



HAL
open science

L'ARN-polymérase des virus influenza et ses interactions avec la chromatine

Jessica Morel

► **To cite this version:**

Jessica Morel. L'ARN-polymérase des virus influenza et ses interactions avec la chromatine. Virologie. Université Paris-Saclay, 2022. Français. NNT : 2022UPASL019 . tel-04861209

HAL Id: tel-04861209

<https://theses.hal.science/tel-04861209v1>

Submitted on 2 Jan 2025

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.

L'ARN-polymérase des virus Influenza et ses interactions avec la chromatine

Influenza virus RNA-polymerase and its interactions with chromatin

Thèse de doctorat de l'université Paris-Saclay

École doctorale n°577, Structure et Dynamique des Systèmes Vivants (SDSV)
Spécialité de doctorat : Sciences de la Vie et de la Santé
Graduate School : Life Sciences and Health
Réfèrent : Université de Versailles Saint-Quentin-en-Yvelines (UVSQ)

Thèse préparée dans l'unité de recherche **VIM (Université Paris-Saclay, UVSQ, INRAE)**,
sous la direction de **Bernard DELMAS**, directeur de recherche,
la co-direction de **Christian MUCHARDT**, directeur de recherche

Thèse soutenue à Jouy-en-Josas, le 23 mars 2022, par

Jessica MOREL

Composition du Jury

Marie-Anne RAMEIX-WELTI Professeure des universités, Université de Versailles Saint-Quentin-en-Yvelines (UVSQ)	Présidente
Darren HART Directeur de recherche, Université Grenoble Alpes	Rapporteur & Examineur
Catherine ISEL-GRIFFITHS Chargée de recherche, Université de Paris	Rapporteur & Examinatrice
Pierre LAFAYE Chargé de recherche, Université de Paris	Examineur
Bernard DELMAS Directeur de recherche, Université de Versailles Saint-Quentin-en-Yvelines (UVSQ)	Directeur de thèse

Titre : L'ARN-polymérase des virus Influenza et ses interactions avec la chromatine

Mots clés : Influenza, ARN-polymérase, transcription, vol de coiffe, chromatine, nanoligands

Résumé : L'objectif général de ce travail de thèse a été de mieux comprendre le fonctionnement de l'ARN-polymérase des virus Influenza (FluPol) qui assure transcription et réplication du génome viral dans le noyau de la cellule hôte. Les virus Influenza ne disposant pas de la capacité à synthétiser une coiffe à l'extrémité 5' des ARNm viraux, ils utilisent les ARNm cellulaires coiffés pour amorcer la transcription virale par un mécanisme appelé vol de coiffe. Ainsi, FluPol se fixe sur la coiffe d'un ARNm cellulaire pour le cliver 10-15 nucléotides en aval et générer l'amorce permettant de transcrire l'ARN génomique en ARNm. Par ailleurs, FluPol interagit avec le domaine non structuré de l'ARN-polymérase II cellulaire (Pol II) constitué de 52 répétitions du motif YSPTSPS. La phosphorylation de la sérine5 de ce motif, marque de l'initiation de la transcription cellulaire, apparaît être un prérequis à l'interaction FluPol - Pol II. Cette interaction permet vraisemblablement à FluPol de cibler les étapes d'initiation. Le premier objectif de ce travail de thèse a été d'identifier des interactions additionnelles entre FluPol et Pol II potentiellement impliquées dans le vol de coiffe. Une interaction entre le domaine N-terminal de la sous-unité

PB2 de FluPol et la sous-unité RPB4 de Pol II a ainsi été identifiée. Cette interaction est conservée chez les virus de type A, B et C. Dans une seconde partie, ce travail a permis de mettre en évidence par des approches de séquençage à haut-débit que FluPol reste fixée à Pol II tout au long de la transcription cellulaire et non uniquement à son initiation. Ces approches ont également permis de mettre en évidence un ciblage préférentiel de FluPol pour les gènes ayant un rôle dans l'immunité innée, participant ainsi au contrôle de la réponse innée de l'hôte. Enfin, afin de mieux comprendre le fonctionnement de FluPol, des nanoligands de deux types (des VHH et des protéines artificielles) spécifiques de FluPol ont été sélectionnés. Un VHH s'est avéré bloquer efficacement l'activité enzymatique de FluPol et reconnaître les sous-unités PA et PB1 de FluPol. À terme, ces nouveaux outils d'étude devraient permettre d'élucider la dynamique des mécanismes moléculaires associés à la transcription et la réplication virale par la caractérisation des fonctions altérées de FluPol et des domaines ciblés à l'échelle atomique.

Title : Influenza virus RNA-polymerase and its interactions with chromatin

Keywords : Influenza, RNA-polymerase, transcription, cap-snatching, chromatin, nanoligands

Abstract : The general goal of this PhD work was to better understand the mechanisms of the RNA-polymerase of Influenza viruses (FluPol), which carries on transcription and replication of the viral genome in the nucleus of host cells. Since Influenza viruses are not capable of synthesizing a cap at the 5' termini of their mRNAs, they use the ones of cellular capped mRNAs to prime viral transcription by a mechanism referred to as cap-snatching. Thus, FluPol binds to the cap of a cellular mRNA and cleaves it 10-15 nucleotides downstream to generate the primer allowing the transcription of genomic RNA into mRNA. Moreover, it was shown that FluPol interacts with the unstructured domain of the cellular RNA-polymerase II (Pol II) composed of 52 repeats the motif YSPTSPS. Phosphorylation on the serine 5 of this motif, a sign of the cellular transcription initiation, appears to be a prerequisite to the FluPol – Pol II interaction. In all likelihood, this interaction allows FluPol to target transcription initiation stages. The first goal of this thesis work was to identify other interactions between FluPol and Pol II potentially involved in cap-

snatching. An interaction between the N-terminal domain of the PB2 subunit of FluPol and the RPB4 subunit of Pol II was identified. This interaction is conserved among types A, B and C viruses. In another section, this work also uncovered by high-throughput sequencing that FluPol binds to Pol II all the way through transcription and not just during the initiation stage. These experiments also brought out that FluPol targeted innate immune response genes, participating in host shut-off. Finally, in order to better understand FluPol's working, two types of nanoligands (VHH and artificial proteins) specific to FluPol were selected. One VHH happened to efficiently block the enzymatic activity of FluPol and recognize the PA and PB1 subunits of FluPol. Eventually, these novel tools could help elucidate the dynamic of molecular mechanisms associated with viral transcription and replication by characterizing the altered functions of FluPol and the targeted regions at an atomic scale.

Remerciements

Tout d'abord, je tiens à remercier mon directeur de thèse, Bernard Delmas, qui m'a accueillie dans l'équipe grippe et sans qui cette aventure n'aurait pas été possible. Merci de m'avoir fait confiance depuis le début. Un grand merci également à Christian Muchardt, mon co-directeur de thèse ainsi qu'à Eric Batsché pour l'accueil qu'ils m'ont réservé et leur investissement dans ce projet. Eric, merci pour toute l'énergie que tu as déployée pour essayer de me transmettre ton savoir et pour m'accompagner dans les dernières (difficiles) étapes de rédaction.

Merci à Sabine Riffault de m'avoir accueillie au sein de la VIM.

Un grand merci aux membres du jury qui ont accepté d'évaluer ce travail de thèse : Darren Hart, Catherine Isel-Griffiths, Marie-Anne Rameix-Welti et Pierre Lafaye.

Un grand merci également aux membres de mon comité de suivi : Valérie Gaudin, Catherine Demeret et Delphine Sitterlin. Merci pour vos précieux conseils et votre bienveillance. Je souhaite adresser un merci tout particulier à Delphine qui a suivi tout mon parcours et a été la première à me donner ma chance en me donnant l'opportunité de découvrir un laboratoire de virologie et de conforter mon attrait pour cette discipline.

Un immense merci aux nombreux collaborateurs que j'ai eu la chance de côtoyer au cours de ces trois années. Ma thèse n'aurait pas été aussi enrichissante sans vous.

Je remercie Gwenaëlle André-Leroux pour la modélisation de l'interaction entre PB2 et RPB4. Tu as su répondre à nos questions et ton aide nous a permis d'avancer et surtout clôturer le plus rapidement possible cet axe avant de rendre le manuscrit.

Merci à Alain Roussel, Magali Grange et Anaïs Gaubert de l'AFMB de Marseille pour la sélection des VHH.

Merci à Philippe Minard, Agathe Urvoas et Marielle Valério-Lepiniec de m'avoir accueillie dans votre équipe et de m'avoir accompagnée dans la sélection des α Rep. Un merci tout particulier à Agnès Mesneau. J'ai beaucoup aimé travailler avec toi. Ta bonne humeur, ton sourire et nos conversations sont les souvenirs que je conserverai de ce mois passé à vos côtés.

Un très grand merci à Thibaut Crépin et Amélie Donchet de m'avoir accueillie chaleureusement à Grenoble alors que je débutais tout juste ma thèse, de m'avoir fait découvrir cette ville mais aussi le monde de la biochimie et de la purification de protéines qui m'étaient jusqu'alors inconnus.

Je souhaite également remercier Mickaël Costallat qui a essayé de m'inculquer les bases du codage et de la bio-informatique en une semaine ! Un challenge difficile à relever mais tu t'y es attelé avec bénévolence.

Bien entendu, merci à tous les collègues de la VIM et particulièrement de l'équipe grippe / coronavirus. Merci à Didier Chevret, Marie Galloux, Thomas Clark, Christophe Chevalier, Ronan Le Goffic, Olivier Leymarie, Nathalie Lejal, Nicolas Meunier, Audrey Saint-Albin, Léna Kleij, Clara Bourgon, Ophélie Gard, Molidia Mao Elise Bruder, Soraya Dinant et Joelle Mettier.

Je souhaite remercier expressément Laura Sedano qui m'a formée à mon arrivée à la VIM. Tout cela est en partie grâce à toi. Merci de m'avoir transmis ta rigueur et merci pour ton appui technique et les nombreux clonages que tu as fait pour que l'on puisse progresser plus rapidement. Merci également à Bruno Da Costa pour son appui technique. Merci de m'avoir proposée spontanément ton aide à maintes reprises et pour ton accompagnement pendant les infections. J'ai beaucoup apprécié

les conversations que nous avons eu dans le P3, que ce soit celles où tu me transmettais ton savoir ou celles où nous refaisons le monde. Merci à Quentin Viboud pour sa bonne humeur et son coaching de carrière. Merci à Cynthia Calzas et Stéphanie Thebault pour toutes ces conversations, que ce soit sur l'avenir professionnel ou nos expériences respectives au Canada. J'espère que nous garderons le contact. Merci également à Mélissa Bessonne. Ça a été un plaisir de pouvoir t'épauler à ton arrivée dans l'équipe. Je te transmets le flambeau, à toi maintenant d'assurer la relève. Je ne doute pas que tu feras une thèse à la hauteur de tes espérances. Soit patiente et, tu verras, tout finira par arriver (souvent à la toute fin !).

A Hortense Decool, j'ai été ravie d'avoir pu passer ces presque quatre années avec toi. Ensemble on a ri, on a pleuré, on s'est épaulées et consolées dans les moments difficiles, parlé sentiments, été confidentes, bref, on a vécu cette aventure folle qu'est la thèse côte à côte. J'espère que notre amitié subsistera au-delà. Je te souhaite tout le meilleur du monde et un excellent post-doc en Allemagne (qui sait, je trouverai peut-être le temps de vous rendre visite à Raph et toi).

A mes amies rencontrées sur les bancs de l'université : Pauline Michel-Flutot et Marion Pilard, je suis heureuse de vous avoir rencontrées et que nous ayons su garder ce lien malgré nos chemins qui divergent de plus en plus. J'espère que nous saurons continuer, malgré la distance géographique encore plus grande qui est amenée à nous séparer. A Margaux Gardet, Virginie Menez et Johanna Canovas, je suis reconnaissante de notre quatuor improbable. Merci pour ces folles soirées entre filles. J'espère qu'il y en aura d'autres !

Pour finir, un immense merci à toute ma famille qui m'a toujours soutenue dans ce long parcours. Particulièrement mes parents et surtout ma mère qui m'a toujours poussée à donner le meilleur de moi-même. Maman, la cheminée est enfin sur le toit de la maison. Et enfin, un profond merci à mon Namour, mon roc, mon phare dans la tempête. Je ne peux imaginer ce qu'aurait été ces trois dernières années sans toi à mes côtés. J'ai hâte que nous commencions ce nouveau chapitre. Ensemble(s).

Table des matières

LISTE DES ABREVIATIONS.....	7
INTRODUCTION.....	9
I. CARACTERISTIQUES DES VIRUS INFLUENZA.....	11
1) <i>Taxonomie</i>	11
2) <i>Structure de la particule virale</i>	12
3) <i>Organisation du génome</i>	14
4) <i>Structure de la RNP</i>	16
II. EPIDEMIOLOGIE DES VIRUS INFLUENZA.....	18
1) <i>Ecologie</i>	18
2) <i>Evolution</i>	19
a. <i>Dérive antigénique</i>	19
b. <i>Cassure antigénique</i>	20
3) <i>Pandémies</i>	20
III. CYCLE VIRAL.....	21
1) <i>Adhésion et entrée du virus dans la cellule hôte</i>	22
2) <i>Libération des RNP dans le cytoplasme</i>	23
3) <i>Import des RNP au noyau</i>	25
4) <i>Transcription et réplication du génome</i>	26
5) <i>Export des RNP</i>	26
6) <i>Assemblage et bourgeonnement</i>	27
IV. L'ARN-POLYMERASE ARN-DEPENDANTE FLUPOL.....	29
1) <i>Description</i>	29
2) <i>Transport au noyau</i>	31
3) <i>Réplication du génome viral</i>	32
4) <i>Transcription du génome viral et vol de coiffe</i>	34
5) <i>Interactions avec les facteurs cellulaires</i>	35
a. <i>Interaction avec l'ARN-polymérase II</i>	35
b. <i>Interaction avec ANP32</i>	38
c. <i>Autres interactions et ciblage à la chromatine</i>	38
6) <i>Structure(s) de FluPol</i>	39
V. CONSEQUENCES DE L'INFECTION SUR LA CHROMATINE DE L'HOTE.....	42
OBJECTIF GÉNÉRAL.....	45
PARTIE I - ÉTUDE DES INTERACTIONS ENTRE FLUPOL ET POL II.....	47
OBJECTIFS.....	49
I. ARTICLE : "THE INFLUENZA VIRUS RNA-POLYMERASE AND THE HOST RNA-POLYMERASE II: RPB4 IS TARGETED BY A PB2 DOMAIN THAT IS INVOLVED IN VIRAL TRANSCRIPTION".....	51
II. EXPERIENCES COMPLEMENTAIRES.....	67
1) <i>Recherche des acides aminés impliqués dans l'interaction PB2 – RPB4</i>	67
2) <i>Modélisation de l'interaction PB2 – RPB4</i>	68
CONCLUSION.....	73
PARTIE II - ÉTUDE DES INTERACTIONS ENTRE FLUPOL ET LA CHROMATINE.....	78
OBJECTIFS.....	80
ARTICLE : "INFLUENZA A VIRUS RNA-POLYMERASE TARGETS THE CHROMATIN OF INNATE IMMUNE RESPONSE GENES".....	82

PARTIE III - SÉLECTION ET PROPRIÉTÉS DE NANOLIGANDS SPÉCIFIQUES DE FLUPOL.....	101
INTRODUCTION	103
1) <i>Les VHH</i>	<i>103</i>
2) <i>Les αRep</i>	<i>107</i>
3) <i>Comment sélectionner des VHH ou des αRep spécifiques?</i>	<i>109</i>
OBJECTIFS	111
RESULTATS	112
1) <i>Les formes de FluPol utilisées pour la sélection de VHH et αRep spécifiques</i>	<i>112</i>
<i>Purification des formes de FluPol</i>	<i>113</i>
2) <i>Criblage des nanoligands.....</i>	<i>116</i>
<i>Obtention de VHH spécifiques de FluPol</i>	<i>116</i>
<i>Obtention d'αRep spécifiques de FluPol</i>	<i>117</i>
3) <i>Propriétés des VHH et αRep sur l'activité de FluPol.....</i>	<i>119</i>
4) <i>Spécificité des VHH et αRep pour les sous-unités de FluPol</i>	<i>122</i>
<i>Marquage de l'amine terminale du VHH16 avec le fluorophore Alexa488.....</i>	<i>122</i>
<i>Analyse par western blot de la spécificité d'interaction du VHH16.....</i>	<i>123</i>
<i>Analyse de la spécificité du VHH16 par test de complémentation</i>	<i>124</i>
5) <i>Mesure de l'affinité entre le VHH16 et FluPol.....</i>	<i>125</i>
6) <i>Effet du VHH16 sur l'assemblage et le transport du dimère PA-PB1 au noyau</i>	<i>128</i>
<i>Effet du VHH16 sur l'assemblage de FluPol.....</i>	<i>128</i>
<i>Activité du VHH16 sur le transport du dimère PA-PB1 au noyau et son interaction avec l'importine RanBP5.....</i>	<i>129</i>
7) <i>Caractérisation structurale des VHH</i>	<i>133</i>
<i>Purification des VHH</i>	<i>133</i>
<i>Cristallisation des VHH.....</i>	<i>133</i>
<i>Structure du VHH18.....</i>	<i>135</i>
CONCLUSIONS.....	138
CONCLUSION GÉNÉRALE & PERSPECTIVES.....	141
INTERACTIONS ENTRE FLUPOL ET POL II	143
INTERACTION ENTRE FLUPOL ET LE COMPARTIMENT CHROMATINIEN	145
DES NANO-OUTILS POUR DISSEQUER LE FONCTIONNEMENT DE FLUPOL.....	147
MATÉRIELS ET MÉTHODES	149
RÉFÉRENCES	159

Liste des abréviations

°C	Degrès celcius
µm	Micromètre
Å	Ångström
AA	Acide aminé
ADN	Acide désoxyribonucléique
ADNc	ADN complémentaire
ANP32	Acidic nuclear phosphoproteins of 32 kDa
ARN	Acide ribonucléique
ARNe	ARN issu des <i>enhancers</i>
ARNm	ARN messenger
ARNv	ARN génomique viral
AS	Acide sialique
BLI	Bio-layer interferometry (interférométrie de biocouche)
CBC	Cap-binding complex (complexe liant la coiffe des ARNm)
CDR	Complementary determining region (région déterminant la reconnaissance des anticorps)
CFP	Cyan fluorescent protein (protéine fluorescente cyan)
CH	Partie constante des chaînes lourdes d'un anticorps
ChIP-seq	Chromatin immunoprecipitation sequencing (immunoprécipitation de chromatine)
CL	Partie constante des chaînes légères d'un anticorps
CPSF	Cleavage and polyadenylation specificity factor
CTD	Domaine C-terminal de la sous-unité RPB1 de l'ARN-polymérase II
DTT	Dithiothréitol
EEC	Early elongation checkpoint (point de contrôle de début d'élongation)
FluPol	ARN-polymérase des virus Influenza
FT	Flow through (partie de l'échantillon non retenue sur la colonne)
GFP	Green fluorescent protein (protéine fluorescente verte)
HA	Hémagglutinine
HEK 293T	Human embryonic kidney 293 cells containing the SV40 virus T-antigen (cellules rénales embryonnaires 293 contenant l'antigène T du virus SV40)
h.p.i.	Heures post-infection
IAV	Virus Influenza de type A
IBV	Virus Influenza de type B
ICV	Virus Influenza de type C
K _d	Constante de dissociation à l'équilibre
kDa	Kilodalton
LC	Low concentration (échantillon à faible concentration)
Luc	Luciférase
Luc1	Moitié N-terminale de la luciférase
Luc2	Moitié C-terminale de la luciférase
M	Molaire
M1	Protéine de matrice 1
M2	Protéine de matrice 2
mNET-seq	Mammalian native elongation transcript sequencing (séquençage des transcrits en cours d'élongation)

MPT	Modification post-traductionnelle
MW	Molecular weight (échelle de poids moléculaires)
NA	Neuraminidase
NELF	Negative elongation factor
NEP	Nuclear export protein
NES	Nuclear export signal (signal d'export hors du noyau)
ng	Nanogramme
NLR	Normalized luminescence ratio (ratio de luminescence normalisé)
NLS	Nuclear localization signal (signal d'import nucléaire)
nm	Nanomètre
nM	Nanomolaire
NP	Nucléoprotéine
nt	Nucléotides
ORF	Open reading frame (cadre ouvert de lecture)
PA	Polymérase acide
PB1	Polymérase basique 1
PB2	Polymérase basique 2
PEI	Polyéthylèneimine
pH	Potentiel hydrogène
Pol II	ARN-polymérase II
pTEV	Protéase du TEV
RNP	Ribonucléoprotéine
RNPc	Ribonucléoprotéine complémentaire
RNPg	Ribonucléoprotéine génomique
Ser2P	Sérine 2 du CTD phosphorylée
Ser5P	Sérine 5 du CTD phosphorylée
Sng	Surnageant
TEV	Tobacco etch virus (virus de la gravure du tabac)
TSS	Transcription starting site (site d'initiation de la transcription)
TTS	Transcription termination site (site de terminaison de la transcription)
UTR	Untranslated region (région non transcrite)
VH	Partie variable des chaînes lourdes d'un anticorps
VL	Partie variable des chaînes légères d'un anticorps
Wash	Lavage
WT	Wild type (sauvage, natif)
YFP	Yellow fluorescent protein (protéine fluorescente jaune)

INTRODUCTION

I. Caractéristiques des virus Influenza

1) Taxonomie

D'après le comité international de taxonomie des virus (ICTV), les virus Influenza appartiennent au groupe des virus à ARN de polarité négative. Comme partiellement représenté dans la figure 1, ils font partie du règne des *Riboviria*, du phylum *Negarnaviricota*, de la classe des *Insthoviricetes*, de l'ordre des *Articulavirales* et de la famille des *Orthomyxoviridae* comprenant sept genres (les *Alpha-*, *Beta-*, *Gamma-*, *Deltainfluenzavirus*, les *Isavirus*, les *Quaranjavirus* et les *Thogotovirus*) et neuf espèces (les virus Influenza A, B, C et D, les isavirus de saumon, les quaranjavirus Johnston Atoll et Quarafil ainsi que les thogotovirus Dhori et Thogoto) (International Committee on Taxonomy of Viruses (ICTV), 2020).

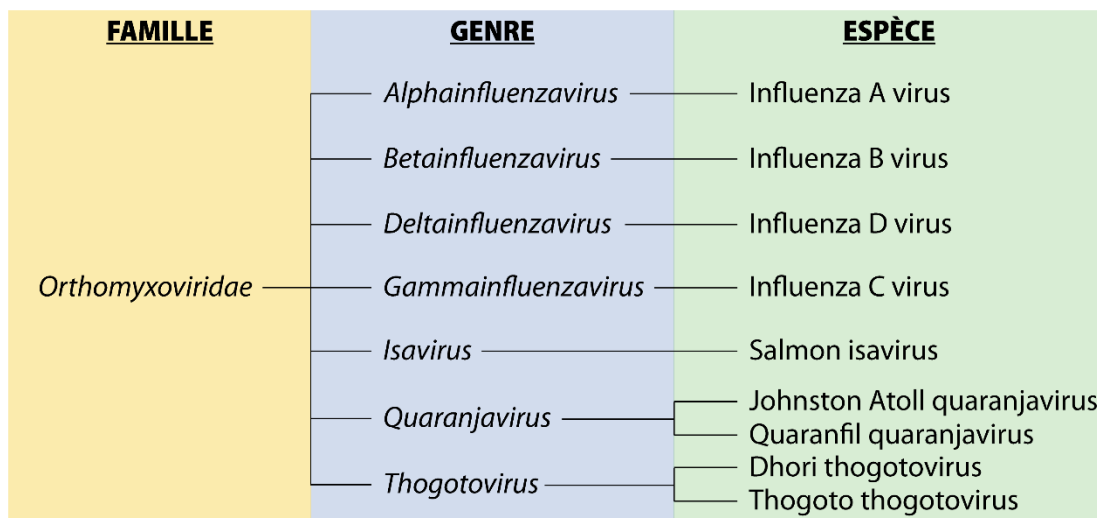


Figure 1 : Arbre phylogénétique de la famille des *Orthomyxoviridae*
D'après la classification de l'ICTV en 2020

Au vue de la grande diversité des virus Influenza, une nomenclature spécifique a été mise en place par l'Organisation mondiale de la santé (OMS) dans les années 80 afin de pouvoir répertorier précisément les souches existantes (World Health Organization (WHO), 1980). Ainsi, chaque souche est répertoriée selon cinq critères : le type d'influenza (A, B, C ou D), l'hôte d'origine pour les souches d'origine animale, l'origine géographique, le numéro d'isolat et l'année (Figure 2). Pour les virus Influenza de type A (IAV), la nomenclature est encore plus détaillée et prend en compte les propriétés antigéniques des deux glycoprotéines de surface : l'hémagglutinine (HA) et la neuraminidase (NA). Dix-huit sous-types de HA (H1 à H18) et onze sous-types de NA (N1 à N11) sont recensés à ce jour. Toutes

les combinaisons possibles entre ces deux glycoprotéines n'ont pas été détectées et certains sous-types sont inféodés à un type d'organisme. À titre d'exemple, les sous-types H17/H18 et N10/N11 n'ont pour l'instant été observés que chez les chauves-souris (Tong *et al*, 2012, 2013). Seuls les sous-types H1, H2, H3, N1 et N2 sont connus pour infecter l'Homme sur un mode épidémique.

TYPE / HÔTE D'ORIGINE / ORIGINE GÉOGRAPHIQUE / NUMÉRO DE SOUCHE / ANNÉE (SOUS-TYPE HXNX)

A / Artibeus lituratus / Brazil / 2301 / 2012 (H18N11)

B / Swine / Hong Kong / 2850 / 2016

A / Adachi / 2 / 1957 (H2N2)

Figure 2 : Taxonomie des virus Influenza déterminée par l'OMS

Trois exemples de nomenclature de souches sont donnés, la première provenant d'une chauve-souris, la seconde d'un porc et la troisième d'origine humaine.

Cette étude étant majoritairement consacrée aux IAV, le reste de ce manuscrit portera principalement sur ce type d'Influenza.

2) Structure de la particule virale

Les IAV sont classiquement représentés comme une particule de forme sphérique de 120 nm de diamètre. Cette forme correspond aux formes du virus cultivé en laboratoire. Cependant, comme illustré en figure 3 A et B, les IAV sont en réalité pléomorphes (Chambers *et al*, 1943) et peuvent adopter différentes formes, notamment une forme de bacille ou filamenteuse de 80 à 100 nm de diamètre pouvant atteindre dans certains cas plusieurs dizaines de μm de long ! Toutefois, l'organisation structurale de ces particules reste sensiblement la même.

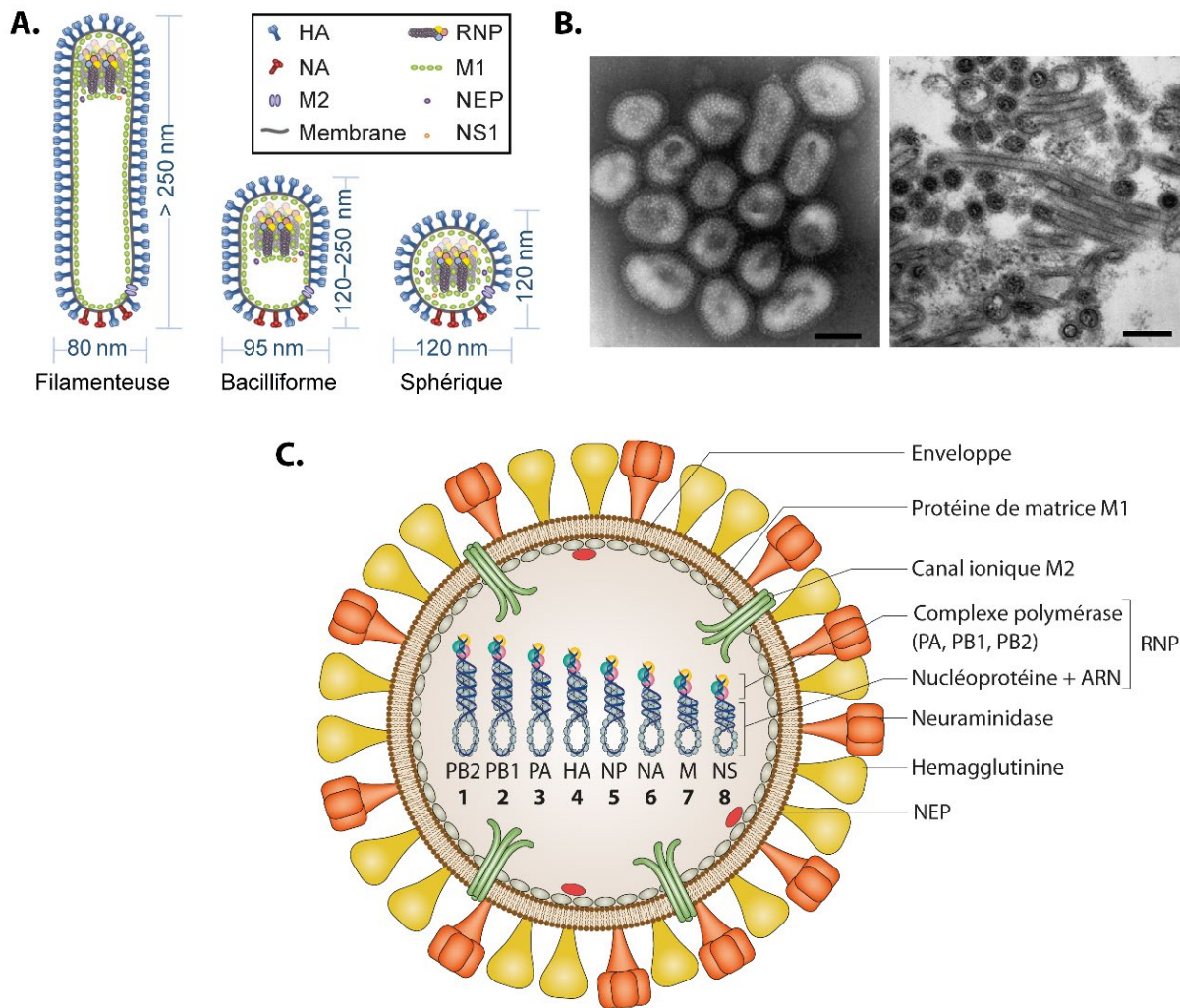


Figure 3 : Morphologie des particules virales des virus Influenza

A. Schématisation des différentes formes que les particules virales peuvent adopter (adapté de (Dadonaite *et al*, 2016)). **B.** Photo de particules virales par microscopie électronique (Nakajima *et al*, 2010). La photo de gauche est issue d'un échantillon de virus A/H1N1pdm09 cultivé en laboratoire alors que le même virus issu d'une coupe de poumon est présenté à droite. Les barres d'échelle représentent 100 et 200 nm respectivement. **C.** Représentation schématique de l'organisation d'une particule virale (adapté de (Krammer *et al*, 2018))

Les IAV sont des virus enveloppés d'une membrane lipidique provenant de la membrane plasmique de la cellule hôte, acquise au cours du bourgeonnement du virus (Figure 3C). À sa surface sont présentes les deux glycoprotéines : HA qui se lie au récepteur cellulaire et la NA qui permet le détachement et la libération des particules virales nouvellement produites. L'hémagglutinine est cinq fois plus présente que la neuraminidase (Getie-Kehtie *et al*, 2013; Moulès *et al*, 2011). La protéine de matrice M1 tapisse l'intérieur de la membrane virale et joue un rôle majeur dans la morphologie de la particule puisque c'est elle qui détermine la forme sphérique ou filamenteuse (Elleman & Barclay, 2004; Vijaykrishnan *et al*, 2013). M1 est l'une des protéines les plus abondantes dans la particule. La protéine M2 est enchâssée dans l'enveloppe et forme un canal ionique. Bien que non essentielle, M2

permet au génome viral d'être libéré dans le cytoplasme. Les protéines PA (polymérase acide), PB1 (polymérase basique 1) et PB2 (polymérase basique 2) constituent les trois sous-unités de l'ARN polymérase ARN-dépendante (FluPol) qui réplique et transcrit le génome viral (ARNv) tandis que la nucléoprotéine NP est liée au génome afin de le protéger. Avec l'ARNv, ces quatre protéines forment une ribonucléoprotéine (RNP). Enfin, la protéine NEP, associée à M1, permet l'export des RNP hors du noyau.

3) Organisation du génome

Les IAV possèdent un génome ARN segmenté de 13,5 kb. Pour être fonctionnelle, chaque particule virale contient ainsi huit segments d'ARN génomique (McGeoch *et al*, 1976) simple brin de polarité négative longs de 890 à 2341 nucléotides (Palese & Shaw, 2007), pouvant coder jusqu'à 19 protéines (Vasin *et al*, 2014) grâce à différents mécanismes (Figure 4B). Ces derniers permettent de coder un plus grand nombre de protéines virales tout en conservant une taille de génome relativement faible, permettant ainsi de tenir dans la particule virale. La segmentation du génome apporte à la fois des avantages et des inconvénients au virus. En effet, la segmentation apporte un avantage évolutif aux IAV car il permet des réassortiments entre différentes souches virales infectant une même cellule hôte (cf section II). Ce phénomène évolutif est appelé cassure antigénique (ou « *antigenic shift* ») et peut favoriser les franchissements de barrière d'espèce en cas d'infection entre souches d'origines animales différentes.

Les segments génomiques sont numérotés de 1 à 8 par ordre de taille décroissant. Tous les segments ont une même organisation générale avec des séquences non traduites (UTR) de part et d'autre du ou des cadre(s) ouvert(s) de lecture (ORF) (Figure 4A). La taille de ces séquences varie d'un segment à l'autre mais les 12 premiers (U12) et 13 derniers (U13) nucléotides correspondant aux séquences promotrices de FluPol sont conservés entre segments mais également entre souches virales (Desselberger *et al*, 1980; Robertson, 1979).

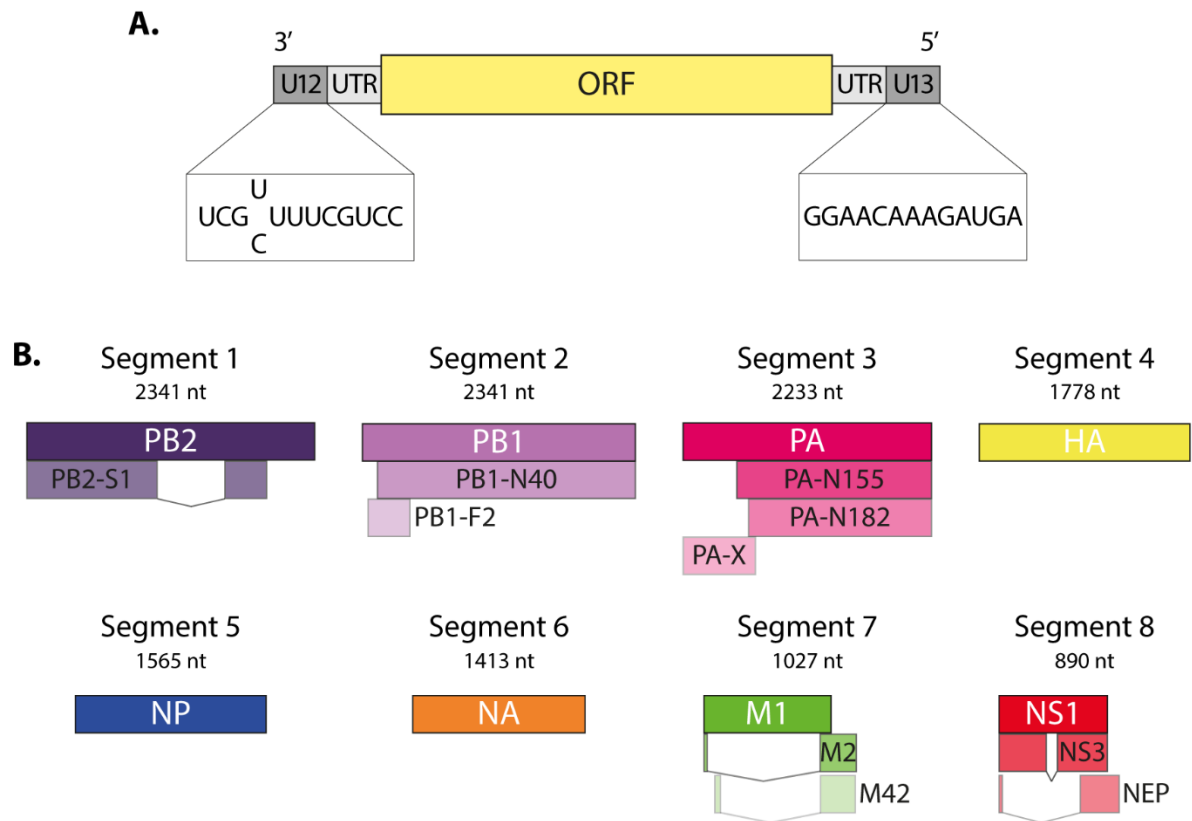


Figure 4 : Organisation génomique des IAV

A. Schéma de l'organisation des segments viraux. **B.** Schéma récapitulatif du protéome des IAV. La taille des segments est indiquée en nucléotides. Adapté de (ViralZone).

Ainsi, le segment 1 code pour la sous-unité PB2 de FluPol, mais également pour la protéine PB2-S1 (Yamayoshi *et al*, 2016) grâce à un mécanisme d'épissage alternatif de son ARNm (Figure 4B). Cette dernière inhibe la réponse immunitaire interféron mais pourrait également inhiber l'activité de FluPol en interagissant de façon compétitive avec PB1 (Yamayoshi *et al*, 2016). Le segment 2 code pour la sous-unité PB1 mais également pour deux autres protéines grâce à l'existence de cadres de lecture alternatifs, PB1-F2 et PB1-N40. PB1-N40 n'est pas essentielle à la viabilité du virus et sa fonction n'est pas encore connue (Wise *et al*, 2009). En revanche, PB1-F2, la plus petite protéine du virus, est un facteur de virulence bien qu'elle ne soit pas présente chez toutes les souches (Chen *et al*, 2001). Le segment 3 est celui qui code pour le plus grand nombre de protéines. Tout d'abord, ce segment code pour la troisième sous-unité de FluPol, PA. Ce segment code également pour les protéines PA-N155 et PA-N182 par un mécanisme d'initiation alternative de la transcription (Muramoto *et al*, 2013). Leur rôle précis n'est pas encore clairement défini mais elles semblent jouer un rôle dans la réplication virale. Enfin, un décalage ribosomique permet la synthèse d'une dernière protéine appelée PA-X qui a un rôle de facteur de virulence et de répression de l'expression des gènes de l'hôte (Jagger *et al*, 2012). Le segment 4 code pour la HA, le segment 5 pour la NP et le 6 pour la NA. Le segment 7 code pour trois protéines différentes. Premièrement, ce segment code pour la protéine de matrice M1. Par un

mécanisme d'épissage alternatif, deux autres protéines sont également produites : le canal ionique M2 ainsi que la protéine M42 qui pourrait assurer une fonction similaire à M2 (Wise *et al*, 2012). Enfin, le segment 8 code pour la protéine non structurale NS1 aux rôles multiples (discutés dans les revues (Hale *et al*, 2008; Rosário-Ferreira *et al*, 2020)), facteur de virulence qui inhibe la réponse immunitaire de l'hôte ainsi que le facteur d'export nucléaire NEP par un processus d'épissage alternatif (Lamb & Choppin, 1979). Une troisième protéine appelée NS3 peut également être exprimée par épissage alternatif bien qu'elle ne soit pas retrouvée dans toutes les souches (Selman *et al*, 2012). Son rôle n'est pas encore clairement identifié mais il semblerait que cette protéine pourrait jouer un rôle dans l'adaptation à l'hôte et le franchissement de barrière d'espèce. Des analyses informatiques ont également identifié de nouvelles ORF dans le sens inverse des ORF précédemment décrites. L'ORF la plus décrite est présente sur le segment 8 qui pourrait coder une autre protéine appelée NSP8 (*Negative strand polypeptide*) (Clifford *et al*, 2009; Zhirnov *et al*, 2007). Cette protéine a pu être exprimée *in silico* (Zhirnov *et al*, 2017) ou *in cellulo* (Hickman *et al*, 2018; Zhirnov & Klenk, 2010) mais n'a, à ce jour, jamais été directement détectée en contexte infectieux (Hickman *et al*, 2018; Zhirnov & Isaeva, 2019; Zhong *et al*, 2003). D'autres ORF pourraient également exister pour les segments 1, 2, 3, 5 et 7 (Zhirnov, 2020).

4) Structure de la RNP

Comme décrit brièvement précédemment, les RNP virales, unités de répllication du virus, sont constituées des segments génomiques du virus, associés à un grand nombre de copies de NP et fixés par leurs extrémités 5' et 3' strictement conservées à un exemplaire de FluPol afin de le protéger (Figure 5A). Une NP recouvre environ 24 nucléotides avec une zone d'interaction directe de 12 nucléotides (Williams *et al*, 2018). Les NP sont liées à l'ARN de façon non uniforme (Lee *et al*, 2017; Williams *et al*, 2018) et sans spécificité de séquence. En effet, cette interaction se fait au niveau des groupements phosphates du squelette de l'ARN viral (Baudin *et al*, 1994; Scholtissek & Becht, 1971) ce qui permettrait aux bases azotées de rester accessibles et fonctionnelles. Certaines régions de l'ARNv libres forment ainsi des structures secondaires protubérantes. En effet, l'interaction entre ces dernières pourrait notamment jouer un rôle dans la sélection des segments lors de l'emballage. Cependant, la fixation de NP le long de l'ARN n'est pas aléatoire (Lee *et al*, 2017; Le Sage *et al*, 2018). Chaque segment possède ainsi un profil spécifique de fixation de NP, qui est partiellement conservé chez les IAV (mais non conservé entre IAV et IBV) (Le Sage *et al*, 2018).

Les RNP présentent une structure hélicoïdale d'un diamètre fixe de 12 nm mais de longueur variable (30 à 120 nm) en fonction des segments génomiques (Compans *et al*, 1972; Sugita *et al*, 2013; Arranz *et al*, 2012; Moeller *et al*, 2012). Les extrémités 3' et 5' des segments génomiques sont liées à FluPol *via* les motifs U12 et U13 (Fodor *et al*, 1994, 1995, 1993; Pritlove *et al*, 1995). L'ARN de ces régions n'est donc pas lié à NP mais s'apparie entre elles pour former une structure double brin (Fodor *et al*, 1993; Hsu *et al*, 1987; Reich *et al*, 2014). Le reste de l'ARN, lié à NP, forme une structure hélicoïdale flexible dont la cohésion est assurée par NP avec une boucle à l'autre extrémité de la RNP (Compans *et al*, 1972).

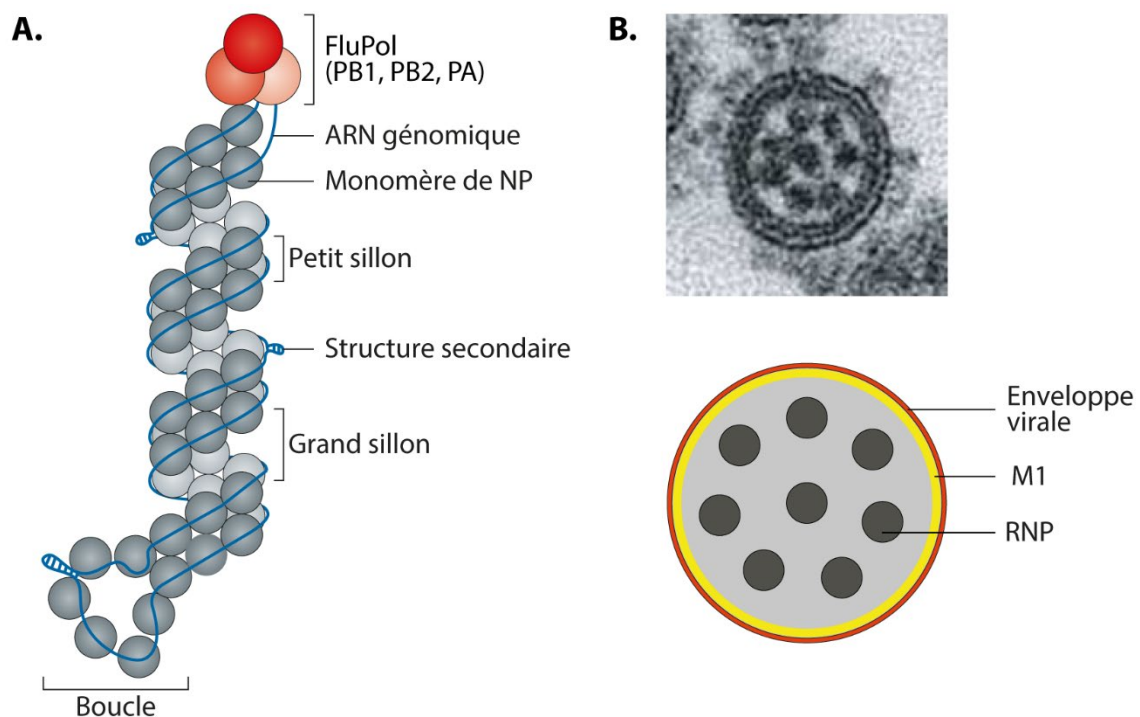


Figure 5 : Organisation des RNP

A. Schéma de la structure d'une RNP. Adapté de (Eisfeld *et al*, 2015).

B. Photo en microscopie électronique (haut) d'une coupe transversale de particule virale de IAV souche A/WSN/1933(H1N1) (Noda *et al*, 2006) et schéma (bas) de la répartition « 7+1 » des RNP au sein de la particule virale. La barre d'échelle représente 50 nm.

Une particule virale fonctionnelle contient donc huit RNP en son sein, structures allongées et parallèles entre elles (Arranz *et al*, 2012; Moeller *et al*, 2012). En effet, les RNP ne sont pas « libres » dans la particule mais sont au contraire ordonnées selon un assemblage dit « 7+1 » avec une RNP centrale autour de laquelle seraient disposées les sept autres RNP (Figure 5B). Cette organisation structurale semblerait maintenue grâce à un réseau d'interactions physiques entre les RNP centrale et périphériques (Noda *et al*, 2012). Cependant, toutes les particules virales ne comprennent pas les huit RNP. En effet, environ 20% des virions produits (proportion variant en fonction de la souche virale

considérée) ne contiennent que 5 à 7 RNP (Nakatsu *et al*, 2016). Le rôle fonctionnel de ces particules reste encore à être élucidé (pour revue (Diefenbacher *et al*, 2018)).

II. Epidémiologie des virus Influenza

Les virus Influenza sont les agents pathogènes responsables de la grippe chez l'Homme, infectant jusqu'à 10% de la population mondiale par an et causant 5 millions de cas graves et jusqu' à 650 000 décès par an (Iuliano *et al*, 2018). Provoquant fièvre, toux, maux de têtes, douleurs musculaires et articulaires ou encore troubles digestifs, l'infection se résorbe souvent au bout d'une semaine. Néanmoins, certaines personnes à risque comme les personnes âgées ou les femmes enceintes peuvent développer des formes plus graves de la maladie pouvant mener au décès.

1) **Ecologie**

Les virus Influenza ont un spectre d'hôte extrêmement large puisqu'ils peuvent infecter aussi bien la faune sauvage que les animaux d'élevage ou l'Homme (Figure 6). Ces virus peuvent donc être zoonotiques. Les oiseaux sauvages aquatiques, particulièrement les canards, constituent le réservoir des IAV. Ayant la capacité de se multiplier dans un grand nombre d'espèces aviaires et de mammifères, les IAV peuvent parfois acquérir la capacité à franchir les barrières d'espèce et représentent donc une menace constante pour l'Homme. Infectant l'épithélium respiratoire, la transmission des IAV se fait par voie aérienne chez l'homme mais également par les fomites (Asadi *et al*, 2020). Chez les oiseaux, le virus se multiplie dans l'épithélium du tractus digestif. La transmission est donc oro-fécale chez eux (Webster, 1997). Seuls les IAV, les IBV et les ICRV ont été répertoriés comme pouvant infecter l'Homme. Dans le cas des IDV, seul des anticorps ont été détectés chez l'Homme (White *et al*, 2016).

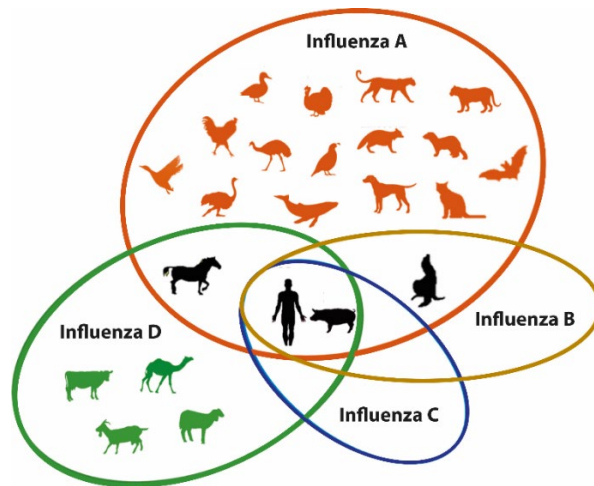


Figure 6 : Représentation de la diversité du spectre d'hôte des virus Influenza.
Adapté de (Kuchipudi & Nissly, 2018).

Ce sont principalement les IAV et IBV qui circulent le plus chez l'Homme, les IAV étant les plus fréquents avec les souches H1N1 et H3N2. La plupart du temps, l'épidémie de grippe suit un cycle saisonnier, étant peu présente dans l'hémisphère Nord en été et atteignant un pic épidémique en plein hiver (souvent en janvier ou février). L'inverse est observé dans l'hémisphère Sud. En revanche, le virus circule de façon continue dans les zones tropicales et équatoriales (Viboud *et al*, 2006). Les virus Influenza circulent donc de façon continue à l'échelle de la planète. Le décalage du pic épidémique entre les hémisphères Nord et Sud est exploité pour prédire les souches qui circuleront la saison suivante, permettant la confection de vaccins adaptés commercialisés tous les ans avant la saison hivernale. Leur efficacité peut se révéler médiocre si les virus circulants ont suffisamment divergé des souches ayant servies à l'élaboration des vaccins. Plusieurs inhibiteurs dirigés contre différentes protéines du virus ont été développés mais de plus en plus de souches acquièrent des résistances. Il est donc nécessaire de chercher continuellement de nouveaux inhibiteurs ciblant les virus Influenza.

2) Evolution

L'évolution des souches d'Influenza peut être très rapide et est médiée par deux mécanismes différents : la dérive antigénique (ou *antigenic drift*) et la cassure antigénique (ou *antigenic shift*).

a. Dérive antigénique

Le mécanisme de dérive antigénique est produit par la polymérase virale. En effet, cette ARN-polymérase ne présente pas d'activité de relecture et de correction des erreurs lors de l'incorporation des nucléotides comme certaines polymérases. En cas d'erreur d'incorporation, une nouvelle mutation

apparaît donc, qui peut être silencieuse ou non. Le taux de mutation observé varie en fonction des segments génomiques en fonction de la pression de sélection exercée mais peut atteindre une mutation incorporée par cycle de réplication (Webby & Webster, 2001; Koelle *et al*, 2006; Fitch *et al*, 1991). En cas de mutation non silencieuse, un avantage évolutif peut être apporté bien que dans la majorité des cas, celle-ci aura plutôt un effet délétère (Wille & Holmes, 2020). Ce phénomène d'évolution est donc relativement lent. Toutefois, l'apparition d'une simple mutation peut parfois suffire pour échapper à la reconnaissance des anticorps (Zambon, 1999) et entraîner des résistances aux molécules antivirales (Treanor, 2004).

b. Cassure antigénique

Comme beaucoup de virus à ARN, les virus Influenza franchissent fréquemment les barrières d'espèce (Steel & Lowen, 2014; Lowen, 2018, 2017). Cependant, dans la plupart des cas, ce phénomène reste sporadique (cul-de-sac épidémiologique) et ne conduit pas à une diffusion épidémique. Contrairement à la dérive antigénique, le phénomène de cassure antigénique est rapide et brutal. Il est dû à la nature segmentée du génome des IAV. Ainsi, en cas de co-infection d'une même cellule par deux souches différentes, des réassortiments entre segments génomiques peuvent avoir lieu, menant à la formation d'une nouvelle particule virale contenant du matériel génétique en provenance des deux souches parentales (Zambon, 1999). Cependant, ces réassortiments sont des phénomènes complexes et régulés. Toutes les combinaisons ne sont pas possibles en raison des mécanismes de sélection des segments lors de l'empaquetage (White & Lowen, 2018). Ce mécanisme peut ainsi se produire entre deux souches d'Influenza d'origine différente au sein d'un hôte intermédiaire tel que le porc. De cette façon, des réassortiments peuvent avoir lieu entre un virus aviaire et un virus d'origine porcine par exemple comme ce fut le cas en 2009 avec la pandémie de grippe porcine (Smith *et al*, 2009). La cassure antigénique peut donc faciliter le franchissement de barrière d'espèce en favorisant l'adaptation à un nouvel hôte (Garten *et al*, 2009). Ce franchissement de barrière d'espèce peut avoir de graves conséquences puisque le virus circule alors dans une population ne possédant pas d'immunité contre cette souche.

3) Pandémies

Selon l'OMS, un virus pandémique correspond à un virus émergent circulant dans une population naïve d'un point de vue immunologique, en dehors du schéma épidémique habituel. Ces virus ont en général une grande capacité de transmission qui leur permet une propagation rapide à l'échelle planétaire. Ils doivent également causer une maladie sévère avec une forte morbidité et/ou

mortalité. Comme mentionné précédemment, les virus Influenza peuvent occasionnellement devenir pandémiques suite à un réassortiment et se propager hors du schéma saisonnier habituel. Plusieurs pandémies ont déjà eu lieu depuis le début du XX^e siècle (Figure 7), toutes causées par une souche de IAV. La première et la plus meurtrière a été la pandémie de grippe dite « espagnole » en 1918 avec plus de cinquante millions de morts estimés. Cette souche est l'ancêtre de tous les IAV pandémiques et endémiques ayant circulés jusqu'à ce jour (Morens *et al*, 2009). Ces virus pandémiques peuvent acquérir une virulence exacerbée qui s'atténuera naturellement après avoir circulé dans la population, et finira par revêtir un caractère saisonnier. De par cette virulence et leur capacité à infecter l'Homme, la présence de virus aviaires de type H5N1 et H7N9 est systématiquement recherchée dans les élevages aviaires. Ces virus font donc l'objet d'une surveillance constante par les instances nationales et internationales de santé, que ce soit chez l'Homme ou chez l'animal.

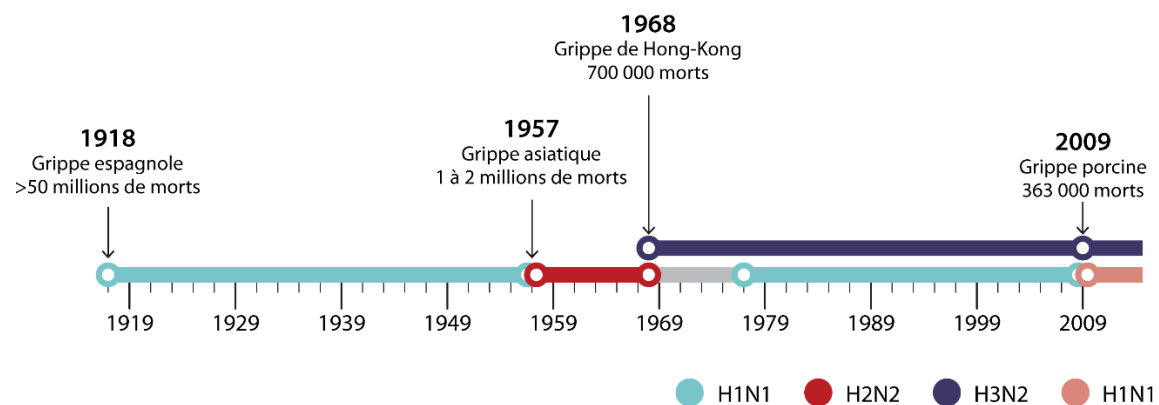


Figure 7 : Chronologie des pandémies de grippe au cours de l'Histoire.

Les zones colorées représentent les années durant lesquelles la souche a circulé lors d'épidémies saisonnières. Adapté de (Taubenberger *et al*, 2019).

Les IAV représentent donc un coût économique et sanitaire important à l'échelle planétaire. En effet, rien qu'aux Etats-Unis, le coût moyen économique annuel estimé pour le système de santé et la société américaine est de 11,2 milliards de dollars (Putri *et al*, 2018).

III. Cycle viral

Contrairement à la plupart des autres virus à ARN simple brin dont le cycle se déroule entièrement dans le cytoplasme, les virus Influenza répliquent et transcrivent leur génome dans le noyau de la cellule hôte (Figure 8). Cette particularité, qui adjoint une difficulté supplémentaire pour le virus dont les RNP doivent parvenir à franchir l'enveloppe nucléaire, a également plusieurs avantages (Hutchinson & Fodor, 2012). Premièrement, répliquer le génome dans le noyau de la cellule

hôte permet une encapsidation par la NP avant l'arrivée du génome au cytoplasme, ce qui évite une détection par le système immunitaire (Rehwinkel *et al*, 2010). De plus, pour la transcription du génome, les IAV volent la coiffe des ARNm cellulaires par un phénomène dit de « *cap-snatching* » (détaillé en section IV.4). Réaliser la transcription dans le noyau permet une meilleure efficacité du vol de coiffe mais également une prise en charge des ARNm viraux par les machineries cellulaires d'épissage ou d'export nucléaire (Engelhardt & Fodor, 2006). Enfin, cette proximité avec la chromatine permet un « *shut-off* » performant de la cellule hôte notamment via la dégradation de l'ARN-polymérase II (Pol II) (Rodriguez *et al*, 2007; Vreede & Fodor, 2010; Vreede *et al*, 2010).

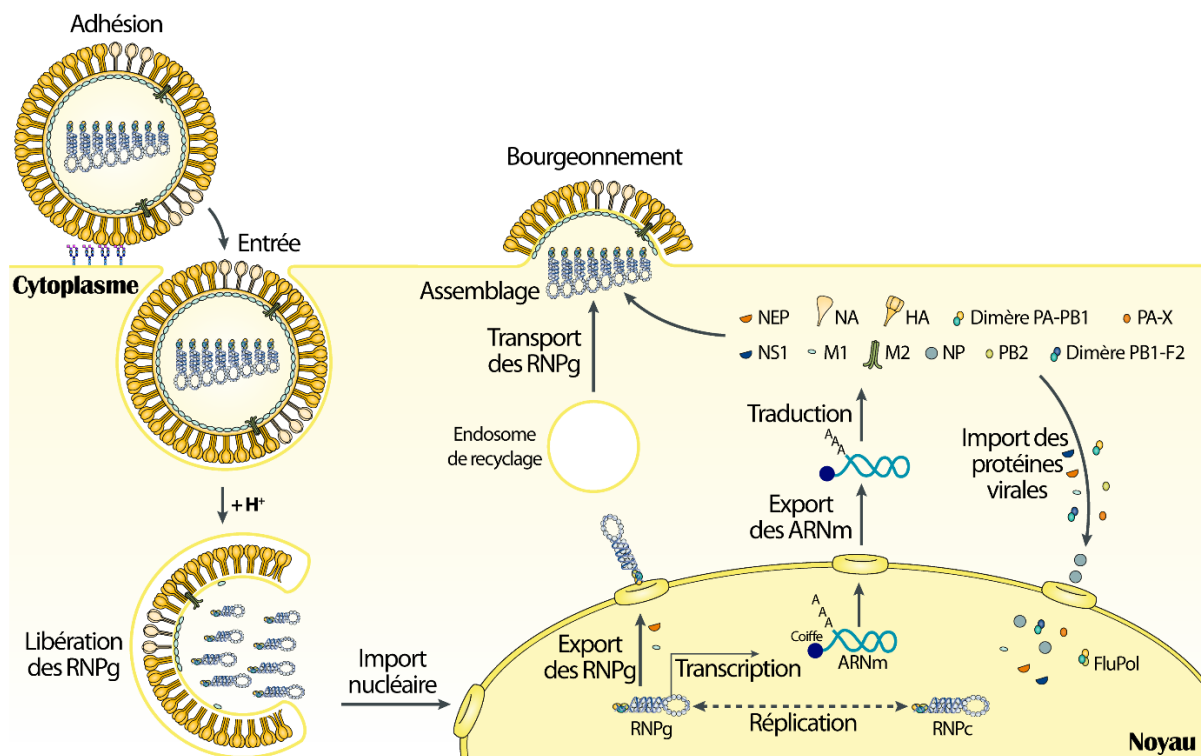


Figure 8 – Représentation schématique du cycle viral des IAV

Adapté de (Krammer *et al*, 2018).

RNPg = RNP génomique i.e. constituée d'un ARNv. RNPc = RNP complémentaire i.e. constituée d'un segment d'ARN complémentaire à l'ARNv

1) Adhésion et entrée du virus dans la cellule hôte

Dans un premier temps, la particule virale se fixe à la surface de la cellule avant de pouvoir y pénétrer. C'est l'hémagglutinine (HA) présente en surface de la particule virale qui rend possible cette fixation en reconnaissant les acides sialiques présents sur les glycolipides et glycoprotéines exprimés à la membrane plasmique (Gottschalk, 1959; Wiley & Skehel, 1987). HA reconnaît plus spécifiquement l'acide N-acétylneuraminique (Neu5Ac). La HA des virus Influenza humains reconnaît préférentiellement les Neu5Ac liés en position α -2,6 abondamment présent sur les cellules épithéliales du tractus

respiratoire supérieur (Connor *et al*, 1994). La HA des virus aviaires interagit plutôt avec les Neu5Ac liés en position α -2,3 présents de façon prépondérante dans le tractus digestif des oiseaux et le tractus inférieur chez l'Homme. Les acides sialiques α -2,3 et α -2,6 sont abondamment exprimés dans la trachée chez le porc. Ce dernier peut donc servir d'hôte aux virus humains et aviaires, menant parfois à des réassortiments entre deux virus d'origine différente (cf section II.2.b). La fixation de HA à son récepteur déclenche le processus d'entrée de la particule dans la cellule hôte, principalement par la voie d'endocytose clathrine-dépendante (Wang & Jiang, 2009; Matlin *et al*, 1981). Cependant, d'autres mécanismes ont été décrits comme voies alternatives possibles : l'endocytose clathrine-indépendante (cavéoles) (Nunes-Correia *et al*, 2004) ou la macropinocytose. Cette dernière semble particulièrement privilégiée par les particules filamenteuses (Rossman *et al*, 2012). De cette façon, la particule virale est internalisée dans la voie des endosomes.

2) Libération des RNP dans le cytoplasme

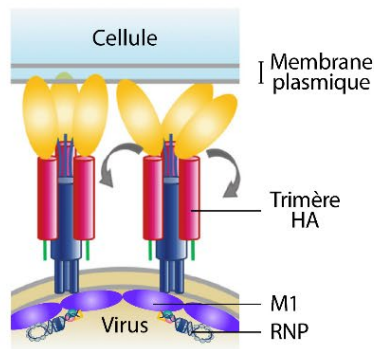
La libération du génome viral dans le cytoplasme de la cellule est une étape indispensable à la réplication virale. Au cours de sa progression vers la région périnucléaire, l'endosome mature et s'acidifie graduellement (le pH baissant de 6,5 à 4,5). Cette acidification a plusieurs conséquences sur la particule virale.

D'un côté, elle entraîne plusieurs changements de conformation de la protéine HA (Stegmann *et al*, 1990; White & Wilson, 1987; Doms & Helenius, 1986; Skehel *et al*, 1982; Russell, 2021). Premièrement, les têtes globulaires vont s'effacer afin de pouvoir permettre l'exposition de son peptide de fusion (Figure 9). Ce dernier s'ancre dans la membrane de l'endosome (Harter *et al*, 1989; Brunner, 1989). La HA réalise ensuite un changement de conformation en fermeture éclair qui lui permet de se replier sur elle-même. Ce nouveau changement de conformation a pour conséquence le rapprochement des deux membranes (virale et endosomale). Une fois accolées, les membranes fusionnent progressivement pour former un pore par lequel les RNP peuvent rejoindre le cytoplasme.

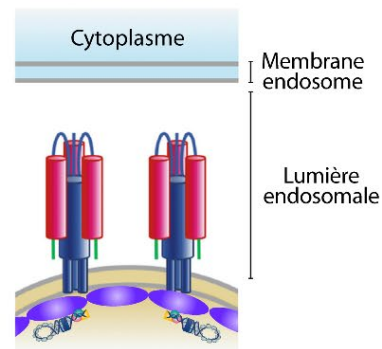
D'autre part, le pH atteint au niveau de l'endosome tardif (pH d'environ 5) provoque l'ouverture du canal ionique M2 et l'influx de protons à l'intérieur de la particule virale (Cady *et al*, 2009). Cette acidification interne provoque un changement de conformation de M1, qui maintient normalement les RNP en position (Bui *et al*, 1996). L'interaction M1 - RNP étant abolie, les RNP sont donc libres dans la particule virale.

L'ensemble de ces événements concomitants permet ainsi aux RNP de se détacher de la particule virale et d'être libérées dans le cytoplasme de la cellule.

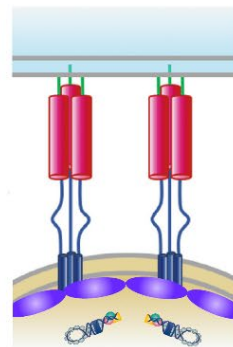
1. Liaison HA-AS



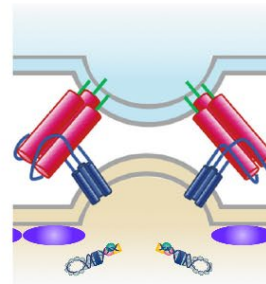
2. Pré-fusion



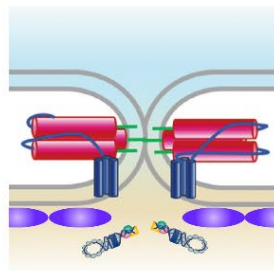
3. Ancrage du peptide fusion



4. Rapprochement des membranes



5. Hémifusion



6. Formation du pore

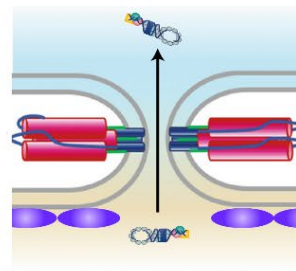


Figure 9 : Illustration du mécanisme de fusion membranaire entre l'endosome et l'enveloppe virale médié par HA

1. Le trimère de HA se fixe à la membrane plasmique via les AS. 2. L'acidification progressive de l'endosome entraîne un changement de conformation de la HA : les têtes globulaires (en jaune sur 1.) se rabattent. 3. Le changement de conformation de HA se poursuit avec le déploiement du peptide de fusion qui s'ancre dans la membrane de l'endosome. Parallèlement, l'ouverture de M2 acidifie l'intérieur de la particule et la liaison M1-RNP est rompue. 4. La HA se replie sur elle-même par un mécanisme en fermeture éclair pour rapprocher les membranes virale et endosomale. 5. Les deux membranes commencent à fusionner. 6. La fusion des membranes est complète : un pore est formé. Les RNP sont libérées dans le cytoplasme. Adapté de (OuYang *et al*, 2018).

3) Import des RNP au noyau

Les RNP relarguées dans le noyau sont encore associées entre elles, notamment via la protéine M1. En 2019, l'importine- β TNPO1 a été identifiée comme nouvel interactant de M1. Cette interaction dans le cytoplasme permet ainsi de retirer les dernières M1 encore fixées aux RNP mais également de désassembler les RNP, entraînant leur dissociation (Miyake *et al*, 2019).

Une fois les RNP libres dans la région périnucléaire, elles sont prises en charge par le système d'import nucléaire afin d'entrer au noyau pour y initier les étapes de transcription / réplication. En effet, les RNP sont trop volumineuses pour pouvoir diffuser librement au travers de l'enveloppe nucléaire. Le signal d'adressage des RNP au noyau (NLS) est porté par l'extrémité N-ter des NP associées au génome (O'Neill *et al*, 1995). Ce NLS est reconnu par une importine- α (Arranz *et al*, 2012; Moeller *et al*, 2012)(Figure 10). En 2016, la protéine chaperonne Hsp40 a été identifiée comme jouant un rôle dans l'import des RNP au noyau (Batra *et al*, 2016). Son interaction avec NP est indispensable pour à une bonne interaction NP – importine- α . L'importine- α est ensuite prise en charge par une importine- β qui va permettre le transfert du complexe au noyau en passant par le complexe du pore nucléaire. Une fois dans le noyau, l'importine- β se dissocie du complexe par l'intervention de la protéine Ran-GTP.

