, F. (2003).** Crystal structure of the M1 protein-binding domain of the influenza A virus nuclear export protein (NEP/NS2). *The EMBO journal* **22**, 4646-4655.
- Akey, C. W. & Radermacher, M. (1993).** Architecture of the *Xenopus* nuclear pore complex revealed by three-dimensional cryo-electron microscopy. *The Journal of cell biology* **122**, 1-19.
- Alber, F., Dokudovskaya, S., Veenhoff, L. M., Zhang, W., Kipper, J., Devos, D., Suprapto, A., Karni-Schmidt, O., Williams, R., Chait, B. T., Sali, A. & Rout, M. P. (2007).** The molecular architecture of the nuclear pore complex. *Nature* **450**, 695-701.
- Albertini, A. A., Wernimont, A. K., Muziol, T., Ravelli, R. B., Clapier, C. R., Schoehn, G., Weissenhorn, W. & Ruigrok, R. W. (2006).** Crystal structure of the rabies virus nucleoprotein-RNA complex. *Science (New York, NY)* **313**, 360-363.
- Albo, C., Valencia, A. & Portela, A. (1995).** Identification of an RNA binding region within the N-terminal third of the influenza A virus nucleoprotein. *Journal of virology* **69**, 3799-3806.
- Area, E., Martin-Benito, J., Gastaminza, P., Torreira, E., Valpuesta, J. M., Carrascosa, J. L. & Ortin, J. (2004).** 3D structure of the influenza virus polymerase complex: localization of subunit domains. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 308-313.
- Arrese, M. & Portela, A. (1996).** Serine 3 is critical for phosphorylation at the N-terminal end of the nucleoprotein of influenza virus A/Victoria/3/75. *Journal of virology* **70**, 3385-3391.
- Arts, G. J., Fornerod, M. & Mattaj, I. W. (1998).** Identification of a nuclear export receptor for tRNA. *Curr Biol* **8**, 305-314.
- Arzt, S., Baudin, F., Barge, A., Timmins, P., Burmeister, W. P. & Ruigrok, R. W. (2001).** Combined results from solution studies on intact influenza virus M1 protein and from a new crystal form of its N-terminal domain show that M1 is an elongated monomer. *Virology* **279**, 439-446.
- Askjaer, P., Bachi, A., Wilm, M., Bischoff, F. R., Weeks, D. L., Ogniewski, V., Ohno, M., Niehrs, C., Kjems, J., Mattaj, I. W. & Fornerod, M. (1999).** RanGTP-regulated interactions of CRM1 with nucleoporins and a shuttling DEAD-box helicase. *Molecular and cellular biology* **19**, 6276-6285.
- Bancroft, C. T. & Parslow, T. G. (2002).** Evidence for segment-nonspecific packaging of the influenza A virus genome. *Journal of virology* **76**, 7133-7139.
- Baudin, F., Bach, C., Cusack, S. & Ruigrok, R. W. (1994).** Structure of influenza virus RNP. I. Influenza virus nucleoprotein melts secondary structure in panhandle RNA and exposes the bases to the solvent. *The EMBO journal* **13**, 3158-3165.
- Baudin, F., Petit, I., Weissenhorn, W. & Ruigrok, R. W. (2001).** In vitro dissection of the membrane and RNP binding activities of influenza virus M1 protein. *Virology* **281**, 102-108.

- Bayliss, R., Littlewood, T. & Stewart, M. (2000).** Structural basis for the interaction between FxFG nucleoporin repeats and importin-beta in nuclear trafficking. *Cell* **102**, 99-108.
- Beck, M., Forster, F., Ecke, M., Plitzko, J. M., Melchior, F., Gerisch, G., Baumeister, W. & Medalia, O. (2004).** Nuclear pore complex structure and dynamics revealed by cryoelectron tomography. *Science (New York, NY)* **306**, 1387-1390.
- Beck, M., Lucic, V., Forster, F., Baumeister, W. & Medalia, O. (2007).** Snapshots of nuclear pore complexes in action captured by cryo-electron tomography. *Nature* **449**, 611-615.
- Biswas, S. K., Boutz, P. L. & Nayak, D. P. (1998).** Influenza virus nucleoprotein interacts with influenza virus polymerase proteins. *Journal of virology* **72**, 5493-5501.
- Blaas, D., Patzelt, E. & Kuechler, E. (1982).** Identification of the cap binding protein of influenza virus. *Nucleic acids research* **10**, 4803-4812.
- Bornholdt, Z. A. & Prasad, B. V. (2006).** X-ray structure of influenza virus NS1 effector domain. *Nature structural & molecular biology* **13**, 559-560.
- Boulo, S., Akarsu, H., Ruigrok, R. W. & Baudin, F. (2007).** Nuclear traffic of influenza virus proteins and ribonucleoprotein complexes. *Virus research* **124**, 12-21.
- Bouloy, M., Plotch, S. J. & Krug, R. M. (1978).** Globin mRNAs are primers for the transcription of influenza viral RNA in vitro. *Proceedings of the National Academy of Sciences of the United States of America* **75**, 4886-4890.
- Bruns, K., Studtrucker, N., Sharma, A., Fossen, T., Mitzner, D., Eissmann, A., Tessmer, U., Roder, R., Henklein, P., Wray, V. & Schubert, U. (2007).** Structural characterization and oligomerization of PB1-F2, a proapoptotic influenza A virus protein. *The Journal of biological chemistry* **282**, 353-363.
- Bui, M., Myers, J. E. & Whittaker, G. R. (2002).** Nucleo-cytoplasmic localization of influenza virus nucleoprotein depends on cell density and phosphorylation. *Virus research* **84**, 37-44.
- Burmeister, W. P., Ruigrok, R. W. & Cusack, S. (1992).** The 2.2 Å resolution crystal structure of influenza B neuraminidase and its complex with sialic acid. *The EMBO journal* **11**, 49-56.
- Catimel, B., Teh, T., Fontes, M. R., Jennings, I. G., Jans, D. A., Howlett, G. J., Nice, E. C. & Kobe, B. (2001).** Biophysical characterization of interactions involving importin-alpha during nuclear import. *The Journal of biological chemistry* **276**, 34189-34198.
- Chapman-Smith, A. & Cronan, J. E., Jr. (1999).** The enzymatic biotinylation of proteins: a post-translational modification of exceptional specificity. *Trends in biochemical sciences* **24**, 359-363.
- Chen, W., Calvo, P. A., Malide, D., Gibbs, J., Schubert, U., Bacik, I., Basta, S., O'Neill, R., Schickli, J., Palese, P., Henklein, P., Bennink, J. R. & Yewdell, J. W. (2001).** A novel influenza A virus mitochondrial protein that induces cell death. *Nature medicine* **7**, 1306-1312.
- Chen, Z., Li, Y. & Krug, R. M. (1999).** Influenza A virus NS1 protein targets poly(A)-binding protein II of the cellular 3'-end processing machinery. *The EMBO journal* **18**, 2273-2283.
- Chi, N. C., Adam, E. J. & Adam, S. A. (1995).** Sequence and characterization of cytoplasmic nuclear protein import factor p97. *The Journal of cell biology* **130**, 265-274.
- Chi, N. C., Adam, E. J. & Adam, S. A. (1997).** Different binding domains for Ran-GTP and Ran-GDP/RanBP1 on nuclear import factor p97. *The Journal of biological chemistry* **272**, 6818-6822.

- Chien, C. Y., Tejero, R., Huang, Y., Zimmerman, D. E., Rios, C. B., Krug, R. M. & Montelione, G. T. (1997).** A novel RNA-binding motif in influenza A virus non-structural protein 1. *Nature structural biology* **4**, 891-895.
- Cingolani, G., Petosa, C., Weis, K. & Muller, C. W. (1999).** Structure of importin-beta bound to the IBB domain of importin-alpha. *Nature* **399**, 221-229.
- Cohen, F. E. & Prusiner, S. B. (1998).** Pathologic conformations of prion proteins. *Annual review of biochemistry* **67**, 793-819.
- Compans, R. W., Content, J. & Duesberg, P. H. (1972).** Structure of the ribonucleoprotein of influenza virus. *Journal of virology* **10**, 795-800.
- Conti, E. & Kuriyan, J. (2000).** Crystallographic analysis of the specific yet versatile recognition of distinct nuclear localization signals by karyopherin alpha. *Structure* **8**, 329-338.
- Conti, E., Uy, M., Leighton, L., Blobel, G. & Kuriyan, J. (1998).** Crystallographic analysis of the recognition of a nuclear localization signal by the nuclear import factor karyopherin alpha. *Cell* **94**, 193-204.
- Cook, A., Bono, F., Jinek, M. & Conti, E. (2007).** Structural biology of nucleocytoplasmic transport. *Annual review of biochemistry* **76**, 647-671.
- Cox, N., Xu, X., Bender, C., Kendal, A., Regnery, H., Hemphill, M. & Rota, P. (1993).** "Evolution of hemagglutinin in epidemic variants and selection of vaccine viruses". *Options for the control of influenza II*, (Hannoun, HKC, Kendal, A and Ruben, F eds) Elsevier, Amsterdam.
- Cros, J. F., Garcia-Sastre, A. & Palese, P. (2005).** An unconventional NLS is critical for the nuclear import of the influenza A virus nucleoprotein and ribonucleoprotein. *Traffic (Copenhagen, Denmark)* **6**, 205-213.
- Cros, J. F. & Palese, P. (2003).** Trafficking of viral genomic RNA into and out of the nucleus: influenza, Thogoto and Borna disease viruses. *Virus research* **95**, 3-12.
- Daly, J. M., Wood, J. M. & Robertson, J. S. (1998).** "Cocirculation and divergence of human influenza viruses". *Textbook of influenza*, (Nicholson, KG, Webster, RG, Hay, AJ, eds) Blackwell Science Ltd, Oxford 168-177.
- Davey, J., Dimmock, N. J. & Colman, A. (1985).** Identification of the sequence responsible for the nuclear accumulation of the influenza virus nucleoprotein in *Xenopus* oocytes. *Cell* **40**, 667-675.
- de Jong, M. D., Simmons, C. P., Thanh, T. T., Hien, V. M., Smith, G. J., Chau, T. N., Hoang, D. M., Chau, N. V., Khanh, T. H., Dong, V. C., Qui, P. T., Cam, B. V., Hado, Q., Guan, Y., Peiris, J. S., Chinh, N. T., Hien, T. T. & Farrar, J. (2006).** Fatal outcome of human influenza A (H5N1) is associated with high viral load and hypercytokinemia. *Nature medicine* **12**, 1203-1207.
- de la Luna, S., Fortes, P., Beloso, A. & Ortin, J. (1995).** Influenza virus NS1 protein enhances the rate of translation initiation of viral mRNAs. *Journal of virology* **69**, 2427-2433.
- de la Luna, S., Martin, J., Portela, A. & Ortin, J. (1993).** Influenza virus naked RNA can be expressed upon transfection into cells co-expressing the three subunits of the polymerase and the nucleoprotein from simian virus 40 recombinant viruses. *The Journal of general virology* **74 ( Pt 3)**, 535-539.
- de la Luna, S., Martinez, C. & Ortin, J. (1989).** Molecular cloning and sequencing of influenza virus A/Victoria/3/75 polymerase genes: sequence evolution and prediction of possible functional domains. *Virus research* **13**, 143-155.

- Denning, D. P., Patel, S. S., Uversky, V., Fink, A. L. & Rexach, M. (2003).** Disorder in the nuclear pore complex: the FG repeat regions of nucleoporins are natively unfolded. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 2450-2455.
- Desselberger, U., Racaniello, V. R., Zazra, J. J. & Palese, P. (1980).** The 3' and 5'-terminal sequences of influenza A, B and C virus RNA segments are highly conserved and show partial inverted complementarity. *Gene* **8**, 315-328.
- Dingwall, C., Robbins, J., Dilworth, S. M., Roberts, B. & Richardson, W. D. (1988).** The nucleoplasmin nuclear location sequence is larger and more complex than that of SV-40 large T antigen. *The Journal of cell biology* **107**, 841-849.
- Elton, D., Medcalf, E., Bishop, K. & Digard, P. (1999a).** Oligomerization of the influenza virus nucleoprotein: identification of positive and negative sequence elements. *Virology* **260**, 190-200.
- Elton, D., Medcalf, L., Bishop, K., Harrison, D. & Digard, P. (1999b).** Identification of amino acid residues of influenza virus nucleoprotein essential for RNA binding. *Journal of virology* **73**, 7357-7367.
- Elton, D., Simpson-Holley, M., Archer, K., Medcalf, L., Hallam, R., McCauley, J. & Digard, P. (2001).** Interaction of the influenza virus nucleoprotein with the cellular CRM1-mediated nuclear export pathway. *Journal of virology* **75**, 408-419.
- Enami, K., Sato, T. A., Nakada, S. & Enami, M. (1994).** Influenza virus NS1 protein stimulates translation of the M1 protein. *Journal of virology* **68**, 1432-1437.
- Enami, M., Sharma, G., Benham, C. & Palese, P. (1991).** An influenza virus containing nine different RNA segments. *Virology* **185**, 291-298.
- Fanara, P., Hodel, M. R., Corbett, A. H. & Hodel, A. E. (2000).** Quantitative analysis of nuclear localization signal (NLS)-importin alpha interaction through fluorescence depolarization. Evidence for auto-inhibitory regulation of NLS binding. *The Journal of biological chemistry* **275**, 21218-21223.
- Feldherr, C. M., Kallenbach, E. & Schultz, N. (1984).** Movement of a karyophilic protein through the nuclear pores of oocytes. *The Journal of cell biology* **99**, 2216-2222.
- Fodor, E., Crow, M., Mingay, L. J., Deng, T., Sharps, J., Fechter, P. & Brownlee, G. G. (2002).** A single amino acid mutation in the PA subunit of the influenza virus RNA polymerase inhibits endonucleolytic cleavage of capped RNAs. *Journal of virology* **76**, 8989-9001.
- Fodor, E., Pritlove, D. C. & Brownlee, G. G. (1994).** The influenza virus panhandle is involved in the initiation of transcription. *Journal of virology* **68**, 4092-4096.
- Fodor, E. & Smith, M. (2004).** The PA subunit is required for efficient nuclear accumulation of the PB1 subunit of the influenza A virus RNA polymerase complex. *Journal of virology* **78**, 9144-9153.
- Fontes, M. R., Teh, T., Jans, D., Brinkworth, R. I. & Kobe, B. (2003).** Structural basis for the specificity of bipartite nuclear localization sequence binding by importin-alpha. *The Journal of biological chemistry* **278**, 27981-27987.
- Fontes, M. R., Teh, T. & Kobe, B. (2000).** Structural basis of recognition of monopartite and bipartite nuclear localization sequences by mammalian importin-alpha. *Journal of molecular biology* **297**, 1183-1194.
- Fornerod, M., Ohno, M., Yoshida, M. & Mattaj, I. W. (1997).** CRM1 is an export receptor for leucine-rich nuclear export signals. *Cell* **90**, 1051-1060.
- Fujii, K., Fujii, Y., Noda, T., Muramoto, Y., Watanabe, T., Takada, A., Goto, H., Horimoto, T. & Kawaoka, Y. (2005).** Importance of both the coding and the segment-specific noncoding regions of the influenza A virus NS segment for its efficient incorporation into virions. *Journal of virology* **79**, 3766-3774.

- Fujii, Y., Goto, H., Watanabe, T., Yoshida, T. & Kawaoka, Y. (2003).** Selective incorporation of influenza virus RNA segments into virions. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 2002-2007.
- Gabriel, G., Dauber, B., Wolff, T., Planz, O., Klenk, H. D. & Stech, J. (2005).** The viral polymerase mediates adaptation of an avian influenza virus to a mammalian host. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 18590-18595.
- Gabriel, G., Herwig, A. & Klenk, H. D. (2008).** Interaction of polymerase subunit PB2 and NP with importin alpha1 is a determinant of host range of influenza A virus. *PLoS pathogens* **4**, e11.
- Garcia-Sastre, A., Egorov, A., Matassov, D., Brandt, S., Levy, D. E., Durbin, J. E., Palese, P. & Muster, T. (1998).** Influenza A virus lacking the NS1 gene replicates in interferon-deficient systems. *Virology* **252**, 324-330.
- GEIG (2008).** Groupe d'Expertise et d'Information sur la Grippe. <http://www.grippe-geig.com/>.
- Gibbs, J. S., Malide, D., Hornung, F., Bennink, J. R. & Yewdell, J. W. (2003).** The influenza A virus PB1-F2 protein targets the inner mitochondrial membrane via a predicted basic amphipathic helix that disrupts mitochondrial function. *Journal of virology* **77**, 7214-7224.
- Gilchrist, D., Mykytka, B. & Rexach, M. (2002).** Accelerating the rate of disassembly of karyopherin.cargo complexes. *The Journal of biological chemistry* **277**, 18161-18172.
- Goldfarb, D. S., Corbett, A. H., Mason, D. A., Harreman, M. T. & Adam, S. A. (2004).** Importin alpha: a multipurpose nuclear-transport receptor. *Trends in cell biology* **14**, 505-514.
- Gonzalez, S., Zurcher, T. & Ortin, J. (1996).** Identification of two separate domains in the influenza virus PB1 protein involved in the interaction with the PB2 and PA subunits: a model for the viral RNA polymerase structure. *Nucleic acids research* **24**, 4456-4463.
- Gorlich, D., Pante, N., Kutay, U., Aebi, U. & Bischoff, F. R. (1996).** Identification of different roles for RanGDP and RanGTP in nuclear protein import. *The EMBO journal* **15**, 5584-5594.
- Gorlich, D., Vogel, F., Mills, A. D., Hartmann, E. & Laskey, R. A. (1995).** Distinct functions for the two importin subunits in nuclear protein import. *Nature* **377**, 246-248.
- Green, T. J., Zhang, X., Wertz, G. W. & Luo, M. (2006).** Structure of the vesicular stomatitis virus nucleoprotein-RNA complex. *Science (New York, NY)* **313**, 357-360.
- Gregoriades, A. (1980).** Interaction of influenza M protein with viral lipid and phosphatidylcholine vesicles. *Journal of virology* **36**, 470-479.
- Guilligay, D., Tarendeau, F., Resa-Infante, P., Coloma, R., Crepin, T., Sehr, P., Lewis, J., Ruigrok, R. W., Ortin, J., Hart, D. J. & Cusack, S. (2008).** The structural basis for cap binding by influenza virus polymerase subunit PB2. *Nature structural & molecular biology* **15**, 500-506.
- Hale, B. G., Randall, R. E., Ortin, J. & Jackson, D. (2008).** The multifunctional NS1 protein of influenza A viruses. *The Journal of general virology* **89**, 2359-2376.
- Hara, K., Schmidt, F. I., Crow, M. & Brownlee, G. G. (2006).** Amino acid residues in the N-terminal region of the PA subunit of influenza A virus RNA polymerase play a critical role in protein stability, endonuclease activity, cap binding, and virion RNA promoter binding. *Journal of virology* **80**, 7789-7798.

- Hatada, E., Hasegawa, M., Mukaigawa, J., Shimizu, K. & Fukuda, R. (1989).** Control of influenza virus gene expression: quantitative analysis of each viral RNA species in infected cells. *Journal of biochemistry* **105**, 537-546.
- He, X., Zhou, J., Bartlam, M., Zhang, R., Ma, J., Lou, Z., Li, X., Li, J., Joachimiak, A., Zeng, Z., Ge, R., Rao, Z. & Liu, Y. (2008).** Crystal structure of the polymerase PA(C)-PB1(N) complex from an avian influenza H5N1 virus. *Nature* **454**, 1123-1126.
- Hinshaw, J. E., Carragher, B. O. & Milligan, R. A. (1992).** Architecture and design of the nuclear pore complex. *Cell* **69**, 1133-1141.
- Hodel, M. R., Corbett, A. H. & Hodel, A. E. (2001).** Dissection of a nuclear localization signal. *The Journal of biological chemistry* **276**, 1317-1325.
- Holsinger, L. J., Nichani, D., Pinto, L. H. & Lamb, R. A. (1994).** Influenza A virus M2 ion channel protein: a structure-function analysis. *Journal of virology* **68**, 1551-1563.
- Honda, A., Ueda, K., Nagata, K. & Ishihama, A. (1987).** Identification of the RNA polymerase-binding site on genome RNA of influenza virus. *Journal of biochemistry* **102**, 1241-1249.
- Hood, F. E. & Clarke, P. R. (2007).** RCC1 isoforms differ in their affinity for chromatin, molecular interactions and regulation by phosphorylation. *Journal of cell science* **120**, 3436-3445.
- Horimoto, T. & Kawaoka, Y. (2005).** Influenza: lessons from past pandemics, warnings from current incidents. *Nature reviews* **3**, 591-600.
- Hsu, M. T., Parvin, J. D., Gupta, S., Krystal, M. & Palese, P. (1987).** Genomic RNAs of influenza viruses are held in a circular conformation in virions and in infected cells by a terminal panhandle. *Proceedings of the National Academy of Sciences of the United States of America* **84**, 8140-8144.
- Huang, T. S., Palese, P. & Krystal, M. (1990).** Determination of influenza virus proteins required for genome replication. *Journal of virology* **64**, 5669-5673.
- Huarte, M., Falcon, A., Nakaya, Y., Ortin, J., Garcia-Sastre, A. & Nieto, A. (2003).** Threonine 157 of influenza virus PA polymerase subunit modulates RNA replication in infectious viruses. *Journal of virology* **77**, 6007-6013.
- Imamoto, N., Shimamoto, T., Takao, T., Tachibana, T., Kose, S., Matsubae, M., Sekimoto, T., Shimonishi, Y. & Yoneda, Y. (1995).** In vivo evidence for involvement of a 58 kDa component of nuclear pore-targeting complex in nuclear protein import. *The EMBO journal* **14**, 3617-3626.
- Inglis, S. C., Barrett, T., Brown, C. M. & Almond, J. W. (1979).** The smallest genome RNA segment of influenza virus contains two genes that may overlap. *Proceedings of the National Academy of Sciences of the United States of America* **76**, 3790-3794.
- Jefferson, T., Demicheli, V., Rivetti, D., Jones, M., Di Pietrantonj, C. & Rivetti, A. (2006).** Antivirals for influenza in healthy adults: systematic review. *Lancet* **367**, 303-313.
- Kalderon, D., Roberts, B. L., Richardson, W. D. & Smith, A. E. (1984).** A short amino acid sequence able to specify nuclear location. *Cell* **39**, 499-509.
- Kemler, I., Whittaker, G. & Helenius, A. (1994).** Nuclear import of microinjected influenza virus ribonucleoproteins. *Virology* **202**, 1028-1033.
- Kimura, N., Nishida, M., Nagata, K., Ishihama, A., Oda, K. & Nakada, S. (1992).** Transcription of a recombinant influenza virus RNA in cells that can express the influenza virus RNA polymerase and nucleoprotein genes. *The Journal of general virology* **73** ( Pt 6), 1321-1328.
- Kingsbury, D. W. & Webster, R. G. (1969).** Some Properties of Influenza Virus Nucleocapsids. *Journal of virology* **4**, 219-225.

- Kistner, O., Muller, K. & Scholtissek, C. (1989).** Differential phosphorylation of the nucleoprotein of influenza A viruses. *The Journal of general virology* **70** ( Pt 9), 2421-2431.
- Klump, K., Ruigrok, R. W. & Baudin, F. (1997).** Roles of the influenza virus polymerase and nucleoprotein in forming a functional RNP structure. *The EMBO journal* **16**, 1248-1257.
- Kobayashi, M., Toyoda, T., Adyshev, D. M., Azuma, Y. & Ishihama, A. (1994).** Molecular dissection of influenza virus nucleoprotein: deletion mapping of the RNA binding domain. *Journal of virology* **68**, 8433-8436.
- Kobe, B. (1999).** Autoinhibition by an internal nuclear localization signal revealed by the crystal structure of mammalian importin alpha. *Nature structural biology* **6**, 388-397.
- Kochs, G., Garcia-Sastre, A. & Martinez-Sobrido, L. (2007).** Multiple anti-interferon actions of the influenza A virus NS1 protein. *Journal of virology* **81**, 7011-7021.
- Kohler, A. & Hurt, E. (2007).** Exporting RNA from the nucleus to the cytoplasm. *Nat Rev Mol Cell Biol* **8**, 761-773.
- Kohler, M., Fiebler, A., Hartwig, M., Thiel, S., Prehn, S., Kettritz, R., Luft, F. C. & Hartmann, E. (2002).** Differential expression of classical nuclear transport factors during cellular proliferation and differentiation. *Cell Physiol Biochem* **12**, 335-344.
- Kosugi, S., Hasebe, M., Tomita, M. & Yanagawa, H. (2008).** Nuclear Export Signal Consensus Sequences Defined Using a Localization-Based Yeast Selection System. *Traffic (Copenhagen, Denmark)*.
- Krug, R. M. (1981).** Priming of influenza viral RNA transcription by capped heterologous RNAs. *Current topics in microbiology and immunology* **93**, 125-149.
- Kudo, N., Matsumori, N., Taoka, H., Fujiwara, D., Schreiner, E. P., Wolff, B., Yoshida, M. & Horinouchi, S. (1999).** Leptomycin B inactivates CRM1/exportin 1 by covalent modification at a cysteine residue in the central conserved region. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 9112-9117.
- Kuersten, S., Ohno, M. & Mattaj, I. W. (2001).** Nucleocytoplasmic transport: Ran, beta and beyond. *Trends in cell biology* **11**, 497-503.
- Kutay, U., Bischoff, F. R., Kostka, S., Kraft, R. & Gorlich, D. (1997).** Export of importin alpha from the nucleus is mediated by a specific nuclear transport factor. *Cell* **90**, 1061-1071.
- Kutay, U., Lipowsky, G., Izaurralde, E., Bischoff, F. R., Schwarzmaier, P., Hartmann, E. & Gorlich, D. (1998).** Identification of a tRNA-specific nuclear export receptor. *Molecular cell* **1**, 359-369.
- Lamb, R. A. & Choppin, P. W. (1979).** Segment 8 of the influenza virus genome is unique in coding for two polypeptides. *Proceedings of the National Academy of Sciences of the United States of America* **76**, 4908-4912.
- Lamb, R. A. & Krug, R. M. (2001).** Orthomyxoviridae: the viruses and their replication. *Fields Virology 4th edition, (Knipe DM, Howley PM, eds) Philadelphia: Lippincott Williams & Wilkins.*
- Lamb, R. A., Lai, C. J. & Choppin, P. W. (1981).** Sequences of mRNAs derived from genome RNA segment 7 of influenza virus: colinear and interrupted mRNAs code for overlapping proteins. *Proceedings of the National Academy of Sciences of the United States of America* **78**, 4170-4174.
- Lanford, R. E. & Butel, J. S. (1984).** Construction and characterization of an SV40 mutant defective in nuclear transport of T antigen. *Cell* **37**, 801-813.
- Laue, T. M., Shah, B. D., Ridgeway, T. M. & Pelletier, S. L. (1992).** In Analytical Ultracentrifugation in Biochemistry and Polymer Science. *The Royal Society of Chemistry, (Cambridge Harding, S E, Rowe, A, J, and Horton, J C, eds)*, 90-125.

