



**HAL**  
open science

# Recherche de signaux d’empaquetage spécifiques du génomme des virus influenza A

Marie Gerber

► **To cite this version:**

Marie Gerber. Recherche de signaux d’empaquetage spécifiques du génome des virus influenza A. Virologie. Université de Strasbourg, 2016. Français. NNT : 2016STRAJ062 . tel-01565024

**HAL Id: tel-01565024**

**<https://theses.hal.science/tel-01565024>**

Submitted on 19 Jul 2017

**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.



**Université de Strasbourg**

**École doctorale Vie et Santé (414)**

**Architecture et Réactivité de l'ARN, UPR 9002 du CNRS, Strasbourg**

**Thèse** présentée par

**Marie GERBER**

soutenue le **27 septembre 2016**

pour obtenir le grade de **Docteur de l'Université de Strasbourg**

Discipline : Sciences du Vivant

Spécialité : Aspects moléculaires et cellulaires de la biologie

## **Recherche de signaux d'empaquetage spécifiques du génome des virus influenza A**

**THÈSE dirigée par :**

**Dr Catherine ISEL**

UPR 9002, Strasbourg

**Dr Roland MARQUET**

UPR 9002, Strasbourg

**RAPPORTEURS :**

**Dr Bernard DELMAS**

UR 0892, Jouy-en-Josas

**Dr Daniel MARC**

UMR 1282, Nouzilly

**EXAMINATEURS :**

**Dr Thibaut CREPIN**

UMR 5075, Grenoble

**Pr David GILMER**

UPR 2357, Strasbourg



**Marie GERBER**

## **Recherche de signaux d’empaquetage spécifiques du génome des virus influenza A**

Les huit ARN viraux (ARNv) génomiques des *influenzavirus* de type A sont sélectivement empaquetés dans les virions, vraisemblablement sous la forme d’un complexe supramoléculaire maintenu par des appariements ARN/ARN entre signaux d’empaquetage.

Afin d’améliorer la compréhension des règles qui gouvernent ce mécanisme, nous avons déterminé le réseau d’interactions formé entre les ARNv de la souche modèle de laboratoire A/Puerto Rico/8/34 (H1N1). Nous avons ensuite défini, au nucléotide près et par deux approches *in vitro*, les séquences des ARNv impliquées dans la formation de certaines interactions. Le rôle fonctionnel de deux d’entre elles a été testé en contexte infectieux.

Nous avons également poursuivi l’étude d’une interaction identifiée au laboratoire entre deux ARNv de la souche A/Moscou/10/99 (H3N2) circulante. Enfin, nous avons collaboré avec le Dr L. Brown (Australie) sur l’étude du rôle d’une interaction entre deux ARNv de la souche A/Udorn/307/72 (H3N2).

Mots clés : *Influenzavirus A*, ARN viraux, empaquetage du génome, interaction intermoléculaire ARN/ARN, génétique inverse

## **Identification of specific packaging signals in influenza A viruses genome**

The genome of influenza A viruses comprises eight viral RNAs (vRNAs) likely to be selectively packaged into progeny virions as organized supramolecular complexes where vRNAs are held together by base pairing within the packaging signals.

To better understand the rules governing this mechanism, we investigated the vRNA/vRNA interaction network of the A/Puerto Rico/8/34 (H1N1) strain. We then identified, at the nucleotide level, using two *in vitro* approaches, the vRNA sequences involved in several of the interactions. Two of them were functionally tested in an infectious context.

We also studied an interaction previously identified in the laboratory between two vRNAs belonging to the circulating A/Moscow/10/99 (H3N2) strain. Finally, we collaborated with Dr L. Brown (Australia) in order to assess the role of an interaction between two vRNAs of the A/Udorn/307/72 (H3N2) strain.

Keywords : influenza A virus, viral RNA, genome packaging, intermolecular RNA/RNA interaction, reverse genetics



# TABLE DES MATIÈRES

<b>REMERCIEMENTS</b> .....	<b>1</b>
<b>ABRÉVIATIONS</b> .....	<b>3</b>
<b>INTRODUCTION</b> .....	<b>7</b>
A. LES <i>INFLUENZAVIRUS A</i> .....	7
I. Généralités .....	7
II. Morphologie et structure des particules virales .....	8
III. Classification et nomenclature des <i>Influenzavirus</i> .....	9
IV. Génome des <i>Influenzavirus A</i> .....	10
1. Organisation génétique des ARNv .....	10
2. Protéines codées par le génome viral .....	13
V. Cycle viral des <i>Influenzavirus A</i> .....	14
1. Attachement du virus à la surface de la cellule cible.....	15
2. Entrée du virus, fusion membranaire et libération du génome dans le cytoplasme .....	16
3. Import des RNPv dans le noyau .....	17
4. Transcription et réplication du génome viral.....	18
4.1. La ribonucléoprotéine virale .....	19
a. La protéine NP .....	19
b. Le complexe ARN polymérase.....	20
c. La ribonucléoprotéine virale.....	22
4.2. Transcription du génome et export des ARNm.....	24
4.3. Réplication du génome viral .....	27
4.4. Transition transcription/réplication .....	27
5. Export des RNPv dans le cytoplasme .....	28
6. Assemblage, bourgeonnement et libération des particules virales.....	30
B. ÉVOLUTION DES <i>INFLUENZAVIRUS A</i> , TRANSMISSION DU VIRUS ENTRE DIFFÉRENTS RÉSERVOIRS ET ADAPTATION À L'HÔTE .....	34
I. Mécanismes moléculaires permettant l'évolution des <i>Influenzavirus A</i> .....	34
1. Dérive antigénique.....	34
2. Cassure antigénique.....	34
II. Notion de « barrières d'espèces ».....	35
III. Infections sporadiques de l'Homme par des virus aviaires.....	36
IV. Transmissibilité des virus par les aérosols .....	37
C. VACCINS ET TRAITEMENTS ANTIGRIPPAUX .....	39
I. Vaccins antigrippaux saisonniers.....	39
II. Traitements antigrippaux .....	42
1. Inhibiteurs du canal ionique M2 .....	42
2. Inhibiteurs de la neuraminidase .....	43
D. EMPAQUETAGE DE GÉNOMES SEGMENTÉS ET RÉASSORTIMENT GÉNÉTIQUE, DEUX PROCESSUS ÉTROITEMENT LIÉS.....	44
I. Empaquetage des génomes viraux segmentés.....	44
1. Empaquetage du génome des <i>Influenzavirus A</i> .....	44
1.1. La génétique inverse, un outil déterminant dans l'étude des mécanismes d'empaquetage .....	45

1.2. Article de revue : emballage sélectif du génome des <i>Influenzavirus A</i> .....	46
1.3. Apports récents en faveur d'un mécanisme d'emballage sélectif et impliquant des interactions ARN/ARN.....	66
a. Rôle des régions non codantes dans l'emballage sélectif.....	66
b. Importance de la compatibilité entre signaux d'emballage .....	66
c. Importance des interactions ARN/ARN pour un co-emballage sélectif .....	67
1.4. Conclusions .....	68
2. Emballage du génome d'autres virus à ARN segmenté .....	69
2.1. Le virus de la bursite infectieuse : un exemple d'emballage aléatoire .....	69
2.2. Le virus de l'immunodéficience humaine, les <i>Reoviridae</i> et le bactériophage phi6 : des exemples d'emballage sélectif .....	69
a. Le virus de l'immunodéficience humaine.....	69
b. Emballage du génome des <i>Reoviridae</i> .....	70
i. « Bluetongue » virus .....	70
ii. Rotavirus.....	71
c. Emballage du génome du bactériophage phi6 .....	71
II. Le réassortiment génétique entre souches d' <i>Influenzavirus A</i> .....	72
1. Des événements de réassortiment génétique à l'origine de pandémies .....	72
2. Les conséquences d'un mécanisme d'emballage sélectif du génome sur le réassortiment génétique.....	73
<b>PROBLÉMATIQUE &amp; SUJET DE RECHERCHE.....</b>	<b>77</b>
<b>MATERIELS &amp; MÉTHODES .....</b>	<b>81</b>
A. CLONAGE DES SÉQUENCES DES SEGMENTS VIRAUX D' <i>INFLUENZAVIRUS A</i> HUMAINS ET MUTAGÈSE .....	81
I. Souches virales utilisées.....	81
II. Plasmides utilisés .....	81
1. Plasmide d'expression <i>in vitro</i> pUC2000 .....	81
2. Plasmide d'expression <i>in cellula/in viro</i> pHW2000.....	83
III. Mutagénèse dirigée .....	84
B. TECHNIQUES D'ÉTUDE <i>IN VITRO</i> DES INTERACTIONS ENTRE LES ARNV PAR RETARD DE MIGRATION SUR GEL .....	86
I. Digestion des plasmides pUC2000.....	86
II. Transcription <i>in vitro</i> et purification des ARNV sur colonne de gel filtration.....	86
III. Co-transcription <i>in vitro</i> des ARNV deux à deux et retard de migration sur gel d'agarose .....	87
IV. Co-incubation deux à deux des ARNV purifiés et retard de migration sur gel.....	88
V. Co-incubation deux à deux des ARNV en présence d'oligodésoxyribonucléotides antisens et retard de migration sur gel.....	89
C. CARTOGRAPHIE CHIMIQUE DES ARNV.....	90
I. Modifications chimiques des ARNV au DMS .....	91
II. Détection des modifications par élongation d'amorce par la transcriptase inverse .....	91
1. Marquage radioactif à l'extrémité 5' des oligodésoxyribonucléotides .....	91
2. Extension d'amorce par la transcriptase inverse .....	91
3. Séquençage par la transcriptase inverse (méthode de Sanger) .....	92
4. Analyse par électrophorèse sur gel de polyacrylamide dénaturant .....	92
D. IDENTIFICATION DES NUCLÉOTIDES IMPORTANTS POUR L'ÉTABLISSEMENT DES INTERACTIONS ENTRE ARNV PAR MIMÉ.....	93
I. Préparation de la banque d'ARNV NP.....	93
1. Mutagénèse aléatoire de l'ADN NP .....	93

2. Transcription <i>in vitro</i> de la banque d'ARNv NP et purification sur gel .....	94
II. Séparation des ARNv de la banque par retard de migration sur gel .....	95
III. Préparation de la banque en vue d'un séquençage Illumina.....	95
1. Transcription inverse et quantification par qPCR .....	95
2. Amplification de l'ADNc db par PCR .....	96
3. Fragmentation aléatoire de l'ADNc .....	97
4. « End polishing » : réparation des extrémités des ADNc.....	97
5. Ajout d'une adénosine aux extrémités des ADNc.....	97
6. Ligation des adaptateurs aux extrémités des ADNc .....	97
7. Conversion des adaptateurs en ADN double brin et quantification par qPCR .....	98
8. Amplification finale de la banque .....	98
IV. Séquençage de la banque et analyse des données.....	99
E. PRODUCTION DE PARTICULES VIRALES D' <i>INFLUENZAVIRUS A</i> ET CARACTÉRISATION .....	99
I. Entretien des lignées cellulaires.....	99
II. Production de virus influenza A .....	100
1. Transfection de cellules HEK293T/T17 .....	100
2. Amplification des virus par infection de cellules MDCK .....	102
III. Détermination des caractéristiques des virus produits.....	103
1. Détermination du titre viral .....	103
2. Cinétique de réplication des virus .....	104
IV. Expériences de génétique inverse compétitive.....	104
<b>RÉSULTATS &amp; DISCUSSION.....</b>	<b>107</b>
A. ÉTUDES SUR LA SOUCHE HUMAINE PR8.....	107
I. Établissement du réseau d'interactions entre les ARNv de la souche PR8 par retard de migration sur gel.....	107
1. Expériences de retard de migration sur gel après co-transcription de paires d'ARNv .....	107
2. Expériences de retard de migration sur gel après co-incubation de paires d'ARNv purifiés.....	110
3. Conclusion.....	112
4. Discussion .....	115
II. Identification des régions des ARNv de la souche PR8 impliquées dans certaines interactions ARN/ARN mises en évidence <i>in vitro</i> .....	120
1. Approche « classique » .....	120
1.1. Interaction entre les ARNv PB1 et M.....	120
a. Mutagenèse de délétion, inhibition par des oligonucléotides antisens et bioinformatique ..	120
b. Cartographie chimique en solution sur l'ARNv PB1 .....	125
c. Conclusion .....	128
1.2. Interaction entre les ARNv PA et NP .....	128
1.3. Interaction entre les ARNv PB1 et PA.....	134
1.4. Autres interactions du réseau .....	139
2. Approche MIME <i>in vitro</i> .....	141
2.1. Mise en place du protocole MIME.....	143
a. Préparation de la banque d'ARNv NP.....	143
b. Séparation des ARNv de la banque NP par retard de migration sur gel d'agarose.....	144
c. Préparation de la banque en vue d'un séquençage Illumina .....	146
2.2. Analyse des données de séquençage.....	151
2.3. Interaction entre les ARNv PB2 et NP .....	160
2.4. Conclusion .....	164

3. Discussion .....	166
3.1. Mise en œuvre de méthodes reposant sur des interactions ARN/ARN détectables <i>in vitro</i> .....	166
3.2. Des séquences d'interaction situées tout le long des RNPv .....	167
<b>III. Effet des interactions identifiées <i>in vitro</i> sur la réplication virale et le co-empaquetage des ARNv .....</b>	<b>171</b>
1. Étude de l'interaction entre les ARNv PB1 et M de la souche PR8 .....	171
1.1. Production des virus sauvage et mutants par génétique inverse et amplification des virus sur cellules MDCK.....	171
1.2. Caractérisation des cinétiques de réplication virale sur cellules MDCK .....	173
1.3. Étude de l'impact de l'interaction entre les ARNv PB1 et M sur leur co-empaquetage par génétique inverse compétitive .....	175
1.4. Conclusion .....	179
1.5. Discussion.....	181
2. Étude de l'interaction entre les ARNv PB2 et NP de la souche PR8.....	182
2.1. Production des virus sauvage et mutants par génétique inverse et amplification des virus sur cellules MDCK.....	182
2.2. Caractérisation des cinétiques de réplication virale sur cellules MDCK .....	183
2.3. Étude de l'impact de l'interaction entre les ARNv PB2 et NP sur leur co-empaquetage par génétique inverse compétitive .....	184
2.4. Conclusion .....	185
<b>B. ÉTUDE D'INTERACTIONS ENTRE ARNV D'IVA HUMAINS DE SOUS-TYPE H3N2.....</b>	<b>187</b>
<b><i>I. Interaction entre les ARNv NA et M de la souche Moscou .....</i></b>	<b>187</b>
1. Contexte de l'étude .....	187
2. Impact des mutations <i>trans</i> -compensatoires des ARNv NA et M sur la formation du complexe, <i>in vitro</i> .....	188
3. Impact des mutations <i>trans</i> -compensatoires des ARNv NA et M sur la réplication virale .....	191
3.1. Production des virus sauvage et mutants par génétique inverse et amplification des virus sur cellules MDCK .....	191
3.2. Caractérisation des cinétiques de réplication virale sur cellules MDCK .....	192
a. Étude du couple d'ARNv mc .....	192
b. Étude du couple d'ARNv mc 2 .....	193
4. Conclusion.....	195
<b><i>II. Étude de l'interaction entre les ARNv PB1 et NA de la souche Udorn .....</i></b>	<b>195</b>
1. Contexte de l'étude .....	195
2. Interaction <i>in vitro</i> entre les ARNv PB1 et NA Udorn.....	196
3. Conclusion .....	199
4. Article associé.....	200
<b>CONCLUSIONS &amp; PERSPECTIVES .....</b>	<b>219</b>
A. CONCLUSIONS.....	219
B. PERSPECTIVES.....	221
<b>ANNEXE : SELECTIVE PACKAGING OF THE INFLUENZA A GENOME AND CONSEQUENCES FOR GENETIC REASSORTMENT .....</b>	<b>227</b>
<b>BIBLIOGRAPHIE .....</b>	<b>237</b>

## REMERCIEMENTS

Pour commencer, je tiens à remercier les membres du jury d'avoir accepté d'évaluer mon travail de thèse, ainsi que le Pr. Éric Westhof de m'avoir accueillie au sein de son UPR. J'en profite également pour remercier tous les membres de l'UPR 9002.

Je remercie mes directeurs de thèse Catherine et Roland pour votre encadrement et vos conseils tout au long de ces quatre années. Catherine, merci pour ta bienveillance et ta disponibilité (à Strasbourg, à Paris et même au Canada !).

Un grand merci à tous les membres de l'équipe « Rétrovirus et virus à ARN » d'avoir contribué à cette thèse ! Je remercie particulièrement les « grippeux » : Anne, merci pour ton aide, ton soutien (moral et technique) et ton côté bricolo (notre MacGyver !). Redmond, merci pour ta grande folie ... et tes nombreuses explications !! Maud, merci pour ta gentillesse et tes encouragements pendant la rédaction (et ne décime pas tous les poulets !). Damien, merci pour ton soutien, je te souhaite une thèse pleine de réussite ... et des kg de NP purifiée ! Au milieu de tous les « grippeux » du 433, merci Serena pour tes conseils et ton sourire ! Merci également aux plus jeunes (!!): Camille, je te souhaite plein de bonnes choses pour ton avenir loin de Strasbourg (au soleil ?!). Elodie, courage pour cette dernière année ! Noé, je te souhaite le meilleur dans ta thèse. Enfin, merci aux stagiaires (« grippeuses » bien sûr !), Maïwenn, Zeinab, Angela et Justine.

Je remercie tous les membres du laboratoire « Virologie et pathologie humaine » de Lyon. Votre accueil a été chaleureux à chaque séjour. Je garde un très bon souvenir de mes (nombreuses) venues, Lyon est presque devenu mon deuxième chez-moi ! Je remercie particulièrement Manuel, Vincent et Olivier pour votre disponibilité ; Matthieu, Thomas et Andrès pour vos conseils et votre aide technique, pendant mes séjours et après. Je n'oublie pas non plus toutes les personnes qui ne m'ont pas mises à la porte le soir et bien sûr merci aux personnes qui ont subi mes grosses commandes de cellules !!

Merci Audrey pour tes visites lyonnaises ! Fourvière et les tartelettes aux pralines vont nous manquer ! Je remercie également les amis du master pour nos sympathiques soirées. Je vous souhaite à tous de vous épanouir dans vos carrières scientifiques.

Pour finir, je tiens à remercier ma famille qui m'a soutenue tout au long de mes études (même si mes recherches restent encore autant incompréhensibles pour vous !).

Gaétan, ces quatre dernières années n'ont pas toujours été faciles, mais grâce à ton soutien nous avons surmonté les épreuves. Merci pour tes nombreuses visites à l'IBMC (même si c'était principalement pour me voler du chocolat !!). Je te remercie également pour ta contribution directe dans ma thèse : clonages, figures et corrections du manuscrit ! Et enfin, merci d'avoir été la seule personne à qui j'ai eu l'occasion de parler pendant les longues journées de rédaction !

## ABRÉVIATIONS

A, C, G, T, U : adénine, cytidine, guanine, thymine, uracile  
ATP, CTP, GTP, UTP : adenosine-, cytosine-, guanosine-, uridine 5'-triphosphate

Å : Angström

AcNa : acétate de sodium

ADN : acide désoxyribonucléique

ADNc : ADN complémentaire

**A** AMV : *avian myeloblastosis virus*

APS : ammonium persulfate

ARN : acide ribonucléique

ARNc : ARN complémentaire

ARNm : ARN messenger

ARNt : ARN de transfert

ARNv : ARN viral

AS : acide sialique

---

BET : bromure d'éthidium

**B** BGH : *bovine growth hormone*

BSA : *bovine serum albumin*

BTV : *bluetongue virus*

---

CFA : *cell-free assembly*

Ci, mCi,  $\mu$ Ci : curie, millicurie, microcurie

CMV : cytomégalovirus

**C** cpm : coups par minute

CRM1 : *chromosome region maintenance 1*

Ct : *cycle threshold*

CTD : *C-terminal domain*

---

Da, kDa : Dalton, kiloDalton

db : double brin

**D** DMEM : *Dulbecco's modified Eagle medium*

DMS : diméthyl sulfate

DNase : désoxyribonucléase

DTT : dithiothréitol

---

E. coli : *Escherichia coli*

ECP : effet cytopathogène

**E** EDTA : éthylène diamine tétra-acétate de sodium

EMEM : *Eagle's minimum essential medium*

England : A/Finch/England/2051/91 (H5N2)

ESCRT : *endosomal sorting complex required for transport*

---

FDA : *Food and Drug Administration*

**F** FPLC : *fast performance liquid chromatography*

FRET : *Förster resonance energy transfer*

---

g : unité de gravitation

Gal : galactose

**G** GDP : guanosine diphosphate

GFP : *green fluorescent protein*

---

H	<p>h : heure            HA (ou H) : hémagglutinine            HBSS : <i>Hanks balanced salt solution</i>            HEK : <i>Human Embryonic Kidney</i></p>
I	<p>IBDV : <i>infectious bursal disease virus</i>            ISAV : <i>infectious salmon anemia virus</i>            IV : <i>influenzavirus</i>                IVA : <i>influenzavirus</i> de type A                IVB : <i>influenzavirus</i> de type B                IVC : <i>influenzavirus</i> de type C</p>
K	<p>kb : kilobase</p>
L	<p>l, ml, µl : litre, millilitre, microlitre            LB : Luria Bertani            LNA : <i>locked nucleic acid</i></p>
M	<p>M, mM, µM, nM : molaire, millimolaire, micromolaire, nanomolaire            m, cm, mm, µm, nm : mètre, centimètre, millimètre, micromètre, nanomètre            mc : mutant <i>trans</i>-compensatoire            MCM : <i>minichromosome maintenance</i>            MDCK : <i>Madin Darby canin kidney</i>            MEM : <i>minimum essential media</i>            MgCl<sub>2</sub> : chlorure de magnésium            MGL1 : <i>macrophage galactose-type lectin 1</i>            milliQ : eau désionisée (Millipore)            MIME : <i>mutational interference mapping experiment</i>            min : minute            M-MuLV : <i>Moloney murine leukemia virus</i>            MOI : <i>multiplicity of infection</i>            mol, mmol, µmol, nmol, pmol : mole, millimole, micromole, nanomole, picomole            Moscou : A/Moscow/10/99 (H3N2)            MTOC : <i>microtubule organizing center</i></p>
N	<p>NA (ou N) : neuraminidase            NCR : <i>non coding region</i>            NEP/NS2 : <i>nuclear export protein / non structural 2</i>            NES : <i>nuclear export signal</i>            NLS : <i>nuclear localisation signal</i>            NP : nucléoprotéine            nt : nucléotide            NTP : nucléoside tri-phosphate                dNTP : désoxynucléoside tri-phosphate            NXF1 : <i>nuclear export factor 1</i></p>
O	<p>OMS : Organisation Mondiale de la Santé            ORF : <i>open reading frame</i></p>
P	<p>PA : <i>protein acid</i>            pb : paire de base            PB1 : <i>protein basic 1</i>            PB2 : <i>protein basic 2</i>            PBS : <i>phosphate buffered saline</i></p>

PCR : *polymerase chain reaction*  
qPCR : PCR quantitative  
PFU : *plaque forming unit*  
PNK : polynucléotide kinase  
PolI : ARN polymérase I  
PolII : ARN polymérase II  
PR8 : A/Puerto Rico/8/34 (H1N1)  
p/v : poids/volume

---

R  
Rab : *ras-related in brain*  
RCC1 : *regulator of chromosome condensation 1*  
RFU : *relative fluorescence unit*  
RNase : ribonucléase  
RNasin : *ribonuclease inhibitor*  
RNP : ribonucléoprotéine  
RNPc : RNP complémentaire  
RNPv : RNP virale  
RT : *reverse transcriptase*  
RTB : *reverse transcriptase buffer*

---

S  
sb : simple brin  
SDS : sodium dodécyl sulfate  
sec : seconde  
SEM : *standard error of mean*  
SHAPE : *selective 2'-hydroxyl acylation analyzed by primer extension*  
smFISH : *single-molecule fluorescence in situ hybridization*

---

T  
tat-SF1 : *tat-specific factor 1*  
TB : Tris - borate  
TBE : Tris - Borate - EDTA  
TBM : Tris - Borate - MgCl<sub>2</sub>  
TCID : *median tissue culture infective dose*  
TEMED : tétraméthyléthylènediamine  
Tm : stabilité thermique  
Tris : tris(hydroxyméthyl)aminométhane  
TGIRT : *thermostable group II intron reverse transcriptase*

---

U  
U : unité  
U12 et U13 : séquences uniques de 12 et 13 nucléotides, respectivement  
Udorn : A/Udorn/307/72 (H3N2)  
UTR : *untranslated region*  
UV : ultraviolet

---

V  
V : volt  
VLP : *virus-like particle*  
vol : volume  
v/v : volume/volume

---

W  
W : watt  
WSN : A/WSN/1933 (H1N1)  
wt : *wild type*



# INTRODUCTION

## A. Les *Influenzavirus A*

### I. Généralités

Les virus de la famille des *Orthomyxoviridae* sont caractérisés par un génome segmenté composé de six à huit molécules d'ARN simple brin de polarité négative (ARNv (-)). Les *Orthomyxoviridae* sont capables d'infecter une panoplie d'hôtes, que ce soient les mammifères, les oiseaux ou les tiques. Les pathologies associées sont également très variables, allant de la fièvre à l'encéphalite. À ce jour, six genres ont été référencés dans cette famille : les *Influenzavirus* (IV) de type A (IVA), les *Influenzavirus* de type B (IVB), les *Influenzavirus* de type C (IVC), les *Thogotovirus*, les *Isavirus* et les *Quaranzavirus* [1]. Récemment, un septième genre d'*Influenzavirus*, de type D, a été proposé suite à l'étude de virus isolés chez des porcs et des bovins [2].

Parmi tous les virus appartenant à la famille des *Orthomyxoviridae*, les *Influenzavirus*, plus connus sous le nom de « virus de la grippe », sont les plus documentés. Leur classement en différents genres est basé sur les propriétés antigéniques des protéines les plus abondantes dans les virus, à savoir la nucléoprotéine NP et la protéine de matrice M1. Les IVC, infectant principalement l'Homme et plus rarement le porc, sont responsables de maladies bénignes des voies respiratoires et ne font donc pas l'objet d'une attention particulière. Les IVB infectent essentiellement l'Homme et exceptionnellement les phoques et engendrent des épidémies de grippe saisonnière. Parmi le genre IV, les virus de type A sont les plus étudiés. Les oiseaux aquatiques sauvages sont considérés comme le réservoir naturel des IVA, mais ces virus infectent également un grand nombre de mammifères dont le porc et l'Homme. Chez ce dernier, les IVA se transmettent par voie aérienne (aérosols libérés par la respiration, l'éternuement et la toux) ou par l'intermédiaire d'objets contaminés, tandis que la transmission entre les oiseaux se fait par voie oro-fécale. Les IVA sont à l'origine d'épidémies de gripes saisonnières, mais peuvent également provoquer des pandémies, humaines ou aviaires, souvent associées à une mortalité et une morbidité élevées [3]. En France, la grippe touche chaque année entre 2 et 8 millions de personnes et est à l'origine de 1 500 à 2 000 morts. En 2015, elle a cependant causé 24 000 décès. À

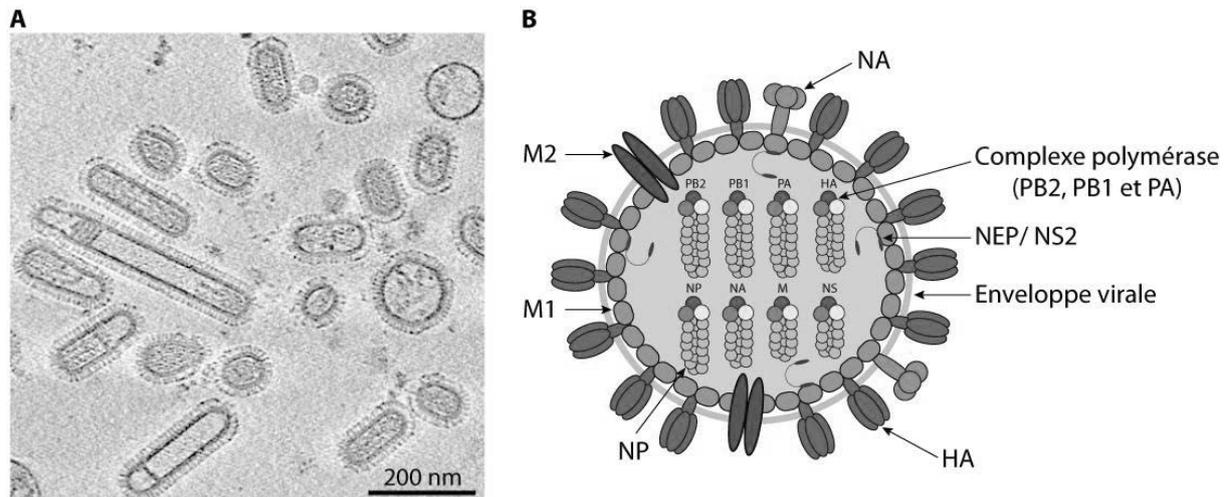
l'échelle mondiale, l'Organisation Mondiale de la Santé (OMS) estime que les infections grippales provoquent entre 3 et 5 millions de cas graves et 250 000 à 500 000 décès par an. Ainsi, ces virus restent, à l'heure actuelle, un problème de santé publique majeur.

La suite du manuscrit se consacrera principalement aux IVA, sur lesquels a porté mon travail de doctorat.

## II. Morphologie et structure des particules virales

Les IVA sont des virus pléomorphiques, adoptant des formes sphériques dont la taille varie entre 80 et 120 nm de diamètre, allongées ou même filamenteuses, dont la longueur est supérieure à 300 nm [4, 5] (figure 1A). Leur enveloppe lipidique, d'origine cellulaire, est acquise lors du bourgeonnement à partir de la cellule hôte. Elle comporte deux types de glycoprotéines antigéniques d'origine virale, l'hémagglutinine (HA) et la neuraminidase (NA), qui forment des spicules (figure 1B). Le nombre et la répartition des spicules varient selon la souche étudiée et la méthode de production des virus. Un ratio de 5:1 entre les protéines HA et NA est souvent retrouvé [6, 7]. Néanmoins, des études de cryo-tomographie électronique sur des particules virales X-31b, qui possèdent les segments HA et NA d'un virus de sous-type H3N2 et les six autres segments d'un virus de sous-type H1N1, produites sur œufs embryonnés, suggèrent que la protéine HA est six à huit fois plus abondante que la protéine NA (environ 300 HA et 40 NA à la surface du virus) [7]. Des résultats similaires ont été observés sur des particules H1N1 produites sur des cellules MDCK (A/Puerto Rico/8/1934 (PR8) et A/WSN/1933 (WSN)) [8]. L'intérieur de l'enveloppe virale est recouverte par les protéines de matrice M1 (figure 1B). Elles interagissent à la fois avec le génome viral et avec les protéines HA, NA et M2 (voir § A.V.6). Ces dernières forment des canaux ioniques qui traversent l'enveloppe lipidique et la matrice. Au sein de virions purifiés, on retrouve également quelques copies de la protéine non-structurale NEP/NS2 en association avec M1.

Le cœur de la particule virale (figure 1B) abrite le génome sous forme de ribonucléoprotéines virales (RNPv). En effet, les huit segments d'ARNv sont associés à de multiples copies de la nucléoprotéine (NP) et au complexe ARN polymérase viral constitué des protéines PB2 (*polymerase basic 2*), PB1 (*polymerase basic 1*) et PA (*polymerase acid*).

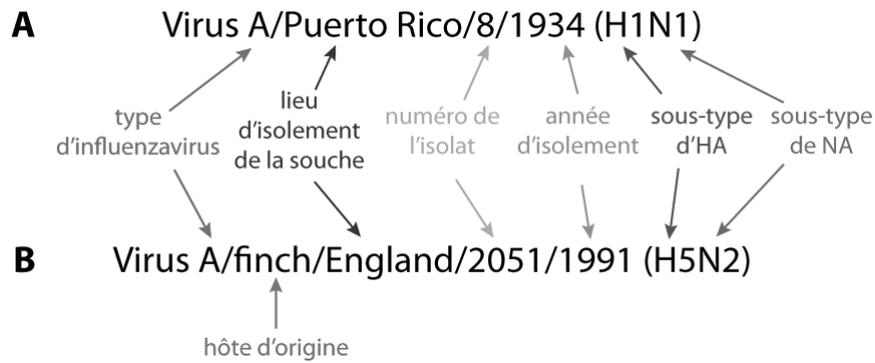


**Figure 1 : Morphologie et structure schématique des IVA. A. Sections de tomogrammes de virions de la souche A/Udorn/72 (H3N2)** (d'après [4]). Les particules virales sont sphériques, allongées ou filamenteuses. **B. Représentation schématique d'une particule virale sphérique d'IVA.** Les glycoprotéines antigéniques HA et NA sont insérées dans l'enveloppe lipidique du virus, dont l'intérieur est recouvert par la protéine de matrice M1. Le canal ionique M2 traverse l'enveloppe virale et la matrice. Le génome viral, constitué de huit ARNv, est présent au centre de la particule virale. Chaque segment est associé à de multiples copies de la nucléoprotéine NP et, à ses extrémités, aux protéines PB2, PB1 et PA formant le complexe ARN polymérase hétérotrimérique.

### III. Classification et nomenclature des *Influenzavirus*

Les IVA sont classés en différents sous-types en fonction de la nature des protéines antigéniques HA et NA. On ne retrouve à la surface d'un virus qu'un seul type de chacune de ces deux protéines. Actuellement, 18 sous-types de protéines HA (H1 à H18) et 11 sous-types de protéines NA (N1 à N11) sont répertoriés [9]. Parmi tous les sous-types d'IVA théoriquement possibles, certaines combinaisons n'ont jamais été détectées et d'autres n'ont été observées que très rarement. Tous les sous-types circulent chez les virus infectant les oiseaux [9], hormis les sous-types H17-H18 et N10-N11 qui ont uniquement été isolés chez des virus infectant les chauves-souris [10, 11]. Chez l'Homme, seuls les sous-types H1, H2, H3, N1 et N2 ont circulé au cours des XX<sup>ème</sup> et XXI<sup>ème</sup> siècles. À ce jour, les pandémies humaines ont uniquement été provoquées par des virus H1N1, H2N2 et H3N2.

Contrairement aux autres virus de la famille des *Orthomyxoviridae*, chaque souche d'IV est répertoriée selon une nomenclature établie par l'OMS. Elle indique le type viral (A, B ou C), l'hôte d'origine (uniquement pour les souches qui n'ont pas été isolées chez l'Homme), le lieu d'isolement de la souche, le numéro de l'isolat et l'année d'isolement (figure 2). S'il s'agit d'un IVA, le sous-type viral est également précisé.



**Figure 2 : Nomenclature des IV.** Exemple d'IV de type A humain (A) et aviaire (B).

## IV. Génome des *Influenzavirus A*

### 1. Organisation génétique des ARNv

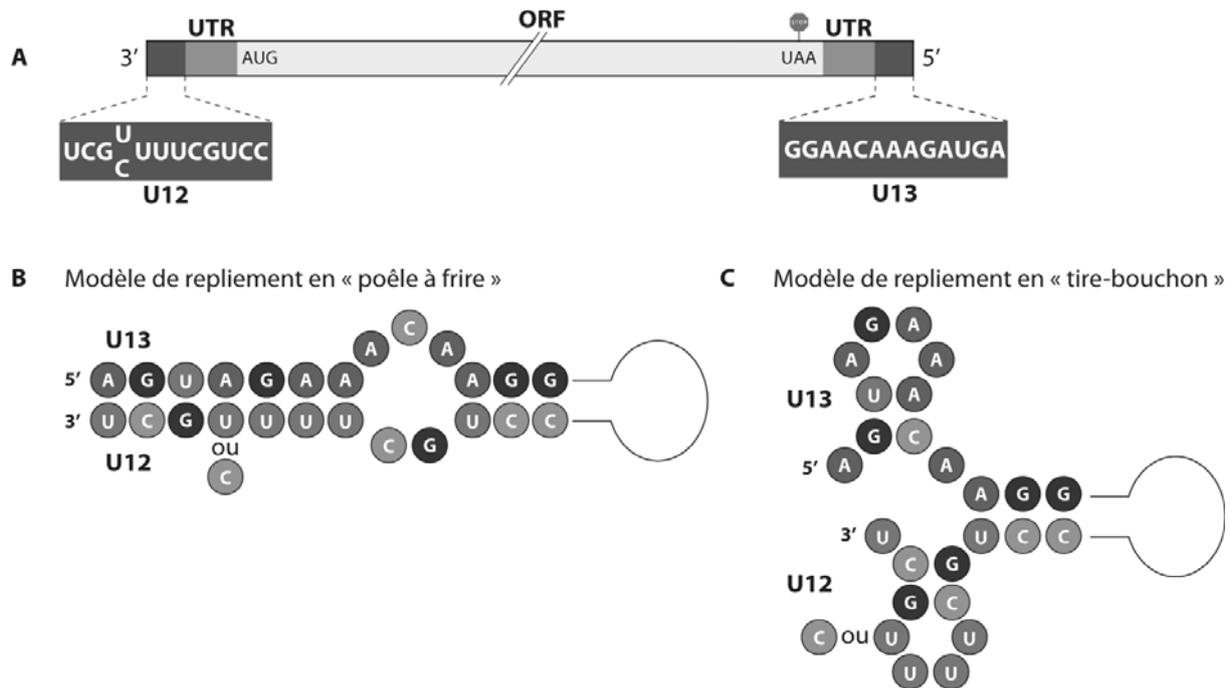
Le génome des IVA est constitué de huit ARNv simple brin de polarité négative, dont la taille est comprise entre 890 et 2341 nucléotides (nt) pour la souche PR8 (H1N1), l'une des souches les plus documentées (tableau 1). Cependant, selon la souche virale, la taille d'un segment donné peut varier de quelques nt. Les ARNv sont numérotés de 1 à 8, selon leur taille décroissante, ou portent le nom de la protéine principale pour laquelle ils codent (tableau 1). D'après la nomenclature conventionnelle, la numérotation des nt des ARNv génomiques de polarité négative se fait de 3' en 5'.

N° du segment	Nom du segment	Longueur du segment (nt)	Nom de la protéine codée	Fonction principale de la protéine	Références pour les protéines auxiliaires
1	PB2	2341	PB2 PB2-S1*	Reconnaissance et liaison de la coiffe Inhibition de la réponse antivirale et de l'activité de la polymérase	[12]
2	PB1	2341	PB1 PB1-F2* PB1-N40*	ARN polymérase ARN-dépendante Activité pro-apoptotique Inconnue	[13], pour revue, voir [14] [15, 16]
3	PA	2233	PA PA-X* PA-N155* PA-N182*	Endonucléase (impliquée dans le "vol de coiffe") Répression de l'expression des gènes transcrits par l'ARN polymérase cellulaire II Probablement impliquée dans la réplication virale Probablement impliquée dans la réplication virale	[17, 18] [19] [19]
4	HA	1778	HA	Liaison au récepteur cellulaire et fusion de la membrane virale et de l'endosome	
5	NP	1565	NP	Liaison aux ARNv et export des RNPv	
6	NA	1413	NA	Activité neuraminidase (clivage entre HA et l'acide sialique)	
7	M	1027	M1 M2 M42*	Protéine de matrice, export des RNPv et rôle dans le bourgeonnement Canal ionique (important pour l'acidification du virion et la libération des RNPv Peut fonctionnellement remplacer M2	[20]
8	NS	890	NS1 NS2-NEP NS3*	Inhibition de la réponse antivirale Export nucléaire des RNPv Apporte un avantage réplicatif	[21]

**Tableau 1 : Segments génomiques des IVA, protéines codées par ces segments et leurs fonctions.** Les protéines auxiliaires sont indiquées par des astérisques.

Bien que les huit segments d'ARNv soient de longueurs différentes, ils possèdent une organisation identique (figure 3A). Une phase de lecture ouverte (ORF, *open reading frame*), en antisens, codant pour une ou plusieurs protéines, est bordée de régions transcrites non-traduites (UTR, *untranslated region*), dont la taille (9-58 nt) est différente pour chaque ARNv. Aux extrémités 3' et 5' de chaque segment se trouvent deux courtes séquences, de 12 nt (U12) et 13 nt (U13), qui sont strictement conservées entre tous les ARNv (intra et inter-souche) [1]. Ces séquences uniques constituent les éléments promoteurs de la transcription et de la réplication et sont indispensables pour initier et réguler les synthèses d'ARN [1]. La complémentarité partielle de ces deux séquences permet leur appariement [22], ce qui confère aux ARNv une structure dite en « poêle à frire » [23–26]. Les extrémités U12 et U13 ainsi appariées sont associées au complexe ARN polymérase ARN-dépendant viral formé des sous-unités PB2, PB1 et PA et le reste de l'ARNv est associé à de multiples copies de la nucléoprotéine NP. Chaque RNPv constitue ainsi une unité de transcription et de réplication indépendante [27–30].

Plusieurs modèles de structures secondaires du promoteur ont été proposés, basés sur des études fonctionnelles. Il semblerait qu'en absence de protéines virales, les séquences promotrices adoptent une structure en « poêle à frire » (*panhandle model*) [29, 31] (figure 3B), alors qu'en présence du complexe ARN polymérase hétérotrimérique, elles adoptent une structure en « tire bouchon » (*corkscrew model*) [32] (figure 3C). Ces deux modèles ont été validés par une récente étude de FRET (*Förster resonance energy transfer*) [26].



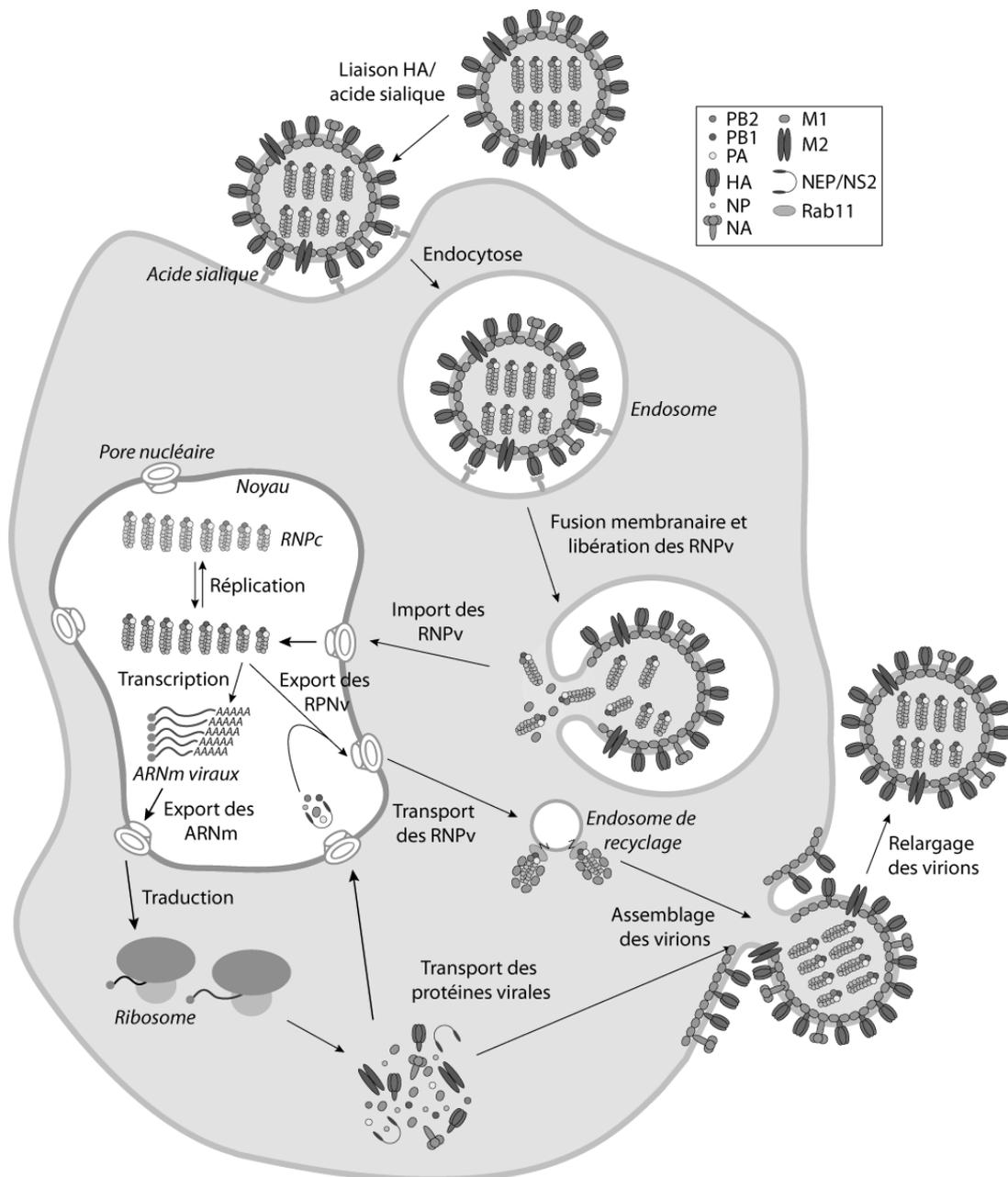
**Figure 3 : Structure des segments génomiques des IVA. A. Organisation génétique commune des huit segments d'ARNv génomiques des IVA.** Chaque gène contient une ou plusieurs séquences codantes centrales (ORF, *open reading frame*), entourées de séquences transcrites non traduites (UTR, *untranslated region*). En 3' et en 5' se trouvent deux séquences uniques, strictement conservées, de 12 nt (U12) et 13 nt (U13). Les ARNv sont représentés selon la nomenclature conventionnelle pour les virus à ARN négatifs, de 3' en 5'. **B & C. Modèles de repliement du promoteur des ARNv des IVA.** En absence de protéines virales (**B**), les séquences U12 et U13 adoptent un repliement en « poêle à frire », tandis qu'en présence du complexe ARN polymérase hétérotrimérique (**C**) le promoteur se replie selon un modèle en « tire-bouchon ».

## 2. Protéines codées par le génome viral

Chaque ARNv constituant le génome des IVA code pour au moins une protéine virale. Dix protéines structurales et non structurales, essentielles, sont communes à toutes les souches. Leurs propriétés, bien définies, sont résumées dans le [tableau 1](#). D'autres protéines, auxiliaires, ont été mises en évidence plus récemment. Elles ne sont pas exprimées par toutes les souches virales et auraient, pour certaines, notamment PB1-F2, des profils d'expression et des rôles différents selon la souche, l'hôte ou le type cellulaire. L'expression de ces protéines s'effectue selon des mécanismes différents (*leaky ribosomal scanning*, réinitiation, décalage du cadre de lecture ouverte ou épissage alternatif) (pour revues, voir [33, 34]).

## V. Cycle viral des *Influenzavirus A*

Contrairement à la majorité des virus à ARN de polarité négative, les étapes de réplication et de transcription des IVA s'effectuent au sein du noyau des cellules infectées. Les principales étapes constituant le cycle viral sont décrites dans la figure 4. Depuis 2013, une carte interactive référence tous les facteurs, cellulaires et viraux, impliqués dans la réplication virale des IVA, ainsi que dans la réponse cellulaire suite à l'infection (<http://www.influenza-x.org/flumap>) [35]. L'ensemble de ces processus fait intervenir près de 1 000 facteurs.

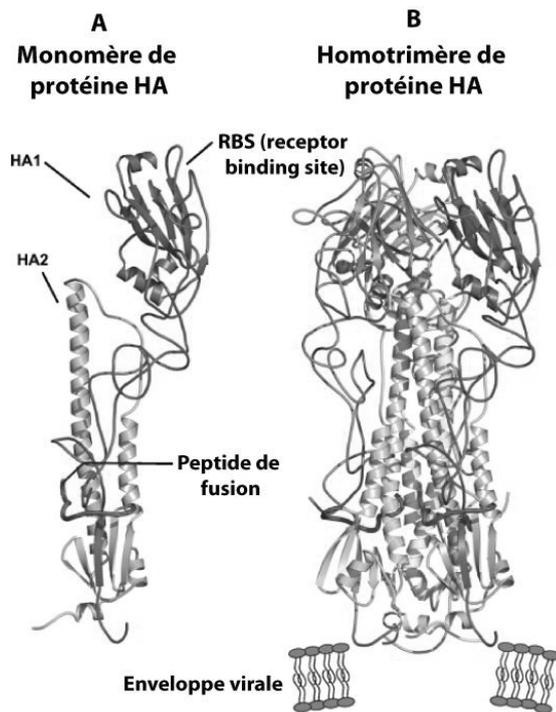


**Figure 4 : Représentation schématique du cycle réplcatif des IVA.** Se rapporter au texte pour les explications.

## 1. Attachement du virus à la surface de la cellule cible

Le cycle répliatif débute par l'attachement des particules virales aux cellules cibles. La protéine HA, présente sous forme homotrimérique à la surface des virions, se lie à des acides N-acétylneuraminiques (acides sialiques) présents à la membrane plasmique des cellules (figure 5). Les virus pourraient également reconnaître d'autres récepteurs. En effet, une étude récente a identifié un récepteur spécifique des IVA, la glycoprotéine MGL1 (*macrophage galactose-type lectin 1*) [36]. Des analyses structurales et fonctionnelles ont également montré que les protéines HA de sous-types H17 et H18, portées par les virus de chauves-souris H17N10 et H18N11, ne se fixent pas à l'acide sialique [11, 37]. Dans ce cas, il semblerait que les protéines HA reconnaissent d'autres glycoprotéines à la surface des cellules [38].

La protéine HA mature est issue du clivage protéolytique d'un précurseur non fonctionnel HA0 et est constituée de deux sous-unités associées par un pont disulfure : HA1 qui contient un domaine RBS (*receptor binding site*) et HA2 possédant un peptide de fusion (voir § A.V.2). Le domaine RBS de chaque monomère est impliqué dans la reconnaissance et l'interaction avec les récepteurs cellulaires de type acide sialique présents au niveau des glycoprotéines ou des glycolipides ancrés dans la membrane cellulaire. Le domaine RBS est directement responsable de la restriction d'hôte. Les protéines HA des virus aviaires se lient préférentiellement à des acides sialiques reliés à un galactose par une liaison de type  $\alpha 2,3$  (SA $\alpha 2,3$ Gal), présents au niveau du tractus digestif, tandis que les virus humains et porcins classiques se lient préférentiellement lorsque la liaison est de type  $\alpha 2,6$  (SA $\alpha 2,6$ Gal). Ces récepteurs sont majoritairement retrouvés chez l'Homme au niveau de l'appareil respiratoire supérieur (nez, pharynx, trachée, bronches). Plusieurs études ont pourtant révélé que des récepteurs de type SA $\alpha 2,3$ Gal sont également présents au niveau de l'appareil respiratoire inférieur (bronchioles et alvéoles) [39, 40]. La localisation de ces récepteurs peut en partie expliquer la faible transmission interhumaine des virus aviaires (H5N1 et H7N7), malgré leur capacité à se répliquer efficacement chez l'Homme. Chez les porcs, considérés comme un « creuset de mélange », les deux types de récepteurs sont présents au niveau de la trachée. Ces animaux sont donc susceptibles d'être infectés à la fois par des virus porcins classiques, mais également par des virus humains et aviaires, ce qui pose incontestablement un problème majeur dans l'émergence de virus pandémiques [41].



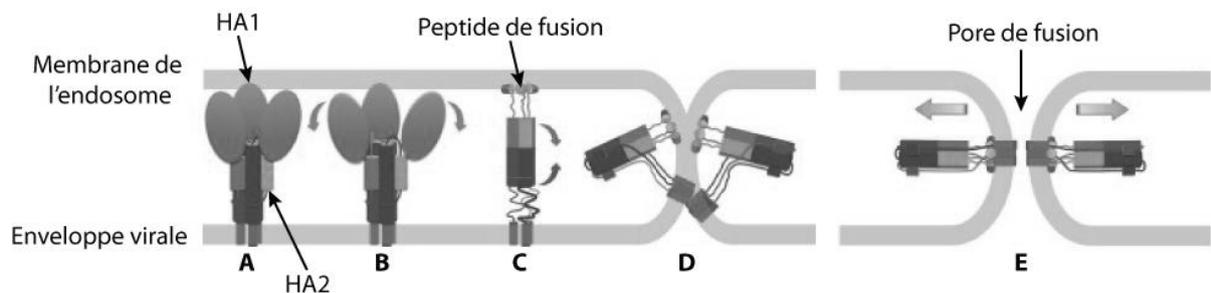
**Figure 5 : Structure tridimensionnelle (A) d'un monomère et (B) d'un homotrimère de la protéine HA des IVA (d'après [42]). Dans sa forme mature, la protéine HA est constituée de deux sous-unités : HA1 qui contient un domaine RBS (*receptor binding site*) et HA2 possédant un peptide de fusion.**

## 2. Entrée du virus, fusion membranaire et libération du génome dans le cytoplasme

L'entrée du virus est réalisée par un mécanisme d'endocytose selon différentes voies : par la voie classique médiée par la clathrine [43], par la voie médiée par la cavéoline [44] ou encore par une voie n'utilisant ni clathrine, ni cavéoline [45]. La voie classique et la voie clathrine/cavéoline-indépendante ne présentent aucune différence d'efficacité vis-à-vis des étapes virales subséquentes [45] et sont utilisées de manière concomitante, même si une légère préférence est observée pour la voie clathrine-dépendante [46]. Les virus filamenteux semblent entrer dans la cellule par un mécanisme de macropinocytose, une forme d'endocytose actine-dépendante initiée par des mouvements à la surface de la membrane qui génèrent de grandes vésicules d'endocytose appelées macropinosomes [47, 48].

Après maturation de l'endosome précoce en endosome tardif, la baisse de pH ( $\approx 5$ ) entraîne un changement conformationnel du canal ionique M2, permettant l'entrée du flux de protons de l'endosome acide dans la particule virale [49]. Les complexes formés entre les RNPv et la protéine de matrice M1 sont alors dissociés. En parallèle, suite à l'acidification, la protéine HA (déjà maturée en HA1 et HA2) subit un changement conformationnel irréversible, rendant accessible le peptide de fusion localisé à l'extrémité N-terminale du

domaine HA2 (figure 6). Celui-ci va ensuite s'orienter en direction de la membrane de l'endosome avec laquelle il va interagir. Un second changement de conformation de la protéine HA rapproche l'enveloppe virale et la membrane de l'endosome et induit ainsi la formation d'un pore de fusion permettant de libérer les RNPv dans le cytoplasme de la cellule hôte.



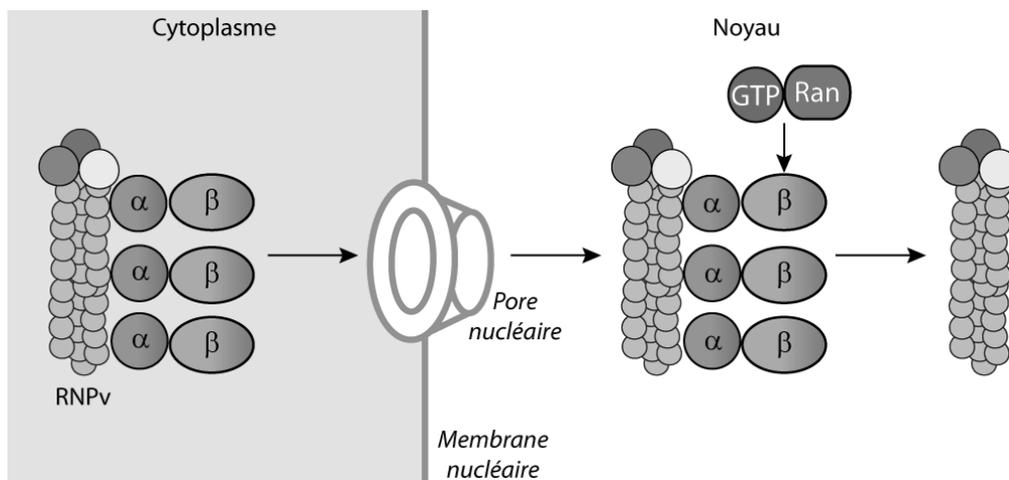
**Figure 6 : Mécanisme de fusion de l'enveloppe virale et de la membrane de l'endosome médié par la protéine virale HA des IVA** (d'après [50]). **A.** Liaison de la protéine HA, *via* le domaine HA1, au récepteur cellulaire de type acide sialique. **B.** L'acidification de l'endosome provoque un changement conformationnel de la protéine HA, permettant l'exposition du peptide de fusion en direction de la membrane de l'endosome. Cette baisse de pH affecte également le domaine HA1, qui reste lié à HA2 (pour plus de clarté, ce domaine n'est plus représenté par la suite). **C.** Le peptide de fusion interagit avec la membrane de l'endosome. **D.** Un changement conformationnel de HA rapproche les deux membranes. **E.** Fusion des membranes et formation d'un pore de fusion.

### 3. Import des RNPv dans le noyau

Suite à la libération du génome dans le cytoplasme, le transport des RNPv s'effectue par diffusion passive jusqu'à la membrane nucléaire et ne fait intervenir ni microtubules, ni filaments (actines ou intermédiaires) [51]. Plusieurs études suggèrent que le complexe associant les huit RNPv pourrait être maintenu dans le cytoplasme et n'être dissocié qu'après l'étape d'import dans le noyau de la cellule hôte [52, 53].

Les quatre protéines liées aux ARNv comportent des signaux de localisation nucléaire (NLS, *nuclear localisation signal*) [54] et sont capables d'entrer dans le noyau, seules (dans le cas de NP et de PB2) ou en complexe (dans le cas de PB1 et PA) [55–57]. Cependant, alors qu'un seul complexe ARN polymérase est lié à chaque RNPv, de multiples protéines NP, possédant plusieurs NLS, sont associées aux ARNv et peuvent promouvoir leur import nucléaire. La séquence d'import non conventionnelle, située en N-terminal [58, 59], semble jouer un rôle prépondérant dans le mécanisme d'import par rapport à la séquence bi-partite conventionnelle localisée entre les résidus 198-216 de la protéine NP [60]. Un troisième NLS,

dont la localisation exacte reste encore à déterminer, pourrait également être impliqué dans l'accumulation nucléaire des RNPv [61]. Le NLS non conventionnel de la protéine NP est reconnu par l'importine  $\alpha$ , qui recrute ensuite l'importine  $\beta$  (figure 7). Celle-ci interagit directement avec les protéines du pore nucléaire, permettant l'import du complexe trimérique. Dans le noyau, l'interaction de la forme activée de la protéine G cellulaire (Ran•GTP) avec l'importine  $\beta$  dissocie le complexe d'import et permet ainsi le relargage des RNPv. L'importine  $\beta$  liée au Ran•GTP est alors exportée dans le cytoplasme et l'hydrolyse du GTP en GDP induit leur dissociation.



**Figure 7 : Modèle pour l'import nucléaire des RNPv.** L'importine  $\alpha$  s'associe à la fois avec le NLS de la protéine NP et avec l'importine  $\beta$ . Le complexe trimérique est ensuite importé dans le noyau *via* la liaison entre l'importine  $\beta$  et les protéines du pore nucléaire. Dans le noyau, l'interaction de la forme activée de la RanGTPase (Ran•GTP) avec l'importine  $\beta$  entraîne la dissociation du complexe et la libération des RNPv.

#### 4. Transcription et réplication du génome viral

Après import, les RNPv diffusent dans le nucléoplasme [62]. La synthèse de nouvelles RNPv nécessite d'une part la transcription du génome viral en ARNm suivie de la synthèse des protéines virales dans le cytoplasme, et d'autre part la réplication du génome par l'intermédiaire d'ARN complémentaires (ARNc). Ces deux mécanismes, bien distincts, où chaque RNPv fonctionne comme une unité de transcription et de réplication indépendante, nécessitent de copier l'ARNv de polarité négative en ARN de polarité positive.

#### 4.1. La ribonucléoprotéine virale

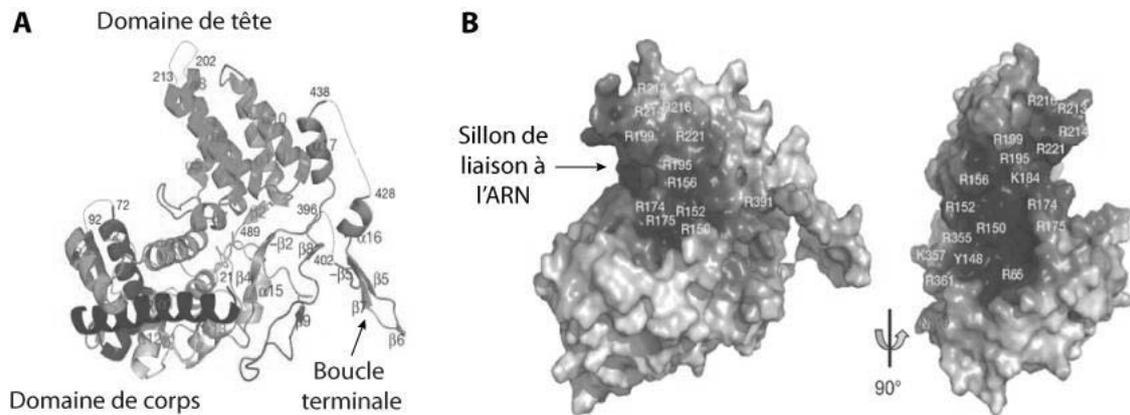
Dans les cellules et dans les virions, chaque ARNV est associé à de multiples copies de la protéine NP et à un complexe ARN polymérase ARN-dépendant, lié au promoteur formé par l'appariement des séquences U12 et U13 [27–30].

##### a. La protéine NP

La protéine NP est une protéine basique de 56 kDa capable de s'oligomériser pour former un squelette protéique hélicoïdal [27, 30, 63–65] stable en absence d'acides nucléiques [64]. L'ARNV s'enroule sur cette architecture protéique *via* son squelette ribose-phosphate à raison d'environ une protéine NP pour 24 nt [66–69], laissant les bases exposées au solvant [24, 70]. La liaison, de forte affinité ( $K_d = 20$  nM), entre la protéine NP et l'ARNV ne présente pas de spécificité de séquence [24, 71–73]. Une étude récente suggère néanmoins que les protéines NP pourraient interagir de manière spécifique avec des séquences localisées en 5' des ARNV [74]. Les structures atomiques à haute résolution des protéines NP d'IVA de sous-type H1N1 [65] et H5N1 [63] sont similaires et révèlent la présence de deux domaines (tête et corps) et d'une boucle terminale flexible (figure 8A).

Au sein des RNPv, les protéines NP interagissent entre elles, avec l'ARNV et avec le complexe ARN polymérase. L'insertion de la boucle terminale d'une protéine NP dans le corps de la protéine NP adjacente permet son oligomérisation [27, 28, 63, 65, 75] et est probablement régulée par la phosphorylation de certains résidus [76–78]. Le domaine de liaison à l'ARNV est localisé dans le premier tiers en N-terminal de la protéine [79, 80], vraisemblablement dans un profond sillon, entre la tête et le corps, qui expose des résidus basiques à la surface de la protéine [63, 65] (figure 8B). D'autres régions tout le long de la protéine sont néanmoins nécessaires pour conférer une liaison de haute affinité [81, 82]. Des études biochimiques ont montré que la protéine NP interagit également avec les sous-unités du complexe ARN polymérase PB1 et PB2 [83, 84], probablement par l'intermédiaire de trois résidus localisés dans sa tête [85]. Des interactions entre la protéine NP et le domaine C-terminal de la protéine PA ont été observées par microscopie électronique mais nécessitent un réarrangement du complexe ARN polymérase, en particulier de la sous-unité PA [27, 30]. Une autre étude suggère que l'extrémité N-terminale de la protéine PA (résidus 1-27) pourrait également interagir avec le domaine de liaison à l'ARN de la protéine NP [86].

Cette interaction pourrait ainsi constituer un « réservoir » de protéines NP afin de protéger le génome viral néo-synthétisé de la dégradation dès sa sortie du site catalytique du complexe ARN polymérase.



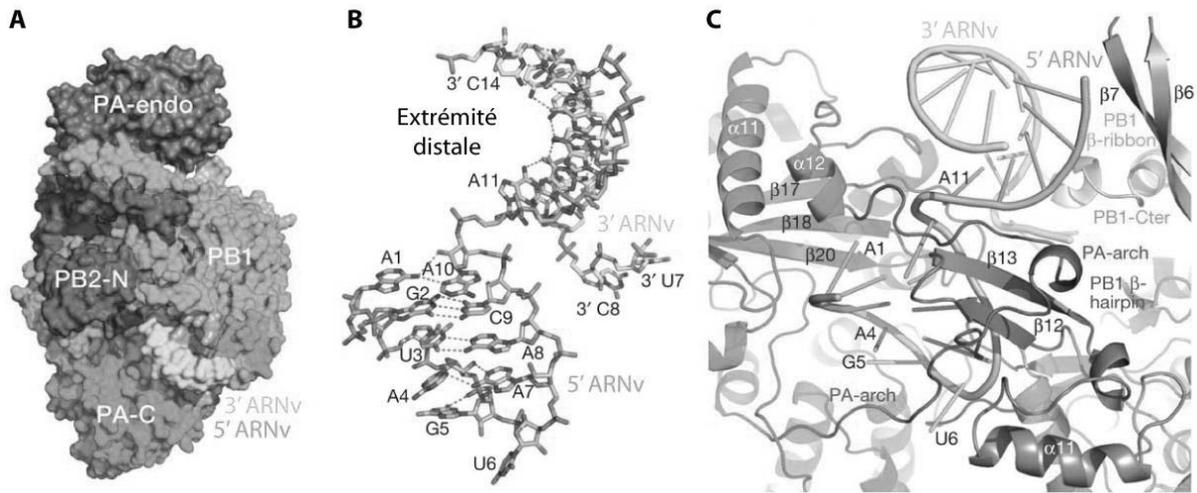
**Figure 8 : Structure de la protéine NP des IVA (adapté de [65]). A. Structure cristallographique de la protéine NP de la souche WSN résolue à 3,2 Å. B. Distribution des potentiels électrostatiques de la protéine NP de la souche WSN. Les potentiels électrostatiques positifs et négatifs sont indiqués en bleu et en rouge, respectivement.**

### b. Le complexe ARN polymérase

Le complexe ARN polymérase ARN-dépendant viral est actif sous forme d'un hétérotrimère de 250 kDa constitué de deux protéines basiques (PB2 et PB1) et d'une protéine acide (PA) (figure 9A) et est responsable de la transcription et de la réplication du génome viral. Jusqu'à récemment, seules des structures à haute résolution de fragments des trois sous-unités étaient disponibles [87–91]. En 2014 et 2015, des études cristallographiques ont conduit à la publication des premières structures du complexe ARN polymérase entier à haute résolution d'un IVA de chauve-souris [92] (figure 9A), d'un IVB humain [93] et d'un IVC humain [94]. Les structures des ARN polymérases obtenues par Pflug *et al.* [92] et Reich *et al.* [93] sont similaires. Le cœur du complexe ARN polymérase est constitué de la sous-unité PB1, du domaine C-terminal de la protéine PA et du domaine N-terminal de la protéine PB2. Des contacts sont également observés entre les sous-unités PB2 et PA. En parallèle de ces travaux, une étude de cryo-microscopie électronique a permis d'obtenir la structure à 4,3 Å du complexe ARN polymérase d'un IVA aviaire H5N1 constitué des protéines PB1 et PA et d'un fragment N-terminal de la protéine PB2 [95].

Le complexe ARN polymérase des IVA adopte des structures différentes en absence [94] ou en présence d'ARNv [92, 93]. Le cœur du complexe est peu impacté lorsque l'ARNv est associé, mais certains domaines des protéines PB2 et PA subissent des réarrangements structuraux. Ainsi, en absence d'ARNv, l'organisation du domaine de liaison à la coiffe et du domaine endonucléase est incompatible avec une activité de vol de coiffe (voir § A.V.4.2) [94], tandis qu'en présence d'ARNv ces deux domaines sont alignés, ce qui est compatible avec une telle activité [92, 93]. Lorsque le complexe est lié à l'extrémité 5' d'un ARNc, le domaine C-terminal de la protéine PB2 subit un réarrangement structural majeur [96]. Le NLS de cette sous-unité s'associe notamment au domaine endonucléase de la sous-unité PA. Le complexe ARN polymérase d'un virus H5N1 [95], présent sous forme de tétramère, pourrait notamment refléter la conformation de l'ARN polymérase lors de la réplication du génome. La fixation de l'ARN (viral ou complémentaire) et probablement d'autres facteurs viraux et cellulaires, induit des changements conformationnels du complexe ARN polymérase, modulant ainsi son activité.

Les structures de l'ARN polymérase d'un IVA [92] et d'un IVB [93], obtenues en présence d'un petit ARN qui mime les séquences terminales 3' et 5' de l'ARNv, ont révélé la conformation du promoteur lié au complexe viral. L'extrémité distale du promoteur forme une structure double-brin, émergeant de l'ARN polymérase, qui implique les nt 11-15 de l'extrémité 5' et les nt 10-14 de l'extrémité 3' (figure 9A). Les nt 1-9 en 3' de l'ARNv se lient à la surface du complexe et interagissent avec des résidus de chacune des trois sous-unités [93]. Les nt 1-10 en 5' de l'ARNv forment une structure en tige-boucle, contenant deux appariements non canoniques A-A (figure 9B), qui s'insère dans une poche formée par les sous-unités PB1 et PA (figure 9C). Cette structure est en adéquation avec le modèle de repliement en « tire-bouchon », hormis les deux appariements non canoniques A-A qui n'avaient pas été prédits par les études précédentes (figure 3C).

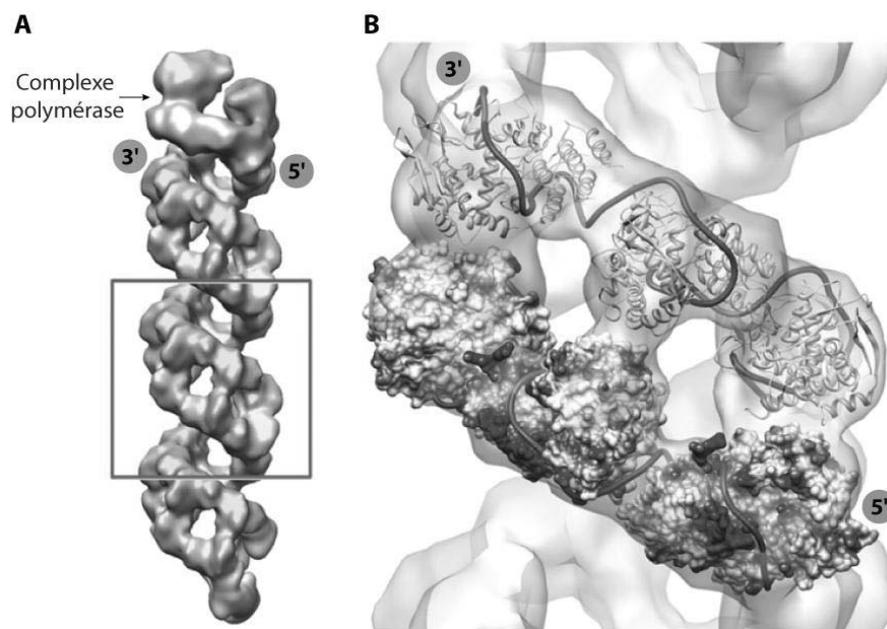


**Figure 9 : Structure de l'ARN polymérase d'un IVA de chauve-souris liée au promoteur de l'ARNv** (adapté de [92]). **A. Structure de l'ARN polymérase en complexe avec le promoteur d'ARNv.** La sous-unité PB1 est représentée en bleu, le domaine N-terminal de la sous-unité PB2 en rouge, les domaines C-terminal et endonucléase de la sous-unité PA en vert. L'extrémité distale du promoteur (vert clair et rose) est à l'interface entre la sous-unité PB1 et le domaine C-terminal de la sous-unité PA. **B. Structure du promoteur d'ARNv.** Les interactions hydrogènes sont représentées en vert et les extrémités 3' et 5' en vert clair et rose, respectivement. **C. Représentation en rubans des domaines des sous-unités PB1 (gris) et PA (violet) qui constituent le site de liaison de la structure en tige-boucle adoptée par l'extrémité 5' du promoteur d'ARNv (rose).** L'extrémité distale du promoteur interagit avec la sous-unité PB1 (orange). L'extrémité 3' du promoteur n'est pas représentée.

### c. La ribonucléoprotéine virale

En microscopie électronique, les RNPv purifiées apparaissent comme des bâtonnets flexibles qui adoptent la structure d'une double-hélice (figure 10A) [66, 97–99]. Leur diamètre est de 13 nm environ et leur longueur varie entre 30 nm et 110 nm [66, 97–99]. En 2009, Coloma *et al.* [28] ont montré, par cryo-microscopie électronique, qu'une RNPv de taille minimale (248 nt), compétente pour la réplication, adopte une forme circulaire avec neuf monomères de protéines NP associées à l'ARNv, dont deux en contact avec le complexe ARN polymérase. Récemment, deux équipes ont reconstruit les structures de RNPv complètes à 20 Å de résolution [27, 30]. Les deux modèles suggèrent que les RNPv adoptent une structure en double hélice avec deux brins antiparallèles maintenus par des interactions entre les protéines NP (figure 10). L'extrémité ouverte des RNPv est liée au complexe ARN polymérase, qui pourrait adopter différentes conformations [27, 92–95] et l'autre extrémité est constituée d'une boucle contenant trois à huit protéines NP [30]. Des différences significatives sont cependant observées dans ces deux modèles, notamment le sens de l'hélice, dû à l'orientation relative des protéines NP des brins opposés, qui s'associent au

niveau de leurs domaines protéiques de tête [30] ou de corps [27], pour former une hélice droite ou gauche, respectivement. Ces différences pourraient être dues à la nature des échantillons, issus dans un cas d'une expression des RNPv après transfection transitoire [30] et dans l'autre de la purification de RNPv natives à partir de virions [27], mais proviennent plus vraisemblablement d'une erreur lors de la construction de l'un de ces modèles. L'ARNv, pour lequel aucune densité électronique n'a été observée, se situerait à la surface du squelette protéique et sa structure ne serait vraisemblablement pas périodique.



**Figure 10 : Structure d'une RNPv d'IVA d'après les données d'Arranz *et al.* (adapté de [100]).**  
**A. Densités électroniques issues d'expériences de cryo-microscopie électronique.** Le complexe ARN polymérase est représenté en gris, les brins antiparallèles de la protéine NP en rouge et bleu et la boucle terminale en jaune. **B. Représentation de l'ARNv matrice (trait rouge) et des monomères de protéines NP à l'intérieur des densités électroniques de la RNPv.** La distribution des potentiels électrostatiques pour les monomères de protéines NP sont représentés dans le brin du bas. Les potentiels électrostatiques positifs et négatifs sont indiqués en bleu et en rouge, respectivement. Dans le brin du haut les monomères de protéines NP sont représentés en rubans.

Des études de cartographie chimique sur les RNPv ont montré que tous les nt sont accessibles aux modifications chimiques : la NP déstructurerait donc les ARNv [24, 70]. À l'inverse, des expériences de cartographie enzymatique suggèrent une persistance des structures secondaires au sein des RNPv [73]. Plusieurs travaux ont montré une stœchiométrie d'environ une protéine NP pour 24 nt [66–69]. Une étude récente de la protéine NP du virus de l'anémie infectieuse du saumon (ISAV, *infectious salmon anemia virus*) a néanmoins révélé que la taille du sillon de la protéine constituant le site de fixation

putatif à l'ARNv ne pourrait accueillir que 12 nt [101], laissant ainsi à nu certaines régions de l'ARNv. Ces résultats sont corroborés par ceux de York *et al.* [102] qui ont purifié des RNPv et des RNPc par affinité, grâce à l'introduction d'un aptamère structuré dans l'ARNv NA, suggérant que des éléments de structure secondaire peuvent émerger des ARNv.

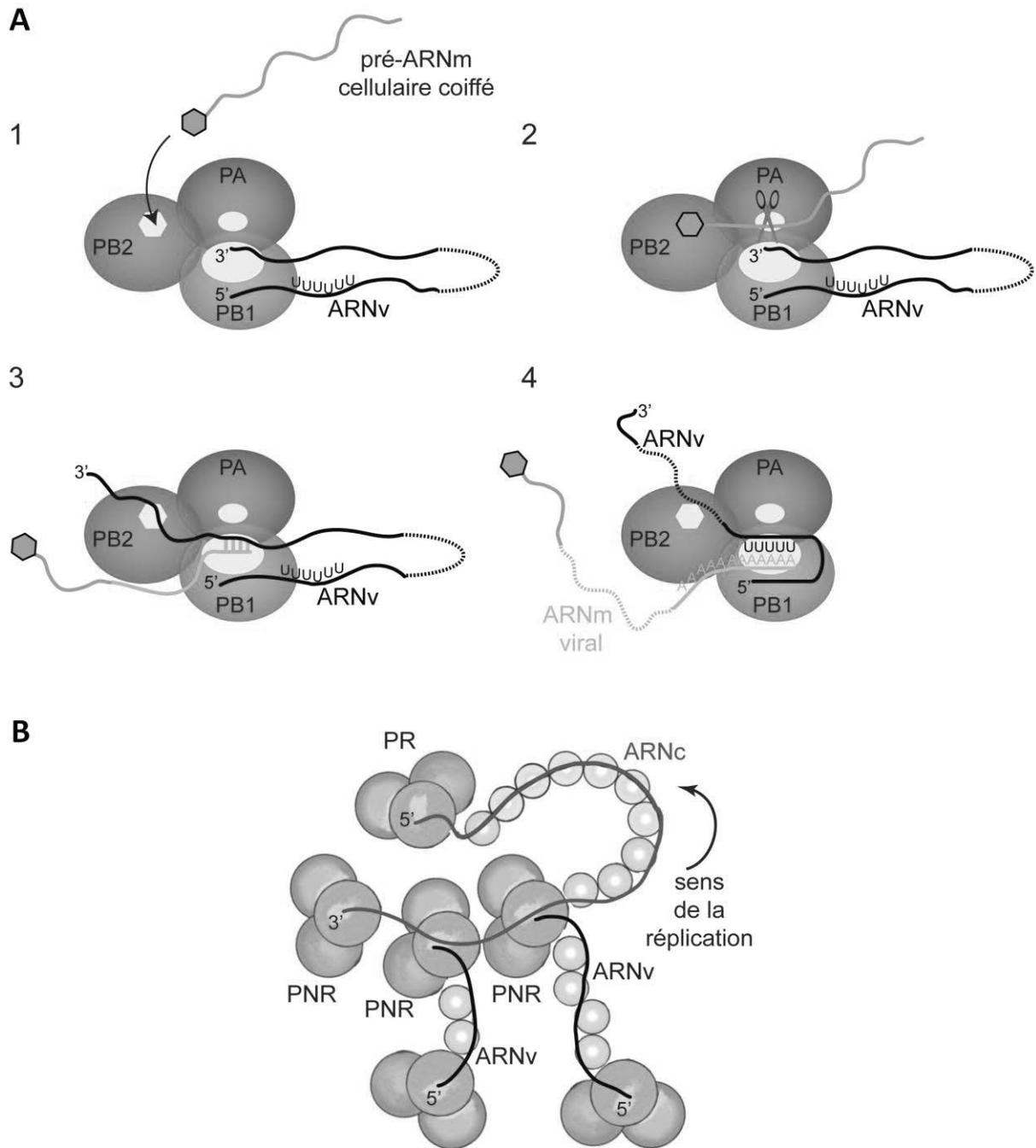
Peu d'informations sur la structure secondaire des ARNv génomiques sont disponibles, la plupart des données structurales étant consacrées aux ARNm viraux de polarité positive, principalement les ARNm M et NS. Plusieurs études ont prédit la présence d'éléments structuraux dans les régions codantes et non codantes des ARNv de polarité négative, qui sont conservés parmi les IVA [103–105]. De manière intéressante, une structure très conservée localisée entre les nt 1436-1530 du segment NP, constituée d'une tige-boucle et d'un pseudo-noeud, est importante pour l'empaquetage de cet ARNv [106–108]. Cette structure semble uniquement être présente dans les segments d'ARNv de polarité négative et non dans les ARNc [104]. De la même manière, une tige-boucle très conservée située entre les nt 219-240 du segment M est importante pour la production de particules virales infectieuses [105]. Ce n'est que récemment que la première structure secondaire d'un segment génomique entier a été obtenue, en combinant des données expérimentales et bioinformatiques [109]. La majorité des nt de l'ARNv NS, transcrit *in vitro* et en absence de protéines virales, sont engagés dans une paire de base (61%). La structure révèle la présence de nombreuses tiges-boucles, stables thermodynamiquement, ainsi que des interactions à longue distance, dont certains motifs sont conservés parmi les IVA. Cependant, *in vivo*, les ARNv sont associés à la protéine NP et au complexe ARN polymérase. Il est donc nécessaire de vérifier l'existence de ces structures secondaires dans un contexte cellulaire.

#### **4.2. Transcription du génome et export des ARNm**

La transcription du génome permet d'obtenir une copie partielle des ARNv matrices, qui porte une coiffe à son extrémité 5' et une queue poly A à son extrémité 3'. L'acquisition de la coiffe se déroule selon un mécanisme partagé par les *Orthomyxoviridae*, les *Bunyaviridae* et les *Arenaviridae*, appelé le vol de coiffe (*cap snatching*) (pour revue, voir [110]) (figure 11A). L'ARN polymérase virale interagit directement avec le CTD (*C-terminal domain*) de l'ARN polymérase II cellulaire [111, 112], ce qui facilite la reconnaissance, par la

sous-unité virale PB2, de la coiffe des pré-ARNm cellulaires en cours de transcription [88]. Une étude a cependant suggéré que les IVA ciblent préférentiellement les coiffes des ARN non codants cellulaires, notamment les petits ARN nucléaires (U1 et U2), ainsi qu'une classe, identifiée récemment, de petits ARN coiffés (18-40 nt) associés au promoteur des gènes cellulaires transcrits par l'ARN polymérase II [113]. Les ARN cellulaires sont ensuite clivés 10 à 15 nt en aval de la coiffe grâce à l'activité endonucléase portée par la protéine PA [87, 114]. Les courts segments d'ARN coiffés ainsi générés servent d'amorce pour initier la transcription des ARNm viraux, *via* le domaine ARN polymérase ARN-dépendant porté par la sous-unité PB1 [92, 93, 115]. L'élongation de la transcription se termine par une étape de bégaiement de l'ARN polymérase virale au niveau d'une courte séquence poly U (5 à 7 résidus uracile) située en 5' de l'ARNv [116, 117]. Ce mécanisme de polyadénylation, non canonique, peut s'expliquer par une gêne stérique lorsque l'extrémité 5' de la matrice, qui reste fixée au complexe ARN polymérase tout au long de la transcription, se rapproche du site catalytique de l'ARN polymérase. Les transcrits viraux subissent également des modifications post-transcriptionnelles, comme la méthylation de certaines adénosines internes [118, 119]. De plus, le transcrit M est épissé en ARNm M2, ARNm M4 et ARNm3, et le transcrit NS est épissé en ARNm NS2 et en ARNm NS3, uniquement retrouvé dans certaines souches d'IVA (pour revue, voir [120]).

Le facteur d'export majeur des ARNm cellulaires est la protéine NXF1 (*nuclear export factor 1*) en complexe avec son co-facteur p15. De nombreuses données expérimentales associent ce même facteur à l'export de tous les ARNm viraux (pour revue, voir [121]). Néanmoins, l'export des ARNm viraux non épissés ou ne possédant pas d'introns d'une part, et les ARNm épissés d'autre part, ne fait pas intervenir les mêmes co-facteurs. Récemment, une étude a révélé que les ARNm codant les protéines PB2, PB1 et PA ne sont pas exportés par une des deux voies principales (NXF1 et CRM1, *chromosome maintenance region 1*), suggérant un mécanisme d'export atypique [122]. Après leur transport dans le cytoplasme, la machinerie cellulaire reconnaît les ARNm viraux, coiffés et polyadénylés, comme des ARNm endogènes et va donc permettre leur traduction.



**Figure 11 : Modèle des étapes de transcription et de réplication des RNPv des IVA (d'après [123]).**

**A. Transcription coiffe-dépendante des RNPv.** Pour plus de clarté, les protéines NP n'ont pas été représentées. (1) La sous-unité PB2 reconnaît la coiffe d'un ARN cellulaire, puis (2) celui-ci est clivé, environ 10-15 nt en aval de la coiffe, par la sous-unité PA. (3) Ce court fragment d'ARN cellulaire coiffé sert ensuite d'amorce pour initier la transcription des ARNm viraux par la sous-unité PB1. (4) Ce processus se termine par le bégaiement de l'ARN polymérase virale au niveau d'une séquence poly U présente à l'extrémité 5' de l'ARNv, permettant la polyadénylation des ARNm viraux.

**B. Réplication des RNPv.** La réplication nécessite la synthèse d'un ARNc de polarité positive par une ARN polymérase hétérotrimérique dite « résidente » (PR). Des complexes hétérotrimériques « non-résidents » (PNR), après avoir été traduits dans le cytoplasme puis importés dans le noyau, se fixent à l'extrémité 3' des RNPc. Les ARNv néo-synthétisés sont associés au complexe ARN polymérase et aux protéines NP pour former des RNPv.

### 4.3. Réplication du génome viral

L'étape de réplication virale débute par la synthèse d'un intermédiaire de réplication de polarité positive, appelé ARNc, complémentaire à la totalité de l'ARNv de polarité négative (pour revue, voir [124]) (figure 11B). Contrairement à l'étape de transcription, l'initiation de la réplication ne nécessite pas d'amorce (pour revue, voir [125]). Les ARNc ne sont ni coiffés, ni polyadénylés et sont associés aux protéines PB2, PB1, PA et NP pour former des RNPC. Ces protéines ont pour rôle de stabiliser les ARNc en empêchant leur dégradation par les nucléases cellulaires [126]. Les RNPC servent ensuite de matrice à l'ARN polymérase virale pour la synthèse de nouveaux ARNv de polarité négative, qui seront empaquetés dans les particules virales néoformées. Le complexe ARN polymérase initie la synthèse de l'ARNv sans amorce, par un mécanisme d'initiation interne au niveau du U en position 4 de l'extrémité 3' de l'ARNc (pour revue, voir [125]).

### 4.4. Transition transcription/réplication

Les phases transcriptionnelles et répliquatives du génome sont régulées dans le temps, la transcription étant prédominante au début de l'infection, tandis que la réplication est différée [127, 128]. Les facteurs impliqués dans le passage entre les deux étapes ne sont pas encore très bien définis, mais un certain nombre de facteurs viraux (l'ARN polymérase, la protéine NP, NS1 et NEP/NS2) et cellulaires (la disponibilité en NTP et en ARN cellulaires coiffés, le facteur d'élongation de la transcription *tat-SF1* (*tat-specific factor 1*), et les hélicases MCM (*minichromosome maintenance*) et UAP56) ont été identifiés (pour revue, voir [124]). La protéine NP pourrait jouer un rôle d'interrupteur, permettant le passage de l'ARN polymérase de la transcription à la réplication. Elle est également essentielle pour l'accumulation des ARNc et leur stabilité [82, 126, 129, 130]. De plus, Perez *et al.* [131] ont proposé que de petits ARNv, correspondant aux extrémités 5' de chacun des segments d'ARNv, sont responsables de la transition entre la transcription et la réplication, de manière segment-spécifique, en se fixant à l'ARN polymérase virale. Des études suggèrent que deux ARN polymérases virales distinctes sont impliquées dans ces deux mécanismes : des ARN polymérases résidentes transcrivent le génome, tandis que la réplication nécessite un complexe ARN polymérase non résident, nouvellement synthétisé et importé dans le noyau [30, 132] (pour revue, voir [125]) (figure 11B). York *et al.* [102] ont néanmoins remis en cause

ce modèle et suggèrent un mécanisme dans lequel une ARN polymérase non résidente, importée dans le noyau, active l'ARN polymérase résidente qui réplique alors le génome viral.

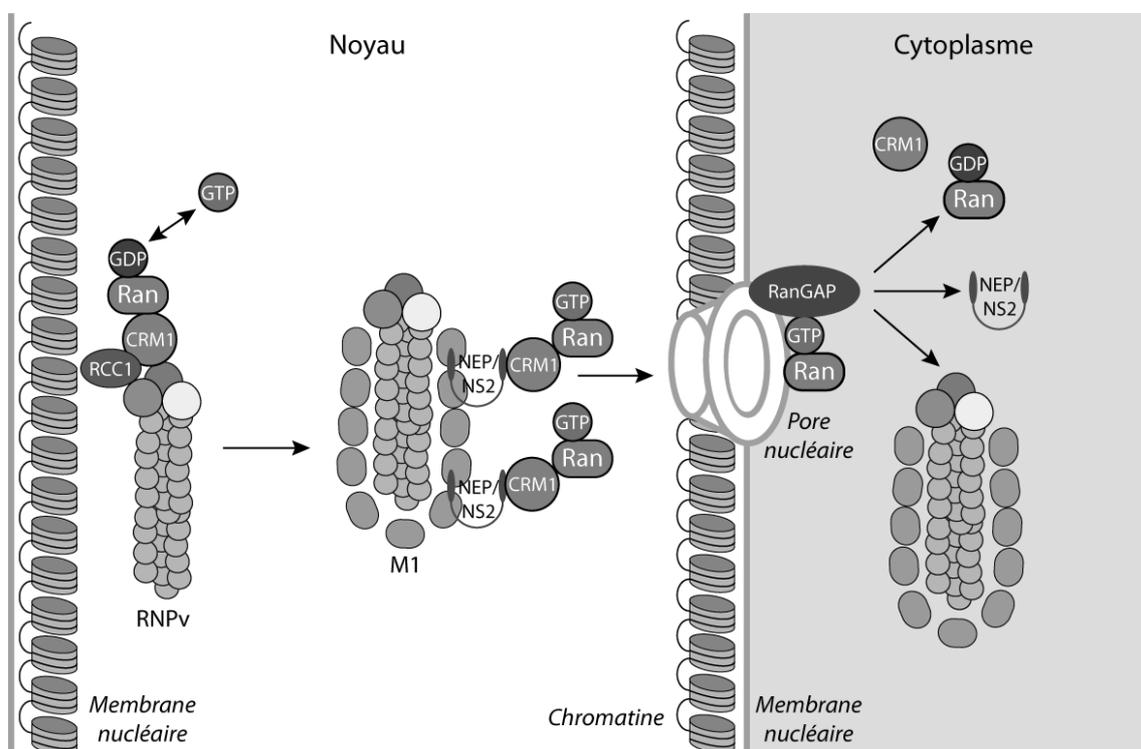
## 5. Export des RNPv dans le cytoplasme

Après l'étape de répllication, le génome viral néo-synthétisé doit être exporté dans le cytoplasme (pour revue, voir [133]). Les protéines PB2, PB1, PA, NP et M1 possèdent toutes un ou plusieurs NLS, qui permet leur import dans le noyau après traduction (pour revue, voir [134]), tandis que la protéine NEP/NS2 entre dans le noyau par diffusion passive [135]. L'ARNv n'est donc pas exporté sous la forme d'ARNv « nu », mais sous forme de RNPv associée aux protéines M1 et NEP/NS2.

Les travaux de Chase *et al.* [136] suggèrent que les RNPv sont fortement associées à la chromatine en périphérie du noyau (figure 12), ce qui permet leur co-localisation avec RCC1 (*regulator of chromosome condensation 1*). Cette protéine est essentielle dans le mécanisme d'export dépendant de CRM1, puisqu'elle charge Ran avec du GTP. La fixation de la protéine virale M1 aux RNPv pourrait promouvoir leur dissociation de la chromatine [136–138]. En parallèle, M1 joue également un rôle d'adaptateur en interagissant à la fois avec les protéines NP [139] associées aux RNPv néo-formées et avec les protéines NEP/NS2 [140] (figure 12). L'export de ces complexes est ensuite possible grâce à deux signaux d'export nucléaire (NES, *nuclear export signal*) dans la protéine NEP/NS2, qui sont reconnus par l'exportine CRM1 activée par une molécule de Ran•GTP [141]. Dans le cytoplasme, la protéine RanGAP hydrolyse le GTP, ce qui induit la dissociation des complexes. Les RNPv restent liées aux protéines M1. Plusieurs modifications post-traductionnelles sont nécessaires afin de réguler le mécanisme d'export, comme la SUMOylation de M1 [142] et la phosphorylation des protéines NP [143] et NEP/NS2 [144], elle-même vraisemblablement régulée par une cascade de phosphorylation initiée par la protéine HA à la membrane cellulaire. Une autre étude suggère que NEP/NS2 interagit directement avec l'ARN polymérase virale [145]. Par ailleurs, un mécanisme d'export impliquant une interaction directe entre la machinerie d'export CRM1-dépendante et la protéine NP n'est pas non plus à exclure [146]. Mühlbauer *et al.* [147] ont montré que le diamètre des pores nucléaires est considérablement augmenté suite à l'infection par un IVA, ce qui facilite donc l'export des

RNPv. Il semblerait également que la diffusion passive des RNPv et la voie d'export active CRM1 soient concomitantes, permettant ainsi l'accumulation rapide des RNPv dans le cytoplasme.

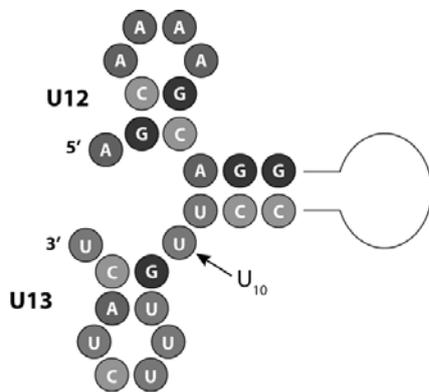
L'utilisation de techniques d'imagerie permet la visualisation des RNPv dans les différents compartiments cellulaires. Alors qu'une étude a récemment suggéré que les RNPv sont exportées du noyau sous la forme de complexes associant plusieurs RNPv [148], d'autres montrent que les RNPv, présentes sous forme d'entités séparées dans le noyau [149], seraient exportées individuellement [52].



**Figure 12 : Modèle d'export nucléaire des RNPv.** Les RNPv sont associées à la chromatine en périphérie du noyau et sont co-localisées avec la protéine RCC1. La protéine NEP/NS2 sert d'adaptateur entre les complexes RNPv-M1 et CRM1-Ran•GTP. Après l'export des RNPv *via* la voie CRM1, RanGAP hydrolyse le GTP, ce qui aboutit à la dissociation des RNPv et des protéines de transport.

Les mécanismes permettant l'export spécifique des RNPv, tandis que les RNPc restent dans le noyau, ne sont pas encore très bien connus. La structure des promoteurs pourrait en partie expliquer cette sélection, en particulier la présence d'une uracile non appariée en position 10 dans le cas des RNPc (comparer la figure 3C et la figure 13) [150]. Des études structurales à haute résolution révèlent que la sous-unité PB2 du complexe ARN polymérase

est très flexible et que certains de ses sous-domaines adoptent des conformations très différentes selon qu'ils se lient aux RNPv ou RNPc [92, 93, 96]. Ces modes de reconnaissances différents des deux promoteurs pourraient notamment expliquer l'export sélectif des RNPv.



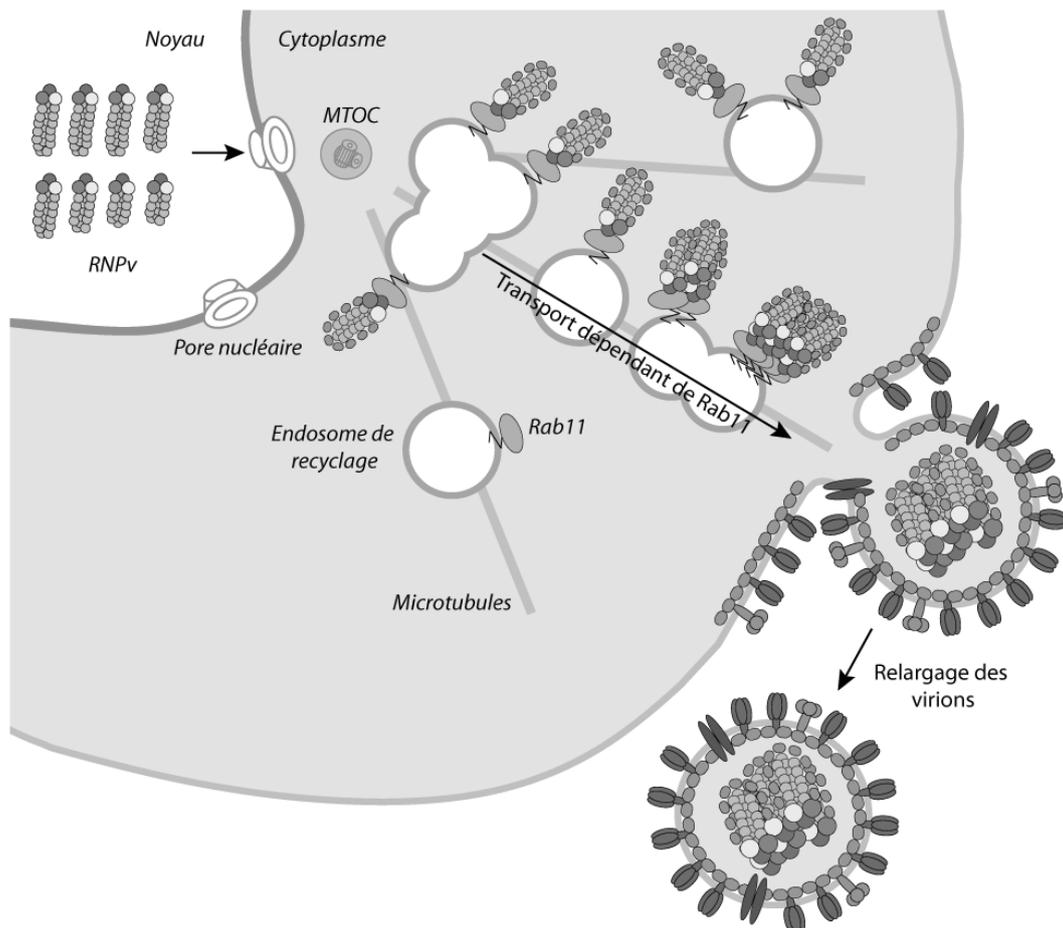
**Figure 13 : Structure du promoteur de l'ARNc selon le modèle de repliement en « tire-bouchon ».** L'uracile en position 10, caractéristique du promoteur de l'ARNc, est indiqué par une flèche.

## 6. Assemblage, bourgeonnement et libération des particules virales

Après export nucléaire, les RNPv s'accumulent près du MTOC (*microtubule organizing center*), en périphérie du noyau (figure 14) [52, 151–153]. La protéine Rab11 (*ras-related in brain*) est cruciale pour un bourgeonnement efficace et une morphogénèse virale correcte (pour revue, voir [154]). Associée aux endosomes de recyclage, Rab11•GTP se lie aux RNPv [151, 155–157], probablement *via* une interaction directe avec la protéine PB2 [151, 155–157]. Le transport de ces complexes jusqu'à la membrane plasmique est réalisé grâce au réseau de microtubules et aux filaments d'actine [151, 155–157].

En périphérie du noyau, plusieurs RNPv sont associées en « sous-complexes » [148], qui sont ensuite transportés jusqu'au site de bourgeonnement au niveau de la membrane plasmique. Ainsi, il est probable que lors de la migration vers la membrane, des événements de co-localisation dynamiques ont lieu entre les RNPv [52, 148]. Les techniques de microscopie ne sont pas encore assez résolutes et ne permettent pas de conclure quant à l'existence d'interactions physiques entre les RNPv en route vers la membrane. Néanmoins, une hypothèse serait que les vésicules Rab11 servent de plateforme, permettant de rapprocher les RNPv et donc de faciliter leur interaction en vue de la formation d'un complexe entre les huit segments, qui sera par la suite incorporé dans les virions (pour revue, voir [158]). Une étude récente a montré que l'infection par un IVA diminue l'efficacité du mécanisme de recyclage et que le temps de migration des endosomes le long des

microtubules est augmenté [159]. Par conséquent, des zones de fortes concentrations en RNPv se forment dans le cytoplasme, ce qui pourrait favoriser l'établissement d'interactions intermoléculaires. Les auteurs proposent néanmoins un mécanisme dans lequel les endosomes se concentrent lors du transport, sans pour autant fusionner.



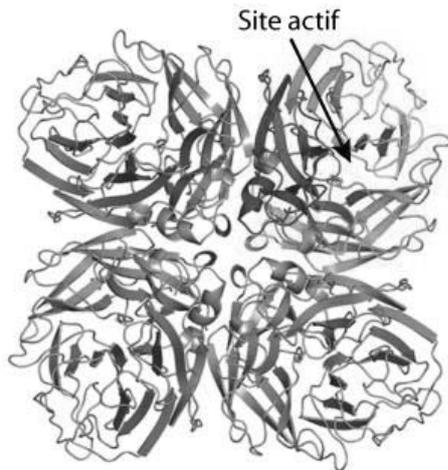
**Figure 14 : Transport et empaquetage des RNPv.** Les RNPv atteignent le cytoplasme *via* la voie d'export nucléaire médiée par CRM1 et s'accumulent près du MTOC (*microtubule organizing center*), en périphérie du noyau. Ces RNPv s'associent à des endosomes de recyclage en interagissant avec la protéine Rab11 (*ras-related in brain*). La migration des endosomes le long du réseau de microtubules, jusqu'au site de bourgeonnement à la membrane plasmique, permet des événements de co-localisation de plusieurs RNPv pour assurer l'empaquetage d'un jeu complet de huit RNPv dans les virions en cours de bourgeonnement. Les particules virales néo-formées sont ensuite relarguées.

Le mécanisme à l'origine de l'initiation du bourgeonnement viral, au niveau de la membrane apicale, n'est pas encore entièrement élucidé (pour revue, voir [160]). L'expression de la protéine HA et son accumulation dans les radeaux lipidiques, des micro-domaines riches en cholestérol et en sphingolipides, permet d'initier la courbure membranaire nécessaire au bourgeonnement et à la production de VLP (*virus-like particle*) [161]. Il en va de même pour l'expression, de manière individuelle, des protéines NA [161,

162], M1 [163–165] et M2 [161], bien que la production des VLP soit significativement augmentée lorsque les protéines HA, NA et/ou M2 sont exprimées simultanément [161, 162, 164]. La protéine M1 est recrutée *via* les queues cytoplasmiques de HA et NA [4, 139, 166, 167]. Suite à l'interaction de la protéine de matrice avec les radeaux lipidiques, un changement conformationnel conduit à la polymérisation de M1 [163, 168] et altère davantage la courbure de la membrane.

L'emballage des RNPv dans les particules virales est probablement réalisé *via* la protéine M1 qui recouvre les RNPv. La délétion des queues cytoplasmiques des protéines HA et NA, capables d'interagir avec M1, s'accompagne d'une diminution de la quantité d'ARNv dans les particules virales [169]. Par ailleurs, une interaction directe entre la queue cytoplasmique de M2 et la protéine M1 a été mise en évidence [170, 171] et la délétion de la queue cytoplasmique de M2 entraîne elle-aussi une diminution de la quantité d'ARNv M emballé [171]. Des travaux récents ont montré qu'une des régions non structurées de la protéine NP, riche en acides aminés chargés positivement et très conservée parmi les IVA humains, porcins et aviaires, est impliquée dans l'oligomérisation de la protéine NP sur l'ARNv pour former des RNPv [172]. Cette région est également impliquée dans l'interaction avec le phosphatidyl-inositol(4,5)bisphosphate (PI(4,5)P2), un composant des radeaux lipidiques de la membrane plasmique essentiel dans l'étape d'emballage du génome viral, permettant l'accumulation des RNPv à la membrane [172]. Pour finir, la scission de la membrane, indépendante des complexes cellulaires ESCRT, recrutés par de nombreux autres virus, nécessite la présence de la protéine M2 [173]. D'autres études ont montré que les protéines HA et NA sont également capables de réaliser cette étape en absence de la protéine M2 (pour revue, voir [174]).

Après bourgeonnement, les virions restent associés à la cellule hôte en raison de l'interaction entre les glycoprotéines virales HA et les acides sialiques cellulaires. Cette liaison est clivée grâce à l'activité enzymatique acylneuraminyl hydrolase (EC 3.2.1.18) portée par la glycoprotéine de surface NA, active sous forme d'homotétramère (figure 15). Ainsi, les virions sont relargués dans le milieu extracellulaire. De manière intéressante, l'acide sialique n'est pas un substrat pour les protéines NA de sous-type H17 et H18, portées par les virus de chauves-souris H17N10 et H18N11 [11, 175, 176].



**Figure 15 : Structure tridimensionnelle de la tête de la protéine NA des IVA (d'après [177]). L'un des monomères est représenté en couleur.**

Les glycoprotéines HA et NA, qui ont des fonctions majeures dans le cycle viral des IVA, ont des rôles antagonistes. En effet, l'infection débute par l'interaction entre la protéine HA et les acides sialiques tandis que le relargage des virions néo-synthétisés nécessite le clivage de cette liaison par la protéine NA. Il doit donc exister une balance fonctionnelle entre les activités des deux protéines afin d'optimiser la réplication virale [178, 179]. Par conséquent, lors de l'adaptation à un nouvel hôte, il est nécessaire que les protéines HA et NA co-évoluent de manière à ce que leur équilibre fonctionnel soit préservé [180, 181]. Par ailleurs, une incompatibilité entre les protéines HA et NA pourrait être à l'origine d'une restriction protéique lors d'un événement de réassortiment génétique [182].

## **B. Évolution des *Influenzavirus A*, transmission du virus entre différents réservoirs et adaptation à l'hôte**

### **I. Mécanismes moléculaires permettant l'évolution des *Influenzavirus A***

Les IVA évoluent principalement par dérive antigénique (*antigenic drift*) et cassure antigénique (*antigenic shift*), ce qui induit leur échappement au système immunitaire de l'hôte.

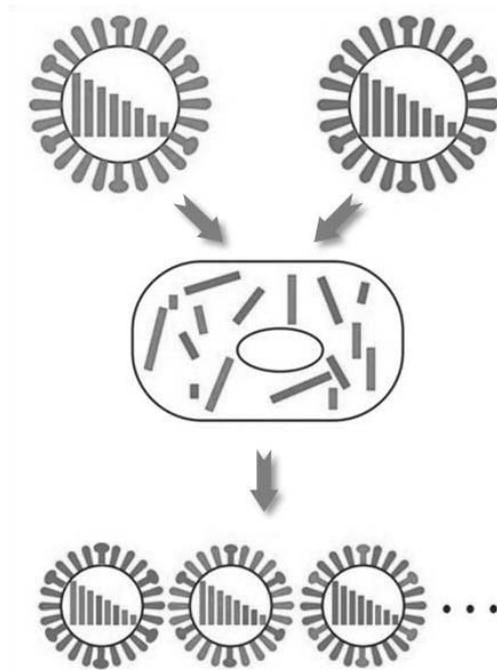
#### **1. Dérive antigénique**

L'évolution des IVA par dérive antigénique est essentiellement liée à la faible fidélité de l'ARN polymérase virale, qui ne possède pas d'activité correctrice. Le taux de mutations est de l'ordre de  $10^{-3}$  à  $10^{-4}$  mutations par site et par an [183], ce qui correspond à l'introduction d'une substitution dans l'ensemble du génome viral à chaque cycle répliatif [184]. La plupart des mutations introduites dans les segments HA et NA, codant pour les protéines de surface, sont neutres, car elles n'affectent pas les propriétés antigéniques des protéines. Néanmoins, certaines substitutions altèrent la conformation des domaines antigéniques, empêchant ainsi leur reconnaissance par les anticorps de l'hôte, ce qui entraîne l'échappement des virus aux défenses de l'organisme [185]. Ce mécanisme permet d'expliquer pourquoi les virus qui circulent au cours des épidémies de gripes saisonnières possèdent des propriétés antigéniques différentes des virus des années précédentes [186].

#### **2. Cassure antigénique**

Lorsqu'une cellule est infectée par au moins deux virus différents, la nature segmentée du génome permet, au moment de l'assemblage de nouvelles particules virales, un mélange et donc un échange de segments d'ARNv provenant de souches différentes (figure 16). Ces événements de réassortiment génétique provoquent l'émergence de virus réassortants possédant des génomes hybrides et permettent au virus d'évoluer rapidement (pour revue, voir [187]). Lorsque des segments HA et/ou NA étrangers (d'origine porcine ou aviaire), sont empaquetés dans des virus humains contre lesquels le système immunitaire de la population humaine est naïf, les virus peuvent être à l'origine de pandémies grippales [188]. Contrairement à la dérive antigénique, la cassure antigénique est donc caractérisée

par un changement soudain dans les déterminants antigéniques HA et/ou NA. Fort heureusement, les évènements provoquant la cassure antigéniques sont relativement rares et ont lieu en moyenne tous les 30 ans dans la population humaine [189] (voir § D.II.1 pour plus de détails).



**Figure 16 : Principe général du réassortiment génétique chez les IVA.** La co-infection de cellules avec deux virus différents permet l'émergence de virus possédant des génomes hybrides.

## II. Notion de « barrières d'espèces »

Les IVA peuvent infecter de nombreux hôtes, notamment les oiseaux et les mammifères (principalement l'Homme et le porc). Néanmoins, tous les sous-types ne sont pas représentés au sein de chaque espèce (voir § A.III). Les virus se transmettent efficacement entre individus d'une même espèce, mais la transmission inter-espèces est plus rare. Ce phénomène est dû à l'existence de « barrières d'espèces », expliquée notamment par la reconnaissance des récepteurs cellulaires par la protéine virale HA [190]. Ces virus n'ont généralement pas la capacité de se transmettre entre individus, mais peuvent acquérir, au cours du temps, des mutations leur conférant cette capacité et ainsi franchir la « barrière d'espèce ». Par ailleurs, des études récentes ont montré que la distribution du type de liaison entre l'acide sialique et le galactose, à la surface des cellules hôtes, est plus complexe qu'initialement démontré. En effet, comme chez le porc, les deux types de récepteurs, SA $\alpha$ 2,6Gal et SA $\alpha$ 2,3Gal, sont présents dans le tractus respiratoire de l'Homme, suggérant qu'il peut être considéré, lui aussi, comme un « creuset de mélange » [191, 192].