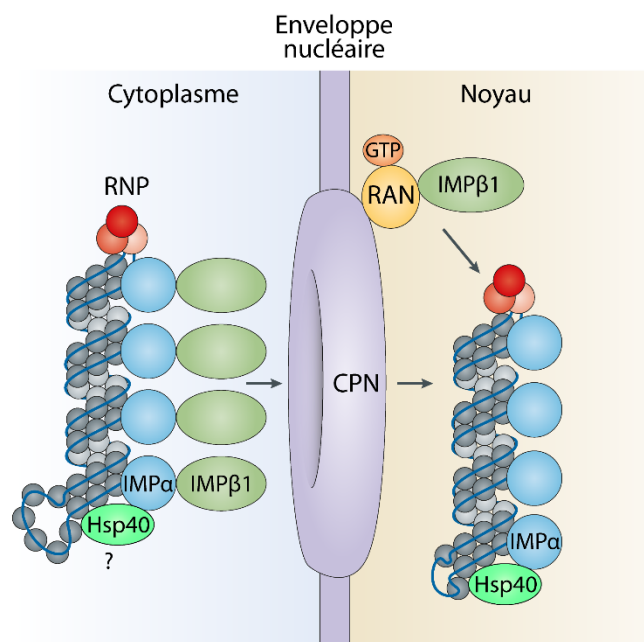


Figure 10 : Modélisation de l'import des RNP au noyau
Adapté de (Eisfeld *et al*, 2015).

4) Transcription et réplication du génome

Afin de pouvoir répliquer son génome à grande échelle, le virus a besoin de produire certaines protéines essentielles à cette synthèse. C'est pourquoi, la toute première étape dans le noyau est la transcription primaire, réalisée par les ARN-polymérase virales liées aux RNP. Les ARNm produits, possédant une coiffe et une queue polyA, sont ensuite exportés hors du noyau via la voie TAP/p15 (Wang *et al*, 2008; York & Fodor, 2013) où ils sont traduits par la machinerie de traduction de la cellule hôte. De cette façon, le complexe polymérase et NP néo-synthétisés ainsi que les protéines M1, NS1 et NEP sont importés au noyau afin de répliquer efficacement le génome. Au cours de la réplication du génome (RNPg), FluPol synthétise un génome viral complémentaire (RNPc), de polarité positive, qui sert ensuite de matrice pour la synthèse des nouveaux segments viraux (figure 11). Les ARN viraux n'étant jamais nus, leur encapsidation par NP et FluPol a lieu simultanément à leur synthèse pour les protéger d'éventuelles dégradations. Une fois les segments génomiques répliqués, ceux-ci peuvent en principe à nouveau servir de matrice pour la transcription bien qu'il ait été observé une diminution graduelle de la transcription virale au cours de l'infection (Shapiro *et al*, 1987).

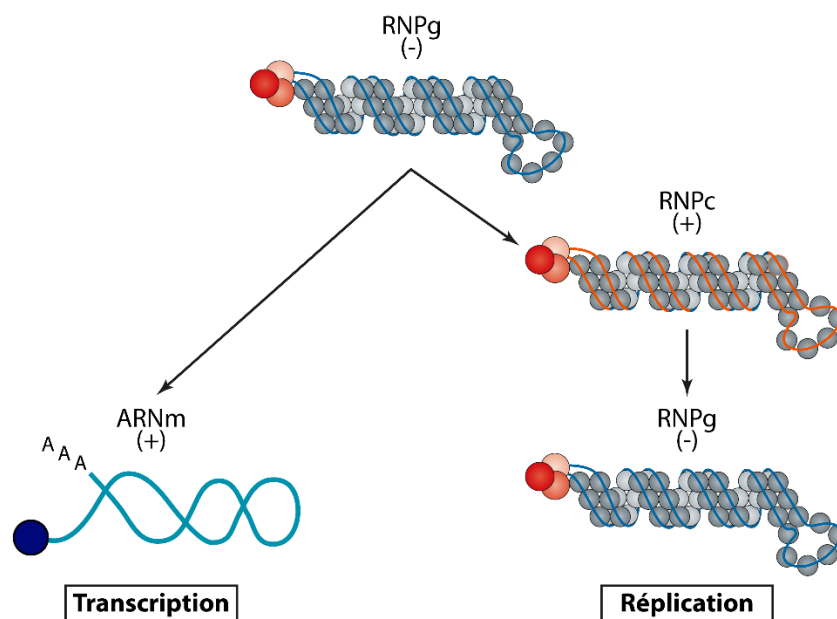


Figure 11 : Comparaison de la réplication versus la transcription du génome viral
La polarité des ARN est indiquée entre parenthèses.

5) Export des RNP

Avant leur export, les RNPg néosynthétisées sont retrouvées associées à la chromatine dans la région périnucléaire (Chase *et al*, 2011; Takizawa *et al*, 2006; Ma *et al*, 2001) grâce à une interaction entre NP et la nucléoline (Terrier *et al*, 2016). Pour les transporter au cytoplasme afin qu'elles soient

incorporées dans les particules virales en formation, les RNP néo-formées utilisent la voie classique d'export nucléaire *via* CRM1 (ou exportine1) (figure 12). L'export des RNPg implique également M1 qui interagit avec les RNP (Noton *et al*, 2007; Bui *et al*, 2000; Sakaguchi *et al*, 2003) et NEP qui possède un signal d'export (NES) reconnu par CRM1 (Neumann *et al*, 2000; O'Neill *et al*, 1998). L'export des RNP nécessite donc la formation d'une chaîne RNP – M1 – Nep –CRM1 (Akarsu *et al*, 2003; Brunotte *et al*, 2014). En effet, aucune interaction directe entre les RNP et CRM1 n'a été démontrée en contexte infectieux. CRM1 est en association avec la GTPase RAN. Une fois le complexe au cytoplasme, la protéine RANGAP hydrolyse le GTP en GDP ce qui permet la dissociation du complexe et le relargage des RNP.

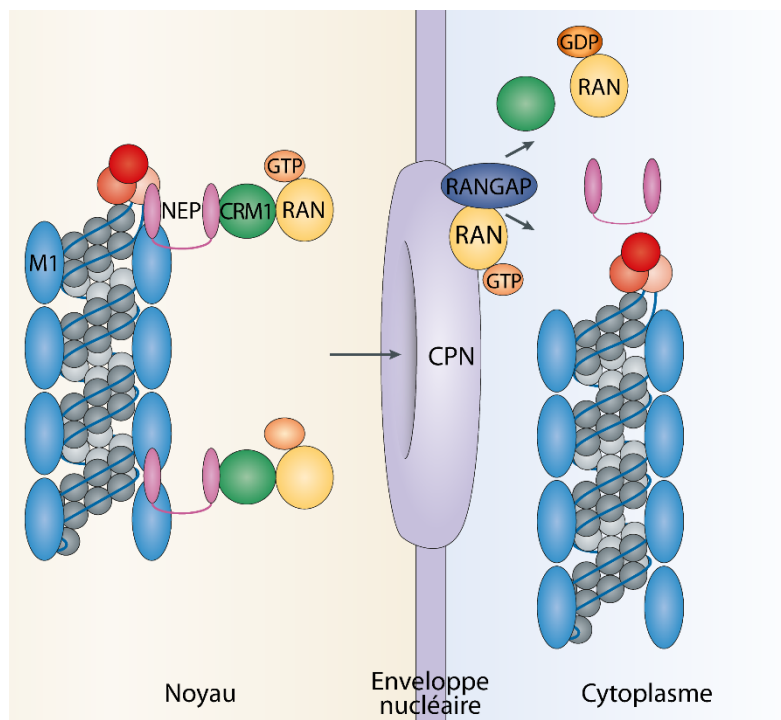


Figure 12 : Modélisation du mécanisme d'exportation des RNPg hors du noyau
Adapté de (Eisfeld *et al*, 2015).

6) Assemblage et bourgeonnement

Après export, les RNP s'accumulent en périphérie du noyau, au niveau du centrosome (MTOC) (Momose *et al*, 2007). Elles utilisent ainsi le réseau de microtubules pour être acheminées à la membrane plasmique au pôle apical. Les RNP s'associent à la protéine RAB11 (Eisfeld *et al*, 2011) présente à la surface des endosomes de recyclage via la sous-unité PB2 de FluPol (Amorim *et al*, 2011; Veler *et al*, 2022). Cependant, un nouveau modèle de transport des RNP a été décrit dans lequel les

RNP s'associent bien à RAB11 mais leur transport se faisant plutôt *via* des vésicules issues du réticulum endoplasmique (de Castro Martin *et al*, 2017). Par conséquent, les RNP seraient transportées jusqu'à la membrane plasmique où l'interaction FluPol-RAB11 est rompue. Les protéines HA, NA et M2 sont déjà présentes pour être intégrées aux nouveaux virions. La protéine M2 pourrait jouer un rôle dans l'empaquetage du génome (McCown & Pekosz, 2005; Chen *et al*, 2008). Les huit segments génomiques sont sélectionnés *via* un mécanisme sélectif (Odagiri & Tashiro, 1997) qui permet l'incorporation de tous les segments pour la production d'une particule fonctionnelle. Ce mécanisme n'est pas encore complètement élucidé (pour revue (Li *et al*, 2021b)). En effet, certaines études suggèrent que la sélection des segments se fait *via* des motifs d'empaquetages présents dans les séquences UTR des segments alors que d'autres suggèrent que ce sont principalement les interactions ARN-NP ou ARN-ARN des régions dépourvues de NP des séquences codantes qui entreraient en jeu (Gerber *et al*, 2014; Fournier *et al*, 2012; Gavazzi *et al*, 2013b, 2013a; Gilbertson *et al*, 2016; Moreira *et al*, 2016; Dadonaite *et al*, 2019).

Des incertitudes subsistent encore également sur le mécanisme de bourgeonnement. De plus amples études sont nécessaires pour affiner la compréhension de ce mécanisme mais un modèle est proposé (pour revue (Rossman & Lamb, 2011)) en figure 13. Les protéines HA et NA présentes au niveau des radeaux lipidiques à la membrane plasmique initient le bourgeonnement en provoquant une courbure de la membrane plasmique. La protéine M1 est ensuite recrutée via les extrémités cytoplasmiques de NA et HA, ce qui pourrait accentuer davantage la courbure de la membrane. Les RNP étant liées à M1 via NP sont ainsi amarrées à la matrice virale afin d'être empaquetées. M1 recrute également la protéine M2 qui permet le bourgeonnement de la particule virale en réalisant la séparation entre membrane plasmique et enveloppe virale (Rossman *et al*, 2010). Enfin, la protéine NA permet la libération de la particule virale en clivant les interactions potentielles entre HA et les acides sialiques présents en surface (Palese & Compans, 1976; Palese *et al*, 1974).

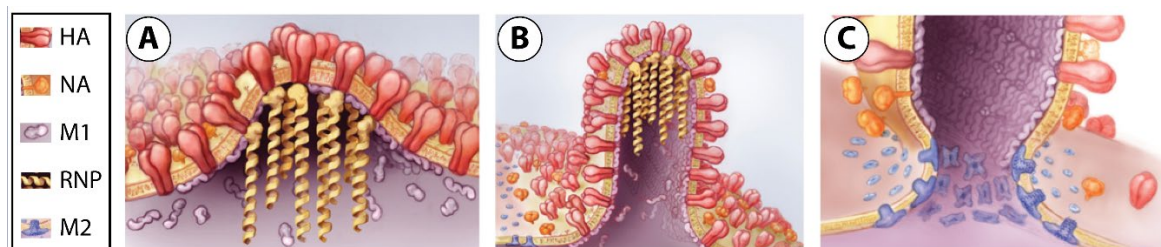


Figure 13 : Modélisation du bourgeonnement d'une particule virale
Adapté de (Rossman & Lamb, 2011).

IV. L'ARN-polymérase ARN-dépendante FluPol

Au sein de la RNP, FluPol interagit avec les deux extrémités de l'ARNg mais possiblement aussi avec NP (Biswas *et al*, 1998; Poole *et al*, 2004; Ng *et al*, 2012; Hsia *et al*, 2018; Miner *et al*, 2021). FluPol est constituée de trois sous-unités dont chacune est essentielle et assure un rôle dédié dans le fonctionnement du complexe. Cet hétérotrimère de 250 kDa est au centre de la réplication virale et assure de multiples fonctions, principalement la transcription et réplication du génome viral, mais aussi des interactions essentielles à d'autres étapes du cycle viral. FluPol joue notamment un rôle important dans l'adaptation du virus à son hôte et dans la barrière d'espèce (Naffakh *et al*, 2008; Gabriel *et al*, 2005) et ne fonctionne qu'en contexte RNP. Contrairement à l'ARN-polymérase II, FluPol ne possède pas de fonction relecture lui permettant de corriger les erreurs d'incorporation réalisées. Un fort taux de mutation est ainsi observé, estimé à une erreur par génome répliqué (Drake, 1993). Ce fort taux de mutation apporte un avantage évolutif aux IAV, appelé « *antigenic drift* » ou dérive antigénique, leur permettant une adaptation rapide à un nouvel hôte et un échappement au système immunitaire.

1) Description

La sous-unité PA est la seule sous-unité de FluPol suffisamment stable pour être exprimée seule en cellule d'insecte. Pendant longtemps, sa fonction est restée inconnue car elle ne possède aucune homologie avec d'autres protéines. Elle est constituée d'un domaine N-terminal portant une fonction endonucléase (Dias *et al*, 2009; Yuan *et al*, 2009) ainsi qu'un domaine C-terminal plus long dont la fonction est encore peu comprise (Figure 14). Ce dernier interagit avec l'extrémité N-terminale de PB1 en formant une pince qui accueille les quinze premiers résidus de PB1 au niveau de son domaine en pouce (He *et al*, 2008; Obayashi *et al*, 2008). La cavité formée est une région conservée au sein des IAV. Les deux domaines de PA sont reliés entre eux par un linker flexible qui entoure le cœur catalytique de FluPol.

La protéine PB1 constitue le cœur catalytique du complexe puisque c'est elle qui assure l'activité polymérase. Elle est également au centre du complexe puisqu'elle interagit avec les sous-unités PA et PB2 par chacune de ses extrémités. PB1 est constituée de trois régions principales : les extrémités N- et C-terminales qui permettent l'interaction avec les autres sous-unités du complexe et une région centrale contenant les motifs conservés des ARN-polymérases des virus à ARN négatif (Müller *et al*, 1994; Poch *et al*, 1989). Elle a une organisation structurale similaire aux ARN-

polymérase, dite en « main droite », avec des domaines en pouce, paume ou doigts délimitant une cavité abritant le site catalytique.

Comme PB1, la sous-unité PB2 ne peut être exprimée seule. Cependant, deux régions ont été identifiées comme étant solubles : une zone centrale (résidus 320-486) (Guilligay *et al*, 2008) qui sera rallongée par la suite (résidus 241-483) (Thierry *et al*, 2016) et une zone en C-terminal (résidus 678-759) (Tarendeau *et al*, 2008). C'est une protéine modulaire comportant quatre domaines principaux. Le tiers N-terminal de PB2 (résidus 1 à 247) est fixé fortement à PB1. La partie C-terminale forme quant à elle des domaines distincts. Tout d'abord, le domaine central, appelé « *cap-binding domain* » fixe la coiffe des ARNm cellulaires. Conservé chez toutes les souches d'Influenza, il est entouré de part et d'autre d'un domaine linker. Le reste du C-ter de PB2 est constitué du domaine 627 qui joue un rôle essentiel dans la restriction d'hôte *via* le résidu 627 ainsi que le domaine NLS qui contient deux séquences NLS. Ce dernier a été le premier domaine de PB2 à être cristallisé (Tarendeau *et al*, 2007). Plus tard, le « *cap-binding domain* » a également été cristallisé lié à la coiffe d'un ARNm (Guilligay *et al*, 2008). Il a ainsi été montré que le mécanisme de liaison de la coiffe était similaire à celui déployé par des protéines ayant des fonctions similaires comme le « *Cap-Binding Complex* » (CBC) ou le facteur de traduction eIF4E. Étonnamment, le « *cap-binding domain* » de PB2 a pourtant une affinité beaucoup plus faible pour la coiffe des ARNm que ces protéines (Marcotrigiano *et al*, 1997; Mazza *et al*, 2002).

Le cœur de FluPol comporte quatre canaux permettant le bon déroulement de la transcription et de la réplication. Les trois sous-unités contribuent à cette organisation. Le site actif de FluPol contient donc un canal d'entrée et de sortie pour la matrice, un canal d'entrée pour les NTP nécessaires à la synthèse de nouveaux produits ainsi qu'un canal de sortie pour les ARN néo-synthétisés. Le domaine endonucléase ainsi que les domaines C-terminaux de PB2 sont reliés au cœur de FluPol par des linkers flexibles qui leurs permettent une mobilité importante (détaillé en section IV.6).

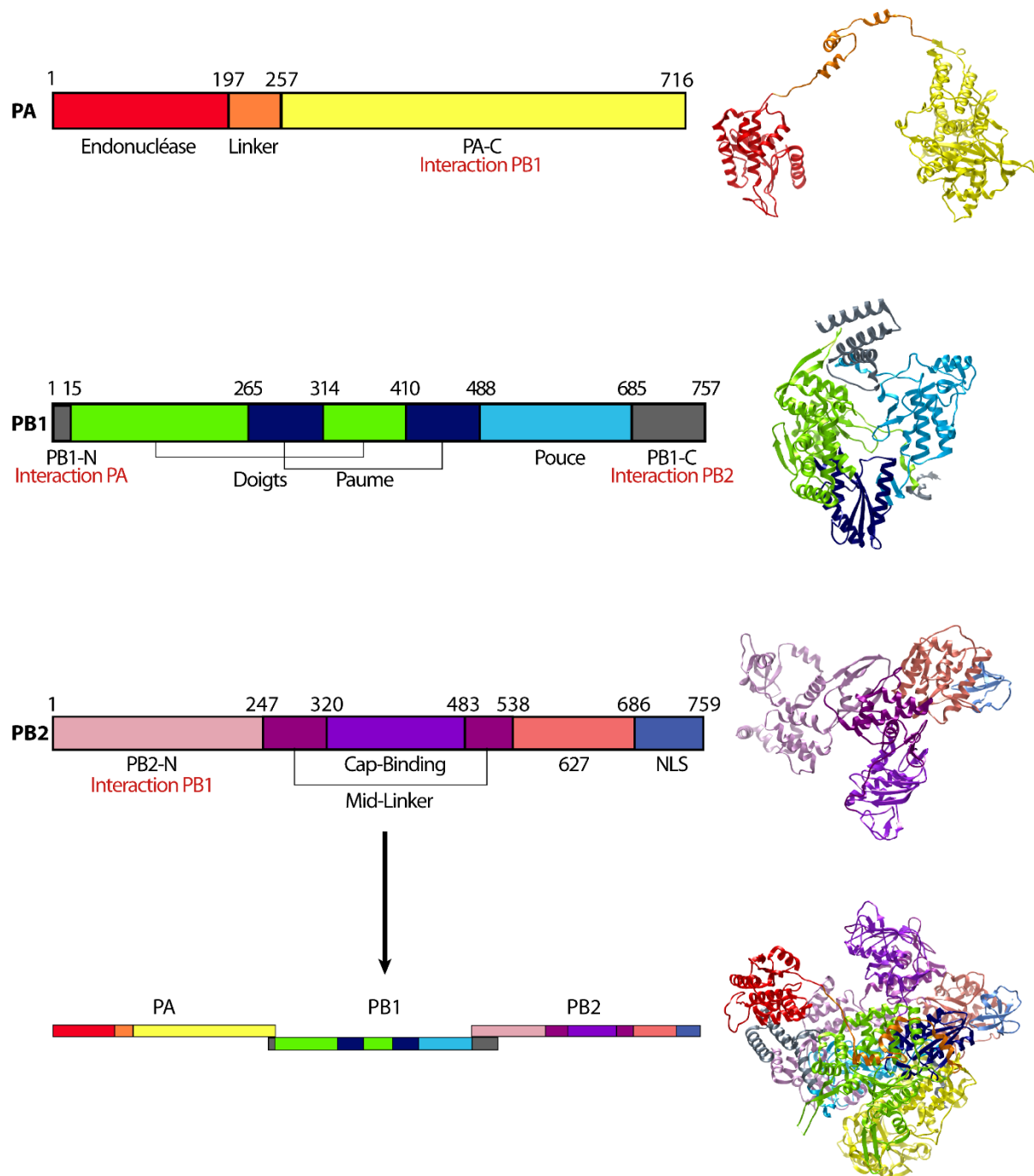


Figure 14 : Organisation structurale des différentes sous-unités de FluPol

Représentation schématique (gauche) et structure (droite) des différents domaines des sous-unités PA, PB1 et PB2 de FluPol de IAV. PDB 6RR7, souche H3N2. D'après (Ruigrok *et al*, 2010; Te Velthuis *et al*, 2021; Pflug *et al*, 2017).

2) Transport au noyau

FluPol liée aux RNP pénètre dans le noyau avec la RNP via le mécanisme d'import précédemment décrits en section III.3. Le transport de FluPol nouvellement synthétisée au noyau est plus particulier puisqu'il se découple en deux étapes distinctes et indépendantes (Huet *et al*, 2010;

Fodor & Smith, 2004; Deng *et al*, 2006a)(Figure 15). En effet, toutes les sous-unités ne migrent pas ensemble au noyau.

D'une part, les sous-unités PA et PB1 s'assemblent dans le cytoplasme avant d'être importées ensemble au noyau. Ainsi, PA ou PB1 doivent être co-exprimées pour s'accumuler au noyau (Fodor & Smith, 2004). Le dimère n'utilise pas la voie d'import classique constituée des importines- α et - β comme c'est le cas des RNP mais interagit principalement avec l'importine- β RanBP5 (Deng *et al*, 2006a; Hutchinson *et al*, 2011), formant un complexe stable avec cette dernière (Swale *et al*, 2016, 2020).

D'autre part, la protéine PB2, stabilisée par la protéine chaperonne Hsp90 (Naito *et al*, 2007), migre seule au noyau en utilisant la voie classique d'import au noyau via les importines- α et β (Tarendeau *et al*, 2007; Resa-Infante *et al*, 2008). Ce n'est qu'une fois que le dimère PA-PB1 et PB2 sont dans le noyau que FluPol est assemblée.

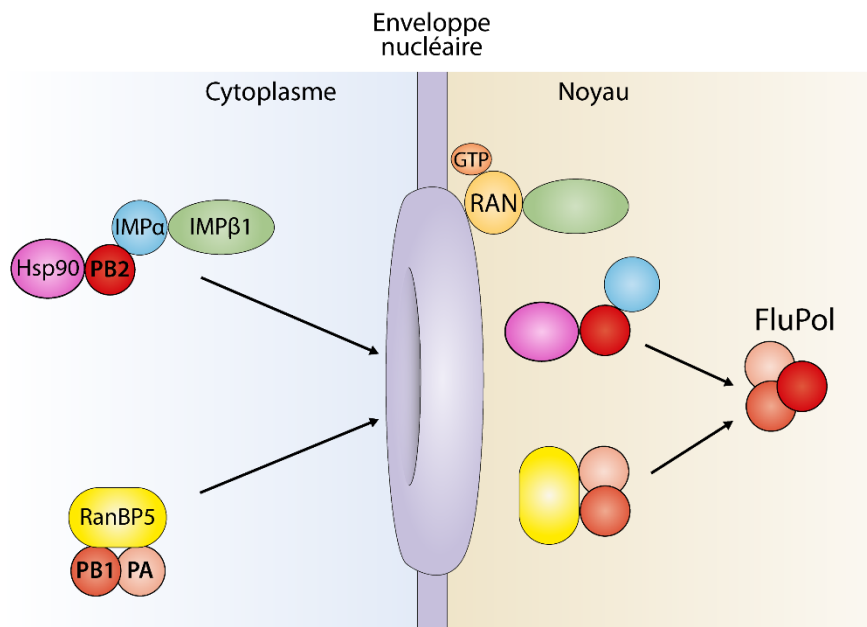


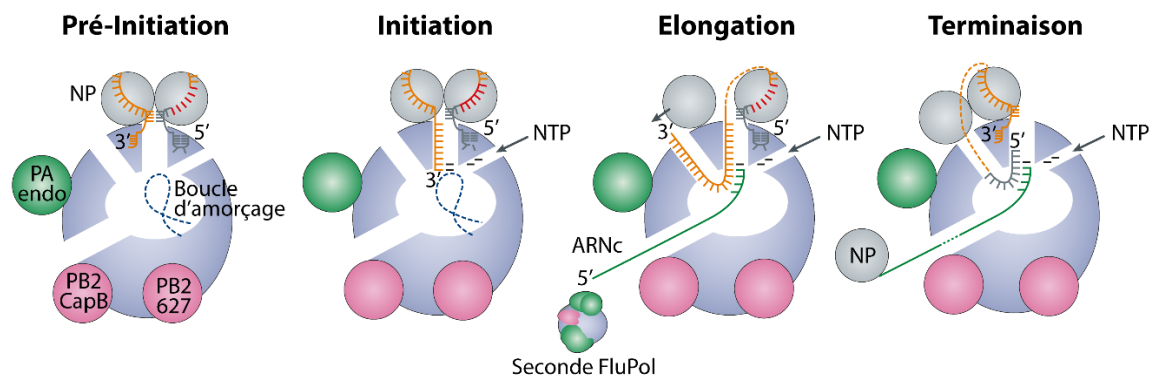
Figure 15 : Schématisation du transport de FluPol au noyau

3) Réplication du génome viral

Comme expliqué précédemment en section III.4, la réplication du génome viral se déroule en deux étapes, avec une première étape de synthèse de RNPc avant de pouvoir synthétiser de nouvelles RNPg. Cette première étape est initiée par l'entrée de l'extrémité 3' du génome matrice dans le site actif de FluPol (Robb *et al*, 2019). Contrairement à la transcription, FluPol n'a pas besoin d'une amorce pour initier la synthèse d'ARN et réalise une initiation *de novo* à partir de l'extrémité 3' de la matrice grâce à la boucle d'amorçage présente dans le site actif qui stabilise le complexe (Te Velhuis *et al*,

2016) pour la réplication du génome (Figure 16A). Lorsque l'ARNc commence à émerger de FluPol, une nouvelle FluPol se fixe à son extrémité 5' et des NP se lient au reste de l'ARN afin de le stabiliser, le protéger de la dégradation et ainsi former une nouvelle RNPC (York *et al*, 2013; Vreede *et al*, 2004). La présence de FluPol et NP néo-synthétisées est donc essentielle pour une accumulation des RNPC. Une fois l'intermédiaire RNPC synthétisé, la synthèse des RNPG peut être réalisée en utilisant cette dernière comme matrice.

A. Synthèse ARNc



B. Synthèse ARNv

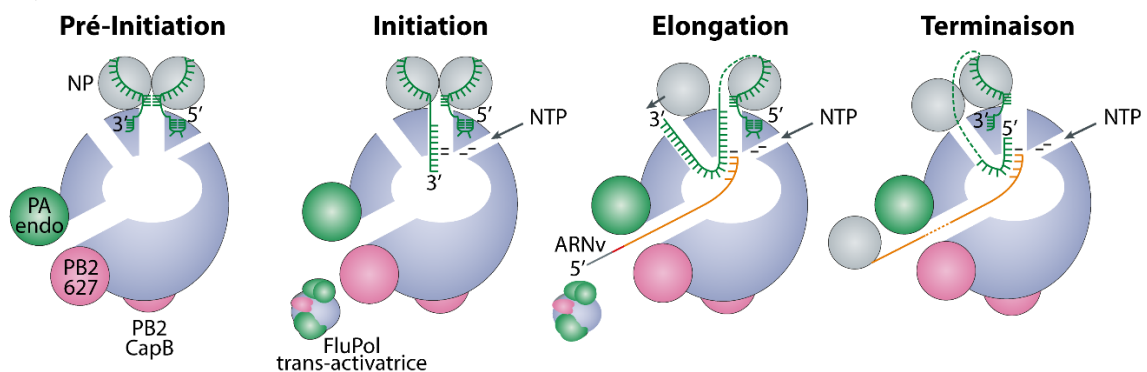


Figure 16 : Modélisation de la synthèse des ARN par FluPol au cours de la réplication

Mécanisme de synthèse des ARNc (A) et des ARNv (B) mis en place par FluPol pour assurer la réplication du génome des IAV. Adapté de (Te Velthuis & Fodor, 2016).

Pour se faire, FluPol met en place un mécanisme différent, indépendant de la boucle d'amorçage (Figure 16B). La dimérisation de FluPol avec une FluPol trans-activatrice est requise pour initier cette réplication. L'extrémité 3' de l'ARNc pénètre dans le site actif où l'incorporation *de novo* des deux premiers nucléotides a lieu au niveau des quatrièmes et cinquièmes nucléotides de l'ARNc. Suite à cette initiation interne, le dinucléotide produit est relocalisé à l'extrémité 3' de l'ARNc (Deng *et al*, 2006b) grâce à la boucle d'amorçage et la FluPol trans-activatrice pour permettre l'élongation de l'ARNg (Oymans & Te Velthuis, 2018; Fan *et al*, 2019a)(non représenté sur le schéma 15B). Une FluPol (peut-être la FluPol trans-activatrice) se fixe à l'extrémité 5' de l'ARNg et NP s'associe à cet ARN pour

former la RNPg. Cette dernière est ainsi prête à être exportée hors du noyau pour être empaquetée dans une nouvelle particule virale.

4) Transcription du génome viral et vol de coiffe

La transcription du génome des IAV est fonctionnellement dépendante de la transcription cellulaire (Lamb & Choppin, 1977; Engelhardt *et al*, 2005). Contrairement à d'autres ARN-polymérases de virus à ARN, FluPol n'a pas la capacité à synthétiser et méthyliser la coiffe de ses propres ARNm (Plotch *et al*, 1978). Afin de les produire, elle a donc recours à un mécanisme particulier appelé « vol de coiffe » (ou « *cap-snatching* ») au cours duquel le complexe hétérotrimérique vole la coiffe des ARNm produits par la cellule hôte (Plotch *et al*, 1981). Ce mécanisme n'est pas exclusif aux *Orthomyxoviridae* car il est également retrouvé chez les *Arenaviridae* (Raju *et al*, 1990) et *Bunyaviridae* (Simons & Pettersson, 1991). Cependant, les virus Influenza sont les seuls à réaliser ce mécanisme dans le noyau de la cellule hôte. Cela s'explique par la façon dont les virus Influenza volent la coiffe puisque ce mécanisme nécessite une interaction directe avec l'ARN-polymérase II (Pol II) de l'hôte (Lukarska *et al*, 2017; Engelhardt *et al*, 2005)(détaillé en section IV.5.a).

Le vol de coiffe a lieu en trois étapes majeures. Après interaction avec Pol II, la sous-unité PB2 de FluPol fixe la coiffe des ARNm de la cellule hôte en cours de synthèse au niveau de son domaine « *cap-binding* » (Guilligay *et al*, 2008)(Figure 17A). Ensuite, la sous-unité PA clive cet ARNm dix à quinze nucléotides en aval de la coiffe *via* son domaine endonucléase (Dias *et al*, 2009). Enfin, l'extrémité 3' résultante est insérée dans le site actif de FluPol où elle s'apparie avec l'extrémité 3' de l'ARNg matrice (Figure 17B) pour initier la transcription. Ainsi, la sous-unité PB1 synthétise l'ARNm viral en utilisant la coiffe volée comme amorce. La séparation du duplex ARNg-ARNm est possible grâce à la Tyrosine 205 de PB2 permettant la sortie de l'ARNm et de l'ARNg via leurs canaux de sortie respectifs (Kouba *et al*, 2019; Wandzik *et al*, 2020). Une queue poly-A est également ajoutée grâce à un mécanisme de bégaïement de FluPol sur une succession de résidus uridines présents à l'extrémité 5' de la matrice (Poon *et al*, 1999; Robertson *et al*, 1981). Les ARNm viraux ainsi produits, coiffés et polyadénylés, sont donc chimériques, avec les premiers nucléotides d'origine cellulaire et peuvent être reconnus par la machinerie de traduction cellulaire dans le cytoplasme. Le réinitiation d'un nouveau cycle de transcription peut se faire rapidement car les extrémités 3' et 5' de l'ARNg restent constamment liées à FluPol même au cours de la transcription (Wandzik *et al*, 2020; Coloma *et al*, 2020).

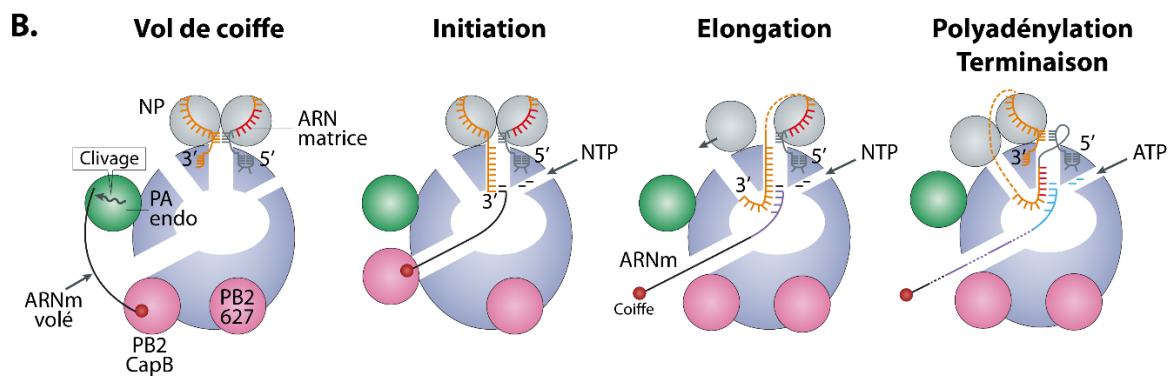
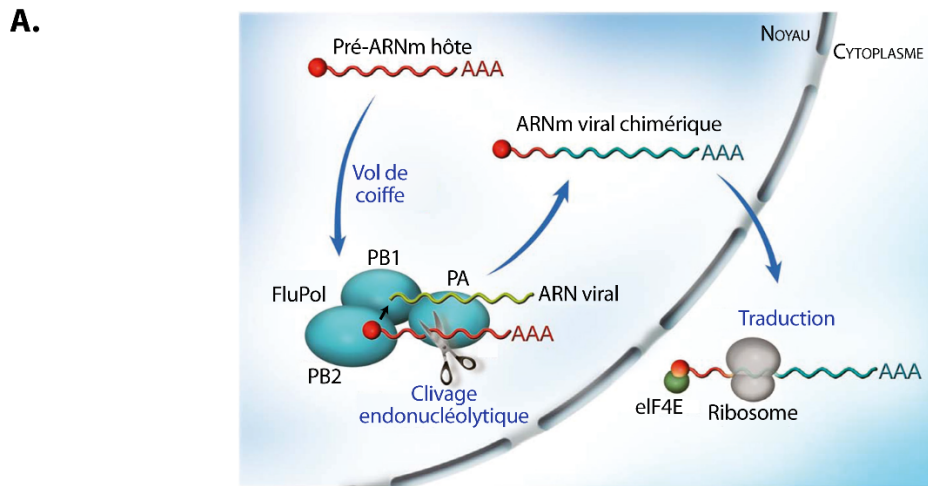


Figure 17 : Détails des mécanismes de transcription mis en place par FluPol

A. Schéma récapitulatif du vol de coiffe. Adapté de (Boivin *et al*, 2010). **B.** Mécanisme de synthèse des ARNm mis en place par FluPol pour assurer la transcription du génome des IAV. Adapté de (Te Velthuis & Fodor, 2016).

5) Interactions avec les facteurs cellulaires

De nombreuses interactions de FluPol avec des partenaires cellulaires au noyau ont été décrites. Les plus pertinentes sont présentées ci-dessous (liste non exhaustive).

a. Interaction avec l'ARN-polymérase II

Pol II est un complexe cellulaire constitué de 12 sous-unités appelées RPB1 à RPB12 par ordre décroissant de taille (Cramer, 2004). Les sous-unités RPB1 et RPB2 composent ainsi le cœur du complexe autour duquel s'agencent les autres sous-unités de taille moindre. RPB1 possède un long domaine C-terminal (CTD) non structuré qui joue un rôle essentiel dans le déroulement de la transcription chez les eucaryotes (Dieci, 2021). En effet, ce long domaine est constitué d'un motif

consensus Tyr-Ser-Pro-Thr-Ser-Pro-Ser répété cinquante-deux fois chez l'Homme. Ce CTD permet notamment la régulation des différentes étapes de transcription en permettant le recrutement des facteurs cellulaires régulateurs, rythmé par les modifications post-traductionnelles de ce CTD (pour revue (Zaborowska *et al*, 2016)). En effet, certains acides aminés de ce CTD peuvent être phosphorylés. C'est le cas des Sérines 2 (Ser2) et 5 (Ser5) de ce motif qui sont phosphorylés à des moments différents de la transcription (Figure 18). Ainsi, la phosphorylation de la Sérine 5 (Ser5P) est principalement présente lors de l'initiation de la transcription (TSS) mais est aussi retrouvée en moindre abondance au niveau des exons. En revanche, la phosphorylation sur la Sérine 2 (Ser2P) est peu présente à l'initiation de la transcription mais augmente peu à peu au cours de l'élongation jusqu'à la fin de la transcription (TTS).

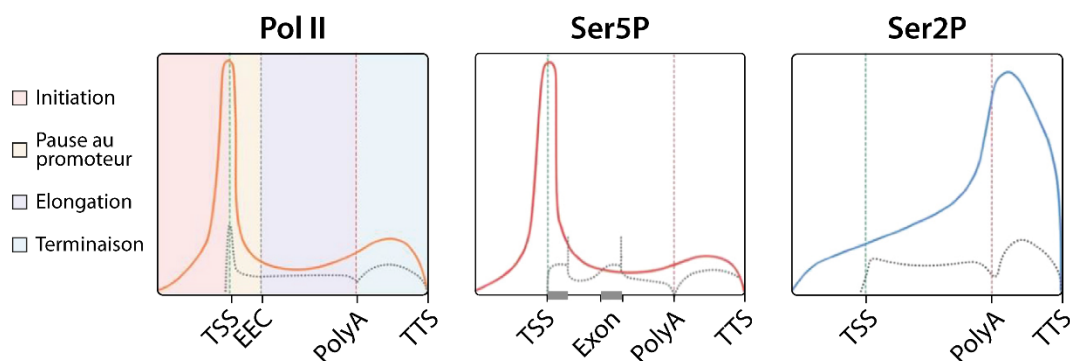


Figure 18 : Représentation des profils de phosphorylations de Pol II et des sérines 2 et 5.

Les traits pleins représentent des données obtenues par immunoprécipitation de chromatine (ChIP-seq) et les traits en pointillés noirs des données obtenues par séquençage des ARN en cours d'élongation (« mammalian Native Elongation Transcript sequencing », mNET-seq). TSS = Initiation de la transcription (« *Transcription Starting Site* »), EEC = Checkpoint de début d'élongation (« *Early Elongation Checkpoint* »), TTS = Initiation de la terminaison (« *Transcription Termination Site* »). Adapté de (Zaborowska *et al*, 2016).

Il a été montré que FluPol interagit directement avec le CTD de Pol II lorsque Ser5 est phosphorylée (Ser5P) et non lorsque Ser2 est phosphorylée (Ser2P)(Engelhardt *et al*, 2005; Martínez-Alonso *et al*, 2016; Lukarska *et al*, 2017). Cela suggère donc que l'interaction FluPol – Pol II ait lieu au niveau du TSS ou en début d'élongation lorsque Pol II est en pause. Cette interaction a été démontrée lorsque FluPol est seule mais également en contexte RNP. De cette façon, FluPol pourrait ainsi se rapprocher de proche en proche de la source des ARNm et faciliter le vol de coiffe (Krischuns *et al*, 2021). L'interaction entre FluPol et Pol II a été caractérisée structuralement pour les Influenza de type A, B et C en utilisant des peptides mimant le CTD, permettant de montrer que cette interaction est entièrement médiée par PA chez les IAV alors qu'elle implique également PB2 pour les IBV et PB1 pour les ICV (Figure 19). Deux sites d'interactions avec le CTD ont été mis en évidence chez ces trois types d'Influenza. Cependant, l'emplacement et la composition de ces sites diffèrent (pour revue (Walker &

Fodor, 2019)). En effet, les sites 1 et 2 identifiés sur la structure de FluPol de IAV sont conservés parmi les différentes souches de IAV et contiennent des acides aminés basiques qui permettent l'interaction avec la Ser5P. Chez les IBV, seul le site 1 est similaire à celui des IAV alors que les deux sites des ICV sont différents de ceux des IAV. Cette interaction est donc conservée chez les différents types d'Influenza mais les modes de fixation de FluPol au CTD ont divergé au cours de l'évolution.

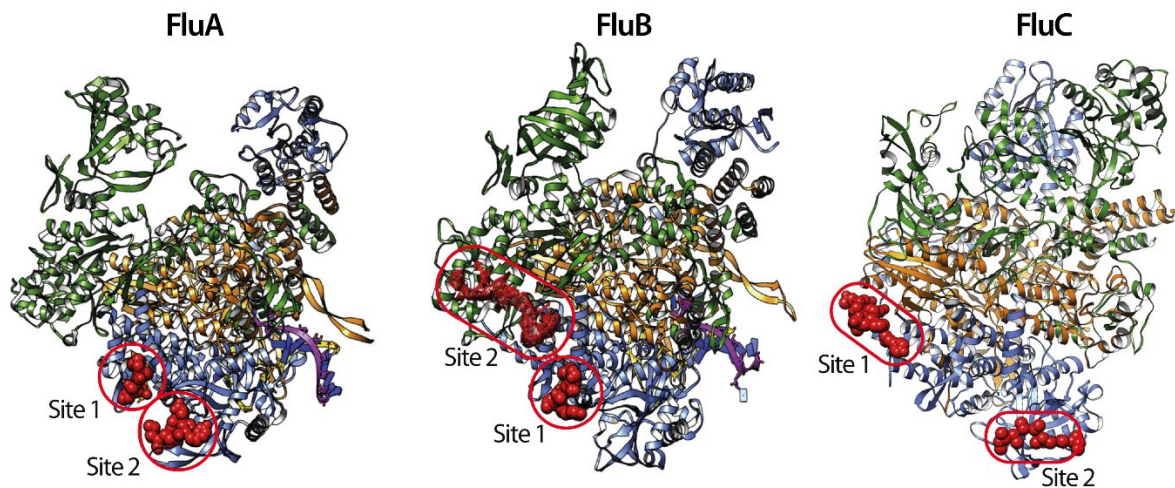


Figure 19 : Structures de FluPol liée à des peptides mimant le CTD de Pol II

Les peptides mimant le CTD sont représentés en rouge. PA est représentée en bleu, PB1 en orange et PB2 en vert. Les deux sites d'interaction pour FluPol de IAV (PDB : 5M3H), d'IBV (PDB : 5M3J) et d'ICV (PDB : 6F5P) sont entourés en rouge. Adapté de (Walker & Fodor, 2019).

Malgré une caractérisation structurale très détaillée, beaucoup de questions subsistent encore autour de l'interaction FluPol – Pol II et du vol de coiffe. En effet, le déroulé précis du vol de coiffe n'est pas encore connu. La liaison de FluPol à la Ser5P ainsi que l'enrichissement de cette dernière au niveau des TSS observé en CHIP-seq suggère que celui-ci aurait lieu lors de l'initiation de la transcription (Chan *et al*, 2006; Walker & Fodor, 2019). Or, pour que le vol de coiffe ait lieu, l'ARN en cours de synthèse doit être coiffé. Cette coiffe est ajoutée dès que l'ARN sort par le canal de sortie de Pol II. L'ajout de la coiffe est couplé à la pause de Pol II 20 à 100 pb après le TSS (Core & Adelman, 2019). Cette pause joue un rôle important dans la régulation des étapes de transcription et permet la préparation à l'élongation, notamment par le recrutement de facteurs d'élongation. Lors de cette pause, Pol II est stabilisée par plusieurs facteurs bloquant sa progression, dont le « *negative elongation factor* » (NELF) (Yamaguchi *et al*, 1999). L'ajout de la coiffe est alors réalisé par une succession d'étapes réalisées par différentes enzymes (Chiu *et al*, 2002; Rasmussen & Lis, 1993; Ghosh & Lima, 2010). Une fois la coiffe ajoutée, celle-ci est fixée par le « *cap-binding complex* » (CBC) (Mazza *et al*, 2001; Izaurralde *et al*, 1994). Or, le CBC a une très forte affinité pour la coiffe (Worch *et al*, 2005), bien supérieure à celle du

« *cap-binding domain* » de PB2 (Guilligay *et al*, 2008). Comment PB2 parvient-elle alors à fixer la coiffe des ARN en cours de synthèse si elle ne peut entrer en compétition avec le CBC ? Cela sous-entend que le vol de coiffe doit avoir lieu dans une fenêtre de temps restreinte, après l'ajout de la coiffe, mais avant liaison au CBC. Plusieurs hypothèses ont été émises où FluPol empêcherait le recrutement du CBC (Krischuns *et al*, 2021). Différents mécanismes ont été évoqués, notamment une séquestration du CBC par FluPol ou une gêne stérique due à la liaison de FluPol au CTD. En effet, le CBC a été identifié comme étant un interactant potentiel de FluPol lors de cribles (pour revue (Krischuns *et al*, 2021)).

b. Interaction avec ANP32

Une autre interaction de grande importance est celle avec les protéines de la famille « *Acidic nuclear phosphoproteins of 32 kDa* » (ANP32). Ces dernières années, plusieurs équipes se sont intéressées à cette interaction qui joue un rôle majeur dans la réplication du génome viral. Décrite pour la première fois en 2015 (Sugiyama *et al*, 2015), il est rapidement apparu qu'ANP32A joue un rôle crucial dans la restriction d'hôte (Long *et al*, 2016). En effet, ANP32A est importante pour la synthèse des ARNg à partir des ARNc en permettant le recrutement d'une FluPol néosynthétisée (Sugiyama *et al*, 2015; Baker *et al*, 2018; Staller *et al*, 2019) et en stabilisant l'interaction entre cette dernière et la FluPol en cours de réplication avec laquelle elle dimérise. ANP32A forme ainsi un pont entre les deux polymérases en interagissant avec le domaine 627 de PB2. Des structures d'ANP32A en complexe avec le domaine C-ter de PB2 ou FluPol complète ont récemment été obtenues (Camacho-Zarco *et al*, 2020; Carrique *et al*, 2020).

c. Autres interactions et ciblage à la chromatine

Plusieurs études ont été menées pour identifier l'interactome de FluPol par différentes méthodes de crible (Munier *et al*, 2013; Tafforeau *et al*, 2011; Watanabe *et al*, 2014). De nombreuses protéines intervenant dans les mécanismes de transcription et de régulation de la chromatine ont ainsi été identifiées comme pouvant être des partenaires potentiels de FluPol (pour revue (Krischuns *et al*, 2021)).

Certaines interactions pourraient faciliter le ciblage de FluPol à la chromatine dans les zones transcriptionnellement actives où Pol II est présente, favorisant ainsi le vol de coiffe. C'est le cas des remodeleurs de la chromatine CHD1 et MORC3 qui ont été identifiés comme étant des interactants de FluPol (Ver *et al*, 2015; Marcos-Villar *et al*, 2016). Ces protéines ayant la capacité de reconnaître les zones transcriptionnellement actives de la chromatine, FluPol pourrait les utiliser pour cibler les zones de transcription où Pol II est présente. FluPol interagit également avec le complexe exosome

responsable du contrôle de la qualité des ARN synthétisés par Pol II (Rialdi *et al*, 2017). Cette interaction pourrait également favoriser le recrutement de FluPol à la chromatine dans les régions promotrices.

D'autres interactions peuvent donner des indices quant au déroulé, au mécanisme et à la localisation du vol de coiffe. En effet, peu de facteurs de transcription généraux ont été identifiés lors de ces cribles. Cependant, beaucoup de facteurs intervenant lors de la pause de Pol II ont été identifiés comme des partenaires potentiels de FluPol. SPT5, un facteur d'élongation permettant le recrutement du complexe synthétisant la coiffe, a par exemple été identifié comme interagissant avec FluPol (Bradel-Tretheway *et al*, 2011). C'est également le cas du complexe P-TEFB responsable de la phosphorylation de la Ser2 (Zhang *et al*, 2010). Ces interactants renforcent l'hypothèse selon laquelle le vol de coiffe aurait lieu lors de la pause de Pol II. Les enzymes réalisant la coiffe (Li *et al*, 2020) ainsi que le CBC (Watanabe *et al*, 2014) ont également été identifiés.

Des études plus poussées sur les interactants de FluPol sont donc nécessaires et permettraient de mieux comprendre le mécanisme du *cap-snatching*, notamment son déroulé et sa localisation.

6) Structure(s) de FluPol

Des dizaines de structures de FluPol de différents types ont été obtenues depuis la fin des années 2000. Elles ont permis de mettre en évidence que les FluPol des différents types d'Influenza sont globalement structurellement conservées malgré des séquences de PA, PB1 et PB2 peu conservées (Yamashita *et al*, 1989; Kemdirim *et al*, 1986). Elles ont aussi dévoilé l'incroyable plasticité de FluPol, qui a la capacité à changer de conformation en fonction des interactions qu'elle peut avoir dans le noyau, lui permettant d'accomplir ses diverses fonctions. En effet, la comparaison de toutes ces structures démontre que FluPol peut être décomposée en deux parties distinctes (Figure 20). Premièrement, une partie fixe qui constitue le cœur (ou « *core* ») de FluPol et est peu mobile, composée du domaine C-terminal de PA, de PB1 et du domaine N-terminal de PB2. Ces domaines sont conservés structurellement entre les souches et ce quel que soit l'activité de FluPol. Peu de variations structurales sont observées au niveau de ce cœur, que FluPol soit liée ou non à un ARNc ou un ARNv à l'exception de modifications localisées au niveau des sites de liaison de l'ARN (Thierry *et al*, 2016). Cependant, le domaine N-terminal de PA composé de l'endonucléase ainsi que le domaine C-terminal de PB2 sont mobiles et forment des domaines périphériques autour du cœur de FluPol. Ces domaines modulaires s'agencent différemment dans l'espace en fonction des liaisons avec les divers ARN ou partenaires cellulaires de FluPol pour favoriser une de ses activités plutôt qu'une autre.

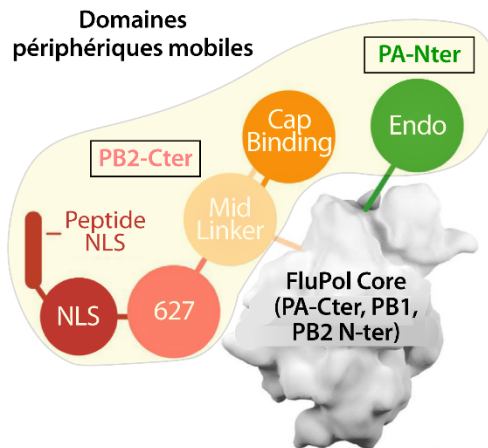


Figure 20 : Schématisation des différentes conformations que FluPol peut adopter au cours des cycles de réplication / transcription

Adapté de (Pflug *et al*, 2017).

En effet, la liaison au CTD stabilise FluPol dans une conformation transcriptionnellement active ce qui permet une efficacité accrue du vol de coiffe (Serna Martin *et al*, 2018)(Figure 21). Ainsi, cette liaison à Pol II entraîne le réarrangement des domaines périphériques de façon à rendre accessible le domaine « *cap-binding* » de PB2 ainsi que le domaine endonucléase de PA qui se font alors face (Pflug *et al*, 2014). Après clivage de l'ARN hôte, le domaine « *cap-binding* » de PB2 pivote de 70° (Reich *et al*, 2014; Pflug *et al*, 2018), éloignant ainsi le fragment d'ARN coiffé de l'endonucléase et permettant son insertion dans le site actif de FluPol via le canal de sortie pour appariement avec l'ARNg (Reich *et al*, 2017).

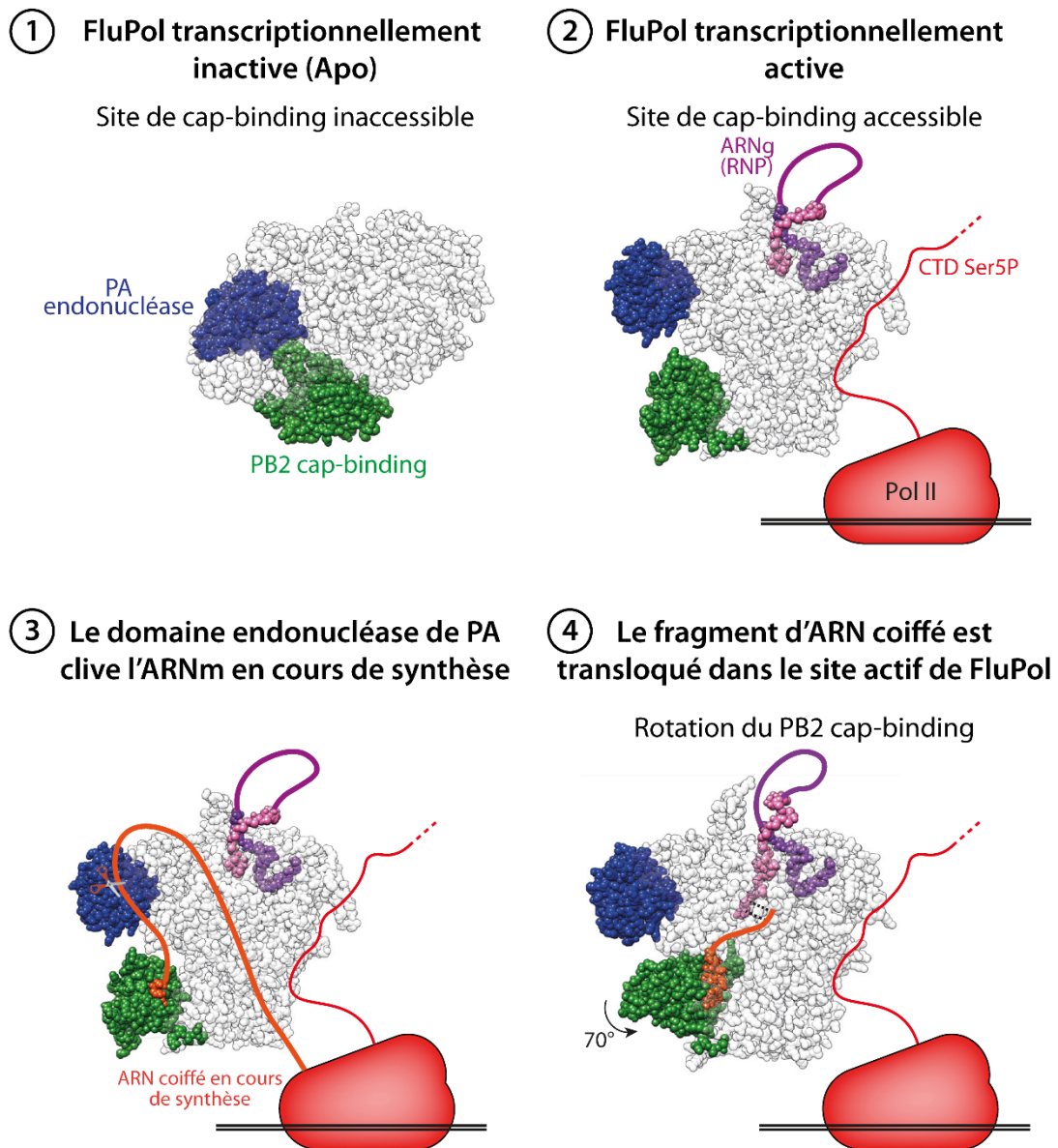


Figure 21 : Illustration des changements de conformation de FluPol au cours du vol de coiffe
PDB dans l'ordre des étapes : 5D98 (ICV), 4WSB (IAV), 6EVK (IAV), 5MSG (IBV). Adapté de (Walker & Fodor, 2019).

Cette année, une nouvelle structure de FluPol a été obtenue (Li *et al*, 2021a). Les auteurs de l'étude proposent que cette structure corresponde à un intermédiaire conformationnel de FluPol, qu'ils ont appelée « *resident state* » qui incarnerait un état transitoire permettant de faire le lien entre la conformation répliation et la conformation transcription (Figure 22). En effet, le domaine « *cap-binding* » de PB2 n'est pas accessible dans cette conformation et le cœur de FluPol est contracté. Les auteurs proposent donc un modèle dans lequel la liaison du CTD à FluPol en « *resident state* » permet un changement de conformation du domaine 627 de PB2 qui libère le domaine « *cap-binding* » et prépare FluPol pour la transcription. Après transcription, la contraction du cœur de FluPol permet un

basculement en conformation inactive. L'accumulation de FluPol néosynthétisées au noyau fait basculer certaines FluPol vers un cycle de réplication en interagissant avec la zone charnière de PB2 (qui fait le lien entre le domaine N-terminal de PB2 faisant parti du cœur et les domaines C-terminaux périphériques). Cette interaction permet la rotation du domaine C-terminal de PB2 et son interaction avec ANP32A qui stabilise FluPol dans sa conformation réplicase avant de terminer son cycle et revenir à l'état intermédiaire.

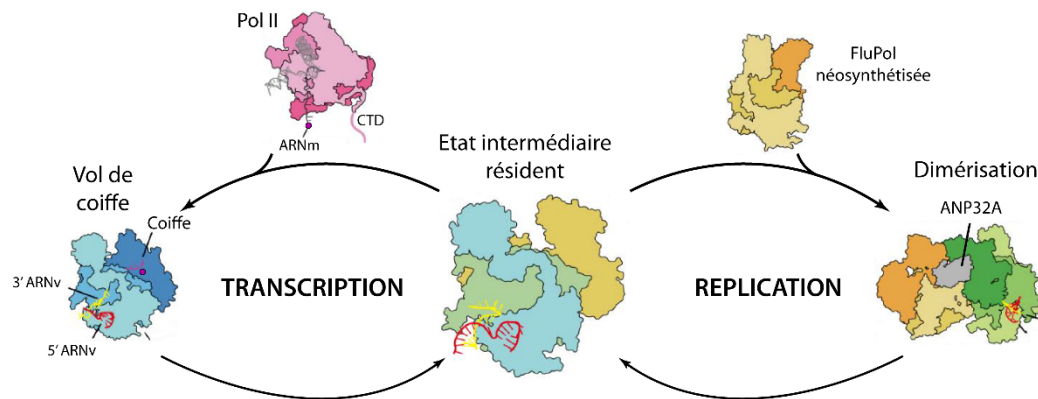


Figure 22 : Modélisation du rôle central de l'état résident dans la transition entre les cycles de réplication et de transcription de FluPol
Adapté de (Li *et al*, 2021a).

V. Conséquences de l'infection sur la chromatine de l'hôte

L'infection de la cellule hôte par les IAV désorganise fortement le fonctionnement de la cellule. En effet, de nombreuses conséquences cellulaires de l'infection ont été répertoriées, à différents niveaux du métabolisme. Le virus met ainsi différents mécanismes en place permettant une optimisation de sa production au détriment du métabolisme cellulaire. Un équilibre délicat doit alors s'installer pour produire du virus en quantité massive mais sans détruire la cellule avant que cela ne soit fait. Une preuve de ce détournement massif est qu'à 8h post-infection, plus de la moitié des ARNm présents dans la cellule sont viraux (Bercovich-Kinori *et al*, 2016). Un déclin progressif de la synthèse des protéines cellulaires est observé au cours de l'infection jusqu'à ce que la cellule ne produise presque plus que des protéines virales (Beloso *et al*, 1992; Inglis, 1982; Zürcher *et al*, 2000). Plusieurs mécanismes sont donc mis en place par les IAV pour optimiser la production virale.

Certaines études ont montré l'impact de FluPol sur la transcription cellulaire. Ainsi, une infection par les IAV inhibe l'élongation de la transcription réalisée par Pol II. Le groupe du Pr. Ervin Fodor a montré par des expériences de ChIP-seq que FluPol inhibe l'élongation de Pol II mais pas

l'initiation de la transcription, ni la reconnaissance du promoteur (Chan *et al*, 2006). Deux mécanismes ont été proposés pour expliquer ce phénomène. Premièrement, le vol de coiffe assuré par FluPol provoquerait une terminaison prématurée de Pol II. En effet, la protéine Xrn2 dégraderait l'ARN non coiffé en cours de synthèse jusqu'au détachement de Pol II de l'ADN, occasionnant une terminaison de la transcription selon le modèle dit « *torpedo model* » (Bauer *et al*, 2018; Levene & Gaglia, 2018)(Figure 23A). La deuxième hypothèse avancée est que la liaison de FluPol au CTD empêcherait le recrutement correct de facteurs d'élongation au niveau du CTD qui joue un rôle important dans l'organisation temporelle de la transcription. Cependant, une étude plus récente menée par ce même groupe n'a pas détecté cette inhibition de l'initiation de l'élongation lors d'expériences de mNET-seq. Ces dernières révèlent plutôt une diminution progressive de la présence de Pol II au fur et à mesure qu'elle s'éloigne du TSS (Bauer *et al*, 2018)(Figure 23B).

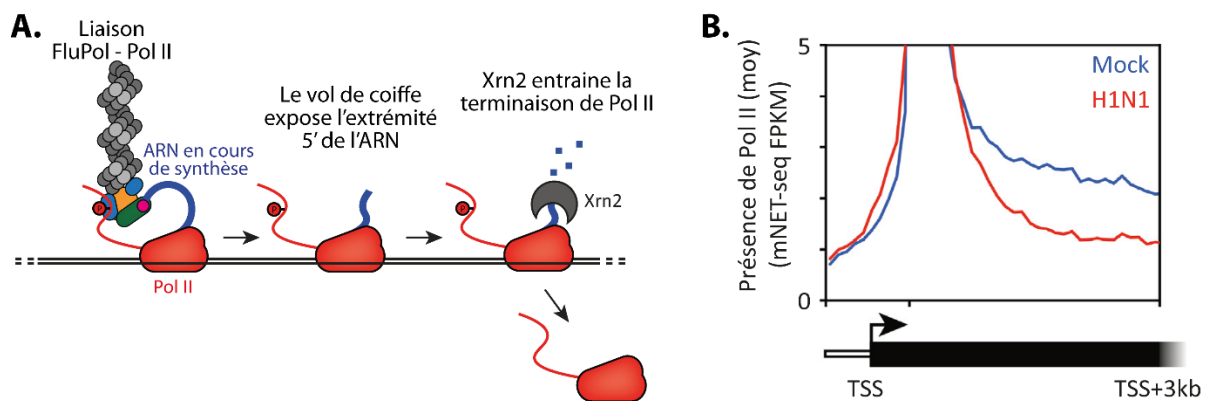


Figure 23 : L'infection par le IAV inhibe l'élongation de la transcription par Pol II

A. Illustration de l'inhibition de l'élongation par le modèle *torpedo*. **B.** Représentation de la diminution progressive de la présence de Pol II dans les gènes. Adapté de (Walker & Fodor, 2019).

En plus de l'inhibition de l'élongation, d'autres mécanismes ont été proposés pour interférer avec la production de protéines cellulaires. Katze et Krug ont ainsi proposé que les transcrits dont la coiffe a été volée étaient dégradés par les nucléases du noyau (Katze & Krug, 1984). NS1 pourrait également inhiber la maturation des ARNm et ainsi empêcher leur export nucléaire (Nemeroff *et al*, 1998; Shimizu *et al*, 1999). La régulation de l'expression des protéines de l'hôte ne se joue donc pas au niveau protéique mais bien au niveau des ARNm. En séquestrant les ARNm cellulaires au noyau, NS1 empêche leur traduction et favorise ainsi la production de protéines virales.

Deux études récentes de 2018 ont montré que l'infection par les IAV inhibe également la terminaison de la transcription (Heinz *et al*, 2018; Zhao *et al*, 2018). Ce défaut de terminaison désorganise complètement la structure de la chromatine (Heinz *et al*, 2018). En effet, la chromatine est organisée en différentes zones dans le noyau en fonction de son niveau d'expression : l'euchromatine étant une région de l'ADN dite « ouverte » car facilement accessible aux facteurs de

transcription et l'hétérochromatine qui est au contraire compactée et pas accessible (Figure 24A). Ainsi, l'ADN forme des boucles qui permettent de regrouper les régions transcriptionnellement actives et inactives entre elles dans certaines zones distinctes du noyau. Cependant, une de ces études a mis en évidence une désorganisation de la chromatine induite par la protéine virale NS1 (Figure 24B). En effet, NS1 séquestrerait le complexe CPSF (« *Cleavage and polyadenylation specificity factor* ») responsable du clivage au niveau du site de polyadénylation (Nemeroff *et al*, 1998), aboutissant à l'inhibition de la terminaison de la transcription. Pol II continuerait donc sa progression le long de la chromatine et la synthèse d'ARNm même après les signaux de polyadénylation, menant ainsi à l'expression d'ARNm aberrants avec une extrémité 3' étendue ou résultant de la fusion de deux gènes. Ce phénomène est appelé « *read-through* ».

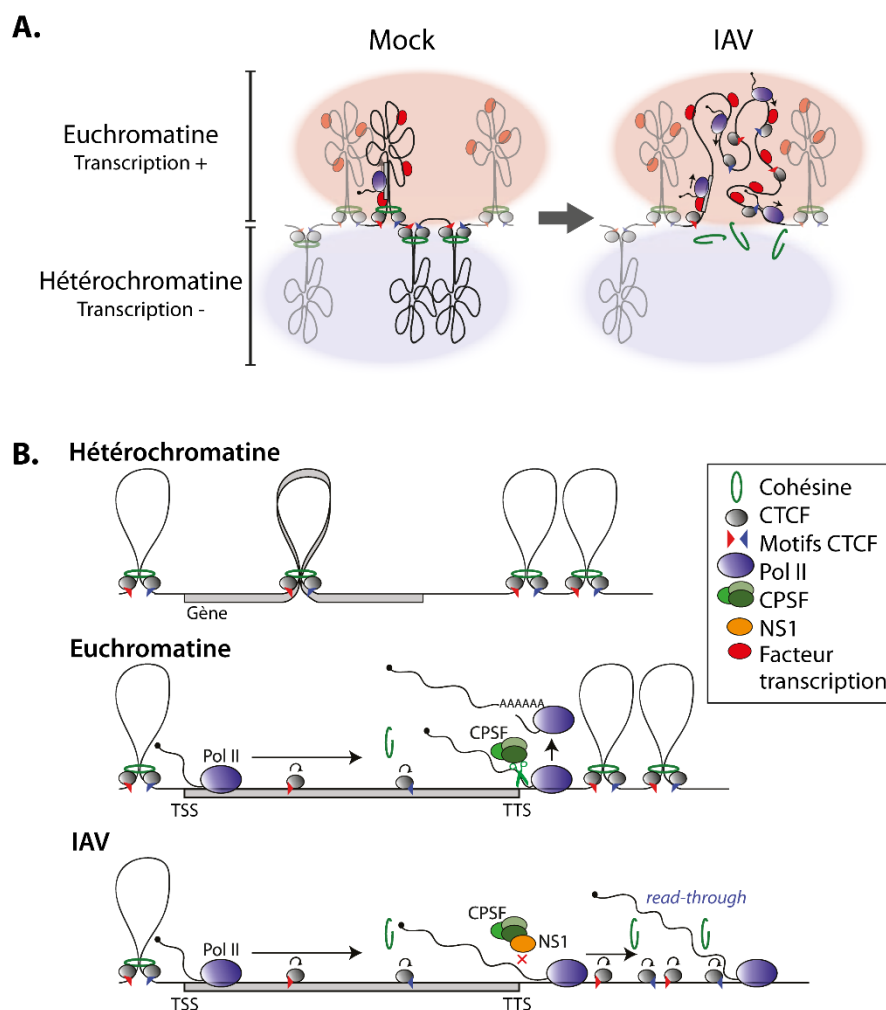


Figure 24 : Schématisation du remodelage de la chromatine induit par l'infection virale

A. Schématisation des mécanismes moléculaires mis en place pour la cohésion de l'ultrastructure de la chromatine et de leur inhibition par les IAV. **B.** Représentation de la désorganisation de la chromatine induite par IAV. Adapté de (Heinz *et al*, 2018).

Tous ces mécanismes mènent ainsi au « *shut-off* » de la cellule hôte en diminuant considérablement la production de protéines cellulaires.

OBJECTIF GÉNÉRAL

L'objectif du travail présenté était de mieux comprendre le fonctionnement de l'ARN-polymérase virale ARN-dépendante (FluPol) dans le noyau de la cellule infectée. Ce travail de thèse se décline en trois parties distinctes :

- Dans une première partie, l'objectif était d'identifier des interactions additionnelles entre FluPol et l'ARN-polymérase II (Pol II) à celle préalablement identifiée entre le CTD de la sous-unité RPB1 de Pol II et FluPol.
- Dans une seconde partie, les interactions de FluPol à la chromatine ont été étudiées pour mieux comprendre les perturbations de l'activité transcriptionnelle cellulaire lors de l'infection.
- Enfin, dans une dernière partie, nous avons cherché à obtenir des nanoligands spécifiques de FluPol. Ces nouveaux outils devaient nous permettre de mieux comprendre les liens entre la flexibilité structurale de FluPol et son fonctionnement.

Partie I

ÉTUDE DES INTERACTIONS ENTRE FLUPOL ET POL II

Objectifs

Afin de réaliser le vol de coiffe, FluPol recrute Pol II. Cette interaction a lieu entre la sous-unité PA de FluPol et le long domaine CTD non structuré de Pol II (Martínez-Alonso *et al*, 2016; Lukarska *et al*, 2017). Les études structurales ont montré que cette interaction est conservée au sein des virus Influenza de type A, B et C bien que la position des sites d'interaction du CTD sur FluPol varie.