- Lee, S. J., Matsuura, Y., Liu, S. M. & Stewart, M. (2005).** Structural basis for nuclear import complex dissociation by RanGTP. *Nature* **435**, 693-696.
- Li, M. L., Rao, P. & Krug, R. M. (2001).** The active sites of the influenza cap-dependent endonuclease are on different polymerase subunits. *The EMBO journal* **20**, 2078-2086.
- Li, Z., Chen, H., Jiao, P., Deng, G., Tian, G., Li, Y., Hoffmann, E., Webster, R. G., Matsuoka, Y. & Yu, K. (2005).** Molecular basis of replication of duck H5N1 influenza viruses in a mammalian mouse model. *Journal of virology* **79**, 12058-12064.
- Liang, Y., Hong, Y. & Parslow, T. G. (2005).** cis-Acting packaging signals in the influenza virus PB1, PB2, and PA genomic RNA segments. *Journal of virology* **79**, 10348-10355.
- Lindstrom, S. E., Hiromoto, Y., Nerome, R., Omoe, K., Sugita, S., Yamazaki, Y., Takahashi, T. & Nerome, K. (1998).** Phylogenetic analysis of the entire genome of influenza A (H3N2) viruses from Japan: evidence for genetic reassortment of the six internal genes. *Journal of virology* **72**, 8021-8031.
- Lipowsky, G., Bischoff, F. R., Izaurralde, E., Kutay, U., Schafer, S., Gross, H. J., Beier, H. & Gorlich, D. (1999).** Coordination of tRNA nuclear export with processing of tRNA. *RNA (New York, NY)* **5**, 539-549.
- Liu, J., Lynch, P. A., Chien, C. Y., Montelione, G. T., Krug, R. M. & Berman, H. M. (1997).** Crystal structure of the unique RNA-binding domain of the influenza virus NS1 protein. *Nature structural biology* **4**, 896-899.
- Liu, Y. & Eisenberg, D. (2002).** 3D domain swapping: as domains continue to swap. *Protein Sci* **11**, 1285-1299.
- Luo, G. X., Luytjes, W., Enami, M. & Palese, P. (1991).** The polyadenylation signal of influenza virus RNA involves a stretch of uridines followed by the RNA duplex of the panhandle structure. *Journal of virology* **65**, 2861-2867.
- Luo, M., Green, T. J., Zhang, X., Tsao, J. & Qiu, S. (2007).** Structural comparisons of the nucleoprotein from three negative strand RNA virus families. *Virology journal* **4**, 72.
- Luytjes, W., Krystal, M., Enami, M., Parvin, J. D. & Palese, P. (1989).** Amplification, expression, and packaging of foreign gene by influenza virus. *Cell* **59**, 1107-1113.
- Ma, K., Roy, A. M. & Whittaker, G. R. (2001).** Nuclear export of influenza virus ribonucleoproteins: identification of an export intermediate at the nuclear periphery. *Virology* **282**, 215-220.
- Marsh, M. & Helenius, A. (2006).** Virus entry: open sesame. *Cell* **124**, 729-740.
- Martin-Benito, J., Area, E., Ortega, J., Llorca, O., Valpuesta, J. M., Carrascosa, J. L. & Ortin, J. (2001).** Three-dimensional reconstruction of a recombinant influenza virus ribonucleoprotein particle. *EMBO reports* **2**, 313-317.
- Mattaj, I. W. & Englmeier, L. (1998).** Nucleocytoplasmic transport: the soluble phase. *Annual review of biochemistry* **67**, 265-306.
- Melen, K., Fagerlund, R., Franke, J., Kohler, M., Kinnunen, L. & Julkunen, I. (2003).** Importin alpha nuclear localization signal binding sites for STAT1, STAT2, and influenza A virus nucleoprotein. *The Journal of biological chemistry* **278**, 28193-28200.
- Mena, I., Jambrina, E., Albo, C., Perales, B., Ortin, J., Arrese, M., Vallejo, D. & Portela, A. (1999).** Mutational analysis of influenza A virus nucleoprotein: identification of mutations that affect RNA replication. *Journal of virology* **73**, 1186-1194.
- Moroianu, J., Hijikata, M., Blobel, G. & Radu, A. (1995).** Mammalian karyopherin alpha 1 beta and alpha 2 beta heterodimers: alpha 1 or alpha 2 subunit binds nuclear localization signal and beta subunit interacts with peptide repeat-containing nucleoporins. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 6532-6536.



- Mukaigawa, J. & Nayak, D. P. (1991).** Two signals mediate nuclear localization of influenza virus (A/WSN/33) polymerase basic protein 2. *Journal of virology* **65**, 245-253.
- Murphy, B. R. & Webster, R. G. (1990).** Orthomyxoviruses. *Virology, second edition, (Field, BN, Knipe DN et al) Raven Press, Ltd, New York* **1**, 1091-1152.
- Nachury, M. V. & Weis, K. (1999).** The direction of transport through the nuclear pore can be inverted. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 9622-9627.
- Naffakh, N., Tomoiu, A., Rameix-Welti, M. A. & van der Werf, S. (2008).** Host restriction of avian influenza viruses at the level of the ribonucleoproteins. *Annual review of microbiology* **62**, 403-424.
- Naim, H. Y. & Roth, M. G. (1993).** Basis for selective incorporation of glycoproteins into the influenza virus envelope. *Journal of virology* **67**, 4831-4841.
- Nath, S. T. & Nayak, D. P. (1990).** Function of two discrete regions is required for nuclear localization of polymerase basic protein 1 of A/WSN/33 influenza virus (H1 N1). *Molecular and cellular biology* **10**, 4139-4145.
- Nayak, D. P., Hui, E. K. & Barman, S. (2004).** Assembly and budding of influenza virus. *Virus research* **106**, 147-165.
- Nemergut, M. E., Mizzen, C. A., Stukenberg, T., Allis, C. D. & Macara, I. G. (2001).** Chromatin docking and exchange activity enhancement of RCC1 by histones H2A and H2B. *Science (New York, NY)* **292**, 1540-1543.
- Nemeroff, M. E., Barabino, S. M., Li, Y., Keller, W. & Krug, R. M. (1998).** Influenza virus NS1 protein interacts with the cellular 30 kDa subunit of CPSF and inhibits 3'end formation of cellular pre-mRNAs. *Molecular cell* **1**, 991-1000.
- Neumann, G., Castrucci, M. R. & Kawaoka, Y. (1997).** Nuclear import and export of influenza virus nucleoprotein. *Journal of virology* **71**, 9690-9700.
- Neumann, G., Hughes, M. T. & Kawaoka, Y. (2000).** Influenza A virus NS2 protein mediates vRNP nuclear export through NES-independent interaction with hCRM1. *The EMBO journal* **19**, 6751-6758.
- Neumann, G. & Kawaoka, Y. (2006).** Host range restriction and pathogenicity in the context of influenza pandemic. *Emerging infectious diseases* **12**, 881-886.
- 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., Walz, T., Wang, J. H. & Shaw, P. C. (2008).** Structure of the influenza virus A H5N1 nucleoprotein: implications for RNA binding, oligomerization, and vaccine design. *FASEB J* **22**, 3638-3647.
- Nieto, A., de la Luna, S., Barcena, J., Portela, A. & Ortin, J. (1994).** Complex structure of the nuclear translocation signal of influenza virus polymerase PA subunit. *The Journal of general virology* **75 ( Pt 1)**, 29-36.
- Nishimura, K., Kim, S., Zhang, L. & Cross, T. A. (2002).** The closed state of a H<sup>+</sup> channel helical bundle combining precise orientational and distance restraints from solid state NMR. *Biochemistry* **41**, 13170-13177.
- Noda, T., Sagara, H., Yen, A., Takada, A., Kida, H., Cheng, R. H. & Kawaoka, Y. (2006).** Architecture of ribonucleoprotein complexes in influenza A virus particles. *Nature* **439**, 490-492.
- Obayashi, E., Yoshida, H., Kawai, F., Shibayama, N., Kawaguchi, A., Nagata, K., Tame, J. R. & Park, S. Y. (2008).** The structural basis for an essential subunit interaction in influenza virus RNA polymerase. *Nature* **454**, 1127-1131.

- Obenauer, J. C., Denson, J., Mehta, P. K., Su, X., Mukatira, S., Finkelstein, D. B., Xu, X., Wang, J., Ma, J., Fan, Y., Rakestraw, K. M., Webster, R. G., Hoffmann, E., Krauss, S., Zheng, J., Zhang, Z. & Naeve, C. W. (2006).** Large-scale sequence analysis of avian influenza isolates. *Science (New York, NY)* **311**, 1576-1580.
- Ohno, M., Segref, A., Bachi, A., Wilm, M. & Mattaj, I. W. (2000).** PHAX, a mediator of U snRNA nuclear export whose activity is regulated by phosphorylation. *Cell* **101**, 187-198.
- Olsen, C. W. (2002).** The emergence of novel swine influenza viruses in North America. *Virus research* **85**, 199-210.
- OMS (2003).** La grippe. <http://www.who.int/mediacentre/factsheets/fs211/fr/>.
- O'Neill, R. E., Jaskunas, R., Blobel, G., Palese, P. & Moroianu, J. (1995).** Nuclear import of influenza virus RNA can be mediated by viral nucleoprotein and transport factors required for protein import. *The Journal of biological chemistry* **270**, 22701-22704.
- O'Neill, R. E., Talon, J. & Palese, P. (1998).** The influenza virus NP (NS2 protein) mediates the nuclear export of viral ribonucleoproteins. *The EMBO journal* **17**, 288-296.
- Ortega, J., Martin-Benito, J., Zurcher, T., Valpuesta, J. M., Carrascosa, J. L. & Ortin, J. (2000).** Ultrastructural and functional analyses of recombinant influenza virus ribonucleoproteins suggest dimerization of nucleoprotein during virus amplification. *Journal of virology* **74**, 156-163.
- Ossareh-Nazari, B., Bachelierie, F. & Dargemont, C. (1997).** Evidence for a role of CRM1 in signal-mediated nuclear protein export. *Science (New York, NY)* **278**, 141-144.
- Osterhaus, A. D., Rimmelzwaan, G. F., Martina, B. E., Bestebroer, T. M. & Fouchier, R. A. (2000).** Influenza B virus in seals. *Science (New York, NY)* **288**, 1051-1053.
- Ostermeier, M. & Lutz, S. (2003).** The creation of ITCHY hybrid protein libraries. *Methods in molecular biology (Clifton, NJ)* **231**, 129-141.
- Otsuka, S., Iwasaka, S., Yoneda, Y., Takeyasu, K. & Yoshimura, S. H. (2008).** Individual binding pockets of importin- $\beta$  for FG-nucleoporins have different binding properties and different sensitivities to RanGTP. *Proceedings of the National Academy of Sciences of the United States of America*.
- Ozawa, M., Fujii, K., Muramoto, Y., Yamada, S., Yamayoshi, S., Takada, A., Goto, H., Horimoto, T. & Kawaoka, Y. (2007).** Contributions of two nuclear localization signals of influenza A virus nucleoprotein to viral replication. *Journal of virology* **81**, 30-41.
- Palese, P. & Shaw, M. (2007).** Orthomyxoviridae: the viruses and their replication. *Fields Virology 5th edition, (Knipe DM, Howley PM, eds) Philadelphia: Lippincott Williams & Wilkins*.
- Pante, N. & Kann, M. (2002).** Nuclear pore complex is able to transport macromolecules with diameters of about 39 nm. *Molecular biology of the cell* **13**, 425-434.
- Patel, S. S., Belmont, B. J., Sante, J. M. & Rexach, M. F. (2007).** Natively unfolded nucleoporins gate protein diffusion across the nuclear pore complex. *Cell* **129**, 83-96.
- Petosa, C., Schoehn, G., Askjaer, P., Bauer, U., Moulin, M., Steuerwald, U., Soler-Lopez, M., Baudin, F., Mattaj, I. W. & Muller, C. W. (2004).** Architecture of CRM1/Exportin1 suggests how cooperativity is achieved during formation of a nuclear export complex. *Molecular cell* **16**, 761-775.
- Pinto, L. H., Holsinger, L. J. & Lamb, R. A. (1992).** Influenza virus M2 protein has ion channel activity. *Cell* **69**, 517-528.

- Plotch, S. J., Bouloy, M. & Krug, R. M. (1979).** Transfer of 5'-terminal cap of globin mRNA to influenza viral complementary RNA during transcription in vitro. *Proceedings of the National Academy of Sciences of the United States of America* **76**, 1618-1622.
- Plotch, S. J., Bouloy, M., Ulmanen, I. & Krug, R. M. (1981).** A unique cap(m7GpppXm)-dependent influenza virion endonuclease cleaves capped RNAs to generate the primers that initiate viral RNA transcription. *Cell* **23**, 847-858.
- Poch, O., Sauvaget, I., Delarue, M. & Tordo, N. (1989).** Identification of four conserved motifs among the RNA-dependent polymerase encoding elements. *The EMBO journal* **8**, 3867-3874.
- Pons, M. W., Schulze, I. T., Hirst, G. K. & Hauser, R. (1969).** Isolation and characterization of the ribonucleoprotein of influenza virus. *Virology* **39**, 250-259.
- Portela, A. & Digard, P. (2002).** The influenza virus nucleoprotein: a multifunctional RNA-binding protein pivotal to virus replication. *The Journal of general virology* **83**, 723-734.
- Rabut, G., Lenart, P. & Ellenberg, J. (2004).** Dynamics of nuclear pore complex organization through the cell cycle. *Current opinion in cell biology* **16**, 314-321.
- Radu, A., Blobel, G. & Moore, M. S. (1995).** Identification of a protein complex that is required for nuclear protein import and mediates docking of import substrate to distinct nucleoporins. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 1769-1773.
- Rao, P., Yuan, W. & Krug, R. M. (2003).** Crucial role of CA cleavage sites in the cap-snatching mechanism for initiating viral mRNA synthesis. *The EMBO journal* **22**, 1188-1198.
- Rexach, M. & Blobel, G. (1995).** Protein import into nuclei: association and dissociation reactions involving transport substrate, transport factors, and nucleoporins. *Cell* **83**, 683-692.
- Richardson, J. C. & Akkina, R. K. (1991).** NS2 protein of influenza virus is found in purified virus and phosphorylated in infected cells. *Archives of virology* **116**, 69-80.
- Robertson, J. S., Schubert, M. & Lazzarini, R. A. (1981).** Polyadenylation sites for influenza virus mRNA. *Journal of virology* **38**, 157-163.
- Rudolph, M. G., Kraus, I., Dickmanns, A., Eickmann, M., Garten, W. & Ficner, R. (2003).** Crystal structure of the borna disease virus nucleoprotein. *Structure* **11**, 1219-1226.
- Ruigrok, R. W., Barge, A., Durrer, P., Brunner, J., Ma, K. & Whittaker, G. R. (2000).** Membrane interaction of influenza virus M1 protein. *Virology* **267**, 289-298.
- Ruigrok, R. W. & Baudin, F. (1995).** Structure of influenza virus ribonucleoprotein particles. II. Purified RNA-free influenza virus ribonucleoprotein forms structures that are indistinguishable from the intact influenza virus ribonucleoprotein particles. *The Journal of general virology* **76 ( Pt 4)**, 1009-1014.
- Russell, C. J. & Webster, R. G. (2005).** The genesis of a pandemic influenza virus. *Cell* **123**, 368-371.
- Sakaguchi, A., Hirayama, E., Hiraki, A., Ishida, Y. & Kim, J. (2003).** Nuclear export of influenza viral ribonucleoprotein is temperature-dependently inhibited by dissociation of viral matrix protein. *Virology* **306**, 244-253.
- Salomon, R., Franks, J., Govorkova, E. A., Ilyushina, N. A., Yen, H. L., Hulse-Post, D. J., Humberd, J., Trichet, M., Rehg, J. E., Webby, R. J., Webster, R. G. & Hoffmann, E. (2006).** The polymerase complex genes contribute to the high virulence of the human H5N1 influenza virus isolate A/Vietnam/1203/04. *The Journal of experimental medicine* **203**, 689-697.

- Scheffzek, K., Klebe, C., Fritz-Wolf, K., Kabsch, W. & Wittinghofer, A. (1995).** Crystal structure of the nuclear Ras-related protein Ran in its GDP-bound form. *Nature* **374**, 378-381.
- Schnell, J. R. & Chou, J. J. (2008).** Structure and mechanism of the M2 proton channel of influenza A virus. *Nature* **451**, 591-595.
- Schuck, P. (2000).** Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophysical journal* **78**, 1606-1619.
- Schulze, I. T. (1972).** The structure of influenza virus. II. A model based on the morphology and composition of subviral particles. *Virology* **47**, 181-196.
- Schwartz, T. U. (2005).** Modularity within the architecture of the nuclear pore complex. *Current opinion in structural biology* **15**, 221-226.
- Sha, B. & Luo, M. (1997).** Structure of a bifunctional membrane-RNA binding protein, influenza virus matrix protein M1. *Nature structural biology* **4**, 239-244.
- Shi, L., Galarza, J. M. & Summers, D. F. (1996).** Recombinant-baculovirus-expressed PB2 subunit of the influenza A virus RNA polymerase binds cap groups as an isolated subunit. *Virus research* **42**, 1-9.
- Shu, L. L., Bean, W. J. & Webster, R. G. (1993).** Analysis of the evolution and variation of the human influenza A virus nucleoprotein gene from 1933 to 1990. *Journal of virology* **67**, 2723-2729.
- Sitterlin, D. (2005).** Quand la virologie dévoile les secrets de l'export des ARN. *Virologie* **9**, 191-200.
- Skehel, J. J. & Waterfield, M. D. (1975).** Studies on the primary structure of the influenza virus hemagglutinin. *Proceedings of the National Academy of Sciences of the United States of America* **72**, 93-97.
- Stoffler, D., Feja, B., Fahrenkrog, B., Walz, J., Typke, D. & Aebi, U. (2003).** Cryo-electron tomography provides novel insights into nuclear pore architecture: implications for nucleocytoplasmic transport. *Journal of molecular biology* **328**, 119-130.
- Subbarao, E. K., London, W. & Murphy, B. R. (1993).** A single amino acid in the PB2 gene of influenza A virus is a determinant of host range. *Journal of virology* **67**, 1761-1764.
- Sun, C., Yang, W., Tu, L. C. & Musser, S. M. (2008).** Single-molecule measurements of importin alpha/cargo complex dissociation at the nuclear pore. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 8613-8618.
- Tarendeau, F., Boudet, J., Guilligay, D., Mas, P. J., Bougault, C. M., Boulo, S., Baudin, F., Ruigrok, R. W., Daigle, N., Ellenberg, J., Cusack, S., Simorre, J. P. & Hart, D. J. (2007).** Structure and nuclear import function of the C-terminal domain of influenza virus polymerase PB2 subunit. *Nature structural & molecular biology* **14**, 229-233.
- Tarendeau, F., Crepin, T., Guilligay, D., Ruigrok, R. W., Cusack, S. & Hart, D. J. (2008).** Host determinant residue lysine 627 lies on the surface of a discrete, folded domain of influenza virus polymerase PB2 subunit. *PLoS pathogens* **4**, e1000136.
- Tashiro, M., Ciborowski, P., Klenk, H. D., Pulverer, G. & Rott, R. (1987).** Role of Staphylococcus protease in the development of influenza pneumonia. *Nature* **325**, 536-537.
- Taubenberger, J. K., Reid, A. H., Lourens, R. M., Wang, R., Jin, G. & Fanning, T. G. (2005).** Characterization of the 1918 influenza virus polymerase genes. *Nature* **437**, 889-893.
- Taylor, R. M. (1949).** Studies on Survival of Influenza Virus Between Epidemics and Antigenic Variants of the Virus. *American journal of public health and the nation's health* **39**, 171-178.

- Terry, L. J. & Wente, S. R. (2007).** Nuclear mRNA export requires specific FG nucleoporins for translocation through the nuclear pore complex. *The Journal of cell biology* **178**, 1121-1132.
- Torreira, E., Schoehn, G., Fernandez, Y., Jorba, N., Ruigrok, R. W., Cusack, S., Ortin, J. & Llorca, O. (2007).** Three-dimensional model for the isolated recombinant influenza virus polymerase heterotrimer. *Nucleic acids research* **35**, 3774-3783.
- Varghese, J. N. & Colman, P. M. (1991).** Three-dimensional structure of the neuraminidase of influenza virus A/Tokyo/3/67 at 2.2 Å resolution. *Journal of molecular biology* **221**, 473-486.
- Varghese, J. N., Laver, W. G. & Colman, P. M. (1983).** Structure of the influenza virus glycoprotein antigen neuraminidase at 2.9 Å resolution. *Nature* **303**, 35-40.
- Vetter, I. R., Arndt, A., Kutay, U., Gorlich, D. & Wittinghofer, A. (1999).** Structural view of the Ran-Importin beta interaction at 2.3 Å resolution. *Cell* **97**, 635-646.
- Vreede, F. T., Jung, T. E. & Brownlee, G. G. (2004).** Model suggesting that replication of influenza virus is regulated by stabilization of replicative intermediates. *Journal of virology* **78**, 9568-9572.
- Wang, P., Palese, P. & O'Neill, R. E. (1997).** The NPI-1/NPI-3 (karyopherin alpha) binding site on the influenza A virus nucleoprotein NP is a nonconventional nuclear localization signal. *Journal of virology* **71**, 1850-1856.
- Wang, W., Riedel, K., Lynch, P., Chien, C. Y., Montelione, G. T. & Krug, R. M. (1999).** RNA binding by the novel helical domain of the influenza virus NS1 protein requires its dimer structure and a small number of specific basic amino acids. *RNA (New York, NY)* **5**, 195-205.
- Ward, A. C., Castelli, L. A., Lucantoni, A. C., White, J. F., Azad, A. A. & Macreadie, I. G. (1995).** Expression and analysis of the NS2 protein of influenza A virus. *Archives of virology* **140**, 2067-2073.
- Ward, C. W., Gleeson, P. A. & Dopheide, T. A. (1980).** Carbohydrate composition of the oligosaccharide units of the haemagglutinin from the Hong Kong influenza virus A/Memphis/102/72. *The Biochemical journal* **189**, 649-652.
- Watanabe, K., Handa, H., Mizumoto, K. & Nagata, K. (1996).** Mechanism for inhibition of influenza virus RNA polymerase activity by matrix protein. *Journal of virology* **70**, 241-247.
- Watanabe, K., Takizawa, N., Katoh, M., Hoshida, K., Kobayashi, N. & Nagata, K. (2001).** Inhibition of nuclear export of ribonucleoprotein complexes of influenza virus by leptomycin B. *Virus research* **77**, 31-42.
- Weber, F., Kochs, G., Gruber, S. & Haller, O. (1998).** A classical bipartite nuclear localization signal on Thogoto and influenza A virus nucleoproteins. *Virology* **250**, 9-18.
- Weis, K. (2002).** Nucleocytoplasmic transport: cargo trafficking across the border. *Current opinion in cell biology* **14**, 328-335.
- Weis, K. (2003).** Regulating access to the genome: nucleocytoplasmic transport throughout the cell cycle. *Cell* **112**, 441-451.
- Weis, W. I., Brunger, A. T., Skehel, J. J. & Wiley, D. C. (1990).** Refinement of the influenza virus hemagglutinin by simulated annealing. *Journal of molecular biology* **212**, 737-761.
- White, C. L., Janakiraman, M. N., Laver, W. G., Philippon, C., Vasella, A., Air, G. M. & Luo, M. (1995).** A sialic acid-derived phosphonate analog inhibits different strains of influenza virus neuraminidase with different efficiencies. *Journal of molecular biology* **245**, 623-634.

- Whittaker, G., Bui, M. & Helenius, A. (1996).** Nuclear trafficking of influenza virus ribonucleoproteins in heterokaryons. *Journal of virology* **70**, 2743-2756.
- Wilson, I. A., Skehel, J. J. & Wiley, D. C. (1981).** Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. *Nature* **289**, 366-373.
- Yamanaka, K., Ishihama, A. & Nagata, K. (1990).** Reconstitution of influenza virus RNA-nucleoprotein complexes structurally resembling native viral ribonucleoprotein cores. *The Journal of biological chemistry* **265**, 11151-11155.
- Yang, Q., Rout, M. P. & Akey, C. W. (1998).** Three-dimensional architecture of the isolated yeast nuclear pore complex: functional and evolutionary implications. *Molecular cell* **1**, 223-234.
- Yasuda, J., Nakada, S., Kato, A., Toyoda, T. & Ishihama, A. (1993).** Molecular assembly of influenza virus: association of the NS2 protein with virion matrix. *Virology* **196**, 249-255.
- Ye, Q., Krug, R. M. & Tao, Y. J. (2006).** The mechanism by which influenza A virus nucleoprotein forms oligomers and binds RNA. *Nature* **444**, 1078-1082.
- Ye, Y. & Godzik, A. (2003).** Flexible structure alignment by chaining aligned fragment pairs allowing twists. *Bioinformatics (Oxford, England)* **19 Suppl 2**, ii246-255.
- Ye, Z., Liu, T., Offringa, D. P., McInnis, J. & Levandowski, R. A. (1999).** Association of influenza virus matrix protein with ribonucleoproteins. *Journal of virology* **73**, 7467-7473.
- Ye, Z., Robinson, D. & Wagner, R. R. (1995).** Nucleus-targeting domain of the matrix protein (M1) of influenza virus. *Journal of virology* **69**, 1964-1970.
- Ye, Z. P., Pal, R., Fox, J. W. & Wagner, R. R. (1987).** Functional and antigenic domains of the matrix (M1) protein of influenza A virus. *Journal of virology* **61**, 239-246.
- Zachariae, U. & Grubmüller, H. (2008).** Importin-beta: structural and dynamic determinants of a molecular spring. *Structure* **16**, 906-915.

# **Annexe**

## Liste des publications

**Le trafic nucléocytoplasmique des protéines et des ribonucléoprotéines du virus de la grippe.**

Akarsu, H., Boulo, S., Ruigrok, R.W.H. et Baudin, F.  
Virologie. 2006 Juillet-Août; Volume 10, Numéro 4:301-309.

**Nuclear traffic of influenza virus proteins and ribonucleoprotein complexes.**

Boulo, S., Akarsu, H., Ruigrok, R.W. and Baudin, F.  
Virus Res. 2007 Mar;124(1-2):12-21. Epub 2006 Nov 1.

**Structure and nuclear import function of the C-terminal domain of influenza virus polymerase PB2 subunit.**

Tarendeau, F., Boudet, J., Guilligay, D., Mas, P.J., Bougault, C.M., Boulo, S., Baudin, F., Ruigrok, R.W., Daigle, N., Ellenberg, J., Cusack, S., Simorre, J.P. & Hart, D.J.  
Nat Struct Mol Biol. 2007 Mar;14(3):229-33. Epub 2007 Feb 25.