De plus, même si les récepteurs de type SA $\alpha$ 2,3Gal sont prédominants chez les poulets, les cailles et certains oiseaux terrestres (piverts, hirondelles...), les récepteurs de type SA $\alpha$ 2,6Gal sont également présents au niveau des tractus respiratoire et gastro-intestinal [193]. Ainsi, les oiseaux peuvent eux aussi servir d'hôtes intermédiaires pour le réassortiment génétique et l'adaptation des IVA aviaires aux mammifères.

Une étude de Chen et *al.* [194] a révélé que 52 acides aminés, principalement localisés dans les protéines PB2, PB1, PA et NP, qui constituent les RNPv, représentent des signatures spécifiques aux virus aviaires ou humains. Ces résidus sont impliqués dans le franchissement de la « barrière d'espèce » puisque leur mutation permet la réplication des virus aviaires chez l'Homme (pour revue, voir [195]). L'acide aminé en position 627 de la protéine PB2 est un exemple remarquable dans ce domaine. Les virus humains possèdent une lysine à cette position, tandis que la majorité des virus aviaires possèdent un acide glutamique. La présence de ce résidu dans la protéine PB2 aviaire empêche la réplication virale dans l'appareil respiratoire supérieur de l'Homme [196] (pour revue, voir [197]). Une étude récente a élucidé ce mécanisme en identifiant une protéine cellulaire, ANP32A, responsable de la faible activité de l'ARN polymérase virale aviaire en cellules de mammifère [198]. Cette protéine est similaire chez les oiseaux et chez l'Homme, hormis une région de 33 acides aminés absente dans la protéine humaine. De manière intéressante, la présence d'une lysine en position 627 de la protéine PB2 aviaire permet une réplication virale efficace en présence de la version courte de la protéine ANP32A humaine. Parallèlement, quelques substitutions dans le domaine RBS de la protéine HA de virus aviaires sont également suffisantes pour induire un changement de spécificité et donc une reconnaissance préférentielle des récepteurs de type SA $\alpha$ 2,6Gal majoritairement présents chez l'Homme [199, 200].

### **III. Infections sporadiques de l'Homme par des virus aviaires**

Si les virus aviaires sont rarement transmis directement à l'Homme sans passer par un hôte intermédiaire, plusieurs événements ont néanmoins été recensés. Le cas le plus marquant a été celui du virus H1N1, qui s'est ensuite efficacement transmis d'Homme à Homme et a causé la mort de 50 à 100 millions de personnes en 1918-1919. Cette pandémie grippale, aussi connue sous le nom de « grippe espagnole », constitue la pandémie la plus

mortelle de l'histoire dans un laps de temps aussi court. L'analyse du génome viral a montré que les huit segments génomiques dérivent d'un virus aviaire [201, 202], suggérant ainsi que le virus a soudainement été transmis de l'oiseau à l'Homme. Cette hypothèse est désormais controversée. En effet, des analyses phylogénétiques suggèrent que le virus H1N1 est issu de deux événements de réassortiment génétique : un virus réassortant issu d'un premier événement, entre des virus humain et porcine, a probablement circulé chez l'Homme ou chez le porc durant plusieurs années, avant un second événement de réassortiment impliquant un virus aviaire [203]. Une autre étude propose qu'un virus circulant humain H1 et un virus aviaire ont réassorti en 1915, conduisant à l'émergence d'un virus possédant le gène H1 humain et les sept autres gènes d'origine aviaire [204].

Plus récemment, d'autres cas de transmission directe de virus aviaires à l'Homme ont été recensés, notamment par des IVA de sous-types H5N1 et H7N9. En 1997, 18 personnes vivant à Hong-Kong ont été infectées avec un virus hautement pathogène H5N1, entraînant la mort de six d'entre elles [205, 206]. Depuis lors, le bilan n'a cessé de s'alourdir, puisque sur les 850 individus infectés à ce jour, 449 sont décédés, principalement en Égypte, en Indonésie et au Vietnam (bulletin du 9 mai 2016 de l'OMS, [http://www.who.int/influenza/human\\_animal\\_interface/EN\\_GIP\\_20160509cumulativenumberc5N1cases.pdf?ua=1](http://www.who.int/influenza/human_animal_interface/EN_GIP_20160509cumulativenumberc5N1cases.pdf?ua=1)). En Chine, depuis 2013, 786 cas d'infection par un virus aviaire H7N9 ont été confirmés, parmi lesquels 307 individus sont décédés (bulletin du 19 mai 2016 de la FAO (*Food and Agriculture Organisation*), [http://www.fao.org/ag/againfo/programmes/en/empres/H7N9/situation\\_update.html](http://www.fao.org/ag/againfo/programmes/en/empres/H7N9/situation_update.html)). Des cas d'infection d'Homme par des virus aviaires de sous-types H6N1, H7N2, H7N3, H7N7, H9N2, H10N7 et H10N8 ont également été reportés [207–214].

#### **IV. Transmissibilité des virus par les aérosols**

Lorsque les virus aviaires passent directement à l'Homme, sans événement de réassortiment génétique, ils se transmettent généralement difficilement entre les individus. La capacité de se transmettre *via* les aérosols est une nécessité pour que la souche soit à l'origine d'une pandémie. Plusieurs études ont révélé que l'introduction de quelques mutations dans le génome d'un virus aviaire hautement pathogène H5N1 lui confère la capacité de se transmettre de cette manière (pour revues, voir [195, 215]). La majorité des substitutions concernent la protéine HA (avec un changement de spécificité de

reconnaissance de l'acide sialique), ainsi que les protéines PB2 et PB1 [216–218]. Une étude similaire sur un virus aviaire hautement pathogène H7N1 a identifié l'apparition de mutations dans les protéines PB2, NP, HA et M1 [219]. Cependant, dans ce cas, la substitution dans la protéine HA n'altère pas la spécificité pour les récepteurs de type SA $\alpha$ 2,3Gal. L'ensemble de ces études prouve que l'introduction d'un nombre limité de mutations dans le génome des IVA aviaires permet leur transmission *via* les aérosols. Dans le cas où cette transmission est efficace, ces virus peuvent être à l'origine de graves pandémies chez l'Homme.

## **C. Vaccins et traitements antigrippaux**

### **I. Vaccins antigrippaux saisonniers**

Les vaccins antigrippaux à destination de l'Homme constituent un moyen de protection efficace contre les IVA. Cependant, en raison de l'évolution constante des souches d'IVA due à la dérive antigénique, il est nécessaire d'administrer, chaque année, une dose de vaccin saisonnier dont la composition est mise à jour annuellement. Celle-ci est recommandée par l'OMS en février pour l'hémisphère nord, en prenant en compte les souches virales ayant circulées durant l'hiver précédent. Lorsque les souches vaccinales correspondent aux souches circulantes, le vaccin confère une protection à 80% des personnes vaccinées [220]. Les données de surveillance sont collectées par cinq grands centres mondiaux de la grippe (Atlanta, Londres, Beijing, Melbourne et Tokyo) et proviennent de plus de 140 centres nationaux répartis dans une centaine de pays.

Les vaccins saisonniers sont des vaccins trivalents correspondant à un mélange de trois souches circulantes contemporaines : deux IVA, de sous-type H1N1 et H3N2, ainsi qu'un IVB. L'OMS a recommandé la composition suivante pour les vaccins antigrippaux trivalents destinés à être utilisés pendant la saison 2016-2017 :

- un virus de la souche A/California/7/2009 (H1N1)pdm09 (la souche responsable de la pandémie de 2009 et incluse dans les vaccins depuis l'hiver 2010-2011) ;
- un virus de la souche A/Hong Kong/4801/2014 (H3N2) ;
- un virus de la souche B/Brisbane/60/2008.

Cependant, depuis les années 1980, deux lignées d'IVB, distinctes du point de vue phylogénétique et antigénique, co-circulent dans la population humaine [222], compliquant donc la prédiction de la lignée d'appartenance de la souche B qui prédominera la saison suivante. Par conséquent, à partir de la saison grippale 2013-2014, l'OMS a recommandé l'utilisation de vaccins quadrivalents contenant une seconde souche d'IVB, qui devraient offrir une protection accrue contre les infections aux virus de type B [223, 224]. Pour la saison 2016-2017, l'OMS préconise un virus de la souche B/Phuket/3073/2013.

La majorité des souches vaccinales sont obtenues par réassortiment génétique entre chacune des deux souches d'IVA et la souche PR8 (H1N1), qui est une souche adaptée à la culture à hauts rendements sur les œufs de poule embryonnés. À partir de 1969, des études ont montré qu'un virus réassortant (X-31b), issu d'un réassortiment génétique entre PR8 et la souche A/Hong Kong/Aichi/68 (H3N2) [225], pouvait aisément être produit dans un système allantoïque [226]. Ce virus réassortant, qui contient les segments HA et NA de la souche H3N2 dans le fond génétique PR8, a constitué la première souche vaccinale antigrippale [227].

Actuellement, des œufs de poule embryonnés sont co-inoculés avec la souche PR8 et l'une des souches d'IVA. À partir de tous les réassortants obtenus, une première sélection permet d'isoler ceux qui possèdent les protéines antigéniques HA et NA d'intérêt épidémiologique et par conséquent les segments HA et NA étrangers. Par la suite, seuls les virus présentant des caractéristiques répliquatives élevées sont conservés. Les souches vaccinales sont ensuite distribuées aux industriels pour débiter leur production à grande échelle. Pour finir, les virus sont inactivés chimiquement, généralement avec du formaldéhyde ou du  $\beta$ -propiolactone, puis fractionnés à l'aide de détergents (*split vaccine*). Un autre type de vaccin (*subunit vaccine*), nécessite la purification des protéines HA et NA, issues de ces virus fragmentés.

Depuis peu, un vaccin antigrippal tétravalent à base de virus vivants atténués est disponible en France (Fluenz Tetra® (MedImmune)) et recommandé chez les enfants et adolescents de 2 à 17 ans. Contrairement aux autres vaccins, il est administré par voie intranasale. Les souches vaccinales sont également obtenues par réassortiment génétique, mais la souche PR8 est remplacée par des souches donneuses IVA ou IVB thermosensibles et adaptées au froid. Ainsi, les virus réassortants atténués peuvent se répliquer au niveau du nasopharynx qui est à une température relativement basse (25°C à 33°C) et donc initier une réponse immunitaire plus proche de celle provoquée par une infection naturelle d'IV, mais ils sont incapables de se répliquer dans le tractus respiratoire inférieur, à une température plus élevée.

L'ensemble du procédé de fabrication des vaccins antigrippaux saisonniers, qu'ils soient inactivés ou vivants atténués, comporte de nombreuses étapes. Le délai entre le

moment où la composition vaccinale est déterminée et la commercialisation des vaccins est d'environ cinq à six mois.

Récemment, des supports de production alternatifs pour les semences vaccinales ont été développés. En effet, divers inconvénients accompagnent la production sur des œufs de poule, notamment leur disponibilité en grande quantité, le risque de contamination de virus aviaires dans les élevages de poules, mais aussi l'apparition de mutations adaptatives liées à la réplication virale sur les œufs [228]. Dans cette optique, plusieurs systèmes cellulaires ont été mis au point (cellules MDCK ou Vero, des cellules rénales canines ou de singe vert africain, respectivement) (pour revue, voir [229]). À ce jour, deux vaccins produits en culture cellulaire ont été autorisés par les autorités sanitaires européennes. Le vaccin contre la grippe saisonnière Optaflu (Novartis) est produit à partir de cellules MDCK et a obtenu son autorisation de mise sur le marché en juin 2007 [230], tandis que le vaccin pandémique H1N1 Celvapan (Baxter), produit à partir de cellules Vero, a été approuvé en 2009. Début 2013, le premier vaccin saisonnier produit à partir de cellules d'insectes a été approuvé par la FDA (*Food and Drug Administration*) aux États-Unis ; il s'agit de Flublok (Protein Science) [231].

Ces systèmes alternatifs ont un intérêt particulier pour l'élaboration de vaccins antigrippaux en cas de risque de pandémie car ils permettent de produire des semences vaccinales plus rapidement. En effet, à partir du moment où un virus potentiellement pandémique apparaît, il est difficilement envisageable de patienter six mois avant d'obtenir un vaccin pour protéger la population humaine. Ainsi, la vaccination contre la souche A(H1N1)pdm09 n'a débuté qu'à partir d'octobre 2009 aux États-Unis, après le pic de la seconde vague pandémique, alors que le virus avait été isolé en avril de la même année. Une étude estime qu'environ un million de cas d'infections ont été évités grâce à la campagne de vaccination pandémique, mais que ce chiffre aurait pu être quatre fois plus élevé si le vaccin avait été administré huit semaines plus tôt [232].

En raison de l'évolution permanente des IVA, de nombreuses équipes s'intéressent à l'élaboration d'un vaccin universel, permettant une protection à l'encontre des différents sous-types (pour revue, voir [233]). Ainsi, une dose vaccinale unique serait suffisante pour assurer une protection immunitaire pendant plusieurs années. Les travaux les plus

prometteurs concernent la protéine HA. En 2015, deux études ont été publiées à ce sujet [234, 235]. Yassine *et al.* [235] ont lié des nanoparticules à la tige de la protéine HA d'une souche H1N1, qui est très conservée entre les différents sous-types. La vaccination avec cette construction a montré que les souris produisent des anticorps à large spectre qui les protègent efficacement contre une souche H5N1. La protection n'est cependant que partielle chez les furets. Dans la seconde étude [234], les chercheurs ont développé une tige HA ayant la séquence de sous-type H1, reconnue par les anticorps neutralisants humains à large spectre. Après vaccination, les anticorps ont protégé des souris contre des virus H5N1, tandis que des singes ont partiellement été protégés. D'autres protéines virales peuvent également être des cibles potentielles pour l'élaboration d'un vaccin grippal universel, comme les protéines M (notamment M1 et M2e, le domaine extracellulaire de M2 localisé en N-terminal), NP et NA (pour revue, voir [236]).

## **II. Traitements antigrippaux**

Il existe actuellement deux familles d'antiviraux approuvés par la FDA dans les traitements contre les infections aux IVA.

### **1. Inhibiteurs du canal ionique M2**

Les adamantanes (amantadine (Mantadix®) et rimantadine (Flumadine®)) inhibent le canal ionique M2, empêchant ainsi la libération du génome viral dans le cytoplasme de la cellule infectée, soit en se fixant au niveau d'un canal ionique, ce qui provoque son obstruction [237], soit en stabilisant la conformation fermée du canal [238]. Ces deux molécules ont respectivement été autorisées pour un usage clinique depuis les années 1960 et 1980 et leur utilisation à grande échelle a entraîné le développement d'une résistance de la majorité des souches circulantes contemporaines d'IVA. Une étude récente a identifié la fréquence et la distribution des mutations dans le segment M2 de plus de 31 000 souches résistantes aux adamantanes ayant circulées dans le monde entre 1902 et 2013 [239]. Les résultats ont révélé que, de manière générale, 45% des souches sont résistantes et la majorité d'entre elles possèdent la mutation S31N, qui joue donc probablement un rôle primordial dans le mécanisme de résistance. D'autres mutations sont présentes de manière plus limitée (V27A est retrouvée dans environ 1% des souches résistantes, alors que la prévalence des substitutions L26F, A30T, G34E et L38F est inférieure à 0,2%). De manière

surprenante, certains sous-types de HA sont plus susceptibles de développer une résistance aux adamantanes (H1, H3, H5, H7, H9 et H17), tandis que pour d'autres, l'apparition de telles mutations est plus rare (H2, H4, H6, H10 et H11), voire inexistante (H8, H12–H16) [239]. Depuis 2003, le nombre de souches H1N1 et H3N2 possédant la mutation de résistance S31N n'a cessé de croître et actuellement la résistance concerne la majorité des souches circulantes [240]. De même, les souches résistantes H5N1 sont de plus en plus fréquentes depuis 2000 et aujourd'hui, dans certains pays asiatiques, la totalité des souches circulantes sont résistantes. Le virus pandémique A(H1N1)pdm09 est également résistant vis-à-vis de ces molécules [240].

## **2. Inhibiteurs de la neuraminidase**

À l'heure actuelle, en raison de la résistance croissante vis-à-vis des adamantanes, les traitements antigrippaux sont dirigés contre la neuraminidase (oseltamivir (Tamiflu®) et zanamivir (Relenza®)). Ces molécules sont des analogues structuraux des acides sialiques et se fixent au niveau du site actif de la protéine NA [241], inhibant ainsi de manière réversible son activité enzymatique. Les virions néo-formés s'agrègent donc à la surface des cellules infectées et leur diffusion au sein de l'organisme est ralentie. Certaines souches virales développent des résistances liées à des mutations au niveau des acides aminés du site actif et/ou d'acides aminés structuraux, qui stabilisent le site catalytique. Fin 2007, des virus saisonniers H1N1 ont acquis la substitution H275Y dans la protéine NA, leur conférant une résistance à l'oseltamivir [242]. Des souches (H1N1)pdm09 possédant cette mutation ont également été isolées, mais de manière plus localisée [243, 244]. Plus de 99% des souches virales circulantes H3N2 et (H1N1)pdm09 collectées entre mai 2014 et mai 2015 dans le monde (environ 13 000 virus) sont sensibles aux inhibiteurs de la neuraminidase [245]. À ce jour, les molécules ciblant cette protéine virale sont donc les seuls traitements efficaces contre les infections aux IVA. Néanmoins, pour que le Tamiflu® soit efficace, il est nécessaire de débiter le traitement au plus tard 48h après le début des symptômes grippaux, ce qui constitue un inconvénient majeur.

Actuellement, d'autres stratégies antivirales (dirigées contre la protéine NP et le complexe ARN polymérase) sont en cours de développement (pour revue, voir [246]).

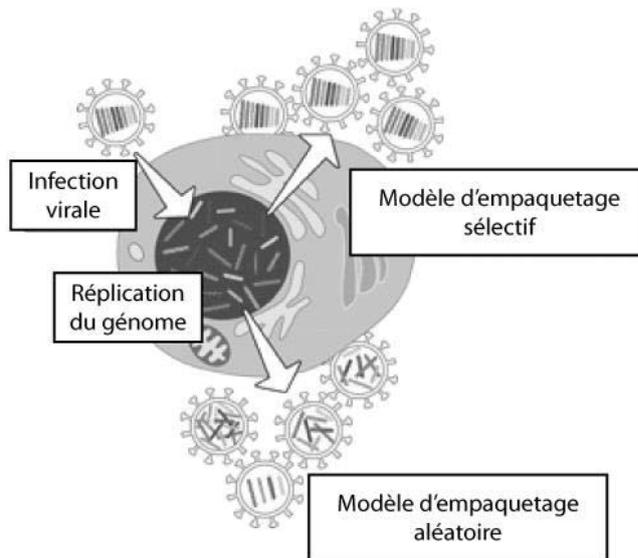
## **D. Empaquetage de génomes segmentés et réassortiment génétique, deux processus étroitement liés**

La segmentation d'un génome viral permet aux virus d'évoluer rapidement par réassortiment génétique, comme décrit précédemment (§ B.I.2). Néanmoins, la nature segmentée du génome complique incontestablement son empaquetage au sein des virions. Chez l'Homme, la transmission des IVA se fait principalement *via* les aérosols, qui contiennent néanmoins peu de particules virales. L'infection est donc réalisée à faible multiplicité d'infection (MOI, *multiplicity of infection*), c'est-à-dire qu'en début d'infection une cellule ne sera pas co-infectée et qu'il n'y aura donc pas de *trans*-complémentation entre particules défectives. Par conséquent, pour être pleinement compétents d'un point de vue répliatif et infectieux, les virus doivent posséder au moins une copie de chacun des huit ARNv, qui codent tous pour au moins une protéine essentielle (tableau 1). Il est donc indispensable que les ARNv génomiques soient empaquetés de manière à assurer la présence d'un jeu complet de huit segments dans les particules virales.

### **I. Empaquetage des génomes viraux segmentés**

#### **1. Empaquetage du génome des *Influenzavirus A***

Deux modèles pour l'empaquetage des génomes grippaux ont longtemps été proposés : un modèle d'empaquetage aléatoire et un modèle d'empaquetage sélectif (figure 17). Les deux modèles nécessitent de discriminer les ARNv de polarité négative des ARNm viraux de polarité positive et des ARN cellulaires. Le second modèle doit également discriminer les huit segments d'ARNv (-) entre eux et implique donc la présence de signaux d'empaquetage spécifiques et différents sur chaque ARNv. De nombreuses données structurales, génétiques, biochimiques et virologiques accumulées au cours des deux dernières décennies, plaident désormais en faveur d'un empaquetage sélectif des segments génomiques. Néanmoins, les mécanismes moléculaires permettant de réaliser cette sélection spécifique commencent à peine à être dévoilés. Au cours de mon doctorat, j'ai participé à l'écriture de deux revues concernant les données récentes privilégiant l'existence d'un mécanisme d'empaquetage spécifique des IVA [33, 247]. La plus détaillée de ces revues est incluse dans mon introduction (§ D.I.1.2). Les résultats les plus récents ont été ajoutés à la suite.



**Figure 17 : Modèles possibles du mécanisme d'empaquetage du génome des IVA** (adapté de [248]). Deux modèles ont été proposés : un mécanisme d'empaquetage sélectif, qui implique l'incorporation de huit segments d'ARNv différents, et un mécanisme aléatoire, dans lequel les ARNv génomiques sont empaquetés aléatoirement dans les virions, générant une très faible proportion de virions infectieux.

### 1.1. La génétique inverse, un outil déterminant dans l'étude des mécanismes d'empaquetage

La production de virus influenza par génétique inverse a débuté en 1989, lorsque l'équipe de P. Palese a transfecté des cellules eucaryotes avec des RNPv, reconstituées à partir de protéines PB2, PB1, PA et NP purifiées et d'un ARNv recombinant transcrit *in vitro* [249]. L'infection de ces mêmes cellules avec un virus « helper » WSN a permis la réplication des RNPv et leur empaquetage dans les particules virales. Il a ensuite fallu attendre 1999 pour que les premiers IVA soient générés entièrement à partir d'ADNc codant les huit segments d'ARNv [250]. Dans ce système, des cellules sont co-transfectées avec huit plasmides permettant l'expression des ARNv génomiques et avec au minimum quatre autres plasmides permettant l'expression des protéines virales essentielles à la réplication (PB2, PB1, PA et NP). Par la suite, d'autres systèmes de génétique inverse ont été développés, nécessitant la transfection de trois à seize plasmides (pour revue, voir [251]). Plus récemment, Chen *et al.* [252] ont proposé une nouvelle stratégie de production d'IVA basée sur la co-transfection d'amplicons linéaires construits par PCR. Le système le plus utilisé actuellement pour générer des IVA fait intervenir un jeu de huit plasmides d'expression bidirectionnels co-transfectés dans des cellules humaines HEK293T [253]. L'ADNc codant chacun des segments génomiques est cloné dans un plasmide pHW2000 et est à la fois sous la dépendance du promoteur PolII, permettant l'expression des ARNv de polarité négative et du promoteur PolIII, pour la synthèse des ARNm viraux et donc des protéines virales.

La génétique inverse s'est révélée être un outil clef pour la compréhension de nombreux aspects de la biologie des IVA, en particulier l'empaquetage du génome viral. Le contour des signaux d'empaquetage a initialement été délimité de manière grossière par l'introduction de délétions dans les plasmides permettant l'expression des ARNv sous le contrôle du promoteur Poll [108, 254–259]. De la mutagénèse ciblée a ensuite permis de déterminer plus précisément les séquences impliquées dans l'empaquetage spécifique de certains ARNv [107, 254, 258, 260–262]. Dans les deux cas, l'efficacité avec laquelle des ARNv mutants sont incorporés dans des VLP ou dans des virions infectieux peut être mesurée par une variété de méthodes (pour revue, voir [263]).

## **1.2. Article de revue : empaquetage sélectif du génome des *Influenzavirus A***

# Empaquetage sélectif du génome des influenza virus A

Marie Gerber<sup>1</sup>  
Vincent Moulès<sup>2,3</sup>  
Cyrille Gavazzi<sup>1</sup>  
Roland Marquet<sup>1</sup>  
Catherine Isel<sup>1</sup>

<sup>1</sup> Architecture et réactivité de l'ARN, Université de Strasbourg, CNRS, IBMC, 15 rue René-Descartes, 67084 Strasbourg, France  
<r.marquet@ibmc-cnrs.unistra.fr>  
<c.isel@ibmc-cnrs.unistra.fr>

<sup>2</sup> Virologie et pathologie humaine, Université Lyon 1, EA4610, Faculté de médecine RTH Laennec, 69008 Lyon, France

<sup>3</sup> VirNext, Université Lyon 1, EA4610, Faculté de médecine RTH Laennec, 69008 Lyon, France

**Résumé.** Le génome des influenza virus de type A (IVA) est constitué de huit segments, tous essentiels à la réplication virale. L'observation des virions par microscopie électronique et la reconstruction tridimensionnelle de l'intérieur des particules virales par tomographie électronique, associées à de nombreuses études génétiques, biochimiques et virologiques, ont montré que l'empaquetage du génome se fait de manière sélective plutôt qu'aléatoire. Les mécanismes moléculaires de ce processus d'assemblage sélectif commencent à être élucidés : les huit ARN viraux forment vraisemblablement un complexe supramoléculaire maintenu par des appariements entre les signaux d'empaquetage présents sur chacun d'entre eux, lesquels diffèrent selon les souches d'IVA. La reconstruction tridimensionnelle de ribonucléoprotéines virales et leur visualisation, par imagerie, dans les cellules, apportent de nouvelles informations concernant la localisation et l'assemblage spatio-temporel de ces complexes supramoléculaires, avant leur empaquetage dans les particules virales bourgeonnantes. L'ensemble de ces données améliore notre compréhension des règles qui gouvernent l'empaquetage du génome des IVA et permettent d'envisager une optimisation de la production des semences vaccinales. Les potentialités de réassortiment génétique entre différents virus, source de graves pandémies, semblent également être affectées par les règles qui guident la sélection et l'assemblage du génome viral.

**Mots clés :** influenza virus A, empaquetage des ARN viraux, interactions ARN/ARN, complexe supramoléculaire

**Abstract.** Electron microscopy of influenza A virus (IAV) and three-dimensional reconstruction of their interior by electron tomography, combined with genetic, biochemical and virology assays, has revealed that genome packaging of IAVs is a selective process, the molecular mechanisms of which start to be unveiled. The eight genomic viral RNAs (vRNAs) most likely form a supramolecular complex maintained by base-pairings within the strain-dependent packaging signals of each vRNA. Visualization of viral ribonucleoproteins inside cells also brought new insights about spatio-temporal assembly of the supramolecular complexes, prior to their incorporation into budding virions. Altogether, these data improve our understanding of the rules governing packaging of the IAV genome and offer clues for optimization of vaccine seeds production. Genetic reassortment events between different IAVs, which can lead to severe pandemics, are probably also affected by the rules that govern genome packaging.

**Key words:** influenza A virus, packaging, RNA/RNA interactions, supramolecular complex

## Introduction

Les influenza virus de type A (IVA) sont les agents étiologiques des épidémies grippales saisonnières et

occasionnellement de graves pandémies humaines et zoonotiques. L'Organisation mondiale de la santé (OMS) considère que ces virus restent un problème de santé publique majeur, puisque les infections saisonnières par les IVA sont associées à une mortalité et une morbidité élevées, avec 250 000 à 500 000 décès dans le monde chaque année

**Tirés à part :** R. Marquet

(<http://www.who.int/mediacentre/factsheets/fs211/en/>). La première pandémie bien documentée, la grippe espagnole de 1918, de sous-type H1N1, responsable d'environ 50 millions de morts, reste la plus meurtrière. Elle a été suivie par les pandémies de grippe asiatique de 1957 (H2N2) et de Hong-Kong en 1968 (H3N2) [1]. La première pandémie grippale du XXI<sup>e</sup> siècle, la grippe mexicaine de 2009 d'origine porcine, a eu elle aussi un impact considérable sur la santé et l'économie [2]. Les cas récents d'infections occasionnelles d'humains par des virus hautement pathogènes d'origine aviaire (*Highly Pathogenic Avian Influenza* [HPAI]), dont le virus H7N9 en Chine en 2013, rappellent le danger constant que font encourir les IVA à la population et prouvent que nous ne sommes pas à l'abri de nouvelles pandémies. Ainsi, les virus humains ne sont pas les seuls à faire l'objet d'une surveillance très active à travers le globe. Des HPAI de sous-type H5N1 sont répandus en Asie, en Europe et en Afrique [3], même s'ils sont uniquement endémiques dans les élevages de volailles de quelques pays. Cette vaste dissémination a été à l'origine de cas occasionnels de transmission inter-espèce chez les mammifères et de 650 cas d'infections, dont 386 mortelles, chez l'homme ([http://www.who.int/influenza/human\\_animal\\_interface/EN\\_GIP\\_20140124CumulativeNumberH5N1cases.pdf](http://www.who.int/influenza/human_animal_interface/EN_GIP_20140124CumulativeNumberH5N1cases.pdf) et [3]). Le franchissement de la barrière d'espèce, s'il reste rare pour l'instant, est l'occasion pour le virus de s'adapter

à son nouvel hôte. En effet, des mutations sur l'HA détournent la fixation préférentielle d'un virus H5N1 HPAI aux acides sialiques de type  $\alpha$ -2,3 des cellules épithéliales du tractus digestif des oiseaux et du système respiratoire inférieur des hommes vers les acides sialiques de type  $\alpha$ -2,6 des cellules épithéliales du système respiratoire supérieur humain, favorisant ainsi une transmissibilité par aérosols [4-8]. Ces mutations, partiellement retrouvées chez le virus émergent H7N9, favorisent la transmission interhumaine des virus aviaires et pourraient ainsi être à l'origine de très graves pandémies [9].

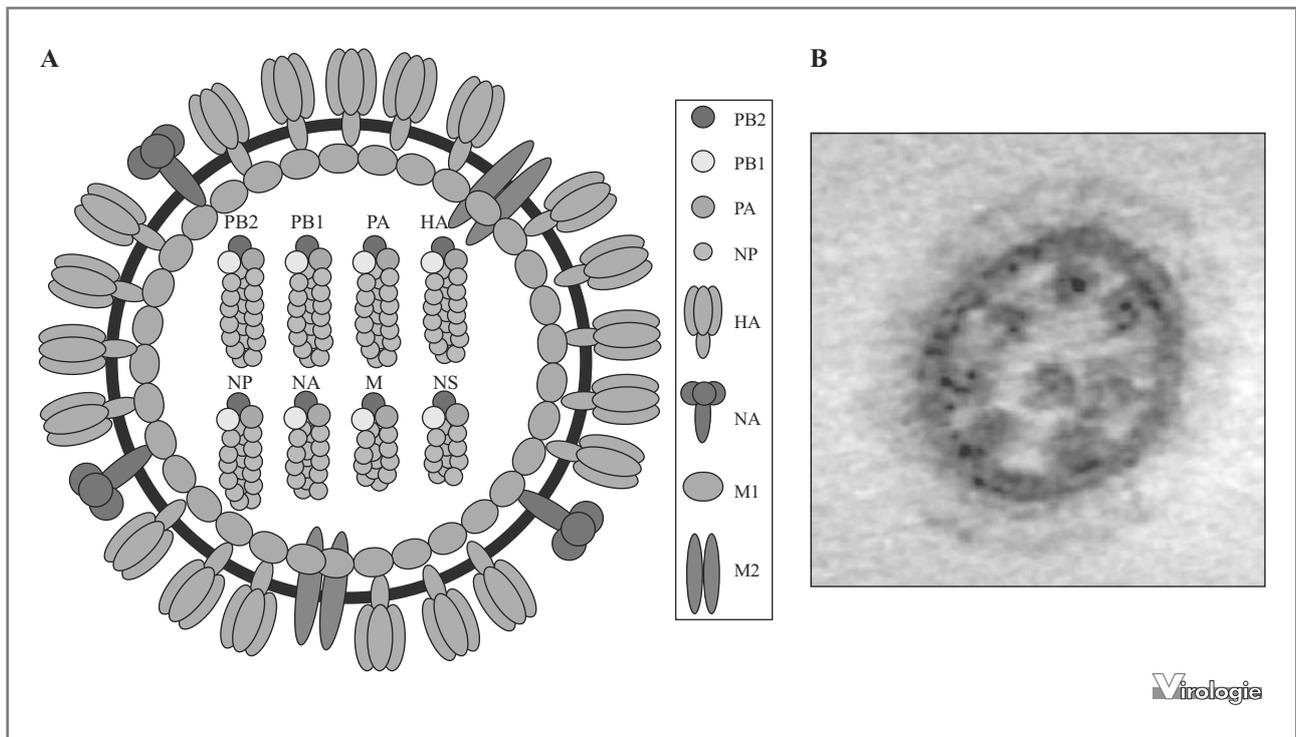
## Composition et structure des influenza virus de type A

Les IVA appartiennent à la famille des *Orthomyxoviridae*, qui comporte les trois genres des *Influenzavirus* (*A*, *B* et *C*), les *Thogotovirus*, un *Isavirus* et un nouveau genre, nommé *Quarjavirus* [10]. Une des caractéristiques des IVA est leur génome multipartite composé de huit segments d'ARN viraux (ARNv) simple brin de polarité négative dont la taille varie de 890 à 2341 nucléotides (nt). Les segments sont numérotés de 1 à 8 ou nommés d'après la protéine principale pour laquelle ils codent (*tableau 1*). Au centre de virions enveloppés, filamenteux ou sphériques [11] (*figure 1*), se trouvent les huit ARNv recouverts par

**Tableau 1** Segments génomiques des influenza virus de type A (IVA), protéines codées par ces segments et leurs fonctions.

ARNv	Nom de la protéine	Fonction principale de la protéine
1	PB2	Reconnaissance et liaison de la coiffe
2	PB1 PB1-F2* N40*	ARN polymérase ARN-dépendante Activité pro-apoptotique Inconnue
3	PA PA-X* PA-N155* PA-N182*	Endonucléase (impliquée dans le « vol de coiffe ») Répression de l'expression des gènes transcrits par l'ARN polymérase cellulaire II Probablement impliquée dans la réplication virale Probablement impliquée dans la réplication virale
4	HA	Liaison au récepteur cellulaire et fusion de la membrane virale et de l'endosome
5	NP	Liaison aux ARNv et export des RNPv
6	NA	Activité neuraminidase (clivage entre HA et l'acide sialique)
7	M1 M2 M43*	Protéine de matrice, export des RNPv et rôle dans le bourgeonnement Canal ionique (important pour l'acidification du virion et la libération des RNPv) Peut fonctionnellement remplacer M2
8	NS1 NS2-NEP NS3*	Inhibition de la réponse antivirale Export nucléaire des RNPv Apporte un avantage réplcatif

Les protéines auxiliaires sont indiquées par des astérisques.



**Figure 1. A)** Représentation schématique d'une particule virale d'un influenza A virus (IVA). Les protéines antigéniques HA et NA sont insérées dans l'enveloppe lipidique du virus. La protéine de matrice M1 recouvre l'intérieur de l'enveloppe. Le canal ionique M2 traverse l'enveloppe virale et la matrice. Les huit ARNv, sous forme de RNPv, au centre de la particule, constituent le génome viral et sont associés aux protéines du complexe polymérase hétérotrimérique (PB2, PB1 et PA) et recouverts de nucléoprotéine NP. **B)** Organisation des RNPv au sein de virions de la souche humaine A/Moscow/10/99 (H3N2). Coupe transversale d'un virion, visualisé par microscopie électronique. Les huit RNPv adoptent l'arrangement canonique en « 7 + 1 ».

la nucléoprotéine NP et associés au complexe polymérase hétérotrimérique constitué des sous-unités PB1, PB2 et PA. La sous-unité PB1 porte l'activité polymérase [12], la sous-unité PB2 est impliquée dans la fixation de la coiffe des ARNm cellulaires [13] et la sous-unité PA possède une activité endonucléase [14, 15]. Ces deux dernières sous-unités sont impliquées dans le mécanisme de vol de coiffe des ARNm cellulaires pour la synthèse des ARNm viraux. L'hémagglutinine (HA) et la neuraminidase (NA) forment des spicules à la surface de l'enveloppe des virions. L'HA est nécessaire à l'entrée du virus par reconnaissance des récepteurs cellulaires et fusion de la membrane virale avec celle de l'endosome. La NA permet la libération des virions néoformés dans le milieu extracellulaire. Les protéines de matrice M1 recouvrent l'intérieur de l'enveloppe et les protéines M2, qui traversent l'enveloppe lipidique et la matrice, forment des canaux ioniques responsables de l'acidification de l'intérieur du virus à partir de l'endosome et la libération des ribonucléoprotéines virales (RNPv) dans la cellule. L'ARNv 8 code deux protéines, NS1 et

NS2/NEP. La protéine NS1, absente des virions, est un facteur de virulence pléiotropique qui inhibe à la fois la production d'interféron et l'activité de certains gènes induits par l'interféron (pour revue voir [16]). NS1 inhibe également la maturation et l'export nucléaire des ARNm cellulaires et pourrait détourner la machinerie de traduction cellulaire pour favoriser l'expression des protéines virales (pour revue voir [16]). La protéine NS2, rebaptisée NEP (*Nuclear Export Protein*), présente en petite quantité dans les virions, est impliquée dans l'export des RNPv du noyau vers le cytoplasme. Plus récemment, NEP a été impliquée dans le processus de bourgeonnement. Enfin, NEP favorise également la transition entre l'étape précoce de transcription des gènes viraux et l'étape de réplication, au cours de laquelle sont produites les RNPv (pour revue voir [17]). Les deux glycoprotéines antigéniques définissent le sous-type des IVA : il existe 18 sous-types de protéine HA (H1 à H18) et 11 sous-types de protéine NA (N1 à N11) [18]. Seuls quelques sous-types de HA (H1, H2 et H3) et de NA (N1 et N2) ont été retrouvés parmi les virus humains, tandis que les sous-types H1-H16 et

---

N1-N9 sont tous présents chez les virus infectant les oiseaux aquatiques sauvages, réservoir naturel des IVA [19]. Les virus H17N10 et H18N11 ont été identifiés très récemment chez des chauves-souris [18, 20].

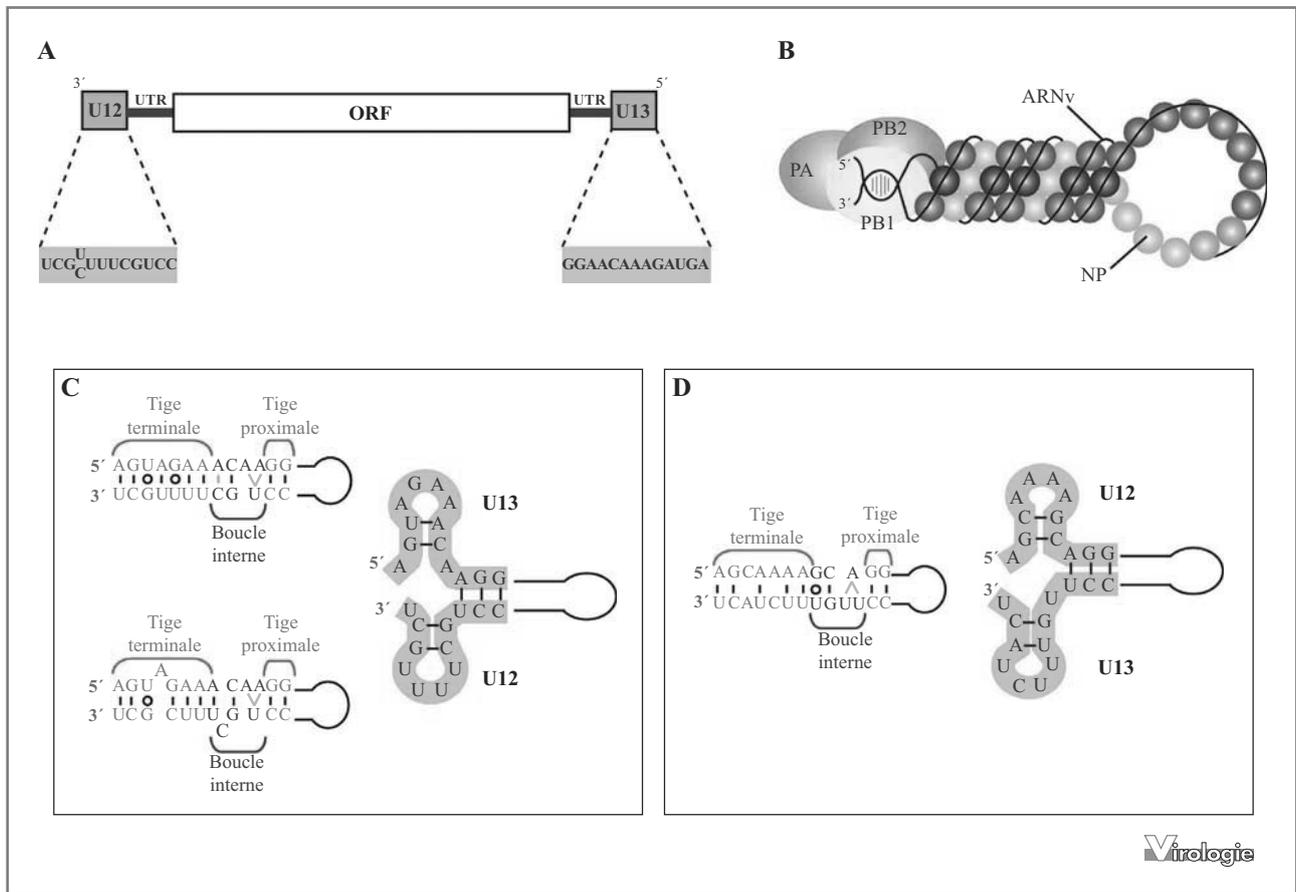
En plus des protéines essentielles (*tableau 1*), le génome des IVA code aussi des protéines auxiliaires récemment découvertes, qui ne sont pas requises pour la réplication virale [21-28] (*tableau 1*). La protéine PB1-F2, présente chez la plupart des IVA, est codée par un cadre de lecture alternatif de l'ARNv 2 et traduite par un mécanisme de « *leaky ribosomal scanning* ». Cette protéine ne semble pas être incorporée dans les virions [29], mais elle a de nombreuses fonctions au sein des cellules infectées (pour revue voir [30]). En effet, elle régule l'activité du complexe polymérase [31], semble impliquée dans l'induction de l'apoptose des cellules immunitaires suite à l'infection par un IVA [22], augmente la pathogénèse virale *in vivo* en dérégulant les cytokines et en induisant l'inflammation, augmente le risque d'infections secondaires bactériennes [32] et empêche la transcription de certains interférons en se liant à des protéines cellulaires [33-35]. De nombreuses études révèlent que l'expression et le rôle de PB1-F2 diffèrent selon la souche, l'hôte ou le type cellulaire (pour revue voir [30]). L'ARNv 2 code également une troisième protéine, appelée N40, qui est un variant de PB1 tronqué des 39 acides aminés en N-terminal [27]. L'initiation de sa traduction est possible par un mécanisme de ré-initiation qui permet de court-circuiter les codons d'initiation précédents [26]. Son rôle n'a toujours pas été identifié clairement, mais il est toutefois distinct des fonctions de la protéine PB1, puisque N40 ne permet pas la transcription *in vitro* des ARNv et n'est pas capable d'interagir avec PA, cette interaction étant indispensable pour l'import nucléaire et l'activité de la polymérase [27]. La protéine PA-X, identifiée en 2012, est codée par l'ARNv 3 et possède 191 acides aminés en N-terminal dérivés de PA, alors que son extrémité C-terminale provient d'un décalage du cadre ouvert de lecture [23]. PA-X réprime l'expression des gènes transcrits par l'ARN polymérase II cellulaire dans des systèmes rapporteurs et réduit la pathogénie dans un modèle murin [23]. L'ARNv 3 code également deux autres protéines accessoires, PA-N155 et PA-N182, traduites respectivement à partir des 11<sup>e</sup> et 13<sup>e</sup> codons AUG [24], tous deux très conservés parmi les IVA. Les deux protéines ont été détectées dans des cellules infectées par des IVA d'hôtes différents, suggérant que leur expression est universelle parmi les IVA. Elles n'ont pas d'activité polymérase mais semblent jouer un rôle important dans le cycle répliatif [24]. Wise *et al.* ont également mis en évidence la protéine M42, dérivant de l'ARNv 7, dont l'ectodomaine et la localisation intracellulaire sont distincts de la protéine M2, mais qui peut toutefois compléter fonctionnellement M2 *in vitro* et *in vivo* [28].

Enfin, l'épissage alternatif de l'ARNv 8 génère un ARNv codant la protéine NS3, tronquée d'une séquence centrale par rapport à la protéine NS1, qui augmente la réplication virale [25].

## Organisation génétique des ribonucléoprotéines virales

Les ARNv des IVA, bien qu'ils soient de tailles différentes, sont tous organisés de la même manière (*figure 2A*) : une phase de lecture ouverte centrale, en orientation antisens, codant une ou plusieurs protéines, est bordée de régions non traduites (*UnTranslated Regions* [UTR]) relativement courtes (9 à 58 nt), différentes pour chaque ARNv. Aux extrémités 3' et 5' de chaque ARNv se trouvent deux séquences, de 12 (Uni12, U12) et 13 (Uni13, U13) nt respectivement, qui sont conservées entre les ARNv d'une même souche et entre les différentes souches virales [19, 36-40]. Les séquences U12 et U13, partiellement complémentaires, constituent le promoteur indispensable à la transcription et la réplication [19, 36-39] (*figure 2B, C*). Celui-ci s'apparie en imposant aux ARNv une structure dite en « poêle à frire » en l'absence de protéines [41-43] (*figure 2C*). Cependant, des expériences de génétique [44] et de biophysique [43] montrent qu'en présence des protéines constituant le complexe polymérase, le promoteur adopte une structure en « tire-bouchon » (*figure 2C*). Les séquences en 3' et en 5' de l'ARNv se lient d'ailleurs à deux domaines distincts de la sous-unité PB1 de l'ARN polymérase virale, engendrant l'activation subséquente des sous-unités PB2 et PA [45]. Le reste de l'ARNv est recouvert par la nucléoprotéine NP pour former une RNPv (*figure 2B*) fonctionnant comme une unité indépendante pour la réplication et la transcription [36, 46-48].

La segmentation du génome des IVA favorise leur évolution en permettant le réassortiment génétique, c'est-à-dire l'échange de segments d'ARNv lorsque deux ou plusieurs virus co-infectent une même cellule. Cependant, elle complique indéniablement l'assemblage des virions, puisque, pour être pleinement compétent d'un point de vue répliatif et infectieux, un virion doit posséder au moins un exemplaire de chacun des huit ARNv. Au cours de la dernière décennie, de nombreux arguments sont venus étayer l'existence d'un mécanisme sélectif, plutôt qu'aléatoire, d'empaquetage du génome des IVA. Un assemblage sélectif nécessite la présence de signaux spécifiques sur chaque ARNv permettant de les différencier les uns des autres. Malgré l'identification progressive de signaux d'empaquetage sur chacun des huit ARNv de deux souches humaines de sous-type H1N1, ce n'est que depuis peu que les



**Figure 2.** Organisation des ribonucléoprotéines virales des influenza virus de type A (IVA). **A)** Organisation génétique commune des huit gènes des IVA. Chaque gène comporte une ou plusieurs séquences codantes centrales (ORF : *Open Reading Frame*), entourées de séquences transcrites non traduites (UTR : *UnTranslated Regions*) et de deux séquences uniques (U), conservées, de 12 (U12) et 13 (U13) nt, en 3' et en 5' respectivement. Les ARNv sont représentés selon la nomenclature conventionnelle pour les virus à ARN négatifs, de 3' en 5'. **B)** Organisation tridimensionnelle schématisée des RNPv d'après des données récentes de cryo-microscopie électronique [46, 48]. L'ARNv est représenté par un trait noir. Ses extrémités 3' et 5' partiellement complémentaires sont appariées et reconnues par le complexe polymérase hétérotrimérique, constitué des trois sous-unités PB1, PB2 et PA. Le reste de l'ARNv est associé à la nucléoprotéine NP. **C, D)** Structure des promoteurs des ARNv (C) et ARNc (D) des IVA. **C)** Les deux appariements de gauche, correspondant aux deux variants possibles de U12, représentent le modèle de repliement en « poêle à frire », en absence de protéines virales [36]. L'appariement de droite correspond au modèle de repliement en « tire-bouchon », en présence des protéines virales [43, 44]. **D)** L'appariement de gauche correspond à la représentation de la structure du promoteur selon le modèle en « poêle à frire », celui de droite à la représentation selon le modèle en « tire-bouchon ». Les appariements Watson-Crick sont représentés par des traits noirs, les interactions wobble G-U par un cercle et les interactions non canoniques par des traits gris.

mécanismes moléculaires gouvernant l'empaquetage des ARNv dans les virions commencent à être élucidés. Cette revue décrit les dernières avancées dans ce domaine.

### Un mécanisme d'empaquetage du génome sélectif, plutôt qu'aléatoire

Lors de l'empaquetage des ARNv dans les virions, les IVA doivent :

- discriminer les ARNv des autres ARN présents dans la cellule (les ARNv complémentaires (ARNc) de polarité positive, les ARN messagers viraux et la multitude d'ARN cellulaires) ;
- incorporer une copie de chaque ARNv.

Historiquement, deux modèles ont été proposés pour expliquer l'empaquetage des huit ARNv au sein de particules virales répliquatives et infectieuses [49]. Le modèle d'empaquetage aléatoire repose sur l'existence d'un mécanisme qui discrimine les ARNv (-) des autres ARN, mais pas

les huit ARNv entre eux. Ce modèle implique par conséquent l'existence de signaux communs à tous les ARNv. La probabilité d'assembler un jeu complet de huit ARN génomiques dans un modèle aléatoire (0,24 %) [50, 51] ne correspond pas à la proportion de virions infectieux, estimée entre 1 et 10 % [52]. Ce constat, ainsi que de nombreux autres éléments, détaillés ultérieurement, plaident désormais en faveur de l'existence de signaux spécifiques sur chaque ARNv et d'un empaquetage sélectif du génome des IVA [49, 53].

#### *Arguments génétiques et biochimiques en faveur d'un empaquetage sélectif*

En absence de *trans*-complémentation, la capacité répliquative d'un virus et son pouvoir infectieux requièrent la présence des huit segments d'ARNv. Selon des travaux récents, le défaut d'expression du jeu complet des protéines virales observé dans une grande majorité des cellules infectées par des IVA à faible multiplicité d'infection (*multiplicity of infection* [m.o.i]) serait dû à un défaut d'empaquetage des huit segments de gènes [54]. Cette observation pourrait cependant également s'expliquer par un défaut d'expression des gènes, en raison de cassures survenues au niveau des segments d'ARNv ou d'un dysfonctionnement du complexe polymérase. En effet, l'analyse par migration électrophorétique d'ARNv provenant de virions purifiés a permis de conclure à la présence de quantités équimolaires de chaque ARNv [55, 56]. Ces résultats ont été confortés par un taux équimolaire des huit segments observés dans les particules virales, malgré la différence de rapports entre ces ARNv dans les cellules infectées [57, 58]. De plus, d'après des études génétiques, les IVA possèderaient un génome haploïde [51, 59, 60]. Enfin, plus récemment, l'hybridation *in situ* en fluorescence (FISH) sur des virus uniques a démontré que la majorité des particules virales incorpore une seule copie de chaque ARNv [61]. En accord avec ces résultats, deux pseudo-ARNv, contenant la séquence codant soit pour la GFP soit pour la DsRed et bordée des séquences d'empaquetage terminales d'un même ARN (par exemple celui de NA) entrent en compétition pour leur empaquetage dans les virions avec genèse de particules virales possédant une copie unique du segment NA rapporteur et donc codant pour une seule des deux protéines fluorescentes [62].

#### *Vers l'identification des signaux d'empaquetage : les ARN défectifs interférents*

Des virus influenza défectifs interférents ou virus DI (*Defective Interfering*) ont été isolés à partir de cultures à haute m.o.i., mais aussi suite à des infections naturelles [63-72]. Ces virus DI non infectieux incorporent des ARN, appelés ARN interférents défectifs (ARN DI), dont la taille

est inférieure à celle des ARNv parentaux dont ils sont issus. En effet, ils sont dépourvus d'une grande partie des séquences codantes centrales mais conservent de manière systématique les UTR ainsi que des portions variables de 100 à 300 nt des deux extrémités de la région codante des ARNv parentaux dont ils dérivent [64-75] (*figure 3*). La grande majorité des ARN DI identifiés proviennent des ARNv 1, 2 et 3, mais des ARN DI dérivés de chacun des cinq autres segments ont également été observés (*figure 3*). De manière importante, les ARN DI sont répliqués de manière efficace en raison de leur taille réduite et entrent spécifiquement en compétition avec leurs ARNv parentaux pour l'empaquetage dans les virions [66, 67, 69, 71]. L'étude des ARN DI a apporté les premiers éléments étayant l'existence de signaux d'empaquetage spécifiques à chaque segment au sein des 100 à 300 nt de leurs extrémités.

#### *Identification de régions d'empaquetage, en cis, sur chaque ARNv (figure 4)*

L'utilisation des outils de la virologie moléculaire moderne, tout particulièrement la génétique inverse [76], sur la souche de laboratoire A/WSN/33 (H1N1) (WSN) et la souche A/Puerto Rico/8/34 (H1N1) (PR8) utilisée pour la production des vaccins, a permis ensuite de délimiter les régions des segments nécessaires à l'empaquetage de chacun des huit ARNv au sein des particules virales.

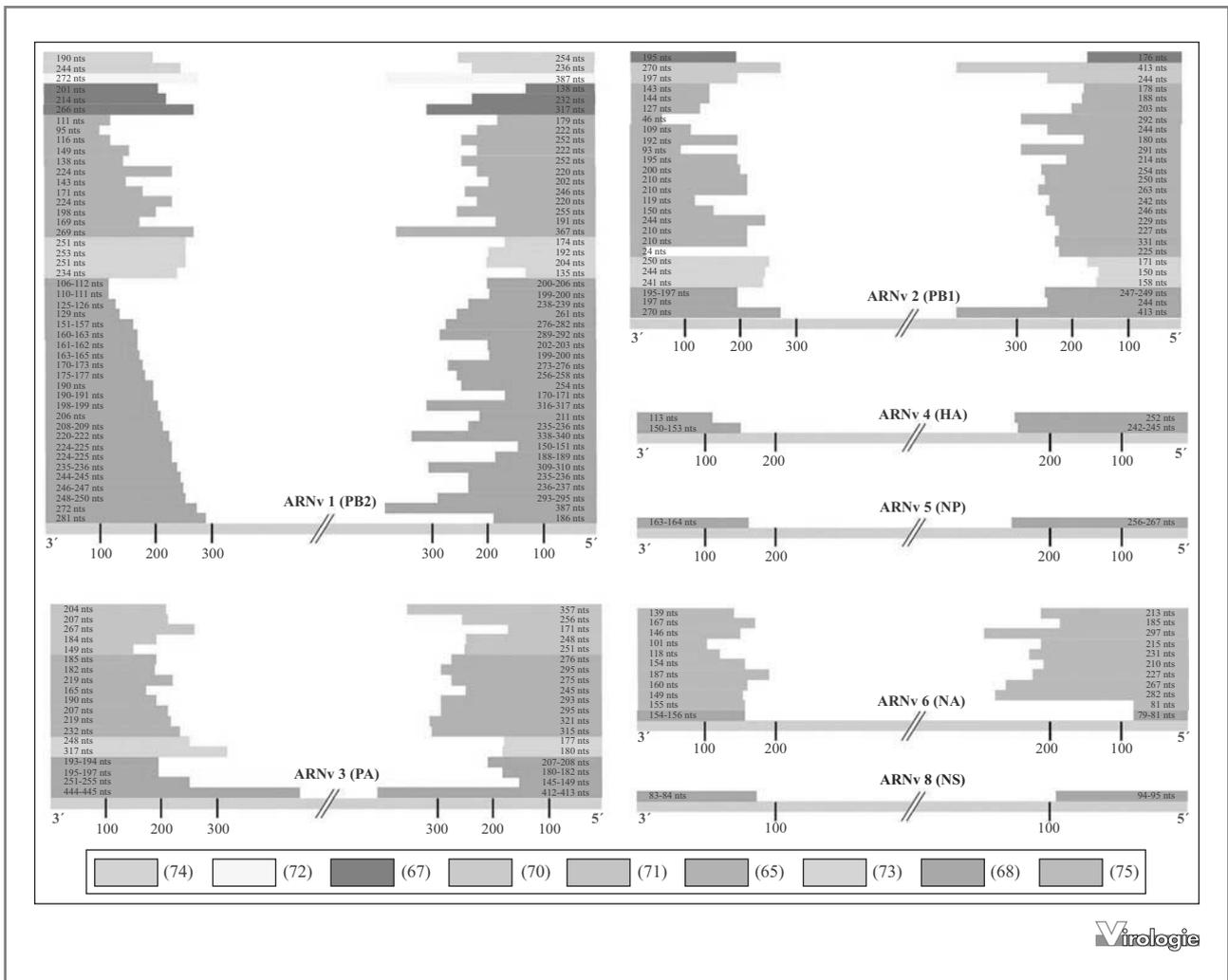
#### **Rôle des régions non codantes**

##### *Les régions U12 et U13, communes à tous les ARNv*

La discrimination, au moment de l'empaquetage, des ARNv (-) de tous les autres ARN non viraux, implique vraisemblablement des signaux communs à tous les segments d'ARNv des IVA. Les seules séquences qui répondent à ce critère sont les séquences conservées U12 et U13. Les équipes de P. Palese et Y. Kawaoka ont depuis longtemps démontré l'importance des régions non codantes (*Non Coding Regions* [NCR]), comprenant U12, U13 et les UTR, à chacune des extrémités des ARNv, pour leur empaquetage au sein de pseudo-particules virales [77, 78].

##### *Les régions non codantes spécifiques à chaque ARNv (tableau 2)*

En présence d'un virus « *helper* », les NCR de l'ARNv 8, entourant de part et d'autre une région codant un gène rapporteur, sont suffisantes pour la transcription, la réplication et l'empaquetage dans les virions de l'ARNv chimère [77]. Des résultats similaires ont également été obtenus pour les ARNv 1 [79], 3 [79], 4 [80, 81] et 5 [81]. Il faut cependant noter que les NCR seules ne permettent pas l'empaquetage efficace d'un segment chimère portant la



**Figure 3.** Régions conservées des ARN déficients interférents issus des ARNv des influenza virus de type A (IVA). Les tailles des régions conservées en 3' et 5' des ARN DI sont indiquées. Seuls des exemples des ARN DI générés par une délétion centrale unique au niveau de l'ARNv parental sont représentés ici. D'après [65, 67, 68, 70-75].

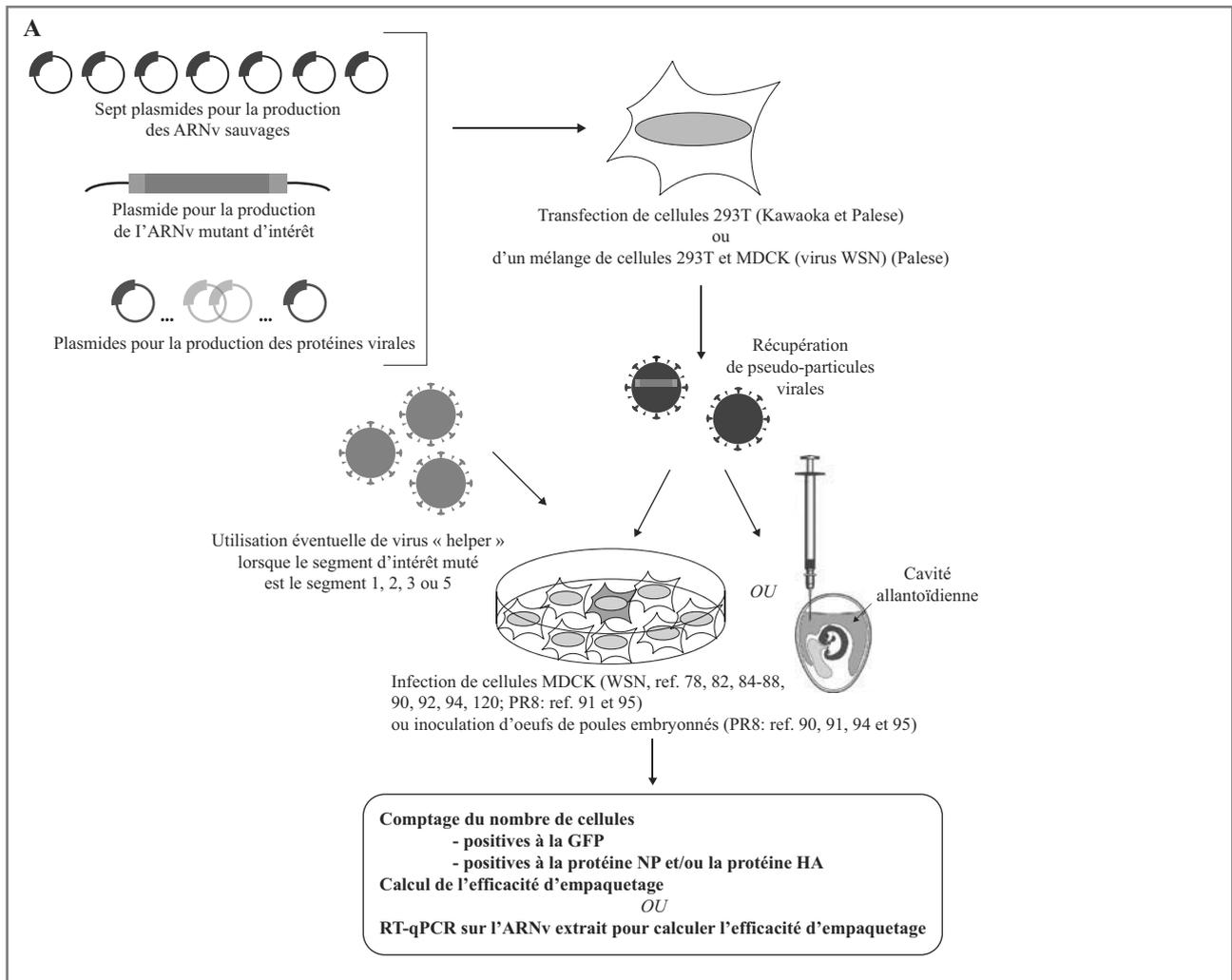
GFP en position centrale des segments 1-3 [78], 4 [82], 5 [83], 6 [84] et 8 [85], la seule exception étant l'ARNv 7 [86].

### Rôle des régions codantes

#### Régions codantes aux extrémités des ARNv (tableau 2)

Dans une étude pionnière, Fujii *et al.* [84] ont démontré qu'en plus des NCR de l'ARNv 6 de la souche WSN, les 183 nt en 3' et les 157 nt en 5' des séquences codantes, flanquant un gène rapporteur central, contribuent à l'empaquetage efficace d'un ARNv chimère dans les virions (figure 4A). Dans ces régions, les 23 nt en 3' de la séquence codante sont particulièrement importants.

Par la suite, l'utilisation de gènes rapporteurs a permis la délimitation des régions minimales nécessaires à l'empaquetage dans les virions de tous les segments (figure 4). Ainsi, dans le cas de l'ARNv 1 de la souche WSN, et d'après l'équipe de Kawaoka, en plus des NCR, les 30 nt en 3' de la région codante sont cruciaux [78], tandis que, d'après Parslow *et al.*, les 80 nt en 5' de cette même région [87] sont importants et plus particulièrement les nt 2248-2253, dont des mutations silencieuses affectent l'empaquetage [88] (figure 4B). En parallèle, d'après d'autres travaux, les 132 nt à chaque extrémité de l'ARNv 1 de la souche WSN sont importants et un signal essentiel à la réplication et/ou l'empaquetage de l'ARNv chevauche les régions codantes et les NCR en 5' [89].



**Figure 4.** Signaux d'empaquetage des ARNv des virus influenza humains A/WSN/33 (H1N1) (WSN) et A/Puerto Rico/8/34 (H1N1) (PR8). **A)** Stratégies expérimentales de détermination des signaux d'empaquetage. Ce schéma résume les stratégies employées par les groupes de Kawaoka, Palese, Parslow et Digard pour déterminer les efficacités d'incorporation des ARNv [78, 82, 84-88, 90-92, 94, 95, 120]. Des cellules en culture (293T ou un mélange de 293T et MDCK) sont transfectées avec des plasmides permettant l'expression des ARNv sauvages sous le contrôle du promoteur Pol I (bleu) et l'expression de l'ARNv muté d'intérêt (portant généralement la séquence codante de la GFP, en vert, à la place de la séquence codante d'origine ainsi que des portions variables des séquences codantes et/ou des UTR, en gris). Les protéines virales sauvages, dont le nombre varie selon les publications, mais qui incluent au minimum PB2, PB1, PA et NP, sont exprimées à partir de plasmides différents (marron) transfectés simultanément. Les surnageants contenant les pseudo-particules virales sont ensuite récoltés et utilisés pour infecter des cellules MDCK (WSN et PR8) ou sont injectés dans la cavité allantoïdienne d'oeufs de poule embryonnés (PR8). Dans le cas où les ARNv mutés sont ceux codant les protéines impliquées dans la formation des RNPv, des virus dits « helper » sont généralement utilisés pour apporter ces protéines sous forme fonctionnelle dans les cellules infectées. Lorsque la région centrale codante de l'ARNv d'intérêt a été remplacée par la GFP, l'efficacité d'incorporation de l'ARNv d'intérêt est généralement déterminée en divisant le nombre de cellules exprimant la GFP par la somme du nombre de cellules exprimant les protéines HA ou NP et du nombre de cellules exprimant la GFP. Alternativement, l'ARNv est extrait des virions et quantifié par RT-qPCR. **B, C)** Régions codantes et non codantes importantes pour l'empaquetage des ARNv des virus WSN (**B**) et PR8 (**C**). Les séquences U12 et U13 sont représentées par des rectangles noirs aux extrémités de chaque ARNv. Les flèches correspondent aux sites de démarrage de la traduction des protéines principales et les points noirs aux codons stop. Le code couleur pour les séquences importantes pour l'empaquetage des ARNv correspond aux publications ayant décrit ces séquences.

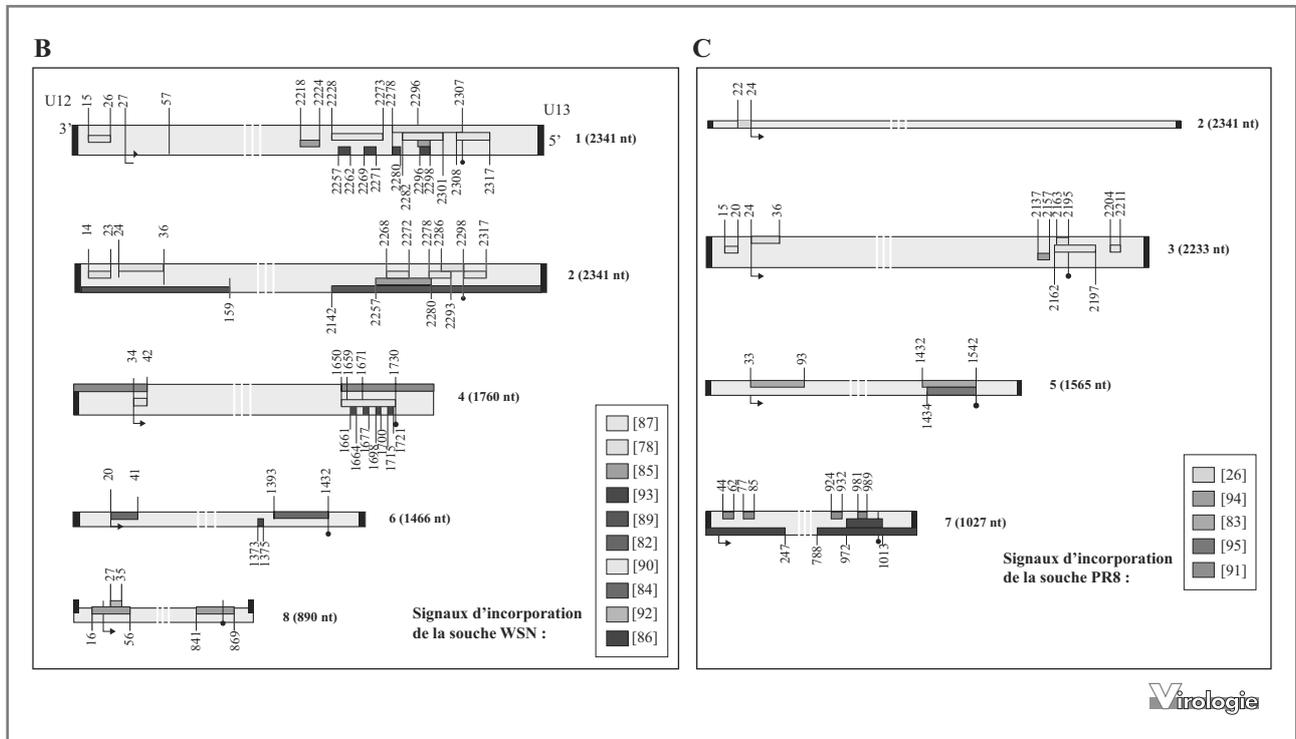


Figure 4. (Suite)

Dans le cas de l'ARNv 2 de la souche WSN, d'après l'équipe de Kawaoka, les NCR et les 12 nt en 3' et en 5' des régions codantes sont suffisants pour diriger l'empaquetage efficace d'un gène rapporteur [78]. D'après l'équipe de Parslow, des régions plus étendues (33 à 66 nt en 3' et 40 nt en 5') sont nécessaires [87] ; de plus des mutations ponctuelles des NCR et des régions codantes en 5', en particulier des nt 2 280 à 2 289, engendrent des effets beaucoup plus marqués que des mutations dans la région 3' de l'ARNv [88] (figure 4B).

Concernant l'ARNv 3 de la souche WSN, 12 et 21 nt respectivement en 3' et en 5' des régions codantes sont, en plus des NCR, importants pour l'empaquetage d'un gène rapporteur selon le groupe de Kawaoka [78]. Ces zones sont plus étendues d'après les travaux de Parslow et *al.*, puisqu'elles recouvrent 15 à 66 nt en 3' et 30 à 40 nt en 5' des régions codantes [87] (figure 4B). L'introduction de courtes délétions ou de mutations ponctuelles, en particulier au niveau des nt 2166-2172, en 5', a un effet majeur, contrairement à des délétions ou mutations en 3', dont l'effet reste plus modéré [88].

Dans le cas de l'ARNv 4 de la souche WSN, les résultats diffèrent aussi selon les publications. En effet, initialement, en plus des NCR, seuls 9 et 80 nt des extrémités 3' et 5' des régions codantes, respectivement, interviennent de manière significative dans l'empaquetage d'un gène

rapporteur [82]. Pourtant, d'après des travaux ultérieurs, 45 nt en 3' de la région codante sont absolument requis [90] (figure 4B). Il est à noter que des mutations silencieuses en 3' des régions codantes sur les ARNv 4 des souches PR8 et WSN n'affectent pas leur empaquetage de façon significative alors que des mutations en 5', surtout celles des nt 1 659-1 673, l'inhibent plus fortement (Figure 4B, C). Ces séquences sont par ailleurs très conservées entre virus aviaires, humains et porcins de sous-type H1 [90].

Concernant l'ARNv 5 de la souche WSN, les 60 et 120 nt, respectivement en 3' et en 5' de la région codante, permettent, avec les NCR, l'empaquetage d'un gène rapporteur [83] (figure 4B).

Les signaux d'empaquetage de l'ARNv 7 de la souche PR8 sont également localisés aux extrémités 3' et 5' de ses séquences codantes [91]. Des régions plus larges ont été délimitées dans le cas de l'ARNv 7 de la souche WSN, englobant, en plus des NCR, 222 et 220 nt en 3' et en 5', respectivement, des régions codantes [86]. Dans la partie 5' de la région codante, les nt 981 à 989 jouent vraisemblablement un rôle important [86] (figure 4B).

Enfin, dans les cas de l'ARNv 8 de la souche WSN, ce sont les 30 nt aux deux extrémités de la région codante qui dirigent, avec les NCR, l'empaquetage de l'ARNv au sein des virions [85]. De manière étonnante, il semblerait

**Tableau 2** Pourcentage d’empaquetage dans les particules virales de gènes rapporteurs flanqués des régions non codantes sans ou avec une partie des régions codantes présentes aux extrémités des ARNv.

ARNv	Importance des régions non codantes		Importance des extrémités des régions codantes		
	Empaquetage d’un gène rapporteur (%)	Références	Constructions	Empaquetage d’un gène rapporteur (%)	Références
1 (PB2)	0	[78]	(132)GFP(132) (120)GFP(120) (455)GFP(230)	27,5 69,7 70	[93] [78] [87]
2 (PB1)	1	[78]	(120)GFP(120) (433)GFP(627)	62,3 85	[78] [87]
3 (PA)	5,6	[78]	(120)GFP(120) (474)GFP(239)	57,3 91	[78] [87]
4 (HA)	3,9	[82]	(468)GFP(513)	42,8	[82]
5 (NP)	10	[83]	(120)GFP(120)	76	[83]
6 (NA)	1,1	[84]	(183)GFP(157)	91	[84]
7 (M)	32,1	[86]	(222)GFP(220)	47,7	[86]
8 (NS)	0,2-0,5	[85]	(150)GFP(160)	62-74	[85, 92]

(X)GFP(Y) signifie que le gène rapporteur de la GFP est flanqué des régions non codantes aux extrémités de l’ARNv ainsi que de X et Y nt respectivement en 3’ et 5’ de sa région codante.

que le signal d’empaquetage dans la région codante en 3’ de l’ARNv puisse tolérer une forte variabilité de séquence [92] (*figure 4B*).

#### *Un rôle pour les régions codantes centrales ?*

L’analyse des pourcentages d’empaquetage des gènes rapporteurs montre que, très souvent, ceux-ci restent relativement peu élevés, même lorsqu’une grande partie des extrémités des régions codantes est conservée (*tableau 2*). Ces résultats portent à croire que, dans certains cas, les régions centrales des gènes, remplacées par la GFP dans les systèmes expérimentaux utilisés, pourraient également comporter des signaux nécessaires à l’empaquetage des ARNv dans les virions.

#### **Importance des séquences conservées**

Les NCR et les extrémités des régions codantes de l’ARNv 8 de la souche WSN [85] et des ARNv grippaux en général sont plus conservées que les régions centrales. Cette observation a conduit à proposer que les codons très conservés au sein des différentes souches d’IVA pourraient constituer des signaux d’empaquetage. Cette proposition a été étayée par la constatation que les séquences importantes pour l’empaquetage du génome sont localisées dans des régions au moins partiellement conservées pour les ARNv 2, 3 et 4 [88, 90]. La démarche inverse, consistant à identifier, sur chacun des huit segments d’ARNv d’IVA de sous-types différents, les régions regroupant plusieurs codons portant un nombre de mutations synonymes plus faible que prévu par la redondance du code génétique, et à rechercher leur

rôle dans la réplication, a été entreprise par l’équipe de P. Digard [93]. Les codons conservés sont, en accord avec ce qui précède, majoritairement localisés aux extrémités des séquences virales et le degré de conservation diminue lorsque l’on s’en éloigne [93]. Ces résultats ont été confortés par une analyse parallèle sur les ARNv codant les protéines du complexe polymérase [94]. Pour les ARNv 1, 2, 3 et 6 de la souche WSN (et pour l’ARNv 3 de la souche PR8) [87, 88, 94], les ARNv 4 des souches WSN et PR8 [90] et les ARNv 7 [91] et 5 [95] de la souche PR8, la mutation synonyme des codons conservés affecte généralement plus fortement l’empaquetage du génome que celle des codons non conservés (*figure 4*). Cette règle n’est pas absolue, car la mutation de certains codons peu conservés de l’ARNv 8 de la souche WSN affecte son empaquetage [85, 93] et celle de certains codons conservés de l’ARNv 5 de la souche PR8 n’a pas d’effet [95].

De manière intéressante, le pourcentage d’empaquetage des ARNv diminue fortement suite à l’introduction de mutations ponctuelles dans des zones définies préalablement comme étant des signaux d’empaquetage alors que d’autres mutations, dans la même zone, n’ont pas d’impact [85, 88, 90, 91, 93-95] (*figure 4*). Il semble donc que les signaux d’empaquetage au niveau des extrémités des ARNv soient discontinus. Enfin, malgré un pourcentage d’identité variant entre 94 et 97 % selon les segments, certains codons importants pour l’empaquetage de l’ARNv 3 de la souche WSN ne le sont pas dans le cas de la souche PR8, suggérant que les signaux d’empaquetage ne soient pas tous conservés parmi les différentes souches [94].

### Existence de réseaux d'interactions

De manière intéressante, certaines mutations dans les régions codantes d'un ARNv donné affectent non seulement son propre emballage dans les virions, mais également celui d'autres ARNv [90, 91, 94, 95], suggérant l'existence d'un réseau d'interactions entre les ARNv [49]. Par ailleurs, certains ARNv sont plus importants que d'autres dans le mécanisme d'emballage des segments d'ARNv et de l'assemblage des virions [78, 91, 94, 95], ce qui laisserait penser que l'emballage des ARNv est un processus hiérarchique [49]. Ainsi, l'absence du segment 1 de la souche WSN affecte plus fortement l'emballage du génome que celle des segments 2 et 3 [78]. Les séquences conservées dans les régions 5' des ARNv 1 et 2 sont importantes pour l'emballage des segments 2, 3, 5, 7 et 8 et des segments 1, 5 et 7 des souches WSN et PR8, respectivement [94] ; au contraire, des mutations de séquences conservées de l'ARNv 3 n'ont que peu d'effet sur l'emballage des autres ARNv des souches WSN et PR8 [94]. Par ailleurs, des éléments spécifiques dans le segment 7 sont cruciaux pour l'emballage de tous les autres ARNv de la souche PR8 [95] et le segment 5 est indispensable à l'insertion de l'ARNv 3 dans les virions PR8 [95]. Enfin, une étude systématique sur la souche PR8 a révélé que les segments 1, 3, 5 et 7 jouent un rôle plus important que les segments 2, 4 et 8 dans le processus d'emballage du génome [96]. Il est intéressant de noter que les segments 4 et 6, codant respectivement les protéines antigéniques HA et NA, ne semblent pas jouer de rôle majeur dans l'emballage des autres ARNv. Cette observation pourrait refléter un processus évolutif qui faciliterait l'échange, par réassortiment génétique, des segments de gènes codant les protéines de surface, cibles de la réponse humorale.

Il faut également souligner que les délétions/mutations des signaux d'emballage n'entraînent jamais une perte totale de la réplication virale. Ces résultats suggèrent l'existence potentielle de redondances au niveau des signaux d'emballage et/ou la présence éventuelle de signaux dans les régions centrales. Ces régions ne sont en effet jamais testées lorsque des gènes rapporteurs sont utilisés dans les systèmes expérimentaux de génétique inverse.

### Des réseaux d'interactions entre les ARNv des influenza virus de type A maintenus par des interactions directes

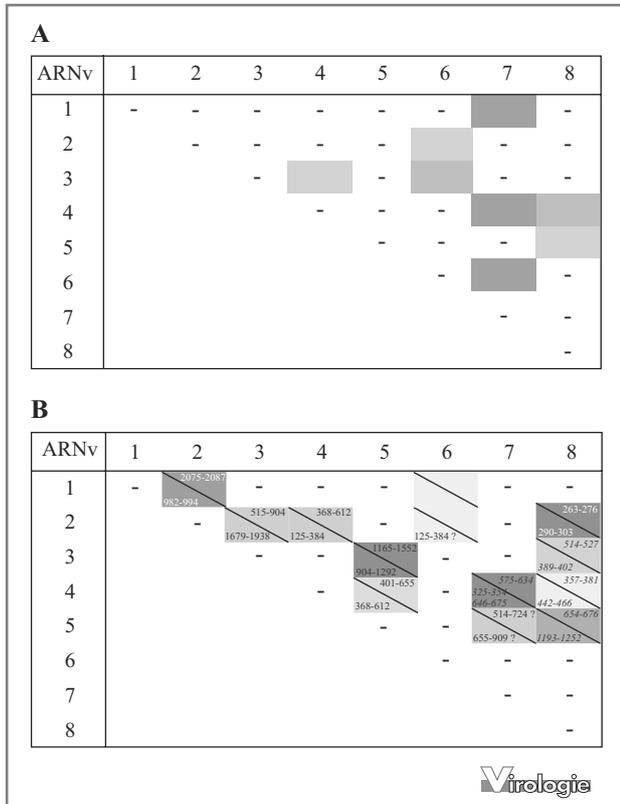
Même si l'existence, sur chaque ARNv, de signaux spécifiques dirigeant l'emballage du génome des IVA dans les particules virales est clairement démontrée, les mécanismes moléculaires responsables de ce processus d'assemblage sélectif étaient jusqu'à peu encore méconnus. Un proces-

sus sélectif implique l'existence d'interactions indirectes ou directes entre les ARNv [49]. Aucune protéine virale ou cellulaire capable de reconnaître spécifiquement les signaux d'emballage présents sur les ARNv n'ayant été mise en évidence à ce jour [97], c'est l'hypothèse d'interactions directes entre les segments d'ARNv, par des appariements de type Watson-Crick, qui a été privilégiée [84].

Des études *in vitro* récentes ont montré, par retard de migration électrophorétique sur gels natifs, que les huit ARNv d'un virus humain H3N2 (A/Moscow/10/99) contemporain et circulant sont impliqués dans un réseau d'interactions unique où chaque ARNv interagit avec au moins un partenaire [98] (*figure 5A*). De manière surprenante, le réseau d'interactions d'un virus aviaire H5N2 (A/Finch/England/2051/91), déterminé par la même approche, est radicalement différent [98, 99] (*figure 5B*).

La délétion des nt identifiés comme des signaux d'emballage sur les ARNv 4, 7 et 8 chez les souches WSN et/ou PR8 [49, 86] affecte fortement les interactions intermoléculaires formées *in vitro* entre les ARNv 6 et 7, 4 et 8 et 5 et 7 de H3N2, suggérant qu'au moins certaines des interactions détectées *in vitro* sont importantes pour le processus d'assemblage sélectif [98]. De façon inattendue, dans le cas du virus aviaire H5N2, la majorité des interactions ARNv/ARNv impliquent des séquences localisées dans des régions plus centrales des ARNv et n'ayant pas été référencées précédemment comme des signaux d'emballage [99] (*figure 5B*). Ces données expérimentales sont en accord avec un modèle théorique basé sur les forces répulsives entre les RNPv contrecarrées par l'appariement de bases entre les ARNv [100]. Ce modèle prédit que la seule caractéristique commune entre des virus ayant évolué de façon indépendante est la répulsion entre les segments.

Dans le cas du virus aviaire H5N2, la majorité des interactions détectées *in vitro* ont été délimitées au nt près, permettant l'analyse de leurs rôles fonctionnels *in cellula/in viro* grâce à l'introduction de mutations silencieuses *trans-complémentaires* sur les ARNv partenaires. Ces mutations sont conçues pour restaurer l'interaction entre les deux ARNv partenaires [99, 101]. Dans le cas de l'interaction entre les ARNv 2 et 8, les virus possédant les deux mutations *trans-complémentaires* (*figure 6A*) se répliquent aussi efficacement que les virus sauvages, tandis que les virus simples mutants présentent un défaut de réplication [101]. Ces derniers présentent également un taux nettement plus élevé de particules virales sans RNPv visibles par rapport aux virus sauvage et double-mutant [101]. Enfin, des expériences de génétique inverse compétitive à 9 plasmides ont montré que lorsque deux ARNv sauvage et mutant étaient en compétition, en présence de leur partenaire sauvage ou mutant (*figure 6B*), les ARNv capables d'interagir étaient incorporés de manière préférentielle au



**Figure 5.** Réseaux d'interactions intermoléculaires des virus humain A/Moscow/10/99 (H3N2) et aviaire A/Finch/England/2051/91 (H5N2). **A)** Réseau d'interactions intermoléculaires entre les ARNv de la souche humaine H3N2. Les rectangles correspondent à une interaction déterminée par retard de migration électrophorétique. L'intensité de la couleur bleue est proportionnelle au pourcentage en masse des ARNv impliqués dans les complexes intermoléculaires [98, 102]. **B)** Réseau d'interactions intermoléculaires entre les ARNv de la souche aviaire H5N2. Les rectangles correspondent à une interaction déterminée par retard de migration électrophorétique. L'intensité de la couleur rouge est proportionnelle au pourcentage en masse des ARNv impliqués dans les complexes intermoléculaires [99, 101]. La numérotation blanche correspond aux interactions ARN/ARN identifiées au nt près et validées *in viro* et/ou *in vitro* par des mutants *trans*-complémentaires. La numérotation bleue italique correspond aux interactions ARN/ARN délimitées par une stratégie antisens. La numérotation noire correspond aux interactions ARN/ARN délimitées par de grandes délétions. Les « ? » correspondent à des délétions dont les effets sont très peu marqués.

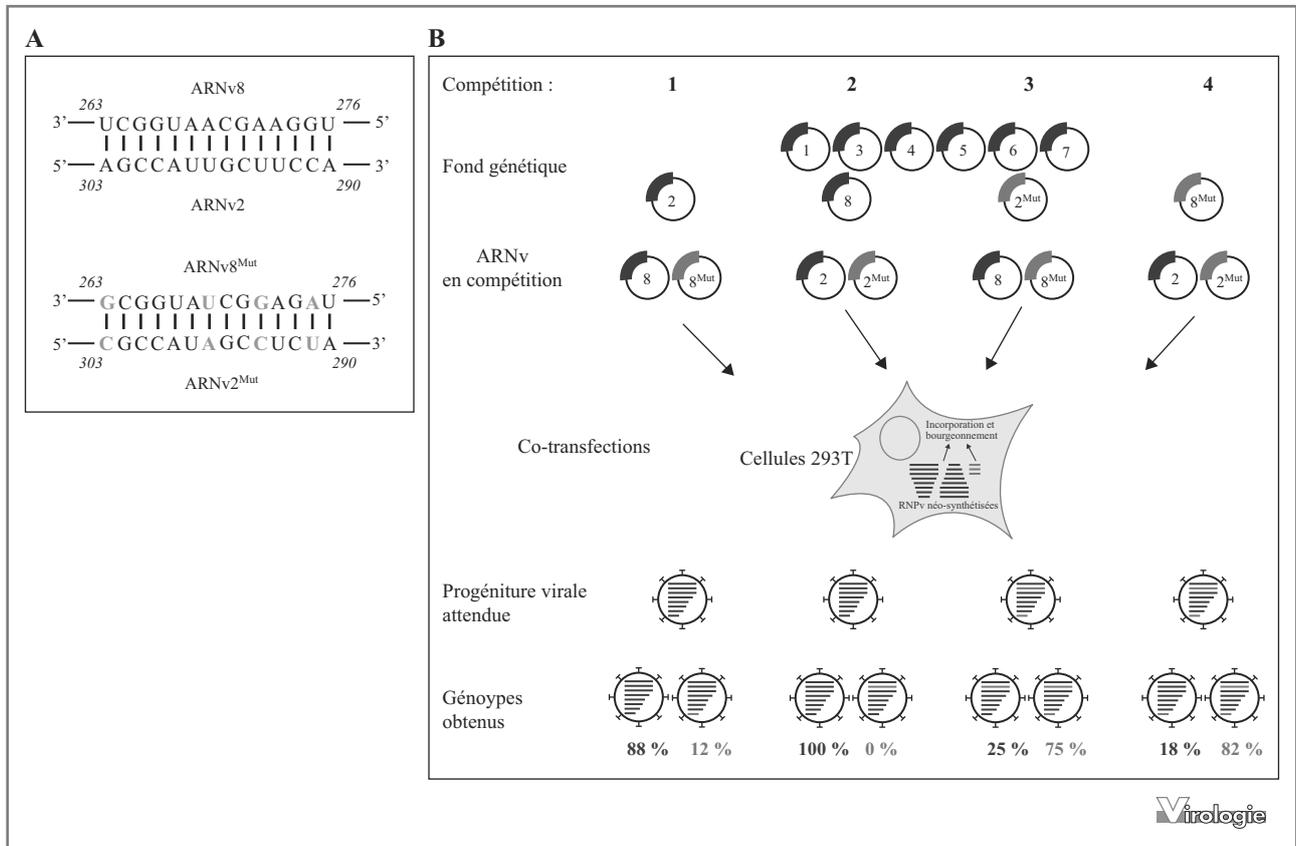
sein des virions [101] (*figure 6C*). L'ensemble de ces données prouve que l'interaction entre les ARNv 2 et 8, au niveau des nt 290-303 et 263-276 respectivement, est cruciale pour la réplication virale et le co-empaquetage sélectif des deux ARNv dans les virions [101]. Ces travaux constituent la première démonstration de l'existence et du rôle fonctionnel d'une interaction directe entre ARNv dans l'assemblage de particules infectieuses d'IVA. Combinés avec la mise en évidence *in vitro* de réseaux d'interactions

entre les ARNv [98, 99, 102], ces résultats suggèrent fortement que les huit ARNv sont incorporés dans les virions sous forme d'un complexe supramoléculaire maintenu par des interactions directes entre ARNv. L'identification de régions d'interaction/empaquetage dans les zones centrales des ARNv du virus aviaire H5N2 mais pas dans le cas des souches WSN et PR8, pourrait être inhérente aux méthodes employées dans les derniers cas, à savoir le remplacement des zones centrales par un gène rapporteur pour la production de virus par génétique inverse, ou à la focalisation sur les régions fortement conservées, localisées aux extrémités des segments [93]. Les différences entre réseaux et séquences impliquées dans la formation de ceux-ci pourraient cependant aussi être liées à des spécificités d'hôte et/ou de sous-type, qui restent encore à déterminer. Incidemment, les séquences impliquées dans l'interaction entre les ARNv 2 et 8 de la souche A/Finch/England/2051/91 ne sont que partiellement conservées parmi les virus H5N2 aviaires et non conservées chez les IVA circulants dans la population humaine [101].

### Impact de la structure des RNPv sur les interactions ARN/ARN

Dans les cellules et dans les virions, chaque ARNv est associé à de multiples copies de la protéine NP et au complexe polymérase, lié aux extrémités des ARNv [36, 46-48]. Sauf exception, les expériences de retardement sur gel réalisées sur les ARNv de deux souches d'IVA [98, 99, 101, 102] l'ont été en absence de protéines, même si l'interaction entre les ARNv 6 et 7 du virus humain H3N2 a été confirmée en présence de NP [98]. Il est par conséquent légitime de se demander si toutes les interactions déterminées *in vitro* existent réellement *in viro* et, inversement, si toutes les interactions qui s'établissent *in viro* peuvent être identifiées par les systèmes expérimentaux *in vitro*.

Observés au microscopie électronique [103, 104], les RNPv purifiées apparaissent comme des bâtonnets flexibles qui adoptent la structure d'une double-hélice droite. La protéine NP forme un squelette protéique hélicoïdal [46, 48, 105-107], pouvant se mettre en place en présence d'ARN non viral, qui persiste lorsque l'ARN est dégradé [106] et sur lequel s'enroule l'ARNv. La reconstruction, par cryo-microscopie électronique, d'une RNPv de taille minimale, compétente pour la réplication [47], a montré la présence d'un anneau nucléoprotéique comportant neuf monomères de protéines NP, dont deux associés au complexe polymérase [47]. En 2012, deux équipes ont reconstruit, toujours par cryo-microscopie électronique, les structures de RNPv complètes [46, 48], à 20 Å de résolution. Dans les deux modèles, les RNPv adoptent une structure en double hélice

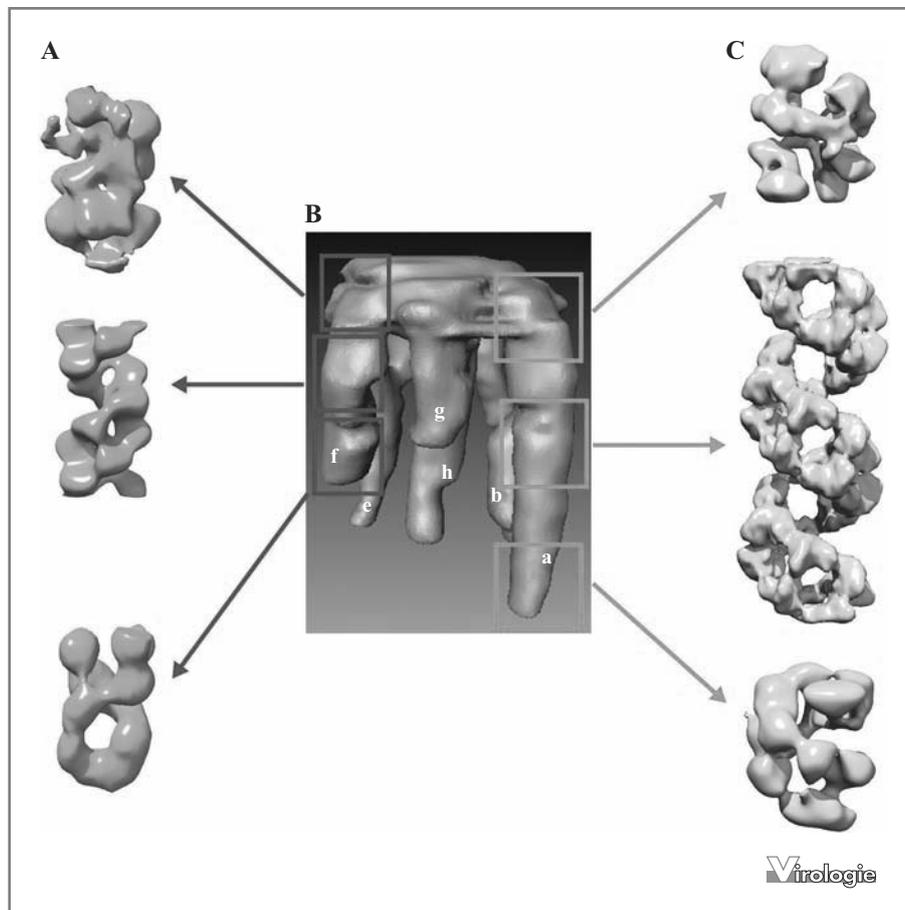


**Figure 6.** Rôle fonctionnel d'une interaction ARNV/ARNv. **A**) Complémentarité entre les ARNV 2 et 8 sauvages de la souche A/Finch/England/2051/91 (H5N2) et entre les mutants *trans*-complémentaires. **B**) Compétition entre les ARNV 2 et 8 sauvages et mutants lors de leur empaquetage dans les virions néo-synthétisés. Neufs plasmides, incluant ceux codant les segments sauvages 1 et 3-7 (fond génétique) et une combinaison de segments 2 et 8 sauvages (WT) et mutants (ARNv en compétition) ont été utilisés pour transfecter des cellules HEK 293T.

avec deux brins antiparallèles maintenus par des interactions entre les protéines NP (*figure 7A, C*). Le complexe polymérase, qui pourrait adopter deux conformations [46], se trouve à l'extrémité ouverte des RNPv, l'autre extrémité étant constituée d'une boucle contenant trois à huit molécules de NP [48]. Il existe cependant des différences significatives entre les deux modèles, parmi lesquelles le sens de l'hélice, dû à l'orientation relative des protéines NP des brins opposés, qui s'associent au niveau de leurs domaines protéiques de tête [48] ou du corps [46], pour former une hélice droite ou gauche, respectivement. Ces différences pourraient provenir de la nature différente des échantillons, issus dans un cas d'une expression des RNPv après transfection transitoire [48] et dans l'autre de la purification de RNPv natives à partir de virions [46]. Les deux groupes s'accordent cependant sur le fait que l'ARNv se trouverait à la surface du squelette protéique. De fait, comme aucune densité correspondant à l'ARNv n'a été observée dans les cartes de densité, les

ARNv n'adopteraient donc pas de structure périodique [46, 48].

Les premières études de cartographie chimique sur les RNPv ont suggéré que la NP déroulerait tous les motifs structurés des ARNV [108, 109], alors que des expériences de cartographie enzymatique suggèrent au contraire la persistance de structures secondaires au sein des RNPv [110]. Le rapport NP/RNPv dans des extraits de RNPv purifiées est d'une protéine NP pour 24 nt d'ARNv [103, 111, 112]. Cependant, l'étude récente de la protéine NP du virus de l'anémie infectieuse du saumon (ISAV) a révélé que cette protéine ne lie directement que 12 nt d'ARN, en accord avec la taille du sillon de la protéine constituant le site de fixation putatif de l'ARNv [40], laissant ainsi à nu certaines régions de l'ARNv. En cohérence avec cette idée, des RNPv et des RNPc ont été purifiées à l'aide d'une stratégie de purification par affinité, basée sur l'introduction d'un aptamère artificiel structuré dans l'ARNv 6 [113]. La réussite d'une telle expérience suggère que des éléments de structure



**Figure 7.** Organisation tridimensionnelle des RNPv au sein des influenzavirus de type A (IVA) et structure d'une RNPv. Panneau central **(B)** : Reconstruction tridimensionnelle de l'intérieur de particules virales d'un virus humain de sous-type H3N2 [98, 102]. Les huit RNPv émergent d'une plate-forme commune au pôle apical du virus. Les longueurs différentes des RNPv ont permis leur identification [98, 102]. Panneaux **(A)** [48] et **(C)** [46] : densités électroniques issues d'expériences de cryo-EM, des extrémités polymérase (haut), de la partie hélicoïdale centrale (centre) et de la boucle terminale (bas) de RNPv. Ces densités électroniques proviennent des entrées EMD-2209, EMD-2211, et EMD-2212 de la « EM Data Bank » et ont été représentées à l'aide du logiciel « UCSF Chimera Software ».

secondaire peuvent saillir des ARNv des IVA. L'ensemble de ces données rend vraisemblable l'existence de régions localement dépourvues de protéines NP et potentiellement structurées, disponibles pour la formation d'un complexe supramoléculaire maintenu par des interactions spécifiques entre ARNv.

### Organisation tridimensionnelle non aléatoire des RNPv au sein des virions

Selon l'observation de particules virales sphériques et filamenteuses par tomographie et microscopie électronique à transmission, les RNPv adoptent une organisation conservée dite en « 7+1 » avec une RNPv centrale entourée des sept autres RNPv [114-116] (*figure 1B*).

Au sein des particules virales bourgeonnantes, les huit RNPv s'ordonnent parallèlement les unes aux autres et sont suspendues à partir de l'extrémité distale des virions [115, 117, 118]. Des organisations différentes ont été détectées dans le cas de grandes particules sphériques, suggérant que l'agencement en « 7+1 » pourrait être perturbé après bourgeonnement et relargage des virions [114]. La reconstruction 3D, par tomographie électronique, de l'intérieur de virions humains H3N2 montre que des contacts entre RNPv pourraient s'établir au niveau d'une plateforme apicale suffisamment épaisse pour héberger les signaux d'emballage se situant aux extrémités des ARNv [98, 102] (*figure 7B*). Pour la souche WSN en revanche, des contacts ont été observés tout le long des RNPv [118]. Ces résultats contradictoires pourraient en partie être expliqués par des divergences techniques.

En effet, Noda *et al.* [118] ont analysé des coupes virtuelles de virions plus fines que celles analysées dans le cas du virus A/Moscow/10/99 (H3N2) [98, 102]. Cependant, Fournier *et al.* [98, 102] et Gavazzi *et al.* [99], utilisant le même microscope et les mêmes paramètres expérimentaux, ont observé de multiples contacts centraux entre les RNPv des virions A/Finch/England/2051/91 (H5N2), contrairement aux virions A/Moscow/10/99 (H3N2).

Si les RNPv sont assemblées en un complexe supramoléculaire maintenu par des appariements de bases, ceux-ci devraient fortement contraindre l'organisation interne des RNPv au sein des particules virales. Les données d'expériences de retard de migration sur gel, *in vitro*, alliées à l'identification de certaines RNPv en fonction de leur taille, par tomographie, a permis de proposer la coexistence d'un nombre limité de dispositions de RNPv au sein des virus de la souche humaine H3N2 [98, 102]. Des conclusions identiques ont été obtenues lors de l'étude structurale de virions de la souche WSN [118], pour lesquels les agencements de RNPv sont toutefois différents de ceux observés dans le cas du virus H3N2. De plus, une étude récente de Sugita *et al.* [119] suggère que même si la majorité des RNPv est incorporée avec le complexe polymérase du côté de l'extrémité bourgeonnante des virions, certaines d'entre elles sont incorporées avec le complexe polymérase du côté basal du virus. Ces nouvelles données, si elles se confirment, complexifient le processus de sélection des RNPv, en augmentant le nombre de possibilités d'interactions entre les ARNv. L'ensemble de ces données permet de proposer l'existence d'un agencement non aléatoire, différent selon les souches, des RNPv dans les particules virales, au sein d'un complexe supramoléculaire maintenu par des appariements directs de type Watson-Crick entre les ARNv. Ces appariements constitueraient les signaux d'emballage du génome et différencieraient, dans leur nature et leur localisation, selon les souches virales.

## Un mécanisme d'emballage spécifique en deux étapes ?

Dans le modèle d'emballage sélectif des ARNv des IAV, la machinerie impliquée doit :

- discriminer les ARNv de polarité négative des ARNv de polarité positive (ARNm et ARNc, ces derniers étant également associés à la NP) et des ARN cellulaires ;
- discriminer les huit ARNv entre eux pour assurer l'emballage d'un jeu complet d'ARNv génomiques.

Plusieurs résultats suggèrent que des signaux différents entrent en jeu pour permettre ces deux niveaux de discrimination. Des signaux communs à tous les segments d'ARNv des IVA permettraient d'assurer la discrimination entre les ARNv et les ARN non viraux. Les séquences conservées

U12 et U13 pourraient jouer ce rôle [77, 78]. Récemment, le groupe de Y. Kawaoka a démontré que les NCR de l'ARNv 5 permettent son emballage dans des pseudo-particules virales alors que les régions d'emballage (incluant les NCR et une partie des régions codantes en 3' et 5' [83]) servent de signal pour assurer l'emballage d'un jeu complet de huit RNPv au sein des virions [120]. Bien que le caractère général de ces résultats reste encore à démontrer, ils suggèrent que les séquences U12 et U13 discriminent les ARN cellulaires et les ARN viraux de polarité positive des ARNv de polarité négative, alors que les interactions ARN/ARN spécifiques permettent la discrimination entre les huit ARNv.

D'un point de vue mécanistique et spatio-temporel, la sélection des ARNv pour l'emballage au sein des virions pourrait avoir lieu lors de l'export nucléaire, lors du transport dans le cytoplasme et/ou lors du bourgeonnement des particules virales. Dans le noyau, M1 interagit à la fois avec les RNPv néo-synthétisées et avec la protéine NS2 qui recrute la machinerie d'export en interagissant avec Crm1, permettant l'export des complexes ainsi formés vers le cytoplasme [121]. Cette voie d'export n'est utilisée ni par les ARN cellulaires, ni par les ARNm viraux. Les segments d'ARNc, de polarité positive, s'assemblent également sous forme de RNPc [113] mais ne sont toutefois pas exportés : la discrimination entre les RNPv et RNPc repose sur une conformation différente des appariements U12-U13 (*figure 2B, C*) due au fait qu'un nt de U13 reste non apparié dans les RNPc. Le complexe polymérase, qui par ailleurs adopte des conformations différentes lorsqu'il est fixé aux RNPv ou aux RNPc, reconnaît vraisemblablement ces deux conformations de manière différentielle [81]. L'export nucléaire pourrait par conséquent constituer le premier niveau de sélection, de manière indirecte, par le biais de reconnaissances différentielles des RNPv et RNPc par le complexe polymérase.

Le transport des RNPv de la région péri-nucléaire à la membrane apicale implique le système de recyclage des endosomes dépendant de la protéine Rab-11, ainsi que la présence d'un réseau microtubulaire intact [122-126]. Chou *et al.* [127] ont récemment démontré que lors de la phase tardive de l'infection, différents ARNv des IVA sont transportés ensemble, au sein du cytoplasme, associés avec les vésicules comprenant la protéine Rab-11 : celle-ci sert probablement de plateforme pour maintenir les ARNv à proximité les uns des autres. Des interactions spécifiques ARN/ARN pourraient s'établir pendant le transport pour former un réseau entre certains ARNv qui constitueraient le cœur du complexe supramoléculaire d'assemblage viral. Un tel noyau, regroupant un nombre limité d'ARNv, différents selon les souches virales, pourrait constituer un « point de contrôle » pour un assemblage correct des huit RNPv, avant qu'elles n'atteignent la membrane plasmique apicale

où a lieu le bourgeonnement. Cela expliquerait pourquoi, comme détaillé précédemment, tous les ARNV ne jouent pas le même rôle au moment de l'empaquetage du génome. De manière très importante, cette hypothèse a été confirmée très récemment par l'observation d'un export groupé de plusieurs RNPv [128]. Ces sous-ensembles de RNPv, qui n'impliquent pas les huit segments génomiques, fusionnent ensuite durant leur cheminement vers la membrane plasmique, lors d'évènements de co-localisation dynamiques. Ces évènements impliquent, comme démontré précédemment, les endosomes de recyclage comprenant la protéine Rab-11 [128].

## Conclusion

L'existence d'un mécanisme d'empaquetage sélectif des ARNV d'IVA au sein des particules virales est maintenant établie. Les données récentes de microscopie électronique, de tomographie, de virologie, de génétique et de biochimie ont permis de mieux comprendre ce processus. Ces travaux suggèrent que les huit ARNV sont incorporés sous forme d'un complexe supramoléculaire maintenu par des appariements de type Watson-Crick entre les ARNV. Ce complexe serait en partie ou totalement formé durant le transport des RNPv jusqu'au site de bourgeonnement. D'autres études sont néanmoins nécessaires pour confirmer ces résultats et donc valider ce modèle. De manière surprenante, ces études révèlent que les interactions ARN/ARN, ainsi que l'organisation interne des particules virales, diffèrent selon les souches d'IVA. Cette variabilité a probablement un impact majeur sur le réassortiment génétique, puisque les interactions ARN/ARN formant les réseaux d'interactions des deux virus parentaux vont influencer le génotype des virus réassortants [53, 129]. À l'heure actuelle, notre connaissance des règles gouvernant le co-empaquetage d'ARNv de deux souches parentales dans une même particule virale est limitée. La compréhension de ces règles est un véritable défi pour le futur, car elles constituent le point de départ pour comprendre les règles gouvernant le réassortiment génétique.

**Remerciements.** Ce travail est soutenu par le Centre national de la recherche scientifique et VirNext. Les auteurs remercient également le Dr F. Jossinet pour son aide concernant la *figure 7*.

**Liens d'intérêts :** les auteurs déclarent ne pas avoir de lien d'intérêt en rapport avec cet article.

## Références

1. Schrauwen EJ, Fouchier RA. Host adaptation and transmission of influenza A viruses in mammals. *Emerg Microbes Infect* 2014 ; 3(e9).

2. Simonsen L, Spreeuwenberg P, Lustig R, *et al*. Global mortality estimates for the 2009 influenza pandemic from the GLaMOR Project: a modeling study. *PLoS Med* 2013 ; 10 : e1001558.
3. Swayne DE. Impact of vaccines and vaccination on global control of avian influenza. *Avian Dis* 2012 ; 56 (4 Suppl) : 818-28.
4. Chen LM, Blixt O, Stevens J, *et al*. *In vitro* evolution of H5N1 avian influenza virus toward human-type receptor specificity. *Virology* 2012 ; 422 : 105-13.
5. Herfst S, Schrauwen EJ, Linster M, *et al*. Airborne transmission of influenza A/H5N1 virus between ferrets. *Science* 2012 ; 336 : 1534-41.
6. Imai M, Watanabe T, Hatta M, *et al*. Experimental adaptation of an influenza H5 HA confers respiratory droplet transmission to a reassortant H5 HA/H1N1 virus in ferrets. *Nature* 2012 ; 486 : 420-8.
7. Zhang H, Hale BG, Xu K, Sun B. Viral and host factors required for avian H5N1 influenza A virus replication in mammalian cells. *Viruses* 2013 ; 5 : 1431-46.
8. Linster M, van Boheemen S, de Graaf M, *et al*. Identification, characterization, and natural selection of mutations driving airborne transmission of A/H5N1 virus. *Cell* 2014 ; 157 : 329-39.
9. Ke Y, Wang Y, Liu S, *et al*. High severity and fatality of human infections with avian influenza A(H7N9) infection in China. *Clin Infect Dis* 2013 ; 57 : 1506-7.
10. Presti RM, Zhao G, Beatty WL, *et al*. Johnston Atoll, and Lake Chad viruses are novel members of the family Orthomyxoviridae. *J Virol* 2009 ; 83 : 11599-606.
11. Moules V, Terrier O, Yver M, *et al*. Importance of viral genomic composition in modulating glycoprotein content on the surface of influenza virus particles. *Virology* 2011 ; 414 : 51-62.
12. Obayashi E, Yoshida H, Kawai F, *et al*. The structural basis for an essential subunit interaction in influenza virus RNA polymerase. *Nature* 2008 ; 454 : 1127-31.
13. Guilligay D, Tarendeau F, Resa-Infante P, *et al*. The structural basis for cap binding by influenza virus polymerase subunit PB2. *Nat Struct Mol Biol* 2008 ; 15 : 500-6.
14. Dias A, Bouvier D, Crepin T, *et al*. The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit. *Nature* 2009 ; 458 : 914-8.
15. Yuan P, Bartlam M, Lou Z, *et al*. Crystal structure of an avian influenza polymerase PA(N) reveals an endonuclease active site. *Nature* 2009 ; 458 : 909-13.
16. Hale BG, Randall RE, Ortin J, Jackson D. The multifunctional NS1 protein of influenza A viruses. *J Gen Virol* 2008 ; 89(Pt 10) : 2359-76.
17. Paterson D, Fodor E. Emerging roles for the influenza A virus nuclear export protein (NEP). *PLoS Pathog* 2012 ; 8 : e1003019.
18. Tong S, Zhu X, Li Y, *et al*. New world bats harbor diverse influenza A viruses. *PLoS Pathog* 2013 ; 9 : e1003657.
19. Palese PA, Shaw ML. Orthomyxoviridae: the viruses and their replication. In : Knipe DM, Howley PM, eds. *Fields virology*. Philadelphia, PA : Lippincott Williams & Wilkins Publishers, 2007, p. 1647-89.
20. Tong S, Li Y, Rivailler P, *et al*. A distinct lineage of influenza A virus from bats. *Proc Natl Acad Sci U S A* 2012 ; 109 : 4269-74.
21. Shi M, Jagger BW, Wise HM, Digard P, Holmes EC, Taubenberger JK. Evolutionary conservation of the PA-X open reading frame in segment 3 of influenza A virus. *J Virol* 2012 ; 86 : 12411-3.
22. Chen W, Calvo PA, Malide D, *et al*. A novel influenza A virus mitochondrial protein that induces cell death. *Nat Med* 2001 ; 7 : 1306-12.
23. Jagger BW, Wise HM, Kash JC, *et al*. An overlapping protein-coding region in influenza A virus segment 3 modulates the host response. *Science* 2012 ; 337 : 199-204.
24. Muramoto Y, Noda T, Kawakami E, Akkina R, Kawaoka Y. Identification of novel influenza A virus proteins translated from PA mRNA. *J Virol* 2013 ; 87 : 2455-62.