Pol II est un complexe constitué de 12 sous-unités (RPB1 à RPB12). Le CTD de Pol II, constitué de 52 répétitions du motif YSPTSPS est non structuré et semble assez éloigné du cœur de Pol II et du canal de sortie des ARNm néosynthétisés. Une interaction supplémentaire avec d'autres sous-unités de Pol II permettrait non seulement un meilleur ancrage de FluPol à Pol II mais également une plus grande proximité aux ARNm en cours de synthèse. Nous avons donc émis l'hypothèse que d'autres interactions pouvaient potentiellement exister entre ces deux complexes polymérases.

Dans le but de cribler de nouvelles interactions entre les complexes FluPol et Pol II, nous avons choisi d'utiliser le système de complémentation binaire « split-luciférase » qui permet de révéler une interaction entre deux protéines exprimées *in cellulo* en mesurant une activité luminescence émise par la luciférase (Munier *et al*, 2013). Chaque protéine à tester est ainsi fusionnée à une moitié N- ou C-terminale de luciférase. Si l'interaction a bien lieu, l'activité de la luciférase est restaurée et peut être mesurée. A l'inverse, si l'interaction est absente, aucun signal de luciférase n'est émis. Après criblage, la ou les interactions identifiées seront validées et caractérisées. Ces expériences sont au cœur de l'article présenté ci-dessous qui va prochainement être soumis pour publication. Des résultats complémentaires non inclus dans ce manuscrit sont détaillés à la suite.

- I. Article : “The influenza virus RNA-polymerase and the host RNA-polymerase II: RPB4 is targeted by a PB2 domain that is involved in viral transcription”

Article

The Influenza Virus RNA-Polymerase and the Host RNA-Polymerase II: RPB4 Is Targeted by a PB2 Domain That Is Involved in Viral Transcription

Jessica Morel ¹, Laura Sedano ¹, Nathalie Lejal ¹, Bruno Da Costa ¹, Eric Batsché ², Christian Muchardt ² and Bernard Delmas ^{1,*}

¹ Unité de Virologie et Immunologie Moléculaires, INRAE, Université Paris-Saclay, 78350 Jouy-en-Josas, France; jessica.morel@inrae.fr (J.M.); laura.sedano@inrae.fr (L.S.); nathalie.lejal@inrae.fr (N.L.); bruno.dacosta@inrae.fr (B.D.C.)

² Biological Adaptation and Ageing (B2A), Institut de Biologie Paris-Seine (IBPS), Sorbonne Université, CNRS, 75005 Paris, France; eric.batsche@sorbonne-universite.fr (E.B.); christian.muchardt@sorbonne-universite.fr (C.M.)

* Correspondence: bernard.delmas@inrae.fr

Abstract: Influenza virus transcription is catalyzed by the viral RNA-polymerase (FluPol) through a cap-snatching activity. The snatching of the cap of cellular mRNA by FluPol is preceded by its binding to the flexible C-terminal domain (CTD) of the RPB1 subunit of RNA-polymerase II (Pol II). To better understand how FluPol brings the 3'-end of the genomic RNAs in close proximity to the host-derived primer, we hypothesized that FluPol may recognize additional Pol II subunits/domains to ensure cap-snatching. Using binary complementation assays between the Pol II and influenza A FluPol subunits and their structural domains, we revealed an interaction between the N-third domain of PB2 and RPB4. This interaction was confirmed by a co-immunoprecipitation assay and was found to occur with the homologous domains of influenza B and C FluPols. The N-half domain of RPB4 was found to be critical in this interaction. Punctual mutants generated at conserved positions between influenza A, B, and C FluPols in the N-third domain of PB2 exhibited strong transcriptional activity defects. These results suggest that FluPol interacts with several domains of Pol II (the CTD to bind Pol II), initiating host transcription and a second transcription on RPB4 to locate FluPol at the proximity of the 5'-end of nascent host mRNA.

Keywords: influenza virus; RNA-polymerase; RNA-polymerase II; protein-protein interaction; PPI; cap snatching; transcription; binary complementation assay



Citation: Morel, J.; Sedano, L.; Lejal, N.; Da Costa, B.; Batsché, E.; Muchardt, C.; Delmas, B. The Influenza Virus RNA-Polymerase and the Host RNA-Polymerase II: RPB4 Is Targeted by a PB2 Domain That Is Involved in Viral Transcription. *Viruses* **2022**, *14*, 518. <https://doi.org/10.3390/v14030518>

Academic Editor: Ayato Takada

Received: 1 February 2022

Accepted: 1 March 2022

Published: 3 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 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 (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Influenza A viruses (IAVs) are important viral respiratory pathogens of humans and are members of the Orthomyxoviridae family. These viruses contain a negative-sense single-stranded segmented RNA genome (reviewed in [1]). The three largest segments encode the subunits of the viral RNA-dependent RNA-polymerase (FluPol), including the two basic proteins PB1 and PB2 and the acidic subunit PA [2]. Unlike what has been observed for many other RNA viruses, the IAV genome is transcribed and replicated in the nucleus of infected cells, wherein the FluPol nucleotide polymerization activity carries out both replication and transcription. During the latter, an additional “cap-snatching” function is implemented to steal short 5'-capped RNA primers from host mRNAs [3].

In the viral particle, the eight viral RNA genomic segments are packaged together with numerous copies of the viral nucleoprotein (NP) and of FluPol in viral ribonucleoprotein complexes (vRNPs). After binding to sialic acid on the cell surface, virions are endocytosed and fused with the endosomal membrane. vRNPs are then released into the cytoplasm and transported to the nucleus where they start synthesizing viral mRNAs to produce viral

proteins. Thus, FluPol transcribes each of the vRNA genome segments, producing positive sense viral mRNA with a 5'-terminal N7-methyl guanosine (m7G) cap and a 3'-polyA tail. FluPol which does not possess capping activity binds to m7G on nascent host RNAs using a cap-binding domain on PB2. Then, it cleaves the host mRNA approximately 10- to 15 nucleotides downstream of the 5'-cap, using the PA endonuclease. The resulting short 5' capped RNA fragment acts as a primer to initiate transcription of the viral genome segments. Polyadenylation is carried out by a FluPol stuttering at a 5'-proximal poly(U) polyadenylation signal present on each genomic vRNA [4]. The complete FluPol transcription activity (initiation, elongation, polyadenylation, and FluPol recycling) has been characterized at near atomic resolution [5,6].

Cap snatching by FluPol requires interaction(s) with the host transcriptional machinery [7–9]. The host RNA-polymerase II (Pol II), a complex of 12 subunits (RPB1–12), synthesizes cellular mRNAs (and diverse noncoding RNAs) which are 5'-m7G capped. Its catalytic center is borne by the two largest subunits (RPB1 and RPB2), while smaller subunits (RPB3 to RPB12) are arrayed at the periphery [10,11]. RPB1 in mammals has a large mobile C-terminal domain (CTD) composed of 52 heptad repeats (consensus sequence Tyr-Ser-Pro-Thr-Ser-Pro-Ser) that recruit factors required for RNA splicing and termination. Residues of the CTD are subjected to post-translational modifications, the most studied being the phosphorylation of Ser2 and/or Ser5. The Ser5P CTD is recognized by the nuclear machinery involved in early steps of mRNA transcription, but remains detectable throughout the gene body [12]. FluPol interacts with RPB1 by binding to its CTD, and more specifically to heptad repeats carrying the Ser5P modification [13]. This interaction has been structurally defined with CTD mimic Ser5P peptides and FluPols of influenza A, B and C viruses [14,15]. These studies document a bipartite binding site on the CTD, while the involved FluPol surfaces appear imperfectly conserved between influenza virus subtypes.

In this study, to better define the molecular mechanisms underlying the FluPol cap-snatching activity, we explored whether FluPol may interact with Pol II via domains or subunits other than the CTD of RPB1. We thus used binary complementation assays to systematically investigate effective interactions between FluPol and Pol II subunits. This approach allowed for identification and validation of an interaction between the third N-ter domain of PB2 and RPB4. This interaction was conserved among FluPol virus types, and mutations in the implicated domain of the viral protein were found to cripple viral transcription. These results suggest that during viral transcription, FluPol associates with the regulatory CTD domain of RPB1, while also interacting with RPB4, at a site where nascent capped RNAs emerge from Pol II. We speculate that this second complementary anchoring position may participate in the docking of FluPol onto nascent RNAs as a step preliminary to the “cap-snatching”.

2. Materials and Methods

2.1. Cells and Virus

HEK-293T cells were maintained in Dulbecco's modified Eagle's Medium (DMEM) supplemented with 10% fetal calf serum (FCS), 2 mM L-glutamine, 100 IU/mL penicillin, 100 mg/mL streptomycin at 37 °C and 5% CO₂. Wild-type (wt) and PB2 mutant viruses were generated by reverse genetics using the 12-plasmid reverse genetics system kindly provided by G. Brownlee [16]. Site-directed mutagenesis was carried out on the PB2 gene by using the QuikChange II site-directed mutagenesis kit (Agilent Technologies France, Les Ulis, France). The viruses were prepared as previously described [17]. Briefly, a 1-day coculture of HEK-293T and MDCK cells (seeding of 3×10^5 and 4×10^5 cells, respectively, in P6 plates) was transfected with a plasmid mixture (0.25 µg per plasmid) using Fugene HD (Promega France, Charbonnières-les-bains, France) according to the manufacturer's recommendations. At 48 h post-transfection, cell supernatants were harvested and used to inoculate MDCK cells for the production of rescued virus stocks. The FluPol genes of the recombinant viruses were sequenced to validate the presence of engineered mutations.

2.2. Protein-Protein Interaction Assays

2.2.1. Split-Luciferase Complementation Assay

A first protein-protein interaction assay was based on the complementation of two trans-complementing fragments of *Gaussia princeps* luciferase (Gluc), Luc1 and Luc2, as described in Cassonnet et al. [18]. Interaction-mediated luciferase activity was measured in cultured cells transiently expressing a protein fused to Gluc1 and another one fused to Gluc2. Codon-optimized of human Pol II cDNAs encoding subunits RPB3, RPB4, RPB5, RPB6, RPB7, RPB8, RPB9, RPB10, RPB11 and RPB12 (all with a HA-tag at their N-terminus) were cloned in pCI-LL-Luc1 and pCI-LL-Luc2 vectors [19], thus resulting to the in frame fusion of the HA-tagged Pol II open reading frame subunits with the Luc1/Luc2 moieties. In these constructs, Pol II subunits are separated from the Luc1/Luc2 moieties by a 10-amino acid long linker (sequence GGGSGGGGS). The plasmid encoding the RPB2 subunit under the control of the CMV promoter (kindly provided by Benoit Coulombe) was used as a template to insert at the RPB2 3'-end the LL-Luc1/Luc2 moieties. The plasmid pFLAG-Pol-II WT encoding human RPB1 (Addgene number 35175, deposited by Benjamin Blencowe) was used to generate two pCI-LL-Luc1/Luc2 derivatives encoding 4 repeats ($4 \times$ YSPTSPS) of the CTD tail fused to Luc1 or Luc2. FluPol PA, PB1 and PB2 subunits cDNA cloned in pCI-LL-Luc1/Luc2 vectors were kindly provided by Nadia Naffakh [19]. Influenza B and C PB2 cDNAs (which were kindly provided by Wendy Barclay and Sylvie van der Werf, respectively), were used to subclone in frame their 5'-domains with Luc1 in pCI-LL-Luc2 with the Gibson Assembly Master Mix Kit (New England BioLabs, Ipswich, Massachusetts, USA). Twenty-six pCI-LL-Luc1/Luc2 derivatives encoding cellular proteins irrelevant in FluPol-Pol II interaction were kindly provided by Caroline Demeret. Point mutations and deletions in FluPol cDNAs were generated using Q5 Site-Directed Mutagenesis Kit (New England BioLabs, Ipswich, MA, USA). HEK-293T cells were seeded at a concentration of 10^5 cells per well in 48-well plates. Twenty-four hours post seeding, cells were transfected in duplicate/triplicate with the indicated combinations of pCI-derived plasmids by using polyethylenimine (Polyscience Inc., Le Perray-en-Yvelines, France). In each well, 150 ng of each pCI-LL-Luc1/Luc2 derivative are co-transfected with 150 ng of an empty pCI vector. In several assays, the empty pCI vector was replaced by pCI vectors encoding wild-type FluPol subunits. Twenty-four hours post-transfection, cells were lysed using the *Renilla* luciferase assay buffer (Promega France, Charbonnières-les-bains, France) for 30 min at room temperature. Next, the luciferase enzymatic activity recovered by the assembly of the Luc1 and Luc2 subunits was measured using *Renilla* luciferase assay reagent and a Tecan Infinite 200 PRO luminometer (*Renilla* luminescence counting program; integration time of 10 s after injection of 50 μ L of the reagent). Normalized luminescence ratios (NLRs) were calculated using pGluc1 and pGluc2 plasmids only encoding *Gaussia* luciferase moieties as described previously by Cassonnet et al. [18]. The NLR for a given interacting protein pair A-B was calculated by dividing the luminescent signal by the sum of the luminescent measured in control cells as indicated in Figure 1A,B. Significance of NLR signals from the noise was determined using GraphPad to identify outliers with a threshold of 1%.

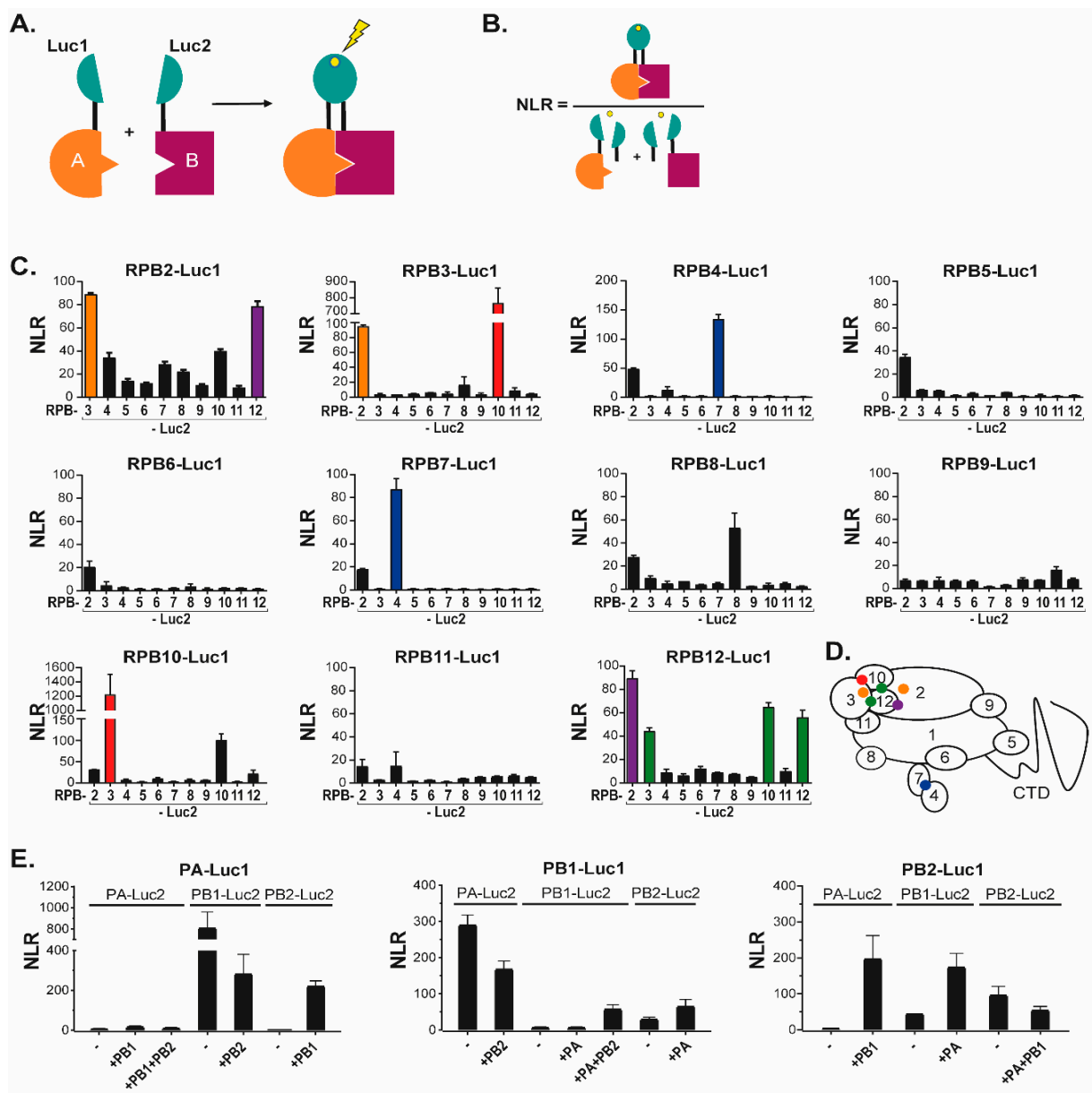


Figure 1. The split-luciferase interaction assay and its validation with Pol II and FluPol subunits. (A) Schematic of the principle of split-luciferase complementation assay. If protein A and protein B interact, the activity of luciferase is restored and luminescence will be emitted in presence of its substrate. (B) The normalized luminescence ratio (NLR) is calculated as described. (C) NLR values obtained for pairs of Luc1/Luc2-tagged Pol II subunits co-expressed into HEK-293T cells. Twenty-four hours after transfection of plasmids driving Pol II subunits expression, cells were lysed and luminescence was measured. Colors indicate reciprocated interactions, orange for RPB2-RPB3 interaction, red for RPB3-RPB10 interaction, blue for RPB4-RPB7 and violet for RPB2-RPB12. The green color indicates one-way interaction detection, RPB12 with RPB3 and RPB10. (D) Interactions are represented by a colored dot on a scheme of Pol II and its subunits as defined in (C). (E) NLR values obtained for pairs of Luc1/Luc2-tagged FluPol subunits plasmids transfected into HEK293T cells with/without untagged FluPol subunits. Data are mean \pm s.d. n = two technical replicates representative of three independent experiments.

2.2.2. Bimolecular Fluorescence Complementation (BiFC) Assay

BiFC constructs were based on the complementation of two trans-complementing fragments of Venus fluorescent protein, Venus1 (amino acids 1 to 155) and Venus2 (amino acids 156 to 239), as described by Shyu et al. [20]. Interaction-mediated fluorescence is observed in cultured cells when a protein A fused to Venus1 interacts with a protein B fused to Venus2. Thus, Pol II and FluPol subunits cDNAs were cloned using adequate restriction sites in pCI-LL-Venus1 and -Venus2 vectors, thus resulting to the in frame fusion of the cloned open reading frames followed by a 10-amino acid long linker (sequence GGGSGGGGS) with the Venus1/2 moieties. Cells were seeded on coverslips in P12-wells and transfected 24 h later with combinations of pCI-LL-Venus1/2 derivatives (1 µg/plasmid). Twenty-four hours after transfection, cells were fixed using 4% paraformaldehyde for 10 min and permeabilized for 15 min using 0.1% Triton X-100. Nuclei were stained with DAPI. The excitation peak used to reveal Venus fluorescence was 515 nm and its emission was revealed at 528 nm. Images were captured with a confocal microscope and processed with the ImageJ software.

2.3. Minireplicon and Transcription Assays

2.3.1. Minireplicon Assay

HEK-293T cells were seeded in P96 wells (5×10^4 cells/well) and transfected one day later with plasmids driving expression of wild-type or mutated forms of the PA, PB1, PB2, and NP viral proteins with the plasmid pPolI-WSN-NA-firefly luciferase as previously described [21]. Plasmid pMAX-GFP (Lonza, Walkersville, MD, USA) was used as an internal control for transfection efficiency and data normalization. As a negative control, 293T cells were transfected with polyethylenimine (PEI) with the same plasmids minus the PA expression plasmid. Forty-eight hours post-transfection, cells were lysed in 150 µL PLB buffer (30 mM Tris pH7.9, 10 mM MgCl₂, 1% Triton X-100, 20% Glycerol, 1 mM DTT). Luminescence activity was measured with Luciferase Assay System (Promega) on a Tecan Infinite M200Pro luminometer according to the manufacturer's instructions. Replicon activity was quantified by a ratio between the luminescence and the fluorescence signals.

2.3.2. Transcription Assay

To quantify transcription activity of FluPol mutants, the same procedure than described in the minireplicon assay was used, except that the wild-type PA subunit was replaced by the PA C95A mutant that is deficient for replication [22].

2.4. Co-Immunoprecipitation Assay

HEK-293T cells plated in 6-well plates were transiently transfected with 1 µg of pci-PB2(36-247)-FLAGtag and/or with 1 µg pci-HAtag-RPB4 using PEI. Twenty-four hours post-transfection, cells were rinsed and scrapped at 4 °C into 1 mL of 50 mM Tris-HCl pH7.4, 2 mM EDTA, 150 mM NaCl, 1% Triton X-100, 10% glycerol supplemented with protease inhibitor (Thermo Fisher Scientific, Les Ulis, France). After 1 h, this material was centrifugated (20 min, $12,000 \times g$ at 4 °C). Supernatants were collected and incubated under gentle agitation overnight at 4 °C with 80 µL of a 1:1 slurry of sepharose-IgG beads (GE Healthcare, Chicago, IL, USA) previously coated with anti-FLAG M5 antibody (Sigma-Aldrich, St. Louis, MO, USA). The beads were washed three times with 1 mL of lysis Buffer and once with PBS, then treated for 2 min at 100 °C in Laemmli's denaturing buffer plus 5% 2-mercaptoethanol, and centrifugated. Resulting supernatants were subjected to SDS-PAGE followed by a western blot (WB) analysis. An anti-HA antibody conjugated with peroxidase (Roche Diagnostics GmbH, Mannheim, Germany) and an anti-FLAG antibody conjugated with peroxidase (Sigma-Aldrich, Darmstadt, Germany) were used to reveal immunoprecipitated products in the WB.

3. Results

3.1. Novel Interactions between FluPol and Pol II Subunits

To identify novel sites of interactions between FluPol and Pol II subunits, we used the previously described *Gaussia princeps* luciferase-based complementation assay [18,23]. In this assay, an interaction between two proteins each fused to either the Luc1 or the Luc2 segments of the *Gaussia* luciferase enzyme, results in reconstitution of a functional luciferase activity, which can be quantified by addition of substrate (Figure 1A). To favor proper folding of the FluPol and Pol II subunits, the luciferase segments were fused onto their C terminus. For each sample, the specificity of the interaction over background was estimated by calculating a normalized luminescent ratio (NLR—Figure 1B).

To validate the assay, we first fused each Pol II subunit (from RPB2 to RPB12) with either Luc1 or Luc2, then co-expressed the Luc1-fusions with each Luc2-fusion (Figure 1C). This approach identified strong and reciprocal contacts between RPB2 and RPB3, RPB2 and RPB4, RPB2 and RPB12, RPB3 and RPB10, and RPB4 and RPB7. Interaction signal was also identified between RPB12 and RPB3 or RPB10, albeit in a non-reciprocal manner. Finally, RPB8, RPB10, and RPB12 were found to interact with themselves. Confrontation of these observations with the 3D-structure of Pol II [24] illustrated that the luciferase complementation assay was very efficient at detecting direct interactions between Pol II subunits, while it was oblivious to interactions mediated by the Pol II complex, possibly due to the shortness of the GGGSGGGS linker (summarized by the color code in Figure 1D).

We next used a similar approach to monitor interactions between FluPol subunits (Figure 1E). In these experiments, co-expressed Luc1/Luc2-tagged forms of PA and PB1 yielded NLR signal both in the presence, and in the absence of untagged PB2. In contrast, interaction of PB2 with PA or PB1 was detected only when the three FluPol subunits were co-expressed. As for Pol II, these interactions between FluPol subunits were in agreement with the known 3D-structures of the viral enzyme and also matched earlier observations on the sequential assembly of in the cytoplasm and in the nucleus [11,25,26].

Finally, we investigated interactions between FluPol subunits (PA, PB1 or PB2) of the WSN strain (H1N1) and Pol II subunits including RPB2 to RPB12 (Figure 2A,C), and a fragment of the RPB1 CTD. As expected from structural analyses [14], the RPB1 CTD was found to interact with PA alone and with the assembled FluPol (Figure 2B). With other Pol II subunits, the range of NLR values was 1- to 2-logs of magnitude lower than those observed for internal Pol II and FluPol interactions (Figure 2C). For each tagged FluPol subunits, RPB3 and RPB4 (and to a less extent RPB7 and RPB8) yielded the highest NLR signals when in the presence of an intact FluPol complex. These data were indicative of proximity interactions between FluPol and Pol II not accounted for by RPB1 CTD (and might possibly involve RPB3 and RPB4). We noted further that, for the PA bait, the FluPol-Pol II interaction-signals benefited from co-expressed untagged PB1 but not from coexpression of untagged PB2. Likewise, NLR signals obtained with tagged PB1 and Pol II subunits remained unaffected upon co-expression of untagged PA and PB2. Finally, interactions of the PB2 bait with Pol II subunits were affected by the presence of other FluPol subunits only for RPB4 and RPB7. These data showed that interactions between FluPol and Pol II subunits were not strictly dependent on the assembly of FluPol into an intact complex, and suggested that the assay is amenable to analyze individually the implication of structurally-defined protein domains of FluPol.

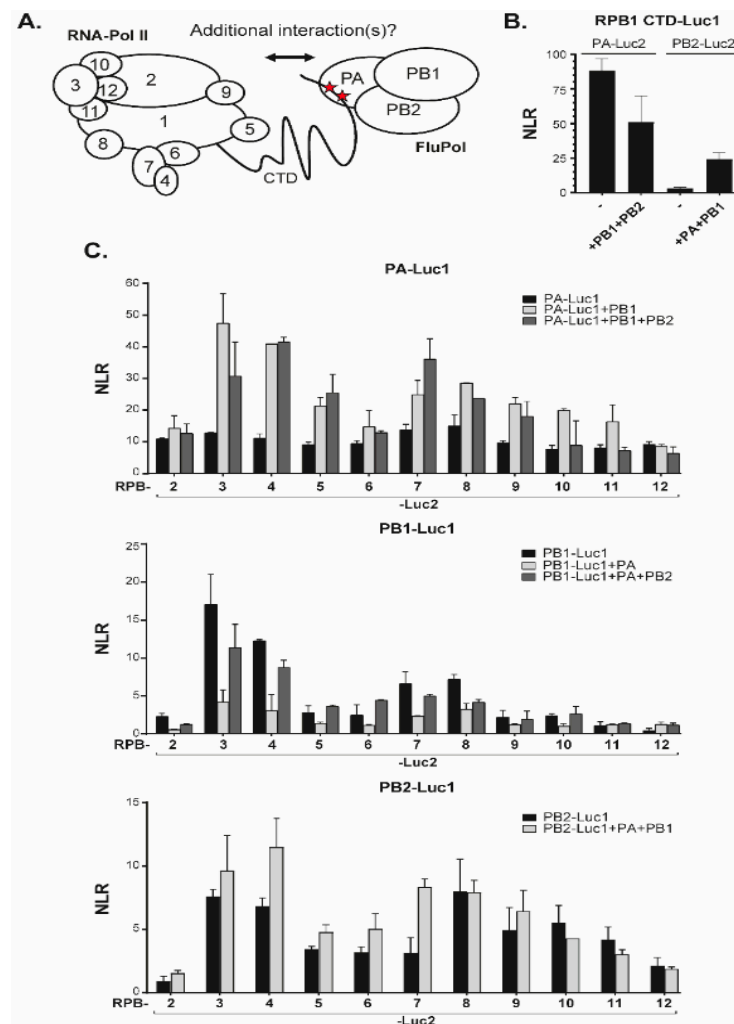


Figure 2. Split-luciferase assays between FluPol and Pol II subunits. **(A)** Schematic of FluPol and Pol II with their subunits and the interaction between the CTD of Pol II with FluPol. Red stars indicate the YSPSPS CTD motif (with the phosphorylated Serine 5). **(B)** A plasmid encoding a CTD stretch of RPB1 fused the Luc1 moiety was co-transfected with plasmids encoding PA or PB2 fused to the Luc2 moiety. Plasmids encoding untagged PA, PB1, or PB2 were also added when indicated to produce complete FluPol. Twenty-four hours post-transfection, cells were lysed and luminescence signal was measured. **(C)** Pairs of plasmids encoding a Luc1-FluPol subunit with a Luc2-Pol II subunit were transfected into HEK 293T along with/without untagged FluPol subunits. NLR signals were quantified as described in Figure 1B. Data are mean \pm s.d. $n =$ two technical replicates representative of two or three independent assays.

3.2. The N-Terminal Region of PB2 Mediates Interaction with RPB4

FluPol can be subdivided in functional and structural protein domains. It associates an invariant core comprising the PB1 subunit, stabilized by the PA-linker and the PA-C [amino acid 197 to the C-terminus] domain from the PA subunit, and the PB2-N [amino acid 1 to 247] from the PB2 subunit. This core is then associated with flexibly-linked peripheral domains including the PA endonuclease domain and the two-third C-terminal domain of PB2] [27]. To map the domains of FluPol principally involved in interactions with Pol II, we generated constructs driving expression of corresponding sub-fragments of PA and PB2. Most of these domains were shown to acquire their native fold when expressed without other FluPol components [2]. In complementation assays with the Pol II subunits, the PA-derived constructs co-expressed with PB1 did not reveal any clear association with Pol II subunits, all combinations yielding very low NLR signals (Figure 3A). In contrast,

co-expression of the core-associated PA domains with a full set of Pol II subunits revealed a strong interaction (NLR score at 85) between PB2-N (residues 1 to 247) and RPB4 which was not observed with other Pol II subunits (Figure 3B). Furthermore, the PB2-C (residues 247 to 759) construct also yielded high NLR scores (approximately 30) with RPB5 and RPB11.

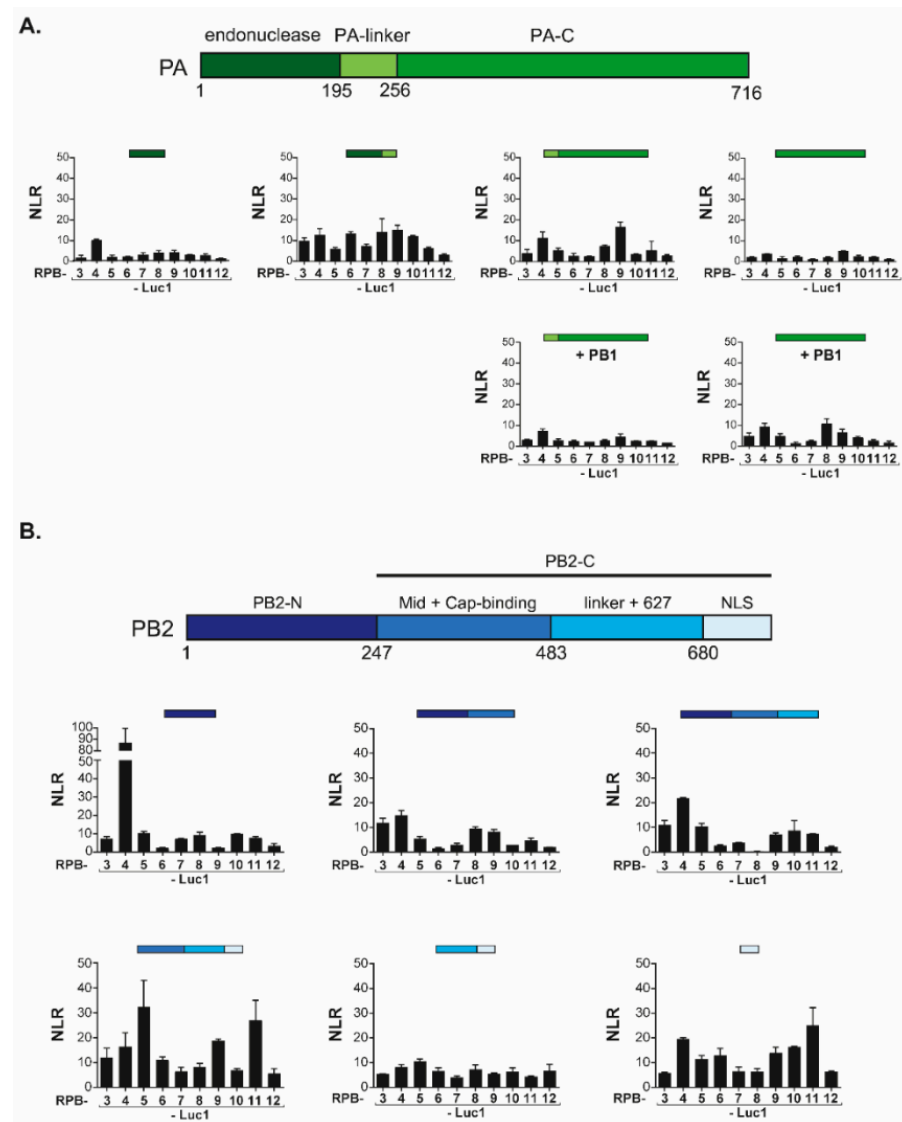


Figure 3. Split-luciferase assays between PA and PB2 domains of FluPol and Pol II subunits. (A) Luc1-PA domains and Luc2-subunits of Pol II were co-expressed into HEK293T cells along with PB1 or an empty plasmid. Twenty-four hours post-transfection, cells were lysed and luminescence was measured. NLR were determined as shown in Figure 1B. (B) Luc1-PB2 domains and Luc2-subunits of Pol II were co-expressed into HEK293T cells with NLR determined as shown in Figure 1B. Data are mean \pm s.d. n = two technical replicates.

To validate the specificity of the PB2-N/RPB4 interaction, we next challenged the PB2-N construct with a set of proteins unrelated to Pol II. In these experiments, neither PB2-N, nor a shorter PB2(36–247) construct lacking 35 N-ter residues interacting with PB1 [28], displayed any interaction with the mock targets (Figure 4A). Together, these results were indicative of a specific contact between host Pol II subunit RPB4 and the PB2-N domain while also suggesting auxiliary contacts between RPB5 and RPB11 and the C-terminal regions of PB2.

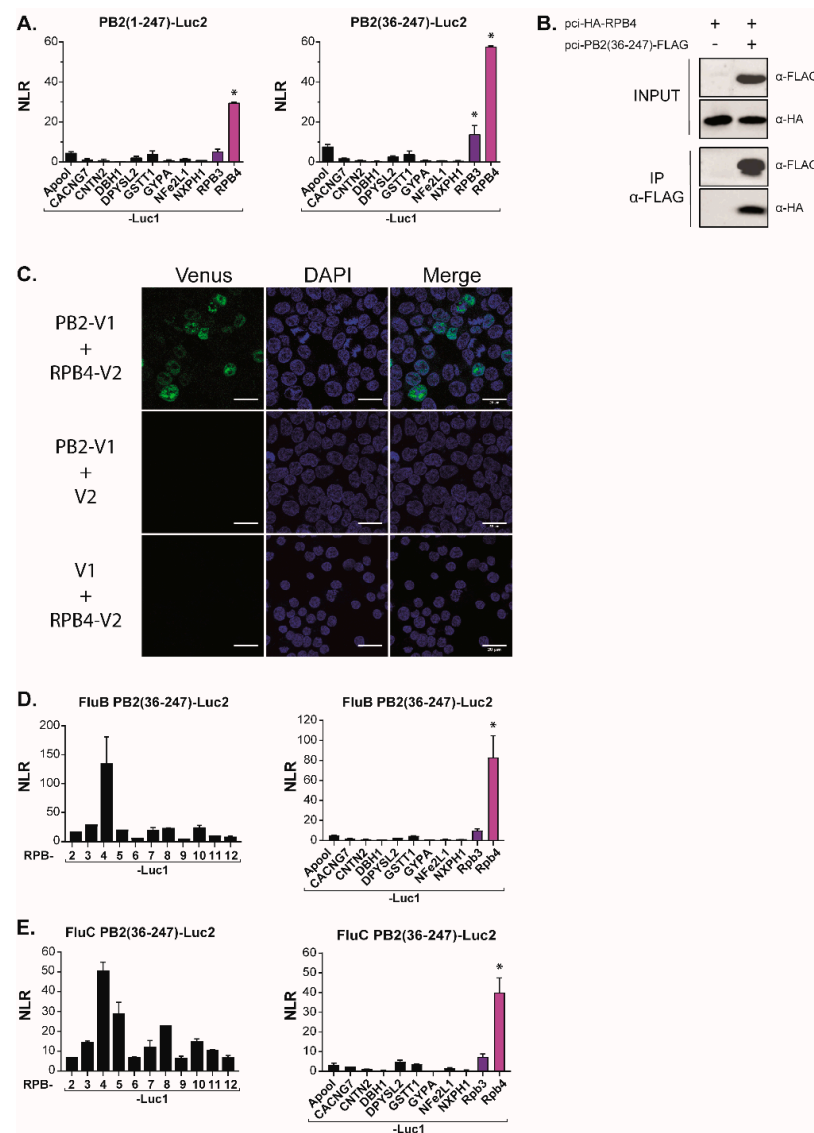


Figure 4. PB2-N domain interacts with the Pol II subunit RPB4. (A) Split-luciferase assays between PB2-N forms (Luc2-PB2(1–247) and Luc2-PB2(36–247) and Luc1-tagged irrelevant proteins (Apool for MICOS complex subunit MIC27, UniProtKB ID Q6UXV4; CACNG7 for Voltage-dependent calcium channel gamma-7 subunit, P62955; CNTN2 for Contactin-2, Q02246; DBH1 for dopamine beta-hydroxylase, P09172; DPYSL2 for dihydropyrimidinase-related protein 2, Q16555; GSTT1 for glutathione S-transferase theta-1, P30711; GYP A for glycophorin-A, P02724; NFE2L1 for endoplasmic reticulum membrane sensor NFE2L1, Q14494 and NXP1 for Neurexophilin-1 for P58417). Data are mean \pm s.d. $n =$ two technical replicates. (B) Coimmunoprecipitation assay. Plasmids encoding HA-RPB4 and PB2(36–247)-FLAG were transfected in HEK 293T cells. Twenty-four hours post-transfection, cells were lysed, PB2(36–247)-FLAG was immunoprecipitated and RPB4 was revealed by Western Blotting. (C) Split-Venus complementation assay between PB2 and RPB4. HEK 293T cells were co-transfected with expression plasmids encoding PB2-Venus1 (PB2-V1), RPB4-Venus 2 (RPB4-V2) and/or with plasmids encoding Venus1 (V1) and Venus2 (V2) moieties. Cells were fixed 24 h later, nuclei were marked with DAPI and fluorescence emitted by reconstituted Venus was measured by confocal microscopy. Bars = 20 μ m. (D,E) The interaction between PB2(36–247) and RPB4 is conserved in influenza types B and C. Plasmids encoding luc2-tagged forms of PB2(36–247) from influenza B (D) or influenza C (E) viruses were co-expressed with Luc1-tagged Pol II subunits in HEK 293T cells. Twenty-four hours post transfection, cells were lysed and luminescence was measured. NLR values were calculated as described in Figure 1. Data are means \pm s.d. $n =$ two technical replicates representative of two independent assays. * indicates outliers with a $p < 0.01$.

3.3. The PB2RBP4 Interaction Domain Is Essential for the Virus Life Cycle

We next wished to probe the robustness and the biological relevance of the PB2/RBP4 interaction. Co-immunoprecipitation-western assays with epitope-tagged forms of RBP4 and of the PB2(36–247) construct allowed for further narrowing down the region of PB2 interacting with RBP4 while also demonstrating that the interaction resisted the relatively stringent conditions characterizing these assays (150 mM NaCl, 1% Triton X-100; Figure 4B). To further test the interaction under *in vivo* conditions, we carried out a second binary complementation assay, with PB2 and RBP4, respectively, tagged with the N- and C-moieties of the Venus fluorescent protein (Figure 4C). Co-transfection of these constructs into HEK293T cells resulted in a strong nuclear fluorescence signal, establishing that PB2 can interact with RBP4 under these conditions while also illustrating that this interaction occurs in the appropriate cellular compartment.

To investigate whether the PB2/RBP4 interaction was confirmed among different influenza virus types, we also constructed a luciferase complementation assay with the PB2 of influenza B and C viruses. In this assay, the influenza B PB2(36–247) construct interacted with RBP4 but not with any other Pol II subunit (nor with any of the unrelated target proteins described above (Figure 4D)). Likewise, the influenza C construct was associated with RBP4. However, it also yielded a signal with additional Pol II subunits, possibly due to indirect interactions (Figure 4E). In this assay, this influenza C PB2 domains did not display any interaction with the non-relevant control proteins.

To further map regions involved in the PB2/RBP4 interaction, we next expressed a series of truncations of PB2-N. When tested in the split-luciferase system, several discrete regions of the PB2 amino acids, stretching from amino acid 130 to 247, yielded a signal with RBP4 (Figure 5). Reciprocally, truncations of RBP4 identified the N-ter moiety of this protein (amino acids 1–72) as essential for the interaction (Figure 5C). Interestingly, when mapped on the 3D-structures of FluPol and RBP4 [26,29], the protein regions identified in these two sets of deletion mutants matched a series of helices located at the surface on the two proteins at a position highly compatible with an implication in protein–protein interactions.

Finally, to address the biological function of the PB2 (130–247) region, we engineered a series of fourteen point-mutants on residues conserved among influenza A, B, and C viruses (Figure 6A). The activity of these mutants was then assessed in two assays, respectively measuring either the resultant of both replication and transcription (R+T) or just transcription (T) (Figure 6B,C). In the latter assay, wild-type PA was replaced by a C95A mutant, rendering FluPol defective for replication [22]. Both assays relied on a pPol1-WSN-NA-firefly luciferase plasmid producing a viral RNA (vRNA) in which the coding sequence of the NA gene was replaced by the firefly luciferase gene. This plasmid was transfected into HEK-293T cells together with expression vectors for the appropriate PA construct, the wild-type PB1 and NP constructs, and the various PB2 mutants. The luciferase enzymatic activity measured in extracts of transfected cells and reflecting either R+T or T activities dependent on the used PA construct (WT or C95A mutant), revealed that most of the PB2 mutations, with the exception of R144A and E188A, strongly decreased FluPol activity in both assays. These observations indicate that this region of PB2 and its conserved residues are critical for viral transcription (Figure 6C). Finally, to definitively establish the importance of the PB2-N region in the virus cycle, we selected residues 115, 142, and 144 for alanine substitution. Using plasmid-driven reverse genetics, we successfully recovered virus mutants for each of these substitutions. While the R144A virus was found to replicate at high titers (1.2×10^7 pfu/mL in cell culture medium after transfection), the Y115A and the R142A mutants replicated poorly (6×10^4 pfu/mL), in accordance with the results obtained in the replication/transcription assays.

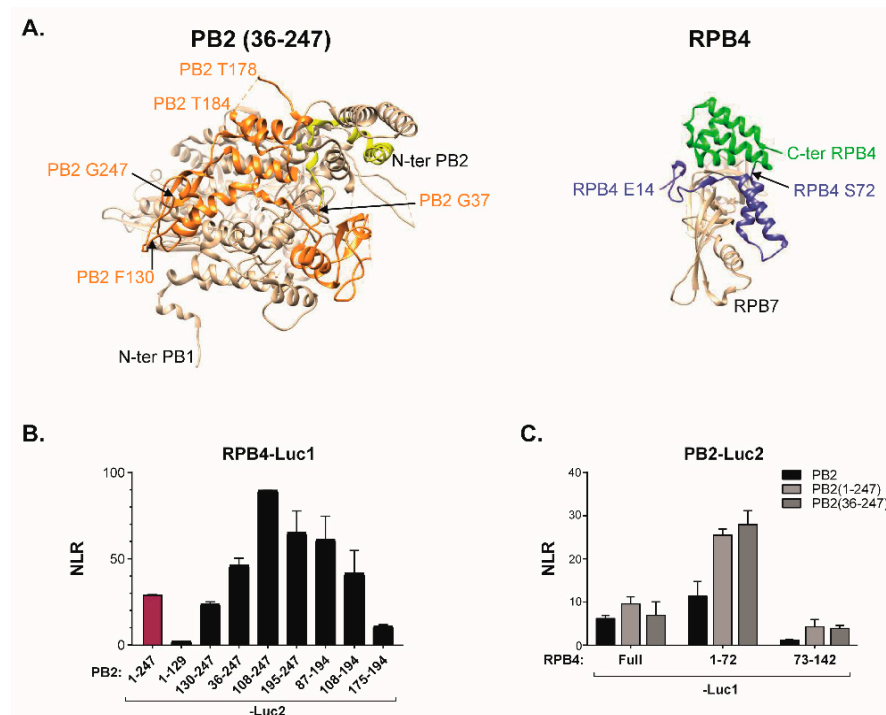


Figure 5. Identification of the PB2 and RPB4 amino acids stretches involved in the FluPol-RPB4 interaction. (A) Structures of the 36–247 domain of PB2 with the PB1 adjacent domain (PDB ID: 6QPF) and of the RPB4/RPB7 dimer (PDB ID: 2C35). (B,C) Split-luciferase complementation assay. (B) Luc2-tagged stretches of PB2(1–247) and Luc1-RPB4 were co-transfected into HEK293T cells. 24 h after transfection, cells were lysed and luminescence was measured. (C) The Luc1-tagged N- or C-moieties of RPB4 were co-expressed with the Luc2-tagged PB2 subunit or with N-ter derived forms of PB2 into HEK293T cells. After cell lysis, luminescence was measured and NLR calculated as in Figure 1B. Data are mean ± s.d. n = two technical replicates representative of three independents assays.

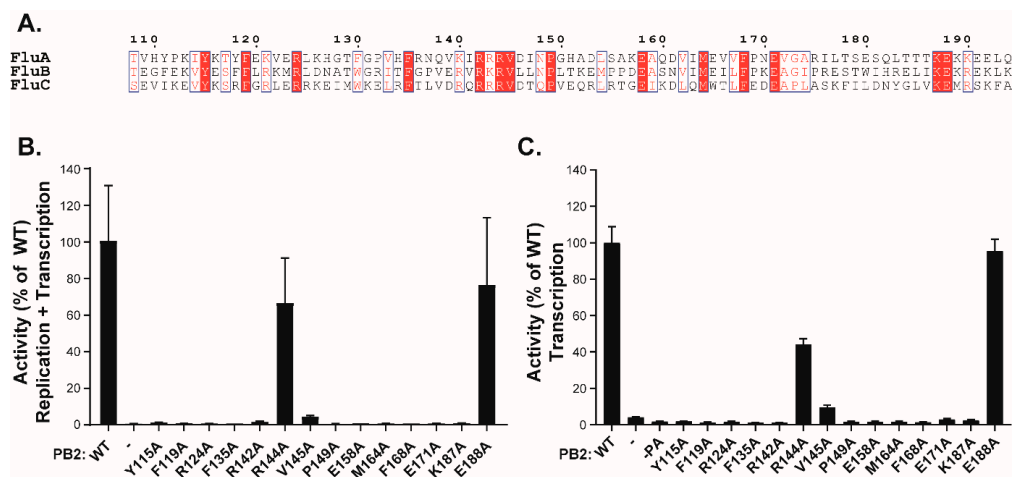


Figure 6. Transcription/replication activity of FluPol with PB2 mutants. (A) Multi-alignment with the PB2 domains (residues 108 to 194) of influenza A, B, and C viruses made with ESPrict [30]. Strictly conserved residues are white on a red background, and partially conserved residues are red. (B) Plasmids expressing NP, PB1, PA, and wild type or PB2 mutants were cotransfected in 293T cells together with the reporter plasmid WSN-NA-firefly luciferase, allowing the quantification of the polymerase activity. In panel (C), the same procedure than in (B) was used, except that the PA C95A mutant deficient in replicase activity [15] was used instead of the wt PA subunit. Data are mean ± n = four technical replicates representative of two independent assays.

4. Discussion

The interaction between FluPol and the Pol II complex through the CTD of its RPB1 subunit has been well-established both functionally and structurally [9]. The interplay between the two polymerases has been recognized to allow cap-snatching on nascent mRNAs then priming of the viral mRNA synthesis. However, the question concerning whether the docking of FluPol to the nascent capped RNA relied solely on RBP1 or involved additional protein-protein interactions had never been elucidated. In this study, we used *in vitro* binary complementation assays to screen for auxiliary interactions between FluPol and Pol II subunits. We found that, in addition to RPB1, FluPol associated with RPB4, one of the twelve Pol II subunits, which is structurally connected to the core of Pol II via RPB7. This interaction, which was confirmed in co-immunoprecipitation assays, was mapped to the N-terminal third of PB2 and was found to be conserved among influenza B and C subtypes. Together, these observations suggest a critical role for multiple FluPol-Pol II interactions in the life cycle of IAVs.

The RPB4/RPB7 heterodimer forms a stalk-like protrusion from the Pol II core complex below a mobile clamp domain, adjacent to the CTD of RPB1 and the RNA exit channel, a position with potential for interactions with multiple components such as CTD-bound factors and modifying enzymes and emerging nascent RNAs [31–37]. In fact, the RPB4/RPB7 stalk has been proposed to act as an additional landing platform for transcription-related factors, some of which also interact with the CTD [38,39]. It is therefore tempting to speculate that the binding of FluPol to RPB4 synergizes with the CTD-mediated interaction to target nascent capped RNAs.

Despite low sequence similarity between the N-terminal third of PB2 among influenza A, B and C viruses, these homologous domains were all found to interact with the RPB4 subunit. When mutating conserved residues in this domain, twelve of fourteen substitutions deeply impacted the transcription activity of FluPol. This documented the functional implication of the N-ter PB2 domain in reaching an efficient interaction with the host transcription machinery. Therefore, we propose a model (Figure 7) in which FluPol first targets Pol II via its PA-CTD interaction and then binds the RPB4 subunit via PB2 to edge closer to the RNA exit tunnel of Pol II. This would bring the cap-binding domain of PB2 in close proximity to nascent mRNAs, thereby allowing the cap-snatching by FluPol to occur.

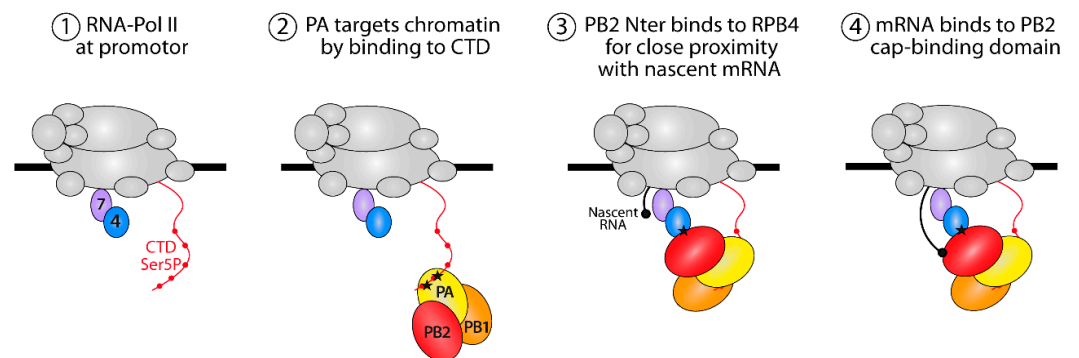


Figure 7. Proposed model for the interaction interplay between FluPol and Pol II. Pol II is represented in gray with RPB4 and RPB7 subunits in blue and lavender, respectively. FluPol first recognizes the YSPTSpPS motifs (with the serine phosphorylated at position 5 indicated by red dot) in the CTD of RPB1 to bind the RPB4 subunit located at the vicinity of the Pol II pore which extrudes nascent mRNA during transcription. PB2 binds the cap (black dot) of the mRNA by its cap-binding domain. Black asterisks indicate FluPol binding sites on the Pol II CTD and RPB4.

What could the time-line of the FluPol-RPB4/RPB7 dimer interaction be? Nascent transcript capping occurs immediately after the emergence of the 5'-end triphosphate from the RNA exit tunnel and is concomitant with promoter-proximal pausing of RNA Pol II bound to NELF and DSIF [40,41]. Shortly after cap completion, the interaction of the

cap-binding complex (CBC) with a NELF subunit correlates with its targeting to the paused RNA Pol II complex and increases by 100–200-fold its affinity for the modified 5' end of the transcript. Since the CBC has an affinity for the cap structure that is higher than that of FluPol, we propose that FluPol binding to Pol II via RPB1 CTD and RPB4/RPB7 interferes with CBC recruitment or with its access to nascent 5'-capped Pol II transcripts. Alternatively, FluPol may inhibit the binding of NELF to Pol II or its association with CBC. These hypotheses are currently under investigation and will allow for better understanding of how FluPol efficiently controls capping of viral mRNAs.

Author Contributions: Conceptualization, B.D. and C.M.; methodology, B.D. and J.M.; validation, J.M., N.L. and B.D.; formal analysis, J.M. and B.D.; investigation, J.M., L.S., N.L., B.D.C. and E.B.; resources, C.M. and B.D.; data curation, J.M. and B.D.; writing—original draft preparation, J.M. and B.D.; writing—review and editing, J.M., C.M. and B.D.; visualization, J.M.; supervision, B.D.; project administration, B.D.; funding acquisition, C.M. and B.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the French Agence Nationale de la Recherche (ANR), grant number ANR-17-CE18-0006-01 and the INRAE Animal Health Division.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data supporting reported results can be provided by B.D.

Acknowledgments: We thank G. Brownlee for the gift of the reverse genetics system of the IAV WSN strain, Benoit Coulombe for the gift of a plasmid encoding human RPB2, Benjamin Blencowe for the plasmid encoding human RPB1 (Addgene number 35175), Nadia Naffakh for the gift of plasmids pCI-Luc1/Luc2, Wendy Barclay for a plasmid encoding the PB2 subunit of influenza B, Sylvie Van der Werf for a plasmid encoding the influenza C PB2 subunit and Caroline Demeret for plasmids encoding irrelevantLuc1/Luc2 tagged proteins. We thank Nicolas Meunier for his help in statistical analyses. C.M. and B.D. acknowledge support of the ANR-17-CE18-0006-01 program. J.M. acknowledges fellowships of the ANR program and the Animal Health Division of INRAE.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study, in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Palese, P.; Shaw, M.L. Orthomyxoviridae: The Viruses and Their Replication. In *Fields Virology*, 6th ed.; Wolters Kluwer Health/Lippincott Williams & Wilkins: Philadelphia, PA, USA, 2013; Volume 1, pp. 1151–1185.
2. Boivin, S.; Cusack, S.; Ruigrok, R.W.H.; Hart, D.J. Influenza A Virus Polymerase: Structural Insights into Replication and Host Adaptation Mechanisms. *J. Biol. Chem.* **2010**, *285*, 28411–28417. [[CrossRef](#)] [[PubMed](#)]
3. Te Velthuis, A.J.W.; Fodor, E. Influenza Virus RNA Polymerase: Insights into the Mechanisms of Viral RNA Synthesis. *Nat. Rev. Microbiol.* **2016**, *14*, 479–493. [[CrossRef](#)] [[PubMed](#)]
4. Poon, L.L.; Pritlove, D.C.; Fodor, E.; Brownlee, G.G. Direct Evidence That the Poly(A) Tail of Influenza A Virus MRNA Is Synthesized by Reiterative Copying of a U Track in the Virion RNA Template. *J. Virol.* **1999**, *73*, 3473–3476. [[CrossRef](#)] [[PubMed](#)]
5. Krischuns, T.; Lukarska, M.; Naffakh, N.; Cusack, S. Influenza Virus RNA-Dependent RNA Polymerase and the Host Transcriptional Apparatus. *Annu. Rev. Biochem.* **2021**, *90*, 321–348. [[CrossRef](#)] [[PubMed](#)]
6. Wandzik, J.M.; Kouba, T.; Karupphasamy, M.; Pflug, A.; Drncova, P.; Provaznik, J.; Azevedo, N.; Cusack, S. A Structure-Based Model for the Complete Transcription Cycle of Influenza Polymerase. *Cell* **2020**, *181*, 877–893.e21. [[CrossRef](#)] [[PubMed](#)]
7. Lamb, R.A.; Choppin, P.W. Synthesis of Influenza Virus Polypeptides in Cells Resistant to Alpha-Amanitin: Evidence for the Involvement of Cellular RNA Polymerase II in Virus Replication. *J. Virol.* **1977**, *23*, 816–819. [[CrossRef](#)] [[PubMed](#)]
8. Engelhardt, O.G.; Smith, M.; Fodor, E. Association of the Influenza A Virus RNA-Dependent RNA Polymerase with Cellular RNA Polymerase II. *JVI* **2005**, *79*, 5812–5818. [[CrossRef](#)]
9. Walker, A.P.; Fodor, E. Interplay between Influenza Virus and the Host RNA Polymerase II Transcriptional Machinery. *Trends Microbiol.* **2019**, *27*, 398–407. [[CrossRef](#)]
10. Cramer, P.; Bushnell, D.A.; Kornberg, R.D. Structural Basis of Transcription: RNA Polymerase II at 2.8 Angstrom Resolution. *Science* **2001**, *292*, 1863–1876. [[CrossRef](#)]
11. Cramer, P. Organization and Regulation of Gene Transcription. *Nature* **2019**, *573*, 45–54. [[CrossRef](#)]

12. Nojima, T.; Gomes, T.; Grosso, A.R.F.; Kimura, H.; Dye, M.J.; Dhir, S.; Carmo-Fonseca, M.; Proudfoot, N.J. Mammalian NET-Seq Reveals Genome-Wide Nascent Transcription Coupled to RNA Processing. *Cell* **2015**, *161*, 526–540. [[CrossRef](#)] [[PubMed](#)]
13. Martínez-Alonso, M.; Hengrung, N.; Fodor, E. RNA-Free and Ribonucleoprotein-Associated Influenza Virus Polymerases Directly Bind the Serine-5-Phosphorylated Carboxyl-Terminal Domain of Host RNA Polymerase II. *J. Virol.* **2016**, *90*, 6014–6021. [[CrossRef](#)] [[PubMed](#)]
14. Lukarska, M.; Fournier, G.; Pflug, A.; Resa-Infante, P.; Reich, S.; Naffakh, N.; Cusack, S. Structural Basis of an Essential Interaction between Influenza Polymerase and Pol II CTD. *Nature* **2017**, *541*, 117–121. [[CrossRef](#)]
15. Serna Martin, I.; Hengrung, N.; Renner, M.; Sharps, J.; Martínez-Alonso, M.; Masiulis, S.; Grimes, J.M.; Fodor, E. A Mechanism for the Activation of the Influenza Virus Transcriptase. *Molecular. Cell* **2018**, *70*, 1101–1110.e4. [[CrossRef](#)]
16. Fodor, E.; Devenish, L.; Engelhardt, O.G.; Palese, P.; Brownlee, G.G.; García-Sastre, A. Rescue of Influenza A Virus from Recombinant DNA. *J. Virol.* **1999**, *73*, 9679–9682. [[CrossRef](#)]
17. Le Goffic, R.; Bouguyon, E.; Chevalier, C.; Vidic, J.; Da Costa, B.; Leymarie, O.; Bourdieu, C.; Decamps, L.; Dhorne-Pollet, S.; Delmas, B. Influenza A Virus Protein PB1-F2 Exacerbates IFN-Beta Expression of Human Respiratory Epithelial Cells. *J. Immunol.* **2010**, *185*, 4812–4823. [[CrossRef](#)] [[PubMed](#)]
18. Cassonnet, P.; Rolloy, C.; Neveu, G.; Vidalain, P.-O.; Chantier, T.; Pellet, J.; Jones, L.; Muller, M.; Demeret, C.; Gaud, G.; et al. Benchmarking a Luciferase Complementation Assay for Detecting Protein Complexes. *Nat. Methods* **2011**, *8*, 990–992. [[CrossRef](#)]
19. Da Costa, B.; Sausset, A.; Munier, S.; Ghounaris, A.; Naffakh, N.; Le Goffic, R.; Delmas, B. Temperature-Sensitive Mutants in the Influenza A Virus RNA Polymerase: Alterations in the PA Linker Reduce Nuclear Targeting of the PB1-PA Dimer and Result in Viral Attenuation. *J. Virol.* **2015**, *89*, 6376–6390. [[CrossRef](#)] [[PubMed](#)]
20. Shyu, Y.J.; Liu, H.; Deng, X.; Hu, C.-D. Identification of New Fluorescent Protein Fragments for Bimolecular Fluorescence Complementation Analysis under Physiological Conditions. *Biotechniques* **2006**, *40*, 61–66. [[CrossRef](#)]
21. Leymarie, O.; Jouvion, G.; Hervé, P.-L.; Chevalier, C.; Lorin, V.; Lecardonnel, J.; Da Costa, B.; Delmas, B.; Escriou, N.; Le Goffic, R. Kinetic Characterization of PB1-F2-Mediated Immunopathology during Highly Pathogenic Avian H5N1 Influenza Virus Infection. *PLoS ONE* **2013**, *8*, e57894. [[CrossRef](#)]
22. Nilsson-Payant, B.E.; Sharps, J.; Hengrung, N.; Fodor, E. The Surface-Exposed PA51-72-Loop of the Influenza A Virus Polymerase Is Required for Viral Genome Replication. *J. Virol.* **2018**, *92*, e00687-18. [[CrossRef](#)] [[PubMed](#)]
23. Munier, S.; Rolland, T.; Diot, C.; Jacob, Y.; Naffakh, N. Exploration of Binary Virus–Host Interactions Using an Infectious Protein Complementation Assay. *Mol. Cell Proteom.* **2013**, *12*, 2845–2855. [[CrossRef](#)] [[PubMed](#)]
24. Cramer, P. RNA Polymerase II Structure: From Core to Functional Complexes. *Curr. Opin. Genet. Dev.* **2004**, *14*, 218–226. [[CrossRef](#)]
25. Huet, S.; Avilov, S.V.; Ferbitz, L.; Daigle, N.; Cusack, S.; Ellenberg, J. Nuclear Import and Assembly of Influenza A Virus RNA Polymerase Studied in Live Cells by Fluorescence Cross-Correlation Spectroscopy. *JVI* **2010**, *84*, 1254–1264. [[CrossRef](#)]
26. 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](#)]
27. Pflug, A.; Lukarska, M.; Resa-Infante, P.; Reich, S.; Cusack, S. Structural Insights into RNA Synthesis by the Influenza Virus Transcription-Replication Machine. *Virus Res.* **2017**, *234*, 103–117. [[CrossRef](#)] [[PubMed](#)]
28. Sugiyama, K.; Obayashi, E.; Kawaguchi, A.; Suzuki, Y.; Tame, J.R.H.; Nagata, K.; Park, S.-Y. Structural Insight into the Essential PB1-PB2 Subunit Contact of the Influenza Virus RNA Polymerase. *EMBO J.* **2009**, *28*, 1803–1811. [[CrossRef](#)] [[PubMed](#)]
29. Meka, H. Crystal Structure and RNA Binding of the Rpb4/Rpb7 Subunits of Human RNA Polymerase II. *Nucleic Acids Res.* **2005**, *33*, 6435–6444. [[CrossRef](#)] [[PubMed](#)]
30. Robert, X.; Gouet, P. Deciphering Key Features in Protein Structures with the New ENDscript Server. *Nucleic Acids Res.* **2014**, *42*, W320–W324. [[CrossRef](#)]
31. Allepuz-Fuster, P.; Martínez-Fernández, V.; Garrido-Godino, A.I.; Alonso-Aguado, S.; Hanes, S.D.; Navarro, F.; Calvo, O. Rpb4/7 Facilitates RNA Polymerase II CTD Dephosphorylation. *Nucleic Acids Res.* **2014**, *42*, 13674–13688. [[CrossRef](#)]
32. Allepuz-Fuster, P.; O'Brien, M.J.; González-Polo, N.; Pereira, B.; Dhoondia, Z.; Ansari, A.; Calvo, O. RNA Polymerase II Plays an Active Role in the Formation of Gene Loops through the Rpb4 Subunit. *Nucleic Acids Res.* **2019**, *47*, 8975–8987. [[CrossRef](#)] [[PubMed](#)]
33. Kimura, M.; Suzuki, H.; Ishihama, A. Formation of a Carboxy-Terminal Domain Phosphatase (Fcp1)/TFIIF/RNA Polymerase II (Pol II) Complex in *Schizosaccharomyces Pombe* Involves Direct Interaction between Fcp1 and the Rpb4 Subunit of Pol II. *Mol. Cell Biol.* **2002**, *22*, 1577–1588. [[CrossRef](#)]
34. Mitsuzawa, H.; Kanda, E.; Ishihama, A. Rpb7 Subunit of RNA Polymerase II Interacts with an RNA-Binding Protein Involved in Processing of Transcripts. *Nucleic Acids Res.* **2003**, *31*, 4696–4701. [[CrossRef](#)] [[PubMed](#)]
35. Chen, Y.; Yamaguchi, Y.; Tsugeno, Y.; Yamamoto, J.; Yamada, T.; Nakamura, M.; Hisatake, K.; Handa, H. DSIF, the Paf1 Complex, and Tat-SF1 Have Nonredundant, Cooperative Roles in RNA Polymerase II Elongation. *Genes Dev.* **2009**, *23*, 2765–2777. [[CrossRef](#)] [[PubMed](#)]
36. Armache, K.-J.; Mitterweger, S.; Meinhart, A.; Cramer, P. Structures of Complete RNA Polymerase II and Its Subcomplex, Rpb4/7. *J. Biol. Chem.* **2005**, *280*, 7131–7134. [[CrossRef](#)]
37. Calvo, O. RNA Polymerase II Phosphorylation and Gene Looping: New Roles for the Rpb4/7 Heterodimer in Regulating Gene Expression. *Curr. Genet.* **2020**, *66*, 927–937. [[CrossRef](#)] [[PubMed](#)]

38. Choder, M. Rpb4 and Rpb7: Subunits of RNA Polymerase II and Beyond. *Trends Biochem. Sci.* **2004**, *29*, 674–681. [[CrossRef](#)]
39. Richard, S.; Gross, L.; Fischer, J.; Bendalak, K.; Ziv, T.; Urim, S.; Choder, M. Numerous Post-Translational Modifications of RNA Polymerase II Subunit Rpb4/7 Link Transcription to Post-Transcriptional Mechanisms. *Cell Rep.* **2021**, *34*, 108578. [[CrossRef](#)]
40. Vos, S.M.; Pöllmann, D.; Caizzi, L.; Hofmann, K.B.; Rombaut, P.; Zimniak, T.; Herzog, F.; Cramer, P. Architecture and RNA Binding of the Human Negative Elongation Factor. *eLife* **2016**, *5*, e14981. [[CrossRef](#)]
41. Yamaguchi, Y.; Takagi, T.; Wada, T.; Yano, K.; Furuya, A.; Sugimoto, S.; Hasegawa, J.; Handa, H. NELF, a Multisubunit Complex Containing RD, Cooperates with DSIF to Repress RNA Polymerase II Elongation. *Cell* **1999**, *97*, 41–51. [[CrossRef](#)]

II. Expériences complémentaires

1) Recherche des acides aminés impliqués dans l'interaction PB2 – RPB4

Dans la continuité des expériences menées, l'effet des mutations en alanine des acides aminés conservés du domaine 108-194 de PB2 a été testé en système split-luciférase afin de mesurer leur impact sur l'interaction PB2 – RPB4 (Figure 25). Le domaine PB2(36-247) couplé à Luc2 portant ces mutations ponctuelles a donc été co-exprimé avec RPB4 couplée à Luc1 en cellules HEK293T. La plupart des mutations n'impacte pas l'interaction. Seul le mutant M164A montre une diminution d'environ 50% de l'interaction, mais celle-ci n'est pas significative ($p=0,0592$, One-way ANOVA). De plus, dans la structure cristallographique de FluPol, ce résidu M164 est situé à l'interface entre les trois sous-unités PA, PB1 et PB2, particulièrement en interaction avec les résidus S702 et Y705 de la sous-unité PB1 de FluPol, ce qui le rend peu disponible pour être engagé dans une autre interaction.

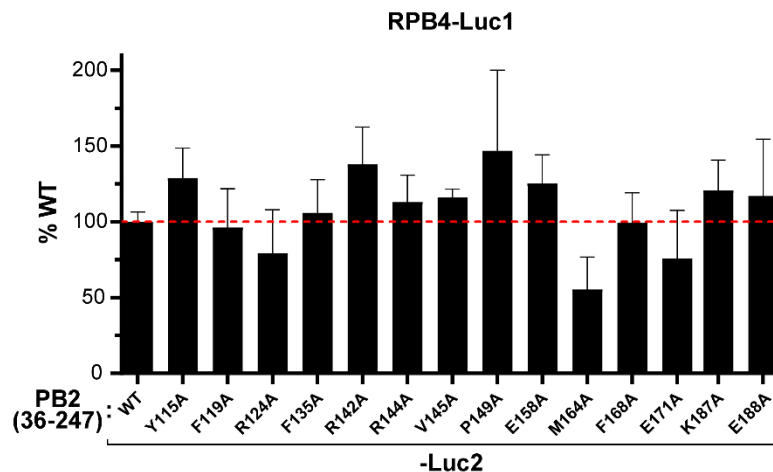


Figure 25 : Test de l'effet des mutations des résidus conservés du domaine PB2(108-194) de PB2 sur l'interaction PB2 – RPB4

Les cellules HEK293T ont été transfectées avec les plasmides pci-RPB4-Luc1 et pci-PB2(36-247)-Luc2 portant les différentes mutations des acides aminés étudiés en triplicata. Vingt-huit heures post-transfection, les cellules sont lysées, le substrat ajouté et la luminescence émise mesurée. Les résultats sont présentés sous forme de moyenne +/- écart-type et sont exprimés en pourcentage par rapport au PB2 WT dont le niveau est représenté par la ligne en pointillés rouge.