**Influenza virus Nucleoprotein oligomerisation state correlates with RNA binding affinity.**

Boulo, S., Lotteau, V., Ruigrok, R.W. and Baudin, F.  
*Soumis.*

**Characterization of Influenza virus Nucleoprotein bound to nuclear transporter importin alpha 5.**

Boulo, S., Akarsu, H., Rybin, V., Lotteau, V., Ruigrok, R.W., Müller, C., and Baudin, F.  
*En préparation.*



**Le trafic nucléocytoplasmique des  
protéines et des ribonucléoprotéines du  
virus de la grippe.**

Akarsu, H., Boulo, S., Ruigrok, R.W.H. et Baudin, F.

Virologie. 2006 Juillet-Août; Volume 10, Numéro 4:301-309.

# Le trafic nucléocytoplasmique des protéines et des ribonucléoprotéines du virus de la grippe

H. Akarsu  
S. Boulo  
R.W.H. Ruigrok  
F. Baudin

Institut de virologie moléculaire et structurale,  
FRE 2854 CNRS-UJF,  
c/o EMBL,  
BP181,  
38042 Grenoble Cedex 9  
<baudin@embl.fr>

**Résumé.** Le virus de la grippe est un virus à ARN négatif et représente l'un des rares virus à ARN à se répliquer dans le noyau. L'ARN viral est associé à 4 protéines virales, formant des particules ribonucléoprotéiques (RNP). Après entrée dans la cellule, les RNP sont dissociées de la protéine virale de matrice M1 suite à l'acidification de l'endosome et sont activement importées dans le noyau de la cellule. Après traduction des ARN messagers viraux, les protéines nécessaires pour l'assemblage de nouvelles RNP (la nucléoprotéine et les trois sous-unités du complexe de polymérase) sont importées dans le noyau. De plus, une fraction de la protéine M1 nouvellement synthétisée ainsi que NEP (protéine d'export nucléaire) sont également importées. Les signaux nucléaires de localisation sur toutes ces protéines virales et leurs interactions avec le système cellulaire de transport sont discutés. Dans le noyau, la protéine M1 se fixe aux RNP nouvellement assemblées et à NEP. NEP contient le signal d'export nucléaire nécessaire au transport des RNP vers le cytoplasme en vue du bourgeonnement de nouvelles particules virales. Il semble y avoir un jeu complexe d'exposition et de protection des signaux de transport nucléaires qui mène à un transport unidirectionnel des RNP vers le noyau au début du processus d'infection et à un export unidirectionnel des RNP à la fin de l'infection.

**Mots clés :** virus de la grippe, réplication, transport nucléaire, interactions virus-hôte

**Abstract.** Influenza virus is a negative strand RNA virus and is one of the rare RNA viruses to replicate in the nucleus. The viral RNA is associated with 4 viral proteins to make ribonucleoprotein particles (RNPs). After cell entry the RNPs are dissociated from the viral matrix protein in the low pH of the endosome and are actively imported into the cell nucleus. After translation of viral mRNAs, the proteins necessary for the assembly of new RNPs (the nucleoprotein and the three subunits of the polymerase complex) are also imported into the nucleus. Apart from these four proteins, part of the newly made matrix protein is also imported and the NEP (nuclear export protein) enters the nucleus probably through diffusion. The nuclear localisation signals on all these viral proteins and their interaction with the cellular transport system are discussed. In the nucleus, the matrix protein binds to the newly assembled RNPs and NEP then binds to the matrix protein. NEP contains the nuclear export signal necessary for transport of the RNPs to the cytoplasm for the budding of new virus particles. There appears to be an intricate ballet in exposing and hiding nuclear transport signals which leads to a unidirectional transport of the RNPs to the nucleus at the start of the infection process and an opposite unidirectional export of RNPs at the end of the infection.

**Key words:** influenza virus, replication, nuclear transport, virus-host interactions

**Tirés à part :** F. Baudin

Les virus à ARN négatifs, tous enveloppés, sont responsables de pathologies très répandues chez l'homme : oreillons, rougeole, bronchiolite (*Paramyxoviridae*), rage (*Rhabdoviridae*), grippe (*Orthomyxoviridae*). D'autres, moins communs en Occident, sont à l'origine de fièvres hémorragiques foudroyantes comme les virus de Lassa (*Arenaviridae*), Ebola (*Filoviridae*) ou encore le virus de la vallée du Rift (*Bunyaviridae*). En ce qui concerne la grippe, il s'agit d'un virus à large spectre d'hôtes, avec pour réservoir naturel les oiseaux sauvages aquatiques. Le passage d'une espèce à l'autre, par accumulation de multiples mutations ou réassortiment entre différents virus dans une cellule doublement infectée, conduit à l'apparition de nouvelles souches potentiellement très virulentes pour l'homme et donc à de graves pandémies cycliques. L'extrême virulence de ces nouvelles souches de grippe peut être liée d'une part à une résistance du virus aux réponses cytokines de l'hôte, associée à un choc inflammatoire, d'autre part à une réplication pulmonaire beaucoup plus forte que celle des virus dans des épidémies humaines régulières [1].

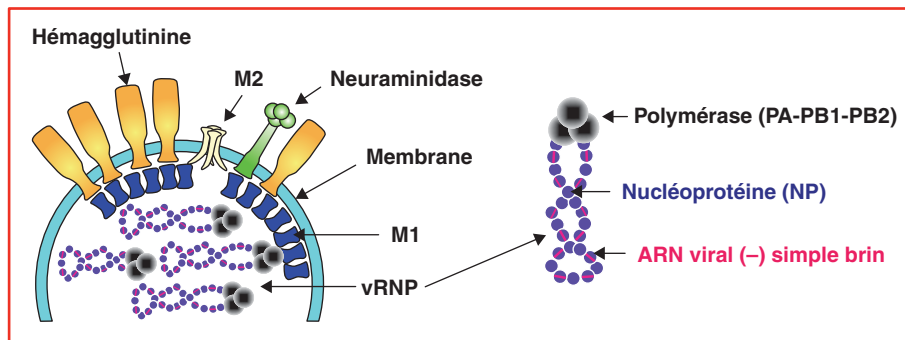
Les *Orthomyxoviridae* – tout comme les *Arena* et *Bunyaviridae* – possèdent un génome segmenté. Celui du virus de la grippe (ou influenza) est segmenté en 8 molécules d'ARN dont la taille varie de 890 à 2 341 nucléotides. Chacune des molécules d'ARN virale est recouverte par la nucléoprotéine (NP) de poids moléculaire 56 kDa. La nucléoprotéine fixe environ 24 nucléotides par monomère [2]. Elle possède des domaines spécifiques pour l'interaction avec l'ARN viral et d'autres pour les interactions NP-NP [3, 4] (figures 1 et 2). Le complexe de la polymérase virale (PA, PB1, PB2) se fixe aux deux extrémités des molécules d'ARN, formant une molécule d'ARN circulaire non covalent (figure 1). L'ensemble (ARN, NP, polymérase) forme une particule ribonucléoprotéique virale, communément appelée vRNP. Ces 8 vRNP renferment l'information nécessaire à l'expression des protéines virales et représentent des unités indépendantes pour la transcription et la réplication.

L'originalité du virus de la grippe tient au fait qu'il est de ces rares virus à ARN dont les étapes de transcription et de réplication sont nucléaires. En effet, contrairement aux virus à ADN qui effectuent en général leurs étapes de transcription et de réplication dans le noyau de la cellule hôte (avec quelques exceptions comme le virus de la vaccine), les virus à ARN les effectuent tous dans le cytoplasme, à l'exception des Bornavirus et du virus de la grippe. Les virus dont le génome passe dans le noyau tirent en général profit de la machinerie nucléaire. Les virus à ADN utilisent l'ARN polymérase II pour la fabrication de leur ARNm. Les virus à ARN, eux, doivent apporter leur propre ARN polymérase ARN-dépendante. En revanche, ils empruntent à la cellule hôte la machinerie d'épissage des ARN messagers (ARNm), nécessaire à leurs besoins. Les segments 7 et 8 du génome du virus de la grippe codent chacun pour deux protéines différentes produites à partir de l'épissage des ARNm viraux correspondants sans séquence intronique : M1 et M2 pour le segment 7, NS1 et NEP pour le segment 8. Le virus de la grippe utilise également les ARN messagers cellulaires présents dans le noyau pour en utiliser la coiffe nécessaire à la synthèse de ses ARNm viraux. Ce phénomène est connu sous le terme de *cap-snatching*.

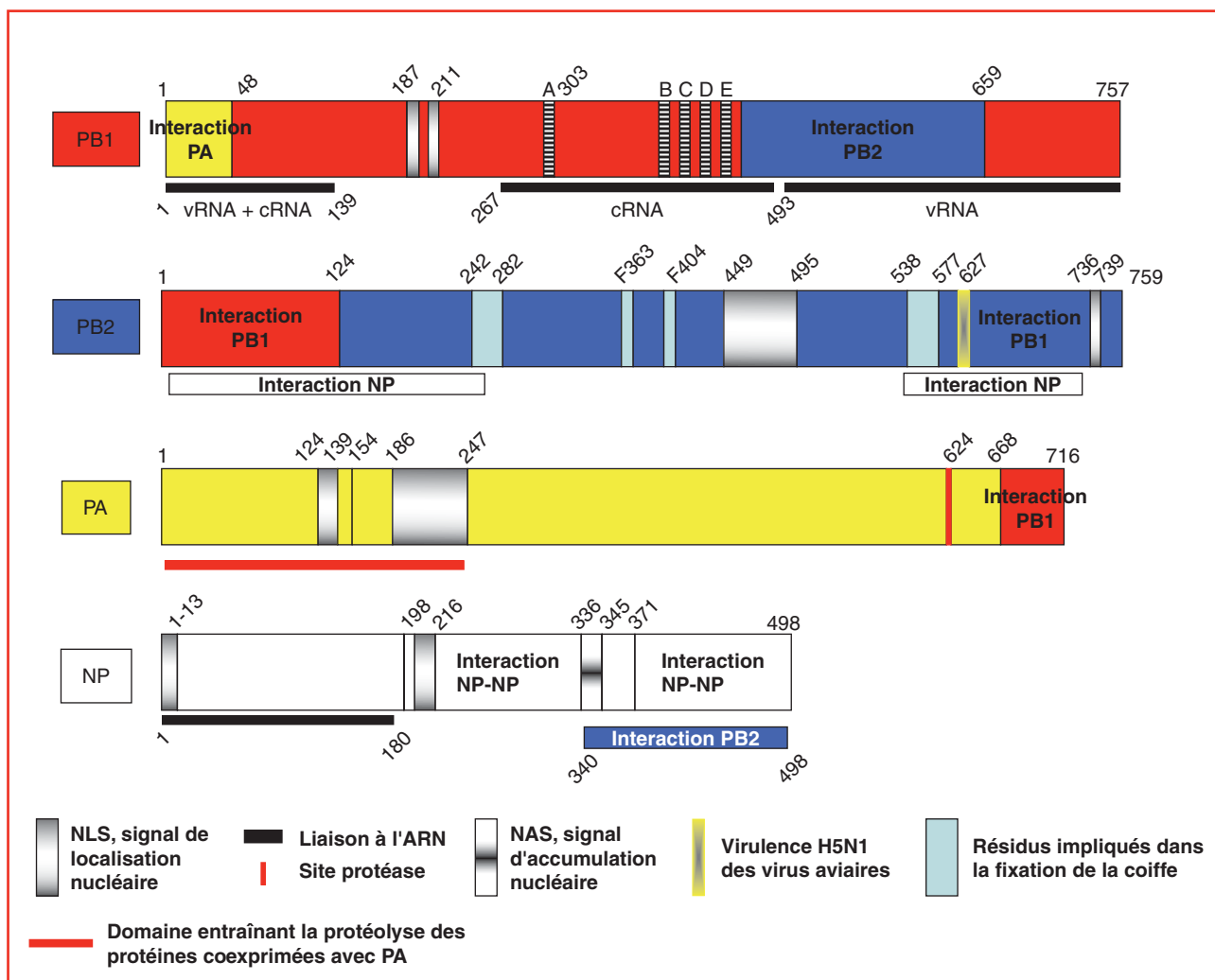
Ces étapes nucléaires de réplication et de transcription impliquent que le génome voyage entre le cytoplasme et le noyau. Dans cette revue, nous détaillerons les signaux connus d'import et d'export nucléaire des protéines du virus de la grippe.

### Le transport nucléaire, import et export

Les protéines ou complexes ARN-protéines substrats d'import ou d'export nucléaire sont identifiés grâce à des séquences d'adressage spécifiques – séquence de localisation nucléaire (NLS) pour l'import, séquence d'export



**Figure 1.** Représentation schématique du virus de la grippe (gauche) et des particules ribonucléoprotéiques (RNP) virales (droite). Les deux protéines de surface hémagglutinine et neuraminidase sont ancrées dans la membrane lipidique. La protéine M1 contacte la membrane lipidique ainsi que les RNP virales présentes à l'intérieur du virion. Les RNP contiennent l'information génétique (ARN segmenté) recouverte par la nucléoprotéine ainsi que le complexe de la polymérase virale.



**Figure 2.** Les trois sous-unités de la polymérase et la nucléoprotéine sont représentées avec leurs différents domaines NLS ainsi que ceux de leurs activités biologiques respectives. PB1 interagit avec PB2 et PA et elle possède deux signaux de localisation nucléaire (NLS) et trois sites de fixation pour l'ARN viral (ARNv, brin négatif) et l'ARN complémentaire (ARNc, brin positif, intermédiaire de la réplication). Les cinq barres rayées (A-E) indiquent les cinq motifs typiques des ARN polymérase-ARN-dépendantes [51]. PB2 interagit avec PB1 et avec la nucléoprotéine. La protéine possède plusieurs résidus impliqués dans la fixation de la coiffe des ARN messagers cellulaires, utilisés comme amorce pour la transcription et un NLS (aa 736-739) ainsi qu'un domaine dont la délétion entraîne une distribution cytoplasmique (449-495). Le résidu 627 est un déterminant de restriction d'hôte [52]. PA interagit seulement avec PB1 et possède également deux motifs NLS. La nucléoprotéine a des domaines d'autopolymérisation, un domaine d'interaction avec PB2 et un domaine d'interaction avec l'ARN viral. La molécule possède deux NLS et un domaine servant à l'accumulation nucléaire (NAS).

nucléaire (NES) pour l'export – qui permettent leur reconnaissance par des transporteurs (importines ou exportines) [5]. Une fois le cargo fixé à son transporteur, la translocation du complexe est assurée à travers le pore nucléaire.

### Le pore nucléaire

Le pore nucléaire représente l'unique voie de passage entre le noyau et le cytoplasme. Il en existe plusieurs centaines par noyau, leur nombre variant suivant le type et l'état

cellulaires. Chaque pore est en réalité un immense complexe protéique (NPC pour *nuclear pore complex*) de 125 MDa environ chez les mammifères et composé d'une centaine de protéines appelées nucléoporines [6]. Le pore nucléaire apparaît comme une structure cylindrique avec une symétrie d'ordre 8. Il est constitué d'un transporteur ou canal central de 10 nm de diamètre au repos, relié par 8 bras à 2 anneaux, l'un cytoplasmique et l'autre nucléaire. Des fibrilles décorent ces anneaux, formant une structure dite en « panier de basket » du côté nucléoplasmique. Si les ions et

les molécules de poids moléculaire inférieur à 40 kDa peuvent diffuser librement à travers ce pore nucléaire, la translocation de toute molécule de poids moléculaire supérieur à 40 kDa doit s'opérer de façon active à travers le canal central dont le diamètre peut s'élargir jusqu'à 30 nm.

Certaines des nucléoporines possèdent des domaines comportant de nombreuses répétitions d'acides aminés phénylalanine (F) et glycine (G) reconnues par les importines  $\beta$  (cf. *infra*). Une fois dans le compartiment d'arrivée, les substrats transportés sont libérés et les récepteurs sont recyclés.

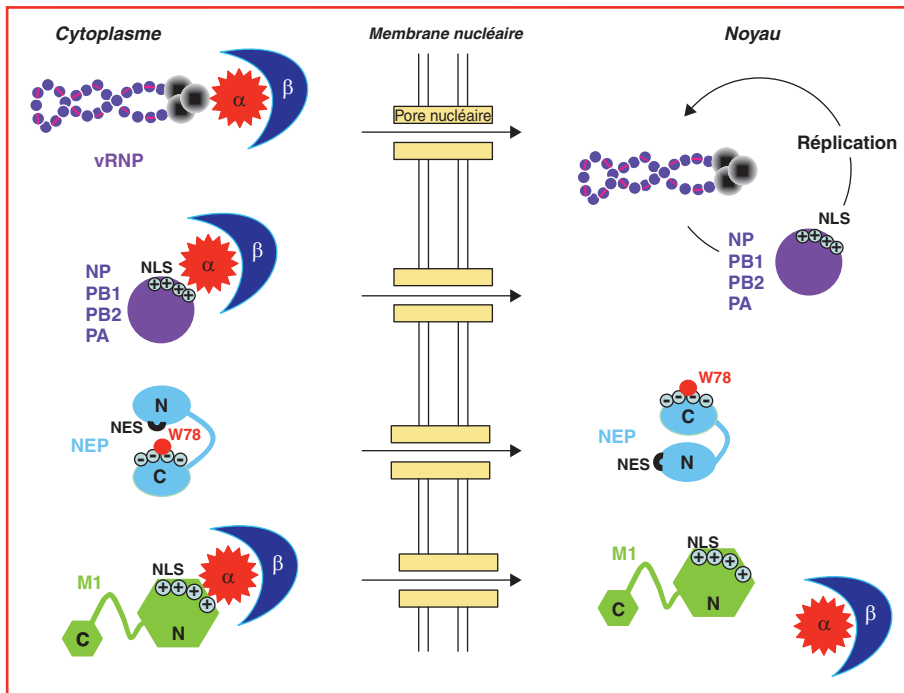
**Les importines**

Les molécules qui vont du cytoplasme à travers les NPC vers le noyau sont prises en charge par des protéines appelées importines (ou karyophérines) [7]. Ces importines sont classées en deux catégories,  $\alpha$  et  $\beta$ , elles-mêmes regroupées dans la superfamille des importines  $\beta$  caractérisée par une structure particulière composée de motifs répétitifs, les *armadillos* [8, 9]. La famille des importines  $\alpha$  reconnaît le NLS du cargo [10]. Quand le cargo se fixe sur l'importine  $\alpha$ , celle-ci est reconnue par l'importine  $\beta$  qui permet la fixation au NPC [11], ce qui conduit à la translocation du cargo dans le noyau (figure 3). Cet événement séquentiel n'est pas le seul moyen d'import nucléaire actif [12] mais il représente néanmoins le modèle le plus répandu dans la littérature, utilisé également par le génome du virus de la grippe.

Il existe plusieurs importines  $\alpha$  qui reconnaissent différents cargos mais il existe une seule importine  $\beta$ . Ainsi, l'importine  $\alpha$  représente la molécule adaptatrice entre le cargo et le système d'import nucléaire [13]. Le signal de localisation nucléaire (NLS) correspond au domaine d'interaction avec l'importine  $\alpha$  [10]. Le motif NLS est caractérisé par une forte concentration de résidus aminés basiques (lysine et/ou arginine) et peut être monopartite ou bipartite. Brièvement, le NLS monopartite comporte une séquence linéaire de type KK/RxK et le NLS bipartite est constitué de 2 groupements basiques de 2 à 3 résidus basiques chacun et séparés par environ 10 acides aminés [14]. Il existe d'autres séquences faisant office de NLS qui sont différentes de ces séquences consensus et que l'on appelle NLS non conventionnel.

**Les exportines**

Les protéines responsables de l'export nucléaire sont appelées exportines ; elles interagissent avec le pore nucléaire. Comme dans le cas des importines, il existe différentes exportines qui reconnaissent différents cargos [15]. Il existe également parfois des molécules adaptatrices entre le cargo et l'exportine. CRM1 (*chromosome region maintenance protein 1*), aussi appelé exportin 1 [16, 17], est l'un des transporteurs très utilisés. L'agent antifongique issu de *Streptomyces*, la leptomycine B, est un inhibiteur spécifique et irréversible de CRM1 par liaison covalente à un



**Figure 3.** Processus d'import des vRNP et des protéines nouvellement synthétisées vers le noyau. Les transporteurs cellulaires sont représentés par les importines  $\alpha$  et  $\beta$ .

résidu de cystéine dans la région centrale de CRM1 [18]. Son utilisation permet de caractériser spécifiquement cette voie de transport. CRM1 appartient à la superfamille des importines  $\beta$  [19, 20] et représente le transporteur généralement utilisé pour l'export nucléaire de protéines, d'ARN et de ribonucléoprotéines. La liaison de RanGTP au complexe d'export CRM1-exportine-adaptateur-cargo autorise la translocation à travers le pore nucléaire (figure 4). Une fois dans le cytoplasme, ces complexes sont désassemblés grâce à l'hydrolyse du GTP de Ran par RanGAP (*Ran GTPase activating protein*) associé aux cofacteurs RanBP1 et RanBP2 (*Ran binding proteins*) localisés sur les fibrilles cytoplasmiques du pore. De ce fait, la concentration de Ran-GDP est prédominante dans le cytoplasme alors que Ran-GTP est la forme prédominante dans le noyau et c'est cette distribution asymétrique des formes GDP et GTP de Ran entre le cytoplasme et le noyau qui orchestre la direction du transport.

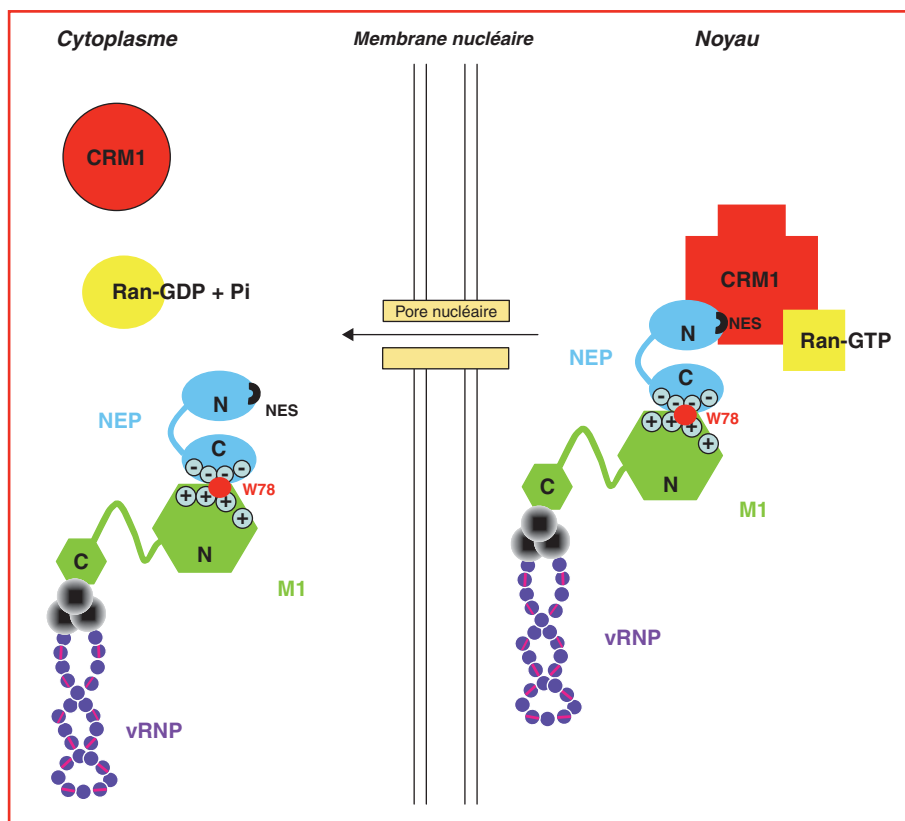
Le NES reconnu par CRM1 est une courte séquence peptidique riche en leucines. La séquence consensus est LxxxLxxLxL, bien que certains autres résidus hydrophobes puissent remplacer les leucines. Après reconnaissance

du motif NES, CRM1 se lie à Ran-GTP [19] et l'ensemble est exporté.

Pour l'export de certains ARN, CRM1 interagit avec des exportines-adaptateurs qui contiennent un motif NES [6]. Ainsi, CRM1 est impliqué dans l'export des snRNA dont le complexe CBP20-CBP80 liant leur coiffe en 5' est reconnu par le stabilisateur PHAX [21]. L'ARNm non épissé du VIH1 [22] suit également la voie d'export CRM1. Sur l'ARNm viral, la séquence nucléotidique appelée RRE (*rev responsive element*) se lie à la protéine virale rev qui contient à la fois un signal NLS et un signal NES. La séquence NES de rev se lie à CRM1, CRM1 se lie à Ran-GTP et l'ensemble est alors exporté au travers des NPC.

### L'import nucléaire des protéines et des vRNP du virus de la grippe

Pour infecter une cellule, le virus de la grippe se fixe sur des récepteurs à la surface de la cellule et entre par endocytose. L'acidification de l'endosome conduit à un changement



**Figure 4.** Processus d'export des vRNP en fin d'infection virale. Seule la protéine NEP permet l'export de ses partenaires. Le transporteur cellulaire est représenté par CRM1 associé à Ran-GTP lors de la formation du complexe ternaire.

dans la conformation de l'hémagglutinine qui va induire la fusion entre la membrane virale et la membrane endosomale. Alors que l'intérieur de l'endosome devient acide, l'intérieur du virion le devient aussi grâce aux canaux à protons formés par les tétramères de la protéine M2. Cette acidification provoque la dissociation des vRNP de la couche de protéine matricielle M1 qui tapisse la face interne de la membrane virale [23].

Une fois dans le cytoplasme, les 8 vRNP sont importées dans le noyau pour la production des ARN messagers du virus. Les RNP ont un diamètre de 15 nm et une longueur variable de 50 à 100 nm en fonction du segment d'ARN viral encapsidé [3]. Il n'a toujours pas été établi si les vRNP entraînent séparément ou groupées. Les vRNP contiennent l'information nécessaire pour leur import nucléaire. En effet, des vRNP injectées seules dans le cytoplasme de cellules de mammifère sont rapidement accumulées dans le noyau [24]. Pour l'instant, le motif directement responsable de la translocation nucléaire des vRNP n'a pas été défini. Une vRNP étant composée d'une molécule d'ARN simple brin complexée à des nucléoprotéines et à la polymérase, il existe donc plusieurs candidats potentiels portant les signaux d'import de la vRNP (*figure 2*). Ces protéines présentent toutes des séquences NLS car les protéines traduites dans le cytoplasme doivent également rejoindre le noyau pour l'assemblage des vRNP nouvellement produites. En effet, la nucléoprotéine possède deux motifs NLS et une séquence d'accumulation nucléaire (NAS) et chacune des trois sous-unités de la polymérase possède des signaux NLS (*figure 2*).