25. Selman M, Dankar SK, Forbes NE, Jia JJ, Brown EG. Adaptive mutation in influenza A virus non-structural gene is linked to host switching and induces a novel protein by alternative splicing. *Emerg Microbes Infect* 2012; 1(e42).
26. Wise HM, Barbezange C, Jagger BW, *et al.* Overlapping signals for translational regulation and packaging of influenza A virus segment 2. *Nucleic Acids Res* 2011; 39: 7775-90.
27. Wise HM, Foeglein A, Sun J, *et al.* A complicated message: Identification of a novel PB1-related protein translated from influenza A virus segment 2 mRNA. *J Virol* 2009; 83: 8021-31.
28. Wise HM, Hutchinson EC, Jagger BW, *et al.* Identification of a novel splice variant form of the influenza A virus M2 ion channel with an antigenically distinct ectodomain. *PLoS Pathog* 2012; 8: e1002998.
29. Krumbholz A, Philipps A, Oehring H, *et al.* Current knowledge on PB1-F2 of influenza A viruses. *Med Microbiol Immunol* 2011; 200: 69-75.
30. Chakrabarti AK, Pasricha G. An insight into the PB1F2 protein and its multifunctional role in enhancing the pathogenicity of the influenza A viruses. *Virology* 2013; 440: 97-104.
31. Mazur I, Anhlán D, Mitzner D, Wixler L, Schubert U, Ludwig S. The proapoptotic influenza A virus protein PB1-F2 regulates viral polymerase activity by interaction with the PB1 protein. *Cell Microbiol* 2008; 10: 1140-52.
32. McAuley JL, Hornung F, Boyd KL, *et al.* Expression of the 1918 influenza A virus PB1-F2 enhances the pathogenesis of viral and secondary bacterial pneumonia. *Cell Host Microbe* 2007; 2: 240-9.
33. Dudek SE, Wixler L, Nordhoff C, *et al.* The influenza virus PB1-F2 protein has interferon antagonistic activity. *Biol Chem* 2011; 392: 1135-44.
34. Reis AL, McCauley JW. The influenza virus protein PB1-F2 interacts with IKKbeta and modulates NF-kappaB signalling. *PLoS ONE* 2013; 8: e63852.
35. Varga ZT, Ramos I, Hai R, *et al.* The influenza virus protein PB1-F2 inhibits the induction of type I interferon at the level of the MAVS adaptor protein. *PLoS Pathog* 2011; 7: e1002067.
36. Elton D, Digard P, Tiley L, Ortin J. Structure and function of the influenza virus RNP. In: Kawaoka Y, ed. *Current Topics in Influenza Virology*. Norfolk: Horizon Scientific Press, 2006, p. 1-36.
37. Gultyaev AP, Fouchier RA, Olsthoorn RC. Influenza virus RNA structure: unique and common features. *Int Rev Immunol* 2010; 29: 533-56.
38. Resa-Infante P, Jorba N, Coloma R, Ortin J. The influenza virus RNA synthesis machine: advances in its structure and function. *RNA Biol* 2011; 8: 207-15.
39. Ruijgrok RW, Crepin T, Hart DJ, Cusack S. Towards an atomic resolution understanding of the influenza virus replication machinery. *Curr Opin Struct Biol* 2010; 20: 104-13.
40. Zheng W, Olson J, Vakharia V, Tao YJ. The crystal structure and RNA-binding of an orthomyxovirus nucleoprotein. *PLoS Pathog* 2013; 9: e1003624.
41. Bae SH, Cheong HK, Lee JH, Cheong C, Kainosho M, Choi BS. Structural features of an influenza virus promoter and their implications for viral RNA synthesis. *Proc Natl Acad Sci U S A* 2001; 98: 10602-7.
42. Cheong HK, Cheong C, Lee YS, Seong BL, Choi BS. Structure of influenza virus panhandle RNA studied by NMR spectroscopy and molecular modeling. *Nucleic Acids Res* 1999; 27: 1392-7.
43. Tomescu AI, Robb NC, Hengrung N, Fodor E, Kapanidis AN. Single-molecule FRET reveals a corkscrew RNA structure for the polymerase-bound influenza virus promoter. *Proc Natl Acad Sci U S A* 2014; 111: E3335-42.
44. Crow M, Deng T, Addley M, Brownlee GG. Mutational analysis of the influenza virus cRNA promoter and identification of nucleotides critical for replication. *J Virol* 2004; 78: 6263-70.
45. Li ML, Ramirez BC, Krug RM. 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* 1998; 17: 5844-52.
46. Arranz R, Coloma R, Chichon FJ, *et al.* The structure of native influenza virion ribonucleoproteins. *Science* 2012; 338: 1634-7.
47. Coloma R, Valpuesta JM, Arranz R, Carrascosa JL, Ortin J, Martin-Benito J. The structure of a biologically active influenza virus ribonucleoprotein complex. *PLoS Pathog* 2009; 5: e1000491.
48. Moeller A, Kirchdoerfer RN, Potter CS, Carragher B, Wilson IA. Organization of the influenza virus replication machinery. *Science* 2012; 338: 1631-4.
49. Hutchinson EC, von Kirchbach JC, Gog JR, Digard P. Genome packaging in influenza A virus. *J Gen Virol* 2010; 91(Pt 2): 313-28.
50. Enami M, Sharma G, Benham C, Palese P. An influenza virus containing nine different RNA segments. *Virology* 1991; 185: 291-8.
51. Nakajima K, Sugiura A. Three-factor cross of influenza virus. *Virology* 1977; 81: 486-9.
52. Donald HB, Isaacs A. Counts of influenza virus particles. *J Gen Microbiol* 1954; 10: 457-64.
53. Gerber M, Isel C, Moules V, Marquet R. Selective packaging of the influenza A genome and consequences for genetic reassortment. *Trends Microbiol* 2014.
54. Brooke CB, Ince WL, Wrammert J, *et al.* Most influenza A virions fail to express at least one essential viral protein. *J Virol* 2013; 87: 3155-62.
55. Hatada E, Hasegawa M, Mukaigawa J, Shimizu K, Fukuda R. Control of influenza virus gene expression: quantitative analysis of each viral RNA species in infected cells. *J Biochem* 1989; 105: 537-46.
56. McGeoch D, Fellner P, Newton C. Influenza virus genome consists of eight distinct RNA species. *Proc Natl Acad Sci U S A* 1976; 73: 3045-9.
57. Bergmann M, Muster T. The relative amount of an influenza A virus segment present in the viral particle is not affected by a reduction in replication of that segment. *J Gen Virol* 1995; 76(Pt 12): 3211-5.
58. Smith GL, Hay AJ. Replication of the influenza virus genome. *Virology* 1982; 118: 96-108.
59. Laver WG, Downie JC. Influenza virus recombination. I. Matrix protein markers and segregation during mixed infections. *Virology* 1976; 70: 105-17.
60. Lubeck MD, Palese P, Schulman JL. Nonrandom association of parental genes in influenza A virus recombinants. *Virology* 1979; 95: 269-74.
61. Chou YY, Vafabakhsh R, Doganay S, Gao Q, Ha T, Palese P. One influenza virus particle packages eight unique viral RNAs as shown by FISH analysis. *Proc Natl Acad Sci U S A* 2012; 109: 9101-6.
62. Inagaki A, Goto H, Kakugawa S, Ozawa M, Kawaoka Y. Competitive incorporation of homologous gene segments of influenza A virus into virions. *J Virol* 2012; 86: 10200-2.
63. Chambers TM, Webster RG. Defective interfering virus associated with A/Chicken/Pennsylvania/83 influenza virus. *J Virol* 1987; 61: 1517-23.
64. Davis AR, Nayak DP AR. Sequence relationships among defective interfering influenza viral RNAs. *Proc Natl Acad Sci U S A* 1979; 76: 3092-6.
65. Duhaut SD, Dimmock NJ. Heterologous protection of mice from a lethal human H1N1 influenza A virus infection by H3N8 equine defective interfering virus: comparison of defective RNA sequences isolated from the DI inoculum and mouse lung. *Virology* 1998; 248: 241-53.
66. Duhaut SD, Dimmock NJ. Defective segment 1 RNAs that interfere with production of infectious influenza A virus require at least 150 nucleotides of 5' sequence: evidence from a plasmid-driven system. *J Gen Virol* 2002; 83(Pt 2): 403-11.
67. Duhaut SD, McCauley JW. Defective RNAs inhibit the assembly of influenza virus genome segments in a segment-specific manner. *Virology* 1996; 216: 326-37.

68. Jennings PA, Finch JT, Winter G, Robertson JS. Does the higher order structure of the influenza virus ribonucleoprotein guide sequence rearrangements in influenza viral RNA? *Cell* 1983; 34: 619-27.
69. Nakajima K, Ueda M, Sugiura A. Origin of small RNA in von Magnus particles of influenza virus. *J Virol* 1979; 29: 1142-8.
70. Nayak DP, Sivasubramanian N, Davis AR, Cortini R, Sung J. Complete sequence analyses show that two defective interfering influenza viral RNAs contain a single internal deletion of a polymerase gene. *Proc Natl Acad Sci U S A* 1982; 79: 2216-20.
71. Odagiri T, Tominaga K, Tobita K, Ohta S. An amino acid change in the non-structural NS2 protein of an influenza A virus mutant is responsible for the generation of defective interfering (DI) particles by amplifying DI RNAs and suppressing complementary RNA synthesis. *J Gen Virol* 1994; 75(Pt 1): 43-53.
72. Sivasubramanian N, Nayak DP. Defective interfering influenza RNAs of polymerase 3 gene contain single as well as multiple internal deletions. *Virology* 1983; 124: 232-7.
73. Noble S, Dimmock NJ. Characterization of putative defective interfering (DI) A/WSN RNAs isolated from the lungs of mice protected from an otherwise lethal respiratory infection with influenza virus A/WSN (H1N1): a subset of the inoculum DI RNAs. *Virology* 1995; 210: 9-19.
74. Winter G, Fields S, Ratti G. The structure of two subgenomic RNAs from human influenza virus A/PR/8/34. *Nucleic Acids Res* 1981; 9: 6907-15.
75. Yang P, Bansal A, Liu C, Air GM. Hemagglutinin specificity and neuraminidase coding capacity of neuraminidase-deficient influenza viruses. *Virology* 1997; 229: 155-65.
76. Hoffmann E, Neumann G, Kawaoka Y, Hobom G, Webster RG. A DNA transfection system for generation of influenza A virus from eight plasmids. *Proc Natl Acad Sci U S A* 2000; 97: 6108-13.
77. Luytjes W, Krystal M, Enami M, Parvin JD, Palese P. Amplification, expression, and packaging of foreign gene by influenza virus. *Cell* 1989; 59: 1107-13.
78. Muramoto Y, Takada A, Fujii K, et al. Hierarchy among viral RNA (vRNA) segments in their role in vRNA incorporation into influenza A virions. *J Virol* 2006; 80: 2318-25.
79. Odagiri T, Tashiro M. Segment-specific noncoding sequences of the influenza virus genome RNA are involved in the specific competition between defective interfering RNA and its progenitor RNA segment at the virion assembly step. *J Virol* 1997; 71: 2138-45.
80. Neumann G, Hobom G G. Mutational analysis of influenza virus promoter elements *in vivo*. *J Gen Virol* 1995; 76(Pt 7): 1709-17.
81. Tchatalbachev S, Flick R, Hobom G. The packaging signal of influenza viral RNA molecules. *RNA* 2001; 7: 979-89.
82. Watanabe T, Watanabe S, Noda T, Fujii Y, Kawaoka Y. Exploitation of nucleic acid packaging signals to generate a novel influenza virus-based vector stably expressing two foreign genes. *J Virol* 2003; 77: 10575-83.
83. Ozawa M, Fujii K, Muramoto Y, et al. Contributions of two nuclear localization signals of influenza A virus nucleoprotein to viral replication. *J Virol* 2007; 81: 30-41.
84. Fujii Y, Goto H, Watanabe T, Yoshida T, Kawaoka Y. Selective incorporation of influenza virus RNA segments into virions. *Proc Natl Acad Sci U S A* 2003; 100: 2002-7.
85. Fujii K, Fujii Y, Noda T, et al. Importance of both the coding and the segment-specific noncoding regions of the influenza A virus NS segment for its efficient incorporation into virions. *J Virol* 2005; 79: 3766-74.
86. Ozawa M, Maeda J, Iwatsuki-Horimoto K, et al. Nucleotide sequence requirements at the 5' end of the influenza A virus M RNA segment for efficient virus replication. *J Virol* 2009; 83: 3384-8.
87. Liang Y, Hong Y, Parslow TG. cis-Acting packaging signals in the influenza virus PB1, PB2, and PA genomic RNA segments. *J Virol* 2005; 79: 10348-55.
88. Liang Y, Huang T, Ly H, Parslow TG, Liang Y. Mutational analyses of packaging signals in influenza virus PA, PB1, and PB2 genomic RNA segments. *J Virol* 2008; 82: 229-36.
89. Dos Santos Afonso E, Escricou N, Leclercq I, van der Werf S, Naffakh N. The generation of recombinant influenza A viruses expressing a PB2 fusion protein requires the conservation of a packaging signal overlapping the coding and noncoding regions at the 5' end of the PB2 segment. *Virology* 2005; 341: 34-46.
90. Marsh GA, Hatami R, Palese P. Specific residues of the influenza A virus hemagglutinin viral RNA are important for efficient packaging into budding virions. *J Virol* 2007; 81: 9727-36.
91. Hutchinson EC, Curran MD, Read EK, Gog JR, Digard P. Mutational analysis of cis-acting RNA signals in segment 7 of influenza A virus. *J Virol* 2008; 82: 11869-79.
92. Fujii K, Ozawa M, Iwatsuki-Horimoto K, Horimoto T, Kawaoka Y. Incorporation of influenza A virus genome segments does not absolutely require wild-type sequences. *J Gen Virol* 2009; 90(Pt 7): 1734-40.
93. Gog JR, Afonso Edos S, Dalton RM, et al. Codon conservation in the influenza A virus genome defines RNA packaging signals. *Nucleic Acids Res* 2007; 35: 1897-907.
94. Marsh GA, Rabadan R, Levine AJ, Palese P. Highly conserved regions of influenza A virus polymerase gene segments are critical for efficient viral RNA packaging. *J Virol* 2008; 82: 2295-304.
95. Hutchinson EC, Wise HM, Kudryavtseva K, Curran MD, Digard P. Characterisation of influenza A viruses with mutations in segment 5 packaging signals. *Vaccine* 2009; 27: 6270-5.
96. Gao Q, Chou YY, Doganay S, Vafabakhsh R, Ha T, Palese P. The influenza A virus PB2, PA, NP, and M segments play a pivotal role during genome packaging. *J Virol* 2012; 86: 7043-51.
97. Moisy D, Naffakh N. Interactions moléculaires entre les ribonucléoprotéines des virus influenza de type A et la cellule hôte. *Virologie* 2013; 17: 278-93.
98. Fournier E, Moules V, Essere B, et al. A supramolecular assembly formed by influenza A virus genomic RNA segments. *Nucleic Acids Res* 2012; 40: 2197-209.
99. Gavazzi C, Isel C, Fournier E, et al. An *in vitro* network of intermolecular interactions between viral RNA segments of an avian H5N2 influenza A virus: comparison with a human H3N2 virus. *Nucleic Acids Res* 2013; 41: 1241-54.
100. Venev SV, Zeldovich KB. Segment self-repulsion is the major driving force of influenza genome packaging. *Phys Rev Lett* 2013; 110: 098104.
101. Gavazzi C, Yver M, Isel C, et al. A functional sequence-specific interaction between influenza A virus genomic RNA segments. *Proc Natl Acad Sci U S A* 2013; 110: 16604-9.
102. Fournier E, Moules V, Essere B, et al. Interaction network linking the human H3N2 influenza A virus genomic RNA segments. *Vaccine* 2012; 30: 7359-67.
103. Compans RW, Content J, Duesberg PH. Structure of the ribonucleoprotein of influenza virus. *J Virol* 1972; 10: 795-800.
104. Pons M, Hirst GK. The single- and double-stranded RNA's and the proteins of incomplete influenza virus. *Virology* 1969; 38: 68-72.
105. Ng AK, Zhang H, Tan K, et al. Structure of the influenza virus A H5N1 nucleoprotein: implications for RNA binding, oligomerization, and vaccine design. *FASEB J* 2008; 22: 3638-47.
106. Ruigrok RW, Baudin F. Structure of influenza virus ribonucleoprotein particles. II. Purified RNA-free influenza virus ribonucleoprotein forms structures that are indistinguishable from the intact influenza virus ribonucleoprotein particles. *J Gen Virol* 1995; 76(Pt 4): 1009-14.
107. Ye Q, Krug RM, Tao YJ. The mechanism by which influenza A virus nucleoprotein forms oligomers and binds RNA. *Nature* 2006; 444: 1078-82.
108. Baudin F, Bach C, Cusack S, Ruigrok RW. Structure of influenza virus RNP. I. Influenza virus nucleoprotein melts secondary structure in

- 
- panhandle RNA and exposes the bases to the solvent. *EMBO J* 1994; 13: 3158-65.
- 109.** Klumpp K, Ruigrok RW, Baudin F. Roles of the influenza virus polymerase and nucleoprotein in forming a functional RNP structure. *EMBO J* 1997; 16: 1248-57.
- 110.** Yamanaka K, Ishihama A, Nagata K. Reconstitution of influenza virus RNA-nucleoprotein complexes structurally resembling native viral ribonucleoprotein cores. *J Biol Chem* 1990; 265: 11151-5.
- 111.** Martin-Benito J, Area E, Ortega J, *et al.* Three-dimensional reconstruction of a recombinant influenza virus ribonucleoprotein particle. *EMBO Rep* 2001; 2: 313-7.
- 112.** Ortega J, Martin-Benito J, Zurcher T, Valpuesta JM, Carrascosa JL, Ortin J. Ultrastructural and functional analyses of recombinant influenza virus ribonucleoproteins suggest dimerization of nucleoprotein during virus amplification. *J Virol* 2000; 74: 156-63.
- 113.** York A, Hengrung N, Vreede FT, Huiskonen JT, Fodor E. Isolation and characterization of the positive-sense replicative intermediate of a negative-strand RNA virus. *Proc Natl Acad Sci U S A* 2013; 110: E4238-45.
- 114.** Harris A, Cardone G, Winkler DC, *et al.* Influenza virus pleiomorphy characterized by cryoelectron tomography. *Proc Natl Acad Sci U S A* 2006; 103: 19123-7.
- 115.** Noda T, Sagara H, Yen A, *et al.* Architecture of ribonucleoprotein complexes in influenza A virus particles. *Nature* 2006; 439: 490-2.
- 116.** Oxford JSH, D. J. Orthomyxoviridae. In: Steven MVNAC, ed. *Perspective in medical virology: animal virus structure*. New York: Elsevier, 1987, p. 213-32.
- 117.** Calder LJ, Wasilewski S, Berriman JA, Rosenthal PB. Structural organization of a filamentous influenza A virus. *Proc Natl Acad Sci U S A* 2010; 107: 10685-90.
- 118.** Noda T, Sugita Y, Aoyama K, *et al.* Three-dimensional analysis of ribonucleoprotein complexes in influenza A virus. *Nat Commun* 2012; 3: 639.
- 119.** Sugita Y, Sagara H, Noda T, Kawaoka Y. The configuration of viral ribonucleoprotein complexes within the influenza A virion. *J Virol* 2013; 87: 12879-84.
- 120.** Goto H, Muramoto Y, Noda T, Kawaoka Y. The genome-packaging signal of the influenza A virus genome comprises a genome incorporation signal and a genome-bundling signal. *J Virol* 2013; 87: 11316-22.
- 121.** Hutchinson EC, Fodor E. Transport of the influenza virus genome from nucleus to nucleus. *Viruses* 2013; 5: 2424-46.
- 122.** Amorim MJ, Bruce EA, Read EK, *et al.* A Rab11- and microtubule-dependent mechanism for cytoplasmic transport of influenza A virus viral RNA. *J Virol* 2011; 85: 4143-56.
- 123.** Avilov SV, Moisy D, Munier S, Schraidt O, Naffakh N, Cusack S. Replication-competent influenza A virus that encodes a split-green fluorescent protein-tagged PB2 polymerase subunit allows live-cell imaging of the virus life cycle. *J Virol* 2011; 86: 1433-48.
- 124.** Bruce EA, Stuart A, McCaffrey MW, Digard P. Role of the Rab11 pathway in negative-strand virus assembly. *Biochem Soc Trans* 2012; 40: 1409-15.
- 125.** Eisfeld AJ, Kawakami E, Watanabe T, Neumann G, Kawaoka Y. RAB11A is essential for transport of the influenza virus genome to the plasma membrane. *J Virol* 2011; 85: 6117-26.
- 126.** Momose F, Sekimoto T, Ohkura T, *et al.* Apical transport of influenza A virus ribonucleoprotein requires Rab11-positive recycling endosome. *PLoS ONE* 2011; 6: e21123.
- 127.** Chou YY, Heaton NS, Gao Q, Palese P, Singer RH, Lionnet T. Colocalization of different influenza viral RNA segments in the cytoplasm before viral budding as shown by single-molecule sensitivity FISH analysis. *PLoS Pathog* 2013; 9: e1003358.
- 128.** Lakdawala SS, Wu Y, Wawrzusins P, *et al.* Influenza A virus assembly intermediates fuse in the cytoplasm. *PLoS Pathog* 2014; 10: e1003971.
- 129.** Essere B, Yver M, Gavazzi C, *et al.* Critical role of segment-specific packaging signals in genetic reassortment of influenza A viruses. *Proc Natl Acad Sci U S A* 2013; 110: E3840-8.

### **1.3. Apports récents en faveur d'un mécanisme d'empaquetage sélectif et impliquant des interactions ARN/ARN**

Depuis la publication de notre article, d'autres travaux sont venus étayer l'existence d'un mécanisme d'empaquetage sélectif, basé sur la présence de signaux spécifiques à chaque ARNv et permettant l'incorporation d'un jeu complet de huit ARNv au sein des particules virales.

#### **a. Rôle des régions non codantes (NCR, *non coding region*) dans l'empaquetage sélectif**

La longueur et la séquence des NCR (qui comprennent les séquences U12, U13 et les UTR), localisées aux extrémités 3' et 5' des segments HA et NA, sont spécifiques à chaque sous-type [264]. En présence d'un gène HA hybride, dans lequel les NCR et la région codante de HA proviennent de deux sous-types différents, les virus possèdent un défaut de répllication, attribuable à une diminution de l'empaquetage de l'ARNv HA, tandis que l'incorporation des autres segments n'est pas diminuée de manière significative. De plus, les NCR en 3' de l'ARNv HA jouent un rôle plus important dans l'empaquetage du segment HA que celles en 5', ce qui est en accord avec les travaux de Fujii *et al.* [255]. Ces données suggèrent que les NCR permettraient l'empaquetage sélectif de l'ARNv HA dans les particules virales, tandis que les nt dans la région codante seraient responsables des interactions intermoléculaires qui maintiennent le complexe supramoléculaire [264]. Un tel découplage entre les signaux a également été proposé dans le cas de l'empaquetage de l'ARNv NP de la souche WSN [265].

#### **b. Importance de la compatibilité entre signaux d'empaquetage**

Des virus chimériques capables d'empaqueter les segments HA et/ou NA d'un IVB dans un fond génétique d'IVA ne peuvent être générés, par génétique inverse, que lorsque les segments HA et/ou NA sont flanqués des séquences d'empaquetage des segments homologues d'un IVA [266]. Ces données suggèrent que si des interactions ARN/ARN sont effectivement requises pour un empaquetage sélectif, les segments chimériques obtenus interagiraient avec des segments issus du virus de sous-type A qui constituent le fond génétique. De la même manière, des virus de chauve-souris ne peuvent être produits par

génétique inverse, dans des cellules de mammifère ou de souris, que lorsque les protéines HA et NA de la souche influenza A/SC35M (H7N7), adaptée à la réplication chez les souris, sont exprimées en présence des six autres gènes du virus de chauve-souris H17N10. Ces virus chimériques ne peuvent cependant émerger que lorsque les segments génomiques HA et NA de la souche A/SC35M (H7N7) sont flanqués des séquences d’empaquetage du virus de chauve-souris [267]. Des résultats similaires ont été obtenus par Zhou *et al.* [268].

### **c. Importance des interactions ARN/ARN pour un co-empaquetage sélectif**

L’analyse des génotypes de souches vaccinales issues d’évènements de réassortiment génétique entre les souches humaines PR8 et H3N2 a révélé que la majorité des virus réassortants comportant les segments HA et NA du virus H3N2 dans le fond génétique PR8 contiennent également le segment PB1 du virus H3N2 [269–271]. De la même manière, lors du réassortiment génétique entre PR8 et la souche A/Udorn/307/72 (H3N2) (abrégée Udorn), une souche modèle de grippe saisonnière, le segment PB1 est retrouvé dans 75% des virions qui possèdent les segments HA et NA Udorn [270]. Néanmoins, cette combinaison de gènes ne confère pas d’avantage répliatif aux virus. Ces résultats soulignent donc que la sélection du segment PB1 Udorn n’est pas liée à la capacité répliatif des virions obtenus, mais à la présence des segments HA et NA Udorn.

Afin d’expliquer le mécanisme à l’origine de cette co-ségrégation, un système de génétique inverse compétitive à neuf plasmides a été développé et fait intervenir les plasmides exprimant les segments HA et NA Udorn, les six autres gènes de la souche PR8 et un plasmide exprimant le gène PB1 Udorn [272]. Dans les expériences de compétitions entre les segments PB1 des deux origines, les virions produits contiennent principalement le segment PB1 Udorn. Des expériences similaires ont montré que la co-ségrégation du segment PB1 est due au segment NA, et non au segment HA. L’utilisation de gènes PB1 PR8 chimériques, flanqués des séquences d’empaquetage (66 nt en 3’ et 83 nt en 5’) de l’ARNv PB1 Udorn, a révélé que l’empaquetage de l’ARNv PB1 PR8 est gouverné par les signaux identifiés dans la littérature [16, 257, 259, 261]. Néanmoins, la co-sélection des segments NA et PB1 Udorn n’est pas guidée par ces séquences d’empaquetage, mais implique une interaction entre l’ARNv NA et une région codante centrale de l’ARNv PB1 [272]. Nous avons

entrepris une collaboration avec l'équipe du Dr L. Brown dans le but de tester l'interaction entre les ARNv PB1 et NA de la souche Udorn et de délimiter plus précisément les séquences responsables de l'interaction (voir Résultats et discussion, § B.II).

#### 1.4. Conclusions

Les données disponibles à ce jour ne laissent plus de doute quant à l'existence d'un mécanisme d'empaquetage sélectif du génome des IVA au sein des particules virales. Le modèle actuel repose sur la présence, au site de bourgeonnement, d'un complexe supramoléculaire associant les huit ARNv, vraisemblablement *via* des interactions directes entre les ARNv. Ces interactions, qui impliquent des séquences spécifiques sur chaque segment, ne seraient que partiellement conservées entre les souches d'IVA [273]. Un tel complexe s'établit probablement durant le transport des RNPv vers la membrane plasmique. Cependant, de nombreuses questions restent encore à élucider concernant :

i) Le rôle général des interactions ARN/ARN dans le mécanisme d'empaquetage. En effet, la présence de signaux aux extrémités de chaque ARNv, cruciaux pour l'incorporation d'un jeu complet des huit segments d'ARN génomiques, est établie. Ce processus a été défini récemment par le groupe de M. Schwemmler comme le processus d'empaquetage du génome (*genome packaging*) [158]. Néanmoins, les mécanismes guidant la formation du complexe supramoléculaire entre les ARNv (*genome bundle* selon la nomenclature de M. Schwemmler) ou de sous-complexes (*sub-bundle*) sont encore vagues. Il est à noter que le sens des termes *genome packaging* et *genome bundling* utilisés par le groupe de M. Schwemmler diffère de celui qui leur a initialement été donné par le groupe de Y. Kawaoka [265], mais la nomenclature de M. Schwemmler nous semble moins sujette à confusion. S'il est tentant de postuler que les séquences impliquées dans l'empaquetage au site de bourgeonnement et celles impliquées dans la formation des sous-complexes d'ARNv dans le cytoplasme sont les mêmes, certaines données vont à l'encontre de cette hypothèse. Ainsi, alors que les ARNv PB1 et PA co-localisent dans le cytoplasme [148], au sein d'un même sous-complexe, des mutations qui affectent l'empaquetage de l'un des ARNv n'ont pas d'effet sur celui de l'autre [262]. Pour répondre à cette question, les mutations qui affectent l'empaquetage des ARNv PB1 et PA devraient être testées dans le contexte de la co-localisation lors du transport des RNPv vers la membrane plasmique.

ii) L'existence d'interactions physiques entre les RNPv dans les sous-complexes. En effet, les techniques d'imagerie ne permettent pas encore d'atteindre une résolution suffisante pour discriminer entre une simple proximité, qui pourrait résulter d'un regroupement de vésicules de recyclage Rab11 positives [159] (voir § A.V.6), et une réelle interaction entre RNPv.

iii) L'implication du mécanisme d'empaquetage dans les événements de réassortiment génétique (voir § D.II.2 et [247]).

## **2. Empaquetage du génome d'autres virus à ARN segmenté**

### **2.1. Le virus de la bursite infectieuse : un exemple d'empaquetage aléatoire**

Contrairement aux IVA, le virus de la bursite infectieuse (IBDV, *infectious bursal disease virus*), un virus aviaire de la famille des *Birnaviridae*, a adopté un mécanisme d'empaquetage aléatoire [274]. Alors que son génome est constitué de deux ARN double brin (db), l'IBDV incorpore jusqu'à quatre segments génomiques [275], ce qui augmente la probabilité d'empaqueter au moins une copie de chaque ARNv. De plus, l'infectiosité des particules virales augmente avec le nombre d'ARNv incorporé. Malgré un mécanisme d'empaquetage aléatoire, la présence de signaux en *cis* sur chaque ARNv est nécessaire pour discriminer les ARNv des ARN cellulaires. Ces signaux sont localisés dans les NCR en 5' et 3', mais également dans la région codante [274]. Contrairement aux IVA, il semblerait que le processus d'empaquetage ne nécessite pas d'interaction entre les deux ARNv [274].

### **2.2. Le virus de l'immunodéficience humaine, les *Reoviridae* et le bactériophage phi6 : des exemples d'empaquetage sélectif**

L'existence d'un mécanisme d'empaquetage sélectif n'est pas propre aux IVA. D'autres virus à ARN segmenté utilisent également un mécanisme pour sélectionner et incorporer un jeu complet d'ARNv.

#### **a. Le virus de l'immunodéficience humaine**

Peu de virus possèdent un génome polyploïde. Les *Retroviridae*, auxquels appartient le virus de l'immunodéficience humaine, constituent un exemple de virus possédant un

génomique diploïde. Dans ce cas, la dimérisation des deux copies d'ARN simple brin (sb) de polarité positive, qui est initiée par une interaction boucle-boucle, est étroitement couplée à son empaquetage spécifique. La polyprotéine virale Gag joue un rôle central dans la sélection de l'ARN génomique dans le cytoplasme des cellules infectées et l'assemblage des particules virales à la membrane plasmique. Cette sélection repose sur une interaction hautement spécifique entre la protéine Gag et une région de 110 nt, très structurée, située dans la 5' UTR de l'ARN génomique et qui contient le site de dimérisation de l'ARN génomique [276, 277].

### **b. Empaquetage du génome des *Reoviridae***

Le génome des *Reoviridae* est constitué de neuf à douze molécules d'ARN db. Deux exemples d'empaquetage sélectif sont explicités ci-dessous : celui des « bluetongue » virus et celui des rotavirus.

#### **i. « Bluetongue » virus**

Les « bluetongue » virus (BTV) sont transmis par les insectes et sont associés à une morbidité et une mortalité élevées chez les ruminants sauvages et domestiques, principalement les moutons. Leur génome est composé de dix ARN db, classés en trois groupes selon leurs tailles (S1-S3 : grands ARN, S4-S6 : ARN intermédiaires, S7-S10 : petits ARN) et codant chacun pour au moins une protéine virale. Ce sont les ARN, provenant de la transcription du génome, qui sont empaquetés dans les capsides virales, où se déroule la synthèse des ARN génomiques db. La région 3' UTR varie de 24 à 113 nt (la longueur étant spécifique à chaque segment, mais conservée entre les sérotypes), alors que la région en 5' est généralement plus courte. Afin d'identifier les signaux d'empaquetage indépendamment des signaux gouvernant les autres processus impliqués dans la réplication virale (transcription, traduction...), l'équipe de P. Roy a développé une technique d'assemblage de particules virales infectieuses *in vitro* (CFA, *cell-free assembly*), en présence des protéines structurales et des dix ARN viraux sb [278]. En combinant cette technique à des essais d'interactions ARN/ARN, il a été démontré que les ARNv du BTV forment un complexe, probablement *via* des interactions directes entre ARNv, qui est ensuite incorporé dans la capsid virale [279]. La formation de ce complexe semble séquentielle : elle est initiée par les segments les plus petits, S10 (*via* sa 3' UTR) et S7 [280], qui vont ensuite lier les ARN de

tailles intermédiaires, puis les grands ARN [279]. Récemment, en combinant bioinformatique et expériences *in vitro* de retard de migration sur gel, Boyce *et al.* [281] ont identifié deux interactions intermoléculaires, conservées (de six à onze pb), entre les ARN S5 et S10 d'une part, et S1 et S7 d'autre part, requises pour la production de virions infectieux générés par génétique inverse. De manière intéressante, le motif en tige-boucle de l'ARN S10 impliqué dans l'interaction avec l'ARN S5 [281], est localisé dans la même région 3' UTR nécessaire à l'initiation de la formation du complexe avec les ARN S7, S8 et S9 [280]. Par ailleurs, des motifs structuraux résultant de l'interaction entre les extrémités 5' et 3' des segments S4 et S8 sont suffisants pour leur empaquetage, *in vitro* et *in viro* [282].

## ii. Rotavirus

Les rotavirus sont la première cause mondiale de diarrhée aiguë sévère des enfants et infectent également les oiseaux. Leur génome est constitué de onze ARN db, mais ce sont les ARN sb de polarité positive qui sont incorporés dans les virions et servent ensuite de matrice à la synthèse des ARN génomiques db. L'interaction des extrémités 5' et 3' des ARN sb, incluant les UTR et une partie des ORF, confère aux ARN sb une structure globale en « poêle à frire » [283–285]. Les régions terminales, très conservées, sont spécifiques à chaque segment et contiennent probablement des tiges-boucles stables, servant de signaux d'empaquetage (pour revue, voir [286]). Ainsi, par bioinformatique, une interaction potentielle a été identifiée entre deux tiges-boucles (*kissing-loop complex*) [287]. Cependant, les données ne permettent actuellement pas de favoriser l'hypothèse d'un assemblage du génome avant empaquetage, comme c'est vraisemblablement le cas des IVA, ou d'une incorporation individuelle de chaque segment dans les particules pré-formées, comme dans le cas du bactériophage phi6 (voir ci-dessous). Néanmoins, les données penchent davantage vers la première hypothèse (pour revue, voir [286]).

## c. Empaquetage du génome du bactériophage phi6

Le bactériophage phi6 est un virus enveloppé appartenant à la famille des *Cystoviridae* et infectant les bactéries gram-négatives *Pseudomonis syringae*. Son génome est composé de trois segments d'ARN db (S, M et L) codant pour treize protéines. Les trois ARN, sous leur forme sb, comportent des séquences *pac* (200 nt) localisées dans les régions 5' UTR, différentes pour chaque ARN, qui s'organisent en tiges-boucles et ne chevauchent

pas les régions codantes [288]. Ces séquences sont nécessaires pour la sélection et l’empaquetage des ARNv, de manière spécifique et régulée, par la polymérase virale contenue dans une capsid préformée [289]. Des études structurales ont révélé qu’après l’incorporation du segment S, un changement conformationnel du « complexe polymérase » (composé de quatre protéines virales) de la capsid expose un site de fixation pour le segment M [290, 291]. Après l’empaquetage de cet ARN, un second changement expose le site de fixation du segment L. L’expansion de la capsid virale, due à l’introduction des trois segments, active la polymérase virale qui synthétise les brins d’ARN complémentaires.

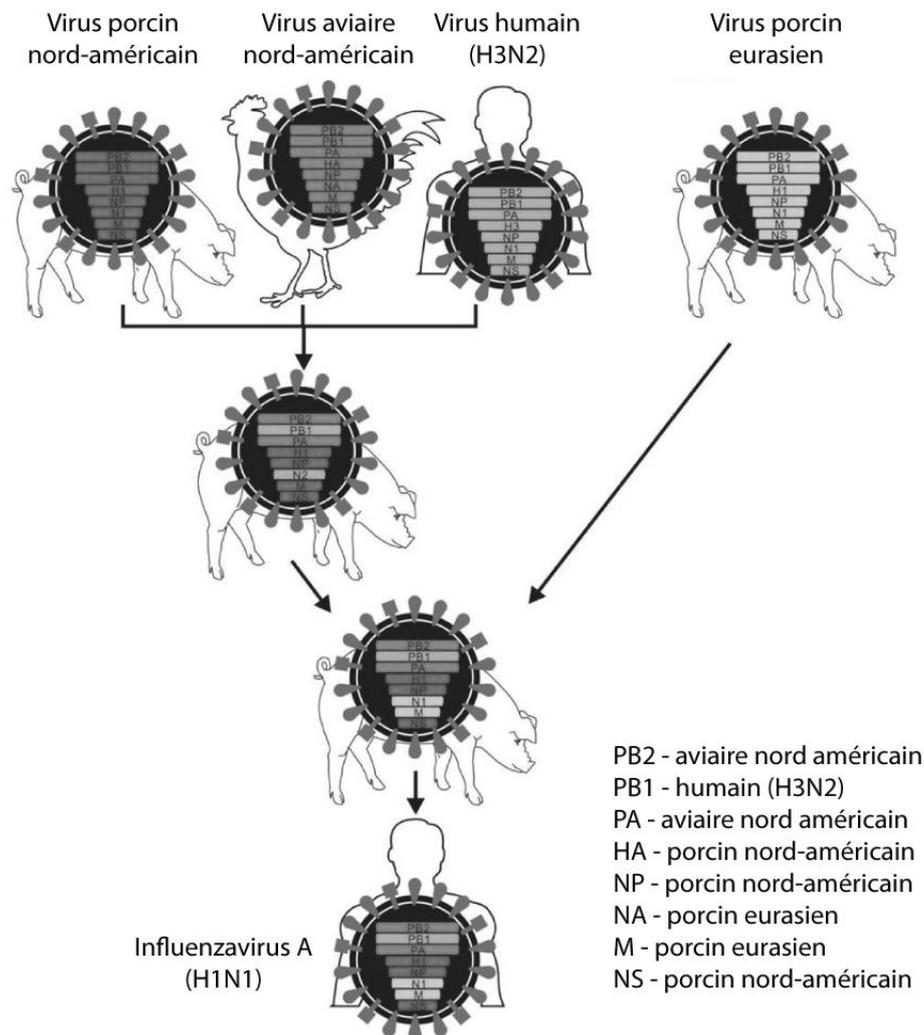
## **II. Le réassortiment génétique entre souches d’*Influenzavirus A***

### **1. Des évènements de réassortiment génétique à l’origine de pandémies**

Le réassortiment génétique est fréquent entre des virus d’un même sous-type infectant le même hôte et est notamment responsable de l’évolution des virus saisonniers circulants chez l’Homme [292–299]. À l’inverse, les évènements de réassortiments sont plus rares entre deux virus humains de sous-types H1N1 et H3N2, qui co-circulent dans la population humaine depuis 1977 [300–303]. De la même manière, le réassortiment génétique est rare entre des virus circulant chez des hôtes différents et a principalement lieu chez le porc. Certains virus réassortants, portant des HA et NA pour lesquelles la population humaine est immunologiquement naïve, peuvent être à l’origine de pandémies, à condition qu’ils se transmettent efficacement entre les individus (voir § B.IV).

Ainsi, au cours du XX<sup>ème</sup> siècle, plusieurs souches pandémiques ont émergé par réassortiment génétique : un virus H2N2 en 1957, entraînant la mort de deux millions de personnes (grippe asiatique) et un virus H3N2 responsable de plus d’un million de morts en 1968 (grippe de Hong Kong) [9]. Plus récemment, un virus H1N1 est apparu en 2009 suite à plusieurs évènements de réassortiment génétique entre des virus porcins, aviaire et humain [304, 305] (figure 18). D’après l’OMS, il a été à l’origine d’environ 18 000 décès cliniquement confirmés dans le monde. Une étude plus récente a estimé que ce nombre est largement sous-estimé et que la pandémie a probablement causé la mort de 150 000 à 250 000 personnes [306]. Cette estimation est proche de la mortalité imputée aux virus saisonniers, mais contrairement à ces derniers, le virus H1N1 a principalement infecté des personnes de

moins de 65 ans, tandis que les personnes nées avant 1957 présentaient une certaine immunité vis-à-vis de ce sous-type viral.



**Figure 18 : Schéma des événements de réassortiments génétiques ayant conduit à l'émergence d'un IVA de sous-type H1N1 à l'origine de la pandémie de 2009** (adapté de [3]). À la fin des années 1990, le réassortiment génétique chez le porc, entre une souche porcine, aviaire et humaine, a provoqué l'émergence d'un virus triple réassortant de sous-type H1N2. Un second événement de réassortiment génétique avec une souche porcine eurasienne, en 2009, a permis l'émergence d'un virus de sous-type H1N1, qui s'est rapidement transmis au sein de la population humaine.

## 2. Les conséquences d'un mécanisme d'empaquetage sélectif du génome sur le réassortiment génétique

Historiquement, la co-infection d'animaux ou de cellules en culture a été la première méthode à être utilisée pour étudier les événements de réassortiment génétique. Généralement, l'inoculation est réalisée à haute MOI (avec un ratio d'au moins trois particules virales pour une cellule), ce qui favorise l'infection d'une même cellule avec les

différentes souches virales. Néanmoins, les virus réassortants se répliquent souvent moins bien que les virus parentaux et il est donc difficile de les détecter et d'analyser leur génotype de manière subséquente. La génétique inverse a été adaptée pour étudier le réassortiment génétique. Les expériences de génétique inverse « compétitive » sont basées sur la co-transfection de neuf plasmides au minimum. Dans ce cas, au moins deux plasmides codent pour un même segment génomique provenant de deux souches virales différentes. Ainsi, le(s) segment(s) issu(s) des souches virales différentes sont en compétition lors de la sélection et de l'empaquetage des ARNv dans les particules virales. Depuis peu, cet outil a également été mis à profit pour étudier le co-empaquetage sélectif de segments génomiques au sein des virions [272, 273, 307, 308].

Théoriquement, lorsque deux virus co-infectent la même cellule, 256 génotypes différents (les deux génotypes parentaux et 254 nouvelles combinaisons) peuvent émerger. Néanmoins, à ce jour, la panoplie complète de ces 254 combinaisons génétiques différentes n'a jamais été observée lors de l'analyse de réassortiments naturels ou expérimentaux [309–311]. Dans des études de réassortiments génétiques naturels ou expérimentaux et dans des expériences de génétique inverse compétitive, le taux de réassortiment entre virus divergents est généralement faible et un nombre de génotype limité a été observé [307, 310–312]. De plus, certains gènes co-ségrègent [272, 307, 310, 313], ce qui suggère que le réassortiment génétique est non aléatoire et biaisé [272, 307] (pour revue, voir [247]).

Des incompatibilités entre les protéines virales peuvent expliquer l'absence de certaines combinaisons génomiques. De telles incompatibilités fonctionnelles ont été documentées, notamment entre les protéines virales formant les RNPv [311, 314], entre les protéines PB2 et PA de la souche humaine H3N2 et des souches H1N1 ou H5N1 [312] ou entre les protéines M et NS2 des souches H3N2 et H5N1 [314, 315]. Des incompatibilités entre les protéines virales et les protéines cellulaires sont également documentées (pour revue, voir [316]). Ainsi, des protéines de l'hôte interagissent tout au long du cycle viral avec les protéines qui forment les RNPv. Certains facteurs cellulaires sont notamment impliqués dans la transcription et la réplication du génome (ARN polymérase II, l'hélicase MCM), l'import nucléaire (importines  $\alpha$ ) ou dans l'export nucléaire (CRM1).

Néanmoins, certains virus, alors qu'ils n'émergent pas naturellement ou dans des expériences de réassortiment *in vitro*, peuvent être produits par génétique inverse forcée (c'est à dire en absence de compétition entre segments de deux souches virales) [272, 307]. Ces résultats suggèrent l'existence d'incompatibilités (ou de compatibilités sub-optimales) au niveau génomique, probablement entre les signaux d'empaquetage. Cette hypothèse est étayée par des travaux récents montrant que lorsque le réassortiment génétique n'est pas biaisé par la grande divergence et/ou des différences de capacités répliquatives entre souches parentales, il est très efficace [317]. En effet, 88% des virus obtenus contiennent un génotype réassortant après co-infection de cellules en culture et ce chiffre varie entre 46% et 86% dans le cas de la co-infection de cochons d'inde.

La première étape du réassortiment génétique consiste en un co-empaquetage, dans la même particule virale, des ARNv provenant des deux IVA parentaux. Le mécanisme garantissant l'empaquetage d'un jeu complet de huit ARNv a donc, très certainement, une implication directe dans le réassortiment génétique, en limitant notamment la panoplie de combinaisons génomiques possibles. Une analyse phylogénétique a d'ailleurs montré que l'interaction entre les ARNv PB1 et NS du virus aviaire H5N2, identifiée *in vitro* et dont le rôle fonctionnel a été validé, n'est conservée que dans la moitié des souches aviaires de sous-type H5N2 et ne l'est pas du tout parmi les IVA de sous-type H1N1 et H3N2 circulants dans la population humaine [273].

Des expériences de réassortiment génétique entre les souches humaine A/Moscow/10/99 (H3N2) (abrégée Moscou) et aviaire A/Finch/England/2051/91 (abrégée England) ont été réalisées au laboratoire et se sont focalisées sur les réassortants possédant le gène HA England. Après co-transfection ou co-infection, le pourcentage de réassortants mesuré sur un total de 200 clones ne s'élève qu'à 21% et 23%, respectivement, suggérant que le réassortiment génétique n'est pas favorisé entre des virus génétiquement différents [307]. De façon intéressante, le gène HA England ne s'intègre jamais seul dans le fond génétique H3N2, mais est toujours accompagné des segments PA et M England. Cependant, l'introduction des séquences terminales d'empaquetage du gène HA Moscou dans le squelette HA England permet une incorporation efficace du segment HA chimère dans le fond génétique humain en absence de tout autre gène d'origine aviaire, suggérant que les signaux d'empaquetage sont importants pour le réassortiment génétique. En accord avec

ces résultats, l'incorporation forcée du segment M England dans le fond génétique humain suffit au co-empaquetage du gène HA England. De manière importante, l'introduction de cinq mutations silencieuses au niveau du signal d'empaquetage en 3' de l'ARNv M Moscou, le rendant identique à l'extrémité 3' du gène M England, affecte fortement le génotype des virus réassortants et permet notamment l'empaquetage du gène HA England dans le fond génétique du virus Moscou [307]. Ces données ont constitué la première preuve quant à l'importance des interactions intermoléculaires ARN/ARN pour le réassortiment génétique et montrent que des incompatibilités ou des compatibilités sous-optimales, au niveau génomique, entre signaux d'empaquetage, sont un facteur limitant de ce processus.

Par la suite, les travaux de Cobbin *et al.* [272] ont montré que l'inclusion très fréquente du gène PB1 d'une souche humaine H3N2 dans des virus réassortants possédant aussi les gènes HA et NA humain dans un fond génétique PR8 ne s'explique pas par un avantage répliatif conféré par la protéine PB1 mais bien par un co-empaquetage sélectif, guidé par une interaction entre les segments PB1 et NA de la souche H3N2. De la même manière, l'absence de réassortiment entre les IVA et IVB, malgré des génomes, des cycles répliatifs et des protéines aux fonctions enzymatiques similaires, pourrait s'expliquer, entre autres, par des incompatibilités entre signaux d'empaquetage. En effet, des IVA chimériques, générés grâce à la présence des séquences d'empaquetage du genre A aux extrémités des segments HA et NA du genre B, sont alors capables de réassortir avec une souche circulante d'IVA [266].

Un des défis dans le domaine est de comprendre les règles qui gouvernent l'empaquetage du génome des IVA ainsi que celles qui régissent le réassortiment génétique, clairement biaisé, entre souches virales. Ces règles sont vraisemblablement étroitement liées. Leur compréhension permettrait d'améliorer l'évaluation des risques d'émergence de nouveaux IVA et le processus de production des vaccins.

## PROBLÉMATIQUE & SUJET DE RECHERCHE

Le génome des IVA est constitué de huit ARNv de polarité négative. La nature segmentée de leur génome représente un atout majeur pour l'évolution des IVA, par réassortiment génétique. *A contrario*, pour être infectieux, les virions doivent empaqueter au moins une copie de chaque ARNv, ce qui complique indéniablement l'étape d'assemblage viral. Actuellement, la multitude d'arguments génétiques, biochimiques, virologiques et structuraux ne laisse plus de doute quant à l'existence d'un mécanisme d'empaquetage sélectif permettant l'incorporation d'un jeu complet de huit ARNv. Ce modèle implique la présence de signaux distincts et spécifiques sur chaque segment, agissant en *cis* et en *trans*, situés aux extrémités des ARNv.

La plupart des travaux destinés à identifier les signaux d'empaquetage ont permis de délimiter plus ou moins précisément, *in cellula/in viro*, des régions sur les ARNv de deux souches de laboratoire (A/Puerto Rico/8/1934 et A/WSN/1933) de sous-type H1N1. Ces travaux ont également révélé qu'un défaut d'empaquetage d'un ARNv donné pouvait s'accompagner de défauts dans l'empaquetage d'autres segments génomiques. Ces résultats ont conduit notre laboratoire et d'autres à émettre l'hypothèse de l'existence d'un réseau entre les ARNv, maintenu par des interactions directes ARN/ARN. Pour étayer cette hypothèse, le laboratoire a développé une stratégie *in vitro* afin d'identifier les nt impliqués dans les interactions entre les segments génomiques. Les résultats obtenus sur un virus humain A/Moscow/10/99 (H3N2) et un virus aviaire A/Finch/England/2051/91 (H5N2) confortent un modèle selon lequel les huit ARNv sont sélectionnés et empaquetés sous la forme de complexes supramoléculaires, maintenus par des interactions directes entre les signaux d'empaquetage des ARNv. Nos études ont également montré que les séquences impliquées dans les interactions entre ARNv et leur co-empaquetage ne sont pas conservées entre les deux souches étudiées et que des incompatibilités entre signaux d'empaquetage de virus différents limitent vraisemblablement la panoplie de virus réassortants capables d'émerger.

Afin de i) progresser dans la compréhension des règles qui gouvernent l'empaquetage du génome des IVA au sein des virions, ii) définir l'existence potentielle de

spécificité d'hôte et/ou de sous-type, iii) délimiter précisément les séquences impliquées dans les interactions entre ARNv et iv) tester le rôle fonctionnel d'interactions, identifiées *in vitro*, dans un contexte infectieux, nous avons étudié la souche modèle de laboratoire de sous-type H1N1, A/Puerto Rico/8/1934, dont les signaux d'empaquetage sont relativement peu définis dans la littérature. Un autre intérêt de cette souche humaine est son utilisation pour la production des vaccins saisonniers, qui sont générés par réassortiment génétique entre PR8 et les souches virales circulantes. La compréhension des règles qui gouvernent l'empaquetage du génome de la souche PR8 présente par conséquent un intérêt indéniable.

Pour atteindre ces objectifs, nous avons dans un premier temps, par des expériences de retard de migration sur gel, déterminé le réseau d'interactions formé entre les huit ARNv de la souche PR8. Sa comparaison aux deux autres réseaux précédemment définis au laboratoire nous a donné des pistes quant à la problématique de l'existence de potentielles spécificités d'hôte et/ou de sous-type dans l'établissement du réseau.

Dans un second temps, nous avons voulu définir les régions des ARNv impliquées dans la formation de ce réseau, afin d'identifier, pour certains couples de partenaires, les séquences d'interaction au nt près. Deux méthodes ont été mises en œuvre :

i) une approche « classique » combinant l'utilisation d'ARNv portant de grandes délétions, d'oligonucléotides antisens et de la bioinformatique. Cette approche avait été validée, dans un contexte infectieux, sur une interaction ARN/ARN du virus aviaire England.

ii) une approche plus innovante, destinée à identifier les nt importants pour l'établissement des interactions ARN/ARN de façon **exhaustive, efficace** et **non-biaisée**. Une telle stratégie, nommée MIME pour « Mutational Interference Mapping Experiment », avait été développée au laboratoire par le Dr R. Smyth et a fait ses preuves dans une étude similaire sur le VIH-1 afin d'identifier les séquences importantes pour la reconnaissance du génome du VIH-1 par la polyprotéine Gag impliquée dans la formation des particules virales lors du bourgeonnement. L'approche MIME allie la mutagenèse aléatoire au séquençage à haut débit.

Parmi les interactions identifiées *in vitro* par ces deux approches, certaines ont été validées *in vitro* par l'introduction de mutations *trans*-compensatoires dans chacun des ARNv partenaires. Une validation du rôle de ces partenariats établis *in vitro*, pour une

réplication virale optimale et le co-empaquetage sélectif des paires d'ARNv, a ensuite été entreprise en collaboration avec le laboratoire du Dr M. Rosa-Calatrava (CIRI, Lyon).

Des travaux menés précédemment au laboratoire avaient permis d'identifier, *in vitro* et au nt près, la séquence d'interaction entre les ARNv NA et M de la souche Moscou. Cette interaction avait été validée *in vitro*, mais les résultats préliminaires de l'étude *in viro* restaient à confirmer. J'ai repris cette étude, construit plusieurs nouveaux couples de mutants *trans*-compensatoires et silencieux qui ont été testés *in vitro*, puis dans un contexte infectieux.

Enfin, lors de la dernière année de mon doctorat, nous avons également eu l'opportunité de collaborer avec l'équipe du Pr L. Brown de l'Université de Melbourne qui s'intéresse à la souche modèle de grippe saisonnière A/Udorn/307/72 (H3N2). Des travaux antérieurs avaient mis en évidence une co-ségrégation de deux ARNv de la souche Udorn, les ARNv NA et PB1, dans un fond génétique PR8, *via* une région codante centrale du segment PB1. Notre objectif était de tester, dans notre système *in vitro*, les interactions entre les ARNv NA Udorn et PB1, sauvages et mutants, afin d'évaluer si les résultats obtenus *in viro* peuvent être expliqués par des aptitudes différentes des ARNv à interagir entre eux.



## MATERIELS & MÉTHODES

### A. Clonage des séquences des segments viraux d'*Influenzavirus A* humains et mutagenèse

#### I. Souches virales utilisées

Dans ce travail, nous avons étudié trois souches humaines d'IVA :

- A/Puerto Rico/8/1934 (H1N1) (PR8), isolée en 1934 et considérée comme une souche modèle de laboratoire. La souche PR8 se réplique très bien sur les œufs et sert de fond génétique pour l'élaboration des vaccins saisonniers.
- A/Moscow/10/99 (H3N2) (Moscou), une souche contemporaine circulante. Elle est entrée dans la composition des sémences vaccinales entre 2000-2001 et 2003-2004.
- A/Udorn/307/72 (H3N2) (Udorn), une souche modèle de laboratoire de grippe saisonnière.

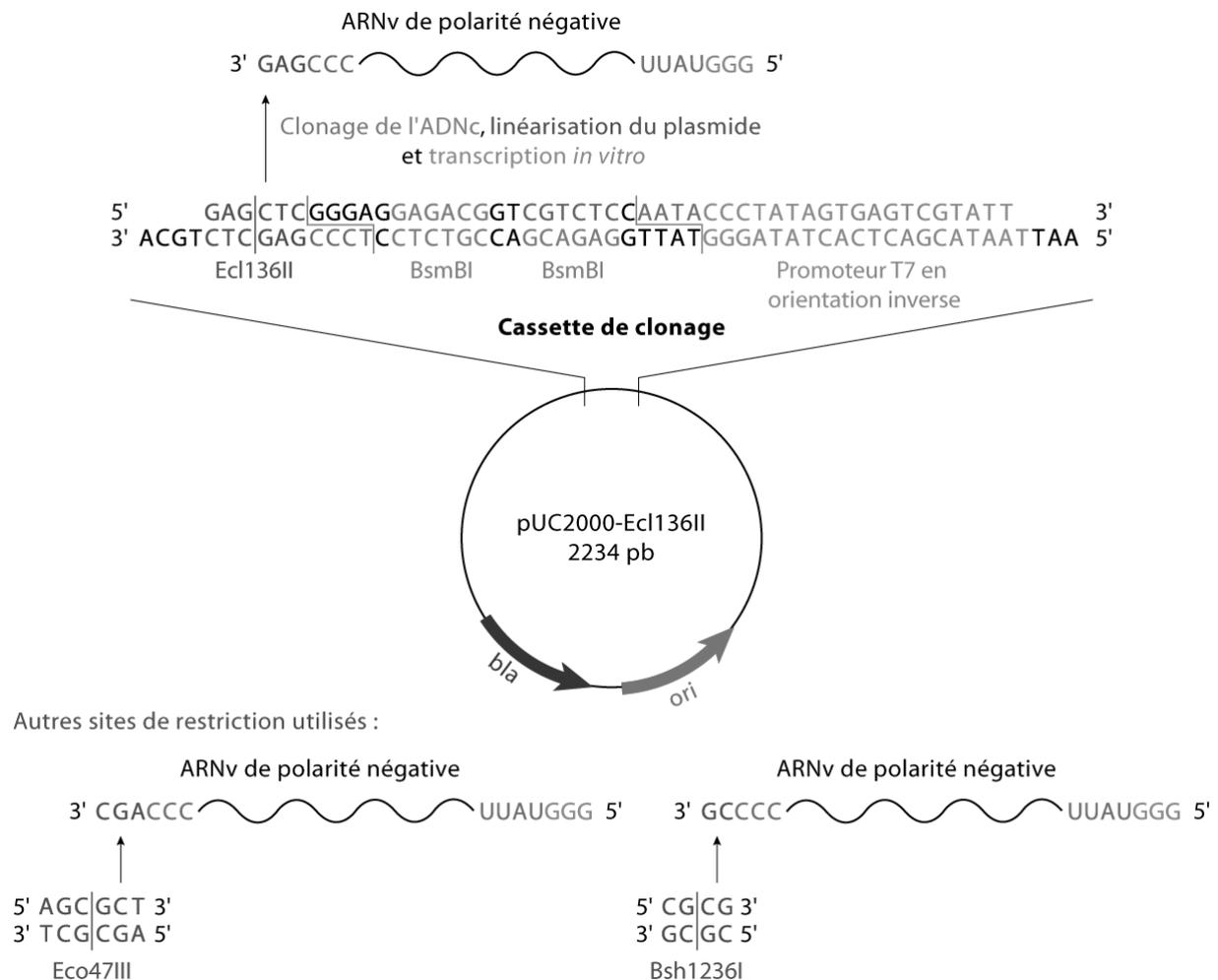
#### II. Plasmides utilisés

Les séquences permettant l'expression des ARNv des souches PR8, Moscou et Udorn ont été clonées dans les plasmides pUC2000 (utilisés pour les expériences *in vitro*) et pHW2000 (utilisés pour les expériences *in cellula/in viro*). Pour des informations détaillées sur la construction des plasmides et les clonages, réalisés avant mon arrivée au laboratoire, voir la thèse de C. Gavazzi (UdS, 2011).

##### 1. Plasmide d'expression *in vitro* pUC2000

Le vecteur pUC2000 dérive du vecteur commercial pUC19 (GenBank Accession # : L09137) (figure 19). La cassette de clonage a été modifiée pour en construire une nouvelle possédant : un promoteur T7 en orientation antisens ; deux sites de reconnaissance pour l'enzyme de restriction BsmBI, permettant de cloner la séquence d'un segment d'IVA ; un site de restriction unique (Eco47III, Bsh1236I ou Ecl136II selon la cassette concernée) (tableau 2) afin de linéariser le plasmide pour arrêter la transcription *in vitro* des ARNv de polarité négative. Le plasmide contient également le gène *bla*, codant une  $\beta$ -lactamase qui confère une résistance à l'ampicilline, ainsi qu'une origine de réplication ori.

L'ARNv de polarité négative transcrit par l'ARN polymérase du bactériophage T7 (voir § B.II) possède quelques nt supplémentaires en 3' et en 5', qui correspondent à la moitié du site de restriction utilisé pour arrêter la transcription, aux séquences restantes des sites de clonage BsmBI et aux trois G du promoteur T7 (figure 19).



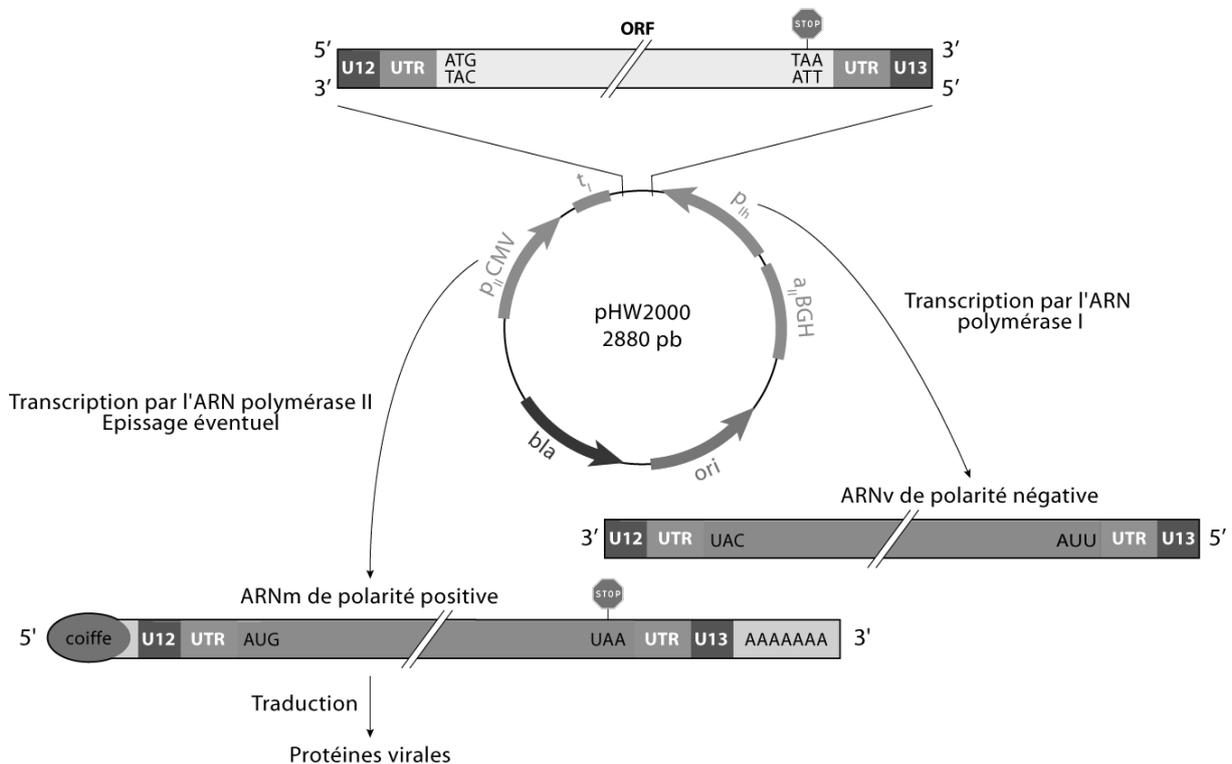
**Figure 19 : Carte du vecteur pUC2000, dérivé du vecteur commercial pUC19** (GenBank Accession # : L09137) (pour plus d'informations voir la thèse de C. Gavazzi, UdS, 2011). Le schéma n'est pas réalisé à l'échelle. Le vecteur pUC2000-Ecl136II est représenté et comporte le gène *bla*, localisé aux coordonnées 2032-1172, qui code une  $\beta$ -lactamase conférant une résistance à l'ampicilline et une origine de répllication *ori* entre les nt 1001-413. La cassette de clonage possède un promoteur T7 en orientation antisens, deux sites de reconnaissance pour l'enzyme de restriction BsmBI, nécessaires au clonage, et un site de restriction (Ecl136II, Eco47III ou Bsh1236I) situé en 5' de l'insert. En plus du site de restriction Bsh1236I localisé au niveau de la cassette de clonage, le vecteur pUC2000 possède six autres sites Bsh1236I (situés aux coordonnées 198, 398, 979, 1039, 1802 et 2134). Le vecteur ne contient aucun autre site de restriction Ecl136II et Eco47III. Après linéarisation du plasmide, la transcription avec l'ARN polymérase du bactériophage T7 permet d'obtenir un ARNv de polarité négative qui contient quelques nt supplémentaires (en gris et mauve sur les représentations schématiques des ARNv) en 3' et en 5'.

ARNv	Souche PR8	Souche Moscou	Souche Udorn
PB2	Ecl136II	/	/
PB1	Bsh1236I	/	Bsh1236I
PA	Bsh1236I	/	/
HA	Bsh1236I	/	/
NP	Bsh1236I	/	/
NA	Ecl136II	Bsh1236I	Ecl136II
M	Bsh1236I	Ecl136II	/
NS	Ecl136II	/	/

**Tableau 2 : Sites de restriction présents dans la cassette de clonage des vecteurs pUC2000, en 5' de chaque segment sauvage ou mutant.**  
/ : segment non étudié dans ce travail.

## 2. Plasmide d'expression *in cellula/in viro* pHW2000

Le système de génétique inverse utilisé fait intervenir huit plasmides d'expression bidirectionnels pHW2000 où chaque segment est cloné entre deux promoteurs en orientation inverse (figure 20). Le promoteur PolI ( $p_{ih}$ ), associé à un terminateur PolI ( $t_i$ ), est encadré par le promoteur PolII du cytomégalovirus humain ( $p_{iiCMV}$ ) et le signal de polyadénylation isolé du gène codant l'hormone de croissance bovine ( $a_{iiBGH}$ ). Ainsi, après transfection, la transcription par l'ARN polymérase I cellulaire permet la synthèse des ARN viraux génomiques, de polarité négative, tandis que la transcription par l'ARN polymérase II cellulaire permet l'expression des ARNm viraux qui sont ensuite traduits en protéines virales.



**Figure 20 : Schéma représentant le vecteur pHW2000 et ses produits de transcription.** Le promoteur  $p_{ih}$ , en antisens, permet la synthèse de l'ARNv de polarité négative, tandis que le promoteur  $p_{iiCMV}$  permet la synthèse des ARNm viraux, qui seront ensuite traduits en protéines virales.

### III. Mutagenèse dirigée

La mutagenèse dirigée par PCR a été réalisée selon la stratégie de mutagenèse QuickChange® (Stratagene) à partir des plasmides pUC2000 ou pHW2000 contenant les séquences des gènes d'IVA. L'utilisation d'amorces spécifiques permet d'éliminer une partie de la séquence (tableau 3) ou d'insérer des mutations ponctuelles (tableau 4). Dans les deux cas, tout le plasmide est amplifié, hormis éventuellement la région à éliminer. Les plasmides obtenus permettent, après transcription *in vitro*, de générer des ARNv porteurs de délétions ou de substitutions.

ARNv	Région délétée	Amorce sens (de 5' vers 3')	Amorce antisens (de 5' vers 3')
PB1	13-163	GGGGAGCGAAAGCAGGATCAGTACTCAGAAAAGG	CCTTTTCTGAGTACTGATCCTGCTTCGCTCCCC
	163-414	GGTACTGTCAACAGGACACTGGACTCTAAATGAAAAC	GGTTTCTATTAGAGTCCAGTGTCTGTTGACAGTATCC
	415- 665	CAAGGCCGACAGACCTATGACAAGGAGTTATCTAATTAGAGC	GCTCTAATTAGATAACTCCTTGTCTAGGTCTGTGCGCCTTG
	666- 916	GAAGCAGAGATTGAACAATTTCTTCCACCATCACTG	CAGTGATGGTGAAGAAATTGTTCAATCTCTGCTTC
	917- 1168	CCAATTCTCAGGACACCGAACATGATTCAACAAGAAAAG	CTTCTTTCTGTTGAATCATGTTCCGGTGTCTGAGAATTGG
	1169- 1420	CGATTTGAAATATTTTATTTATCGAACCTGTAAGC	GCTTACAGGTTGATAAATGAAATATTTCAAATCG
	1421- 1673	CAAGCCGGAGTCGACAGGTGTTTCATCAAAGATTACAGG	CCTGTAATCTTTGATGAACACCTGTGACTCCGGCTTG
	1674- 1925	GCTCAAATGGCCCTTCAGTTAAAGAAATTGAATCAATG	CATTGATTCAATTTCTTAAACTGAAGGGCCATTTGAGC
	1926 – 2177	CTGAACCCATTTGTCAGCCAGGTTCCAGAGCCCGAATTG	CAATTCGGGCTCTGAAACCTGGCTGACAAATGGGTTCCAG
	2178-2328	GTATGGTGGAGGCTATCCTTGTCTACTAATAC	GTATTAGTAGAAAACAAGGATAGCCTCCACCATAC
PA	13-162	GGGAGCGAAAGCAGGATGTATTAGATTTTC	GAAAATCTGAATACATCCTGCTTCGCTCCC
	163-432	CTTGGAAGTATGCTTCAAATTAATCTGAGAAAAC	GTTTTCTCAGATTTAATTTGAAGCATACTTCCAAG
	433-702	CTGAAAAGGCCAATGAAAATTTTAGAGCCTATG	CATAGGCTCTAAAATTTTATTGGCCTTTTCCAG
	703-972	GAACCTCTCCAGCCTTTGGAAGGAACCAATG	CATTGGGTTCTTCCAAGGCTGGAGAAGTTC
	973-1272	GAGAACATTCTTGGACTGACAGATTCAAGCTG	CAGCTTGAATCTGTGAGTCAAAGAATGTTCTC
	1273-1542	GTTTAAACAAGGCATGCGAAGGAAGATCCCACTTAAG	CTTAAGTGGGATCTTCTTCGATGCCTTGTAAAC
	1543-1812	GTATGGTTTCATCATAAAGAAGCTGAGTCCTCTG	CAGAGGACTCAGCTTCTTTATGATGAAACCATAC
	1813-2070	CAAATTGAGAGTATGATTCTGGGGGGCTATATG	CATATAGCCCCCAACAATGATACTCTCAATTTG
	2071-2220	CCTGGGACCTTTGATCCTTGTCTACTAATAC	GTATTAGTAGAAAACAAGGATCAAAGGTCCCAGG
	497-506	GAAGAAATGGCCACAACACTCTCGATGAAG	CTTCATCGAGAGTGTGTTGGCCATTTCTTC
	1049-1058	GGAAGCAAGTACGACAGGACATTGAG	CTCAATGTCTCGTACTTGTCTCC
	1059-1065	GTAAGGCAAGTATTGAGAATGAGG	CCTCATTCTCAATAGTTCTGCCAGTAC
M	13-112	GGGAGCAAAGCAGGGATGTCTTGCAGGG	CCCTGCAAAGACATCCCTGCTTTTGCTCCC
	915-1014	CCTTCTACGGAAGGACCTTGTCTACTAATAC	GTATTAGTAGAAAACAAGGCTTCCGTAGAAGG

**Tableau 3 : Amorces de mutagenèse dirigée utilisées pour réaliser des délétions au niveau des segments de la souche PR8.**

Souche d'IVA	Nom du mutant	Interaction étudiée	Amorce sens (de 5' vers 3')	Amorce antisens (de 5' vers 3')
PR8	PB2 mc 1	PB2/NP	CAAAATTGGGGAGTAGAGCCTATCGACAATGTG	CACATTGTCGATAGGCTCTACTCCCAATTTTG
	PB2 mc 2		CAAAATTGGGGGTAGAGCCCATCGACAATGTG	CACATTGTCGATGGGCTCTACACCCAATTTTG
	PB1 mc	PB1/M	CCCTTACTGGAGATCCACCCTACAGCCATGGGAC	GTCCCATGGCTGTAGGGTGGATCTCCAGTATAAGGG
	NP mc 1	PB2/NP	GCTCTGTATGCAAGGCTCTACTCTCCCTAGGAGGTC	GACCTCTAGGGAGAGTAGAGCCTGCATCAGAGAGC
	NP mc 2		GCTCTGTATGCAGGGCTCTACACTCCCTAGGAGGTC	GACCTCTAGGGAGTGTAGAGCCTGCATCAGAGAGC
	M mc	PB1/M	GCTTTAAATACGGACTGAAGGGTGGACCTTCTACGG AAGGAG	CTCCTCCGTAGAAGGTCCACCTTCAGTCCGTATTTA AAGC
Moscou	NA mc 1	NA/M	TGTTTTGTGGCACCTCCGGAACATATGGAACAGGC	GCCTGTTCCATATGTTCCGGAGGTGCCACAAAACA
	NA mc 2		TGTTTTGTGGCACTTCCGGTACATATGGAACAGGC	GCCTGTTCCATATGTACCGGAAGTGCACAAAACA
	NA mc 3		TGTTTTGTGGCACTTCAGGAACATATGGAACAGGC	GCCTGTTCCATATATACCGTAAGTGCACAAAACA
	M mc 1	NA/M	TTCTACGGAAGGAGTTCGGAGTCTATGAGGGAAG	TTCTACGGAAGGAGTTCGGAGTCTATGAGGGAAG
	M mc 2		TTCTACGGAAGGAGTACCGGAATCTATGAGGGAAG	CTTCCCTCATAGATTCCGGTACTCCTCCGTAGAA
	M mc 3		TTCTACGGAAGGAGTTCCTGAATCTATGAGGGAAG	CTTCCCTCATAGATTGAGGAACCTCCTCCGTAGAA

**Tableau 4 : Amorces de mutagenèse dirigée utilisées pour réaliser des mutations ponctuelles au niveau des segments de la souche PR8 ou Moscou. mc : mutant *trans*-compensatoire.**

Selon les mutations réalisées, cinq conditions de PCR différentes sont utilisées à partir de 40 ng de plasmide matrice (tableau 5). Après 4 min de dénaturation à 94°C, 25 à 30 cycles d'amplification sont réalisés (1 min de dénaturation à 94°C, 1 min d'hybridation des amorces à 55°C, 1 min d'élongation par kb à 72°C), suivis par 10 min d'une élongation finale à 72°C.

	A	B	C	D	E
Tampon HF (NEB)	1x		/		
Tampon GC (NEB)	/		1x		/
Tampon ThermoPol (NEB)	/				1x
MgCl <sub>2</sub> (NEB)	/	50 nmol	/	50 nmol	/
dNTP (Promega)	10 nmol				
Amorce sens	25 pmol				
Amorce antisens	25 pmol				
Phusion High-Fidelity DNA Pol (NEB)	1 U				/
VentR® DNA Pol (NEB)	/				2 U

**Tableau 5 : Conditions de PCR utilisées pour l'obtention des mutants selon la stratégie de mutagenèse QuickChange®.**

L'ADN plasmidique matrice est ensuite digéré par 10 U d'enzyme de restriction DpnI (Fermentas) pendant 2-3 heures (h) à 37°C. Cette endonucléase coupe après les adénines des sites 5' GATC 3' lorsque l'ADN est méthylé, ce qui est le cas de l'ADN matrice sauvage. Pour finir, 100 µl de bactéries *E. coli* DH5α chimio-compétentes, préparées au laboratoire, sont transformées avec 2 µl de produits de PCR digérés par DpnI. Le mélange est incubé

pendant 30 min sur la glace. Après un choc thermique de 1 min à 42°C, les bactéries sont incubées 10 min sur la glace, puis 30 min à 37°C après addition de 1 ml de LB. Après centrifugation pendant 30 secondes (s) à 15700 *g* à température ambiante, 1 ml de surnageant est retiré et le culot bactérien est remis en suspension dans les 100 µl de milieu restant. Après étalement sur boîtes de Pétri LB-Agar contenant 0,1 mg/ml d'ampicilline, les boîtes sont incubées à 37°C pendant la nuit.

## **B. Techniques d'étude *in vitro* des interactions entre les ARNv par retard de migration sur gel**

### **I. Digestion des plasmides pUC2000**

Les plasmides pUC2000 contenant les séquences codant les ARNv sauvages ou mutés sont linéarisés pendant une nuit à 37°C avec 0,4 U d'enzyme de restriction par µg d'ADN (Bsh1236I ou Ecl136II (Fermentas) selon les cas) (tableau 2), en présence du tampon de l'enzyme 1x (les tampons R ou Ecl136II, respectivement). Après une extraction au phénol-chloroforme (v/v), la phase aqueuse, contenant l'ADN, est additionnée de 0,1 vol d'AcNa 3 M (pH 5,6) et de 3 vol d'éthanol absolu. Le tout est précipité pendant 30 min dans un bain carboglace-éthanol, puis centrifugé pendant 30 min à 15700 *g* à 4°C. Les culots sont ensuite lavés deux fois avec de l'éthanol 80%, séchés 5 min sous vide au SpeedVac et remis en suspension dans de l'eau milliQ (Millipore). La linéarisation du plasmide est vérifiée par migration de 0,8 µg d'ADN digéré et non digéré sur un gel d'agarose 0,8% (p/v)/TBE 0,5x (Tris 50 mM, acide borique 44,5 mM) (Euromedex) contenant 0,4 µg/ml de BET (Roth).

### **II. Transcription *in vitro* et purification des ARNv sur colonne de gel filtration**

Pour chaque transcription, 30 à 50 µg d'ADN plasmidique linéarisé sont incubés 3h à 37°C dans un volume final de 300 µl contenant 40 mM de tampon Tris-HCl (pH 8 à 37°C), 15 mM de MgCl<sub>2</sub>, 50 mM de NaCl, 4 mM de chacun des quatre NTP (pH 7), 0,01% (v/v) de Triton X-100, 400 U de Rnasin® Rnase inhibitor, 1,7 mM de spermidine, 5 mM de DTT et 3 µl d'ARN polymérase du bactériophage T7 produite et purifiée au laboratoire à partir de la souche surproductrice BL21/pAR1219. Après une incubation de 2h à 37°C en présence de 50 U de DNase I recombinante Rnase-free (Roche) et du tampon de l'enzyme concentré 1x, une extraction au phénol-chloroforme (v/v) est réalisée. L'ARN, contenu dans la phase aqueuse,

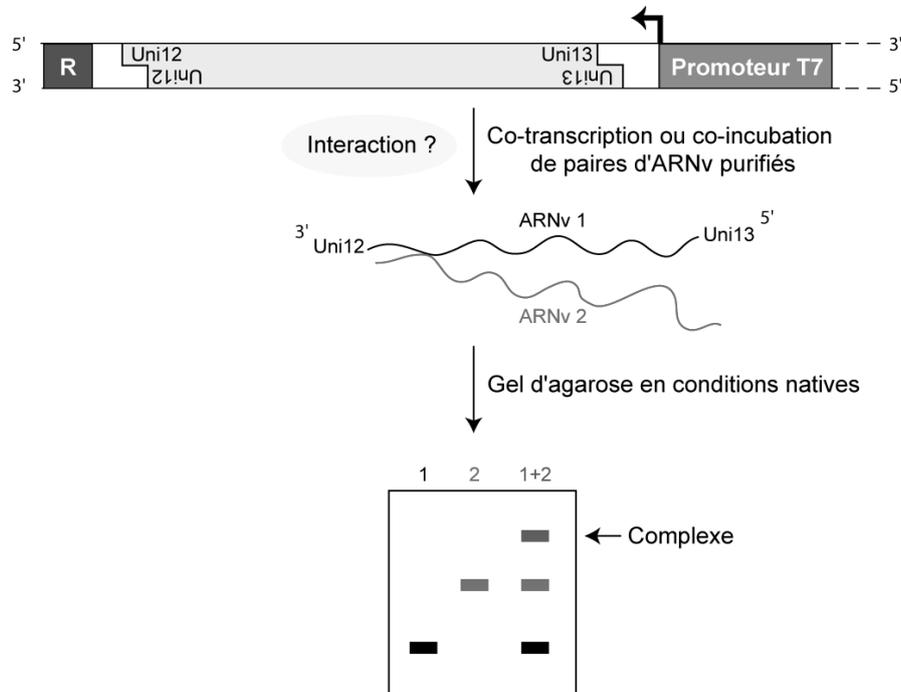
est additionné de 0,1 vol d'AcNa 3 M (pH 5,6) et 2,5 vol d'éthanol absolu, puis mis à précipiter 30 min dans un bain carboglace-éthanol. Après centrifugation (15700 *g* à 4°C pendant 30 min), les culots sont lavés deux fois avec 100 µl d'éthanol 80%, séchés à l'air libre et remis en suspension dans 500 µl d'eau milliQ (Millipore).

Les ARN sont ensuite purifiés par HPLC sur une colonne de gel filtration TSK G2000SW (Tosoh Bioscience) installée sur un système ÄKTA (GE Healthcare), dans un tampon contenant 0,2 M AcNa (pH 6,5) et 1% de méthanol à un débit de 1 ml/min. Après ajout de 3 vol d'éthanol absolu et 10 µl d'AcNa 3 M (pH 5,6), les fractions (300 µl) contenant les ARN sont précipitées 30 min dans un bain carboglace-éthanol. Après centrifugation (30 min à 15700 *g* et à 4°C), les culots sont lavés deux fois à l'éthanol 80%, séchés à l'air libre et remis en suspension dans de l'eau milliQ (Millipore). La pureté et l'intégrité des ARNv sont contrôlées sur gel de polyacrylamide 8% en conditions dénaturantes.

### **III. Co-transcription *in vitro* des ARNv deux à deux et retard de migration sur gel d'agarose (figure 21)**

Pour chaque co-transcription, 250 ng de chacun des plasmides linéarisés (voir § B.I) sont incubés 3h à 37°C dans un volume final de 10 µl contenant 40 mM de tampon Tris-HCl (pH 8 à 37°C), 15 mM de MgCl<sub>2</sub>, 50 mM de NaCl, 4 mM de chacun des quatre NTP (pH 7), 0,05% (v/v) de Triton X-100, 1 U de Rnasin® Rnase inhibitor, 1 mM de spermidine, 5 mM de DTT, 0,5 µg de BSA et 0,167 µl d'ARN polymérase du bactériophage T7. Ensuite, 5 µl de solution de dépôt (glycérol 50 % (v/v), xylène-cyanol 0,1 % (p/v), bleu de bromophénol 0,1 % (p/v)) sont ajoutés aux échantillons, qui sont déposés sur un gel d'agarose 0,8% (p/v)/TBM 0,5x (Tris 50 mM, acide borique 44,5 mM, MgCl<sub>2</sub> 0,1 mM) contenant 0,4 µg/ml de BET (Roth). Après 4h de migration à 4°C et à 150 V dans le tampon TBM, le gel est trans-illuminé aux UV et analysé avec un ChemiDoc (Bio-Rad) associé au logiciel Quantity One ou avec un GelDoc (Bio-Rad) associé au logiciel ImageLab (Bio-Rad). La fraction en masse (%) des ARNv correspondant à chaque bande est déterminée à l'aide des logiciels Image Gauge (Fujifilm) ou ImageLab (Bio-Rad), et le pourcentage d'ARNv impliqué dans le complexe intermoléculaire est calculé en divisant la fraction en masse de l'hétérodimère par la somme des fractions en masse des ARNv monomériques et hétérodimériques issus de la même piste. Dans les expériences de retard de migration sur gel en présence d'ARNv mutés

(portant des délétions ou des mutations ponctuelles) ou d'oligodésoxyribonucléotides antisens (voir § B.V), le pourcentage de chaque complexe hétérodimérique est normalisé par rapport au complexe formé entre les deux ARNv sauvages ou en l'absence d'oligodésoxyribonucléotide.



**Figure 21 : Schéma de la stratégie *in vitro* de retard de migration sur gel développée au laboratoire dans le but de mettre en évidence des interactions ARN/ARN.** Les séquences permettant l'expression des ARNv sont clonées sous le contrôle du promoteur de l'ARN polymérase du phage T7. Les ARN sont co-transcrits par paires ou transcrits séparément, purifiés et incubés deux à deux. D'éventuelles interactions ARN/ARN sont ensuite visualisées par retard de migration sur gel d'agarose en conditions natives. R : site de restriction.

#### IV. Co-incubation deux à deux des ARNv purifiés et retard de migration sur gel (figure 21)

Pour les expériences de co-incubation, 2 pmoles de chacun des deux ARNv purifiés, sauvages ou mutants, sont mélangés et le volume est complété à 8  $\mu$ l avec de l'eau milliQ (Millipore). Le tout est ensuite dénaturé 2 min à 90°C et refroidi rapidement pendant 2 min sur la glace. Après avoir ajouté 2  $\mu$ l d'un tampon de faible salinité concentré 5x (concentration finale : 50 mM cacodylate de sodium pH 7,5, 40 mM KCl, 0,1 mM MgCl<sub>2</sub>), les échantillons sont incubés pendant 30 min à 55°C. Les échantillons sont ensuite traités comme précédemment (§ B.III).

## V. Co-incubation deux à deux des ARNv en présence d'oligodésoxyribonucléotides antisens et retard de migration sur gel

Dans le but de déterminer plus précisément le rôle de courtes régions des ARNv dans l'établissement des interactions intermoléculaires observées lors des expériences de retard de migration sur gel, nous avons utilisé des oligodésoxyribonucléotides de synthèse (Integrated DNA Technologies) complémentaires à ces régions (tableau 6). Le protocole est le même que précédemment (§ B.IV), à l'exception de l'ajout de 10 pmol d'oligodésoxyribonucléotides dans le tube contenant les ARNv, avant la dénaturation.

Interaction étudiée	Nom de l'oligonucléotide	Séquence (de 5' vers 3')
<b>PB1/M</b>	Oligo anti-PB1 13-42	CAAACCATTGGAATGGATGTCAATCCGACC
	Oligo anti-PB1 43-73	TTACTTTTCTTAAAAGTGCCAGCACAAAATG
	Oligo anti-PB1 74-103	CTATAAGCACAACTTCCCTTATACTGGAG
	Oligo anti-PB1 104-137	ACCCTCCTTACAGCCATGGGACAGGAACAGGATA
	Oligo anti-PB1 138-163	CACCATGGATACTGTCAACAGGACAC
	Oligo anti-M 155-183	GAATGGCTAAAGACAAGACCAATCCTGTCAC
	Oligo anti-M 327-352	GGAAGCTCAAGAGGGAGATAACATTC
	Oligo anti-M 592-614	GGAGCAAATGGCTGGATCGAGTG
	Oligo anti-M 885-908	GGACTGAAAGGAGGGCCTTCTACG
<b>PA/NP</b>	Oligo anti-PA 973-1008	TGGAAGGAACCAATGTTGTTAAACCACACGAAAAG
	Oligo anti-PA 1009-1046	GGAATAAATCAAATTATCTTCTGTCATGGAAGCAAGT
	Oligo anti-PA 1047-1083	ACTGGCAGAACTGCAGGACATTGAGAATGAGGAGAAA
	Oligo anti-PA 1084-1120	ATTCCAAGACTAAAAATATGAAGAAAACAAGTCAGC
	Oligo anti-PA 1121-1157	TAAAGTGGGCACTTGGTGAGAACATGGCACCAGAAAA
	Oligo anti-PA 1158-1195	GGTAGACTTTGACGACTGTAAGATGTAGGTGATTTGA
	Oligo anti-PA 1196-1233	AGCAATATGATAGTGATGAACCAGAATTGAGGTGCCTT
	Oligo anti-PA 1234-1272	GCAAGTTGATTGAGAATGAGTTTAAACAAGGCATGCGAA
	Oligo anti-PA 1034-1058	CATGGAAGCAAGTACTGGCAGAACT
	Oligo anti-PA 1059-1087	GCAGGACATTGAGAATGAGGAGAAAATTC
	Oligo anti-NP 358-383	ATGAGAGAACTCATCCTTTATGACAA
	Oligo anti-NP 581-600	CAGGTGCTGCAGTCAAAGGA
	Oligo anti-NP 710-732	TGTGCAACATTCTCAAAGGGAAA
	Oligo anti NP 857-877	CTCACAAGTCTGCCTGCCTG
Oligo anti-NP 1040-1060	CATGCCATTCTGCCGATTTG	
<b>PB1/PA</b>	Oligo anti-PB1 1674-1709	GTTTCATCAAAGATTACAGGTACACGTACCGATGCCA
	Oligo anti-PB1 1710-1745	TATAGGTGACACACAAAATACAAACCCGAAGATCATT
	Oligo anti-PB1 1746-1781	TGAAATAAAGAACTGTGGGAGCAAACCCGTTCCAA
	Oligo anti-PB1 1782-1817	AGCTGGACTGCTGGTCTCCGACGGAGGCCAAATTT
	Oligo anti-PB1 1818-1853	ATACAACATTAGAAATCTCCACATTCTGAAGTCTG

PB1/PA	Oligo anti-PB1 1854-1889	CCTAAAATGGGAATTGATGGATGAGGATTACCAGGG
	Oligo anti-PB1 1890-1925	GCGTTTATGCAACCCACTGAACCCATTTGTCAGCCA
	Oligo anti-PB1 1911-1940	CCCATTTGTCAGCCATAAAGAAATTGAATC
	Oligo anti-PB1 1838-1866	ACATTCCTGAAGTCTGCCTAAAATGGGAA
	Oligo anti-PA Anti 445-473	GAGAAAACACACATCCACATTTTCTCGTT
	Oligo anti-PA Anti 475-491	ACTGGGGAAGAAATGGCCACAA
	Oligo anti-PA Anti 490-514	GCCACAAAGGCAGACTACACTCTCG
	Oligo anti-PA Anti 494-524	CAAAGGCAGACTACACTCTCGATGAAGAAAG
	Oligo anti-PA Anti 506-532	ACACTCTCGATGAAGAAAGCAGGGCTA
	Oligo anti-PA Anti 533-558	GGATCAAACAGACTATTACCATA
	Oligo anti-PA Anti 555-578	CATAAGACAAGAAATGGCCAGCAG
	Oligo anti-PA Anti 635-664	GGTTTGAATCACAGGAACAATGCGTAAGC
	Oligo anti-PA Anti 684-706	GCCGAACCTTCTCCAGCCTTGAAA

**Tableau 6 : Séquences des oligodésoxyribonucléotides antisens utilisés pour caractériser les interactions entre les ARNv de la souche PR8.**

Afin de vérifier l'hybridation spécifique des oligodésoxyribonucléotides sur les ARNv, 10 pmol d'oligodésoxyribonucléotides sont marqués radioactivement à l'extrémité 5'-OH en présence de 2,5 pmol de [ $\gamma$ -<sup>32</sup>P]ATP à une activité spécifique de 6 000 Ci/mmol et de 10 U de polynucléotide kinase du bactériophage T4 (Fermentas), dans le tampon réactionnel A (Fermentas) concentré 1x, pendant 1h à 37°C. Les oligodésoxyribonucléotides sont ensuite purifiés en utilisant des colonnes chromatographiques Micro Bio-Spin® 6 (Bio-Rad) selon les recommandations du fabricant. Puis, 20 000 cpm de chaque oligodésoxyribonucléotide, dosé sur un compteur « LS6500 multipurpose scintillation » en mode Cerenkov, sont incubés avec les ARNv, toujours selon le protocole de co-incubation décrit précédemment. Après migration, le gel d'agarose est fixé pendant 10 min dans une solution d'acide trichloroacétique 10% (Roth), puis séché sous vide pendant 1h à température ambiante. Une « Imaging Plate » Fujifilm (lue par un FLA-5000 de Fujifilm) est ensuite utilisée pour révéler les bandes radioactives.

### C. Cartographie chimique des ARNv

À l'aide de la stratégie *in vitro* décrite précédemment (utilisation d'ARNv mutés et d'oligodésoxyribonucléotides antisens), l'interaction entre les ARNv PB1 et M a été déterminée au nt près. Dans le but de vérifier que les nt identifiés sur le segment PB1 sont effectivement impliqués dans l'hétérodimérisation, des expériences de cartographie chimique sont réalisées sur l'ARNv PB1 seul ou en présence de son partenaire M.

## **I. Modifications chimiques des ARNv au DMS**

Dans chaque expérience, trois conditions sont testées : i) l'ARNv PB1 seul, ii) les ARNv PB1 et M en quantités équimolaires (2 pmoles de chaque segment) et iii) un excès molaire de l'ARNv M de trois fois (2 pmoles du segment PB1 et 6 pmoles du segment M).

Après co-incubation des ARNv comme décrit précédemment (§ B.IV), ils sont modifiés avec 1 µl de diméthyl sulfate (DMS) dilué au 450<sup>ème</sup> (concentration finale 0,023 M) dans de l'éthanol absolu pendant 1, 3 ou 5 min à 55°C. Une réaction contrôle est réalisée en présence d'éthanol absolu. Les réactions sont arrêtées avec 3 vol d'éthanol absolu et 0,1 vol d'AcNa 3 M (pH 5,6). Après avoir ajouté 1 µg de glycogène dans chaque tube et des quantités variables d'ARNt total de levure pour équilibrer les quantités d'ARN dans toutes les conditions testées, les ARN sont mis à précipiter dans un bain carboglace-éthanol pendant 1h. Après centrifugation (30 min à 15700 g et à 4°C), les culots sont lavés deux fois avec de l'éthanol 80%, séchés 5 min sous vide au SpeedVac et repris dans 7 µl d'eau milliQ (Millipore).

## **II. Détection des modifications par élongation d'amorce par la transcriptase inverse**

### **1. Marquage radioactif à l'extrémité 5' des oligodésoxyribonucléotides**

L'amorce utilisée, complémentaire aux positions 54-78 de l'ARNv PB1, obtenue par synthèse chimique (Integrated DNA Technologies), est marquée à l'extrémité 5'-OH par transfert du phosphate γ radioactif du  $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ . L'oligodésoxyribonucléotide (100 pmoles) est incubé pendant 30 min à 37°C dans le tampon kinase 1x, en présence de 100 µCi de  $[\gamma\text{-}^{32}\text{P}]\text{ATP}$  (3000 Ci/mme) et de 15 U de polynucléotide kinase du bactériophage T4 (Fermentas), dans un volume final de 20 µl. La purification se fait ensuite par élution au travers d'une colonne chromatographique Micro Bio-Spin® 6 (Bio-Rad) selon les recommandations du fabricant. Les oligonucléotides marqués sont ensuite comptés sur un compteur « LS6500 multipurpose scintillation » en mode Cerenkov.

### **2. Extension d'amorce par la transcriptase inverse**

Après avoir ajouté 400 000 cpm d'amorce radioactive à l'ARNv modifié repris dans l'eau, le tout est incubé 2 min à 90°C puis refroidit rapidement à -80°C. Après ajout de 2 µl

de tampon RTB 5x (250 mM Tris-HCl (pH 8 à 37°C), 200 mM KCl, 30 mM MgCl<sub>2</sub>) et décongélation, l'élongation se fait en présence de 2 U de transcriptase inverse (RT) du virus du myéloblastome aviaire (AMV) (MP Biomedicals) et d'un mélange équimolaire de chacun des quatre dNTP (2,5 mM), dans 25 mM Tris HCl (pH 8,3), 50 mM KCl, 2 mM DTT et 5 mM MgCl<sub>2</sub>, pendant 30 min à 42°C. Les fragments d'ADNc ainsi synthétisés sont précipités par 500 µl d'éthanol et 200 µl d'AcNa 0,3 M (pH 5,6), centrifugés (30 min à 15700 g et à 4°C), lavés deux fois avec de l'éthanol 80% et séchés 5 min sous vide au SpeedVac. Enfin, les culots sont comptés au compteur à scintillation en mode Cerenkov et repris dans un volume de bleu-formamide (formamide 95% (v/v), xylène-cyanol 0,025% (p/v), bleu de bromophénol 0,025% (p/v)) ajusté à chaque échantillon de manière à déposer sur le gel la même quantité de radioactivité dans le même volume.

### **3. Séquençage par la transcriptase inverse (méthode de Sanger)**

L'hybridation de l'amorce radioactive est réalisée comme précédemment (§ C.II.2). L'élongation est ensuite effectuée en présence de 2 U de RT d'AMV, et de l'un des quatre ddNTP pendant 30 min à 42°C. Afin que l'incorporation soit statistique, la concentration de chaque ddNTP est de 2,5 µM et celle du dNTP correspondant à la même base est de 25 µM, tandis que les autres dNTP sont à 100 µM. Les échantillons sont précipités par 400 µl d'éthanol et 100 µl d'AcNa 0,3 M (pH 5,6), puis traités comme décrit ci-dessus (§ C.II.2).

### **4. Analyse par électrophorèse sur gel de polyacrylamide dénaturant**

Les échantillons, repris dans du bleu-formamide, sont dénaturés 2 min à 90°C, puis déposés sur un gel bisacrylamide/acrylamide (19/1) 8%/Urée 8 M/TBE 1x, le gel ayant subi une pré-électrophorèse de 30 min à 75 W. Après 1h30 de migration à 75 W, le gel est séché sous vide à 80°C et une « Imaging Plate » Fujifilm (lue par un FLA-5000 de Fujifilm) est utilisée pour révéler les bandes radioactives. Enfin, l'intensité de modification des nt de l'ARNv PB1, en absence ou en présence de l'ARNv M est quantifiée par le logiciel Image Gauge (Fujifilm) et l'intensité de modification des nt de l'ARNv PB1 seul est ramenée à 100%.

## D. Identification des nucléotides importants pour l'établissement des interactions entre ARNv par MIME

Nom de l'amorce	Séquence (de 5' vers 3')	Étape lors de laquelle l'amorce est utilisée
Amorce NP 1-16	AGCAAAAGCAGGGTAG	PCR mutagène et PCR (pour l'amplification)
Amorce NP-T7 1554-1565	TAATACGACTCACTATAGGGAGTAGAAACAAGG	
Amorce NP LNA 1-16	A{G}CA{A}AA{G}CAGGGTAG {} = LNA	Transcription inverse
Amorce NP 703-730	GAAAGAATGTGCAACATTCTCAAAGGGA	Transcription inverse, qPCR et PCR
Amorce NP 853-873	CAGGCAGGACTTGTGAGCAAC	qPCR et PCR
Amorce NP 1-28	AGCAAAAGCAGGGTAGATAATCACTCAC	qPCR et PCR
Amorce NP 1530-1565	AGTAGAAACAAGGGTATTTTTCTTAATTGTCGTAC	qPCR et PCR
IlluminaMAs	GATCGGAAGAGCACACGTCT	Ligation des adaptateurs
IlluminaMAa	ACACTCTTCCCTACACGACGCTCTCCGATCT	
Illumina PE 1.0	AATGATACGGCGACCACCGAGATCTACTACTCTTCCCTACACGACGC TCTCCGATCT	Conversion des adaptateurs, qPCR et PCR finale
Index Illumina	CAAGCAGAAGACGGCATACGAGATNNNNNNGTGACTGGAGTTCA GACGTGTGCTCTCCGATCT	

Tableau 7 : Amorces utilisées dans les expériences de MIME. LNA : *locked nucleic acid*.

## I. Préparation de la banque d'ARNv NP

### 1. Mutagenèse aléatoire de l'ADN NP

Pour obtenir la banque d'ADN NP, où chaque position nucléotidique est mutée aléatoirement, deux étapes de PCR mutagènes successives sont réalisées à l'aide du kit GeneMorph® II Random Mutagenesis (Agilent Technologies). D'après le fournisseur, la mutazyme II® contenue dans ce kit introduit des mutations de manière moins biaisée que les protocoles standards de mutagenèse aléatoire, les quatre nucléotides étant mutés de manière équivalente.

Dans un premier temps, pour obtenir la *banque 1*, 50 ng de plasmide pUC2000 NP sont utilisés comme matrice en présence de 0,25 pmol des amorces NP 1-16 et NP-T7 1554-1565 (qui permet l'ajout du promoteur T7) (Integrated DNA Technologies) (tableau 7), du tampon mutazyme concentré 1x, de 0,2 mM d'un mélange équimolaire de chacun des quatre dNTP et de 0,5 U de mutazyme II® (Agilent Technologies). La réaction est préparée dans 25 µl, puis le mélange est réparti dans deux tubes de PCR. Après une étape de dénaturation de 2 min à 95°C, 35 cycles d'amplification sont réalisés (15 sec de dénaturation à 95°C, 30 sec d'hybridation des amorces à 50°C et 2 min d'élongation à 72°C), suivis par une élongation finale de 10 min à 72°C. Après avoir déposé le produit de PCR sur un gel d'agarose low melting 0,8% (p/v) (Fisher BioReagents®)/TBE 0,5x (Euromedex) (0,4 µg/ml de

BET (Roth)), l'ADN est purifié à l'aide du kit Nucleospin® Gel and PCR clean-up (Macherey-Nagel) d'après les recommandations du fournisseur. Cette première banque d'ADN NP sert ensuite de matrice pour la réalisation de deux réactions de PCR mutagènes, réalisées en parallèle, afin d'obtenir les *banques 2a* et *2b*. Après avoir déposé les produits de PCR sur un gel d'agarose low melting 0,8% (p/v) (Fisher BioReagents®)/TBE 0,5x (Euromedex) (0,4 µg/ml de BET (Roth)), ces deux banques sont purifiées à l'aide du kit Nucleospin® Gel and PCR clean-up (Macherey-Nagel). Le taux de mutations introduit dans les banques d'ADN NP *2a* et *2b* est vérifié après clonage dans le vecteur pJET (CloneJET™ PCR cloning kit, ThermoScientific), selon le protocole du fournisseur, et séquençage de 27 clones par la méthode de Sanger (GATC Biotech).

Par la suite, les *banques 2a* et *2b* sont amplifiées par PCR, toujours séparément, à partir de 50 ng d'ADN matrice, en présence de 0,5 pmol de chaque amorce (NP 1-16 et NP-T7 1554-1565), de tampon taq concentré 1x, de 0,8 mM d'un mélange équimolaire de chacun des quatre dNTP et de 2,5 U de taq polymérase (produite et purifiée au laboratoire). La réaction est préparée dans 100 µl, puis le mélange est réparti dans quatre tubes. Après une étape de dénaturation de 4 min à 94°C, 25 cycles d'amplification sont réalisés (1 min à 94°C, 1 min à 50°C et 2 min à 72°C), suivis d'une élongation finale de 10 min à 72°C. Après avoir déposé les produits de PCR sur un gel d'agarose low melting 0,8% (p/v) (Fisher BioReagents®)/TBE 0,5x (Euromedex) (0,4 µg/ml de BET (Roth)) et purifié l'ADN à l'aide du kit Nucleospin® Gel and PCR clean-up (Macherey-Nagel), les deux banques sont rassemblées pour constituer la banque d'ADN NP finale sur laquelle la suite du protocole est réalisée.

## **2. Transcription *in vitro* de la banque d'ARNv NP et purification sur gel**

La transcription *in vitro* à partir de 200 ng de produit de PCR purifié est réalisée à l'aide du kit MEGAscript® T7 (Ambion) en présence de 7,5 mM de chacun des quatre NTP, du tampon commercial 1x et de 2 µl du mélange d'enzyme dans un volume final de 20 µl. Le mélange est incubé 4h à 37°C avant de rajouter 2 U de TURBO DNase et de placer le tout 15 min à 37°C. Après une extraction au phénol-chloroforme (v/v), une précipitation en présence de 0,1 vol d'AcNa 3 M (pH 5,6) et 2,5 vol d'éthanol absolu (30 min dans un bain carboglace-éthanol) et une centrifugation (30 min à 15700 g et à 4°C), l'ARNv est lavé deux fois à l'EtOH 80% avant de subir un second traitement DNase. Enfin, le culot est repris dans de l'eau

milliQ (Millipore) avant d'être déposé sur un gel d'agarose low melting 0,8% (p/v) (Fisher BioReagents®)/TBE 0,5x (Euromedex) (0,4 µg/ml de BET (Roth)). La purification de l'ARNv se fait à nouveau à l'aide du kit Nucleospin® Gel and PCR clean-up (Macherey-Nagel), mais en utilisant le tampon de solubilisation NTC (spécifique des acides nucléiques sb). L'échantillon est ensuite traité comme précédemment (extraction au phénol-chloroforme, précipitation, centrifugation et lavages du culot). Pour finir, le culot est séché à l'air libre, puis repris dans de l'eau milliQ (Millipore).

## **II. Séparation des ARNv de la banque par retard de migration sur gel**

Les ARNv sauvages PB2, PB1, PA et NP sont transcrits et purifiés comme indiqué précédemment (§ B.II) Deux pmoles de chaque ARNv sauvage ou de l'ARNv NP transcrit à partir de la banque d'ADN sont utilisées pour les expériences de co-incubation et retard de migration sur gel (voir § B.IV). À l'issue de la migration des échantillons sur un gel d'agarose low melting 0,8% (p/v) (Fisher BioReagents®)/TBM 0,5x (0,4 µg/ml de BET (Roth)), chaque ARNv correspondant à l'ARNv NP, soit non complexé, soit complexé à ses partenaires, est découpé du gel, puis purifié à l'aide du kit Nucleospin® Gel and PCR clean-up (Macherey-Nagel) en utilisant le tampon de solubilisation NTC. Après une extraction au phénol-chloroforme (v/v) et une précipitation en présence de 0,1 vol d'AcNa 3 M (pH 5,6) et 2,5 vol d'éthanol absolu, les échantillons sont centrifugés (30 min à 15700 g et à 4°C). Les ARNv sont ensuite lavés deux fois à l'EtOH 80% et séchés à l'air libre. Pour finir, les culots sont repris dans de l'eau milliQ (Millipore).

## **III. Préparation de la banque en vue d'un séquençage Illumina**

### **1. Transcription inverse et quantification par qPCR**

Afin de ne pas perdre d'information, notamment à l'extrémité 5', la totalité des 1565 nt de l'ARNv NP doit être rétro-transcrite. Pour cela, chaque échantillon est rétro-transcrit avec deux amorces : NP LNA 1-16 (Eurogentec) et NP 703-730 (Integrated DNA Technologies) (tableau 7). Pour plus de clarté, à partir de cette étape, les ADNc obtenus avec ces deux amorces sont nommés « segments U12 » et « segments U13 » et correspondent respectivement aux « moitiés » 3' et 5' de l'ARNv. Les ARNv sont chauffés 2 min à 90°C après ajout de 10 pmol de l'une des deux amorces, puis incubés 2 min sur la glace. La réaction de

transcription inverse est ensuite réalisée dans un volume final de 40 µl en présence de tampon RT AMV 1x, de 25 mM de chacun des quatre dNTP, d'1 U de RT AMV (MP Biomedicals) et de 4 U de RNasin® Rnase inhibitor, pendant 20 min à 42°C, 30 min à 50°C, puis 10 min à 60°C. Les ADNc sb sont ensuite purifiés à l'aide du kit Nucleospin® Gel and PCR clean-up (Macherey-Nagel) en utilisant le tampon de solubilisation NTC. La quantification par qPCR des ADNc sb U12 et U13 est réalisée sur un aliquot dilué de chaque produit de rétro-transcription en présence de 0,8 mM de chacun des quatre dNTP, de tampon HF concentré 1x, de 0,15 U de Phusion, de 0,625 µM de réactif Eva Green (Biotium) et des amorces NP 1-28 et NP 853-873 (350 nM) ou des amorces NP 703-730 et NP 1530-1565 (500 nM), respectivement (tableau 7). Après une dénaturation initiale de 4 min à 94°C, 40 cycles sont réalisés (1 min de dénaturation à 94°C, 1 min d'hybridation des amorces à 55°C, 1 min d'élongation à 72°C) avec l'appareil CFX Touch™ Real-Time PCR Detection System (Biorad). Les résultats sont ensuite analysés avec le logiciel CFX Manager™ (Biorad) et permettent de normaliser les échantillons, de sorte à obtenir la même quantité d'ADNc db U12 et U13 après l'amplification par PCR (ci-dessous).