Des mutations additionnelles ont été testées, notamment sur des positions non conservées mais proches de la M164 ainsi que sur RPB4 mais aucune n'a inhibé l'interaction PB2 – RPB4 (résultats non présentés). Afin d'aiguiller notre recherche des résidus qui pourraient être impliqués dans cette interaction, une étude de modélisation *in silico* de cette interaction a été menée.

2) Modélisation de l'interaction PB2 – RPB4

Parallèlement aux expériences menées *in cellulo*, nous avons cherché à modéliser l'interaction entre PB2 et RPB4 (Figure 26A) *in silico* en suivant la démarche décrite en figure 26B. Pour se faire, nous avons fait appel au Dr. Gwenaëlle André-Leroux, une spécialiste en modélisation de l'unité MalAGE à l'INRAE de Jouy-en-Josas. Une première étape d'exploration globale des interactions possibles a été entreprise entre les protagonistes. Celle-ci a été effectuée sans *a priori*, c'est-à-dire sans donner de zones d'interaction supposées, dans le but d'avoir une exploration conformationnelle la plus exhaustive possible et ne pas biaiser les résultats *in silico* avec nos données expérimentales. Ainsi, les interactions entre FluPol (choisie dans sa conformation active pour la transcription) et RPB4 ont été explorées alors que chacune de ces protéines est considérée en corps rigide. Le domaine « core » structurellement invariable de FluPol constitué du domaine N-terminal de PB2 (résidus 1 à 247) identifié dans l'interaction avec RPB4, de PB1 et du domaine C-terminal de PA a été retenu pour le premier criblage. Ainsi, les domaines périphériques mobiles (le domaine N-terminal de PA portant l'activité endonucléase et le domaine C-terminal de PB2) n'ont pas été intégrés à l'analyse. Vingt modèles de complexes « core » de FluPol - RPB4 sont obtenus suite à cette étape d'exploration avec le serveur pour « docking » de protéines GramMX (Tovchigrechko & Vakser, 2006). Lorsque les modèles d'interaction sont redondants (i.e. à une distance de moins de 2 Å), le complexe donnant le meilleur score d'interaction GramMX est conservé. A ce stade, nous avons donc des complexes montrant des sites d'interactions distincts, aussi bien sur PA que sur PB1 ou le domaine N-terminal de PB2. FluPol interagissant avec RPB4 en complexe avec RPB7, les modèles offrant des sites d'interactions incompatibles avec FluPol entière et RPB4 en complexe avec RPB7 sont éliminés. Seuls quatre modèles FluPol – RPB4 ne comportant aucune gêne stérique ni avec les domaines périphériques de FluPol, ni avec RPB7 ont été conservés. Ils ont servi de modèle de départ pour une optimisation grâce à un affinement local effectué avec Rosetta (Lyskov *et al*, 2013).

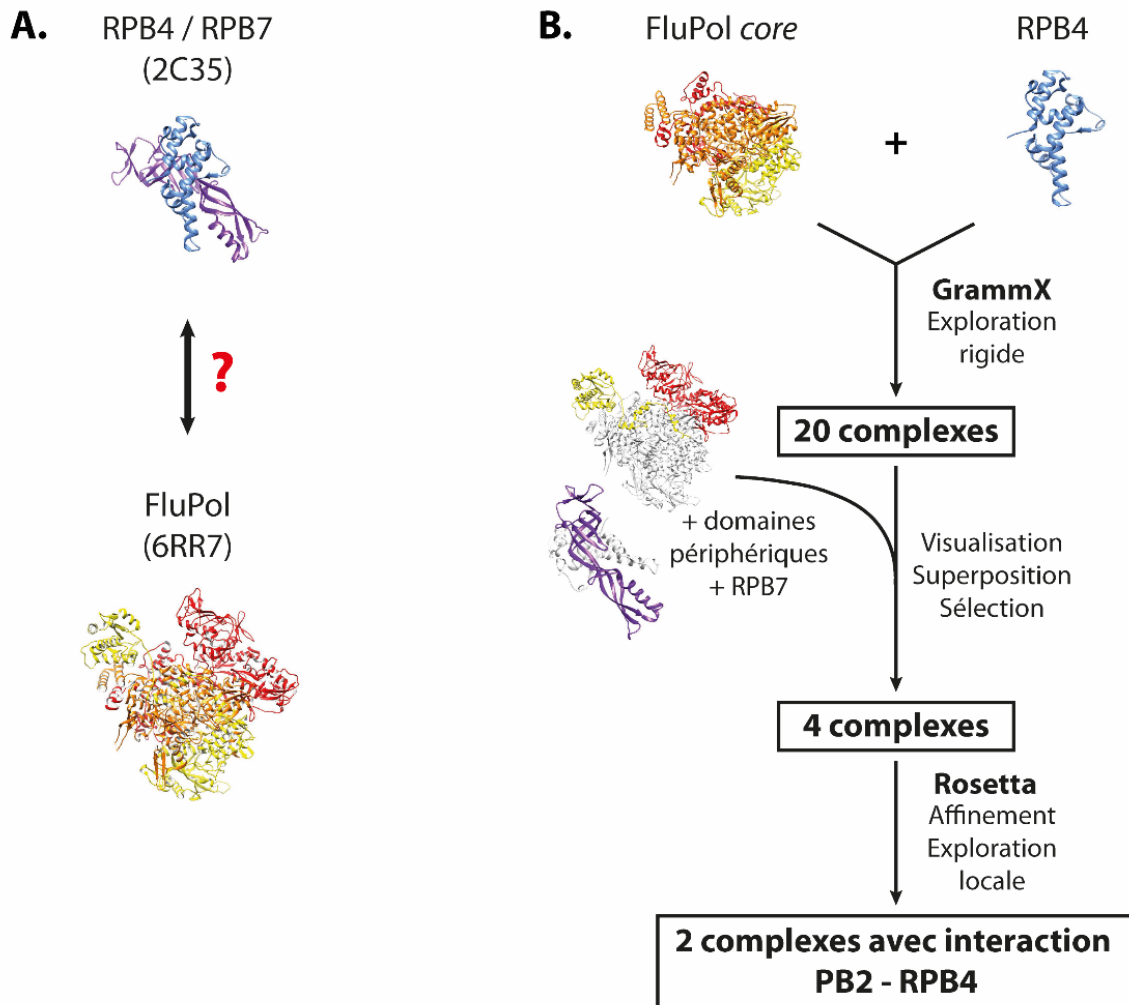


Figure 26 : Démarche de modélisation

- A.** Les possibilités d'interaction entre le dimère RPB4 / RPB7 (PDB 2C35) et FluPol d'un IAV en conformation « transcription » (PDB 6RR7) sont explorées.
- B.** Détails de la démarche suivie pour modéliser l'interaction RPB4 / FluPol

Après analyse, deux modèles de RPB4 dont l'interaction engage PB2 sont obtenus : les modèles 12 et 17 présentés sur la Figure 27. On peut ainsi observer que les zones d'interaction prédites sont proches mais restent relativement différentes pour les deux modèles. RPB4 montre une interface de contact sensiblement identique impliquant les boucles N-terminales 14-30 et 90-94, et particulièrement les résidus polaires et chargés E14, K24 et K91 conservés dans les deux modèles. Ces deux modèles d'interaction sont aussi à l'interface entre PB2 et PB1. Cependant, les zones d'interaction côté FluPol sont différentes, sauf pour le résidu R584 de PB1 qui est impliqué dans les deux modèles.

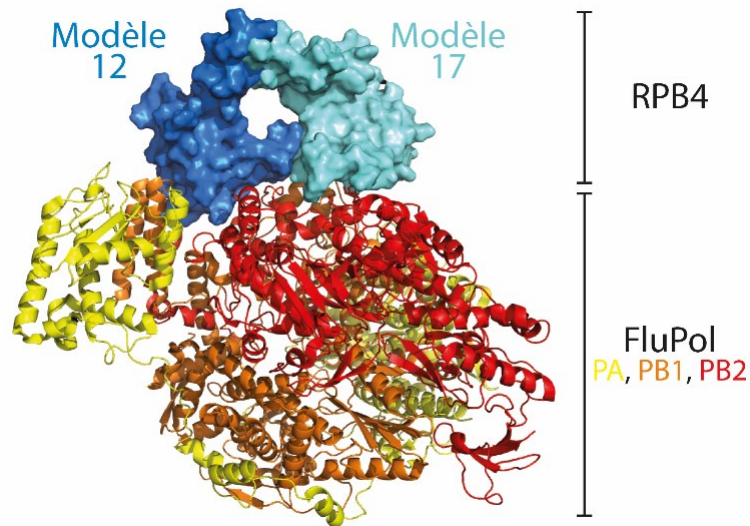


Figure 27 : Modélisation de l'interaction FluPol – RPB4 *in silico*

Structures de FluPol (PDB 6RR7) et de RPB4 (PDB 2C35) dans les modèles 12 et 17. FluPol est représentée en rubans selon le code couleur suivant : PA en jaune, PB1 en orange et PB2 en rouge. La sous-unité RPB4 est représentée en surface dans différentes nuances de bleu selon le modèle (bleu foncé pour le modèle 12 et cyan pour le modèle 17).

Dans le modèle 12 (Figure 28), RPB4 interagit avec le résidu M105 et la zone 194-198 de PB2 mais également avec quelques résidus de PB1 (résidus R584 et R734) présents de part et d'autre du site d'interaction. Dans ce modèle, RPB4 est positionnée parallèlement à la surface de FluPol, n'engendrant donc pas de gêne stérique. Une partie de l'interaction est médiée par la partie N-terminale de RPB4, en particulier la Lysine 24 au niveau de PB2. Cette région de RPB4 pourrait également interagir avec PB1. Cependant, une partie importante du domaine C-terminal de RPB4 interagit également avec PB2, ce qui semble peu en accord avec nos données expérimentales qui montrent une perte totale d'interaction lorsque RPB4 (73-142) est exprimée (Figure 5 de l'article).

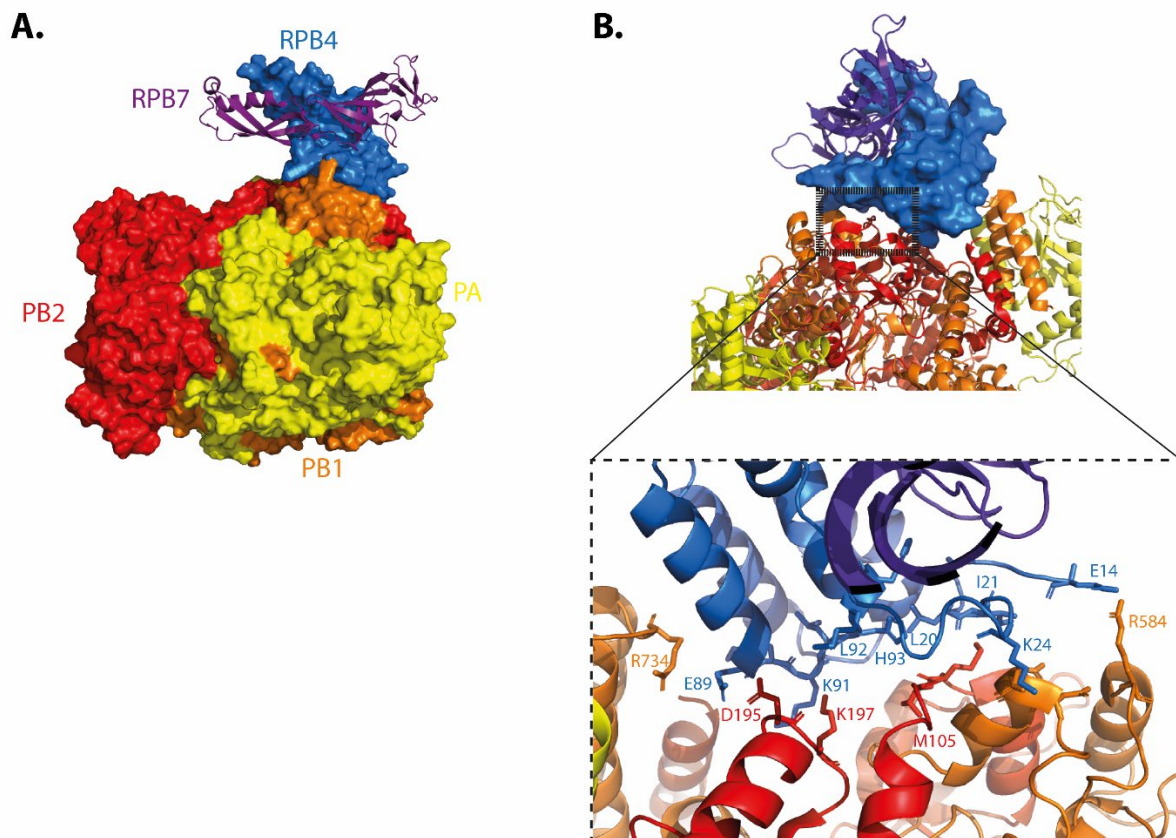


Figure 28 : Modélisation de l'interaction – Modèle 12

A. Structures de FluPol (PDB 6RR7) et du dimère RPB4 / RPB7 (PDB 2C35) dans le modèle 12. FluPol et RPB4 sont représentées en surface alors que RPB7 est représentée en rubans.

B. Zoom sur l'interface d'interaction PB2 – RPB4. FluPol et RPB4 sont représentées en surface alors que RPB7 est représentée en rubans. Dans l'encadré, toutes les protéines sont représentées en rubans et les positions des acides aminés impliqués dans l'interaction sont indiquées.

Code couleur : PA en jaune, PB1 en orange et PB2 en rouge ; RPB4 en bleu foncé et RPB7 en violet.

Dans le modèle 17 (Figure 29), l'interaction de RPB4 se fait conjointement entre PB2 et PB1. Comme montré avec le zoom, l'extrémité N-terminale de RPB4 est principalement engagée dans l'interaction avec PB2, ce qui est plus en accord avec nos données expérimentales. L'interaction au niveau de PB2 est localisée entièrement sur la zone 120-130 sur la face latérale d'une hélice- α . PB1 semble également bien impliquée dans l'interaction, particulièrement avec l'extrémité C-terminale de RPB4.

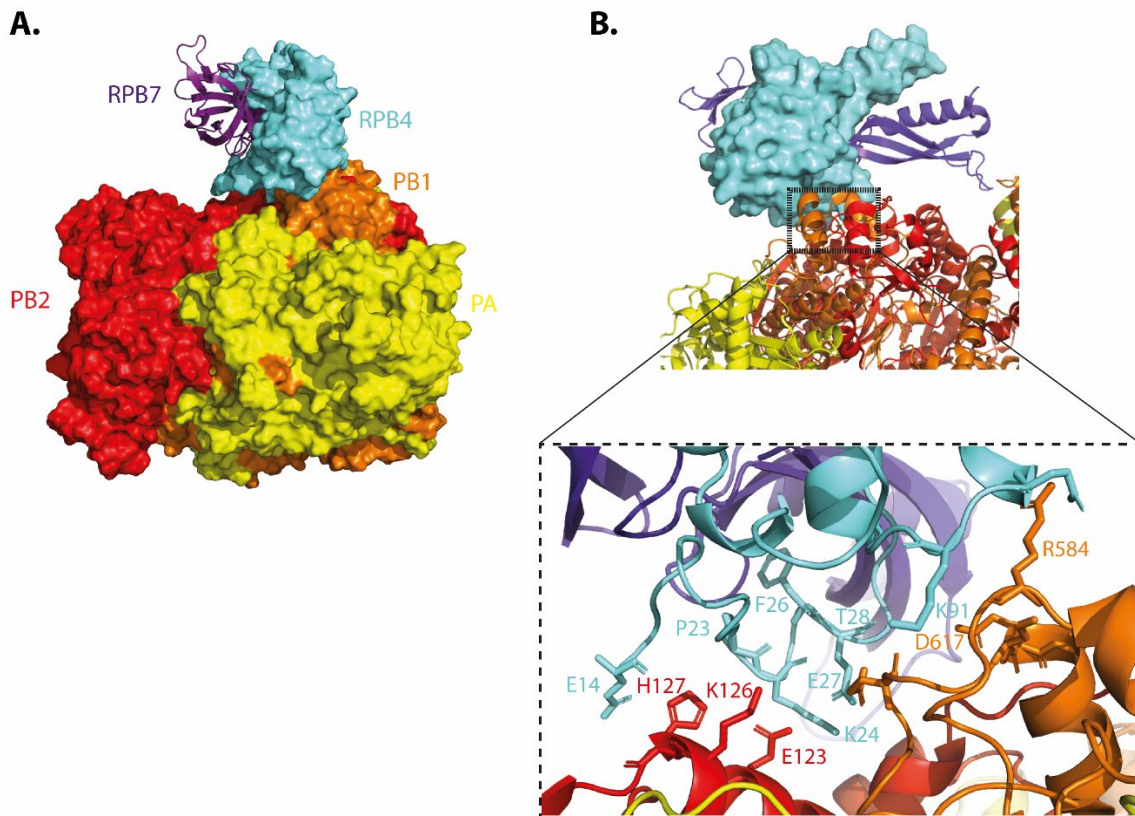


Figure 29 : Modélisation de l'interaction – Modèle 17

A. Structures de FluPol (PDB 6RR7) et du dimère RPB4 / RPB7 (PDB 2C35) dans le modèle 17. FluPol et RPB4 sont représentée en surface alors que RPB7 est représentée en rubans.

B. Zooms sur l'interface d'interaction PB2 – RPB4. FluPol et RPB7 sont représentées en rubans alors que RPB4 est représentée en surface. Dans l'encadré, toutes les protéines sont représentées en rubans et les positions des acides aminés impliqués dans l'interaction sont indiquées.

Code couleur : PA en jaune, PB1 en orange et PB2 en rouge ; RPB4 en cyan et RPB7 en violet.

Conclusion

Ces travaux ont permis de mettre en évidence une interaction entre les complexes FluPol et Pol II impliquant le domaine N-terminal de PB2 et la sous-unité RPB4 de Pol II.

Cette interaction a été mise en évidence par un crible en split-luciférase puis confirmée par co-immunoprécipitation et localisée au noyau *via* le système split-Venus. Une caractérisation de cette interaction a été menée *in cellulo* pour essayer de déterminer précisément les zones d'interaction sur un plan moléculaire, ainsi que les résidus engagés dans cette interaction. La stratégie a donc été de déterminer le plus petit domaine d'interaction de PB2 possible puis de muter les résidus conservés au sein de ce domaine. Le potentiel d'interaction de ces mutants a ensuite été mesuré en split-luciférase. Le mutant M164A a montré une diminution de l'interaction PB2 – RPB4 mais ce résidu est déjà engagé dans une interaction avec PB1. De nombreux mutants ont été testés dans la région voisine de ce résidu mais sans résultat. Devant la difficulté d'une recherche à l'aveugle des acides aminés qui pourraient être impliqués, des essais de modélisation de cette interaction *in silico* ont été réalisés.

Deux modèles d'interaction entre le domaine N-terminal de PB2 et RPB4 ont été obtenus et doivent encore être confrontés *in cellulo*. Les modèles 12 et 17 ciblent tous les deux le domaine 36-247 de PB2 identifié dans cette étude, avec des zones d'interaction voisines. Cependant, le domaine C-terminal de RPB4 est également impliqué dans l'interaction avec FluPol dans ces deux modèles alors que nos données suggèrent que seul le domaine N-terminal est impliqué, ou en tous cas, que celui-ci est suffisant pour qu'il y ait interaction. Le modèle 17 semble donc le modèle le plus compatible avec nos données car le domaine C-terminal de RPB4 n'est pas engagé dans l'interaction avec PB2 mais plutôt avec PB1. De plus, pour chacun des modèles, RPB4 n'interagit pas uniquement avec PB2 mais également avec PB1. Une interaction complémentaire de RPB4 avec PB1 n'est cependant pas un motif d'exclusion des modèles car les trois sous-unités de FluPol sont étroitement entrelacées. De plus, le dimère RPB4 / RPB7 interagit également avec PA et PB1 sur le crible des interactions entre FluPol et Pol II.

Si l'on positionne les sites d'interactions du CTD de Pol II sur PA (Sites 1 et 2, Figure 30), on remarque que RPB4 du modèle 17 est plus proche de ces derniers que le modèle 12. Cette observation est intéressante et peut avoir son importance car une structure de Pol II de *Schizosaccharomyces pombe* (*S. pombe*) a révélé une interaction entre le CTD et l'hélice $\alpha 2$ de RPB4, ce dernier permettant une stabilisation du CTD (PDB 3H0G, (Spåhr *et al*, 2009)). La sous-unité RPB4 ayant une structure conservée entre *S. pombe* et l'Homme, cette région correspondrait aux résidus 33 à 50 de RPB4

humaine. Il semblerait donc que le CTD et RPB4 aient une forte proximité qui pourrait être représentée dans le modèle 17.

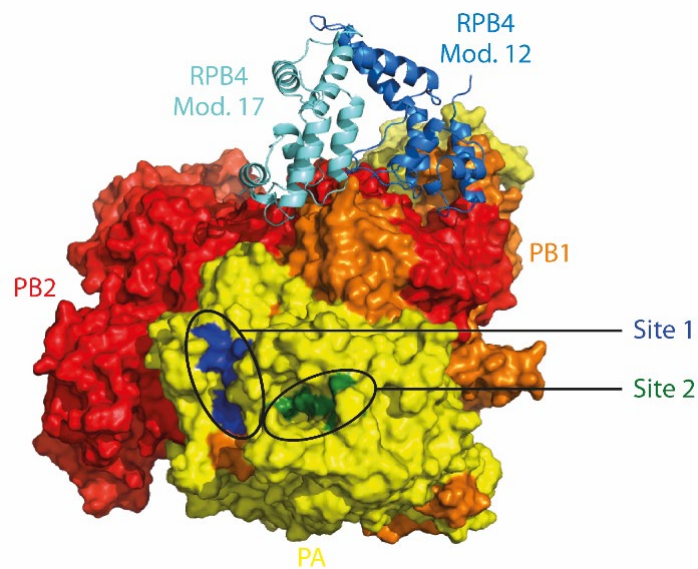


Figure 30 : Comparaison des localisations de RPB4 par rapport aux sites d'interactions de PA avec le CTD de Pol II

Les sites d'interaction 1 (en bleu) et 2 (en vert) sont placés sur la structure de FluPol de IAV (PDB 6RR7). RPB4 des modèles 12 et 17 est également positionnée.

Dans les deux modèles obtenus, aucun des résidus conservés du domaine 108-194 que nous avons testé ne sont engagés dans l'interaction FluPol – RPB4. L'alignement de différentes séquences de PB2 d'Influenza de types A, B et C infectant l'Homme (Figure 31) montre qu'aucun des résidus identifiés dans les modélisations n'est conservé chez ces trois types d'Influenza. Seuls les résidus E123 et K126 impliqués dans le modèle 17 sont conservés chez les Influenza de type A et C. Si l'un de ces modèles est avéré, les modes de fixation de RPB4 à PB2 seraient donc différents entre les PB2 de différents types comme cela a été montré dans le cas de l'interaction entre PA et le CTD de Pol II.

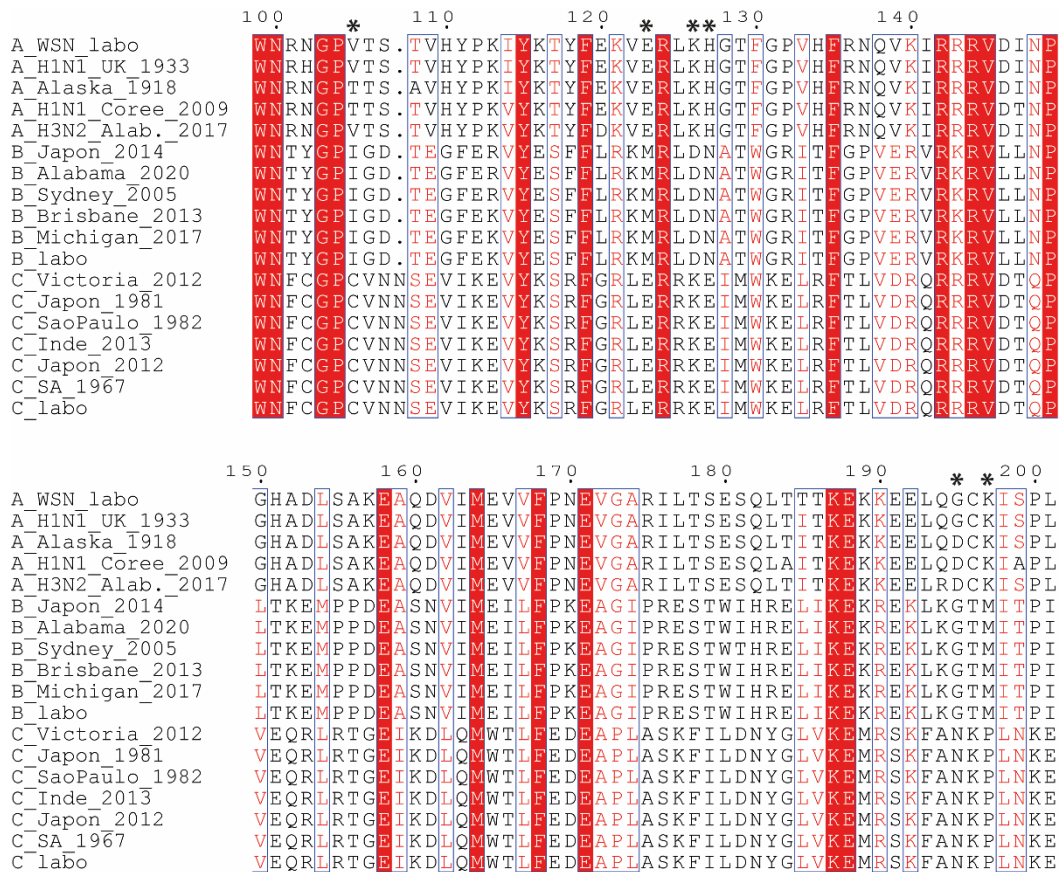


Figure 31 : Aligment des séquences de PB2 de virus Influenza de type A, B et C humains

Les résidus impliqués dans l'interaction PB2 – RPB4 dans nos modélisations sont indiqués par une étoile. La figure a été réalisée à l'aide du logiciel Esprout.

Plusieurs facteurs compliquent la modélisation de l'interaction entre PB2 et RPB4. Premièrement, FluPol est un complexe d'une grande plasticité et le vol de coiffe est un processus très dynamique impliquant une succession de changements de conformation. Dans cette séquence temporelle, la perspective de réalisation de cette interaction dans le bon laps de temps ne semble pas aisée. De plus, dans les structures existantes de Pol II, les trente premiers acides aminés de RPB4 humaine sont vraisemblablement non structurés ou ne sont pas résolus. Cela représente environ 40% de la zone d'interaction préalablement identifiée. Le logiciel de modélisation n'a donc pas pu intégrer ces données pour réaliser une prospection complète des interactions possibles.

L'interaction identifiée entre le domaine N-terminal de PB2 et RPB4 permet néanmoins de proposer un nouveau modèle pour assurer le vol de coiffe par FluPol. Cette interaction pourrait permettre de rapprocher FluPol du canal de sortie des ARNm de Pol II en fin d'initiation ou au début de l'étape d'élongation de la transcription. La conservation de cette interaction au sein des principaux types d'Influenza souligne son importance dans la réplication virale. La liaison de FluPol à Pol II se ferait donc en deux étapes : tout d'abord l'interaction du domaine C-terminal de PA avec le CTD de Pol II

permettrait de rapprocher de proche en proche FluPol du « *core* » de Pol II, puis, une interaction du domaine N-terminal de PB2 avec la partie N-terminale de la sous-unité RPB4 permettrait de rapprocher et ancrer FluPol sous le canal de sortie des ARN de Pol II. Ajoutons qu'en sus de leurs nombreuses fonctions, le dimère RPB4/RPB7 aurait un rôle de stabilisation des ARNm (Garrido-Godino *et al*, 2016). Il a également été proposé qu'en sus du CTD de Pol II, le dimère RPB4 / RPB7, à l'instar de FluPol, permet à divers facteurs de transcription cellulaires de se fixer sur Pol II et de moduler son activité transcriptionnelle (Richard *et al*, 2021; Choder, 2004).

Partie II

ÉTUDE DES INTERACTIONS ENTRE FLUPOL ET LA CHROMATINE

Objectifs

Pour augmenter l'efficacité du vol de coiffe, il est fort probable que les RNPg ciblent spécifiquement les zones transcriptionnellement actives riches en Pol II dans le noyau. Pol II en cours de transcription est présente dans des zones spécifiques du noyau où la chromatine est ouverte et accessible. Ainsi, il a été montré que FluPol interagit avec certains remodeleurs de la chromatine (CHD1 et MORC3) reconnaissant les régions transcriptionnellement actives (Marcos-Villar *et al*, 2016; Ver *et al*, 2015). CHD1 et MORC3 font partie des éléments permettant le ciblage des RNP aux zones ouvertes de la chromatine où se trouve Pol II. FluPol interagit également avec le complexe exosome (Rialdi *et al*, 2017) présent dans les régions transcriptionnellement actives (Andrulis *et al*, 2002; Hieronymus *et al*, 2004). Ces interactions augmentent l'efficacité de transcription de FluPol. Cependant, il n'existe pas plus de détails sur le ciblage de FluPol une fois présente dans ces zones de transcription actives. Seul son recrutement sur les promoteurs de quelques gènes comme la β -actine, la dihydrofolate réductase, EXT1, POGZ et MGST3 ont pour le moment été documentées (Chan *et al*, 2006; Rialdi *et al*, 2017). Des données manquaient encore pour évaluer la répartition de FluPol sur l'ensemble du génome des cellules infectées.

Plusieurs éléments suggèrent que le vol de coiffe aurait lieu au niveau des promoteurs des gènes mais aucune preuve formelle ne fait de lien direct entre vol de coiffe par FluPol (i.e. transcription virale) et temporalité de la transcription cellulaire de Pol II. En effet, plusieurs études tendent à montrer, ou du moins suggérer, que le vol de coiffe aurait lieu à l'initiation de la transcription.

La « coiffe » (ou 5'-cap) est une modification de l'extrémité 5' des ARNm qui consiste en l'ajout d'une guanosine méthylée et en la méthylation des deux premiers nucléotides. Cette modification permet la stabilisation des transcrits, l'export vers le cytoplasme et la traduction par les ribosomes. L'ajout de la coiffe sur les ARNm par le système cellulaire de synthèse de la coiffe a lieu lors de la pause au promoteur (Ho & Shuman, 1999). En effet, le recrutement du complexe enzymatique synthétisant la coiffe est médié par la phosphorylation de la Ser5 du CTD de Pol II. Or, l'interaction préalablement décrite de FluPol avec Pol II est également médiée par la phosphorylation de cette même Ser5, marque de l'initiation de la transcription. Des études séquençant les 10 à 15 nucléotides des ARNm viraux issus du vol de coiffe ont montré que la moitié d'entre eux provient des régions promotrices (Sikora *et al*, 2014) où FluPol y est enrichie (Bauer *et al*, 2018). D'autres encore ont montré que FluPol choisissait préférentiellement des ARN en cours de synthèse pour le vol de coiffe (Gu *et al*, 2015).

L'analyse des ARN associés à Pol II par la technique mNET-seq a montré que les segments viraux étaient mille fois plus enrichis sur la Pol II phosphorylée en Ser5 que sur l'ensemble des Pol II

(Bauer *et al*, 2018). Ceci est donc également en accord avec le vol de coiffe lors de la pause de la Pol II aux promoteurs. Toutefois, Ser5P est présente au niveau des TSS mais également au niveau d'autres régions de la chromatine comme les exons (Nojima *et al*, 2015) ou les amplificateurs (« *enhancer* ») (Andersson *et al*, 2014). De plus, certaines études suggèrent que FluPol cible plutôt les ARN non codants plutôt que les ARNm (Gu *et al*, 2015; Rialdi *et al*, 2017). En effet, des ARN sont également synthétisés au niveau des amplificateurs actifs (ARNe) et sont coiffés. Bien que les enhancers actifs soient quatre à dix fois plus abondants que les promoteurs (Shen *et al*, 2012; Zhu *et al*, 2013) et pourraient ainsi constituer une source importante de coiffes, une analyse plus récente des chimères transcrits cellulaires – ARNg viral montre que les ARNe ne représentent qu'une infime partie des sources de coiffe par rapport aux gènes codants (Rialdi *et al*, 2017). La question se pose donc de savoir ce qui favorise l'utilisation des transcrits codants comme source de coiffe plutôt que d'autres points d'initiation transcriptionnelle *a priori* plus nombreux.

Devant le manque de données sur le comportement de ces deux protagonistes, nous avons voulu investiguer plus précisément le ciblage de FluPol à la chromatine, notamment pour répondre aux questions spécifiques suivantes :

- 1) Quel est le comportement de FluPol à la chromatine ?
- 2) FluPol cible-t-elle des gènes particuliers le long de la chromatine ?

Pour répondre à ces questions, des expériences d'immunoprécipitation de chromatine à l'aide d'anticorps anti-FluPol en contexte infectieux suivies de séquençage haut-débit ont été réalisées pour analyser la position de FluPol le long du génome. Les résultats obtenus sont présentés et discutés dans le manuscrit ci-dessous.

Article (non publié) : “Influenza A virus RNA-polymerase targets the chromatin of innate immune response genes”

Ma contribution à ses travaux a été la réalisation de l’immunoprécipitation de FluPol portant le tag V5 sur la sous-unité PA ainsi que l’extraction des acides nucléiques qui lui étaient associés. J’ai également participé à l’analyse des données issues du séquençage de ces acides nucléiques.

Influenza A virus RNA Polymerase targets the chromatin of innate immune response genes

Jia Yi^{1,3}, Jessica Morel^{4,5}, Mickaël Costallat^{1,2}, Nathalie Lejal⁵, Bernard Delmas⁵, Christian Muchardt^{1,2}, Eric Batsché^{1,2*},

1 Epigenetics and RNA metabolism in human diseases. Institut de Biologie Paris-Seine. Université Sciences Sorbonne. 7-9 Quai Saint Bernard, 75005 Paris, France

2 CNRS UMR8256 - Biological Adaptation and Ageing, Paris, France

3 Doctoral School Life science complexity (ED515), University Sorbonne, Paris, France

4 Doctoral School Structure and dynamics of living systems (ED577), University Paris-Saclay

5 INRAE, UVSQ, VIM, Université Paris-Saclay, F-78352 Jouy-en-Josas, France

* corresponding author. Tel: +33 144273477; E-mail: eric.batsche@sorbonne-universite.fr

Abstract

The influenza A virus (IAV) profoundly affects host cell nuclear processes to accommodate for its transcription. In particular, the IAV RNA polymerase (FluPol) associates with host genes to snatch capped 5'-ends from nascent messenger RNAs, that are then used to prime viral transcription. The extent of FluPol-host gene interaction and its actual consequence on host transcription are still poorly characterized. We show here, using genome-wide approaches, that the PA, PB1, and PB2 subunits of FluPol establish contacts with both chromatin and RNA transcripts at promoters, but also within the coding region of genes. These interactions also reach into downstream regions where recruitment of FluPol subunits correlates with previously reported transcription termination defects induced by IAV, indicative of an implication of the FluPol in this viral strategy to limit host cell defense by interfering with transcription re-initiation. The latter is further suggested by a bias in the FluPol genomic targets, enriched in genes associated with anti-viral defense. Together, our observations suggest that FluPol-chromatin binding contributes to targeted dampening of host immune response pathways.

Introduction

Influenza A virus (IAV) is an RNA virus whose replication occurs in the nucleus of infected cells (Herz et al. 1981) and its own transcription is dependent of the cell RNA-polymerase II (RNAPII) activity (Amorim et al., 2007). This feature, atypical for RNA viruses arise from the activity of its RNA viral polymerase complex (FluPol) formed by the intricate association of the three subunits PA, PB1, and PB2. Indeed the FluPol requires RNA primers originating from the endonucleotidic cleavage of the 5'-end of capped cellular transcripts by the PA subunit to produce viral transcripts (Plotch et al., 1979; Dias et al., 2009; Reich et al., 2014). This process called “cap snatching” allows FluPol to produce viral mRNAs that can be translated into proteins from chimeric transcripts containing the 10-15 nucleotides downstream of the cap from host cell RNAs and the rest composed of viral sequences generated by copying the negative sense genomic vRNA(-).

Several observations suggest that FluPol steals capped RNAs directly at the promoters of genes. Firstly, FluPol interacts with RNAPII phosphorylated at the serine-5 (S5p) on the C-terminal domain (CTD) (Engelhardt et al., 2005; Martínez-Alonso et al., 2016; Lukarska et al., 2017), a phosphorylated form enriched at promoters (Komarnitsky et al., 2000; Boehm et al., 2003). Binding of FluPol to the RNAPII S5p-CTD activates the capping enzyme to proceed to the first step of the mRNA maturation (Ho and Shuman, 1999). Deep-sequencing of the 10-15 nucleotides, which prime the viral mRNAs indicates that about half of these heterogenous sequences matches with the 5-ends of polyadenylated mRNAs (Sikora et al., 2014). Analysis by mNET-seq of viral RNAs co-immunoprecipitated with the RNAPII further showed a 1000-fold enrichment in S5p-RNAPII compared to the total pool of RNAPII, suggesting that FluPol generating the vRNA segments requires the association with the S5p-RNAPII (Bauer et al., 2018). FluPol was also found to interact with CHD1 (Marcos-Villar et al., 2016), a chromatin remodeler involved in transcriptional

elongation, binding to histone 3 trimethylated at lysine 4 (H3K4me3), an epigenetic hallmark of active promoters (Park et al., 2014). Finally, the interaction between FluPol and the nuclear RNA exosome complex was reported to favor the recruitment to promoters (Rialdi et al., 2017).

These data suggest that the cap-snatching occurs in the early phase of transcription elongation and provides an attractive explanation for the coupling of the vRNAs synthesis with the activity of the RNAPII, ensuring FluPol to find the capped-RNA primers.

However, the question remains unclear whether FluPol catches the 5'-end termini of host cells on promoters or also at other positions where S5p-RNAPII can be found. Indeed, S5p-RNAPII is also recruited to enhancers to produce the Capped eRNAs. Because active enhancers are more abundant than promoters (ENCODE Project Consortium et al., 2012; Shen et al., 2012; Zhu et al., 2013), these may be an alternative source of Capped RNAs for FluPol.

Initial approaches to study the “cap snatching” process have suggested that the virus preferentially targets the most abundant, highly transcribed; host mRNAs (Sikora et al., 2014). A more recent study based on an enrichment strategy by hybridization with vRNA likewise identified the protein-coding mRNAs as the main source of Capped extremities (Rialdi et al., 2017). However, in both studies, the limitations of the RNA sequencing may have prevented the detection of less abundant sites of RNA production, while the localization of the initiation sites may be biased by the very short human sequences snatched by the virus (10-15 nucleotides) that cannot always be mapped unambiguously on the human genome. More recently, the IAV was shown to cause a global RNAPII termination defects, by inhibiting the termination complexes (Bauer et al., 2018; Zhao et al., 2018) which, in turn, disrupts the cohesin/CTCF-anchored chromatin loops (Heinz et al., 2018). These chromatin structures are extensively modified by the IAV infection, resulting in an opening of the topologically associating domains (TAD) to allow RNAPII long run-through events in intergenic regions.

We have investigated whether IAV transcription could also use other sources of capped RNAs, such as eRNAs from enhancers, for example. To this end, we aim at identifying additional chromatin loci where cap-snatching might occur by performing a series of chromatin immunoprecipitation (ChIP) using different antibodies recognizing the FluPol subunits PA, PB1 and PB2. We also generated viruses with the FluPol PA subunit labeled with a highly efficient V5 epitope to perform chromatin immunoprecipitations. The nascent transcripts associated to the chromatin at location recruiting FluPol were then analyzed by deep-sequencing. This unprecedented approach surprisingly revealed that FluPol remains linked to the nascent RNAs (probably through the RNAPII association) during phases of transcriptional elongation

and termination. Together, these observations suggested that FluPol participates in inhibiting host gene transcriptional termination, a mechanism that facilitates evasion of the host-cell defense.

Results

High-throughput sequencing of chromatin-associated RNAs in IAV-infected cells revealed deep transcriptional perturbations

To directly examine the effect of influenza virus infection on host transcriptional activity, rather than its combined effect of transcription and RNA maturation, we carried out sequencing of RNA bound to the chromatin in human lung epithelial A549 cells infected at a high MOI (MOI = 5) with an H1N1 influenza virus (A/WSN/33). We observed that at 6 hours post infection (hpi), when focusing of the 3/4 most expressed genes, ~40 % were differentially regulated as evaluated by the DESeq2 package. (**Figure 1A and SupFig S1A**). This dysregulation was more extensive than those reported previously for this early times of infection (Bauer et al., 2018; Heinz et al., 2018; Zhao et al., 2018; Chung et al., 2018; Wang et al., 2020; Zhou et al., 2017).

The strong perturbation of the cellular transcriptome by IAV favors heterogeneous transcript distributions between samples (Bauer et al., 2018; Zhao et al., 2018), and the performance of statistical normalization may be affected (Robinson et al., 2010; Tarazona et al., 2011)(Conesa et al., 2016). Therefore, we used an alternative normalization method based on a set of reference genes. We reasoned that, after just 6 h of infection, pools of transcripts with long half-time would remain essentially unaffected. We first confirmed that mRNAs identified as long-lived in HeLa cells (Tani et al., 2012) preserved this characteristics in A549 cells, since there was a robust correlation between RNA life-span and their amount. Indeed, in both A549 chromatin-associated RNAs and total RNA extract, a higher proportion of shorter-lived RNA was among the weakest expressed genes, and inversely a higher proportion of longer-lived RNA was in the highest expressed genes (**SupFig S1B**). Thus, considering that the 20% most expressed transcripts (> 300 rpkm) with long half-lives (> 15 h), we confirmed that approximately 40% of the 3/4 most expressed genes were differentially regulated (**Fig. 1B, SupFig S1C**), but with only a 75% overlap with the genes identified by DESeq2 pipeline (**Fig. 1C**).

Earlier studies have shown that IAV-dependent transcriptional termination defects allowed the production of “downstream-of-gene” transcripts (DoGs) from all active genes (Bauer et al., 2018; Zhao et al., 2018). This phenomenon, that may largely contribute to the bias we corrected, was well observed in chromatin-bound RNAs, with a general decrease in apparent exon coverage associated with an increase in the abundance of reads beyond the site of polyadenylation (**SupFig S1D and Fig. 1D**). The relevance of chromatin associated RNAs for the observation of DoGs was illustrated by their efficient vi-

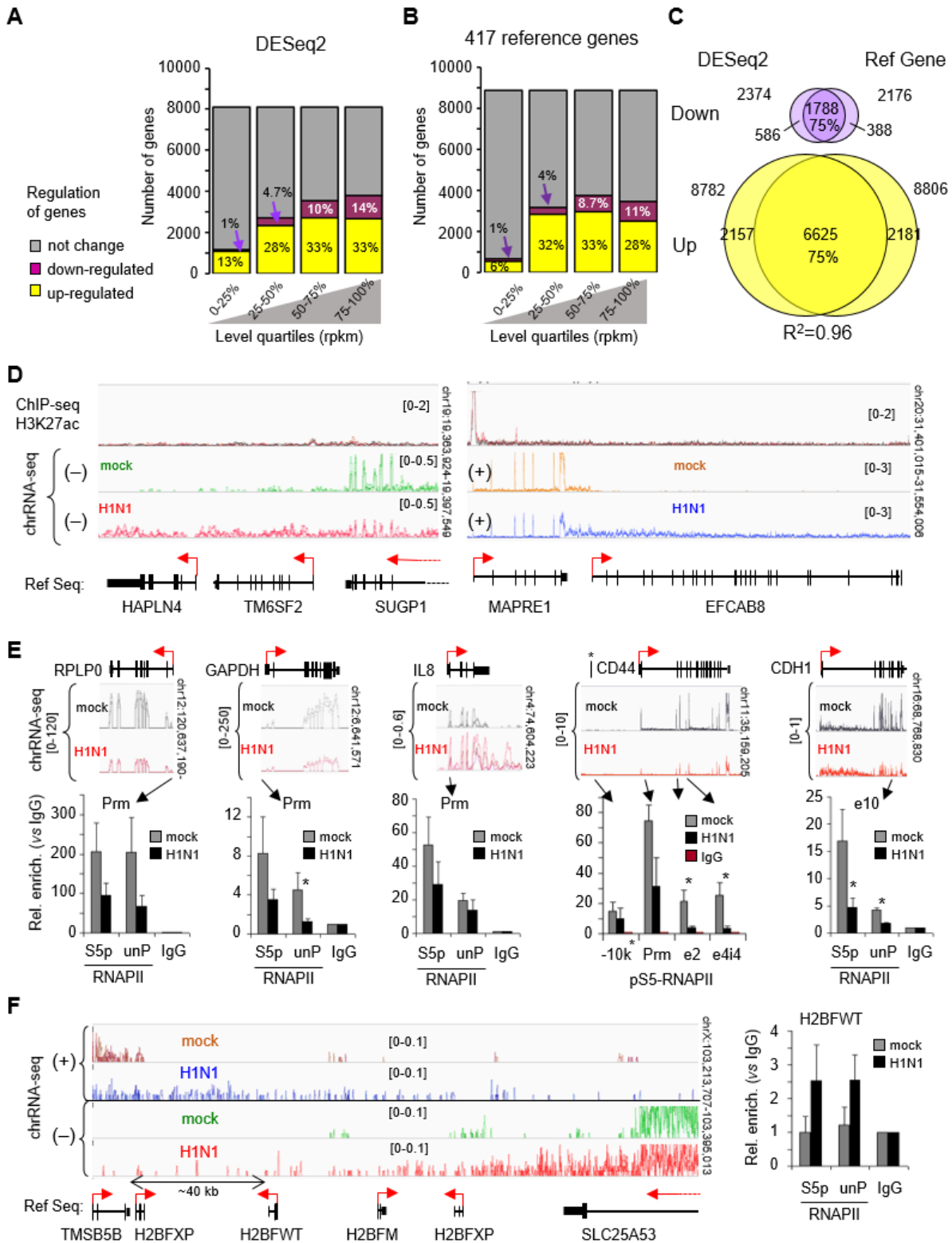


Figure 1 : RNAs associated to the chromatin illustrate cell transcription modifications induced by 6 h of infection by IAV H1N1 in A549 cells

A) Differential RNA levels between H1N1-infected cells versus non-infected cells (mock) using DESeq2 package. The genes were considered regulated for a fold change >1.5 and the adjusted p -value <0.05 . Genes were ranked in quartiles as indicated. **B)** Differential RNA levels using the normalization based on the average variation of 424 reference genes expressing higher level ($>10 \log_2$ cpm) of transcripts with long half-life (>15 h), (SupFig S1B) from a paired t test (two-tails). Genes were ranked in quartiles from the mean of the two conditions. **C)** Venn-diagram for up and down regulated genes as evaluated by the two methods. R^2 is the Pearson's correlation coefficient comparing the two methods. **D)** Examples illustrating the variation in the distribution of reads in H1N1 infected cells from IGV visualization of indicated loci. The RNA-seq independent triplicate were overlaid in the same color as indicated, for each strand orientation. The top track shows the enrichment H3K27ac on the chromatin of infected or mock cells. The scale range of the tracks were indicated in brackets. The bottom track (Ref Seq) indicates the position of the genes and their orientation by the red arrows. **E, F)** ChIP assays with antibodies against S5-phosphorylated (S5p) or non-phosphorylated (unP) RNAPII of infected and non-infected cells. Immunoprecipitated DNA was quantified by qPCR with primers targeting indicated regions. Enrichment are expressed relatively to the signal obtained for ChIP with negative IgG control. Values are means (\pm deviation) of at least three independent experiments. Statistical significance of the differential levels between infected and non infected cells were evaluated using Student's t -test (two-tailed), with $P < 0.05$ (*).

sualization. Even on genes reported previously to not producing these RNAs (LY6E, APOL1, DEFB1 and IFI6) (Zhao et al., 2018), it was possible to observe transcripts produced by RNAPII terminal run-through (SupFig S1E). On the other hand, in agreement with Bauer et al. (Bauer et al., 2018), histone genes did not produce DoGs (SupFig S1F), due to the alternative maturation mechanism active at these genes and not relying on the canonical polyadenylation complex (Kolev and Steitz, 2005).

To gain further understanding of the mechanisms perturbing transcription, we carried out RNAPII ChIP-seq on the infected cells. This revealed that the apparent decrease in exon coverage in the RNA-seq data was also concomitant with an apparent decrease in RNAPII accumulation at the promoters of highly expressed genes not affected by the infection (GAPDH, RPLP0), at induced genes (IL8), and on the transcribed region of moderately expressed genes (CD44, CDH1) (Fig. 1E). Inversely, we noted an increased accumulation of RNAPII, possibly as a consequence of a defective termination at nearby genes, at loci corresponding to

genes not expressed in A549 cells, such as H2BFWT, even after IAV infection (Fig. 1F). Re-analysis of publicly available ChIP-seq data from primary human monocyte-derived macrophages (MDMs) infected with IAV and harvested at 6 hpi (Heinz et al., 2018) revealed a very similar decreased RNAPII recruitment at expressed genes coupled to its increased accumulation on intergenic regions and at genes not expressed in these cells (SupFig S1G).

Quantification of chromatin-bound RNAs confirms that the IAV infection induced defects of transcriptional termination.

We next re-examined the DoGs in the light of their improved detection provided by the sequencing of chromatin-associated RNA. Visual examination in a genome browser showed that numerous genes produced DoGs in the absence of IAV infection, illustrating the intrinsic imperfections of the termination machinery (see examples Fig. 1D, SupFig. S1D,E). As production of DoGs has been proposed to participate in the global reduction in host cell transcription, we examined the

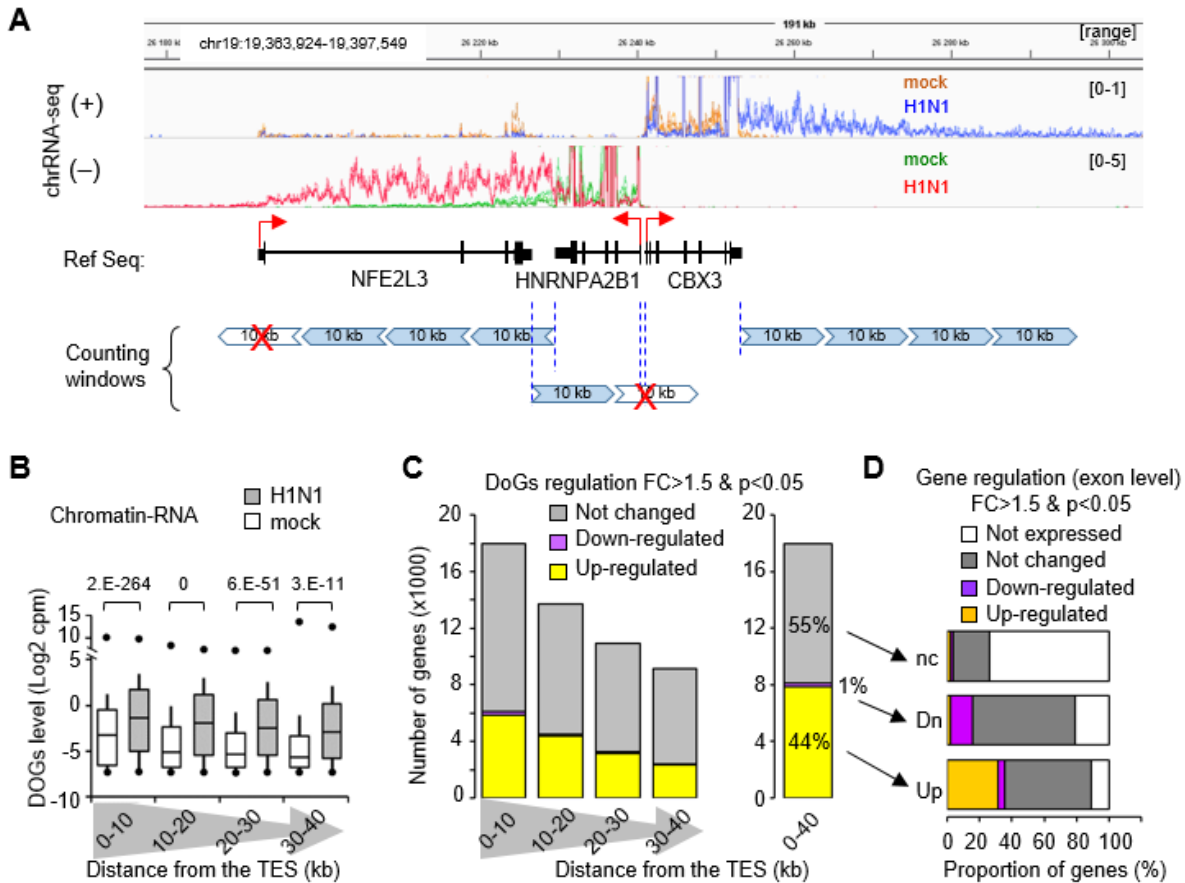


Figure 2 : chromatin RNAs were enriched in DoGs within the 3' end region of the genes. **A)** Strategy to quantify the 3' end transcripts downstream to the termination sites (TES). Oriented reads were counted in 10 kb windows downstream to the TES extending until 40 kb. The downstream-10 kb windows which overlapped with the transcription start site of another Ensembl gene in both orientation were not taken in account (red crosses). **B)** For each selected genes the indicated 10-kb windows from the TES were quantified for their DoGs levels evaluated in Log2(cpm). Data are displayed in box plot where the bottom line is the 1st centile, the box corresponds to the second and third quartile, the top line to the 9th centile, and the central line in the box is the median. Small circles indicate the minimum and the maximum. **C)** Counts of genes for which a 3' end downstream 10-kb windows exhibited differential transcription level. These windows were considered modified for a p-value < 0.05 (paired t test, two tailed) and a fold change > 1.5. The right panel is the same quantification for all pooled windows. The transcription downstream to TES covering the 40 kb was considered upregulated when the majority of 10-kb windows were upregulated (yellow), and the inverse for the down-regulation (magenta). The 40 kb region was considered unchanged (grey) in case of equality or when the 10 kb windows were themselves found unchanged, or when there was not a significant differential levels of DoGs on the four 10 kb-windows, as evaluated by a paired t-test (two-tailed, p < 0.05). **D)** For each category of 3' end downstream region, the regulation of the genes (as evaluated in Fig. 1B) is indicated in proportion.

correspondence between gene expression and production of DoGs. In the aim to evaluate how the DoGs were modified in infected cells, oriented read density were measured in 10 kb windows downstream to the TES extending until 40 kb (**Figure 2A**). This DoGs quantification showed a robust overall upregulation of these RNA species even at sites located up to 40kb away from the TES of genes (**Fig. 2B**). We estimated that 44%

of these downstream regions were upregulated by more than 1.5 times, while almost none were down-regulated (**Fig. 2C**). Among the upregulated DoGs, a majority (~55%) were related to genes unaffected in their expression in the presence of IAV (**Fig. 2D**), consistent with an interference between the biology of this virus and the host termination machinery. Yet, upregulated DoGs also originated from up-regulated genes (approximately

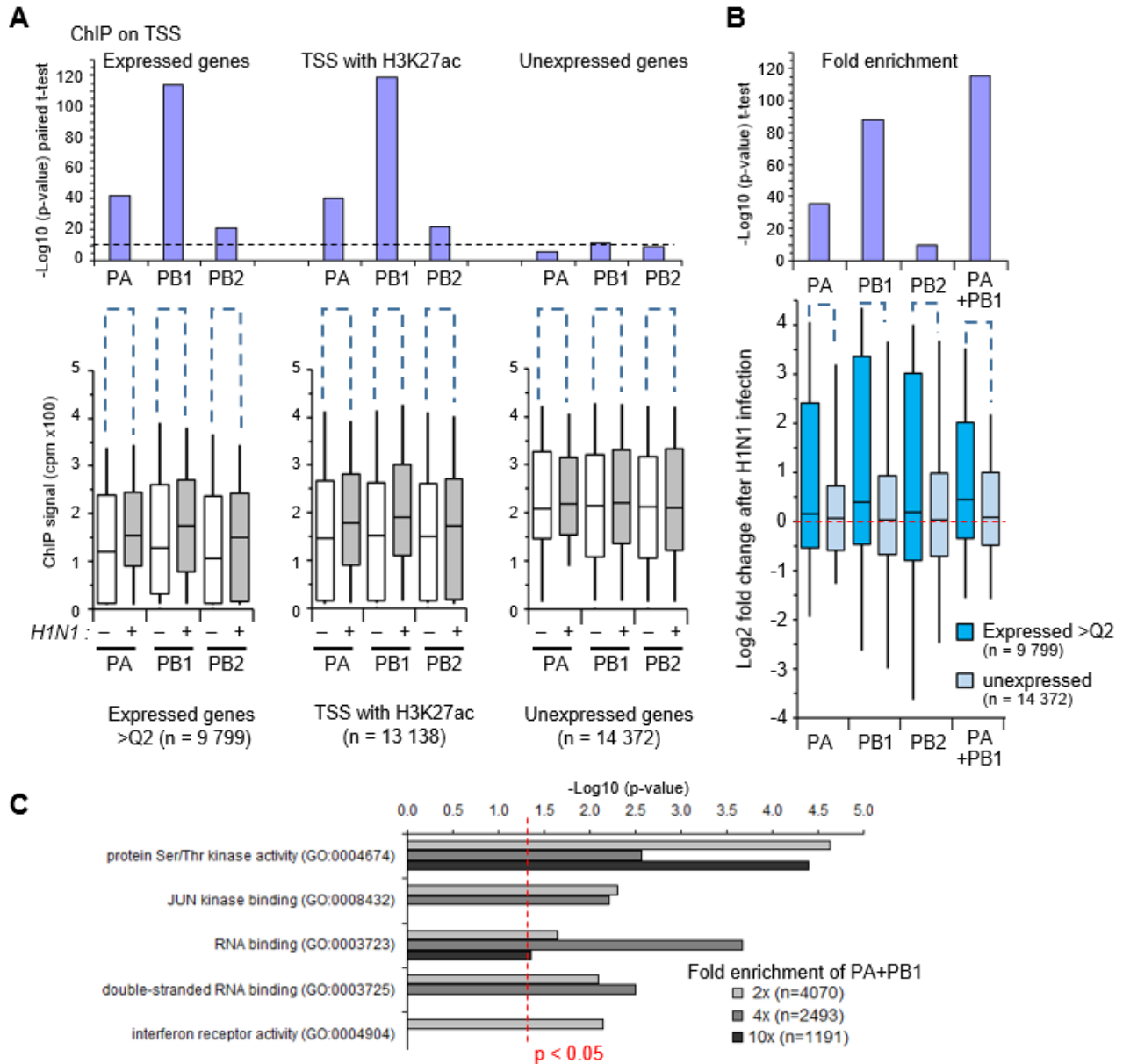


Figure 3 : Enrichment of FluPol subunits on the chromatin. Chromatin of non-infected (-) or infected (+) A549 cells were immunoprecipitated with indicated subunits of the FluPol and analysed by deep-sequencing. For each 500 bp-windows centered on annotated transcriptional start sites (TSS), the DNA enrichment were the means of replicates expressed in cpm (multiplied by 100 to facilitate the representation). Only TSS separated from other TSS by 5 kb were considered to avoid potential misleading with close neighbors or overlapped genes. **A**) Bottom box plots illustrate the distribution of the TSS from genes expressed more than the median (> Q2), of TSS with a H3K27ac signal greater than the median of the H3K27ac signals, or of TSS from non-expressed genes. The bottom and top lines of the box plots indicate the 1st and the 9th centiles. The upper graphs show the p-values evaluated on log2 normalized cpm by a paired t-test (two-tailed) assessing the enrichment induced by the H1N1 infection on each TSS in indicated categories. The dotted line indicate the p-value threshold based on the unexpressed genes. **B**) Fold change of enrichment for indicated TSS categories were the averages of TSS differences between log2 cpm from infected and non-infected cells. Statistical analysis has been carried out the log2 fold change using Student t-test (two-tailed) to compare TSS from unexpressed genes (light blue) and genes expressed more than median (blue). The dashed red line indicates the absence of change. The p-values were indicated above the box plots. In each case, the number (n) of tested TSS were indicated. PA+PB1 analysis were conducted by combining their ChIP-seq signals. **C**) Pathway analysis of genes on which PA+PB1 were enriched 2, 4, and 10 times on TSS between infected and mock cells. P-values of Gene Ontology (GO) molecular function 2021 were evaluated through Enrichr, and the threshold of $p < 0.05$ is indicated by the dotted red line. The number (n) of genes submitted to the analysis were indicated for each fold enrichment category..

30%), suggesting that increased production of DoGs may also be caused by gene activation, and that production of DoGs does not *per se* prevent gene activation, at least at early times of infection. Finally, we noted that approximately 10% of the upregulated DoGs, as well as 70% of the non-significantly modified DoGs, that have been associated with unexpressed genes seemed produced by genes not expressed in the A549 cells. Manual verification indicated that these cases corresponded to genes entirely traversed by DoGs originating from neighboring genes.

Genome wide Analysis revealed that FluPol is recruited to gene promoters

We next examine the genome-wide recruitment of FluPol, using antibodies directed against each of its PA, PB1, and PB2 subunits. When tested on extracts from infected cells, these antibodies all resulted in a co-immunoprecipitation of phosphorylated RNAPII (**Sup Fig S3A**), confirming earlier observations documenting an interaction of the FluPol with the pS5-RNAPII.

Chromatin immunoprecipitation assays carried out with these antibodies. The relatively high background signal suggested that FluPol established only indirect contacts with the chromatin. However, examination of the ChIP signals on 500 bp-windows centered on promoter TSSs, showed a significantly increased signal upon viral infection, which was more obvious on expressed genes than on unexpressed genes, validating the specificity of these antibodies (**Fig. 3A**). To avoid interference from neighboring genes, we focused on genes located at least 5 kb away from other genes. This was confirmed also by the increased signal observed at active promoters harboring H3K27ac marks, that was more important than at TSS of unexpressed genes (compare middle graph to the left one, **Fig 3A**). The differential TSS enrichment induced by the infection for each subunit was significantly greater on expressed than on unexpressed genes (above the red line, **Fig. 3B**). The PB2 antibody seemed to be the less efficient at discriminating the specific signal induced by the infection on expressed genes versus unexpressed genes (**Fig 3A, B**). Finally, the combined signal for the PA and PB1 subunits was more obviously enriched upon infection on expressed genes than on non-expressed genes, and more than the subunits taken individually (right bars in **Fig 3B**). In contrast, no enrichment upon H1N1 infection was detected in introns of genes, with no significant difference between expressed and silenced genes (**Suppl. Fig. S3B**). Together, these observations were consistent with a recruitment of the FluPol subunits to sites of transcriptional initiation.

We next examined whether the recruitment of FluPol was governed only by transcription, or whether its recruitment also involved site specificity. To that end, we carried out gene ontology analysis on the genes strongly recruiting PA+PB1 (2-, 4-, and 10-fold enrichment between infected and mock cells). This revealed a significant enrichment in genes associated with kinases involved in cell signaling, in RNA binding proteins especially of the double-

stranded RNA recognition, as well as receptors to IFN (**Fig. 3C**). Furthermore, we observed a particular enrichment of the FluPol on promoters of OASL, OAS1, OAS2, DDX58/RIG-I, DDX60, ILF2, and IFNAR1/2, INFR1/2 (**Sup. Table 1**), suggesting that FluPol specifically targets genes involved in the innate cell-defense and possibly interferes with their expression.

Chromatin-bound RNAs associated to the PA subunit revealed that FluPol is associated to the RNAPII in elongation.

As an alternative approach to study FluPol recruitment to the host genome, we next implemented an RNA-ChIP approach to explore the range of nascent host RNA species associating with the viral polymerase. As available antibodies were of insufficient quality for this more demanding approach, we generated a V5-tagged version of the PA subunit of FluPol. The resulting recombinant H1N1-derived IAV retained its ability to infect the A549 cells. Immunoprecipitation with an anti-V5 antibody under conditions used for RNA-ChIP (see Material & Methods) on cells infected with the V5-tagged virus, resulted in efficient immunoprecipitation of the tagged-PA (**Fig. 4A** and **SupFig. S4A**).

DNA ChIP experiments conducted with the V5 antibodies confirmed the presence of the engineered FluPol on promoters shown above to recruit the native FluPol, including moderately-to-strongly expressed genes (RPLP0, GAPDH, CD44), and at an induced gene (IL8), and only in infected cells (**Fig. 4B**). Interestingly, V5-PA was also detected with the coding region of CD44 and CDH1, at in regions covered by DoGs, but with no proper transcriptional activity, as illustrated by the H2BFWT/H2BW1 gene, silent in the A549 cells, but crossed by run-away RNAPII (**Fig. 4B**).

We then used the V5-tagged virus and anti-V5 antibody to carry out RNA ChIP-seq. This approach revealed that chromatin-RNAs associating with V5-PA distributed throughout the coding region of a large set of expressed genes (a total of 15299 genes, corresponding to 55% of the examined genes; see example of the OAS genes **Fig. 4C**).

Quantification confirmed that the immunoprecipitated RNAs were preferentially originating from the most expressed genes, although genes expressed at low levels were also contributing (**Fig 4D**).

Specificity of the immunoprecipitation was verified with non-immune IgG, yielding background signal not correlated with gene expression levels (**SupFig. S4B**). In addition, PA-bound RNAs covering the gene bodies concerned genes with an equivalent proportion of differentially expressed genes than what it was observed in the input (compare **Fig 4D** to **1B**), suggesting that FluPol-RNA association could be a simple consequence of the highest transcriptional activity in these genes. Importantly, we also identified a small set of genes (1809 genes) at which V5-PA-bound RNAs was enriched only in matured mRNA sequences (only exonic sequences) (**Fig. 4E**). This set of genes was enriched in up-regulated

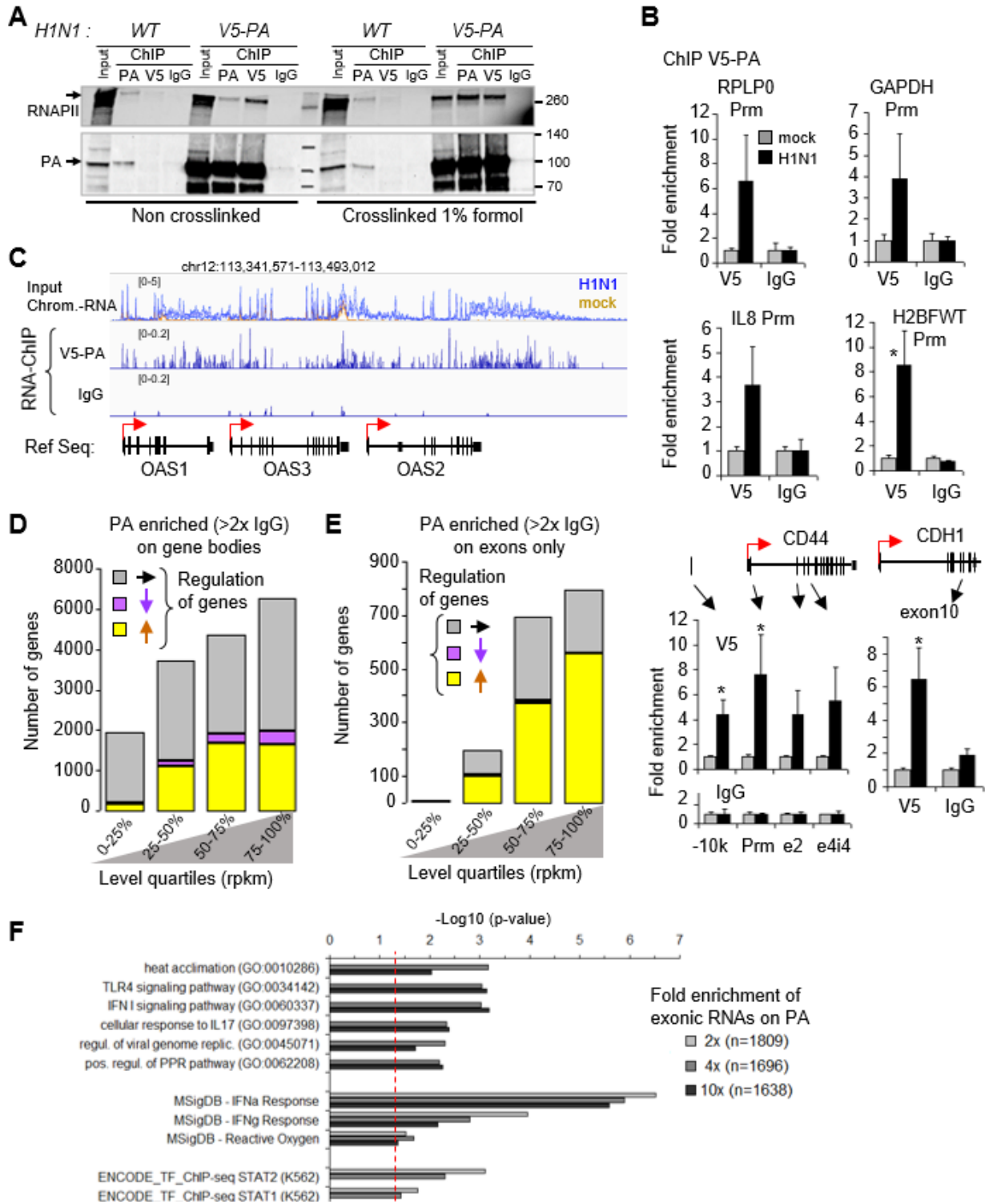


Figure 4 : FluPol subunits on chromatin is linked to nascent RNAs from host-cells **A**) A549 cells infected cells by WT IAV or V5-PA engineered IAV were crosslinked by 1% formaldehyde or not. Chromatin fraction were subjected to immunoprecipitation with PA, V5 antibodies or non immun IgG. Eluates and 10% of extracts were resolved by western blot. The lower part of the membranes was revealed with the PA antibodies and the upper part with phospho-RNAPII antibodies. **B**) ChIP assays with V5 and IgG antibodies of infected and non-infected cells. Immunoprecipitated DNA was quantified by qPCR with primers targeting indicated regions. Enrichment are expressed relatively to the signal obtained for ChIP in uninfected cells. Values are means (\pm dev.) of three independent experiments. Statistical significance of the differential levels were evaluated using Student's t-test (two-tailed), with $P < 0.05$ (*) **C**) Example of PA enriched RNAs bound to the chromatin with IG. The RNA-seq independent triplicate were overlaid in the same color as indicated, for the (+) strand. The scale range of the tracks were indicated in brackets. The bottom track (Ref Seq) indicates the position of the genes and their orientation by the red arrows. **D**) RNAs from ChIP by V5-PA or by non-immun IgG in infected cells were analysed by deep-sequencing. RNAs covering the gene bodies enriched with PA were evaluated versus IgG and have been considered significant when the average of the changes subtracted with the triplicate deviation is superior of two times to the IgG. The number of enriched genes that were upregulated (yellow), downregulated (magenta) or unmodified (gray) were plotted in function of gene expression as determined **Fig. 1B** and ranked in quartiles. **E**) The same analysis as in **C**) was conducted for RNAs covering the exons only. **F**) Pathway analysis of genes on which PA were enriched 2, 4, and 10 times on exonic RNAs between infected and mock cells. P-values of Gene Ontology (GO) biological function 2021, Molecular Signatures Database 2020 (MSigDB), and ENCODE ChIP-seq 2015 were evaluated through Enrichr, and the threshold of $p < 0.05$ is indicated by the dotted red line. The number (n) of genes submitted to the analysis were indicated for each fold enrichment category.

genes) at which V5-PA-bound RNAs was enriched only in matured mRNA sequences (only exonic sequences) (Fig. 4E). This set of genes was enriched in up-regulated genes, an enrichment not observed with non-immune IgG (SupFig. S4C). Most of the PA-bound exonic RNAs corresponded to upregulated genes in all quartiles of expression level (compare Fig 4E to 1B), suggesting a particular gene targeting of FluPol.