#### Les NLS des sous-unités de la polymérase

La sous-unité PB2 possède un motif NLS situé aux acides aminés 736-739 (K R K R) (*figure 2*). De plus, la délétion de la région comprenant les acides aminés 449 à 495 entraîne une distribution cytoplasmique périnucléaire de PB2, suggérant une étape de transit de la protéine avant sa translocation nucléaire [25]. Quand la sous-unité PB1 est exprimée en l'absence d'autres protéines virales, elle s'accumule dans le noyau grâce au signal présent entre les résidus 180 à 252 [26]. Néanmoins, pour être efficacement transloquée dans le noyau, PB1 requiert la coexpression de la sous-unité PA [27]. La sous-unité PA possède deux régions impliquées dans sa translocation nucléaire : la région située entre les acides aminés 124 à 139 et la région située entre les acides aminés 186 à 247 [28]. De plus, la mutation ponctuelle du résidu 154 abolit totalement le transport [28]. Même si ces trois sous-unités de la polymérase possèdent des signaux indépendants pour leur translocation nucléaire, il est possible que la formation du complexe ternaire PA-PB1-PB2 soit réalisée dans le cytoplasme avant le transport dans le noyau. La *figure 2* montre aussi les domaines d'interaction entre les sous-unités du complexe

de la polymérase et avec la NP ainsi que les domaines d'interaction avec l'ARN.

#### La nucléoprotéine

La nucléoprotéine isolée interagit *in vivo* [29, 30] et *in vitro* [31] avec plusieurs importines  $\alpha$  humaines ( $\alpha 1$ ,  $\alpha 3$  et  $\alpha 5$ ). Les nombreuses études menées sur les signaux de transport de nucléoprotéines ont mis en évidence la présence de plusieurs de ces séquences sur la nucléoprotéine : un motif NLS dit « non conventionnel » en N-terminal (acides aminés 1 à 13) [30, 32] et un NLS classique bipartite entre les acides aminés 198-216 [33]. Par ailleurs, des études utilisant des œufs de xénope ont mis en évidence la présence d'une séquence dite d'accumulation/rétention nucléaire (NAS/NRS) entre les acides aminés 336 à 345 qui serait à l'origine de l'accumulation de nucléoprotéine dans le noyau [34]. Le motif situé entre les acides aminés 1-13 serait crucial pour l'import de la nucléoprotéine libre [30, 35, 36]. De plus, le transport de la nucléoprotéine est régulé par la phosphorylation de la sérine présente en position 3 de la nucléoprotéine [32, 37]. En effet, la phosphorylation de la sérine 3 entraîne une rétention cytoplasmique de la protéine [32].

### L'export nucléaire des vRNP

#### Les protéines impliquées dans l'export des vRNP

Après la réplication de l'ARN viral et la reconstitution des vRNP, ces vRNP nouvellement synthétisées empruntent la voie d'export nucléaire vers le cytoplasme pour l'assemblage des nouveaux virions à la membrane cellulaire. Ce transport est spécifique de CRM1 car il est inhibé par la leptomycine B [38, 39]. En plus des protéines nécessaires pour l'assemblage des nouvelles RNP, deux autres protéines virales, M1 et NEP, se retrouvent localisées dans le noyau. M1 possède un motif NLS classique [40] (*figure 3*) et sa présence dans le noyau est nécessaire pour déclencher le processus d'export des vRNP [24, 41]. Différentes études ont montré que la protéine NEP est également impliquée dans le processus d'export nucléaire des vRNP. En effet, la micro-injection d'anticorps anti-NEP bloque l'export des vRNP [42] et la création d'un virus recombinant sans protéine NEP entraîne la rétention nucléaire des vRNP, leur export étant restauré par l'apport en *trans* de la protéine NEP [43]. NEP possède un motif NES [44] mis en évidence par des essais d'interaction directe avec CRM1 [43] et par des essais plus spécifiques d'hydrolyse de RanGTP en RanGDP lors de la formation du complexe ternaire [45]. Une séquence NLS n'a pas été montrée pour NEP et il est possible que la protéine entre dans le noyau par diffusion sans intervention d'importines (*figure 3*).

Le fait qu'elle soit visualisée par immunofluorescence dans les deux compartiments cellulaires [46, 32] et qu'elle soit

capable de se fixer à CRM1 *in vitro* [38] laisse penser que la nucléoprotéine possédait également un motif NES. Cependant, la formation *in vitro* du complexe NP-CRM1 en présence de RanGTP ne provoque aucune activité de l'hydrolyse de RanGTP en RanGDP, montrant que, si telle interaction existe, elle n'est pas spécifique [45].

### Régulation de l'export des vRNP

Depuis longtemps, la protéine M1 a été désignée comme l'élément déclencheur de l'export des vRNP [24, 47]. Cette protéine consiste en deux domaines : un domaine aminoterminal (résidus 1-164) qui contient le NLS, et un domaine C-terminal (résidus 165-252). Les deux domaines, bien repliés et riches en hélices  $\alpha$ , sont connectés par une liaison flexible [48, 49]. Le NLS est une région à la surface du domaine N-terminal possédant une forte charge positive [49] (figure 5). Par son domaine C-terminal (165-252), M1 fixe les vRNP [48] exposant ainsi son domaine N-terminal pour sa fixation consécutive à NEP (figure 4).

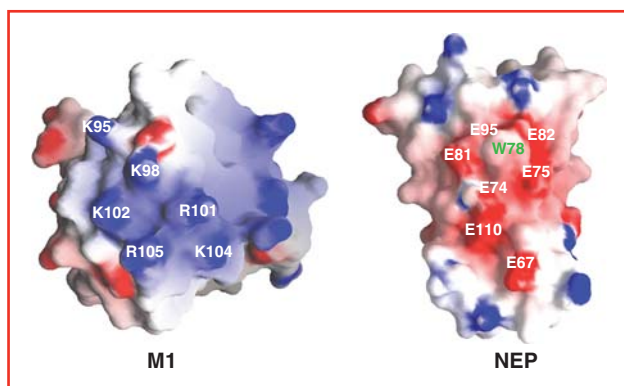
M1 et NP n'interagissent pas avec CRM1, contrairement au domaine (1-54) de NEP qui contient le NES [45]. D'après des résultats *in vitro*, M1 et NEP constituent l'interface entre les vRNP et le système d'export cellulaire CRM1-RanGTP. NEP se fixe sur le domaine N-terminal de la protéine M1, plus précisément sur le signal NLS situé autour des résidus 101 à 105. La protéine NEP engage son résidu tryptophane 78 (W78) pour sa liaison à la protéine M1. Cet acide aminé W78 de NEP est entouré de glutamates chargés négativement pouvant se lier au motif NLS [45] (figure 5). L'interaction entre NEP et M1 est abolie lorsque le résidu tryptophane 78 de la protéine NEP est muté en sérine ainsi que lorsque les résidus arginine et lysine du motif NLS sont mutés sur la protéine M1, les structures secondaires et tridimensionnelles de NEP et M1 restant par ailleurs conservées [45, 50]. Il est surprenant de constater

que cette interaction entre tryptophane et acides aminés basiques se rapproche de celle rencontrée entre protéine cargo à NLS et protéine d'import cellulaire importine  $\alpha$  [10]. Le complexe viral vRNP-M1-NEP est exporté du noyau en utilisant le système d'export cellulaire CRM1-RanGTP. La protéine directement responsable de la liaison à CRM1 est NEP [43, 45]. L'existence et la présentation du motif NES de NEP pourraient permettre une régulation de l'export des vRNP par phosphorylation ou par interaction avec un partenaire viral. Lorsque NEP est cytoplasmique, son domaine N-terminal flexible contenant le NES serait positionné contre le domaine C-terminal plus compact [45], rendant le NES inaccessible. Une fois NEP entrée dans le noyau, la liaison de son domaine C-terminal avec M1 pourrait exposer le motif NES du domaine N-terminal, rendant le motif NES accessible pour l'interaction avec CRM1.

La protéine M1 pourrait également représenter l'élément discriminatoire des deux espèces ribonucléoprotéiques existant dans le noyau : les cRNP et les vRNP. En se fixant uniquement aux vRNP, M1 permettrait de désigner les vRNP comme substrat de sortie, les intermédiaires répliqués cRNP restant dans le noyau. Des expériences de fixation de M1 aux cRNP seraient nécessaires pour valider cette hypothèse. De plus, en se fixant aux vRNP, la protéine M1 pourrait ainsi éviter le réimport intempestif des vRNP exportées dans le cytoplasme pour leur incorporation au complexe d'assemblage viral au niveau de la membrane plasmique.

### Conclusion

Nous avons décrit les processus de l'import nucléaire des protéines, la formation des vRNP, l'association de NEP et M1 sur ce complexe et l'export définitif et irréversible des vRNP : les vRNP transportées vers le cytoplasme ne peuvent pas être réimportées dans le noyau quand M1 est fixée. Elles peuvent seulement entrer dans le noyau quand elles sont dépourvues de M1 et le seul moment où M1 n'est pas associée avec les RNP est juste avant l'entrée des RNP dans le cytoplasme au sein de la cellule nouvellement infectée. L'acidification de l'intérieur du virus quand il réside dans l'endosome est la condition *sine qua non* qui autorise la libération de M1 des RNP, les rendant aptes à leur import nucléaire. Ainsi, c'est la protéine M1 qui joue le rôle de chef d'orchestre pour la localisation des RNP. L'étape d'acidification de l'intérieur du virus est l'un des talons d'Achille du virus et bloquer le canal de protons dans la membrane virale constitue une bonne cible antivirale. En fait, ces médicaments ont été développés depuis longtemps (amantadine et rimantidine), mais le virus échappe facilement à ces agents bloquant l'infection par mutation de sa protéine M2, constituant du canal à protons. L'autre cible



**Figure 5.** Surfaces électrostatiques des deux protéines M1 et NEP. À noter le motif NLS présent sur M1 et le résidu W78 sur NEP. Dans le complexe M1-NEP, la surface basique (bleue) de M1 interagit avec la surface acide (rouge) de NEP.



antivirale pourrait être les étapes d'import et d'export en inhibant la fixation des protéines virales aux protéines partenaires de l'hôte.

**Remerciements.** Le travail décrit dans cette publication a été financé en partie par le programme européen Virgil (*European Vigilance Network for the Management of Antiviral Drug Resistance*).

## Références

- Horimoto T, Kawaoka Y. Influenza : lessons from past pandemics, warnings from current incidents. *Nat Rev Microbiol* 2005 ; 8 : 591-600.
- Compans RW, Content J, Duesberg PH. Structure of the ribonucleoprotein of influenza virus. *J Virol* 1972 ; 4 : 795-800.
- Elton D, Medcalf L, Bishop K, Harrison D, Digard P. Identification of amino acid residues of influenza virus nucleoprotein essential for RNA binding. *J Virol* 1999 ; 73 : 7357-67.
- 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 : 1009-14.
- Sitterlin D. Quand la virologie dévoile les secrets de l'export des ARN. *Virologie* 2005 ; 9 : 191-200.
- Schwartz TU. Modularity within the architecture of the nuclear pore complex. *Curr Opin Struct Biol* 2005 ; 15 : 221-6.
- Imamoto N, Shimamoto T, Takao T, et al. *In vivo* evidence for involvement of a 58 kDa component of nuclear pore-targeting complex in nuclear protein import. *EMBO J* 1995 ; 14 : 3617-26.
- Cingolani G, Petosa C, Weis K, Müller CW. Structure of importin beta bound to the IBB domain of importin- $\alpha$ . *Nature* 1999 ; 399 : 221-9.
- 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 : 193-204.
- Weis K, Mattaj IW, Lamond AI. Identification of hSRP1  $\alpha$  as a functional receptor for nuclear localization sequences. *Science* 1995 ; 268 : 1049-53.
- Bayliss R, Littlewood T, Stewart M. Structural basis for the interaction between FxFG nucleoporin repeats and importin- $\beta$  in nuclear trafficking. *Cell* 2000 ; 102 : 99-108.
- Miyamoto Y, Imamoto N, Sekimoto T, et al. Differential modes of nuclear localization signal (NLS) recognition by three distinct classes of NLS receptors. *J Biol Chem* 1997 ; 272 : 26375-81.
- Conti E, Muller CW, Stewart M. Karyopherin flexibility in nucleocytoplasmic transport. *Curr Opin Struct Biol* 2006 ; 16 : 1-8.
- Hodel MR, Corbett AH, Hodel AE. Dissection of a nuclear localization signal. *J Biol Chem* 2001 ; 276 : 1317-25.
- Fukuda M, Asano S, Nakamura T, et al. CRM1 is responsible for intracellular transport mediated by the nuclear export signal. *Nature* 1997 ; 390 : 308-11.
- Weis K. Nucleocytoplasmic transport : cargo trafficking across the border. *Curr Opin Cell Biol* 2002 ; 3 : 328-35.
- Ossareh-Nazari B, Bachelier F, Dargemont C. Evidence for a role of CRM1 in signal-mediated nuclear protein export. *Science* 1997 ; 278 : 141-4.
- Kudo N, Matsumori N, Taoka H, et al. Leptomycin B inactivates CRM1/exportin 1 by covalent modification at a cysteine residue in the central conserved region. *Proc Natl Acad Sci USA* 1999 ; 96 : 9112-7.
- Askjaer P, Bachi A, Wilm M, et al. RanGTP-regulated interactions of CRM1 with nucleoporins and a shuttling DEAD-box helicase. *Mol Cell Biol* 1999 ; 19 : 6276-85.
- Petosa C, Schoehn G, Askjaer P, et al. Architecture of CRM1/Exportin1 suggests how cooperativity is achieved during formation of a nuclear export complex. *Mol Cell* 2004 ; 16 : 761-75.
- Ohno M, Segref A, Bachi A, Wilm M, Mattaj IW. PHAX, a mediator of U snRNA nuclear export whose activity is regulated by phosphorylation. *Cell* 2000 ; 101 : 187-98.
- Fischer U, Huber J, Boelens WC, Mattaj IW, Lührmann R. The HIV-1 Rev activation domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs. *Cell* 1995 ; 82 : 475-83.
- Martin K, Helenius A. Nuclear transport of influenza virus ribonucleoproteins : the viral matrix protein (M1) promotes export and inhibits import. *Cell* 1991 ; 67 : 117-30.
- Kemler I, Whittaker G, Helenius A. Nuclear import of microinjected influenza virus ribonucleoproteins. *Virology* 1994 ; 202 : 1028-33.
- Mukaigawa J, Nayak DP. Two signals mediate nuclear localization of influenza virus (A/WSN/33) polymerase basic protein 2. *J Virol* 1991 ; 65 : 245-53.
- Nath S, Nayak DP. Function of two discrete regions is required for nuclear localization of polymerase basic protein 1 of A/WSN/33 influenza virus (H1N1). *Mol Cell Biol* 1990 ; 10 : 4139-45.
- 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 : 9144-53.
- Nieto A, de la Luna S, Barcena J, Portela A, Ortin J. Complex structure of the nuclear translocation signal of influenza virus polymerase PA subunit. *J Gen Virol* 1994 ; 75 : 29-36.
- O'Neill RE, Palese P. NPI-1, the human homolog of SRP-1, interacts with influenza virus nucleoprotein. *Virology* 1995 ; 206 : 116-25.
- 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 : 1850-6.
- 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 : 28193-200.
- Neumann G, Castrucci MR, Kawaoka Y. Nuclear import and export of influenza virus nucleoprotein. *J Virol* 1997 ; 71 : 9690-700.
- 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.
- Davey J, Dimmock NJ, Colman A. Identification of the sequence responsible for the nuclear accumulation of the influenza virus nucleoprotein in *Xenopus* oocytes. *Cell* 1985 ; 40 : 667-75.
- Bullido R, Gomez-Puertas P, Albo C, Portela A. Several protein regions contribute to determine the nuclear and cytoplasmic localization of the influenza A virus nucleoprotein. *J Gen Virol* 2000 ; 81 : 135-42.
- Cros JF, Garcia-Sastre A, Palese P. An unconventional NLS is critical for the nuclear import of the influenza A virus nucleoprotein and ribonucleoprotein. *Traffic* 2005 ; 6 : 205-13.
- Arrese M, Portela A. Serine 3 is critical for phosphorylation at the N-terminal end of the nucleoprotein of influenza virus A/Victoria/3/75. *J Virol* 1996 ; 70 : 3385-91.
- 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 : 408-19.
- Watanabe K, Takizawa N, Katoh M, Hoshida K, Kobayashi N, Nagata K. Inhibition of nuclear export of ribonucleoprotein complexes of influenza virus by leptomycin B. *Virus Res* 2001 ; 77 : 31-42.
- Ye Z, Robinson D, Wagner RR. Nucleus-targeting domain of the matrix protein (M1) of influenza virus. *J Virol* 1995 ; 69 : 1964-70.
- Sakaguchi A, Hirayama E, Hiraki A, Ishida Y, Kim J. Nuclear export of influenza viral ribonucleoprotein is temperature-dependently inhibited by dissociation of viral matrix protein. *Virology* 2003 ; 306 : 244-53.
- Yasuda J, Nakada S, Kato A, Toyoda T, Ishihama A. Molecular assembly of influenza virus : association of the NEP protein with virion matrix. *Virology* 1993 ; 196 : 249-55.
- Neumann G, Hughes MT, Kawaoka Y. Influenza A virus NEP protein mediates vRNP nuclear export through NES-independent interaction with hCRM1. *EMBO J* 2000 ; 19 : 6751-86.

44. O'Neill RE, Talon J, Palese P. The influenza virus NEP (NS2 protein) mediates the nuclear export of viral ribonucleoproteins. *EMBO J* 1998 ; 17 : 288-96.
45. 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/NEP). *EMBO J* 2003 ; 22 : 4646-55.
46. Whittaker G, Bui M, Helenius A. Nuclear trafficking of influenza virus ribonucleoproteins in heterokaryons. *J Virol* 1996 ; 70 : 2743-56.
47. Bui M, Wills EG, Helenius A, Whittaker GR. Role of the influenza virus M1 protein in nuclear export of viral ribonucleoproteins. *J Virol* 2000 ; 74 : 1781-6.
48. Baudin F, Petit I, Weissenhorn W, Ruigrok RW. In vitro dissection of the membrane and RNP binding activities of influenza virus M1 protein. *Virology* 2001 ; 281 : 102-8.
49. Arzt S, Baudin F, Barge A, Timmins P, Burmeister WP, Ruigrok RW. Combined results from solution studies on intact influenza virus M1 protein and from a new crystal form of its N-terminal domain show that M1 is an elongated monomer. *Virology* 2001 ; 279 : 439-46.
50. Arzt S, Petit I, Burmeister WP, Ruigrok RW, Baudin F. Structure of a knockout mutant of influenza virus M1 protein that has altered activities in membrane binding, oligomerisation and binding to NEP (NEP). *Virus Res* 2004 ; 99 : 115-9.
51. Poch O, Sauvaget I, Delarue M, Tordo N. Identification of four conserved motifs among the RNA-dependent polymerase encoding elements. *EMBO J* 1989 ; 8 : 3867-74.
52. Shinya K, Hamm S, Hatta M, Ito H, Ito T, Kawaoka Y. PB2 amino acid at position 627 affects replicative efficiency, but not cell tropism, of Hong Kong H5N1 influenza A viruses in mice. *Virology* 2004 ; 320 : 258-66.

**Nuclear traffic of influenza virus proteins and  
ribonucleoprotein complexes.**

Boulo, S., Akarsu, H., Ruigrok, R.W. and Baudin, F.

Virus Res. 2007 Mar;124(1-2):12-21. Epub 2006 Nov 1.

Review

# Nuclear traffic of influenza virus proteins and ribonucleoprotein complexes

Sébastien Boulo, Hatice Akarsu, Rob W.H. Ruigrok, Florence Baudin\*

*Institut de Virologie Moléculaire et Structurale, FRE 2854 CNRS-UJF, BP 181, 38042 Grenoble cedex 9, France*

Received 25 July 2006; received in revised form 29 September 2006; accepted 30 September 2006

Available online 1 November 2006

## Abstract

Influenza virus is a negative strand RNA virus and is one of the rare RNA viruses to replicate in the nucleus. The viral RNA is associated with 4 viral proteins to form ribonucleoprotein particles (RNPs). After cell entry the RNPs are dissociated from the viral matrix protein in the low pH of the endosome and are actively imported into the cell nucleus. After translation of viral mRNAs, the proteins necessary for the assembly of new RNPs (the nucleoprotein and the three subunits of the polymerase complex) are also imported into the nucleus. Apart from these four proteins, part of the newly made matrix protein is also imported and the nuclear export protein (NEP) enters the nucleus probably through diffusion. Finally, NS1 also enters the nucleus in order to regulate a number of nuclear processes. The nuclear localization signals on all these viral proteins and their interaction with the cellular transport system are discussed. In the nucleus, the matrix protein binds to the newly assembled RNPs and NEP then binds to the matrix protein. NEP contains the nuclear export signal necessary for transport of the RNPs to the cytoplasm, necessary for the budding of new virus particles. There appears to be an intricate ballet in exposing and hiding nuclear transport signals which leads to a unidirectional transport of the RNPs to the nucleus at the start of the infection process and an opposite unidirectional export of RNPs at the end of the infection. © 2006 Elsevier B.V. All rights reserved.

**Keywords:** Influenza virus; Nuclear transport; Nucleoprotein; Polymerase; NEP; M1

## Contents

1. Introduction .....	12
2. Nuclear import and export .....	13
2.1. The nuclear pore complex .....	13
2.2. Importins .....	14
2.3. Exportins .....	14
3. Nuclear import of influenza virus RNPs and of newly translated viral proteins .....	15
3.1. Incoming RNPs .....	15
3.2. Nuclear import of newly made viral proteins .....	16
4. Nuclear export of newly assembled RNPs .....	17
5. Multiple roles for the M1 protein .....	18
Acknowledgements .....	19
References .....	19

## 1. Introduction

Negative strand RNA viruses are enveloped viruses that cause a large variety of human diseases such as measles, bronchiolitis, mumps (viruses of the family of the *Paramyxoviridae*), rabies (*Rhabdoviridae*) and influenza (*Orthomyxoviridae*).

\* Corresponding author. Tel.: +33 476207287; fax: +33 476209400.  
E-mail address: [baudin@embl.fr](mailto:baudin@embl.fr) (F. Baudin).

Other viruses that have animal hosts can also infect humans and cause hemorrhagic fevers such as *Filoviridae*, *Arenaviridae* and *Bunyaviridae*. Influenza viruses are viruses of aquatic birds that can adapt to other animal hosts by sequential accumulation of mutations or by reassortment between a bird virus and a virus infecting another type of animal when the same cell is simultaneously infected by two viruses. Both types of adaptation can lead to new human virus strains that can be extremely virulent and lethal. This virulence can be related to an incompatibility between the host innate immune system leading to a large overproduction of “toxic” cytokines and to the fact that these avian type viruses may replicate in the lower respiratory tract rather than in the upper tract (Van Riel et al., 2006; Shinya et al., 2006).

The influenza viruses, like the *Arena*- and *Bunyaviridae*, have a segmented genome. The influenza A virus genome is split up into 8 RNA segments, varying in length between 890 and 2341 nucleotides. The RNA fragments are all covered by the viral nucleoprotein (NP). Each copy of this protein of 56 kDa binds to about 24 nucleotides of viral RNA (Compans et al., 1972). NP contains domains that interact with the viral RNA and other domains are involved in NP–NP oligomerisation (Elton et al., 1999a,b; Ruigrok and Baudin, 1995) (see Figs. 1 and 2). The trimeric polymerase complex (PB1, PB2 and PA) binds to both ends of the viral RNA forming a non-covalent circular complex (Fig. 1) (Klumpp et al., 1997). The complex consisting of viral RNA, the polymerase complex and the stoichiometrically bound NP is called the ribonucleoprotein (RNP) complex. The 8 RNPs represent independent units for transcription and replication inside the host cell.

Most DNA viruses replicate in the nucleus of the infected cell (with some exceptions like pox viruses) where they use the cellular RNA polymerase II to make mRNAs that can be translated in the cytoplasm. Since RNA viruses cannot use cellular Pol II (because it is dependent on a DNA matrix) most RNA viruses replicate in the cytoplasm. Influenza virus is one of the very few RNA viruses to replicate in the nucleus. This is probably due to the fact that at least two of the genes have full-length and spliced mRNAs leading to different proteins: segment 7 codes for M1 and M2 and segment 8 for NS1 and NEP (previously NS2). For that purpose, splicing of viral mRNAs probably uses components of the cellular splicing machinery present in the nucleus. The nuclear replication of influenza virus implies

extensive nuclear transport in both directions. In this review we will detail what is known about the various nuclear transport signals carried by the viral proteins and RNP. These signals are exposed or hidden at different moments during the viral cycle.

## 2. Nuclear import and export

Proteins or protein–RNA complexes are imported into the nucleus or exported from the nucleus to the cytoplasm due to the exposition of specific signals (presence of a specific constellation of amino acid residues at the surface of the proteins). For nuclear import a protein needs to expose a Nuclear Localization Signal (NLS) and for export a protein needs to have a Nuclear Export Signal (NES). These signals allow specific interactions with cellular nuclear transport complexes that interact with the nuclear pore complex.

### 2.1. The nuclear pore complex

Molecules can only get into or out of the nucleus through the nuclear pore complex (NPC). There are several hundreds of these complexes per nucleus but their exact number depends on the cell type and also on the position in the cell cycle. Each pore is a very large molecular complex (around 125 MDa in mammalian cells) containing some hundred different proteins called nucleoporins (Schwartz, 2005). The NPC consists of a cylindrical structure with an 8-fold symmetry and a central diameter of about 10 nm. It consists of a central canal spanning the double nuclear membrane with fibrils on each side. At the nuclear side of the pore the fibrils end with a ring structure and form what is called the nuclear basket. Ions and proteins with molecular weights of less than about 40 kDa can freely diffuse through the NPC and go into and out of the nucleus without energy consumption. However, transport of larger proteins can only be accomplished through the interaction of transporter molecules and is an active process. The diameter of the NPC can increase to some 30 nm in order to let large complexes through. Some of the nucleoporins have sequence repeats containing phenylalanine (F) and glycine (G) residues. These repeats are specifically recognized by proteins of the importin  $\beta$  family (see below). Once transport has been completed, the cargo is released and the transporters are recycled.

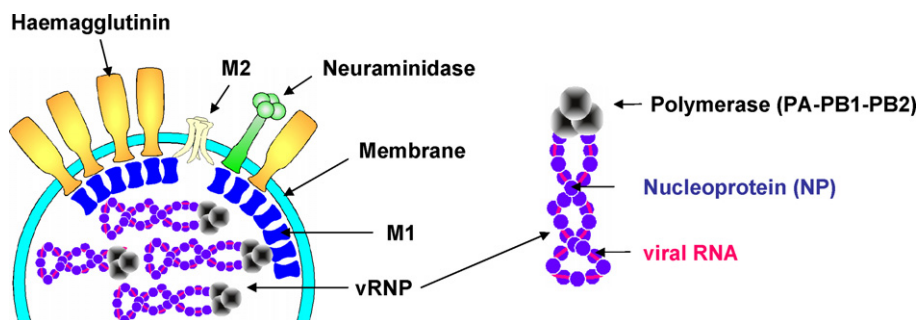


Fig. 1. Schematic representation of an influenza virion and a ribonucleoprotein particle (RNP) containing the viral RNA, nucleoprotein and the heterotrimeric polymerase complex.