## **2. Amplification de l'ADNc db par PCR**

Une fraction non diluée de chaque ADNc sb est amplifiée par PCR afin d'obtenir des ADNc db. Les conditions sont les mêmes que celles de la qPCR (§ D.III.1), hormis le réactif Eva Green. Après avoir effectué 13 cycles d'amplification pour les échantillons U12 et 14 cycles pour les échantillons U13, une étape d'élongation finale de 10 min à 72°C termine le programme. Les produits de PCR db U12 et U13 sont ensuite déposés sur un gel d'agarose low melting 0,8% (p/v) (Fisher BioReagents®)/TBE 0,5x (Euromedex) (0,4 µg/ml de BET, Roth), puis purifiés à l'aide du kit Nucleospin® Gel and PCR clean-up (Macherey-Nagel). Les échantillons ainsi purifiés sont ensuite précipités en présence de 0,1 vol d'AcNa 3 M (pH 5,6) et 3 vol d'éthanol absolu pendant 30 min dans un bain carboglace-éthanol. Après centrifugation (30 min à 15700 g et à 4°C), les culots sont lavés deux fois à l'EtOH 80%, séchés 5 min sous vide au Speedvac, puis repris dans de l'eau milliQ (Millipore).

### **3. Fragmentation aléatoire de l'ADNc**

Les ADNc db U12 et U13 sont incubés 5 min sur la glace en présence de tampon fragmentase 1x (NEB) et de 0,1 mg/ml de BSA (NEB), avant de rajouter 1 µl de NEBNext dsDNA Fragmentase (NEB), dans un volume total de 25 µl. Les échantillons sont incubés 35 min à 37°C. Ils sont ensuite déposés sur un gel d'agarose low melting 0,8% (p/v) (Fisher BioReagents®)/TBE 0,5x (Euromedex) (0,4 µg/ml de BET, Roth) et les fragments d'ADN dont la taille est comprise entre 100 et 500 pb sont purifiés à l'aide du kit Nucleospin® Gel and PCR clean-up (Macherey-Nagel).

### **4. « End polishing » : réparation des extrémités des ADNc**

L'ADN est réparé en présence du tampon T4 ADN ligase 1x (NEB), de 0,2 mM de chacun des quatre dNTP, d'1 mM d'ATP, de 0,5 µl d'ADN ligase d'*E. coli* (du kit NEBNext dsDNA fragmentase), de 4,5 U de T4 ADN polymérase (NEB) et de 25 U de T4 polynucléotide kinase (NEB) dans un volume final de 50 µl, pendant 1h à 20°C. Les enzymes sont ensuite inactivées pendant 30 min à 75°C.

### **5. Ajout d'une adénosine aux extrémités des ADNc**

Une adénosine est ensuite ajoutée aux extrémités des fragments d'ADNc en incubant le milieu réactionnel précédent (§ D.III.4) pendant 1h à 37°C en présence de 5 U de polymérase de Klenow (dépourvue d'activité exonucléase 3'-5') et de 1 µl de dATP 100 mM. À nouveau, les enzymes sont inactivées pendant 30 min à 75°C.

### **6. Ligation des adaptateurs aux extrémités des ADNc**

Dans un premier temps, les deux adaptateurs Illumina (IlluminaMAs et IlluminaMAa) (tableau 7) sont mélangés dans le tampon ligase 1x (NEB), chauffés à 95°C pendant 1 min, puis refroidis lentement jusqu'à température ambiante, afin de les pré-hybrider. Ensuite, 1 µl de ces adaptateurs (10 µM) est ajouté au milieu précédent (§ D.III.5), ainsi que 9 µl de tampon T4 ADN ligase 10x (Life Technologies), 28 µl de PEG 4000 (NEB) et 1 µl de T4 ADN ligase (Life Technologies). Après une incubation de 2h à 20°C, les adaptateurs non ligués sont éliminés en purifiant les échantillons à l'aide du kit Nucleospin® Gel and PCR clean-up (Macherey-Nagel) en utilisant le tampon NTI dilué au ½.

## **7. Conversion des adaptateurs en ADN double brin et quantification par qPCR**

Les adaptateurs Illumina, ligués aux échantillons, sont convertis en ADN db par PCR (30 sec à 98°C suivis de 5 cycles : 15 sec à 98°C, 30 sec à 63°C et 1 min à 72°C) avec 10 pmol d'amorce Illumina PE 1.0 et 10 pmol des index Illumina (tableau 7 ; NNNNNN correspond à l'une des séquences des index pour le multiplexing) et en présence du tampon HF 1x, de 0,3 U de Phusion et de 0,2 mM de dNTP. Dans le but d'éliminer les adaptateurs non utilisés, les échantillons sont déposés sur un gel d'agarose low melting 0,8% (p/v) (Fisher BioReagents®)/TBE 0,5x (Euromedex) (0,4 µg/ml de BET (Roth)), puis les fragments dont la taille est comprise entre 300 et 600 pb sont purifiés avec le kit Nucleospin® Gel and PCR clean-up (Macherey-Nagel). Les deux adaptateurs ligués et amplifiés faisant une taille totale de 109 pb, ces fragments correspondent à des échantillons d'environ 200 à 500 pb. Les banques obtenues sont ensuite quantifiées par qPCR à l'aide des amorces Illumina PE PCR 1.0 et d'un mélange des 30 index avec le SYBR Green qPCR Master Mix (Life technologies), selon le protocole du fournisseur, avec l'appareil CFX Touch™ Real-Time PCR Detection System (Biorad). Quarante cycles d'amplification sont effectués (3 min à 95°C, 3 sec à 95°C, 30 sec à 60°C).

## **8. Amplification finale de la banque**

Les échantillons sont ensuite normalisés à partir des résultats de la qPCR précédente (§ D.III.7), puis amplifiés avec les amorces Illumina PE PCR 1.0 et chacun des index Illumina dans le même milieu réactionnel que précédemment (§ D.III.7). Après 30 sec de dénaturation à 98°C, 19 cycles sont effectués pour les échantillons provenant de la fragmentation de U12 et 22 cycles pour les échantillons provenant de la fragmentation de U13, dont chacun comporte 15 sec de dénaturation à 98°C, 30 sec d'hybridation des amorces à 63°C et 30 sec d'élongation à 72°C. Les échantillons U12 et U13 sont ensuite rassemblés, puis une dernière sélection des fragments d'ADN de 300 à 600 pb est effectuée sur un gel d'agarose low melting ultrapur 0,8% (p/v) (MetaPhor™ Agarose, Lonza) /TBE 0,5x (Euromedex) (0,4 µg/ml de BET (Roth)) afin d'éliminer les adaptateurs non ligués.

#### **IV. Séquençage de la banque et analyse des données**

La banque ainsi obtenue est séquençée dans une piste de HiSeq2500 en mode « paired-end 100 » à la plateforme « Biopuces et séquençage » à l'IGBMC à Illkirch. Les données de séquençage sont ensuite analysées en collaboration avec le Dr R. Smyth (IBMC, UPR 9002) à l'aide des outils bioinformatiques décrits dans Smyth *et al.* [277] et du logiciel MIMeAnTo [318].

#### **E. Production de particules virales d'*Influenzavirus A* et caractérisation**

##### **I. Entretien des lignées cellulaires**

Dans ce projet deux lignées cellulaires ont été utilisées : les cellules humaines HEK293T/17 (*Human Embryonic Kidney 293T* ; ATCC® CRL-11268TM) et les cellules canines MDCK (*Madin-Darby Canine Kidney* ; ATCC® CRL-2936TM).

Les cellules HEK293T/17 sont cultivées à 37°C en présence de 5% de CO<sub>2</sub> en flasques T75 dans du milieu DMEM (*Dulbecco's modified Eagle medium*) contenant 1 g/l de glucose (Lonza) supplémenté avec 2 mM de L-glutamine dans 0,85% de NaCl (Lonza), 10% (v/v) de sérum de veau fœtal (Gibco) et en présence de 100 U/ml de pénicilline/streptomycine (Lonza) (défini comme le milieu d'entretien). Le passage s'effectue tous les trois à quatre jours : après avoir lavé le tapis cellulaire avec du PBS (*phosphate-buffered saline*, sans Ca<sup>2+</sup> et Mg<sup>2+</sup>, Lonza), les cellules sont décollées par 0,25% de trypsine (sans Ca<sup>2+</sup> et Mg<sup>2+</sup>, Lonza) et diluées dans du PBS avant d'être remises en suspension dans du milieu DMEM supplémenté. Le comptage à l'aide d'une cellule de Malassez permet d'ensemencer une nouvelle flasque T75 avec 2,5 millions de cellules dans 30 ml de milieu DMEM supplémenté. Dans le but de préparer des cellules HEK293T/17 destinées à être transfectées, des boîtes de Pétri BD BioCoat™ Collagène I de 60 mm de diamètre (BD Biosciences, surface de 21,3 cm<sup>2</sup>) sont ensemencées 24h avant transfection avec 0,8 million de cellules dans 3 ml de milieu d'entretien.

Les cellules MDCK sont cultivées à 37°C en présence de 5% de CO<sub>2</sub> en flasques T175 dans du milieu UltraMDCK sans sérum (Lonza) supplémenté avec 2 mM de L-Glutamine dans 0,85% de NaCl (Lonza) et 100 U/ml de pénicilline/streptomycine (Lonza). Le passage

s'effectue tous les trois à quatre jours : le tapis cellulaire est lavé avec du HBSS (*Hanks balanced salt solution*, sans rouge de phénol, Ca<sup>2+</sup> et Mg<sup>2+</sup>, Lonza), puis 50 ml de HBSS sont ajoutés et la flasque est incubée pendant 1h30 à 37°C (5% de CO<sub>2</sub>). Lorsque les cellules se sont toutes arrondies, le HBSS est enlevé et remplacé par 2 ml d'une solution contenant 0,025% (v/v) de trypsine dans du HBSS (Lonza) et 0,0198% (v/v) de versene (Lonza), avant d'incuber la flasque durant 10 min à 37°C (5% de CO<sub>2</sub>). Lorsque les cellules sont décollées, 18 ml de milieu UltraMDCK sont ajoutés. La numération à l'aide d'une cellule de Malassez permet ensuite d'ensemencer une nouvelle flasque T175 avec 3 millions de cellules dans 75 mL de milieu d'entretien. Différents supports, dans le but d'amplifier ou de caractériser les virus, sont également ensemencés avec 1,3 millions de cellules dans 24 ml par plaque de 6, 12, 24, 48 ou 96 puits, ou des flasques T25 avec 1 million de cellules dans 10 ml de milieu d'entretien.

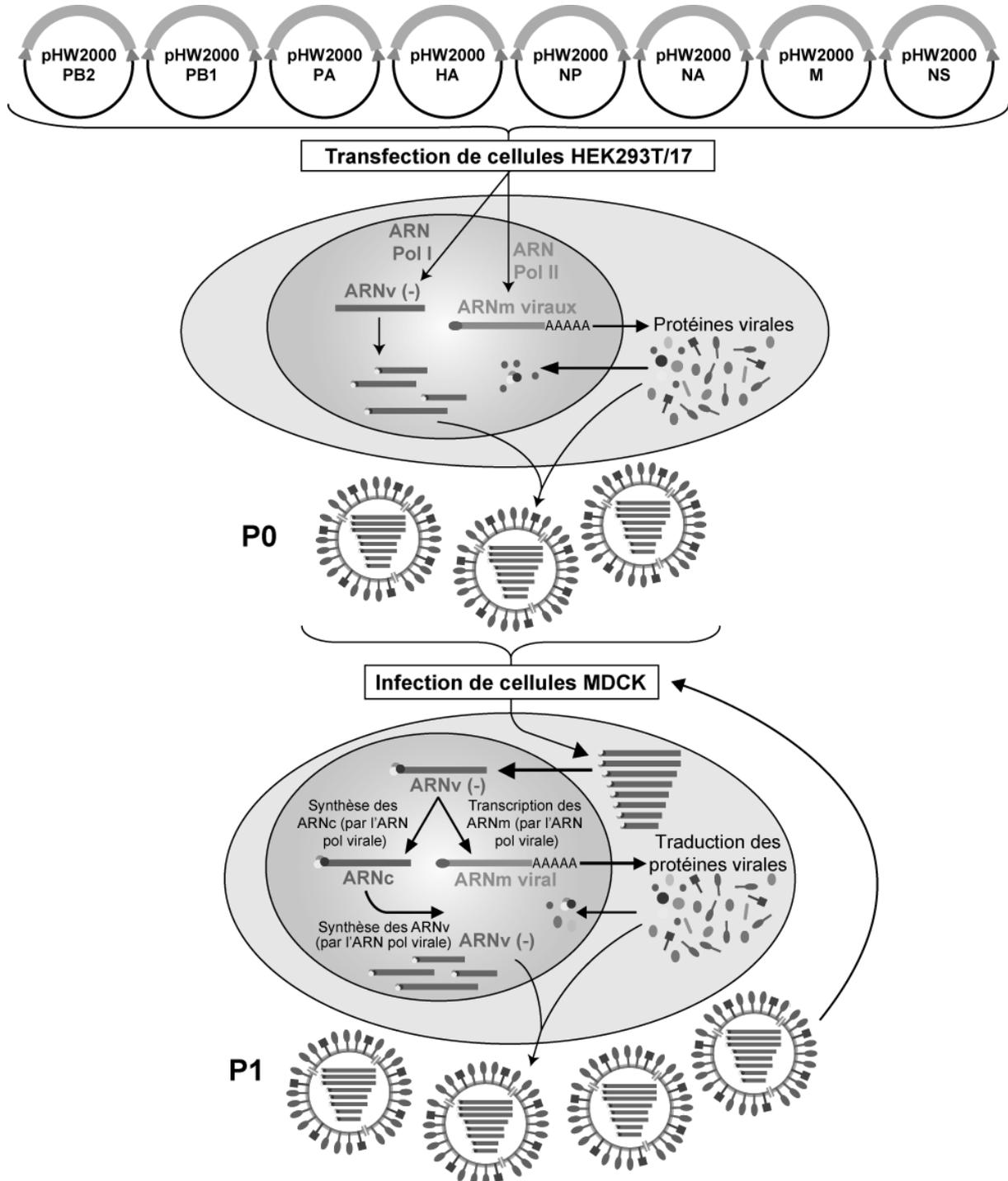
## **II. Production de virus influenza A**

### **1. Transfection de cellules HEK293T/T17**

Le système de génétique inverse utilisé [253] fait intervenir huit plasmides d'expression bidirectionnels (figure 22). Chaque plasmide contient la séquence d'ADNc d'un segment viral cloné entre le promoteur Poll humain et le promoteur PolIII du CMV. Dans les expériences de génétique inverse compétitive, un neuvième plasmide est transfecté, correspondant à un segment muté qui va entrer en compétition avec le segment sauvage pour la sélection et l'empaquetage dans les virions.

Les cellules HEK293T/17 ont été ensemencées 24h avant la transfection dans des boîtes de Pétri BD BioCoat™ Collagène I de 60 mm de diamètre (BD Biosciences) et sont à 40-50% de confluence. Dans un premiers temps, les plasmides à transférer sont mélangés (1 µg de chaque), puis 84 µl d'OptiMEM (Life Technologies) et 2 µl de Trans-IL LT1 (Mirus) par µg de plasmide sont ajoutés et le tout est mélangé au vortex. Après une incubation de 20 min à température ambiante, 884 µl d'OptiMEM sont ajoutés. Le tapis cellulaire est ensuite rincé avec 1 ml de PBS (sans Ca<sup>2+</sup> et Mg<sup>2+</sup>, Lonza), puis le mélange de transfection est distribué au goutte-à-goutte sur les cellules. Après une incubation de 4h à 37°C (5% de CO<sub>2</sub>), le milieu de transfection est remplacé par 1 ml d'OptiMEM supplémenté avec 0,3% de BSA (Roche). Vingt-quatre heures post-transfection, 1 ml d'OptiMEM supplémenté avec 100

U/ml de pénicilline/streptomycine (Lonza) et 1 µg/ml de trypsine acétylée issue de pancréas bovin dans du PBS 1x (Sigma-Aldrich) est ajouté. Le surnageant contenant les particules virales est récolté 48h à 72h après la transfection et centrifugé 1 min à 4000 g à température ambiante. Cette fraction appelée P0 est ensuite utilisée pour infecter des cellules MDCK.



**Figure 22 : Description du système de génétique inverse utilisé pour la production des IVA.** Les séquences des ADNc de chaque segment viral sont clonées dans un vecteur d'expression bidirectionnel pHW2000 (voir § A.II.2). Les particules virales obtenues après la co-transfection de cellules humaines HEK293T/17 avec les huit plasmides sont ensuite amplifiées dans des cellules canines MDCK. Le stock viral est constitué après plusieurs cycles d'infection.

## 2. Amplification des virus par infection de cellules MDCK

Dans le but d'amplifier les virus obtenus par génétique inverse, les fractions P0 sont utilisées pour infecter des cellules MDCK confluentes (figure 22), généralement ensemencées dans des plaques 24 puits (BD Biosciences, surface de 1,9 cm<sup>2</sup> par puits) dans du milieu UltraMDCK (Lonza). À partir de cette étape, les cellules sont cultivées dans un milieu d'infection contenant du EMEM (*Eagle's minimum essential medium*, Lonza) supplémenté avec 2 mM de L-Glutamine dans 0,85% de NaCl (Lonza), 100 U/ml de pénicilline/streptomycine (Lonza) et 1 µg/ml de trypsine acétylée issue de pancréas bovin dans du PBS 1x (Sigma-Aldrich). Les cellules sont lavées, puis 350 µl de surnageant dilué dans du milieu de culture (généralement au 1/4, 1/10 et 1/100) sont utilisés pour infecter les cellules. Après 1h d'incubation à 37°C sous agitation douce, 600 µl de milieu d'infection sont ajoutés, puis les cellules sont incubées à 37°C (5% de CO<sub>2</sub>).

Aux alentours de 48h après l'infection, les cellules MDCK s'arrondissent et se détachent du tapis cellulaire : ce sont des effets cytopathogènes (ECP), caractéristiques de l'infection par des IVA. Le délai d'apparition des ECP est variable selon les souches étudiées. Des débuts d'ECP sont généralement visibles 24h post-infection pour la souche PR8 et 48h pour la souche Moscou. La fraction P1, résultant d'un premier passage sur les cellules MDCK, permet de vérifier la présence de virus dans la fraction P0 provenant de la co-transfection de cellules HEK293T/T17. Après apparition des ECP, le surnageant peut être récolté pour être congelé à -80°C ou être dilué (dilutions en cascade de 10<sup>-3</sup> à 10<sup>-7</sup> par exemple) pour effectuer un second cycle d'infection de cellules MDCK, dans des plaques 12 puits généralement (BD Biosciences, surface de 3,9 cm<sup>2</sup> par puits). Dans ce cas, le protocole est le même que précédemment, mais 450 µl de surnageant dilué dans du milieu de culture sont utilisés pour l'infection, puis après 1h de contact, le milieu d'infection est remplacé par 2 ml de milieu de culture. Après 72h, le surnageant est récolté, centrifugé 1 min à 4000 g à température ambiante, aliquoté, puis congelé à -80°C, pour constituer le stock de virus.

### III. Détermination des caractéristiques des virus produits

#### 1. Détermination du titre viral

Le titre infectieux d'un virus peut être exprimé en PFU (*plaque forming unit*) ou en DICT<sub>50</sub> (ou TCID<sub>50</sub>, *tissue culture infective dose 50*). La première méthode permet de déterminer le nombre de particules virales capables de former des plages de lyse, tandis que la seconde est une méthode statistique qui évalue la quantité de particules virales infectieuses nécessaires pour induire un ECP dans 50 % des cellules inoculées.

Pour le titrage en PFU, des plaques 6 puits (BD Biosciences, surface de 9,6 cm<sup>2</sup> par puits) contenant des cellules MDCK à confluence sont lavées avec du milieu EMEM, puis 600 µl de milieu d'infection contenant le virus dilué (dilutions au 10<sup>ème</sup> de 10<sup>-4</sup> à 10<sup>-8</sup> par exemple) sont déposés dans chaque puits. Après 1h d'incubation à 37°C sous agitation douce, le surnageant est retiré et 1 ml de milieu semi-solide contenant 1,1% (v/v) de Difco Noble agar (Becton, Dickinson and Company) dans du milieu MEM (*minimum essential media*, Gibco) complémenté avec 100 U/ml de pénicilline/streptomycine et 1 µg/ml de trypsine acétylée issue de pancréas bovin dans du PBS 1x (Sigma-Aldrich) est ajouté dans chaque puits. Après solidification du milieu à température ambiante, les cellules sont incubées à 37°C (5% de CO<sub>2</sub>). Après quatre jours, les plages de lyse sont comptées dans chaque puits afin de calculer le titre infectieux.

$$\text{Titre infectieux (PFU/ml)} = \frac{\text{nombre de plages de lyse dans un puits}}{\text{dilution du virus à ce puits} \times \text{volume d'infection en ml}}$$

Le titrage en DICT<sub>50</sub> se fait en plaque 96 puits (BD Biosciences, surface de 34 mm<sup>2</sup> par puits) sur des cellules MCDK à confluence. Le milieu de culture est totalement remplacé par 150 µl de milieu EMEM, puis 50 µl de milieu d'infection contenant le virus dilué (dilutions au 10<sup>ème</sup> de 10<sup>-1</sup> à 10<sup>-9</sup> par exemple) sont déposés dans chaque puits. Généralement huit puits sont infectés pour une dilution donnée. Les cellules sont ensuite incubées quatre jours à 37°C (5 % de CO<sub>2</sub>) avant de vérifier la présence d'ECP. Pour chaque dilution, le pourcentage de puits avec ECP est défini, puis le titre infectieux est calculé selon la méthode statistique de Reed & Muench :

$$\text{Titre infection (DICT}_{50}/\text{ml)} = 10^{((-\log \text{ de la première dilution avec un pourcentage de puits positifs inférieur à } 50\% + \text{DP}) + 1,3)}$$

$$\text{Avec DP} = \frac{\text{valeur du pourcentage d'ECP inférieur à 50\%-50}}{\text{valeur du pourcentage d'ECP inférieur à 50\%-valeur du pourcentage d'ECP supérieur à 50\%}}$$

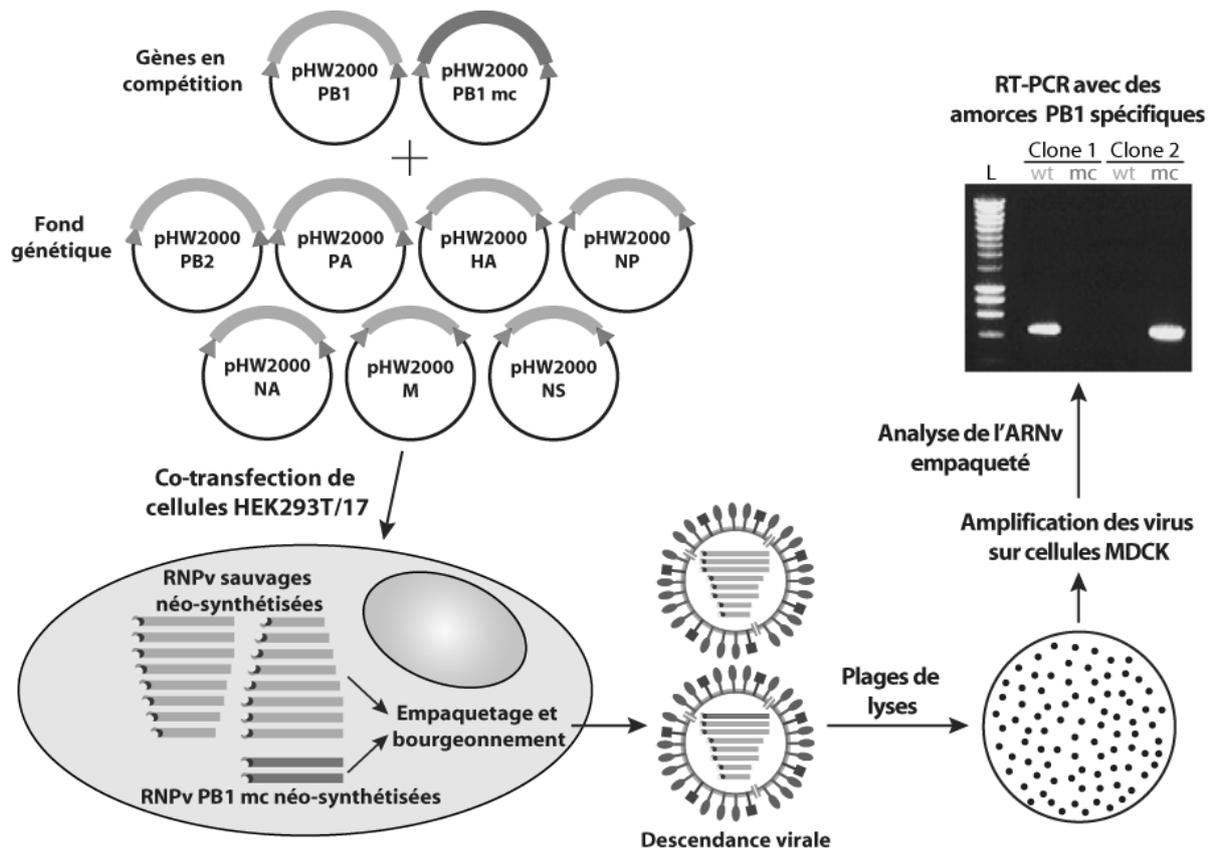
## 2. Cinétique de réplication des virus

Des cellules MDCK ensemencées dans des flasques T25 (5 millions de cellules à confluence) sont infectées à des MOI variables ( $10^{-7}$  et  $10^{-8}$  dans le cas de la souche PR8,  $10^{-4}$  dans le cas de la souche Moscou). Le volume de virus à déposer est soit calculé à partir du titre infectieux exprimé en PFU/ml (souche PR8), soit en  $\text{DICT}_{50}/\text{ml}$  (souche Moscou). Les cellules sont lavées avec le milieu d'infection, puis 1 ml de ce milieu contenant le virus est déposé sur le tapis cellulaire. Après 1h d'incubation à 37°C sous agitation douce, le surnageant est retiré, les cellules sont lavées avec du milieu, puis 11 ml de milieu d'infection sont ajoutés et les cellules sont incubées à 37°C (5% de  $\text{CO}_2$ ). Des aliquots du surnageants sont récoltés au temps zéro, ainsi qu'à différents temps post-infection et sont congelés à -80°C. Les échantillons sont ensuite titrés en  $\text{DICT}_{50}$  dans le but de déterminer le titre infectieux à chaque temps prélevé et ainsi tracer la courbe de réplication des virus étudiés.

## IV. Expériences de génétique inverse compétitive

Les virus produits en génétique inverse peuvent être isolés en plages de lyse afin d'être analysés individuellement. Dans le cadre des expériences de compétition entre deux ARNv, nous avons utilisé cette technique afin d'identifier quel segment a été empaqueté dans les particules virales (figure 23).

Des cellules HEK293T/17 sont co-transfectées comme précédemment (§ E.II.1), la seule différence étant l'utilisation d'un plasmide supplémentaire, à savoir neuf plasmides au total. Quarante-huit heures post-transfection, le surnageant est récolté et dilué au  $10^{\text{ème}}$  dans du milieu d'infection (de  $10^{-1}$  à  $10^{-3}$  par exemple). Le protocole est le même que pour le titrage selon la technique des plages de lyse (voir § E.III.1), la seule différence étant que le volume d'infection est de 800  $\mu\text{l}$ . Après 72h à 96h, les plages de lyse sont dénombrées et repiquées pour infecter des cellules MDCK en plaque 48 puits (BD Biosciences, surface de 1  $\text{cm}^2$  par puits) dans un volume final de 500  $\mu\text{l}$ . Le surnageant de chaque puits est ensuite récolté 48h post-infection, centrifugé 1 min à 4000  $g$ , puis congelé à -20°C.



**Figure 23 : Schéma de la stratégie expérimentale de génétique inverse compétitive.** Les cellules HEK293/T17 sont co-transfectées avec neuf plasmides (dans cet exemple, huit plasmides permettant l'expression des ARNv sauvages, en orange, et un plasmide permettant l'expression de l'ARNv PB1 muté mc, en rouge). Le surnageant est ensuite utilisé pour infecter des cellules MDCK. Les plages de lyse sont repiquées et amplifiées sur des cellules MDCK. Enfin, les virus sont récoltés, les ARN extraits et la version, sauvage ou mutée, du segment d'intérêt (l'ARNv PB1 dans cet exemple) est déterminé par RT-PCR. L : échelle de poids moléculaire ; wt : amorces permettant l'amplification de l'ADNc PB1 sauvage ; mc : amorces permettant l'amplification de l'ADNc PB1 muté.

Par la suite, les ARN sont extraits à partir de 140 µl de surnageant avec le kit QIAamp Viral RNA Mini Kit (Qiagen) selon les recommandations du fabricant. Puis 10 µl d'ARN extrait et 300 ng d'oligonucléotide complémentaire à la séquence U12 (5' AGCAAAGCGG 3') sont incubés à 70°C pendant 5 min. La transcription inverse est réalisée en présence de 11 U de M-MuLV RT (*Moloney murine leukemia virus*) (NEB), du tampon de l'enzyme 1x et de 0,3 mM d'un mélange équimolaire de chacun des quatre dNTP (Eurogentec) dans un volume final de 20 µl pendant 1h à 42°C. L'ensemble est ensuite incubé à 65°C durant 5 min. Afin d'identifier quel ARNv a été empaqueté dans les virions, une PCR est réalisée avec des amorces spécifiques permettant d'amplifier l'ADNc sauvage ou muté (tableau 8). L'ADNc (5 µl) est amplifié en présence de tampon colorless GoTaq® Flexi 1x, de 200 ng de chaque amorce, de 0,4 mM d'un mélange équimolaire de chacun des quatre dNTP (Eurogentec), d'1,25 mM de

MgCl<sub>2</sub> (NEB) et de 0,75 U de GoTaq® (Promega) dans un volume final de 25 µl. Après 5 min de dénaturation à 94°C, 30 cycles d'amplification sont réalisés (1 min de dénaturation à 94°C, 1 min d'hybridation des amorces à 68°C, 1 min d'élongation à 72°C) et sont suivis de 5 min d'élongation à 72°C. Les échantillons sont ensuite déposés sur un gel d'agarose 0,8% (p/v)/TAE 1x contenant du SYBR® Green 1x (Sigma).

Interaction étudiée	ADNc amplifié	Amorce sens (de 5' vers 3')	Amorce antisens (de 5' vers 3')	Taille de l'amplicon
PB2/NP	PB2 wt	TATGCAAGGCTGCAATGG	CATCACATTGTCGATAGGTTCA	446 pb
	PB2 mc 1		CATCACATTGTCGATAGGCTCT	
PB1/M	PB1 wt	CCCTTATACTGGAGACCCTCCT	GCCTTCCAGACTCATTGGC	422 pb
	PB1 mc	CCCTTATACTGGAGATCCACCC		
PB2/NP	NP wt	TGATTGGTGAATTGGACG	CCTCCTAGGGAGAGTTGAA	431 pb
	NP mc 1		CCTCCTAGGGAGAGTAGAG	
PB1/M	M wt	CTGACTCCCAGCATCGGTC	CTTCCGTAGAAGGCCCTCCT	424 pb
	M mc		CTTCCGTAGAAGGTCCACCC	

**Tableau 8 : Amorces de PCR utilisées pour identifier les ARNv empaquetés lors des expériences de compétitions dans le cadre de l'étude des interactions entre les ARNv PB1 et M et les ARNv PB2 et NP de la souche PR8.** Les segments PB1 et M mutés sont ceux qui possèdent les trois mutations *trans*-compensatoires et silencieuses décrites dans la figure 31A et les segments PB2 et NP mutés sont ceux qui possèdent les deux mutations décrites dans la figure 55A.

## RÉSULTATS & DISCUSSION

### A. Études sur la souche humaine PR8

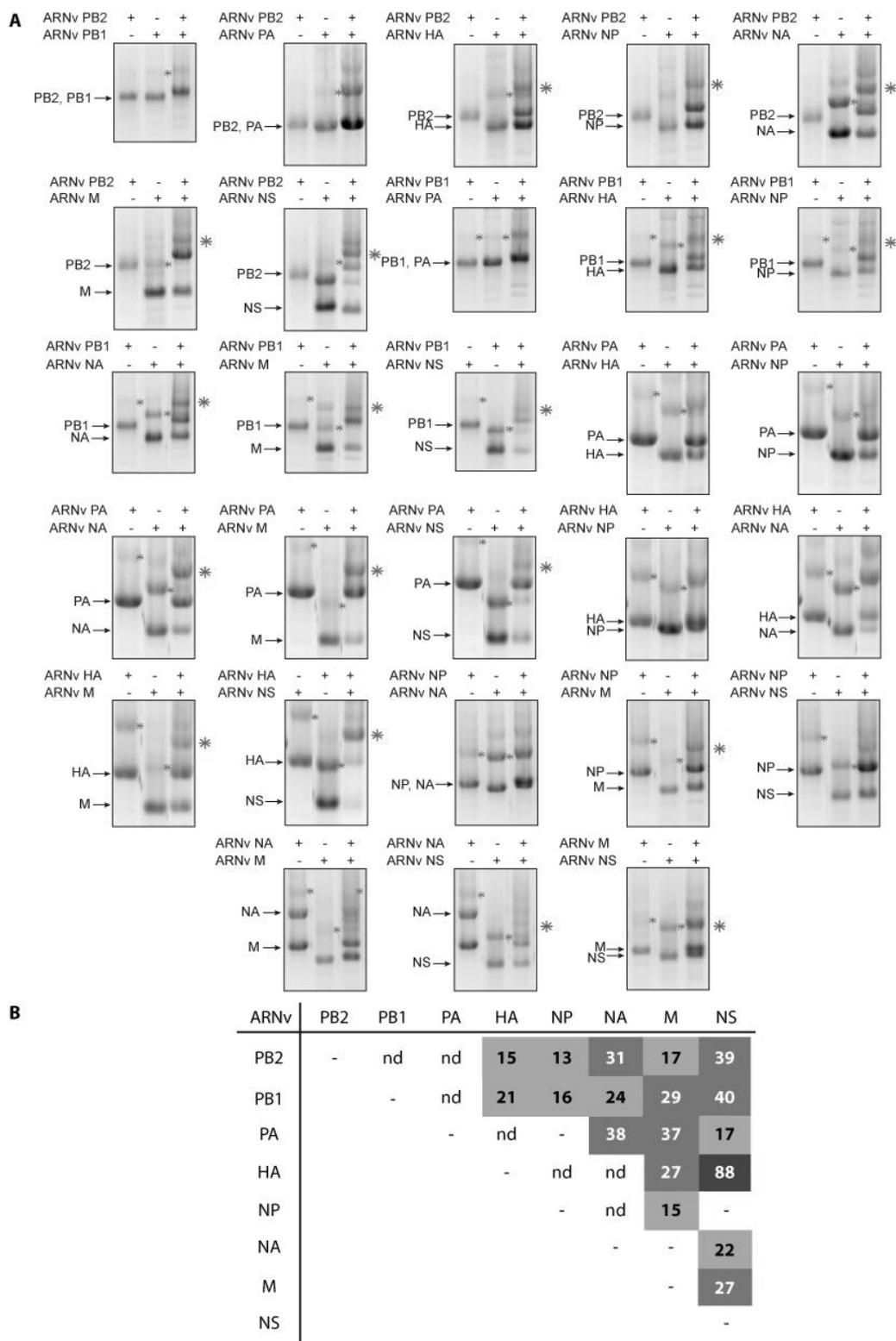
#### I. Établissement du réseau d'interactions entre les ARNv de la souche PR8 par retard de migration sur gel

À l'aide de la stratégie développée au laboratoire, nous avons cherché à établir le réseau d'interactions formé entre les huit ARNv de la souche PR8, *in vitro*. Auparavant, la co-transcription de paires d'ARNv a permis de définir le réseau d'interactions entre les segments de la souche humaine Moscou (H3N2) [308], tandis que dans les mêmes conditions expérimentales les segments de la souche aviaire England (H5N2) interagissent de manière aspécifique (thèse C. Gavazzi, UdS, 2011). Cependant, pour ce virus aviaire, un réseau d'interactions spécifique a été mis en évidence en co-incubant par paires les ARNv préalablement purifiés [319]. De plus, le rôle fonctionnel, pour une réplication optimale et un co-empaquetage sélectif, de l'interaction détectée *in vitro* entre les ARNv PB1 et NS de la souche England, a été établi *in viro* [273], validant ainsi la stratégie *in vitro* utilisée pour l'identification des interactions ARN/ARN.

##### 1. Expériences de retard de migration sur gel après co-transcription de paires d'ARNv

Dans un premier temps, nous avons co-transcrit toutes les paires d'ARNv possibles et recherché des interactions potentielles en déposant les produits de co-transcription sur gel d'agarose en conditions natives (figure 24A). Dans ces conditions, nous avons constaté que la plupart des ARNv ont tendance à homodimériser, ce qui complique l'identification des hétérodimères lorsque les deux ARNv en présence possèdent des tailles relativement similaires. Dans les autres cas, la fraction en masse des ARNv impliqués dans les complexes hétérodimériques a été quantifiée (figure 24B). Sur 28 interactions théoriquement possibles, 18 ont été quantifiées. La moitié d'entre elles sont de faible intensité (entre 15% et 24% de complexe), huit sont d'intensité moyenne (entre 27% et 40% de complexe) et seule l'interaction entre les ARNv HA et NS est très forte (88% de complexe). Il paraît peu probable que la majorité de ces interactions, d'intensité moyenne, puissent s'établir au sein d'un

complexe supramoléculaire tel que celui visualisé par cryo-tomographie électronique [308, 320] et puissent jouer un rôle fonctionnel. En effet, un tel réseau impliquerait la formation d'un grand nombre d'interactions entre des RNPs non adjacentes et contraindrait la formation du complexe, alors que des études suggèrent une certaine flexibilité dans l'organisation des RNPs les uns par rapport aux autres entre virions appartenant à la même souche [308, 320]. Il semblerait donc que les conditions expérimentales utilisées pour mettre en évidence des interactions entre les ARNs de la souche humaine Moscou ne soient pas adaptées à l'étude de la souche humaine PR8.

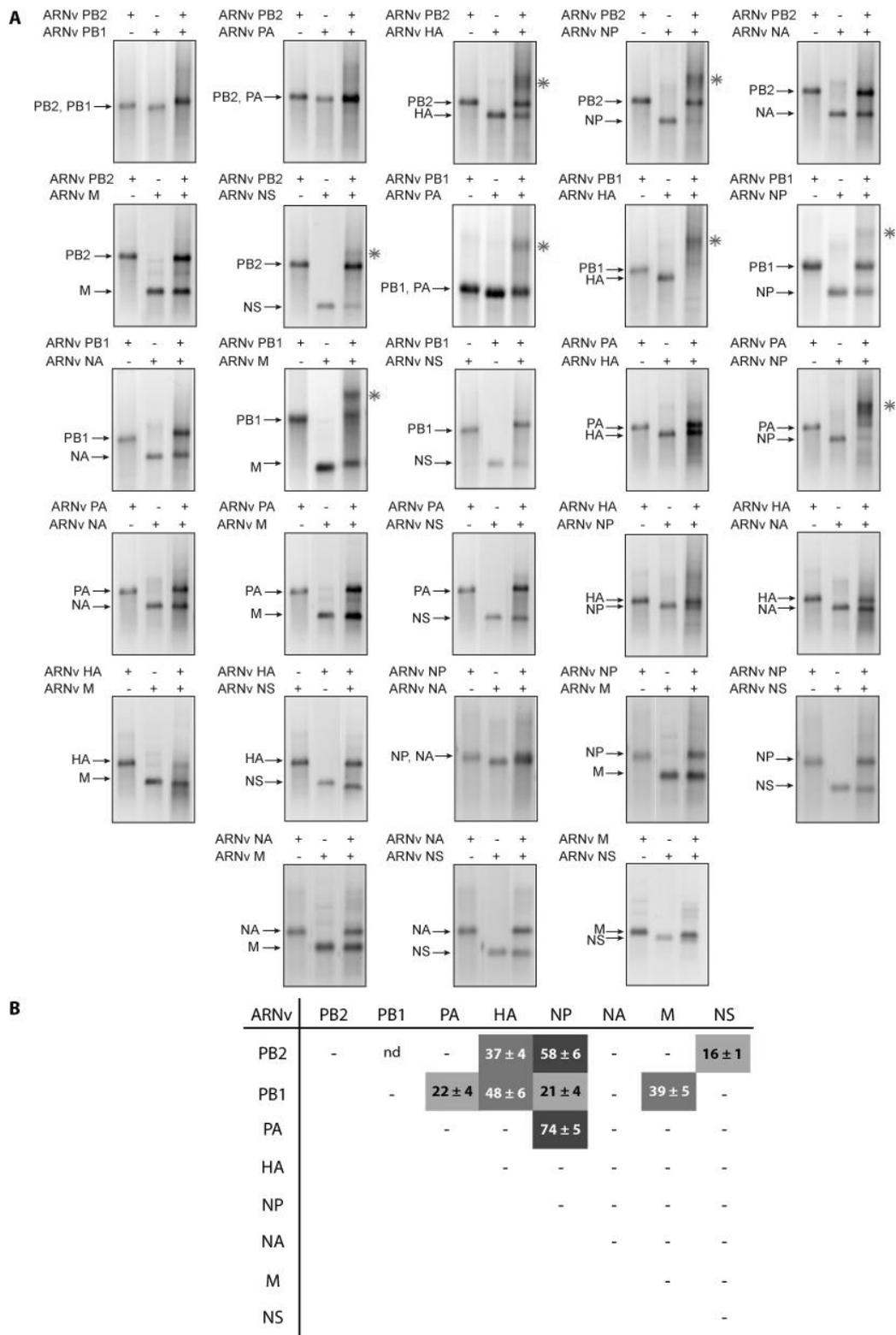


**Figure 24 : Réseau d'interactions *in vitro* entre les ARNV de la souche PR8 co-transcrits deux à deux.**  
**A.** Analyse des interactions ARN/ARN entre le segment PB1 et les sept autres ARNV par électrophorèse sur gel d'agarose en conditions natives après co-transcription par paires des ARNV. Les homodimères d'ARNv sont indiqués par des astérisques bleus et les hétérodimères par des astérisques rouges. **B.** Quantification de la proportion de complexe formé (%) par rapport à la quantité totale d'ARNv par piste. - : non détectable ou moins de 10%. nd : non déterminé, dans le cas où il n'a pas été possible de déterminer si les complexes formés sont des homodimères ou des hétérodimères.

## 2. Expériences de retard de migration sur gel après co-incubation de paires d'ARNv purifiés

Après avoir transcrit les huit ARNv de la souche PR8 à l'aide de l'ARN polymérase du phage T7 et purifié les segments sur colonne de gel filtration (voir Matériels et Méthodes, § B.II), nous avons co-incubé les segments viraux deux à deux à différentes températures et dans différents tampons. Aucune interaction n'a été identifiée en co-incubant les segments à 37°C ou 42°C, que ce soit dans un tampon de faible salinité contenant 50 mM de cacodylate de sodium pH 7,5, 40 mM de KCl et 0,1 mM de MgCl<sub>2</sub> (le tampon utilisé pour le virus aviaire England [319]) ou dans un tampon de plus haute salinité (50 mM de cacodylate de sodium pH 7,5, 300 mM de KCl, 5 mM de MgCl<sub>2</sub>; le tampon utilisé pour la dimérisation d'ARN rétroviraux [321]) (résultats non présentés). Nous avons également dénaturé les ARNv séparément dans l'eau avant de les co-incuber, à 37°C ou 42°C, dans le tampon de plus haute salinité. Enfin, nous avons tenté de co-incuber les segments dans le milieu de transcription, dépourvu d'ARN polymérase du phage T7, à 37°C. Dans ces trois derniers cas, aucune interaction entre les ARNv n'était détectable (résultats non présentés).

Finalement, des interactions entre les ARNv de la souche PR8 ont été mises en évidence en co-incubant les ARNv à 55°C, dans le tampon de faible salinité (figure 25A). Dans ces conditions, qui correspondent à celles qui ont permis d'établir le réseau du virus England [319], les ARNv homodimérisent très peu et nous identifions huit partenariats ARN/ARN. La fraction pondérale de complexe varie de 16% à 74% (figure 25B). Trois interactions sont de faible intensité (entre 16% et 22% de complexe), trois sont d'intensité moyenne (entre 37% et 48% de complexe) et deux sont de forte intensité (58% et 74% de complexe). Il semblerait donc que cette condition expérimentale soit la plus appropriée pour identifier des interactions spécifiques entre les ARNv de l'IVA PR8. Dans ce réseau, les segments les plus longs (PB2, PB1, PA, HA et NP) sont impliqués dans la plupart des interactions intermoléculaires. L'ARNv PB1 est impliqué dans quatre interactions, tandis que les ARNv PB2 et NP en établissent chacun trois. Le segment NP est engagé dans les deux interactions les plus fortes, avec les ARNv PB2 et PA (58% et 74% de complexe, respectivement), et pourrait donc jouer un rôle important dans l'établissement du réseau. Le segment NA n'a pas de partenaire dans le réseau d'interactions déterminé *in vitro*.



**Figure 25 : Réseau d'interactions *in vitro* entre les ARNv de la souche PR8 co-incubés deux à deux.**  
**A.** Analyse des interactions ARN/ARN par électrophorèse sur gel d'agarose en conditions natives après co-incubation des ARNv par paires, à 55°C, dans un tampon de faible salinité. Les monomères d'ARNv sont indiqués par des flèches et les hétérodimères par des astérisques rouges.  
**B.** Quantification de la proportion de complexe formé (%) par rapport à la quantité totale d'ARNv par piste (moyenne de 4 expériences indépendantes ± SEM). - : non détectable ou moins de 10%. nd : non déterminé, dans le cas où il n'a pas été possible de déterminer si les complexes formés sont des homodimères ou des hétérodimères.

### 3. Conclusion

La comparaison des trois réseaux d'interactions désormais établis au laboratoire (celui des souches humaines Moscou (H3N2) [308] et PR8 (H1N1) (figure 25B) et de la souche aviaire England (H5N2) [319]), nous donne des pistes quant à l'existence d'une éventuelle spécificité d'hôte et/ou de sous-type pour l'établissement des interactions ARN/ARN formant le réseau (figure 26). En effet, nous constatons que les trois virus testés, de sous-types différents, ne partagent pas le même réseau. Il en va de même pour les deux virus humains, de sous-types différents. Les virus England et PR8 partagent néanmoins trois interactions (entre les ARNv PB1 et PA, PB1 et HA, PA et NP). Il semblerait donc que les réseaux établis ne soient pas conservés parmi les IVA infectant un même hôte. À ce stade, la possibilité d'une spécificité liée au sous-type des souches restait à vérifier.

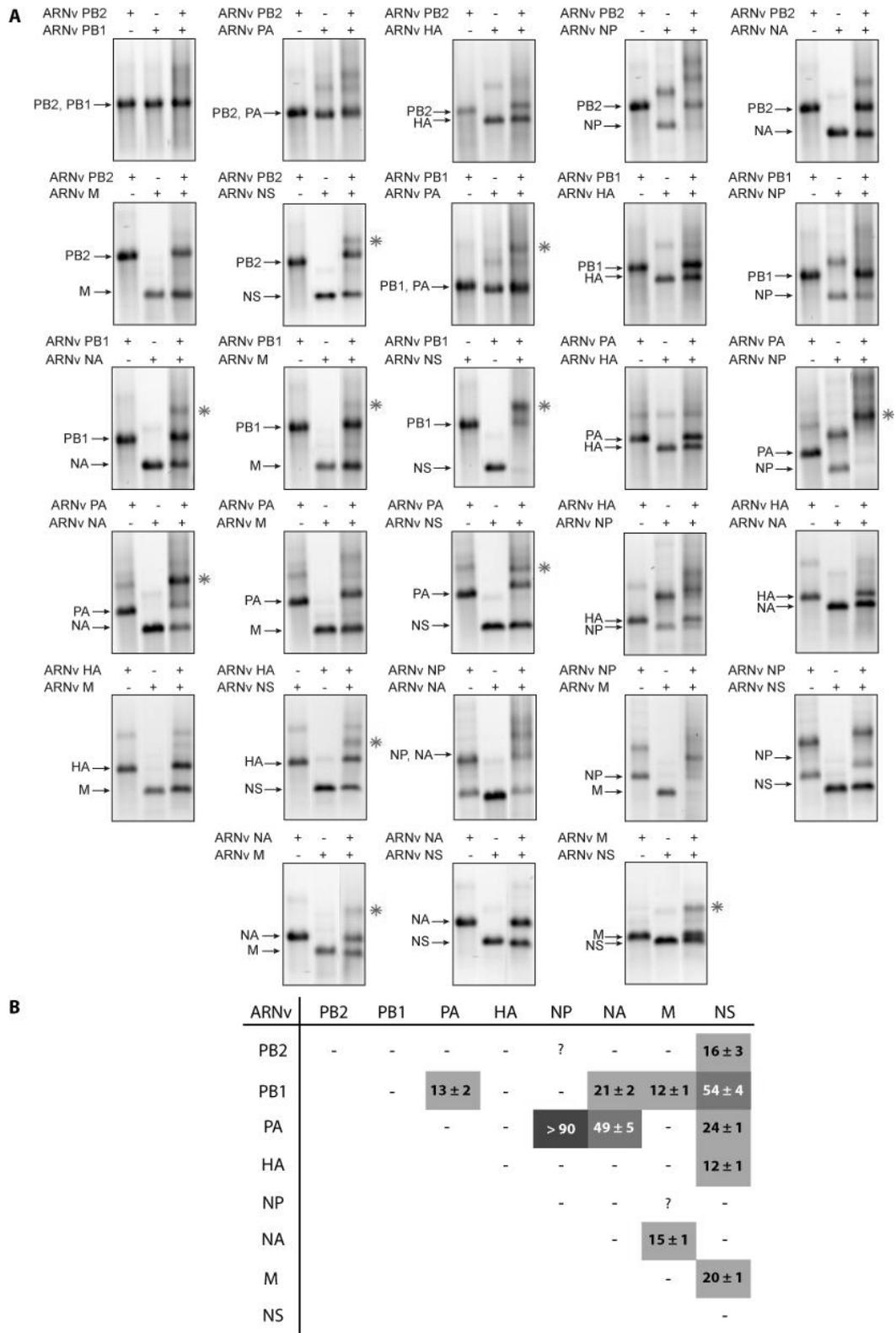
ARNv	PB2	PB1	PA	HA	NP	NA	M	NS
PB2	-	X		X	X	X	X	X
PB1		-	X	X	X	X	X	X
PA			-	X	X	X	X	X
HA				-	X	X	X	X
NP					-	X	X	X
NA						-	X	X
M							-	X
NS								-

**Figure 26 : Comparaison des trois réseaux d'interactions établis au laboratoire.** Les croix indiquent les interactions mises en évidence entre les ARNv de la souche humaine Moscou (H3N2) (rouge), aviaire England (H5N2) (violet) et humaine PR8 (H1N1) (bleu).

Dans ce but, nous avons récemment déterminé, en collaboration avec A. Printz-Schweigert au sein de notre équipe, le réseau d'interactions entre les ARNv de la souche humaine WSN (H1N1) (figure 27), une souche modèle de laboratoire dont les signaux d'empaquetage ont partiellement été définis dans la littérature (pour revues, voir [33, 247]). La comparaison de la séquence nucléotidique des ARNv des souches PR8 et WSN révèle un pourcentage d'identité variant entre 94% et 97%. Les interactions entre les segments d'ARNv de la souche WSN ont été identifiées dans les mêmes conditions que celles choisies précédemment pour déterminer le réseau de la souche PR8 (co-incubation des ARNv purifiés à 55°C dans un tampon de faible salinité). Onze interactions intermoléculaires ont été identifiées, dont la fraction pondérale de complexe varie entre 12% et plus de 90% (figure

27B). La plupart des interactions sont néanmoins de faible intensité (entre 12% et 24% de complexe), tandis que deux interactions sont de forte intensité (49% et 54%) et une interaction est de très forte intensité (plus de 90% de complexe). Les ARNv de petites tailles semblent jouer un rôle important, notamment l'ARNv NS qui est impliqué dans cinq interactions intermoléculaires, tandis que les ARNv NA et M ont chacun deux partenaires. Le segment PB1 est également crucial dans ce réseau puisqu'il interagit avec quatre autres segments, PA, NA, M et NS. L'ARNv NA est impliqué dans trois interactions ARN/ARN, dont celle avec le segment PA qui constitue l'une des interactions la plus forte du réseau (49% de complexe). De manière intéressante, l'interaction entre les ARNv PA et NP est la plus forte dans les deux réseaux de virus humains H1N1 : 74% de complexe pour la souche PR8 (figure 25B) et plus de 90% pour la souche WSN (figure 27B).

Néanmoins, de façon surprenante, les réseaux d'interactions entre les ARNv des souches PR8 (figure 25B) et WSN (figure 27B) sont radicalement différents. Par conséquent, il semblerait que l'établissement du réseau d'interactions entre les ARNv des IVA ne soit ni dicté par l'hôte d'origine de la souche, ni par le sous-type viral, tout du moins dans les conditions expérimentales testées.



**Figure 27 : Réseau d'interactions *in vitro* entre les ARNV de la souche WSN co-incubés deux à deux.**  
**A.** Analyse des interactions ARN/ARN par électrophorèse sur gel d'agarose en conditions natives après co-incubation des ARNV par paires, à 55°C, dans un tampon de faible salinité. Les monomères d'ARNv sont indiqués par des flèches et les hétérodimères par des astérisques rouges.  
**B.** Quantification de la proportion de complexe formé (%) par rapport à la quantité totale d'ARNv par piste (moyenne de 3 expériences indépendantes ± SEM). - : non détectable ou moins de 10%. ? : interaction détectée de manière non-reproductible.

#### 4. Discussion

Dans cette première partie, nous avons déterminé les réseaux d'interactions formés entre les huit ARNv des virus humains PR8 (figure 25B) et WSN (figure 27B), tous deux de sous-type H1N1. La comparaison de ces réseaux avec les deux réseaux établis précédemment au laboratoire (celui des souches humaine Moscou (H3N2) [308] et aviaire England (H5N2) [319]) nous permet de conclure que dans nos conditions expérimentales il n'existe ni de spécificité d'hôte, ni de spécifié de sous-type dans la formation du réseau supramoléculaire d'assemblage entre les ARNv grippaux.

Cependant, dans notre système *in vitro*, nous élaborons le réseau d'interactions entre des ARNv nus, qui ne sont ni associés au complexe ARN polymérase, ni aux protéines NP. Ainsi, notre réseau n'a de sens que si l'association de ces protéines aux ARNv expose les régions impliquées dans l'interaction et permet donc la formation du complexe dans le cytoplasme des cellules infectées et au site de bourgeonnement. L'absence du complexe ARN polymérase ne semble pas être déterminante pour la structure, étant donné que l'ARNv en présence des protéines NP peut adopter des structures semblables à des RNPv [73]. L'absence de protéines NP est plus critique, puisque celles-ci forment le squelette protéique [27, 30, 64] sur lequel s'enroule l'ARNv. Plusieurs études ont néanmoins montré que l'ARNv est présent à la surface des protéines NP [27, 30, 63, 65] et que leur présence augmente l'accessibilité des bases de l'ARNv à des sondes chimiques de structure [24]. La présence de protéines NP devrait donc faciliter les interactions entre les ARNv plutôt que les restreindre. Il est cependant probable que la région de l'ARNv fixée dans le sillon de la protéine ne puisse pas interagir avec un autre ARNv en raison de gênes stériques. De plus, le laboratoire a montré que la présence de la protéine NP n'affecte pas la spécificité de l'interaction entre les ARNv NA et M de la souche Moscou [308]. Enfin, l'interaction entre les ARNv PB1 et NS de la souche England, identifiée par l'approche « classique » *in vitro* que nous venons de détailler, joue un rôle crucial dans l'empaquetage du génome viral [273], étayant notre approche d'identification des interactions ARN/ARN, en absence de protéines virales.

L'existence d'un segment directeur (« master » segment), qui pourrait être situé au centre du complexe supramoléculaire et qui interagirait avec la plupart ou l'ensemble des sept autres segments génomiques, même si cette seconde hypothèse est peu probable [308,

320], a été proposée par Hutchinson *et al.* [322]. En effet, plusieurs études ont montré que certains ARNv sont plus importants que d'autres dans le mécanisme d'empaquetage du génome et de l'assemblage des particules virales, dans la souche PR8 [107, 260] ou WSN [258, 259, 262, 323]. D'après ces travaux, les segments PB2 (virus WSN) [259, 262] ou M (virus WSN et PR8) [260, 323] pourraient constituer de potentiels candidats pour ce rôle central, étant donné que des mutations dans les signaux d'empaquetage de ces ARNv ont un impact considérable sur l'empaquetage d'autres segments. De manière intéressante, l'introduction de la même mutation dans l'ARNv PA n'a pas le même impact sur l'empaquetage du génome des souches WSN ou PR8 [262]. Récemment, l'importance fonctionnelle des signaux d'empaquetage de chacun des huit segments de la souche PR8 a été évaluée [324], démontrant ainsi que les segments PB2, PA, NP et M jouent un rôle plus important que les autres segments lors de l'empaquetage du génome. Dans le réseau *in vitro* que nous avons déterminé, ces quatre segments sont impliqués dans au moins une interaction de moyenne ou de forte intensité (figure 25B) : l'ARNv PB2 interagit avec les ARNv HA (37% de complexe) et NP (58% de complexe), tandis que l'ARNv NP interagit également fortement avec l'ARNv PA (74% de complexe). L'ARNv M interagit avec l'ARNv PB1 (39% de complexe). Les résultats de Gao *et al.* [324] sont donc compatibles avec la « force » des interactions que nous avons détectées *in vitro*. Cependant, les ARNv HA et PB1 jouent également un rôle important dans le réseau de la souche PR8, puisqu'ils forment 48% de complexe. De plus, les ARNv les plus longs pourraient jouer un rôle central dans la mise en place du réseau d'interactions. En effet, les segments PB2, PB1, PA, HA et NP sont impliqués dans deux à quatre interactions intermoléculaires (figure 25B). Les ARNv PB2, PB1 ou NP pourraient donc jouer le rôle de segments directeurs, puisqu'ils établissent chacun trois ou quatre interactions ARN/ARN. Dans le cas du réseau de la souche WSN, seules trois interactions sont de forte ou très forte intensité (figure 27B) : entre les ARNv PA et NP (plus de 90% de complexe), les ARNv PA et NA (49% de complexe) et les ARNv PB1 et NP (54% de complexe), tandis que les huit autres interactions sont de faible intensité. Les ARNv PB1 et NS sont ceux qui établissent le plus d'interactions dans ce réseau (quatre et cinq interactions, respectivement), tandis que les ARNv PA, NA et M sont chacun impliqués dans trois interactions. L'un de ces cinq ARNv pourrait donc constituer le « master » segment de ce virus. Les travaux de Muramoto *et al.* [259] ont néanmoins montré que l'ARNv PB2 joue un rôle plus important dans l'empaquetage du génome que les segments PB1 et PA. Les

virus PR8 et WSN, qui ne partagent pas les mêmes réseaux d'interactions, semblent donc également posséder des « master » segments différents, alors que les séquences nucléotidiques des ARNv des deux souches présentent 96% d'identité. Dans le réseau Moscou, les ARNv HA et M établissent chacun trois interactions intermoléculaires (15% à 39% et 27% à 54% de complexe, respectivement) [308], tandis que dans le réseau England, les ARNv PB1 et NS sont impliqués dans quatre (9% à 82% de complexe) et trois interactions (8% à 82% de complexe), respectivement [319]. L'ensemble de ces données suggère donc que chaque virus pourrait posséder son propre réseau, dont l'assemblage serait guidé par des « master » segments différents selon la souche.

Dans nos expériences, nous avons fixé un seuil de détection des interactions à 10%. En dessous de cette valeur, nous avons considéré que l'interaction n'était pas significative. Dans le réseau PR8, le segment NA est le seul pour lequel aucune interaction n'a été détectée au dessus de 10% (figure 25). En se basant sur ce seuil, l'ARNv NA de la souche England étudiée précédemment au laboratoire est également orphelin, puisque les interactions qu'il établit avec ses deux partenaires, les ARNv PB2 et PB1, sont de 7% et 9%, respectivement [319]. À l'inverse, l'ARNv NA de la souche Moscou est impliqué dans trois interactions du réseau, deux de faible intensité, avec les ARNv PB1 et PA (15% et 23% de complexe, respectivement), et celle avec le segment M qui est la plus forte (54% de complexe) [308]. Dans le réseau de la souche WSN, l'ARNv NA possède également trois partenaires, les ARNv PB1 et M, avec lesquels il établit des partenariats de faible intensité (21% et 15% de complexe, respectivement), et l'ARNv PA, avec lequel il forme un complexe de plus forte intensité (49% de complexe) (figure 27B). Il n'est pas exclu que l'absence de partenaire de l'ARNv NA puisse être expliquée par nos conditions expérimentales, qui ne révéleraient pas toutes les interactions entre ARNv. Inversement, il est possible également que des interactions mises en évidence dans ces conditions ne soient pas représentatives des interactions qui s'établissent dans la cellule infectée.

Il est intéressant de noter que les travaux visant à identifier les signaux d'empaquetage des segments génomiques n'ont pas alloué aux segments HA et NA, codant les protéines antigéniques, un rôle majeur dans l'empaquetage des ARNv [255, 256, 258, 324, 325]. Cette observation pourrait refléter un processus évolutif qui faciliterait l'échange des segments HA et NA lors du réassortiment génétique. Des expériences de génétique

inverse compétitive entre les segments NA des souches PR8 et (H1N1)pdm09, dans le fond génétique PR8, ont par ailleurs montré que la moitié des virus obtenus contiennent le segment d'origine pandémique, suggérant que l'empaquetage de cet ARNv n'est pas dicté par son interaction avec un autre partenaire (M. Rosa-Calatrava, communication personnelle). À l'inverse, l'ARNv HA, qui, comme le segment NA ne semble pas primordial pour l'empaquetage [256, 258, 324, 325], est impliqué dans deux interactions de forte intensité dans notre réseau *in vitro* (37% de complexe avec l'ARNv PB2 et 48% avec l'ARNv PB1) (figure 25B). L'implication de ce segment dans la formation des réseaux *in vitro* des souches Moscou, England et WSN est également variable. En effet, le segment HA est impliqué dans trois interactions dans le réseau Moscou [308], de 15% intensité avec l'ARNv PA et de moyenne intensité avec les ARNv M (35% de complexe) et NS (39% de complexe), dans quatre interactions dans le réseau England [319], dont une très forte avec l'ARNv M (66% de complexe), tandis qu'il n'établit qu'une seule interaction de faible intensité dans le réseau WSN (12% de complexe avec l'ARNv NS) (figure 27B). En accord avec ces résultats, une étude au laboratoire a montré que l'interaction entre les segments HA et M de la souche England est cruciale pour leur co-empaquetage dans le fond génétique Moscou [307].

La reconstitution tridimensionnelle de l'intérieur des particules virales par tomographie électronique permet d'observer l'organisation des RNPv les unes par rapport aux autres. Des travaux sur la souche WSN (H1N1) [320] ont révélé que la disposition relative des huit RNPv est différente à l'intérieur des 30 particules virales analysées. Certaines contraintes existent pourtant, puisque la position centrale est toujours occupée par une des RNPv longues (PB2, PB1, PA ou HA), qui pourraient ainsi constituer les « master » segments de ce virus. De plus, les trois autres RNPv longues, situées à la périphérie, ne sont jamais contiguës. Dans le cas de la souche H3N2, la combinaison des données biochimiques d'interactions entre ARNv et de la tomographie a permis d'assigner une identité à la plupart des RNPv [308]. Parmi les quatre particules virales analysées, seules deux combinaisons ont été retrouvées (sur les 5760 combinaisons possibles) et l'ARNv HA a été identifié comme étant le segment central. L'organisation des RNPv n'est donc pas strictement conservée parmi les particules virales d'une même souche, même si certaines dispositions sont favorisées. Des études de tomographie électronique sont en cours (en collaboration avec le

laboratoire du Dr P. Bron, Montpellier), afin d'obtenir une reconstruction tridimensionnelle des RNPv à l'intérieur de particules virales bourgeonnantes de la souche PR8. Ces résultats, combinés à nos données *in vitro*, devraient permettre de proposer un modèle de l'organisation des huit RNPv à l'intérieur des virions. La détection, ou non, d'une plateforme contenant les complexes ARN polymérase, située à l'extrémité bourgeonnante des virions, comme dans le cas du virus Moscou [308] sera un autre résultat majeur de cette étude. Dans le cas des segments qui partagent la même longueur, il est également envisageable de construire des segments chimériques, de longueurs variables, conçus pour être facilement identifiables au sein des virions. En parallèle, la tomographie électronique sur des virus réassortants possédant des segments HA et NA étrangers dans un fond génétique PR8 permettra d'évaluer les perturbations éventuelles liées à l'introduction de segments étrangers sur l'organisation interne des RNPv.

## **II. Identification des régions des ARNv de la souche PR8 impliquées dans certaines interactions ARN/ARN mises en évidence *in vitro***

### **1. Approche « classique »**

Après avoir défini le réseau d'interactions entre les huit segments de la souche humaine PR8, il s'agissait d'identifier précisément les régions des ARNv impliquées dans ces interactions, si possible au nt près. Pour cela, nous avons suivi le schéma expérimental établi pour l'identification au nt près de séquences impliquées dans certaines interactions ARN/ARN du réseau aviaire England [273, 319]. Nous avons i) construit et testé des ARNv portant de grandes délétions, ii) utilisé une stratégie d'inhibition des interactions à l'aide de petits oligonucléotides antisens complémentaires aux ARNv, iii) appliqué une recherche de complémentarité de séquence par bioinformatique, sur des régions définies des ARNv et iv) validé *in vitro* l'existence d'interactions entre ARNv à l'aide de mutants *trans-compensatoires*. Ainsi, en combinant ces outils, nous avons déterminé les régions d'interaction de plusieurs partenariats ARN/ARN. Dans le cas de l'interaction entre les segments PB1 et M nous avons délimité, au nt près, les séquences responsables de l'hétérodimérisation.

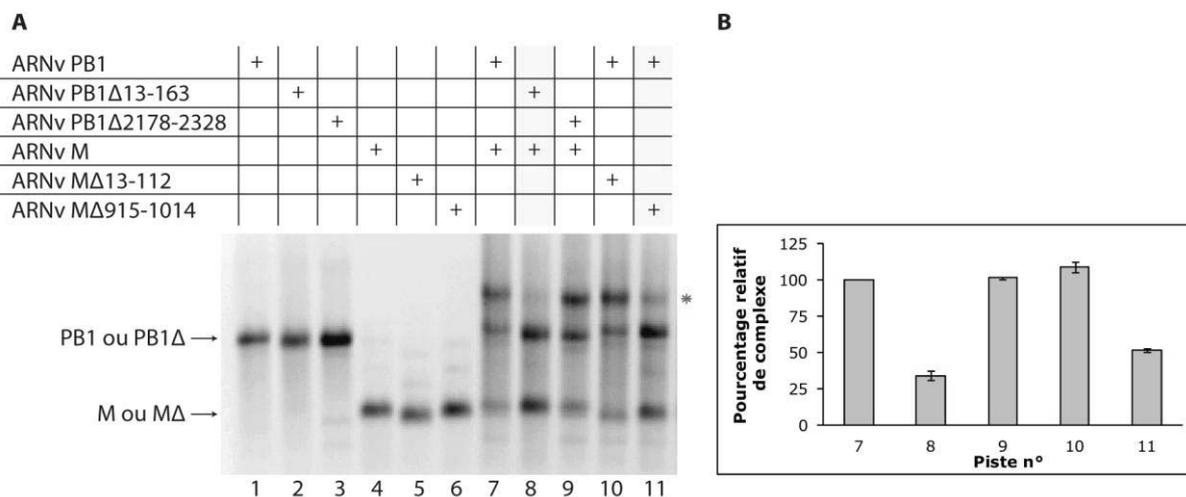
#### **1.1. Interaction entre les ARNv PB1 et M**

##### **a. Mutagenèse de délétion, inhibition par des oligonucléotides antisens et bioinformatique**

D'après la littérature, les signaux d'empaquetage des souches H1N1 PR8 et WSN sont localisés aux extrémités des ARNv, généralement dans les 100 nt attenants aux séquences U12 et U13 (pour revues, voir [33, 247]). L'hypothèse du laboratoire étant que les séquences impliquées dans les interactions entre ARNv se superposent au moins partiellement aux signaux d'empaquetage, nous avons construit des ARNv délétés de grandes régions terminales (150 ou 100 nt, respectivement pour les segments PB1 et M), tout en conservant les séquences U12 et U13. Les segments mutés ont été testés en présence de leur partenaire sauvage et le pourcentage de chaque complexe intermoléculaire obtenu en présence des segments mutés a été normalisé, pour faciliter la comparaison des données, par rapport au complexe formé entre les ARNv PB1 et M sauvages. Toutes les données concernant les ARNv

mutants et l'inhibition par les oligonucléotides antisens sont présentées de cette manière tout au long du manuscrit.

La co-incubation des ARNv sauvages en présence de leur partenaire muté a montré que les ARNv PB1 $\Delta$ 13-163 et M $\Delta$ 915-1014 diminuent respectivement l'interaction de 66% (34% de complexe restant) et 48% (52% de complexe restant) (figure 28, pistes 8 et 11). À l'inverse, les ARNv PB1 $\Delta$ 2178-2328 et M $\Delta$ 13-112 n'affectent pas la formation du complexe (figure 28, pistes 9 et 10). Ces résultats suggèrent que l'interaction entre les ARNv PB1 et M fait intervenir des nt localisés entre les positions 13-163 de l'ARNv PB1 et les positions 915-1014 de l'ARNv M.

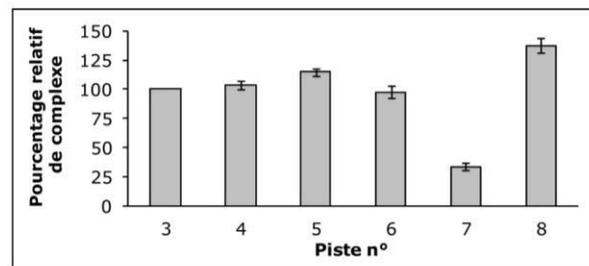
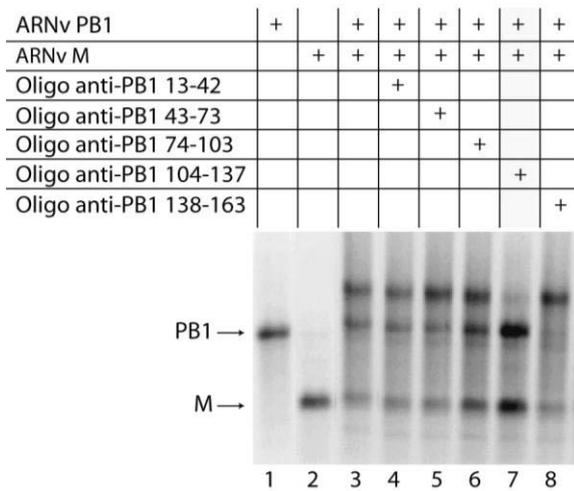


**Figure 28 : Identification des régions des ARNv PB1 et M impliquées dans l'hétérodimérisation à l'aide d'ARNv portant des délétions au niveau de leurs extrémités. A.** Analyse des interactions ARN/ARN par électrophorèse sur gel d'agarose en conditions natives après co-incubation des ARNv PB1 et M sauvages et des ARNv portant de grandes délétions au niveau de leurs extrémités. Les monomères d'ARNv sont indiqués par des flèches et les hétérodimères par un astérisque rouge. **B.** Quantification de la proportion de complexe présent dans chaque piste (moyenne de 3 expériences indépendantes  $\pm$  SEM). Les complexes intermoléculaires obtenus dans les pistes 8-11 sont normalisés par rapport au complexe formé entre les ARNv PB1 et M sauvages (piste 7).

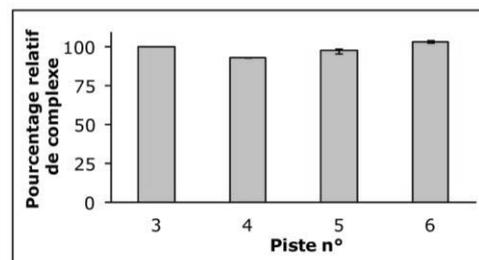
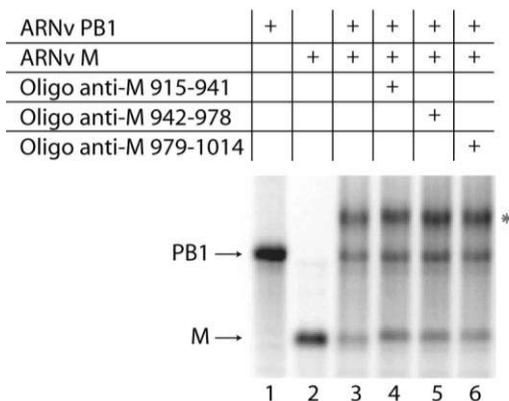
La séquence d'interaction a ensuite été délimitée plus précisément par l'ajout d'oligonucléotides antisens d'une trentaine de nt lors de la co-incubation des deux ARNv sauvages (voir Matériels et Méthodes, tableau 6). Si l'oligonucléotide ajouté cible efficacement plusieurs nt engagés dans l'interaction, l'hétérodimérisation des deux ARNv devrait en être affectée. Ces oligonucléotides sont complémentaires aux régions 13-163 du segment PB1 et 915-1014 du segment M, précédemment définies comme étant impliquées dans la formation du complexe (figure 28). Parmi les cinq oligonucléotides antisens testés et

complémentaires à l'ARNv PB1, le seul à inhiber l'interaction avec l'ARNv M (34% d'interaction restante), de la même manière que la délétion de la région 13-163 de l'ARNv PB1, s'hybride aux nt 104-137 (figure 29A, piste 7). Cependant, aucun oligonucléotide complémentaire à la région 915-1014 de l'ARNv M ne diminue l'interaction avec le segment PB1 (figure 29B).

### A



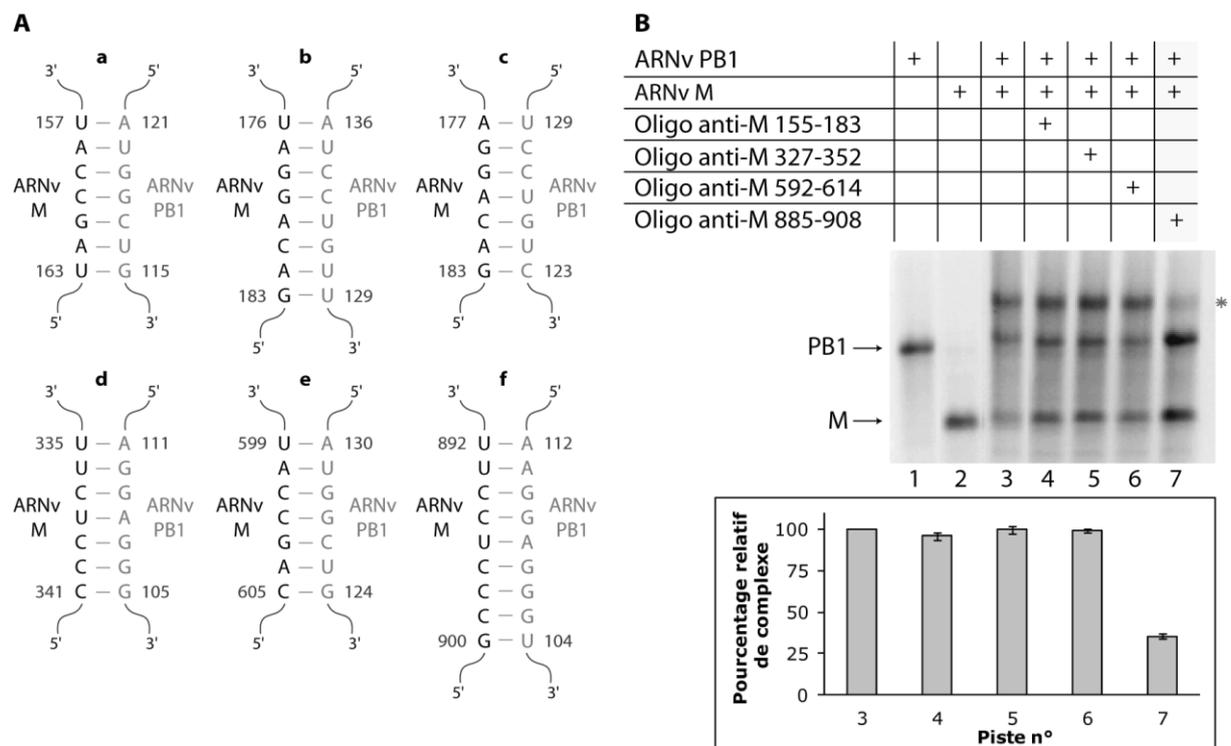
### B



**Figure 29 : Identification des régions des ARNv PB1 et M impliquées dans l'hétérodimerisation à l'aide d'oligonucléotides antisens complémentaires aux ARNv.** Analyse des interactions ARN/ARN par électrophorèse sur gel d'agarose en conditions natives après co-incubation des ARNv PB1 et M sauvages en présence d'oligonucléotides complémentaires à de courtes séquences de (A) l'ARN PB1 et de (B) l'ARNv M. Les monomères d'ARNv sont indiqués par des flèches et les hétérodimères par un astérisque rouge. La quantification de la proportion de complexe présent dans chaque piste (moyenne de 2 à 4 expériences indépendantes  $\pm$  SEM) a été réalisée.

Nous avons ensuite cherché, à l'aide de l'outil bioinformatique Google (<http://bibiserv.techfak.uni-bielefeld.de/google>) [326], des appariements antiparallèles possibles entre les nt 104-137 de l'ARNv PB1 et la totalité de l'ARNv M. Un grand nombre d'interactions a été proposé. Nous avons donc éliminé les interactions qui présentaient des

répétitions de U et plusieurs appariements « wobble » G-U. Ces deux critères nous ont permis de sélectionner six interactions potentielles (figure 30A). Nous avons testé quatre oligonucléotides antisens qui ciblent ces régions de l'ARNv M (l'oligonucléotide complémentaire aux nt 155-183 permet de tester les interactions a, b et c) (voir Matériels et Méthodes, tableau 6). L'oligonucléotide complémentaire aux nt 885-908 diminue l'interaction de la même manière que l'oligonucléotide complémentaire au segment PB1 (35% d'interaction restante) (figure 30B, piste 7), tandis que les autres oligonucléotides n'inhibent pas la formation du complexe (figure 30B, pistes 4-6). D'après l'ensemble de ces résultats, il semblerait que l'interaction entre les segments PB1 et M fasse intervenir les nt 104-112 de l'ARNv PB1 et les nt 892-900 de l'ARNv M (figure 30A, interaction f).

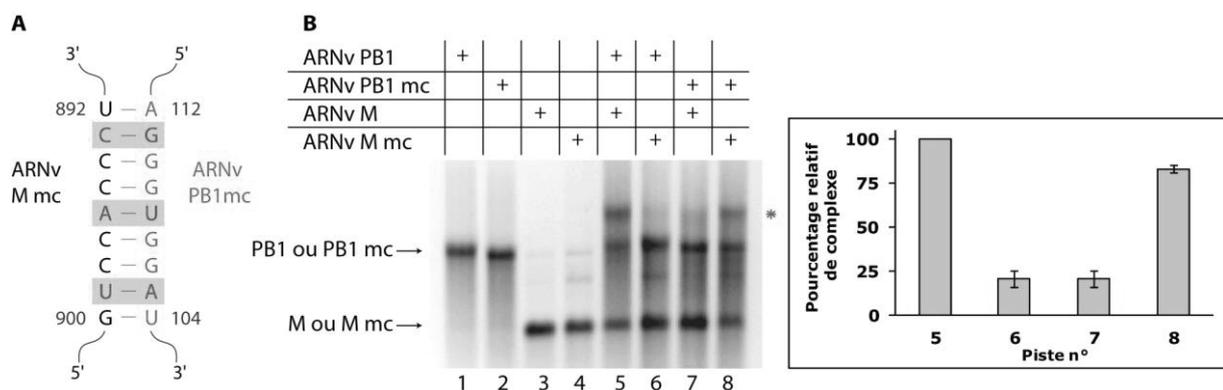


**Figure 30 : Identification de la région de l'ARNv M impliquée dans l'hétérodimérisation avec l'ARNv PB1 à l'aide d'oligonucléotides antisens complémentaires à l'ARNv M.** **A.** Interactions potentielles entre les ARNv PB1 et M proposées par l'outil bioinformatique Guugle. **B.** Analyse des interactions ARN/ARN par électrophorèse sur gel d'agarose en conditions natives après co-incubation des ARNv PB1 et M sauvages en présence d'oligonucléotides complémentaires à de courtes séquences de l'ARNv M. Les monomères d'ARNv sont indiqués par des flèches et les hétérodimères par un astérisque rouge. La quantification de la proportion de complexe présent dans chaque piste (moyenne de 4 expériences indépendantes  $\pm$  SEM) a été réalisée.

Les résultats de co-incubation entre l'ARNv PB1 sauvage et les ARNv M portant des délétions aux extrémités suggèrent que les nt 915-1014 de l'ARNv M sont impliqués dans

l'hétérodimérisation (figure 28, piste 11). Néanmoins, contrairement à ce qui était attendu, l'utilisation d'oligonucléotides complémentaires à cette région ne diminue pas l'interaction formée entre les deux ARNv (figure 29B). La bioinformatique combinée à l'utilisation d'oligonucléotides antisens ont permis de délimiter la séquence d'interaction aux nt 892-900 du segment M (figure 30B, piste 7). Cette séquence est localisée quinze nt avant la région 915-1014 de l'ARNV M initialement délétée. Ainsi, l'effet inhibiteur que nous avons observé avec le mutant de délétion 915-1024 pourrait être dû à un changement conformationnel de l'ARNv M qui affecterait la structure de la région 892-900 et l'empêcherait d'interagir avec l'ARNv PB1.

Afin de valider les séquences de chaque ARNv impliquées dans l'interaction entre les segments PB1 et M, trois mutations *trans*-compensatoires ont été introduites dans chaque ARNv (ARNv PB1 mc et ARNv M mc) (figure 31A). Les mutations, situées dans les régions codantes des deux segments, ont été conçues de manière à être silencieuses afin d'être testées par la suite dans un contexte infectieux. Les ARNv PB1 et M mutés interagissent tous deux faiblement avec leur ARNv partenaire sauvage (21% de complexe formé dans les deux cas) (figure 31B, pistes 6 et 7), tandis que la co-incubation des deux segments mutés restaure l'interaction à 83% par rapport à la situation sauvage (figure 31B, piste 8).

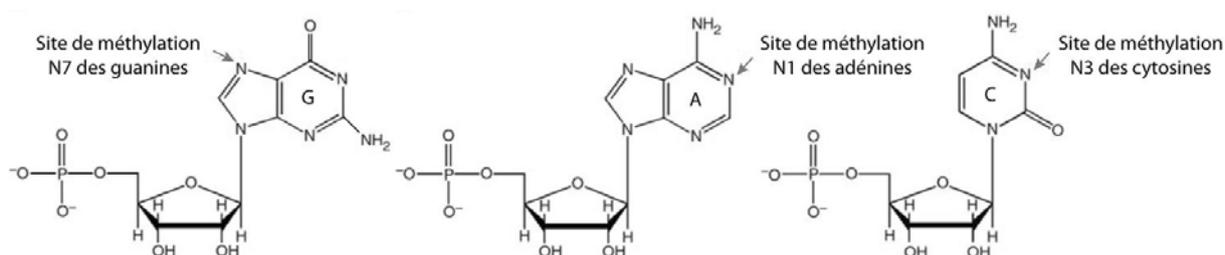


**Figure 31 : Impact de mutations *trans*-compensatoires au niveau des régions impliquées dans l'interaction entre les ARNv PB1 et M sur la formation du complexe.** **A.** Description des mutations *trans*-compensatoires et silencieuses introduites au niveau de la séquence d'interaction potentielle des ARNv PB1 et M (ARNv PB1 mc et ARNv M mc). **B.** Analyse des interactions ARN/ARN par électrophorèse sur gel d'agarose en conditions natives après co-incubation des ARNv PB1 et M sauvages et des ARNv PB1 et M portant des mutations *trans*-compensatoires. Les monomères d'ARNv sont indiqués par des flèches et les hétérodimères par un astérisque rouge. La quantification de la proportion de complexe présent dans chaque piste (moyenne de 4 expériences indépendantes  $\pm$  SEM) a été réalisée.

## b. Cartographie chimique en solution sur l'ARNv PB1

Les molécules d'ARN adoptent des structures secondaires et tertiaires qui sont principalement dues à l'établissement d'interactions hydrogènes (appariements canoniques Watson-Crick ou liaisons non canoniques) et à des interactions d'empilement. L'existence de ces interactions peut être détectée grâce à des sondes chimiques de structure, qui testent l'accessibilité de différents groupements fonctionnels impliqués dans le maintien des structures de l'ARN (pour revue, voir [327]). Ainsi, lorsque les nt ne sont pas appariés, certaines positions Watson-Crick des bases peuvent être modifiées par les sondes. À l'inverse, lorsque les bases sont appariées, ces positions ne sont plus accessibles aux modifications chimiques. Les modifications sont ensuite détectées par extension d'amorces radioactives ou fluorescentes par la transcriptase inverse, qui sera bloquée en 3' d'une position modifiée. Ces sondes chimiques permettent également de détecter l'existence d'appariements intermoléculaires entre deux ARN, à condition que les régions impliquées dans l'appariement dans les deux segments soient accessibles à la modification chimique en absence de leur partenaire [328].

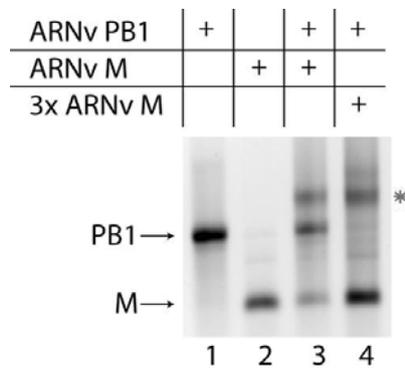
Nous avons appliqué la méthode de cartographie chimique en solution à l'ARNv PB1 afin de valider l'interaction avec son ARNv partenaire M. Nous avons principalement utilisé le DMS qui modifie les positions N7 des guanines, N1 des adénines et N3 des cytosines (figure 32).



**Figure 32 : Sites de modifications au DMS.** La modification par le DMS implique une méthylation des positions N7 des guanines, détectées après scission du double brin d'ARN, et des positions N1 des adénines et N3 des cytosines, détectées par extension d'amorce.

Afin de se placer dans des conditions statistiques de modifications, à savoir au maximum une modification par molécule d'ARNv, des gammes de concentration en DMS (concentrations finales de 0,021 M, 0,023 M et 0,026 M) et de temps de modification (1, 3 et 5 min) ont été réalisées (résultats non présentés). Des conditions de modifications

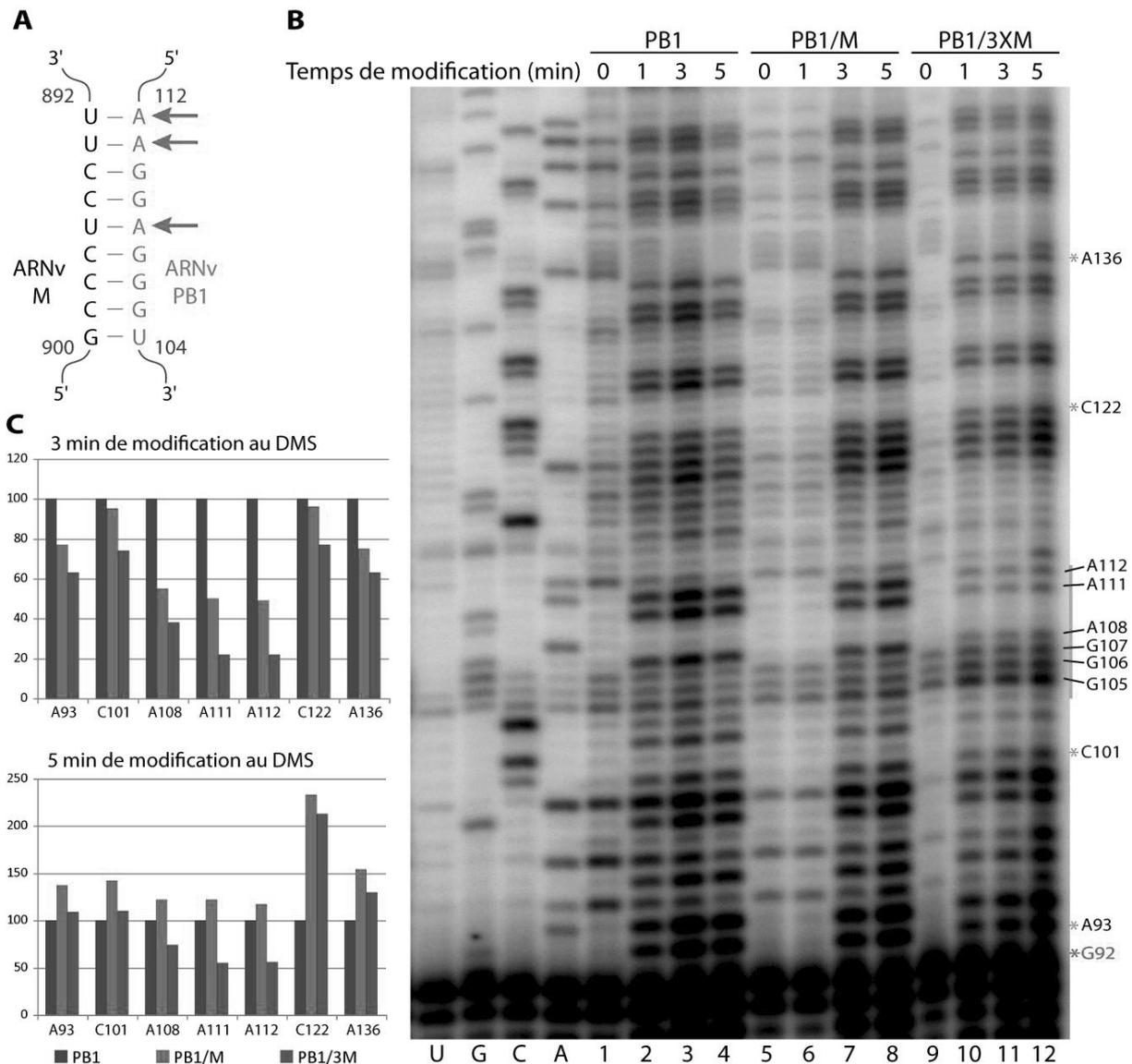
≈ statistiques ont été obtenues pour une concentration en DMS de 0,023 M. Pour éviter de cartographier une population hétérogène d'ARNv (monomérique et hétérodimérique), nous avons co-incubé l'ARNv PB1 en présence d'un excès molaire de trois fois du segment M, ce qui nous assurait de déplacer la totalité de l'ARNv PB1 vers la forme complexée (figure 33).



**Figure 33 : Complexes entre les ARNv PB1 et M visualisés sur gel d'agarose en conditions natives.** L'ARNv PB1 est incubé en présence d'une quantité équimolaire ou d'un excès molaire de trois fois de son partenaire M. Les monomères d'ARNv sont indiqués par des flèches et les hétérodimères par un astérisque rouge.

Au début de la région d'interaction présumée du segment PB1 avec l'ARNv M, une pause au niveau des nt 105-GGG-107 est nettement visible (figure 34B), ce qui laisse à penser que cette région est très structurée dans l'ARNv libre. De plus, cette pause s'intensifie en présence de l'ARNv M, suggérant qu'une interaction stable se forme entre les trois G du segment PB1 et les trois C en positions 897-899 du segment M (figure 34A). Au-delà de cette pause, le DMS modifie les adénines en positions 108, 111 et 112 (figure 34A). Lorsque l'ARNv PB1 est co-incubé avec son partenaire, ces trois nt devraient être moins réactifs, car protégés de la méthylation en raison de leur interaction avec l'ARNv M. La quantification du gel (figure 34B) a permis de représenter l'intensité de modification de certains nt, qu'ils soient dans la zone d'interaction putative ou en dehors (figure 34C). Les trois adénines 108, 111 et 112 de l'ARNv PB1 sont protégées en présence du partenaire M, qu'il soit présent en quantité équimolaire (figure 34C, histogrammes rouges) ou en excès (figure 34C, histogrammes verts). Ceci est particulièrement visible pour des temps d'incubation de 3 min, puisque les réactivités de ces adénines diminuent lorsque la quantité d'ARNv M augmente, pour respectivement atteindre 38%, 22% et 22% par rapport aux mêmes nt en absence du partenaire (figure 34C, graphique du haut). Cependant, la protection est plus faible après 5 min de modification au DMS (figure 34C, graphique du bas). À l'inverse, les réactivités des nt en dehors de cette zone d'interaction, servant de contrôles, sont peu affectées par l'ajout de l'ARNv M. En effet, après 3 min de modification au DMS, la diminution de la réactivité est bien plus faible que celle observée pour les trois

adénines 108, 111 et 112 (figure 34C, graphique du haut). Ces mêmes nt sont plus réactifs après 5 min de modification (figure 34C, graphique du bas).



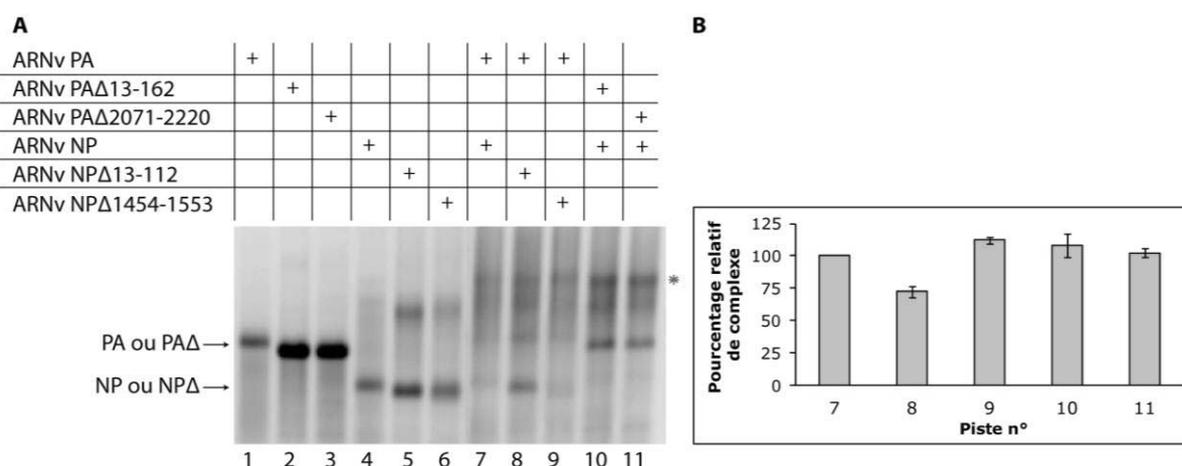
**Figure 34 : Cartographie chimique en solution sur une région de l'ARNv PB1 comprenant la séquence d'interaction potentielle avec l'ARNv M. A.** Interaction potentielle entre les ARNv PB1 et M. Les flèches rouges indiquent les cibles de la méthylation par le DMS sur l'ARNv PB1 sauvage. **B.** Électrophorèse sur gel de polyacrylamide dénaturant des ADNc obtenus par transcription inverse de l'ARNv PB1 seul, en présence d'une quantité équimolaire ou d'un excès de trois fois de l'ARNv M. Les pistes 1, 5 et 9 correspondent aux ARNv non modifiés. Dans les autres pistes, l'ARNv PB1 a été modifié au DMS pendant 1, 3 ou 5 min. U, G, C et A correspondent à un séquençage par la technique de Sanger réalisé en parallèle, sur l'ARNv PB1. La barre orange à droite du gel indique les nt impliqués dans l'interaction potentielle avec le segment M. **C.** Comparaison de l'intensité de modification des nt A108, A111 et A112 et de nt à l'extérieur de la région d'interaction de l'ARNv PB1 à l'aide du logiciel Image Gauge. L'intensité de modification des nt sur l'ARNv PB1 seul est ramenée à 100% (histogrammes bleus) et est comparée à celle des nt de l'ARNv PB1 en présence d'une quantité équimolaire (histogrammes rouges) ou d'un excès de trois fois (histogrammes vert) de l'ARNv M, après 3 et 5 min en présence de DMS. L'intensité de chaque modification sur une piste a été normalisée par rapport à la pause en position G92.

### c. Conclusion

À l'aide de l'approche « classique » *in vitro* précédemment développée au laboratoire et appliquée au cas de l'interaction entre les ARNv PB1 et M de la souche PR8, nous avons identifié la séquence, au nt près, responsable de l'hétérodimérisation entre les deux segments. La construction de mutants *trans*-compensatoires a ensuite permis de valider l'implication de ces nt dans la formation du complexe *in vitro*. Par cartographie chimique en solution nous avons montré qu'en présence de l'ARNv M, les adénines 108, 111 et 112 sont protégées de la méthylation et semblent donc interagir avec des nt du segment M. La cartographie chimique de l'ARNv partenaire M, ainsi que celles des mutants *trans*-compensatoires, permettra de confirmer cette hypothèse. À noter qu'avec Z. Wehbe (Master 2) nous avons testé d'autres sondes dans ce cadre (CMCT, kethoxal et sondes de SHAPE (*selective 2'-hydroxyl acylation analysed by primer extension*)). Bien qu'intéressants et encourageants, les résultats sont trop préliminaires pour être présentés ici.

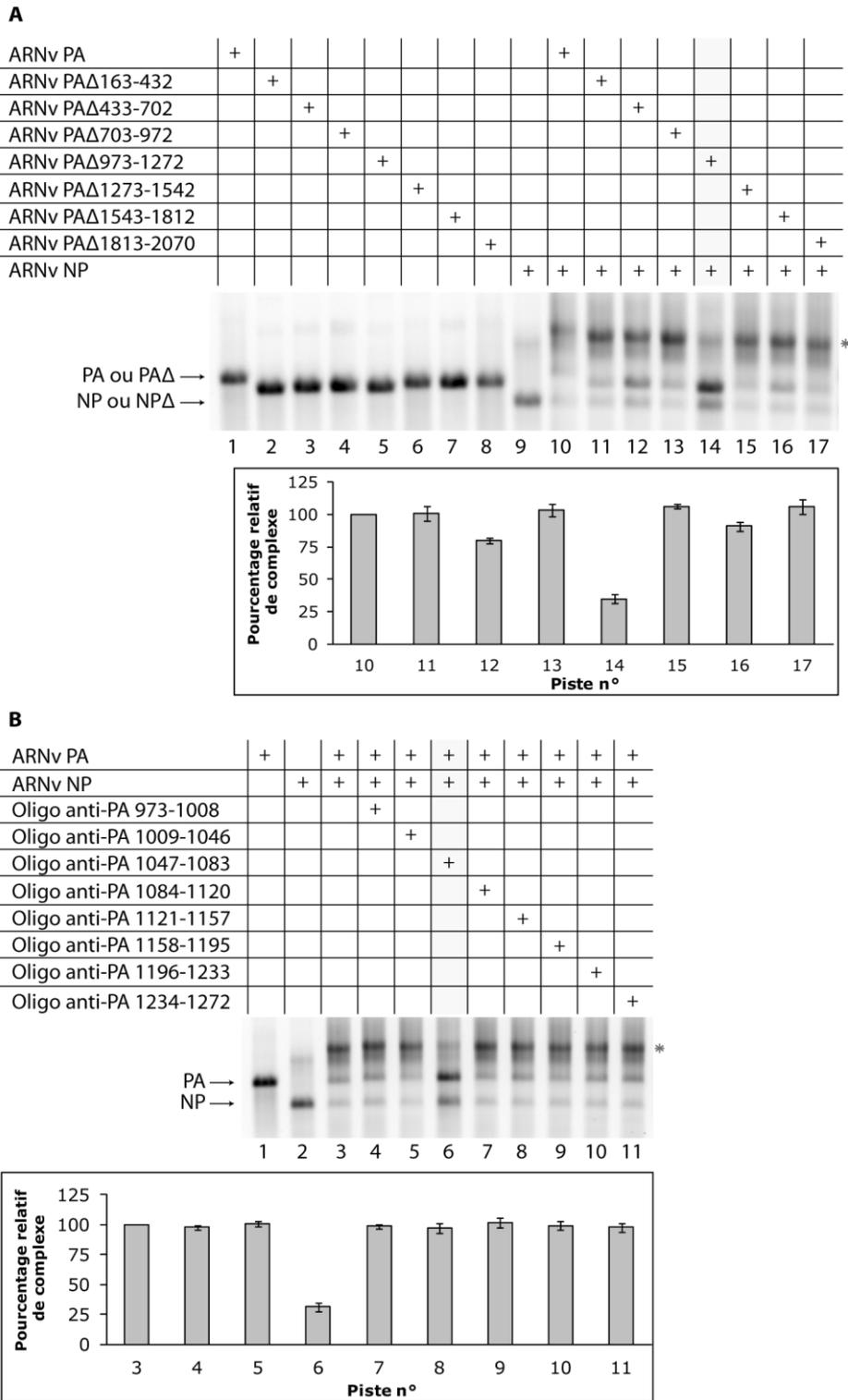
#### 1.2. Interaction entre les ARNv PA et NP

L'interaction entre les ARNv PA et NP est la plus forte des huit interactions du réseau de la souche PR8 (74% d'interaction) (figure 25B). Des délétions de 150 nt aux extrémités 3' ou 5' de l'ARNv PA n'ont pas révélé d'effet majeur sur l'hétérodimérisation (figure 35).



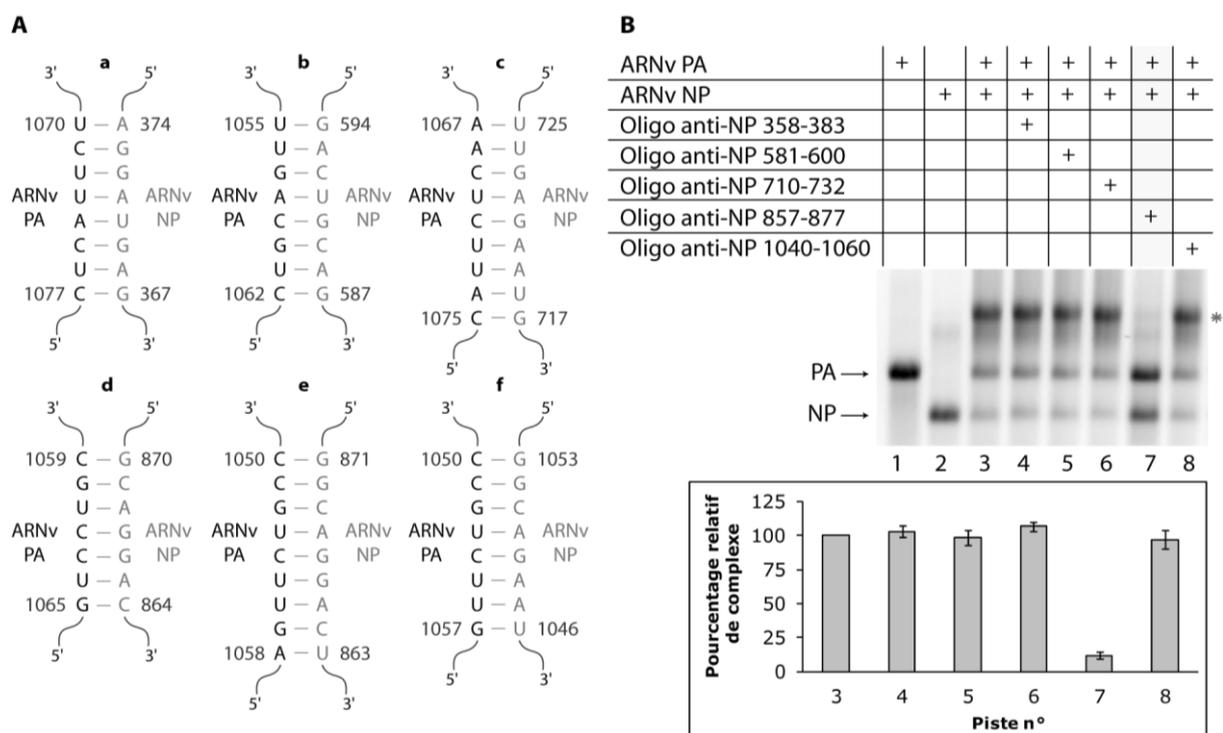
**Figure 35 : Effet de délétions terminales dans les ARNv PA et NP sur l'hétérodimérisation.** **A.** Analyse des interactions ARN/ARN par électrophorèse sur gel d'agarose en conditions natives après co-incubation des ARNv PA et NP sauvages et des ARNv portant de grandes délétions au niveau de leurs extrémités. Les monomères d'ARNv sont indiqués par des flèches et les hétérodimères par un astérisque rouge. **B.** Quantification de la proportion de complexe présent dans chaque piste (moyenne de 3 expériences indépendantes  $\pm$  SEM).

Nous avons par conséquent voulu tester si, comme dans le cas de la souche aviaire England [319], les séquences d'interaction sont situées dans des régions plus centrales des ARNv. De grandes délétions ont ainsi été introduites dans l'ARNv PA de manière à couvrir la totalité du segment. La délétion des nt 973-1272 de l'ARNv PA diminue l'interaction de 65% par rapport au complexe sauvage (figure 36A, piste 14), tandis que les autres mutations n'ont pas ou très peu d'impact sur l'interaction (figure 36A, pistes 11-13 et 15-17). Pour délimiter plus précisément la région du segment PA impliquée dans l'interaction avec l'ARNv NP, nous avons ensuite utilisé des oligonucléotides antisens complémentaires à la région 973-1272. Dans ces expériences, l'ajout de l'oligonucléotide qui cible les nt 1047-1083 inhibe l'hétérodimérisation de la même façon que le mutant de délétion (diminution de l'interaction de 69%) (figure 36B, piste 6).



**Figure 36 : Identification de la région de l'ARNv PA impliquée dans l'hétérodimérisation avec l'ARNv NP à l'aide d'ARNv PA délétés de manière systématique et d'oligonucléotides antisens complémentaires à l'ARNv PA.** Analyse des interactions ARN/ARN par électrophorèse sur gel d'agarose en conditions natives après (A) co-incubation de l'ARNv NP sauvage et des ARNv PA portant de grandes délétions systématiques ou après (B) co-incubation des ARNv PA et NP sauvages en présence d'oligonucléotides complémentaires à de courtes séquences de l'ARNv PA. Les monomères d'ARNv sont indiqués par des flèches et les hétérodimères par un astérisque rouge. La quantification de la proportion de complexe présent dans chaque piste (moyenne de 3 expériences indépendantes  $\pm$  SEM) a été réalisée.

Avec l'outil bioinformatique Google nous avons alors cherché une complémentarité antiparallèle entre les nt 1047-1083 du segment PA et la totalité de l'ARNv NP. Parmi le grand nombre d'interactions proposées, nous en avons sélectionné six qui respectaient nos critères de sélection (voir § A.II.1.1.a) (figure 37A). Nous avons testé leur existence, *in vitro*, en ajoutant des oligonucléotides antisens complémentaires au segment NP (voir Matériels et Méthodes, tableau 6). L'interaction entre les deux partenaires est fortement diminuée en présence de l'oligonucléotide ciblant les nt 857-877 de l'ARNv NP (12% de complexe restant) (figure 37B, piste 7), tandis que les autres oligonucléotides n'inhibent pas la formation du complexe (figure 37B, pistes 4-6 et 8).