Gene ontology analysis on this subset of genes indicated a significant enrichment in genes involved in host-cell defense, associated with Toll-like receptor (TLR) and IFN pathways (Fig. 4F). This enrichment was not observed when analyzing the entirety of the 8824 upregulated genes. Thus, this approach confirmed the IFN α / γ pathway as a particular target of the FluPol recruitment, and suggested that binding of FluPol to mature mRNA may participate in the viral strategy to dampen the host-cell defense machinery. Finally, we noted an enrichment of the V5-PA subunit on the transcripts bound to the chromatin matched genes recruiting FluPol at their promoters, as documented by our DNA-ChIP assay, including the OAS genes, DDX58/60, IFIT2/3/M1/M2, and ISG15/20 genes (Sup. Table 2). This suggests that cap-snatching at promoters requires or causes a metastable association between the FluPol and the messenger RNAs, lasting at least through some of the pre-mRNA maturation steps.

Chromatin-bound RNAs associated to the PA subunit revealed the FluPol is associated to the RNAPII to termination sites and downstream regions.

We next used the V5-PA RNA ChIP-seq to examine in more details the association of the FluPol with with DoGs. The distribution of RNAs bound to the V5-PA subunit suggested an enrichment at the 3' end of genes (TES) and at regions covered by DoGs (Fig. 4C). This was verified quantitatively by counting reads in the regions extending 40kb downstream of the TES of all genes, in bins of 10 kb (Left panel, Fig. 5A). When consolidating reads over the entirety of the 40kb-regions, the median density of V5-PA-associated RNAs was 2-fold higher than that observed with the non-immune IgG (right panel, Fig. 5A). Among V5-PA-associated DoGs, there were 3 to 4 times more upregulated than unmodified DoGs, and virtually none diminished (compare brown and black bars, Fig. 5B). At note, DoGs identified as upregulated in the V5-PA RNA ChIP-seq data outnumbered those identified with the RNA-seq data (Fig. 3C), and V5-PA-bound DoGs were greatly enriched at DoGs upregulated in the RNA-seq (47%-55%, vs 6%-12% in unmodified DoGs and 0-7% in downregulated DoGs). This strongly suggested that the FluPol is an active player in the mechanism causing the production of DoGs in infected cells and may participate in the viral complex involved in the inhibition termination processes. To further document the presence of the FluPol in the DoG regions, we reassess the DNA ChIP-seq data with the anti-FluPol subunits, to quantify enrichment of individual subunits at TESs and at downstream regions. In

order, to discriminate between TESs and potentially overlapping TSSs, we selected genes separated from each other by at least 10 kb. A significant enrichment in downstream regions of expressed genes versus unexpressed genes was mainly observed for PA and PB1 (Fig 5C). The recruitment of these FluPol subunits was in average strongest at regions proximal to the TES (compare quantifications at 500bp, 1kb, 10kb and 40 kb away from the TES), suggesting that coding regions and proximal terminal downstream region are preferential sites for FluPol binding to nascent transcripts.

Gene ontology analysis on genes at which the FluPol is enriched on the 3'-ends revealed an enrichment in genes involved in the cell-defense (pattern recognition receptors (PRR) and NF-kB pathways), in signaling pathways, in oxidative stress associated to cell death, and in the degradation of mRNA 3' end (Fig 5D). Genes combining an enrichment in FluPol on their TSS, their exons, and on their TES were few, yet again significantly enriched in genes involved in the cell defense (PLSCR1; IRF5; DDX60; SOD2, CXCL3; SOD2; RELA).

Discussion

We have shown that infection of cells with IAV causes a significant redistribution of transcripts into the chromatin compartment (Fig. 1C) which is a consequence of the inhibition of the transcriptional termination process. This phenomenon is caused by a significant increase in DoGs that can be easily observed by analysis of chromatin-associated RNAs in ChIP experiments (Fig. 2). This inhibition of gene termination results in a general decrease of the RNAPII at promoters and in gene bodies, and an increase in intergenic regions (Fig 1D-F, S1D-S1G).

By using the presence of the 3 subunits of the viral RNA polymerase in the chromatin of infected cells, we showed that FluPol is not only associated with promoters where it is thought to snatch the caps of RNAPII-initiated RNAs. In fact, our observations suggest that FluPol remains associated with RNAPII during elongation, at transcriptional termination sites and also in downstream regions where RNAPII continues to run in intergenic regions (Fig 3, 4, 5). FluPol was not only bound to the DNA of transcriptionally active genes but remains associated with host cells nascent transcripts (Fig 4C-E, 5AB), with a concentration to DoGs. This is the first time to our knowledge that FluPol is observed along the gene bodies and beyond. The association with DoGs suggests that FluPol could, in addition to NS1, also participate in the process of transcriptional termination inhibition.

Analysis of the regulatory pathways defined by the genes for which FluPol is most strongly enriched indicates that FluPol particularly targets genes of the anti-viral innate response. The participation of FluPol in the inhibition of transcriptional termination particularly on cellular defense genes would add a function of FluPol, in addition to cap-snatching, to participate in the viral escape strategy.

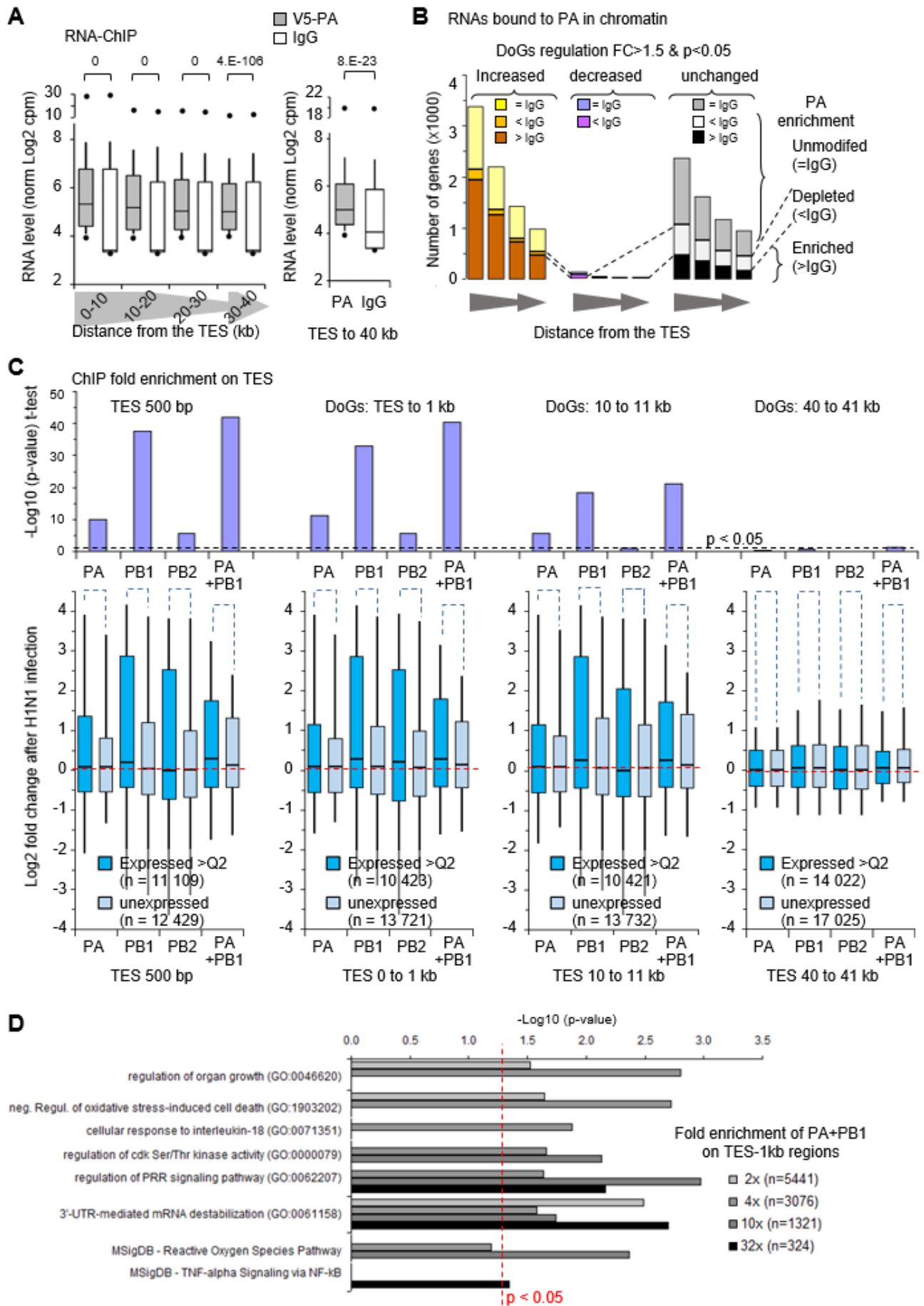


Figure 5 : Enrichment of FluPol subunits on chromatin downstream to the 3'ends of genes. **A)** RNA bound to the chromatin of infected A549 cells and immunoprecipitated with indicated antibodies were analysed by deep-sequencing. For each selected genes the indicated 10-kb windows from the TES were quantified for their DoGs levels evaluated in Log2(cpm). Data are displayed in box plot as in **Fig 2B**. The right panel is the same quantification for all pooled windows. **B)** Counts of DoG regions enriched with V5-PA and their differential transcription level upon infection. These regions were considered enriched by PA (>IgG, in brown and black) when the average of the ChIP V5-PA signals minus the deviation is greater than twice the level of the control IgG. The regions were considered depleted in PA (<IgG, orange, violet, white) when the PA signal plus the deviation was weaker than twice the IgG signal. DoG levels were considered modified for a p-value<0.05 (paired t test, two tailed) and a fold change >1.5. **C)** Fold change of enrichment for categories of indicated terminal region relative to TES. These were the averages differences between log2 cpm from infected and non-infected cells. Statistical analysis has been carried out the log2 fold change using Student t-test (two-tailed) to compare DoGs regions from unexpressed genes (light blue) and genes expressed more than median (blue). The dashed red line indicates the absence of change. The p-values were indicated above the box plots. The dashed line indicate the limit p-value=0.05. In each case, the number (n) of tested TSS were indicated. PA+PB1 analysis were conducted by combining their ChIP-seq signals. **D)** Pathway analysis of genes on which PA+PB1 were enriched 2, 4, 10 and 32 times on 1 kb windows downstream to the TES between infected and mock cells. P-values of Gene Ontology (GO) biological process 2021 and MSigDB were evaluated through Enrichr, and the threshold of p<0.05 is indicated by the dotted red line. The number (n) of genes submitted to the analysis were indicated for each fold enrichment category.

Transcriptional changes during viral infection confound the analysis of differential gene expression.

It was known the IAV-induced terminal defects allowed the production of “downstream-of-gene” transcripts (DoGs) (Bauer et al., 2018; Zhao et al., 2018). These reads beyond the poly-A termination signal are very generally transient RNA species, and are too rare to be observed at high levels by RNA-seq, but not with data from IAV-infected cells. The DoGs and the subsequent pervasive transcription become so high that they negatively affect the power of mRNA quantification by RNA-seq. This lies on the modification the distribution of reads among the transcriptome which significantly affects the quantification of gene expression (Robinson et al., 2010; Tarazona et al., 2011). The change in transcript distribution in the chromatin fraction partially defeated the assessment of differential gene expression by algorithms, such as DeSeq2, validated to measure mRNA expression changes. The method we proposed using reference genes assumed to be constant, in our case here based on the lifetime of mRNAs while the biological triggering phenomenon (viral infection) was in a much shorter time frame, demonstrated that there is an assessment uncertainty with DeSeq2 for 25% of genes predicted to be differentially expressed (**Fig 1C**). Our results highlighted the importance of the choice of normalization methods in the case of biological phenomena, such as here the viral infection by IAV, which disrupt the distribution of RNAs in the genome, in addition to the expression level of genes. Chromatin RNA enrichment allowed the observation of many transcriptional changes, quite in agreement with those obtained by more complicated techniques such as NET-Seq (Bauer et al., 2018). Part of these differences could lie on the RNAPII passage through silent genes, for which the promoters are inactive (**Fig. 1D, 1F, S1D**).

The distribution of FluPol chromatin association indicates that the site of Cap snatching is not necessarily the TSS.

ChIP experiments with several antibodies directed against FluPol subunits allowed us to show that FluPol localizes

5C). To our knowledge this is the first time that FluPol has been localized on a large scale across the genome.

Although we noted a certain inefficiency in ChIP of antibodies directed against the PA, PB1 subunits and even worse with the anti-PB2 antibody, statistical processing of the enrichment data allowed us to show a specific targeting effect of FluPol on transcriptionally active versus silent genes. These results are also confirmed by the detection of cellular RNAs associated with ChIP by an anti-V5-tag antibody recognizing a virus modified to express a V5-tagged version of PA (**Fig 5A B**). Interactions between FluPol and the Ser5 phosphorylated form of RNAPII (Engelhardt et al., 2005; Martínez-Alonso et al., 2016; Lukarska et al., 2017), as well as interaction with the chromatin reader CHD1 that recognizes the promoter-specific H3K4me3 marks (Marcos-Villar et al., 2016) would suggest that the process of cap-snatching must occur when nascent RNAs exit from RNAPII paused at promoters (Walker and Fodor, 2019). We have no good reason to question this hypothetical model. However, studying the localization of FluPol on chromatin greatly expands the windows of opportunity where this reaction might occur.

PA-associated RNAs are more often found up-regulated and more significantly involved in the cellular response when they were exonic RNAs than when considering RNAs spanning the entire gene body (**Fig 4DE**). This suggests that spliced RNAs would be associated with PA, in the same time with DoGs at the 3'-end of genes. We propose a model where FluPol is recruited to TSSs but would remain attached to RNAPII during elongation. Due to the coupling between transcription and RNA splicing, maturing RNAs arrive at the 3' end with RNAPII and FluPol. At the TES level, termination inhibition causes transcription to continue after the poly-adenylation site of the genes where FluPol still remains associated with the elongating complex carrying spliced RNAs (**Figure 6**).

The association of FluPol with RNAPII in elongation could be part of the viral strategy of escape from cellular defenses.

What might be the advantage of FluPol remaining attached to the transcriptional complex throughout gene

transcription? On the one hand, FluPol recruited to the TSS preferentially on genes involved in antiviral defense (Fig 3C), could also participate in termination inhibition of these same genes (Fig 5D). This could be an alternative pathway to the cap-snatching process by which FluPol would inhibit the expression of defense genes. On the other hand, by remaining associated with the elongating transcriptional complex, FluPol may ensure that its own transcripts are in close proximity to the factors necessary for the splicing of genomic segments 7 and 8 (which are partially spliced to generate M2 and NS2 viral proteins, respectively), or for nuclear export (Ferhadian et al., 2018). The increase in DoGs found very often associated with PA also suggests that FluPol may also play a role in the stability of these usually very unstable RNAs. Recently, we have shown that RNA DoGs in intergenic regions are enriched in repeat regions that promote cellular senescence via the PRR pathway (Mullani et al., 2021). This increase in DoGs is caused by a decrease in the expression of certain exosome RNA subunits. Because the intergenic regions crossed by DoGs are rich in repeated sequences (SINE, LINE, LTR), this pervasive transcription can result in the formation of double-stranded RNAs recognized by intracellular PRRs (Chow et al., 2018). In this regard, it is interesting to recall that FluPol has been shown to be associated with the nuclear exosome RNA which plays a positive role in viral multiplication (Rialdi et al., 2017). Loss of function of the nuclear RNAexosome makes viral multiplication less efficient possibly by decreasing the recruitment of FluPol to the TSS. It can be noticed in the results of this study that depletion of the RNA exosome, in addition to the decrease in FluPol at the TSS (5'end), instead caused an increase in FluPol in the gene terminal region (3' region). This confirms our data on the presence of FluPol on termination sites of gene. It is therefore conceivable that FluPol may have an inhibitory action on the activity of the exosome RNA which therefore promotes the increased presence of of DoGs (and intergenic RNAs) usually targeted by rapid degradation. The robust increase of DoGs during viral infection combined with the transcriptional repression of genes involved in the PRR pathway could allow the saturation of the double-stranded RNA detection system and thus leave the way clear for viral processes, such as the formation of cellular:viral RNA hybrids, essential for the viral transcription.

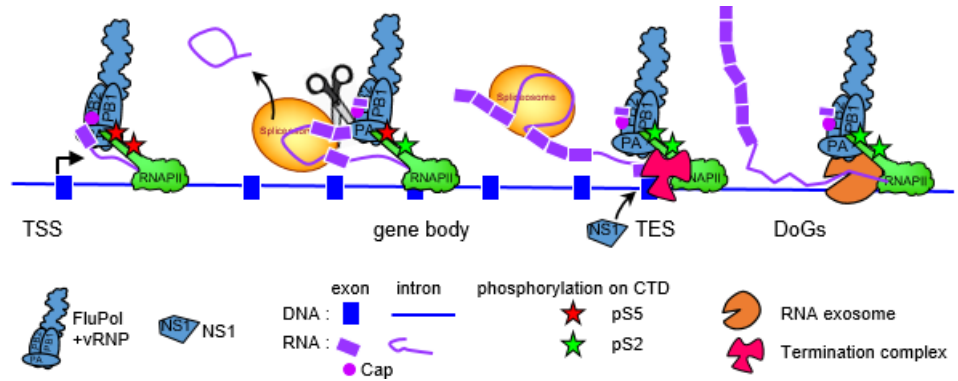


Figure 6 Working model on the functional association of FluPol with RNAPII. The FluPol is recruited first at the level of TSS and PB2 can bind the cap of the nascent RNA. During elongation the FluPol may stay linked to the RNAPII either through its interaction with the S5p which remain present all along the elongation process or either the link by the nascent transcript. The splicing of the nascent transcript occur during the elongation, producing intermediate of spliced forms. At the 3'end of the gene of the termination process is inhibited by NS1. At this position the FluPol could participate to the termination inhibition and also to the inhibition of the RNA exosome promoting the stabilization of DoGs. The cutting reaction of the nascent RNA by the PA subunit (scissors) could occur at anytime of these steps. The nascent RNA remains linked to the RNAPII in elongation until the 3'end region when the RNAPII continues after the TES because of the termination inhibition. This proposed model could explain why the FluPol is found enriched to DoGs and exonic RNAs on the chromatin.

Material and Methods

H1N1 virus construction and preparation

Amplification of WSN H1N1 // Construction of V5-tag-PA virus WSN H1N1

Antibodies :

Antibodies were purchased from Diagenode for H3K27ac (C15410196) ; H3K4me1 (C15410037) from Abcam for RNAPII (8WG16) , RNAPII pS5 (ab5131) for Western blot, RNAPII pS5 (CTD 4H8) and for ChIP, histone H3 (Ab1791), from ThermoFisher for PA (PA5-31315), PB1 (PA5-34914) and PB2 (PA5-32220); and from Bethyl for anti-V5 epitope (A190-120A).

Cell culture, H1N1 infection

A549 cells (CCL-185) were purchased from ATCC. Cells were maintained in Dulbecco's modified Eagle's medium (Gibco) supplemented with 7% (v/v) fetal bovine serum (Thermo Scientific) and 100 U/ml penicillin-streptomycin (Gibco). A549 pellet of cells, infected (MOI=5, h.p.i=6h) with V5-tag-PA (virusV5) or normal virus WSN H1N1. Cells have been not fixed or fixed by 0.5%, 1%, 2% formaldehyde in PBS solution during 10 min at room temperature prior to harvest. Cells were pelleted and frozen at -80°C prior performing the extracts.

Chromatin immunoprecipitation (ChIP and RNA-ChIP) and RNA-chromatin extraction

ChIP assays were performed as previously described (Ameyar-Zazoua et al., 2012; Mauger et al., 2015). Cells were crosslinked during 10 min at room temperature in PBS with 1% (v/v) formaldehyde, followed by a 5 min wash in PBS containing 125 mM Glycine before harvest.

Resuspended cells (10-20E6 in 1.5 mL μ tube) were incubated on ice 5 min in 1 mL of chilled buffer A (0.25% TRITON, 10 mM TRIS pH8, 10 mM EDTA, 0.5 mM EGTA), 30 min in 1 mL of buffer B (250 mM NaCl, 50 mM TRIS pH8, 1 mM EDTA, 0.5 mM EGTA), and then resuspended in 100 to 200 μ L of buffer C (1% SDS, 10 mM TRIS pH8, 1 mM EDTA, 0.5 mM EGTA) at room temperature. All buffers were extemporately supplemented by 0.5 U/ μ L RNasin (Promega), CComplete protease and PhosSTOP phosphatase inhibitors (Roche). Cell suspensions in 0.6 mL μ tube were sonicated in water bath at 4°C during 12 min (15 sec. ON, 15 sec. OFF) with a Bioruptor apparatus (Diagenode) setted on high power and then clarified by 10 min centrifugation at 10 000 rpm, 4°C. Shearing of the DNA was checked after reversing the crosslinking on agarose gel electrophoresis to be around 300-500 bp. Sheared chromatin was quantified by optical density (260 nm) and diluted 10-fold in IP buffer to final concentrations: 1% TRITON, 0.1% NaDeoxycholate, 0.1% SDS, 150 mM NaCl, 10 mM TRIS pH8, 1 mM EDTA, 0.5 mM EGTA, 1 U/ μ L RNasin, 1X protease and pshotase inhibitors). For ChIP-PCR assays 30 μ g of chromatin was incubated with 2 μ g of antibodies. For ChIP-seq or RNA-ChIP-seq assays, 200 μ g of chromatin was incubated with 10 μ g of antibodies at 4°C during 16 h. 20 μ L of saturated magnetic beads coupled to protein G (Dynabeads) per μ g of antibody were used to recover the immuno-complexes. After 2 h of incubation the bound complexes were washed extensively 10 min at room temperature on a wheel in the following wash buffers: WBI (1% TRITON, 0.1% NaDOC, 150 mM NaCl, 10 mM TRIS pH8), WBII (1% TRITON, 1% NaDOC, 150 mM KCl, 10 mM TRIS pH8), WBIII (0.5% TRITON, 0.1% NaDOC, 250 mM NaCl, 10 mM TRIS pH8), WBIV (0.5% Igepal CA630 (Sigma), 0.5% NaDOC, 250 mM LiCl, 10 mM TRIS pH8, 1 mM EDTA), WBV (0.1% Igepal, 150 mM NaCl, 20 mM TRIS pH8, 1 mM EDTA), WBVI (0.001% Igepal, 10 mM TRIS pH8). 20-100 μ g of sheared chromatin used as input and ChIP beads were then boiled 10 min in 100 μ L H₂O containing 10% (V/W) chelex resin (BioRad), followed by Proteinase K (0.2 mg/mL)-digestion for 30 min at 55°C, and then finally incubated 10 min at 100°C. 0.5 μ L was used for quantitative real-time PCR assays. For RNA-chromatin extraction, supernatant were purified by acidic phenol/chloroform method and ethanol-precipitation with 5 μ g of Glycoblue (Ambion) as carrier. After resuspension these input, as well as the eluate from RNA-ChIP were treated by 5 U of TURBO-DNase (Ambion) 1 h a 37°C prior purification by acidic phenol/chloroform method and ethanol-precipitation with 5 μ g of Glycoblue (Ambion) as carrier. These samples were then checked for the absence of DNA contamination by qPCR with 1 μ L of eluate and primers targeting genomic repetitive elements. Another round of DNase

digestion and purification was performed in case of qPCR signal < 35 Ct.

Library preparation and high throughput sequencing

1 to 5 nanograms of immunoprecipitated DNA were used to build DNA libraries. Libraries were prepared following the instruction from NEBNext® Ultra™ II DNA Library Prep Kit for Illumina® (NEB #E7645L) and indexes (NEB #7335S; #E7500S; #E7710S). No sized selections were applied. The quality and size of the libraries were assessed using the 2100 Bioanalyzer (Agilent). The amount of the DNA used for library preparation and the amount of DNA libraries were measured using Qubit (Invitrogen). Indexed libraries were pooled according to the Illumina calculator (<https://support.illumina.com/help/pooling-calculator/pooling-calculator.htm>) and deep sequenced by the Illumina platform provided by Novogene (HiSeq-SE50). RNA from IgG ChIP triplicate were pooled to ensure sufficient material equivalent to individual V5-PA ChIP. Libraries from ChIPed RNA and input chromatin-RNA were performed by the GenomEast platform, a member of the 'France Genomique' consortium (ANR-10-INBS-0009) (IGBMC, Illkirch, France). RNA-Seq libraries were generated from 100 to 300 ng of RNA using TruSeq Stranded Total RNA Library Prep Gold kit and TruSeq RNA Single Indexes kits A and B (Illumina, San Diego, CA), according to manufacturer's instructions. Briefly, ribosomal RNA was removed by Ribo-Zero beads. After chemical fragmentation depleted RNA were reverse-transcribed with random primers. After second strand synthesis by DNA Polymerase I and RNase H and blunting, the cDNA were ligated to adapters and amplified by PCR for 12 cycles. Residual primers were removed by AMPure XP beads and the libraries were sequenced on Illumina Hiseq 4000 sequencer as Paired-End 100 base reads. Image analysis and base calling were performed using RTA 2.7.7 and bc12fastq 2.17.1.14.

Data availability

ChIP-seq, RNA-ChIP-seq, RNA-seq data have been deposited in the NCBI Gene Expression Omnibus database under GEO accession number GSExxxxxx.

Protein extraction for immunoprecipitation.

Pellet of frozen cells corresponding to 10E6 cells were thawed in ice and resuspended in 500 μ L of Buffer A with protease inhibitors (Roche). 150 μ L of suspension were supplemented with 180 μ L of buffer B and sonicated 5 min at 4°C by BioRuptor (high power, 15 sec. ON / 15 sec. OFF). Clarified supernatants were quantified by DO (260 nm). 20 μ g of protein extract were incubated with 1 μ g of indicated antibodies during 16 h at 4°C on a wheel and then 2 h with 30 μ L of anti-rabbit Dynabeads. Beads were washed 5 times in wash buffer WBV, resuspended

in laemmli buffer with 100 mM DTT and boiled 20 min and supernatant was analyzed by western blot.

Western blot

Denaturated proteins were resolved by SDS-PAGE (4–12% Criterion XT Bis-Tris Protein Gel; Bio-Rad), and transferred to nitrocellulose membrane (Bio-Rad). Blocked membrane were incubated with 1/1000 diluted first antibody as indicated in figures in PBS with 0.1% Tween-20 and 5% (V/W) non-fat milk, washed 10 min 3 times in PBS-tween(0.1%), incubated 1 h at room temperature with 1/3000 anti-Mouse StarBright blue700 (BioRad) or 1/2000 True blot HRP anti-rabbit antibodies. After 4x10 min washes in PBS-tween(0.1%), the membranes were revealed by chemiluminescence (for HRP), and quantified using Chemidoc MP imaging system (Bio-Rad).

Real-time quantitative PCR:

1 μ L of ChIP eluate was used for quantitative real-time PCR (qPCR) in 10 μ L reactions with Brilliant III Ultra Fast SYBR-Green Mix (Agilent) using a Stratagene MX3005p system. The analysis of qPCR was performed using the MxPro software. The sequences of primers used for PCR are in **SupTab. S3**. Statistical analysis and graphs were produced with Microsoft Excel.

Bioinformatic analysis

The human genome of reference considered for this study was hg19 homo sapiens primary assembly from Ensembl. After reads alignment, the SAM files were then converted to BAM files and sorted by coordinate and indexed using samtools (v1.7) (Danecek et al., 2021). For ChIP-seq data from this study or from those in MDM cells (GSE103477) (Heinz et al. 2018), reads were mapped to hg 19 using bowtie2 (v2.3.4) (Langmead and Salzberg, 2012) (parameters: -N 0 -k 1 -very-sensitive-local). We then selected reads with a MAPQ equal or higher than 30 corresponding to uniquely mapped reads for further analysis. After indexing MACS2 (v.2.1.1) (Zhang et al., 2008) was used to call peaks (parameters: -p 0.05). For RNA-chromatin and RNA-ChIP-seq data, mapping was carried out with STAR (v2.6.0b) (Dobin et al., 2013) (parameters: --outFilterMismatchNmax 1 --outSAMmultNmax 1 --outMultimapperOrder Random --outFilterMultimapNmax 30). Data observations were performed with the Integrative Genomics Viewer software (Robinson et al., 2011) using the CPM normalized BigWig files produced from the BAM files by Deeptools suite (parameter: --normalizeUsing CPM) (Ramírez et al., 2016)

Evaluation of gene expression from chromatin-RNA (our study) or from total RNA extract of A549 cells mock or infected by normal H1N1 at 12 hpi (GSE103604) (Zhao et al. 2018) were performed by using featureCounts (v1.28.1) (Liao et al., 2014) from the Rsubreads package (parameter: -s 2). These read counts were analysed using the DESeq2 (Love et al., 2014) package in order to test

for the differential gene expression (DGE). The normalization, the dispersion estimation and the statistical analysis were performed with DESeq2 using the default parameters. Raw *P*-values were adjusted for multiple testing according to the Benjamini and Hochberg (BH) procedure and genes with an adjusted *P*-value lower than 0.05 and a fold change greater than 50% were considered differentially expressed. The raw count from featureCount were also used to produce normalized gene levels using reference genes as follow : counts covering mRNAs were converted in rpkm by considering the library total count of each sample and the length of each gene as defined in Ensembl. The average level of genes with half-life greater than 15 h as defined in (Tani et al., 2012) and that were more expressed than 75% of all expressed genes (corresponding 417 genes with level higher than 300 rpkm for chromatin-RNA, and 281 genes with more than 17 rpkm in total RNA extract from Zhao et al. 2018) was used to calculate a corrective index applied on the gene levels in infected cells. The maximum mRNA level between mock and infected cells were used to calculate the correlation between the level quartiles and the life-span quartile (**SupFig S1B**) and to draw the MA-plot (**SupFig S1C**). Genes with a *P*-value from a paired t test ($n=3$) lower than 0.05 and a fold change greater than 50% were considered differentially expressed.

Signal evaluation of ChIP-seq, RNA-ChIPseq and RNA-seq on defined windows were performed as followed. Multiple regions based on the Ensembl annotations (hg19, version 87) were generated using the bedtools suite (v2.27.1) (Quinlan and Hall, 2010), 500 bp around the center of the intron that is closest to the center of the gene, 500 bp around the TSS, 500 bp around the TES, from TES to 1 kb downstream gene, from TES to 10 kb downstream gene and 40 kb to 41 kb downstream gene. In the case of the downstream region of genes, only regions retained were those not overlapping and distant by at least 10 kb with any other gene from the TES to the end of the counting region. For TSS regions, windows retained were those distant by at least 10 kb from both extremities of other genes. Consecutive 10 kb windows were generated after TES of genes with a maximum of 40 kb as long as there was no overlap with downstream genes oriented in the same direction. All previously generated counting tracks were converted to the SAF format and then used as annotation file in featureCounts (v1.6.1) (parameter: -s 2 for RNA-ChIP-seq or -s 0 for ChIP-seq). Regions with a *P*-value from a paired t test ($n=3$) lower than 0.05 and a fold change greater than 50% were considered differentially enriched between infected and mock cells. In the case of V5-PA RNA-ChIP the relative enrichment in infected cells was evaluated in comparison with non-immun IgG RNA-ChIP as explained in the legend of **Fig. 4D**.

Statistical analysis and graphs were produced with Microsoft Excel. Pathway analysis were performed with Enrichr (Kuleshov et al., 2016).

Acknowledgements

The authors thank our colleagues C. Rachez and S. Thouroude for helpful discussions. A special thanks also to Catherine Bodin for technical assistance and to Edith Ollivier for administrative assistance.

Fundings

J.Y. is part of the Pasteur - Paris University (PPU) International PhD Program; European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie [665807]. Ministry of Research (MENRT; to J.B.); Centre National de Recherche Scientifique (CNRS; to E.B. and C.M.); Agence Nationale de la Recherche [ANR-11-BSV8-0013] ANR FluChromatin; REVIVE-Investissement d'Avenir (to J.Y., M.C., C.M.). Funding for open access charge: CNRS recurrent funding;

Conflict of interest statement. None declared.

References

- Ameyar-Zazoua, M., Rachez, C., Souidi, M., Robin, P., Fritsch, L., Young, R., Morozova, N., Fenouil, R., Descostes, N., Andrau, J.-C., et al. (2012). Argonaute proteins couple chromatin silencing to alternative splicing. *Nat. Struct. Mol. Biol.* *19*, 998–1004.
- Amorim, M.-J., Read, E.K., Dalton, R.M., Medcalf, L., and Digard, P. (2007). Nuclear export of influenza A virus mRNAs requires ongoing RNA polymerase II activity. *Traffic Cph. Den.* *8*, 1–11.
- Bauer, D.L.V., Tellier, M., Martínez-Alonso, M., Nojima, T., Proudfoot, N.J., Murphy, S., and Fodor, E. (2018). Influenza Virus Mounts a Two-Pronged Attack on Host RNA Polymerase II Transcription. *Cell Rep.* *23*, 2119–2129.e3.
- Boehm, A.K., Saunders, A., Werner, J., and Lis, J.T. (2003). Transcription factor and polymerase recruitment, modification, and movement on dhsp70 in vivo in the minutes following heat shock. *Mol. Cell. Biol.* *23*, 7628–7637.
- Chow, K.T., Gale, M., and Loo, Y.-M. (2018). RIG-I and Other RNA Sensors in Antiviral Immunity. *Annu. Rev. Immunol.* *36*, 667–694.
- Chung, M., Cho, S.Y., and Lee, Y.S. (2018). Construction of a Transcriptome-Driven Network at the Early Stage of Infection with Influenza A H1N1 in Human Lung Alveolar Epithelial Cells. *Biomol. Ther.* *26*, 290–297.
- Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A., Szczesniak, M.W., Gaffney, D.J., Elo, L.L., Zhang, X., et al. (2016). A survey of best practices for RNA-seq data analysis. *Genome Biol.* *17*, 13.
- Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., et al. (2021). Twelve years of SAMtools and BCFtools. *GigaScience* *10*, giab008.
- Dias, A., Bouvier, D., Crépin, T., McCarthy, A.A., Hart, D.J., Baudin, F., Cusack, S., and Ruigrok, R.W.H. (2009). The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit. *Nature* *458*, 914–918.
- Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinforma. Oxf. Engl.* *29*, 15–21.
- ENCODE Project Consortium, Dunham, I., Kundaje, A., Aldred, S.F., Collins, P.J., Davis, C.A., Doyle, F., Epstein, C.B., Frietze, S., Harrow, J., et al. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature* *489*, 57–74.
- Engelhardt, O.G., Smith, M., and Fodor, E. (2005). Association of the Influenza A Virus RNA-Dependent RNA Polymerase with Cellular RNA Polymerase II. *J. Virol.* *79*, 5812–5818.
- Ferhadian, D., Contrant, M., Printz-Schweigert, A., Smyth, R.P., Paillart, J.-C., and Marquet, R. (2018). Structural and Functional Motifs in Influenza Virus RNAs. *Front. Microbiol.* *9*, 559.
- Heinz, S., Texari, L., Hayes, M.G.B., Urbanowski, M., Chang, M.W., Givarkes, N., Rialdi, A., White, K.M., Albrecht, R.A., Pache, L., et al. (2018). Transcription Elongation Can Affect Genome 3D Structure. *Cell* *174*, 1522–1536.e22.
- Ho, C.K., and Shuman, S. (1999). Distinct roles for CTD Ser-2 and Ser-5 phosphorylation in the recruitment and allosteric activation of mammalian mRNA capping enzyme. *Mol. Cell* *3*, 405–411.
- Kolev, N.G., and Steitz, J.A. (2005). Symplekin and multiple other polyadenylation factors participate in 3'-end maturation of histone mRNAs. *Genes Dev.* *19*, 2583–2592.
- Komarnitsky, P., Cho, E.J., and Buratowski, S. (2000). Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes Dev.* *14*, 2452–2460.
- Kuleshov, M.V., Jones, M.R., Rouillard, A.D., Fernandez, N.F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S.L., Jagodnik, K.M., Lachmann, A., et al. (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* *44*, W90–97.
- Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* *9*, 357–359.
- Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinforma. Oxf. Engl.* *30*, 923–930.
- Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* *15*, 550.
- Lukarska, M., Fournier, G., Pflug, A., Resa-Infante, P., Reich, S., Naffakh, N., and Cusack, S. (2017). Structural basis of an essential interaction between influenza polymerase and Pol II CTD. *Nature* *541*, 117–121.
- Marcos-Villar, L., Pazo, A., and Nieto, A. (2016). Influenza Virus and Chromatin: Role of the CHD1 Chromatin Remodeler in the Virus Life Cycle. *J. Virol.* *90*, 3694–3707.
- Martínez-Alonso, M., Hengrung, N., and 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*, 6014–6021.

Mauger, O., Klinck, R., Chabot, B., Muchardt, C., Allemand, E., and Batsché, E. (2015). Alternative splicing regulates the expression of G9A and SUV39H2 methyltransferases, and dramatically changes SUV39H2 functions. *Nucleic Acids Res.* *43*, 1869–1882.

Mullani, N., Porozhan, Y., Mangelinck, A., Rachez, C., Costallat, M., Batsché, E., Goodhardt, M., Cenci, G., Mann, C., and Muchardt, C. (2021). Reduced RNA turnover as a driver of cellular senescence. *Life Sci. Alliance* *4*.

Park, D., Shivram, H., and Iyer, V.R. (2014). Chd1 co-localizes with early transcription elongation factors independently of H3K36 methylation and releases stalled RNA polymerase II at introns. *Epigenetics Chromatin* *7*, 32.

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

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinforma. Oxf. Engl.* *26*, 841–842.

Ramirez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* *44*, W160–165.

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

Rialdi, A., Hultquist, J., Jimenez-Morales, D., Peralta, Z., Campisi, L., Fenouil, R., Moshkina, N., Wang, Z.Z., Laffleur, B., Kaake, R.M., et al. (2017). The RNA Exosome Syncs IAV-RNAPII Transcription to Promote Viral Ribogenesis and Infectivity. *Cell* *169*, 679–692.e14.

Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. *Nat. Biotechnol.* *29*, 24–26.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* *26*, 139–140.

Shen, Y., Yue, F., McCleary, D.F., Ye, Z., Edsall, L., Kuan, S., Wagner, U., Dixon, J., Lee, L., Lobanenkova, V.V., et al. (2012). A map of the cis-regulatory sequences in the mouse genome. *Nature* *488*, 116–120.

Sikora, D., Rocheleau, L., Brown, E.G., and Pelchat, M. (2014). Deep sequencing reveals the eight facets of the influenza A/HongKong/1/1968 (H3N2) virus cap-snatching process. *Sci. Rep.* *4*, 6181.

Tani, H., Mizutani, R., Salam, K.A., Tano, K., Ijiri, K., Wakamatsu, A., Isogai, T., Suzuki, Y., and Akimitsu, N. (2012). Genome-wide determination of RNA stability reveals hundreds of short-lived noncoding transcripts in mammals. *Genome Res.* *22*, 947–956.

Tarazona, S., García-Alcalde, F., Dopazo, J., Ferrer, A., and Conesa, A. (2011). Differential expression in RNA-seq: A matter of depth. *Genome Res.* *21*, 2213–2223.

Walker, A.P., and Fodor, E. (2019). Interplay between Influenza Virus and the Host RNA Polymerase II Transcriptional Machinery. *Trends Microbiol.* *27*, 398–407.

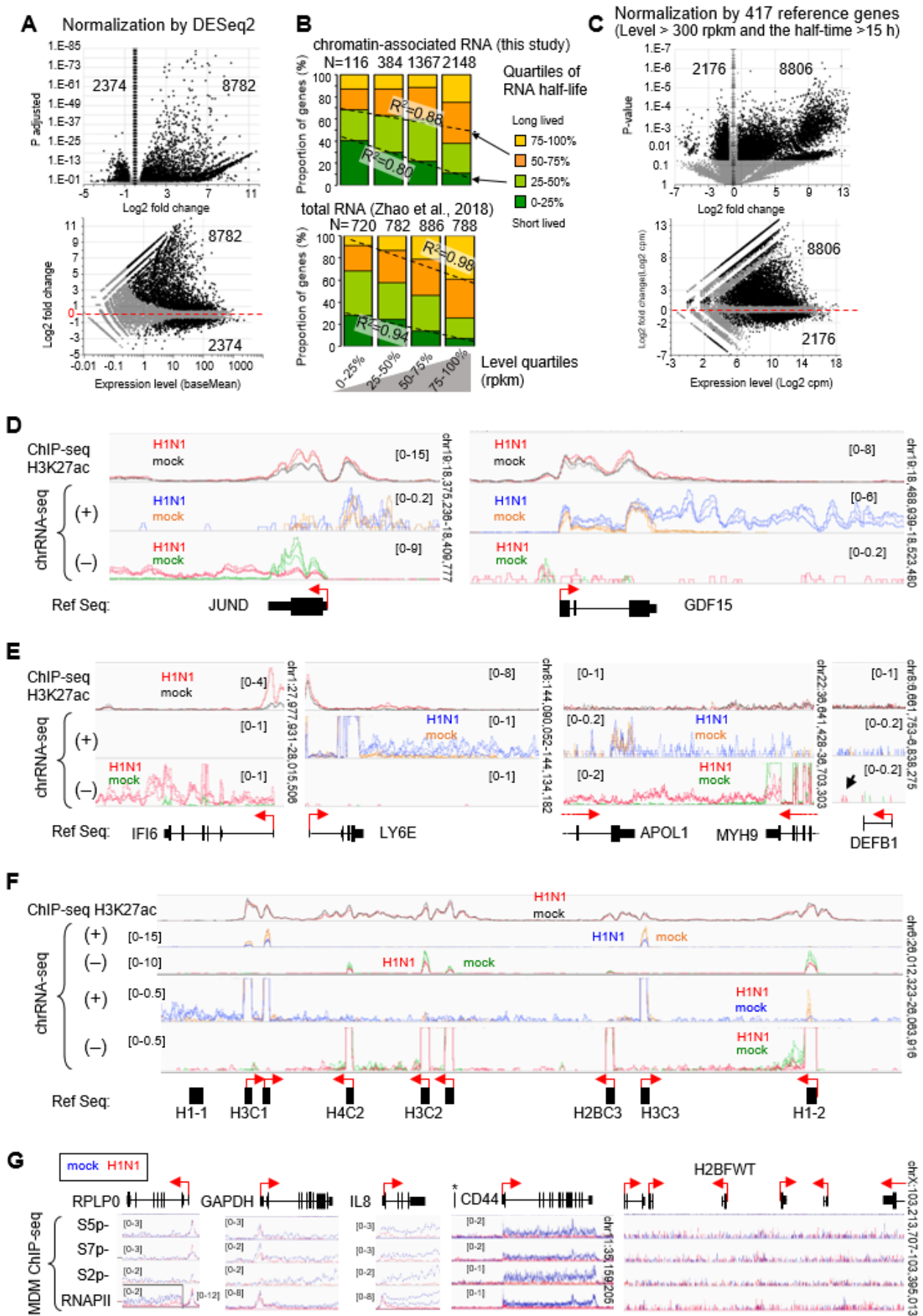
Wang, C., Forst, C.V., Chou, T., Geber, A., Wang, M., Hamou, W., Smith, M., Sebra, R., Zhang, B., Zhou, B., et al. (2020). Cell-to-Cell Variation in Defective Virus Expression and Effects on Host Responses during Influenza Virus Infection. *MBio* *11*.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoutte, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* *9*, R137.

Zhao, N., Sebastiano, V., Moshkina, N., Mena, N., Hultquist, J., Jimenez-Morales, D., Ma, Y., Rialdi, A., Albrecht, R., Fenouil, R., et al. (2018). Influenza virus infection causes global RNAPII termination defects. *Nat. Struct. Mol. Biol.* *25*, 885–893.

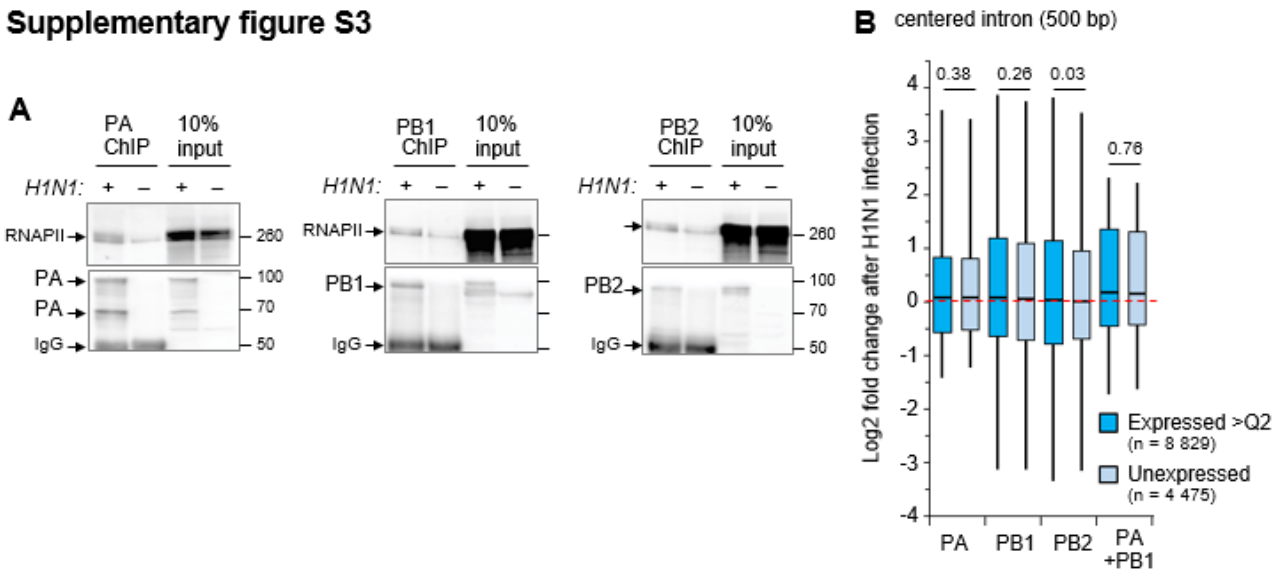
Zhou, B., Li, J., Liang, X., Yang, Z., and Jiang, Z. (2017). Transcriptome profiling of influenza A virus-infected lung epithelial (A549) cells with lacticresinol-4- β -D-glucopyranoside treatment. *PLOS ONE* *12*, e0173058.

Zhu, J., Adli, M., Zou, J.Y., Verstappen, G., Coyne, M., Zhang, X., Durham, T., Miri, M., Deshpande, V., De Jager, P.L., et al. (2013). Genome-wide chromatin state transitions associated with developmental and environmental cues. *Cell* *152*, 642–654.

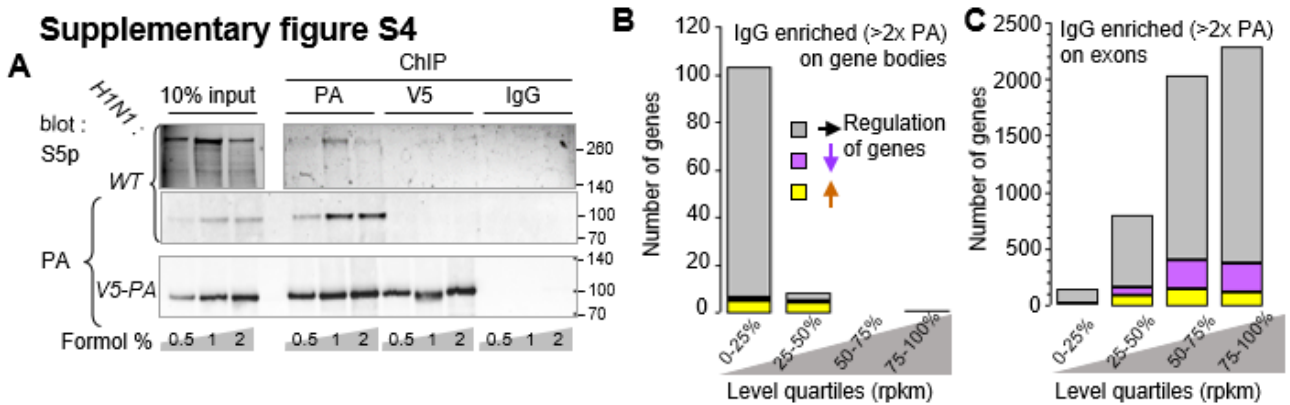


Supplementary figure S1 : Differential gene expression at 6 h in H1N1-infected A549 cells versus non-infected cells

.A) Differential RNA levels between H1N1-infected cells versus non-infected cells using DESeq2 package. Variations have been considered significant (black dot) for a mean fold change (FC)>1.5 and the adjusted p-value <0.05. The number of black dots were indicated. **B)** The top graph indicates the proportion of genes ranked by their level of RNA associated to the chromatin in A549 cells and their half-life as evaluated in HeLa cells (Tani et al., 2012). Among the 4678 genes investigated for their RNA life span, the number of genes expressed in each quartile in A549 cells are indicated on the top of graph. Pearson's coefficient of determination (R^2) were calculated between minimal RNA level of each quartile (rpkm) and the percentage of genes with the half-time indicated by the dotted lines and arrow. The bottom graph is the same analysis performed with the mean of RNA levels from total RNA extract in uninfected and 12-hpi infected A549 cells (Zhao et al. 2018). **C)** Differential levels of RNA associated to the chromatin using the normalization based on the average variation of 424 reference genes expressing higher level (> 10 log₂ cpm) of transcripts with long half-life (> 15 h). Variations have been considered significant (black dot) for a mean fold change is >1.5 and the p-value <0.05 calculated by a paired T-test (two-tailed) on the log₂(cpm). **D, E, F)** Examples illustrating the variation in the distribution of reads in H1N1 infected cells from IGV visualization of indicated loci. The RNA-seq independent triplicate were overlaid in the same color as indicated, for each strand orientation. The top track shows the enrichment H3K27ac on the chromatin of infected or mock cells. The scale range of the tracks were indicated in brackets. **G)** Distribution of indicated phosphorylated (on Serine5, 7 and 2) and whole RNAPII on genes analysed **Fig 1E**. ChIP-seq assays with antibodies targeting the indicated RNAPII were conducted in infected (red line) or uninfected MDM cells (blue line) were used from GSE103477 datasets (Heinz et al., 2018). The range scale are indicated in brackets

Supplementary figure S3

Supplementary figure S3 : immunoprecipitation efficiencies of Antibodies directed against the Flu Pol subunits. **A)** Infected (+) or uninfected (-) cells were extracted and subjected to immunoprecipitation with the indicated antibodies used in ChIP assay. Eluates and 10% of the extracts (input) were resolved by western blot. The lower part of the membranes (50 to 120 kDa) was revealed with the corresponding antibodies and the upper part (120 kDa-top) with RNAPII antibodies. **B)** The same analysis, as in **fig. 3B**, has been conducted for the averages of intronic coverage. For each gene the more central intron has been used to defined a centered 500 bp windows of counting. Only windows separated by 10 kb from TSS were considered.

Supplementary figure S4

Supplementary figure S4 : immunoprecipitation efficiencies of Antibodies directed against the Flu Pol subunits. **A)** A549 cells infected cells by WT IAV or V5-PA engineered IAV were crosslinked by indicated % of formaldehyde or not. Chromatin fraction were subjected to immunoprecipitation with PA, V5 antibodies or non immun IgG. Eluates and 10% of extracts were resolved by western blot. The lower part of the membranes was revealed with the PA antibodies and the upper part with phospho-RNAPII antibodies. **B, C)** The same analysis as described for **figure 4D, E** was carried out but this time evaluating the signal enrichment for IgG control compared to V5-PA

Partie III

SÉLECTION ET PROPRIÉTÉS DE NANOLIGANDS SPÉCIFIQUES DE FLUPOL

Introduction

Comme discuté précédemment dans ce manuscrit, FluPol interagit avec de nombreuses protéines au cours du cycle viral, que ce soit avec des protéines virales et/ou avec des partenaires cellulaires. Chacune de ces interactions joue un rôle primordial dans le bon déroulement de la prolifération virale. En effet, la liaison de FluPol avec la protéine virale NP (Biswas *et al*, 1998; Poole *et al*, 2004) permet une bonne cohésion des RNP (Klumpp *et al*, 1997) et l’empaquetage d’un nombre suffisant de FluPol par particule virale (une FluPol par segment viral). Les interactions avec des protéines cellulaires ne jouent pas un moindre rôle. En réalité, ces interactions sont essentielles à toutes les fonctions de FluPol. Par exemple, le dimère PA-PB1 est transporté au noyau par l’intermédiaire de l’importine- β RanBP5 (Deng *et al*, 2006a) alors que c’est l’interaction avec des importines- α qui permet à PB2 d’entrer au noyau (Resa-Infante *et al*, 2008). L’interaction avec ANP32A serait également nécessaire en permettant la dimérisation de FluPol, créant ainsi une plateforme pour son activité réplicase (Carrique *et al*, 2020), alors que l’interaction de FluPol avec l’ARN-polymérase II permet son activité transcriptase via le vol de coiffe (Lukarska *et al*, 2017; Plotch *et al*, 1981).

Identifier et caractériser la fonction de ces interactions est nécessaire pour mieux comprendre le fonctionnement de FluPol dans le noyau de la cellule, mais elles restent difficiles à reconnaître de par leur caractère souvent transitoire. L’obtention de nouveaux outils comme des nanoligands spécifiques de FluPol potentiellement capables de bloquer ou d’interférer avec ces interactions peut donc permettre de mieux comprendre le fonctionnement de FluPol. Au cours de ce projet, deux types de nanoligands protéiques spécifiques de FluPol ont été obtenus: des VHH (ou « nanobodies ») et des protéines artificielles nommées α Rep. Ces outils ont été utilisés pour analyser leur capacité à interférer avec le fonctionnement de FluPol.

1) Les VHH

Les anticorps conventionnels (ou IgG) sont composés de deux chaînes lourdes reliées par deux ponts disulfures et deux chaînes légères (Figure 32A). Les chaînes lourdes présentent une partie constante constituée de trois domaines (CH1, CH2 et CH3) et une partie variable (VH). Les chaînes légères suivent un schéma similaire avec un domaine variable (VL) et un domaine constant (CL) et sont reliées aux chaînes lourdes de façon non covalente. Les domaines variables sont responsables de la reconnaissance de l’antigène et de la spécificité de l’anticorps pour sa cible.

Les animaux de la famille des camélidés produisent des anticorps conventionnels, mais également un autre type d'anticorps constitués uniquement de chaînes lourdes (Hamers-Casterman *et al*, 1993) dans des proportions variables en fonction des espèces (environ 45% du sérum pour les lamas, 75% pour les dromadaires et chameaux). En raison d'une mutation au niveau d'un site d'épissage, ces anticorps sont dépourvus de domaine CH1 ainsi que de la chaîne légère en partie attachée (Nguyen *et al*, 1999; Woolven *et al*, 1999). Un VHH correspond ainsi au domaine variable de ces anticorps particuliers.

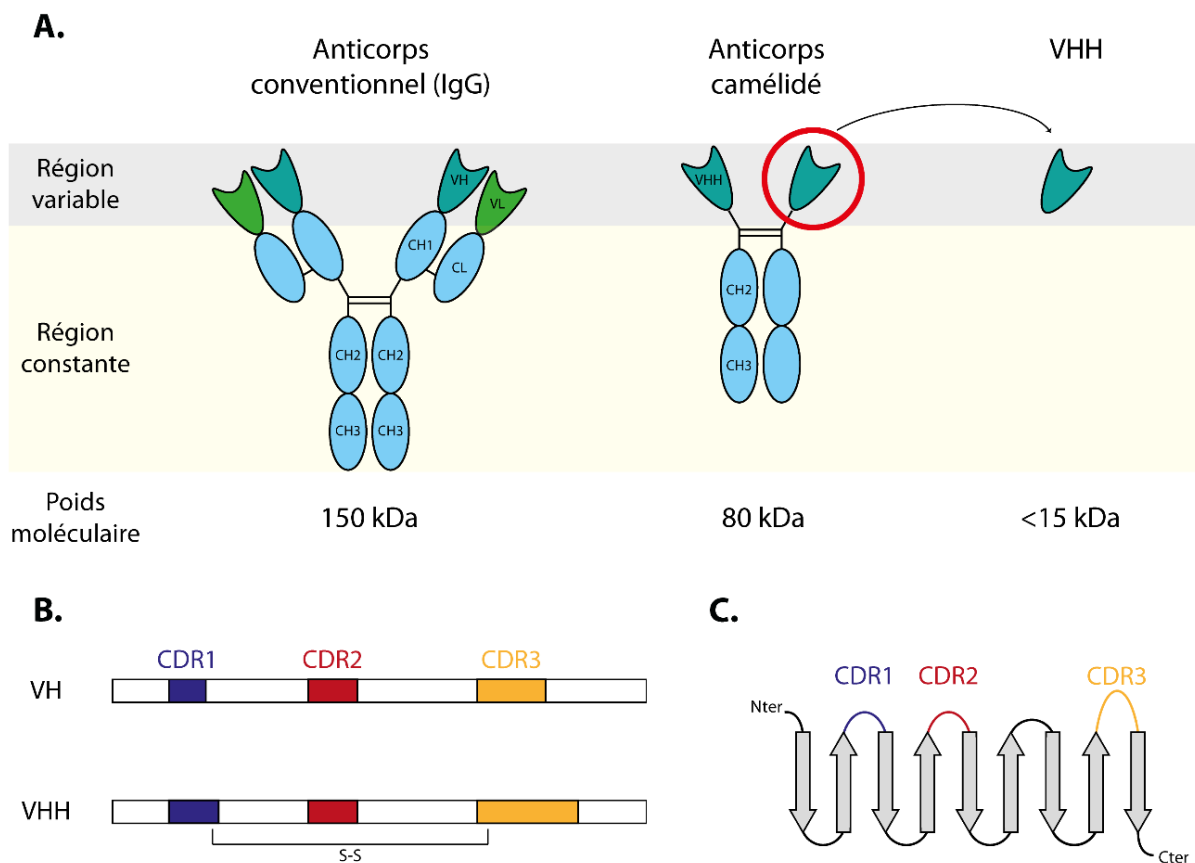


Figure 32 : Qu'est-ce qu'un VHH ?

A. Contrairement aux anticorps dits conventionnels (ou IgG), les anticorps de camélidés sont dépourvus de chaînes légères, ce qui leur confère une plus petite taille. Le VHH correspond à la partie variable d'un anticorps de camélidé et n'est donc constitué que d'une chaîne lourde. Son poids moléculaire est ainsi inférieur à 15 kDa. **B.** La structure du VHH est similaire à celle de la région variable de la chaîne lourde d'une IgG. La reconnaissance spécifique de l'antigène s'opère grâce à trois régions variables appelées CDR (Complementary Determining Region). Généralement, les CDR 1 et 3 sont plus long chez le VHH que le VH. Ils sont également connectés par un pont disulfure qui stabilise la conformation. Le CDR3 est souvent le CDR le plus long chez les VHH. **C.** La structure du VHH est constituée d'un enchaînement de brins β organisés de façon antiparallèle. Les trois CDR forment des jonctions entre deux brins β et sont situés d'un même côté pour former le site de reconnaissance de l'antigène.

Les VHH présentent une architecture comparable à celle des VH. La reconnaissance de l'antigène se fait grâce à trois CDR (Complementary Determining Region). Leur comparaison montre que le CDR3 des VHH est plus long que ceux d'un VH d'un anticorps conventionnel (Muyldermans *et al*, 1994; Vu *et al*, 1997) et est parfois stabilisé par un pont disulfure qui le relie au CDR1 (Figure 32B). Ainsi, le VHH est formé d'un enchaînement de brins β reliés entre eux par des boucles (Figure 32C). Les CDR1, 2 et 3 forment donc des boucles situées à un même pôle de la protéine.

Ces VHH possèdent donc des propriétés uniques qui en font une alternative intéressante aux IgG conventionnels. Premièrement, malgré leur taille dix fois inférieure à celle des IgG (moins de 15 kDa contre 150 kDa) et l'absence de chaîne légère, les VHH peuvent avoir une affinité de l'ordre du nanomolaire voire du picomolaire, similaire à celle des IgG. Cette petite taille permet aussi aux VHH de reconnaître des structures non accessibles à des IgG et d'atteindre des cibles enfouies dans des poches (Conrath *et al*, 2009; De Genst *et al*, 2006; Desmyter *et al*, 1996, 2001; Lauwereys *et al*, 1998; Transue *et al*, 1998). En outre, ces molécules sont très stables, y compris dans des conditions extrêmes. En effet, les VHH résistent la plupart du temps à des températures jusqu'à 80°C et peuvent retrouver leur conformation originelle ainsi que leur activité après un choc thermique pouvant aller jusqu'à 90°C (Ewert *et al*, 2002; Lauwereys *et al*, 1998; van der Linden *et al*, 1999; Pérez *et al*, 2001). Ils sont également souvent résistants aux hautes pressions, aux détergents et aux pH acides ou basiques (Dolk *et al*, 2005; Dumoulin *et al*, 2002). Contrairement aux IgG, les VHH ne nécessitent pas une expression en cellules eucaryotes ce qui permet une expression facile et peu coûteuse en *Escherichia coli* avec un fort rendement. Ils peuvent également être produits ou fonctionner dans des compartiments cellulaires tels que le réticulum endoplasmique ou le cytoplasme (Klooster *et al*, 2009). Enfin, les VHH peuvent facilement et rapidement être multimérisés pour augmenter leur affinité à l'aide de linkers ou de protéines cargo. Des protéines multimériques (bivalente, trivalente, ... voire décavalentes) sont régulièrement décrites (Els Conrath *et al*, 2001; Groot *et al*, 2006; Hmila *et al*, 2010; Hultberg *et al*, 2011; Stewart *et al*, 2007; Stone *et al*, 2007; Wichgers Schreur *et al*, 2020; Zhang *et al*, 2004)! Des constructions multispécifiques sont également réalisées.

Depuis la première génération de VHH ciblant un virus en 2002 (un VHH ciblant un bactériophage (Ledebøer *et al*, 2002)), de nombreux VHH dirigés contre des pathogènes ont été développés, notamment contre les virus Influenza. Le tableau 1 recense tous les VHH générés contre les virus Influenza à ce jour. L'hémagglutinine a été la première protéine ciblée et reste celle contre laquelle le plus de VHH ont été sélectionnés, notamment dans le but de mettre au point un antiviral universel pour tous les types d'Influenza. Seul un VHH ciblant FluPol a été récemment produit (Fan *et al*, 2019b). Ce VHH inhibe la réplication du génome en se liant à l'interface de dimérisation de FluPol.

Tableau 1 : VHH/nanobodies spécifiques des protéines du virus Influenza

Cible	Référence	Souche	Observations
HA	(Hultberg <i>et al</i> , 2011)	H5N1	Multimérisation (trivalent) Neutralise clades 1 et 2 et souche NIBRG14
	(Ibañez <i>et al</i> , 2011)	H5N1	Multimérisation (bivalent) Cible site antigénique B Protection en souris
	(Hufton <i>et al</i> , 2014) (Gaiotto & Hufton, 2016) (Del Rosario <i>et al</i> , 2020)	H1N1 pdm09	Multimérisation (bivalent) Cross-réactivité (H5N1, H2N2, H9N2) Cible domaine conservé HA2 Humanisé (fusion Fc) Protection en souris (vecteur AAV)
	(Laursen <i>et al</i> , 2018)	H1N1 H3N2 H2 H7 FluB	Multimérisation (tétravalent x2) Humanisé (fusion Fc) Cible les épitopes conservés Protection en souris (vecteur AAV)
	(Ramage <i>et al</i> , 2019)	FluB	Cross-réactivité Cible les épitopes conservés
	(Gaiotto <i>et al</i> , 2021)	H7	Multimérisation (bivalent) Cross-réactivité (H7)
	M2	(Wei <i>et al</i> , 2011)	H3N2
(De Vlioger <i>et al</i> , 2019)		FluA	Cible M2e Fusion à VHH anti- récepteur Fc (FcγR) de souris pour activation immunité effectrice Protection des souris (injection intranasale)
Virus entier	(Zhu <i>et al</i> , 2014b)	H3N2	Diagnostic Billes magnétiques couplées au VHH
	(Zhu <i>et al</i> , 2014a)	H5N1	Diagnostic Double sandwich ELISA
	(Schmidt <i>et al</i> , 2016)	FluA	Ciblent uniquement NP en intracellulaire Bloquent import des vRNP au noyau ou transcription Protection en cellules A549
NA	(Cardoso <i>et al</i> , 2014)	H5N1	Multimérisation (bivalent) Fusion Fc souris Protection en souris (injection intranasale)
NP	(Ashour <i>et al</i> , 2015) (Hanke <i>et al</i> , 2016)	WSN	Bloque l'import des vRNP au noyau, réplication et transcription Même site d'interaction que la protéine antivirale Mx
	(Du <i>et al</i> , 2019)	FluA porc	Diagnostic ELISA
FluPol	(Fan <i>et al</i> , 2019b)	H5N1	Bloque la réplication par liaison à l'interface de dimérisation

2) Les α Rep

Dans l'objectif de sélectionner des nanoligands pouvant se fixer à une cible (généralement une protéine), l'équipe « Ingénierie des protéines et modélisation » du Professeur Philippe Minard à l'I2BC d'Orsay (Urvoas *et al*, 2010) a conçu et construit des banques de phages codant pour une famille de protéines artificielles. Ces protéines sont structurellement proches de protéines eucaryotes ou bactériennes constituées de répétitions d'un même motif élémentaire de vingt à quarante acides aminés de long (Kajava, 2001). Au sein de ce motif, certains acides aminés sont très conservés (ceux impliqués dans l'architecture de la protéine) alors que d'autres sont hypervariables. Ces résidus sont regroupés sur une même face concave de la protéine, formant ainsi une macro-surface pouvant interagir avec d'autres protéines. De telles protéines ont été naturellement sélectionnées au cours de l'évolution pour assurer une large gamme de fonctions. C'est le cas des protéines riches en répétitions de leucines (« *Leucin rich repeat proteins* ») à l'origine de la diversité du système immunitaire adaptatif chez les agnathes, poissons sans mâchoire (Pancer *et al*, 2004, 2005; Pancer & Cooper, 2006). En effet, cette diversité est obtenue par duplication, recombinaison puis diversification de ces motifs répétés riches en leucines. Ce processus évolutif de diversification peut être reproduit *in vitro* afin de créer des banques de protéines artificielles ayant un même motif répété (Binz *et al*, 2003; Binz & Plückthun, 2005).