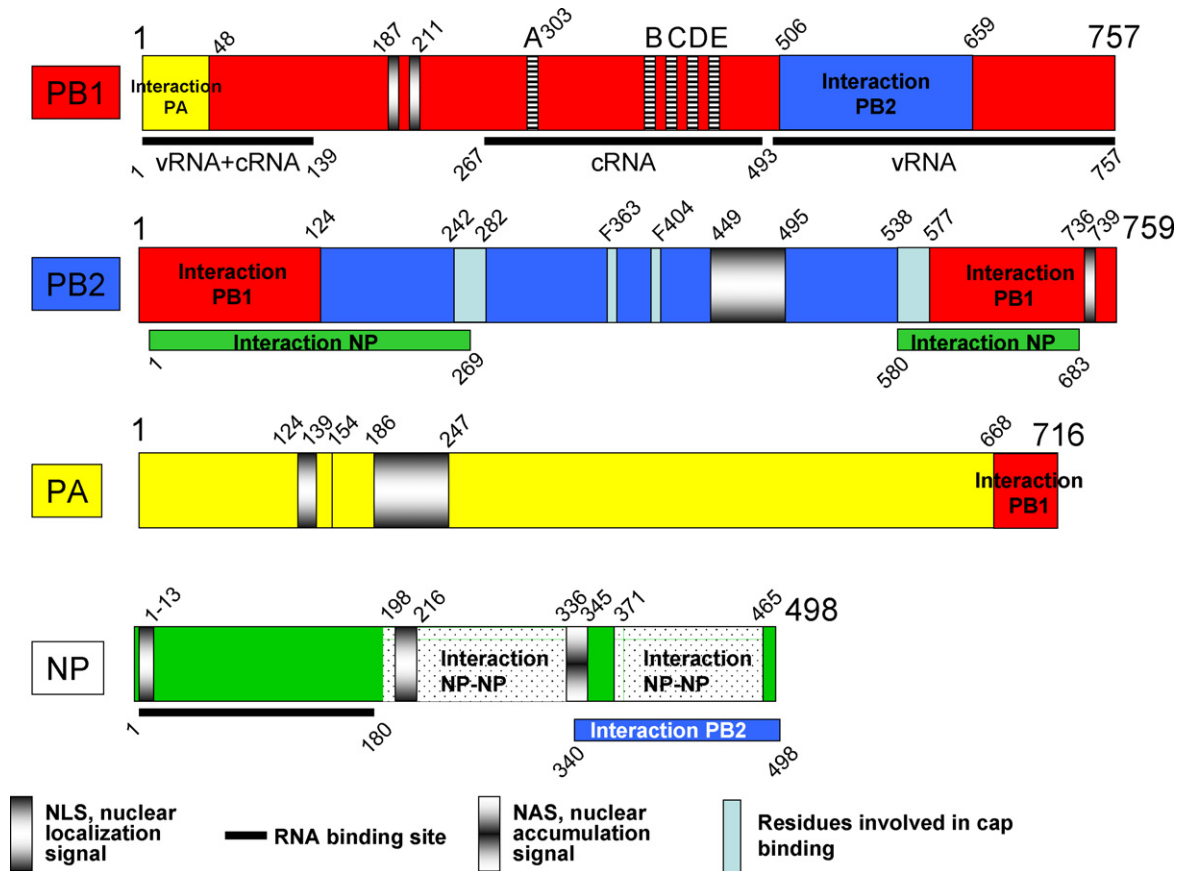


Fig. 2. Protein-protein and protein–RNA interactions for the polymerase subunits and nucleoprotein. The figure indicates the interactions between the different proteins in a color code. The information for these interactions was assembled from Perez and Donis (1995), Gonzalez et al. (1996), Biswas and Nayak (1996), Ohtsu et al. (2002), Toyoda et al. (1996), Poole et al. (2004), Kawaguchi et al. (2005), Elton et al. (1999b), Biswas et al. (1998). The figure also indicates the domains of PB1 and NP that interact with RNA (Albo et al., 1995; Gonzalez and Ortin, 1999; Kobayashi et al., 1994; Elton et al., 1999a) and the domains or residues on PB2 that interact with the cap (Honda et al., 1999; Fechter et al., 2003; Li et al., 2001). The domains A–E are the conserved sequences found in all RNA polymerases (Poch et al., 1989; Biswas and Nayak, 1994). Finally, all known nuclear localization and accumulation signals are indicated, the references for these can be found in Table 1.

## 2.2. Importins

Molecules that have to enter the nucleus are carried in by importins (also called karyopherins) (Imamoto et al., 1995). There are two types of importins,  $\alpha$  and  $\beta$ , both of which are part of the importin  $\beta$  super-family. The proteins belonging to this family are characterized by a structure made up of a series of armadillo repeats (Cingolani et al., 1999; Conti et al., 1998). Importin  $\alpha$  proteins recognize the NLS on the cargo proteins (Weis et al., 1995). Once the cargo is bound by importin  $\alpha$ , the complex is recognized and bound by importin  $\beta$  that subsequently binds to the fibrils of the NPC and is responsible for the actual translocation (Fig. 3) (Bayliss et al., 2000). The schematic drawing in Fig. 3 is not the only way molecules can be transported into the nucleus (Miyamoto et al., 1997) but it is the most common way and the one that the influenza virus proteins take. There are several importin  $\alpha$  isoforms that recognize different variants of NLSs but they are all recognised by a single importin  $\beta$ . Therefore, these isoforms could be seen as adaptors for different signals to the importin  $\beta$  molecule (Conti et al., 2006). NLS motifs consist of grouped positively charged amino acids (Arg-R and Lys-K) and can be mono- or bipartite. Briefly, monopartite

signals have the sequence KK/RxK and bipartite signals consist of two groups of 2–3 basic amino acids separated by about 10 residues (Hodel et al., 2001). It is also possible that the sequence is not linear, in which case the basic residues may form a conformational epitope on the surface of the protein.

## 2.3. Exportins

The proteins that take cargo out of the nucleus are called exportins. Like for the importins, there exist exportins that directly recognize the various nuclear export signals and this class of adaptor molecules binds to the second exportin molecule that interacts with the nuclear pore. Exportin1 or Chromosome Region Maintenance protein 1 (CRM1) mediates the export from the nucleus of numerous proteins carrying a nuclear export signal (NES) (Fukuda et al., 1997; Ossareh-Nazari et al., 1997). The NES recognized by CRM1 is a short peptide sequence rich in leucines. The consensus sequence is LxxxLxxLxL but other hydrophobic residues can replace the leucines. CRM1 is the regular exportin for proteins, RNAs and ribonucleoproteins complexes (Askjaer et al., 1999; Petosa et al., 2004). It also contains armadillo repeats and belongs to the importin  $\beta$  super-

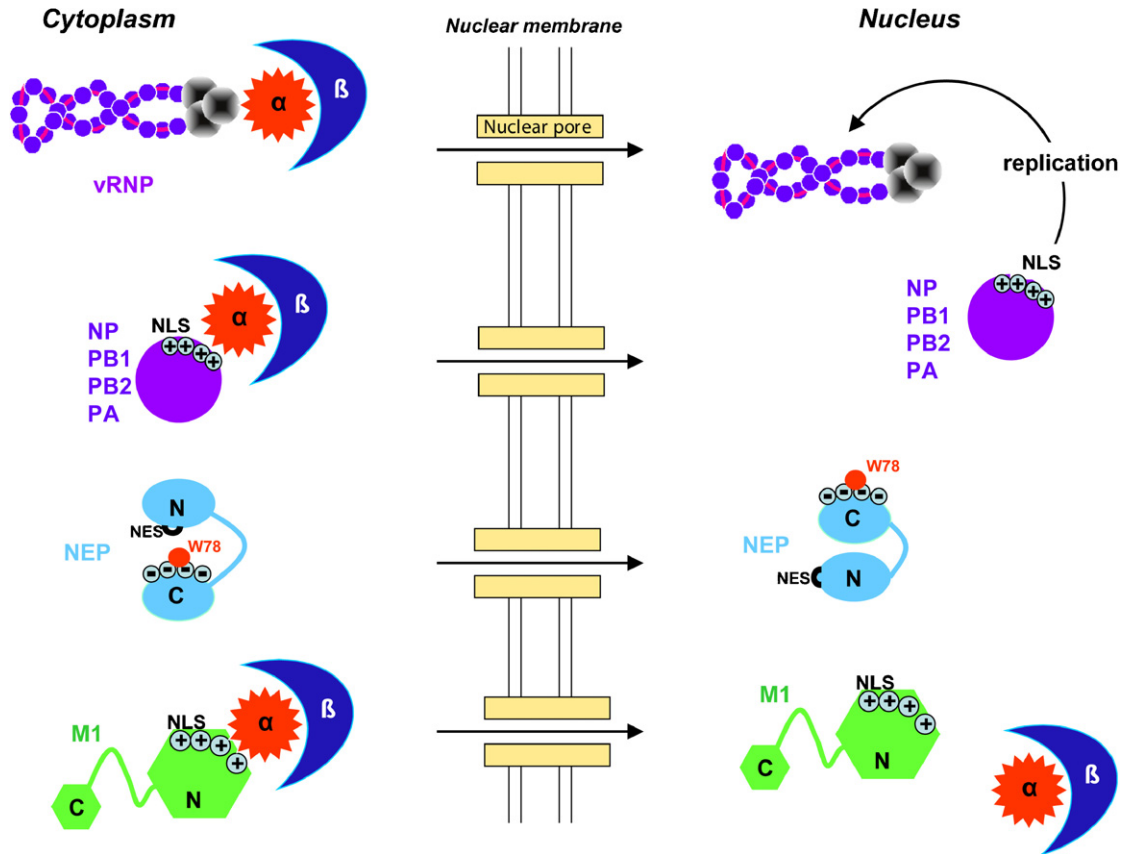


Fig. 3. Schematic representation of the nuclear import of influenza virus proteins. The nuclear import of NS1 is not represented here.

family. The anti-mould compound secreted by *Streptomyces*, Leptomycin B, is a specific and irreversible inhibitor of CRM1 and binds covalently to a cysteine residue in a central domain of CRM1 (Kudo et al., 1999). This substance is used to identify nuclear export pathways that involve CRM1 since not all nuclear export is mediated by CRM1, alternative pathways exist (Weis, 2002).

CRM1 cannot leave the nucleus with its cargo by itself but needs to be associated with Ran (a small GTPase) that has to be associated with GTP (Askjaer et al., 1999) (Fig. 4). As soon as the complex of cargo/adaptor-exportin/CRM1/RanGTP has left the nucleus, GTP is hydrolysed by Ran through interaction with RanGAP (Ran GTPase Activating Protein) which is associated with cofactors RanBP1 and RanBP2 (Ran Binding Proteins) that are localized on the fibrils on the cytoplasmic side of the NPC. After hydrolysis of GTP, the conformation of Ran changes and the complex falls apart and the exportins and RanGDP go back to the nucleus but the cargo remains in the cytoplasm.

For the export of certain RNAs, CRM1 interacts with proteins that recognise those RNA molecules and that expose their NES upon binding of the RNA. The snRNAs involved in splicing need to be transported to the cytoplasm in order to form snRNP complexes. These snRNPs contain a regular cap at their 5'-end which is bound by the nuclear cap-binding complex (CBC). The bound CBC is recognized by PHAX that exposes its NES upon binding to CBC and PHAX then associates with CRM1 (Ohno et al., 2000). One of the most famous examples is the way that non-

spliced HIV is transported to the cytoplasm (Fischer et al., 1995). This RNA contains a sequence motif called Rev Response Element (RRE) which binds to Rev that contains the NES capable of binding to CRM1.

### 3. Nuclear import of influenza virus RNPs and of newly translated viral proteins

#### 3.1. Incoming RNPs

When influenza virus infects a cell, the virion binds to receptors on the surface of the cell and is internalised through receptor-mediated endocytosis. The acidification of the endosome leads to two essential steps necessary for cell entry. First, protonation of the haemagglutinin leads to a conformational change that catalyses the fusion between the viral and the endosomal membrane leading to entry of the viral genome into the cytoplasm (Skehel and Wiley, 2000). Second, before membrane fusion takes place, the interior of the virus particle is also acidified by the action of a proton channel in the viral membrane made up by a tetramer of M2 (Hay et al., 1985; Pinto et al., 1992). This acidification is necessary for the depolymerisation of the M1 protein layer underneath the membrane and for the disassociation of M1 from the RNPs (Martin and Helenius, 1991; Zhirnov, 1992).

After entry into the cytoplasm, the RNPs are imported into the nucleus for the production of viral messenger RNAs. The

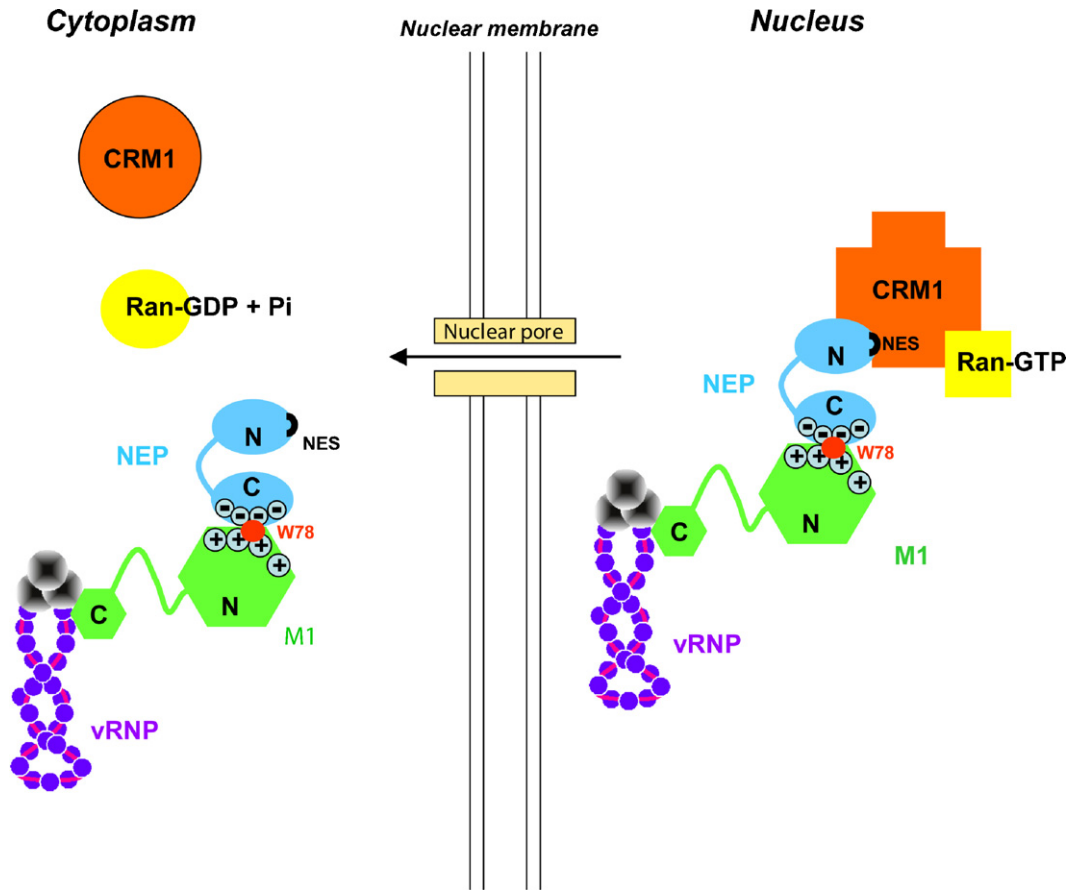


Fig. 4. Schematic representation of the nuclear export of newly assembled influenza virus RNPs. The different shapes of CRM1 and Ran in the nucleus compared to those in the cytoplasm indicate the conformational changes of these proteins upon hydrolysis of GTP and subsequent dissociation of the complex from its cargo.

RNPs have a diameter of about 15 nm and a length between 50 and 100 nm depending on the length of the RNA segment (Compans et al., 1972). It is not known whether the RNPs are transported as a group or whether each complex enters the nucleus independently through different pores. The RNPs contain the information to enter the nucleus and when purified RNPs are injected into the cytoplasm they rapidly accumulate in the nucleus (Kemler et al., 1994). However, it is not known which NLS is used for this import. All individual protein components of the RNP (NP and the three polymerase subunits) all have individual NLSs (see below) and we do not know which of these signals in the assembled RNP are used for import.

### 3.2. Nuclear import of newly made viral proteins

The proteins that make up the RNPs together with the viral RNA are translated from mRNAs in the cytoplasm and imported into the nucleus. The positions of the NLS signals on these proteins are indicated in Fig. 2 and Table 1. This figure summarises all information that is known today on these molecules, NLS, functional domains, and domains involved in the interaction between these proteins and with the RNA.

PB2 contains a linear NLS between residues 736 and 739 (Table 1, Fig. 2). Apart from this NLS, deletion of amino acids 449–495 leads to a perinuclear distribution of the protein with-

out actual nuclear import (Mukaigawa and Nayak, 1991). When PB1 is expressed alone it is imported into the nucleus due to signals present between residues 187 and 211 (Fig. 2) (Nath and Nayak, 1990). However, co-expression of PA is needed for efficient import of PB1 (Fodor and Smith, 2004). PA contains two NLSs, the first located between residues 124 and 139 and the second between 186 and 247 (Nieto et al., 1994). Apart from these, a point mutation of residue 154 totally prevents nuclear import (Nieto et al., 1994). Even though each polymerase subunit contains its own nuclear localization signals, it is possible that trimerization of the polymerase occurs in the cytoplasm before nuclear import. The transport of monomers, PB1-PA dimers or trimers could be part of a mechanism regulating different phases of the infection (Deng et al., 2005).

Nucleoprotein binds to a number of human importins  $\alpha$ , ( $\alpha$ 1,  $\alpha$ 3 and  $\alpha$ 5) both in vivo and in vitro (O'Neill and Palese, 1995; Wang et al., 1997; Melen et al., 2003). There have been quite a number of studies on the nuclear import signals of nucleoprotein and at least two NLSs have been identified; a non-conventional NLS at the very N-terminal 13 residues of the molecule (O'Neill and Palese, 1995; Neumann et al., 1997) and a classical bipartite signal between residues 198 and 216 (Fig. 2) (Weber et al., 1998). The N-terminal NLS is supposed to be important for the import of newly made protein (Wang et al., 1997; Bullido et al., 2000; Cros et al., 2005) which is inhibited by phosphorylation on Ser-3



Table 1  
List of nuclear localization, export and accumulation signals, in influenza virus proteins

	References
Nuclear localization signals	
NP	
1MASQGTRKRSYxxM <sub>13</sub>	Wang et al. (1997), Neumann et al. (1997)
198KR13xRKTR <sub>216</sub>	Weber et al. (1998)
PB1	
187KRRRVRDNMTKKMVTQRTIG- K <sub>13</sub> QR <sub>211</sub>	Nath and Nayak (1990)
PB2	
449–495	Mukaigawa and Nayak (1991)
736KRRR <sub>739</sub>	
PA	
124RREVHIYYLEKANKIK <sub>139</sub> E154 to Gly: abolishes nuclear import 186–247	Nieto et al. (1994)
NS1	
34RLRR <sub>38</sub> 216PKQKRK <sub>221</sub>	Greenspan et al. (1998)
M1	
95KAVKLYRKLKR <sub>105</sub>	Ye et al. (1995)
NEP	
Not known if NLS exists	
Nuclear accumulation signal	
NP	
336AAFEDLRVL <sub>345</sub>	Davey et al. (1985), Weber et al. (1998)
Nuclear export signals	
NS1	
138FDRLETLL <sub>147</sub>	Li et al. (1998)
NEP	
11DILLRMSKMQLES <sub>23</sub>	O'Neill et al. (1998)

(Neumann et al., 1997; Arrese and Portela, 1996). Studies on the transport of NP in *Xenopus* eggs have shown the presence of yet another type of signal; a nuclear accumulation/retention signal between residues 336 and 345 (Davey et al., 1985) (Fig. 2), necessary for the retention of NP in the nucleus until it binds to viral RNA and is exported from the nucleus as newly made RNP (see below).

There are two other viral proteins that go to the nucleus and that are important for the nuclear export of newly made RNPs. First, most of the viral matrix protein remains in the cytoplasm where it binds to lipids and the cytoplasmic tails of the viral glycoproteins while they are transported to the cellular membrane where virus budding takes place (Barman et al., 2001; Schmitt and Lamb, 2005). However, many years ago it was recognised that a second population of M1 did not bind to lipids in the cytoplasm but was transported to the nucleus (Hay and Skehel, 1975). Nuclear presence of M1 is necessary for export of the RNPs (Kemler et al., 1994; Sakaguchi et al., 2003). Second, another protein is also needed inside the nucleus for export of RNPs; Nuclear Export Protein (NEP). This protein, formerly called NS2, is a splicing product from the smallest gene segment of the virus that also codes for an unspliced product, NS1. A

specific antibody against NEP blocks nuclear export of the RNPs (Yasuda et al., 1993) and a recombinant virus that lacks NEP shows nuclear retention of the RNPs that can be overcome by providing NEP in “trans” (Neumann et al., 2000). NEP does not seem to have an NLS but, because of its size (14.5 kDa), it is possible that it diffuses freely through the nuclear pore. Finally, the last protein that is imported into the nucleus but that is not involved in the export of RNPs is NS1 (Table 1). This 26 kDa protein antagonizes the cellular immune response during the viral life cycle and is essential for a viable infection (Min and Krug, 2006).

#### 4. Nuclear export of newly assembled RNPs

After replication of the RNA and assembly of the new RNPs, these complexes have to leave the nucleus in order to be incorporated into new virus particles that bud at the cell membrane. Nuclear export of RNPs depends on CRM1 because it can be inhibited by leptomycin B (Elton et al., 2001; Watanabe et al., 2001). The protein elements of RNPs do not seem to have nuclear export sequences. Although it has been reported that NP can bind to CRM1 in vitro (Elton et al., 2001), the complex NP/CRM1 does not induce the hydrolysis of GTP by Ran in the presence of RanGAP, suggesting that if an interaction of NP and CRM1 exists, this is not of the kind that can lead to nuclear export (Akarsu et al., 2003). As mentioned above, for export of RNPs, both viral M1 and NEP proteins are needed (Neumann et al., 2000; O'Neill et al., 1998; Bui et al., 2000). M1 is a two-domain protein. The N-terminal domain (residues 1–164) contains the NLS (see Table 1) and the C-terminal domain (165–252) has affinity for RNP (Baudin et al., 2001). However, M1 does not have the necessary NES motif for nuclear export. This motif is found on the N-terminal domain of NEP (residues 1–54; Fig. 4) (O'Neill et al., 1998). NEP can bind in vitro to CRM1 (Neumann et al., 2000) and when incubated with CRM1, RanGTP and RanGAP, the GTP is indeed hydrolyzed as evidence for a biologically relevant interaction (Akarsu et al., 2003).

The atomic structure of the C-terminal half of NEP shows one face of the protein that is very acidic with in the middle of the glutamic acid side-chains an exposed tryptophan, W78 (Figs. 4 and 5). This tryptophan and the glutamic acids are involved in binding to M1 (Akarsu et al., 2003) and binding is lost when W78 is mutated to a serine. This surface of NEP binds to the positively charged NLS surface of M1 (Figs. 4 and 5). Interestingly, the proposed binding of M1 with NEP involving complementary charges plus a strong hydrophobic contribution resembles the binding of a peptide with a linear NLS to importin  $\alpha$  (Conti et al., 1998). Therefore, in the nucleus, M1 is proposed to bind to RNP with its C-terminal domain, leaving its N-terminal domains with the NLS free for binding to NEP that, by binding to the NLS, hides it at the same time. Then, NEP could bind to M1 with its C-terminal domain leaving its N-terminal domain free for interaction with CRM1 which can then bind RanGTP and the complex can be exported (Akarsu et al., 2003), see Fig. 4.

There remain a number of questions to be answered. First, it is not known how many M1-NEP molecules bind to the RNP.

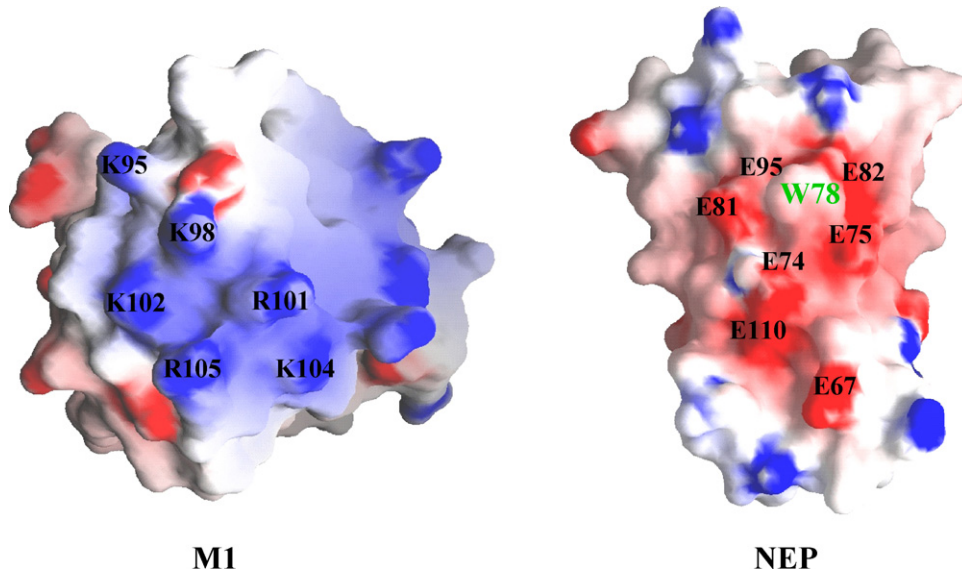


Fig. 5. Surface potentials of the interacting domains of M1 and NEP. The positively charged residues on M1 and the negatively charged residues on NEP are indicated, plus tryptophan 78 of NEP that is important for binding to M1.

There are about 200 copies of NEP in a mature virus particle (Yasuda et al., 1993) which could mean that some 20–30 M1-NEP complexes bind per RNP. Large complexes such as RNPs may indeed need more than one CRM1 molecule for transport. If the interaction between CRM1 and the nucleoporins is needed for transport one can expect that CRM1 needs to be present all along the RNP.

Second, only minus-strand RNPs or vRNPs are exported from the nucleus and not plus-strand or complementary RNPs (Shapiro et al., 1987). If so many M1-NEP complexes bind per RNP it cannot just be the conformation of the viral polymerase that determines export. It is possible that another aspect differs between vRNPs and cRNPs. M1 has a strong affinity for purified NP in the absence of RNA and the polymerase complex (Baudin et al., 2001) but we do not know if it also has affinity for the polymerase itself. For the moment it is not known whether there is a difference in phosphorylation status between v and cRNPs nor whether M1 can bind to cRNPs.

Third, it is not clear why the virus needs two extra proteins for export, rather than a single protein. The M1-NEP couple also forms a possibility for regulation of transport. Both proteins need to be present in high enough concentrations for transport to occur and post-translational modifications may also be used for regulation. Indeed, hyper-phosphorylation of M1 leads to its nuclear retention (Whittaker et al., 1995).