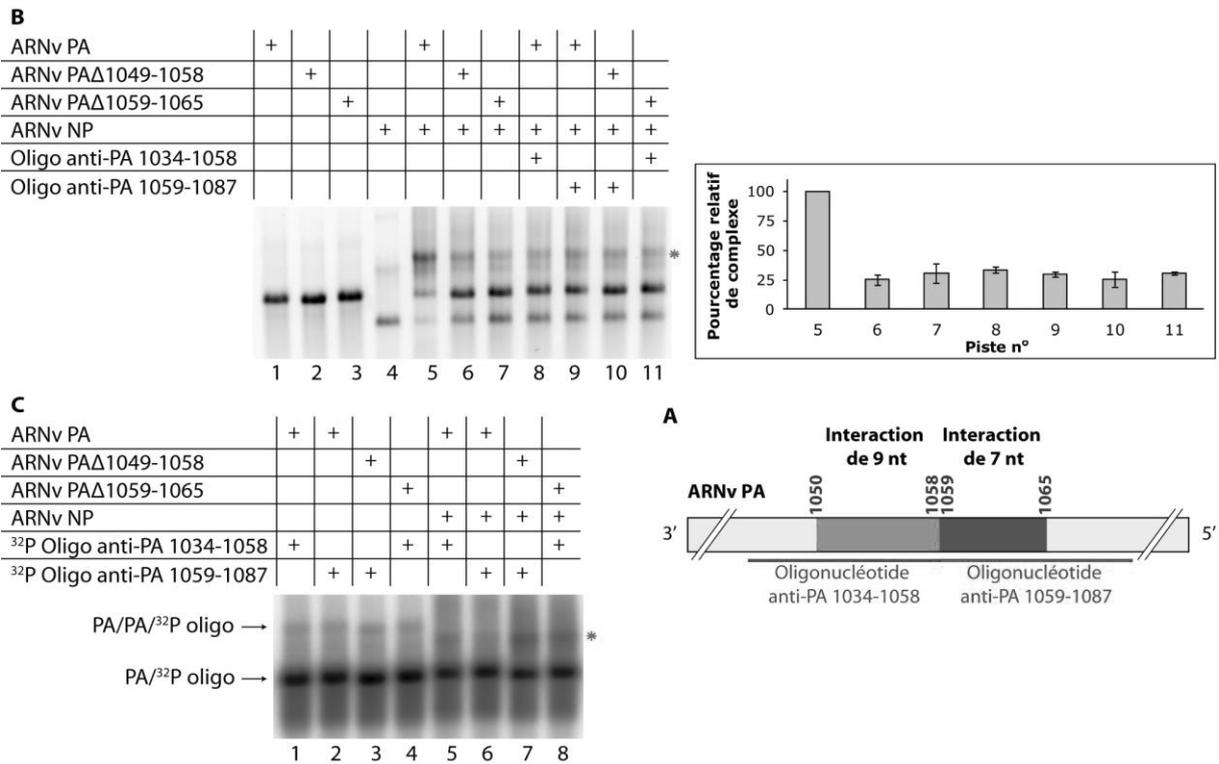


**Figure 37 : Identification de la région de l'ARNv NP impliquée dans l'hétérodimérisation avec l'ARNv PA à l'aide d'oligonucléotides antisens complémentaires à l'ARNv NP. A.** Interactions potentielles entre les ARNv PA et NP proposées par l'outil bioinformatique Google. **B.** Analyse des interactions ARN/ARN par électrophorèse sur gel d'agarose en conditions natives après co-incubation des ARNv PA et NP sauvages en présence d'oligonucléotides complémentaires à de courtes séquences de l'ARNv NP. Les monomères d'ARNv sont indiqués par des flèches et les hétérodimères par un astérisque rouge. La quantification de la proportion de complexe présent dans chaque piste (moyenne de 6 expériences indépendantes  $\pm$  SEM) a été réalisée.

La recherche bioinformatique de complémentarité antiparallèle entre les nt 1047-1083 de l'ARNv PA et 857-877 de l'ARNv NP a révélé l'existence de deux séquences potentielles d'interactions, de sept (figure 37A, interaction d) et neuf pb (figure 37A,

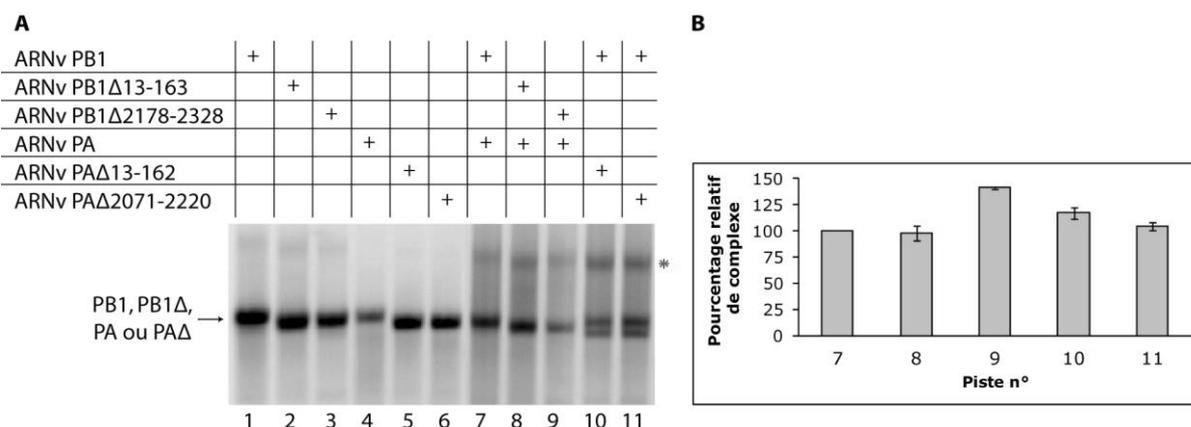
interaction e). Les deux interactions font intervenir les mêmes nt de l'ARNv NP, tandis que les nt de l'ARNv PA sont contigus. L'ajout d'oligonucléotides antisens complémentaires à l'une ou l'autre de ces deux séquences (oligonucléotides anti-PA 1034-1058 et anti-PA 1059-1087) ou l'utilisation d'ARNv PA délétés (PA $\Delta$ 1049-1058 et PA $\Delta$ 1059-1065) diminuent la formation du complexe entre 67% et 75% (figure 38A & B, pistes 6-9). Ces résultats ne permettent donc pas de favoriser l'existence de l'une ou l'autre des deux interactions. Nous avons ensuite co-incubé l'ARNv NP sauvage avec chaque ARNv PA portant une délétion et en présence d'un oligonucléotide complémentaire à la seconde zone d'interaction potentielle sur l'ARNv PA (figure 38A & B, pistes 10 et 11). À nouveau, il reste 25% et 31% de complexe dans ces conditions. Nous avons vérifié que les deux oligonucléotides antisens s'hybrident aussi bien à l'ARNv PA sauvage qu'aux ARNv délétés des nt 1049-1058 ou des nt 1059-1087, en absence et en présence du partenaire NP, en les marquant radioactivement et en testant leur hybridation par retard de migration sur gel d'agarose en conditions natives (Figure 38C).

Nous avons par conséquent identifié les régions impliquées dans l'interaction entre les segments PA et NP, sans toutefois réussir à définir précisément quels nt sont responsables de l'hétérodimérisation. En effet, nos travaux ont mis en évidence l'existence de deux séquences d'interaction possibles, qui font intervenir les mêmes nt de l'ARNv NP et des nt contigus sur l'ARNv PA (figure 37A, interactions d et e). Il serait envisageable que les deux séquences sur l'ARNv PA puissent être impliquées, alternativement, dans la formation de l'hétérodimère, permettant ainsi à l'une de prendre le relais si l'autre est altérée. Cependant, dans ce cas, nous nous attendions à observer une inhibition quasi-totale de la formation du complexe lorsqu'un des ARNv PA mutants est co-incubé avec l'oligonucléotide antisens ciblant la seconde séquence d'interaction potentielle (figure 38B). Cette hypothèse reste donc encore à vérifier. Nous pourrions, comme dans le cas de l'interaction entre les ARNv PB1 et M (voir § A.II.1.1.b), réaliser une cartographie chimique en solution sur l'ARNv PA afin d'identifier les nt protégés de la modification en présence de l'ARNv NP. En parallèle, nous pourrions tester l'existence de ces deux interactions potentielles à l'aide d'ARNv PA et NP portant des mutations *trans*-compensatoires.



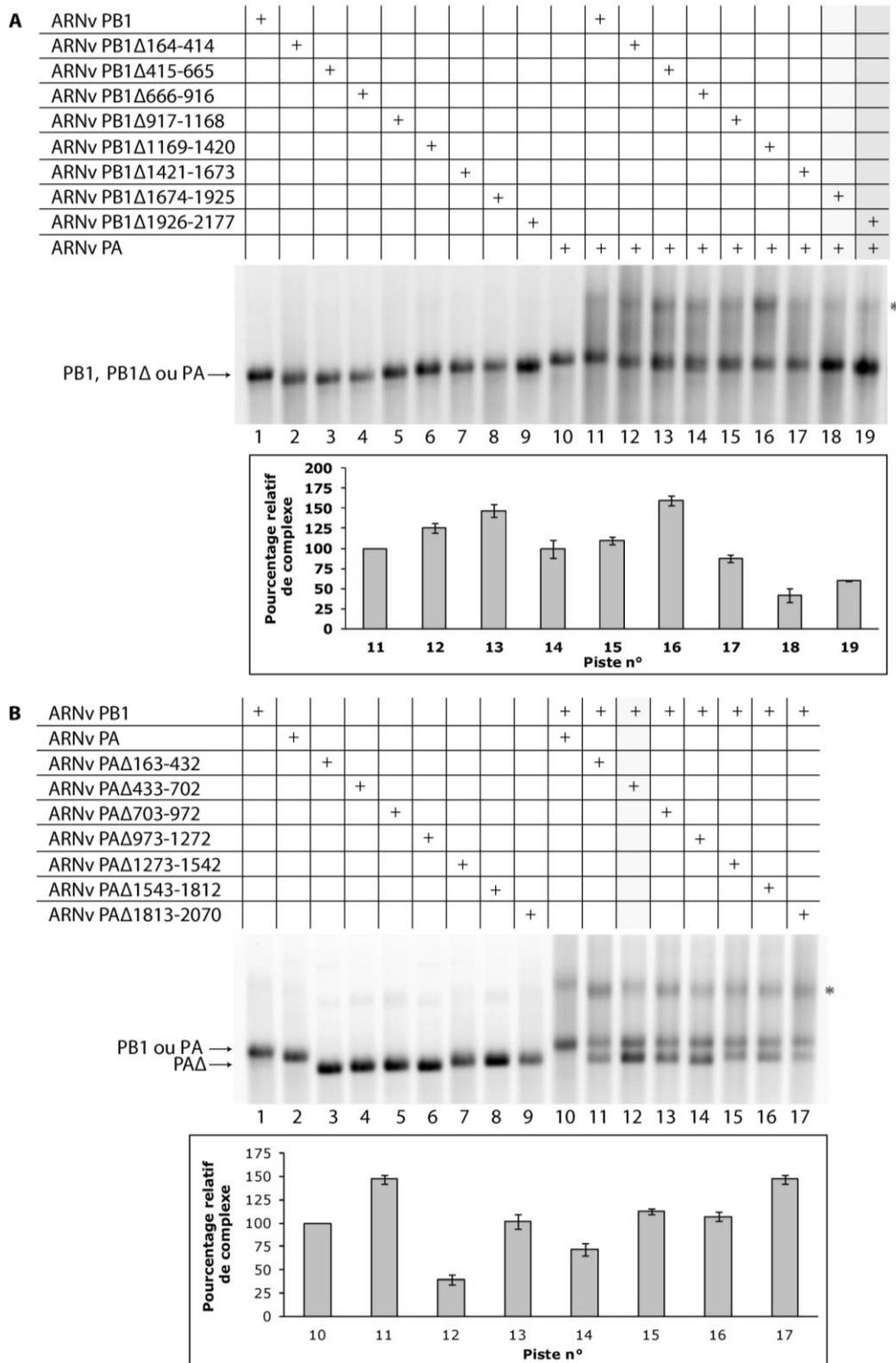
### 1.3. Interaction entre les ARNv PB1 et PA

L'interaction entre les segments PB1 et PA a également été étudiée dans le but d'identifier les nt impliqués dans l'hétérodimérisation. Des délétions de 150 nt aux extrémités 3' ou 5' des deux partenaires n'ont pas eu d'impact sur la formation du complexe (figure 39).



**Figure 39 : Effet de délétions terminales dans les ARNv PB1 et PA sur l'hétérodimérisation.** **A.** Analyse des interactions ARN/ARN par électrophorèse sur gel d'agarose en conditions natives après co-incubation des ARNv PB1 et PA sauvages et des ARNv portant de grandes délétions au niveau de leurs extrémités. Les monomères d'ARNv sont indiqués par une flèche et les hétérodimères par un astérisque rouge. **B.** Quantification de la proportion de complexe présent dans chaque piste (moyenne de 4 expériences indépendantes  $\pm$  SEM).

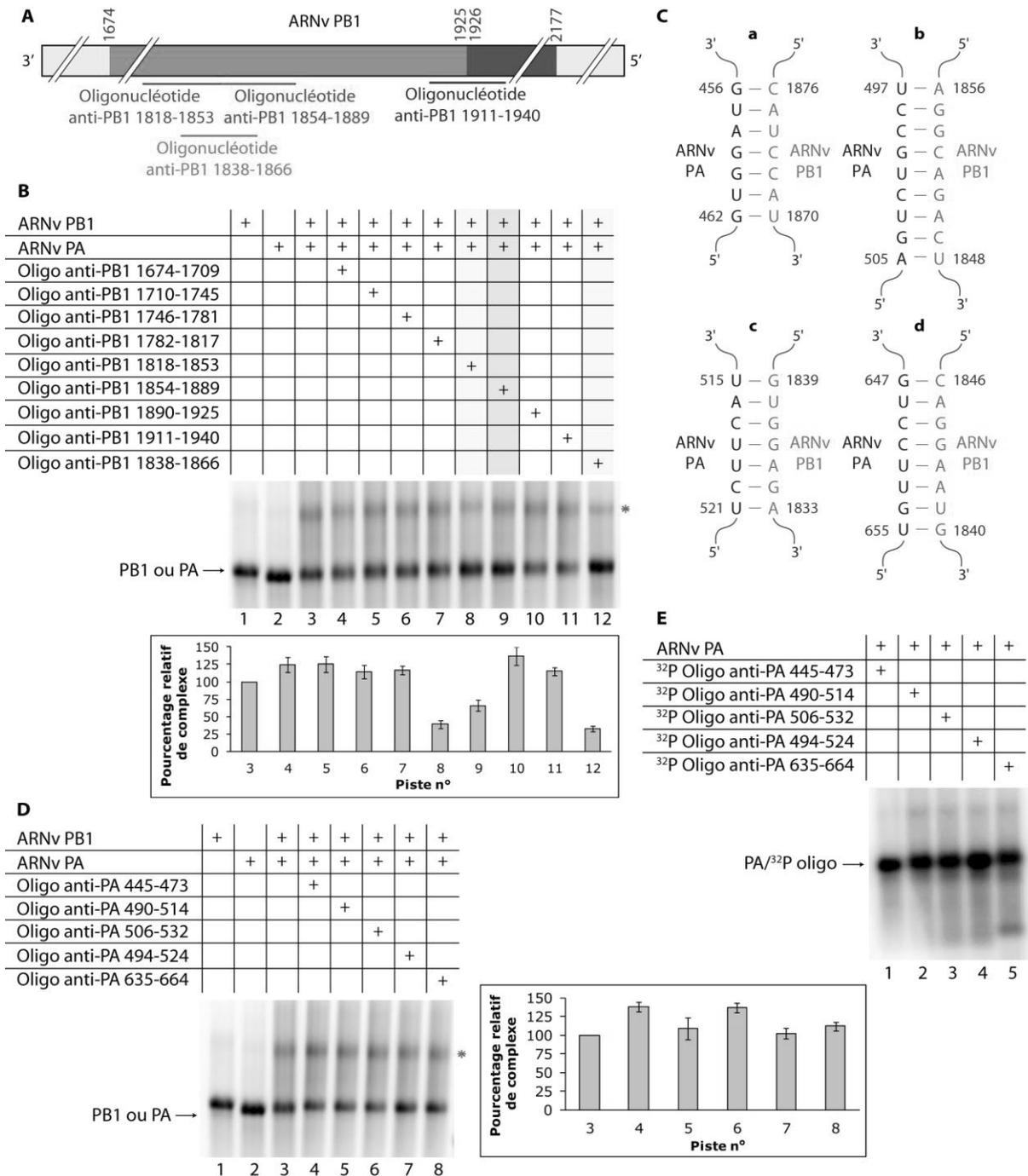
Par la suite, nous avons construit des ARNv PB1 et PA portant de grandes délétions internes et testé leurs effets sur la formation du complexe en présence de l'ARNv partenaire sauvage. Les nt 1674-1925 du segment PB1 et les nt 433-702 du segment PA sont nécessaires à l'interaction (respectivement 42% et 39% de complexe restant) (figure 40A, piste 18 et figure B, piste 12). L'hétérodimérisation est également affectée lors de la délétion des nt 1926-2177 de l'ARNv PB1, mais dans une moindre mesure (60% de complexe restant) (figure 40A, piste 19).



**Figure 40 : Identification des régions des ARNv PB1 et PA impliquée dans l'hétérodimérisation à l'aide d'ARNv délétés de manière systématique.** Analyse des interactions ARN/ARN par électrophorèse sur gel d'agarose en conditions natives (A) après co-incubation de l'ARNv PA sauvage et des ARNv PB1 portant de grandes délétions systématiques ou (B) après co-incubation de l'ARNv PB1 sauvage et des ARNv PA portant de grandes délétions systématiques. Les monomères d'ARNv sont indiqués par des flèches et les hétérodimères par un astérisque rouge. La quantification de la proportion de complexe présent dans chaque piste (moyenne de 4 à 5 expériences indépendantes  $\pm$  SEM) a été réalisée.

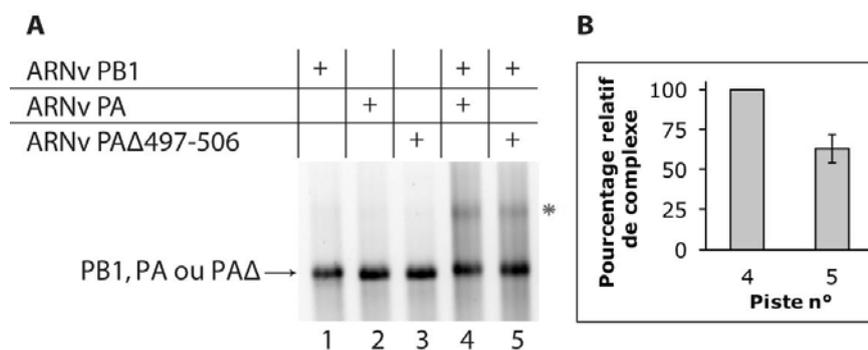
Ensuite, nous avons vérifié si l'ajout d'oligonucléotides antisens complémentaires à la région 1674-1925 de l'ARNv PB1 a un impact sur l'interaction entre les segments PB1 et PA (voir Matériels et Méthodes, tableau 6). Étant donné que la délétion de la région située en aval (nt 1926-2177) diminue également l'hétérodimérisation, nous avons voulu nous assurer que la séquence d'interaction n'est pas localisée de part et d'autre du nt 1925. L'oligonucléotide antisens qui cible les nt 1911-1940 du segment PB1 (figure 41A) n'inhibe pas la formation du complexe entre les ARNv PB1 et PA (115% de complexe) (figure 41B, piste 11). En revanche, la diminution de l'interaction en présence de l'oligonucléotide antisens complémentaire aux nt 1818-1853 de l'ARNv PB1 est similaire à celle obtenue en présence de l'ARNv PB1 dépourvu des nt 1674-1925 (39% de complexe restant) (figure 41B, piste 8). Dans une moindre mesure, l'oligonucléotide complémentaire aux nt 1854-1889 du segment PB1 diminue également l'interaction (66% de complexe restant) (figure 41B, piste 9). Pour vérifier si la séquence d'interaction chevauche éventuellement ces deux régions, nous avons ajouté un oligonucléotide qui cible les nt 1838-1866 du segment PB1 (figure 41A). La formation du complexe est inhibée de la même façon qu'avec l'oligonucléotide complémentaire à la région 1818-1853 (33% de complexe restant) (figure 41B, piste 12). L'interaction avec le segment PA semble par conséquent faire intervenir les nt 1838-1866 de l'ARNv PB1. L'inhibition de l'hétérodimérisation en présence de l'ARNv PB1 dépourvu de la région 1926-2177 (60% de complexe restant) (figure 40A, piste 19) pourrait ainsi être dû à un changement conformationnel de cet ARNv muté, qui affecte la structure de la région 1838-1866.

La recherche, par bioinformatique, de complémentarités antiparallèles entre les nt 1838-1866 de l'ARNv PB1 et les nt 433-702 de l'ARNv PA a mis en évidence plusieurs interactions potentielles. Après avoir appliqué nos critères de sélection (voir § A.II.1.1.a), nous avons testé l'existence de quatre d'entre elles (figure 41C) à l'aide d'oligonucléotides antisens complémentaires au segment PA (voir Matériels et Méthodes, tableau 6). Cependant, aucun de ces oligonucléotides n'a eu d'effet sur la formation du complexe entre les segments PB1 et PA (figure 41D), bien que leur hybridation au segment PA ait été validée à l'aide d'oligonucléotides marqués radioactivement (figure 41E).



**Figure 41 : Identification des régions des ARNv PB1 et PA impliquées dans l'hétérodimerisation à l'aide d'oligonucléotides antisens complémentaires aux deux ARNv.** **A.** Représentation schématique de l'ARNv PB1 et des oligonucléotides antisens utilisés. Le schéma n'est pas réalisé à l'échelle. **B & D.** Analyse des interactions ARN/ARN par électrophorèse sur gel d'agarose en conditions natives après co-incubation des ARNv PB1 et PA sauvages en présence d'oligonucléotides antisens complémentaires à de courtes séquences de **(B)** l'ARNv PB1 ou de **(D)** l'ARNv PA. Les monomères d'ARNv sont indiqués par une flèche et les hétérodimères par un astérisque rouge. La quantification de la proportion de complexe présent dans chaque piste (moyenne de 3 à 7 expériences indépendantes  $\pm$  SEM) a été réalisée. **C.** Interactions potentielles entre les ARNv PB1 et PA proposées par l'outil bioinformatique Google. **E.** Hybridation spécifique des oligonucléotides antisens complémentaires à l'ARNv PA. Chaque oligonucléotide, marqué radioactivement en 5' au <sup>32</sup>P, a été incubé en présence de l'ARNv PA dans les mêmes conditions que celles utilisées pour tester l'interaction entre les ARNv, afin de vérifier son hybridation spécifique au segment PA.

Parmi les interactions proposées par Google, celle entre les nt 497-505 de l'ARNv PA et les nt 1848-1856 de l'ARNv PB1 nous semblait pourtant être un bon candidat (figure 41C, interaction b). En effet, cette interaction est constituée de neuf appariements canoniques Watson-Crick, dont la moitié sont des appariements G-C, alors que les autres interactions putatives présentent davantage d'appariements Wobble et sont plus courtes. Les deux oligonucléotides complémentaires au segment PA ciblant ces nt (oligonucléotides anti-PA 490-514 et anti-PA 494-524) n'ayant pas d'impact sur la formation du complexe (figure 41D, pistes 5 et 7), nous avons construit un ARNv PA dépourvu des nt 497-506, potentiellement impliqués dans l'interaction avec le segment PB1 (figure 41C, interaction b). L'hétérodimérisation entre les ARNv PB1 et PA n'est que modérément affectée par la présence de cette délétion dans l'ARNv PA (63% de complexe restant par rapport au complexe formé entre les ARNv sauvages) (figure 42).



**Figure 42 : Effet de la délétion des nt 497-506 de l'ARNv PA sur l'hétérodimérisation avec l'ARNv PB1.** **A.** Analyse des interactions ARN/ARN par électrophorèse sur gel d'agarose en conditions natives après co-incubation des ARNv PB1 et PA sauvages et de l'ARNv PA dépourvu des nt 497-506. Les monomères d'ARNv sont indiqués par une flèche et les hétérodimères par un astérisque rouge. **B.** Quantification de la proportion de complexe présent dans chaque piste (moyenne de 2 expériences indépendantes  $\pm$  SEM).

Ainsi, bien que nous ayons délimité les régions responsables de l'interaction entre les segments PB1 et PA, nous n'avons à ce jour pas identifié précisément quels sont les nt des deux partenaires impliqués dans la formation du complexe. Il semblerait que l'interaction fasse intervenir les nt 497-505 de l'ARNv PB1 et les nt 497-506 de l'ARNv PA. La délétion de ces nt sur l'ARNv PA n'est cependant pas suffisante pour inhiber l'hétérodimérisation de manière importante. Une hypothèse serait que d'autres nt, situés à proximité ou dans d'autres régions de l'ARNv PA, interviennent également dans la formation de ce complexe.

#### 1.4. Autres interactions du réseau

Les résultats de l'identification des régions impliquées dans les huit interactions qui s'établissent entre les ARNv de la souche PR8 sont résumés dans la figure 43. Les expériences concernant les cinq interactions qui n'ont pas été détaillées (entre les ARNv PB2 et HA, PB2 et NP, PB2 et NS, PB1 et HA, PB1 et NP) ont permis d'identifier la région impliquée dans l'hétérodimérisation sur l'un des deux segments (sauf dans le cas de l'interaction entre les ARNv PB2 et NP), tandis que nos tentatives pour identifier la séquence complémentaire sur le second partenaire sont restées infructueuses. Plusieurs types de difficultés ont été rencontrés : i) l'utilisation d'ARNv portant de grandes délétions n'a pas toujours permis d'identifier la région responsable de la formation du complexe, ii) l'utilisation d'oligonucléotides antisens complémentaires à certaines régions des ARNv, elles-mêmes délimitées à l'aide d'ARNv mutés, n'a pas toujours inhibé la formation des complexes hétérodimériques, iii) les résultats obtenus n'ont pas toujours été reproductibles.

- **Interaction entre les segments PB2 et HA** : la région du segment HA impliquée dans l'interaction a été délimitée à l'aide d'ARNv mutés et d'oligonucléotides antisens, tandis que la majorité des délétions systématiques de l'ARNv PB2 ont inhibé la formation du complexe (30%-60% de complexe restant par rapport à l'interaction sauvage) et n'ont donc pas permis de privilégier une région en particulier.

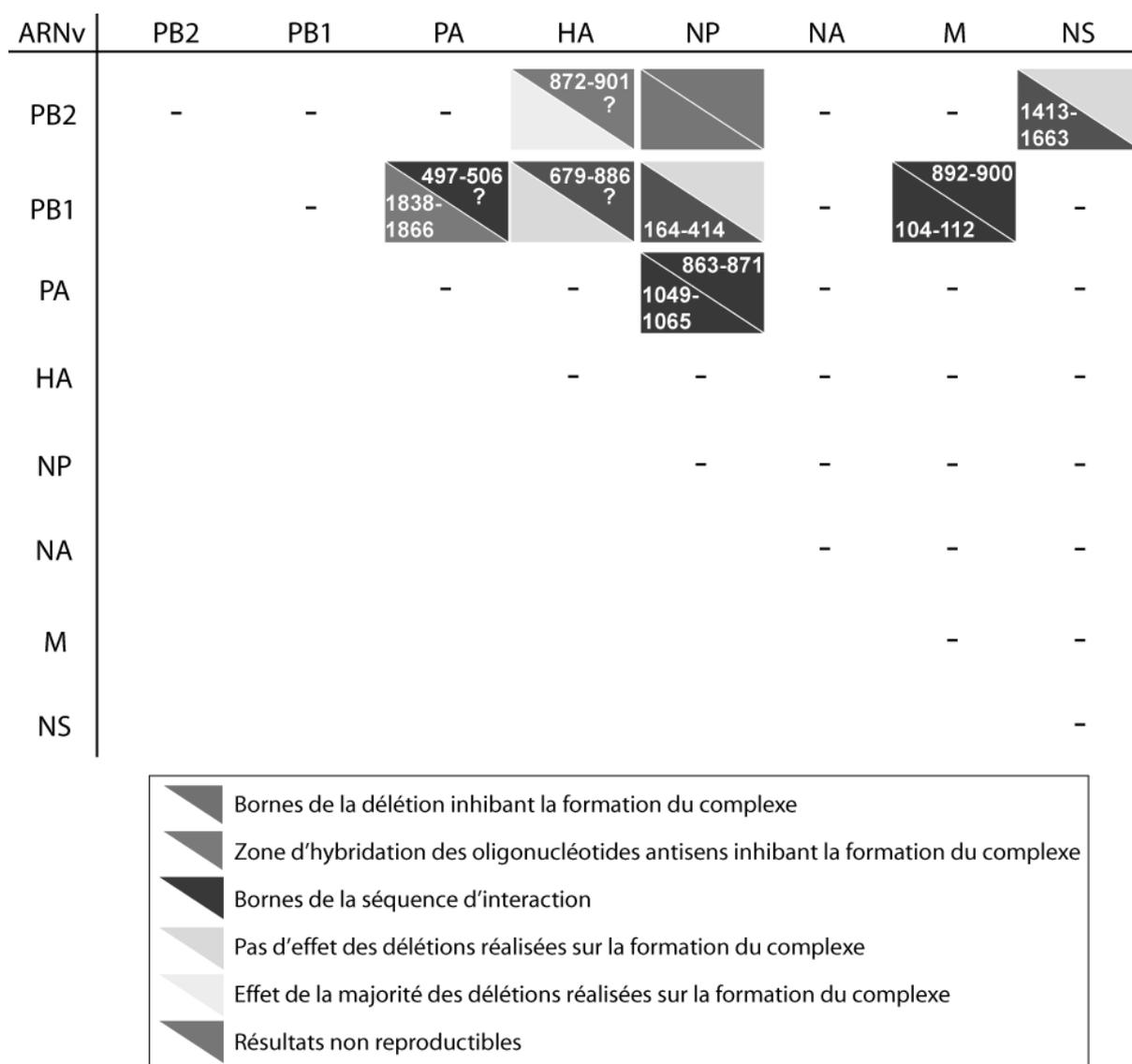
- **Interaction entre les segments PB2 et NP** : les résultats ont été peu reproductibles, que ce soit lors de l'utilisation d'ARNv NP et PB2 mutés ou d'oligonucléotides antisens complémentaires au segment PB2, tandis que ceux complémentaires au segment NP n'ont pas eu d'effet.

- **Interaction entre les segments PB2 et NS** : la région d'interaction de l'ARNv PB2 a été identifiée à l'aide d'ARNv portant de grandes délétions, mais les oligonucléotides antisens n'ont pas inhibé l'hétérodimérisation. Dans ce cas, le test systématique de leur efficacité d'hybridation à l'ARNv PB2 devra être réalisé. De la même manière, les délétions terminales de l'ARNv NS n'ont pas eu d'effet. Nous allons donc introduire des délétions systématiques dans ce segment afin d'évaluer leur impact sur la formation du complexe.

- **Interaction entre les segments PB1 et HA** : aucune des délétions systématiques introduites dans l'ARNv PB1 n'a inhibé la formation de l'hétérodimère, tandis que la région d'interaction a été identifiée sur l'ARNv HA à l'aide d'ARNv mutés. La région n'a cependant pas été délimitée plus précisément car les oligonucléotides antisens complémentaires à l'ARNv HA n'ont pas eu d'effet sur la formation du complexe. Il faudra par conséquent vérifier leur efficacité d'hybridation au segment HA.

- **Interaction entre les segments PB1 et NP** : la région d'interaction de l'ARNv PB1 a été déterminée à l'aide d'ARNv mutés, tandis que les oligonucléotides antisens n'ont pas eu d'impact sur la formation du complexe. Par ailleurs, les délétions terminales du segment NP n'ont pas eu d'effet. Nous allons introduire des délétions systématiques dans cet ARNv afin d'évaluer leur impact sur la formation du complexe.

Pour conclure, cette approche « classique » a permis d'identifier un certain nombre de partenariats entre les ARNv de la souche d'IVA PR8, mais peu ont été délimités au nt près (figure 43). Son utilisation au laboratoire, sur les souches Moscou, England et PR8, montre donc qu'elle n'est pas toujours couronnée de succès.



**Figure 43 : Résumé des régions des ARNv de la souche PR8 impliquées dans les interactions intermoléculaires identifiées à 55°C dans un tampon de faible salinité.** Les positions des nt des ARNv sont données dans le sens 3' vers 5'. Certains résultats, marqués du symbole « ? », restent à confirmer.

## 2. Approche MIME *in vitro*

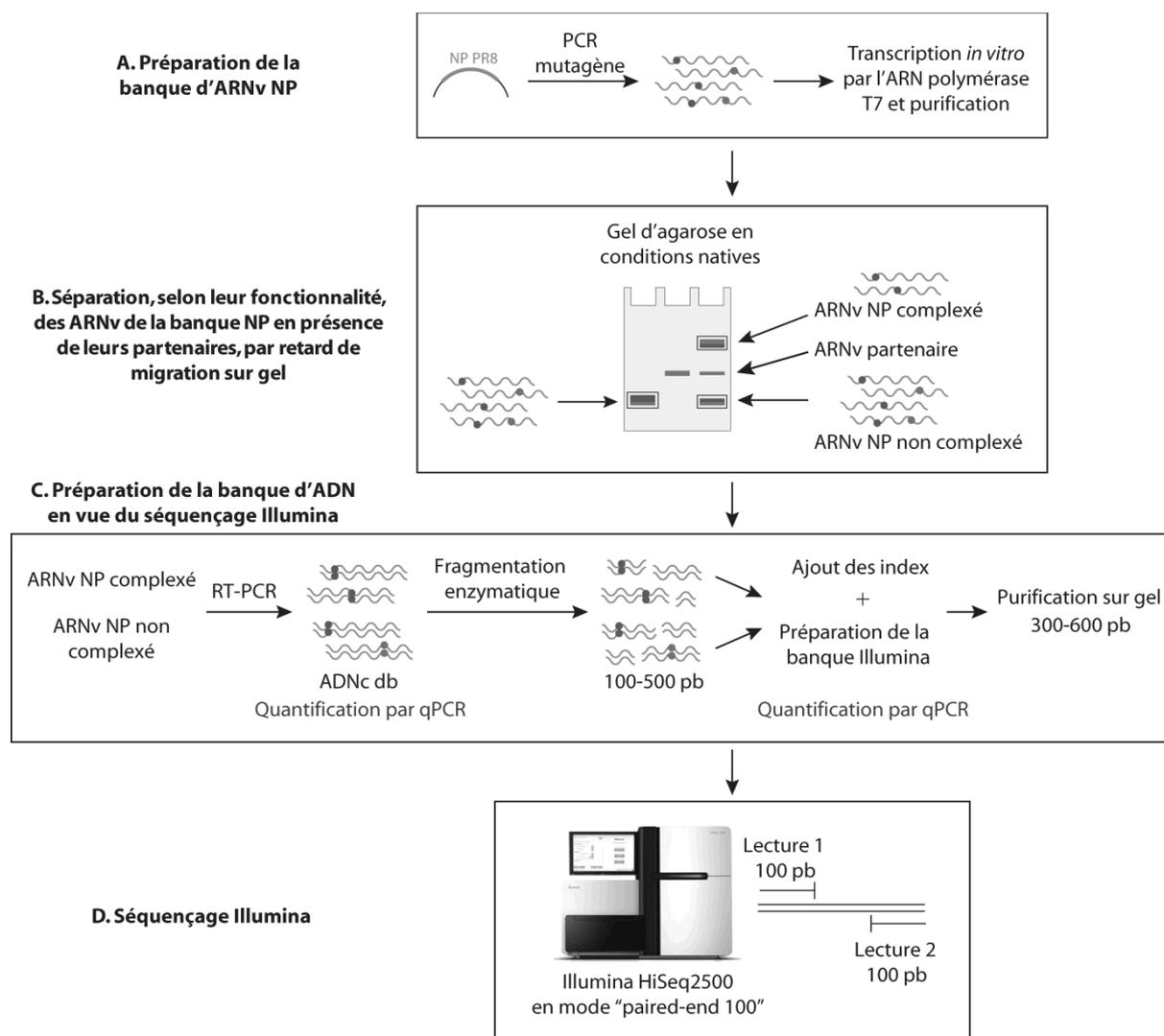
La stratégie *in vitro* développée au laboratoire et appliquée ici à la souche PR8, avait été validée *in vivo* sur une interaction ARN/ARN d'un virus aviaire England (H5N2) [273, 319]. Comme nous venons de le voir, cette approche est néanmoins longue et fastidieuse : elle nécessite la construction d'un grand nombre de mutants de délétion afin d'identifier, dans un premier temps, sur chacun des partenaires, une vaste région (de 100 à 250 nt) impliquée dans les interactions. Cette région doit ensuite être affinée, pour délimiter la séquence, au nt près, responsable de l'interaction. De plus, à ce jour, peu d'informations sont disponibles

concernant la structure secondaire des ARNv génomiques et les méthodes de prédictions de structure secondaire sont difficilement applicables à des ARN aussi longs. De ce fait, l'introduction à l'aveugle de délétions dans un ARNv n'est pas toujours neutre quant à son effet sur la structure secondaire de celui-ci. Ainsi, l'impact de la délétion sur la formation du complexe peut parfois être dû à la modification de la structure de l'ARNv, plutôt qu'à l'absence des nt responsables de l'hétérodimérisation.

Il serait par conséquent avantageux de disposer d'une méthode permettant de déterminer de manière exhaustive, efficace et non-biaisée, tous les nt des ARNv impliqués dans les interactions ARN/ARN. Une telle stratégie, nommée MIME pour « Mutational Interference Mapping Experiment », a récemment été développée au laboratoire par le Dr R. Smyth afin d'identifier les séquences et les structures importantes pour la reconnaissance du génome du VIH-1 par la polyprotéine virale Pr55<sup>Gag</sup> impliquée dans la formation des particules virales lors du bourgeonnement [277]. L'approche MIME est basée sur l'introduction de mutations aléatoires au sein d'un ARN cible, la sélection des ARN mutés selon leur fonctionnalité et la détermination de la fréquence de chaque mutation à chaque position de l'ARN dans les populations d'ARN fonctionnels et non fonctionnels par séquençage à haut débit.

Dans le but d'identifier par cette méthode les nt impliqués dans la formation de complexes entre les ARNv de la souche PR8, nous nous sommes intéressés au segment NP, qui possède trois partenaires dans le réseau d'interactions *in vitro* que nous avons précédemment établi (figure 25B). Deux de ces interactions sont parmi les plus fortes du réseau (48% de complexe avec l'ARNv PB2 et 74% de complexe avec l'ARNv PA). De plus, parmi tous les ARNv des IVA, le segment NP présente une taille « intermédiaire », ce qui nous semblait avantageux dans le cadre de la mise au point et de la validation de l'approche MIME pour la mise en évidence d'interactions ARN/ARN, avant de l'appliquer aux segments d'ARNv plus grands.

## 2.1. Mise en place du protocole MIME



**Figure 44 : Schéma des principales étapes du protocole MIME *in vitro* pour l'identification des nt de l'ARNv NP impliqués dans l'interaction avec ses ARNv partenaires.** Les positions nucléotidiques mutées qui sont impliquées dans l'interaction sont représentées en rouge et celles qui ne sont pas importantes pour l'hétérodimérisation sont représentées en mauve. Se référer au texte pour l'explication de chaque étape.

### a. Préparation de la banque d'ARNv NP (figure 44A)

En vue de l'expression de la banque d'ARNv NP, nous avons introduit des mutations ponctuelles aléatoires tout le long de l'ADN codant pour l'ARNv NP à l'aide de l'ADN polymérase mutazyme II commerciale. D'après le fournisseur, cette enzyme génère un spectre de mutations moins biaisé que celui obtenu avec d'autres protocoles standards de mutagenèse aléatoire. En effet, les taux de mutations des purines et des pyrimidines sont équivalents, tandis que d'autres enzymes mutent préférentiellement les purines, comme

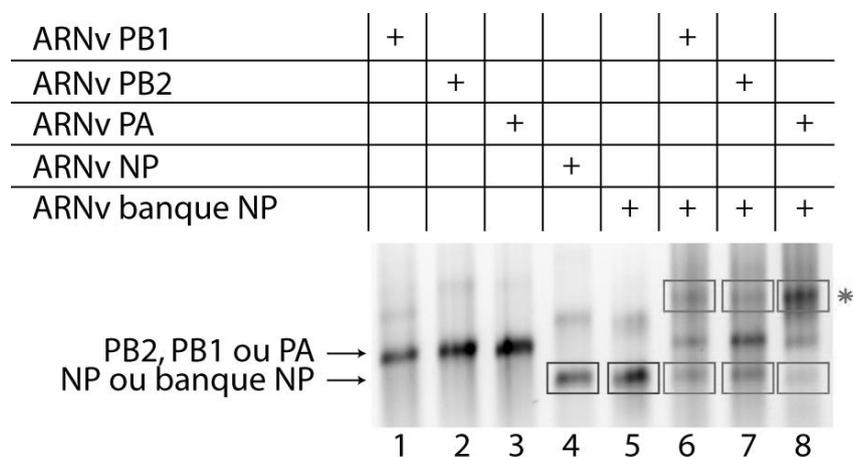
l'ADN polymérase mutazyme I, ou les pyrimidines, comme l'ADN polymérase Taq. Le taux de mutations peut être modulé en variant la quantité initiale d'ADN matrice, le nombre de cycles de PCR effectués (généralement de 20 à 30 cycles) ou le nombre de PCR successives.

Deux à trois PCR mutagènes successives ont été nécessaires dans l'étude de Smyth *et al.* [277] afin d'obtenir un taux de mutations moyen de 0,6%. Dans le cas de la banque d'ADN NP, nous souhaitons obtenir un taux de mutations compris entre 0,5% et 1%. Par conséquent, deux PCR successives de mutagenèse aléatoire ont été réalisées (voir Matériels et Méthodes, § D.I.1). Afin de moduler le taux de mutations introduites lors de la seconde PCR mutagène, nous avons effectué deux réactions en parallèle, la seule différence étant la quantité d'ADN NP matrice provenant de la première PCR mutagène (1 µl de la *banque 1* a été utilisé pour obtenir la *banque 2a* et 0,5 µl pour la *banque 2b*). Ainsi, dans le cas de la *banque 2b*, nous nous attendions à potentiellement augmenter le nombre de mutations introduites dans chaque molécule d'ADN NP. Afin de déterminer le taux de mutations des *banques 2a* et *2b*, nous avons cloné une fraction de chacune des deux banques dans le vecteur pJET, puis séquencé 17 et 10 clones différents, respectivement. Le taux de mutations moyen obtenu pour chaque banque est approximativement de 0,5% (soit en moyenne huit mutations sur l'ensemble du segment d'ADN NP). La quantité de matrice utilisée lors de la seconde PCR mutagène n'a donc pas influencé le nombre de mutations introduites. La même observation avait été faite dans le cas des banques utilisées au laboratoire dans l'étude sur le VIH-1 [277]. À ce jour, nous n'avons pas trouvé d'explication à ce phénomène. Afin d'obtenir la banque la plus diverse possible, les banques *2a* et *2b* ont été rassemblées pour constituer la banque d'ADN NP finale sur laquelle la suite du protocole a été réalisée.

#### **b. Séparation des ARNv de la banque NP par retard de migration sur gel d'agarose (figure 44B)**

La banque d'ARNv NP, après transcription *in vitro* et purification sur gel d'agarose, a été co-incubée avec les ARNv partenaires PB2, PB1 et PA, dans les conditions expérimentales choisies pour l'établissement du réseau d'interactions entre les ARNv de la souche PR8 (voir Matériels et Méthodes, § B.IV) (figure 45). La migration des échantillons sur un gel d'agarose en conditions natives a permis de séparer les différentes populations d'ARNv selon leur fonctionnalité. En effet, les segments qui sont mutés à des positions nucléotidiques

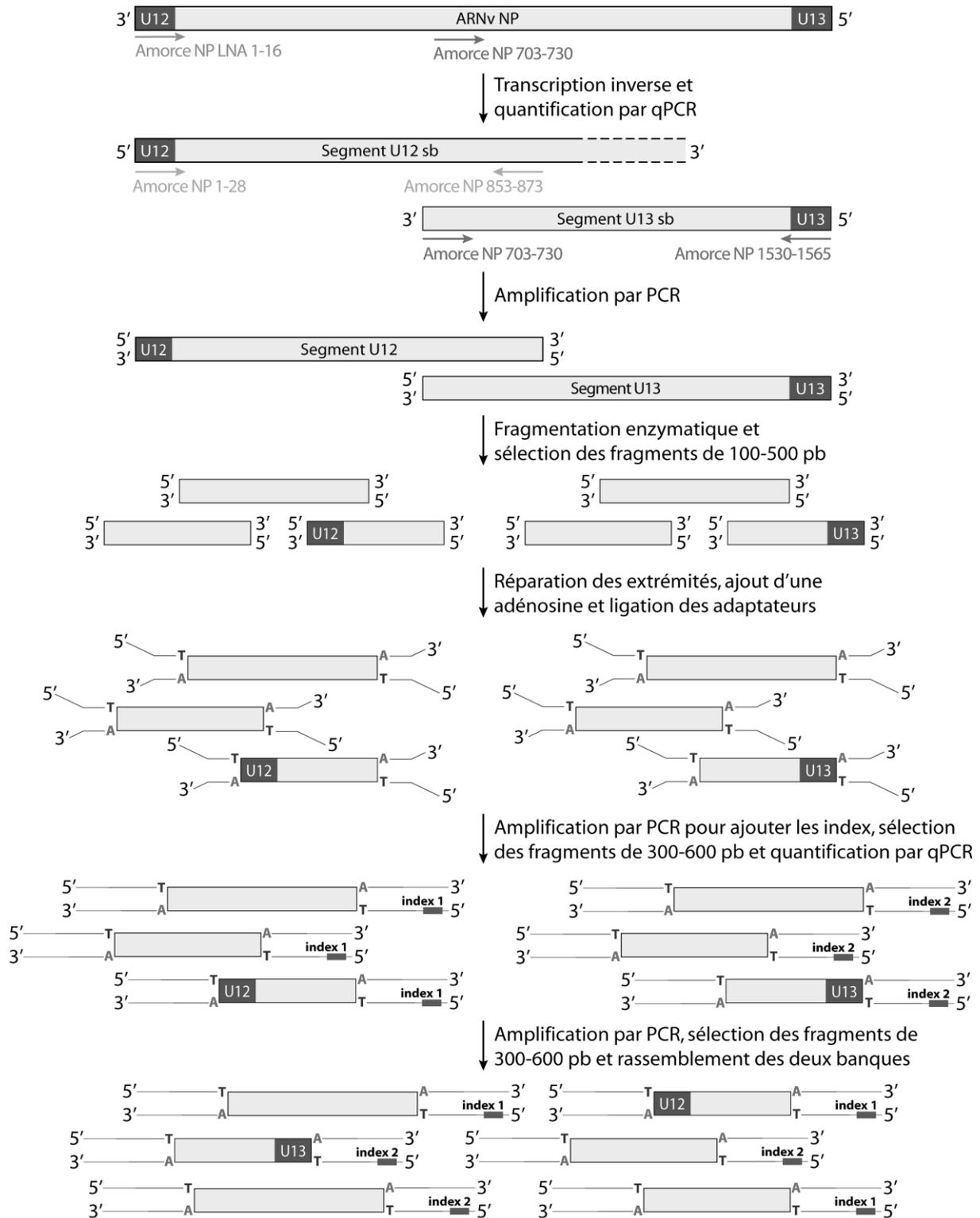
impliquées dans l'hétérodimérisation avec l'ARNv partenaire ne sont plus capables d'interagir et se trouvent donc dans la population d'ARNv NP non complexés. À l'inverse, les ARNv dont les mutations n'ont aucun effet sur l'interaction entre ARNv se trouvent à la fois dans les populations d'ARNv libres et complexés. Après migration, nous avons récupéré les bandes correspondant aux ARNv NP non complexés et aux ARNv NP engagés dans un complexe avec les ARNv partenaires PB2, PB1 et PA (figure 45, rectangles oranges). L'ARNv NP sauvage et la banque d'ARNv NP servent de contrôle et ont été traités en parallèle des autres échantillons (figure 45, rectangles mauves). L'ARNv NP sauvage permet d'identifier les mutations introduites lors des étapes de préparation de la banque et du séquençage Illumina et la banque d'ARNv NP permet de déterminer le taux de mutations obtenu après la seconde PCR mutagène de manière plus précise. En effet, cette valeur sera calculée à partir de l'ensemble des molécules de la banque NP (voir § A.II.2.2), tandis que la valeur obtenue précédemment avait été calculée à partir des résultats du séquençage de 27 clones (voir § A.II.2.1.a). De plus, elle permet également de déterminer la fréquence de chacune des trois mutations possibles à chaque position de l'ARNv (voir § A.II.2.2).



**Figure 45 : Séparation fonctionnelle des ARNv NP capables ou non d'interagir avec leur partenaires PB2, PB1 ou PA par électrophorèse sur gel d'agarose en conditions natives après co-incubation des ARNv PB2, PB1 et PA sauvages et de la banque d'ARNv NP.** Les monomères d'ARNv sont indiqués par une flèche et les hétérodimères par un astérisque rouge. Les bandes d'ARNv récupérées du gel qui correspondent aux contrôles et aux échantillons sont encadrées en mauve et en orange, respectivement.

### c. Préparation de la banque en vue d'un séquençage Illumina (figure 44C)

Les six échantillons d'ARNv NP et les deux contrôles, élués du gel (figure 45), ont ensuite été rétrotranscrits par la RT d'AMV (figure 46). L'utilisation d'une amorce complémentaire aux seize premiers nt de l'ARNv NP a révélé que seule une faible proportion d'ADNc était de pleine longueur (résultats non présentés). Afin d'optimiser l'étape de rétrotranscription, nous avons augmenté la stabilité thermique de cette amorce en ajoutant des nt modifiés (LNA) [329, 330]. Pour ce faire, nous avons utilisé deux amorces, contenant chacune trois LNA à des positions différentes. En parallèle, nous avons testé la RT TGIRT<sup>®</sup>-III (*thermostable group II intron reverse transcriptase*, Ingex) qui est plus processive, fidèle et thermostable en comparaison des RT rétrovirales [331]. Cependant, ces optimisations n'ont pas permis d'augmenter significativement la proportion d'ADNc de pleine longueur (résultats non présentés). Nous avons par conséquent opté pour l'utilisation d'une de ces amorces modifiées et d'une seconde amorce, complémentaire aux nt 703-730 (voir Matériels et Méthodes, tableau 7). La réalisation de deux réactions en parallèle permet ainsi d'obtenir deux segments d'ADNc sb correspondant aux « moitiés » 3' et 5' de l'ARNv NP. Pour plus de clarté, ces deux segments sont respectivement appelés « segments U12 sb » et « segments U13 sb ».



**Figure 46 : Schéma des principales étapes de la préparation de la banque en vue d'un séquençage Illumina.** Le schéma n'est pas réalisé à l'échelle. Se référer au texte pour l'explication de chaque étape.

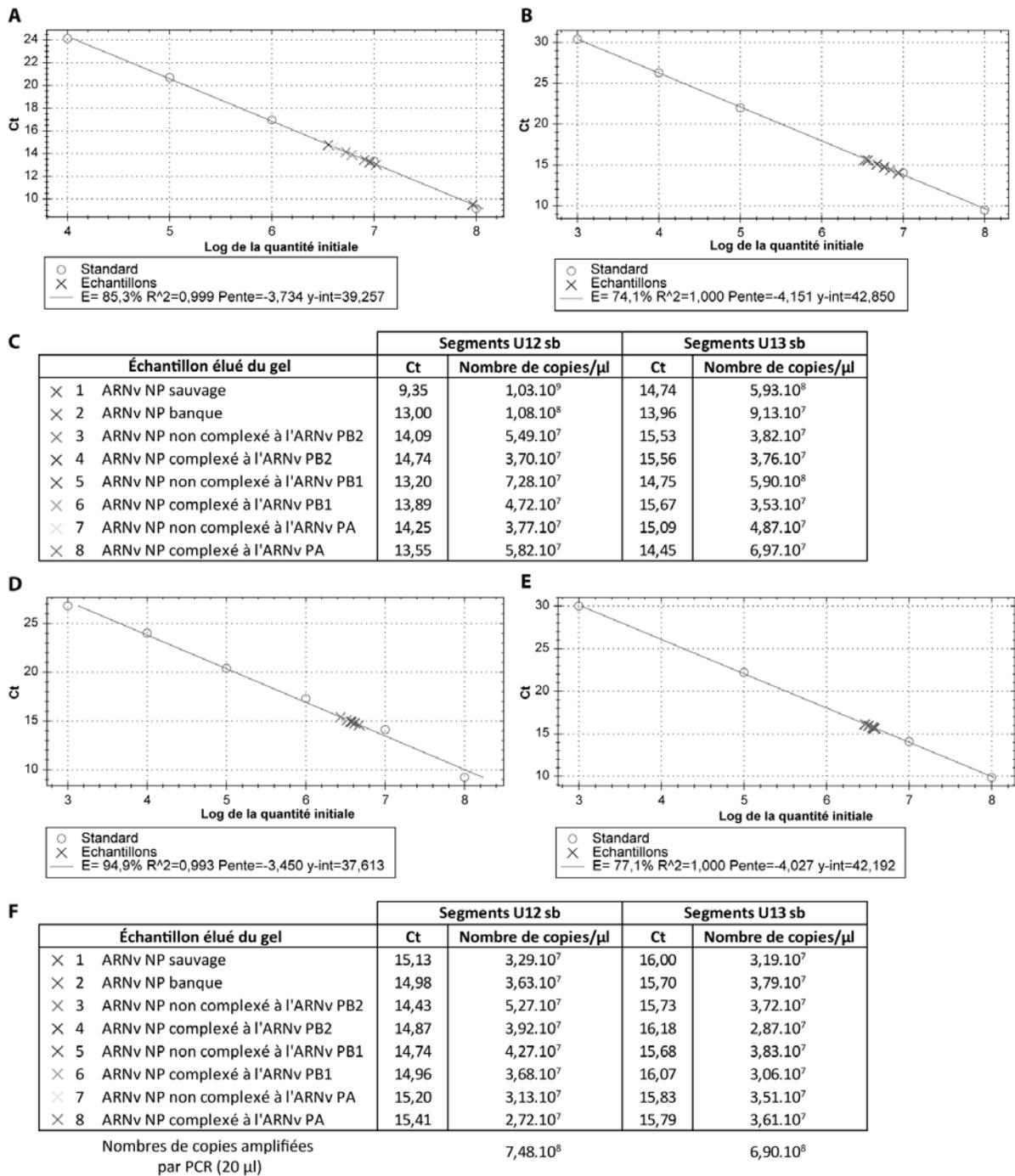
Ces échantillons ont ensuite été quantifiés par qPCR (figure 47A-C). La normalisation des segments U12 sb d'une part, et U13 sb d'autre part (figure 47D-F), est nécessaire afin de

poursuivre chacune des analyses avec approximativement le même nombre de copies de chaque « moitié » du segment NP.

L'amplification par PCR des huit segments U12 sb a été réalisée à partir de  $7,5 \cdot 10^8$  copies de chaque échantillon, tandis que l'amplification des échantillons U13 sb a été réalisée à partir de  $6,9 \cdot 10^8$  copies de chaque échantillon (figure 47F). Afin de ne pas perdre d'information, notamment au niveau des nt situés à la fin du segment U12 et au début du segment U13, les deux segments d'ADNc db obtenus se chevauchent sur 170 nt (figure 46). Ainsi, les segments U12 et U13 correspondent respectivement aux nt 1-873 et 703-1565 de l'ARNv NP.

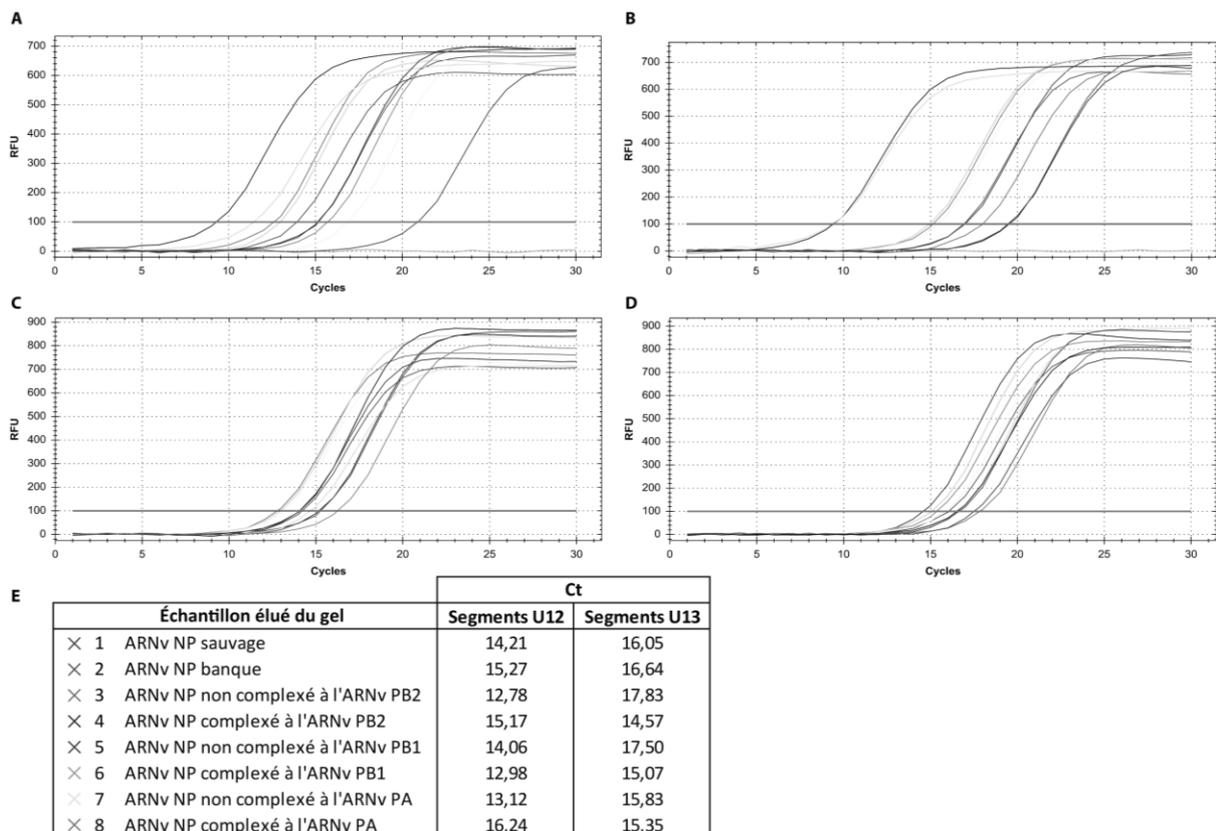
La technologie de séquençage Illumina HiSeq2500 permet de séquencer 100 pb à chacune des extrémités des molécules, alors que les segments U12 et U13 ont une taille de 875 et 862 pb, respectivement. Afin de s'affranchir de cette limitation, il est nécessaire de fragmenter les échantillons, de manière aléatoire, avant la préparation de la banque. Dans notre cas, les ADNc ont été fragmentés à l'aide du kit NEBNext dsDNA Fragmentase (NEB). Ce kit contient deux enzymes, la première générant des brèches dans l'ADN db, reconnues par la seconde enzyme qui coupe l'ADNc au niveau du brin opposé, créant ainsi des coupures db 5'-phosphate et 3'-OH. Nous avons ensuite récupéré uniquement les fragments d'ADNc db dont la taille est comprise entre 100 et 500 pb. Ainsi, il sera possible d'obtenir des données de séquençage pour chacun des nt des segments U12 et U13 [277].

La préparation de la banque Illumina comporte ensuite différentes étapes (figure 46) : i) les extrémités provenant de l'étape de fragmentation enzymatique sont réparées afin d'obtenir des bouts francs, ii) une adénosine est ajoutée en 3' de chaque brin, permettant ainsi iii) la ligation des adaptateurs grâce au T sortant présent à leurs extrémités 3'. Par la suite, une étape d'amplification par PCR est nécessaire pour ajouter des index à chaque échantillon. En effet, la technologie Illumina permet de séquencer plusieurs échantillons dans la même piste de séquençage en identifiant chacun d'eux à l'aide d'un index composé de six nt (tableau 9). Après la PCR, il est essentiel de se débarrasser de la majorité des adaptateurs qui n'ont pas été ligués aux ADN, qui seraient préférentiellement lus au moment du séquençage, en sélectionnant les fragments de 300 à 600 pb.



**Figure 47 : Quantification par qPCR des échantillons d'ADNc NP correspondant aux segments U12 sb et U13 sb obtenus après rétrotranscription des ARNv NP séparés sur gel d'agarose, avant (A-C) et après normalisation (D-F).** La droite standard (en rouge) a été obtenue à partir de  $10^3$  à  $10^8$  copies d'ARNv NP rétrotranscrit avec les amorces RT NP 2 (A & D) ou NP 703-730 (B & E). **A.** Quantification des segments U12 sb, dilués au  $10^{\text{ème}}$ , obtenus après rétrotranscription des échantillons 1-8 avec l'amorce RT NP 2. **B.** Quantification des segments U13 sb, obtenus après rétrotranscription des échantillons 1-8 avec l'amorce NP 703-730. L'échantillon 1 est dilué au  $100^{\text{ème}}$ , les échantillons 2-8 au  $10^{\text{ème}}$ . **C.** Tableau récapitulatif indiquant les Ct (*cycle threshold*) et le nombre de copies/ $\mu$ l correspondant à chaque échantillon quantifié par qPCR. **D.** Quantification des échantillons normalisés U12, dilués au  $10^{\text{ème}}$ . **E.** Quantification des segments normalisés U13, dilués au  $10^{\text{ème}}$ . **F.** Tableau récapitulatif indiquant les Ct et le nombre de copies/ $\mu$ l correspondant à chaque échantillon normalisé quantifié par qPCR.

Les segments U12 et U13, auxquels ont été ajoutés les index, ont ensuite été quantifiés par qPCR et normalisés (figure 48), de manière à avoir, dans la banque finale, approximativement le même nombre de copies de chacun des échantillons. En effet, si un échantillon est sur-représenté dans la banque, alors, après séquençage, le nombre de lectures associées à cet échantillon sera également sur-représenté. Dans notre cas, nous avons environ 3,4 Ct de différence, soit un facteur dix entre le nombre de copies des huit échantillons présents dans chacune des deux banques (segments U12 ou U13) (figure 48E).



**Figure 48 : Quantification par qPCR des segments U12 et U13, avant (A & B) et après (C & D) normalisation des échantillons. A & B. Les contrôles sans matrice sont représentés en orange. E. Tableau récapitulatif indiquant les Ct de chaque échantillon normalisé quantifié par qPCR.**

Enfin, les segments U12 et U13 ont été amplifiés par PCR et les seize échantillons ont été rassemblés. Une dernière purification a été réalisée afin d'éliminer les adaptateurs libres. La banque finale, contenant des fragments de 300 à 600 pb, a ensuite été séquencée dans une piste de HiSeq2500 en mode « paired-end 100 » à la plateforme « Biopuces et séquençage » à l'IGBMC à Illkirch (figure 44D).

## 2.2. Analyse des données de séquençage

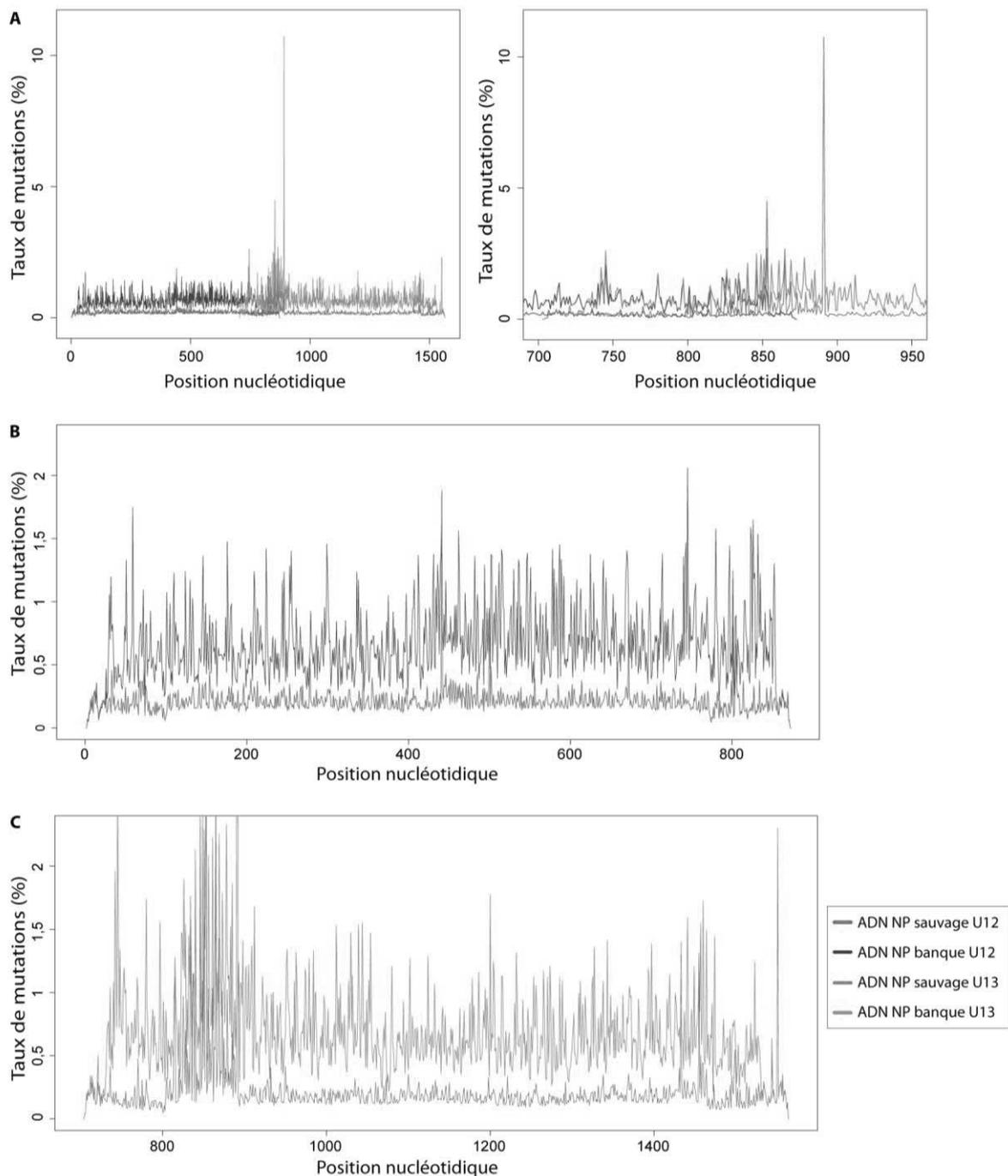
Le séquençage de la banque a généré 253 millions de lectures, dont 30% correspondent aux adaptateurs qui persistent malgré la réalisation de plusieurs étapes de purification sur gel d'agarose, tout au long du protocole de préparation de la banque. Il reste toutefois 176 millions de lectures correspondant aux échantillons, ce qui est amplement suffisant pour l'analyse subséquente. En effet, d'après cette valeur, chacun des nt de la banque séquencée a été lu en moyenne un million de fois.

Grâce aux index ajoutés lors de la préparation de la banque (voir Matériels et Méthodes, § D.III.7), nous avons associé chaque lecture à son échantillon initial élué du gel (tableau 9).

Index	Code	Échantillon élué du gel	Nombre de lectures	Nombre de lectures alignées	Nombre de bases analysées
1	CGATGT	ARNv NP sauvage - segment U12	4 599 182	4 317 234	96 814 945
2	ATCACG	ARNv NP banque - segment U12	5 370 351	5 275 897	272 644 110
3	TTAGGC	ARNv NP non complexé à l'ARNv PB2 - segment U12	6 364 736	6 255 168	341 341 515
4	TGACCA	ARNv NP complexé à l'ARNv PB2 - segment U12	2 888 286	2 685 311	149 677 129
5	ACAGTG	ARNv NP non complexé à l'ARNv PB1 - segment U12	5 644 195	5 326 637	226 178 422
6	GCCAAT	ARNv NP complexé à l'ARNv PB1 - segment U12	2 834 269	2 723 844	138 967 306
7	CAGATC	ARNv NP non complexé à l'ARNv PA - segment U12	4 189 693	3 681 904	233 379 379
8	ACTTGA	ARNv NP complexé à l'ARNv PA - segment U12	3 199 757	2 966 636	149 298 922
9	GGCTAC	ARNv NP sauvage - segment U13	3 783 198	3 491 389	153 136 586
10	CTTGTA	ARNv NP banque - segment U13	3 497 501	3 386 152	161 066 947
11	AGTCAA	ARNv NP non complexé à l'ARNv PB2 - segment U13	2 458 287	1 931 449	115 475 896
12	AGTTCC	ARNv NP complexé à l'ARNv PB2 - segment U13	1 118 570	736 430	41 203 250
13	ATGTCA	ARNv NP non complexé à l'ARNv PB1 - segment U13	2 768 564	2 475 844	92 861 807
14	CCGTCC	ARNv NP complexé à l'ARNv PB1 - segment U13	1 769 605	1 687 965	76 107 978
15	GTAGAG	ARNv NP non complexé à l'ARNv PA - segment U13	1 161 772	1 124 013	38 116 971
16	GTCCGC	ARNv NP complexé à l'ARNv PA - segment U13	3 032 948	2 909 910	145 189 752

**Tableau 9 : Nombre total de lectures et de nt analysés pour chacun des seize échantillons, identifiés avec un index de six nt.**

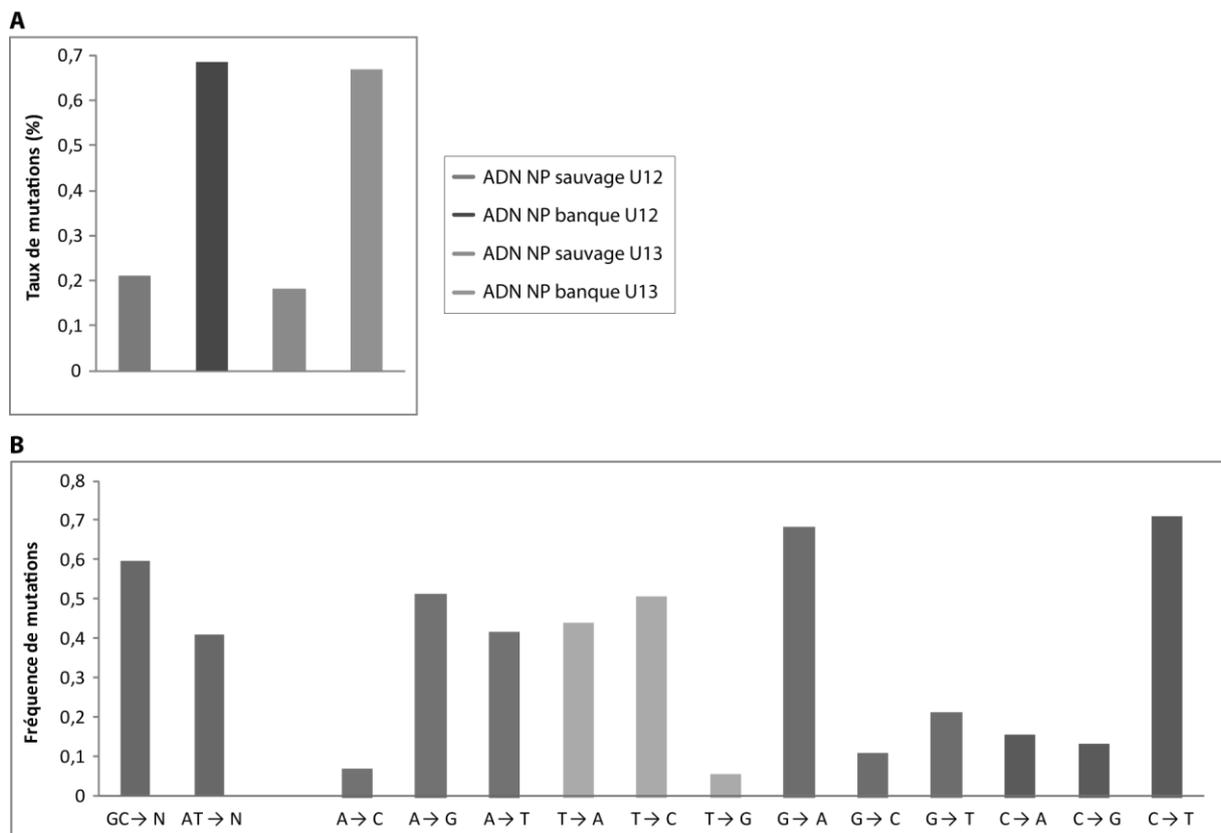
Dans un premier temps, le taux de mutations de chacun des 1565 nt a été représenté graphiquement pour les échantillons contrôles correspondant à l'ADN NP sauvage et à la banque d'ADN NP séquencés (figure 49). Il est à noter que ces ADN proviennent respectivement de l'ARNv NP sauvage et de la banque d'ARNv NP élués du gel (figure 45, rectangles mauves). Ce taux de mutations est homogène tout le long du segment U12 (figure 49A & B). Les données correspondant aux nt 820-900 du segment U13 sont plus surprenantes, puisqu'un pic important est retrouvé dans les deux contrôles. Néanmoins, ce pic n'est pas visible à la fin du segment U12, qui correspond aux nt 1-873 du segment NP (figure 49A). Par conséquent, il semblerait que ce problème vienne des étapes de transcription inverse ou de la préparation de la banque en vue du séquençage, plutôt que de l'analyse bioinformatique des données de séquençage, même si à ce jour nous n'en avons pas trouvé l'origine exacte. Par conséquent, il faudra rester prudent quant à l'interprétation des données du segment U13 entre les nt 820-900. Cette figure montre également que de manière générale le taux de mutations de l'ADN NP sauvage est inférieur au taux de mutations de la banque d'ADN. Lorsqu'à une position particulière, ce taux était trop proche du taux de mutation de la banque, cette position a été exclue de l'analyse (positions A865 de la figure 52C ou U556 de la figure 53A). Ainsi, nous avons pu analyser la majorité des nt du segment NP.



**Figure 49 : Représentation du taux de mutations des nt 1-1565 de l'ADN NP sauvage et de la banque d'ADN NP séquencés, provenant respectivement de l'ARNv NP sauvage et de la banque d'ARNv NP élués du gel. A.** Taux de mutations des nt 1-1565 (panneau de gauche) ou des nt 700-950 (panneau de droite). **B.** Taux de mutations des nt 1-873 (segment U12). **C.** Taux de mutations des nt 703-1565 (segment U13).

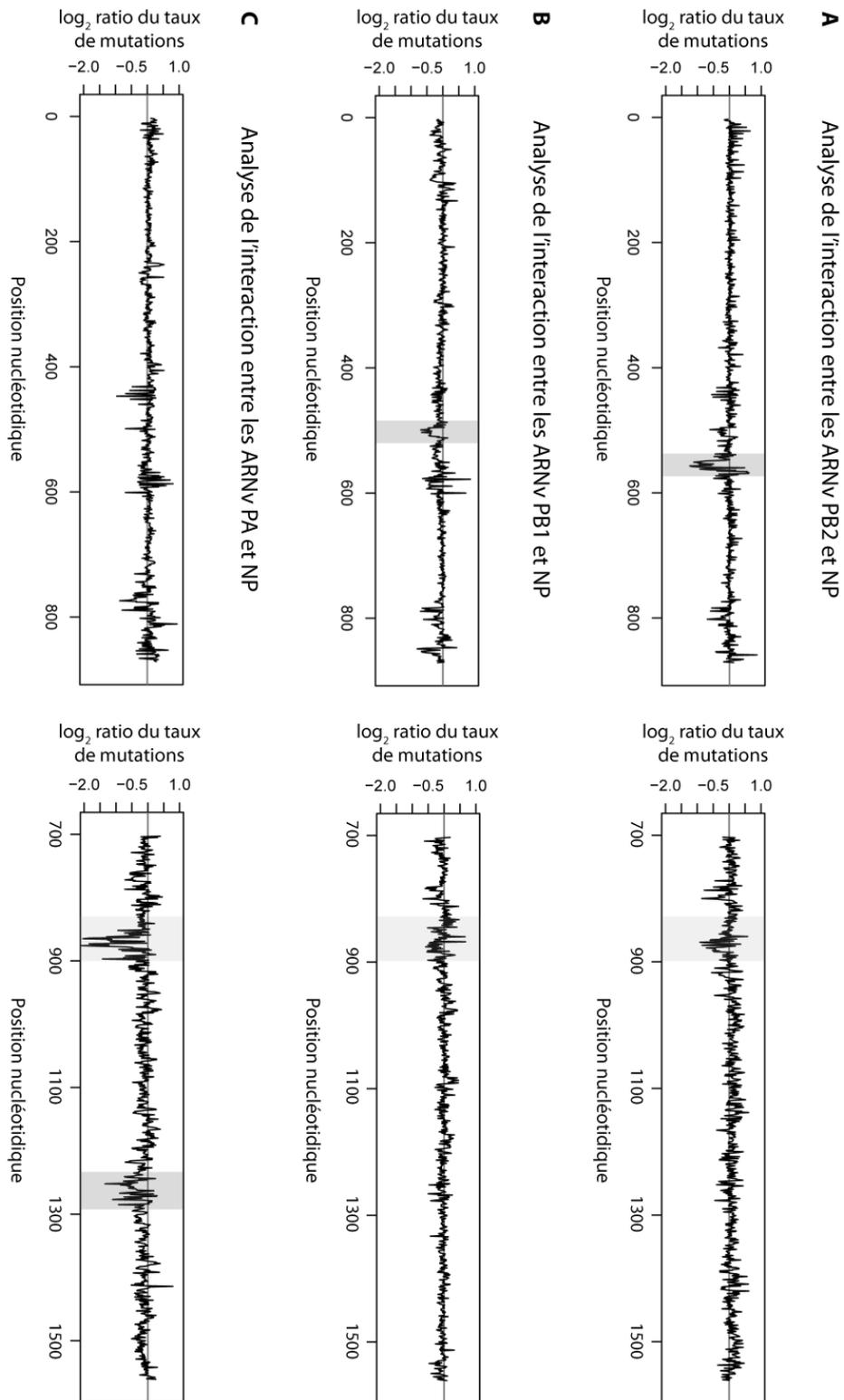
Dans un second temps, nous avons calculé le taux de mutations de l'ADN NP sauvage et de la banque d'ADN NP (figure 50A). Pour l'ADN NP sauvage, il est de 0,2%, ce qui représente approximativement trois mutations par molécule. Celles-ci ont probablement été introduites au moment des étapes de préparation des banques en vue du séquençage

(rétrotranscription, amplification par PCR ...) et du séquençage Illumina. Dans le cas de la banque d'ADN NP, le taux de mutations est de 0,67%, ce qui représente approximativement dix mutations par molécule. En tenant compte du taux de mutations lié aux erreurs expérimentales, nous pouvons en conclure que le taux de mutations de la banque avoisine les 0,48% (huit mutations par molécule). En sachant que chaque nt a été lu en moyenne un million de fois, nous estimons que chaque nt muté a été lu environ 5 000 fois. Nous avons également déterminé, pour chaque position, la fréquence de chacune des trois mutations possibles introduites par l'ADN polymérase Mutazyme II (figure 50B). En accord avec le fournisseur, les fréquences de mutations des purines et des pyrimidines sont proches (0,6% et 0,4%, respectivement) (figure 50B). Certaines mutations sont tout de même favorisées ( $A \leftrightarrow G$ ,  $A \leftrightarrow T$  et  $T \leftrightarrow C$ ), tandis que d'autres sont plus rares ( $A \leftrightarrow C$ ,  $T \leftrightarrow G$  et  $C \leftrightarrow G$ ).



**Figure 50 : Représentation (A) du taux de mutations des segments U12 et U13, correspondant à l'ADN NP sauvage et à la banque d'ADN NP séquençés, provenant respectivement de l'ARNv NP sauvage et de la banque d'ARNv NP élués du gel et (B) de la fréquence de mutations de chacune des trois mutations possibles à chaque position.**

Les populations d'ARNv NP non complexés et complexés ont ensuite été comparées. La figure 51 représente en  $\log_2$ , pour chacun des 1565 nt, le ratio entre le taux de mutations à une position dans les échantillons contenant les ARNv NP complexés et le taux de mutations à cette même position dans les échantillons contenant les ARNv NP libres. Les valeurs négatives correspondent à des mutations moins fréquentes dans les complexes que dans l'ARNv NP non complexé. Ces mutations empêchent par conséquent l'ARNv d'interagir avec ses partenaires et les nucléotides sauvages présents à ces positions sont donc importants pour la formation des complexes. À nouveau, les données du segment U13 comportent davantage de bruit de fond que celles du segment U12, principalement entre les positions 820 et 900 (figure 51, panneaux de droite).



**Figure 51 : Représentation, pour les nt 1-1565 de l'ARNv NP, du log<sub>2</sub> du ratio entre le taux de mutations à chaque position dans les échantillons contenant les ARNv NP complexés et le taux de mutations à la même position dans les échantillons contenant les ARNv NP non complexés, dans le cas des interactions entre les ARNv PB2 et NP (A), PB1 et NP (B), PA et NP (C). Rectangle violet : positions nucléotidiques qui pourraient être impliquées dans la formation du complexe ; rectangle bleu : régions avec un bruit de fond important dans le taux de mutations (voir figure 49). Se référer au texte pour l'analyse des résultats.**

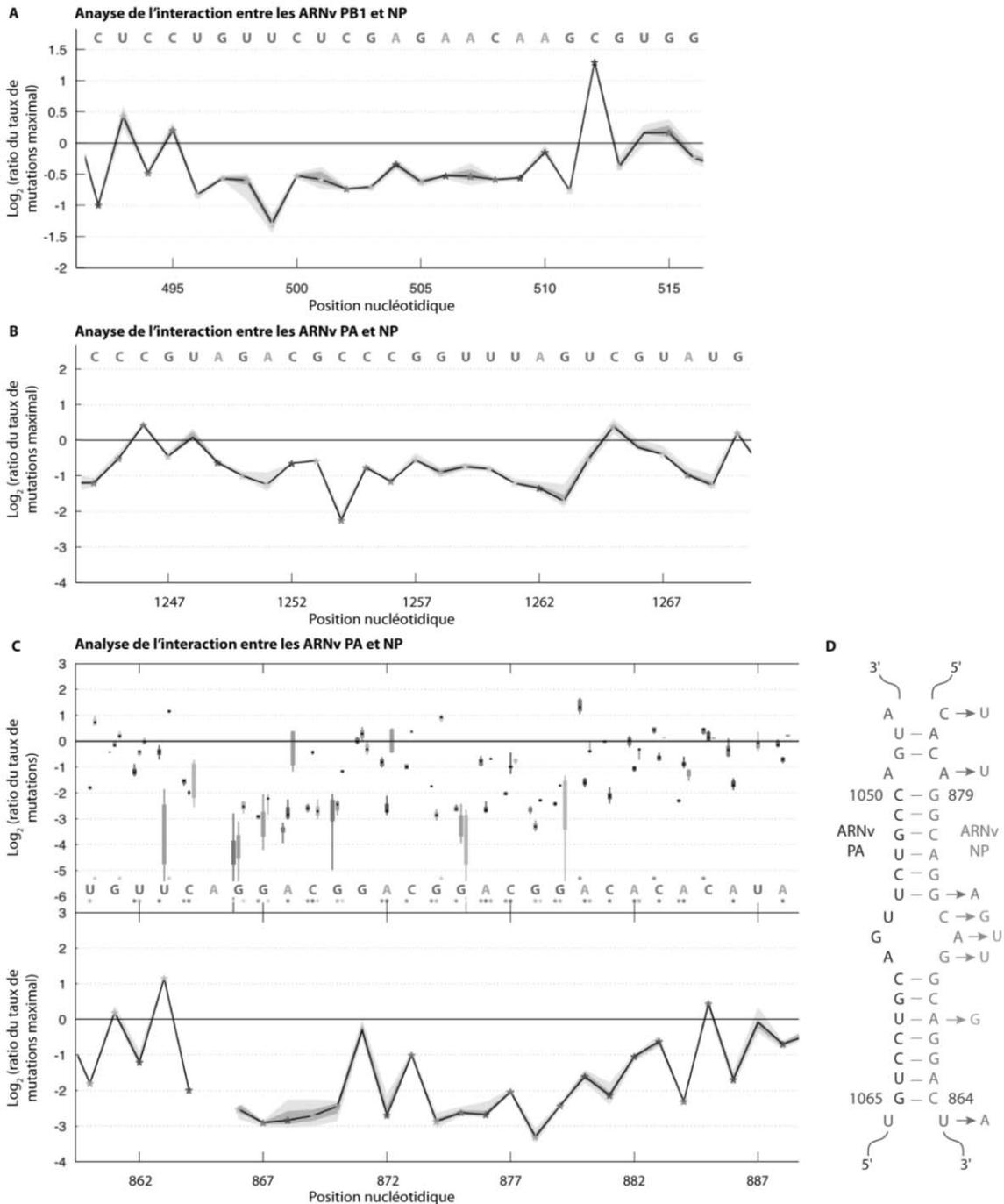
Nous avons cherché à identifier des zones où le ratio du taux de mutations de plusieurs nt contigus est négatif. Cependant, il n'est pas exclu que des nt isolés et dispersés sur l'ARNv NP puissent également avoir un rôle dans l'hétérodimérisation. Dans le cadre de l'interaction entre les ARNv NP et PB2, la région 549-563 de l'ARNv NP, située dans le segment U12, satisfait à ce critère (figure 51A, panneau de gauche). L'implication de ces nt dans l'hétérodimérisation a fait l'objet d'une étude plus approfondie et est détaillée dans la partie suivante (§ A.II.2.3). En parallèle de cette région, trois autres ont également attiré notre attention :

i) La région 496-511 de l'ARNv NP dans le cadre de l'interaction avec le segment PB1 (figure 51B, panneau de gauche et figure 52A). La recherche bioinformatique de complémentarité de séquence entre ces nt et les nt 164-414 de l'ARNv PB1 (identifiés comme étant impliqués dans la formation du complexe par l'approche *in vitro* « classique » (figure 43)) a révélé une seule interaction potentielle de sept pb, qui nous semble cependant peu probable car peu stable ( $T_m = 37,5^\circ\text{C}$ ). Nous avons par conséquent recherché une complémentarité avec la totalité de l'ARNv PB1, qui a permis de proposer plusieurs interactions, que nous pourrions discriminer à l'aide d'oligonucléotides antisens complémentaires aux deux ARNv.

ii) La région 1250-1263 de l'ARNv NP dans le cadre de l'interaction avec le segment PA (figure 51C, panneau de droite et figure 52B). La recherche bioinformatique de complémentarité de séquence entre ces nt et la totalité de l'ARNv PA a révélé plusieurs interactions potentielles. L'existence de ces interactions n'a pas été testée *in vitro*, mais des oligonucléotides antisens seront utilisés afin de les discriminer.

iii) La région 864-884 de l'ARNv NP dans le cadre de l'interaction avec le segment PA (figure 51C, panneau de droite et figure 52C). Ces nt sont localisés dans la zone présentant un bruit de fond important au début du segment U13 (figure 49). Néanmoins, l'effet de ces nt sur la formation du complexe est plus marqué que dans le cas des interactions entre les ARNv NP et PB2 (figure 51A, panneau de droite) et entre les ARNv NP et PB1 (figure 51B, panneau de droite), ce qui suggère que ces nt pourraient avoir un impact sur la formation du complexe. De manière intéressante, par l'approche *in vitro* « classique » nous avons identifié deux interactions potentielles entre les ARNv PA et NP, qui impliquent les nt 863-

871 du segment NP (voir § A.II.1.2, figures 37 et 38). Il avait cependant été impossible de les discriminer à l'aide d'oligonucléotides antisens et d'ARNv mutés. Nous avons analysé les données issues du MIME avec le logiciel MIMEAnTo [318] (figure 52C). Dans cette analyse, seuls les nt qui ont été lus au moins 50 000 fois ont été pris en compte. Les astérisques correspondent à des P-value inférieures à 0,05 et l'effet de ces nt sur la formation du complexe est donc considéré comme significatif. Le graphique du haut représente, pour chaque nt, l'effet sur la formation du complexe de toutes les mutations qui ont été introduites, tandis que le graphique du bas représente la mutation qui a le plus d'effet. Ces données, combinées à des analyses bioinformatiques (prédiction de structures secondaires et recherche de complémentarités de séquences), nous amènent à proposer une autre interaction potentielle, entre les nt 864-879 de l'ARNv NP et les nt 1050-1065 de l'ARNv PA (figure 52D). Ainsi, certaines mutations, notamment celles qui permettent d'étendre l'interaction (U → A en position 863, A → U en position 880, C → U en position 883), favorisent la formation du complexe. La substitution G → A en position 874 favorise également la formation du complexe. D'après l'interaction que nous proposons, les mutations A → U en position 872, G → U en position 871 et C → G en position 873 devraient stabiliser l'interaction entre les deux ARNv en éliminant la boucle interne, alors qu'elles ont peu d'effet sur l'hétérodimérisation. Le bruit de fond dans cette région ne nous permet cependant pas de valider notre hypothèse avec certitude. Elle expliquerait néanmoins les résultats obtenus en présence des oligonucléotides antisens, puisque quelle que soit la séquence de l'ARNv PA ciblée (les nt 1034-1058 ou les nt 1059-1087), l'interaction est inhibée (voir § A.II.1.2). De la même manière, l'interaction est abolie lorsque les nt 1049-1058 ou 1059-1065 sont délétés. D'après ces résultats, il semblerait donc que l'ensemble des nt 1050-1065 soient nécessaires pour la formation du complexe. Il convient désormais de tester, *in vitro*, l'interaction entre les ARNv NP et PA à l'aide de mutants *trans*-compensatoires et silencieux. L'interaction entre les ARNv PA et NP représente un intérêt particulier dans notre étude, puisqu'elle constitue l'interaction la plus forte du réseau de la souche d'IVA PR8 (74% de complexe) (figure 25B). Il serait par conséquent intéressant d'évaluer son rôle dans un contexte infectieux, à la fois sur la réplication virale et sur l'empaquetage du génome.

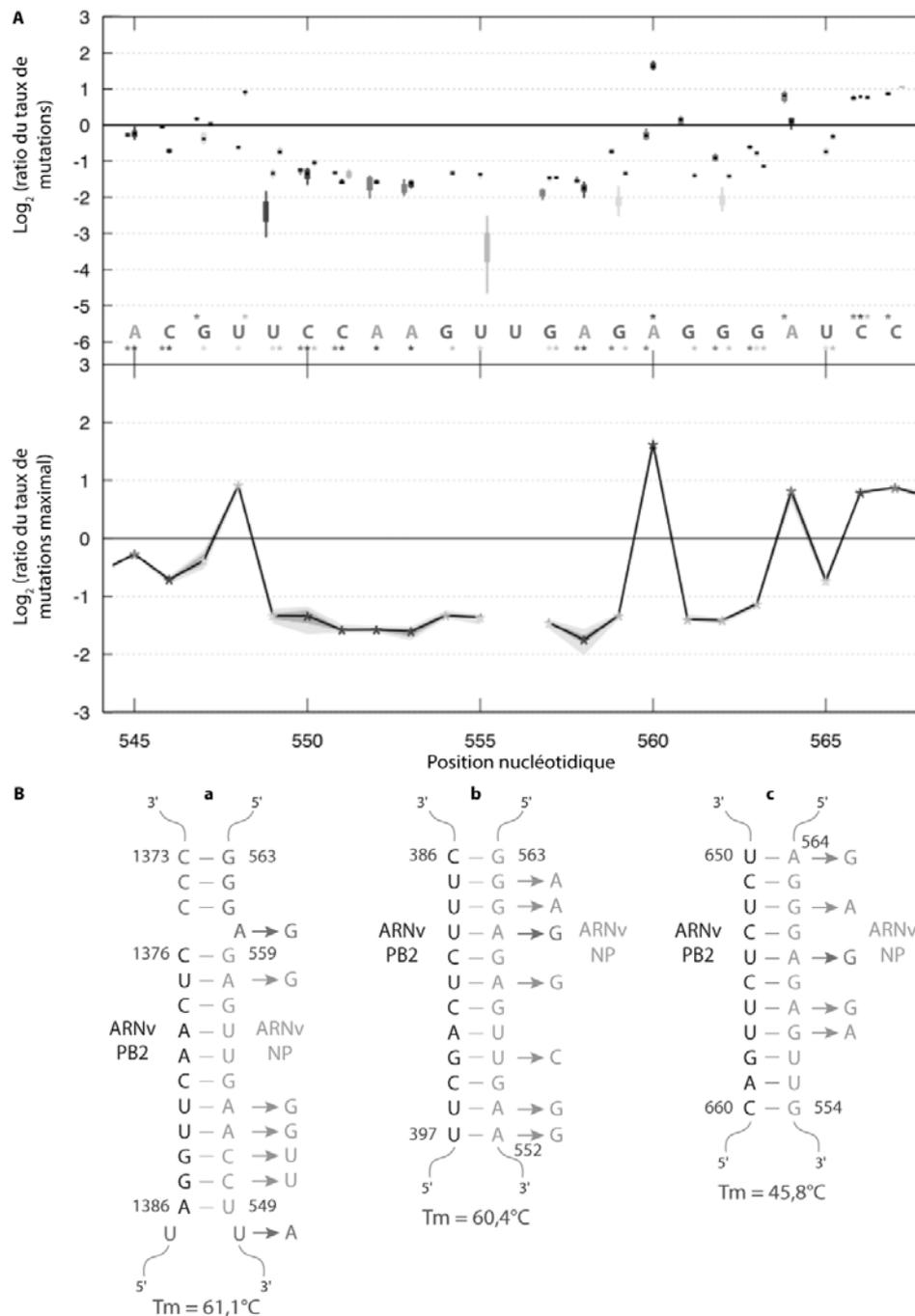


**Figure 52 : Nt de l'ARNv NP potentiellement impliqués dans des interactions avec les ARNv PB1 et PA.** L'analyse a été réalisée avec le logiciel MIMeAnTo. Représentation, pour les nt de l'ARNv NP, du  $\log_2$  du ratio entre le taux de mutations à cette position dans les échantillons contenant les ARNv NP complexés et le taux de mutations à cette même position dans les échantillons contenant les ARNv NP non complexés, dans le cas de l'interaction entre les ARNv PB1 et NP (**A**) et entre les ARNv PA et NP (**B & C**). Représentation des mutations qui ont l'effet maximal sur la formation des complexes (**A, B & C, graphique du bas**) ou de l'effet de chacune des mutations trouvées à cette position (**C, graphique du haut**). **D.** Interaction potentielle entre les segments PA et NP. Cette interaction est proposée à partir des données de MIME et après analyse bioinformatique. Les nt impliqués dans l'interaction sont représentés en noir (ARNv PA) et bleu clair (ARNv NP), tandis que les nt adjacents capables de former des appariements qui étendraient le complexe intermoléculaire sont représentés en rouge. Les mutations en bleu foncé favorisent la formation du complexe, tandis que la mutation en vert ne le favorise pas et celles en orange sont neutres.

### 2.3. Interaction entre les ARNv PB2 et NP

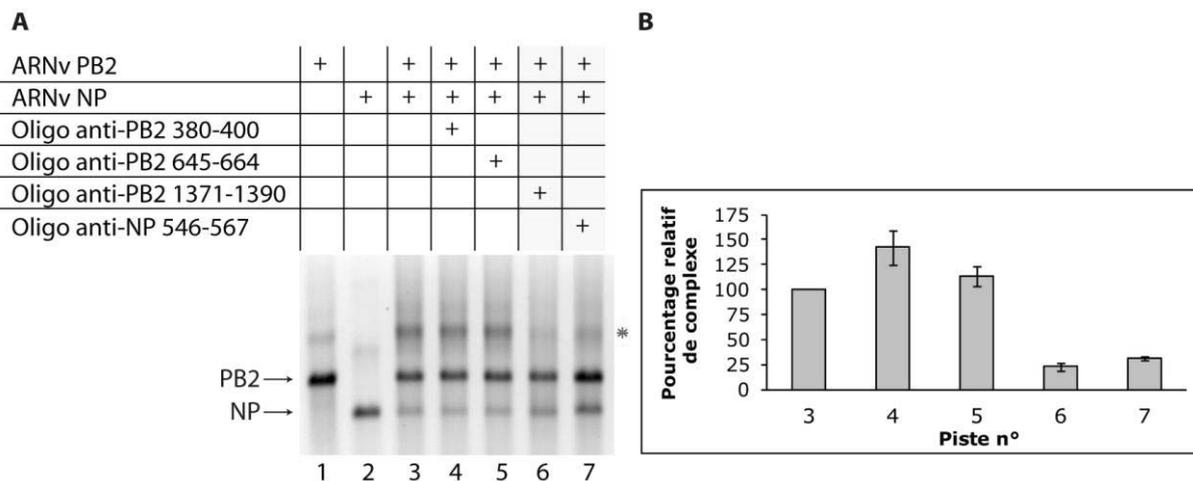
Nous avons approfondi l'étude de l'interaction entre les ARNv PB2 et NP, qui est l'une des interactions les plus fortes dans le réseau PR8 (58%) (figure 25B). Les résultats des expériences de MIME montrent que lorsque les nt 549-563 sont mutés, ils sont sous-représentés dans la population d'ARNv complexés à l'ARNv PB2 (figure 53A, panneau de gauche). Nous avons analysé les données concernant cette région avec le logiciel MIMeAnTo (figure 53). À nouveau, le graphique du haut représente, pour chaque nt, l'effet sur la formation du complexe de toutes les mutations qui ont été introduites, tandis que le graphique du bas représente la mutation qui a le plus d'effet.

La recherche de séquences antiparallèles et complémentaires entre les nt 549-563 de l'ARNv NP et la totalité du segment PB2 a mis en évidence un grand nombre d'interactions potentielles. Après avoir appliqué nos critères de sélection (voir § A.II.1.1.a), nous en avons sélectionné trois (figure 53B). L'une d'entre elles, entre les nt 1376-1386 du segment PB2 et les nt 549-559 du segment NP, est la plus en adéquation avec les résultats de MIME (figure 53B, interaction a). En effet, lorsque les trois G situés en 5' de cette région (561-563) sont mutés en A, l'effet sur la formation du complexe est maximal (figure 53). Il semblerait que ces trois G interagissent avec les trois C de l'ARNv PB2 (1373-1375). De plus, la mutation du A en position 560 de l'ARNv NP favorise l'établissement d'un hétérodimère, mais uniquement lors de l'introduction d'un G (figure 53). Cette mutation pourrait éventuellement stabiliser l'interaction entre les deux ARNv en décalant les appariements et en supprimant le nt non apparié. De même, la mutation U → A en position 548 de l'ARNv NP favorise la formation du complexe (figure 53), ce qui se traduit probablement par l'établissement d'un appariement avec le U en position 1387 de l'ARNv PB2. Néanmoins, les données ne coïncident pas totalement avec notre hypothèse : les mutations A → G en positions 552, 553 et 558 devraient maintenir l'appariement (pb U-G) et devraient donc peu affecter la formation du complexe. Cependant, d'après les résultats, la présence d'un G diminue l'interaction avec l'ARNv PB2 dans les trois cas. De la même manière, d'après l'interaction que nous proposons, les mutations C → U en positions 550 et 551 devraient permettre l'établissement d'une pb G-U, alors que les données montrent que la formation du complexe est affectée par l'introduction des trois nt possibles à ces positions (figure 53).



**Figure 53 : Nt de l'ARNv NP potentiellement impliqués dans des interactions avec l'ARNv PB2.** L'analyse a été réalisée avec le logiciel MIMeAnTo. **A.** Représentation, pour les nt 545-567 du segment NP, du  $\text{log}_2$  du ratio entre le taux de mutations à cette position dans les échantillons contenant les ARNv NP complexés et le taux de mutations à cette même position dans les échantillons contenant les ARNv NP non complexés, dans le cas de l'interaction entre les ARNv PB2 et NP. Représentation des mutations qui ont l'effet maximal sur la formation des complexes (**graphique du bas**) ou de l'effet de chacune des mutations trouvées à cette position (**graphique du haut**). **B.** Interactions potentielles entre les segments PB2 et NP proposées à partir des données de MIMe et après analyse bioinformatique. Les nt impliqués dans l'interaction sont représentés en noir (ARNv PB2) et bleu clair (ARNv NP), tandis que les nt adjacents capables de former des appariements qui étendraient le complexe intermoléculaire sont représentés en rouge. Les mutations en bleu foncé favorisent la formation du complexe, les mutations en vert ne le favorisent pas et la mutation en orange est neutre. La stabilité thermique ( $T_m$ ) de chaque interaction est indiquée.

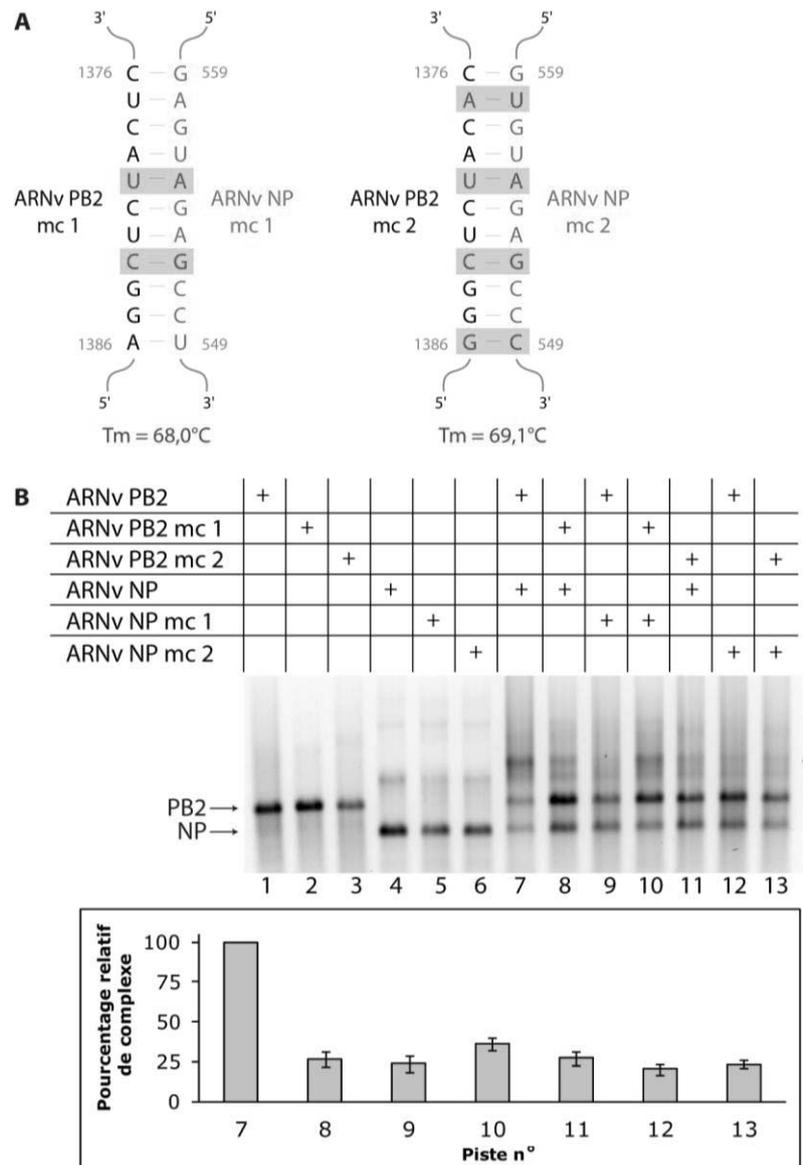
Afin de tester l'implication des nt 549-563 de l'ARNv NP dans la formation du complexe avec le segment PB2, nous avons testé, par retard de migration sur gel, si l'ajout d'oligonucléotides antisens affecte l'hétérodimérisation. Nous avons utilisé un oligonucléotide complémentaire aux nt 546-567 de l'ARNv NP, ainsi que trois oligonucléotides anti-PB2 complémentaires aux trois séquences d'interaction potentielles (figure 53B) (voir Matériels et Méthodes, tableau 6). L'hétérodimérisation entre les segments PB2 et NP est fortement diminuée en présence des oligonucléotides qui ciblent les nt 1371-1390 de l'ARNv PB2 (figure 54, piste 6) ou les nt 546-567 de l'ARNv NP (figure 54, piste 7) (respectivement 23% et 31% de complexe restant par rapport au complexe sauvage). Les deux autres oligonucléotides antisens complémentaires au segment PB2 n'ont pas d'effet sur l'hétérodimérisation (figure 54, pistes 4 et 5). Nous pouvons en conclure que les nt 1376-1386 de l'ARNv PB2 et 549-559 de l'ARNv NP sont responsables de l'interaction *in vitro* entre ces deux segments.



**Figure 54 : Identification des régions des ARNv PB2 et NP impliquées dans l'hétérodimérisation à l'aide d'oligonucléotides antisens complémentaires aux deux ARNv. A.** Analyse des interactions ARN/ARN par électrophorèse sur gel d'agarose en conditions natives après co-incubation des ARNv PB2 et NP sauvages en présence d'oligonucléotides antisens complémentaires à de courtes séquences des ARNv PB2 et NP. Les monomères d'ARNv sont indiqués par des flèches et les hétérodimères par un astérisque rouge. **B.** Quantification de la proportion de complexe présent dans chaque piste (moyenne de 3 expériences indépendantes  $\pm$  SEM).

Par la suite, deux couples de mutants *trans*-compensatoires ont été construits, contenant deux ou quatre substitutions (figure 55A). Les mutations, situées dans les régions codantes des deux segments, ont été conçues de manière à être silencieuses afin d'être testées par la suite dans un contexte infectieux. Pour les deux couples, *in vitro*, les ARNv PB2

et NP mutés interagissent peu avec l'ARNv partenaire sauvage (entre 20% et 27% de complexe par rapport au complexe formé entre les ARNv sauvages) (figure 55B, pistes 8, 9, 11 et 12). Cependant, l'interaction n'est pas restaurée significativement lorsque les ARNv mutés sont co-incubés (36% et 24% de complexe) (figure 55B, pistes 10 et 13).



**Figure 55 : Impact de mutations *trans*-compensatoires au niveau des régions impliquées dans l'interaction entre les ARNv PB2 et NP sur la formation du complexe. A.** Description des mutations *trans*-compensatoires et silencieuses introduites au niveau de la séquence d'interaction entre les ARNv PB2 et NP (ARNv PB2 mc 1, ARNv PB2 mc 2, ARNv NP mc 1 et ARNv NP mc 2). Les ARNv PB2 mc 1 et NP mc 1 ont été mutés au niveau de deux nt centraux, tandis que les ARNv PB2 mc 2 et NP mc 2 ont été mutés au niveau de quatre nt. La stabilité thermique ( $T_m$ ) de chaque interaction est indiquée. **B.** Analyse des interactions ARN/ARN par électrophorèse sur gel d'agarose en conditions natives après co-incubation des ARNv PB2 et NP sauvages et des ARNv PB2 et NP portant des mutations *trans*-compensatoires. Les monomères d'ARNv sont indiqués par des flèches et les hétérodimères par un astérisque rouge. La quantification de la proportion de complexe présent dans chaque piste (moyenne de 3 expériences indépendantes  $\pm$  SEM) a été réalisée.

## 2.4. Conclusion

L'approche *in vitro* « classique » appliquée à l'interaction entre les ARNv PB2 et NP n'avait pas permis d'identifier la séquence responsable de l'hétérodimérisation (résultats non présentés) (voir § A.II.1.4). En effet, les résultats étaient peu reproductibles, les ARNv PB2 et NP délétés n'ayant pas le même effet sur l'interaction dans toutes les expériences réalisées. De la même manière, l'inhibition de la formation du complexe par les oligonucléotides antisens complémentaires à l'ARNv PB2 était très variable, tandis que les délétions terminales de l'ARNv NP n'ont pas eu d'effet sur la formation du complexe (résultats non présentés). Par l'approche MIME *in vitro*, nous avons identifié plusieurs nt, localisés entre les positions 549 et 563 de l'ARNv NP, dont la mutation affecte la formation de l'hétérodimère (figure 53A). L'utilisation de la bioinformatique et d'oligonucléotides antisens (figure 54) ont conduit à l'identification de la séquence d'interaction entre les deux partenaires (figure 53B, interaction a). Néanmoins, *in vitro*, l'utilisation de mutants *trans*-compensatoires n'a pas totalement validé l'existence de cette interaction. En effet, bien que l'ARNv muté interagisse peu avec son partenaire sauvage, l'interaction en présence des deux partenaires *trans*-compensatoires, mc ou mc 2, n'est pas restaurée (figure 55).

Afin d'expliquer l'absence de restauration, nous avons comparé les stabilités thermiques de l'interaction sauvage entre les ARNv PB2 et NP (figure 53B, interaction a) et entre les ARNv mutés (figure 55A) à l'aide de l'outil bioinformatique DINAMelt (<http://unafold.rna.albany.edu>) [332]. L'interaction sauvage possède une  $T_m$  de 61,1°C, tandis que les deux interactions mutées sont plus stables ( $T_m = 68^\circ\text{C}$  et  $69,1^\circ\text{C}$  pour les interactions mc 1 et mc 2, respectivement). Une perte de stabilité ne peut donc pas expliquer l'absence de restauration de l'interaction *in vitro*. Les structures secondaires prédites des séquences d'interaction des deux ARNv sauvages (<http://unafold.rna.albany.edu/?q=mfold/RNA-Folding-Form>) [333], suggèrent qu'elles peuvent adopter des structures locales en tiges-boucles, peu stables, qui pourraient potentiellement initier la formation de l'interaction intermoléculaire en formant un complexe boucle-boucle (*kissing loop complex*). L'existence de telles interactions a été suggérée au laboratoire dans le cas du complexe formé entre les ARNv PB1 et NS de la souche d'IVA aviaire England [273] et démontrée dans la dimérisation du génome du VIH-1 [334]. Il est cependant intéressant de constater que ces structures locales sont

thermodynamiquement plus stables dans les ARNv mutés mc 2, ce qui pourrait expliquer pourquoi ces ARNv ne restaurent pas l'interaction. Les structures locales et les stabilités thermodynamiques des ARNv PB2 et NP sont cependant moins affectées par l'introduction des deux mutations du couple mc.

Lors de l'étude d'autres interactions (notamment entre les ARNv PB1 et M de la souche PR8 (résultats non présentés) et entre les ARNv NA et M de la souche Moscou (voir § B.I), nous avons remarqué que la restauration de l'interaction entre les deux ARNv mutés est dépendante du type de mutations introduites. Afin d'évaluer l'impact des mutations sur la restauration de l'interaction entre les ARNv PB2 et NP, nous pourrions par conséquent construire d'autres couples de mutants *trans*-compensatoires et silencieux, en réalisant des combinaisons des quatre mutations introduites dans le couple mc 2 (hormis la combinaison du couple mc 1) (figure 55A).

Malgré l'absence de restauration de l'interaction entre les ARNv PB2 et NP portant des mutations *trans*-complémentaires, nous avons tout de même décidé de poursuivre l'étude de cette interaction dans le contexte de la réplication virale (voir § A.III.2).