Les protéines artificielles construites par le groupe du Pr. Minard et nommées α Rep sont constituées de répétitions d'un motif de 31 acides aminés (motif HEAT) formé de deux hélices α , flanquées d'un domaine N-Cap et d'un domaine C-Cap (Figure 33A). L'acronyme HEAT tient son origine dans les quatre premières protéines identifiées contenant ce motif : Huntingtin, le facteur d'élongation 3 (elongation factor 3), la sous-unité A de la phosphatase 2A (PP2A) et la kinase TOR1 (Andrade & Bork, 1995). Au sein de ce motif, six acides aminés sont variables (représentés en rouge sur la structure) et définissent la surface concave de la protéine. Chaque α Rep est donc unique, de par les acides aminés définissant sa surface concave et par le nombre de répétitions de son motif élémentaire (Figure 33B).

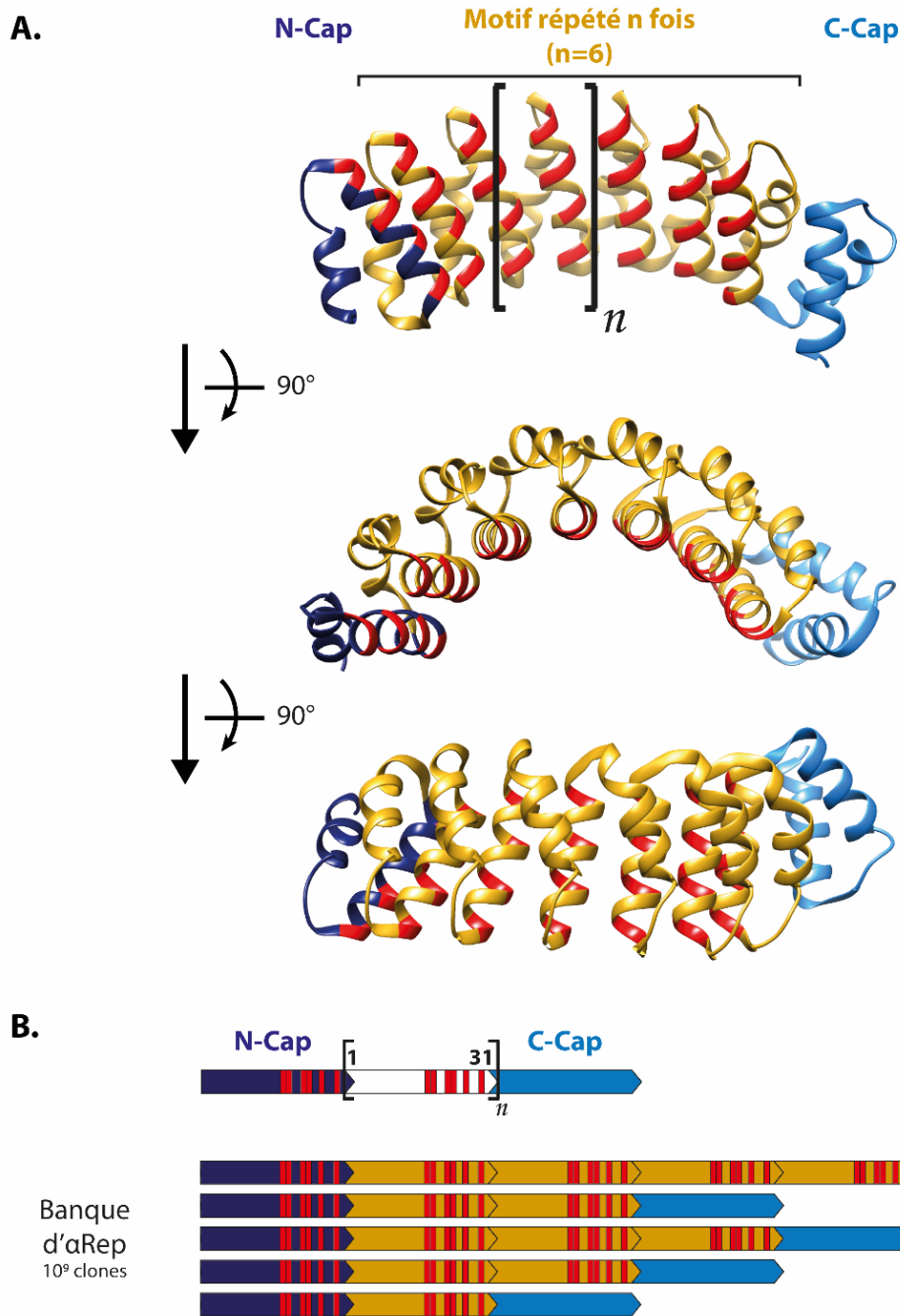


Figure 33 : Qu'est-ce qu'une α Rep ?

A. Structure de l' α Rep b-GFP-A qui reconnaît spécifiquement la GFP (PDB 4XL5). Le N-Cap est coloré en bleu marine, les motifs répétés en doré et le C-Cap en bleu clair. Les positions hypervariables sont représentées en rouge. **B.** Illustration de la diversité obtenue dans la banque d' α Rep.

Pour créer cette famille de protéines artificielles nommées α Rep, ces auteurs sont partis de la séquence primaire de la protéine Mth187 dont la structure (PDB 1TE4) comporte une extrémité N-terminale désordonnée, deux paires d'hélices α correspondant au motif répété de 31 acides aminés ainsi qu'une dernière paire d'hélices α qui constitue l'extrémité C-terminale. Ce motif répété a été comparé à une banque de donnée afin d'obtenir une séquence consensus avec les meilleurs résultats. Une séquence constituée de cinq répétitions de ce motif consensus a ensuite été à nouveau comparé

à cette banque de données afin d'identifier les positions conservées et les positions hypervariables au sein de ce motif. Ainsi, les positions 18, 19, 22, 23, 26 et 30 (en rouge sur la figure 33A) ont été identifiées comme étant hypervariables et situées sur la surface concave du motif (i.e. la seconde hélice α). Ces positions ont donc été sélectionnées pour introduire de la variabilité dans la banque. Pour le choix du N-Cap, le N-Cap de la protéine UniprotKB A0B7C6 a été retenu car cette protéine possède un N-Cap structuré en double hélice ainsi qu'une séquence consensus similaire à celle de Mth187 bien qu'un peu plus longue (34 acides aminés au lieu de 31). Le C-Cap de Mth187 a été conservé car il est déjà structuré en paire d'hélice α . Au final, une banque constituée de 10^9 clones différents a été générée.

3) Comment sélectionner des VHH ou des α Rep spécifiques?

Le processus de sélection final des VHH et des α Rep est très similaire. Dans les deux cas, il consiste à cribler une banque de phages M13 par « phage-display ». Pour la réalisation de la banque de phage-VHH, un lama ou un dromadaire est préalablement immunisé avec l'antigène d'intérêt et la banque est réalisée à partir des ARNm de lymphocytes B circulants. Pour la constitution de la banque de phage- α Rep, des minigènes codant le motif structural de 31 acides aminés et portant de la variabilité sur les six acides aminés définissant leur surface concave sont assemblés. Les ADNc produits à partir des ARNm de lymphocytes ou les gènes d' α Rep assemblés sont fusionnés à un gène de phage codant pour une protéine de capsid. Ainsi, les VHH et les α Rep sont exprimés en surface des phages, leur conférant la capacité à se fixer spécifiquement sur une cible. Les banques de phages utilisées (VHH ou α Rep) sont définies par 10^9 phages indépendants (Figure 34). Soulignons également l'existence de banques de phage-VHH synthétiques permettant de s'affranchir de l'étape d'immunisation et de construction de la banque de phage « post-immunisation » (Moutel *et al*, 2016).

Une fois les banques de phages construites, les nanoligands spécifiques de la protéine cible sont sélectionnés par « phage display ». Dans un premier temps, la cible est fixée à un support. Puis, les phages sont mis en contact avec la cible. De cette façon, les phages exprimant des nanoligands spécifiques de la cible y seront fixés. Après plusieurs lavages, les phages fixés sont élués puis amplifiés. Ces phages sont ensuite à nouveau mis au contact de la cible pour un nouveau tour de sélection. Par conséquent, seuls les phages exprimant un nanoligand spécifique de la cible sont amplifiés. Après n tours de sélection, des bactéries sont infectées par les phages puis étalées sur boîte afin d'obtenir des plages de lyse. Les plages sont amplifiées individuellement puis sélectionnés par phage-ELISA de façon à éliminer les phages non spécifiques de la cible. Les clones donnant un signal ELISA sont alors séquencés pour futures études fonctionnelles des nanoligands sélectionnés.

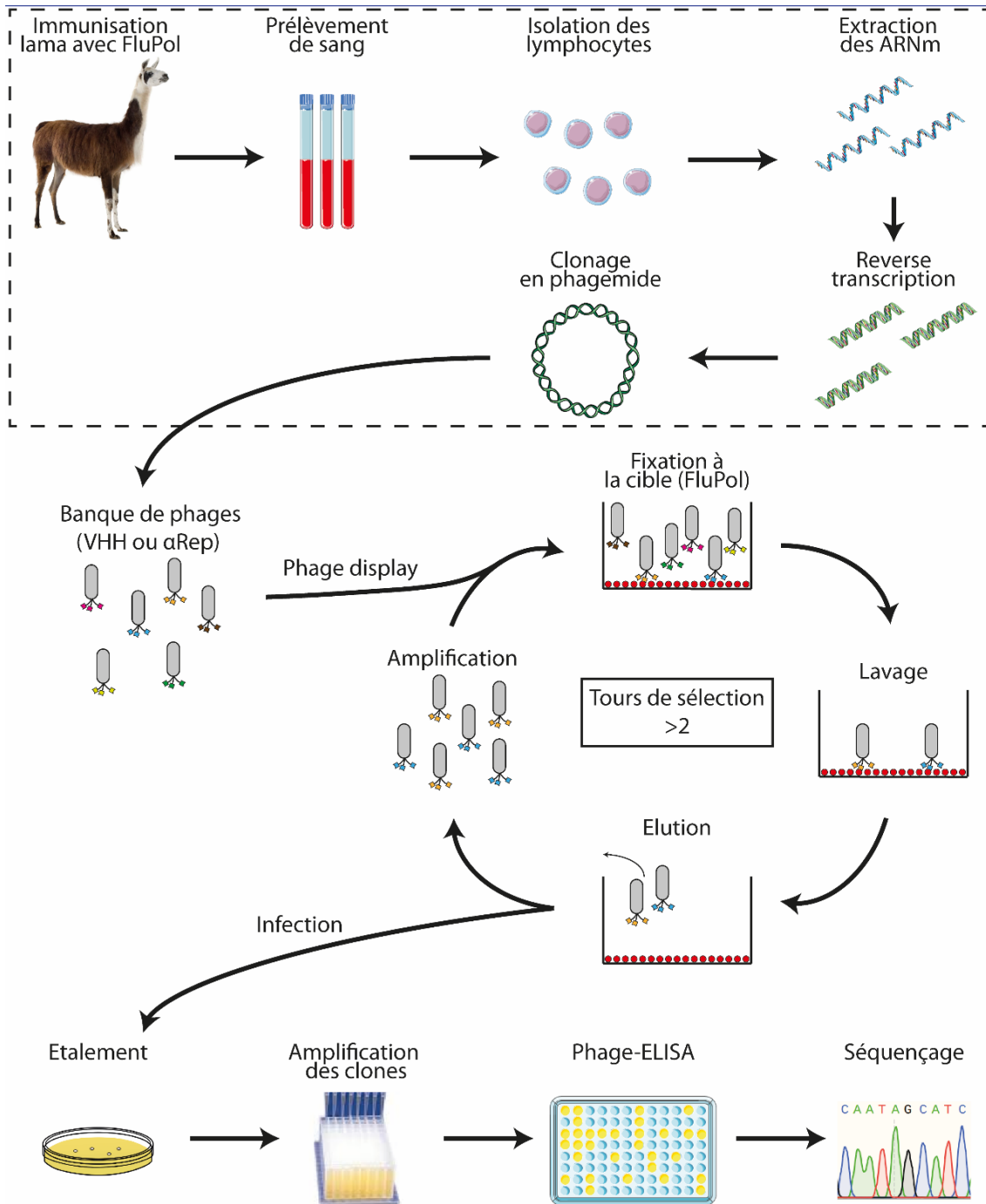


Figure 34 : Processus de sélection des nanoligands

Pour sélectionner des nanoligands, il faut avoir une banque de phages exprimant ces nanoligands à leur surface. Pour les α Rep, la banque est déjà construite. Pour les VHH, des étapes préliminaires sont nécessaires pour obtenir cette banque (encadré en pointillés). Dans un premier temps, un lama a été immunisé avec FluPol puis son sang a été prélevé afin d'en récolter les lymphocytes. Les ARNm en sont extraits puis « reverse transcrits » pour être clonés en phagemide et pouvoir produire la banque. Une fois la banque obtenue, les nanoligands spécifiques de FluPol sont sélectionnés par phage display. Ainsi, FluPol est fixée à un support puis les phages sont mis en contact avec. Après un lavage extensif, les phages fixés sont élués et amplifiés. Ce processus est répété plusieurs fois afin d'amplifier les phages spécifiques à FluPol. Enfin, des bactéries sont infectées par ces phages puis étalées pour amplifier des clones. Ces derniers sont ensuite testés par phage-ELISA. Les clones positifs sont alors séquencés pour obtenir la séquence du nanoligand sélectionné.

Objectifs

Afin de mieux étudier le fonctionnement de FluPol et comprendre ses différentes fonctions et les interactions que le complexe provoque au cours du cycle viral, nous avons voulu élaborer de nouveaux outils qui lui sont spécifiques. Nous avons donc souhaité obtenir des VHH et α Rep spécifiques de FluPol qui pourraient avoir un effet inhibiteur sur l'activité de cette dernière et étudier leur(s) mode(s) d'action.

Mes objectifs pour cette dernière partie de ma thèse étaient donc, dans un premier temps, de sélectionner des nanoligands spécifiques de FluPol puis, dans un second temps, de les caractériser sur le plan biochimique mais également fonctionnel.

Résultats

1) Les formes de FluPol utilisées pour la sélection de VHH et α Rep spécifiques

Dans l'objectif d'obtenir des VHH et des α Rep spécifiques de l'ARN-polymérase (FluPol) des virus influenza de type A, les domaines et sous-unités de FluPol ont été exprimés sous la forme d'une polyprotéine comme décrit dans (Swale *et al*, 2016). Les cadres de lecture codants pour les sous-unités de FluPol sont fusionnés ainsi qu'avec ceux de la protéase du « Tobacco Etch Virus » (pTEV) et de la protéine fluorescente cyan (CFP) dans un cadre de lecture unique dans la séquence suivante : TEV-HisTagPA-PB1-PB2-CFP (Figure 35). Chaque cadre de lecture des sous-unités de FluPol est séparé par une séquence d'ADN codant pour un peptide comprenant un court linker composé de sérine et glycine et du site de clivage à la pTEV. Les constructions ont été optimisées par l'emploi de codons préférentiellement utilisés pour expression recombinante en cellules d'insectes en utilisant le système MultiBac (Bieniossek *et al*, 2012, 2008; Trowitzsch *et al*, 2010). Lors de sa synthèse, la polyprotéine est clivée par pTEV, permettant un assemblage stœchiométrique des sous-unités de FluPol. Ainsi, deux formes tronquées de FluPol de la souche A/Victoria/3/1975(H3N2) permettant d'exprimer des quantités voisines du milligramme ont été exprimées, à savoir PA(197-716) - PB1 - PB2(1-116) [Forme A] et PA(197-716) – PB1(1-686) [Forme B] pour purification des produits d'assemblage de FluPol via un His-tag placé en N-terminal de la sous-unité PA. Ainsi, la forme A est dépourvue du domaine endonucléase de PA et des domaines « cap-binding », 627 et « NLS » de PB2 alors que la forme B également dépourvue du domaine endonucléase de PA contient PB1 sans son domaine d'interaction à PB2.

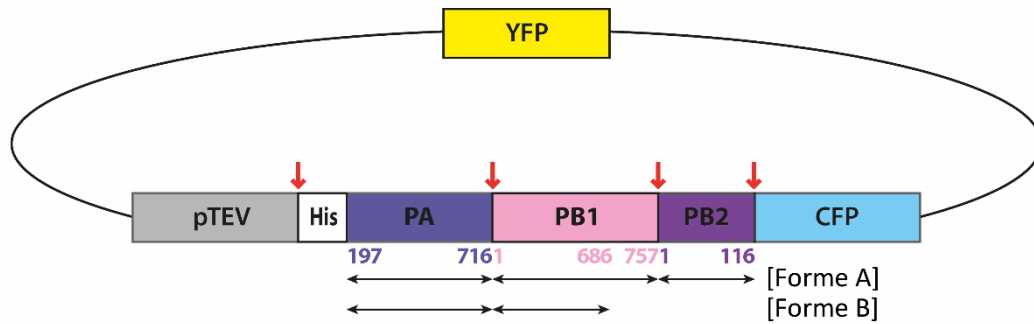


Figure 35 : Schéma du système d'expression de FluPol en cellules d'insecte

Les formes A et B de FluPol (flèches noires) sont exprimées sous forme d'une polyprotéine qui est co-traductionnellement clivée par la Tobacco Etch Virus (TEV) protéase suivant les sites de clivage de l'enzyme (flèches oranges). Ainsi, les séquences des sous-unités de FluPol accolées et flanquées des sites de clivage à pTEV et de la Cyan Fluorescent Protein (CFP) sont insérées dans le génome d'un baculovirus. La production de protéine peut donc être mesurée en temps réel grâce à la fluorescence émise par la CFP. La cassette de la Yellow Fluorescent Protein (YFP) est également insérée dans le génome du baculovirus afin de mesurer la bonne prolifération du baculovirus lors de l'expression.

Purification des formes de FluPol

Suite à l'expression des deux constructions dans le système baculovirus - cellules d'insecte, les formes A et B ont été purifiées par chromatographie d'affinité au nickel suivie d'une chromatographie d'affinité sur héparine.

La Figure 36 décrit la purification de la forme A. Les fractions éluées de la colonne de nickel et dans lesquelles les sous-unités PA et PB1 sont abondamment représentées (fractions 3 à 8) ont été rassemblées (Figure 36A) pour purification sur colonne héparine. Le chromatogramme révèle un double pic (Figure 36B) avec un pic plus élevé correspondant à une forme plus pure de FluPol (Figure 36C). Les fractions 25 à 28 correspondant à ce dernier pic ont été rassemblées pour le criblage des VHH. La pTEV représente un contaminant majeur de ces fractions.

La forme B (Figure 37) est produite en plus grande quantité et possède une meilleure stabilité que la forme A (Swale *et al*, 2016). Suite à une chromatographie d'affinité sur résine de nickel (Figure 37A), les fractions 3 à 10 résultantes ont été rassemblées pour purification sur résine d'héparine afin d'éliminer certains contaminants. Cette fois, le chromatogramme obtenu présente un pic majeur associé à un épaulement (Figure 37B). Après analyse des fractions par électrophorèse SDS-Page, les fractions 23 à 27 contenant la forme B ont été rassemblées pour criblage des α Rep (Figure 37C). Le rendement de production et de purification a été multiplié par un facteur 4 entre les formes A et B.

Ainsi, 2,3 mg de la forme A de FluPol ont été purifiés pour sélection de VHH spécifiques et 8,8 mg de la forme B de FluPol ont été produits pour criblage de la banque d' α Rep.

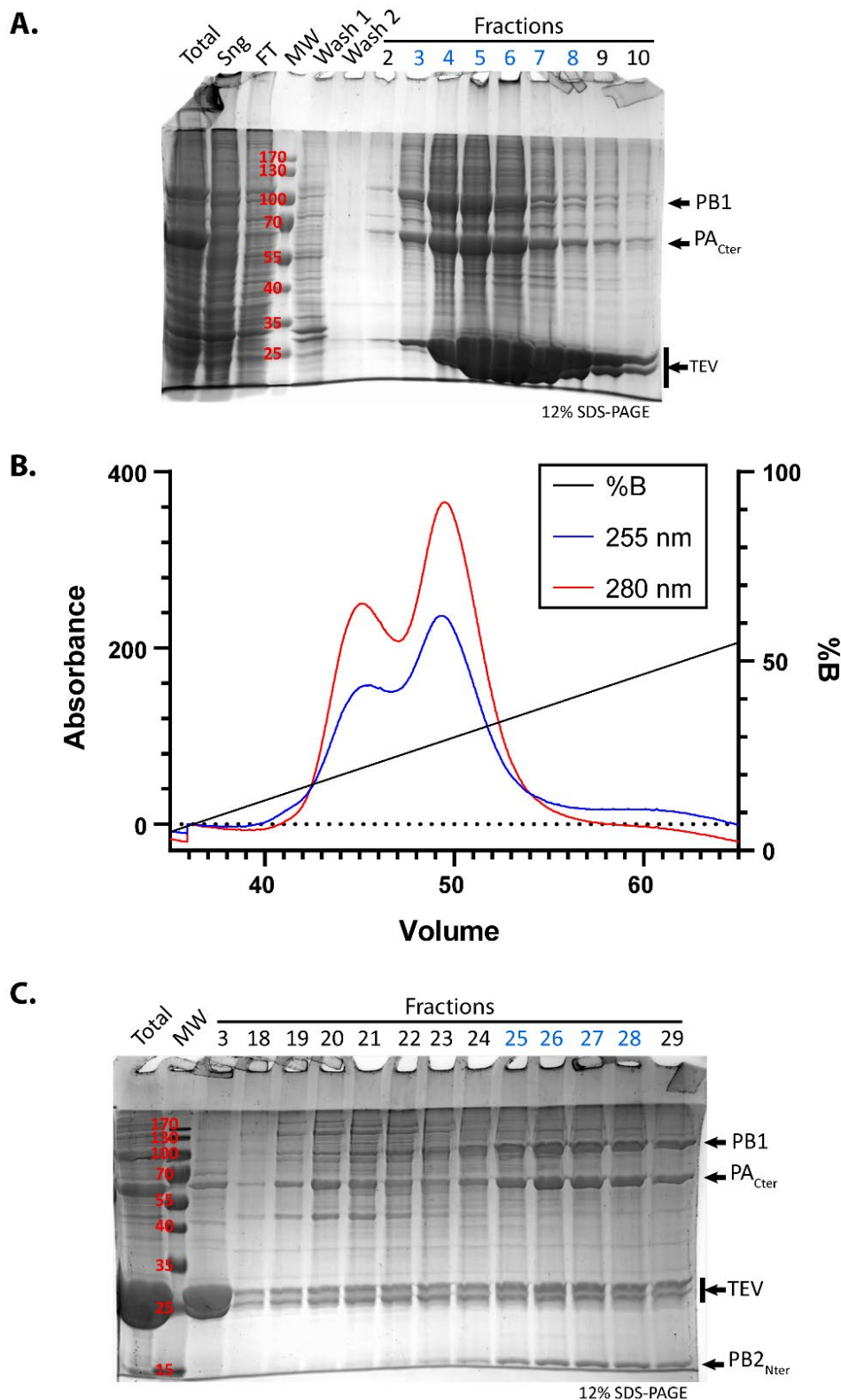


Figure 36 : Purification de l'hétérotrimère PA(197-716) - PB1 - PB2(1-120)

La forme A de FluPol est d'abord purifiée par chromatographie d'affinité (nickel). Les différentes fractions sont observées sur gel SDS-Page (**A**) puis rassemblées pour une étape de purification par chromatographie échangeuse d'ions sur résine (héparine) dont le chromatogramme est présenté en (**B**). Les fractions sont ensuite analysées sur gel SDS-Page (**C**).

Sng = surnagent, FT = Flow Through, Wash = lavage, MW = marqueur de poids moléculaires.

Les fractions en bleu correspondent aux fractions rassemblées.

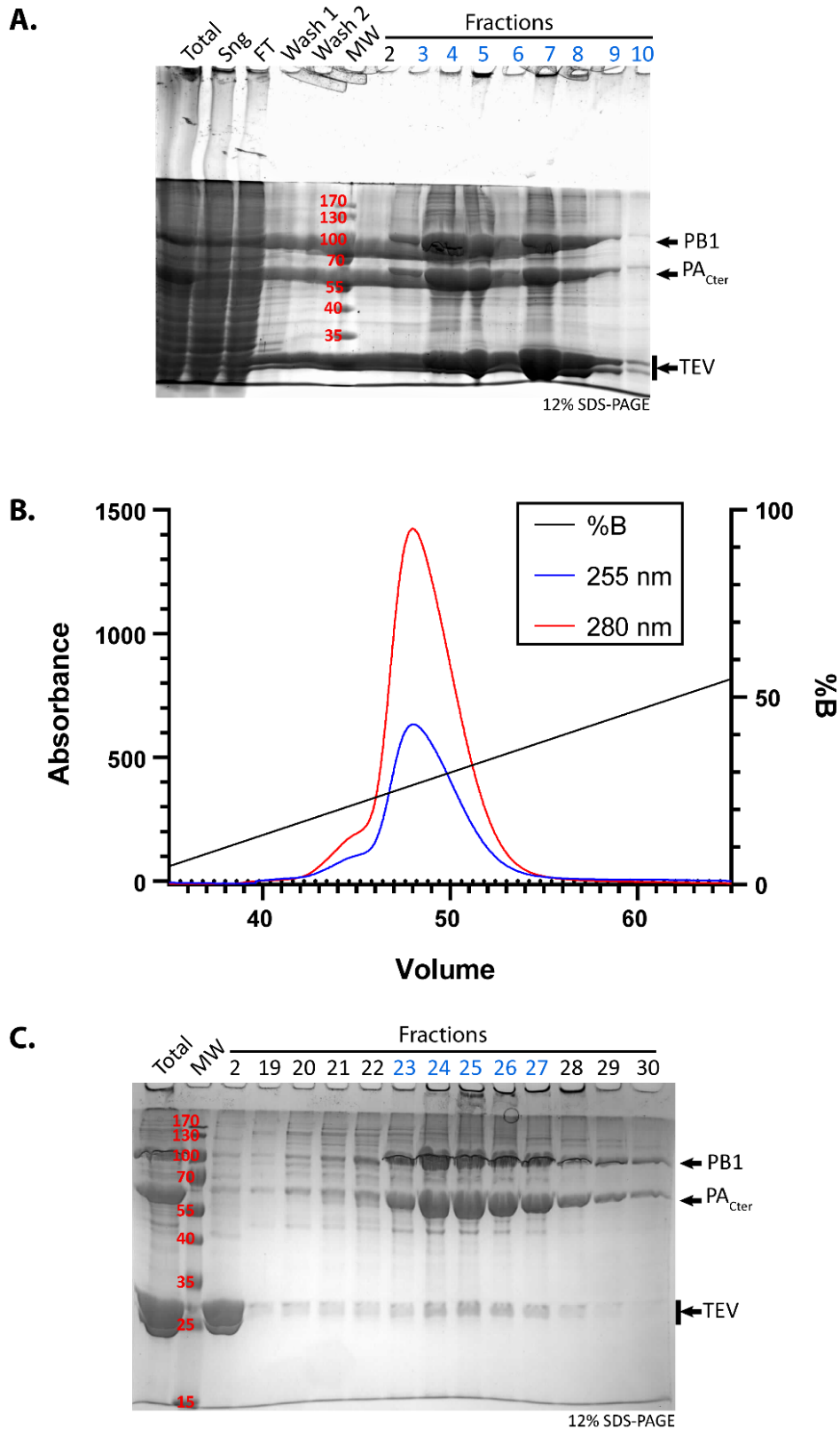


Figure 37 : Purification de l'hétérodimère PA(197-716) - PB1(1-686)

A. Première étape de purification par chromatographie d'affinité (nickel). Les différentes fractions sont observées sur gel SDS-Page. **B.** Chromatogramme de la purification par chromatographie échangeuse d'ions (héparine). **C.** Analyse des fractions de la chromatographie échangeuse d'ions sur gel SDS-Page Sng = surnageant, FT = Flow Through, Wash = lavage, MW = marqueurs de poids moléculaire. Les fractions en bleu correspondent aux fractions rassemblées.

2) Criblage des nanoligands

Obtention de VHH spécifiques de FluPol

L'obtention de VHH spécifiques de la forme A de FluPol a été réalisée par la plateforme « Sélection et production de VHH/nanobodies de camélidés » de l'AFMB à Marseille. Après immunisation d'un lama et sélection par « phage display », sept phages codant pour des VHH uniques et spécifiques de FluPol ont été obtenus. Leurs séquences sont présentées dans la figure 38. Ces VHH ont des tailles comprises entre 117 et 127 acides aminés, des points isoélectriques compris entre 4,3 et 9,5 et des masses prédites comprises entre 12,7 et 13,8 kDa (Tableau 2). Pour les tests fonctionnels en cellules eucaryotes, un épitope HA-Tag (YPYDVPDYG) a été fusionné à l'extrémité C-terminale de chacun des VHH. Pour l'expression en *E. coli*, une étiquette polyhistidine (His-tag) a été insérée en C-terminal de chacun des VHH à des fins de purification.

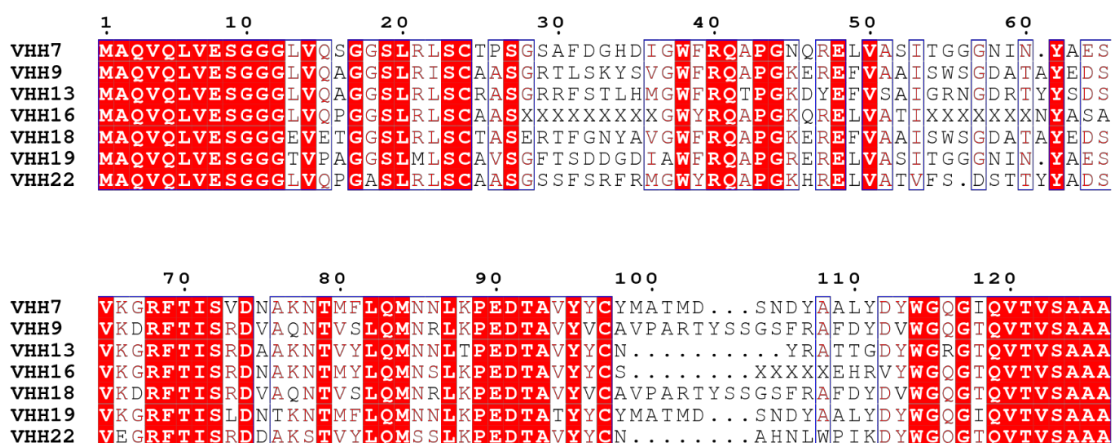


Figure 38 : Séquence des VHH spécifiques de FluPol

Alignement des séquences des sept VHH sélectionnés par phage-ELISA. Les acides aminés conservés sont surlignés en rouge. Figure réalisée avec ESPript (Robert & Gouet, 2014).

Tableau 2 : Tableau récapitulatif des propriétés physico-chimiques des VHH sélectionnés. Données obtenues par ProtParam.

Numéro du VHH	Nombre d'acides aminés	Poids moléculaire (kDa)	pHi
7	123	13,3	4,6
9	127	13,7	7,7
13	117	13,0	9,5
16	119	12,7	9,0
18	127	13,8	4,9
19	123	13,3	4,3
22	118	13,0	7,8

Dans le but de générer des VHH incapables de reconnaître FluPol, deux VHH ont été raccourcis de sept acides aminés dans leur domaine CDR3. Leurs séquences sont présentées en figure 39.

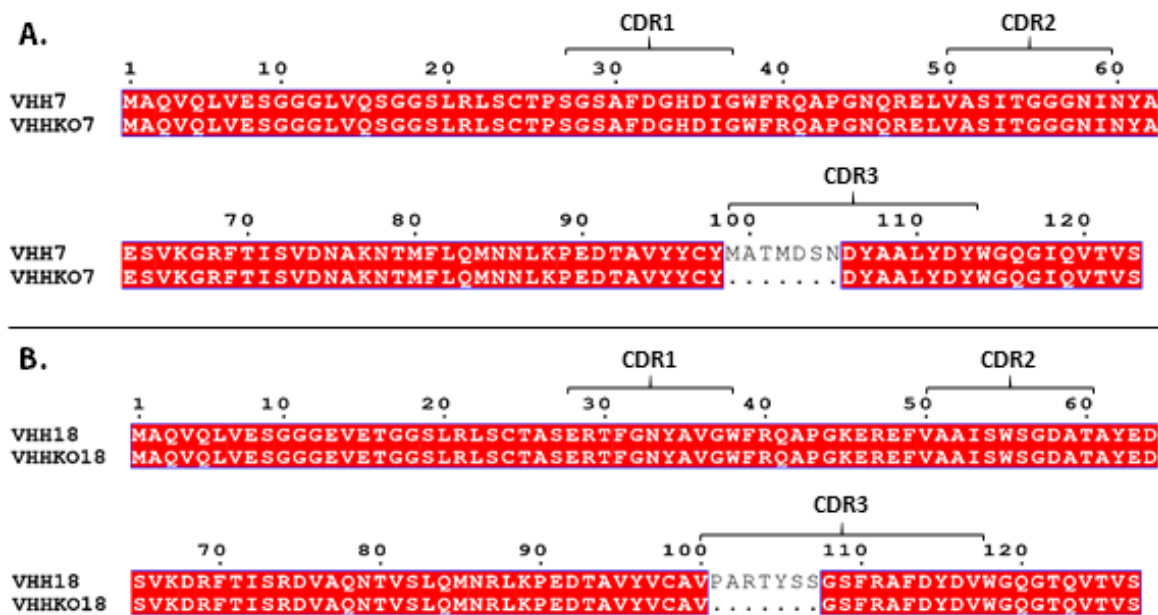


Figure 39 : Séquence des VHH

Alignement des séquences des VHH KO avec leurs homologues WT. Les acides aminés conservés sont surlignés en rouge. Figure réalisée avec ESPript.

Obtention d' α Rep spécifiques de FluPol

L'obtention d' α Rep spécifiques de la forme B de FluPol a été réalisée par criblage d'une banque de phages codant pour des protéines artificielles appelées α Rep (Urvoas et al. 2010). Le criblage a été réalisé par « phage display » suivant une procédure identique à la sélection des VHH. Quatre α Rep de séquences différentes ont été sélectionnées (Tableau 3). Chacune est unique de par sa taille (de 2, 4, 9 ou de plus de 10 motifs répétés de 31 résidus flanqués d'un domaine N-terminal et d'un domaine C-terminal). Chacun de ses motifs possède une séquence spécifique de 6 acides aminés formant la surface concave de l' α Rep. Chaque α Rep possède un 6xHis-tag en N-terminal de sa séquence pour purification après expression en *E. coli* ou pour immunodétection en cellules eucaryotes.

Tableau 3 : Tableau récapitulatif des quatre α Rep sélectionnées par phage-ELISA

Les séquences indiquées correspondent à la séquence des protéines exprimées en bactéries et en cellules eucaryotes. Le His-tag fusionné aux α Rep a été surligné en gras. Les acides aminés variables de chaque motif répété sont en rouge. La séquence soulignée correspond aux acides aminés exogènes. Données obtenues par ProtParam. AA = acides aminés.

N° α Rep	Séquence	Nombre d'AA	Nombre de repeat	Poids Moléculaire (kDa)	pHi
E3	<u>MRGSHHHHHH</u> TDPEKVE MYIKNLQDDSGLVRNAAATALGKI GDERAVEPLIKALKDEDWLVRYGAAFALGKI GDERAVEPLIKALKDEDDEVRLAAAGALGQI GDERAVEPLIKALKDEDYSVRENAAWALGKI GGERVRAAMEKLAETGTGFARKVAVNYLETHKSLIS	170	4	18,6	5,8
F9	<u>MRGSHHHHHH</u> TDPEKVE MYIKNLQDDSLAVRSNAADALGKI GDERAVEPLIKALKDEDSFVRYGAALALGKI RDERAVEPLIKALKDEDSGFVRQRAALALGKI GDERAVEPLIKALKDEDSFVRYGAALALGKI RDERAVEPLIKALKDEDSGFVRQRAALALGKI GDERAVEPLIKALKDEDSFVRYGAALALGKI RDERAVEPLIKALKDEDSGFVRQRAALALGKI GDERAVEPLIKALKDEDSFVRYGAALALGKI RDERAVEPLIKALKDEDSGFVRQRAALALGKI GDERAVEPLIKALKDESDVRYAAANALGKI GGERVRAAMEKLAETGTGFARKVAVNYLETHKSLIS	356 ?	10 ?	39,0	6,7
H3	<u>MRGSHHHHHH</u> TDPEKVE MYIKNLQDDSPTVRYTAASALGKI GDERAVEPLIKALKDEDYAVRYSAAKALGKI GGERVRAAMEKLAETGTGFARKVAVNYLETHKSLIS	108	2	11,9	8,9
H7	<u>MRGSHHHHHH</u> TDPEKVE MYIKNLQDDSNVRYSAATALGKI RDERAVEPLIKALKDEDSGVRYAAAALGKI GDERAVEPLIKALKDEDSGVRRAAAQALGKI GDERAVEPLIKALKDESDENVRWSAALALGKI GDERAVEPLIKALKDESDYAVRRNAAKALGKI GDERAVEPLIKALKDEDSRFVRTAAAKALGKI GDERAVEPLIKALKDESDLVREDAARALGKI GDERAVEPLIKALKDESDVRRAAAALGKI GDERAVEPLIKALKDESDWLVRLSAARALGKI GGERVRAAMEKLAETGTGFARKVAVNYLETHKSLIS	325	9	35,6	6,7

3) Propriétés des VHH et α Rep sur l'activité de FluPol

Afin de caractériser les propriétés des VHH et des α Rep sur l'activité de FluPol, un criblage fonctionnel a tout d'abord été réalisé. La capacité des VHH et α Rep à inhiber ou bloquer l'activité polymérase de FluPol a été testée dans un système où FluPol, les VHH ou les α Rep sont intracellulairement co-exprimés. Pour se faire, un plasmide produisant un segment génomique NA modifié (pNA-LUC) dans lequel la séquence codant pour la neuraminidase virale est remplacée par le gène de la luciférase de la luciole a été utilisé. Ce plasmide co-transfecté avec des plasmides assurant l'expression de PA, PB1, PB2 et NP permet de quantifier une activité luciférase qui reflète globalement les activités répliqueuse et transcriptase de FluPol exprimée de façon transitoire.

La figure 40A présente les résultats obtenus avec les sept VHH spécifiques de FluPol. Le VHH16 bloque totalement l'activité de FluPol quel que soit la quantité de plasmide codant utilisé. L'inhibition de l'activité de FluPol par les VHH9 et 18 est dépendante de la quantité de plasmide codant utilisé. Les VHH7, 13, 19 et 22 n'ont pas d'effet inhibiteur dépendant de leur concentration, tout comme les VHH délétés partiellement de leur domaine CDR3 KO7 et KO18. La co-expression d'une protéine cellulaire (RBP7) ne montre aucun effet dépendant de la concentration de plasmide sur l'activité de FluPol.

Les résultats obtenus avec les quatre α Rep sont présentés figure 40B. L' α Rep E3 inhibe jusqu'à 50% l'activité de FluPol aux deux quantités de plasmide les plus élevées. L' α Rep H3 n'a aucune activité inhibitrice de FluPol. Étonnamment, les α Rep F9 et H7 augmentent l'activité de FluPol, en particulier, un facteur trois est obtenu pour H7 pour une quantité de plasmide codant de 50 ng. Nous notons que l' α Rep inhibiteur E3 est constitué de quatre motifs répétés (ce qui constitue une taille moyenne dans les cribles d' α Rep réalisés par l'équipe du Pr P. Minard) alors que les deux α Rep facilitateurs sont de grande taille (constitués de 9 et 10 motifs répétés), une observation qui peut peut-être marquer un effet stabilisateur de l'activité de FluPol. Les trois α Rep utilisées en contrôle et ne ciblant pas FluPol n'interfèrent pas ou peu avec l'activité de FluPol.

Pour résumer ces expériences, les VHH 9, 16 et 18 inhibent l'activité ARN-polymérase de FluPol, et ce de façon dépendante de la concentration de plasmide codant utilisé. Le VHH16 a l'activité inhibitrice de FluPol la plus marquée. Parmi les α Rep testés, seul l' α Rep E3 a une activité inhibitrice détectable, deux d'entre elles pouvant augmenter l'activité de FluPol. Pour la suite de notre étude, les VHH 9, 16, 18, 19 et 22 ainsi que l' α Rep E3 ont été sélectionnés.

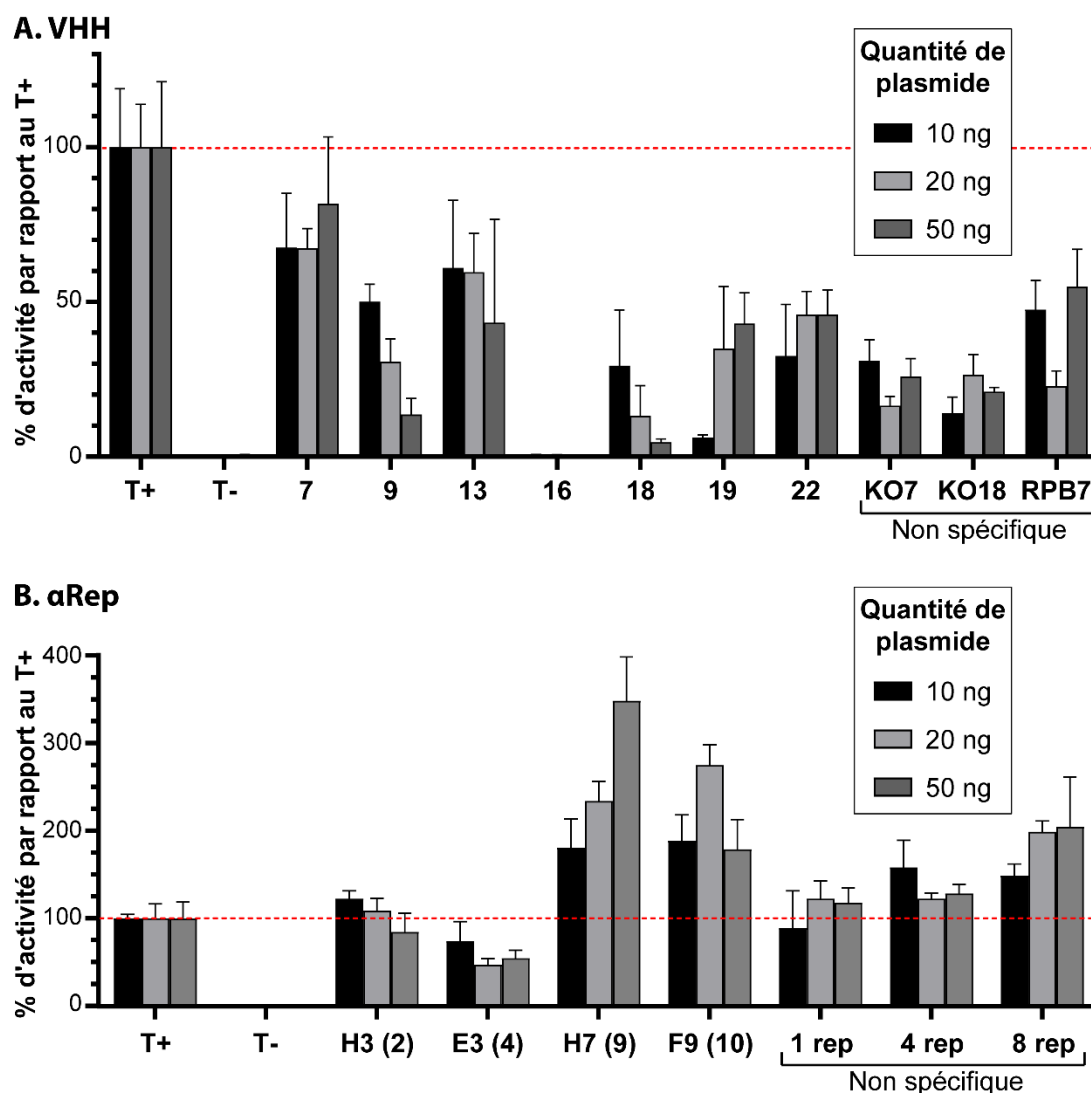


Figure 40 : Inhibition de l'activité de FluPol par les différents VHH et αRep obtenus

Des cellules 293T ont été transfectées avec des plasmides codant pour un système minigénome classique permettant de mesurer l'activité de FluPol ainsi qu'avec trois quantités différentes de chaque plasmide codant pour les VHH, des VHH témoins (KO7 et KO18), une sous-unité de Pol II (RPB7) (A) ou pour les αRep et trois αRep non spécifiques de FluPol de tailles similaires (B). Les cellules sont lysées 48h post-transfection puis la luminescence est révélée par l'ajout de substrat de la luciférase et mesure de la luminescence résultante. Le nombre de répétition du motif des αRep est indiqué entre parenthèses. Les résultats sont présentés sous forme de moyenne +/- écart-type.

Afin de déterminer si l'activité des VHH et de l'αRep E3 sur FluPol était associée à une inhibition de la transcription virale et non de la résultante « réplication + transcription », le mutant de FluPol dépourvu d'activité réplicase a été généré (Nilsson-Payant *et al*, 2018). Ce mutant de FluPol est constitué des sous-unités PB1, PB2 et du mutant de PA C95A. Ce mutant est décrit comme incapable de former des homodimères de FluPol et d'assurer la réplication virale. Ainsi, la cotransfection du plasmide pNA-LUC avec les plasmides codant pour PB1, PB2, PA C95A et NP permet de quantifier l'activité transcriptionnelle seule. La figure 41 montre que le VHH16 inhibe complètement l'activité

transcriptase et les activités réplicase + transcriptase de FluPol. Les VHH 18 et 9 bloquent plus efficacement l'activité transcriptase de FluPol (à respectivement 90% et 80% d'inhibition) que la résultante des activités réplicase + transcriptase (à 80 et 60% d'inhibition), ce qui suggère que ces deux VHH bloquent essentiellement l'activité transcriptionnelle, et non l'activité réplicase, de FluPol. Les VHH 19 et 22 ainsi que l' α Rep E3 n'ont pas d'effet inhibiteur marqué de l'activité transcriptase de FluPol.

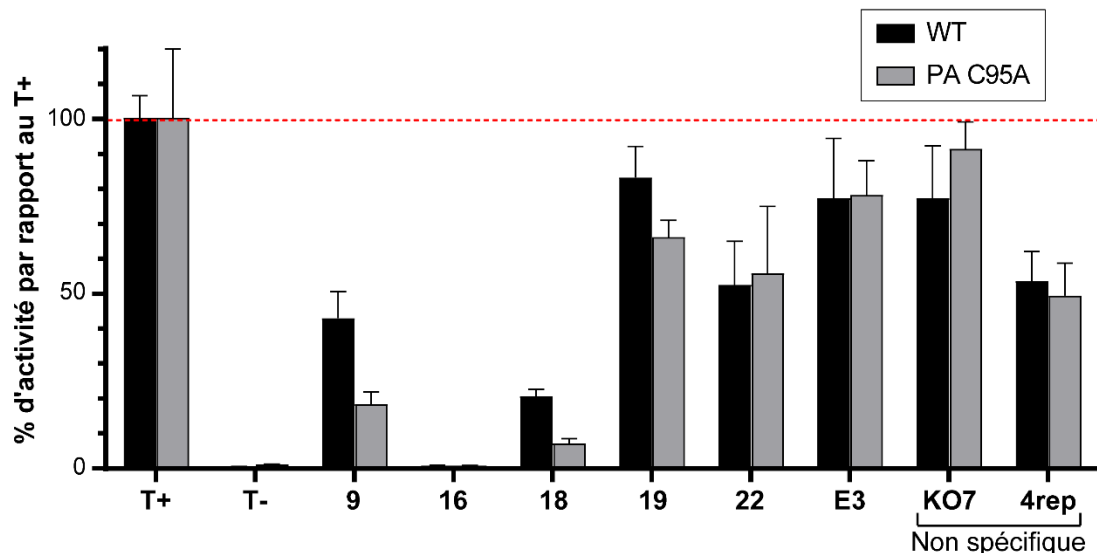


Figure 41 : Inhibition de l'activité transcriptase de FluPol par différents VHH et α Rep obtenus

Des cellules 293T ont été transfectées avec des plasmides codant pour un système minigénome classique natif (WT) ou portant la mutation PA C95A qui rend FluPol KO dans son activité réplicase. 20 ng de chaque plasmide codant pour les VHH, α Rep et des VHH et α Rep non spécifiques ont été ajoutés. Les cellules sont lysées 48h post-transfection puis la luminescence est révélée par l'ajout de substrat de la luciférase et mesure de la luminescence émise. Les résultats sont présentés sous forme de moyenne +/- écart-type.

Afin de mieux caractériser l'activité du VHH16 sur FluPol, une gamme de concentration du plasmide pci-VHH16 a été utilisée pour co-transfection avec les plasmides d'expression de PA, PB1, PB2, NP et pNA-LUC (Figure 42). Le plasmide pci-VHH KO7 et un plasmide d'expression sans insert ont été utilisés à des concentrations identiques à celles utilisées pour pci-VHH16. Dans cet essai, chacun des plasmides codant pour PA, PB1, PB2, NP ou portant le segment NA modifié sont transfectés à des quantités comprises entre 50 et 132 ng. Les résultats montrent que le VHH16 a une activité spécifique forte et dépendante de la concentration de plasmide sur les fonctions réplicase + transcriptase de FluPol. À 2,5 ng de plasmide, le VHH16 inhibe spécifiquement 75% de l'activité de FluPol.

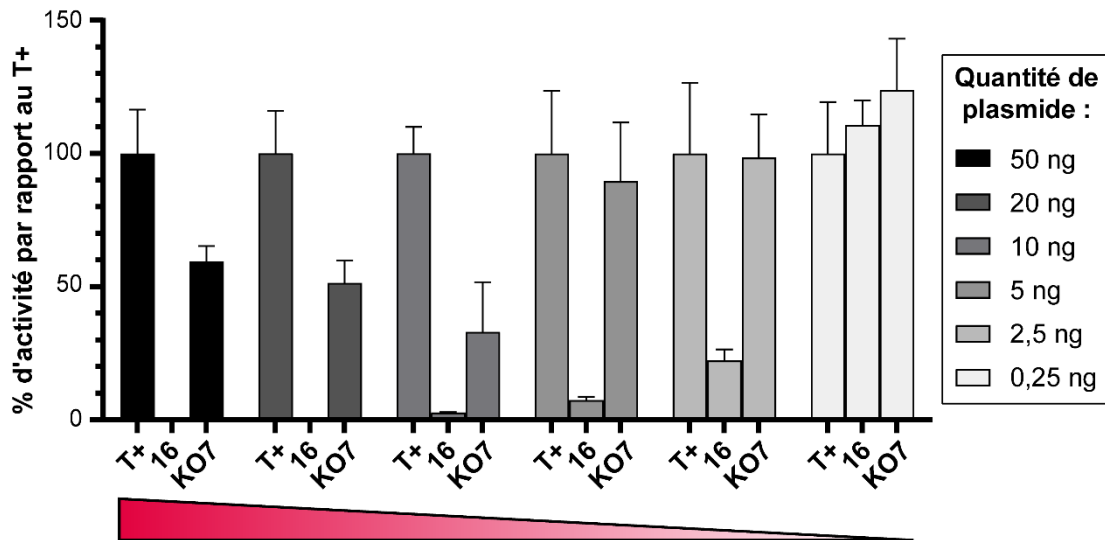


Figure 42 : Inhibition de l'activité de FluPol par le VHH16

Des cellules 293T ont été transfectées avec des plasmides codant pour un système minigénome ainsi que différentes quantités de chaque plasmide codant pour le VHH16, le VHH témoin VHH KO7 ou un plasmide vide. Les cellules sont lysées 48h post-transfection puis la luminescence est révélée par l'ajout de substrat de la luciférase et mesure de la luminescence résultante. Les résultats sont présentés sous forme de moyenne +/- écart-type.

4) Spécificité des VHH et α Rep pour les sous-unités de FluPol

Pour rappel, les VHH ont été sélectionnés contre une forme trimérique de FluPol (forme A) et les α Rep contre une forme dimérique (forme B). Afin d'identifier la sous-unité de FluPol reconnue par ces différents nanoligands (PA, PB1 ou PB2 dans le cas des VHH ; PA ou PB1 pour les α Rep), plusieurs stratégies pouvaient être envisagées. Dans un premier temps, un essai d'identification de la sous-unité reconnue par le VHH16 par Western blot a été réalisée. A cette fin, nous avons choisi de marquer le VHH avec un fluorophore pour identifier la sous-unité reconnue.

Marquage de l'amine terminale du VHH16 avec le fluorophore Alexa488

Cette procédure permet de greffer le composé Alexa488 - ester succinimidyle aux fonctions amines d'une protéine par une réaction de substitution nucléophile. Le pH auquel a lieu la réaction permet de pouvoir déterminer quelle fonction amine est ciblée (l'amine terminale du VHH, les amines des lysines et arginines ou toutes les fonctions amines de la protéine). En se plaçant à un pH de 7,2, l'amine terminale du VHH est spécifiquement ciblée. Un milligramme de VHH16 a été utilisé pour réaliser ce marquage.

Après marquage, l'échantillon a été déposé en quantité croissante sur un gel d'acrylamide pour électrophorèse SDS-Page. La figure 43A montre que le VHH a été marqué à l'Alexa488 pour utilisation en Western blot.

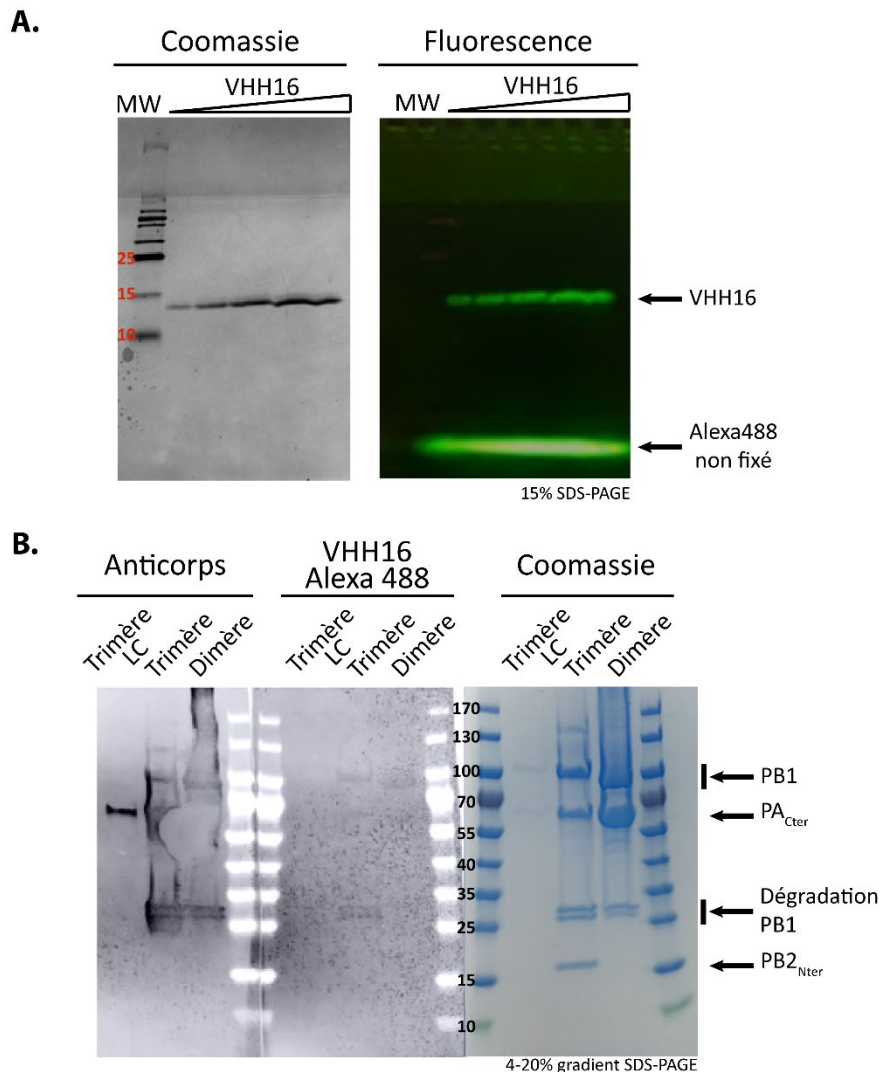


Figure 43 : Le VHH16 reconnaît le dimère PA-PB1

A. Vérification du marquage Alexa488 du VHH16 par gel SDS-Page et révélation de la fluorescence à la lumière bleue. **B.** Migration sur gel SDS-Page puis révélation au bleu de coomassie ou par Western Blot avec anticorps « commerciaux » anti-PA, PB1 et PB2 ou le VHH16-Alexa488
LC = low concentration

Analyse par western blot de la spécificité d'interaction du VHH16

Le Western blot a été réalisé avec les deux formes de FluPol purifiées (A et B). L'analyse en SDS-PAGE et coloration au bleu de coomassie permet d'identifier l'ensemble des sous-unités ou domaines de FluPol présents dans les différentes formes de FluPol (Figure 43B). Des bandes de dégradation attribuées à PB1 à environ 30 kDa sont également visibles. L'utilisation d'anticorps

polyclonaux anti-PA et -PB1 permet de révéler l'ensemble des bandes. La révélation par le VHH16 ne permet pas de caractériser formellement la sous-unité reconnue. Si PB1 et ses bandes de dégradation semblent reconnues par le VHH, une fine bande attribuée à PA est également révélée. Ce résultat semble suggérer que l'épitope reconnu par le VHH16 est sensible à la dénaturation par le SDS ou qu'il est dépendent de l'assemblage des trois sous-unités. Nous avons donc choisi d'analyser la spécificité d'interaction du VHH par une seconde approche expérimentale.

Analyse de la spécificité du VHH16 par test de complémentation

Le choix s'est porté sur l'utilisation d'un test de complémentation pour quantifier l'interaction entre le VHH et chacune des sous-unités de FluPol. Pour se faire, le cadre de lecture codant pour le VHH16 a été fusionné à celui codant pour le domaine C-terminal de la luciférase de *Gaussia princeps* (Luc2) alors que les sous-unités de FluPol PA, PB1 et PB2 ont été fusionnées à son domaine N-terminal Luc1. Si le VHH se fixe sur FluPol ou une de ses sous-unités, un signal luciférase doit être détecté.

La figure 44 montre que l'interaction entre le VHH16 avec les sous-unités de FluPol est très forte lorsque PA+PB1 ou PA+PB1+PB2 sont co-exprimées (NLR voisins de 600 et 400). Si PA ou PB1 sont exprimées seules avec le VHH16, aucune interaction n'est détectée avec PB1 et un NLR voisin de 25 est obtenu pour PA seule. Ce résultat suggère que le VHH16 reconnaît un site chevauchant partiellement PA et PB1. La co-expression de PB2 avec le dimère PA-PB1 ne bloque pas la fixation du VHH16. Notons également que la fusion de Luc1 aux sous-unités PB1 et PB2 ne permet pas de révéler l'interaction avec le VHH16, un résultat vraisemblablement dû à une longueur trop courte du linker utilisé entre ces sous-unités de FluPol et les domaines Luc1 et Luc2.

L'ensemble de ces données suggèrent que le VHH16 reconnaît un épitope chevauchant les deux chaînes polypeptidiques de PA et PB1.

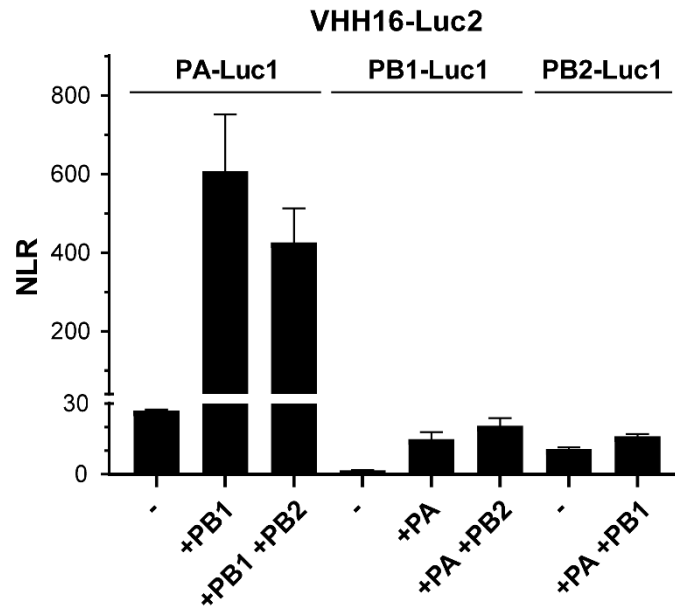


Figure 44 : Le VHH16 reconnaît spécifiquement FluPol

Des cellules 293T ont été transfectées avec 150 ng de chaque plasmide codant la Luc1/2 et de 75 ng de chaque plasmide codant la sous-unité PA, PB1 ou PB2. Les cellules sont lysées 24h post-transfection puis la luminescence est révélée par l'ajout de substrat de la luciférase et mesure de la luminescence émise. Les résultats sont représentés sous forme de moyenne +/- écart-type.

5) Mesure de l'affinité entre le VHH16 et FluPol

Afin de mieux caractériser l'interaction entre le VHH 16 et FluPol, son affinité a été mesurée par des expériences d'interférométrie de biocouche (BLI).

Pour se faire, le VHH 16 portant un Strep-tag II a été exprimé dans le périplasma de bactéries, puis purifié en plusieurs étapes (Figure 45). Le VHH a ainsi d'abord été purifié par chromatographie d'affinité sur résine Streptactine (Figure 45A). L'expression dans le compartiment périplasmique ainsi que l'utilisation d'une résine Streptactine permet l'obtention d'un échantillon avec très peu de contaminants. Les fractions 27 à 30 ont ainsi été rassemblées puis purifiées par chromatographie d'exclusion de taille afin d'éliminer toute trace de contaminants (Figure 45B et C). Le chromatogramme présente un pic bien défini et homogène. Les fractions 15 à 18 ont été rassemblées suite à l'analyse SDS-Page. Concernant FluPol, ce sont les formes A et B précédemment décrites qui ont été utilisées pour mesurer l'affinité.

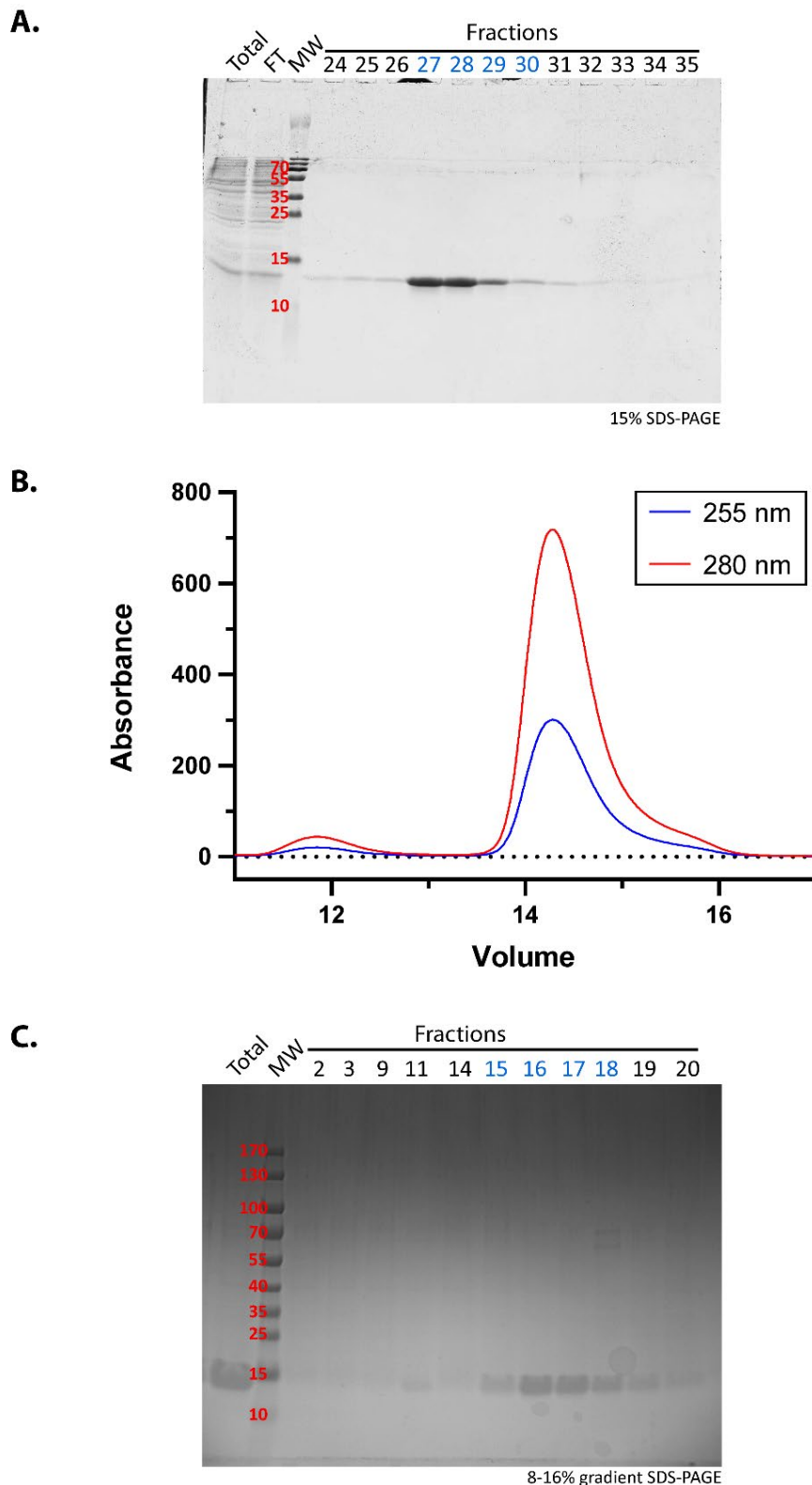


Figure 45 : Purification du VHH16-Strep

Pour une première étape de purification, le VHH16-Strep est purifié par chromatographie d'affinité (Streptactine). Les différentes fractions sont observées sur gel SDS-Page (**A**). Les fractions d'intérêt sont rassemblées pour purification par chromatographie d'exclusion stérique dont le chromatogramme est présenté en (**B**). Enfin, ces fractions sont analysées sur gel SDS-Page (**C**). Sng = surnagent, FT = Flow Through, Wash = lavage, MW = marqueur de poids moléculaires. Les fractions en bleu correspondent aux fractions rassemblées.

Les résultats de mesures d'interactions réalisées sur la plateforme SPR/BLI de l'ISBG (URA3518; Grenoble) sont présentés dans la figure 46. Dans une première étape, l'association du VHH16 avec le dimère de FluPol a été mesurée pendant cinq minutes. Il apparaît que la liaison du VHH à FluPol se fait de façon lente. En effet, on n'observe pas de phase exponentielle dans l'association et aucun plateau n'est atteint malgré les fortes concentrations testées (jusqu'à 400 nM de FluPol, résultats non montrés). Dans un second temps, la dissociation du VHH et de FluPol a été mesurée. Aucune dissociation n'est observée. Le VHH16 reste donc fixé à FluPol. Un profil similaire a été obtenu avec la forme trimérique de FluPol (résultats non montrés).

Suite à ces résultats, la constante de dissociation à l'équilibre K_d a été calculée. Pour les deux formes de FluPol testées, le K_d est égal à des concentrations de l'ordre du picomolaire, en dessous de la limite de détection de l'appareil utilisé. En conclusion, le VHH16 a un K_d estimé à des concentrations sous-nanomolaires pour les deux formes de FluPol testées.