## 5. Multiple roles for the M1 protein

As all matrix proteins from negative strand RNA viruses (but also of retroviruses) for virus budding M1 needs to bind to the cytoplasmic tails of the glycoproteins, to lipid rafts, to the protein-encapsidated RNA genome and it has to be able to polymerize and pull the lipid rafts together in order to exclude most host cell membrane proteins from the virus particle (Schmitt and Lamb, 2005; Timmins et al., 2004). For the final membrane

pinching step M1 also has to bind to proteins from the ESCRT pathway. Here we have described a number of other functions for influenza virus M1, nuclear import of itself and nuclear export of RNPs.

The nuclear localization signal of M1 was initially identified as the following sequence  $_{100}\text{RKLKR}_{105}$  by Ye et al. (1995), see Table 1. When the crystal structure of the N-terminal domain of M1 was determined (Sha and Luo, 1997; Arzt et al., 2001, 2004) it appeared that this sequence was part of a larger positively charged surface area of the protein:  $_{95}\text{KAVKLYRKLKR}_{105}$  (Figs. 4 and 5). The same motif is also involved in binding to negatively charged membranes (Baudin et al., 2001) and binding to NEP. Most recently it was described by Hui et al. (2006) that residues  $_{100}\text{YRKL}_{103}$  of this motif also form the late domain of M1 that is involved in binding to ESCRT proteins. Finally, this positively charged surface binds, *in vitro*, also to RNA (Elster et al., 1997) but we do not know if RNA-binding has any biological relevance. Many proteins with a nuclear localization sequence bind nucleic acids through their NLS (LaCasse and Lefebvre, 1995). It is possible that this surface with multiple activities is not always exposed. As mentioned above, this surface is hidden when NEP binds to it, allowing nuclear export without re-importation. Both N- and C-terminal domains of M1 are rich in  $\alpha$ -helices and are well folded. The connection between the two domains may be through a flexible peptide (Baudin et al., 2001; Arzt et al., 2001). However, the connecting sequence also contains the zinc-binding sequence of M1 (Elster et al., 1994). Inside the mature virus particle there are two types of M1 molecules; those that have zinc bound and those that have not (Elster et al., 1994). The apposition of the two domains must be influenced by the presence of a bound zinc ion and may have an influence of the exposition of the NLS. If this were the case, the biological activity of M1 may be regulated by zinc binding. It is interesting to note that there are about as many M1 molecules with bound zinc as there are NEP copies in the mature virus particle. As

mentioned above, phosphorylation has also an influence on the biological activity of M1.

Finally, M1 has a strong affinity for NP and RNP (Baudin et al., 2001). When it is stuck to RNP it is difficult to remove it. In the test tube this is only possible when the pH of the buffer medium is lowered to about 5 (Zhirnov, 1992). The only moment when these conditions occur in the viral life cycle is at the beginning of the infection when the interior of the virus is acidified while it resides inside the endosome, just before membrane fusion. Only then are RNPs stripped off their M1 allowing them to enter the nucleus. When M1 binds to RNP in a later stage, they are exported from the nucleus through binding of NEP and CRM1 and they cannot reenter the nucleus (Martin and Helenius, 1991; Whittaker et al., 1996).

### Acknowledgements

We thank Dr. Christoph Müller (EMBL, Grenoble) and Nadia Naffakh (Pasteur Institute, Paris) for critical comments on the text. This work was financed in part by the EU FP6 project: VIRGIL (European Vigilance Network for the Management of Antiviral Drug Resistance).

### References

- Albo, C., Valencia, A., Portela, A., 1995. Identification of an RNA binding region within the N-terminal third of the influenza A virus nucleoprotein. *J. Virol.* 69, 3799–3806.
- Akarsu, H., Burmeister, W.P., Petosa, C., Petit, I., Muller, C.W., Ruigrok, R.W.H., Baudin, F., 2003. Crystal structure of the M1 protein-binding domain of the influenza A virus nuclear export protein (NEP/NEP). *EMBO J.* 22, 4455–4464.
- Askjaer, P., Bachi, A., Wilm, M., Bischoff, F.R., Weeks, D.L., Ogniewski, V., Ohno, M., Niehrs, C., Kjems, J., Mattaj, I.W., Fornerod, M., 1999. RanGTP-regulated interactions of CRM1 with nucleoporins and a shuttling DEAD-box helicase. *Mol. Cell. Biol.* 19, 6276–6285.
- Arrese, M., Portela, A., 1996. Serine 3 is critical for phosphorylation at the N-terminal end of the nucleoprotein of influenza virus A/Victoria/3/75. *J. Virol.* 70, 3385–3391.
- Arzt, S., Baudin, F., Barge, A., Timmins, P., Burmeister, W.P., Ruigrok, R.W.H., 2001. Combined results from solution studies on intact influenza virus M1 protein and from a new crystal form of its N-terminal domain show that M1 is an elongated monomer. *Virology* 279, 439–446.
- Arzt, S., Petit, I., Burmeister, W.P., Ruigrok, R.W.H., Baudin, F., 2004. Structure of a knockout mutant of influenza virus M1 protein that has altered activities in membrane binding, oligomerisation and binding to NEP (NEP). *Virus Res.* 99, 115–119.
- Barman, S., Ali, A., Hui, E.K., Adhikary, L., Nayak, D.P., 2001. Transport of viral proteins to the apical membranes and interaction of matrix protein with glycoproteins in the assembly of influenza viruses. *Virus Res.* 77, 61–69.
- Baudin, F., Petit, I., Weissenhorn, W., Ruigrok, R.W.H., 2001. In vitro dissection of the membrane and RNP binding activities of influenza virus M1 protein. *Virology* 281, 102–108.
- Bayliss, R., Littlewood, T., Stewart, M., 2000. Structural basis for the interaction between FxFG nucleoporin repeats and importin-beta in nuclear trafficking. *Cell* 102, 99–108.
- Biswas, S.K., Nayak, D.P., 1994. Mutational analysis of the conserved motifs of influenza A virus polymerase basic protein 1. *J. Virol.* 68, 1819–1826.
- Biswas, S.K., Nayak, D.P., 1996. Influenza virus polymerase basic protein 1 interacts with influenza virus polymerase basic protein 2 at multiple sites. *J. Virol.* 70, 6716–6722.
- Biswas, S.K., Boutz, P.L., Nayak, D.P., 1998. Influenza virus nucleoprotein interacts with influenza virus polymerase proteins. *J. Virol.* 72, 5493–5501.
- Bui, M., Wills, E.G., Helenius, A., Whittaker, G.R., 2000. Role of the influenza virus M1 protein in nuclear export of viral ribonucleoproteins. *J. Virol.* 74, 1781–1786.
- Bullido, R., Gomez-Puertas, P., Albo, C., Portela, A., 2000. Several protein regions contribute to determine the nuclear and cytoplasmic localization of the influenza A virus nucleoprotein. *J. Gen. Virol.* 81, 135–142.
- Cingolani, G., Petosa, C., Weis, K., Müller, C.W., 1999. Structure of importin beta bound to the IBB domain of importin- $\alpha$ . *Nature* 399, 221–229.
- Compans, R.W., Content, J., Duesberg, P.H., 1972. Structure of the ribonucleoprotein of influenza virus. *J. Virol.* 4, 795–800.
- Conti, E., Uy, M., Leighton, L., Blobel, G., Kuriyan, J., 1998. Crystallographic analysis of the recognition of a nuclear localization signal by the nuclear import factor karyopherin  $\alpha$ . *Cell* 94, 193–204.
- Conti, E., Muller, C.W., Stewart, M., 2006. Karyopherin flexibility in nucleocytoplasmic transport. *Curr. Opin. Struct. Biol.* 16, 1–8.
- Cros, J.F., Garcia-Sastre, A., Palese, P., 2005. An unconventional NLS is critical for the nuclear import of the influenza A virus nucleoprotein and ribonucleoprotein. *Traffic* 6, 205–213.
- Davey, J., Dimmock, N.J., Colman, A., 1985. Identification of the sequence responsible for the nuclear accumulation of the influenza virus nucleoprotein in *Xenopus* oocytes. *Cell* 40, 667–675.
- Deng, T., Sharps, J., Fodor, E., Brownlee, G.G., 2005. In vitro assembly of PB2 with a PB1-PA dimer supports a new model of assembly of influenza A virus polymerase subunits into a functional trimeric complex. *J. Virol.* 79, 8669–8674.
- Elster, C., Fourest, E., Baudin, F., Larsen, K., Cusack, S., Ruigrok, R.W.H., 1994. A small percentage of influenza virus M1 protein contains zinc but zinc does not influence in vitro M1-RNA interaction. *J. Gen. Virol.* 75, 37–42.
- Elster, C., Larsen, K., Gagnon, J., Ruigrok, R.W., Baudin, F., 1997. Influenza virus M1 protein binds to RNA through its nuclear localization signal. *J. Gen. Virol.* 78, 1589–1596.
- Elton, D., Medcalf, L., Bishop, K., Harrison, D., Digard, P., 1999a. Identification of amino acid residues of influenza virus nucleoprotein essential for RNA binding. *J. Virol.* 73, 7357–7367.
- Elton, D., Medcalf, E., Bishop, K., Digard, P., 1999b. Oligomerization of the influenza virus nucleoprotein: identification of positive and negative sequence elements. *Virology* 260, 190–200.
- Elton, D., Simpson-Holley, M., Archer, K., Medcalf, L., Hallam, R., McCauley, J., Digard, P., 2001. Interaction of the influenza virus nucleoprotein with the cellular CRM1-mediated nuclear export pathway. *J. Virol.* 75, 408–419.
- Fechter, P., Mingay, L., Sharps, J., Chambers, A., Fodor, E., Brownlee, G.G., 2003. Two aromatic residues in the PB2 subunit of influenza A RNA polymerase are crucial for cap binding. *J. Biol. Chem.* 278, 20381–20388.
- Fischer, U., Huber, J., Boelens, W.C., Mattaj, I.W., Lüthmann, R., 1995. The HIV-1 Rev activation domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs. *Cell* 82, 475–483.
- Fodor, E., Smith, M., 2004. The PA subunit is required for efficient nuclear accumulation of the PB1 subunit of the influenza A virus RNA polymerase complex. *J. Virol.* 78, 9144–9153.
- Fukuda, M., Asano, S., Nakamura, T., Adachi, M., Yoshida, M., Yanagida, M., Nishida, E., 1997. CRM1 is responsible for intracellular transport mediated by the nuclear export signal. *Nature* 390, 308–311.
- Gonzalez, S., Zurcher, T., Ortin, J., 1996. Identification of two separate domains in the influenza virus PB1 protein involved in the interaction with the PB2 and PA subunits: a model for the viral RNA polymerase structure. *Nucleic Acids Res.* 24, 4456–4463.
- Gonzalez, S., Ortin, J., 1999. Distinct regions of influenza virus PB1 polymerase subunit recognize vRNA and cRNA templates. *EMBO J.* 18, 3767–3775.
- Greenspan, D., Palese, P., Krystal, M., 1998. Two nuclear location signals in the influenza virus NS1 nonstructural protein. *J. Virol.* 62, 3020–3026.
- Hay, A.J., Skehel, J.J., 1975. Studies on the synthesis of influenza virus proteins. In: Mahy, B.W.J., Barry, R.D. (Eds.), *Negative Strand Viruses*. Academic Press, London, pp. 635–655.
- Hay, A.J., Wolstenholme, A.J., Skehel, J.J., Smith, M.H., 1985. The molecular basis of the specific anti-influenza action of amantadine. *EMBO J.* 4, 3021–3024.
- Hodel, M.R., Corbett, A.H., Hodel, A.E., 2001. Dissection of a nuclear localization signal. *J. Biol. Chem.* 276, 1317–1325.

- Honda, A., Mizumoto, K., Ishihama, A., 1999. Two separate sequences of PB2 subunit constitute the RNA cap-binding site of influenza virus RNA polymerase. *Genes Cells* 4, 475–485.
- Hui, E.K., Barman, S., Tang, D.H., France, B., Nayak, D.P., 2006. YRKL sequence of influenza virus M1 functions as the L domain motif and interacts with VPS28 and Cdc42. *J. Virol.* 80, 2291–2308.
- Imamoto, N., Shimamoto, T., Takao, T., Tachibana, T., Kose, S., Matsubae, M., Sekimoto, T., Shimonishi, Y., Yoneda, Y., 1995. In vivo evidence for involvement of a 58 kDa component of nuclear pore-targeting complex in nuclear protein import. *EMBO J.* 14, 3617–3626.
- Kawaguchi, A., Naito, T., Nagata, K., 2005. Involvement of influenza virus PA subunit in assembly of functional RNA polymerase complexes. *J. Virol.* 79, 732–744.
- Kemler, I., Whittaker, G., Helenius, A., 1994. Nuclear import of microinjected influenza virus ribonucleoproteins. *Virology* 202, 1028–1033.
- Klumpp, K., Ruigrok, R.W.H., Baudin, F., 1997. Roles of the influenza virus polymerase and nucleoprotein in forming a functional RNP structure. *EMBO J.* 16, 1248–1257.
- Kobayashi, M., Toyoda, T., Adyshev, D.M., Azuma, Y., Ishihama, A., 1994. Molecular dissection of influenza virus nucleoprotein: deletion mapping of the RNA binding domain. *J. Virol.* 68, 8433–8436.
- Kudo, N., Matsumori, N., Taoka, H., Fujiwara, D., Schreiner, E.P., Wolff, B., Yoshida, M., Horinouchi, S., 1999. Leptomycin B inactivates CRM1/exportin 1 by covalent modification at a cysteine residue in the central conserved region. *Proc. Natl. Acad. Sci. U.S.A.* 96, 9112–9117.
- LaCasse, E.C., Lefebvre, Y.A., 1995. Nuclear localisation signals overlap DNA- or RNA-binding domains in nucleic acid-binding proteins. *Nucl. Acids Res.* 23, 1647–1656.
- Li, Y., Yamakita, Y., Krug, R.M., 1998. Regulation of a nuclear export signal by an adjacent inhibitory sequence: the effector domain of the influenza virus NS1 protein. *Proc. Natl. Acad. Sci. U.S.A.* 95, 4864–4869.
- Li, M.L., Rao, P., Krug, R.M., 2001. The active sites of the influenza cap-dependent endonuclease are on different polymerase subunits. *EMBO J.* 20, 2078–2086.
- Martin, K., Helenius, A., 1991. Nuclear transport of influenza virus ribonucleoproteins: the viral matrix protein (M1) promotes export and inhibits import. *Cell* 67, 117–130.
- Melen, K., Fagerlund, R., Franke, J., Kohler, M., Kinnunen, L., Julkunen, I., 2003. Importin  $\alpha$  nuclear localization signal binding sites for STAT1, STAT2, and influenza A virus nucleoprotein. *J. Biol. Chem.* 278, 28193–28200.
- Min, J.Y., Krug, R.M., 2006. The primary function of RNA binding by the influenza A virus NS1 protein in infected cells: Inhibiting the 2'-5' oligo (A) synthetase/RNase L pathway. *PNAS* 103, 7100–7105.
- Miyamoto, Y., Imamoto, N., Sekimoto, T., Tachibana, T., Seki, T., Tada, S., Enomoto, T., Yoneda, Y., 1997. Differential modes of nuclear localization signal (NLS) recognition by three distinct classes of NLS receptors. *J. Biol. Chem.* 272, 26375–26381.
- Mukaigawa, J., Nayak, D.P., 1991. Two signals mediate nuclear localization of influenza virus (A/WSN/33) polymerase basic protein 2. *J. Virol.* 65, 245–253.
- Nath, S.T., Nayak, D.P., 1990. Function of two discrete regions is required for nuclear localization of polymerase basic protein 1 of A/WSN/33 influenza virus (H1N1). *Mol. Cell. Biol.* 10, 4139–4145.
- Neumann, G., Castrucci, M.R., Kawaoka, Y., 1997. Nuclear import and export of influenza virus nucleoprotein. *J. Virol.* 71, 9690–9700.
- Neumann, G., Hughes, M.T., Kawaoka, Y., 2000. Influenza A virus NEP protein mediates vRNP nuclear export through NES-independent interaction with hCRM1. *EMBO J.* 19, 6751–6786.
- Nieto, A., de la Luna, S., Barcena, J., Portela, A., Ortin, J., 1994. Complex structure of the nuclear translocation signal of influenza virus polymerase PA subunit. *J. Gen. Virol.* 75, 29–36.
- Ohno, M., Segref, A., Bachi, A., Wilm, M., Mattaj, I.W., 2000. PHAX, a mediator of U snRNA nuclear export whose activity is regulated by phosphorylation. *Cell* 101, 187–198.
- Ohtsu, Y., Honda, Y., Sakata, Y., Kato, H., Toyoda, T., 2002. Fine mapping of the subunit binding sites of influenza virus RNA polymerase. *Microbiol. Immunol.* 46, 167–175.
- O'Neill, R.E., Palese, P., 1995. NPI-1, the human homolog of SRP-1, interacts with influenza virus nucleoprotein. *Virology* 206, 116–125.
- O'Neill, R.E., Talon, J., Palese, P., 1998. The influenza virus NEP (NS2 protein) mediates the nuclear export of viral ribonucleoproteins. *EMBO J.* 17, 288–296.
- Ossareh-Nazari, B., Bachelier, F., Dargemont, C., 1997. Evidence for a role of CRM1 in signal-mediated nuclear protein export. *Science* 278, 141–144.
- Perez, D.R., Donis, R.O., 1995. A 48-amino-acid region of influenza A virus PB1 protein is sufficient for complex formation with PA. *J. Virol.* 69, 6932–6939.
- Petosa, C., Schoehn, G., Askjaer, P., Bauer, U., Moulin, M., Steuerwald, U., Soler-Lopez, M., Baudin, F., Mattaj, I.W., Muller, C.W., 2004. Architecture of CRM1/Exportin1 suggests how cooperativity is achieved during formation of a nuclear export complex. *Mol. Cell.* 16, 761–775.
- Pinto, L.H., Holsinger, L.J., Lamb, R.A., 1992. Influenza virus M2 protein has ion channel activity. *Cell* 69, 517–528.
- Poch, O., Sauvaget, I., Delarue, M., Tordo, N., 1989. Identification of four conserved motifs among the RNA-dependent polymerase encoding elements. *EMBO J.* 8, 3867–3874.
- Poole, E., Elton, D., Medcalf, L., Digard, P., 2004. Functional domains of the influenza A virus PB2 protein: identification of NP- and PB1-binding sites. *Virology* 321, 120–133.
- Ruigrok, R.W., Baudin, F., 1995. Structure of influenza virus ribonucleoprotein particles. II. Purified RNA-free influenza virus ribonucleoprotein forms structures that are indistinguishable from the intact influenza virus ribonucleoprotein particles. *J. Gen. Virol.* 76, 1009–1014.
- Sakaguchi, A., Hirayama, E., Hiraki, A., Ishida, Y., Kim, J., 2003. Nuclear export of influenza viral ribonucleoprotein is temperature-dependently inhibited by dissociation of viral matrix protein. *Virology* 306, 244–253.
- Schmitt, A.P., Lamb, R.A., 2005. Influenza virus assembly and budding at the viral budzone. *Adv. Virus Res.* 64, 383–416.
- Schwartz, T.U., 2005. Modularity within the architecture of the nuclear pore complex. *Curr. Opin. Struct. Biol.* 15, 221–226 (review).
- Sha, B., Luo, M., 1997. Structure of a bifunctional membrane-RNA binding protein, influenza virus matrix protein. *Nat. Struct. Biol.* 4, 239–244.
- Shapiro, G.I., Gurney Jr., T., Krug, R.M., 1987. Influenza virus gene expression: control mechanisms at early and late times of infection and nuclear-cytoplasmic transport of virus-specific RNAs. *J. Virol.* 61, 764–773.
- Shinya, K., Ebina, M., Yamada, S., Ono, M., Kasai, N., Kawaoka, Y., 2006. Avian flu: influenza virus receptors in the human airway. *Nature* 440, 435–436.
- Skehel, J.J., Wiley, D.C., 2000. Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. *Annu. Rev. Biochem.* 69, 531–569.
- Timmins, J., Ruigrok, R.W.H., Weissenhorn, W., 2004. Structural studies on the Ebola virus matrix protein VP40 indicate that matrix proteins of enveloped RNA viruses are analogues but not homologous. *FEMS Microbiol. Lett.* 233, 179–186.
- Toyoda, T., Adyshev, D.M., Kobayashi, M., Iwata, A., Ishihama, A., 1996. Molecular assembly of the influenza virus RNA polymerase: determination of the subunit-subunit contact sites. *J. Gen. Virol.* 77, 2149–2157.
- Van Riel, D., Munster, V.J., de Wit, E., Rimmelzwaan, G.F., Fouchier, R.A., Osterhaus, A.D., Kuiken, T., 2006. H5N1 virus attachment to lower respiratory tract. *Science* 312, 319.
- Wang, P., Palese, P., O'Neill, R.E., 1997. The NPI-1/NPI-3 (karyopherin  $\alpha$ ) binding site on the influenza A virus nucleoprotein NP is a nonconventional nuclear localization signal. *J. Virol.* 71, 1850–1856.
- Watanabe, K., Takizawa, N., Katoh, M., Hoshida, K., Kobayashi, N., Nagata, K., 2001. Inhibition of nuclear export of ribonucleoprotein complexes of influenza virus by leptomycin B. *Virus Res.* 77, 31–42.
- Weber, F., Kochs, G., Gruber, S., Haller, O., 1998. A classical bipartite nuclear localization signal on Thogoto and influenza A virus nucleoproteins. *Virology* 250, 9–18.
- Weis, K., Mattaj, I.W., Lamond, A.I., 1995. Identification of hSRP1 as a functional receptor for nuclear localization sequences. *Science* 268, 1049–1053.
- Weis, K., 2002. Nucleocytoplasmic transport: cargo trafficking across the border. *Curr. Opin. Cell Biol.* 14, 328–335.
- Whittaker, G., Kemler, I., Helenius, A., 1995. Hyperphosphorylation of mutant influenza virus matrix protein, M1, causes its retention in the nucleus. *J. Virol.* 69, 439–445.

- Whittaker, G., Bui, M., Helenius, A., 1996. Nuclear trafficking of influenza virus ribonucleoproteins in heterokaryons. *J. Virol.* 70, 2743–2756.
- Yasuda, J., Nakada, S., Kato, A., Toyoda, T., Ishihama, A., 1993. Molecular assembly of influenza virus: association of the NEP protein with virion matrix. *Virology* 196, 249–255.
- Ye, Z., Robinson, D., Wagner, R.R., 1995. Nucleus-targeting domain of the matrix protein (M1) of influenza virus. *J. Virol.* 69, 1964–1970.
- Zhirnov, O.P., 1992. Isolation of matrix protein M1 from influenza viruses by acid-dependent extraction with nonionic detergent. *Virology* 186, 324–330.

**Structure and nuclear import function of the  
C-terminal domain of influenza virus  
polymerase PB2 subunit.**

Tarendeau, F., Boudet, J., Guilligay, D., Mas, P.J., Bougault, C.M., Boulo, S., Baudin, F.,  
Ruigrok, R.W., Daigle, N., Ellenberg, J., Cusack, S., Simorre, J.P. & Hart, D.J.

Nat Struct Mol Biol. 2007 Mar;14(3):229-33. Epub 2007 Feb 25.

# Structure and nuclear import function of the C-terminal domain of influenza virus polymerase PB2 subunit

Franck Tarendeau<sup>1,5</sup>, Julien Boudet<sup>2,5</sup>, Delphine Guilligay<sup>1</sup>, Philippe J Mas<sup>1</sup>, Catherine M Bougault<sup>2</sup>, Sébastien Boulo<sup>3</sup>, Florence Baudin<sup>3</sup>, Rob W H Ruigrok<sup>3</sup>, Nathalie Daigle<sup>4</sup>, Jan Ellenberg<sup>4</sup>, Stephen Cusack<sup>1</sup>, Jean-Pierre Simorre<sup>2</sup> & Darren J Hart<sup>1</sup>

**The trimeric influenza virus polymerase, comprising subunits PA, PB1 and PB2, is responsible for transcription and replication of the segmented viral RNA genome. Using a novel library-based screening technique called expression of soluble proteins by random incremental truncation (ESPRIT), we identified an independently folded C-terminal domain from PB2 and determined its solution structure by NMR. Using green fluorescent protein fusions, we show that both the domain and the full-length PB2 subunit are efficiently imported into the nucleus dependent on a previously overlooked bipartite nuclear localization sequence (NLS). The crystal structure of the domain complexed with human importin  $\alpha 5$  shows how the last 20 residues unfold to permit binding to the import factor. The domain contains three surface residues implicated in adaptation from avian to mammalian hosts. One of these tethers the NLS-containing peptide to the core of the domain in the unbound state.**

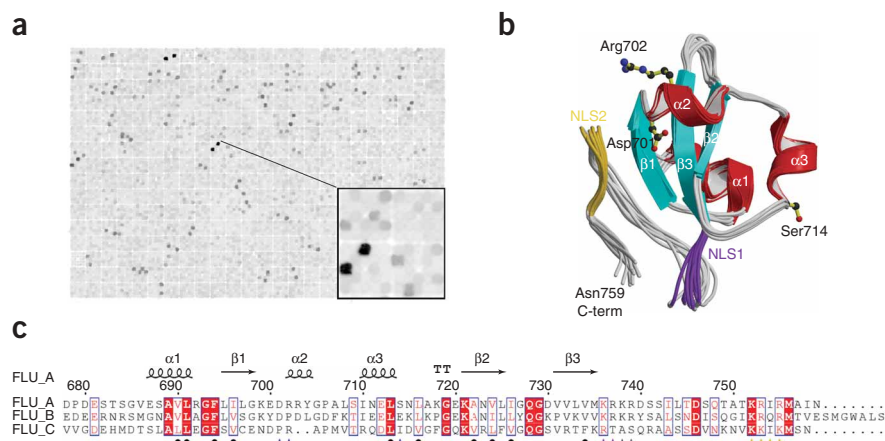
The eight segments of the influenza virus RNA genome are packaged into ribonucleoprotein particles (RNPs) containing the nucleoprotein (NP) and the trimeric RNA-dependent RNA polymerase complex, which comprises subunits PA, PB1 and PB2. The polymerase operates in two distinct modes: in the first, it transcribes virally encoded genes, using a cap-snatching mechanism to prime transcription and ensure proper 5' capping of viral messenger RNA; in the second, it replicates full-length viral RNA to produce first positive-strand complementary RNA and then progeny viral RNA<sup>1</sup>. The PB1 subunit binds the conserved 5' and 3' ends of the viral RNA<sup>2–4</sup> and carries both the polymerase active site and the endonuclease activity that cleaves host-cell mRNA bound by the cap-binding PB2 subunit<sup>5</sup>. PB1 residues implicated in the endonuclease and polymerase active sites have been identified<sup>5</sup>, although the location of the cap-binding site of PB2 remains controversial<sup>5–7</sup>. The polymerase is active in the nucleus and NLSs have been identified on PB1 (ref. 8), PB2 (ref. 9) and PA<sup>10</sup>. Recent results suggest that cytoplasmically expressed PB1 and PA may be imported by the importin RanBP5 as a subcomplex, which then assembles with separately imported PB2 (refs. 11,12). The host proteins mediating PB2 nuclear import have not been characterized, although it seems to be chaperoned by Hsp90 during transport<sup>13</sup>.