Par l'approche de MIME, nous avons également identifié d'autres nt de l'ARNv NP qui semblent impliqués dans l'interaction avec les ARNv PB1 et PA (figures 51 et 52). Il convient donc désormais de poursuivre l'analyse des données et d'évaluer l'impact des nt 496-511 et 1250-1263 sur la formation des complexes avec les ARNv partenaires PB1 et PA, respectivement. Nous allons vérifier l'existence de certaines interactions proposées par l'outil bioinformatique Google à l'aide d'oligonucléotides antisens complémentaires aux ARNv. Par la suite, des mutants *trans*-compensatoires et silencieux, dans la mesure du possible, seront construits afin de valider les interactions *in vitro*. Pour finir, dans le cas de mutations silencieuses, le rôle fonctionnel de ces interactions sera testé en culture cellulaire. Les données suggèrent également que les nt 866-884 de l'ARNv NP sont impliqués dans l'interaction avec l'ARNv PA (figure 51C, panneau de droite et figure 52C). Cette région se situe néanmoins dans la zone présentant un bruit de fond important (figure 49) et il est donc pour le moment difficile de conclure quant à la réelle implication de ces nt. Les résultats semblent tout de même plus spécifiques que ceux obtenus, dans la même région, dans le cas des interactions avec les ARNv PB2 et PB1. De manière intéressante, ces nt correspondent

en partie à ceux identifiés par l'approche *in vitro* « classique » (voir § A.II.1.2). Rappelons que deux séquences d'interaction potentielles avaient été proposées entre les segments PA et NP et qu'il avait été impossible de les discriminer à l'aide d'ARNv mutés et d'oligonucléotides antisens. Les données issues du MIME, combinées à des analyses bioinformatiques (prédiction de structures secondaires et recherche de complémentarités de séquences), mènent vers une interaction potentielle, contenant une boucle interne, entre les nt 864-879 de l'ARNv NP et les nt 1050-1065 de l'ARNv PA (figure 52D). Il s'agit désormais de vérifier notre hypothèse en construisant des mutants *trans*-compensatoires et silencieux, qui seront ensuite testés *in vitro* et *in viro*.

Les données concernant l'interaction entre les ARNv PB2 et NP, et potentiellement entre les ARNv PA et NP, si son existence se confirme, suggèrent que l'approche MIME *in vitro* est tout à fait appropriée pour identifier des nt fonctionnels des ARNv, à savoir impliqués dans des interactions ARN/protéines ou ARN/ARN. Dans notre cas, les ARNv ont été sélectionnés selon leur capacité à interagir, *in vitro*, avec leurs ARNv partenaires. Nous avons démontré la faisabilité du protocole appliqué à l'ARNv NP de l'IVA PR8. Dans une seule expérience, nous avons testé toutes les positions nucléotidiques du segment NP, ce qui aurait été impossible avec l'approche « classique ».

### **3. Discussion**

#### **3.1. Mise en œuvre de méthodes reposant sur des interactions ARN/ARN détectables *in vitro***

Historiquement, au laboratoire, nous avons utilisé une méthode « classique » combinant i) l'introduction de grandes délétions dans les ARNv, ii) l'inhibition de la formation des complexes ARN/ARN par des oligonucléotides antisens, iii) l'identification subséquente, par bioinformatique, de séquences complémentaires sur les ARNv partenaires et iv) la validation *in vitro* des interactions par la construction de mutants *trans*-compensatoires sur les ARNv partenaires (voir § A.II.1). Cette méthode, bien qu'ayant démontré sa capacité à identifier des nt impliqués dans des interactions fonctionnelles entre ARNv [273] est longue à mettre en œuvre. De plus, les ARNv adoptent des structures secondaires et tertiaires nécessaires à leur fonctionnalité. Il est par conséquent difficile d'introduire des mutations qui ne perturbent pas la structure des ARNv. De petits éléments

de structures secondaires ont également été prédits dans les ARNv génomiques des IVA [103–105] et la structure du segment NS entier, en absence de protéines virales, a été obtenue récemment [109]. Une connaissance plus approfondie de la structure des ARNv au sein des RNPv nous permettrait de réaliser des mutations plus rationnelles.

Durant ma thèse, nous avons également mis en place une nouvelle approche afin d'identifier les nt de l'ARNv NP qui sont importants pour la formation des complexes intermoléculaires, *in vitro* (voir § A.II.2). L'approche MIME est basée sur i) l'introduction de mutations aléatoires au sein d'un ARN cible, ii) la sélection des ARN mutés selon leur fonctionnalité et iii) la détermination de la fréquence de chaque mutation à chaque position de l'ARN dans les populations d'ARN fonctionnels et non fonctionnels par séquençage à haut débit. Dans ce cas, l'impact sur la structure de l'ARNv est donc plus limité que lors de l'introduction de grandes délétions. Cette approche est par conséquent moins biaisée et plus rapide, puisque toutes les positions sont testées dans une seule expérience. De manière importante, le MIME a été validé pour l'étude d'interactions ARN/protéines [277] et durant ma thèse nous l'avons appliqué, pour la première fois, à l'étude d'interactions ARN/ARN.

Ces deux méthodes ont cependant des points communs, puisqu'elles impliquent l'utilisation d'ARNv en absence de protéines NP et du complexe ARN polymérase. De plus, la sélection fonctionnelle des ARNv est basée sur le retard de migration sur gel d'agarose natif. L'application de ces deux méthodes au même partenariat devrait par conséquent conduire à l'identification d'une même séquence d'interaction. Il semblerait effectivement que certains nt de l'ARNv NP, impliqués dans l'interaction avec l'ARNv PA, aient été identifiés par les deux approches, même si l'existence de l'interaction, *in vitro* et *in viro*, reste à confirmer.

### **3.2. Des séquences d'interaction situées tout le long des RNPv**

D'après la littérature, les signaux d'emballage des souches humaines PR8 et WSN (H1N1) sont localisés aux extrémités de chaque segment génomique (pour revues, voir [33, 247]). Nous avons émis l'hypothèse que ces signaux se superposent aux séquences impliquées dans les appariements ARN/ARN qui guident l'assemblage des complexes supramoléculaires. Les travaux au laboratoire concernant les souches humaine Moscou (H3N2) [308] et aviaire England (H5N2) [319] n'ont que partiellement validé ce postulat. En effet, si certaines interactions de la souche humaine Moscou s'établissent effectivement, *in*

*vitro, via* des régions terminales, d'autres font intervenir des régions plus centrales [308]. La mutation des codons très conservés S71 et R73 de la protéine M2, qui affecte la réplication virale, l'assemblage des virions et l'empaquetage du génome de la souche PR8 [260], a également un rôle dans la réplication virale et l'empaquetage de l'ARNv M de la souche Moscou [308]. Dans le cas de la souche aviaire England, les segments interagissent *via* des régions plus centrales des ARNv [319].

Au sein de particules virales Moscou, les RNPv sont suspendues parallèlement les unes par rapport aux autres à partir d'une plateforme vraisemblablement constituée des complexes ARN polymérase de chaque segment et contenant les régions terminales des ARNv (200-300 nt), qui hébergent les signaux d'empaquetage [308]. Aucun contact entre RNPv n'a été identifié en dehors de cette zone. À l'inverse, une étude de Noda *et al.* [320] suggère l'existence d'interactions ARN/ARN tout le long des segments de la souche WSN, en dehors des signaux d'empaquetage. L'existence de contacts entre RNPv en dehors des extrémités a également été visualisée par tomographie dans le cas de la souche England [319]. Ce résultat est conforté par la démonstration de l'importance de l'interaction entre les segments PB1 et NS, qui implique les régions centrales des segments, pour une réplication virale optimale et le co-empaquetage des ARNv [273]. Trois autres études suggèrent que les régions centrales de certains ARNv sont impliquées dans le mécanisme d'empaquetage du génome [259, 325, 335]. Récemment, Cobbin *et al.* [272] ont déterminé que la co-ségrégation des segments PB1 et NA de la souche Udorn est guidée par une interaction entre l'ARNv NA et une région codante centrale de l'ARNv PB1.

L'orientation parallèle ou antiparallèle des RNPv à l'intérieur des particules virales n'est pas encore très bien définie. Les résultats du laboratoire sont en faveur d'une organisation parallèle des RNPv [273, 308], mais des structures secondaires locales pourraient permettre une orientation antiparallèle des nt impliqués dans l'interaction, comme dans le cas de la dimérisation du génome du VIH-1 [334, 336]. Une étude de Sugita *et al.* [337] suggère que même si la majorité des RNPv sont incorporées avec le complexe ARN polymérase du côté de l'extrémité bourgeonnante des virions, certaines d'entre elles pourraient être incorporées avec le complexe ARN polymérase du côté basal du virus, augmentant ainsi le nombre de possibilités d'interactions entre les ARNv.

Dans le cas de la souche humaine PR8, nous avons montré que la majorité des régions engagées dans les interactions intermoléculaires, *in vitro*, sont situées à distance des extrémités des ARNv. En effet, seule l'interaction entre les ARNv PB1 et M fait intervenir deux régions terminales, situées à 100-130 nt des extrémités. Les régions des ARNv PB1 et PA impliquées dans l'interaction sont toutes deux situées à environ 500 nt des extrémités. Ces deux interactions, qui nécessitent une orientation parallèle des RNPv, impliquent des régions situées à la même distance des extrémités des ARNv et pourraient donc s'établir facilement *in viro*. Les données *in vitro* concernant les interactions entre les ARNv PB2 et NP et entre les ARNv PA et NP révèlent que les nt responsables de l'hétérodimérisation ne se trouvent pas à la même distance des extrémités. L'existence de ces deux interactions paraît cependant moins évidente, même si une orientation parallèle, plutôt qu'antiparallèle, semble plus probable. Néanmoins, si dans le cas de la souche PR8, la plateforme identifiée dans les virus Moscou n'est pas conservée [308] et que les complexes ARN polymérase ne sont pas sur le même plan, ces deux interactions pourraient s'établir dans les particules virales. De plus, l'existence de structures secondaires pourrait rapprocher les deux séquences, afin de faciliter leur interaction.

Dans les expériences *in vitro*, les délétions ou les oligonucléotides antisens complémentaires aux ARNv qui perturbent la formation des complexes hétérodimériques ont souvent un effet partiel. L'existence d'interactions multipartites pourrait expliquer ce phénomène et serait compatible avec les données de la littérature qui documentent l'existence de signaux d'empaquetage multipartites et discontinus sur les ARNv [107, 258, 260–262, 338].

Récemment, le processus d'incorporation d'un jeu complet des huit segments génomiques des IVA, nécessitant la présence de signaux aux extrémités de chaque ARNv, a été défini comme le processus d'empaquetage du génome (*genome packaging*) par le groupe de M. Schwemmlé [158]. Les mécanismes moléculaires qui permettent la formation du complexe supramoléculaire (*genome bundle*) ou de sous-complexes (*sub-bundle*) restent encore à élucider. Certaines études suggèrent que les sous-complexes sont associés de manière spécifique [52, 148], tandis que Vale-Costa *et al.* [159] ont montré que des zones de fortes concentrations en RNPv se forment dans le cytoplasme, ce qui pourrait favoriser l'établissement d'interactions intermoléculaires, afin de former des sous-complexes, non

spécifiques, entre les RNPv. Il est envisageable que les signaux d’empaquetage soient impliqués dans l’établissement d’interactions ARN/ARN afin de former le complexe supramoléculaire entre les huit ARNv. Deux études concernant les segments PB1 et PA suggèrent cependant que les deux mécanismes seraient partiellement indépendants. En effet, des mutations qui affectent l’empaquetage de l’un des ARNv de la souche WSN n’ont pas d’effet sur l’incorporation du segment partenaire [262], tandis que ces deux segments co-localisent dans le cytoplasme [148], au sein d’un même sous-complexe. Dans notre étude, nous avons déterminé que les régions d’interaction entre les ARNv PB1 et PA sont situées à environ 500 nt des extrémités. Ces régions pourraient par conséquent constituer les signaux nécessaires au *genome bundling*, tandis que les régions terminales, délimitées par Marsh *et al.* [262], sont nécessaires à l’empaquetage du génome. De manière générale, nous avons montré que pour la souche PR8 les régions impliquées dans les interactions ARN/ARN, *in vitro*, sont majoritairement localisées au dehors des signaux d’empaquetage identifiés dans la littérature (pour revues, voir [33, 247]) et pourraient donc être impliquées dans le *genome bundling*.

### **III. Effet des interactions identifiées *in vitro* sur la réplication virale et le co-empaquetage des ARNv**

Les interactions entre les ARNv PB1 et M et entre les ARNv PB2 et NP de la souche PR8 ont été déterminées au nt près. Les séquences impliquées dans la première interaction ont été délimitées à l'aide de l'approche *in vitro* « classique » (voir § A.II.1.1), tandis que les séquences impliquées dans la seconde l'ont été par l'approche MIME *in vitro* (voir § A.II.2.3), récemment développée au laboratoire. Dans les deux cas, nous avons introduit des mutations *trans*-compensatoires et silencieuses dans chacun des partenaires (figures 31A et 55A), afin de valider les interactions, *in vitro*. Nous avons ensuite voulu évaluer le rôle fonctionnel de ces deux interactions dans un contexte infectieux, en étudiant l'impact de ces mêmes mutations sur la réplication virale et sur le co-empaquetage des ARNv au sein des virions. Pour étudier ces deux aspects, nous avons :

i) comparé les cinétiques de réplication de virus sauvage, simples mutants (où l'un des deux segments est muté de manière à détruire l'interaction entre les ARNv) et double mutant (où les deux segments sont mutés de manière à restaurer l'interaction entre les ARNv) ;

ii) déterminé le génotype de virus obtenus par génétique inverse compétitive à neuf plasmides, où les segments sauvages et mutés sont en compétition en présence de leurs partenaires sauvage ou muté.

Pour réaliser ces expériences, j'ai effectué trois séjours au cours de ma thèse (huit mois au total) au sein de l'équipe « Virologie et pathologie humaine » (CIRI) à Lyon dirigée par le Dr Manuel Rosa-Calatrava.

#### **1. Étude de l'interaction entre les ARNv PB1 et M de la souche PR8**

##### **1.1. Production des virus sauvage et mutants par génétique inverse et amplification des virus sur cellules MDCK**

Dans le cadre de cette étude, quatre virus différents ont été produits par génétique inverse « forcée » à partir des plasmides de génétique inverse (pHW2000) sauvages (le virus PR8) ou en présence des plasmides de génétique inverse PB1 ou/et M dans lesquels les trois mutations *trans*-compensatoires et silencieuses décrites dans la figure 31A ont été introduites par mutagénèse dirigée (les virus simples mutants PB1 mc et M mc, et le virus

double mutant PB1 mc/M mc) (tableau 11). Après deux passages successifs sur cellules MDCK, les stocks viraux obtenus ont été titrés en PFU et en DICT<sub>50</sub> (tableau 11).

Virus	Composition génomique								Titre viral	
	ARNv PB2	ARNv PB1	ARNv PA	ARNv HA	ARNv NP	ARNv NA	ARNv M	ARNv NS	PFU/ml	DICT <sub>50</sub> /ml
PR8	wt	wt	wt	wt	wt	wt	wt	wt	2.10 <sup>6</sup>	6,3.10 <sup>8</sup>
PB1 mc	wt	mc	wt	wt	wt	wt	wt	wt	3.10 <sup>6</sup>	9,3.10 <sup>8</sup>
M mc	wt	wt	wt	wt	wt	wt	mc	wt	2,1.10 <sup>6</sup>	7,4.10 <sup>8</sup>
PB1 mc/M mc	wt	mc	wt	wt	wt	wt	mc	wt	3,5.10 <sup>6</sup>	1,1.10 <sup>9</sup>

**Tableau 11 : Composition génomique et titre viral des virus sauvage et mutants produits par génétique inverse « forcée » dans le cadre de l'étude de l'interaction entre les ARNv PB1 et M de la souche PR8.** wt : segment sauvage de la souche PR8, mc : segment possédant trois mutations *trans*-compensatoires et silencieuses.

De manière assez surprenante, pour les quatre virus, les titres en PFU et en DICT<sub>50</sub> diffèrent d'un facteur d'environ 300. La première méthode permet de déterminer le nombre de particules virales capables de former des plages de lyse, tandis que la seconde est une méthode statistique qui évalue la quantité de particules virales infectieuses nécessaires pour induire un ECP dans 50% des cellules inoculées. Les deux protocoles présentent des différences qui peuvent expliquer les variations des titres, notamment le temps de contact entre les virus et les cellules MDCK, les milieux de culture (MEM ou EMEM) et surtout le fait que le milieu soit semi-solide ou non : le milieu semi-solide utilisé pour les titres en PFU restreint les mouvements des virus à une zone plus réduite et limite également l'accessibilité à la trypsine, nécessaire à la maturation de la protéine HA. Dans les deux cas, la méthode de détection, comptage des plages ou visualisation des ECP, est également différente. De manière générale, le titre en PFU est sous-estimé, car certaines particules virales ne vont pas produire de plage, bien qu'elles soient tout de même infectieuses. De plus, il est possible d'avoir deux plages de lyse très proches, qui seront donc considérées comme une seule particule virale infectieuse. Mathématiquement, le facteur de conversion pour passer du titre en PFU au titre en DICT<sub>50</sub> est de 1,43 (titre en PFU/ml = 0,7 x titre en DICT<sub>50</sub>/ml). Cette valeur est cependant rarement observée expérimentalement. Dans notre cas, l'application du facteur de conversion ne permet pas de faire coïncider les titres viraux mesurés par les deux méthodes.

En parallèle du titrage, les ARNV de chacun des stocks de virus ont été extraits et nous avons partiellement séquencé les ARNV PB1 et M afin de vérifier que la séquence d'interaction définie sur chacun des ARNV n'a pas été mutée au cours de la réplication. En effet, des mutations peuvent être introduites en raison de la faible fidélité de l'ARN polymérase virale, soit de manière aléatoire (généralement une substitution par génome viral et par cycle réplcatif), soit de manière plus « spécifique », dans le cas par exemple des virus simples mutants qui pourraient acquérir des mutations rétablissant l'interaction entre les deux segments, dans le cas où l'interaction identifiée est importante *in cellula/in viro*.

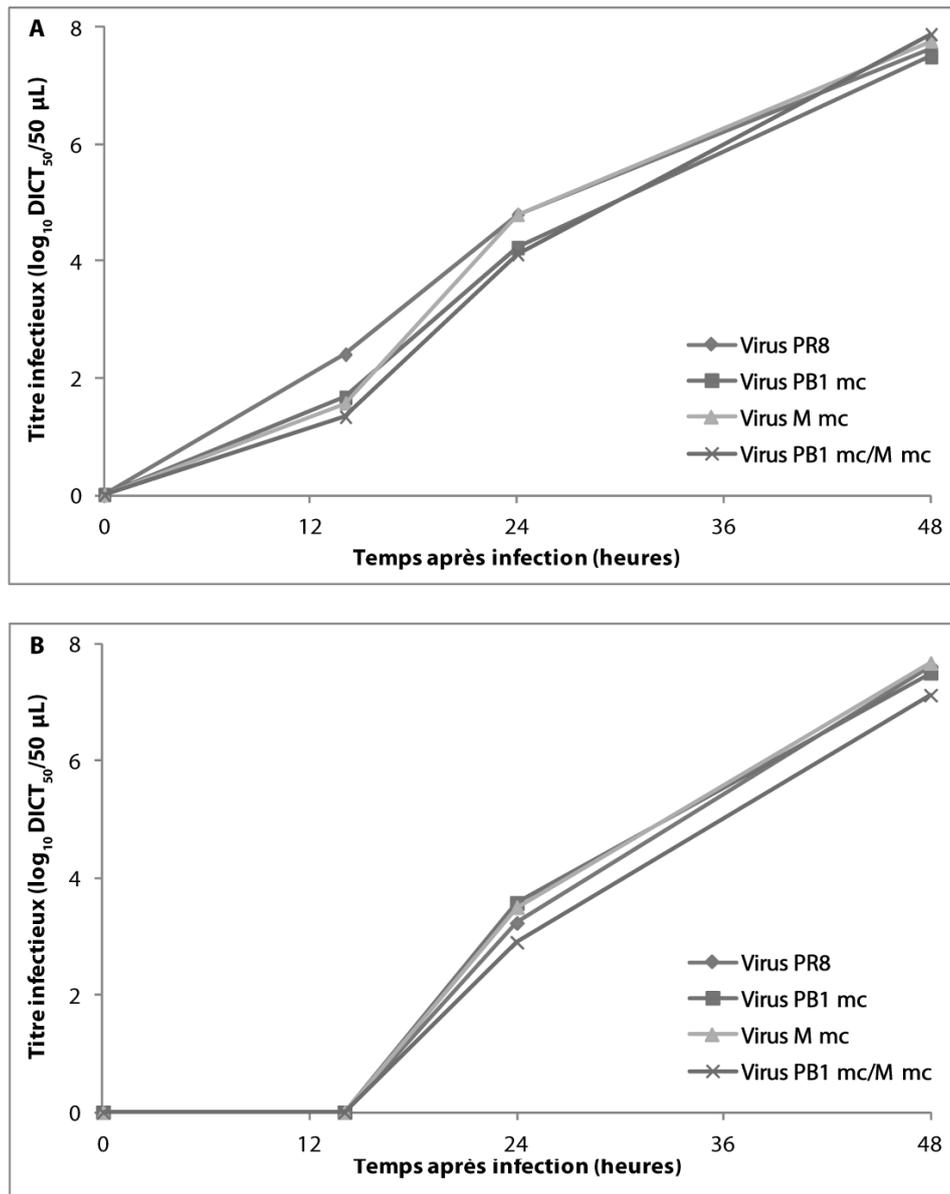
## 1.2. Caractérisation des cinétiques de réplication virale sur cellules MDCK

Des cellules MDCK à confluence ont ensuite été infectées avec les virus PR8, PB1 mc, M mc ou PB1 mc/M mc afin de comparer leurs cinétiques de réplication virale. Si l'interaction que nous avons identifiée *in vitro* joue un rôle *in viro*, nous nous attendons à ce que les virus simples mutants se répliquent moins bien que les virus sauvage et double mutant dans lequel l'interaction est restaurée [273].

Dans le but d'identifier les conditions expérimentales adéquates pour suivre les cinétiques de réplication de ces virus, nous avons testé deux MOI différentes, où une particule virale est présente pour  $10^7$  ou  $10^8$  cellules ( $\text{MOI} = 10^{-7}$  ou  $10^{-8}$ , respectivement). Pour calculer le volume de chaque stock viral nécessaire à l'infection, nous avons utilisé le titre viral exprimé en PFU/ml (tableau 11). Après infection, des fractions aliquotes des surnageants de culture ont été prélevées au cours du temps, puis le titre viral a été déterminé par la méthode de  $\text{DICT}_{50}$ . Nous avons privilégié le titrage en  $\text{DICT}_{50}$  plutôt qu'en PFU, car étant donné le nombre élevé d'échantillons à titrer c'est une méthode plus rapide et plus simple à mettre en œuvre. Pour chaque échantillon viral titré nous avons infecté huit puits d'une plaque 96 puits contenant des cellules MDCK confluentes avec chacune des six dilutions réalisées (soit 48 puits infectés par échantillon). Étant donné que pour chacun des stocks de virus la différence de titres obtenus en PFU et en  $\text{DICT}_{50}$  est similaire (environ 300 fois) (tableau 11), nous avons supposé qu'une différence significative de réplication virale sera visible par titrage en  $\text{DICT}_{50}$ .

Pour chacun des virus, le titre infectieux augmente au cours du temps (figure 56). Nous n'avons pas analysé le surnageant prélevé à 64h, mais nous pensons qu'à 48h le titre viral maximal est presque atteint (entre  $10^{7,1}$  et  $10^{7,9}$  DICT<sub>50</sub>/50  $\mu$ l). En effet, le titre viral 48h après infection est similaire pour les cinétiques réalisées à une MOI de  $10^{-7}$  (figure 56A) et à une MOI de  $10^{-8}$  (figure 56B). L'infection avec chacun des quatre virus à une MOI de  $10^{-7}$  (figure 56A) montre que 14h après infection le virus PR8 atteint un titre viral légèrement plus élevé ( $10^{2,4}$  DICT<sub>50</sub>/50  $\mu$ l) que les trois autres virus (entre  $10^{1,33}$  et  $10^{1,67}$  DICT<sub>50</sub>/50  $\mu$ l), ce qui représente au maximum 1,1 log de différence. À 24h, cette différence s'amenuise ( $10^{4,8}$  DICT<sub>50</sub>/50  $\mu$ l pour les virus PR8 et M mc ;  $10^{4,23}$  DICT<sub>50</sub>/50  $\mu$ l pour le virus PB1 mc et  $10^{4,12}$  DICT<sub>50</sub>/50  $\mu$ l pour le virus PB1 mc/M mc), jusqu'à totalement disparaître à 48h (entre  $10^{7,5}$  et  $10^{7,88}$  DICT<sub>50</sub>/50  $\mu$ l). Le titrage des surnageants récoltés 14h après infection avec une MOI de  $10^{-8}$  n'ont révélé aucun ECP, même dans la première dilution déposée. Par conséquent, ce point a été représenté, par défaut, comme étant égal à zéro (figure 56B). En considérant la gamme de dilutions utilisée dans cet essai nous pouvons néanmoins affirmer que le titre de ces virus est inférieur à  $10^{0,5}$  DICT<sub>50</sub>/50  $\mu$ l. Une détermination plus précise du titre aurait pu être effectuée en réalisant une nouvelle série de titrages en diluant le virus au 1/2, et non au 1/10 comme réalisé classiquement. Dans cette seconde cinétique, les titres viraux sont très proches 24h (entre  $10^{2,9}$  et  $10^{3,57}$  DICT<sub>50</sub>/50  $\mu$ l) et 48h après infection (entre  $10^{7,12}$  et  $10^{7,67}$  DICT<sub>50</sub>/50  $\mu$ l) (entre 0,6 et 0,7 log de différences). Il semblerait que pour la souche PR8, une MOI de  $10^{-8}$  soit trop faible pour analyser la réplication virale. Ainsi, une MOI de  $10^{-7}$  paraît plus appropriée pour suivre les cinétiques de réplication.

La réplication virale n'est jamais totalement abolie lors de l'introduction de mutations dans les séquences d'empaquetage [16, 108, 254, 258, 260, 272, 273, 325, 339]. Il est généralement admis qu'une différence de titre viral significative du point de vue réplicatif doit atteindre 1,5 à 2 log. Nos données suggèrent donc que les caractéristiques répliquatives des virus simples ou double mutants ne sont pas significativement différentes de celles du virus sauvage. Par conséquent, l'interaction entre les segments PB1 et M, si elle existe, n'a pas d'effet important sur la réplication virale, tout du moins dans les conditions expérimentales testées.



**Figure 56 : Impact de mutations *trans*-compensatoires au niveau des régions impliquées dans l'interaction entre les ARNv PB1 et M de la souche PR8 sur la réplication virale.** Les cellules MDCK ont été infectées à une MOI de 10<sup>-7</sup> (A) ou 10<sup>-8</sup> (B). Le titre viral est représenté en log<sub>10</sub> DICT<sub>50</sub>/50 µl. Aucun ECP n'a été observé 14h après infection pour la cinétique réalisée à une MOI de 10<sup>-8</sup>, le titre viral a donc été représenté, par défaut, égal à zéro.

### 1.3. Étude de l'impact de l'interaction entre les ARNv PB1 et M sur leur co-empaquetage par génétique inverse compétitive

L'interaction entre les ARNv PB1 et M, identifiée *in vitro*, pourrait exister au sein des particules virales, sans pour autant, si elle n'est pas centrale à l'établissement du réseau, affecter significativement les capacités répliquatives du virus. Si cette interaction existe *in vivo*, elle pourrait en effet intervenir dans le co-empaquetage sélectif des ARNv partenaires. Afin

de vérifier cette hypothèse, nous avons réalisé des expériences de génétique inverse compétitive à neuf plasmides. Dans cette approche, les ARNV sauvage et muté d'un même segment peuvent entrer en compétition lors de la sélection et de l'empaquetage dans les particules virales, en présence de leur partenaire sauvage ou muté. Si l'interaction identifiée est importante pour ce mécanisme, nous nous attendons à favoriser la production de virus possédant les deux segments capables d'interagir efficacement, à savoir les deux ARNV sauvages ou les deux ARNV mutés [273].

Afin de tester cette hypothèse, quatre transfections ont été réalisées (tableau 12) : les deux premières évaluent la compétition entre les segments PB1 et PB1 mc ou les segments M et M mc dans le fond génétique sauvage, tandis que les deux suivantes permettent la compétition entre ces mêmes segments, mais en présence du partenaire muté. Les virus produits ont été isolés par plages de lyse et leur génotype a été déterminé partiellement par RT-PCR (voir Matériels et Méthodes, § E.IV). Selon les compétitions, nous avons analysé entre 38 et 72 clones, issus de deux (conditions 1 et 3) ou trois (conditions 2 et 4) expériences de génétique inverse compétitive indépendantes.

Pour les deux compétitions impliquant les ARNV PB1 et PB1 mc, en présence des ARNV M ou M mc (tableau 12, conditions 1 et 3), nous obtenons approximativement 50% de chacun des deux génotypes possibles. Il semblerait par conséquent que la composition génomique des virions obtenus par compétition entre les ARNV PB1 sauvage et mutant n'est pas dépendante de la nature de l'ARNV M présent. Ainsi, deux hypothèses peuvent être formulées :

i) L'interaction entre les ARNV PB1 et M n'existe pas *in viro*. Dans ce cas, la sélection du segment PB1 est guidée par son interaction avec un autre partenaire, qui fait intervenir d'autres nt que ceux mutés dans l'ARNV PB1 mc.

ii) L'interaction entre les ARNV PB1 et M existe *in viro*. Néanmoins, la sélection du segment PB1 pourrait être dictée par son interaction avec un autre partenaire que le segment M. Cette interaction serait, *in viro*, plus « forte » et serait donc prédominante par rapport à l'interaction avec l'ARNV M.

Les résultats concernant les compétitions entre les segments M et M mc sont différents (tableau 12, conditions 2 et 4). En effet, en présence du segment PB1 sauvage (tableau 12, condition 2), 92% des virus analysés contiennent l'ARNv M sauvage. De la même manière, en présence de l'ARNv PB1 mc, nous nous attendions à obtenir une majorité de virus possédant l'ARNv M mc. Néanmoins, les résultats montrent que le segment M sauvage est, à nouveau, préférentiellement empaqueté dans les particules virales (tableau 12, condition 4). Ainsi, lorsque les segments M sauvage et muté sont en compétition pour la sélection et l'empaquetage dans les virions, le segment M sauvage est toujours sélectionné quel que soit le segment PB1 présent (sauvage ou muté). Cette situation ne peut s'expliquer par des capacités répliquatives moindres des virus possédant le segment M mc, puisque nous avons montré précédemment que tous les virus se répliquent de manière similaire (figure 56).

	Condition 1	Condition 2	Condition 3	Condition 4
Fond génétique utilisé				
Plasmides transfectés				
Proportion de chaque génotype obtenu	ARNv PB1 ARNv M <b>48%</b> (23 clones/48)	ARNv PB1 ARNv M <b>92%</b> (66 clones/72)	ARNv PB1 ARNv M mc <b>53%</b> (20 clones/38)	ARNv PB1 mc ARNv M <b>89%</b> (62 clones/70)
	ARNv PB1 mc ARNv M <b>52%</b> (25 clones/48)	ARNv PB1 ARNv M mc <b>8%</b> (6 clones/72)	ARNv PB1 mc ARNv M mc <b>47%</b> (18 clones/38)	ARNv PB1 mc ARNv M mc <b>11%</b> (8 clones/70)
P	0,68	<0,0001	0,76	0,0012

**Tableau 12 : Étude du co-empaquetage des ARNv PB1 ou M, sauvages ou mutés, de la souche PR8 par génétique inverse compétitive.** Les segments sauvages (ADN ou ARNv) sont représentés en orange et les segments (ADN ou ARNv) possédant les mutations *trans*-compensatoires sont représentés en rouge. Les résultats des compétitions sont issus de deux (conditions 1 et 3) ou trois (conditions 2 et 4) expériences de génétique inverse compétitive indépendantes. Pour chacune des compétitions, nous avons calculé, par un test de Student, la probabilité que la distribution des deux ARNv en compétition soit aléatoire.

Pour tenter de favoriser l'interaction *in viro* entre les ARNv PB1 et M, nous nous sommes intéressés aux autres partenaires du segment PB1. Comme nous venons de le suggérer, son empaquetage dans les virions pourrait être guidé par son interaction avec un autre ARNv, différent de l'ARNv M. D'après le réseau d'interactions identifié (figure 25B), l'ARNv PB1 possède trois autres partenaires, dont l'ARNv HA avec lequel il établit l'interaction la plus forte (48% de complexe), tandis que les interactions avec les ARNv partenaires PA et NP sont plus faibles (22% et 21% de complexe, respectivement). Si la « force » d'une interaction *in vitro* détermine l'importance de son implication *in viro*, nous nous attendons à ce que l'interaction entre les ARNv PB1 et HA soit favorisée par rapport à celle entre les ARNv PB1 et M. Nous avons émis l'hypothèse selon laquelle la présence d'un ARNv HA hétérologue, dont la séquence est suffisamment différente de celle de l'ARNv HA PR8 pour que l'interaction avec l'ARNv PB1 ait peu de chances d'être conservée, pourrait favoriser l'interaction entre les segments PB1 et M. Nous avons par conséquent choisi de remplacer, dans nos expériences de génétique inverse compétitive, l'ARNv HA PR8 par celui de la souche Moscou, en sachant que leur identité, au niveau nucléotidique, n'est que de 57%. Les données obtenues par l'approche *in vitro* « classique » ont suggéré que la région de l'ARNv HA responsable de l'interaction avec l'ARNv PB1 est localisée entre les nt 679-886 (voir § A.II.1.4, figure 43), mais son implication reste encore à prouver. Dans cette région, l'identité, au niveau nucléotidique, entre les segments HA PR8 et HA Moscou est de 56%.

Nous avons ensuite réalisé les mêmes expériences de génétique inverse compétitive que précédemment, en omettant la première condition (tableau 13). Nous avons privilégié les compétitions 3 et 4 dans lesquelles, si notre hypothèse est vraie, les deux segments mutés devraient préférentiellement être co-empaquetés, ainsi que la compétition 2, servant de contrôle, dans laquelle les proportions de chaque génotype ne devraient pas différer de celles obtenues précédemment. Dans un premier temps nous avons confirmé, par génétique inverse « forcée », que toutes les combinaisons génétiques attendues étaient viables. Pour chacune des trois compétitions réalisées, nous avons analysé entre 25 et 40 clones, issus d'un seul jeu d'expériences. Comme attendu, dans la compétition entre les segments M ou M mc en présence du segment PB1 sauvage, l'ARNv M est préférentiellement empaqueté dans les particules virales (tableau 13, condition 2). Néanmoins, dans les deux dernières conditions, les proportions de chaque génotypes sont identiques à celles obtenues lors des

expériences de compétition en présence de l'ARNv HA PR8 (tableau 13, conditions 3 et 4), démontrant ainsi que l'introduction du segment HA de la souche Moscou, et la destruction potentielle de l'interaction entre les ARN HA et PB1, n'a aucun impact sur le profil de sélection des ARNv PB1 et M, sauvages ou mutés.

	Condition 1	Condition 2	Condition 3	Condition 4
Fond génétique utilisé				
Plasmides transfectés				
Proportion de chaque génotype obtenu		ARNv PB1 ARNv M <b>85%</b> (34 clones/40)	ARNv PB1 ARNv M mc <b>48%</b> (12 clones/25)	ARNv PB1 mc ARNv M <b>93%</b> (37 clones/40)
		ARNv PB1 ARNv M mc <b>15%</b> (6 clones/40)	ARNv PB1 mc ARNv M mc <b>52%</b> (13 clones/25)	ARNv PB1 mc ARNv M mc <b>7%</b> (3 clones/40)

**Tableau 13 : Étude du co-empaquetage des ARNv PB1 ou M, sauvages ou mutés, de la souche PR8 par génétique inverse compétitive.** Les segments sauvages (ADN ou ARNv) sont représentés en orange et les segments (ADN ou ARNv), possédant trois mutations *trans*-compensatoires sont représentés en rouge. Le plasmide en mauve code pour l'ARNv HA de la souche Moscou. Les résultats sont issus d'un seul jeu d'expériences de génétique inverse compétitive.

#### 1.4. Conclusion

Les séquences d'interaction entre les segments PB1 et M ont été déterminées au nt près et validées, *in vitro*, à l'aide de mutants *trans*-compensatoires et silencieux. Nous avons ensuite testé le rôle de cette interaction dans un contexte infectieux. Contrairement à ce qui était attendu, les caractéristiques répliquatives des virus sauvage et mutants ne se sont pas révélées significativement différentes. Les expériences de génétique inverse compétitive entre les segments M ont montré que quel que soit le segment PB1 présent (sauvage ou muté), le segment M sauvage est sélectionné de manière préférentielle. À l'inverse, la composition génomique des virions obtenus par compétition entre les ARNv PB1 sauvage et

muté n'est pas dépendante de la nature de l'ARNv M présent. Par conséquent, il semblerait que cette interaction ne soit pas essentielle pour la réplication virale et pour l'empaquetage des ARNv dans les virions. Comme nous venons de l'évoquer, nous envisageons plusieurs hypothèses qui pourraient expliquer ces résultats :

i) L'interaction entre les ARNv PB1 et M n'existe pas *in viro*. L'empaquetage du segment PB1 pourrait par conséquent être dicté par son interaction avec un autre partenaire, qui implique des nt différents de ceux mutés dans l'ARNv PB1 mc. Dans ce cas, la séquence identifiée sur l'ARNv M pourrait être impliquée dans une interaction, *in viro*, avec un partenaire, différent de l'ARNv PB1, qui n'aurait pas été identifié par la stratégie de co-incubation des ARNv *in vitro*, et qui expliquerait donc sa sélection préférentielle quel que soit le segment PB1 présent. La probabilité que la séquence de M qui interagit *in vitro* avec PB1 interagisse « par hasard » avec un autre partenaire en culture cellulaire est toutefois très faible.

ii) L'interaction entre les ARNv PB1 et M existe *in viro*. L'empaquetage du segment PB1 pourrait néanmoins être guidé par son interaction avec un partenaire différent de l'ARNv M. Cette interaction serait, *in viro*, plus importante en terme de « force » et serait donc prédominante par rapport à l'interaction avec l'ARNv M. Selon le réseau d'interactions que nous avons établi, cet autre partenaire pourrait être le segment PA, HA ou NP. Nous avons vérifié si la présence d'un ARNv HA hétérologue pouvait déstabiliser l'interaction entre les ARNv PB1 et HA, et ainsi favoriser l'établissement de l'interaction entre les ARNv PB1 et M. Néanmoins, les proportions de chaque génotype obtenu ne se sont pas révélées significativement différentes.

D'autres études sont nécessaires afin de valider l'une de ces hypothèses. De manière intéressante, la substitution de trois nt dans l'ARNv M, qui ne modifie pas la séquence protéique, affecte considérablement son empaquetage lorsqu'il est en compétition avec le segment sauvage. Même si pour le moment nous n'avons pas toutes les réponses, il paraît fort probable que l'établissement d'une interaction intermoléculaire ARN/ARN soit responsable de la sélection préférentielle du segment M sauvage.

## 1.5. Discussion

L'interaction entre les ARNv PB1 et M a également été identifiée, *in vitro*, dans la souche WSN (H1N1). Elle est cependant de faible intensité (12% de complexe) (figure 33), tandis qu'elle est de plus forte intensité dans la souche PR8 (39% de complexe) (figure 31). L'analyse des séquences des ARNv PB1 et M du virus WSN révèle que l'interaction pourrait également s'établir. Un seul nt diffère entre les deux séquences des ARNv M, mais conserve l'appariement (pb G-U). Il est néanmoins difficile d'évaluer l'implication de ces nt dans la formation du complexe, *in vitro*, étant donné qu'il est de faible intensité. Par ailleurs, Kimble *et al.* [340] ont analysé le génotype de virions réassortants obtenus par co-transfection de cellules HEK293T avec des plasmides codant les gènes d'un virus de sous-type H9N2 et les gènes d'un virus pandémique de sous-type H1N1, suivie par l'infection de furets. Les données révèlent une co-ségrégation des segments PB1 et M issus de la même souche virale. La séquence de l'interaction entre les ARNv PB1 et M de la souche PR8, que nous avons identifiée et validée *in vitro*, n'est cependant pas conservée dans la souche pandémique H1N1. Deux substitutions sont présentes dans l'ARNv PB1 et une dans l'ARNv M, en comparaison aux ARNv de la souche PR8, mais une seule permet de conserver l'appariement avec son partenaire. L'établissement de cette interaction entre les ARNv PB1 et M de la souche pandémique H1N1 semble par conséquent peu probable. Nous pourrions réaliser des expériences de retard de migration sur gel afin de tester si dans notre système les ARNv PB1 et M de la souche pandémique sont capables d'interagir. Si c'est le cas, nous pourrions identifier les séquences d'interaction, au nt près, entre les deux ARNv, et comparer ces séquences à celles de la souche PR8. En effet, la co-ségrégation des ARNv PB1 et M de la souche pandémique est vraisemblablement guidée par l'existence d'une interaction ARN/ARN. Nous pourrions ainsi envisager d'identifier une interaction entre deux ARNv, *in viro*, pour ensuite appliquer notre approche *in vitro* afin de délimiter précisément la séquence responsable de l'hétérodimérisation. Il serait également intéressant de vérifier, par génétique inverse compétitive, si les segments PB1 et M de la souche pandémique sont préférentiellement co-empaquetés dans le fond génétique de la souche PR8 ou dans tout autre fond génétique. L'expérience inverse, où les segments PB1 et M du virus PR8 sont incorporés dans le fond génétique pandémique, pourrait aussi être réalisée.

## **2. Étude de l'interaction entre les ARNv PB2 et NP de la souche PR8**

L'étude *in vitro* a montré que les mutations *trans*-compensatoires et silencieuses dans les ARNv PB2 et NP (figure 55A) diminuent fortement l'interaction avec les ARNv partenaires sauvages. L'interaction n'est néanmoins pas restaurée en présence des deux segments mutés (voir § A.II.2.3). Ces résultats confirment que les séquences identifiées sont impliquées dans l'interaction entre les segments PB2 et NP, *in vitro*, sans toutefois prouver qu'elles sont effectivement appariées. Afin de valider cette interaction dans un contexte infectieux, nous avons comparé les cinétiques de réplication virales des virus sauvage et simples mutants (mc 1 et mc 2), sans inclure les doubles mutants. Pour évaluer l'impact de cette interaction sur le co-empaquetage des ARNv partenaire par génétique inverse compétitive, nous nous sommes focalisés sur le couple mc 1.

### **2.1. Production des virus sauvage et mutants par génétique inverse et amplification des virus sur cellules MDCK**

Dans le cadre de cette étude, cinq virus différents ont été produits par génétique inverse « forcée » à partir des plasmides de génétique inverse (pHW2000) sauvages (le virus PR8) ou en présence des plasmides de génétique inverse PB2 ou NP dans lesquels deux (virus PB2 mc 1, NP mc 1) ou quatre (virus PB2 mc 2, NP mc 2) mutations *trans*-compensatoires et silencieuses décrites dans la figure 55A ont été introduites par mutagenèse dirigée (tableau 14). Après deux passages successifs sur les cellules MDCK, les stocks viraux obtenus ont été titrés en PFU et DICT<sub>50</sub> (tableau 14). Dans ce cas, les différences entre les titres en PFU et DICT<sub>50</sub> pour les mêmes virus varient entre 130 fois et 400 fois. Les titrages en PFU ont été effectués sur deux aliquots indépendants de chaque stock viral et les résultats obtenus étaient relativement peu variables, tandis que les titrages en DICT<sub>50</sub> n'ont été réalisés qu'une seule fois. Par conséquent, les titres en PFU de ces cinq virus nous semblent plus précis que les titres en DICT<sub>50</sub>.

Virus	Composition génomique								Titre viral	
	ARNv PB2	ARNv PB1	ARNv PA	ARNv HA	ARNv NP	ARNv NA	ARNv M	ARNv NS	PFU/ml	DICT <sub>50</sub> /ml
PR8	wt	wt	wt	wt	wt	wt	wt	wt	1,0.10 <sup>6</sup>	1,3.10 <sup>8</sup>
PB2 mc 1	mc 1	wt	wt	wt	wt	wt	wt	wt	6,3.10 <sup>5</sup>	2,6.10 <sup>8</sup>
PB2 mc 2	mc 2	wt	wt	wt	wt	wt	wt	wt	1,2.10 <sup>6</sup>	2.10 <sup>8</sup>
NP mc 1	wt	wt	wt	wt	mc 1	wt	wt	wt	1,3.10 <sup>6</sup>	4,3.10 <sup>8</sup>
NP mc 2	wt	wt	wt	wt	mc 2	wt	wt	wt	1,8.10 <sup>6</sup>	3,2.10 <sup>8</sup>

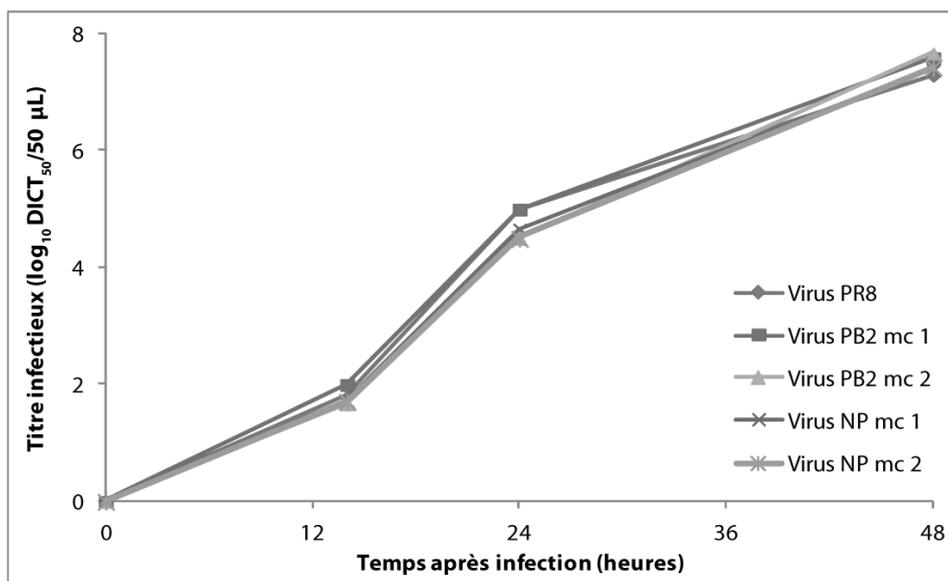
**Tableau 14 : Composition génomique et titre viral des virus sauvage et mutants produits par génétique inverse « forcée » dans le cadre de l'étude de l'interaction entre les ARNv PB2 et NP de la souche PR8.** wt : segment sauvage de la souche PR8, mc 1 : segment possédant deux mutations *trans*-compensatoires, mc 2 : segment possédant quatre mutations *trans*-compensatoires.

Les ARNv ont été extraits de chaque stock viral et rétrotranscrits, puis les segments PB2 et NP ont été séquencés partiellement dans le but de vérifier qu'aucune mutation n'a été introduite par l'ARN polymérase virale au niveau de la séquence d'interaction.

## 2.2. Caractérisation des cinétiques de réplication virale sur cellules MDCK

Des cellules MDCK à confluence ont ensuite été infectées à une MOI de 10<sup>-7</sup> afin de déterminer les cinétiques de réplication des virus PR8 sauvage et simples mutants. À nouveau, nous avons utilisé le titre viral exprimé en PFU/ml pour calculer le volume de chaque stock viral nécessaire à l'infection (tableau 14). Après infection, des fractions aliquotes des surnageants de culture ont été prélevées au cours du temps, puis le titre viral a été déterminé par la méthode de DICT<sub>50</sub>.

Les titres infectieux pour les trois virus augmentent au cours du temps, jusqu'à atteindre un titre viral proche de 10<sup>7,5</sup> DICT<sub>50</sub>/50 µl, 48h après infection (figure 57). Aux trois temps prélevés, les titres obtenus sont très proches entre les virus sauvage et mutants : 14h après infection, le titre varie entre 10<sup>1,67</sup> et 10<sup>2</sup> DICT<sub>50</sub>/50 µl ; 48h après infection, il varie entre 10<sup>4,5</sup> et 10<sup>5</sup> DICT<sub>50</sub>/50 µl ; 48h après infection, il varie entre 10<sup>7,29</sup> et 10<sup>7,67</sup> DICT<sub>50</sub>/50 µl. Nous avons donc une différence de titres allant de 0,3 à 0,5 log, ce qui n'est pas significatif. Nous pouvons en conclure que les capacités répliquatives des virus étudiés ne sont pas significativement différentes. Par conséquent, cette interaction, si elle existe, n'a pas d'effet sur la réplication virale dans les conditions expérimentales testées.



**Figure 57 : Impact de mutations *trans*-compensatoires au niveau des régions impliquées dans l'interaction entre les ARNv PB2 et NP de la souche PR8 sur la réplication virale.** Les cellules MDCK ont été infectées à une MOI de  $10^{-7}$ . Le titre viral est représenté en  $\log_{10}$  DICT<sub>50</sub>/50  $\mu$ L.

### 2.3. Étude de l'impact de l'interaction entre les ARNv PB2 et NP sur leur co-empaquetage par génétique inverse compétitive

Comme dans le cas de l'interaction entre les ARNv PB1 et M, nous pouvons émettre l'hypothèse que l'interaction entre les ARNv PB2 et NP, qui n'affecte pas les capacités répliquatives du virus, puisse malgré tout exister au sein des particules virales. Afin de tester cette possibilité, nous avons étudié son implication dans le co-empaquetage des ARNv partenaires par génétique inverse compétitive à neuf plasmides. Dans ce but, quatre transfections ont été réalisées (tableau 15) : les deux premières évaluent la compétition entre les segments PB2 et PB2 mc 1 ou les segments NP et NP mc 1 dans le fond génétique sauvage, tandis que les deux suivantes permettent la compétition entre ces mêmes segments, mais en présence du partenaire muté. Pour chacune des compétitions, nous avons analysé entre 21 et 24 clones, issus d'un seul jeu d'expériences. Quelle que soit la compétition considérée, nous obtenons environ 50% de chacun des deux génotypes possibles. Ainsi, nous n'observons pas de co-ségrégation des ARNv partenaires PB2 et NP. Par conséquent, l'interaction identifiée entre les segments PB2 et NP ne semble pas être importante pour le co-empaquetage des ARNv au sein des virions.

	Condition 1	Condition 2	Condition 3	Condition 4
Fond génétique utilisé				
Plasmides transfectés				
Proportion de chaque génotype obtenu	<p>ARNv PB2</p> <p>ARNv NP</p> <p><b>45%</b> (10 clones/22)</p>	<p>ARNv PB2</p> <p>ARNv NP</p> <p><b>39%</b> (9 clones/23)</p>	<p>ARNv PB2</p> <p>ARNv NP mc 1</p> <p><b>38%</b> (8 clones/21)</p>	<p>ARNv PB2 mc 1</p> <p>ARNv NP</p> <p><b>58%</b> (14 clones/24)</p>
	<p>ARNv PB2 mc 1</p> <p>ARNv NP</p> <p><b>55%</b> (12 clones/22)</p>	<p>ARNv PB2</p> <p>ARNv NP mc 1</p> <p><b>61%</b> (14 clones/23)</p>	<p>ARNv PB2 mc 1</p> <p>ARNv NP mc 1</p> <p><b>62%</b> (13 clones/21)</p>	<p>ARNv PB2 mc 1</p> <p>ARNv NP mc 1</p> <p><b>42%</b> (10 clones/24)</p>

**Tableau 15 : Étude du co-empaquetage des ARNv PB2 ou NP, sauvages ou mutés, de la souche PR8 par génétique inverse compétitive.** Les segments sauvages (ADN ou ARNv) sont représentés en orange et les segments (ADN ou ARNv) possédant deux mutations *trans*-compensatoires sont représentés en rouge. Les résultats sont issus d'un seul jeu d'expériences.

## 2.4. Conclusion

Les séquences d'interaction, *in vitro*, entre les segments PB2 et NP ont été déterminées au nt près. Les mutants *trans*-compensatoires et silencieux n'ont que partiellement validés l'interaction. Son étude dans un contexte infectieux a montré que les cinétiques de réplication virale des virus sauvage et simples mutants ne sont pas significativement différentes. De la même manière, les expériences de compétition entre les segments PB2 ou NP, sauvages ou mutés, n'ont pas révélé de co-ségrégation des ARNv partenaires. L'ensemble de ces résultats suggère i) que l'interaction identifiée *in vitro* entre ces deux ARNv n'existe pas *in viro* ou ii) que cette interaction n'a aucun effet sur l'empaquetage des ARNv. D'autres études sont encore nécessaires afin de valider l'une ou l'autre de ces hypothèses. Nous pourrions notamment réaliser des expériences de cartographie chimique en solution à l'aide de sondes chimiques telles que le DMS ou des sondes de SHAPE, capables de pénétrer dans les particules virales. Ainsi, la comparaison des intensités de modifications des nt des ARNv PB2 et NP potentiellement impliqués dans l'interaction, provenant de virions sauvages PR8 et de virions simples mutants (où

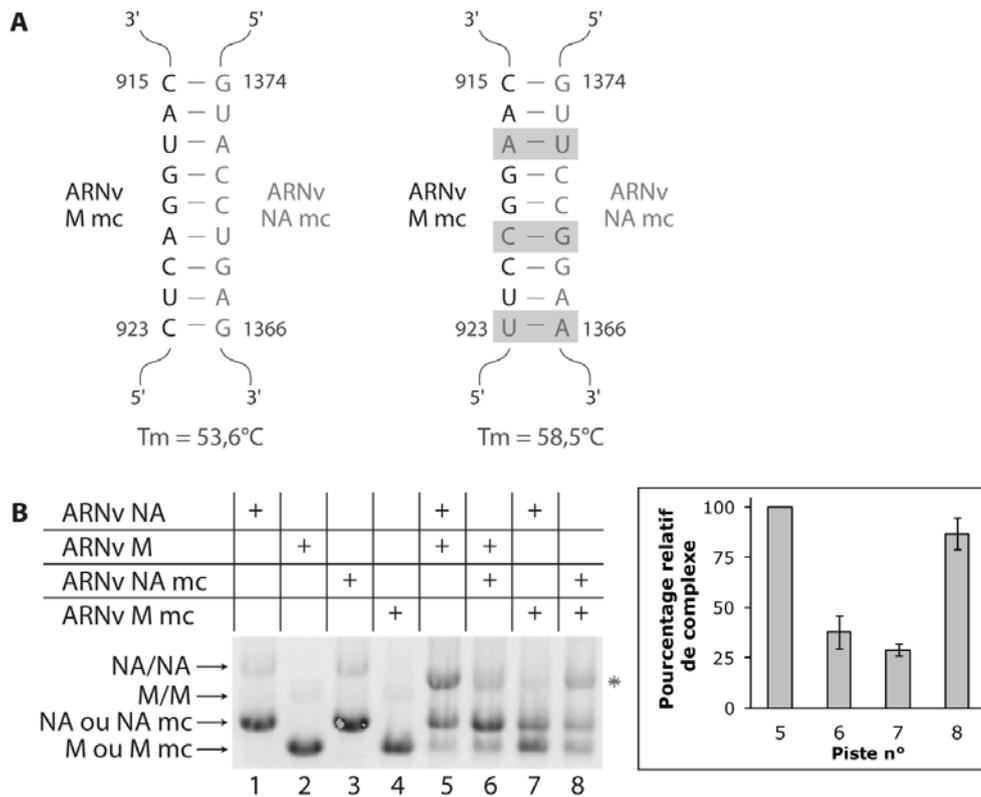
l'interaction devrait donc être abolie), devrait nous permettre d'écarter l'une de ces hypothèses. Une autre solution serait de comparer les profils de réactivité des nt au sein de particules virales sauvages (nt potentiellement appariés) à ceux des nt de RNPv intactes, extraites des virions mais dissociées les unes des autres (nt non appariés). De telles expériences de cartographie chimiques en solution sur des particules virales intactes ont déjà été réalisées au laboratoire et visaient à identifier des nt de l'ARNv M de la souche Moscou potentiellement impliqués dans une interaction avec l'ARNv partenaire NA.

## B. Étude d'interactions entre ARNv d'IVA humains de sous-type H3N2

### I. Interaction entre les ARNv NA et M de la souche Moscou

#### 1. Contexte de l'étude

L'interaction entre les ARNv NA et M de la souche humaine Moscou avait été identifiée au laboratoire par la méthode *in vitro* « classique » (utilisation d'ARNv portant de grandes délétions, inhibition de l'interaction par des oligonucléotides antisens et analyse bioinformatique) (figure 58A, interaction de gauche). Rappelons que dans le cas de la souche Moscou, les complexes sont visualisés après co-transcription des ARNv [308], la co-incubation mettant en évidence des interactions de faibles intensités, dont la spécificité n'a pas été étudiée (thèse C. Gavazzi, UdS, 2011). Par la suite, le Dr C. Isel a introduit trois mutations *trans*-compensatoires dans chacun des ARNv partenaires (figure 58A, interaction de droite). Les mutations, situées dans les régions codantes des deux segments, ont été conçues de manière à être silencieuses afin d'être testées dans un contexte infectieux. Les ARNv NA et M mutés (ARNv NA mc et M mc) interagissent faiblement avec l'ARNv partenaire sauvage (38% et 29% de complexe restant par rapport à l'interaction sauvage) (figure 58B, pistes 6 et 7), tandis que la co-transcription des deux segments mutés restaure l'interaction à 87% (figure 58B, piste 8). Afin d'évaluer le rôle de cette interaction, identifiée et validée *in vitro*, sur la réplication virale, les cinétiques de réplication des virus sauvage, simples mutants et double mutant avaient été comparées. Comme attendu, les virus simples mutants se répliquent moins bien que le virus sauvage. Cependant, le virus possédant les segments NA et M mutés présente également des défauts de réplication. Ces résultats, qui seront présentés plus en détails plus loin (§ B.I.3.2.a), suggèrent donc que l'interaction déterminée *in vitro* entre les ARNv NA et M est importante pour la réplication virale, mais qu'elle pas restaurée *in viro/in cellula* lorsque les mutations *trans*-compensatoires sont présentes simultanément dans les ARNv NA et M.



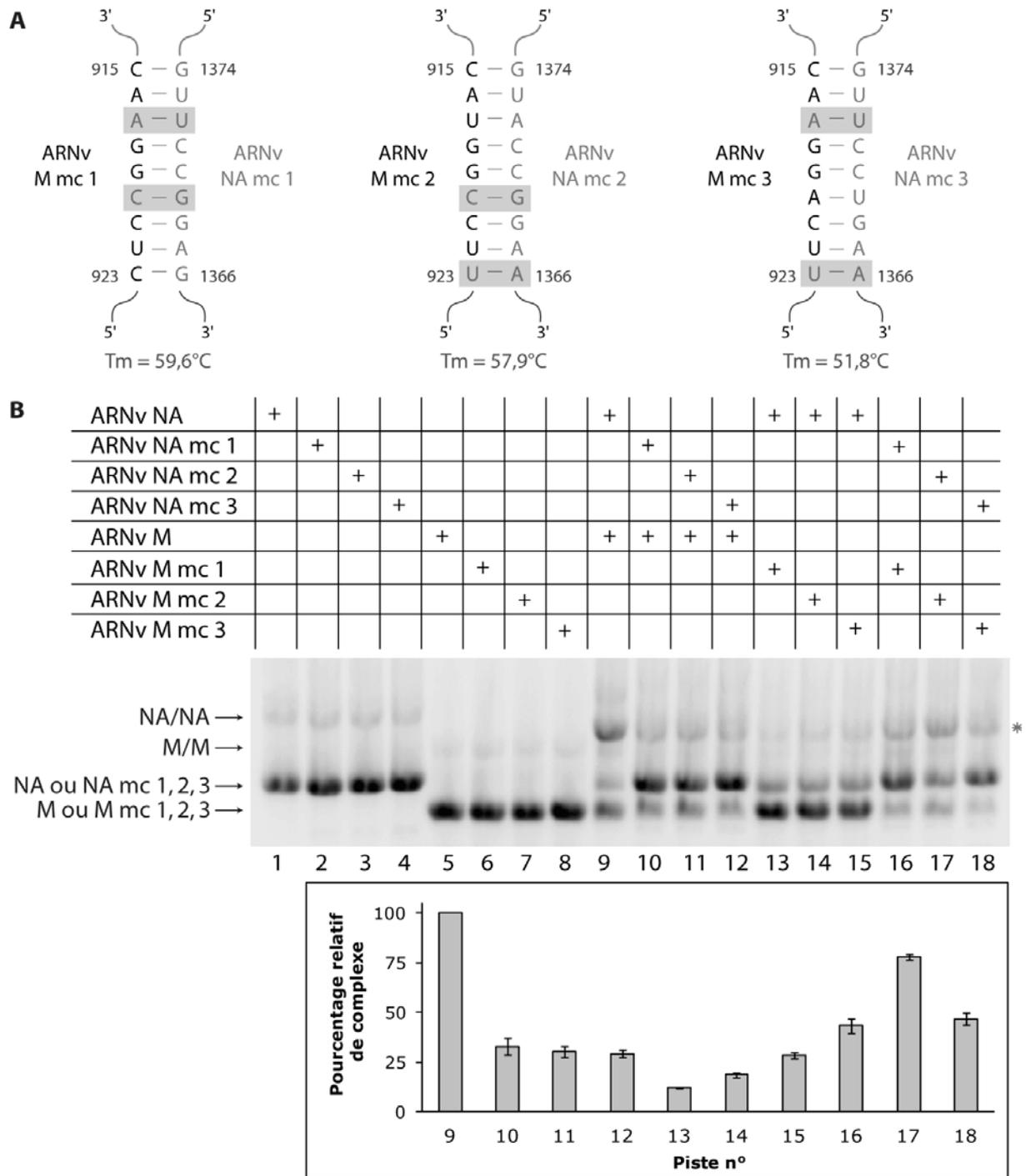
**Figure 58 : Impact de mutations *trans*-compensatoires au niveau des régions impliquées dans l'interaction entre les ARNv NA et M de la souche Moscou sur la formation du complexe.**  
**A.** Interaction potentielle entre les ARNv NA et M et description des trois mutations *trans*-compensatoires et silencieuses introduites (ARNv NA mc et ARNv M mc). La stabilité thermique (Tm) de chaque interaction est indiquée. **B.** Analyse des interactions ARN/ARN par électrophorèse sur gel d'agarose en conditions natives après co-transcription des ARNv NA et M sauvages et des ARNv NA et M portant des mutations *trans*-compensatoires. Les monomères d'ARNv sont indiqués par des flèches et les hétérodimères par un astérisque rouge. La quantification de la proportion de complexe présent dans chaque piste (moyenne de 3 expériences indépendantes ± SEM) a été réalisée.

## 2. Impact des mutations *trans*-compensatoires des ARNv NA et M sur la formation du complexe, *in vitro*

Durant ma thèse, nous avons donc cherché à identifier un autre couple de mutants *trans*-compensatoires et silencieux qui restaure l'interaction *in vitro* entre les deux partenaires mutés, avec l'espoir que cette interaction sera aussi restaurée en contexte infectieux. Dans un premier temps, nous avons comparé les stabilités thermiques des couples sauvage (ARNv NA et M) et muté (ARNv NA mc et M mc) (figure 58A). L'interaction entre les deux ARNv mutés (Tm = 58,5°C) est plus stable que l'interaction sauvage (Tm = 53,6°C) et une perte de stabilité ne peut donc pas expliquer l'absence de restauration de l'interaction entre les mutants NA mc et M mc en culture cellulaire. Sur la base des trois mutations qui avaient été introduites (figure 58A, interaction de droite), nous avons réalisé

des combinaisons de deux mutations *trans*-compensatoires, permettant d'obtenir les ARNv mutés NA mc 1, NA mc 2, NA mc 3 et les ARNv mutés M mc 1, M mc 2, M mc 3 (figure 59A). Les stabilités thermiques de ces trois nouveaux couples sont variables : les couples mc 1 et mc 2 sont plus stables que l'interaction sauvage ( $T_m = 59,6^\circ\text{C}$  et  $57,9^\circ\text{C}$ , respectivement), tandis que le couple mc 3 est moins stable que l'interaction sauvage ( $T_m = 51,8^\circ\text{C}$ ). Nous avons ensuite testé les trois couples par retard de migration sur gel, *in vitro*. Quel que soit le mutant considéré, l'interaction avec le segment partenaire sauvage est diminuée (entre 12% et 33% de complexe restant) (figure 59B, pistes 10-15). La co-transcription des ARNv du couple mc 2 restaure l'interaction à 78% (figure 59B, piste 16), tandis que les ARNv des deux autres couples (mc 1 et mc 3) restaurent l'interaction moins efficacement (respectivement 50% et 48% de complexe) (figure 59B, pistes 17 et 18). Par conséquent, le seul couple qui réponde à notre critère de sélection *in vitro* est le couple d'ARNv NA mc 2 et M mc 2, qui sera évalué dans un contexte infectieux (voir § B.I.3.2.b).

La forte variabilité du taux d'hétérodimérisation (48% à 78%) (figures 58 et 59) entre les différents mutants *trans*-compensatoires suggère que la formation du complexe *in vitro* est fortement dépendante du choix des mutations introduites, qui inclut à la fois le type de mutations (transitions ou transversions), ainsi que l'emplacement des nt au sein de la séquence d'interaction. De manière intéressante, les couples qui restaurent le mieux l'interaction, mc (87% de complexe) (figure 58) et mc 2 (78% de complexe) (figure 59), ont une stabilité thermique similaire ( $T_m = 58,5^\circ\text{C}$  et  $57,9^\circ\text{C}$ , respectivement), mais néanmoins supérieure au couple sauvage ( $T_m = 53,6^\circ\text{C}$ ). *A contrario*, que la stabilité soit nettement supérieure à l'interaction sauvage ( $T_m = 59,6^\circ\text{C}$  pour le couple mc 1) ou plus faible ( $T_m = 51,8^\circ\text{C}$  pour le couple mc 3), ne semble pas jouer sur la formation du complexe (43% et 47% de complexe, respectivement) (figure 59).



**Figure 59 : Impact de différentes mutations *trans*-compensatoires au niveau des régions impliquées dans l'interaction entre les ARNv NA et M de la souche Moscou sur la formation du complexe.** **A.** Description des deux mutations *trans*-compensatoires et silencieuses introduites au niveau de la séquence d'interaction des ARNv NA et M. Les trois couples construits (ARNv NA mc 1, NA mc 2, NA mc 3, M mc 1, M mc 2 et M mc 3) possèdent une combinaison de deux mutations sur les trois présentes dans le mutant *trans*-compensatoire initial mc (figure 58A). La stabilité thermique ( $T_m$ ) de chaque interaction est indiquée. **B.** Analyse des interactions ARN/ARN par électrophorèse sur gel d'agarose en conditions natives après co-transcription des ARNv NA et M sauvages et des ARNv NA et M portant des mutations *trans*-compensatoires. Les monomères d'ARNv et les homodimères sont indiqués par des flèches et les hétérodimères par un astérisque rouge. La quantification de la proportion de complexe présent dans chaque piste (moyenne de 3 à 4 expériences indépendantes  $\pm$  SEM) a été réalisée.

### **3. Impact des mutations *trans*-compensatoires des ARNv NA et M sur la réplication virale**

Dans un second temps, nous avons voulu évaluer l'impact des mutations *trans*-compensatoires et silencieuses introduites dans les segments NA et M sur la réplication virale. Nous nous sommes intéressés au couple mc (figure 58A) afin de confirmer les résultats préliminaires obtenus précédemment, ainsi qu'au couple mc 2, qui restaure l'interaction à un niveau plus proche de la situation sauvage que les deux autres couples construits (figure 59A). Dans les deux cas, nous avons comparé les cinétiques de réplication de virus sauvage, simples mutants (où l'un des deux segments est muté de manière à détruire l'interaction entre les ARNv) et double mutant (où les deux segments sont mutés, de manière à restaurer l'interaction entre les ARNv). J'ai initié ces expériences au cours de mon dernier séjour au laboratoire « Virologie et pathologie humaine » (CIRI, Lyon), mais elles n'ont pas pu être menées à terme lors de ce séjour et ont été poursuivies par T. Julien et le Dr A. Pizzorno après mon départ.

#### **3.1. Production des virus sauvage et mutants par génétique inverse et amplification des virus sur cellules MDCK**

Dans le cadre de cette étude, huit virus différents ont été produits par génétique inverse « forcée » à partir des plasmides de génétique inverse (pHW2000) sauvages (les virus Moscou) ou en présence des plasmides de génétique inverse NA ou/et M dans lesquels trois ou deux mutations *trans*-compensatoires et silencieuses décrites dans les figures 58A et 59A ont été introduites par mutagénèse dirigée (les virus simples mutants NA mc, NA mc 2, M mc et M mc 2, et les virus doubles mutants NA mc/M mc et NA mc 2/M mc 2) (tableau 16). Deux virus sauvages Moscou ont été générés, car les virus mutants des couples mc et mc 2 n'ont pas été produits sur le même lot de cellules HEK293T et MDCK. Ainsi, chaque jeu de virus mutés est associé à un virus sauvage Moscou (tableau 16). Après deux passages successifs sur cellules MDCK, les stocks viraux obtenus ont été titrés en DICT<sub>50</sub> (tableau 16). Dans ce cas, le titrage en PFU n'a pas été réalisé, car nous avons utilisé les mêmes conditions que lors de l'étude préliminaire des cinétiques de réplication des virus du couple d'ARNv mc, où le volume nécessaire à l'infection avait été déterminé à partir du titre en DICT<sub>50</sub>.

Virus	Composition génomique								Titre viral
	ARNv PB2	ARNv PB1	ARNv PA	ARNv HA	ARNv NP	ARNv NA	ARNv M	ARNv NS	DICT <sub>50</sub> /ml
Moscou	wt	wt	wt	wt	wt	wt	wt	wt	4,3.10 <sup>7</sup>
NA mc	wt	wt	wt	wt	wt	mc	wt	wt	3,2.10 <sup>7</sup>
M mc	wt	wt	wt	wt	wt	wt	mc	wt	3,2.10 <sup>7</sup>
NA mc/M mc	wt	wt	wt	wt	wt	mc	mc	wt	1,1.10 <sup>7</sup>
Moscou	wt	wt	wt	wt	wt	wt	wt	wt	2,6.10 <sup>7</sup>
NA mc 2	wt	wt	wt	wt	wt	mc 2	wt	wt	2.10 <sup>7</sup>
M mc 2	wt	wt	wt	wt	wt	wt	mc 2	wt	3,2.10 <sup>7</sup>
NA mc 2/M mc 2	wt	wt	wt	wt	wt	mc 2	mc 2	wt	3,2.10 <sup>7</sup>

**Tableau 16 : Composition génomique et titre viral des virus sauvage et mutants produits par génétique inverse « forcée » dans le cadre de l'étude de l'interaction entre les ARNv NA et M de la souche Moscou.** wt : segment sauvage de la souche Moscou, mc : segment possédant trois mutations *trans*-compensatoires, mc 2 : segment possédant deux mutations *trans*-compensatoires.

En parallèle, les ARNv de chacun des stocks de virus ont été extraits, et nous avons partiellement séquencé les ARNv NA et M afin de vérifier que la séquence d'interaction définie sur chacun des ARNv n'a pas été mutée au cours de la réplication.

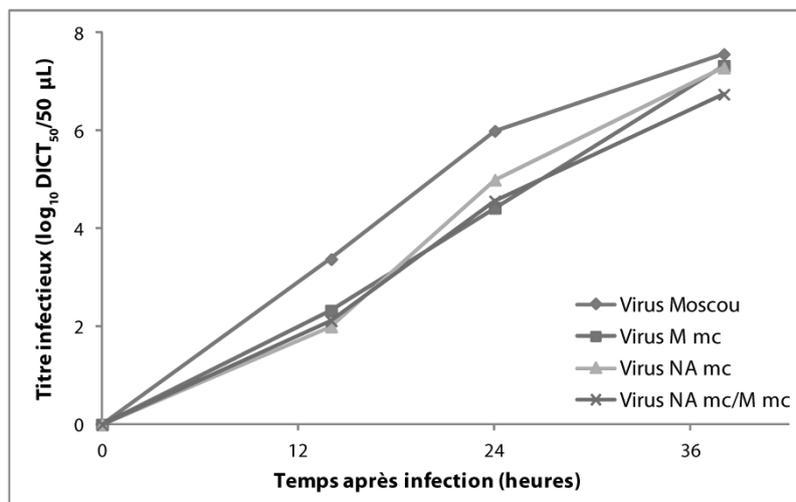
### 3.2. Caractérisation des cinétiques de réplication virale sur cellules MDCK

#### a. Étude du couple d'ARNv mc

En premier lieu, nous avons cherché à confirmer les résultats des cinétiques de réplication qui avaient été obtenus. Des cellules MDCK à confluence ont été infectées avec les virus Moscou, NA mc, M mc ou NA mc/M mc afin de comparer leurs cinétiques de réplication virale. Contrairement au virus PR8, les cinétiques de la souche Moscou sont généralement réalisées à une MOI de 10<sup>-4</sup> (un virus pour 10<sup>4</sup> cellules) et le volume de chaque stock viral nécessaire à l'infection est calculé à partir du titre viral exprimé en DICT<sub>50</sub>/ml (tableau 16). Après infection, des fractions aliquotes des surnageants de culture ont été prélevées au cours du temps, puis le titre viral a été déterminé par la méthode de DICT<sub>50</sub>.

Pour chacun des virus, le titre infectieux augmente au cours du temps (figure 60). Le titrage des surnageants récoltés 14h après infection montre que le virus Moscou atteint un titre viral plus élevé (10<sup>3,4</sup> DICT<sub>50</sub>/50 µl) que les virus simples et double mutants (entre 10<sup>2</sup> et 10<sup>2,3</sup> DICT<sub>50</sub>/50 µl). De la même manière, 24h après infection, le titre du virus sauvage est de 10<sup>6</sup> DICT<sub>50</sub>/50 µl, tandis que le titre des trois autres virus varie entre 10<sup>4,4</sup> et 10<sup>5</sup> DICT<sub>50</sub>/50 µl. Néanmoins, 38h après infection, les titres des virus simples mutants (10<sup>7,3</sup> DICT<sub>50</sub>/50 µl) sont

proches de celui du virus sauvage ( $10^{7,6}$  DICT<sub>50</sub>/50  $\mu$ l). Seul le titre du virus double mutant reste légèrement plus faible ( $10^{6,8}$  DICT<sub>50</sub>/50  $\mu$ l). Ainsi, 14h et 24h après infection, le titre entre le virus sauvage et les trois autres virus varie entre 1,4 et 1,6 log au maximum. Nous sommes à la limite de ce qui est considéré comme significatif. Ces résultats suggèrent que les virus simples et double mutants possèdent des défauts réplcatifs, en accord avec les résultats précédemment obtenus au laboratoire. Il sera néanmoins essentiel de répéter ces expériences afin de conclure sur les capacités réplcatives de chacun des virus.



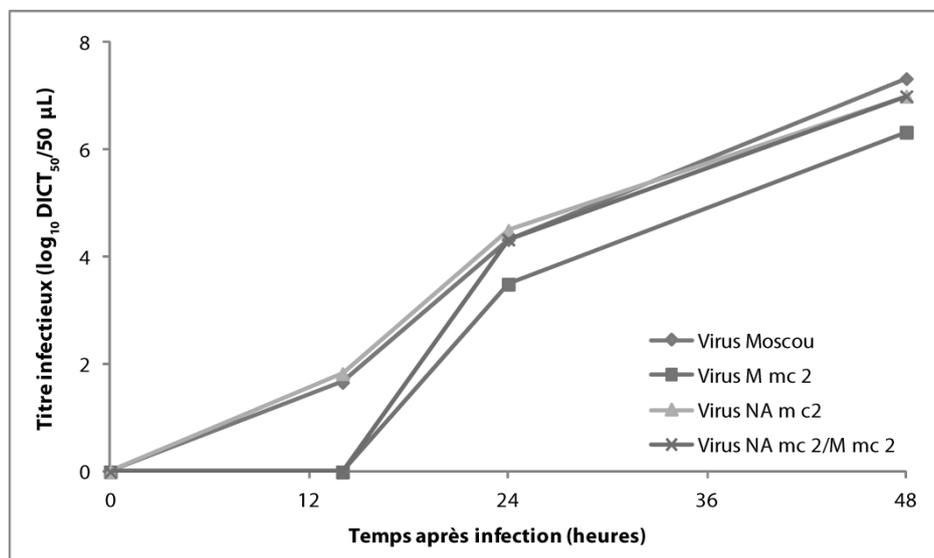
**Figure 60 : Impact de mutations *trans*-compensatoires au niveau des régions impliquées dans l'interaction entre les ARNv NA et M de la souche Moscou sur la réplcation virale.** Les virus mutants possèdent trois mutations *trans*-compensatoires et silencieuses. Les cellules MDCK ont été infectées à une MOI de  $10^{-4}$ . Le titre viral est représenté en  $\log_{10}$  DICT<sub>50</sub>/50  $\mu$ l.

### b. Étude du couple d'ARNv mc 2

Dans un deuxième temps, nous avons cherché à évaluer si l'introduction de deux mutations sur les trois que contenait le couple mc permet de restaurer l'interaction *in viro*, si elle existe, et donc d'obtenir un phénotype sauvage. Ainsi, des cellules MDCK à confluence ont été infectées avec les virus Moscou, NA mc 2, M mc 2 ou NA mc 2/M mc 2 à une MOI de  $10^{-4}$  afin de comparer leurs cinétiques de réplcation virale. Après infection, des fractions aliquotes des surnageants de culture ont été prélevées au cours du temps, puis le titre viral a été déterminé par la méthode de DICT<sub>50</sub>.

Pour chacun des virus, le titre infectieux augmente au cours du temps (figure 61). Le titrage des surnageants prélevés 14h après infection avec les virus M mc 2 ou NA mc 2/M mc

2 n'ont révélé aucun ECP, même dans la première dilution déposée. Par conséquent, ce point a été représenté, par défaut, comme étant égal à zéro (figure 61). Considérant la gamme de dilutions utilisée dans cet essai nous pouvons néanmoins affirmer que le titre de ces virus est inférieur à  $10^{0,5}$  DICT<sub>50</sub>/50 µl. Une détermination plus précise du titre de ces deux virus peut être effectuée en réalisant une nouvelle série de titrages en diluant le virus au 1/2. Il semblerait qu'à ce premier point de prélèvement les titres des virus sauvage et NA mc 2 ( $10^{1,67}$  et  $10^{1,83}$  DICT<sub>50</sub>/50 µl, respectivement) soient plus élevés que ceux des deux autres virus, même s'il est difficile de conclure avant d'avoir déterminé le titre des virus M mc 2 et NA mc 2/M mc 2. Au cours de la cinétique, le titre du virus simple mutant NA mc 2/M mc 2 est similaire à celui des virus Moscou et NA mc 2 (entre  $10^{4,33}$  et  $10^{4,5}$  DICT<sub>50</sub>/50 µl à 24h et entre  $10^7$  et  $10^{7,33}$  DICT<sub>50</sub>/50 µl à 48h), tandis que le titre du virus simple mutant M mc 2 reste plus faible ( $10^{3,5}$  DICT<sub>50</sub>/50 µl à 24h et  $10^{6,33}$  DICT<sub>50</sub>/50 µl à 48h). À ces deux temps, les titres ne sont néanmoins pas significativement différents (entre 0,7 et 0,8 log). D'après ces résultats, les capacités répliquatives des virus NA mc 2 et NA mc 2/M mc 2 sont semblables à celle du virus Moscou. Le virus simple mutant M mc 2 possède un léger défaut répliquatif, qui n'est toutefois pas significatif.



**Figure 61 : Impact de mutations *trans*-compensatoires au niveau des régions impliquées dans l'interaction entre les ARNv NA et M de la souche Moscou sur la répllication virale.** Les virus mutants possèdent deux mutations *trans*-compensatoires et silencieuses. Les cellules MDCK ont été infectées à une MOI de  $10^{-4}$ . Le titre viral est représenté en  $\log_{10}$  DICT<sub>50</sub>/50 µl. Aucun ECP n'a été observé 14h après infection pour les virus M mc 2 et NA mc 2/M mc 2, le titre viral a donc été représenté, par défaut, égal à zéro.

#### 4. Conclusion

D'après l'ensemble de ces données, la présence d'un seul partenaire possédant les trois mutations *trans*-compensatoires et silencieuses NA mc ou M mc affecte la réplication virale de manière modeste, vraisemblablement en déstabilisant l'interaction, *in viro*, entre les ARNv NA et M. La présence des deux ARNv mutés ne restaure pourtant pas des capacités répliquatives sauvages. À l'inverse, l'introduction de deux de ces mutations dans les ARNv ne semble pas suffisante pour détruire significativement l'interaction, *in viro*, entre les deux partenaires. Les résultats des cinétiques des mutants mc 2 sont encore préliminaires et restent à confirmer. De plus, il est primordial d'évaluer l'impact des mutations des couples mc et mc 2 dans l'empaquetage des ARNv partenaires, par génétique inverse compétitive. Si les trois mutations du couple mc sont importantes pour la formation du complexe entre les ARNv NA et M, nous nous attendons à favoriser la production de virus possédant les deux segments capables d'interagir efficacement, à savoir les deux ARNv sauvages ou les deux ARNv mutés. Néanmoins, si les deux mutations du couple mc 2 ne sont pas suffisantes pour abolir l'interaction entre les deux segments, la présence de l'ARNv sauvage (ou muté) ne devrait pas guider l'empaquetage dans les virions du partenaire sauvage (ou muté).

Pour conclure, les résultats *in viro* de réplication virale ne reflètent pas les résultats *in vitro* de retard de migration sur gel. D'autres études sont nécessaires afin de conclure sur le rôle des trois nt mutés dans le couple d'ARNv NA mc et M mc dans la formation du complexe.

## II. Étude de l'interaction entre les ARNv PB1 et NA de la souche Udorn

### 1. Contexte de l'étude

Une étude de réassortiment génétique entre un virus modèle saisonnier A/Udorn/307/72 (H3N2) et le virus PR8 a révélé que la majorité des virus réassortants qui possèdent les segments HA et NA Udorn contiennent également le segment PB1 Udorn, malgré l'absence d'avantage répliquatif conféré par cette combinaison de gènes [270]. Afin de comprendre quels mécanismes permettent à ce virus d'émerger, l'équipe du Dr L. Brown a réalisé des expériences de génétique inverse compétitive (voir Introduction, § D.I.1.3.c et figure 4 de la revue [263]) pour analyser les fréquences d'incorporation de segments de

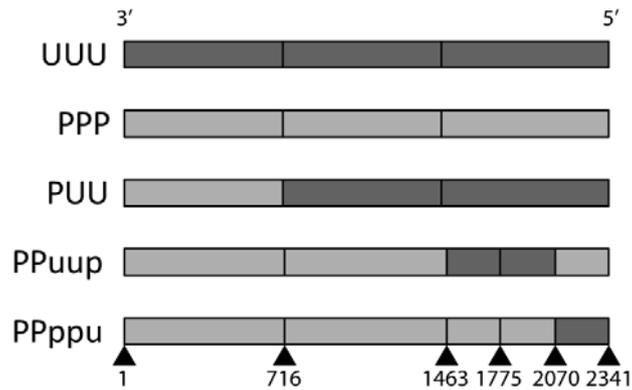
gènes viraux [272]. Ces travaux ont montré que cette constellation génomique particulière s'explique par un co-empaquetage des ARNv NA et PB1, dans lequel HA n'intervient pas, guidé par une région centrale, codante, de l'ARNv PB1 Udorn et non *via* les signaux d'empaquetage identifiés dans la littérature [272].

Dans le but d'identifier la région de l'ARNv PB1 impliquée dans le phénomène de co-ségrégation observé, des gènes PB1 chimériques possédant, par tiers, des séquences Udorn ou PR8 (figure 1 de l'article), ont été testés en génétique inverse compétitive. Ces expériences ont révélé que le dernier tiers de PB1 (nt 1464-2341) est responsable de la sélection préférentielle du segment PB1 en présence du segment NA Udorn (Ud-NA) (figure 4 de l'article). Afin de délimiter plus finement la région de l'ARNv PB1 impliquée dans cette interaction, des chimères PR8/Udorn dans ce dernier tiers de PB1 ont été construites (figure 5 de l'article) et testées (figure 7 de l'article). Les résultats montrent que les chimères possédant la séquence de l'ARNv PB1 Udorn entre les positions 1776 et 2070 sont sélectionnées de manière préférentielle lorsque le segment Ud-NA est présent.

Durant ma dernière année de thèse, nous avons collaboré avec l'équipe du Dr L. Brown afin de vérifier si la région impliquée dans la co-ségrégation des ARNv PB1 et NA dans un contexte infectieux peut être assimilée à une région d'interaction ARN/ARN. Le but de notre travail était de tester, dans notre système *in vitro*, les interactions entre les ARNv NA Udorn et PB1 Udorn, sauvages et mutés, afin d'évaluer si les résultats obtenus *in viro* peuvent être expliqués par des aptitudes différentes des ARNv à interagir entre eux. Nous voulions également identifier plus précisément la région du segment PB1 impliquée dans ce mécanisme.

## **2. Interaction *in vitro* entre les ARNv PB1 et NA Udorn**

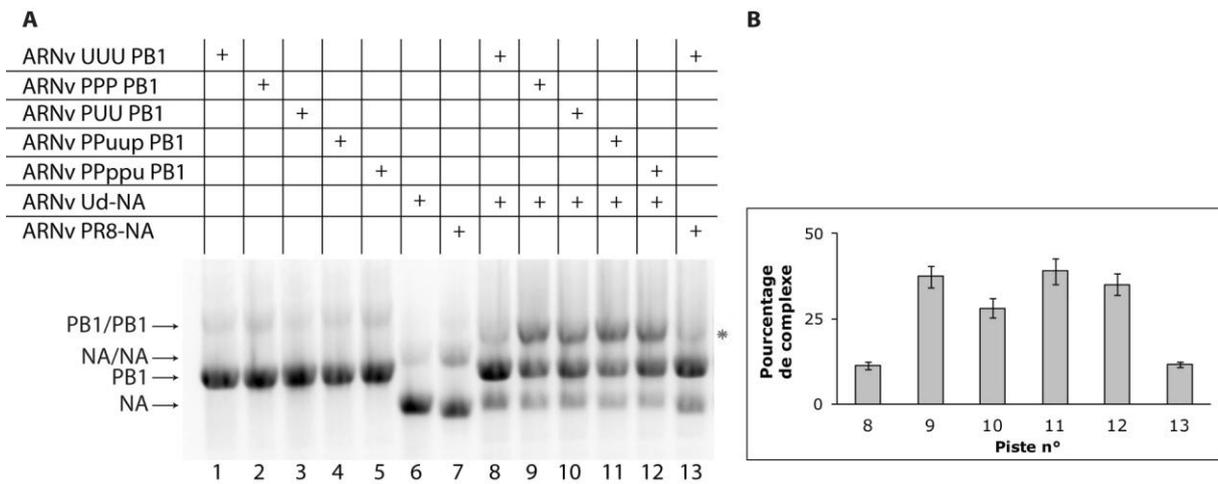
Nous disposons au laboratoire des ARNv sauvages PB1 et NA, correspondant aux séquences de la souche Udorn ou PR8 (respectivement UUU PB1, PPP PB1, Ud-NA et PR8-NA) et de chimères du segment PB1. Treize chimères au total ont été construites par l'équipe du Dr L. Brown et contiennent des combinaisons des régions codantes de l'ARNv PB1 des souches Udorn et PR8 (figures 1 et 5 de l'article). Trois de ces chimères (figure 62) nous ont été fournies et ont été utilisées pour les tests *in vitro*.



**Figure 62 : Représentation schématique des chimères des ARNv PB1 testées *in vitro*.** Le nom des chimères indique l'origine de chaque section de la séquence de l'ARNv PB1 : U correspond à la séquence de Udorn (bleu) et P à la séquence de PR8 (orange) ; les majuscules indiquent l'origine du tiers et les minuscules l'origine de chaque petite section du dernier tiers de l'ARNv PB1. Les bornes de chaque section sont indiquées.

Il nous a fallu en premier lieu définir les conditions expérimentales les mieux adaptées (co-transcription d'ARNv ou co-incubation d'ARNv préalablement purifiés) à la détection de l'interaction entre les segments PB1 et NA de la souche Udorn.

En co-transcription, les ARNv UUU PB1 et Ud-NA interagissent faiblement (11% de complexe) (figure 63, piste 8). Lorsque le segment UUU PB1 est co-transcrit avec un ARNv NA hétérologue (PR8-NA ou Moscou-NA), le complexe atteint également 11% (figure 63, piste 13 ; résultat non présenté pour Moscou-NA). De manière surprenante, les interactions sont plus fortes entre le segment Ud-NA et l'ARNv PPP PB1 (37% de complexe) (figure 63, piste 9) ou les chimères de PB1 (28% à 39% de complexe) (figure 63, pistes 10 à 12), que celle observée entre les segments homologues Ud-NA et UUU PB1 (figure 63, piste 8). Dans ces conditions, tous les ARNv homodimérisent.

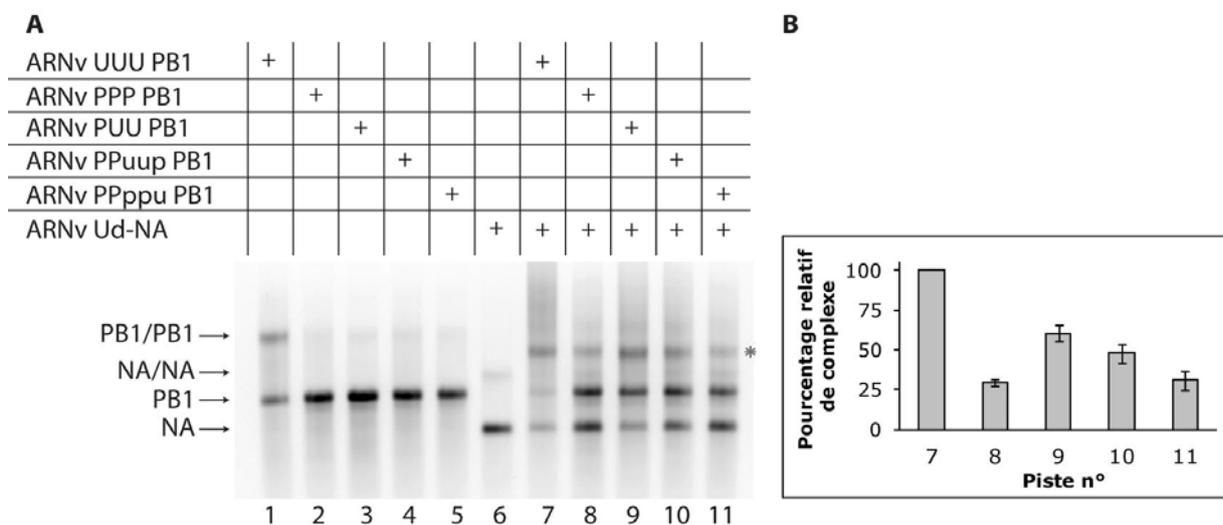


**Figure 63 : Interactions ARN/ARN entre les ARNv NA et PB1, sauvages ou chimériques, des souches Udorn et PR8, après co-transcription des ARNv par paires. A.** Analyse des interactions ARN/ARN par électrophorèse sur gel d'agarose en conditions natives après co-transcription des ARNv PB1, sauvages ou chimériques, et des ARNv NA sauvages. Les monomères et les homodimères d'ARNv sont indiqués par des flèches et les hétérodimères par un astérisque rouge. **B.** Quantification de la proportion de complexe (%) par rapport à la quantité totale d'ARNv par piste (moyenne de 4 à 6 expériences indépendantes  $\pm$  SEM).

Nous avons ensuite co-incubé les ARNv préalablement purifiés dans les mêmes conditions que celles choisies pour identifier le réseau d'interactions entre les ARNv de la souche PR8 (co-incubation à 55°C dans un tampon de faible salinité). Les ARNv UUU PB1 et Ud-NA forment 65% de complexe dans ces conditions (figure 64, piste 7), soit une interaction considérablement plus forte que celle observée en co-transcription (figure 63, piste 8). Les ARNv homodimérisent également dans ces conditions expérimentales, principalement le segment UUU PB1 (figure 64, piste 1), mais cet homodimère est peu visible lorsque l'ARNv est incubé en présence de Ud-NA (figure 64, piste 7). Nous avons donc poursuivi l'étude des interactions entre les ARNv PB1 et NA des souches PR8 et Udorn par co-incubation des ARNv purifiés.

Le pourcentage d'interaction entre les ARNv UUU PB1 et Ud-NA constitue notre référence par rapport à laquelle les autres interactions ont été normalisées. Comme attendu, les segments PPP PB1 et NA-Ud interagissent peu comparé à notre référence (30% de complexe) (figure 64, piste 8). L'ARNv PUU PB1, qui était majoritairement sélectionné (67%) par rapport à l'ARNv UUU PB1 dans les études de génétique inverse compétitive à neuf plasmides (figure 4A de l'article), est également celui qui interagit le mieux *in vitro* avec l'ARNv NA comparé à notre référence (60% de complexe) (figure 64, piste 9). Étant donné

que les ARNv possédant le dernier tiers de PB1 sont préférentiellement sélectionnés *in viro*, nous avons ensuite comparé la capacité des ARNv PB1 PPUup et PPpu à interagir *in vitro* avec l'ARNv Ud-NA. En accord avec les résultats *in viro*, la chimère PPUup, qui contient les nt 1464-2070 de Udorn, interagit davantage avec Ud-NA (48% de complexe comparé à notre référence) (figure 64, piste 10) que la chimère PPpu (31% de complexe comparé à notre référence) (figure 64, piste 11). L'ensemble de ces données suggère que les nt 1776-2070 du segment PB1 Udorn sont responsables de l'interaction avec l'ARNv partenaire NA, ce qui expliquerait la co-ségrégation des deux ARNv dans les expériences de réassortiment génétique et de génétique inverse compétitive à neuf plasmides.



**Figure 64 : Analyse des interactions ARN/ARN entre les ARNv NA et PB1, sauvages ou chimériques, des souches Udorn et PR8, après co-incubation des ARNv. A.** Analyse des interactions ARN/ARN par électrophorèse sur gel d'agarose en conditions natives après co-incubation des ARNv PB1, sauvages ou chimériques, et des ARNv NA sauvages à 55°C dans un tampon de faible salinité. Les monomères et les homodimères d'ARNv sont indiqués par des flèches et les hétérodimères par un astérisque rouge. **B.** Quantification de la proportion de complexe présent dans chaque piste (moyenne de 4 à 5 expériences indépendantes  $\pm$  SEM).

### 3. Conclusion

L'ensemble des résultats obtenus dans le cadre de notre collaboration suggère que, dans certains cas, les interactions ARN/ARN détectées *in vitro* pourraient prédire les interactions entre les RNPv au sein du complexe supramoléculaire, qui se forme avant ou pendant le bourgeonnement des particules virales. L'interaction entre les ARNv PB1 et NA avait également été observée pour les souches Moscou [308] et England [319]. Cependant, dans ces deux conditions, le pourcentage de complexe (15% et 9%, respectivement) est bien

plus faible que celui atteint, en co-incubation, entre les segments UUU PB1 et Ud-NA (65% de complexe) (figure 64, piste 7). Ainsi, par le passé, il a été difficile d'établir l'importance de cette interaction. Néanmoins, dans cette étude, 52% des ARNv PB1 et NA forment des hétérodimères (n=5, 47%-65% de complexe) (figure 64), ce qui indique une forte interaction entre ces deux segments.

Afin de caractériser cette interaction de manière plus précise, nous avons cherché une complémentarité de séquence antiparallèle entre les nt 1776-2070 du segment PB1 et la totalité du segment Ud-NA, tout en analysant les structures secondaires prédites. Les séquences des ARNv PB1 et NA qui ont co-ségrégé dans les souches pandémiques H2N2 et H3N2 ont également été prises en compte. Les appariements proposés suite à l'application de toutes ces contraintes varient entre 7 et 22 pb. Afin d'identifier la séquence d'interaction, au nt près, entre les ARNv PB1 et NA, il convient désormais de tester l'effet d'oligonucléotides antisens complémentaires aux deux ARNv sur la formation du complexe. Ce travail est en cours.

#### **4. Article associé**

Article

# Influenza NA and PB1 Gene Segments Interact during the Formation of Viral Progeny: Localization of the Binding Region within the PB1 Gene

Brad Gilbertson <sup>1</sup>, Tian Zheng <sup>1</sup>, Marie Gerber <sup>2</sup>, Anne Printz-Schweigert <sup>2</sup>, Chi Ong <sup>3</sup>, Roland Marquet <sup>2</sup>, Catherine Isel <sup>2,4</sup>, Steven Rockman <sup>1,3</sup> and Lorena Brown <sup>1,\*</sup>

<sup>1</sup> Department of Microbiology and Immunology, The University of Melbourne at the Peter Doherty Institute of Infection and Immunity, Parkville 3010, Victoria, Australia; bg@unimelb.edu.au (B.G.); tianchi.zheng@gmail.com (T.Z.); Steve.Rockman@seqirus.com (S.R.)

<sup>2</sup> Architecture et Réactivité de l'ARN, Université de Strasbourg, CNRS, IBMC, 15 rue René Descartes, Strasbourg 67084, France; m.gerber@ibmc-cnrs.unistra.fr (M.G.); A.schweigert@ibmc-cnrs.unistra.fr (A.P.-S.); r.marquet@ibmc-cnrs.unistra.fr (R.M.); c.isel@ibmc-cnrs.unistra.fr (C.I.)

<sup>3</sup> Seqirus, 63 Poplar Rd, Parkville 3052, Victoria, Australia; Chi.Ong@seqirus.com

<sup>4</sup> Unité de Génétique Moléculaire des Virus à ARN, Département de virologie, Institut Pasteur, Paris 75005, France

\* Correspondence: lorena@unimelb.edu.au; Tel.: +61-3-8344-3865

Academic Editor: Eric O. Freed

Received: 28 June 2016; Accepted: 16 August 2016; Published: 20 August 2016

**Abstract:** The influenza A virus genome comprises eight negative-sense viral RNAs (vRNAs) that form individual ribonucleoprotein (RNP) complexes. In order to incorporate a complete set of each of these vRNAs, the virus uses a selective packaging mechanism that facilitates co-packaging of specific gene segments but whose molecular basis is still not fully understood. Recently, we used a competitive transfection model where plasmids encoding the A/Puerto Rico/8/34 (PR8) and A/Udorn/307/72 (Udorn) PB1 gene segments were competed to show that the Udorn PB1 gene segment is preferentially co-packaged into progeny virions with the Udorn NA gene segment. Here we created chimeric PB1 genes combining both Udorn and PR8 PB1 sequences to further define the location within the Udorn PB1 gene that drives co-segregation of these genes and show that nucleotides 1776–2070 of the PB1 gene are crucial for preferential selection. In vitro assays examining specific interactions between Udorn NA vRNA and purified vRNAs transcribed from chimeric PB1 genes also supported the importance of this region in the PB1-NA interaction. Hence, this work identifies an association between viral genes that are co-selected during packaging. It also reveals a region potentially important in the RNP-RNP interactions within the supramolecular complex that is predicted to form prior to budding to allow one of each segment to be packaged in the viral progeny. Our study lays the foundation to understand the co-selection of specific genes, which may be critical to the emergence of new viruses with pandemic potential.

**Keywords:** influenza virus; packaging; reassortment; RNA-RNA interaction; gene segments; viral polymerase; competitive transfection

## 1. Introduction

The influenza A virus (IAV) genome comprises eight individual ribonucleoprotein (RNP) complexes each with a single-stranded negative-sense viral RNA (vRNA) segment encoding one or more proteins. In each RNP complex, the vRNA is encapsidated by nucleoprotein (NP) and the 3' and 5' termini show inverted complementarity and associate to form a double stranded structure referred to as the panhandle [1]. On each segment, the termini are attached to the viral RNA-dependent

RNA polymerase complex consisting of three subunits: polymerase basic protein 1 (PB1), polymerase basic protein 2 (PB2) and polymerase acidic protein (PA) [2,3]. Each RNP is thus an independent unit for transcription and replication of the associated vRNA [4,5].

Each gene segment encodes at least one essential viral protein, therefore complicating packaging of the viral genome, as the virus must incorporate one copy of each vRNA to be fully infectious. Electron microscopy [6] and tomography [7–9] have repeatedly revealed that the eight RNPs of budding IAVs are organized in a specific '7 + 1' pattern, with seven RNPs surrounding a central segment. Recently, single-molecule fluorescent in situ hybridization analysis directly showed that only one copy of each RNA segment is packaged into most viral particles [10]. Similarly, in a reverse genetics system that included two virus-like RNAs encoding different reporter proteins within identical vRNA segments, plaques formed from progeny virions displayed only one or the other reporter protein but not both, further showing that individual influenza virions incorporate single, not multiple, copies of each vRNA segment [11]. The mechanism underlying a now well established selective rather than random packaging process has recently started to be unveiled. *Cis*-acting packaging sequences, located in both the 3' and 5' noncoding regions as well as the terminal coding regions of each RNA segment have been identified in all eight vRNA segments, for the A/WSN/33 (H1N1) and A/Puerto Rico 8/1934 (H1N1) (PR8) strains (for reviews, see [12,13]). Mutations introduced into these sequences dramatically reduce the incorporation efficiency of both the mutated vRNA and of other vRNAs into progeny virions, highlighting the importance of these regions and suggesting the existence of intersegment interactions critical for proper packaging [14–16]. Direct vRNA/vRNA interactions as guides for selective genome packaging [12] warranted investigation. Using an in vitro RNA/RNA interaction assay, the eight vRNA segments from the human H3N2 (A/Moscow/10/99) and avian H5N2 (A/Finch/England/205/91) viruses were shown to be organized into a limited number of strain-specific supramolecular assemblies maintained by a network of intermolecular interactions where each vRNA segment interacted with at least one other partner [17,18]. In the case of H5N2, sequences involved in vRNA/vRNA interactions were found to be non-conserved and not exclusively located at the termini but rather spread along the viral RNPs (vRNPs) [18].

While the segmented nature of the viral genome requires a highly sophisticated packaging strategy, it provides IAVs with a significant evolutionary advantage. Indeed, it allows genetic reassortment when two or more influenza viruses co-infect the same cell, creating hybrid viruses containing a mixture of segments from the two parental strains. This process is a major contributing factor in the emergence of novel pandemic strains, whereby IAVs acquire novel hemagglutinin (HA) and neuraminidase (NA) to which the human population is essentially naïve. Although the theoretical number of genotypes that can emerge from natural or experimental reassortment is 256 ( $2^8$ ), the whole panel of different gene combinations has never been observed [19–21] and certain genes tend to co-segregate [20,22–24], suggesting that genetic reassortment is biased [13,22,24]. Documented incompatibilities between viral proteins or viral and host proteins explain the absence of some gene constellations [21,25–28] and viruses engineered to differ slightly in gene but not protein sequences have been found to reassort very efficiently and without genetic bias [29]. However, it is possible that incompatibilities at the genomic level, most likely between packaging signals, may also shape the final constellation of genes. The ability to rescue efficiently replicating viruses by reverse genetics that do not emerge during reassortment would support this. Hence, we have postulated that the drivers of packaging of influenza gene segments can also impact on the reassortment process. Accordingly, Essere et al. [22] recently demonstrated that during in vitro reassortment between a human H3N2 and an avian H5N2 virus, the HA and M gene segments co-segregate and that genetic reassortment is limited by sub-optimal compatibilities between packaging signals. Similarly, it was observed that H3N2 vaccine seeds, derived by classical reassortment with the egg-adapted PR8 H1N1 virus and selected with antisera to PR8 to yield progeny containing H3N2 HA and NA, frequently contain the PB1 gene from the H3N2 strain [30–32]. This finding was recapitulated using the model seasonal H3N2 virus Udorn and is noteworthy because, despite its preferential selection in the reassortment progeny, reverse genetics

experiments showed that the virus containing the Udorn PB1, in addition to Udorn HA and NA, did not replicate as well as the corresponding virus containing the PR8 PB1 segment [32]. This indicated that inclusion of the Udorn PB1 gene was not driven by selection of the fittest virus.

To study this phenomenon a competitive nine-plasmid transfection model was developed, where plasmids encoding Udorn HA and NA and the other six genes of PR8 were provided together with an additional Udorn PB1. When given the choice of PB1 genes, packaged progeny viruses predominantly contained the Udorn PB1. Similar experiments linked the preferential co-segregation of Udorn PB1 to the Udorn NA and not HA segment [24]. Therefore it was postulated that incorporation of H3N2 PB1 into vaccine seeds was likely due to this same co-selection mechanism with potential pairing with the H3N2 NA segment. Additional competitive transfection experiments using chimeric PB1 genes in which the 3' and 5' packaging sequences were swapped between PR8 and Udorn PB1 genes revealed that incorporation of PR8 PB1 was governed by previously described packaging sequences present at the 3' end of the vRNA. However, the co-selection of Udorn NA and PB1 segments was not directed by these terminal packaging sequences but rather through interactions involving the internal coding region of the PB1 gene [24].

In this study, we investigate interactions between the PB1 and NA genes that potentially explain why H3N2 NA and PB1 co-segregate. We have created additional chimeric PB1 genes to further delineate the regions within the central coding area of the Udorn PB1 gene that interacts with the Udorn NA gene. This work has identified a region of the PB1 gene that influences the association between viral genes to direct their co-selection during reassortment. This region potentially encompasses sequences driving key RNP-RNP interactions within the supramolecular complex that is predicted to form prior to budding to allow one of each segment to be packaged in the viral progeny.

## 2. Materials and Methods

### 2.1. Cells and Media

Human embryonic kidney (293T) cells and Madin-Darby canine kidney (MDCK) cells were sourced from an existing collection in the Department of Microbiology and Immunology, The University of Melbourne, Australia. 293T cells were grown in Dulbecco's modified Eagle's medium (DMEM) (Life Technologies, Mulgrave, Victoria, Australia) and MDCK cells were grown in RPMI-1640 (Sigma Aldrich, Castle Hill, NSW, Australia). Both media were supplemented with 10% heat-inactivated fetal calf serum (FCS, Life Technologies), 2 mM L-glutamine (Sigma Aldrich), 2 mM sodium pyruvate (Thermo Fisher, Scoresby, Victoria, Australia), 24 µg/mL gentamicin (Pfizer, West Ryde, NSW, Australia), 50 µg/mL streptomycin (Life Technologies) and 50 IU/mL penicillin (Life Technologies). Co-cultures of 293T cells and MDCK cells for transfection were established in Opti-MEM (Life Technologies) with 50 µg/mL streptomycin and 50 IU/mL penicillin.

### 2.2. Construction of Reverse-Engineered Influenza Viruses

Individual gene segments from PR8 and Udorn viruses were reverse transcribed and cloned into pHW2000 plasmids (St. Jude Children's Hospital, Memphis, TN, USA). Viruses were reverse engineered using the eight-plasmid system developed by Hoffmann et al. [33]. The reverse genetics-derived viruses consisted of a parental or chimeric PB1 gene in a genetic background comprising six gene segments from PR8 virus (PB2, PA, HA, NP, M, NS) and the NA gene segment derived from Udorn virus. This backbone virus is referred to as PR8(Ud-NA) and those expressing chimeric PB1 segments collectively as PR8(Ud-NA, chimeric PB1) or specifically by the chimeric PB1 name e.g., PPU or PPpup. The origin of each third of the PB1 gene is denoted by P (PR8) or U (Udorn). Uppercase nomenclature has been used to indicate larger thirds of the PB1; lowercase nomenclature used to indicate the smaller three sections of the last third of the PB1. Rescued viruses were amplified from transfection supernatants in 10-day embryonated hen's eggs. Infectious allantoic fluid was titrated for virus content by plaque formation in MDCK cells [34] and stored at −80 °C.

### 2.3. Construction of Plasmids Expressing Chimeric PB1 Gene Segments

Briefly, plasmids expressing chimeric PB1 genes were created by the construction of fragments of PR8 and Udorn PB1 genes representing approximate thirds of the respective gene and had a 20 nucleotide overlap with the adjacent sequence, conserved in both virus strains. These were amplified by PCR using section-specific primers and joined together by another round of PCR using segment-specific primers containing AarI restriction sites [35]. Full-length chimeric products were then cloned into the pHW2000 expression vector. All clones were confirmed by full-length sequencing. The resulting chimeric PB1 genes were designated PPU, PUU, PUP, UUP, UPP and UPU (representing the PR8 or Udorn origin of each third of the gene). This same approach was used to construct additional chimeras in which the 5' third was further subdivided into approximate thirds to yield PPppu, PPpuu, PPpup, PPuup, PPpup and PPupu (where the lowercase letters designate the origin of thirds within the last 5' section). The primer sequences are available upon request.

### 2.4. Minigenome Assay for Polymerase Activity

A  $\beta$ -lactamase (BLA) reporter assay [36] was used to compare the intrinsic activities of viral RNP complexes as previously described [32]. Briefly, the pCAGGS-BLA reporter plasmid (2 ng/well), in which the BLA gene is flanked by influenza virus segment untranslated regions (UTRs; derived from H1 HA), was transfected into 293T cells, together with 2 ng of four pHW2000 plasmids expressing either the NP gene or one of the three influenza virus polymerase genes (PB1, PB2, or PA); the PB1 being either parental or chimeric. Transfected cells were incubated at 37 °C in 5% CO<sub>2</sub>. At 24 h post transfection, the  $\beta$ -lactamase produced was detected, after lysis of cells, by the addition of the LyticBLazer-FRET B/G substrate (Life Technologies) according to the manufacturer's instructions. Fluorescence was read using a ClarioStar microplate reader (BMG Labtech, Mornington, Victoria, Australia). The relative polymerase activity was calculated as follows: (530 nm/460 nm ratio of the sample)/(530 nm/460 nm ratio of untransfected control).