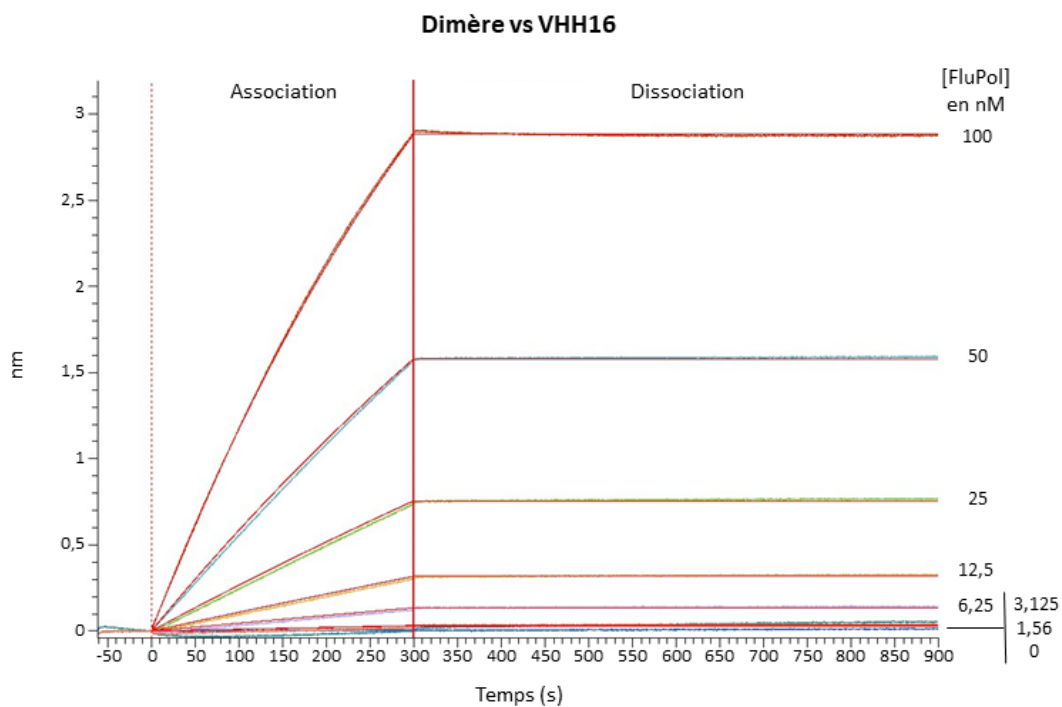


Figure 46 : Mesure d'affinité entre le VHH16 et FluPol par interférométrie de biocouche

Le graphique présente les résultats obtenus en BLI pour la mesure de l'affinité entre le VHH16-Strep et le dimère PA-PB1. Le VHH16-Strep a été lié sur les biocapteurs à une concentration de $1,5 \mu\text{g}\cdot\text{mL}^{-1}$ qui ont ensuite été mis en contact avec des concentrations croissantes de FluPol allant de 0 à 100 nM.

6) Effet du VHH16 sur l'assemblage et le transport du dimère PA-PB1 au noyau

Après avoir montré que le VHH16 reconnaît spécifiquement le dimère PA-PB1 avec une forte affinité et inhibe très fortement l'activité polymérase de FluPol, une question se pose : comment le VHH16 inhibe-t-il si fortement FluPol ?

Effet du VHH16 sur l'assemblage de FluPol

Le VHH16 reconnaissant le dimère PA-PB1, nous nous sommes naturellement demandé s'il pouvait gêner l'assemblage intracellulaire du complexe PA-PB1 et l'association de PB2 au dimère.

Pour répondre à cette question, l'influence du VHH16 et de deux autres nanoligands (le VHH7 et l' α Rep F9) sur l'interaction entre PA et PB1 a été testée en système split-luciférase. Ainsi, les sous-unités PA et PB1 respectivement fusionnées aux domaines Luc1 et Luc2 ont été exprimées en cellules HEK 293T en présence ou non des nanoligands. La figure 47A montre que le VHH16 diminue de 40% mais n'abolit pas l'interaction entre PA et PB1 alors que le VHH7 et l' α Rep F9 n'ont aucun effet.

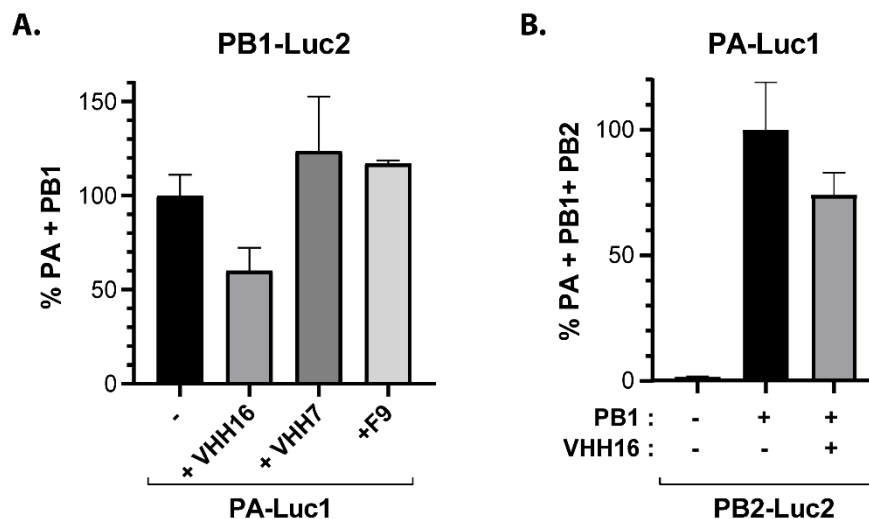


Figure 47 : Le VHH16 n'empêche pas l'assemblage de FluPol

Des cellules 293T ont été transfectées avec 150 ng de plasmide PA/PB1-Luc1/2 et 150 ng de plasmide vide ou codant pour un VHH ou l' α Rep F9 (A) ou 150 ng de plasmide PA/PB2-Luc1/2 et 75 ng de plasmide codant pour PB1 et/ le VHH16 (B). Les cellules sont lysées 24h post-transfection puis la luminescence est révélée par l'ajout de substrat de la luciférase et mesure du signal émis. Les résultats sont présentés sous forme de moyenne +/- écart-type.

Dans un second temps, l'impact du VHH16 sur la fixation de PB2 au dimère PA-PB1 a été également testé en système split-luciférase. Ainsi, les sous-unités PA et PB2 respectivement fusionnées aux domaines Luc1 et Luc2 ont été exprimées en cellules HEK 293T en présence ou non de PB1 et du VHH16. La figure 47B montre que le VHH16 diminue d'environ 30% mais n'abolit pas l'interaction entre le dimère et PB2. Ce résultat suggère donc que l'activité du VHH16 sur FluPol n'est pas associée à un blocage de l'association entre PA et PB1.

Globalement, les résultats de ces deux expériences montrent que le VHH16 n'abolit pas l'assemblage du dimère ou de PB2 à PA-PB1.

Activité du VHH16 sur le transport du dimère PA-PB1 au noyau et son interaction avec l'importine RanBP5

Il avait été précédemment établi que les sous-unités de FluPol néosynthétisées au cours du cycle viral s'associent entre elles en deux étapes pour assurer transcription et réplication virale dans le noyau (Deng *et al*, 2005; Fodor & Smith, 2004). PA et PB1 s'associent dans le cytoplasme pour être transportées au noyau par l'importine- β RanBP5 (Deng *et al*, 2006a), alors que PB2 s'associe seule à une importine- α pour migrer au noyau (Resa-Infante *et al*, 2008). PB2 s'associe ensuite au dimère PA+PB1 pour constituer une FluPol fonctionnelle. Dans le but de déterminer si le VHH16 entrave le transport du dimère PA-PB1 au noyau, deux expériences ont été réalisées.

La protéine PB1 fusionnée à la GFP a été utilisée comme rapporteur pour localiser le dimère PA+PB1 dans la cellule et visualiser son transport au noyau (Figure 48). En premier lieu, la localisation cellulaire du VHH16 exprimé seul a été étudiée. Le VHH16 se distribue dans toute la cellule, y compris au noyau. Exprimée seule ou en présence du VHH16, PB1 reste localisée au cytoplasme comme attendu (Fodor & Smith, 2004). Lorsque PA est co-exprimée avec PB1, le dimère PA-PB1 est révélé au noyau. Lorsque le VHH16 est co-exprimé avec PB1 et PA, le dimère PA-PB1 est principalement révélé au noyau avec une certaine proportion encore présente dans le cytoplasme. Le VHH16 ne semble donc pas bloquer le transport du dimère PA-PB1 au noyau.

Afin de confirmer que les nanoligands spécifiques de FluPol ne bloquent pas l'interaction entre l'importine RanBP5 et le dimère PA-PB1, PA et l'importine RanBP5 ont été co-exprimés sous forme de protéines de fusion PA-Luc1 et RanBP5-Luc2 avec PB1 et les différents nanoligands de FluPol. (Figure 49). Les résultats montrent que le VHH13 et l' α Rep H7 diminuent d'environ 40% l'interaction du dimère avec RanBP5. En revanche, le VHH16 n'empêche pas l'interaction du dimère PA-PB1 avec RanBP5.

En conclusion, la forte action inhibitrice du VHH16 ne semble pas due à une inhibition de l'assemblage du complexe polymérase ou à un blocage de son acheminement au noyau. Ces résultats suggèrent que les changements conformationnels profonds de FluPol lors de son fonctionnement comme transcriptase et réplacase sont entravés lorsque le VHH16 est fixée à FluPol dans le noyau. Une alternative à cette proposition serait la capacité du VHH16 à bloquer une interaction essentielle de FluPol avec un partenaire cellulaire.

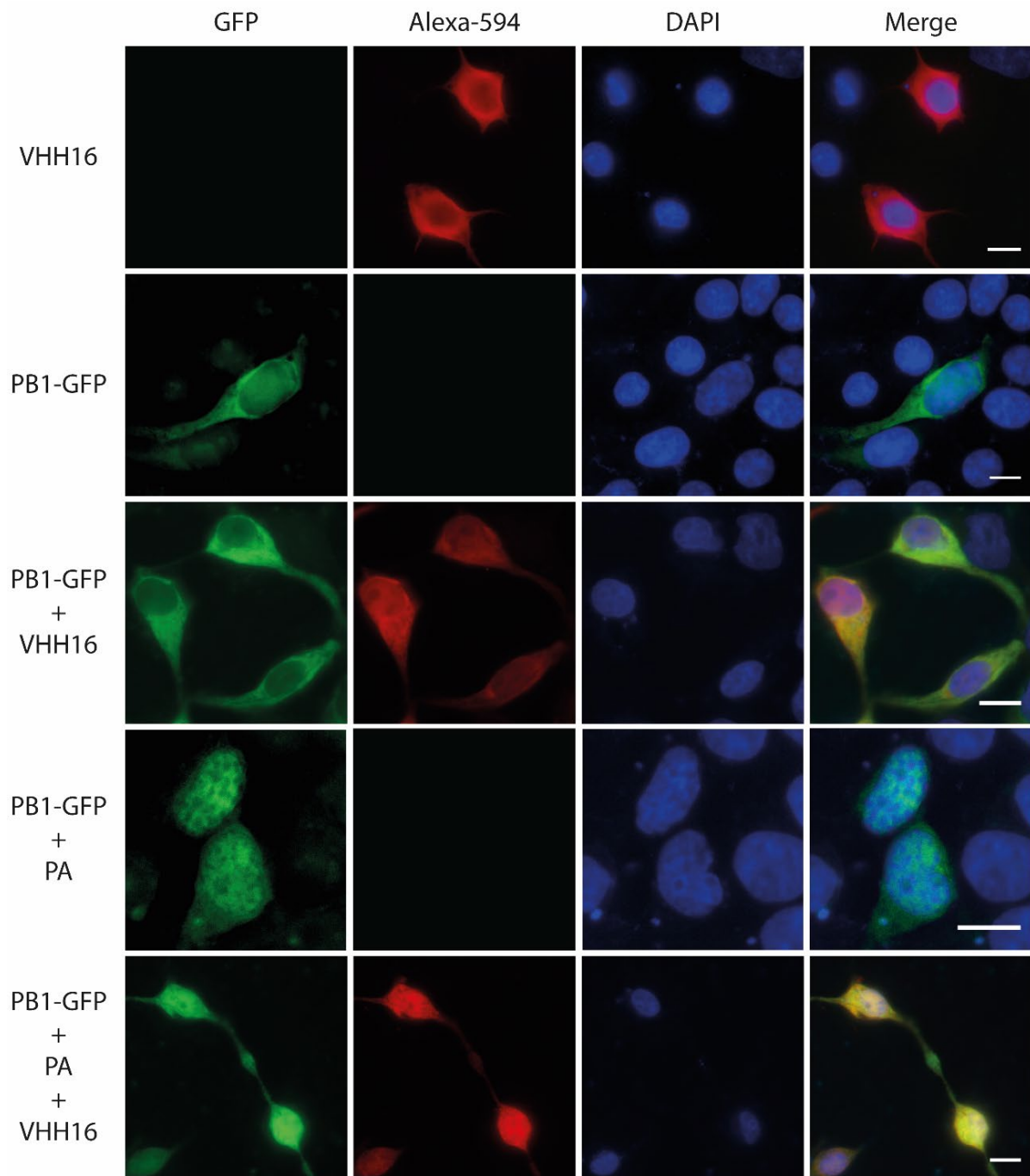


Figure 48 : Le VHH16 n'empêche pas le transport du dimère PA-PB1 au noyau

Des cellules 293T ont été transfectées avec 0,5 µg d'un plasmide codant pour la sous-unité PB1 fusionnée à la GFP, la sous-unité PA et le VHH16. 28h post-transfection, les cellules sont fixées, le VHH16 est révélé avec un anticorps couplé à l'Alexa-594 et les noyaux sont colorés au DAPI. Les photos ont été réalisées avec un microscope à épifluorescence ZEISS à l'aide d'un objectif à immersion x63. L'échelle représente 10 µm.

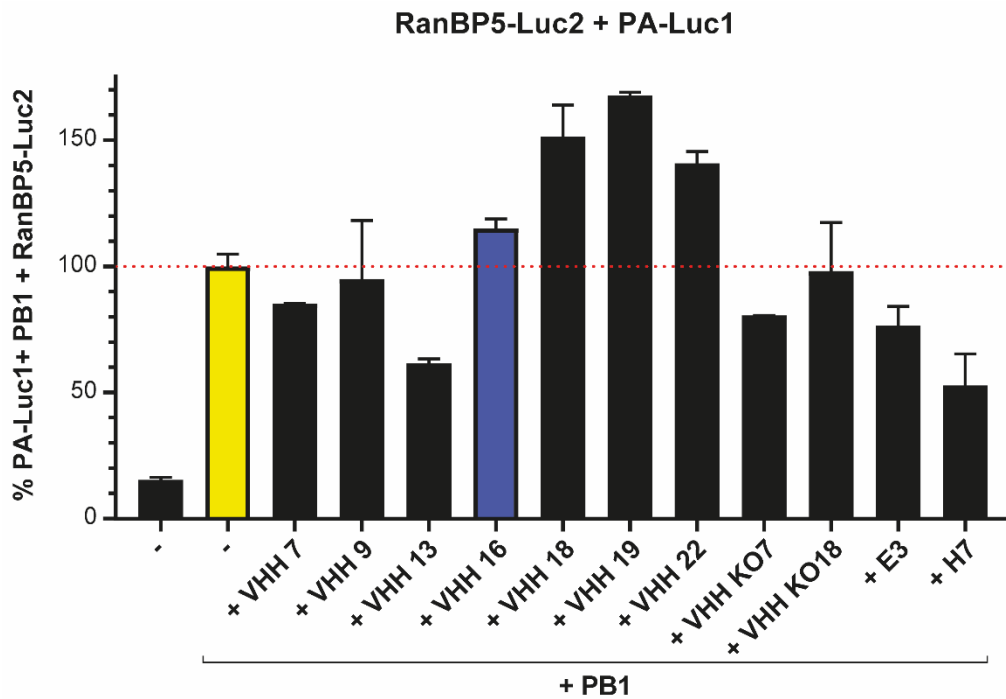


Figure 49 : Le VHH16 ne bloque pas l'interaction de PA-PB1 avec RanBP5

Des cellules 293T ont été transfectées avec 150 ng de chaque plasmide codant la Luc1/2 et de 75 ng de chaque plasmide codant PB1, un VHH ou une α Rep. Les cellules sont lysées 24h post-transfection puis la luminescence est révélée par l'ajout de substrat de la luciférase et mesure de la luminescence résultante. Les résultats sont représentés sous forme de moyenne +/- écart-type en pourcentage du signal sans ajout de VHH ou α Rep (en jaune).

7) Caractérisation structurale des VHH

En préalable à une étude structurale des complexes FluPol ou de ses sous-unités avec des VHH ou des α Rep d'intérêt, des essais de détermination de la structure des VHH ont été menés par cristallographie et diffraction des rayons X.

Purification des VHH

Les séquences codantes pour chacun des VHH spécifiques de FluPol ont été placées en aval d'une séquence codant pour un peptide signal pour assurer leur translocation dans le domaine périplasmique de la bactérie et en amont d'une séquence codant pour un His-tag. Après extraction périplasmique, les VHH ont été purifiés par chromatographie d'affinité sur résine nickel suivie d'une chromatographie d'exclusion de taille. A titre d'exemple, la figure 50A présente la purification par chromatographie d'affinité sur colonne de nickel du VHH18. Dans un second temps, les fractions contenant le VHH sont rassemblées pour purification par chromatographie d'exclusion de taille afin d'éliminer de potentiels contaminants. La figure 50B et C montre que le VHH18 est purifié à homogénéité. Des résultats similaires ont été obtenus pour les autres VHH. Les VHH ont été concentrés afin de réaliser des tests de cristallisation.

Cristallisation des VHH

Dans le but de trouver les conditions idéales de cristallisation, 576 conditions différentes ont été testées pour les VHH 7 (10 mg.mL⁻¹), 16 (7 mg.mL⁻¹), 18 (14 mg.mL⁻¹), 19 (30 mg.mL⁻¹), et 22 (10 et 27,5 mg.mL⁻¹) par la plateforme HTX (EMBL Grenoble). Seul le VHH18 a donné des résultats positifs. Pour ce dernier, les premiers cristaux sont apparus au bout de deux jours (Figure 51 gauche) pour la condition 0,1M Hepes pH7,5, 2M Ammonium sulfate, 2% PEG 400. Ces cristaux ont ensuite été reproduits à la main en 0,1M Hepes pH7,5, 1,8M Ammonium sulfate, 2% PEG 400 (Figure 51 droite).

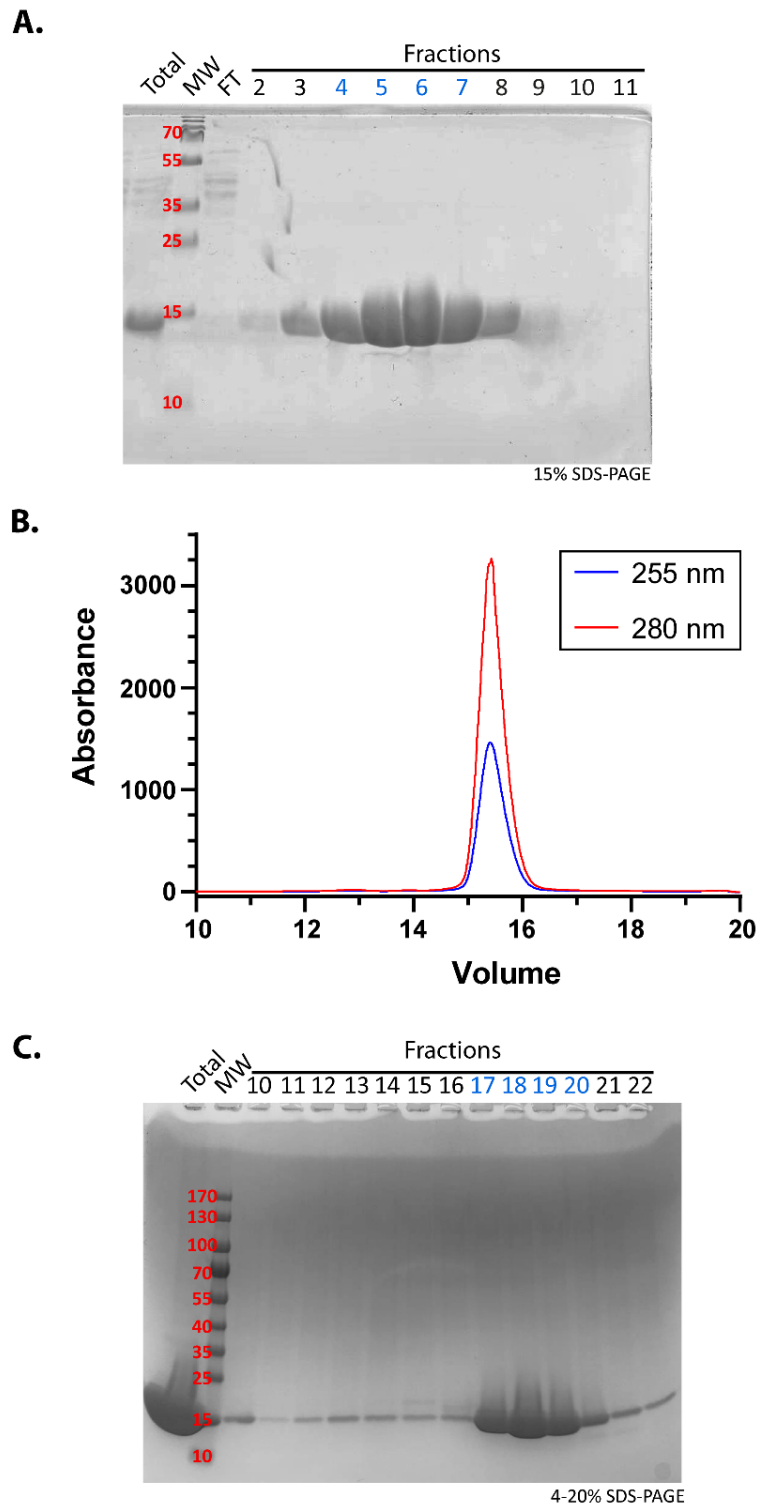


Figure 50 : Purification du VHH18

Dans un premier temps, la fraction périplasmique du lysat bactérien est purifiée par chromatographie d'affinité (résine nickel). Les différentes fractions sont observées sur gel SDS-Page **(A)** puis rassemblées pour purification par chromatographie d'exclusion de taille **(B)**. Les fractions sont analysées sur gel SDS-Page **(C)** puis les fractions adéquates sont rassemblées pour l'obtention de cristaux.

Sng = surnagent, FT = Flow Through, Wash = lavage, MW = marqueur de poids moléculaires.

Les fractions en bleu correspondent aux fractions rassemblées.

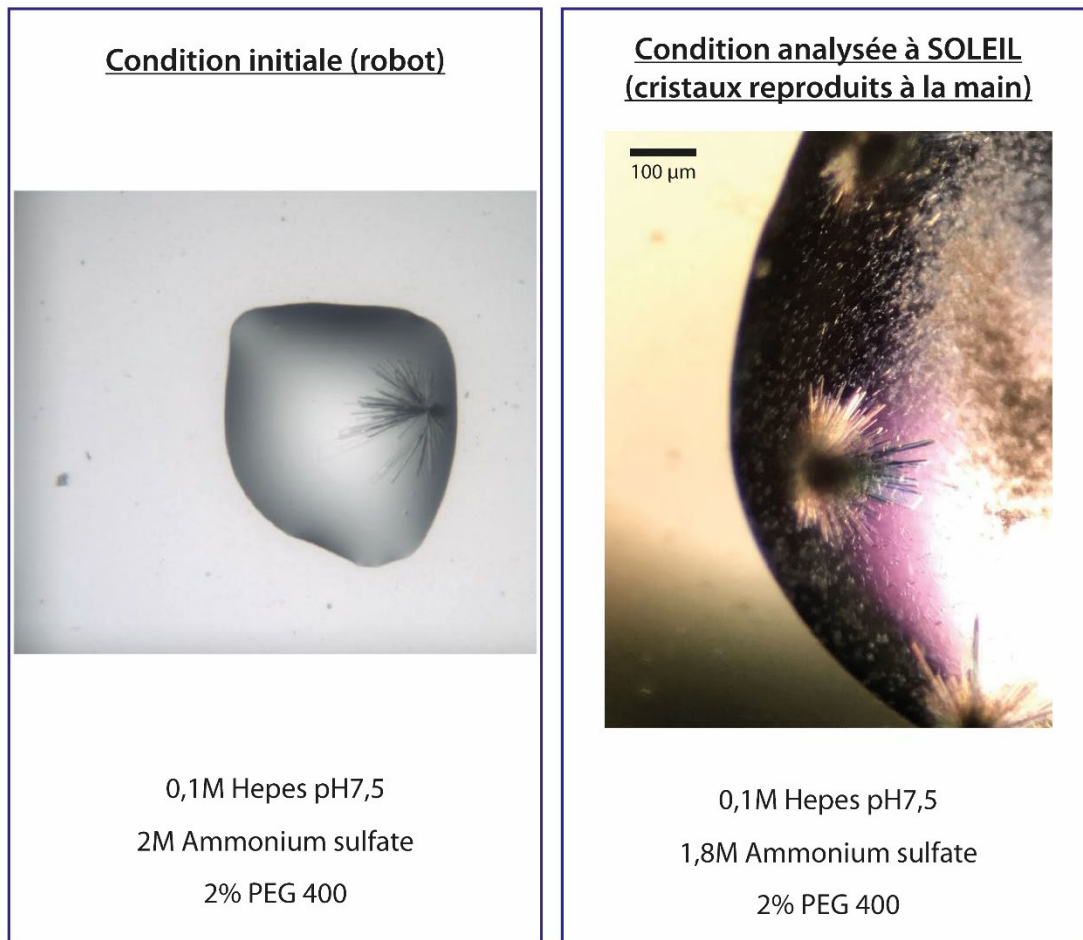


Figure 51 : Obtention de cristaux du VHH18 pour analyse structurale

Goutte de cristallisation ayant permis l'obtention de cristaux à 20 °C. **Gauche** : Photo d'une goutte obtenue à partir de la plateforme HTX (EMBL Grenoble). **Droite** : Goutte reproduite à la main ayant permis l'obtention de cristaux.

Structure du VHH18

Les cristaux obtenus manuellement ont ensuite été analysés par diffraction des rayons X au synchrotron SOLEIL. Les données de diffraction sont reportées dans le tableau 4. Le traitement des données de diffraction a permis d'obtenir un jeu de données à 1,65 Å de résolution. La structure a été résolue par remplacement moléculaire. La structure du VHH18 est présentée dans la figure 52A avec les 3 régions variables (CDR) représentées en rouge ainsi qu'un détail de la carte de densité électronique reflétant la haute résolution de la structure. La structure obtenue est fortement similaire aux autres structures connues de VHH. La principale différence réside dans le CDR3. L'alignement des séquences des VHH générés dans ce travail montre que le CDR3 est la partie la plus variable, que ce soit en taille ou en séquence (Figure 52B).

Tableau 4 : Données de diffraction et statistiques de traitement des données pour les cristaux du VHH18.

Les valeurs entre parenthèses correspondent aux données à haute résolution.

VHH18	
Données de diffractions	
source X-ray	SOLEIL (PROXIMA-1)
Longueur d'onde (Å)	0.97918
Groupe d'espace	<i>C2</i>
<i>a, b, c</i> (Å)	37.23, 101.28, 64.65
<i>α, β, γ</i> (°)	90, 105.84, 90
Résolution (Å)	36 - 1.65 (1.75 - 1.65)
R _{fact} ^(a) (%)	6.4 (79.3)
I/σI	12.2 (1.3)
Complétude (%)	97.9 (90.0)
Réflexions totales	129 281 (13 623)
Réflexions uniques	27 247 (4 031)
Multiplicité	4.74 (3.37)
CC(1/2)	0.998 (0.57)
Affinement	
Nbres de réflexions	25 929
R _{crist} / R _{libre}	18.3 / 23.3
Nbres d'atomes	2 078
Déviations standarts	
Longueur de la liaison (Å)	0.009
Angle de la liaison (°)	1.54

$$^{(a)} R_{\text{fact}} = \frac{\sum_{hkl} \sum_j |I_{hkl,j} - \langle I_{hkl} \rangle|}{\sum_{hkl} \sum_j I_{hkl,j}}$$

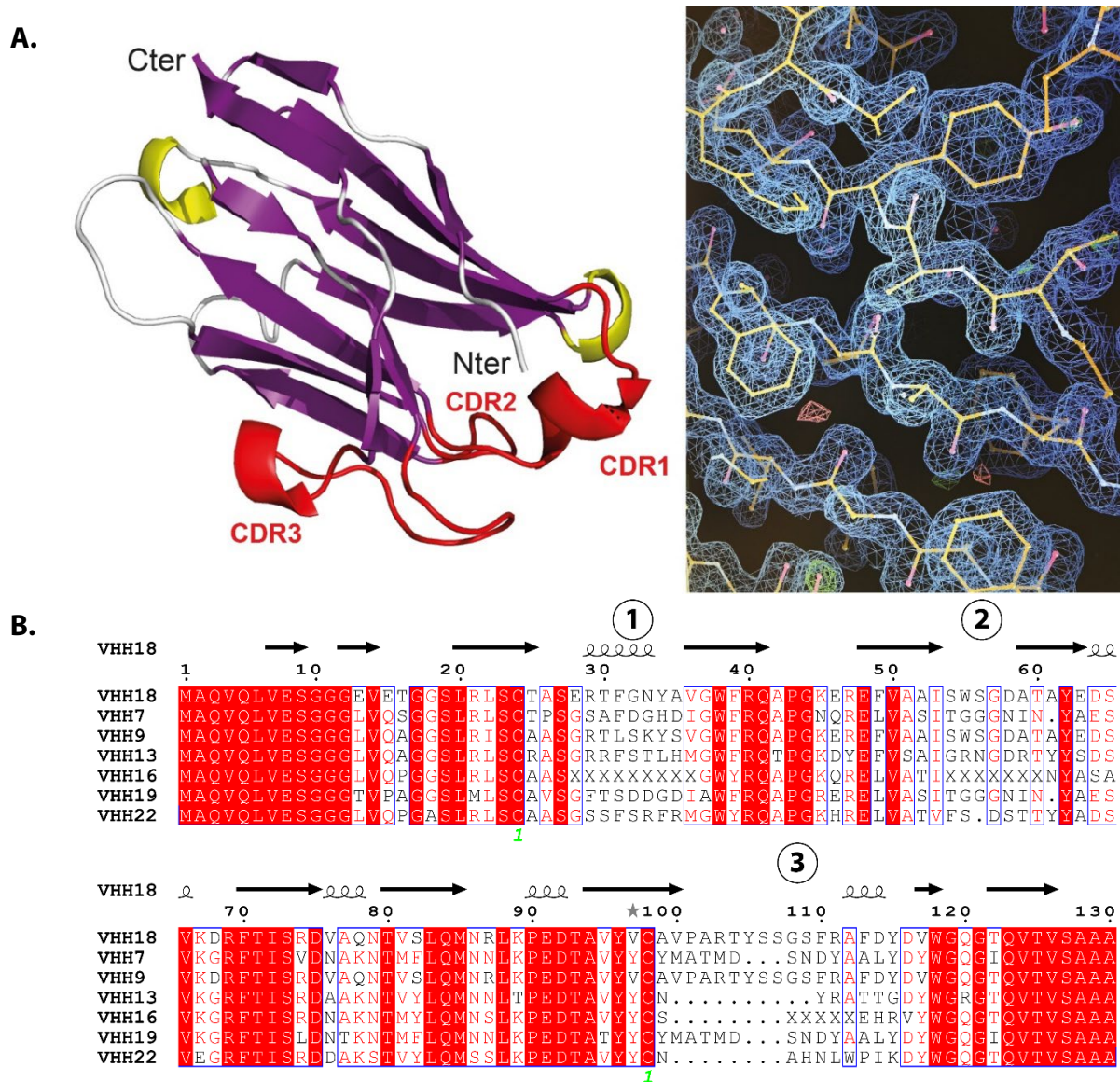


Figure 52 : Structure du VHH18

A. Structure du VHH18 sur laquelle les trois CDR sont représentés en rouge, les hélices α en jaune, les brins β en violet et détails de la carte de densité électronique. **B.** Alignement de la séquence des différents VHH de cette étude sur la séquence et la structure du VHH18. Les homologies de séquences sont surlignées en rouge. Les hélices représentent les hélices α , les flèches noires les brins β et les « 1 » verts les ponts disulfures. Les numéros des différents CDR sont indiqués dans les cercles. Figure réalisée avec ESPript.

Conclusions

Les VHH et α Rep s'avèrent être des outils pertinents, originaux et uniques pour étudier le fonctionnement de FluPol. Au cours de ce projet, sept VHH et quatre α Rep dirigés contre FluPol ont été générés. Le VHH16 est le nanoligand qui s'est particulièrement détaché du lot car il bloque l'activité de FluPol dans un test cellulaire de réplication et de transcription virale. Aux concentrations testées, plusieurs autres nanoligands inhibent partiellement l'activité de FluPol. Cette inhibition est dose-dépendante. En outre, ce VHH a une très forte affinité pour FluPol, de l'ordre du sub-nanomolaire. Le site de fixation du VHH16 sur FluPol n'a pas pu être clairement identifiée sur le dimère PA-PB1 et PB2 ne contribue pas à son interaction avec FluPol. Les sous-unités PA et PB1 semblent contribuer toutes deux au site de liaison du VHH16. Le mode d'action du VHH16 n'a pas encore été élucidé. Cependant, il ne semble pas bloquer la formation du dimère PA-PB1 ni son transport au noyau et/ou son interaction avec l'importine RanBP5. Des travaux additionnels sont donc nécessaires pour élucider le mode d'action du VHH16.

La structure du VHH18 a été résolue à 1,65 Å. Les essais de cristallisation du VHH16 ont échoués. Il serait intéressant d'obtenir la structure du complexe FluPol-VHH16 afin de pouvoir déterminer précisément la zone de FluPol ciblée. Cela permettrait également d'orienter les recherches pour comprendre le mode d'action de ce VHH sur FluPol et si cette inhibition est directe (i.e. bloque une fonction de FluPol) ou empêche une interaction essentielle au fonctionnement de FluPol.

Pour conclure ces travaux, les séquences des VHH 16 et 18 ont été comparées au VHH Nb8205 spécifiques de FluPol produit par le groupe du Professeur Ervin Fodor (Fan *et al*, 2019b). Ce nanobody reconnaît le domaine C-terminal de PA ainsi que le sous-domaine formant le pouce de PB1. La fixation de Nb8205 à FluPol empêche sa dimérisation. Les alignements de séquence sont présentés dans la figure 53. Ces deux VHH sont différents de Nb8205. Le VHH16 est plus court, ce qui se traduit par une très grande différence de séquence au niveau du CDR3, en taille mais également en séquence avec seulement un résidu glycine en commun. Les deux autres CDR restent assez différents (environ 50% d'identité/fonction conservée, résultat non visible sur la figure pour confidentialité de la séquence). Le VHH16 possède donc une séquence primaire très différente du Nb8205. Le VHH18 a une taille plus proche du Nb8205 que le VHH16 (deux acides aminés de plus que le Nb8205). Plus spécifiquement, son CDR3 est de taille similaire bien que très différent en termes d'identité de séquence. Les CDR 1 et 2 sont également assez similaires à ceux du Nb8205. Le VHH18 est donc également très différent du Nb8205. Les VHH 16 et 18 sélectionnés au cours de ce travail sont donc uniques et très différents du seul VHH anti-FluPol (Nb8205) décrit à ce jour. On peut donc estimer que nos VHH peuvent reconnaître une zone différente de celle reconnue par le Nb8205 et ont donc un mécanisme d'action bien différent.

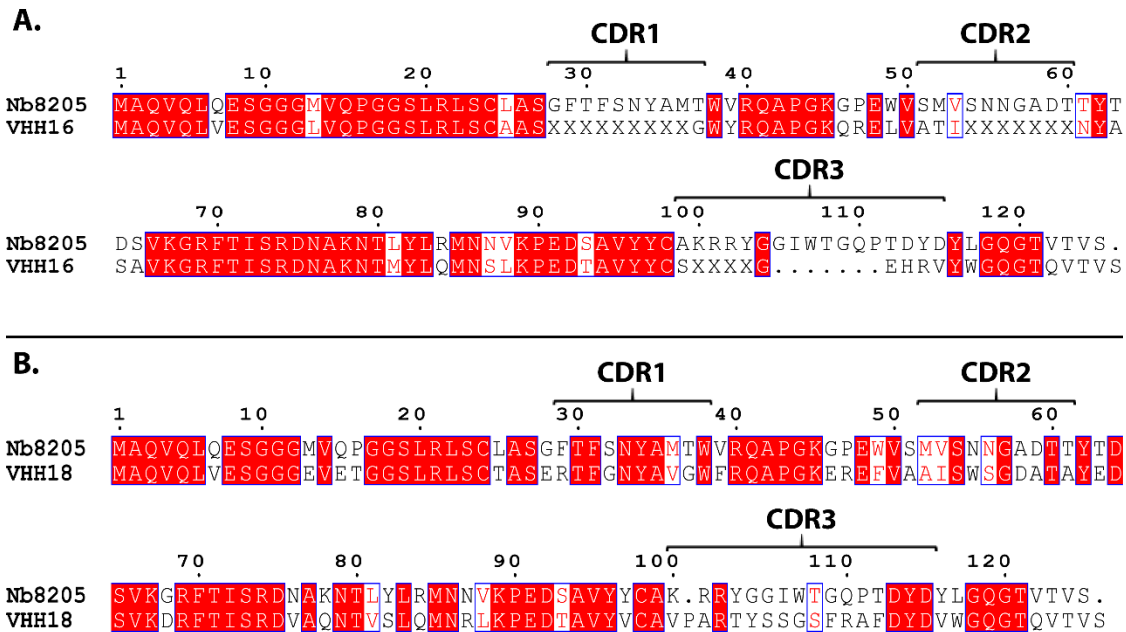


Figure 53 : Comparaison des VHH16 et 18 au Nb8205 généré par le groupe de Fodor

A. Alignement de séquence du Nb8205 avec le VHH16. **B.** Alignement de séquence du Nb8205 avec le VHH18. Figure réalisée avec ESPript.

Très récemment, le groupe du Pr. Ervin Fodor vient de publier de nouveaux travaux sur des VHH ciblant une FluPol de IAV (Keown *et al*, 2022). Dans cette dernière étude, une vingtaine de VHH générés contre FluPol ont été caractérisés structuralement afin d'identifier les zones reconnues sur FluPol. Une grande partie des VHH obtenus dans cette étude cible PA. Une autre partie importante cible les différents sous-domaines du domaine C-terminal de PB2. Seul un VHH cible le domaine N-terminal de PB2 en ciblant également PB1. Aucun de ces VHH ne cible uniquement PB1, au cœur du complexe polymérase. Après analyse structurale, le potentiel d'inhibition de ces nanobodies a été étudié par différentes techniques *in vitro* puis *in cellulo* en système minigénome ou en contexte infectieux. Le VHH ayant le plus grand pouvoir d'inhibition est le Nb8207 qui cible PB1 et le domaine N-terminal de PB2 car il bloque le canal de sortie des ARN de FluPol.

Une analyse des séquences a été réalisée pour comparer nos VHH à ceux décrits dans ce papier. Après alignement des séquences des PDB, aucun de nos VHH n'est identique aux VHH générés par l'équipe anglaise. La plus grande identité de séquence est de 76% entre le VHH16 et le Nb8189 qui reconnaît l'interface PA(615-630) – PB1(10-15). Ce nanobody inhibe la transcription primaire en empêchant une conformation correcte du domaine 627 de PB2 via son interaction avec PA. Bien que les CDR 1 et 2 présentent des similarités, le CDR3 est cependant très différent entre ces deux VHH. Il

se pourrait donc que le VHH16 reconnaisse une zone différente et ait également un mode d'action différent. Le mode d'action de nos VHH reste donc encore à être élucidé.

CONCLUSION GÉNÉRALE
&
PERSPECTIVES

Ce travail de thèse visait à mieux comprendre le fonctionnement de FluPol dans le noyau, notamment à mieux caractériser le comportement de FluPol lors du vol de coiffe et sa façon de cibler la chromatine. Trois aspects ont été abordés : premièrement, une étude pour identifier et caractériser d'autres interactions potentielles pouvant exister entre les complexes FluPol et Pol II, une seconde pour étudier les interactions de FluPol à la chromatine et enfin, dans une dernière partie, pour développer de nouveaux outils pour mieux étudier le fonctionnement de FluPol.

Interactions entre FluPol et Pol II

L'interaction de FluPol avec Pol II est nécessaire à la transcription du génome viral (Lamb & Choppin, 1977; Engelhardt & Fodor, 2006). Une première interaction avait été identifiée entre le domaine C-terminal de PA de FluPol et le CTD non structuré de Pol II. Il a été montré que cette interaction est régulée par la présence d'une phosphorylation au niveau de la sérine 5 (Ser5P) du CTD, connue pour être un marqueur de l'initiation de la transcription (Lukarska *et al*, 2017; Engelhardt *et al*, 2005). Nous avons fait l'hypothèse que l'interaction de FluPol avec le CTD apporte un ciblage sur Pol II initiant la transcription, mais qu'une ou plusieurs interactions additionnelles sont nécessaires pour assurer un positionnement adéquat de FluPol pour se lier à la coiffe des ARNm en début de synthèse.

La méthode choisie pour identifier des interactions additionnelles entre FluPol et Pol II consiste à réaliser des tests de complémentation binaire entre chacune des sous-unités de FluPol avec chacune des douze sous-unités de Pol II *in cellulo*. Chacune des protéines ou domaines analysés dans le test est préalablement fusionné avec la moitié N- ou C-terminale de la luciférase de *Gaussia princeps*, l'interaction entre le couple de sous-unités étant révélée par émission de photons. Après avoir validé cette approche entre sous-unités de FluPol et entre sous-unité de Pol II, une nouvelle interaction a été identifiée entre la sous-unité RPB4 de Pol II et le tiers N-terminal de PB2 d'Influenza A. Cette interaction a été validée par des expériences de co-immunoprécipitation ainsi que par un test de complémentation binaire split-Vénus. De plus, cette interaction a été retrouvée entre RPB4 et ce tiers N-terminal de PB2 des virus Influenza B et C. La nouvelle interaction identifiée permet à FluPol de se fixer à proximité du canal de sortie des ARNm néosynthétisés par Pol II.

Un système de reconstitution de l'activité ARN-polymérase de FluPol *in cellulo* a été construit. La co-expression des sous-unités PA, PB1 et PB2 et de NP en présence d'un segment génomique dans lequel un cadre de lecture codant pour une protéine virale a été remplacé par un gène codant pour la *Firefly* luciférase permet de quantifier l'activité enzymatique de FluPol. Une série de mutations générées dans le tiers N-terminal de PB2 a permis d'établir l'importance relative de ce domaine dans

l'activité transcriptase de FluPol. Ainsi, douze mutants sur les quatorze construits dans ce domaine de PB2 ont perdu toute activité transcriptase.

En prenant appui sur les structures atomiques de Pol II (de son dimère RPB4 / RPB7 situé en position distale de son cœur catalytique) et de FluPol, des modèles *in silico* de l'interaction FluPol / Pol II ont été générés à l'aide des serveurs GramMX et Rosetta. Dans une démarche de modélisation à l'aveugle, les interactions entre RPB4 et le domaine « *core* » de FluPol (i.e. sans le domaine endonucléase de PA et les domaines périphériques de PB2) ont été explorées dans un premier temps à l'aide de GramMX. Les vingt modèles ayant les meilleurs scores ont été superposés aux structures complètes de FluPol du dimère RPB4 / RPB7 de façon à vérifier l'absence de gênes stériques entre les complexes. Suite à cette étape, les quatre modèles restants ont été affinés localement à l'aide du logiciel Rosetta. Enfin, seuls les modèles dans lesquels RPB4 interagissait avec PB2 ont été retenus. A l'issue de cette démarche, deux modèles ont été obtenus avec des zones d'interaction voisines, à l'interface entre PB2 et PB1.

Ces modèles doivent à présent être confrontés à l'expérimentation *in vitro* / *in cellulo*. Pour se faire, les résidus identifiés dans chacun des modèles seront mutés en alanine puis leur(s) effet(s) sur l'interaction sera(ont) testé(s) en système split-luciférase ou par co-immunoprécipitation. Si l'un des modèles est validé, ces mutations seront également testées en système minigénome WT et PA C95A pour mesurer leur impact sur l'activité enzymatique générale ou transcriptase de FluPol puis des virus portant ces mutations seront générés par génétique inverse pour appréhender leur effet sur la capacité du virus à se répliquer. De plus, il serait pertinent de refaire des essais de modélisation en partant de structures de FluPol d'IBV ou d'ICV afin de voir si les zones d'interaction sont identiques ou alors sont modifiées. A terme, il serait attrayant d'obtenir une structure tridimensionnelle du complexe FluPol en interaction avec RPB4 (ou le dimère RPB4 / RPB7) afin de mieux comprendre les modalités de cette interaction. Ceci implique cependant que cette interaction soit stable et de relativement forte affinité pour être stabilisée et donc étudiée.

Enfin, il serait également intéressant de cibler la zone de PB2 impliquée dans l'interaction avec RPB4 à l'aide d'un nanoligand et mesurer son effet inhibiteur. Cependant, il est peu probable que l'un des VHH identifiés dans ce travail (cf partie III) cible cette zone car ils ont été produits à partir d'une forme de FluPol ne possédant que les 120 premiers résidus de PB2. En outre, parmi les VHH générés par l'équipe du Pr. Ervin Fodor, seul le Nb8207 cible le domaine N-terminal de PB2 (Keown *et al*, 2022). Malheureusement, ce VHH ciblerait le canal de sortie des ARN de FluPol et non nos zones d'intérêt. Il conviendrait donc de générer de nouveaux VHH pour obtenir des nanoligands spécifiques de cette zone.

Interaction entre FluPol et le compartiment chromatinien

Dans la deuxième partie de ce projet, des expériences de CHIP-seq et de RNA-seq en contexte infectieux ont été menées pour mieux comprendre le ciblage de FluPol à la chromatine ainsi que son impact sur cette dernière.

Analyse non biaisée de l'impact de l'infection sur l'activité transcriptionnelle de la cellule

En accord avec des études précédentes, le séquençage des ARN cellulaires au cours de l'infection montre une modification du niveau d'expression des gènes dans le compartiment chromatinien *via* une diminution du nombre de transcrits. Des études précédentes avaient déjà documenté cette redistribution de l'expression (Bauer *et al*, 2018; Heinz *et al*, 2018; Zhao *et al*, 2018). Pol II est ainsi moins présente aux promoteurs et dans le corps des gènes mais est plus présente dans les régions intergéniques.

Nous avons voulu vérifier si ces changements des niveaux de transcrits et leur assignement aux gènes dont ils sont issus n'étaient pas dus à un biais dans la méthode de traitement des résultats de séquençage. En effet, comme démontré précédemment par plusieurs études, la protéine virale NS1 inhibe la terminaison de la transcription. Pol II se retrouve donc dans des régions intergéniques où elle n'est habituellement pas présente. Ce phénomène entraîne plusieurs biais dans l'estimation du niveau d'expression des gènes de différentes façons, le nombre de « *reads* » totaux de séquençage étant le même pour les échantillons infectés ou non. Pol II poursuivant la transcription y compris dans les zones intergéniques, un certain nombre de « *reads* » sont donc alignés sur ces régions intergéniques. Le nombre de « *reads* » présents sur des régions normalement exprimées est donc proportionnellement diminué. Ces gènes sont donc comptabilisés comme étant sous-exprimés alors qu'ils ne sont en réalité pas impactés par l'infection. *A contrario*, l'absence de terminaison de la transcription induisant une continuité de transcription de Pol II, celle-ci peut donc transcrire un gène présent en aval, proche du site de terminaison de la transcription. Ce gène apparaît donc comme étant surexprimé dans l'échantillon infecté alors qu'il est juste le produit de la continuité de synthèse du gène en amont.

Pour contourner ce problème, nous avons choisi d'utiliser une méthode de normalisation basée sur les gènes de référence. Nous avons ainsi émis le postulat que le niveau d'expression des transcrits à longue durée de vie (> quinze heures) ne devait pas être impacté à 6 h.p.i.. Ces gènes ont donc été utilisés pour normaliser les niveaux d'expression dans nos échantillons infectés, permettant ainsi de mesurer l'impact réel de l'infection sur l'expression des gènes.

Les cibles de FluPol

Nos expériences de CHIP-seq au cours desquelles FluPol est immunoprécipitée nous ont permis de visualiser l'emplacement de FluPol sur la chromatine au cours de l'infection. L'interaction préalablement décrite de FluPol avec le CTD portant la Ser5P suggère une interaction avec Pol II en phase d'initiation ou en début d'élongation de la transcription. Nous montrons en réalité que FluPol interagit avec Pol II tout au long de la transcription, y compris dans les régions intergéniques, au-delà du site de terminaison de la transcription. FluPol reste donc fixée à Pol II, possiblement même après le vol de coiffe. Cette présence de FluPol après les sites de terminaison de la transcription (TES) suggère un rôle de FluPol dans l'inhibition de la terminaison, actuellement principalement attribué à NS1. Cette hypothèse est en accord avec les conclusions de l'étude de Bauer *et al* qui avait montré qu'un virus n'exprimant pas NS1 inhibait tout de même la terminaison de la transcription (Bauer *et al*, 2018). Ce phénomène pourrait donc bien être dû à FluPol.

Nos travaux nous permettent de proposer le modèle suivant : l'interaction de FluPol avec le motif répété du CTD de Pol II porteur d'une Ser5P et avec RPB4 a lieu lors de l'initiation de la transcription cellulaire ou en début d'élongation. FluPol reste fixée sur Pol II lors de la transcription sans terminaison médiée par NS1 et le complexe Pol II / FluPol. Cette liaison continue de FluPol à Pol II pourrait également permettre aux transcrits de FluPol d'être plus facilement pris en charge par les machineries cellulaires pour leur épissage puis leur export hors du noyau.

Enfin, l'étude des ARNm liés à FluPol nous a également permis de mettre en évidence un enrichissement significatif au niveau des gènes ayant un rôle dans la réponse immunitaire innée antivirale. FluPol aurait ainsi un rôle supplémentaire dans le « *shut-off* » de l'hôte en ciblant particulièrement les gènes de la réponse interféron.

De plus en plus d'études utilisent des technologies de séquençage à haut débit pour percer le fonctionnement de FluPol à la chromatine. Or, la plupart de ces études sont difficilement comparables car réalisées dans des conditions différentes (MOI, techniques utilisées, temps d'infection...), ce qui mène souvent à des conclusions différentes voire discordantes. Ainsi, Chan *et al* concluent à une inhibition de l'élongation par l'infection alors que Bauer *et al* n'observent pas ce phénomène mais plutôt une diminution progressive de Pol II le long des gènes (Bauer *et al*, 2018; Chan *et al*, 2006). Or, l'étude de 2006 a été réalisée à 3 h.p.i., pendant la phase de transcription primaire alors que la seconde étude analyse l'impact de l'infection à 12 h.p.i., en phase tardive de l'infection (soit plutôt en phase de réplication du génome). Ces études se placent donc à des moments très différents du cycle.

Il apparaît donc très difficile de déterminer avec précision le moment où FluPol vole la coiffe des ARN cellulaires. Pour essayer de réduire la fenêtre temporelle du vol de coiffe, des expériences

d'immunoprécipitation successives pourraient peut-être être menées. Ainsi, l'immunoprécipitation des différentes formes phosphorylées de Pol II concomitante à l'immunoprécipitation de FluPol permettrait de récupérer uniquement les FluPol liées à Pol II et de définir une fenêtre temporelle qui pourrait être vérifiée par le séquençage de l'ADN co-immunoprécipité. Parallèlement, les ARN co-immunoprécipités pourraient être isolés et la proportion d'ARN coiffés ou non pourrait être déterminée par l'utilisation d'anticorps anti-coiffe. Une autre alternative pourrait être l'utilisation de la microscopie à super-résolution en temps réel sur molécule unique en marquant FluPol et les ARN en cours de synthèse.

Des nano-outils pour disséquer le fonctionnement de FluPol

Dans la dernière partie de ce travail, deux types de nanoligands protéiques spécifiques de FluPol ont été sélectionnés par phage-display : des nanobody / VHH et des protéines synthétiques appelées α Rep. Alors que les VHH ont été sélectionnés après immunisation de FluPol à un lama, les α Rep ont été obtenues par criblage d'une banque de phage préalablement construite. Sept VHH et quatre α Rep dirigés contre FluPol ont été générés. Dans un premier temps, leur capacité à inhiber l'activité de FluPol *in cellulo* par suivi d'un gène rapporteur marqueur de la réplication / transcription virale a été quantifiée afin de sélectionner les nanoligands ayant le plus grand pouvoir inhibiteur. Le VHH16 étant le nanoligand avec la plus forte activité inhibitrice, il a été sélectionné pour la suite de cette étude. Ce nanoligand présentant une très forte affinité pour FluPol, de l'ordre du sub-nanomolaire, il reconnaît très efficacement le dimère PA – PB1, et dans une moindre mesure PA exprimée seule. Le VHH16 n'interfère pas sur le transport des sous-unités de FluPol au noyau ni sur son assemblage.

Pour prolonger la caractérisation du VHH16 et des autres nanoligands sélectionnés ainsi que leurs propriétés biologiques, des essais seront réalisés pour déterminer leur pouvoir neutralisant en contexte infectieux par expression en intracellulaire. Afin de mieux caractériser leurs activités, des études structurales seront entreprises pour identifier les domaines de FluPol reconnus. Il pourrait être également envisagé de sélectionner des nanoligands spécifiques des domaines mobiles de FluPol. A plus long terme, il pourra être envisagé de cribler des drogues à vocation antivirale capables de reconnaître les domaines de FluPol identifiés comme neutralisables par nos nanoligands.

L'ensemble des résultats obtenus contribue à une meilleure connaissance du fonctionnement de FluPol au sein de la cellule infectée.

MATÉRIELS ET MÉTHODES

Partie 1

Split-luciférase

Les cellules HEK-293T (ATCC) ont étéensemencées à $0,1.10^6$ cellules par puits en plaque 48 puits. Vingt heures plus tard, les cellules sont transfectées en triplicata à l'aide de polyéthylèneimine (PEI ; 2,5 μ L de Pei / μ g d'ADN ; Polyscience Inc.) avec 150 ng des plasmides pci vide, pci-RPB4-Luc1 et pci-PB2(36-247)-Luc2 sauvage ou portant les différentes mutations des acides aminés étudiés. Vingt-huit heures post-transfection, les cellules sont lysées dans du « *Renilla Luciferase Assay Lysis Buffer* » pendant 30 minutes à 200 rpm et la luminescence est révélée à l'aide du kit « *Renilla Luciferase Assay System* » (Promega) selon les indications du fournisseur. La luminescence est mesurée avec un luminomètre Tecan Infinite 200 PRO. Les données ont été analysées sur Microsoft Excel (Microsoft Corporation, 2018) et mises en forme avec le logiciel GraphPad Prism.

Modélisation *in silico* de l'interaction PB2-RPB4

La modélisation de l'interaction PB2/RPB4 a comporté plusieurs étapes résumées ci-après :

1- Sélection des coordonnées de départ : Les coordonnées cristallographiques de la sous-unité RPB4 de Pol II (code PDB 2C35 chaîne A résidus 14-142 (Meka, 2005)) et du domaine FluPol de IAV lié à une coiffe d'ARNm (code PDB 6RR7 (Fan *et al*, 2019a)) à laquelle les domaines périphériques ont été préalablement retirés sont sélectionnées. Ce domaine "core" est composé du domaine C-terminal de PA (résidus 257-716), de PB1 complète (résidus 1-756) et du domaine N-terminal de PB2 (résidus 1-248).

2- Exploration globale en corps rigides avec GrammX : L'interaction entre les entités RPB4 et core_FluPol, considérées chacune comme corps rigide, est explorée de manière stochastique globale par GrammX (GRAMM-X Protein-Protein Docking Web Server v.1.2.0), un web server de recherche stochastique d'interaction protéine/protéine (Tovchigrechko & Vakser, 2006). Cette première exploration globale se fait sans *a priori* d'interface. Les paramètres par défaut sont gardés et vingt modèles sont obtenus.

3- Clusterisation et sélection : Les vingt modèles sont visualisés sous PyMOL (Schrödinger, LLC, 2015) et les complexes redondants sont éliminés.

4- Sélection après vérification des données biologiques : a) Les modèles sélectionnés à ce stade sont superposés à FluPol entière. Les modèles de GrammX qui donnent des interactions volumiques incompatibles avec les domaines périphériques de FluPol sont éliminés. b) Ces modèles sont ensuite superposés à RPB4 en complexe avec RPB7. De manière similaire, ces modèles

Core_FluPol/RPB4 qui donnent des interactions volumiques incompatibles avec RPB7 sont éliminés. A ce stade, quatre complexes sont retenus.

5- Exploration stochastique locale avec Rosetta : Les quatre complexes sont affinés localement en utilisant le logiciel Rosetta (Rosetta online server) d'exploration d'interactions protéine/protéine locales, puis optimisés et classés (Lyskov *et al*, 2013).

6- Visualisation finale : Une inspection visuelle des complexes optimisés permet de retenir deux modèles qui donnent des interactions entre RPB4 et le domaine N-terminal de PB2 de FluPol où le premier peut être en complexe avec RPB7 et le second peut être entier. Deux complexes sont retenus et leur interface est caractérisée avec l'outil Pisa (PDBePISA (Proteins, Interfaces, Structures and Assemblies)) (Krissinel & Henrick, 2007).

Partie 3

Purification de FluPol

Après expression des complexes PA(197-716) - PB1 - PB2(1-116) [Forme A] et PA(197-716) – PB1(1-686) [Forme B] en cellules d'insecte, ces dernières ont été centrifugées à 800 rpm pendant 15 minutes. Les culots ont été repris dans un tampon de lyse (50 mM Hepes pH 7,5, 300 mM NaCl, 1 mM β -mercaptoéthanol) supplémenté avec une pastille d'antiprotéase (cOmplete, EDTA-free ; Roche). Ils ont ensuite été plusieurs fois plongés successivement dans l'azote liquide puis décongelés à 27°C. Du glycérol a été ajouté pour atteindre une concentration finale de 10 %. Les échantillons ont ensuite été centrifugés pendant 30 minutes à 17 000 rpm. Le surnageant a été récupéré, de l'imidazole a été ajouté pour atteindre une concentration de 15 mM, mis en contact avec 3 mL de résine NiNTA (QIAGEN) puis l'ensemble a été agité sur roue à 4°C pendant 30 minutes avant d'être centrifugé pendant 10 minutes à 2 000 rpm. La résine a été reprise dans le tampon d'équilibration (50 mM Hepes pH 7,5, 300 mM NaCl, 1 mM β -mercaptoéthanol, 15 mM imidazole, 10 % glycérol) afin d'être transférée dans une colonne vide puis lavée avec du tampon d'équilibration dans un premier temps puis avec un tampon à forte concentration en sel (50 mM Hepes pH 7,5, 1 M NaCl, 1 mM β -mercaptoéthanol, 15 mM imidazole, 10 % glycérol). Après un dernier lavage de la résine avec le tampon d'équilibration, FluPol est éluée avec un tampon 20 mM Hepes pH 7,5, 300 mM NaCl, 1 mM β -mercaptoéthanol, 500 mM imidazole, 10 % glycérol. Les différentes fractions ont été analysées en gel SDS-Page et les fractions contenant FluPol sont rassemblées. La forme A a été dialysée sur la nuit dans un tampon 20 mM Hepes pH 7,5, 300 mM NaCl, 5 mM β -mercaptoéthanol, 10 % glycérol puis les échantillons ont été chargés sur une colonne héparine (HiTrap™ Heparin HP 5 mL ; GE Healthcare) reliée à un automate de purification (NGC ; BioRad) et équilibrée avec le tampon A (20 mM Hepes pH 7,5 , 300 mM NaCl, 1 mM

β -mercaptoéthanol, 10 % glycérol). L'élution est effectuée par un gradient (sur 50 mL à un débit de 1 mL.min⁻¹) entre le tampon A et le tampon B (20 mM Hepes pH 7,5, 2 M NaCl, 1 mM β -mercaptoéthanol, 10 % glycérol). Son suivi est effectué par analyse des absorbances à 215, 260 et 280 nm en temps réel. Les fractions d'intérêt ont été analysées en gel SDS-Page et celles contenant FluPol sont rassemblées.

Sélection des VHH

Afin de générer des nanobodies spécifiques de FluPol, un lama a reçu cinq injections de 0,3 mg de la forme A de FluPol diluée dans de l'adjuvant de Freund complet (première injection) ou incomplet à sept jours d'intervalle. Cinq jours après la dernière immunisation, les lymphocytes ont été collectés, les ARN isolés et les ADNc synthétisés par transcription inverse. Les régions codant pour les domaines variables des chaînes lourdes des anticorps ont été amplifiées par PCR puis clonées dans un vecteur phagemide pHENA pour construction d'une banque de phages. FluPol a été adsorbée sur une plaque Nunc MaxiSorp™ (Invitrogen™) et les nanobodies spécifiques de FluPol ont été sélectionnés par phage-display comme décrit dans (Duhoo *et al*, 2017). Après deux tours de sélection, les séquences des VHH retenus ont été analysées et sept VHH uniques ont été identifiés. Leurs séquences ont été sous-clonées par digestion enzymatique (NcoI / NotI) en plasmide d'expression bactérien pET22 comportant un signal pelB d'export dans l'espace périplasmique ou par recombinaison à l'aide du kit Gibson Assembly® Master Mix (NEB) en plasmide d'expression eucaryote pci pour étude des caractéristiques des différents nanobodies.

Sélection des α REP

Afin de sélectionner des α Rep spécifiques de FluPol [forme B], la banque de phages 2.1 (Guellouz *et al*, 2013) a été criblée par phage-ELISA. Brièvement, les puits de plaques ELISA ont été tapissés de Streptavidine puis incubés avec de la biotine-Tris-NTA afin de pouvoir y fixer FluPol via son His-tag. La banque de phages, préalablement amplifiée en infectant des bactéries XL1 Blue à l'aide du phage helper M13-K1, a été incubée en présence de TTBS - 4% BSA afin de limiter les interactions non spécifiques. Les phages ont été incubés en présence de FluPol [forme B] pendant deux heures à 300 rpm puis vingt lavages ont été effectués pour éliminer les phages non fixés. Les phages fixés ont été élués par élution acide à l'aide de glycine 0,1M pH2,5 pendant 10 minutes à 300 rpm. L'acidité a ensuite été neutralisée avec du Tris 1M et les phages récoltés amplifiés en infectant des bactéries XL1 Blue cultivées en milieu 2YT supplémenté en glucose (1%, p/v), ampicilline (200 mg.mL⁻¹) et tétracycline (12,5 mg.mL⁻¹). Ces étapes ont été répétées à nouveau deux fois en faisant trente lavages afin de bien éliminer les phages non spécifiques.

Après le troisième tour de sélection, les colonies obtenues ont été individuellement récupérées et amplifiées puis cultivées toute une nuit à 30°C en plaque P96 Deepwell après infection par le phage helper K1. La plaque est ensuite centrifugée pendant une heure à 2000 g à 4°C et les phages récupérés. La spécificité de ces phages a ensuite été testée par phage-ELISA contre FluPol ou le Tris-NTA. La présence des phages est révélée par un anticorps HRP/Anti-M13 Monoclonal Conjugate (#27-9421-01, Cytiva) dilué au 1/5000^e à l'aide de substrat BM Blue POD soluble (Roche) conformément aux instructions du fournisseur.

Les bactéries contenant les phagemides des phages exprimant les α Rep spécifiques de FluPol ont été amplifiées et leur ADN extrait à l'aide du kit « NucleoSpin Plasmid » (Macherey Nagel) puis envoyé au séquençage. Après analyse des séquences, quatre clones ont été obtenus. Leur séquence a été sous-clonée par digestion enzymatique (BamHI / HindIII) dans un plasmide d'expression en bactérie pQE81 (Quiagen) puis en plasmide d'expression eucaryote pCI, dont le site NheI a été préalablement déplacé, par digestion EcoRI / NheI pour analyse fonctionnelle.

Minigénome

Les cellules HEK-293T (ATCC) ont étéensemencées à $0,05 \cdot 10^6$ cellules par puits en plaque 96 puits. Vingt heures plus tard, les cellules sont transfectées en quadruplicate à l'aide de polyéthylèneimine (PEI ; 2,5 μ L de Pei / μ g d'ADN ; Polyscience Inc.) avec 50 ng de plasmides pcDNA3 codant pour les sous-unités PA, PB1 et PB2, 82,5 ng de plasmide pcDNA3-NP et 132 ng de plasmide minigénome pPoll-WSN-NA-Luciférase codant pour la firefly luciférase dont la séquence est flanquée des régions UTR du segment NA en présence de :

- soit 10, 20 ou 50 ng de plasmide pci codant pour un des nanoligands (Figure 40)
- soit 20 ng de plasmide pci codant pour un des nanoligands (Figure 41). Dans cette expérience, le plasmide pcDNA3-PA a également été remplacé par une version mutée de ce plasmide portant la mutation C95A.
- soit une quantité décroissante allant de 0,25 à 50 ng de pci-VHH16 ou pci-VHH KO7 (Figure 42)

Cinquante heures post-transfection, les cellules sont lysées avec du tampon PLB (30 mM Tris pH7,9, 10mM MgCl₂, 1% Triton X-100, 20% glycérol, 1mM DTT) pendant 30 minutes à 200 rpm et la luminescence est révélée à l'aide du kit « *Luciferase Assay System* » (Promega) selon les indications du fournisseur. La luminescence est mesurée avec un luminomètre Tecan Infinite 200 PRO. Les données ont été analysées sur Microsoft Excel (Microsoft Corporation, 2018) et mises en forme avec le logiciel GraphPad Prism.

Marquage Alexa Fluor® 488 du VHH16

Le marquage réalisé permet de greffer le fluorophore Alexa Fluor® 488 à l'amine terminale du VHH16 et de pouvoir ainsi révéler sa présence par fluorescence. Pour se faire, 6 µL d'Ester NHS Alexa Fluor™ 488 (ester succinimidyle, #A20000, Invitrogen™) sont ajoutés à 3 mL de VHH16 à 29 µM (soit 0,4 mg.mL⁻¹) en 20mM Hepes pH7,2 100mM NaCl. La réaction est arrêtée une heure plus tard par ajout de Tris 1M pH7,4 à 1mM Tris final. L'efficacité de la réaction est vérifiée par chromatographie d'exclusion de taille (sur colonne G25) en mesurant l'absorbance à 488 nm. Cette étape permet également d'éliminer le fluorophore non fixé. Une quantité croissante de la protéine est également déposée sur gel 15% acrylamide puis migrée en SDS-Page. Le gel est fixé dans du tampon de décoloration et de l'eau puis la fluorescence est révélée à la lumière bleue. Le gel est ensuite coloré au Bleu de Coomassie pour estimer la proportion de protéine marquée par rapport à la quantité totale de protéine.

Western Blot

Les formes A et B de FluPol précédemment purifiées ont été déposées sur un gel 12% acrylamide pour analyse SDS-Page. Les protéines ont ensuite été transférées sur membrane de nitrocellulose. La membrane a été bloquée pendant 1 heure dans un tampon PBS - 0.1 % Tween-20 contenant 5 % de poudre de lait avant d'être incubée 1 heure à température ambiante dans un tampon PBS – 0,1% Tween – 5% lait contenant le VHH16 Alexa Fluor® 488, un anticorps anti-PB1 produit en lapin provenant du laboratoire de Juan Ortin (Madrid) et un anti-His tag produit en ascites de souris (#27-4710-01, GE Healthcare). La membrane a ensuite été lavée trois fois avec du tampon PBS - 0,1% Tween puis incubée 1 heure à température ambiante en présence des anticorps secondaires dilués dans le tampon PBS-Tween en suivant les recommandations du fournisseurs : IgG anti souris produits en chèvre marqués avec l'Alexa Fluor® 532 (#A11002, Life Technologies) dilué au 1/1000^e et IgG anti lapin produits en chèvre marqués avec l'Alexa Fluor® 633 (#A21070, Life Technologies) dilué au 1/2000^e. Après trois lavages avec du tampon PBS – 0,1%Tween, la membrane a été séchée et la fluorescence révélée avec un scanner à fluorescence (Amersham Typhoon, GE Healthcare).

Split-luciférase

Les cellules HEK-293T (ATCC) ont étéensemencées à 0,1.10⁶ cellules par puits en plaque 48 puits. Vingt heures plus tard, les cellules sont transfectées en triplicata à l'aide de polyéthylèneimine (PEI ; 2,5 µL / µg d'ADN ; Polyscience Inc.) avec :

- soit 150 ng des plasmides pci-PA/PB1/PB2-Luc1 et pci-VHH16-Luc2 complétés avec 150 ng de plasmide pci vide ou 75 ng de pci vide + 75 ng de pci-PA/PB1 ou 75 ng de pci codant pour chaque sous-unités complémentaires de FluPol (Figure 44).

- soit 150 ng des plasmides pci-PA-Luc1 et pci-PB1/PB2-Luc2 complétés de 150 ng de pci-VHH16, pci-VHH7, pci-F9 ou complétés avec 75 ng de pci vide ou pci-PB1 et 75 ng de pci-VHH16 (Figure 47).

- soit 150 ng des plasmides pci-PA-Luc1 et pci-RanBP5-Luc2 complétés de 75 ng de pci vide ou pci-PB1 et de 75 ng d'un plasmide pci codant pour un nanoligand (Figure 49).

Vingt-huit heures post-transfection, les cellules sont lysées pendant 30 minutes à 200 rpm et la luminescence est révélée à l'aide du kit « *Renilla Luciferase Assay* » (Promega) selon les indications du fournisseur. La luminescence est mesurée avec un luminomètre Tecan Infinite 200 PRO. Les données ont été analysées sur Microsoft Excel (Microsoft Corporation, 2018) et mises en forme avec le logiciel GraphPad Prism.