Current concern about the adaptation of highly pathogenic H5N1 avian influenza strains to humans has highlighted the need to understand the genetic and molecular determinants of virulence and

transmissibility of influenza viruses<sup>14</sup>. Recent studies have focused attention on polymerase subunit and NP point mutations that accompany natural or laboratory-selected adaptations from avian to mammalian hosts, notably in the case of the 1918 pandemic strain<sup>15–18</sup>. However, the functional implications of these mutations—for instance, whether they affect polymerase activity, assembly, stability or interactions with host factors—are poorly understood. This is partly due to the lack of detailed structural information on the polymerase, although low-resolution negative-stain electron microscopic reconstructions of a mini-RNP<sup>19</sup> and of the polymerase trimer<sup>20</sup> are available. The intractability of the polymerase to structure determination results partly from the fact that bacterially expressed subunits are insoluble and yields of active enzyme produced in insect cells or mammalian cells are prohibitively low. Additionally, biochemical or bioinformatic knowledge of the domain structure of the polymerase subunits is limited. As a new approach to overcome these problems, we devised a directed evolution-type screen called ESPRIT to isolate bacterial clones expressing soluble fragments of polymerase subunits from an essentially complete library of terminal truncation variants. Here we present the structural and functional characterization of a PB2 C-terminal domain identified by this screen. This is the first successful application of a new soluble expression screening strategy that should be of general use in finding soluble domains in otherwise insoluble proteins for both structural and functional studies.

<sup>1</sup>European Molecular Biology Laboratory (EMBL) Grenoble Outstation, 6 rue Jules Horowitz, BP181, 38042 Grenoble Cedex 9, France. <sup>2</sup>Institut de Biologie Structurale Jean-Pierre Ebel UMR 5075 CNRS-CEA-UJF, 41 rue Jules Horowitz, 38027 Grenoble Cedex 1, France. <sup>3</sup>Unit of Virus Host Cell Interactions UMR 5233 UJF-EMBL-CNRS, 6 rue Jules Horowitz, BP181, 38042 Grenoble Cedex 9, France. <sup>4</sup>EMBL Gene Expression Programme, Meyerhofstrasse 1, D-69117 Heidelberg, Germany. <sup>5</sup>These authors contributed equally to this work. Correspondence should be addressed to D.J.H. (hart@embl.fr) or S.C. (cusack@embl.fr).

Received 13 December 2006; accepted 31 January 2007; published online 25 February 2007; doi:10.1038/nsmb1212



**Figure 1** Identification and solution NMR structure of PB2 C-terminal domain. **(a)** Protein expression screen of 26,880 random deletion constructs of the *pb2* gene. Stable expression of soluble protein results in efficient *in vivo* labeling of a C-terminal biotin acceptor peptide. Detection is by fluorescent streptavidin and fluorimaging. **(b)** Ribbon diagram of the ten lowest-energy NMR structures superimposed using backbone heavy atoms (r.m.s. deviation 0.94 Å for 224 atoms). Indicated are Asp701, Arg702 and Ser714, which are implicated in cross-species transmission, and basic regions corresponding to the minor (purple) and major (gold) sites of the bipartite NLS. **(c)** Primary sequence alignment comparing influenza (FLU) A, B and C strains. Indicated are residues implicated in cross-species transmission (blue triangles), the minor site (purple triangles) and major site (gold triangles) of the bipartite NLS, conserved buried hydrophobic residues (black ovals) and secondary structure elements.

## RESULTS

### Isolation of a soluble domain by random construct screening

An exonuclease protocol<sup>21</sup> was used to generate a comprehensive, unbiased set of 5' *pb2* gene deletions, to which N-terminal methionine and lysine were genetically fused. A library of approximately 26,880 constructs with a seven-fold oversample of each clone was then arrayed robotically onto nitrocellulose membranes and screened for soluble protein expression in a colony format using a C-terminal biotin acceptor peptide<sup>22</sup>. This short peptide is biotinylated *in vivo* by endogenous BirA and has been extensively exploited as a tool for affinity purification and immobilization using avidin. Here we used efficient biotinylation of expressed variants as a quantitative readout for expression of stable, soluble protein. Colony array blots were prepared by *in situ* lysis and expression of biotinylated proteins was detected with a fluorescent streptavidin conjugate and fluorimeter (Fig. 1a). Intense positive clones were sequenced to identify the truncation boundaries, revealing that many were predicted to express C-terminal protein fragments of 80–110 amino acid residues. Two clones in particular, corresponding to residues 678–759 (designated DPDE, from the first four residues) and 661–759 (TTKR), were

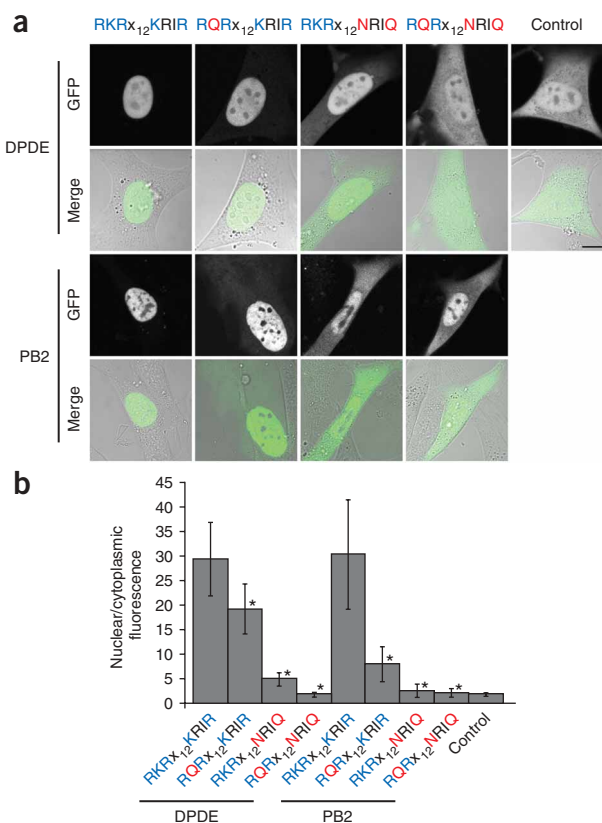
**Figure 2** Nuclear import of PB2 C-terminal domain and full-length PB2 subunit is directed by a bipartite NLS. **(a)** Steady-state subcellular localization of PB2 C-terminal domain (DPDE) or full-length PB2 fused to enhanced GFP in NIH 3T3 cells. Constructs have either the wild-type NLS (737-RKR<sub>x12</sub>KRIR-755) or the point mutations shown in red. Images show 2- $\mu$ m single confocal sections of GFP fluorescence (GFP) and merged DIC and fluorescence (Merge). Contrast and brightness have been adjusted for display purposes. Scale bar, 10  $\mu$ m. **(b)** Ratio of nuclear to cytoplasmic concentration of GFP-fusion proteins, as a measure of nuclear import. Nuclear and cytoplasmic mean fluorescence intensities were background-subtracted and normalized to differences in total cell fluorescence intensity for each mutant. Mean nuclear/cytoplasmic fluorescence intensities are shown with s.d. ( $n = 12$  cells for each mutant). Asterisks mark significant differences from wild-type construct ( $P < 0.05$ , Student's *t*-test).

expressed in soluble form at high levels. Multimilligram quantities of each protein were purified and characterized as monodisperse by gel filtration.

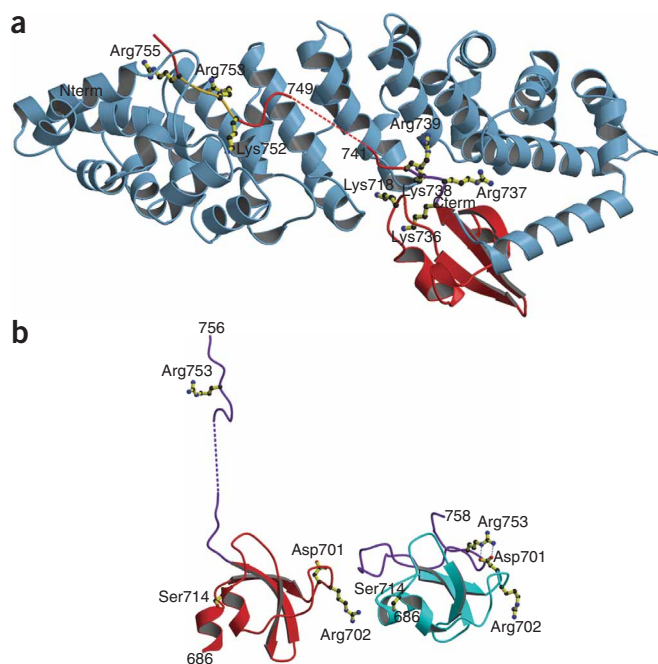
### NMR solution structure of the DPDE domain

The <sup>1</sup>H-<sup>15</sup>N HSQC NMR spectrum of DPDE shows a good dispersion of the amide resonances at 10 °C, characteristic of a folded domain (Supplementary Fig. 1 online). TTKR differs only in the extent of disordered N-terminal peptide. Measurements on a sample of uniformly <sup>13</sup>C,<sup>15</sup>N-labeled DPDE led to the assignment of backbone chemical shifts for all residues except Asp678, Gly693 and Ser741. At 25 °C, the intensities of the resonances of the N-terminal (678–685) and C-terminal (744–759) regions were reduced, as were those of Arg702, Gln728, Gly729 and Lys738. Measurements of <sup>15</sup>N-backbone relaxation at 10 °C and analysis of data showed that, with the exception of Lys738, all temperature-sensitive residues belong to flexible parts of the molecule (data not shown). The lack of <sup>1</sup>H-<sup>1</sup>H NOEs for residues 678–684 confirms the absence of a stable conformation for the N-terminal extremity.

Structure calculations were therefore performed for residues 685–759 using NMR restraints measured at 10 °C. Corresponding structural statistics are presented in Supplementary Table 1 online.







**Figure 3** X-ray structure of the PB2 C-terminal domain complexed with importin  $\alpha 5$ . **(a)** Ribbon diagram showing DPDE (red) bound to human importin  $\alpha 5$  (blue), comprising ten armadillo repeats. The C-terminal helix of the importin is unpacked and mediates domain-swap dimer formation in the crystal. The bipartite NLS at the C terminus of DPDE binds classically within the superhelical groove of importin  $\alpha 5$ . Basic residues Arg737, Lys738 and Arg739 from the minor site (purple) interact with the C-terminal armadillo repeats; Lys752, Arg753 and Arg755 from the major site (gold) interact with the N-terminal armadillo repeats. Lys736 does not interact with importin  $\alpha 5$  but makes intramolecular hydrogen bonds in DPDE, perhaps preventing further unfolding of the C terminus. Lys718 makes three hydrogen bonds with importin  $\alpha 5$ . **(b)** Comparison of the PB2 domain structure in complexed (red) and free solution state (cyan) demonstrates unfolding of residues 736–759 (purple) upon binding to importin  $\alpha 5$ . Residue Asp701, important in host specificity and virulence, forms a salt bridge with Arg753 of the major NLS motif and tethers the C terminus to the core of the domain in the unbound state. Residues Arg702 and Ser714 are also implicated in interspecies transmission. Note different orientations of the N-terminal helix of DPDE in the two structures.

The solution structure of the C-terminal domain of PB2 reveals a compact  $\alpha$ - $\beta$  domain with a hydrophobic core formed by a three-stranded antiparallel  $\beta$ -sheet, onto which a short amphipathic  $\alpha$ -helix packs (Fig. 1b). Except for the flexible loop 726–730 connecting the  $\beta 2$  and  $\beta 3$  strands and the last 20 residues at the C terminus, the structure was determined with a high resolution (r.m.s. deviation of  $0.42 \pm 0.12$  Å, calculated for the 81 backbone atoms of the secondary structure elements). Regions of lower structural resolution are correlated with enhanced dynamics (Supplementary Fig. 2 online). The PB2 C-terminal domain seems to be a novel structural motif, as no similar protein with a Z-score  $> 2$  was found using Dali<sup>23</sup>. It is likely that the equivalent domains of PB2 from influenza B and C strains have similar structures because of their relatively high sequence conservation, particularly of buried hydrophobic residues (Fig. 1c).

### Importin $\alpha 5$ binding and live-cell localization studies

Efficient nuclear localization of PB2 is believed to depend on an internal region (residues 449–495) and a classical monopartite basic motif, 736-KRKR-739 (ref. 9), which is present in DPDE. Mutation<sup>9</sup> of the basic motif, deletion<sup>24</sup> of a C-terminal region containing it or deletion of the internal domain<sup>9</sup> all abrogate nuclear localization. In the NMR structure, 736-KRKR-739 immediately follows strand  $\beta 3$ , and all the basic side chains are surface exposed, except that of Lys736, which is partly buried. To investigate the nuclear localization properties of the C-terminal domain, we studied its interaction with human importin  $\alpha 5$  ( $\alpha$  importin), the classical eukaryotic nuclear import adaptor. We first showed that DPDE forms a stable complex with human importin  $\alpha 5$  *in vitro* (Supplementary Fig. 3 online), suggesting that the domain may be capable of directing nuclear import. We then performed live-cell localization studies using a green fluorescent protein (GFP)-DPDE fusion, which was found almost exclusively in the nucleus ( $29 \pm 8$  times more nuclear than cytoplasmic), whereas a GFP control remained equilibrated between the cytoplasm and nucleus (Fig. 2). As previous results showed that the monopartite motif was itself insufficient for nuclear import<sup>9</sup>, we examined

sequence alignments for other residues that might be implicated (Fig. 1c). These show that although the monopartite NLS motif is conserved in influenza B, only two basic residues occur in influenza C (KRTA). However, a second basic motif, 752-KRIR-755, conserved in influenza B and C strains, suggested the presence of a classical bipartite NLS (KR $x_{12-15}$ KR $x$ R, where  $x$  is any residue and  $x_{12-15}$  is any sequence 12–15 residues long)<sup>25</sup>. To test this, the GFP-DPDE fusion was mutated in the putative bipartite NLS (Fig. 2). Mutating the first ('minor') part of the NLS (K738Q) mildly but significantly diminished nuclear import ( $19 \pm 5$  times more nuclear;  $P = 0.0004$  compared with wild-type), in agreement with studies on the whole PB2 protein<sup>9</sup>. Mutating the second ('major') part (K752N R755Q) greatly diminished nuclear import ( $5 \pm 1$  times more nuclear;  $P = 1.6 \times 10^{-8}$  compared with wild-type). Combining minor (K738Q) and major (K752N R755Q) NLS mutations abolished import to levels of GFP alone ( $2 \pm 1$  times more nuclear;  $P = 9.5 \times 10^{-8}$  compared with wild-type and  $P = 0.86$  compared with GFP). Similar data were obtained with a second mutation in the minor region (R737Q; data not shown) and with full-length PB2-GFP fusions (Fig. 2). Consistent results were also obtained when DPDE domains bearing the same NLS mutants were analyzed for importin  $\alpha 5$  complex formation *in vitro* (Supplementary Fig. 3). Thus, the C-terminal PB2 sequence 738-KR $x_{12}$ KRIR-755 functions as an efficient, classical, bipartite nuclear import signal, with a similar response to mutations as that of the retinoblastoma protein (KR $x_{11}$ KKLR) or phosphoprotein NIN2 (KR $x_{12}$ KKSK)<sup>25,26</sup>. Although this NLS is functional in the isolated domain as well as the full-length PB2 subunit, we cannot exclude the possibility that other sequences present in PB2 are responsible for the very low residual nuclear localization seen with the double NLS mutant of the whole protein.

### Cocrystal structure of DPDE with human importin $\alpha 5$

To provide further evidence for this revised mechanism of PB2 nuclear import, we cocrystallized DPDE with residues 66–512 of human importin  $\alpha 5$  (lacking the autoinhibitory N-terminal region) and determined the X-ray structure at 2.2-Å resolution (see Methods and Supplementary Fig. 4 online). The structure of importin  $\alpha 5$  comprises ten armadillo repeats and is similar to that of yeast karyopherin  $\alpha$  (PDB 1BK5), with an r.m.s. deviation of 2 Å (for 400 C $\alpha$  positions of 422 residues aligned) and 56% sequence identity. This is the first structure of a human importin  $\alpha$  family member and the first of a complex of importin with a complete folded domain, rather than simply an NLS-containing peptide. DPDE interacts with

importin  $\alpha 5$  with its C-terminal residues beyond Lys736 in an extended conformation, permitting binding of the bipartite NLS to two distinct regions within the superhelical groove of importin  $\alpha 5$  (Fig. 3a). The mode of NLS binding is similar to that previously observed for bipartite NLS peptides binding yeast and mouse  $\alpha$  importins<sup>25,27</sup>. Strong interactions of the side chains of three basic residues in both the minor (737-RKR-739) and major (752-KRIR-755) sites, within discrete pockets of the importin, are supplemented by additional hydrogen bonds, notably from the NE1 atoms of importin tryptophans Trp149, Trp191, Trp234 and Trp360 to the DPDE NLS peptide main chain carbonyl oxygens (Supplementary Fig. 5 online). Residues 742–747 of the long (12-residue) linker are not visible in the electron density, nor are the C-terminal residues 758–759, presumably owing to flexibility. The side chain of Lys736 of DPDE makes strong intradomain hydrogen bonds with the main chains of Ser714 and Leu716 in both the free and bound states of the DPDE domain (Supplementary Fig. 5), consistent with it not being crucial for nuclear import<sup>9</sup>. The folded region of the DPDE domain packs against importin  $\alpha 5$  with a total buried surface area of about 916 Å<sup>2</sup>; in particular, Lys718 makes three hydrogen bonds with Gly284, Asn286 and Thr325 of importin  $\alpha 5$  (Fig. 3a).

## DISCUSSION

The structure of the core of the DPDE domain (residues 693–736) observed in solution by NMR is very similar to the structure in the crystalline state in complex with importin  $\alpha 5$  (r.m.s. deviation of C $\alpha$  positions of 1.87 Å for 41 residues aligned), although the X-ray structure shows more secondary structure (Fig. 3b). The comparison reveals how the region 736–759 unfolds to allow binding of the bipartite NLS to the import factor (Fig. 3b). In the unbound NMR structure, the carboxyl group of Asp701 forms a salt bridge with Arg753, an important basic residue of the major site of the bipartite NLS motif, thus tethering the C-terminal extremity to the core of the domain. Binding to importin  $\alpha 5$  requires that this interaction be broken, as shown by the structure of the complex (Fig. 3a). A number of recent studies have shown that PB2 residue 701 is important for interspecies transmission. The D701N substitution has been found to contribute considerably to the acquisition of pathogenicity to mice of both H5N1 (ref. 15) and H7N7 (ref. 17) avian strains and has been observed in both fatal and nonfatal cases of H5N1 infecting humans in Vietnam<sup>28</sup>. We note that the mutation D701N would disrupt the salt bridge with Arg753 of the NLS, potentially altering the energetics of importin binding and, in turn, the efficiency of trimeric polymerase assembly. The adjacent residue 702 is also an important determinant of host specificity and, with only rare exceptions, is an arginine in human isolates, including the 1918 H1N1 pandemic virus, and a lysine in avian strains<sup>16</sup>. This residue is completely solvent exposed in both structures, and its function is not yet known. Mutation of a third residue in this domain, Ser714, in common with position 701, affects levels of polymerase activity of reconstituted ribonucleoprotein complexes in mammalian 293T cells; mutants with Asn701 and Arg714 have the highest activity<sup>17</sup>. The fact that these three cross-species-transfer residues<sup>15–17</sup> are exposed to solvent suggests a possible role of this domain in intermolecular interactions with other viral or host proteins; the interaction with importin  $\alpha 5$  and the regulation of nuclear transport of PB2 both could be important in this respect.

## METHODS

**Identification of soluble constructs.** DNA encoding PB2 from strain A/Victoria/3/75(H3N2)<sup>29</sup> fused to a C-terminal biotin acceptor peptide was

truncated from the 5' end with exonuclease III. An *Escherichia coli* BL21 strain was transformed with a pool of 24,036 plasmids, and clones were screened for protein expression by robotic arraying onto nitrocellulose membranes over agar. Colonies were lysed *in situ* and streptavidin conjugates used to detect biotinylated proteins (Supplementary Methods online).

**NMR spectroscopy and structure calculation.** Heteronuclear NMR data for assignment and extraction of structural restraints were collected at 10 °C on a 0.67 mM uniformly <sup>13</sup>C,<sup>15</sup>N-labeled DPDE sample in 50 mM Tris buffer (90% (v/v) H<sub>2</sub>O, 10% (v/v) D<sub>2</sub>O) at pH 7.5 with 200 mM NaCl, 1 mM EDTA and 0.01% (w/v) NaN<sub>3</sub>. Structure calculations including residues 685–759 were performed using CNS<sup>30</sup> and refined in explicit water (Supplementary Methods).

**Crystallography.** Cocrystals of DPDE and human importin  $\alpha 5$  (NSMB Protein AAP35605; residues 66–512) were grown by vapor diffusion with 2-methyl-2,4-pentanediol (MPD) as precipitant. Data to 2.2-Å resolution were measured on beamline ID23-1 at the European Synchrotron Radiation Facility, using X-rays of wavelength 1.072 Å and a temperature of 100 K, and integrated with XDS<sup>31</sup>. Two complexes were located in the asymmetric unit by molecular replacement using PDB 1Q1S (mouse importin  $\alpha$  with NLS peptide) as search model. The map was greatly improved using RESOLVE<sup>32</sup> with two-fold noncrystallographic symmetry averaging, allowing the complete structure (residues 84–508 of importin  $\alpha 5$ , and 686–741 and 748–757 of DPDE) to be constructed and refined using REFMAC<sup>33</sup>. The final *R*-factor was 20.6% (*R*<sub>free</sub> = 24.7%). According to Molprobity (http://molprobity.biochem.duke.edu/), 98.4% and 99.6% of the total of 966 residues are within the favored or allowed regions of the Ramachandran plot, respectively, with four outliers (see Table 1 and Supplementary Methods).

**Nuclear localization.** Mouse Swiss National Institutes of Health embryonic fibroblasts (NIH 3T3) were transfected with plasmids encoding the PB2

**Table 1 Crystallographic data collection and refinement statistics**

	Human importin $\alpha 5$ with influenza A PB2 C-terminal domain
<b>Data collection</b>	
Space group	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
Cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	94.97, 100.55, 151.84
Resolution (Å)	50–2.2 (2.2–2.3)
<i>R</i> <sub>merge</sub>	0.093 (0.694)
<i>I</i> / $\sigma$ <i>I</i>	9.40 (2.41)
Completeness (%)	99.5 (99.9)
Redundancy	4.00 (4.06)
<b>Refinement</b>	
Resolution (Å)	30–2.2 (2.257–2.200)
No. reflections (total / free)	74,089 / 2,969
<i>R</i> <sub>work</sub>	0.206 (0.284)
<i>R</i> <sub>free</sub>	0.247 (0.340)
No. atoms	7,863
Protein	7,597 <sup>a</sup>
Water	266
<i>B</i> -factors (Å <sup>2</sup> )	41.15
Protein (chain A, chain B)	41.9, 48.3
DPDE (chain D, chain E)	62.7, 51.4
Water	33.2
R.m.s. deviations	
Bond lengths (Å)	0.012
Bond angles (°)	1.273

Values in parentheses are for highest-resolution shell. <sup>a</sup>Two molecules in the asymmetric unit.

domain or full-length PB2 fused C-terminally to enhanced GFP. Cells were imaged by quantitative confocal microscopy (**Supplementary Methods**).

**Accession codes.** BioMagResBank (<http://www.bmrb.wisc.edu>): The  $^1\text{H}$ ,  $^{13}\text{C}$  and  $^{15}\text{N}$  chemical shifts for the C-terminal domain of PB2 with Pro679 in the *trans* (major) and *cis* (minor) conformations have been deposited with accession number 7056 and linked to the atomic coordinates of the ten structures of lowest energy in the Protein Data Bank (PDB 2GMO and RCSB 037302). Protein Data Bank: coordinates and structure factors of the PB2 domain–importin  $\alpha 5$  complex have been deposited with accession codes 2JDQ and R2JDQSF, respectively.

Note: Supplementary information is available on the Nature Structural & Molecular Biology website.

#### ACKNOWLEDGMENTS

We thank J. Ortin (Centro Nacional de Biotecnología, CSIC, Madrid) for the *pb2* gene, A. Favier (Institut de Biologie Structurale, Grenoble) for NMR scripts and the EMBL Centre for Molecular and Cellular Imaging for suggestions. Screening for crystals was done by the Partnership for Structural Biology high-throughput crystallization facility. We thank the European Synchrotron Radiation Facility and EMBL Joint Structural Biology group for assistance with the synchrotron beamtime and T. Crepin (EMBL, Grenoble) for help with data collection. Partial funding was provided by the European Commission Framework 5 Integrated Project 'Structural Proteomics in Europe' (SPINE, contract QL-G-CT-2002-00988).

#### AUTHOR CONTRIBUTIONS

D.J.H. conceived the ESPRIT method. D.J.H., F.T. and P.J.M. implemented ESPRIT. D.G. purified wild-type and mutant DPDE for *in vitro* binding studies and crystallization and made double-labeled protein for NMR. C.M.B., J.B. and J.-P.S. performed the NMR measurements and structural analysis. S.B. purified importin  $\alpha 5$  under the supervision of F.B. and cocrystallized it with DPDE. R.W.H.R. and S.C. initiated the influenza polymerase project, and S.C. determined the crystallographic structure. F.T. and N.D. performed nuclear import assays with instrumentation and methodology established by J.E. S.C. and D.J.H. compiled the text, with contributions from all authors.

#### COMPETING INTERESTS STATEMENT

The authors declare competing financial interests (see the Nature Structural & Molecular Biology website for details).