### 2.5. Determination of Viral Replication Kinetics

The replication characteristics of the reverse-engineered viruses were determined by infecting MDCK cells at a multiplicity of infection (MOI) of 0.01 plaque forming units (pfu)/cell. Following 1 h absorption (at  $t = 0$  h) the inoculum was removed and cells were washed and incubated at 37 °C, 5% CO<sub>2</sub> in RPMI-1640 (Sigma Aldrich) supplemented with 2 mM L-glutamine, 2 mM sodium pyruvate, 24  $\mu$ g/mL gentamicin, 50  $\mu$ g/mL streptomycin and 50 IU/mL penicillin and 1  $\mu$ g/mL of TPCK-trypsin (Worthington Biochemical Corporation, Lakewood, NJ, USA). Cell culture supernatants were harvested at various time points post infection and stored at  $-80$  °C for analysis. Viral titres of the supernatants were determined by plaque formation on confluent monolayers of MDCK cells.

### 2.6. Nine-Plasmid Competitive Transfections

These were undertaken using a modified version of the eight-plasmid reverse genetics system [33] as described previously [24]. Plasmids encoding the six internal PR8 genes and the Udorn NA gene were provided with additional competing PB1 genes. One  $\mu$ g of each plasmid was mixed with FuGENE6 transfection reagent (Promega, Alexandria, NSW, Australia) in Opti-MEM and added to a co-culture of 239T and MDCK cells. Six hours post-transfection, media was replaced with Opti-MEM supplemented with 50  $\mu$ g/mL streptomycin and 50 IU/mL penicillin. Twenty-four hours later, TPCK-trypsin (1  $\mu$ g/mL) was added and the supernatant harvested after a further 42 h and stored at  $-80$  °C. To determine the incorporation frequencies of the PB1 gene in progeny viruses, the transfection supernatant was plaque in MDCK cells. Randomly chosen plaques were picked by sampling through the agarose and resuspended in 0.05% Triton-X (Sigma Aldrich). The source of the competing gene segments was identified by gene-specific reverse transcription PCR (RT-PCR). The final data are derived from at least three independent experiments.

### 2.7. Gene-Specific RT-PCR

SensiFast probe no-ROX one-step RT-PCR Kit (Bioline, Alexandria, NSW, Australia) was used for gene identification. Each 20  $\mu$ L reaction was performed using 5  $\mu$ L of plaque-picked virus suspension, 10  $\mu$ L of 2 $\times$  SensiFast No-ROX one-step mix, 0.2  $\mu$ L of reverse transcriptase, 0.4  $\mu$ L of Ribosafe RNase inhibitor, 0.8  $\mu$ L of each 10  $\mu$ M gene-specific forward and reverse primer and 0.08  $\mu$ L of each 25  $\mu$ M gene-specific probe. The reverse transcription reaction was incubated for 10 min at 45 °C. Amplification and detection was performed using a Bio-Rad CFX96 RT-PCR system (Gladesville, NSW, Australia). Reaction conditions, primers (Geneworks, Thebarton, SA, Australia) and TaqMan (Integrated DNA Technologies, Baulkham Hills, NSW, Australia) probe sequences are available on request.

### 2.8. In Vitro Transcription and Electrophoretic Mobility Assay

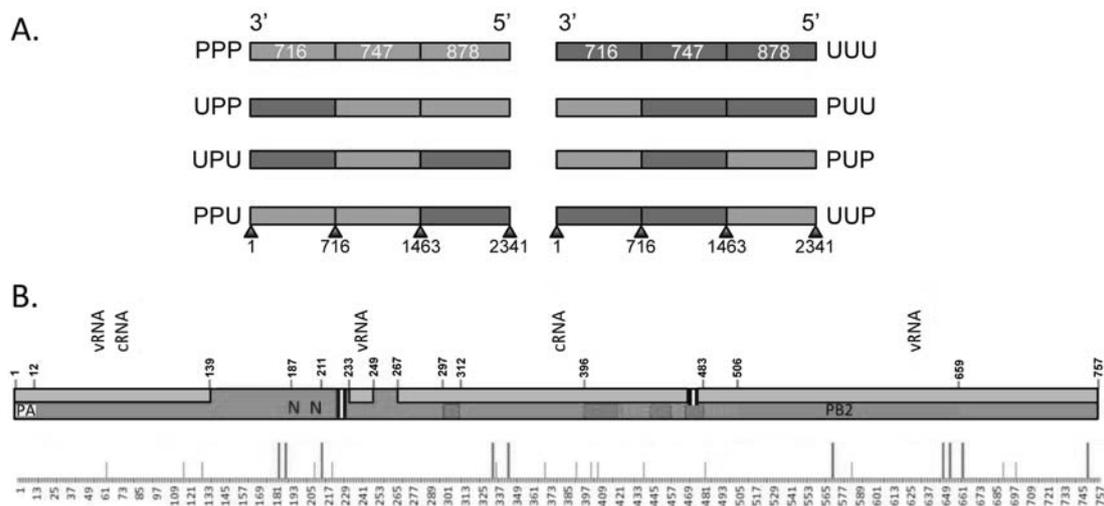
DNA corresponding to the chimeric PB1 or Udorn NA genes was inserted between the BsmBI (ThermoFisher Scientific, Illkirch, France) sites of a pUC2000 vector after excision from the corresponding pHW2000 plasmid. These vectors were obtained as previously described and contain a T7 promoter, the cloning cassette of pHW2000 containing two BsmBI sites and unique Bsh1236I or Ecl36II restriction sites between the PstI and EcoRI sites [9]. In vitro transcription of PB1 and NA vRNAs was performed with 30  $\mu$ g of Bsh1236I (ThermoFisher Scientific) or Ecl36II (New England BioLabs, Evry, France) linearized plasmid, respectively, in a 300  $\mu$ L reaction volume containing 40 mM Tris-HCl pH 7.5, 50 mM NaCl, 15 mM MgCl<sub>2</sub>, 1.7 mM spermidine (Sigma, Saint-Quentin Fallavier France), 5 mM 1,4-dithiothreitol (DTT) (USB, Santa Clara, CA, USA), 4 mM of each nucleotide triphosphate (NTP) (Sigma), 400 U RNasin (Promega), 0.01% Triton X-100 (Sigma) and 3  $\mu$ L of T7 RNA polymerase (made in-house). After 3 h of incubation, samples were treated with 50 U of RNase-free DNase I ((Roche, Meylan, France)) for 2 h at 37 °C, vRNAs were extracted with phenol/chloroform (Carl Roth GmbH, Karlsruhe, Germany), ethanol precipitated and purified on a TSK G2000SW column (Tosoh Bioscience GmbH, Stuttgart, Germany). The integrity of the vRNAs was checked by denaturing polyacrylamide gel electrophoresis. Pairs of purified vRNAs (2 pmol each) were denatured for 2 min at 90 °C in 8  $\mu$ L of water and cooled on ice. Two microliters of 5-fold concentrated buffer (final concentration: 50 mM sodium cacodylate (Sigma), pH 7.5; 40 mM KCl and 0.1 mM MgCl<sub>2</sub>) were added, the samples were incubated for 30 min at 55 °C and then analysed on 0.8% agarose gels containing 0.01% (w/v) ethidium bromide (Carl Roth GmbH) after addition of 5  $\mu$ L of loading buffer [40% (v/v) glycerol (Sigma); 0.05% (w/v) xylene cyanol (USB, Santa Clara, CA, USA); 0.05% bromophenol blue (Sigma)]. Native gel electrophoresis was performed for 4 h at 150 V at 4 °C in a buffer containing 50 mM Tris, 44.5 mM borate and 0.1 mM MgCl<sub>2</sub>. Gels were analyzed using a Gel Doc (Bio-Rad) imager and the ImageLab (Bio-Rad) software. The RNA weight fraction (%) of each band in a lane was determined and the percentage of intermolecular complex was calculated by dividing the weight fraction of the corresponding band by the sum of the weight fractions of all bands in the lane.

## 3. Results

### 3.1. Viruses Expressing Chimeric PB1 Genes Show No Major Differences in Replication in Vitro

To investigate the interaction between the Udorn NA gene and the central coding region of the Udorn PB1 gene, we used overlapping PCR to divide the PB1 gene segment into approximate thirds and create plasmids encoding a series of chimeric PB1 genes containing sections swapped between Udorn PB1 and PR8 PB1. The six possible chimeric PB1s produced by this process are illustrated in Figure 1A and are denoted as UPP, UPU, UUP, PPU, PUP or PUU depending on the PR8 (P) or Udorn (U) origin of each third of the gene. In addition, plasmids encoding entirely PR8 (PPP) and Udorn (UUU) PB1 genes were included as controls. Figure 1B shows a representative schematic of the PB1 protein encoding 757 amino acids, highlighting various functional domains. The first section contains the PA binding domain [37] and two discrete nuclear localisation signals [38]. The middle section contains four regions of high conservation in all RNA-dependent RNA polymerases [39,40],

and the last section contains the PB2 binding domain [37]. Interactions of the PB1 protein with vRNA and complementary RNA (cRNA) span all three sections [41]. Efforts to minimise any potential disruption of the functional domains were made by placing the boundary sites in regions of homology between PR8 and Udorn PB1s. Table 1 summarises the location, length, nucleotide homology and amino acid differences between PR8 and Udorn in respect to the three sections of the chimeric PB1 genes. There are only 25 differences out of 757 amino acids across the entire PR8 and Udorn PB1 proteins despite a relatively low nucleotide homology of 84%.



**Figure 1.** Structure of chimeric PB1 plasmids and functional domains of the PB1 gene. **(A)** Schematic representation of chimeric PB1 plasmid constructs. The origin of each third of the gene is colour-coded, orange represents PR8 (P) and blue represents Udorn (U) sequences. The nucleotide length and boundaries of each section are indicated. Influenza vRNA are numbered from 3' to 5' according to international nomenclature; **(B)** Schematic representation of the functional domains of the PB1 protein highlighting the amino acid differences between Udorn and PR8 PB1. Polymerase acidic protein (PA) binding domain (yellow); Nuclear localization signals (N); conserved RNA-dependent RNA polymerase motifs (red); Polymerase basic protein 2 (PB2) binding domain (blue); interaction with viral RNA (vRNA) and complementary RNA (cRNA) (orange). Amino acid differences between PR8 and Udorn PB1 proteins are shown below the schematic with smaller grey lines indicating the location of conservative amino acid changes and bigger red lines indicating non-conservative amino acid changes.

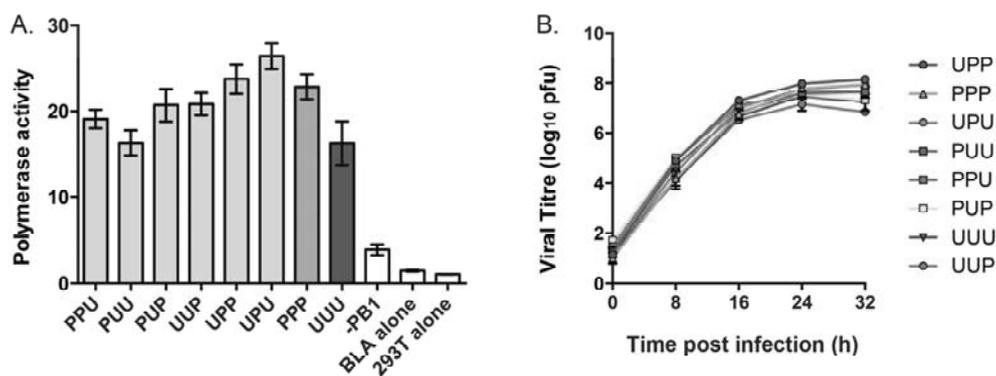
**Table 1.** Nucleotide homologies and amino acid differences between Udorn PB1 and PR8 PB1.

Region of PB1	Nucleotides	Length (nt)	Nucleotide Homology	Amino Acid Differences
Overall	1–2341	2341	83.7%	25
First third	1–716	716	85.2%	8
Middle third	717–1463	747	82.6%	9
Last third	1464–2341	878	83.7%	8

To establish whether the chimeric PB1s were functional, a BLA reporter assay was used in cells cotransfected with plasmids expressing parental or chimeric PB1 in addition to those expressing each of the RNP components (PA, PB2, NP) of PR8 virus. Measurement of the cleavage of substrate by the expressed BLA enzyme (Figure 2A), which correlates with polymerase activity, revealed that an entirely PR8 RNP has a higher polymerase activity than one containing the PB1 from Udorn virus as expected [32]. Notably, all of the chimeric PB1s formed RNP complexes that were functional and were not significantly different from either PR8 or Udorn PB1-containing RNPs.

Eight-plasmid reverse genetics was then used to construct viruses expressing chimeric PB1 genes on the backbone of PR8(Ud-NA) virus, which contains the NA gene from Udorn virus and all other

genes from PR8 virus. This backbone was chosen for all future competitive transfection experiments as it was postulated that the Udorn NA RNP would allow co-selection of specific chimeric PB1 genes that contained the putative PB1-NA nucleotide binding motif. All of the six PR8(Ud-NA) viruses expressing chimeric PB1 genes could be rescued. When the viral growth kinetics of these viruses were assessed in MDCK cells (Figure 2B), no statistically significant differences in the replicative efficiency was observed over the first 16 h between PR8(Ud-NA) virus containing a PR8 PB1 (PPP), its PR8(Ud-NA, PB1) counterpart containing a Udorn PB1 (UUU), or any of the PR8(Ud-NA) viruses expressing chimeric PB1 genes. This meant that the PR8(Ud-NA) backbone would allow future PB1 competitive transfection experiments to be investigated in the absence of differences in the intrinsic replicative capacity of the competing progeny. This formed a key part of the rationale for using the PR8(Ud-NA) backbone as we had previously shown with viruses additionally expressing the Udorn HA that PR8(Ud-HA,NA,PB1) displays vastly inferior replicative efficiency compared to PR8(Ud-HA,NA) [32], making the PR8(Ud-HA,NA) backbone unsuitable to determine outcomes due to gene co-segregation rather than replicative advantage. At the later time points of 24–32 h only one virus, which expressed UUP, showed statistically lower replication than the other viruses (lower compared to UPP, PPP and UPU at 24 h, and compared to all except UUU at 32 h) and it is possible that this chimeric PB1 may not be represented as well in the progeny of competitive transfections due to this slight decrease in replicative capacity.



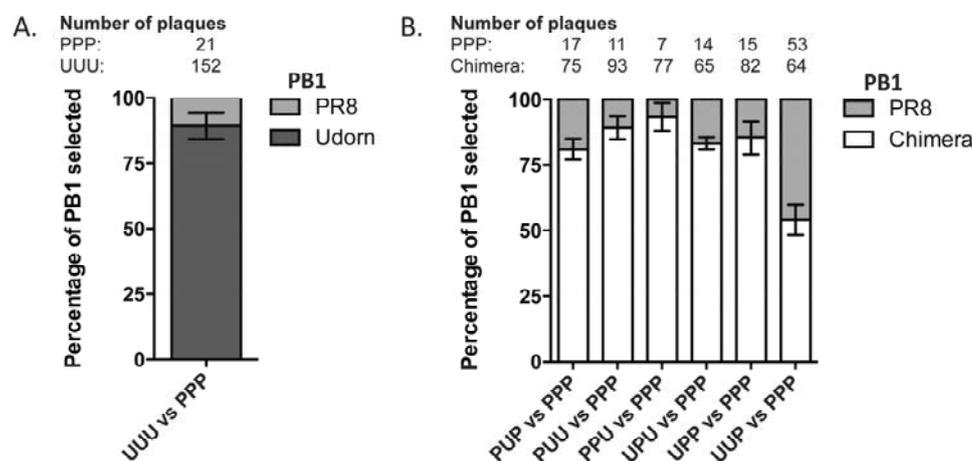
**Figure 2.** Effect of chimeric PB1 genes on polymerase activity and replication of PR8(Ud-NA, chimeric PB1) viruses. (A) Minigenome assays were performed in 293T cells transfected with the pCAGGS-BLA reporter gene, plasmids expressing the PR8 PB2, PA and NP genes, together with the indicated chimeric PB1 gene or wild-type PR8 (PPP; orange) or Udorn (UUU; blue) PB1 genes. The polymerase activity of each ribonucleoprotein (RNP) complex was measured after 90 min and compared to a transfection mixture without a PB1 gene (-PB1), cells transfected with the BLA reporter alone (BLA alone) and untransfected cells (293T alone). The data represents the mean and standard error of two independent experiments; (B) MDCK cells were infected with viruses containing chimeric PB1 genes at a multiplicity of infection (MOI) of 0.01. At the indicated time points, the amount of infectious virus released into the supernatant was determined by plaque assay. The data represent the mean and standard error of three individual experiments.

### 3.2. Nine Plasmid Competitive Transfections Reveal Chimeras Expressing Any Section of Udorn PB1 Are Preferentially Selected Over PR8 PB1 in Progeny Virus

To attempt to define the region of the PB1 driving co-selection of this gene with the Udorn NA we used a nine-plasmid transfection system in which plasmids encoding the genetic backbone of PR8(Ud-NA) were provided together with plasmids encoding different PB1 genes. The resulting viral progeny in the transfection supernatants were then analysed by RT-PCR from plaque-picks to determine the source of the competing PB1 gene. Initially we competed the wildtype PR8 and Udorn PB1 genes to re-affirm data we have published previously [24] showing that the Udorn PB1 gene is

incorporated at a significantly higher frequency than PR8 PB1 ( $p < 0.0001$ ) in the presence of the Udorn NA gene on this background (Figure 3A).

The chimeric PB1 genes were then competed with the PR8 PB1 gene in the same system. At least three independent experiments were performed for each competitive transfection and the frequencies of the PB1 genes selected in each case were averaged across the different experiments (Figure 3B). Chimeric PB1s with a Udorn section in any position were selected at a greater frequency than that of the PR8 PB1 (PUP;  $p = 0.0004$ , PUU;  $p = 0.0002$ , PPU;  $p = 0.0003$ , UPU;  $p < 0.0001$ , UPP;  $p = 0.0014$  (Student's *t*-test)). This was in agreement with selection of Udorn PB1 over PR8 PB1 in PR8(Ud-NA) backbone virus. The exception to this was the chimeric UUP PB1 which was selected at a similar frequency ( $54\% \pm 5\%$ ;  $p = 0.3577$  (Student's *t*-test)) to PPP PB1. This was not surprising as PR8(Ud-NA, UUP PB1) demonstrated a significantly inferior intrinsic replication efficiency compared to PR8(Ud-NA) expressing PPP PB1 in MDCK cells (Figure 2B). As a result, any selection due to the presence of a Udorn section in UUP PB1 could potentially be offset by the replicative advantage of PR8(Ud-NA) in the transfection culture. The chimeric PPU PB1 containing the last section of Udorn had the highest preferential incorporation frequency over PPP PB1 ( $93\% \pm 4\%$ ).

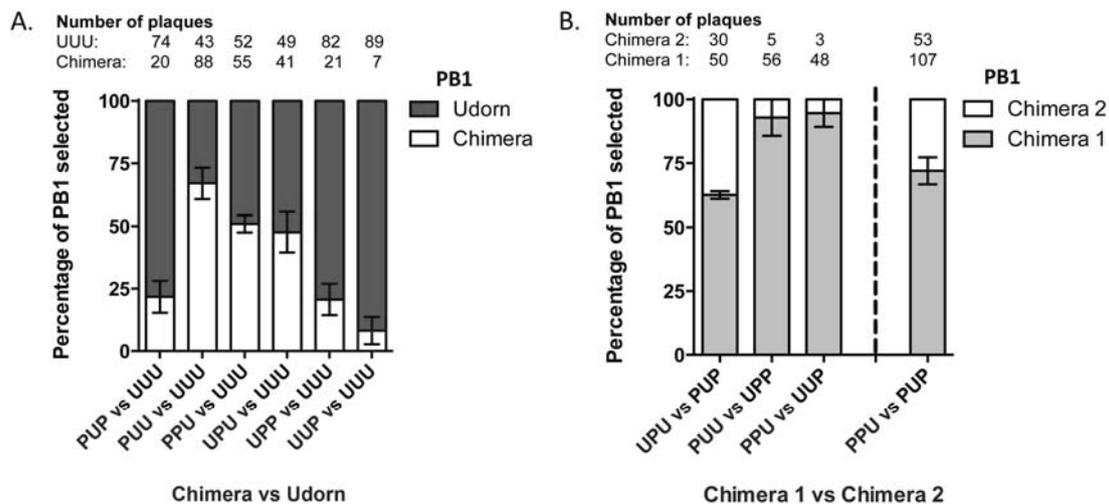


**Figure 3.** Frequency of competing genes incorporated in progeny viruses resulting from nine-plasmid competitive transfection assays. Nine-plasmid competitive transfections were performed on a PR8(Ud-NA) background competing (A) wild-type PR8 (orange) and Udorn (blue) PB1 genes or (B) the wild-type PR8 PB1 gene (PPP; orange) against chimeric PB1 genes (white) as indicated. The origin of the PB1 gene segment from a random selection of progeny plaques was determined by gene-specific RT-PCR. Total numbers of specific plaques determined from each competition are indicated. The data represent the mean and standard error of 3–5 individual experiments.

### 3.3. The Last Third of PB1 RNA Is Driving Preferential Co-Selection of Udorn PB1 and Udorn NA

Chimeric PB1 genes were then competed against the Udorn PB1 gene in the nine-plasmid transfection system. Again, three or more replicate experiments were performed for each competition and the frequency of each chimera averaged (Figure 4A). Unlike the competitions against PR8 PB1, chimeras displayed hierarchical selection against Udorn PB1. Three of the chimeras displayed comparable or even greater preferential selection to UUU PB1. Of these, the chimera selected at the highest frequency was PUU ( $67\% \pm 6\%$ ;  $p = 0.0084$ ), containing the middle and last sections of Udorn, followed by PPU ( $51\% \pm 3\%$ ;  $p = 0.7137$ ) and UPU ( $48\% \pm 6\%$ ;  $p = 0.6964$  (Student's *t*-test)). PUP, UPP and UUP chimeras were all selected at a significantly lower frequency to the wild-type Udorn PB1 ( $<25\%$ ; PUP;  $p = 0.0033$ , UPP  $p = 0.0027$ , UUP;  $p = 0.0004$  (Student's *t*-test)). The chimeric genes that showed equivalent selection to UUU had the last third from Udorn PB1 in common, suggesting this region may be driving the co-selection.

To confirm the importance of the last third of Udorn PB1 in cosegregation with Udorn NA, the different chimeras were then competed head-to-head in nine-plasmid transfections. The initial approach was to compete sets of chimeras that contained ‘complementary’ regions of the PB1. For example, competing UPU against PUP, PUU against UPP and PPU against UUP. In these experiments the PB1 having Udorn first and last sections (UPU) was preferentially selected over the middle (PUP;  $p = 0.0066$ ); Udorn middle and last sections (PUU) was preferentially selected over the first (UPP;  $p = 0.0136$ ); and the Udorn last section (PPU) was preferentially selected over the first and middle (UUP;  $p = 0.0071$  (Student’s *t*-test)) (Figure 4B). Once again the chimeras that were preferentially selected in these competitions all contained the last Udorn PB1 section, although the UPU chimera was not selected as strongly when competed against PUP ( $63\% \pm 1\%$ ) as was PUU ( $93\% \pm 5\%$ ) and PPU ( $95\% \pm 3\%$ ) when competed against their complementary PB1s. This indicates a possible involvement of the middle section or conversely, a negative effect on selection of the first section when paired with the last section (UPU). To confirm whether the middle or last section was more important in driving co-selection with the Udorn NA it was decided to compete the PUP and PPU chimeras head-to-head. Five independent transfections competing these specific chimeras were performed and in all competitions the PPU chimera, containing the last section of the Udorn PB1, was preferentially selected ( $p = 0.0011$  (Student’s *t*-test)) (Figure 4B), indicating the region between nucleotides 1464–2341 of PB1, corresponding to the 5′ end of the vRNA, is the likely driver of co-segregation with Udorn NA and may contain a region by which the two RNPs interact.

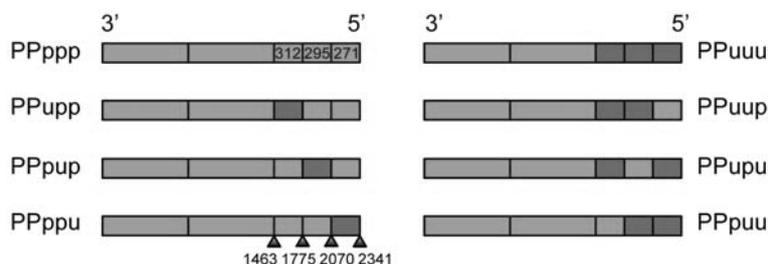


**Figure 4.** Frequency of competing genes incorporated in progeny viruses resulting from nine-plasmid competitive transfection assays. Nine-plasmid competitive transfections were performed on a PR8(Ud-NA) background competing (A) the wild-type Udorn PB1 gene (UUU; blue) against chimeric PB1 genes (white) as indicated or (B) two chimeric PB1 genes as indicated. The origin of the PB1 gene segment from a random selection of progeny plaques was determined by gene-specific RT-PCR. Total numbers of specific plaques determined from each competition are indicated. The data represent the mean and standard error of 2–5 individual experiments.

### 3.4. Viruses Expressing 5′ Chimeric PB1 Genes Show No Major Differences in Replication in Vitro

As the last section of the PB1 was crucial for the PB1-NA gene co-segregation we wished to further dissect this region by creating chimeras of the last 5′ section. Once again the strategy was to divide this section into approximate thirds using overlapping PCR and create a panel of 5′ chimeric PB1 plasmids that we could compete in the nine-plasmid transfection system. A schematic of the 5′ chimeric PB1 plasmids is shown in Figure 5. These 5′ chimeras contained the first and middle thirds of PR8 (P), with the last third divided into three sections and swapped interchangeably between PR8 (p) and Udorn (u), designated PPpup, PPpup, PPppu etc. Uppercase nomenclature has been used to indicate

larger thirds of the PB1; lowercase nomenclature used to indicate the smaller three sections of the last third of the PB1.



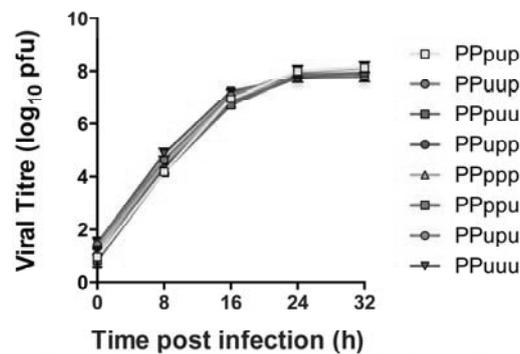
**Figure 5.** Structure of 5' chimeric PB1 plasmids. Schematic representation of 5' chimeric PB1 plasmid constructs showing the origin of each section of the gene; orange represents PR8 (P) and blue represents Udorn (U) sequences. In the nomenclature for the chimeras, uppercase indicates the origin of the larger thirds of the PB1 (all PR8); lowercase indicates the origin of the smaller three sections of the last third of the PB1. The nucleotide length and boundaries of each 5' section is indicated.

Table 2 summarises the location, length, nucleotide homology and amino acid differences between PR8 and Udorn in respect to the three sections within the 5' third of the chimeric PB1 genes. Eight of the 25 amino acid differences across the entire PR8 and Udorn PB1 proteins occur in the last third of the gene. At the nucleotide level the last third of the PB1 gene showed 83.7% homology between Udorn and PR8. Within this third, the middle section displayed the lowest homology between PR8 and Udorn viruses.

**Table 2.** Nucleotide homologies and amino acid differences between Udorn PB1 and PR8 PB1 in the 5' third.

Region within the Last (5') Third of PB1	Nucleotides	Length (nt)	Nucleotide Homology	Amino Acid Differences
Last 5' third overall	1464–2341	878	83.7%	8
First third of 5'	1464–1775	312	83.9%	2
Middle third of 5'	1776–2070	295	79.5%	4
Last third of 5'	2071–2341	271	86.6%	2

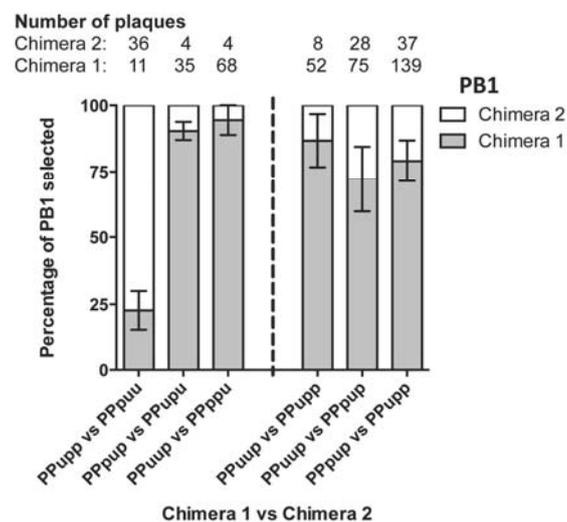
Eight-plasmid reverse genetics was again used to engineer these 5' chimeric PB1 genes onto the backbone of PR8(Ud-NA). All of the PR8(Ud-NA) viruses expressing 5' chimeric PB1 genes were successfully rescued. The viral growth kinetics of these viruses were again assessed in MDCK cells. Viral loads determined from samples of culture supernatants taken at 8 h intervals showed there were no significant differences in the viral replication of the respective 5' chimeric viruses ( $p = 0.1919$  (two-way analysis of variance (ANOVA) repeated measures)) (Figure 6). As a result selection of chimeras containing specific Udorn sections in subsequent competitive transfections could be attributed solely to co-selection of the Udorn region with the Udorn NA without any confounding effects of intrinsic replicative advantage.



**Figure 6.** Growth kinetics of reverse genetics-derived PR8(Ud-NA, 5' chimeric PB1) viruses. Madin-Darby canine kidney (MDCK) cells were infected with viruses containing 5' chimeric PB1 genes at an MOI of 0.01. At the indicated time points, the amount of infectious virus released into the supernatant was determined by plaque assay. The data represent the mean and range of two individual experiments.

### 3.5. The PB1 RNA Segment Containing Nucleotides 1776–2070 Is Driving Preferential Coselection of Udorn PB1 and Udorn NA

Sets of complementary 5' chimeric PB1 genes on a PR8(Ud-NA) backbone were then competed head-to-head in nine-plasmid transfections. Initially we competed PPupp against PPpuu, PPpup against PPupu and PPuup against PPppu. In these experiments, the Udorn middle and last 5' sections (PPpuu) were preferentially selected over the first 5' section (PPupp;  $p = 0.0339$ ); the Udorn middle 5' section (PPpup) was preferentially selected over the first and last 5' sections (PPupu;  $p = 0.0035$ ); and the Udorn first and middle 5' sections (PPuup) were preferentially selected over the last 5' section (PPppu;  $p = 0.0077$  (Student's  $t$ -test)) (Figure 7 left panel).



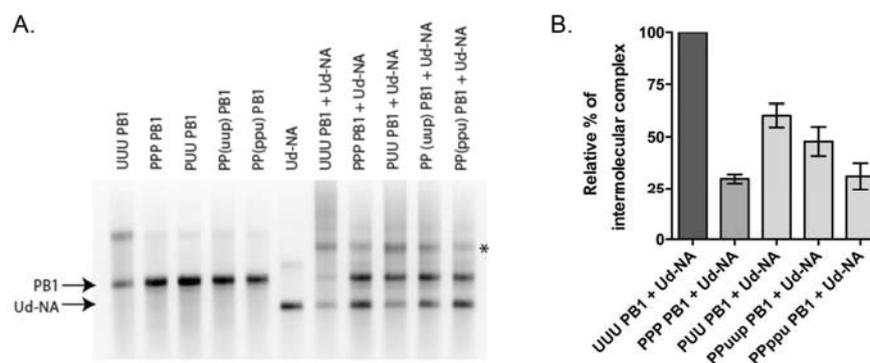
**Figure 7.** Frequency of competing chimeric genes incorporated in progeny viruses resulting from nine-plasmid competitive transfection assays. Nine-plasmid competitive transfections were performed on a PR8(Ud-NA) background competing two 5' chimeric PB1 genes as indicated. The origin of the PB1 gene segment from a random selection of progeny plaques was determined by gene-specific RT-PCR. Total numbers of specific plaques determined from each competition are indicated. The data represent the mean and standard errors of 2–5 individual experiments.

Chimeras that were preferentially selected in these competitions all contained the middle section of the last third of the Udorn PB1 gene. To confirm this observation and to investigate whether the

first 5' section had any co-involvement in selection, a set of additional competitions were performed (Figure 7, right panel), firstly keeping the first 5' section constant and testing the requirement for the middle section (PPuup versus PPpup). In these competitions the chimera that also contained the middle 5' Udorn section was preferentially selected ( $87\% \pm 6\%$ ,  $p = 0.0352$ ) confirming the importance of this component of the gene. Competitions to investigate any co-involvement of the first 5' Udorn region (PPuup versus PPpup) with the middle 5' section, showed that inclusion of the first 5' section was also favoured for selection ( $p = 0.0442$ ). Lastly, to definitively show which region was the most important, the PPpup and PPpup 5' chimeras were competed head-to-head. Five independent transfections competing just the first and middle 5' Udorn regions confirmed that the middle section was indeed paramount for preferential selection ( $p = 0.0008$  (Student's *t*-test)). In summary, the Udorn region spanning nucleotides 1776–2070 of the PB1 appears to be crucial in the co-selection of the Udorn PB1 and NA genes.

### 3.6. RNA-RNA Binding Studies Recapitulate Strong and Weak Interactions Predicted by Nine-Plasmid Competitive Transfections

In order to assess whether the co-selection of the Udorn PB1 and NA genes was driven by vRNA-vRNA interactions, we used an *in vitro* assay that previously allowed the establishment of interaction networks between the vRNAs of the human A/Moscow/10/99 (H3N2) [17] and the avian A/Finch/England/2051/91 (H5N2) [18] influenza strains. Individually transcribed and purified vRNAs were co-incubated by pairs prior to native agarose gel electrophoresis (Figure 8A). The percentage of complex formed between the Udorn PB1 and NA genes was our reference and was set to 100%. As expected, PR8 PB1 vRNA (PPP) interacts weakly with the Udorn NA vRNA as compared to this reference, with only  $30\% \pm 2\%$  of complex formed compared to the homologous PB1-NA Udorn interaction (Figure 8B). The PUU PB1 chimera, which was selected at the highest frequency ( $67\% \pm 6\%$ ) over UUU PB1 in the nine-plasmid transfection system (Figure 4A), was also the one that rescued the *in vitro* vRNA-vRNA interaction the best ( $60\% \pm 5\%$ ) compared to the reference. Since the last third region of PB1 displayed the strongest capacity to select the NA Udorn vRNA, we compared the ability of the PPuup and PPppu PB1 chimeras to interact, *in vitro*, with the Udorn NA vRNA. In agreement with results obtained from the nine-plasmid competitive transfections (Figure 7), the PPuup chimera, containing the middle section of the last 5' region of Udorn PB1, rescued the interaction more efficiently ( $48\% \pm 6\%$ ) than the PPppu PB1 chimera ( $31\% \pm 6\%$ ), further supporting the idea that nucleotides 1776–2070 play an important role in the co-selection of the Udorn PB1 and NA genes through vRNA-vRNA interactions.



**Figure 8.** *In vitro* analysis of the interaction between Udorn NA vRNA and wild-type or chimeric PB1 vRNAs. (A) Analysis by native gel electrophoresis. Monomeric NA or PB1 vRNAs are indicated by arrows and the intermolecular complexes are marked by an asterisk; (B) Relative percentage of intermolecular complex. The weight fraction (%) of the RNA migrating as an intermolecular complex was calculated by dividing the weight fraction of the corresponding band by the sum of the weight fractions of all bands in the lane and normalized to the complex formed between wild-type (UUU) PB1 and Udorn NA vRNAs set to 100%. Data represent mean  $\pm$  standard error ( $n = 3-4$ ).

#### 4. Discussion

This study was performed to locate the key region of the PB1 gene that was responsible for driving co-selection of the Udorn PB1 with the Udorn NA genes in reassortant progeny. The use of nine-plasmid competitive transfection experiments to determine preferential packaging of chimeric and wild-type PB1 RNPs was successful in narrowing down the region of potential interaction with Udorn NA RNP to approximately 1/8 the length of the complete PB1 gene. Our ability to use chimeric genes to compete with wild-type genes was facilitated by the highly conserved nature of the encoded PB1 protein that resulted in no major loss of function of the chimeric proteins nor of the replicative capacity of the viruses expressing chimeric PB1 genes. As such this approach may be applicable to other internal, highly conserved proteins of the virus. Moreover, as the genotype of progeny virions is used as the readout, the findings can be related to RNP-RNP interactions dictating the final gene constellation rather than to other vRNA-vRNA interactions that may occur during the course of replication within the infected cell.

The domain on Udorn PB1 vRNA that we have identified for potential interaction with Udorn NA vRNA, spanning nucleotides 1776–2070 of the PB1 gene, resided in the middle section of the last (5′) third of the gene, outside the identified packaging sequences [42]. Although the first 5′ section of the last third of the gene was not crucial for co-selection, it was favoured in competitions where the middle 5′ section was kept constant (PPuup versus PPpup). This implies that this first 5′ section may contribute to the co-selection of the interacting region, potentially by providing relevant nucleotide sequence across the boundary with the middle 5′ section. The boundaries were designed in nucleotide regions common to PR8 and Udorn PB1 vRNA and this particular junction had 17 nucleotides, from 1759 to 1775, in common. Hence if a non-critical part of the interacting domain resides in the first 5′ section it would conceivably be upstream of nucleotide 1759. Alternatively the first 5′ Udorn section may promote the formation of optimal secondary structures within the middle 5′ Udorn section for interaction with Udorn NA RNP. A third possibility is that this first 5′ section contains another subdominant interacting region with the Udorn NA. If there were several co-segregation signals within the gene, the strength of these signals could be hierarchical such that other interacting signals could potentially compensate for the loss of one in chimeras. This could also potentially explain why nearly all chimeric genes were preferentially selected against the wild-type PR8 gene, despite not always containing the last third of the Udorn PB1 sequence. The concept of multiple interacting regions within genes can be visualized by reference to Noda et al. [8] in supplementary movie 3 showing the 3D model of eight RNPs within a virion. Similarly, preferential selection of the Udorn PB1 gene over that of PR8 was also never 100% in competitive assays. This might indicate that regions of the PR8 PB1 vRNA can also form weak interactions with the Udorn NA or alternatively that the PR8 PB1 gene can be incorporated at a low frequency due to interactions with other genes in the supramolecular assembly complex.

We have previously shown that the eight vRNAs of both human H3N2 (A/Moscow/10/99) and avian H5N2 (A/Finch/England/2051/91) influenza A viruses form a single network of interactions *in vitro* [17,18]. Eight H3N2 and thirteen H5N2 interactions have been reported with only four of these in common between the two viruses. In addition to largely different networks, the sequences that were involved in the specific vRNA-vRNA interactions also differed: while vRNA-vRNA interaction sequences were mainly mapped to terminal regions in the case of the H3N2 strain [17], they essentially resided centrally for the avian H5N2 virus [18]. Similarly, RNP-RNP interactions were also visualized all along the rod-shaped structure of the RNPs in the case of the WSN viral strain [8]. Our previous report that Udorn PB1-NA co-selection is driven by interactions involving the internal coding region of the PB1 [24] supports this, and is consistent with our further refinement of the interacting region to nucleotides 1776–2070 of the PB1, which lies within this central coding region. Whether a similar co-selection of PR8 PB1 with PR8 NA exists to that of the Udorn segments is difficult to ascertain from competitive transfection experiments as the PR8 virus (containing the PR8 PB1) has a significant replicative advantage over its PR8(Ud-PB1) counterpart [24].

Experiments quantifying the ability of a subset of purified Udorn-PR8 chimeric PB1 vRNAs to form intermolecular complexes with Udorn NA vRNA were consistent with the importance of the first and middle sections of the 5' region of the PB1 gene. This suggests that, in some instances at least, the RNA-RNA interactions measured *in vitro* may well predict RNP-RNP interactions within the supramolecular complex formed prior to or during budding. PB1-NA vRNA interactions were previously shown for both the A/Moscow/10/99 (H3N2) [9] and A/Finch/England/2051/91 (H5N2) [18] influenza strains, but these were relatively weak, with  $15\% \pm 2\%$  and  $9\% \pm 1\%$  of the purified RNAs forming intermolecular complexes respectively. This has made it difficult in the past to establish the importance of this interaction. In the present study, however,  $52\% \pm 3\%$  of the purified Udorn PB1 and NA RNA formed intermolecular complexes ( $n = 5$ , range 47%–65%), indicative of a strong association.

Within the 1776–2070 nucleotide domain of the PB1 we identified a number of regions where the PB1 vRNA sequence was highly complementary to that of the NA vRNA sequence, potentially allowing PB1-NA interactions. These regions were between seven and 22 nucleotides in length with complementarity between the two genes ranging from 85.7% up to 100%. At this stage, we cannot conclude whether these regions do indeed interact in RNPs, and there are too many possibilities to test them systematically using site directed mutagenesis and competition of mutated and wild-type genes, or by *in vitro* binding assays. Additional criteria will have to be used in order to prioritize these possible interactions, and will include those regions showing conservation of pairing within pandemic H2N2 and H3N2 isolates.

The presence of highly conserved regions of pairing in the N2 NA and the PB1 genes would explain the gene constellations arising in past pandemic viruses during their creation by reassortment. In the precursors of the 2009 H1N1 pandemic virus, the NA and PB1 genes initially co-segregated with the HA gene from the human H3N2 donor in the formation of the swine H3N2 strain. This occurred again in the formation of the swine H3N2 'triple reassortant' strain, and eventually only the human H3N2 PB1 and NA genes co-segregated to form the swine H1N2 'triple reassortant' strain [43]. Similarly, the PB1, NA and HA genes from the avian H2N2 donor co-segregated, when reassorting with the human H1N1 virus, to produce the 1957 H2N2 'Asian flu' pandemic strain [44,45]. The PB1 of the H1N1 virus circulating in 1957 may, like that of PR8, be devoid of the correct sequence and structural characteristics to co-segregate with the N2 NA, thus favouring the incorporation of the PB1 from the avian donor. In creation of the 1968 H3N2 pandemic virus, although only the PB1 and HA genes were introduced from the avian virus donor into the existing human strain, this is not incompatible with the notion of PB1-NA co-segregation as the incoming PB1 gene from the avian donor may, as in the H2N2 pandemic virus, have had the capacity to pair with the human N2 NA.

Understanding the interactions that are involved in gene co-segregation is essential to fully appreciate the factors that govern the gene constellations of viral progeny likely to result from reassortment of viruses from human and non-human species. Those favouring the incorporation of particular polymerase genes are especially important as these genes can control the speed of replication as well as contain virulence factors that may contribute to pathogenicity. For example, some strains encode a full-length inflammatory PB1-F2 protein in the second reading frame of the PB1 gene that is associated with enhanced disease severity [46,47]. This study elucidates a mechanism of viral gene co-selection in emerging influenza viruses with pandemic potential that may govern their virulence.

**Acknowledgments:** This work was supported by Program Grant 1071916 from the National Health and Medical Research Council of Australia and by the Centre National de la Recherche Scientifique (CNRS).

**Author Contributions:** B.G., L.B., S.R., C.I., M.G., C.O. and R.M. conceived and designed the experiments; B.G., C.I., M.G., A.P.-S. and T.Z. performed the experiments; all authors analyzed the data and contributed to the writing of the paper.

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

## References

1. Hsu, M.T.; Parvin, J.D.; Gupta, S.; Krystal, M.; Palese, P. Genomic RNAs of influenza viruses are held in a circular conformation in virions and in infected cells by a terminal panhandle. *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 8140–8144. [CrossRef] [PubMed]
2. Compans, R.W.; Content, J.; Duesberg, P.H. Structure of the ribonucleoprotein of influenza virus. *J. Virol.* **1972**, *10*, 795–800. [PubMed]
3. Pons, M.W.; Schulze, I.T.; Hirst, G.K.; Hauser, R. Isolation and characterization of the ribonucleoprotein of influenza virus. *Virology* **1969**, *39*, 250–259. [CrossRef]
4. Pflug, A.; Guilligay, D.; Reich, S.; Cusack, S. Structure of influenza A polymerase bound to the viral RNA promoter. *Nature* **2014**, *516*, 355–360. [CrossRef] [PubMed]
5. Eisfeld, A.J.; Neumann, G.; Kawaoka, Y. At the centre: Influenza A virus ribonucleoproteins. *Nat. Rev. Microbiol.* **2015**, *13*, 28–41. [CrossRef] [PubMed]
6. Harris, A.; Cardone, G.; Winkler, D.C.; Heymann, J.B.; Brecher, M.; White, J.M.; Steven, A.C. Influenza virus pleiomorphy characterized by cryoelectron tomography. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 19123–19127. [CrossRef] [PubMed]
7. Noda, T.; Sagara, H.; Yen, A.; Takada, A.; Kida, H.; Cheng, R.H.; Kawaoka, Y. Architecture of ribonucleoprotein complexes in influenza A virus particles. *Nature* **2006**, *439*, 490–492. [CrossRef] [PubMed]
8. Noda, T.; Sugita, Y.; Aoyama, K.; Hirase, A.; Kawakami, E.; Miyazawa, A.; Sagara, H.; Kawaoka, Y. Three-dimensional analysis of ribonucleoprotein complexes in influenza A virus. *Nat. Commun.* **2012**, *3*, 639. [CrossRef] [PubMed]
9. Fournier, E.; Moules, V.; Essere, B.; Paillart, J.C.; Sirbat, J.D.; Isel, C.; Cavalier, A.; Rolland, J.P.; Thomas, D.; Lina, B.; Marquet, R. A supramolecular assembly formed by influenza A virus genomic RNA segments. *Nucleic Acids Res.* **2012**, *40*, 2197–2209. [CrossRef] [PubMed]
10. Chou, Y.Y.; Vafabakhsh, R.; Doganay, S.; Gao, Q.; Ha, T.; Palese, P. One influenza virus particle packages eight unique viral RNAs as shown by FISH analysis. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 9101–9106. [CrossRef] [PubMed]
11. Inagaki, A.; Goto, H.; Kakugawa, S.; Ozawa, M.; Kawaoka, Y. Competitive incorporation of homologous gene segments of influenza A virus into virions. *J. Virol.* **2012**, *86*, 10200–10202. [CrossRef] [PubMed]
12. Hutchinson, E.C.; von Kirchbach, J.C.; Gog, J.R.; Digard, P. Genome packaging in influenza A virus. *J. Gen. Virol.* **2010**, *91*, 313–328. [CrossRef] [PubMed]
13. Gerber, M.; Isel, C.; Moules, V.; Marquet, R. Selective packaging of the influenza A genome and consequences for genetic reassortment. *Trends Microbiol.* **2014**, *22*, 446–455. [CrossRef] [PubMed]
14. Marsh, G.A.; Hatami, R.; Palese, P. Specific residues of the influenza A virus hemagglutinin viral RNA are important for efficient packaging into budding virions. *J. Virol.* **2007**, *81*, 9727–9736. [CrossRef] [PubMed]
15. Marsh, G.A.; Rabadan, R.; Levine, A.J.; Palese, P. Highly conserved regions of influenza A virus polymerase gene segments are critical for efficient viral RNA packaging. *J. Virol.* **2008**, *82*, 2295–2304. [CrossRef] [PubMed]
16. Hutchinson, E.C.; Curran, M.D.; Read, E.K.; Gog, J.R.; Digard, P. Mutational analysis of *cis*-acting RNA signals in segment 7 of influenza A virus. *J. Virol.* **2008**, *82*, 11869–11879. [CrossRef] [PubMed]
17. Fournier, E.; Moules, V.; Essere, B.; Paillart, J.C.; Sirbat, J.D.; Cavalier, A.; Rolland, J.P.; Thomas, D.; Lina, B.; Isel, C.; et al. Interaction network linking the human H3N2 influenza A virus genomic RNA segments. *Vaccine* **2012**, *30*, 7359–7367. [CrossRef] [PubMed]
18. Gavazzi, C.; Isel, C.; Fournier, E.; Moules, V.; Cavalier, A.; Thomas, D.; Lina, B.; Marquet, R. An in vitro network of intermolecular interactions between viral RNA segments of an avian H5N2 influenza A virus: comparison with a human H3N2 virus. *Nucleic Acids Res.* **2013**, *41*, 1241–1254. [CrossRef] [PubMed]
19. Hatta, M.; Halfmann, P.; Wells, K.; Kawaoka, Y. Human influenza A viral genes responsible for the restriction of its replication in duck intestine. *Virology* **2002**, *295*, 250–255. [CrossRef] [PubMed]
20. Jackson, S.; Van Hoven, N.; Chen, L.M.; Maines, T.R.; Cox, N.J.; Katz, J.M.; Donis, R.O. Reassortment between avian H5N1 and human H3N2 influenza viruses in ferrets: A public health risk assessment. *J. Virol.* **2009**, *83*, 8131–8140. [CrossRef] [PubMed]

21. Li, C.; Hatta, M.; Watanabe, S.; Neumann, G.; Kawaoka, Y. Compatibility among polymerase subunit proteins is a restricting factor in reassortment between equine H7N7 and human H3N2 influenza viruses. *J. Virol.* **2008**, *82*, 11880–11888. [CrossRef] [PubMed]
22. Essere, B.; Yver, M.; Gavazzi, C.; Terrier, O.; Isel, C.; Fournier, E.; Giroux, F.; Textoris, J.; Julien, T.; Socratous, C.; et al. Critical role of segment-specific packaging signals in genetic reassortment of influenza A viruses. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, E3840–E3848. [CrossRef] [PubMed]
23. Octaviani, C.P.; Ozawa, M.; Yamada, S.; Goto, H.; Kawaoka, Y. High level of genetic compatibility between swine-origin H1N1 and highly pathogenic avian H5N1 influenza viruses. *J. Virol.* **2010**, *84*, 10918–10922. [CrossRef] [PubMed]
24. Cobbin, J.C.; Ong, C.; Verity, E.; Gilbertson, B.P.; Rockman, S.P.; Brown, L.E. Influenza virus PB1 and neuraminidase gene segments can cosegregate during vaccine reassortment driven by interactions in the PB1 coding region. *J. Virol.* **2014**, *88*, 8971–8980. [CrossRef] [PubMed]
25. Chen, L.M.; Davis, C.T.; Zhou, H.; Cox, N.J.; Donis, R.O. Genetic compatibility and virulence of reassortants derived from contemporary avian H5N1 and human H3N2 influenza A viruses. *PLoS Pathog.* **2008**, *4*, e1000072. [CrossRef] [PubMed]
26. Hara, K.; Nakazono, Y.; Kashiwagi, T.; Hamada, N.; Watanabe, H. Co-incorporation of the PB2 and PA polymerase subunits from human H3N2 influenza virus is a critical determinant of the replication of reassortant ribonucleoprotein complexes. *J. Gen. Virol.* **2013**, *94*, 2406–2416. [CrossRef] [PubMed]
27. Li, C.; Hatta, M.; Nidom, C.A.; Muramoto, Y.; Watanabe, S.; Neumann, G.; Kawaoka, Y. Reassortment between avian H5N1 and human H3N2 influenza viruses creates hybrid viruses with substantial virulence. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 4687–4692. [CrossRef] [PubMed]
28. Octaviani, C.P.; Goto, H.; Kawaoka, Y. Reassortment between seasonal H1N1 and pandemic (H1N1) 2009 influenza viruses is restricted by limited compatibility among polymerase subunits. *J. Virol.* **2011**, *85*, 8449–8452. [CrossRef] [PubMed]
29. Marshall, N.; Priyamvada, L.; Ende, Z.; Steel, J.; Lowen, A.C. Influenza virus reassortment occurs with high frequency in the absence of segment mismatch. *PLoS Pathog.* **2013**, *9*, e1003421. [CrossRef] [PubMed]
30. Bergeron, C.; Valette, M.; Lina, B.; Ottmann, M. Genetic content of Influenza H3N2 vaccine seeds. *PLoS Curr.* **2010**, *2*, RRN1165. [CrossRef]
31. Fulvini, A.A.; Ramanunnair, M.; Le, J.; Pokorny, B.A.; Arroyo, J.M.; Silverman, J.; Devis, R.; Bucher, D. Gene constellation of influenza A virus reassortants with high growth phenotype prepared as seed candidates for vaccine production. *PLoS ONE* **2011**, *6*, e20823. [CrossRef] [PubMed]
32. Cobbin, J.C.; Verity, E.E.; Gilbertson, B.P.; Rockman, S.P.; Brown, L.E. The source of the PB1 gene in influenza vaccine reassortants selectively alters the hemagglutinin content of the resulting seed virus. *J. Virol.* **2013**, *87*, 5577–5585. [CrossRef] [PubMed]
33. Hoffmann, E.; Neumann, G.; Kawaoka, Y.; Hobom, G.; Webster, R.G. A DNA transfection system for generation of influenza A virus from eight plasmids. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 6108–6113. [CrossRef] [PubMed]
34. Tannock, G.A.; Paul, J.A.; Barry, R.D. Relative immunogenicity of the cold-adapted influenza virus A/Ann Arbor/6/60 (A/AA/6/60-ca), recombinants of A/AA/6/60-ca, and parental strains with similar surface antigens. *Infect. Immun.* **1984**, *43*, 457–462. [PubMed]
35. Hoffmann, E.; Stech, J.; Guan, Y.; Webster, R.G.; Perez, D.R. Universal primer set for the full-length amplification of all influenza A viruses. *Arch. Virol.* **2001**, *146*, 2275–2289. [CrossRef] [PubMed]
36. Cavois, M.; De Noronha, C.; Greene, W.C. A sensitive and specific enzyme-based assay detecting HIV-1 virion fusion in primary T lymphocytes. *Nat. Biotechnol.* **2002**, *20*, 1151–1154. [CrossRef] [PubMed]
37. Naffakh, N.; Tomoiu, A.; Rameix-Welti, M.A.; van der Werf, S. Host restriction of avian influenza viruses at the level of the ribonucleoproteins. *Annu. Rev. Microbiol.* **2008**, *62*, 403–424. [CrossRef] [PubMed]
38. 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 (H1N1). *Mol. Cell. Biol.* **1990**, *10*, 4139–4145. [CrossRef] [PubMed]
39. Biswas, S.K.; Nayak, D.P. Mutational analysis of the conserved motifs of influenza A virus polymerase basic protein 1. *J. Virol.* **1994**, *68*, 1819–1826. [PubMed]
40. Poch, O.; Sauvaget, I.; Delarue, M.; Tordo, N. Identification of four conserved motifs among the RNA-dependent polymerase encoding elements. *EMBO J.* **1989**, *8*, 3867–3874. [PubMed]

41. Gonzalez, S.; Ortin, J. Distinct regions of influenza virus PB1 polymerase subunit recognize vRNA and cRNA templates. *EMBO J.* **1999**, *18*, 3767–3775. [CrossRef] [PubMed]
42. Fujii, K.; Fujii, Y.; Noda, T.; Muramoto, Y.; Watanabe, T.; Takada, A.; Goto, H.; Horimoto, T.; Kawaoka, Y. Importance of both the coding and the segment-specific noncoding regions of the influenza A virus NS segment for its efficient incorporation into virions. *J. Virol.* **2005**, *79*, 3766–3774. [CrossRef] [PubMed]
43. Kingsford, C.; Nagarajan, N.; Salzberg, S.L. 2009 Swine-origin influenza A (H1N1) resembles previous influenza isolates. *PLoS ONE* **2009**, *4*, e6402. [CrossRef] [PubMed]
44. Kawaoka, Y.; Krauss, S.; Webster, R.G. Avian-to-human transmission of the PB1 gene of influenza A viruses in the 1957 and 1968 pandemics. *J. Virol.* **1989**, *63*, 4603–4608. [PubMed]
45. Scholtissek, C.; Rohde, W.; Von Hoyningen, V.; Rott, R. On the origin of the human influenza virus subtypes H2N2 and H3N2. *Virology* **1978**, *87*, 13–20. [CrossRef]
46. Chen, W.; Calvo, P.A.; Malide, D.; Gibbs, J.; Schubert, U.; Bacik, I.; Basta, S.; O’Neill, R.; Schickli, J.; Palese, P.; et al. A novel influenza A virus mitochondrial protein that induces cell death. *Nat. Med.* **2001**, *7*, 1306–1312. [CrossRef] [PubMed]
47. McAuley, J.L.; Tate, M.D.; MacKenzie-Kludas, C.J.; Pinar, A.; Zeng, W.; Stutz, A.; Latz, E.; Brown, L.E.; Mansell, A. Activation of the NLRP3 inflammasome by IAV virulence protein PB1-F2 contributes to severe pathophysiology and disease. *PLoS Pathog.* **2013**, *9*, e1003392. [CrossRef] [PubMed]



© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).



## CONCLUSIONS & PERSPECTIVES

### A. Conclusions

Lors de mon doctorat, nous avons étudié la souche modèle de laboratoire d'IVA PR8 adaptée à la croissance sur œufs et donc largement utilisée pour la production des vaccins antigrippaux saisonniers. La compréhension des mécanismes d'empaquetage présente par conséquent un intérêt majeur pour cette souche. Néanmoins, peu d'informations sont disponibles sur les signaux d'empaquetage des ARNv de ce virus, qui ont uniquement été délimités de manière grossière pour les segments PB1, PA, NP et M [16, 107, 108, 260, 262].

Au laboratoire, nous disposons d'un système qui permet d'identifier, *in vitro*, des interactions entre ARNv, dont le rôle fonctionnel doit ensuite être testé en culture cellulaire. Lors de ce travail de thèse, deux approches ont été utilisées :

i) Une approche « classique », longue et fastidieuse, mais dont la pertinence a été démontrée dans le cas de l'étude entre les ARNv PB1 et NS de la souche aviaire England [273]. À l'aide de cette approche, nous avons délimité les séquences, au nt près, responsables de l'hétérodimérisation entre les ARNv PB1 et M. Nous avons également étudié d'autres partenariats, mais les régions précises impliquées dans ces interactions n'ont pas pu être identifiées (voir Résultats et discussion, § A.II.1).

ii) Une approche innovante, récemment développée au laboratoire [277] que nous avons adaptée à l'étude des IVA. Pour la première fois, le MIME a été utilisé dans le but d'étudier des interactions ARN/ARN. Nous l'avons appliqué à l'ARNv NP afin de déterminer de manière exhaustive, efficace et non-biaisée, tous les nt impliqués dans les interactions intermoléculaires. Plusieurs séquences d'interaction potentielles ont été identifiées par cette approche et restent encore à valider *in vitro* et *in viro* (voir Résultats et discussion, § A.II.2).

Dans ces deux approches *in vitro*, la sélection fonctionnelle des ARNv est basée sur le retard de migration sur gel d'agarose natif. Par conséquent, les séquences identifiées devraient se superposer. Il semblerait que les deux méthodes appliquées à l'interaction entre les ARNv PA et NP aient effectivement mis en évidence l'implication des mêmes nt dans la formation du complexe, mais ces résultats restent toutefois à confirmer. Il serait

intéressant d'appliquer le MIME à l'interaction entre les ARNv PB1 et M, dont les séquences d'interaction ont été identifiées et validées, *in vitro*, par l'approche « classique ».

Suite à l'identification des séquences d'interaction, au nt près, par ces approches *in vitro*, il est indispensable de valider leur rôle dans l'empaquetage du génome viral dans un contexte infectieux [273]. Pour ce faire, nous avons comparé les capacités répliquatives de virus qui possèdent un ou deux ARNv porteur(s) de mutations *trans*-compensatoires. Dans la littérature, la réplication virale n'est cependant jamais totalement abolie lors de l'introduction de mutations dans les séquences d'empaquetage et la diminution du titre viral est généralement faible (un à trois log) [16, 108, 254, 258, 260, 272, 273, 325, 339]. Cet effet partiel peut être dû à la nature discontinue et multipartite des signaux d'empaquetage présents sur un ARNv donné [107, 258, 260–262, 338], entraînant la compensation des séquences mutées par d'autres séquences d'empaquetage présentes sur l'ARNv. Il est donc probable que les signaux d'interaction ARN/ARN soient eux aussi multipartites. De plus, la flexibilité dans l'organisation des huit RNPv à l'intérieur des particules virales [308, 320] pourrait également augmenter la probabilité d'empaqueter un jeu complet de huit ARNv. En parallèle, la génétique inverse compétitive est un outil puissant pour évaluer le rôle d'une interaction dans le co-empaquetage des ARNv partenaires [273]. Dans ce cas, la détermination de la fréquence d'incorporation des ARNv sauvage et muté en présence de l'ARNv partenaire sauvage ou muté peut mettre en évidence un biais, lié à l'introduction des mutations, dans l'empaquetage du génome.

Ces deux approches *in vitro* ont permis de valider l'existence et le rôle de l'interaction entre les ARNv PB1 et NS de la souche aviaire England [273]. Nous avons également appliqué cette méthodologie à l'étude des interactions entre les ARNv PB1 et M et entre les ARNv PB2 et NP de la souche PR8, dont les séquences d'interaction ont été identifiées durant ma thèse, et entre les ARNv NA et M de la souche Moscou. Aucune de ces trois interactions n'a totalement été validée dans un contexte infectieux. Néanmoins, la perte de 1,4 log de capacités répliquatives lorsque les ARNv NA et M de la souche Moscou sont mutés suggère tout de même que l'interaction entre les ARNv existe au sein des particules virales, même si les mutations *trans*-compensatoires ne restaurent pas la réplication virale. À l'inverse, il semblerait que les interactions entre les ARNv PB1 et M et entre les ARNv PB2 et NP n'existent pas ou qu'elles n'ont aucun effet sur l'empaquetage des ARNv partenaires. Nos

deux approches *in vitro* sont réalisées en absence de protéines virales et ne reflètent donc probablement pas la structure adoptée par les ARNv au sein des cellules infectées et des virions. Ainsi, certaines interactions identifiées *in vitro* n'existeraient pas *in viro* et à l'inverse, toutes les interactions qui existent dans les virions ne seraient pas mises en évidence.

Durant ma dernière année de thèse, nous avons également collaboré avec l'équipe du Dr L. Brown qui avait mis en évidence une co-ségrégation des ARNv PB1 et NA de la souche humaine Udorn. L'utilisation de notre approche *in vitro* a permis dans ce cas d'étayer l'idée que cette co-ségrégation est vraisemblablement due à une interaction ARN/ARN entre les deux partenaires. Les études *in viro* et *in vitro* menées en parallèle ont conduit à l'identification de la région de l'ARNv PB1 impliquée dans l'interaction avec son partenaire [341]. Ces études se poursuivent actuellement au laboratoire pour une délimitation plus précise des séquences d'interaction sur les deux ARNv partenaires.

## **B. Perspectives**

Des études de tomographie électronique sont d'ores et déjà en cours (en collaboration avec le laboratoire du Dr P. Bron, Montpellier) afin d'obtenir une reconstruction tridimensionnelle des RNPv à l'intérieur de particules virales bourgeonnantes de la souche PR8. En combinant ces résultats à nos données *in vitro*, nous serons capables de proposer un modèle de l'organisation des huit RNPv à l'intérieur des virions. Certains partenariats de la souche PR8 devront néanmoins être validés dans un contexte infectieux afin de légitimer le réseau d'interactions. De plus, la tomographie devrait nous permettre de visualiser d'éventuels contacts tout le long des RNPv, comme dans le cas des virus WSN [320] et England [319], ou à l'inverse, de constater que les contacts se restreignent aux extrémités des segments, comme dans le cas du virus Moscou [308]. Nous pourrions également conclure quant à l'existence d'une plateforme contenant les complexes ARN polymérase, située à l'extrémité bourgeonnante des virions, observée dans les particules virales Moscou [308]. Les données préliminaires suggèrent que la résolution des images est supérieure à celle atteinte lors de l'étude des particules Moscou et England et permettront donc d'obtenir des reconstructions plus détaillées (P. Bron, communication personnelle). Enfin, étant donné que l'ARNv NA n'a pas de partenaire dans notre réseau *in vitro* et que, d'après la littérature, ce segment ne semble pas jouer de rôle majeur dans l'empaquetage

du génome, il serait intéressant de le remplacer par un segment NA étranger et d'évaluer la perturbation, ou non, de l'organisation interne des RNPv du fond génétique PR8.

Lors de mon travail de doctorat, aucune des trois interactions identifiées *in vitro* et testées *in viro* n'a été réellement validée. Une stratégie purement *in vitro* paraît donc trop aléatoire pour identifier des interactions entre ARNv jouant un rôle dans leur empaquetage. Il est par conséquent primordial de réaliser les expériences de MIME *in viro*, afin d'identifier, de manière exhaustive, les nt importants pour la sélection et l'empaquetage des ARNv ainsi que les structures des signaux d'empaquetage dans un contexte infectieux. Dans ce cas, les cellules HEK293T seront co-transfectées avec un nouveau système de génétique inverse à deux plasmides permettant l'expression de sept ARNv sauvages et de toutes les protéines virales et un plasmide exprimant la banque d'ARNv mutés. La machinerie virale réalisera elle-même la sélection, puisque les mutations qui interfèrent avec l'empaquetage seront uniquement retrouvées dans la population d'ARN cellulaires, alors que les mutations qui n'ont pas d'effet se trouveront à la fois dans les cellules et dans les virions. L'analyse comparative des deux populations se fera grâce aux techniques de séquençage à haut débit. Dans le cas du segment NP, les nt importants pour l'empaquetage seront comparés aux nt impliqués dans les interactions ARN/ARN, identifiés par MIME *in vitro*. Ainsi, nous pourrions évaluer si les nt importants pour l'empaquetage du génome se superposent aux nt impliqués dans les interactions ARN/ARN *in vitro*. Actuellement, certaines étapes du MIME *in viro* restent à mettre au point, mais le protocole devrait incessamment être testé sur les segments HA et NA, avant d'être appliqué aux six autres segments.

Une autre stratégie consiste à développer une méthode de recherche des interactions ARN/ARN « à haut-débit » par cartographie chimique. Un des objectifs du laboratoire est de réaliser la cartographie des ARNv en présence de la protéine NP dont l'influence sur la structure secondaire de l'ARNv reste à déterminer. La comparaison de la structure de l'ARNv seul à celle de l'ARNv en présence de la protéine NP permettra de déterminer i) dans quelle mesure sa fixation affecte la structure de l'ARNv, ii) quelle est la proportion de nt de l'ARNv en contact direct avec la NP, iii) si la NP se fixe à certaines séquences/structures particulières des ARNv et iv) si des structures secondaires persistent en présence de la protéine NP, qui pourraient être impliquées dans l'établissement d'interactions intermoléculaires et donc dans la formation du réseau nécessaire à

l’empaquetage. De manière importante, une approche de cartographie au sein de virions natifs, couplée aux techniques de séquençage à haut-débit, est également envisagée. Dans ce cas, la réactivité des nt au sein de virions natifs sera comparée à celle de RNPv extraites des virions par traitement doux aux détergents, préservant leur structure. Ainsi, cette approche devrait permettre d’appréhender toutes les interactions ARN/ARN qui s’établissent au sein des particules virales. Les interactions identifiées devront ensuite être validées, *in vitro* et *in viro*.

Les données de la littérature sont en faveur de l’existence d’un complexe supramoléculaire entre les huit ARNv des IVA qui serait maintenu par des interactions ARN/ARN entre les signaux d’empaquetage. Ce complexe pourrait s’assembler en partie ou totalement durant le transport des RNPv jusqu’au site de bourgeonnement, grâce à des signaux qui pourraient être des signaux de *bundling*, potentiellement différents des signaux d’empaquetage [158]. L’assemblage spatio-temporel de ces complexes devra notamment être élucidé afin de valider ce modèle. Ainsi, il semble crucial d’évaluer si des mutations importantes pour l’empaquetage ont également un rôle dans la formation des sous-complexes et du complexe supramoléculaire, visualisables par hybridation *in situ* en fluorescence sur des molécules uniques (smFISH, *single-molecule fluorescence in situ hybridization*) [52] dans le cytoplasme des cellules infectées. La résolution des microscopes actuels ne permet cependant pas de confirmer s’il existe une interaction physique entre les différentes RNPv en route vers la membrane ou si la co-localisation est uniquement due à la forte concentration de RNPv dans le cytoplasme. La résolution peut cependant être augmentée en combinant des analyses de smFISH avec du FRET ou avec de la microscopie de super résolution.

Le but ultime du laboratoire est de comprendre les règles qui gouvernent l’empaquetage du génome des IVA. Il est probable que ces règles régissent également le réassortiment génétique, clairement biaisé, entre souches virales. Récemment, le laboratoire a amené la première preuve du rôle d’une interaction ARN/ARN dans l’empaquetage et dans le réassortiment génétique [307]. Ainsi, l’identification de toutes les séquences et les structures nécessaires à l’empaquetage du génome constitue une étape cruciale pour comprendre l’émergence de virus réassortants. Les données obtenues par MIME *in vitro*, et dans un avenir proche, par MIME *in viro*, fourniront des clés essentielles

dans ce domaine. Par la suite, il sera possible d'appliquer l'approche MIME à des IVA circulants chez les oiseaux, les porcs et l'Homme, afin d'analyser la conservation de signaux d'empaquetage et des interactions ARN/ARN. Ces connaissances sont primordiales afin d'évaluer quelles combinaisons génomiques sont susceptibles de s'établir. À long terme, il sera par conséquent possible d'améliorer l'évaluation des risques d'émergence de nouveaux IVA, potentiellement pandémiques, de produire les virus par génétique inverse, de tester leur pathogénicité à l'aide de modèles animaux et de produire à l'avance les semences vaccinales, afin de démarrer la production à grande échelle en cas de risque de pandémie.

Enfin, les travaux du laboratoire pourraient également avoir un impact considérable dans la production industrielle des vaccins antigrippaux. À ce jour les vaccins restent le meilleur moyen de se prémunir contre les IVA, mais en raison de leur évolution rapide et constante il est nécessaire d'adapter annuellement la composition vaccinale. La durée nécessaire à la production des vaccins, saisonniers ou pandémiques, est de cinq à six mois, ce qui est trop long, notamment en cas de pandémie grippale. En effet, les souches vaccinales sont le plus souvent obtenues par réassortiment génétique entre la souche PR8 et une souche d'IVA sélectionnée par l'OMS. Cette étape est limitante, car les virus réassortants d'intérêt se répliquent souvent lentement. Les résultats du laboratoire suggèrent que des incompatibilités entre les segments d'ARNv peuvent constituer le mécanisme moléculaire qui empêche l'émergence de virus réassortants dont les capacités répliquatives ne sont pas pour autant altérées. En parallèle, l'étude menée en collaboration avec l'équipe du Dr L. Brown a également démontré que la co-ségrégation des ARNv PB1 et NA de la souche Udorn dans le fond génétique vaccinal PR8 est due à une interaction entre ces deux partenaires. Cette co-ségrégation est majoritaire malgré le fait que ces virus ne se répliquent pas mieux que ceux possédant le segment PB1 de la souche PR8. Par conséquent, l'identification du réseau d'interactions ARN/ARN de la souche PR8 et la caractérisation des séquences impliquées dans les interactions faisant intervenir les ARNv d'intérêt épidémiologique, HA et NA, pourrait permettre d'améliorer l'efficacité de l'étape de réassortiment génétique. En effet, en adaptant les séquences d'interaction des ARNv partenaires de HA et NA au sein de PR8, leur interaction avec les ARNv HA et NA des souches circulantes pourrait être favorisée pour guider leur co-empaquetage dans le fond génétique PR8 (virus 6+2). Si un autre segment de la souche circulante est impliqué dans

l'emballage des segments HA et NA dans le fond génétique PR8, comme dans le cas de l'ARNv PB1 de la souche Udorn, il est également envisageable de produire des souches vaccinales contenant ces trois segments étrangers (virus 5+3).



# Selective packaging of the influenza A genome and consequences for genetic reassortment

Marie Gerber<sup>1\*</sup>, Catherine Isel<sup>1\*</sup>, Vincent Moules<sup>2,3</sup>, and Roland Marquet<sup>1</sup>

<sup>1</sup> Architecture et Réactivité de l'ARN, Université de Strasbourg, Centre National de Recherche Scientifique (CNRS), Institut de Biologie Moléculaire et Cellulaire (IBMC), 15 rue René Descartes, 67084 Strasbourg, France

<sup>2</sup> Virologie et Pathologie Humaine, Université Lyon 1, EA4610, Faculté de Médecine RTH Laennec, 69008 Lyon, France

<sup>3</sup> VirNext, Université Lyon 1, EA4610, Faculté de Médecine RTH Laennec, 69008 Lyon, France

**Influenza A viruses package their segmented RNA genome in a selective manner. Electron tomography, biochemical assays, and replication assays of viruses produced by reverse genetics recently unveiled molecular details of this mechanism, whereby different influenza viral strains form different and unique networks of direct intermolecular RNA–RNA interactions. Together with detailed views of the three-dimensional structure of the viral ribonucleoproteins, these recent advances help us understand the rules that govern genome packaging. They also have deep implications for the genetic reassortment processes, which are responsible for devastating pandemics.**

## General features of influenza A viruses

Influenza A viruses (IAVs) are the causative agents of recurrent influenza epidemics and occasional severe pandemics in humans and animals. According to the World Health Organization (WHO), IAV infections remain a major public health issue as seasonal epidemics claim 250 000 to 500 000 lives annually (<http://www.who.int/mediacentre/factsheets/fs211/en/>). The first well-documented human influenza pandemic, the 1918 Spanish flu, caused the death of ~50 million people, and was followed by the 1957 Asian and the 1968 Hong Kong flu pandemics. The first pandemic of the 21st century emerged in 2009, creating considerable impact on global health and economics [1]. Occasional deadly infections of humans by avian influenza viruses, including the recent human H7N9 cases [2], remind us of the constant threat of the next flu pandemic. The wide dissemination of highly pathogenic avian influenza (HPAI) H5N1 viruses [3] has resulted in intraspecies transmission in mammals (zoonotic infections), even if at a limited scale, and in a number of highly lethal human infections (<http://www.who.int>). This situation, which provides opportunities for the virus to adapt to the new host,

might not evolve, but recent studies have nevertheless demonstrated how HPAI H5N1 viruses could reach aerosol transmissibility and thus be a direct source for pandemics by increasing the ability of sustained human-to-human transmissibility [4].

IAVs are members of the *Orthomyxoviridae* family, comprising three types of influenza viruses (A, B, and C) as well as Thogoto-like viruses, Isavirus, and one tentative genus, *Quarjavirus* [5]. IAVs are characterized by a segmented genome that consists of eight single-stranded negative-sense viral RNA molecules (vRNA1 to vRNA8) varying in length between 890 and 2341 nucleotides (nts). The gene segments are numbered from one to eight or named after the main protein they encode (Table 1). The IAV genome expresses ten essential proteins as well as several more recently characterized accessory proteins [6–12] (Table 1). IAVs are classified into antigenic subtypes based on their surface glycoproteins, with hemagglutinin (HA) proteins falling into classes H1 to H18 and neuraminidase (NA) proteins falling into classes N1 to N11. Although only a limited number of these HAs and NAs have been isolated from viruses infecting humans, the H1–H16 and N1–N9 subtypes are all maintained in aquatic bird populations [13], which constitute the natural reservoir of IAVs, and the H17N10 and H18N11 viruses were recently identified in bats [14].

vRNAs all share the same organization (Figure 1A): a central open reading frame that encodes one or more proteins, in antisense orientation, that is flanked by two short (19–58 nts) untranslated regions (UTRs). The 3' and 5' termini of the vRNAs consist of 12 (Uni12, U12) and 13 (Uni13, U13) nts, respectively, that are highly conserved between vRNAs and between species (reviewed in [13,15–19]). These conserved sequences are partially complementary and anneal to form a hairpin structure essential for transcription and replication (reviewed in [13,15–19]). Each vRNA is associated to the PB2–PB1–PA polymerase complex, which binds to the vRNA termini, and the RNA-binding protein NP [15,20–22] to form a viral ribonucleoprotein (vRNP).

Genome segmentation confers significant evolutionary advantages to the virus by allowing genetic reassortment, that is, the exchange of gene segments when two, or more,

Corresponding author: Marquet, R. ([r.marquet@ibmc-cnrs.unistra.fr](mailto:r.marquet@ibmc-cnrs.unistra.fr)).

Keywords: influenza A virus; packaging; RNA–RNA interactions; supramolecular complex; genetic reassortment.

\*These authors contributed equally to this review.

0966-842X/

© 2014 Elsevier Ltd. All rights reserved. <http://dx.doi.org/10.1016/j.tim.2014.04.001>

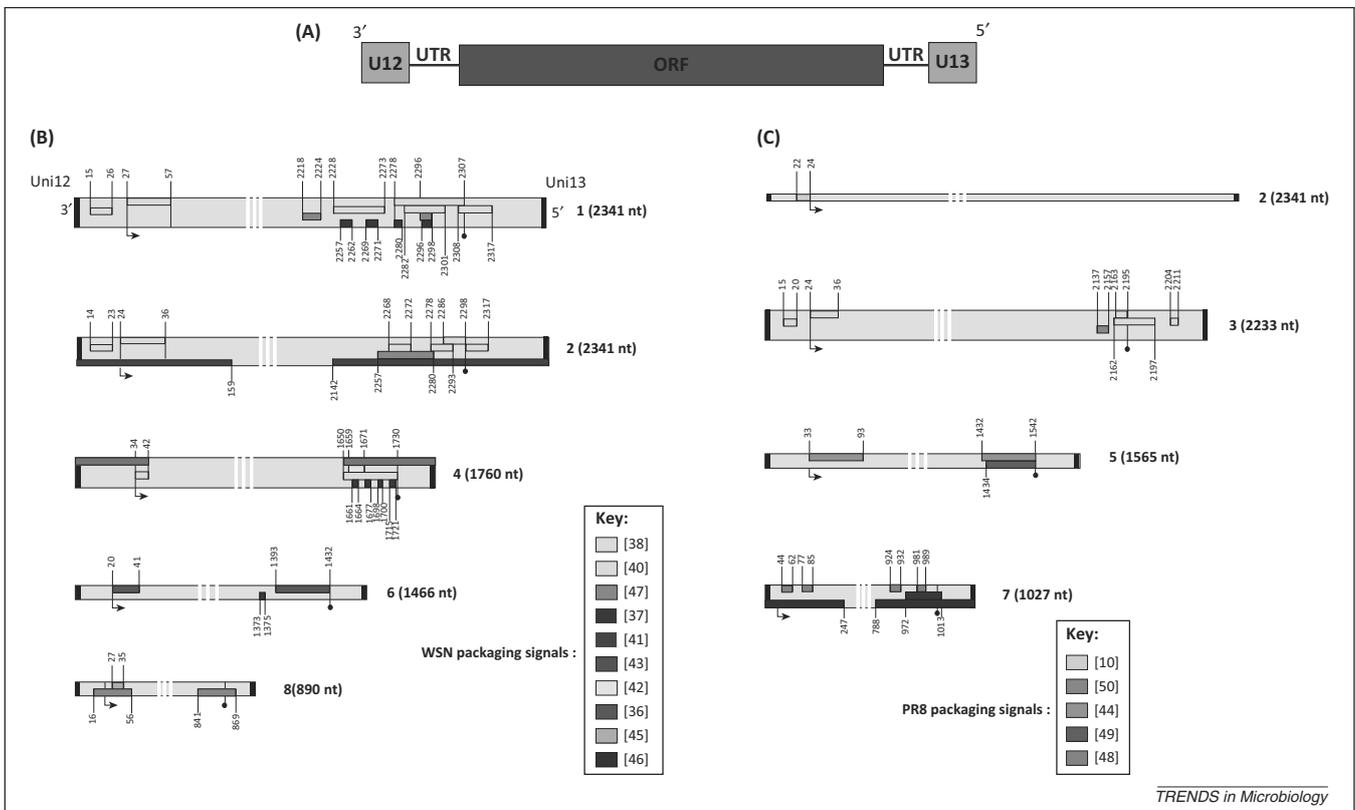
**Table 1. IAVs gene segments, proteins and their functions**

vRNA	Protein name <sup>a</sup>	Main function of the protein
1	PB2	Cellular mRNA cap recognition and binding
2	PB1 PB1-F2* N40*	RNA-dependent RNA polymerase Proapoptotic activity Unknown
3	PA PA-X* PA-N155* PA-N182*	Endonuclease (implicated in cap snatching) Repression of cellular RNA polymerase II-mediated gene expression Probably involved in viral replication Probably involved in viral replication
4	HA	Cellular receptor binding and viral membrane/endosome fusion
5	NP	vRNA binding, vRNP nuclear export, and vRNA replication
6	NA	Neuraminidase activity (cleavage between HA and sialic acid)
7	M1 M2 M43*	Matrix protein, vRNP nuclear export, and virus budding Ionic channel (important for virion acidification and vRNP release) Can functionally replace the M2 protein
8	NS1 NS2-NEP NS3*	Antiviral response inhibition vRNP nuclear export Provides replicative gain-of-function

<sup>a</sup>Auxiliary proteins are marked by asterisks (\*).

viruses coinfect the same cell. However, it undeniably complicates virion assembly, as IAV particles must incorporate at least one copy of each vRNA to be replication-competent and thus fully infectious. Evidence has accumulated in the past 10 years to support a selective packaging of the influenza A genome, implying the existence of specific features on each vRNA, allowing for discrimination

between the eight vRNAs. Despite the identification of *cis*-acting packaging signals in all eight vRNAs, for two human H1N1 strains, the molecular mechanisms that govern packaging have only recently started to be unveiled. This review will highlight recent advances towards the understanding of the selective packaging process of IAVs as well as the implications for genetic reassortment.



**Figure 1.** Mapping of the packaging regions for human A/WSN/33 (H1N1) (WSN) and A/PuertoRico/8/34 (H1N1) (PR8) influenza strains. **(A)** Genetic organization of the influenza A viral gene segments. Abbreviations: U12, unique 12 nucleotide (nt) conserved sequence; U13, unique 13 nt conserved sequence; UTR, untranslated region; ORF, open reading frame. **(B,C)** Coding and noncoding regions important for packaging of influenza WSN (B) and PR8 (C) strains. vRNAs are drawn and numbered according to the conventional representation and nomenclature for negative-strand RNA viruses, from 3' to 5'. The U12 and U13 sequences are represented as black boxes at each end of the vRNAs. The arrows correspond to the translation initiation sites, the black dots to the stop codons. The sequences referenced as important for packaging are color-coded according to the publications that described them [10,36–38,40–46,48–50]. Note that the vRNAs are not drawn to scale. Abbreviation: vRNA, viral RNA.

### Evidence for selective packaging of the influenza A virus genome

During the packaging process, IAVs have: (i) to discriminate the vRNAs from the positive-sense viral complementary RNAs (cRNAs), mRNAs, and the multitude of cellular RNAs; and (ii) to achieve the co-packaging of at least one copy of each gene segment. Historically, two models have been proposed for the incorporation of the eight vRNAs into infectious influenza A particles [23]. A random incorporation model, where the virus discriminates the vRNAs against other RNAs but does not selectively choose one copy of the eight vRNAs, implies the existence of a common feature in all the vRNAs. Numerous data, detailed below, now suggest a specific packaging of the IAV segmented genome, in turn predicting the existence of segment-specific features in each vRNA [23].

#### Genetic and biochemical evidence

For some IAV strains, the number of viral particles able to promote cell killing activity, interferon induction, and interferon suppression is much greater than the number of particles able to form plaques (plaque forming units, PFUs) [24,25]. In line with these studies, it was recently suggested, based on the observation of the expression of incomplete sets of viral proteins, that a large number of influenza virions may not contain eight functional gene segments [26]. Even though this might be the case, the possibility remains that the eight genes are not fully expressed due to failure to transport the entire set of vRNPs to the nucleus, to malfunctions of the polymerase complex, to lethal mutations or to nicks in the vRNA segments. Indeed, early studies using electrophoresis of vRNA revealed equimolar ratios of each of the vRNAs within purified virus [27,28], whilst genetic studies also suggested that IAVs usually bear a haploid genome [29–31]. In addition, equimolar levels of each segment were measured inside the viral particles, even when they were present at different ratios in infected cells [32,33]. More recently, single molecule fluorescence *in situ* hybridization (FISH) analysis directly demonstrated that the majority of viral particles package only one copy of each RNA segment [34]. Indeed, two virus-like RNAs derived from identical vRNA segments (NA or PB2) were found to compete for incorporation into progeny virions that package only single copies of each of the vRNAs [35].

#### Cis-acting packaging regions

Naturally occurring defective interfering RNAs (DI RNAs) derived from every vRNA have been observed under high multiplicity of infection in cell culture and/or in natural influenza infections (reviewed in [23]). DI RNAs carry large deletion(s) in the central coding region whilst systematically preserving the untranslated regions and variable portions of the terminal sequences of the coding region of the segment from which they are derived. Importantly, DI RNAs specifically compete with their parental segment for packaging (reviewed in [23]), strongly suggesting that segment-specific packaging signals are present within ~100 to ~300 nts from each end of the segments.

Usage of reverse genetics, on the laboratory-adapted A/WSN/33 (H1N1) (WSN) strain and the A/PuertoRico/8/34

(H1N1) (PR8) influenza strain used for vaccine production allowed mapping of sequences, in all eight vRNAs, that promote their incorporation into the viral particles (Figure 1B). In a seminal study, Fujii *et al.* [36] demonstrated that sequences in the terminal regions of the coding region drive the incorporation of vRNA6 into virions. Subsequent analysis of reporter constructs derived from vRNA1 [37–40], vRNA2 [10,38–41], vRNA3 [38–40], vRNA4 [42,43], vRNA5 [44], vRNA6 [37], vRNA7 [45,46], and vRNA8 [45,47] also indicated that the minimal sequences needed for efficient packaging extend beyond each UTR, to include nine to 80 nts of adjacent coding sequence at either end of the segment (Figure 1B,C).

Interestingly, mutations in the coding region of a particular vRNA segment also affect, at least in some cases, the incorporation efficiency of other vRNAs [42,48–50], probably reflecting the existence of a network of interactions between the vRNPs [23]. In addition, several studies indicated that not all vRNAs are equally important for packaging and virus assembly [40,42,48–50], suggesting that vRNA packaging is a hierarchical process. A recent systematic study on the PR8 virus showed that vRNA1, vRNA3, vRNA5, and vRNA7 play a more important role than the remaining four vRNAs in the packaging process [51]. Of note, vRNA4 and vRNA6 were never identified as playing a crucial role in packaging: this observation might underline an evolutionary process facilitating exchange of the gene segments encoding the surface antigenic protein by genetic reassortment.

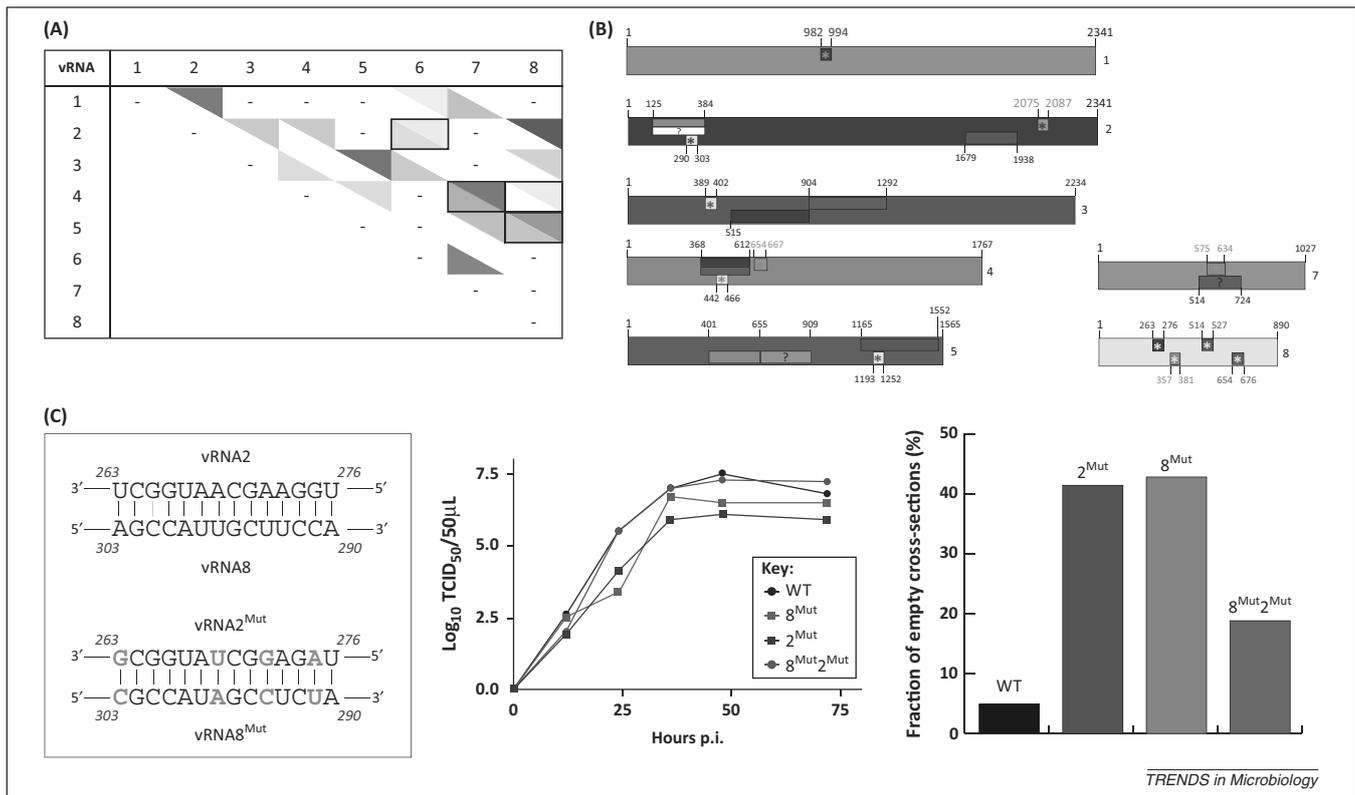
The terminal coding regions of vRNAs are more highly conserved than the central ones, and it has been proposed that clusters of highly conserved codons might constitute packaging signals [37]. Analysis of several of these codons showed that some, but not all, are involved in packaging and indicated that within the defined packaging regions, the signal is discontinuous [37,39,42,48–50] (Figure 1B,C). Notably, some of these codons appeared to be important for vRNA packaging in WSN, but not PR8, indicating that not all packaging signals are conserved [50].

#### Mechanistic insight into the packaging process

Although the existence of segment-specific packaging signals was clearly demonstrated at that point, the molecular basis underlying the selective packaging model was still elusive. A selective packaging mechanism implies indirect or direct interactions between vRNAs [23], the latter hypothesis being the most attractive [36] as no viral or cellular protein specifically recognizing an IAV segment-specific packaging signal has been identified so far.

#### Intermolecular networks between influenza A vRNAs maintained by direct RNA–RNA interactions

The eight vRNAs of a human H3N2 (A/Moscow/10/99) virus were recently shown to be involved *in vitro* in a single interaction network, where each vRNA interacts with at least one other partner [52] (Figure 2A). Surprisingly, the interaction network for an avian H5N2 (A/Finch/England/205/91) virus was found to be radically different [52,53] (Figure 2A). In the case of the H3N2 virus, several deletions of sequences known to encompass packaging signals (in WSN or/and PR8) strongly affected intermolecular



**Figure 2.** Intermolecular vRNA networks and biological function of a vRNA–vRNA interaction. **(A)** Intermolecular network between the eight vRNAs of a human H3N2 virus (blue) and an avian H5N2 (red) virus *in vitro*. The intensity of the blue and red colors is proportional to the weight fraction of the vRNA involved in the intermolecular complexes [52,53]. Interactions that are conserved in the two viruses are boxed. When there is no interaction, there is no triangle. **(B)** vRNA regions of the avian H5N2 vRNAs involved in intermolecular interactions. vRNAs are drawn to scale. Regions delineated by large deletions are presented by rectangles that have the same color as the partner vRNA. Deletions that have limited effects are indicated by question marks. Sequences narrowed down by oligodeoxynucleotides mapping and/or site-directed mutagenesis are indicated by squares containing asterisks [53,55]. **(C)** Importance of the interaction between vRNA2 and vRNA8 for replication and packaging of the avian H5N2 virus. Left panel: complementarity between wild type (WT) vRNA2 and vRNA8 and *trans*-complementary silent mutants vRNA2<sup>Mut</sup> and vRNA8<sup>Mut</sup>. vRNAs are drawn and numbered according to the conventional representation and nomenclature for negative-strand RNA viruses, from 3' to 5'. Middle panel: typical replication curves of WT and mutant viruses. Right panel: percentage of empty electron microscopy cross-sections of WT and mutant viruses (drawn from data in [55]). Abbreviation: vRNA, viral RNA.

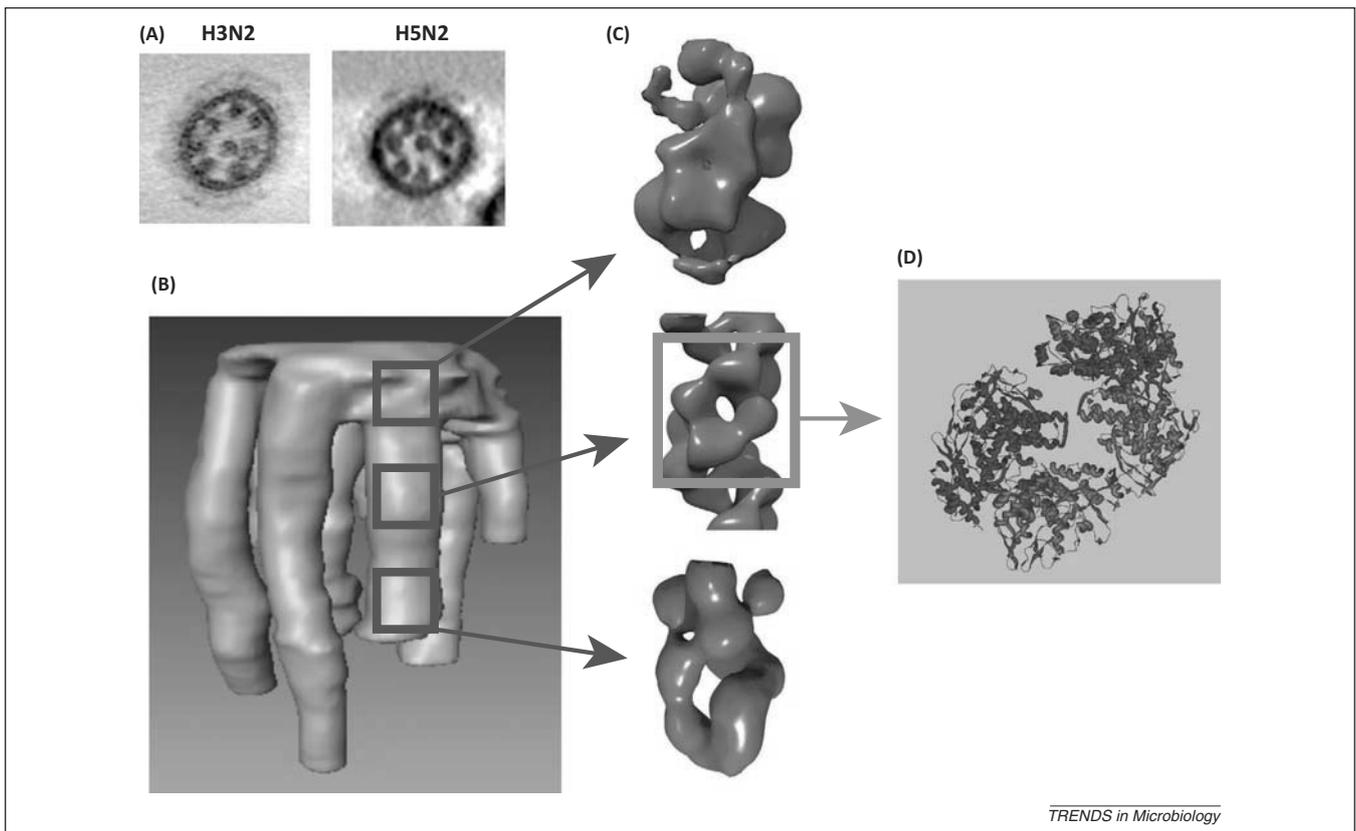
vRNA interactions *in vitro*, suggesting that interactions detected *in vitro* are relevant for the packaging process [52]. Unexpectedly, in the case of the H5N2 virus, several vRNA–vRNA interactions involve sequences not previously recognized as packaging signals and are generally more centrally located [53] (Figure 2B). In line with these observations, a theoretical model based on repulsive forces between identical vRNPs balanced by base pairing between vRNAs [54] predicts that the vRNA–vRNA interactions would not be conserved between independently evolved viruses. Why repulsion between two identical vRNPs would be stronger than between different vRNPs, however, remains unclear.

The precise identification of the nucleotides mediating *in vitro* interactions between several pairs of H5N2 vRNAs opened up the possibility to test their functional role *in cellulolin virio* using *trans*-complementary, silent mutations. Using this strategy, Gavazzi *et al.* proved that the interaction between vRNA2 and vRNA8 is crucial for optimal viral replication and for the selective co-packaging of the interacting vRNAs [55] (Figure 2C). This study demonstrated for the first time the existence and functional role of a direct interaction between two vRNAs. Combined with the demonstration of vRNA interactions

networks *in vitro* [52,53,56], these results strongly suggest that the eight vRNPs are packaged as a single supramolecular complex held together by direct vRNA–vRNA base pairing.

#### Non-random three-dimensional organization of vRNPs inside virus particles

Tomography and transmission electron microscopy of various spherical and filamentous IAVs revealed that vRNPs are usually organized in a ‘7+1’ pattern, with seven vRNPs surrounding a central segment [57–59] (Figure 3A). This ‘7+1’ vRNP pattern is more frequently observed in budding virions than in purified viral particles [56–58], suggesting that the ‘7+1’ array might not always be conserved after virion release. The individual vRNPs are suspended from the interior of the viral envelope at the distal end of the budding virions, in an orderly parallel bundle, oriented perpendicular to the budding tip [58,60,61]. Although contacts between vRNPs were observed to form an apical platform in 3D reconstruction of individual human H3N2 viruses by electron tomography (Figure 3B) [52,56], numerous contacts were also observed all along the vRNPs in WSN virions [61]. These differences might in part have a technical origin as Noda *et al.* analyzed thinner



**Figure 3.** Organization and structure of IAV ribonucleoproteins. **(A)** Typical '7+1' pattern of the vRNPs observed in electron tomography cross-sections of a human H3N2 [56] and an avian H5N2 [53] virus. **(B)** 3D reconstruction of the interior of a human H3N2 viral particle. The eight vRNPs emerge from a common platform at the budding tip of the virions. The different lengths of the vRNPs allowed identification of several of them [56]. **(C)** Cryo-EM electron density of the polymerase end (top), helical part (middle), and loop end (bottom) of a vRNP [22]. These electron densities were drawn from the EM Data Bank entries EMD-2209, EMD-2211, and EMD-2212 using the UCSF Chimera Software. **(D)** Docking of the NP crystal structure into the density map of the helical region [22] drawn from the PDB entry code 2YMN using iMol. NP protomers corresponding to the two strands of the vRNP double helix are shown in red and blue. Abbreviations: IAV, influenza A virus; EM, electron microscopy; PDB, Protein Data Bank; vRNP, viral ribonucleoprotein.

virtual viral cross-sections [61]; however, Fournier *et al.* [52,56] and Gavazzi *et al.* [53], who used the same experimental set-up, reported more inter-vRNP contacts in avian H5N2 as compared to human H3N2 virions.

If vRNPs are assembled as a supramolecular complex maintained by base pairings between them, these interactions should strongly constrain the internal organization of vRNPs within viral particles. Recent electron tomography studies identified some of the vRNPs according to their length [52,56,61] (Figure 3B), with the limitation that identification of individual vRNPs within viral particles by electron tomography is restricted by the fact that several vRNAs have identical or similar lengths and by the difficulty of accurately measuring the vRNP size. Importantly, these studies reported a nonrandom distribution of vRNPs within virions. A few different distributions coexisted in different virions of the same IAV strain and these distributions were different in WSN and human H3N2 viruses. In addition, a recent study suggested that even though most vRNPs are incorporated into the virus particles with their polymerase complex at the budding tip of the virions, some are incorporated upside down [62]. If this behavior turns out not to be an oddity, it changes fairly profoundly the model that places the packaging signals exclusively near the extremities of the vRNAs and

increases the number of possible vRNA–vRNA interactions, outside the vRNA termini.

#### *The impact of the vRNP structure on vRNA–vRNA interactions*

The *in vitro* vRNA–vRNA bandshift experiments discussed above [52,53,55,56] were, for the main part, performed in the absence of viral proteins and it is therefore unclear whether all *in vitro* interactions take place *in virio*, and conversely, whether all *in virio* interactions can be identified *in vitro*. In cells and virions, NP forms the protein scaffold of the helical genomic vRNPs [20,22,63–65], which, strikingly, can be maintained in the presence of unrelated RNA or even in the absence of vRNA [64]. Recently, the structure of full-length native vRNPs isolated from virions [20] or full-length recombinant vRNPs isolated from cells [22] have been reconstructed using cryo-electron microscopy [20,22] (Figure 3C). In both models, the vRNP adopts a double helical structure with two antiparallel strands maintained by interactions between NP molecules (Figure 3C,D). Significant differences exist between the vRNP models proposed by these two groups, including the handedness of the helix [20,22]. However, both groups suggested that the vRNA is located at the surface of the NP scaffold; no density corresponding to the vRNA was

observed in the density maps, suggesting that its structure is not periodic [20,22].

Chemical probing experiments suggested that all secondary structure motifs of vRNAs are completely melted by NP [66,67], in contrast with nuclease probing experiments that indicated the persistence of secondary structure elements [68]. Importantly, although the NP/vRNA ratio in vRNP preparations is ~1 NP to 24 nts [69–71], recent studies of salmon anemia virus (ISAV) NP indicated that an NP protomer directly binds 12 nts of RNA, in keeping with the size of the RNA-binding groove of the ISAV NP inferred from its crystallographic structure [19]. The remaining RNA sequences would form a linker between the NP protomers and might adopt secondary/tertiary structures. Notably, vRNPs or complementary RNPs were recently isolated using an RNA-based affinity purification strategy, by introducing an artificial secondary structural element in the vRNA [72]. The feasibility of such experiments indicates that secondary structure elements can protrude from the influenza vRNPs. From the above-discussed data, one could speculate that part of the influenza vRNA exists in an NP-free state, possibly leaving local secondary and/or tertiary structures available for the formation of organized vRNP supramolecular complexes linked by specific RNA–RNA interactions.

#### *A two level packaging mechanism?*

In the specific packaging model, the packaging machinery must: (i) discriminate vRNAs embedded in vRNPs from positive strand viral RNAs, which include cRNAs also packed into complementary RNPs (cRNPs) and viral mRNAs, but also from cellular RNAs; and (ii) discriminate between the eight vRNAs themselves to ensure packaging a complete set of genomic segments. Past and recent evidence suggest that different signals ensure these two levels of discrimination. To face the first challenge of discriminating vRNAs from all the other nonviral RNAs, signals that are common to all influenza A vRNAs are required. The only sequences that match this criterion are the conserved U12 and U13 sequences, which are selectively recognized by the viral polymerase [15,20–22]. The groups of Palese and Kawaoka both demonstrated the importance of the noncoding regions (NCRs) (comprising U12 and U13 and the UTRs) at both ends of the vRNAs for incorporation into virus-like particles (VLPs) [40,73]. Recently, the Kawaoka group further dissected the mechanism by showing that, for one particular vRNA, the NCRs only permitted incorporation of this vRNA into VLPs, whereas the presence of the previously defined packaging regions (including the NCRs and part of the 3' and 5' coding sequences) serve as a bundling signal to ensure incorporation of a complete set of vRNPs into the virus particles [74]. It is tempting to speculate that the U12 and U13 sequences allow discrimination against cellular RNAs and viral mRNAs, whereas sequence-dependent vRNA–vRNA interactions allow discrimination between the eight vRNAs.

From a mechanistic point of view, selection of the vRNAs for packaging might take place during nuclear export, during cytoplasmic trafficking and/or during budding of the viral particles (Figure 4). Unlike cellular RNAs and viral mRNAs, neosynthesized vRNPs are exported

from the nucleus via a Crm1-dependent pathway. NS2 makes a bridge between Crm1 and M1, which interacts with the vRNPs [75]. The positive-sense cRNA also forms RNPs [72], but these cRNPs are not exported: discrimination between vRNPs and cRNPs relies on a single bulged nucleotide in the base paired vRNA termini [76], possibly via a differential recognition by the viral polymerase, which adopts different conformations when bound to cRNPs or vRNPs. Thus, nuclear export could achieve the first selection level, with U12 and U13, and possibly the polymerase complex, playing a key role in the process.

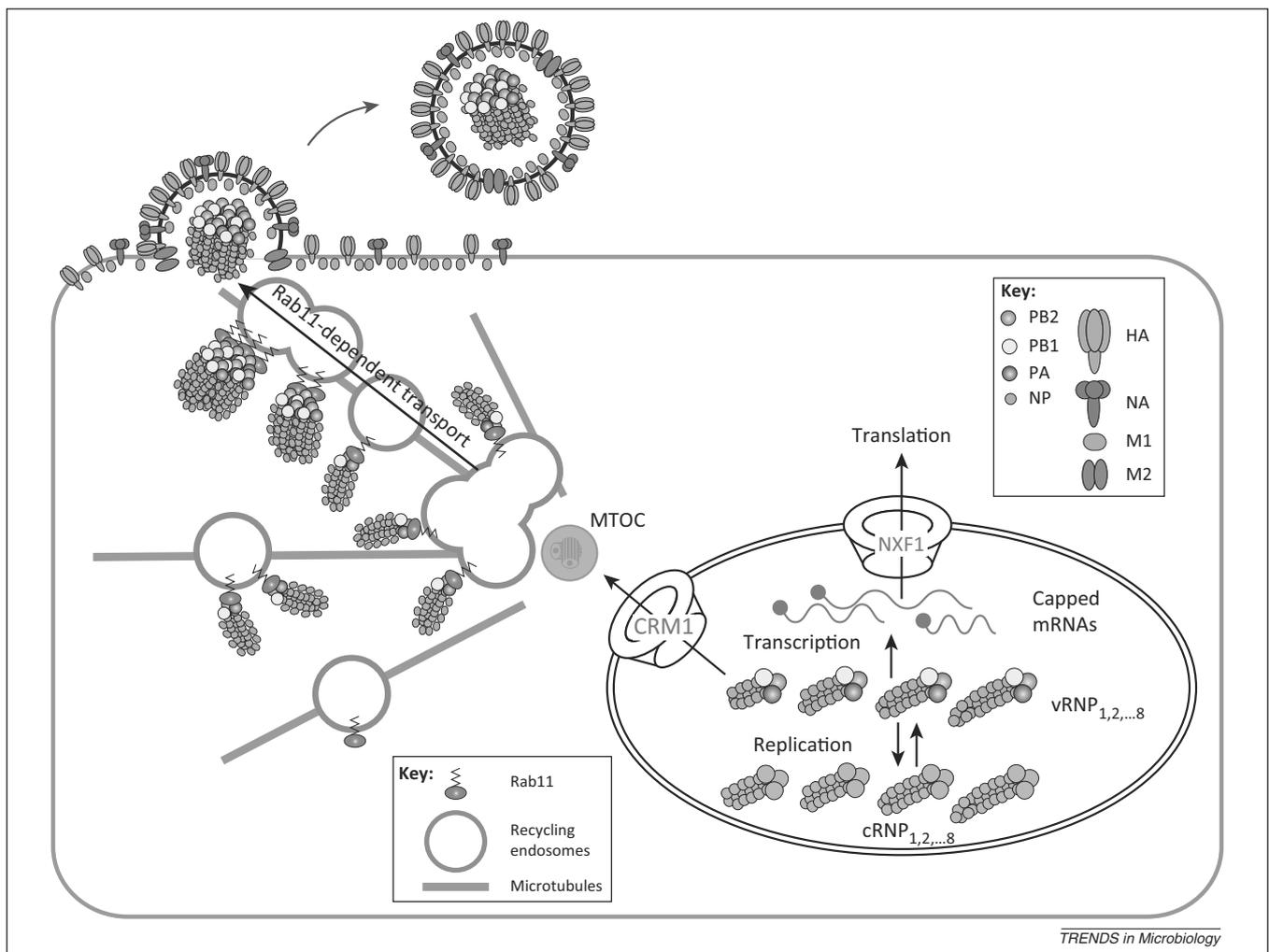
The transport of influenza vRNPs from the perinuclear region to the apical plasma membrane involves the Rab-11 positive recycling endosome system, together with an intact microtubule network for membrane targeting [77–81] (Figure 4). Chou *et al.* recently demonstrated that late during infection, the different influenza vRNAs travel together in the cytoplasm, associated with Rab-11 vesicles [82] that could serve as platforms to gather the vRNAs in close proximity. Specific vRNA–vRNA interactions could take place during transport, and a core of interacting vRNAs might be formed. Such a preformed core complex, made by a limited number of vRNPs, that might not be the same for all viral strains, could constitute a checkpoint for assembly of the correct combination of the eight vRNA fragments, before the vRNPs reach the cell surface where budding takes place. This would explain why, as discussed above, not all vRNAs are equally important for packaging. Notably, the latter hypothesis finds strong support in a very recent publication by Lakdawala *et al.* [83] who showed, using a four-color FISH assay to visualize all eight vRNA segments, that vRNAs are not exported individually from the nucleus, although not all together. Viral RNA subcomplexes then further assemble en route to the plasma membrane via dynamic colocalization events that involve Rab11a-containing recycling endosomes [83].

#### **Implications for genetic reassortment**

The rapid evolution of IAVs is enabled both by genetic drift and genetic shift or genetic reassortment, that is, the exchange of gene segments between two or more viruses that coinfect the same cell. Among the viruses generated by the latter mechanism, some do not encounter any pre-existing immunity in the human population and have therefore a high pandemic potential if they are efficiently transmitted between humans. The first step of this process is the co-packaging of vRNAs from two parental IAVs into the same viral particles. Thus, the packaging mechanism is expected to have deep implications for the genetic reassortment process.

The interactions taking place between the vRNAs of an avian H5N2 virus are not conserved in IAVs circulating in the human population [53,55]. Because interacting vRNAs are preferentially co-packaged [55], it is expected that genetic reassortment between distantly related viruses should be disfavored and the genotype of the reassortant viruses should be nonrandom.

Indeed, when reassortant viruses containing segments from a human and an avian IAV were systematically produced by reverse genetics in an experimental set-up that did not allow competition between segments, 44–58%



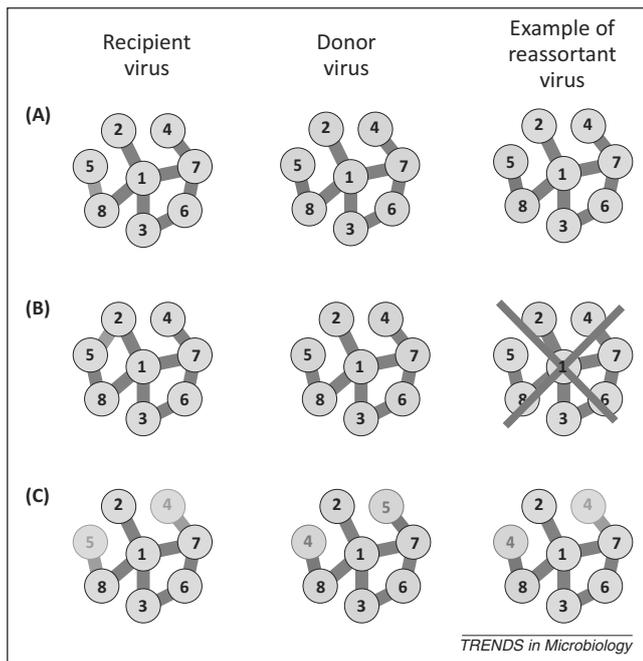
**Figure 4.** Transport, selection, and packaging of vRNPs. Unlike viral and cellular mRNAs, vRNPs are exported via a CRM1-dependent pathway; cRNAs are incorporated into cRNPs that are not exported out of the nucleus. In the cytoplasm, vRNPs concentrate close to the microtubule organizing center (MTOC). They bind recycling endosomes via Rab11 and migrate along the microtubule network. Several vRNP species migrate together, and this process might favor formation of a supramolecular complex held together by base pairing between vRNAs. Abbreviations: cRNA, complementary RNA; vRNA, viral RNA; vRNP, viral ribonucleoprotein.

of the viruses harboring the avian HA segment replicated as efficiently as the parental viruses [84–86]. For the other half of reassortant viruses, replication was probably compromised by incompatibilities between viral proteins forming vRNPs [87,88], by the necessity of keeping the PA and PB2 proteins together for the human H3N2 IAV [89] or by incompatibilities between M and NS2 [84,88]. Remarkably, in natural or experimental competitive situations, the reassortment rate was generally very low and limited numbers of reassortant genotypes were observed [87,89–91]. The fact that reverse genetics generate viruses that replicate efficiently but will not emerge naturally infers the existence of incompatibilities (or suboptimal compatibilities) at the genomic level. In agreement with this hypothesis, genetically similar viruses reassort at high frequency and with no apparent genotype bias [92]. Importantly, Essere *et al.* [90] were recently able to directly demonstrate that suboptimal compatibility between vRNA packaging signals limit genetic reassortment. Indeed, a few silent mutations in the packaging signals were found to profoundly affect the genotype of the reassortant viruses.

Inter-vRNAs interactions are expected to have multiple consequences on genetic reassortment (Figure 5). First, if two segments interact through different (incompatible) sequences in the two parental viruses, they should co-segregate during genetic reassortment, even if the two parental viruses have the same vRNA–vRNA interaction network and the same internal organization (Figure 5A). Second, if two vRNAs interact in one but not in the other parental virus, this interaction might act as a gatekeeper that prevents incorporation of foreign segments in the genetic background in which this interaction exists (Figure 5B). Third, as discussed above the vRNA–vRNA interaction network and internal organization are probably not widely conserved among different IAV strains, and as a consequence most reassortant progeny viral particles probably do not contain a full set of gene segments and are therefore replication-incompetent (Figure 5C).

### Concluding remarks

The existence of a segment-specific packaging process in IAVs is now firmly established. Recent electron microscopy



**Figure 5.** Predicted consequences of inter-vRNAs interactions on genetic reassortment. The recipient and donor viruses are blue and red, respectively. **(A)** Co-segregation of two segments (vRNA5 and vRNA8) that interact through different (incompatible) sequences (blue and red lines) in the two parental viruses, even when the two parental viruses have the same vRNA–vRNA interaction network and the same internal organization. **(B)** An interaction that exists in the recipient virus (between vRNA2 and vRNA5, blue) but not in the donor virus might prevent incorporation of foreign segments in the genetic background in which this interaction exists. **(C)** Co-segregation of two interacting segments (vRNA4 and vRNA8, red) from a donor virus into a recipient virus with a different vRNA–vRNA network and internal organization of the vRNPs results in reassortant viral progeny that does not contain a full set of gene segments and is thus replication-incompetent. Abbreviations: vRNA, viral RNA; vRNP, viral ribonucleoprotein.

### Box 1. Outstanding questions

- **Do vRNA–vRNA interactions play a general role in vRNA packaging?** Is it possible to extend the observation made about the involvement of one particular vRNA–vRNA interaction to other interactions and possibly make it a general rule for different viruses?
- **Where and when in the cell are the vRNP supramolecular complexes assembled?** Is the vRNP supramolecular complex progressively assembled during transport along the microtubule networks in association with Rab-11 positive endosome recycling vesicles?
- **Are the vRNP networks strain- and/or host-specific?**
- **To what extent do incompatibilities at the genomic level bias genetic reassortment? How does the lack of conservation of the vRNA–vRNA interactions, which favor co-packaging of the interacting vRNAs, impact genetic reassortment?**

and tomography, virology, and biochemistry studies provided new mechanistic insight into this process. These studies suggested that the eight vRNPs are incorporated into viral particles as a single supramolecular complex maintained by base pairing between the vRNAs; this complex might be in part or completely formed during transport of the vRNPs to the budding sites. Further studies are needed to demonstrate the general character of the conclusions drawn from these studies and to validate this model (Box 1).

Strikingly, these studies revealed that the vRNA–vRNA interactions and the internal organization of the viral particles differ from one IAV strain to the other. This lack of conservation is expected to have a major impact on genetic reassortment, which is important for generation of both pathogenic reassortant viruses and vaccine seeds. At present, our understanding of the rules underlying co-packaging of vRNAs from two parental strains in a single viral particle is, at best, very limited. Understanding these rules is a formidable challenge for the future, as it is the first step to understand the rules governing genetic reassortment.

### Acknowledgments

This work is supported by the Centre National de la Recherche Scientifique and VirNext. The authors also wish to thank Dr F. Jossinet for his help with Figure 3.

### References

- 1 Simonsen, L. *et al.* (2013) Global mortality estimates for the 2009 influenza pandemic from the GLaMOR Project: a modeling study. *PLoS Med.* 10, e1001558
- 2 Ke, Y. *et al.* (2013) High severity and fatality of human infections with avian influenza A(H7N9) infection in China. *Clin. Infect. Dis.* 57, 1506–1507
- 3 Swayne, D.E. (2012) Impact of vaccines and vaccination on global control of avian influenza. *Avian Dis.* 56, 818–828
- 4 Imai, M. *et al.* (2013) Transmission of influenza A/H5N1 viruses in mammals. *Virus Res.* 178, 15–20
- 5 Presti, R.M. *et al.* (2009) Quarantfil, Johnston Atoll, and Lake Chad viruses are novel members of the family *Orthomyxoviridae*. *J. Virol.* 83, 11599–11606
- 6 Chen, W. *et al.* (2001) A novel influenza A virus mitochondrial protein that induces cell death. *Nat. Med.* 7, 1306–1312
- 7 Jagger, B.W. *et al.* (2012) An overlapping protein-coding region in influenza A virus segment 3 modulates the host response. *Science* 337, 199–204
- 8 Muramoto, Y. *et al.* (2013) Identification of novel influenza A virus proteins translated from PA mRNA. *J. Virol.* 87, 2455–2462
- 9 Selman, M. *et al.* (2012) Adaptive mutation in influenza A virus non-structural gene is linked to host switching and induces a novel protein by alternative splicing. *Emerg. Microbes Infect.* 1, e42
- 10 Wise, H.M. *et al.* (2011) Overlapping signals for translational regulation and packaging of influenza A virus segment 2. *Nucleic Acids Res.* 39, 7775–7790
- 11 Wise, H.M. *et al.* (2009) A complicated message: identification of a novel PB1-related protein translated from influenza A virus segment 2 mRNA. *J. Virol.* 83, 8021–8031
- 12 Wise, H.M. *et al.* (2012) Identification of a novel splice variant form of the influenza A virus M2 ion channel with an antigenically distinct ectodomain. *PLoS Pathog.* 8, e1002998
- 13 Palese, P.A. and Shaw, M.L. (2007) *Orthomyxoviridae: the viruses and their replication*. In *Fields Virology* (5th edn) (Knipe, D.M. and Howley, P.M., eds), pp. 1647–1689. Lippincott Williams & Wilkins
- 14 Tong, S. *et al.* (2013) New world bats harbor diverse influenza A viruses. *PLoS Pathog.* 9, e1003657
- 15 Elton, D. *et al.* (2006) Structure and function of the influenza virus RNP. In *Current Topics in Influenza Virology* (Kawaoka, Y., ed.), pp. 1–36. Horizon Scientific Press
- 16 Gulyaev, A.P. *et al.* (2010) Influenza virus RNA structure: unique and common features. *Int. Rev. Immunol.* 29, 533–556
- 17 Resa-Infante, P. *et al.* (2011) The influenza virus RNA synthesis machine: advances in its structure and function. *RNA Biol.* 8, 207–215
- 18 Ruigrok, R.W. *et al.* (2010) Towards an atomic resolution understanding of the influenza virus replication machinery. *Curr. Opin. Struct. Biol.* 20, 104–113
- 19 Zheng, W. *et al.* (2013) The crystal structure and RNA-binding of an orthomyxovirus nucleoprotein. *PLoS Pathog.* 9, e1003624
- 20 Arranz, R. *et al.* (2012) The structure of native influenza virion ribonucleoproteins. *Science* 338, 1634–1637

- 21 Coloma, R. *et al.* (2009) The structure of a biologically active influenza virus ribonucleoprotein complex. *PLoS Pathog.* 5, e1000491
- 22 Moeller, A. *et al.* (2012) Organization of the influenza virus replication machinery. *Science* 338, 1631–1634
- 23 Hutchinson, E.C. *et al.* (2010) Genome packaging in influenza A virus. *J. Gen. Virol.* 91, 313–328
- 24 Marcus, P.I. *et al.* (2005) Interferon induction and/or production and its suppression by influenza A viruses. *J. Virol.* 79, 2880–2890
- 25 Ngunjiri, J.M. *et al.* (2008) Clonogenic assay of type A influenza viruses reveals noninfectious cell-killing (apoptosis-inducing) particles. *J. Virol.* 82, 2673–2680
- 26 Brooke, C.B. *et al.* (2013) Most influenza A virions fail to express at least one essential viral protein. *J. Virol.* 87, 3155–3162
- 27 Hatada, E. *et al.* (1989) Control of influenza virus gene expression: quantitative analysis of each viral RNA species in infected cells. *J. Biochem.* 105, 537–546
- 28 McGeoch, D. *et al.* (1976) Influenza virus genome consists of eight distinct RNA species. *Proc. Natl. Acad. Sci. U.S.A.* 73, 3045–3049
- 29 Laver, W.G. and Downie, J.C. (1976) Influenza virus recombination. I. Matrix protein markers and segregation during mixed infections. *Virology* 70, 105–117
- 30 Lubeck, M.D. *et al.* (1979) Nonrandom association of parental genes in influenza A virus recombinants. *Virology* 95, 269–274
- 31 Nakajima, K. and Sugiura, A. (1977) Three-factor cross of influenza virus. *Virology* 81, 486–489
- 32 Bergmann, M. and Muster, T. (1995) The relative amount of an influenza A virus segment present in the viral particle is not affected by a reduction in replication of that segment. *J. Gen. Virol.* 76, 3211–3215
- 33 Smith, G.L. and Hay, A.J. (1982) Replication of the influenza virus genome. *Virology* 118, 96–108
- 34 Chou, Y.Y. *et al.* (2012) One influenza virus particle packages eight unique viral RNAs as shown by FISH analysis. *Proc. Natl. Acad. Sci. U.S.A.* 109, 9101–9106
- 35 Inagaki, A. *et al.* (2012) Competitive incorporation of homologous gene segments of influenza A virus into virions. *J. Virol.* 86, 10200–10202
- 36 Fujii, Y. *et al.* (2003) Selective incorporation of influenza virus RNA segments into virions. *Proc. Natl. Acad. Sci. U.S.A.* 100, 2002–2007
- 37 Gog, J.R. *et al.* (2007) Codon conservation in the influenza A virus genome defines RNA packaging signals. *Nucleic Acids Res.* 35, 1897–1907
- 38 Liang, Y. *et al.* (2005) cis-Acting packaging signals in the influenza virus PB1 PB2, and PA genomic RNA segments. *J. Virol.* 79, 10348–10355
- 39 Liang, Y. *et al.* (2008) Mutational analyses of packaging signals in influenza virus PA, PB1, and PB2 genomic RNA segments. *J. Virol.* 82, 229–236
- 40 Muramoto, Y. *et al.* (2006) Hierarchy among viral RNA (vRNA) segments in their role in vRNA incorporation into influenza A virions. *J. Virol.* 80, 2318–2325
- 41 Dos Santos Afonso, E. *et al.* (2005) The generation of recombinant influenza A viruses expressing a PB2 fusion protein requires the conservation of a packaging signal overlapping the coding and noncoding regions at the 5' end of the PB2 segment. *Virology* 341, 34–46
- 42 Marsh, G.A. *et al.* (2007) Specific residues of the influenza A virus hemagglutinin viral RNA are important for efficient packaging into budding virions. *J. Virol.* 81, 9727–9736
- 43 Watanabe, T. *et al.* (2003) Exploitation of nucleic acid packaging signals to generate a novel influenza virus-based vector stably expressing two foreign genes. *J. Virol.* 77, 10575–10583
- 44 Ozawa, M. *et al.* (2007) Contributions of two nuclear localization signals of influenza A virus nucleoprotein to viral replication. *J. Virol.* 81, 30–41
- 45 Fujii, K. *et al.* (2009) Incorporation of influenza A virus genome segments does not absolutely require wild-type sequences. *J. Gen. Virol.* 90, 1734–1740
- 46 Ozawa, M. *et al.* (2009) Nucleotide sequence requirements at the 5' end of the influenza A virus M RNA segment for efficient virus replication. *J. Virol.* 83, 3384–3388
- 47 Fujii, K. *et al.* (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. *J. Virol.* 79, 3766–3774
- 48 Hutchinson, E.C. *et al.* (2008) Mutational analysis of cis-acting RNA signals in segment 7 of influenza A virus. *J. Virol.* 82, 11869–11879
- 49 Hutchinson, E.C. *et al.* (2009) Characterisation of influenza A viruses with mutations in segment 5 packaging signals. *Vaccine* 27, 6270–6275
- 50 Marsh, G.A. *et al.* (2008) Highly conserved regions of influenza A virus polymerase gene segments are critical for efficient viral RNA packaging. *J. Virol.* 82, 2295–2304
- 51 Gao, Q. *et al.* (2012) The influenza A virus PB2, PA, NP, and M segments play a pivotal role during genome packaging. *J. Virol.* 86, 7043–7051
- 52 Fournier, E. *et al.* (2012) Interaction network linking the human H3N2 influenza A virus genomic RNA segments. *Vaccine* 30, 7359–7367
- 53 Gavazzi, C. *et al.* (2013) An in vitro network of intermolecular interactions between viral RNA segments of an avian H5N2 influenza A virus: comparison with a human H3N2 virus. *Nucleic Acids Res.* 41, 1241–1254
- 54 Venev, S.V. and Zeldovich, K.B. (2013) Segment self-repulsion is the major driving force of influenza genome packaging. *Phys. Rev. Lett.* 110, 098104
- 55 Gavazzi, C. *et al.* (2013) A functional sequence-specific interaction between influenza A virus genomic RNA segments. *Proc. Natl. Acad. Sci. U.S.A.* 110, 16604–16609
- 56 Fournier, E. *et al.* (2012) A supramolecular assembly formed by influenza A virus genomic RNA segments. *Nucleic Acids Res.* 40, 2197–2209
- 57 Harris, A. *et al.* (2006) Influenza virus pleiomorphy characterized by cryoelectron tomography. *Proc. Natl. Acad. Sci. U.S.A.* 103, 19123–19127
- 58 Noda, T. *et al.* (2006) Architecture of ribonucleoprotein complexes in influenza A virus particles. *Nature* 439, 490–492
- 59 Oxford, J.S. and Hockley, D.J. (1987) *Orthomyxoviridae*. In *Perspective in Medical Virology: Animal Virus Structure* (Nermut, M.V. and Steven, A.C., eds), pp. 213–232, Elsevier
- 60 Calder, L.J. *et al.* (2010) Structural organization of a filamentous influenza A virus. *Proc. Natl. Acad. Sci. U.S.A.* 107, 10685–10690
- 61 Noda, T. *et al.* (2012) Three-dimensional analysis of ribonucleoprotein complexes in influenza A virus. *Nat. Commun.* 3, 639
- 62 Sugita, Y. *et al.* (2013) The configuration of viral ribonucleoprotein complexes within the influenza A virion. *J. Virol.* 87, 12879–12884
- 63 Ng, A.K. *et al.* (2008) Structure of the influenza virus A H5N1 nucleoprotein: implications for RNA binding, oligomerization, and vaccine design. *FASEB J.* 22, 3638–3647
- 64 Ruigrok, R.W. and Baudin, F. (1995) Structure of influenza virus ribonucleoprotein particles. II. Purified RNA-free influenza virus ribonucleoprotein forms structures that are indistinguishable from the intact influenza virus ribonucleoprotein particles. *J. Gen. Virol.* 76, 1009–1014
- 65 Ye, Q. *et al.* (2006) The mechanism by which influenza A virus nucleoprotein forms oligomers and binds RNA. *Nature* 444, 1078–1082
- 66 Baudin, F. *et al.* (1994) Structure of influenza virus RNP. I. Influenza virus nucleoprotein melts secondary structure in panhandle RNA and exposes the bases to the solvent. *EMBO J.* 13, 3158–3165
- 67 Klumpp, K. *et al.* (1997) Roles of the influenza virus polymerase and nucleoprotein in forming a functional RNP structure. *EMBO J.* 16, 1248–1257
- 68 Yamanaka, K. *et al.* (1990) Reconstitution of influenza virus RNA–nucleoprotein complexes structurally resembling native viral ribonucleoprotein cores. *J. Biol. Chem.* 265, 11151–11155
- 69 Compans, R.W. *et al.* (1972) Structure of the ribonucleoprotein of influenza virus. *J. Virol.* 10, 795–800
- 70 Martin-Benito, J. *et al.* (2001) Three-dimensional reconstruction of a recombinant influenza virus ribonucleoprotein particle. *EMBO Rep.* 2, 313–317
- 71 Ortega, J. *et al.* (2000) Ultrastructural and functional analyses of recombinant influenza virus ribonucleoproteins suggest dimerization of nucleoprotein during virus amplification. *J. Virol.* 74, 156–163
- 72 York, A. *et al.* (2013) Isolation and characterization of the positive-sense replicative intermediate of a negative-strand RNA virus. *Proc. Natl. Acad. Sci. U.S.A.* 110, E4238–E4245
- 73 Luytjes, W. *et al.* (1989) Amplification, expression, and packaging of foreign gene by influenza virus. *Cell* 59, 1107–1113

- 74 Goto, H. *et al.* (2013) The genome-packaging signal of the influenza A virus genome comprises a genome incorporation signal and a genome-bundling signal. *J. Virol.* 87, 11316–11322
- 75 Hutchinson, E.C. and Fodor, E. (2013) Transport of the influenza virus genome from nucleus to nucleus. *Viruses* 5, 2424–2446
- 76 Tchatalbachev, S. *et al.* (2001) The packaging signal of influenza viral RNA molecules. *RNA* 7, 979–989
- 77 Amorim, M.J. *et al.* (2011) A Rab11- and microtubule-dependent mechanism for cytoplasmic transport of influenza A virus viral RNA. *J. Virol.* 85, 4143–4156
- 78 Avilov, S.V. *et al.* (2011) Replication-competent influenza A virus that encodes a split-green fluorescent protein-tagged PB2 polymerase subunit allows live-cell imaging of the virus life cycle. *J. Virol.* 86, 1433–1448
- 79 Bruce, E.A. *et al.* (2012) Role of the Rab11 pathway in negative-strand virus assembly. *Biochem. Soc. Trans.* 40, 1409–1415
- 80 Eisfeld, A.J. *et al.* (2011) RAB11A is essential for transport of the influenza virus genome to the plasma membrane. *J. Virol.* 85, 6117–6126
- 81 Momose, F. *et al.* (2011) Apical transport of influenza A virus ribonucleoprotein requires Rab11-positive recycling endosome. *PLoS ONE* 6, e21123
- 82 Chou, Y.Y. *et al.* (2013) Colocalization of different influenza viral RNA segments in the cytoplasm before viral budding as shown by single-molecule sensitivity FISH analysis. *PLoS Pathog.* 9, e1003358
- 83 Lakdawala, S.S. *et al.* (2014) Influenza A virus assembly intermediates fuse in the cytoplasm. *PLoS Pathog.* 10, e1003971
- 84 Chen, L.M. *et al.* (2008) Genetic compatibility and virulence of reassortants derived from contemporary avian H5N1 and human H3N2 influenza A viruses. *PLoS Pathog.* 4, e1000072
- 85 Li, C. *et al.* (2010) Reassortment between avian H5N1 and human H3N2 influenza viruses creates hybrid viruses with substantial virulence. *Proc. Natl. Acad. Sci. U.S.A.* 107, 4687–4692
- 86 Sun, Y. *et al.* (2011) High genetic compatibility and increased pathogenicity of reassortants derived from avian H9N2 and pandemic H1N1/2009 influenza viruses. *Proc. Natl. Acad. Sci. U.S.A.* 108, 4164–4169
- 87 Li, C. *et al.* (2008) Compatibility among polymerase subunit proteins is a restricting factor in reassortment between equine H7N7 and human H3N2 influenza viruses. *J. Virol.* 82, 11880–11888
- 88 Octaviani, C.P. *et al.* (2011) Reassortment between seasonal H1N1 and pandemic (H1N1) 2009 influenza viruses is restricted by limited compatibility among polymerase subunits. *J. Virol.* 85, 8449–8452
- 89 Hara, K. *et al.* (2013) Co-incorporation of the PB2 and PA polymerase subunits from human H3N2 influenza virus is a critical determinant of the replication of reassortant ribonucleoprotein complexes. *J. Gen. Virol.* 94, 2406–2416
- 90 Essere, B. *et al.* (2013) Critical role of segment-specific packaging signals in genetic reassortment of influenza A viruses. *Proc. Natl. Acad. Sci. U.S.A.* 110, E3840–E3848
- 91 Jackson, S. *et al.* (2009) Reassortment between avian H5N1 and human H3N2 influenza viruses in ferrets: a public health risk assessment. *J. Virol.* 83, 8131–8140
- 92 Marshall, N. *et al.* (2013) Influenza virus reassortment occurs with high frequency in the absence of segment mismatch. *PLoS Pathog.* 9, e1003421

## BIBLIOGRAPHIE

1. **Shaw ML, Palese P** (2013) *Orthomyxoviridae*. *Fields Virology*, eds Knipe DM, Howley PM (Wolters Kluwer Health/Lippincott Williams & Wilkins, Philadelphia, PA), pp 1151–1185. 6th Ed. eds Knipe DM, Howley PM.
2. **Hause BM, Collin EA, Liu R, Huang B, Sheng Z, Lu W, Wang D, Nelson EA, Li F** (2014) Characterization of a Novel Influenza Virus in Cattle and Swine: Proposal for a New Genus in the Orthomyxoviridae Family. *mBio* 5(2):e00031-14-e00031-14.
3. **Neumann G, Noda T, Kawaoka Y** (2009) Emergence and pandemic potential of swine-origin H1N1 influenza virus. *Nature* 459:931–9.
4. **Calder LJ, Wasilewski S, Berriman JA, Rosenthal PB** (2010) Structural organization of a filamentous influenza A virus. *Proc Natl Acad Sci* 107(23):10685–10690.
5. **Moules V, Terrier O, Yver M, Riteau B, Moriscot C, Ferraris O, Julien T, Giudice E, Rolland JP, Erny A, Bouscambert-Duchamp M, Frobert E, Rosa-Calatrava M, Pu Lin Y, Hay A, Thomas D, Schoehn G, Lina B** (2011) Importance of viral genomic composition in modulating glycoprotein content on the surface of influenza virus particles. *Virology* 414:51–62.
6. **Getie-Kehtie M, Sultana I, Eichelberger M, Alterman M** (2013) Label-free mass spectrometry-based quantification of hemagglutinin and neuraminidase in influenza virus preparations and vaccines. *Influenza Other Respir Viruses* 7(4):521–530.
7. **Harris A, Cardone G, Winkler DC, Heymann JB, Brecher M, White JM, Steven AC** (2006) Influenza virus pleiomorphy characterized by cryoelectron tomography. *Proc Natl Acad Sci* 103(50):19123–19127.
8. **Nayak DP, Balogun RA, Yamada H, Zhou ZH, Barman S** (2009) Influenza virus morphogenesis and budding. *Virus Res* 143(2):147–161.
9. **Wright PF, Neumann G, Kawaoka Y** (2013) Orthomyxoviruses. *Fields Virology*, eds Knipe DM, Howley PM (Wolters Kluwer Health/Lippincott Williams & Wilkins, Philadelphia, PA), pp 1186–1243. 6th Ed. eds Knipe DM, Howley PM.
10. **Tong S, Li Y, Rivaller P, Conrardy C, Castillo DA, Chen LM, Recuenco S, Ellison JA, Davis CT, York IA, Turmelle AS, Moran D, Rogers S, Shi M, Tao Y, Weil MR, Tang K, Rowe LA, Sammons S, et al.** (2012) A distinct lineage of influenza A virus from bats. *Proc Natl Acad Sci U A* 109:4269–74.
11. **Tong S, Zhu X, Li Y, Shi M, Zhang J, Bourgeois M, Yang H, Chen X, Recuenco S, Gomez J, others** (2013) New world bats harbor diverse influenza A viruses. *PLoS Pathog* 9(10):e1003657.
12. **Yamayoshi S, Watanabe M, Goto H, Kawaoka Y** (2015) Identification of a Novel Viral Protein Expressed from the PB2 Segment of Influenza A Virus. *J Virol* 90(1):444–456.
13. **Chen W, Calvo PA, Malide D, Gibbs J, Schubert U, Bacik I, Basta S, O'Neill R, Schickli J, Palese P, Henklein P, Bennink JR, Yewdell JW** (2001) A novel influenza A virus mitochondrial protein that induces cell death. *Nat Med* 7:1306–12.
14. **Chakrabarti AK, Pasricha G** (2013) An insight into the PB1F2 protein and its multifunctional role in enhancing the pathogenicity of the influenza A viruses. *Virology* 440(2):97–104.
15. **Wise HM, Foeglein A, Sun J, Dalton RM, Patel S, Howard W, Anderson EC, Barclay WS, Digard P** (2009) A complicated message: Identification of a novel PB1-related protein translated from influenza A virus segment 2 mRNA. *J Virol* 83:8021–31.
16. **Wise HM, Barbezange C, Jagger BW, Dalton RM, Gog JR, Curran MD, Taubenberger JK, Anderson EC, Digard P** (2011) Overlapping signals for translational regulation and packaging of influenza A virus segment 2. *Nucleic Acids Res* 39:7775–90.
17. **Jagger BW, Wise HM, Kash JC, Walters KA, Wills NM, Xiao YL, Dunfee RL, Schwartzman LM, Ozinsky A, Bell GL, Dalton RM, Lo A, Efstathiou S, Atkins JF, Firth AE, Taubenberger JK, Digard P** (2012) An overlapping protein-coding region in influenza A virus segment 3 modulates the host response. *Science* 337:199–204.
18. **Shi M, Jagger BW, Wise HM, Digard P, Holmes EC, Taubenberger JK** (2012) Evolutionary Conservation of the PA-X Open Reading Frame in Segment 3 of Influenza A Virus. *J Virol* 86(22):12411–12413.
19. **Muramoto Y, Noda T, Kawakami E, Akkina R, Kawaoka Y** (2013) Identification of novel influenza A virus proteins translated from PA mRNA. *J Virol* 87:2455–62.

20. **Wise HM, Hutchinson EC, Jagger BW, Stuart AD, Kang ZH, Robb N, Schwartzman LM, Kash JC, Fodor E, Firth AE, Gog JR, Taubenberger JK, Digard P** (2012) Identification of a novel splice variant form of the influenza A virus M2 ion channel with an antigenically distinct ectodomain. *PLoS Pathog* 8:e1002998.
21. **Selman M, Dankar SK, Forbes NE, Jia J-J, Brown EG** (2012) Adaptive mutation in influenza A virus non-structural gene is linked to host switching and induces a novel protein by alternative splicing. *Emerg Microbes Infect* 1.
22. **Desselberger U, Racaniello VR, Zazra JJ, 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(3):315–328.
23. **Bae SH, Cheong HK, Lee JH, Cheong C, Kainosho M, Choi BS** (2001) Structural features of an influenza virus promoter and their implications for viral RNA synthesis. *Proc Natl Acad Sci U S A* 98:10602–7.
24. **Baudin F, Bach C, Cusack S, Ruigrok RW** (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.
25. **Cheong HK, Cheong C, Lee YS, Seong BL, Choi BS** (1999) Structure of influenza virus panhandle RNA studied by NMR spectroscopy and molecular modeling. *Nucleic Acids Res* 27(5):1392–1397.
26. **Tomescu AI, Robb NC, Hengrung N, Fodor E, Kapanidis AN** (2014) Single-molecule FRET reveals a corkscrew RNA structure for the polymerase-bound influenza virus promoter. *Proc Natl Acad Sci U S A* 111:E3335–42.
27. **Arranz R, Coloma R, Chichon FJ, Conesa JJ, Carrascosa JL, Valpuesta JM, Ortin J, Martin-Benito J** (2012) The structure of native influenza virion ribonucleoproteins. *Science* 338:1634–7.
28. **Coloma R, Valpuesta JM, Arranz R, Carrascosa JL, Ortin J, Martin-Benito J** (2009) The structure of a biologically active influenza virus ribonucleoprotein complex. *PLoS Pathog* 5:e1000491.
29. **Elton D, Digard P, Tiley L, Ortin J** (2006) Structure and function of the influenza virus RNP. *Current Topics in Influenza Virology*, ed Kawaoka Y (Horizon Scientific Press, Norfolk), pp 1–36. ed Kawaoka Y.
30. **Moeller A, Kirchdoerfer RN, Potter CS, Carragher B, Wilson IA** (2012) Organization of the influenza virus replication machinery. *Science* 338:1631–4.
31. **Hsu MT, Parvin JD, 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. *Proc Natl Acad Sci U S A* 84:8140–4.
32. **Flick R, Hobom G** (1999) Interaction of influenza virus polymerase with viral RNA in the “corkscrew” conformation. *J Gen Virol* 80(10):2565–2572.
33. **Gerber M, Moulès V, Gavazzi C, Marquet R, Isel C** (2014) Empaquetage sélectif du génome des influenza virus A. *Virologie* 18(5):278–296.
34. **Vasin AV, Temkina OA, Egorov VV, Klotchenko SA, Plotnikova MA, Kiselev OI** (2014) Molecular mechanisms enhancing the proteome of influenza A viruses: An overview of recently discovered proteins. *Virus Res* 185:53–63.
35. **Matsuoka Y, Matsumae H, Katoh M, Einfeld AJ, Neumann G, Hase T, Ghosh S, Shoemaker JE, Lopes TJ, Watanabe T, others** (2013) A comprehensive map of the influenza A virus replication cycle. *BMC Syst Biol* 7(1):1.
36. **Ng WC, Liong S, Tate MD, Irimura T, Denda-Nagai K, Brooks AG, Londrigan SL, Reading PC** (2014) The Macrophage Galactose-Type Lectin Can Function as an Attachment and Entry Receptor for Influenza Virus. *J Virol* 88(3):1659–1672.
37. **Zhu X, Yu W, McBride R, Li Y, Chen L-M, Donis RO, Tong S, Paulson JC, Wilson IA** (2013) Hemagglutinin homologue from H17N10 bat influenza virus exhibits divergent receptor-binding and pH-dependent fusion activities. *Proc Natl Acad Sci U S A* 110(4):1458–1463.
38. **Maruyama J, Nao N, Miyamoto H, Maeda K, Ogawa H, Yoshida R, Igarashi M, Takada A** (2016) Characterization of the glycoproteins of bat-derived influenza viruses. *Virology* 488:43–50.
39. **Matrosovich MN, Matrosovich TY, Gray T, Roberts NA, Klenk H-D** (2004) Human and avian influenza viruses target different cell types in cultures of human airway epithelium. *Proc Natl Acad Sci U S A* 101(13):4620–4624.
40. **Shinya K, Kawaoka Y** (2006) [Influenza virus receptors in the human airway]. *Nature* 56(1):85–89.
41. **Ito T, Couceiro JNS, Kelm S, Baum LG, Krauss S, Castrucci MR, Donatelli I, Kida H, Paulson JC, Webster RG, others** (1998) Molecular basis for the generation in pigs of influenza A viruses with pandemic potential. *J Virol* 72(9):7367–7373.

42. **Amorij J-P, Huckriede A, Wilschut J, Frijlink HW, Hinrichs WLJ** (2008) Development of Stable Influenza Vaccine Powder Formulations: Challenges and Possibilities. *Pharm Res* 25(6):1256–1273.
43. **Matlin KS, Reggio H, Helenius A, Simons K** (1981) Infectious entry pathway of influenza virus in a canine kidney cell line. *J Cell Biol* 91(3):601–613.
44. **Nunes-Correia I, Eulálio A, Nir S, Pedroso de Lima MC** (2004) Caveolae as an additional route for influenza virus endocytosis in MDCK cells. *Cell Mol Biol Lett* 9(1):47–60.
45. **Sieczkarski SB, Whittaker GR** (2002) Influenza Virus Can Enter and Infect Cells in the Absence of Clathrin-Mediated Endocytosis. *J Virol* 76(20):10455–10464.
46. **Rust MJ, Lakadamyali M, Zhang F, Zhuang X** (2004) Assembly of endocytic machinery around individual influenza viruses during viral entry. *Nat Struct Mol Biol* 11(6):567–573.
47. **Rossman JS, Leser GP, Lamb RA** (2012) Filamentous Influenza Virus Enters Cells via Macropinocytosis. *J Virol* 86(20):10950–10960.
48. **de Vries E, Tscherné DM, Wienholts MJ, Cobos-Jiménez V, Scholte F, García-Sastre A, Rottier PJM, de Haan CAM** (2011) Dissection of the Influenza A Virus Endocytic Routes Reveals Macropinocytosis as an Alternative Entry Pathway. *PLoS Pathog* 7(3):e1001329. ed Pekosz A.
49. **Cady SD, Luo W, Hu F, Hong M** (2009) Structure and Function of the Influenza A M2 Proton Channel. *Biochemistry (Mosc)* 48(31):7356–7364.
50. **Harrison SC** (2015) Viral membrane fusion. *Virology* 479–480:498–507.
51. **Martin K, Helenius A** (1991) Transport of incoming influenza virus nucleocapsids into the nucleus. *J Virol* 65(1):232–244.
52. **Chou Y, Heaton NS, Gao Q, Palese P, Singer R, Lionnet T** (2013) Colocalization of Different Influenza Viral RNA Segments in the Cytoplasm before Viral Budding as Shown by Single-molecule Sensitivity FISH Analysis. *PLoS Pathog* 9(5):e1003358. ed Whelan SPJ.
53. **Nevalainen M, Nissinen M, Kaakinen M, Metsikkö K** (2010) Influenza virus infection in multinucleated skeletal myofibers. *Exp Cell Res* 316(11):1784–1794.
54. **Boulo S, Akarsu H, Ruigrok RWH, Baudin F** (2007) Nuclear traffic of influenza virus proteins and ribonucleoprotein complexes. *Virus Res* 124(1–2):12–21.
55. **Deng T, Sharps J, Fodor E, Brownlee GG** (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* 79:8669–74.
56. **Huet S, Avilov SV, Ferbitz L, Daigle N, Cusack S, Ellenberg J** (2010) Nuclear Import and Assembly of Influenza A Virus RNA Polymerase Studied in Live Cells by Fluorescence Cross-Correlation Spectroscopy. *J Virol* 84(3):1254–1264.
57. **MacDonald LA, Aggarwal S, Bussey KA, Desmet EA, Kim B, Takimoto T** (2012) Molecular interactions and trafficking of influenza A virus polymerase proteins analyzed by specific monoclonal antibodies. *Virology* 426(1):51–59.
58. **Neumann G, Castrucci MR, Kawaoka Y** (1997) Nuclear import and export of influenza virus nucleoprotein. *J Virol* 71(12):9690–9700.
59. **Wang P, Palese P, O'Neill RE** (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(3):1850–1856.
60. **Weber F, Kochs G, Gruber S, Haller O** (1998) A classical bipartite nuclear localization signal on Thogoto and influenza A virus nucleoproteins. *Virology* 250(1):9–18.
61. **Bullido R, Gómez-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(1):135–142.
62. **Babcock HP, Chen C, Zhuang X** (2004) Using Single-Particle Tracking to Study Nuclear Trafficking of Viral Genes. *Biophys J* 87(4):2749–2758.
63. **Ng AK, Zhang H, Tan K, Li Z, Liu JH, Chan PK, Li SM, Chan WY, Au SW, Joachimiak A, Walz T, Wang JH, Shaw PC** (2008) Structure of the influenza virus A H5N1 nucleoprotein: implications for RNA binding, oligomerization, and vaccine design. *Faseb J* 22:3638–47.
64. **Ruigrok RW, Baudin F** (1995) Structure of influenza virus ribonucleoprotein particles. II. Purified RNA-free influenza virus ribonucleoprotein forms structures that are indistinguishable from the intact influenza virus ribonucleoprotein particles. *J Gen Virol* 76 ( Pt 4):1009–14.
65. **Ye Q, Krug RM, Tao YJ** (2006) The mechanism by which influenza A virus nucleoprotein forms oligomers and binds RNA. *Nature* 444(7122):1078–1082.

66. **Compans RW, Content J, Duesberg PH** (1972) Structure of the Ribonucleoprotein of Influenza Virus. *J Virol* 10(4):795–800.
67. **Hutchinson EC, Charles PD, Hester SS, Thomas B, Trudgian D, Martínez-Alonso M, Fodor E** (2014) Conserved and host-specific features of influenza virion architecture. *Nat Commun* 5:4816.
68. **Martín-Benito J, Area E, Ortega J, Llorca O, Valpuesta JM, Carrascosa JL, Ortín J** (2001) Three-dimensional reconstruction of a recombinant influenza virus ribonucleoprotein particle. *EMBO Rep* 2(4):313–317.
69. **Ortega J, Martín-Benito J, Zurcher T, Valpuesta JM, Carrascosa JL, Ortín J** (2000) Ultrastructural and functional analyses of recombinant influenza virus ribonucleoproteins suggest dimerization of nucleoprotein during virus amplification. *J Virol* 74:156–63.
70. **Klumpp K, Ruigrok RW, Baudin F** (1997) Roles of the influenza virus polymerase and nucleoprotein in forming a functional RNP structure. *Embo J* 16:1248–57.
71. **Kingsbury DW, Jones IM, Murti KG** (1987) Assembly of influenza ribonucleoprotein in vitro using recombinant nucleoprotein. *Virology* 156(2):396–403.
72. **Scholtissek C, Becht H** (1971) Binding of ribonucleic acids to the RNP-antigen protein of influenza viruses. *J Gen Virol* 10(1):11–16.
73. **Yamanaka K, Ishihama A, Nagata K** (1990) Reconstitution of influenza virus RNA-nucleoprotein complexes structurally resembling native viral ribonucleoprotein cores. *J Biol Chem* 265:11151–5.
74. **Woo J, Yu KL, Lee SH, You JC** (2015) Development of a functional cell-based assay that probes the specific interaction between influenza A virus NP and its packaging signal sequence RNA. *Biochem Biophys Res Commun* 457(2):227–233.
75. **Chan WH, Ng AK, Robb NC, Lam MK, Chan PK, Au SW, Wang JH, Fodor E, Shaw PC** (2010) Functional analysis of the influenza virus H5N1 nucleoprotein tail loop reveals amino acids that are crucial for oligomerization and ribonucleoprotein activities. *J Virol* 84:7337–45.
76. **Chenavas S, Estrozi LF, Slama-Schwok A, Delmas B, Primo CD, Baudin F, Li X, Crépin T, Ruigrok RWH** (2013) Monomeric Nucleoprotein of Influenza A Virus. *PLoS Pathog* 9(3):e1003275.
77. **Hutchinson EC, Denham EM, Thomas B, Trudgian DC, Hester SS, Ridlova G, York A, Turrell L, Fodor E** (2012) Mapping the Phosphoproteome of Influenza A and B Viruses by Mass Spectrometry. *PLoS Pathog* 8(11):e1002993. ed Gack MU.
78. **Turrell L, Hutchinson EC, Vreede FT, Fodor E** (2015) Regulation of Influenza A Virus Nucleoprotein Oligomerization by Phosphorylation. *J Virol* 89(2):1452–1455.
79. **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(6):3799–3806.
80. **Kobayashi M, Toyoda T, Adyshev DM, Azuma Y, Ishihama A** (1994) Molecular dissection of influenza virus nucleoprotein: deletion mapping of the RNA binding domain. *J Virol* 68(12):8433–8436.
81. **Elton D, Medcalf L, Bishop K, Harrison D, Digard P** (1999) Identification of amino acid residues of influenza virus nucleoprotein essential for RNA binding. *J Virol* 73(9):7357–7367.
82. **Medcalf L, Poole E, Elton D, Digard P** (1999) Temperature-sensitive lesions in two influenza A viruses defective for replicative transcription disrupt RNA binding by the nucleoprotein. *J Virol* 73:7349–56.
83. **Biswas SK, Boutz PL, Nayak DP** (1998) Influenza virus nucleoprotein interacts with influenza virus polymerase proteins. *J Virol* 72:5493–501.
84. **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(1):120–133.
85. **Marklund JK, Ye Q, Dong J, Tao YJ, Krug RM** (2012) Sequence in the Influenza A Virus Nucleoprotein Required for Viral Polymerase Binding and RNA Synthesis. *J Virol* 86(13):7292–7297.
86. **Vidic J, Noiray M, Bagchi A, Slama-Schwok A** (2016) Identification of a Novel Complex between the Nucleoprotein and PA(1–27) of Influenza A Virus Polymerase. *Biochemistry (Mosc)*.
87. **Dias A, Bouvier D, Crépin T, McCarthy AA, Hart DJ, Baudin F, Cusack S, Ruigrok RWH** (2009) The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit. *Nature* 458(7240):914–918.
88. **Guilligay D, Tarendeau F, Resa-Infante P, Coloma R, Crepin T, Sehr P, Lewis J, Ruigrok RWH, Ortín J, Hart DJ, Cusack S** (2008) The structural basis for cap binding by influenza virus polymerase subunit PB2. *Nat Struct Mol Biol* 15(5):500–506.

89. **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* 454:1123–6.
90. **Sugiyama K, Obayashi E, Kawaguchi A, Suzuki Y, Tame JR, Nagata K, Park SY** (2009) Structural insight into the essential PB1-PB2 subunit contact of the influenza virus RNA polymerase. *Embo J* 28:1803–11.
91. **Tarendeau F, Crepin T, Guilligay D, Ruigrok RW, Cusack S, Hart DJ** (2008) Host determinant residue lysine 627 lies on the surface of a discrete, folded domain of influenza virus polymerase PB2 subunit. *PLoS Pathog* 4:e1000136.
92. **Pflug A, Guilligay D, Reich S, Cusack S** (2014) Structure of influenza A polymerase bound to the viral RNA promoter. *Nature* 516(7531):355–360.
93. **Reich S, Guilligay D, Pflug A, Malet H, Berger I, Crépin T, Hart D, Lunardi T, Nanao M, Ruigrok RWH, Cusack S** (2014) Structural insight into cap-snatching and RNA synthesis by influenza polymerase. *Nature* 516(7531):361–366.
94. **Hengrung N, El Omari K, Martin IS, Vreede FT, Cusack S, Rambo RP, Vonnrhein C, Bricogne G, Stuart DI, Grimes JM, Fodor E** (2015) Crystal structure of the RNA-dependent RNA polymerase from influenza C virus. *Nature* 527(7576):114–117.
95. **Chang S, Sun D, Liang H, Wang J, Li J, Guo L, Wang X, Guan C, Boruah BM, Yuan L, Feng F, Yang M, Wang L, Wang Y, Wojdyla J, Li L, Wang J, Wang M, Cheng G, et al.** (2015) Cryo-EM Structure of Influenza Virus RNA Polymerase Complex at 4.3 Å Resolution. *Mol Cell* 57(5):925–935.
96. **Thierry E, Guilligay D, Kosinski J, Bock T, Gaudon S, Round A, Pflug A, Hengrung N, El Omari K, Baudin F, Hart DJ, Beck M, Cusack S** (2016) Influenza Polymerase Can Adopt an Alternative Configuration Involving a Radical Repacking of PB2 Domains. *Mol Cell* 61(1):125–137.
97. **Heggeness MH, Smith PR, Ulmanen I, Krug RM, Choppin PW** (1982) Studies on the helical nucleocapsid of influenza virus. *Virology* 118(2):466–470.
98. **Jennings PA, Finch JT, Winter G, Robertson JS** (1983) Does the higher order structure of the influenza virus ribonucleoprotein guide sequence rearrangements in influenza viral RNA? *Cell* 34:619–27.
99. **Pons MW, Schulze IT, Hirst GK** (1969) Isolation and characterization of the ribonucleoprotein of influenza virus. *Virology* 39(2):250–259.
100. **Martín-Benito J, Ortín J** (2013) Chapter Four - Influenza Virus Transcription and Replication. *Advances in Virus Research*, ed Murphy KM and FA (Academic Press), pp 113–137. ed Murphy KM and FA.
101. **Zheng W, Olson J, Vakharia V, Tao YJ** (2013) The crystal structure and RNA-binding of an orthomyxovirus nucleoprotein. *PLoS Pathog* 9:e1003624.
102. **York A, Hengrung N, Vreede FT, Huiskonen JT, Fodor E** (2013) Isolation and characterization of the positive-sense replicative intermediate of a negative-strand RNA virus. *Proc Natl Acad Sci U S A* 110:E4238–45.
103. **Gultyaev AP, Fouchier RA, Olsthoorn RC** (2010) Influenza virus RNA structure: unique and common features. *Int Rev Immunol* 29:533–56.
104. **Gultyaev AP, Tsyganov-Bodounov A, Spronken MI, van der Kooij S, Fouchier RA, Olsthoorn RC** (2014) RNA structural constraints in the evolution of the influenza A virus genome NP segment. *RNA Biol* 11(7):942–952.
105. **Kobayashi Y, Dadonaite B, Doremalen N van, Suzuki Y, Barclay WS, Pybus OG** (2016) Computational and molecular analysis of conserved influenza A virus RNA secondary structures involved in infectious virion production. *RNA Biol* 0(ja):00–00.
106. **Anhlan D, Hrinčius E-R, Scholtissek C, Ludwig S** (2012) Introduction of silent mutations into the NP gene of influenza A viruses as a possible strategy for the creation of a live attenuated vaccine. *Vaccine* 30(30):4480–4489.
107. **Hutchinson EC, Wise HM, Kudryavtseva K, Curran MD, Digard P** (2009) Characterisation of influenza A viruses with mutations in segment 5 packaging signals. *Vaccine* 27:6270–5.
108. **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. *J Virol* 81(1):30–41.
109. **Lenartowicz E, Keszy J, Ruskowska A, Soszynska-Jozwiak M, Michalak P, Moss WN, Turner DH, Kierzek R, Kierzek E** (2016) Self-Folding of Naked Segment 8 Genomic RNA of Influenza A Virus. *PLOS ONE* 11(2):e0148281.

110. **Reguera J, Gerlach P, Cusack S** (2016) Towards a structural understanding of RNA synthesis by negative strand RNA viral polymerases. *Curr Opin Struct Biol* 36:75–84.
111. **Engelhardt OG, Smith M, Fodor E** (2005) Association of the Influenza A Virus RNA-Dependent RNA Polymerase with Cellular RNA Polymerase II. *J Virol* 79(9):5812–5818.
112. **Martínez-Alonso M, Hengrung N, Fodor E** (2016) RNA-Free and Ribonucleoprotein-Associated Influenza Virus Polymerases Directly Bind the Serine-5-Phosphorylated Carboxyl-Terminal Domain of Host RNA Polymerase II. *J Virol* 90(13):6014–6021.
113. **Gu W, Gallagher GR, Dai W, Liu P, Li R, Trombly MI, Gammon DB, Mello CC, Wang JP, Finberg RW** (2015) Influenza A virus preferentially snatches noncoding RNA caps. *RNA* 21(12):2067–2075.
114. **Yuan P, Bartlam M, Lou Z, Chen S, Zhou J, He X, Lv Z, Ge R, Li X, Deng T, Fodor E, Rao Z, Liu Y** (2009) Crystal structure of an avian influenza polymerase PAN reveals an endonuclease active site. *Nature* 458(7240):909–913.
115. **Biswas SK, Nayak DP** (1994) Mutational analysis of the conserved motifs of influenza A virus polymerase basic protein 1. *J Virol* 68:1819–26.
116. **Poon LL, Pritlove DC, Fodor E, Brownlee GG** (1999) Direct evidence that the poly (A) tail of influenza A virus mRNA is synthesized by reiterative copying of a U track in the virion RNA template. *J Virol* 73(4):3473–3476.
117. **Robertson JS, Schubert M, Lazzarini RA** (1981) Polyadenylation sites for influenza virus mRNA. *J Virol* 38(1):157–163.
118. **Krug RM, Morgan MA, Shatkin AJ** (1976) Influenza viral mRNA contains internal N6-methyladenosine and 5'-terminal 7-methylguanosine in cap structures. *J Virol* 20:45–53.
119. **Narayan P, Ayers DF, Rottman FM, Maroney PA, Nilsen TW** (1987) Unequal distribution of N6-methyladenosine in influenza virus mRNAs. *Mol Cell Biol* 7:1572–5.
120. **Dubois J, Terrier O, Rosa-Calatrava M** (2014) Influenza Viruses and mRNA Splicing: Doing More with Less. *mBio* 5(3):e00070-14.
121. **York A, Fodor E** (2013) Biogenesis, assembly, and export of viral messenger ribonucleoproteins in the influenza A virus infected cell. *RNA Biol* 10(8):1274–1282.
122. **Larsen S, Bui S, Perez V, Mohammad A, Medina-Ramirez H, Newcomb LL** (2014) Influenza polymerase encoding mRNAs utilize atypical mRNA nuclear export. *Virology* 461(1):1.
123. **Chenavas S, Monod A, Ruigrok RWH, Crepin T** (2013) Le complexe de transcription et de réplication des virus influenza, une cible pour de nouvelles molécules antivirales. *Virologie* 17(1):6–16.
124. **Resa-Infante P, Jorba N, Coloma R, Ortin J** (2011) The influenza virus RNA synthesis machine: Advances in its structure and function. *RNA Biol* 8(2):207–215.
125. **Fodor E** (2013) The RNA polymerase of influenza A virus: mechanisms of viral transcription and replication. *Acta Virol* 57(2):113–122.
126. **Vreede FT, Jung TE, Brownlee GG** (2004) Model suggesting that replication of influenza virus is regulated by stabilization of replicative intermediates. *J Virol* 78:9568–72.
127. **Shapiro GI, Gurney T, Krug RM** (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(3):764–773.
128. **Vreede FT, Chan AY, Sharps J, Fodor E** (2010) Mechanisms and functional implications of the degradation of host RNA polymerase II in influenza virus infected cells. *Virology* 396(1):125–134.
129. **Beaton AR, Krug RM** (1986) Transcription antitermination during influenza viral template RNA synthesis requires the nucleocapsid protein and the absence of a 5' capped end. *Proc Natl Acad Sci U S A* 83:6282–6.
130. **Portela A, Digard P** (2002) The influenza virus nucleoprotein: a multifunctional RNA-binding protein pivotal to virus replication. *J Gen Virol* 83:723–34.
131. **Perez JT, Varble A, Sachidanandam R, Zlatev I, Manoharan M, Garcia-Sastre A, tenOever BR** (2010) Influenza A virus-generated small RNAs regulate the switch from transcription to replication. *Proc Natl Acad Sci U S A* 107:11525–30.
132. **Jorba N, Coloma R, Ortin J** (2009) Genetic trans-Complementation Establishes a New Model for Influenza Virus RNA Transcription and Replication. *PLoS Pathog* 5(5):e1000462. ed Whelan SPJ.
133. **Eisfeld AJ, Neumann G, Kawaoka Y** (2014) At the centre: influenza A virus ribonucleoproteins. *Nat Rev Microbiol* 13(1):28–41.
134. **Li J, Yu M, Zheng W, Liu W** (2015) Nucleocytoplasmic Shuttling of Influenza A Virus Proteins. *Viruses* 7(5):2668–2682.

135. **Gao S, Wang S, Cao S, Sun L, Li J, Bi Y, Gao GF, Liu W** (2014) Characteristics of Nucleocytoplasmic Transport of H1N1 Influenza A Virus Nuclear Export Protein. *J Virol* 88(13):7455–7463.
136. **Chase GP, Rameix-Welti M-A, Zvirbliene A, Zvirblis G, Götz V, Wolff T, Naffakh N, Schwemmle M** (2011) Influenza virus ribonucleoprotein complexes gain preferential access to cellular export machinery through chromatin targeting. *PLoS Pathog* 7(9):e1002187.
137. **Bui M, Wills EG, Helenius A, Whittaker GR** (2000) Role of the influenza virus M1 protein in nuclear export of viral ribonucleoproteins. *J Virol* 74:1781–6.
138. **Zhirnov OP, Klenk HD** (1997) Histones as a target for influenza virus matrix protein M1. *Virology* 235(2):302–310.
139. **Noton SL, Medcalf E, Fisher D, Mullin AE, Elton D, Digard P** (2007) Identification of the domains of the influenza A virus M1 matrix protein required for NP binding, oligomerization and incorporation into virions. *J Gen Virol* 88:2280–90.
140. **Akarsu H, Burmeister WP, Petosa C, Petit I, Müller CW, Ruigrok RW, Baudin F** (2003) Crystal structure of the M1 protein-binding domain of the influenza A virus nuclear export protein (NEP/NS2). *EMBO J* 22(18):4646–4655.
141. **Neumann G, Hughes MT, Kawaoka Y** (2000) Influenza A virus NS2 protein mediates vRNP nuclear export through NES-independent interaction with hCRM1. *EMBO J* 19(24):6751–6758.
142. **Wu C-Y, Jeng K-S, Lai MM-C** (2011) The SUMOylation of matrix protein M1 modulates the assembly and morphogenesis of influenza A virus. *J Virol* 85(13):6618–6628.
143. **Zheng W, Li J, Wang S, Cao S, Jiang J, Chen C, Ding C, Qin C, Ye X, Gao GF, Liu W** (2015) Phosphorylation Controls the Nuclear-Cytoplasmic Shuttling of Influenza A Virus Nucleoprotein. *J Virol* 89(11):5822–5834.
144. **Reuther P, Giese S, Götz V, Riegger D, Schwemmle M** (2014) Phosphorylation of highly conserved serine residues in the influenza A virus nuclear export protein NEP plays a minor role in viral growth in human cells and mice. *J Virol* 88(13):7668–7673.
145. **Mänz B, Brunotte L, Reuther P, Schwemmle M** (2012) Adaptive mutations in NEP compensate for defective H5N1 RNA replication in cultured human cells. *Nat Commun* 3:802.
146. **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(1):408–419.
147. **Mühlbauer D, Dzieciolowski J, Hardt M, Hocke A, Schierhorn KL, Mostafa A, Müller C, Wisskirchen C, Herold S, Wolff T, Ziebuhr J, Pleschka S** (2015) Influenza Virus-Induced Caspase-Dependent Enlargement of Nuclear Pores Promotes Nuclear Export of Viral Ribonucleoprotein Complexes. *J Virol* 89(11):6009–6021.
148. **Lakdawala SS, Wu Y, Wawrzusin P, Kabat J, Broadbent AJ, Lamirande EW, Fodor E, Altan-Bonnet N, Shroff H, Subbarao K** (2014) Influenza A Virus Assembly Intermediates Fuse in the Cytoplasm. *PLoS Pathog* 10(3):e1003971. ed Heise MT.
149. **Avilov S, Magnus J, Cusack S, Naffakh N** (2016) Time-Resolved Visualisation of Nearly-Native Influenza A Virus Progeny Ribonucleoproteins and Their Individual Components in Live Infected Cells. *PLOS ONE* 11(3):e0149986.
150. **Tchatalbachev S, Flick R, Hobom G** (2001) The packaging signal of influenza viral RNA molecules. *RNA* 7(7):979–989.
151. **Amorim MJ, Bruce EA, Read EKC, Foeglein A, Mahen R, Stuart AD, Digard P** (2011) A Rab11- and Microtubule-Dependent Mechanism for Cytoplasmic Transport of Influenza A Virus Viral RNA. *J Virol* 85(9):4143–4156.
152. **Avilov SV, Moisy D, Munier S, Schraidt O, Naffakh N, Cusack S** (2012) Replication-Competent Influenza A Virus That Encodes a Split-Green Fluorescent Protein-Tagged PB2 Polymerase Subunit Allows Live-Cell Imaging of the Virus Life Cycle. *J Virol* 86(3):1433–1448.
153. **Momose F, Kikuchi Y, Komase K, Morikawa Y** (2007) Visualization of microtubule-mediated transport of influenza viral progeny ribonucleoprotein. *Microbes Infect* 9(12–13):1422–1433.
154. **Bruce EA, Stuart A, McCaffrey MW, Digard P** (2012) Role of the Rab11 pathway in negative-strand virus assembly. *Biochem Soc Trans* 40(6):1409–1415.
155. **Avilov SV, Moisy D, Naffakh N, Cusack S** (2012) Influenza A virus progeny vRNP trafficking in live infected cells studied with the virus-encoded fluorescently tagged PB2 protein. *Vaccine* 30(51):7411–7417.

156. **Eisfeld AJ, Kawakami E, Watanabe T, Neumann G, Kawaoka Y** (2011) RAB11A Is Essential for Transport of the Influenza Virus Genome to the Plasma Membrane. *J Virol* 85(13):6117–6126.
157. **Momose F, Sekimoto T, Ohkura T, Jo S, Kawaguchi A, Nagata K, Morikawa Y** (2011) Apical Transport of Influenza A Virus Ribonucleoprotein Requires Rab11-positive Recycling Endosome. *PLoS ONE* 6(6):e21123. ed Darlix J-LEPH.
158. **Giese S, Bolte H, Schwemmler M** (2016) The Feat of Packaging Eight Unique Genome Segments. *Viruses* 8(6):165.
159. **Vale-Costa S, Alenquer M, Sousa AL, Kellen B, Ramalho J, Tranfield EM, Amorim MJ** (2016) Influenza A virus ribonucleoproteins modulate host recycling by competing with Rab11 effectors. *J Cell Sci* 129(8):1697–1710.
160. **Rossmann JS, Lamb RA** (2011) Influenza virus assembly and budding. *Virology* 411(2):229–236.
161. **Chen BJ, Leser GP, Morita E, Lamb RA** (2007) Influenza Virus Hemagglutinin and Neuraminidase, but Not the Matrix Protein, Are Required for Assembly and Budding of Plasmid-Derived Virus-Like Particles. *J Virol* 81(13):7111–7123.
162. **Lai JCC, Chan WWL, Kien F, Nicholls JM, Peiris JSM, Garcia J-M** (2010) Formation of virus-like particles from human cell lines exclusively expressing influenza neuraminidase. *J Gen Virol* 91(9):2322–2330.
163. **Gómez-Puertas P, Albo C, Pérez-Pastrana E, Vivo A, Portela A** (2000) Influenza virus matrix protein is the major driving force in virus budding. *J Virol* 74(24):11538–11547.
164. **Latham T, Galarza JM** (2001) Formation of wild-type and chimeric influenza virus-like particles following simultaneous expression of only four structural proteins. *J Virol* 75:6154–65.
165. **Wang D, Harmon A, Jin J, Francis DH, Christopher-Hennings J, Nelson E, Montelaro RC, Li F** (2010) The Lack of an Inherent Membrane Targeting Signal Is Responsible for the Failure of the Matrix (M1) Protein of Influenza A Virus To Bud into Virus-Like Particles. *J Virol* 84(9):4673–4681.
166. **Bui M, Whittaker G, Helenius A** (1996) Effect of M1 protein and low pH on nuclear transport of influenza virus ribonucleoproteins. *J Virol* 70(12):8391–8401.
167. **Zhang J, Lamb RA** (1996) Characterization of the membrane association of the influenza virus matrix protein in living cells. *Virology* 225(2):255–266.
168. **Ruigrok R, Baudin F, Petit I, Weissenhorn W** (2001) Role of influenza virus M1 protein in the viral budding process. *Int Congr Ser* (1219):397–404.
169. **Zhang J, Leser GP, Pekosz A, Lamb RA** (2000) The Cytoplasmic Tails of the Influenza Virus Spike Glycoproteins Are Required for Normal Genome Packaging. *Virology* 269(2):325–334.
170. **Chen BJ, Leser GP, Jackson D, Lamb RA** (2008) The influenza virus M2 protein cytoplasmic tail interacts with the M1 protein and influences virus assembly at the site of virus budding. *J Virol* 82:10059–70.
171. **McCown MF, Pekosz A** (2005) The Influenza A Virus M2 Cytoplasmic Tail Is Required for Infectious Virus Production and Efficient Genome Packaging. *J Virol* 79(6):3595–3605.
172. **Kakisaka M, Yamada K, Yamaji-Hasegawa A, Kobayashi T, Aida Y** (2016) Intrinsically disordered region of influenza A NP regulates viral genome packaging via interactions with viral RNA and host PI(4,5)P2. *Virology* 496:116–126.
173. **Rossmann JS, Jing X, Leser GP, Lamb RA** (2010) Influenza Virus M2 Protein Mediates ESCRT-Independent Membrane Scission. *Cell* 142(6):902–913.
174. **Yondola M, Carter C** (2011) Un-“ESCRT”-ed Budding. *Viruses* 3:26–31.
175. **Li Q, Sun X, Li Z, Liu Y, Vavricka CJ, Qi J, Gao GF** (2012) Structural and functional characterization of neuraminidase-like molecule N10 derived from bat influenza A virus. *Proc Natl Acad Sci U S A* 109(46):18897–18902.
176. **Zhu X, Yang H, Guo Z, Yu W, Carney PJ, Li Y, Chen L-M, Paulson JC, Donis RO, Tong S, Stevens J, Wilson IA** (2012) Crystal structures of two subtype N10 neuraminidase-like proteins from bat influenza A viruses reveal a diverged putative active site. *Proc Natl Acad Sci U S A* 109(46):18903–18908.
177. **Yang H, Carney PJ, Mishin VP, Guo Z, Chang JC, Wentworth DE, Gubareva LV, Stevens J** (2016) Molecular Characterizations of Surface Proteins Hemagglutinin and Neuraminidase from Recent H5Nx Avian Influenza Viruses. *J Virol* 90(12):5770–5784. ed Schultz-Cherry S.
178. **Le Briand N, Casalegno J-S, Escuret V, Gaymard A, Lina B, Ottmann M, Frobert E** (2016) La balance HA-NA des virus influenza A (H1N1). *Virologie* 20(1):47–60.
179. **Wagner R, Matrosovich M, Klenk HD** (2002) Functional balance between haemagglutinin and neuraminidase in influenza virus infections. *Rev Med Virol* 12:159–66.

180. **Casalegno J-S, Ferraris O, Escuret V, Bouscambert M, Bergeron C, Linès L, Excoffier T, Valette M, Frobert E, Pillet S, Pozzetto B, Lina B, Ottmann M** (2014) Functional Balance between the Hemagglutinin and Neuraminidase of Influenza A(H1N1)pdm09 HA D222 Variants. *PLOS ONE* 9(8):e104009.
181. **Xu R, Zhu X, McBride R, Nycholat CM, Yu W, Paulson JC, Wilson IA** (2012) Functional Balance of the Hemagglutinin and Neuraminidase Activities Accompanies the Emergence of the 2009 H1N1 Influenza Pandemic. *J Virol* 86(17):9221–9232.
182. **Moules V, Ferraris O, Terrier O, Giudice E, Yver M, Rolland JP, Bouscambert-Duchamp M, Bergeron C, Ottmann M, Fournier E, Traversier A, Boule C, Rivoire A, Lin Y, Hay A, Valette M, Marquet R, Rosa-Calatrava M, Naffakh N, et al.** (2010) In vitro characterization of naturally occurring influenza H3NA-viruses lacking the NA gene segment: Toward a new mechanism of viral resistance? *Virology* 404(2):215–224.
183. **Hanada K, Suzuki Y, Gojobori T** (2004) A large variation in the rates of synonymous substitution for RNA viruses and its relationship to a diversity of viral infection and transmission modes. *Mol Biol Evol* 21(6):1074–1080.
184. **Drake JW, Holland JJ** (1999) Mutation rates among RNA viruses. *Proc Natl Acad Sci U S A* 96(24):13910–13913.
185. **Chambers BS, Parkhouse K, Ross TM, Alby K, Hensley SE** (2015) Identification of Hemagglutinin Residues Responsible for H3N2 Antigenic Drift during the 2014–2015 Influenza Season. *Cell Rep* 12(1):1–6.
186. **Tewawong N, Prachayangprecha S, Vichiwattana P, Korkong S, Klinfueng S, Vongpunsawad S, Thongmee T, Theamboonlers A, Poovorawan Y** (2015) Assessing Antigenic Drift of Seasonal Influenza A(H3N2) and A(H1N1)pdm09 Viruses. *PLOS ONE* 10(10):e0139958.
187. **Taubenberger JK, Kash JC** (2010) Influenza virus evolution, host adaptation, and pandemic formation. *Cell Host Microbe* 7:440–51.
188. **Guan Y, Vijaykrishna D, Bahl J, Zhu H, Wang J, Smith GJD** (2010) The emergence of pandemic influenza viruses. *Protein Cell* 1(1):9–13.
189. **Potter C w.** (2001) A history of influenza. *J Appl Microbiol* 91(4):572–579.
190. **Medina RA, Garcia-Sastre A** (2011) Influenza A viruses: new research developments. *Nat Rev Microbiol* 9:590–603.
191. **Nelli RK, Kuchipudi SV, White GA, Perez BB, Dunham SP, Chang K-C** (2010) Comparative distribution of human and avian type sialic acid influenza receptors in the pig. *BMC Vet Res* 6:4.
192. **Van Poucke SG, Nicholls JM, Nauwynck HJ, Van Reeth K** (2010) Replication of avian, human and swine influenza viruses in porcine respiratory explants and association with sialic acid distribution. *Virology* 7:38.
193. **Guo C-T, Takahashi N, Yagi H, Kato K, Takahashi T, Yi S-Q, Chen Y, Ito T, Otsuki K, Kida H, Kawaoka Y, Hidari KI-PJ, Miyamoto D, Suzuki T, Suzuki Y** (2007) The quail and chicken intestine have sialyl-galactose sugar chains responsible for the binding of influenza A viruses to human type receptors. *Glycobiology* 17(7):713–724.
194. **Chen GW, Chang SC, Mok CK, Lo YL, Kung YN, Huang JH, Shih YH, Wang JY, Chiang C, Chen CJ, Shih SR** (2006) Genomic signatures of human versus avian influenza A viruses. *Emerg Infect Dis* 12:1353–60.
195. **Neumann G, Kawaoka Y** (2015) Transmission of Influenza A Viruses. *Virology* 0:234–246.
196. **Hatta M, Hatta Y, Kim JH, Watanabe S, Shinya K, Nguyen T, Lien PS, Le QM, Kawaoka Y** (2007) Growth of H5N1 influenza A viruses in the upper respiratory tracts of mice. *PLoS Pathog* 3(10):1374–1379.
197. **Cauldwell AV, Long JS, Moncorgé O, Barclay WS** (2014) Viral determinants of influenza A virus host range. *J Gen Virol* 95(6):1193–1210.
198. **Long JS, Giotis ES, Moncorgé O, Frise R, Mistry B, James J, Morisson M, Iqbal M, Vignat A, Skinner MA, Barclay WS** (2016) Species difference in ANP32A underlies influenza A virus polymerase host restriction. *Nature* 529(7584):101–104.
199. **Chen LM, Blixt O, Stevens J, Lipatov AS, Davis CT, Collins BE, Cox NJ, Paulson JC, Donis RO** (2012) In vitro evolution of H5N1 avian influenza virus toward human-type receptor specificity. *Virology* 422:105–13.
200. **Matrosovich MN, Krauss S, Webster RG** (2001) H9N2 Influenza A Viruses from Poultry in Asia Have Human Virus-like Receptor Specificity. *Virology* 281(2):156–162.
201. **Reid AH, Taubenberger JK, Fanning TG** (2004) Evidence of an absence: the genetic origins of the 1918 pandemic influenza virus. *Nat Rev Microbiol* 2(11):909–914.

202. **Taubenberger JK, Baltimore D, Doherty PC, Markel H, Morens DM, Webster RG, Wilson IA** (2012) Reconstruction of the 1918 influenza virus: unexpected rewards from the past. *mBio* 3(5).
203. **Smith GJD, Bahl J, Vijaykrishna D, Zhang J, Poon LLM, Chen H, Webster RG, Peiris JSM, Guan Y** (2009) Dating the emergence of pandemic influenza viruses. *Proc Natl Acad Sci U S A* 106(28):11709–11712.
204. **Worobey M, Han G-Z, Rambaut A** (2014) Genesis and pathogenesis of the 1918 pandemic H1N1 influenza A virus. *Proc Natl Acad Sci U S A* 111(22):8107–8112.
205. **Claas EC, Osterhaus AD, van Beek R, De Jong JC, Rimmelzwaan GF, Senne DA, Krauss S, Shortridge KF, Webster RG** (1998) Human influenza A H5N1 virus related to a highly pathogenic avian influenza virus. *The Lancet* 351(9101):472–477.
206. **Subbarao K, Klimov A, Katz J, Regnery H, Lim W, Hall H, Perdue M, Swayne D, Bender C, Huang J, Hemphill M, Rowe T, Shaw M, Xu X, Fukuda K, Cox N** (1998) Characterization of an avian influenza A (H5N1) virus isolated from a child with a fatal respiratory illness. *Science* 279(5349):393–396.
207. **Arzey GG, Kirkland PD, Arzey KE, Frost M, Maywood P, Conaty S, Hurt AC, Deng Y-M, Iannello P, Barr I, Dwyer DE, Ratnamohan M, McPhie K, Selleck P** (2012) Influenza Virus A (H10N7) in Chickens and Poultry Abattoir Workers, Australia. *Emerg Infect Dis* 18(5):814–816.
208. **Chen H, Yuan H, Gao R, Zhang J, Wang D, Xiong Y, Fan G, Yang F, Li X, Zhou J, Zou S, Yang L, Chen T, Dong L, Bo H, Zhao X, Zhang Y, Lan Y, Bai T, et al.** (2014) Clinical and epidemiological characteristics of a fatal case of avian influenza A H10N8 virus infection: a descriptive study. *The Lancet* 383(9918):714–721.
209. **Fouchier RA, Schneeberger PM, Rozendaal FW, Broekman JM, Kemink SA, Munster V, Kuiken T, Rimmelzwaan GF, Schutten M, van Doornum GJ, others** (2004) Avian influenza A virus (H7N7) associated with human conjunctivitis and a fatal case of acute respiratory distress syndrome. *Proc Natl Acad Sci U S America* 101(5):1356–1361.
210. **Koopmans M, Wilbrink B, Conyn M, Natrop G, van der Nat H, Vennema H, Meijer A, van Steenbergen J, Fouchier R, Osterhaus A, Bosman A** (2004) Transmission of H7N7 avian influenza A virus to human beings during a large outbreak in commercial poultry farms in the Netherlands. *The Lancet* 363(9409):587–593.
211. **Lopez-Martinez I, Balish A, Barrera-Badillo G, Jones J, Nuñez-García TE, Jang Y, Aparicio-Antonio R, Azziz-Baumgartner E, Belser JA, Ramirez-Gonzalez JE, Pedersen JC, Ortiz-Alcantara J, Gonzalez-Duran E, Shu B, Emery SL, Poh MK, Reyes-Teran G, Vazquez-Perez JA, Avila-Rios S, et al.** (2013) Highly Pathogenic Avian Influenza A(H7N3) Virus in Poultry Workers, Mexico, 2012. *Emerg Infect Dis* 19(9).
212. **Ostrowsky B, Huang A, Terry W, Anton D, Brunagel B, Traynor L, Abid S, Johnson G, Kacica M, Katz J, Edwards L, Lindstrom S, Klimov A, Uyeki TM** (2012) Low Pathogenic Avian Influenza A (H7N2) Virus Infection in Immunocompromised Adult, New York, USA, 2003. *Emerg Infect Dis* 18(7):1128–1131.
213. **Peiris M, Yuen K, Leung C, Chan K, Ip P, Lai R, Orr W, Shortridge K** (1999) Human infection with influenza H9N2. *The Lancet* 354(9182):916–917.
214. **Yuan J, Zhang L, Kan X, Jiang L, Yang J, Guo Z, Ren Q** (2013) Origin and Molecular Characteristics of a Novel 2013 Avian Influenza A(H6N1) Virus Causing Human Infection in Taiwan. *Clin Infect Dis* 57(9):1367–1368.
215. **Richard M, Fouchier RAM** (2016) Influenza A virus transmission via respiratory aerosols or droplets as it relates to pandemic potential. *FEMS Microbiol Rev* 40(1):68–85.
216. **Herfst S, Schrauwen EJ, Linster M, Chutinimitkul S, de Wit E, Munster VJ, Sorrell EM, Bestebroer TM, Burke DF, Smith DJ, Rimmelzwaan GF, Osterhaus AD, Fouchier RA** (2012) Airborne transmission of influenza A/H5N1 virus between ferrets. *Science* 336:1534–41.
217. **Imai M, Watanabe T, Hatta M, Das SC, Ozawa M, Shinya K, Zhong G, Hanson A, Katsura H, Watanabe S, Li C, Kawakami E, Yamada S, Kiso M, Suzuki Y, Maher EA, Neumann G, Kawaoka Y** (2012) Experimental adaptation of an influenza H5 haemagglutinin (HA) confers respiratory droplet transmission to a reassortant H5 HA/H1N1 virus in ferrets. *Nature* 486(7403):420–428.
218. **Linster M, van Boheemen S, de Graaf M, Schrauwen EJA, Lexmond P, Mänz B, Bestebroer TM, Baumann J, van Riel D, Rimmelzwaan GF, Osterhaus ADME, Matrosovich M, Fouchier RAM, Herfst S** (2014) Identification, Characterization, and Natural Selection of Mutations Driving Airborne Transmission of A/H5N1 virus. *Cell* 157(2):329–339.
219. **Sutton TC, Finch C, Shao H, Angel M, Chen H, Capua I, Cattoli G, Monne I, Perez DR** (2014) Airborne Transmission of Highly Pathogenic H7N1 Influenza Virus in Ferrets. *J Virol* 88(12):6623–6635.
220. **Demicheli V, Di Pietrantonj C, Jefferson T, Rivetti A, Rivetti D** (2007) Vaccines for preventing influenza in healthy adults. *Cochrane Libr.*

221. Vaccines for preventing influenza in healthy adults - Demicheli - 2007 - The Cochrane Library - Wiley Online Library.
222. **Rota PA, Wallis TR, Harmon MW, Rota JS, Kendal AP, Nerome K** (1990) Cocirculation of two distinct evolutionary lineages of influenza type B virus since 1983. *Virology* 175(1):59–68.
223. **Tisa V, Barberis I, Faccio V, Paganino C, Trucchi C, Martini M, Ansaldo F** (2016) Quadrivalent influenza vaccine: a new opportunity to reduce the influenza burden. *J Prev Med Hyg* 57(1):E28–E33.
224. **Uhart M, Bricout H, Clay E, Largeron N** (2016) Public health and economic impact of seasonal influenza vaccination with quadrivalent influenza vaccines compared to trivalent influenza vaccines in Europe. *Hum Vaccines Immunother*:1–10.
225. **Kilbourne ED** (1969) Future influenza vaccines and the use of genetic recombinants. *Bull World Health Organ* 41(3-4-5):643–645.
226. **Kilbourne ED, Schulman JL, Schild GC, Schloer G, Swanson J, Bucher D** (1971) Correlated Studies of a Recombinant Influenza-Virus Vaccine. I. Derivation and Characterization of Virus and Vaccine. *J Infect Dis* 124(5):449–462.
227. **Leibovitz A, Coultrip RL, Kilbourne ED, Legters LJ, Smith CD, Chin J, Schulman JL** (1971) Correlated Studies of a Recombinant Influenza-Virus Vaccine. IV. Protection against Naturally Occurring Influenza in Military Trainees. *J Infect Dis* 124(5):481–487.
228. **Skowronski DM, Janjua NZ, Serres GD, Sabaiduc S, Eshaghi A, Dickinson JA, Fonseca K, Winter A-L, Gubbay JB, Krajden M, Petric M, Charest H, Bastien N, Kwindt TL, Mahmud SM, Caesele PV, Li Y** (2014) Low 2012–13 Influenza Vaccine Effectiveness Associated with Mutation in the Egg-Adapted H3N2 Vaccine Strain Not Antigenic Drift in Circulating Viruses. *PLOS ONE* 9(3):e92153.
229. **Milián E, Kamen AA** (2015) Current and Emerging Cell Culture Manufacturing Technologies for Influenza Vaccines. *BioMed Res Int* 2015:1–11.
230. **Doroshenko A, Halperin SA** (2009) Trivalent MDCK cell culture-derived influenza vaccine Optaflu (Novartis Vaccines). *Expert Rev Vaccines* 8(6):679–688.
231. **Yang LPH** (2013) Recombinant trivalent influenza vaccine (flublok®): a review of its use in the prevention of seasonal influenza in adults. *Drugs* 73(12):1357–1366.
232. **Borse RH, Shrestha SS, Fiore AE, Atkins CY, Singleton JA, Furlow C, Meltzer MI** (2013) Effects of Vaccine Program against Pandemic Influenza A(H1N1) Virus, United States, 2009–2010. *Emerg Infect Dis* 19(3).
233. **Berlanda Scorza F, Tsvetnitsky V, Donnelly JJ** (2016) Universal influenza vaccines: Shifting to better vaccines. *Vaccine* 34(26):2926–2933.
234. **Impagliazzo A, Milder F, Kuipers H, Wagner MV, Zhu X, Hoffman RMB, van Meersbergen R, Huizingh J, Wanningen P, Verspuij J, de Man M, Ding Z, Apetri A, Kukrer B, Sneekes-Vriese E, Tomkiewicz D, Laursen NS, Lee PS, Zakrzewska A, et al.** (2015) A stable trimeric influenza hemagglutinin stem as a broadly protective immunogen. *Science* 349(6254):1301–1306.
235. **Yassine HM, Boyington JC, McTamney PM, Wei C-J, Kanekiyo M, Kong W-P, Gallagher JR, Wang L, Zhang Y, Joyce MG, Lingwood D, Moin SM, Andersen H, Okuno Y, Rao SS, Harris AK, Kwong PD, Mascola JR, Nabel GJ, et al.** (2015) Hemagglutinin-stem nanoparticles generate heterosubtypic influenza protection. *Nat Med* 21(9):1065–1070.
236. **Subbarao K, Matsuoka Y** (2013) The prospects and challenges of universal vaccines for influenza. *Trends Microbiol* 21(7):350–358.
237. **Stouffer AL, Acharya R, Salom D, Levine AS, Di Costanzo L, Soto CS, Tereshko V, Nanda V, Stayrook S, DeGrado WF** (2008) Structural basis for the function and inhibition of an influenza virus proton channel. *Nature* 451:596–9.
238. **Schnell JR, Chou JJ** (2008) Structure and mechanism of the M2 proton channel of influenza A virus. *Nature* 451:591–5.
239. **Dong G, Peng C, Luo J, Wang C, Han L, Wu B, Ji G, He H** (2015) Adamantane-Resistant Influenza A Viruses in the World (1902–2013): Frequency and Distribution of M2 Gene Mutations. *PLoS ONE* 10(3).
240. **Gubareva LV, Trujillo AA, Okomo-Adhiambo M, Mishin VP, Deyde VM, Sleeman K, Nguyen HT, Sheu TG, Garten RJ, Shaw MW, Fry AM, Klimov AI** (2010) Comprehensive assessment of 2009 pandemic influenza A (H1N1) virus drug susceptibility in vitro. *Antivir Ther* 15(8):1151–1159.

241. **Kati WM, Montgomery D, Carrick R, Gubareva L, Maring C, McDaniel K, Steffy K, Molla A, Hayden F, Kempf D, Kohlbrenner W** (2002) In vitro characterization of A-315675, a highly potent inhibitor of A and B strain influenza virus neuraminidases and influenza virus replication. *Antimicrob Agents Chemother* 46:1014–21.
242. **Kawai N, Ikematsu H, Hirotsu N, Maeda T, Kawashima T, Tanaka O, Yamauchi S, Kawamura K, Matsuura S, Nishimura M, Iwaki N, Kashiwagi S** (2009) Clinical Effectiveness of Oseltamivir and Zanamivir for Treatment of Influenza A Virus Subtype H1N1 with the H274Y Mutation: A Japanese, Multicenter Study of the 2007–2008 and 2008–2009 Influenza Seasons. *Clin Infect Dis* 49(12):1828–1835.
243. **Hurt AC, Hardie K, Wilson NJ, Deng Y-M, Studies GDL Maggi Osbourn, BN, Gehrig N, Kelso A** (2011) Community Transmission of Oseltamivir-Resistant A(H1N1)pdm09 Influenza. *N Engl J Med* 365(26):2541–2542.
244. **Takashita E, Kiso M, Fujisaki S, Yokoyama M, Nakamura K, Shirakura M, Sato H, Odagiri T, Kawaoka Y, Tashiro M** (2015) Characterization of a Large Cluster of Influenza A(H1N1)pdm09 Viruses Cross-Resistant to Oseltamivir and Peramivir during the 2013–2014 Influenza Season in Japan. *Antimicrob Agents Chemother* 59(5):2607–2617.
245. **Hurt AC, Besselaar TG, Daniels RS, Ermetal B, Fry A, Gubareva L, Huang W, Lackenby A, Lee RTC, Lo J, Maurer-Stroh S, Nguyen HT, Pereyaslov D, Rebelo-de-Andrade H, Siqueira MM, Takashita E, Tashiro M, Tilmanis D, Wang D, et al.** (2016) Global update on the susceptibility of human influenza viruses to neuraminidase inhibitors, 2014–2015. *Antiviral Res.*
246. **Yen H-L** (2016) Current and novel antiviral strategies for influenza infection. *Curr Opin Virol* 18:126–134.
247. **Gerber M, Isel C, Moules V, Marquet R** (2014) Selective packaging of the influenza A genome and consequences for genetic reassortment. *Trends Microbiol* 22(8):446–455.
248. **Noda T, Kawaoka Y** (2012) Packaging of influenza virus genome: Robustness of selection. *Proc Natl Acad Sci U S A* 109(23):8797–8798.
249. **Luytjes W, Krystal M, Enami M, Parvin JD, Palese P** (1989) Amplification, expression, and packaging of a foreign gene by influenza virus. *Cell* 59(6):1107–1113.
250. **Neumann G, Watanabe T, Ito H, Watanabe S, Goto H, Gao P, Hughes M, Perez DR, Donis R, Hoffmann E, Hobom G, Kawaoka Y** (1999) Generation of influenza A viruses entirely from cloned cDNAs. *Proc Natl Acad Sci U A* 96:9345–50.
251. **Engelhardt OG** (2013) Many ways to make an influenza virus - review of influenza virus reverse genetics methods: Influenza virus reverse genetics methods. *Influenza Other Respir Viruses* 7(3):249–256.
252. **Chen H, Ye J, Xu K, Angel M, Shao H, Ferrero A, Sutton T, Perez DR** (2012) Partial and Full PCR-Based Reverse Genetics Strategy for Influenza Viruses. *PLoS ONE* 7(9):e46378. ed Brown EG.
253. **Hoffmann E, Neumann G, Kawaoka Y, Hobom G, Webster RG** (2000) A DNA transfection system for generation of influenza A virus from eight plasmids. *Proc Natl Acad Sci U A* 97:6108–13.
254. **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. *J Virol* 79:3766–74.
255. **Fujii Y, Goto H, Watanabe T, Yoshida T, Kawaoka Y** (2003) Selective incorporation of influenza virus RNA segments into virions. *Proc Natl Acad Sci U A* 100:2002–7.
256. **Gao Q, Brydon EW, Palese P** (2008) A seven-segmented influenza A virus expressing the influenza C virus glycoprotein HEF. *J Virol* 82:6419–26.
257. **Liang Y, Hong Y, Parslow TG** (2005) cis-Acting packaging signals in the influenza virus PB1, PB2, and PA genomic RNA segments. *J Virol* 79:10348–55.
258. **Marsh GA, Hatami R, Palese P** (2007) Specific residues of the influenza A virus hemagglutinin viral RNA are important for efficient packaging into budding virions. *J Virol* 81:9727–36.
259. **Muramoto Y, Takada A, Fujii K, Noda T, Iwatsuki-Horimoto K, Watanabe S, Horimoto T, Kida H, Kawaoka Y** (2006) Hierarchy among viral RNA (vRNA) segments in their role in vRNA incorporation into influenza A virions. *J Virol* 80:2318–25.
260. **Hutchinson EC, Curran MD, Read EK, Gog JR, Digard P** (2008) Mutational analysis of cis-acting RNA signals in segment 7 of influenza A virus. *J Virol* 82:11869–79.
261. **Liang Y, Huang T, Ly H, Parslow TG, Liang Y** (2008) Mutational analyses of packaging signals in influenza virus PA, PB1, and PB2 genomic RNA segments. *J Virol* 82:229–36.

262. **Marsh GA, Rabadan R, Levine AJ, Palese P** (2008) Highly conserved regions of influenza A virus polymerase gene segments are critical for efficient viral RNA packaging. *J Virol* 82:2295–304.
263. **Isel C, Munier S, Naffakh N** (2016) Experimental Approaches to Study Genome Packaging of Influenza A Viruses. *Viruses* 8(8):218.
264. **Zhao L, Peng Y, Zhou K, Cao M, Wang J, Wang X, Jiang T, Deng T** (2014) New Insights into the Nonconserved Noncoding Region of the Subtype-Determinant Hemagglutinin and Neuraminidase Segments of Influenza A Viruses. *J Virol* 88(19):11493–11503.
265. **Goto H, Muramoto Y, Noda T, Kawaoka Y** (2013) The genome-packaging signal of the influenza A virus genome comprises a genome incorporation signal and a genome-bundling signal. *J Virol* 87:11316–22.
266. **Baker SF, Nogales A, Finch C, Tuffy KM, Domm W, Perez DR, Topham DJ, Martínez-Sobrido L** (2014) Influenza A and B Virus Intertypic Reassortment through Compatible Viral Packaging Signals. *J Virol* 88(18):10778–10791.
267. **Juozapaitis M, Aguiar Moreira É, Mena I, Giese S, Riegger D, Pohlmann A, Höper D, Zimmer G, Beer M, García-Sastre A, Schwemmle M** (2014) An infectious bat-derived chimeric influenza virus harbouring the entry machinery of an influenza A virus. *Nat Commun* 5:4448.
268. **Zhou B, Ma J, Liu Q, Bawa B, Wang W, Shabman RS, Duff M, Lee J, Lang Y, Cao N, Nagy A, Lin X, Stockwell TB, Richt JA, Wentworth DE, Ma W** (2014) Characterization of Uncultivable Bat Influenza Virus Using a Replicative Synthetic Virus. *PLoS Pathog* 10(10).
269. **Bergeron C, Valette M, Lina B, Ottmann M** (2010) Genetic content of Influenza H3N2 vaccine seeds. *PLoS Curr* 2.
270. **Cobbin JCA, Verity EE, Gilbertson BP, Rockman SP, Brown LE** (2013) The Source of the PB1 Gene in Influenza Vaccine Reassortants Selectively Alters the Hemagglutinin Content of the Resulting Seed Virus. *J Virol* 87(10):5577–5585.
271. **Fulvini AA, Ramanunnair M, Le J, Pokorny BA, Arroyo JM, Silverman J, Devis R, Bucher D** (2011) Gene Constellation of Influenza A Virus Reassortants with High Growth Phenotype Prepared as Seed Candidates for Vaccine Production. *PLOS ONE* 6(6):e20823.
272. **Cobbin JCA, Ong C, Verity E, Gilbertson BP, Rockman SP, Brown LE** (2014) Influenza Virus PB1 and Neuraminidase Gene Segments Can Cosegregate during Vaccine Reassortment Driven by Interactions in the PB1 Coding Region. *J Virol* 88(16):8971–8980.
273. **Gavazzi C, Yver M, Isel C, Smyth RP, Rosa-Calatrava M, Lina B, Moules V, Marquet R** (2013) A functional sequence-specific interaction between influenza A virus genomic RNA segments. *Proc Natl Acad Sci U S A* 110:16604–9.
274. **Mosley Y-YC, Wu CC, Lin TL** (2016) IBDV particles packaged with only segment A dsRNA. *Virology* 488:68–72.
275. **Luque D, Rivas G, Alfonso C, Carrascosa JL, Rodríguez JF, Castón JR** (2009) Infectious bursal disease virus is an icosahedral polypliod dsRNA virus. *Proc Natl Acad Sci* 106(7):2148–2152.
276. **Abd El-Wahab EW, Smyth RP, Mailler E, Bernacchi S, Vivet-Boudou V, Hijnen M, Jossinet F, Mak J, Paillart J-C, Marquet R** (2014) Specific recognition of the HIV-1 genomic RNA by the Gag precursor. *Nat Commun* 5:4304.
277. **Smyth RP, Despons L, Huili G, Bernacchi S, Hijnen M, Mak J, Jossinet F, Weixi L, Paillart J-C, von Kleist M, Marquet R** (2015) Mutational interference mapping experiment (MIME) for studying RNA structure and function. *Nat Methods* 12(9):866–872.
278. **Lourenco S, Roy P** (2011) In vitro reconstitution of Bluetongue virus infectious cores. *Proc Natl Acad Sci* 108(33):13746–13751.
279. **Sung P-Y, Roy P** (2014) Sequential packaging of RNA genomic segments during the assembly of Bluetongue virus. *Nucleic Acids Res* 42(22):13824–13838.
280. **Fajardo T, Sung P-Y, Roy P** (2015) Disruption of Specific RNA-RNA Interactions in a Double-Stranded RNA Virus Inhibits Genome Packaging and Virus Infectivity. *PLoS Pathog* 11(12).
281. **Boyce P, McCrae MA, Kim JT, Boyce M** (2016) Inter-segment complementarity in orbiviruses: a driver for co-ordinated genome packaging in the Reoviridae? *J Gen Virol* 97(5):1145–1157.
282. **Burkhardt C, Sung P-Y, Celma CC, Roy P** (2014) Structural constraints in the packaging of bluetongue virus genomic segments. *J Gen Virol* 95(Pt\_10):2240–2250.
283. **Chen D, Patton JT** (1998) Rotavirus RNA replication requires a single-stranded 3' end for efficient minus-strand synthesis. *J Virol* 72(9):7387–7396.

284. **Li W, Manktelow E, von Kirchbach JC, Gog JR, Desselberger U, Lever AM** (2010) Genomic analysis of codon, sequence and structural conservation with selective biochemical-structure mapping reveals highly conserved and dynamic structures in rotavirus RNAs with potential cis-acting functions. *Nucleic Acids Res* 38(21):7718–7735.
285. **Tortorici MA, Shapiro BA, Patton JT** (2006) A base-specific recognition signal in the 5' consensus sequence of rotavirus plus-strand RNAs promotes replication of the double-stranded RNA genome segments. *RNA N Y N* 12(1):133–146.
286. **McDonald SM, Nelson MI, Turner PE, Patton JT** (2016) Reassortment in segmented RNA viruses: mechanisms and outcomes. *Nat Rev Microbiol* 14(7):448–460.
287. **Suzuki Y** (2014) A possible packaging signal in the rotavirus genome. *Genes Genet Syst* 89(2):81–86.
288. **Pirttimaa MJ, Bamford DH** (2000) RNA secondary structures of the bacteriophage phi6 packaging regions. *RNA* 6(6):880–889.
289. **Mindich L** (2004) Packaging, replication and recombination of the segmented genome of bacteriophage Phi6 and its relatives. *Virus Res* 101:83–92.
290. **Huiskonen JT, Haas F de, Bubeck D, Bamford DH, Fuller SD, Butcher SJ** (2006) Structure of the Bacteriophage phi6 Nucleocapsid Suggests a Mechanism for Sequential RNA Packaging. *Structure* 14(6):1039–1048.
291. **Nemecek D, Cheng N, Qiao J, Mindich L, Steven AC, Heymann JB** (2011) Stepwise Expansion of the Bacteriophage phi6 Procapsid: Possible Packaging Intermediates. *J Mol Biol* 414(2):260–271.
292. **Ghedin E, Sengamalay NA, Shumway M, Zaborsky J, Feldblyum T, Subbu V, Spiro DJ, Sitz J, Koo H, Bolotov P, Dernovoy D, Tatusova T, Bao Y, St George K, Taylor J, Lipman DJ, Fraser CM, Taubenberger JK, Salzberg SL** (2005) Large-scale sequencing of human influenza reveals the dynamic nature of viral genome evolution. *Nature* 437(7062):1162–1166.
293. **Holmes EC, Ghedin E, Miller N, Taylor J, Bao Y, St George K, Grenfell BT, Salzberg SL, Fraser CM, Lipman DJ, Taubenberger JK** (2005) Whole-genome analysis of human influenza A virus reveals multiple persistent lineages and reassortment among recent H3N2 viruses. *PLoS Biol* 3:e300.
294. **Lindstrom SE, Cox NJ, Klimov A** (2004) Genetic analysis of human H2N2 and early H3N2 influenza viruses, 1957–1972: evidence for genetic divergence and multiple reassortment events. *Virology* 328(1):101–119.
295. **Nelson MI, Simonsen L, Viboud C, Miller MA, Taylor J, George KS, Griesemer SB, Ghedin E, Sengamalay NA, Spiro DJ, Volkov I, Grenfell BT, Lipman DJ, Taubenberger JK, Holmes EC** (2006) Stochastic processes are key determinants of short-term evolution in influenza a virus. *PLoS Pathog* 2:e125.
296. **Nelson MI, Viboud C, Simonsen L, Bennett RT, Griesemer SB, St George K, Taylor J, Spiro DJ, Sengamalay NA, Ghedin E, Taubenberger JK, Holmes EC** (2008) Multiple reassortment events in the evolutionary history of H1N1 influenza A virus since 1918. *PLoS Pathog* 4:e1000012.
297. **Rambaut A, Pybus OG, Nelson MI, Viboud C, Taubenberger JK, Holmes EC** (2008) The genomic and epidemiological dynamics of human influenza A virus. *Nature* 453(7195):615–619.
298. **Schweiger B, Bruns L, Meixenberger K** (2006) Reassortment between human A(H3N2) viruses is an important evolutionary mechanism. *Vaccine* 24:6683–90.
299. **Simonsen L, Viboud C, Grenfell BT, Dushoff J, Jennings L, Smit M, Macken C, Hata M, Gog J, Miller MA, Holmes EC** (2007) The Genesis and Spread of Reassortment Human Influenza A/H3N2 Viruses Conferring Adamantane Resistance. *Mol Biol Evol* 24(8):1811–1820.
300. **Chen M-J, La T, Zhao P, Tam JS, Rappaport R, Cheng S-M** (2006) Genetic and phylogenetic analysis of multi-continent human influenza A(H1N2) reassortant viruses isolated in 2001 through 2003. *Virus Res* 122(1–2):200–205.
301. **Cox NJ, Bai ZS, Kendal AP** (1983) Laboratory-based surveillance of influenza A(H1N1) and A(H3N2) viruses in 1980-81: antigenic and genomic analyses. *Bull World Health Organ* 61(1):143–152.
302. **Ellis JS, Alvarez-Aguero A, Gregory V, Lin YP, Hay A, Zambon MC** (2003) Influenza AH1N2 viruses, United Kingdom, 2001-02 influenza season. *Emerg Infect Dis* 9(3):304–310.
303. **Mizuta K, Katsushima N, Ito S, Sanjoh K, Murata T, Abiko C, Murayama S** (2003) A rare appearance of influenza A(H1N2) as a reassortant in a community such as Yamagata where A(H1N1) and A(H3N2) co-circulate. *Microbiol Immunol* 47(5):359–361.

304. **Garten RJ, Davis CT, Russell CA, Shu B, Lindstrom S, Balish A, Sessions WM, Xu X, Skepner E, Deyde V, Okomo-Adhiambo M, Gubareva L, Barnes J, Smith CB, Emery SL, Hillman MJ, Rivaller P, Smagala J, de Graaf M, et al.** (2009) Antigenic and genetic characteristics of swine-origin 2009 A(H1N1) influenza viruses circulating in humans. *Science* 325:197–201.
305. **Smith GJ, Vijaykrishna D, Bahl J, Lycett SJ, Worobey M, Pybus OG, Ma SK, Cheung CL, Raghvani J, Bhatt S, Peiris JS, Guan Y, Rambaut A** (2009) Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. *Nature* 459:1122–5.
306. **Simonsen L, Spreeuwenberg P, Lustig R, Taylor RJ, Fleming DM, Kroneman M, Van Kerkhove MD, Mounts AW, Paget WJ, the GLaMOR Collaborating Teams** (2013) Global Mortality Estimates for the 2009 Influenza Pandemic from the GLaMOR Project: A Modeling Study. *PLoS Med* 10(11):e1001558. ed Hay SI.
307. **Essere B, Yver M, Gavazzi C, Terrier O, Isel C, Fournier E, Giroux F, Textoris J, Julien T, Socratous C, Rosa-Calatrava M, Lina B, Marquet R, Moules V** (2013) Critical role of segment-specific packaging signals in genetic reassortment of influenza A viruses. *Proc Natl Acad Sci U S A* 110:E3840–8.
308. **Fournier E, Moules V, Essere B, Paillart JC, Sirbat JD, Isel C, Cavalier A, Rolland JP, Thomas D, Lina B, Marquet R** (2012) A supramolecular assembly formed by influenza A virus genomic RNA segments. *Nucleic Acids Res* 40:2197–209.
309. **Hatta M, Halfmann P, Wells K, Kawaoka Y** (2002) Human influenza A viral genes responsible for the restriction of its replication in duck intestine. *Virology* 295(2):250–255.
310. **Jackson S, Van Hoven N, Chen LM, Maines TR, Cox NJ, Katz JM, Donis RO** (2009) Reassortment between avian H5N1 and human H3N2 influenza viruses in ferrets: a public health risk assessment. *J Virol* 83:8131–40.
311. **Li C, Hatta M, Watanabe S, Neumann G, Kawaoka Y** (2008) Compatibility among polymerase subunit proteins is a restricting factor in reassortment between equine H7N7 and human H3N2 influenza viruses. *J Virol* 82:11880–8.
312. **Hara K, Nakazono Y, Kashiwagi T, Hamada N, Watanabe H** (2013) Co-incorporation of the PB2 and PA polymerase subunits from human H3N2 influenza virus is a critical determinant of the replication of reassortant ribonucleoprotein complexes. *J Gen Virol* 94:2406–16.
313. **Octaviani CP, Ozawa M, Yamada S, Goto H, Kawaoka Y** (2010) High level of genetic compatibility between swine-origin H1N1 and highly pathogenic avian H5N1 influenza viruses. *J Virol* 84:10918–22.
314. **Octaviani CP, Goto H, Kawaoka Y** (2011) Reassortment between seasonal H1N1 and pandemic (H1N1) 2009 influenza viruses is restricted by limited compatibility among polymerase subunits. *J Virol* 85:8449–52.
315. **Chen LM, Davis CT, Zhou H, Cox NJ, Donis RO** (2008) Genetic compatibility and virulence of reassortants derived from contemporary avian H5N1 and human H3N2 influenza A viruses. *PLoS Pathog* 4:e1000072.
316. **Moisy D, Naffakh N** (2013) Interactions moléculaires entre les ribonucléoprotéines des virus influenza de type A et la cellule hôte. *Virologie* 17(4):278–293.
317. **Marshall N, Priyamvada L, Ende Z, Steel J, Lowen AC** (2013) Influenza virus reassortment occurs with high frequency in the absence of segment mismatch. *PLoS Pathog* 9:e1003421.
318. **Smith MR, Smyth RP, Marquet R, von Kleist M** (2016) MIMeAnTo - Profiling functional RNA in Mutational Interference Mapping Experiments. *Bioinforma Oxf Engl*.
319. **Gavazzi C, Isel C, Fournier E, Moules V, Cavalier A, Thomas D, Lina B, Marquet R** (2013) An in vitro network of intermolecular interactions between viral RNA segments of an avian H5N2 influenza A virus: comparison with a human H3N2 virus. *Nucleic Acids Res* 41:1241–54.
320. **Noda T, Sugita Y, Aoyama K, Hirase A, Kawakami E, Miyazawa A, Sagara H, Kawaoka Y** (2012) Three-dimensional analysis of ribonucleoprotein complexes in influenza A virus. *Nat Commun* 3:639.
321. **Marquet R, Baudin F, Gabus C, Darlix J-L, Mougél M, Ehresmann C, Ehresmann B** (1991) Dimerization of human immunodeficiency virus (type 1) RNA: stimulation by cations and possible mechanism. *Nucleic Acids Res* 19(9):2349–2357.
322. **Hutchinson EC, von Kirchbach JC, Gog JR, Digard P** (2010) Genome packaging in influenza A virus. *J Gen Virol* 91:313–28.
323. **Ozawa M, Maeda J, Iwatsuki-Horimoto K, Watanabe S, Goto H, Horimoto T, Kawaoka Y** (2009) Nucleotide sequence requirements at the 5' end of the influenza A virus M RNA segment for efficient virus replication. *J Virol* 83:3384–8.

324. **Gao Q, Chou YY, Doganay S, Vafabakhsh R, Ha T, Palese P** (2012) The influenza A virus PB2, PA, NP, and M segments play a pivotal role during genome packaging. *J Virol* 86:7043–51.
325. **Watanabe T, Watanabe S, Noda T, Fujii Y, Kawaoka Y** (2003) Exploitation of nucleic acid packaging signals to generate a novel influenza virus-based vector stably expressing two foreign genes. *J Virol* 77:10575–83.
326. **Gerlach W, Giegerich R** (2006) GUUGle: a utility for fast exact matching under RNA complementary rules including G–U base pairing. *Bioinformatics* 22(6):762–764.
327. **Silverman IM, Berkowitz ND, Gosai SJ, Gregory BD** (2016) Genome-Wide Approaches for RNA Structure Probing. *Adv Exp Med Biol* 907:29–59.
328. **Kolb FA, Malmgren C, Westhof E, Ehresmann C, Ehresmann B, Wagner EG, Romby P** (2000) An unusual structure formed by antisense-target RNA binding involves an extended kissing complex with a four-way junction and a side-by-side helical alignment. *Rna* 6:311–24.
329. **Crouzier L, Dubois C, Edwards SL, Lauridsen LH, Wengel J, Veedu RN** (2012) Efficient reverse transcription using locked nucleic acid nucleotides towards the evolution of nuclease resistant RNA aptamers. *PLoS One* 7(4):e35990.
330. **Fratczak A, Kierzek R, Kierzek E** (2009) LNA-modified primers drastically improve hybridization to target RNA and reverse transcription. *Biochemistry (Mosc)* 48(3):514–516.
331. **Mohr S, Ghanem E, Smith W, Sheeter D, Qin Y, King O, Polioudakis D, Iyer VR, Hunicke-Smith S, Swamy S, Kuersten S, Lambowitz AM** (2013) Thermostable group II intron reverse transcriptase fusion proteins and their use in cDNA synthesis and next-generation RNA sequencing. *RNA N Y N* 19(7):958–970.
332. **Markham NR, Zuker M** (2005) DINAMelt web server for nucleic acid melting prediction. *Nucleic Acids Res* 33(Web Server issue):W577–581.
333. **Zuker M** (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* 31(13):3406–3415.
334. **Skripkin E, Paillart JC, Marquet R, Ehresmann B, Ehresmann C** (1994) Identification of the primary site of the human immunodeficiency virus type 1 RNA dimerization in vitro. *Proc Natl Acad Sci U S A* 91(11):4945–4949.
335. **Regan JF, Liang Y, Parslow TG** (2006) Defective assembly of influenza A virus due to a mutation in the polymerase subunit PA. *J Virol* 80:252–61.
336. **Paillart J-C, Shehu-Xhilaga M, Marquet R, Mak J** (2004) Dimerization of retroviral RNA genomes: an inseparable pair. *Nat Rev Microbiol* 2(6):461–472.
337. **Sugita Y, Sagara H, Noda T, Kawaoka Y** (2013) Configuration of Viral Ribonucleoprotein Complexes within the Influenza A Virion. *J Virol* 87(23):12879–12884.
338. **Gog JR, Afonso Edos S, Dalton RM, Leclercq I, Tiley L, Elton D, von Kirchbach JC, Naffakh N, Escriou N, Digard P** (2007) Codon conservation in the influenza A virus genome defines RNA packaging signals. *Nucleic Acids Res* 35:1897–907.
339. **Dos Santos Afonso E, Escriou N, Leclercq I, van der Werf S, Naffakh N** (2005) The generation of recombinant influenza A viruses expressing a PB2 fusion protein requires the conservation of a packaging signal overlapping the coding and noncoding regions at the 5' end of the PB2 segment. *Virology* 341(1):34–46.
340. **Kimble JB, Angel M, Wan H, Sutton TC, Finch C, Perez DR** (2014) Alternative Reassortment Events Leading to Transmissible H9N1 Influenza Viruses in the Ferret Model. *J Virol* 88(1):66–71.
341. **Gilbertson B, Zheng T, Gerber M, Printz-Schweigert A, Ong C, Marquet R, Isel C, Rockman S, Brown L** (2016) Influenza NA and PB1 Gene Segments Interact during the Formation of Viral Progeny: Localization of the Binding Region within the PB1 Gene. *Viruses* 8(8):238.

**Marie GERBER**

## **Recherche de signaux d’empaquetage spécifiques du génome des virus influenza A**

Les huit ARN viraux (ARNv) génomiques des *influenzavirus* de type A sont sélectivement empaquetés dans les virions, vraisemblablement sous la forme d’un complexe supramoléculaire maintenu par des appariements ARN/ARN entre signaux d’empaquetage.

Afin d’améliorer la compréhension des règles qui gouvernent ce mécanisme, nous avons déterminé le réseau d’interactions formé entre les ARNv de la souche modèle de laboratoire A/Puerto Rico/8/34 (H1N1). Nous avons ensuite défini, au nucléotide près et par deux approches *in vitro*, les séquences des ARNv impliquées dans la formation de certaines interactions. Le rôle fonctionnel de deux d’entre elles a été testé en contexte infectieux.

Nous avons également poursuivi l’étude d’une interaction identifiée au laboratoire entre deux ARNv de la souche A/Moscou/10/99 (H3N2) circulante. Enfin, nous avons collaboré avec le Dr L. Brown (Australie) sur l’étude du rôle d’une interaction entre deux ARNv de la souche A/Udorn/307/72 (H3N2).

Mots clés : *Influenzavirus A*, ARN viraux, empaquetage du génome, interaction intermoléculaire ARN/ARN, génétique inverse

## **Identification of specific packaging signals in influenza A viruses genome**

The genome of influenza A viruses comprises eight viral RNAs (vRNAs) likely to be selectively packaged into progeny virions as organized supramolecular complexes where vRNAs are held together by base pairing within the packaging signals.

To better understand the rules governing this mechanism, we investigated the vRNA/vRNA interaction network of the A/Puerto Rico/8/34 (H1N1) strain. We then identified, at the nucleotide level, using two *in vitro* approaches, the vRNA sequences involved in several of the interactions. Two of them were functionally tested in an infectious context.

We also studied an interaction previously identified in the laboratory between two vRNAs belonging to the circulating A/Moscow/10/99 (H3N2) strain. Finally, we collaborated with Dr L. Brown (Australia) in order to assess the role of an interaction between two vRNAs of the A/Udorn/307/72 (H3N2) strain.

Keywords : influenza A virus, viral RNA, genome packaging, intermolecular RNA/RNA interaction, reverse genetics