Published online at <http://www.nature.com/nsmb/>

Reprints and permissions information is available online at <http://npg.nature.com/reprintsandpermissions>

- Noah, D.L. & Krug, R.M. Influenza virus virulence and its molecular determinants. *Adv. Virus Res.* **65**, 121–145 (2005).
- Jung, T.E. & Brownlee, G.G. A new promoter-binding site in the PB1 subunit of the influenza A virus polymerase. *J. Gen. Virol.* **87**, 679–688 (2006).
- Li, M.L., Ramirez, B.C. & Krug, R.M. RNA-dependent activation of primer RNA production by influenza virus polymerase: different regions of the same protein subunit constitute the two required RNA-binding sites. *EMBO J.* **17**, 5844–5852 (1998).
- Gonzalez, S. & Ortin, J. Characterization of influenza virus PB1 protein binding to viral RNA: two separate regions of the protein contribute to the interaction domain. *J. Virol.* **73**, 631–637 (1999).
- Li, M.L., Rao, P. & Krug, R.M. The active sites of the influenza cap-dependent endonuclease are on different polymerase subunits. *EMBO J.* **20**, 2078–2086 (2001).
- Honda, A., Mizumoto, K. & Ishihama, A. Two separate sequences of PB2 subunit constitute the RNA cap-binding site of influenza virus RNA polymerase. *Genes Cells* **4**, 475–485 (1999).
- Fechter, P. *et al.* Two aromatic residues in the PB2 subunit of influenza A RNA polymerase are crucial for cap binding. *J. Biol. Chem.* **278**, 20381–20388 (2003).
- Nath, S.T. & Nayak, D.P. 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.* **10**, 4139–4145 (1990).
- Mukaigawa, J. & Nayak, D.P. Two signals mediate nuclear localization of influenza virus (A/WSN/33) polymerase basic protein 2. *J. Virol.* **65**, 245–253 (1991).
- Nieto, A. *et al.* Nuclear transport of influenza virus polymerase PA protein. *Virus Res.* **24**, 65–75 (1992).
- Fodor, E. & Smith, M. The PA subunit is required for efficient nuclear accumulation of the PB1 subunit of the influenza A virus RNA polymerase complex. *J. Virol.* **78**, 9144–9153 (2004).
- Deng, T. *et al.* Role of ran binding protein 5 in nuclear import and assembly of the influenza virus RNA polymerase complex. *J. Virol.* **80**, 11911–11919 (2006).
- Naito, T., Momose, F., Kawaguchi, A. & Nagata, K. Involvement of hsp90 in assembly and nuclear import of influenza virus RNA polymerase subunits. *J. Virol.* **81**, 1339–1349 (2007).
- Russell, C.J. & Webster, R.G. The genesis of a pandemic influenza virus. *Cell* **123**, 368–371 (2005).
- Li, Z. *et al.* Molecular basis of replication of duck H5N1 influenza viruses in a mammalian mouse model. *J. Virol.* **79**, 12058–12064 (2005).
- Taubenberger, J.K. *et al.* Characterization of the 1918 influenza virus polymerase genes. *Nature* **437**, 889–893 (2005).
- Gabriel, G. *et al.* The viral polymerase mediates adaptation of an avian influenza virus to a mammalian host. *Proc. Natl. Acad. Sci. USA* **102**, 18590–18595 (2005).
- Salomon, R. *et al.* The polymerase complex genes contribute to the high virulence of the human H5N1 influenza virus isolate A/Vietnam/1203/04. *J. Exp. Med.* **203**, 689–697 (2006).
- Martin-Benito, J. *et al.* Three-dimensional reconstruction of a recombinant influenza virus ribonucleoprotein particle. *EMBO Rep.* **2**, 313–317 (2001).
- Area, E. *et al.* 3D structure of the influenza virus polymerase complex: localization of subunit domains. *Proc. Natl. Acad. Sci. USA* **101**, 308–313 (2004).
- Ostermeier, M. & Lutz, S. The creation of ITCHY hybrid protein libraries. *Methods Mol. Biol.* **231**, 129–141 (2003).
- Beckett, D., Kovaleva, E. & Schatz, P.J. A minimal peptide substrate in biotin holoenzyme synthetase-catalyzed biotinylation. *Protein Sci.* **8**, 921–929 (1999).
- Holm, L. & Sander, C. Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* **233**, 123–138 (1993).
- Peralles, B., de la Luna, S., Palacios, I. & Ortin, J. Mutational analysis identifies functional domains in the influenza A virus PB2 polymerase subunit. *J. Virol.* **70**, 1678–1686 (1996).
- Fontes, M.R., Teh, T., Jans, D., Brinkworth, R.I. & Kobe, B. Structural basis for the specificity of bipartite nuclear localization sequence binding by importin- $\alpha$ . *J. Biol. Chem.* **278**, 27981–27987 (2003).
- Zacksenhaus, E., Bremner, R., Phillips, R.A. & Gallie, B.L. A bipartite nuclear localization signal in the retinoblastoma gene product and its importance for biological activity. *Mol. Cell. Biol.* **13**, 4588–4599 (1993).
- Conti, E. & Kuriyan, J. Crystallographic analysis of the specific yet versatile recognition of distinct nuclear localization signals by karyopherin  $\alpha$ . *Structure* **8**, 329–338 (2000).
- de Jong, M.D. *et al.* Fatal outcome of human influenza A (H5N1) is associated with high viral load and hypercytokinemia. *Nat. Med.* **12**, 1203–1207 (2006).
- de la Luna, S., Martinez, C. & Ortin, J. Molecular cloning and sequencing of influenza virus A/Victoria/3/75 polymerase genes: sequence evolution and prediction of possible functional domains. *Virus Res.* **13**, 143–155 (1989).
- Brunger, A.T. *et al.* Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **54**, 905–921 (1998).
- Kabsch, W. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Cryst.* **26**, 795–800 (1993).
- Terwilliger, T.C. Maximum-likelihood density modification. *Acta Crystallogr. D Biol. Crystallogr.* **56**, 965–972 (2000).
- 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 (1997).

**Influenza virus Nucleoprotein  
oligomerisation state correlates with RNA  
binding affinity.**

Boulo, S., Lotteau, V., Ruigrok, R.W. and Baudin, F.

Soumis.

# **Influenza virus Nucleoprotein oligomerisation state correlates with RNA binding affinity**

Sébastien Boulo<sup>1,2</sup>, Vincent Lotteau<sup>3,4,5</sup>, Rob.W.H. Ruigrok<sup>2</sup> and Florence Baudin<sup>1,2\*</sup>

<sup>1</sup>European Molecular Biology Laboratory (EMBL), Meyerhofstrasse 1, 69117 Heidelberg, Germany.

<sup>2</sup>Université Joseph Fourier, UJF-EMBL-CNRS UMR 5233, Unit of Virus Host-Cell Interactions, 6 rue Jules Horowitz, 38042 Grenoble Cedex 9, France.

<sup>3</sup>Inserm, U851, Lyon, France

<sup>4</sup>Université de Lyon, IFR128 BioSciences Lyon-Gerland, Lyon, France

<sup>5</sup>Hospices civils de Lyon, Hôpital de la croix rousse, Laboratoire de virologie, Lyon, France

\* Corresponding author. Mailing address: EMBL, Meyerhofstrasse 1, D-69117 Heidelberg, Germany.

Tel: +49-6221-387-8483. Fax: +49-6221-387-8518. Email: [baudin@embl.de](mailto:baudin@embl.de)

Total words : 1448

## **Abstract**

Nucleoprotein (NP) of influenza virus binds to the viral RNA and is essential for viral replication and transcription. We have studied the polymerization state and RNA binding activities of RNA-free NP isolated from virus, recombinant NP and a monomeric, mutant recombinant NP. Viral NP forms more higher-order polymers than recombinant NP that only forms small oligomers (trimers) whereas mutant NP was indeed found to be monomeric. Viral NP binds to panhandle RNA with a slightly higher affinity than recombinant NP whereas monomeric NP bound with even lower affinity. Thus, the affinity of NP for RNA correlates with its polymerization state.

The 8 viral RNA segments of influenza virus are covered by the viral nucleoprotein (NP) with a stoichiometry of one NP protomer for 24 nucleotides (4, 14) and each ribonucleoprotein complex (RNP) has a trimeric RNA-dependent RNA-polymerase complex attached. Viral replication is not possible without this nucleoprotein, i.e. the nucleoprotein-vRNA complex is the template for the viral polymerase (3). Influenza A virus nucleoprotein (NP) is 498 amino acids long (56 kDa) and is positively charged. NP isolated from virus and free from RNA is able to self-associate to form large oligomeric complexes. The polymerization process is temperature and concentration dependent with non-specific, large aggregates forming at low temperatures and structures that resemble the viral RNP complexes after incubation at 37°C. At low protein concentrations small oligomeric structures are observed that go from trimers to closed ring structures containing around 6-11 NP protomers (15). As expected from its role in the formation of the RNPs, NP has been shown to possess a high affinity for ssRNA with no sequence specificity (2, 5, 10, 16, 18). In viral RNP or in a reconstituted NP-RNA complex, RNA binds to NP through its phosphate-sugar backbone (2). The interaction between NP and RNA has been shown to be cooperative (18).

The atomic structure of recombinant NP free of RNA was recently determined (19). In solution this protein forms a mixture of monomers and small oligomers, mainly trimers and the protein crystallized as a trimer. The oligomerisation state of recombinant NP appears lower than that of NP isolated from virus (compare figures in [19] and [15]). The structure shows three domains, a helical head domain, a helical body domain and a C-terminal tail domain that is involved in domain exchange. The putative RNA-binding site is located between the head and body domains, opposite to the tail domain involved

in NP-NP contacts. A large number of basic residues present in this region could constitute the RNA binding platform. The RNA would be largely exposed on the external surface of nucleoprotein molecules in the RNP; consistent with the fact that influenza virus RNA covered with nucleoprotein is sensitive to RNases (2, 6).

In this paper we compare the polymerization state and RNA binding affinities of three NP preparations: 1: NP<sub>v</sub>; viral NP isolated from influenza A/PR/8/34 virus (as described in reference 2). 2: NP<sub>rec</sub>; recombinant NP (A/PR/8/34 strain) with a cleavable His-Tag at its N-terminal end. After purification this His-Tag was removed with TEV protease. 3: NP<sub>mono</sub>; a recombinant mutant of NP (R416A), constructed by changing the arginine codon AGA for the alanine codon GCA. Purified protein was at least 95% pure according to Coomassie SDS-PAGE (not shown). The R416A mutation is located in the tail domain and destroys a salt bridge with E339 in a neighboring subunit involved in NP-oligomerisation. This mutant was previously shown to be monomeric (19). In our hands NP<sub>mono</sub> eluted from a High load Superdex 200 (Pharmacia) with an elution volume corresponding to a protein of 50 Kd. Sedimentation velocity experiments on NP<sub>mono</sub> (Fig. 1) resulted in a molecular weight of 55 kDa, indicating that it indeed behaves as a monomer in solution. This experiment was done at 20°C in a Beckman Optima XL-A centrifuge fitted with a four-hole AN-60 rotor. The sedimentation coefficient and molecular weight distributions were analyzed by the C(s) method implemented in the Sedfit software package (17). Buffer density and viscosity corrections were made according to the data published by (11).

Subsequently the three NP samples were prepared for negative stain electron microscopy by incubating at 37°C for 10 minutes and staining with a freshly prepared 2%



uranyl-acetate solution. Images were taken with a Morgagni FEI electron microscope at 100 kV with a magnification of 71,000 times. As we have shown before (15), viral NP is highly polymerized, making rings and higher order polymers resembling viral RNPs (Fig. 2 NPv) whereas only very few small oligomers are formed. Recombinant NP also makes oligomers but smaller than those produced by viral NP, with many small oligomeric forms in the background (Fig. 2 NPrec). The R416A mutant NP only produced monomers (Fig. 2 NPmono), as suggested by the sedimentation experiment (Fig. 1).

We then measured the affinity of the three NP samples for panhandle RNA. This RNA comprises the conserved 3' and 5' ends of influenza A vRNA connected by a sequence derived from the NP gene as described in (2). The affinities were tested by filter binding experiments in which a very small amount of radioactively labelled RNA is mixed with increasing concentrations (from  $10^{-9}$  M to  $10^{-4}$  M) of protein. The mixture is then filtered through presoaked nitrocellulose filter as described in (2). RNA complexed with NP was detected by Cerenkov counting in scintillation coulters. Since the amount of RNA used was extremely small (0.03 pmol), the NP concentration at 50% retention corresponds to the apparent dissociation constant  $K_d$ . The results are shown in Fig. 3. Viral NP binds to panhandle RNA with a  $K_d$  of 40 nM comparable to the affinity of 37 nM published in (2). Recombinant NP bound with a slightly lower affinity ( $K_d = 100$  nM) whereas monomeric mutant R416A presented a  $K_d$  of 3  $\mu$ M. There exists an inverse correlation between the polymerization state and the affinity to RNA for the three NP samples. The RNA binding curves are sigmoidal, indicating that RNA binding is cooperative. The curves for binding by NPv and NPrec are steeper than the curve for NPmono, suggesting the absence of a cooperative effect for the monomeric protein (9).

Elton *et al.* (1999) performed UV cross-linking experiments to determine RNA binding by NP mutants. The R416A mutant did not bind RNA. Here we show that RNA binding does occur even if the affinity is low. This was to be expected from the crystal structure since the mutation does not change the putative RNA binding platform. Recently, Ng *et al.* (2008) defined the RNA binding platform on influenza H5N1 by mutation and surface plasma resonance experiments. Based on the EM model of mini RNPs (12) they proposed a continuous binding surface on the NP polymer. Our results presented here confirm that polymeric NP binds RNA stronger than monomeric NP. One of the reasons for this may be found in the crystal structures of recombinant, circular rabies and Vesicular Stomatitis Virus nucleocapsids (1, 8). In these structures nucleotide base stacking occurs between nucleotides bound to adjacent nucleoprotein molecules. Such stacking is impossible with monomeric nucleoprotein and must contribute to the affinity constant of the protein for the RNA.

## **Acknowledgments**

We thank Dr. Vladimir Rybin (EMBL, Heidelberg) for analytical ultracentrifugation experiments and Claire Blachier-Batisse for electron microscopy. This work was funded by the EU FLUPOL contract (SP5B-CT-2007-044263) and the French ANR FLU INTERPOL contract (ANR-06-MIME-014-02). S.B. is supported by Conseil Régional Rhône Alpes (cluster 10 infectiologie).

## **Figure legends**

**Figure 1. Characterization of monomeric nucleoprotein.** Analytical ultracentrifugation results for the R416A NP mutant. The sedimentation coefficient distribution derived from the sedimentation velocity profile using SEDFIT. The protein concentration was 0.8 mg/ml.

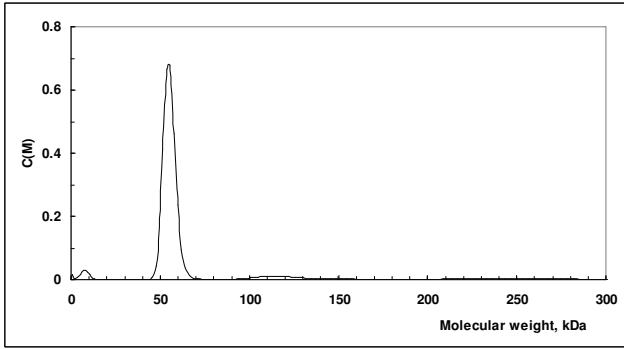
**Figure 2. Electron micrographs of negatively stained influenza virus nucleoprotein** Top, NP purified from the virus. Middle, recombinant NP (NPrec). Bottom, recombinant NP mutant R416A (NPmono). Negative staining was done with 2% uranyl acetate.

**Figure 3. RNA-binding affinity of influenza nucleoprotein.** Titration curves from filter binding assays of panhandle RNA with NPv (—), NPrec (----) and NP R416A mutant (NPmono) (····). The mid-points of the curves represent the K<sub>d</sub> values. The bars represent the standard deviation of the results from 4 experiments.

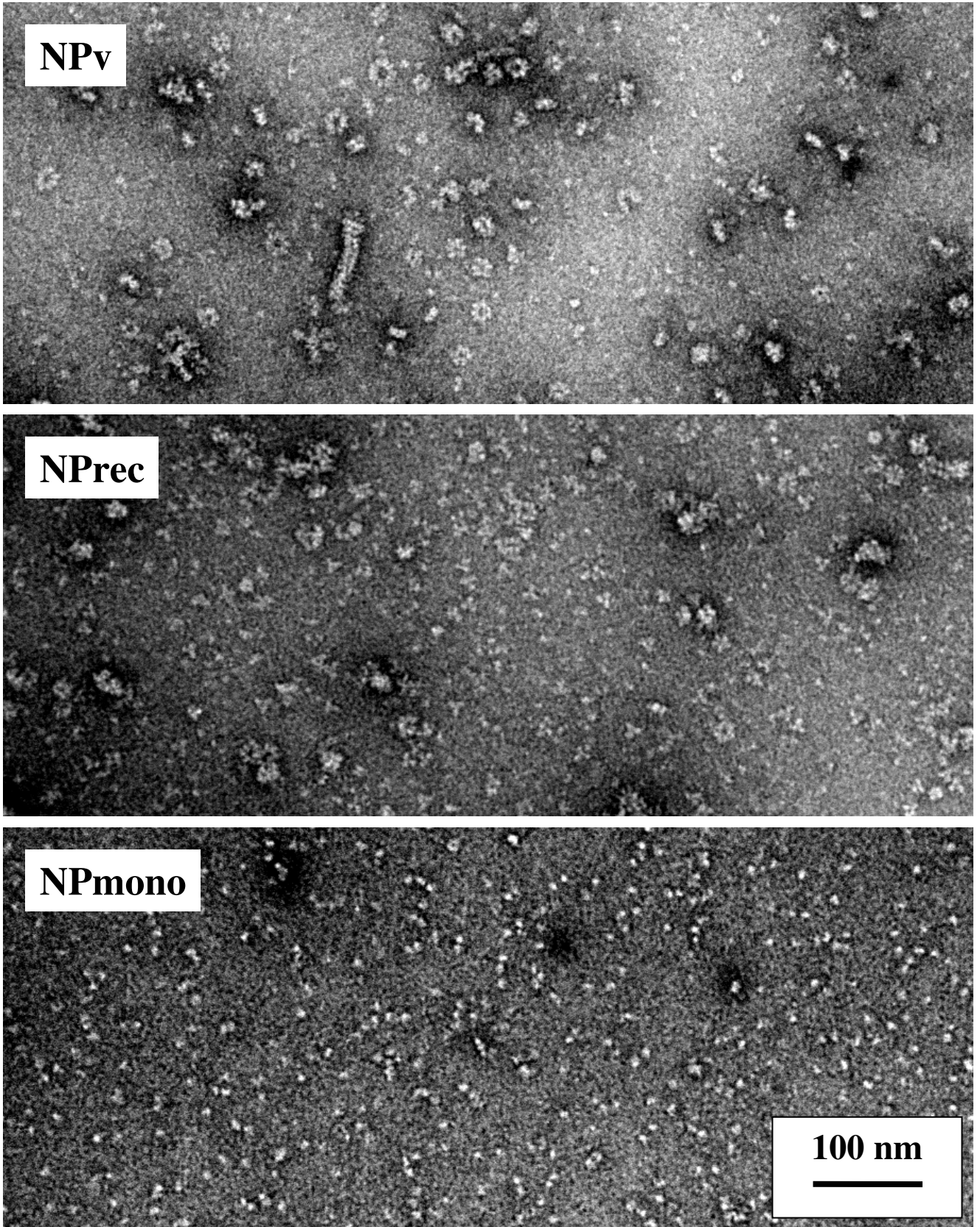
## References

1. **Albertini, A. A., A. K. Wernimont, T. Muziol, R. B. Ravelli, C. R. Clapier, G. Schoehn, W. Weissenhorn, and R. W. Ruigrok.** 2006. Crystal structure of the rabies virus nucleoprotein-RNA complex. *Science* **313**:360-3.
2. **Baudin, F., C. Bach, S. Cusack, and R. W. Ruigrok.** 1994. Structure of influenza virus RNP. I. Influenza virus nucleoprotein melts secondary structure in panhandle RNA and exposes the bases to the solvent. *Embo J* **13**:3158-65.
3. **Bishop, D. H., J. F. Obijeski, and R. W. Simpson.** 1971. Transcription of the influenza ribonucleic acid genome by a virion polymerase. I. Optimal conditions for in vitro activity of the ribonucleic acid-dependent ribonucleic acid polymerase. *J Virol* **8**:66-73.
4. **Compans, R. W., J. Content, and P. H. Duesberg.** 1972. Structure of the ribonucleoprotein of influenza virus. *J Virol* **10**:795-800.
5. **Digard, P., D. Elton, K. Bishop, E. Medcalf, A. Weeds, and B. Pope.** 1999. Modulation of nuclear localization of the influenza virus nucleoprotein through

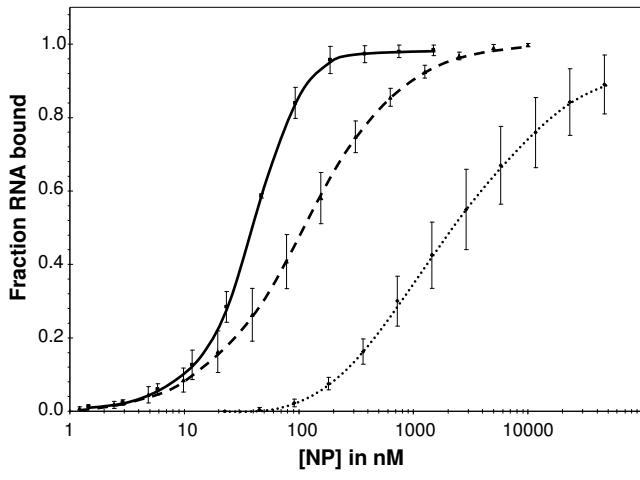
- interaction with actin filaments. *J Virol* **73**:2222-31.
6. **Duesberg, P. H.** 1969. Distinct subunits of the ribonucleoprotein of influenza virus. *J Mol Biol* **42**:485-99.
  7. **Elton, D., L. Medcalf, K. Bishop, D. Harrison, and P. Digard.** 1999. Identification of amino acid residues of influenza virus nucleoprotein essential for RNA binding. *J Virol* **73**:7357-67.
  8. **Green, T. J., X. Zhang, G. W. Wertz, and M. Luo.** 2006. Structure of the vesicular stomatitis virus nucleoprotein-RNA complex. *Science* **313**:357-60.
  9. **Henriet, S., D. Richer, S. Bernacchi, E. Decroly, R. Vigne, B. Ehresmann, C. Ehresmann, J. C. Paillart, and R. Marquet.** 2005. Cooperative and specific binding of Vif to the 5' region of HIV-1 genomic RNA. *J Mol Biol* **354**:55-72.
  10. **Kingsbury, D. W., I. M. Jones, and K. G. Murti.** 1987. Assembly of influenza ribonucleoprotein in vitro using recombinant nucleoprotein. *Virology* **156**:396-403.
  11. **Laue, T. M., B. D. Shah, T. M. Ridgeway, and S. L. Pelletier.** 1992. In *Analytical Ultracentrifugation in Biochemistry and Polymer Science*. The Royal Society of Chemistry, (Cambridge Harding, S. E., Rowe, A., J., and Horton, J. C., eds):90-125.
  12. **Martin-Benito, J., E. Area, J. Ortega, O. Llorca, J.M. Valpuesta, J.L. Carrascosa, and J. Ortin.** 2001. Three-dimensional reconstruction of a recombinant influenza virus ribonucleoprotein particle. *EMBO Rep* **2**:313-7.
  13. **Ng, A. K., H. Zhang, K. Tan, Z. Li, J. H. Liu, P. K. Chan, S. M. Li, W. Y. Chan, S. W. Au, A. Joachimiak, T. Walz, J. H. Wang, and P. C. Shaw.** 2008. Structure of the influenza virus A H5N1 nucleoprotein: implications for RNA binding, oligomerization, and vaccine design. *Faseb J*.
  14. **Ortega, J., J. Martin-Benito, T. Zurcher, J. M. Valpuesta, J. L. Carrascosa, and J. Ortin.** 2000. Ultrastructural and functional analyses of recombinant influenza virus ribonucleoproteins suggest dimerization of nucleoprotein during virus amplification. *J Virol* **74**:156-63.
  15. **Ruigrok, R. W., and F. Baudin.** 1995. Structure of influenza virus ribonucleoprotein particles. II. Purified RNA-free influenza virus ribonucleoprotein forms structures that are indistinguishable from the intact influenza virus ribonucleoprotein particles. *J Gen Virol* **76 ( Pt 4)**:1009-14.
  16. **Scholtissek, C., and H. Becht.** 1971. Binding of ribonucleic acids to the RNP-antigen protein of influenza viruses. *J Gen Virol* **10**:11-6.
  17. **Schuck, P.** 2000. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys J* **78**:1606-19.
  18. **Yamanaka, K., A. Ishihama, and K. Nagata.** 1990. Reconstitution of influenza virus RNA-nucleoprotein complexes structurally resembling native viral ribonucleoprotein cores. *J Biol Chem* **265**:11151-5.
  19. **Ye, Q., R. M. Krug, and Y. J. Tao.** 2006. The mechanism by which influenza A virus nucleoprotein forms oligomers and binds RNA. *Nature* **444**:1078-82.



**Figure 1**



**Figure 2**



**Figure 3**

**Characterization of Influenza virus Nucleoprotein bound to  
nuclear transporter importin alpha 5.**

Boulo, S., Akarsu, H., Rybin, V., Lotteau, V., Ruigrok, R.W.,  
Müller, C., and Baudin, F.

En préparation.



## **RESUME**

Le virus de la grippe est un virus à ARN négatif de la famille des *Orthomyxoviridae* et représente l'un des rares virus à ARN à se répliquer dans le noyau. Les particules ribonucléoprotéiques (RNP) contiennent l'ARN viral protégé par les nucléoprotéines (NP) et associé à l'ARN polymérase virale (sous-unités PB1, PB2 et PA). Nous avons étudié, par des techniques de biochimie et de biophysique, l'interaction de la nucléoprotéine non seulement avec l'ARN viral mais aussi avec son transporteur nucléaire humain, l'importine  $\alpha 5$ . D'autre part, nous avons résolu la structure cristallographique d'un domaine de la polymérase (PB2) en complexe avec l'importine  $\alpha 5$ . L'ensemble des résultats obtenus nous permet de mieux comprendre les interactions mises en jeu entre protéines virales et protéines de l'hôte et ainsi de comprendre pourquoi certains acides aminés présents dans le virus aviaire augmentent la virulence de la grippe chez l'humain.

---

## **TITLE**

Functional and structural studies on influenza virus nucleoprotein and polymerase complexed with its human import factor.

---

## **ABSTRACT**

The virus of influenza is a negative-sense RNA virus belonging to the family of *Orthomyxoviridae* and represents one of the rare RNA viruses to replicate in the nucleus. The ribonucleoprotein particles (RNP) contain the viral RNA, protected by the nucleoproteins (NP) and associated with the RNA viral polymerase (subunits PB1, PB2 and PA). Using biochemical and biophysical tools, we have studied the interaction of the nucleoprotein not only with the viral RNA but also with its human nuclear transporter importin  $\alpha 5$ . Besides, we have solved the crystallographic structure of a domain of the polymerase (PB2) in complex with importin  $\alpha 5$ . Our results enable us to better understand the interactions between viral proteins and host proteins and thus, to understand why some amino-acids in the avian influenza virus increase the virulence of the virus in humans.

---

## **DISCIPLINE**

Biochimie, Biologie structurale

---

**MOTS-CLES :** Virus de la grippe / nucléoprotéine / polymérase / importine / interactions virus hôte / signal de localisation nucléaire / ARN.

**KEYWORDS:** Influenza virus / nucleoprotein / polymerase / importin / virus-host interactions / NLS / RNA.

---

## **INTITULE ET ADRESSE DU LABORATOIRE :**

Unit of Virus Host Cell Interactions (UMR 5233 UJF-EMBL-CNRS)  
6, rue Jules Horowitz, BP181 - 38042 Grenoble Cedex 9 - France.