Immunofluorescence

Les cellules HEK 293T ont étéensemencées sur lamelles puis transfectées avec 0,5 µg de chaque plasmide à l'aide de polyéthylèneimine (PEI). Trente heures post-transfection, les cellules sont fixées en PBS - 4% paraformaldéhyde (PFA) puis perméabilisées avec une solution PBS - 0,2% Triton. Le VHH16 est ensuite révélé après incubations successives avec un anticorps anti-HA tag produit chez le lapin (#H6908, Sigma) dilué au 1/100^e puis avec un anticorps anti-lapin couplé Alexa Fluor® 594 (#A21207, Life technologies) dilué au 1/1000^e. Les noyaux ont été colorés au DAPI. Les lamelles sont montées sur lame avec le milieu de montage ProLong™ Diamond Antifade Mountant (#P36961, invitrogen). Les clichés ont été réalisés avec un microscope inversé à épifluorescence Zeiss (AxioObserver Z1) puis traités avec le logiciel ImageJ.

Purification des VHH

Les VHH fusionnés à un HisTag ont été exprimés dans l'espace périplasmique de bactéries compétentes Rosetta après induction avec 1 mL d'IPTG 1M pour 1 L de culture. Les bactéries ont ensuite été cultivées à 28°C toute la nuit puis centrifugées pendant 30 minutes à 7 650 g et le culot a été repris dans un tampon hypertonique 30 mM Tris-HCl pH8, 20 % w/v sucrose, 1 mM EDTA. Les échantillons ont été incubés à 4°C pendant 30 minutes puis centrifugés à 7 650g pendant 30 minutes. Les surnageants ont été conservés à 4°C et les culots repris dans une solution hypotonique 5 mM MgSO₄ avant d'être incubés à 4°C pendant 30 minutes. Les échantillons ont été à nouveau centrifugés

pendant 30 minutes à 7 650g puis les surnageants ont été rassemblés. Une dernière centrifugation des surnageants pendant 30 minutes à 7650g a été réalisée afin d'éliminer les débris restants.

Les VHH ont été purifiés sur une résine de nickel (HiTrap™ IMAC HP 1mL, Cytiva) reliée à un automate de purification (ÄKTA, Cytiva) et équilibrée avec le tampon 20 mM Tris-HCl pH8, 300 mM NaCl, 40 mM imidazole. Les protéines ont été éluées avec un tampon 20 mM Tris-HCl pH8, 300 mM NaCl, 500 mM imidazole. Les différentes fractions ont été analysées en gel SDS-Page et celles contenant les VHH sont rassemblées, mises à dialyser sur la nuit contre le tampon de gel filtration (20 mM Hepes pH 7,5, 100 mM NaCl). Une étape de purification par exclusion de taille a ensuite été réalisée. Les solutions de VHH ont été concentrées par centrifugation (Amicon® Ultra-4 Ultracel®-3K) avant d'être injectées sur une colonne Superdex™ 75 Increase 10/300 GL (GE Healthcare). Le suivi de cette étape est effectué par analyse des absorbances comme précédemment décrit. Les fractions d'intérêt ont été analysées en gel SDS-Page et celles contenant les VHH sont rassemblées et concentrées par centrifugation jusqu'à l'obtention de concentrations comprises entre 10 et 15 mg.mL⁻¹.

Structure du VHH18 par cristallographie et diffractions des rayons X

Le criblage des conditions de cristallisation a été effectué par la plateforme HTX (EMBL ; Grenoble) avec une solution de protéine concentrée à 14 mg.mL⁻¹. A partir d'une condition identifiée (0,1M Hepes pH 7,5 ; 2 M Ammonium sulfate ; 2 % PEG 400), des cristaux ont été reproduits manuellement par la méthode de la goutte suspendue en criblant la concentration des différents constituants (tampon, sulfate d'ammonium et PEG 400). Ils ont pu être reproduits dans la condition 0,1M Hepes pH 7,5 ; 1.8 M Ammonium sulfate ; 2 % PEG 400. Les cristaux ont été transférés dans une solution cryoprotectante (0,1M Hepes pH 7,5 ; 3 M Ammonium sulfate ; 2 % PEG 400) avant d'être immédiatement congelés dans l'azote liquide.

Les données de diffraction ont été collectées au synchrotron SOLEIL (Gif-sur-Yvette) sur la ligne PROXIMA-1 et traitées avec le programme XDS (PMID : 20124693, 20124692). La structure a été résolue par remplacement moléculaire en utilisant le programme PHASER (PMID : 19461840) de la suite de programme CCP4 (PMID : 21460441) et la structure d'un VHH de la *Protein Data bank* présentant 78 % d'identité de séquence avec celle du VHH18 comme modèle (PDB : 4HEM, PMID : 23530214). Le modèle a été construit manuellement avec le programme COOT (PMID : 20383002) et affiné avec le programme REFMAC (PMID : 15299926) jusqu'à obtenir les données présentées dans le tableau 4.

Interférométrie de biocouches (*Bio-layer interferometry* - BLI)

Le VHH16 fusionné à un StrepTag a été exprimé dans l'espace périplasmique de bactéries compétentes Rosetta après induction avec 1 mM d'IPTG. Les bactéries ont ensuite été cultivées à 28°C toute la nuit, ont été centrifugées pendant 30 minutes à 7 650 g et le culot a été repris dans un tampon hypertonique 20 mM Hepes pH 7,5, 20 % w/v sucrose, 1 mM EDTA. L'échantillon a été incubé à 4°C pendant 30 minutes puis centrifugé à 7 650 g pendant 30 minutes. Le surnageant a été conservé à 4°C et le culot repris dans une solution hypotonique 5 mM MgSO₄ avant d'être incubé à 4°C pendant 30 minutes. L'échantillon a été à nouveau centrifugé pendant 30 minutes à 7 650 g puis les deux surnageants ont été rassemblés. Une dernière centrifugation a été réalisée afin d'éliminer les débris restants. La solution a ensuite été diluée au demi dans un tampon 20 mM Hepes pH7,5, 300 mM NaCl, 1 mM β-mercaptoéthanol et filtrée avant purification par chromatographie d'affinité (colonne utilisée : StrepTrap™ HP 1mL (Cytiva)). La protéine a été éluée avec un gradient de tampon 20 mM Hepes pH 7,5, 300 mM NaCl, 1 mM β-mercaptoéthanol, 2,5 mM desthiobiotine. Dans un second temps, le Strep-VHH16 a été purifié par chromatographie d'exclusion de taille (colonne utilisée : Superdex™ 75 Increase 10/300 GL (GE Healthcare)) avant d'être biotinyllé. La biotinylation a été réalisée avec le kit *EZ-Link NHS-PEG4-Biotin* (Thermo Scientific) en suivant les recommandations du fabricant et le Strep-VHH16 a été conservé dans du tampon PBS avant son utilisation.

Les expériences d'interférométrie de biocouches ont été réalisées sur la plateforme SPR/BLI de l'ISBG (URA3518; Grenoble) avec un OctetRED96e (Sartorius) et des biocapteurs streptavidine. Des expériences pilotes ont validé les échantillons utilisés et ont permis de déterminer les conditions optimales de mesure. Les mesures cinétiques ont ainsi été réalisées en utilisant une concentration fixe de Strep-VHH16 (dilué dans un tampon PBS + 0,02 % Tween-20) à 1,5 µg.mL⁻¹ et des gammes de polymérase virale de 1,56 à 100 nM en cascade dans une solution d'HBS + 0,02 % Tween-20. Les données ont été analysées avec le programme d'analyse HT 11.1 en utilisant le modèle de liaison 1:1 entre le ligand et l'analyte.

RÉFÉRENCES

- Akarsu H, Burmeister WP, Petosa C, Petit I, Müller CW, Ruigrok RWH & Baudin F (2003) Crystal structure of the M1 protein-binding domain of the influenza A virus nuclear export protein (NEP/NS2). *EMBO J* 22: 4646–4655
- 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: 4143–4156
- Andersson R, Gebhard C, Miguel-Escalada I, Hoof I, Bornholdt J, Boyd M, Chen Y, Zhao X, Schmidl C, Suzuki T, *et al* (2014) An atlas of active enhancers across human cell types and tissues. *Nature* 507: 455–461
- Andrade MA & Bork P (1995) HEAT repeats in the Huntington's disease protein. *Nat Genet* 11: 115–116
- Andrulis ED, Werner J, Nazarian A, Erdjument-Bromage H, Tempst P & Lis JT (2002) The RNA processing exosome is linked to elongating RNA polymerase II in *Drosophila*. *Nature* 420: 837–841
- Arranz R, Coloma R, Chichón FJ, Conesa JJ, Carrascosa JL, Valpuesta JM, Ortín J & Martín-Benito J (2012) The structure of native influenza virion ribonucleoproteins. *Science (New York, NY)* 338: 1634–7
- Asadi S, Gaaloul Ben Hnia N, Barre RS, Wexler AS, Ristenpart WD & Bouvier NM (2020) Influenza A virus is transmissible via aerosolized fomites. *Nat Commun* 11: 4062
- Ashour J, Schmidt FI, Hanke L, Cragolini J, Cavallari M, Altenburg A, Brewer R, Ingram J, Shoemaker C & Ploegh HL (2015) Intracellular expression of camelid single-domain antibodies specific for Influenza virus nucleoprotein uncovers distinct features of its nuclear localization. *J Virol* 89: 2792–2800
- Baker SF, Ledwith MP & Mehle A (2018) Differential splicing of ANP32A in birds alters its ability to stimulate RNA synthesis by restricted Influenza polymerase. *Cell Rep* 24: 2581-2588.e4
- Batra J, Tripathi S, Kumar A, Katz JM, Cox NJ, Lal RB, Sambhara S & Lal SK (2016) Human Heat shock protein 40 (Hsp40/DnaJB1) promotes influenza A virus replication by assisting nuclear import of viral ribonucleoproteins. *Sci Rep* 6: 19063
- 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: 3158–3165
- Bauer DLV, Tellier M, Martínez-Alonso M, Nojima T, Proudfoot NJ, Murphy S & Fodor E (2018) Influenza virus mounts a two-pronged attack on host RNA polymerase II transcription. *Cell Reports* 23: 2119-2129.e3
- Beloso A, Martínez C, Valcárcel J, Santarén JF & Ortín J (1992) Degradation of cellular mRNA during influenza virus infection: its possible role in protein synthesis shutoff. *J Gen Virol* 73 (Pt 3): 575–581
- Bercovich-Kinori A, Tai J, Gelbart IA, Shitrit A, Ben-Moshe S, Drori Y, Itzkovitz S, Mandelboim M & Stern-Ginossar N (2016) A systematic view on influenza induced host shutoff. *Elife* 5: e18311

- Bieniossek C, Imasaki T, Takagi Y & Berger I (2012) MultiBac: expanding the research toolbox for multiprotein complexes. *Trends Biochem Sci* 37: 49–57
- Bieniossek C, Richmond TJ & Berger I (2008) MultiBac: multigene baculovirus-based eukaryotic protein complex production. *Curr Protoc Protein Sci* Chapter 5: Unit 5.20
- Binz HK & Plückthun A (2005) Engineered proteins as specific binding reagents. *Curr Opin Biotechnol* 16: 459–469
- Binz HK, Stumpp MT, Forrer P, Amstutz P & Plückthun A (2003) Designing repeat proteins: well-expressed, soluble and stable proteins from combinatorial libraries of consensus ankyrin repeat proteins. *J Mol Biol* 332: 489–503
- Biswas SK, Boutz PL & Nayak DP (1998) Influenza virus nucleoprotein interacts with influenza virus polymerase proteins. *J Virol* 72: 5493–5501
- Boivin S, Cusack S, Ruigrok RWH & Hart DJ (2010) Influenza A virus polymerase: structural insights into replication and host adaptation mechanisms. *J Biol Chem* 285: 28411–28417
- Bradel-Tretheway BG, Mattiaccio JL, Krasnoselsky A, Stevenson C, Purdy D, Dewhurst S & Katze MG (2011) Comprehensive proteomic analysis of influenza virus polymerase complex reveals a novel association with mitochondrial proteins and RNA polymerase accessory factors. *J Virol* 85: 8569–8581
- Brunner J (1989) Testing topological models for the membrane penetration of the fusion peptide of influenza virus hemagglutinin. *FEBS Lett* 257: 369–372
- Brunotte L, Flies J, Bolte H, Reuther P, Vreede F & Schwemmle M (2014) The nuclear export protein of H5N1 influenza A viruses recruits Matrix 1 (M1) protein to the viral ribonucleoprotein to mediate nuclear export. *J Biol Chem* 289: 20067–20077
- Bui M, Whittaker G & Helenius A (1996) Effect of M1 protein and low pH on nuclear transport of influenza virus ribonucleoproteins. *J Virol* 70: 8391–8401
- 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: 1781–1786
- Cady SD, Luo W, Hu F & Hong M (2009) Structure and function of the influenza A M2 proton channel. *Biochemistry* 48: 7356–7364
- Camacho-Zarco AR, Kalayil S, Maurin D, Salvi N, Delaforge E, Milles S, Jensen MR, Hart DJ, Cusack S & Blackledge M (2020) Molecular basis of host-adaptation interactions between influenza virus polymerase PB2 subunit and ANP32A. *Nat Commun* 11: 3656
- Cardoso FM, Ibañez LI, Van den Hoecke S, De Baets S, Smet A, Roose K, Schepens B, Descamps FJ, Fiers W, Muyltermans S, *et al* (2014) Single-domain antibodies targeting neuraminidase protect against an H5N1 influenza virus challenge. *J Virol* 88: 8278–8296
- Carrique L, Fan H, Walker AP, Keown JR, Sharps J, Staller E, Barclay WS, Fodor E & Grimes JM (2020) Host ANP32A mediates the assembly of the influenza virus replicase. *Nature* 587: 638–643

- de Castro Martin IF, Fournier G, Sachse M, Pizarro-Cerda J, Risco C & Naffakh N (2017) Influenza virus genome reaches the plasma membrane via a modified endoplasmic reticulum and Rab11-dependent vesicles. *Nat Commun* 8: 1396
- Chambers LA, Henle W, Lauffer MA & Anderson TF (1943) Studies on the nature of the virus of Influenza: II. The size of the infectious unit in Influenza A. *Journal of Experimental Medicine* 77: 265–276
- Chan AY, Vreede FT, Smith M, Engelhardt OG & Fodor E (2006) Influenza virus inhibits RNA polymerase II elongation. *Virology* 351: 210–217
- 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: e1002187
- 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: 10059–10070
- Chen W, Calvo PA, Malide D, Gibbs J, Schubert U, Bacik I, Basta S, O'Neill R, Schickli J, Palese P, *et al* (2001) A novel influenza A virus mitochondrial protein that induces cell death. *Nat Med* 7: 1306–1312
- Chiu Y-L, Ho CK, Saha N, Schwer B, Shuman S & Rana TM (2002) Tat stimulates cotranscriptional capping of HIV mRNA. *Mol Cell* 10: 585–597
- Choder M (2004) Rpb4 and Rpb7: subunits of RNA polymerase II and beyond. *Trends Biochem Sci* 29: 674–681
- Clifford M, Twigg J & Upton C (2009) Evidence for a novel gene associated with human influenza A viruses. *Virol J* 6: 198
- Coloma R, Arranz R, de la Rosa-Trevín JM, Sorzano COS, Munier S, Carlero D, Naffakh N, Ortín J & Martín-Benito J (2020) Structural insights into influenza A virus ribonucleoproteins reveal a processive helical track as transcription mechanism. *Nat Microbiol* 5: 727–734
- Compans RW, Content J & Duesberg PH (1972) Structure of the ribonucleoprotein of influenza virus. *J Virol* 10: 795–800
- Connor RJ, Kawaoka Y, Webster RG & Paulson JC (1994) Receptor specificity in human, avian, and equine H2 and H3 influenza virus isolates. *Virology* 205: 17–23
- Conrath K, Pereira AS, Martins CE, Timóteo CG, Tavares P, Spinelli S, Kinne J, Flaudrops C, Cambillau C, Muyldermans S, *et al* (2009) Camelid nanobodies raised against an integral membrane enzyme, nitric oxide reductase. *Protein Sci* 18: 619–628
- Core L & Adelman K (2019) Promoter-proximal pausing of RNA polymerase II: a nexus of gene regulation. *Genes Dev* 33: 960–982
- Cramer P (2004) RNA polymerase II structure: from core to functional complexes. *Curr Opin Genet Dev* 14: 218–226

- Dadonaite B, Gilbertson B, Knight ML, Trifkovic S, Rockman S, Laederach A, Brown LE, Fodor E & Bauer DLV (2019) The structure of the influenza A virus genome. *Nat Microbiol* 4: 1781–1789
- Dadonaite B, Vijaykrishnan S, Fodor E, Bhella D & Hutchinson EC (2016) Filamentous influenza viruses. *J Gen Virol* 97: 1755–1764
- De Genst E, Silence K, Decanniere K, Conrath K, Loris R, Kinne J, Muyldermans S & Wyns L (2006) Molecular basis for the preferential cleft recognition by dromedary heavy-chain antibodies. *Proc Natl Acad Sci U S A* 103: 4586–4591
- De Vlioger D, Hoffmann K, Van Molle I, Nerinckx W, Van Hoecke L, Ballegeer M, Creytens S, Remaut H, Hengel H, Schepens B, *et al* (2019) Selective engagement of FcγRIV by a M2e-specific single domain antibody construct protects against Influenza A virus infection. *Front Immunol* 10: 2920
- Del Rosario JMM, Smith M, Zaki K, Risley P, Temperton N, Engelhardt OG, Collins M, Takeuchi Y & Hufton SE (2020) Protection from Influenza by intramuscular gene vector delivery of a broadly neutralizing nanobody does not depend on antibody dependent cellular cytotoxicity. *Front Immunol* 11: 627
- Deng T, Engelhardt OG, Thomas B, Akoulitchev AV, Brownlee GG & Fodor E (2006a) Role of ran binding protein 5 in nuclear import and assembly of the influenza virus RNA polymerase complex. *J Virol* 80: 11911–11919
- 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: 8669–8674
- Deng T, Vreede FT & Brownlee GG (2006b) Different de novo initiation strategies are used by influenza virus RNA polymerase on its cRNA and viral RNA promoters during viral RNA replication. *J Virol* 80: 2337–2348
- Desmyter A, Decanniere K, Muyldermans S & Wyns L (2001) Antigen specificity and high affinity binding provided by one single loop of a camel single-domain antibody. *J Biol Chem* 276: 26285–26290
- Desmyter A, Transue TR, Ghahroudi MA, Thi MH, Poortmans F, Hamers R, Muyldermans S & Wyns L (1996) Crystal structure of a camel single-domain VH antibody fragment in complex with lysozyme. *Nat Struct Biol* 3: 803–811
- 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: 315–328
- 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: 914–918
- Dieci G (2021) Removing quote marks from the RNA polymerase II CTD 'code'. *Biosystems* 207: 104468
- Diefenbacher M, Sun J & Brooke CB (2018) The parts are greater than the whole: the role of semi-infectious particles in influenza A virus biology. *Curr Opin Virol* 33: 42–46

- Dolk E, van der Vaart M, Lutje Hulsik D, Vriend G, de Haard H, Spinelli S, Cambillau C, Frenken L & Verrips T (2005) Isolation of llama antibody fragments for prevention of dandruff by phage display in shampoo. *Appl Environ Microbiol* 71: 442–450
- Doms RW & Helenius A (1986) Quaternary structure of influenza virus hemagglutinin after acid treatment. *J Virol* 60: 833–839
- Drake JW (1993) Rates of spontaneous mutation among RNA viruses. *Proc Natl Acad Sci U S A* 90: 4171–4175
- Du T, Zhu G, Wu X, Fang J & Zhou E-M (2019) Biotinylated single-domain antibody-based blocking ELISA for detection of antibodies against swine Influenza virus. *Int J Nanomedicine* 14: 9337–9349
- Duhoo Y, Roche J, Trinh TTN, Desmyter A, Gaubert A, Kellenberger C, Cambillau C, Roussel A & Leone P (2017) Camelid nanobodies used as crystallization chaperones for different constructs of PorM, a component of the type IX secretion system from *Porphyromonas gingivalis*. *Acta Crystallogr F Struct Biol Commun* 73: 286–293
- Dumoulin M, Conrath K, Van Meirhaeghe A, Meersman F, Heremans K, Frenken LGJ, Muyldermans S, Wyns L & Matagne A (2002) Single-domain antibody fragments with high conformational stability. *Protein Sci* 11: 500–515
- 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: 6117–6126
- Eisfeld AJ, Neumann G & Kawaoka Y (2015) At the centre: influenza A virus ribonucleoproteins. *Nat Rev Microbiol* 13: 28–41
- Elleman CJ & Barclay WS (2004) The M1 matrix protein controls the filamentous phenotype of influenza A virus. *Virology* 321: 144–153
- Els Conrath K, Lauwereys M, Wyns L & Muyldermans S (2001) Camel single-domain antibodies as modular building units in bispecific and bivalent antibody constructs. *J Biol Chem* 276: 7346–7350
- Engelhardt OG & Fodor E (2006) Functional association between viral and cellular transcription during influenza virus infection. *Rev Med Virol* 16: 329–345
- Engelhardt OG, Smith M & Fodor E (2005) Association of the Influenza A virus RNA-dependent RNA polymerase with cellular RNA polymerase II. *JVI* 79: 5812–5818
- Ewert S, Cambillau C, Conrath K & Plückthun A (2002) Biophysical properties of camelid V(HH) domains compared to those of human V(H)3 domains. *Biochemistry* 41: 3628–3636
- Fan H, Walker AP, Carrique L, Keown JR, Serna Martin I, Karia D, Sharps J, Hengrung N, Pardon E, Steyaert J, *et al* (2019a) Structures of influenza A virus RNA polymerase offer insight into viral genome replication. *Nature* 573: 287–290
- Fan H, Walker AP, Carrique L, Keown JR, Serna Martin I, Karia D, Sharps J, Hengrung N, Pardon E, Steyaert J, *et al* (2019b) Structures of influenza A virus RNA polymerase offer insight into viral genome replication. *Nature* 573: 287–290

- Fitch WM, Leiter JM, Li XQ & Palese P (1991) Positive Darwinian evolution in human influenza A viruses. *Proceedings of the National Academy of Sciences* 88: 4270–4274
- Fodor E, Pritlove DC & Brownlee GG (1994) The influenza virus panhandle is involved in the initiation of transcription. *J Virol* 68: 4092–4096
- Fodor E, Pritlove DC & Brownlee GG (1995) Characterization of the RNA-fork model of virion RNA in the initiation of transcription in influenza A virus. *J Virol* 69: 4012–4019
- Fodor E, Seong BL & Brownlee GG (1993) Photochemical cross-linking of influenza A polymerase to its virion RNA promoter defines a polymerase binding site at residues 9 to 12 of the promoter. *J Gen Virol* 74 (Pt 7): 1327–1333
- Fodor E & Smith M (2004) The PA subunit is required for efficient nuclear accumulation of the PB1 subunit of the Influenza A virus RNA polymerase complex. *JVI* 78: 9144–9153
- Fournier E, Moules V, Essere B, Paillart J-C, Sirbat J-D, Isel C, Cavalier A, Rolland J-P, Thomas D, Lina B, *et al* (2012) A supramolecular assembly formed by influenza A virus genomic RNA segments. *Nucleic Acids Res* 40: 2197–2209
- Gabriel G, Dauber B, Wolff T, Planz O, Klenk H-D & Stech J (2005) The viral polymerase mediates adaptation of an avian influenza virus to a mammalian host. *Proc Natl Acad Sci U S A* 102: 18590–18595
- Gaiotto T & Hufton SE (2016) Cross-neutralising nanobodies bind to a conserved pocket in the hemagglutinin stem region identified using yeast display and deep mutational scanning. *PLoS One* 11: e0164296
- Gaiotto T, Ramage W, Ball C, Risley P, Carnell GW, Temperton N, Engelhardt OG & Hufton SE (2021) Nanobodies mapped to cross-reactive and divergent epitopes on A(H7N9) influenza hemagglutinin using yeast display. *Sci Rep* 11: 3126
- Garrido-Godino AI, García-López MC, García-Martínez J, Pelechano V, Medina DA, Pérez-Ortín JE & Navarro F (2016) Rpb1 foot mutations demonstrate a major role of Rpb4 in mRNA stability during stress situations in yeast. *Biochim Biophys Acta* 1859: 731–743
- Garten RJ, Davis CT, Russell CA, Shu B, Lindstrom S, Balish A, Sessions WM, Xu X, Skepner E, Deyde V, *et al* (2009) Antigenic and genetic characteristics of swine-origin 2009 A(H1N1) influenza viruses circulating in humans. *Science* 325: 197–201
- Gavazzi C, Isel C, Fournier E, Moules V, Cavalier A, Thomas D, Lina B & Marquet R (2013a) 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
- Gavazzi C, Yver M, Isel C, Smyth RP, Rosa-Calatrava M, Lina B, Moulès V & Marquet R (2013b) A functional sequence-specific interaction between influenza A virus genomic RNA segments. *Proc Natl Acad Sci U S A* 110: 16604–16609
- Gerber M, Isel C, Moules V & Marquet R (2014) Selective packaging of the influenza A genome and consequences for genetic reassortment. *Trends Microbiol* 22: 446–455

- 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: 521–530
- Ghosh A & Lima CD (2010) Enzymology of RNA cap synthesis. *Wiley Interdiscip Rev RNA* 1: 152–172
- 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: E238
- Gottschalk A (1959) The viruses. *The viruses* 3: 51–61
- GRAMM-X Protein-Protein Docking Web Server v.1.2.0
- Groot AJ, Verheesen P, Westerlaken EJ, Gort EH, van der Groep P, Bovenschen N, van der Wall E, van Diest PJ & Shvarts A (2006) Identification by phage display of single-domain antibody fragments specific for the ODD domain in hypoxia-inducible factor 1alpha. *Lab Invest* 86: 345–356
- 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: 2067–2075
- Guellouz A, Valerio-Lepiniec M, Urvoas A, Chevrel A, Graille M, Fourati-Kammoun Z, Desmadril M, van Tilbeurgh H & Minard P (2013) Selection of specific protein binders for pre-defined targets from an optimized library of artificial helicoidal repeat proteins (alphaRep). *PLoS One* 8: e71512
- Guilligay D, Tarendeau F, Resa-Infante P, Coloma R, Crepin T, Sehr P, Lewis J, Ruigrok RWH, Ortin J, Hart DJ, *et al* (2008) The structural basis for cap binding by influenza virus polymerase subunit PB2. *Nat Struct Mol Biol* 15: 500–506
- Hale BG, Randall RE, Ortín J & Jackson D (2008) The multifunctional NS1 protein of influenza A viruses. *J Gen Virol* 89: 2359–2376
- Hamers-Casterman C, Atarhouch T, Muyldermans S, Robinson G, Hamers C, Songa EB, Bendahman N & Hamers R (1993) Naturally occurring antibodies devoid of light chains. *Nature* 363: 446–448
- Hanke L, Knockenhauer KE, Brewer RC, van Diest E, Schmidt FI, Schwartz TU & Ploegh HL (2016) The antiviral mechanism of an Influenza A virus nucleoprotein-specific single-domain antibody fragment. *mBio* 7: e01569-16
- Harter C, James P, Bächli T, Semenza G & Brunner J (1989) Hydrophobic binding of the ectodomain of influenza hemagglutinin to membranes occurs through the 'fusion peptide'. *J Biol Chem* 264: 6459–6464
- He X, Zhou J, Bartlam M, Zhang R, Ma J, Lou Z, Li X, Li J, Joachimiak A, Zeng Z, *et al* (2008) Crystal structure of the polymerase PA(C)-PB1(N) complex from an avian influenza H5N1 virus. *Nature* 454: 1123–1126
- Heinz S, Texari L, Hayes MGB, Urbanowski M, Chang MW, Givarkes N, Rialdi A, White KM, Albrecht RA, Pache L, *et al* (2018) Transcription elongation can affect genome 3D structure. *Cell* 174: 1522-1536.e22

- Hickman HD, Mays JW, Gibbs J, Kosik I, Magadán JG, Takeda K, Das S, Reynoso GV, Ngudiankama BF, Wei J, *et al* (2018) Influenza A virus negative strand RNA is translated for CD8+ T cell immunosurveillance. *J Immunol* 201: 1222–1228
- Hieronimus H, Yu MC & Silver PA (2004) Genome-wide mRNA surveillance is coupled to mRNA export. *Genes Dev* 18: 2652–2662
- Hmila I, Saerens D, Ben Abderrazek R, Vincke C, Abidi N, Benlasfar Z, Govaert J, El Ayeb M, Bouhaouala-Zahar B & Muyldermans S (2010) A bispecific nanobody to provide full protection against lethal scorpion envenoming. *FASEB J* 24: 3479–3489
- Ho CK & Shuman S (1999) Distinct roles for CTD Ser-2 and Ser-5 phosphorylation in the recruitment and allosteric activation of mammalian mRNA capping enzyme. *Mol Cell* 3: 405–411
- Hsia H-P, Yang Y-H, Szeto W-C, Nilsson BE, Lo C-Y, Ng AK-L, Fodor E & Shaw P-C (2018) Amino acid substitutions affecting aspartic acid 605 and valine 606 decrease the interaction strength between the influenza virus RNA polymerase PB2 '627' domain and the viral nucleoprotein. *PLoS One* 13: e0191226
- 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: 8140–8144
- 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. *JVI* 84: 1254–1264
- Hufton SE, Risley P, Ball CR, Major D, Engelhardt OG & Poole S (2014) The breadth of cross sub-type neutralisation activity of a single domain antibody to influenza hemagglutinin can be increased by antibody valency. *PLoS One* 9: e103294
- Hultberg A, Temperton NJ, Rosseels V, Koenders M, Gonzalez-Pajuelo M, Schepens B, Ibañez LI, Vanlandschoot P, Schillemans J, Saunders M, *et al* (2011) Llama-derived single domain antibodies to build multivalent, superpotent and broadened neutralizing anti-viral molecules. *PLoS One* 6: e17665
- Hutchinson EC & Fodor E (2012) Nuclear import of the influenza A virus transcriptional machinery. *Vaccine* 30: 7353–7358
- Hutchinson EC, Orr OE, Man Liu S, Engelhardt OG & Fodor E (2011) Characterization of the interaction between the influenza A virus polymerase subunit PB1 and the host nuclear import factor Ran-binding protein 5. *J Gen Virol* 92: 1859–1869
- Ibañez LI, De Filette M, Hultberg A, Verrips T, Temperton N, Weiss RA, Vandeveldel W, Schepens B, Vanlandschoot P & Saelens X (2011) Nanobodies with in vitro neutralizing activity protect mice against H5N1 influenza virus infection. *J Infect Dis* 203: 1063–1072
- Inglis SC (1982) Inhibition of host protein synthesis and degradation of cellular mRNAs during infection by influenza and herpes simplex virus. *Mol Cell Biol* 2: 1644–1648
- International Committee on Taxonomy of Viruses (ICTV) (2020) Virus Taxonomy : 2020 Release

- Iuliano AD, Roguski KM, Chang HH, Muscatello DJ, Palekar R, Tempia S, Cohen C, Gran JM, Schanzer D, Cowling BJ, *et al* (2018) Estimates of global seasonal influenza-associated respiratory mortality: a modelling study. *Lancet* 391: 1285–1300
- Izaurrealde E, Lewis J, McGuigan C, Jankowska M, Darzynkiewicz E & Mattaj IW (1994) A nuclear cap binding protein complex involved in pre-mRNA splicing. *Cell* 78: 657–668
- Jagger BW, Wise HM, Kash JC, Walters K-A, Wills NM, Xiao Y-L, Dunfee RL, Schwartzman LM, Ozinsky A, Bell GL, *et al* (2012) An overlapping protein-coding region in influenza A virus segment 3 modulates the host response. *Science* 337: 199–204
- Kajava AV (2001) Review: proteins with repeated sequence--structural prediction and modeling. *J Struct Biol* 134: 132–144
- Katze MG & Krug RM (1984) Metabolism and expression of RNA polymerase II transcripts in influenza virus-infected cells. *Mol Cell Biol* 4: 2198–2206
- Kemdirim S, Palefsky J & Briedis DJ (1986) Influenza B virus PB1 protein; nucleotide sequence of the genome RNA segment predicts a high degree of structural homology with the corresponding influenza A virus polymerase protein. *Virology* 152: 126–135
- Keown JR, Zhu Z, Carrique L, Fan H, Walker AP, Serna Martin I, Pardon E, Steyaert J, Fodor E & Grimes JM (2022) Mapping inhibitory sites on the RNA polymerase of the 1918 pandemic influenza virus using nanobodies. *Nat Commun* 13: 251
- Klooster R, Eman MR, le Duc Q, Verheesen P, Verrips CT, Roovers RC & Post JA (2009) Selection and characterization of KDEL-specific VHH antibody fragments and their application in the study of ER resident protein expression. *J Immunol Methods* 342: 1–12
- Klumpp K, Ruigrok RW & Baudin F (1997) Roles of the influenza virus polymerase and nucleoprotein in forming a functional RNP structure. *EMBO J* 16: 1248–1257
- Koelle K, Cobey S, Grenfell B & Pascual M (2006) Epochal evolution shapes the phylodynamics of inter-pandemic Influenza A (H3N2) in humans. *Science* 314: 1898–1903
- Kouba T, Drncová P & Cusack S (2019) Structural snapshots of actively transcribing influenza polymerase. *Nat Struct Mol Biol* 26: 460–470
- Krammer F, Smith GJD, Fouchier RAM, Peiris M, Kedzierska K, Doherty PC, Palese P, Shaw ML, Treanor J, Webster RG, *et al* (2018) Influenza. *Nat Rev Dis Primers* 4: 3
- Krischuns T, Lukarska M, Naffakh N & Cusack S (2021) Influenza virus RNA-dependent RNA polymerase and the host transcriptional apparatus. *Annu Rev Biochem* 90: 321–348
- Krissinel E & Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372: 774–797
- Kuchipudi SV & Nissly RH (2018) Novel Flu Viruses in Bats and Cattle: ‘Pushing the Envelope’ of Influenza Infection. *Vet Sci* 5: E71
- Lamb RA & Choppin PW (1977) Synthesis of influenza virus polypeptides in cells resistant to alpha-amanitin: evidence for the involvement of cellular RNA polymerase II in virus replication. *Journal of Virology* 23: 816–819

- Lamb RA & Choppin PW (1979) Segment 8 of the influenza virus genome is unique in coding for two polypeptides. *Proc Natl Acad Sci U S A* 76: 4908–4912
- Laursen NS, Friesen RHE, Zhu X, Jongeneelen M, Blokland S, Vermond J, van Eijgen A, Tang C, van Diepen H, Obmolova G, *et al* (2018) Universal protection against influenza infection by a multidomain antibody to influenza hemagglutinin. *Science* 362: 598–602
- Lauwereys M, Arbabi Ghahroudi M, Desmyter A, Kinne J, Hölzer W, De Genst E, Wyns L & Muyldermans S (1998) Potent enzyme inhibitors derived from dromedary heavy-chain antibodies. *EMBO J* 17: 3512–3520
- Le Sage V, Nanni AV, Bhagwat AR, Snyder DJ, Cooper VS, Lakdawala SS & Lee N (2018) Non-uniform and non-random binding of nucleoprotein to Influenza A and B viral RNA. *Viruses* 10: E522
- Ledeboer AM, Bezemer S, de Hiaard JJW, Schaffers IM, Verrips CT, van Vliet C, Düsterhöft EM, Zoon P, Moineau S & Frenken LGJ (2002) Preventing phage lysis of *Lactococcus lactis* in cheese production using a neutralizing heavy-chain antibody fragment from llama. *J Dairy Sci* 85: 1376–1382
- Lee N, Le Sage V, Nanni AV, Snyder DJ, Cooper VS & Lakdawala SS (2017) Genome-wide analysis of influenza viral RNA and nucleoprotein association. *Nucleic Acids Res* 45: 8968–8977
- Levene RE & Gaglia MM (2018) Host shutoff in Influenza A virus: many means to an end. *Viruses* 10: E475
- Li B, Clohisey SM, Chia BS, Wang B, Cui A, Eisenhaure T, Schweitzer LD, Hoover P, Parkinson NJ, Nachshon A, *et al* (2020) Genome-wide CRISPR screen identifies host dependency factors for influenza A virus infection. *Nat Commun* 11: 164
- Li H, Wu Y, Li M, Guo L, Wang Q, Lai Z, Zhang J, Zhang X, Zhu L, Lan P, *et al* (2021a) A resident state allows Influenza polymerase to smoothly switch between transcription and replication cycles. *bioRxiv*: 2021.11.30.470684
- Li X, Gu M, Zheng Q, Gao R & Liu X (2021b) Packaging signal of influenza A virus. *Virology* 18: 36
- van der Linden RH, Frenken LG, de Geus B, Harmsen MM, Ruuls RC, Stok W, de Ron L, Wilson S, Davis P & Verrips CT (1999) Comparison of physical chemical properties of llama VHH antibody fragments and mouse monoclonal antibodies. *Biochim Biophys Acta* 1431: 37–46
- Long JS, Giotis ES, Moncorgé O, Frise R, Mistry B, James J, Morisson M, Iqbal M, Vignal A, Skinner MA, *et al* (2016) Species difference in ANP32A underlies influenza A virus polymerase host restriction. *Nature* 529: 101–104
- Lowen AC (2017) Constraints, drivers, and implications of Influenza A virus reassortment. *Annu Rev Virol* 4: 105–121
- Lowen AC (2018) It's in the mix: Reassortment of segmented viral genomes. *PLoS Pathog* 14: e1007200
- Lukarska M, Fournier G, Pflug A, Resa-Infante P, Reich S, Naffakh N & Cusack S (2017) Structural basis of an essential interaction between influenza polymerase and Pol II CTD. *Nature* 541: 117–121

- Lyskov S, Chou F-C, Conchúir SÓ, Der BS, Drew K, Kuroda D, Xu J, Weitzner BD, Renfrew PD, Sripakdeevong P, *et al* (2013) Serverification of molecular modeling applications: the Rosetta Online Server that Includes Everyone (ROSIE). *PLoS One* 8: e63906
- Ma K, Roy AM & Whittaker GR (2001) Nuclear export of influenza virus ribonucleoproteins: identification of an export intermediate at the nuclear periphery. *Virology* 282: 215–220
- Marcos-Villar L, Pazo A & Nieto A (2016) Influenza virus and chromatin: role of the CHD1 chromatin remodeler in the virus life cycle. *J Virol* 90: 3694–3707
- Marcotrigiano J, Gingras AC, Sonenberg N & Burley SK (1997) Cocrystal structure of the messenger RNA 5' cap-binding protein (eIF4E) bound to 7-methyl-GDP. *Cell* 89: 951–961
- 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: 6014–6021
- 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: 601–613
- Mazza C, Ohno M, Segref A, Mattaj IW & Cusack S (2001) Crystal structure of the human nuclear cap binding complex. *Mol Cell* 8: 383–396
- Mazza C, Segref A, Mattaj IW & Cusack S (2002) Large-scale induced fit recognition of an m(7)GpppG cap analogue by the human nuclear cap-binding complex. *EMBO J* 21: 5548–5557
- 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: 3595–3605
- McGeoch D, Fellner P & Newton C (1976) Influenza virus genome consists of eight distinct RNA species. *Proc Natl Acad Sci U S A* 73: 3045–3049
- Meka H (2005) Crystal structure and RNA binding of the Rpb4/Rpb7 subunits of human RNA polymerase II. *Nucleic Acids Research* 33: 6435–6444
- Microsoft Corporation (2018) Microsoft Excel
- Miner JC, Lappala A, Fenimore PW, Fischer WM, McMahon BH, Hengartner NW, Sanbonmatsu KY & Tung C-S (2021) Modeling the Influenza A NP-vRNA-polymerase complex in atomic detail. *Biomolecules* 11: 124
- Miyake Y, Keusch JJ, Decamps L, Ho-Xuan H, Iketani S, Gut H, Kutay U, Helenius A & Yamauchi Y (2019) Influenza virus uses transportin 1 for vRNP debundling during cell entry. *Nat Microbiol* 4: 578–586
- Moeller A, Kirchdoerfer RN, Potter CS, Carragher B & Wilson IA (2012) Organization of the influenza virus replication machinery. *Science* 338: 1631–1634
- Momose F, Kikuchi Y, Komase K & Morikawa Y (2007) Visualization of microtubule-mediated transport of influenza viral progeny ribonucleoprotein. *Microbes Infect* 9: 1422–1433

- Moreira ÉA, Weber A, Bolte H, Kolesnikova L, Giese S, Lakdawala S, Beer M, Zimmer G, García-Sastre A, Schwemmle M, *et al* (2016) A conserved influenza A virus nucleoprotein code controls specific viral genome packaging. *Nat Commun* 7: 12861
- Morens DM, Taubenberger JK & Fauci AS (2009) The persistent legacy of the 1918 influenza virus. *N Engl J Med* 361: 225–229
- Moulès V, Terrier O, Yver M, Riteau B, Moriscot C, Ferraris O, Julien T, Giudice E, Rolland J-P, Erny A, *et al* (2011) Importance of viral genomic composition in modulating glycoprotein content on the surface of influenza virus particles. *Virology* 414: 51–62
- Moutel S, Bery N, Bernard V, Keller L, Lemesre E, de Marco A, Ligat L, Rain J-C, Favre G, Olichon A, *et al* (2016) NaLi-H1: A universal synthetic library of humanized nanobodies providing highly functional antibodies and intrabodies. *Elife* 5: e16228
- Müller R, Poch O, Delarue M, Bishop DH & Bouloy M (1994) Rift Valley fever virus L segment: correction of the sequence and possible functional role of newly identified regions conserved in RNA-dependent polymerases. *J Gen Virol* 75 (Pt 6): 1345–1352
- Munier S, Rolland T, Diot C, Jacob Y & Naffakh N (2013) Exploration of binary virus–host interactions using an infectious protein complementation assay. *Mol Cell Proteomics* 12: 2845–2855
- 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: 2455–2462
- Muyldermans S, Atarhouch T, Saldanha J, Barbosa JA & Hamers R (1994) Sequence and structure of VH domain from naturally occurring camel heavy chain immunoglobulins lacking light chains. *Protein Eng* 7: 1129–1135
- Naffakh N, Tomoiu A, Rameix-Welti M-A & van der Werf S (2008) Host restriction of avian influenza viruses at the level of the ribonucleoproteins. *Annu Rev Microbiol* 62: 403–424
- Naito T, Momose F, Kawaguchi A & Nagata K (2007) Involvement of Hsp90 in assembly and nuclear import of influenza virus RNA polymerase subunits. *J Virol* 81: 1339–1349
- Nakajima N, Hata S, Sato Y, Tobiome M, Katano H, Kaneko K, Nagata N, Kataoka M, Ainai A, Hasegawa H, *et al* (2010) The first autopsy case of pandemic influenza (A/H1N1pdm) virus infection in Japan: detection of a high copy number of the virus in type II alveolar epithelial cells by pathological and virological examination. *Jpn J Infect Dis* 63: 67–71
- Nakatsu S, Sagara H, Sakai-Tagawa Y, Sugaya N, Noda T & Kawaoka Y (2016) Complete and incomplete genome packaging of Influenza A and B viruses. *mBio* 7: e01248-16
- Nemeroff ME, Barabino SM, Li Y, Keller W & Krug RM (1998) Influenza virus NS1 protein interacts with the cellular 30 kDa subunit of CPSF and inhibits 3' end formation of cellular pre-mRNAs. *Mol Cell* 1: 991–1000
- Neumann G, Hughes MT & Kawaoka Y (2000) Influenza A virus NS2 protein mediates vRNP nuclear export through NES-independent interaction with hCRM1. *EMBO J* 19: 6751–6758
- Ng AK-L, Chan W-H, Choi S-T, Lam MK-H, Lau K-F, Chan PK-S, Au SW-N, Fodor E & Shaw P-C (2012) Influenza polymerase activity correlates with the strength of interaction between nucleoprotein and PB2 through the host-specific residue K/E627. *PLoS One* 7: e36415

- Nguyen VK, Hamers R, Wyns L & Muyldermans S (1999) Loss of splice consensus signal is responsible for the removal of the entire C(H)1 domain of the functional camel IGG2A heavy-chain antibodies. *Mol Immunol* 36: 515–524
- Nilsson-Payant BE, Sharps J, Hengrung N & Fodor E (2018) The surface-exposed PA51-72-loop of the Influenza A virus polymerase is required for viral genome replication. *J Virol* 92: e00687-18
- Noda T, Sagara H, Yen A, Takada A, Kida H, Cheng RH & Kawaoka Y (2006) Architecture of ribonucleoprotein complexes in influenza A virus particles. *Nature* 439: 490–492
- Noda T, Sugita Y, Aoyama K, Hirase A, Kawakami E, Miyazawa A, Sagara H & Kawaoka Y (2012) Three-dimensional analysis of ribonucleoprotein complexes in influenza A virus. *Nat Commun* 3: 639
- Nojima T, Gomes T, Grosso ARF, Kimura H, Dye MJ, Dhir S, Carmo-Fonseca M & Proudfoot NJ (2015) Mammalian NET-seq reveals genome-wide nascent transcription coupled to RNA processing. *Cell* 161: 526–540
- 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: 2280–2290
- 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: 47–60
- Obayashi E, Yoshida H, Kawai F, Shibayama N, Kawaguchi A, Nagata K, Tame JRH & Park S-Y (2008) The structural basis for an essential subunit interaction in influenza virus RNA polymerase. *Nature* 454: 1127–1131
- Odagiri T & Tashiro M (1997) 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* 71: 2138–2145
- O’Neill RE, Jaskunas R, Blobel G, Palese P & Moroianu J (1995) Nuclear import of influenza virus RNA can be mediated by viral nucleoprotein and transport factors required for protein import. *J Biol Chem* 270: 22701–22704
- O’Neill RE, Talon J & Palese P (1998) The influenza virus NEP (NS2 protein) mediates the nuclear export of viral ribonucleoproteins. *EMBO J* 17: 288–296
- OuYang B, Dong Y & Chou JJ (2018) Structural and functional properties of viral membrane proteins. In *Advances in Membrane Proteins: Part I: Mass Processing and Transportation*, Cao Y (ed) pp 147–181. Singapore: Springer Singapore
- Oymans J & Te Velthuis AJW (2018) A mechanism for priming and realignment during Influenza A virus replication. *J Virol* 92: e01773-17
- Palese P & Compans RW (1976) Inhibition of influenza virus replication in tissue culture by 2-deoxy-2,3-dehydro-N-trifluoroacetylneuraminic acid (FANA): mechanism of action. *J Gen Virol* 33: 159–163
- Palese P & Shaw ML (2007) Orthomyxoviridae: the viruses and their replication. In *Fields virology*, 5th ed. pp 1647–1689. Philadelphia : Wolters Kluwer Health/Lippincott Williams & Wilkins

- Palese P, Tobita K, Ueda M & Compans RW (1974) Characterization of temperature sensitive influenza virus mutants defective in neuraminidase. *Virology* 61: 397–410
- Pancer Z, Amemiya CT, Ehrhardt GRA, Ceitlin J, Gartland GL & Cooper MD (2004) Somatic diversification of variable lymphocyte receptors in the agnathan sea lamprey. *Nature* 430: 174–180
- Pancer Z & Cooper MD (2006) The evolution of adaptive immunity. *Annu Rev Immunol* 24: 497–518
- Pancer Z, Saha NR, Kasamatsu J, Suzuki T, Amemiya CT, Kasahara M & Cooper MD (2005) Variable lymphocyte receptors in hagfish. *Proc Natl Acad Sci U S A* 102: 9224–9229
- PDBePISA (Proteins, Interfaces, Structures and Assemblies)
- Pérez JM, Renisio JG, Prompers JJ, van Platerink CJ, Cambillau C, Darbon H & Frenken LG (2001) Thermal unfolding of a llama antibody fragment: a two-state reversible process. *Biochemistry* 40: 74–83
- Pflug A, Gaudon S, Resa-Infante P, Lethier M, Reich S, Schulze WM & Cusack S (2018) Capped RNA primer binding to influenza polymerase and implications for the mechanism of cap-binding inhibitors. *Nucleic Acids Research* 46: 956–971
- Pflug A, Guilligay D, Reich S & Cusack S (2014) Structure of influenza A polymerase bound to the viral RNA promoter. *Nature* 516: 355–360
- Pflug A, Lukarska M, Resa-Infante P, Reich S & Cusack S (2017) Structural insights into RNA synthesis by the influenza virus transcription-replication machine. *Virus Research* 234: 103–117
- Plotch SJ, Bouloy M, Ulmanen I & Krug RM (1981) A unique cap(m7GpppXm)-dependent influenza virion endonuclease cleaves capped RNAs to generate the primers that initiate viral RNA transcription. *Cell* 23: 847–858
- Plotch SJ, Tomasz J & Krug RM (1978) Absence of detectable capping and methylating enzymes in influenza virions. *J Virol* 28: 75–83
- Poch O, Sauvaget I, Delarue M & Tordo N (1989) Identification of four conserved motifs among the RNA-dependent polymerase encoding elements. *EMBO J* 8: 3867–3874
- Poole E, Elton D, Medcalf L & Digard P (2004) Functional domains of the influenza A virus PB2 protein: identification of NP- and PB1-binding sites. *Virology* 321: 120–133
- 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: 3473–3476
- Pritlove DC, Fodor E, Seong BL & Brownlee GG (1995) In vitro transcription and polymerase binding studies of the termini of influenza A virus cRNA: evidence for a cRNA panhandle. *J Gen Virol* 76 (Pt 9): 2205–2213
- Putri WCWS, Muscatello DJ, Stockwell MS & Newall AT (2018) Economic burden of seasonal influenza in the United States. *Vaccine* 36: 3960–3966
- Raju R, Raju L, Hacker D, Garcin D, Compans R & Kolakofsky D (1990) Nontemplated bases at the 5' ends of Tacaribe virus mRNAs. *Virology* 174: 53–59

- Ramage W, Gaiotto T, Ball C, Risley P, Carnell GW, Temperton N, Cheung CY, Engelhardt OG & Hufton SE (2019) Cross-reactive and lineage-specific single domain antibodies against Influenza B hemagglutinin. *Antibodies (Basel)* 8: E14
- Rasmussen EB & Lis JT (1993) In vivo transcriptional pausing and cap formation on three *Drosophila* heat shock genes. *Proc Natl Acad Sci U S A* 90: 7923–7927
- Rehwinkel J, Tan CP, Goubau D, Schulz O, Pichlmair A, Bier K, Robb N, Vreede F, Barclay W, Fodor E, *et al* (2010) RIG-I detects viral genomic RNA during negative-strand RNA virus infection. *Cell* 140: 397–408
- Reich S, Guilligay D & Cusack S (2017) An in vitro fluorescence based study of initiation of RNA synthesis by influenza B polymerase. *Nucleic Acids Res* 45: 3353–3368
- Reich S, Guilligay D, Pflug A, Malet H, Berger I, Crépin T, Hart D, Lunardi T, Nanao M, Ruigrok RWH, *et al* (2014) Structural insight into cap-snatching and RNA synthesis by influenza polymerase. *Nature* 516: 361–366
- Resa-Infante P, Jorba N, Zamarreño N, Fernández Y, Juárez S & Ortín J (2008) The host-dependent interaction of alpha-importins with influenza PB2 polymerase subunit is required for virus RNA replication. *PLoS One* 3: e3904
- Rialdi A, Hultquist J, Jimenez-Morales D, Peralta Z, Campisi L, Fenouil R, Moshkina N, Wang ZZ, Laffleur B, Kaake RM, *et al* (2017) The RNA exosome syncs IAV-RNAPII transcription to promote viral ribogenesis and infectivity. *Cell* 169: 679-692.e14
- Richard S, Gross L, Fischer J, Bendalak K, Ziv T, Urim S & Choder M (2021) Numerous post-translational modifications of RNA polymerase II subunit Rpb4/7 link transcription to post-transcriptional mechanisms. *Cell Rep* 34: 108578
- Robb NC, Te Velthuis AJW, Fodor E & Kapanidis AN (2019) Real-time analysis of single influenza virus replication complexes reveals large promoter-dependent differences in initiation dynamics. *Nucleic Acids Res* 47: 6466–6477
- Robert X & Gouet P (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* 42: W320–W324
- Robertson JS (1979) 5' and 3' terminal nucleotide sequences of the RNA genome segments of influenza virus. *Nucleic Acids Res* 6: 3745–3757
- Robertson JS, Schubert M & Lazzarini RA (1981) Polyadenylation sites for influenza virus mRNA. *J Virol* 38: 157–163
- Rodriguez A, Pérez-González A & Nieto A (2007) Influenza virus infection causes specific degradation of the largest subunit of cellular RNA polymerase II. *J Virol* 81: 5315–5324
- Rosário-Ferreira N, Preto AJ, Melo R, Moreira IS & Brito RMM (2020) The Central Role of Non-Structural Protein 1 (NS1) in Influenza Biology and Infection. *Int J Mol Sci* 21: E1511
- Rosetta online server
- Rossmann JS, Jing X, Leser GP & Lamb RA (2010) Influenza virus M2 protein mediates ESCRT-independent membrane scission. *Cell* 142: 902–913

- Rossman JS & Lamb RA (2011) Influenza virus assembly and budding. *Virology* 411: 229–236
- Rossman JS, Leser GP & Lamb RA (2012) Filamentous influenza virus enters cells via macropinocytosis. *J Virol* 86: 10950–10960
- Ruigrok RW, Crépin T, Hart DJ & Cusack S (2010) Towards an atomic resolution understanding of the influenza virus replication machinery. *Current Opinion in Structural Biology* 20: 104–113
- Russell CJ (2021) Hemagglutinin Stability and Its Impact on Influenza A Virus Infectivity, Pathogenicity, and Transmissibility in Avians, Mice, Swine, Seals, Ferrets, and Humans. *Viruses* 13: 746
- Sakaguchi A, Hirayama E, Hiraki A, Ishida Y & Kim J (2003) Nuclear export of influenza viral ribonucleoprotein is temperature-dependently inhibited by dissociation of viral matrix protein. *Virology* 306: 244–253
- Schmidt FI, Hanke L, Morin B, Brewer R, Brusica V, Whelan SPJ & Ploegh HL (2016) Phenotypic lentivirus screens to identify functional single domain antibodies. *Nat Microbiol* 1: 16080
- Scholtissek C & Becht H (1971) Binding of ribonucleic acids to the RNP-antigen protein of influenza viruses. *J Gen Virol* 10: 11–16
- Schrödinger, LLC (2015) The PyMOL Molecular Graphics System, Version 1.8.
- 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: e42
- Serna Martin I, Hengrung N, Renner M, Sharps J, Martínez-Alonso M, Masiulis S, Grimes JM & Fodor E (2018) A mechanism for the activation of the Influenza virus transcriptase. *Molecular Cell* 70: 1101-1110.e4
- 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: 764–773
- Shen Y, Yue F, McCleary DF, Ye Z, Edsall L, Kuan S, Wagner U, Dixon J, Lee L, Lobanenkov VV, *et al* (2012) A map of the cis-regulatory sequences in the mouse genome. *Nature* 488: 116–120
- Shimizu K, Iguchi A, Gomyou R & Ono Y (1999) Influenza virus inhibits cleavage of the HSP70 pre-mRNAs at the polyadenylation site. *Virology* 254: 213–219
- Sikora D, Rocheleau L, Brown EG & Pelchat M (2014) Deep sequencing reveals the eight facets of the influenza A/HongKong/1/1968 (H3N2) virus cap-snatching process. *Sci Rep* 4: 6181
- Simons JF & Pettersson RF (1991) Host-derived 5' ends and overlapping complementary 3' ends of the two mRNAs transcribed from the ambisense S segment of Uukuniemi virus. *J Virol* 65: 4741–4748
- Skehel JJ, Bayley PM, Brown EB, Martin SR, Waterfield MD, White JM, Wilson IA & Wiley DC (1982) Changes in the conformation of influenza virus hemagglutinin at the pH optimum of virus-mediated membrane fusion. *Proc Natl Acad Sci U S A* 79: 968–972

- Smith GJD, Vijaykrishna D, Bahl J, Lycett SJ, Worobey M, Pybus OG, Ma SK, Cheung CL, Raghwani J, Bhatt S, *et al* (2009) Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. *Nature* 459: 1122–1125
- Spåhr H, Calero G, Bushnell DA & Kornberg RD (2009) Schizosaccharomyces pombe RNA polymerase II at 3.6-Å resolution. *Proc Natl Acad Sci U S A* 106: 9185–9190
- Staller E, Sheppard CM, Neasham PJ, Mistry B, Peacock TP, Goldhill DH, Long JS & Barclay WS (2019) ANP32 proteins are essential for Influenza virus replication in human cells. *J Virol* 93: e00217-19
- Steel J & Lowen AC (2014) Influenza A virus reassortment. *Curr Top Microbiol Immunol* 385: 377–401
- Stegmann T, White JM & Helenius A (1990) Intermediates in influenza induced membrane fusion. *EMBO J* 9: 4231–4241
- Stewart CS, MacKenzie CR & Hall JC (2007) Isolation, characterization and pentamerization of alpha-cobrotoxin specific single-domain antibodies from a naïve phage display library: preliminary findings for antivenom development. *Toxicon* 49: 699–709
- Stone E, Hiramata T, Tanha J, Tong-Sevinc H, Li S, MacKenzie CR & Zhang J (2007) The assembly of single domain antibodies into bispecific decavalent molecules. *J Immunol Methods* 318: 88–94
- Sugita Y, Sagara H, Noda T & Kawaoka Y (2013) Configuration of viral ribonucleoprotein complexes within the influenza A virion. *J Virol* 87: 12879–12884
- Sugiyama K, Kawaguchi A, Okuwaki M & Nagata K (2015) pp32 and APRIL are host cell-derived regulators of influenza virus RNA synthesis from cRNA. *Elife* 4: e08939
- Swale C, Da Costa B, Sedano L, Garzoni F, McCarthy AA, Berger I, Bieniossek C, Ruigrok RWH, Delmas B & Crépin T (2020) X-ray structure of the human karyopherin RanBP5, an essential factor for Influenza polymerase nuclear trafficking. *Journal of Molecular Biology*: S0022283620302540
- Swale C, Monod A, Tengo L, Labaronne A, Garzoni F, Bourhis J-M, Cusack S, Schoehn G, Berger I, Ruigrok RW, *et al* (2016) Structural characterization of recombinant IAV polymerase reveals a stable complex between viral PA-PB1 heterodimer and host RanBP5. *Sci Rep* 6: 24727
- Tafforeau L, Chantier T, Pradezynski F, Pellet J, Mangeot PE, Vidalain P-O, Andre P, Roubourdin-Combe C & Lotteau V (2011) Generation and comprehensive analysis of an Influenza virus polymerase cellular interaction network. *Journal of Virology* 85: 13010–13018
- Takizawa N, Watanabe K, Nouno K, Kobayashi N & Nagata K (2006) Association of functional influenza viral proteins and RNAs with nuclear chromatin and sub-chromatin structure. *Microbes Infect* 8: 823–833
- Tarendeau F, Boudet J, Guilligay D, Mas PJ, Bougault CM, Boulo S, Baudin F, Ruigrok RWH, Daigle N, Ellenberg J, *et al* (2007) Structure and nuclear import function of the C-terminal domain of influenza virus polymerase PB2 subunit. *Nat Struct Mol Biol* 14: 229–233
- Tarendeau F, Crépin T, Guilligay D, Ruigrok RWH, 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

- Taubenberger JK, Kash JC & Morens DM (2019) The 1918 influenza pandemic: 100 years of questions answered and unanswered. *Sci Transl Med* 11: eaau5485
- Te Velthuis AJW & Fodor E (2016) Influenza virus RNA polymerase: insights into the mechanisms of viral RNA synthesis. *Nat Rev Microbiol* 14: 479–493
- Te Velthuis AJW, Grimes JM & Fodor E (2021) Structural insights into RNA polymerases of negative-sense RNA viruses. *Nat Rev Microbiol* 19: 303–318
- Te Velthuis AJW, Robb NC, Kapanidis AN & Fodor E (2016) The role of the priming loop in Influenza A virus RNA synthesis. *Nat Microbiol* 1: 16029
- Terrier O, Carron C, De Chasse B, Dubois J, Traversier A, Julien T, Cartet G, Proust A, Hacot S, Ressenkoff D, *et al* (2016) Nucleolin interacts with influenza A nucleoprotein and contributes to viral ribonucleoprotein complexes nuclear trafficking and efficient influenza viral replication. *Sci Rep* 6: 29006
- Thierry E, Guilligay D, Kosinski J, Bock T, Gaudon S, Round A, Pflug A, Hengrung N, El Omari K, Baudin F, *et al* (2016) Influenza polymerase can adopt an alternative configuration involving a radical repacking of PB2 domains. *Molecular Cell* 61: 125–137
- Tong S, Li Y, Rivallier P, Conrardy C, Castillo DAA, Chen L-M, Recuenco S, Ellison JA, Davis CT, York IA, *et al* (2012) A distinct lineage of influenza A virus from bats. *Proc Natl Acad Sci U S A* 109: 4269–4274
- Tong S, Zhu X, Li Y, Shi M, Zhang J, Bourgeois M, Yang H, Chen X, Recuenco S, Gomez J, *et al* (2013) New world bats harbor diverse influenza A viruses. *PLoS Pathog* 9: e1003657
- Tovchigrechko A & Vakser IA (2006) GRAMM-X public web server for protein–protein docking. *Nucleic Acids Research* 34: W310–W314
- Transue TR, De Genst E, Ghahroudi MA, Wyns L & Muyldermans S (1998) Camel single-domain antibody inhibits enzyme by mimicking carbohydrate substrate. *Proteins* 32: 515–522
- Treanor J (2004) Influenza Vaccine — Outmaneuvering Antigenic Shift and Drift. *New England Journal of Medicine* 350: 218–220
- Trowitzsch S, Bieniossek C, Nie Y, Garzoni F & Berger I (2010) New baculovirus expression tools for recombinant protein complex production. *J Struct Biol* 172: 45–54
- Urvoas A, Guellouz A, Valerio-Lepiniec M, Graille M, Durand D, Desravines DC, van Tilbeurgh H, Desmadril M & Minard P (2010) Design, production and molecular structure of a new family of artificial alpha-helical repeat proteins (α Rep) based on thermostable HEAT-like repeats. *Journal of Molecular Biology* 404: 307–327
- 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
- Veler H, Fan H, Keown J, Sharps J, Fournier M, Grimes JM & Fodor E (2022) The C-terminal domains of the PB2 subunit of the influenza A virus RNA polymerase directly interact with cellular GTPase Rab11a. *J Virol*: jvi0197921

- Ver LS, Marcos-Villar L, Landeras-Bueno S, Nieto A & Ortín J (2015) The cellular factor NXP2/MORC3 is a positive regulator of Influenza virus multiplication. *J Virol* 89: 10023–10030
- Viboud C, Alonso WJ & Simonsen L (2006) Influenza in tropical regions. *PLoS Med* 3: e89
- Vijayakrishnan S, Loney C, Jackson D, Suphamungmee W, Rixon FJ & Bhella D (2013) Cryotomography of budding influenza A virus reveals filaments with diverse morphologies that mostly do not bear a genome at their distal end. *PLoS Pathog* 9: e1003413
- ViralZone Alphainfluenzavirus Molecular Biology
- 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: 125–134
- Vreede FT & Fodor E (2010) The role of the influenza virus RNA polymerase in host shut-off. *Virulence* 1: 436–439
- Vreede FT, Jung TE & Brownlee GG (2004) Model suggesting that replication of influenza virus is regulated by stabilization of replicative intermediates. *J Virol* 78: 9568–9572
- Vu KB, Ghahroudi MA, Wyns L & Muyldermans S (1997) Comparison of llama VH sequences from conventional and heavy chain antibodies. *Mol Immunol* 34: 1121–1131
- Walker AP & Fodor E (2019) Interplay between Influenza virus and the host RNA polymerase II transcriptional machinery. *Trends in Microbiology* 27: 398–407
- Wandzik JM, Kouba T, Karuppasamy M, Pflug A, Drncova P, Provaznik J, Azevedo N & Cusack S (2020) A structure-based model for the complete transcription cycle of Influenza polymerase. *Cell*: S0092867420303895
- Wang H & Jiang C (2009) Influenza A virus H5N1 entry into host cells is through clathrin-dependent endocytosis. *Sci China C Life Sci* 52: 464–469
- Wang W, Cui Z-Q, Han H, Zhang Z-P, Wei H-P, Zhou Y-F, Chen Z & Zhang X-E (2008) Imaging and characterizing influenza A virus mRNA transport in living cells. *Nucleic Acids Res* 36: 4913–4928
- Watanabe T, Kawakami E, Shoemaker JE, Lopes TJS, Matsuoka Y, Tomita Y, Kozuka-Hata H, Gorai T, Kuwahara T, Takeda E, *et al* (2014) Influenza virus-host interactome screen as a platform for antiviral drug development. *Cell Host Microbe* 16: 795–805
- Webby RJ & Webster RG (2001) Emergence of influenza A viruses. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences* 356: 1817–1828
- Webster RG (1997) Influenza virus: transmission between species and relevance to emergence of the next human pandemic. *Arch Virol Suppl* 13: 105–113
- Wei G, Meng W, Guo H, Pan W, Liu J, Peng T, Chen L & Chen C-Y (2011) Potent neutralization of influenza A virus by a single-domain antibody blocking M2 ion channel protein. *PLoS One* 6: e28309
- White JM & Wilson IA (1987) Anti-peptide antibodies detect steps in a protein conformational change: low-pH activation of the influenza virus hemagglutinin. *J Cell Biol* 105: 2887–2896

- White MC & Lowen AC (2018) Implications of segment mismatch for influenza A virus evolution. *J Gen Virol* 99: 3–16
- White SK, Ma W, McDaniel CJ, Gray GC & Lednicky JA (2016) Serologic evidence of exposure to influenza D virus among persons with occupational contact with cattle. *J Clin Virol* 81: 31–33
- Wichgers Schreur PJ, van de Water S, Harmsen M, Bermúdez-Méndez E, Drabek D, Grosveld F, Wernike K, Beer M, Aebischer A, Daramola O, *et al* (2020) Multimeric single-domain antibody complexes protect against bunyavirus infections. *Elife* 9: e52716
- Wiley DC & Skehel JJ (1987) The structure and function of the hemagglutinin membrane glycoprotein of influenza virus. *Annu Rev Biochem* 56: 365–394
- Wille M & Holmes EC (2020) The Ecology and evolution of Influenza viruses. *Cold Spring Harb Perspect Med* 10: a038489
- Williams GD, Townsend D, Wylie KM, Kim PJ, Amarasinghe GK, Kutluay SB & Boon ACM (2018) Nucleotide resolution mapping of influenza A virus nucleoprotein-RNA interactions reveals RNA features required for replication. *Nat Commun* 9: 465
- 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: 8021–8031
- Wise HM, Hutchinson EC, Jagger BW, Stuart AD, Kang ZH, Robb N, Schwartzman LM, Kash JC, Fodor E, Firth AE, *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
- Woolven BP, Frenken LG, van der Logt P & Nicholls PJ (1999) The structure of the llama heavy chain constant genes reveals a mechanism for heavy-chain antibody formation. *Immunogenetics* 50: 98–101
- Worch R, Niedzwiecka A, Stepinski J, Mazza C, Jankowska-Anyszka M, Darzynkiewicz E, Cusack S & Stolarski R (2005) Specificity of recognition of mRNA 5' cap by human nuclear cap-binding complex. *RNA* 11: 1355–1363
- World Health Organization (WHO) (1980) A revision of the system of nomenclature for influenza viruses: a WHO Memorandum*. *Bulletin of the World Health Organization* 58: 585–591
- Yamaguchi Y, Takagi T, Wada T, Yano K, Furuya A, Sugimoto S, Hasegawa J & Handa H (1999) NELF, a multisubunit complex containing RD, cooperates with DSIF to repress RNA polymerase II elongation. *Cell* 97: 41–51
- Yamashita M, Krystal M & Palese P (1989) Comparison of the three large polymerase proteins of influenza A, B, and C viruses. *Virology* 171: 458–466
- Yamayoshi S, Watanabe M, Goto H & Kawaoka Y (2016) Identification of a novel viral protein expressed from the PB2 segment of Influenza A virus. *J Virol* 90: 444–456
- York A & Fodor E (2013) Biogenesis, assembly, and export of viral messenger ribonucleoproteins in the influenza A virus infected cell. *RNA Biol* 10: 1274–1282

- 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: E4238-4245
- Yuan P, Bartlam M, Lou Z, Chen S, Zhou J, He X, Lv Z, Ge R, Li X, Deng T, *et al* (2009) Crystal structure of an avian influenza polymerase PA(N) reveals an endonuclease active site. *Nature* 458: 909–913
- Zaborowska J, Egloff S & Murphy S (2016) The pol II CTD: new twists in the tail. *Nat Struct Mol Biol* 23: 771–777
- Zambon MC (1999) Epidemiology and pathogenesis of influenza. *Journal of Antimicrobial Chemotherapy* 44: 3–9
- Zhang J, Li G & Ye X (2010) Cyclin T1/CDK9 interacts with influenza A virus polymerase and facilitates its association with cellular RNA polymerase II. *J Virol* 84: 12619–12627
- Zhang J, Tanha J, Hiramata T, Khieu NH, To R, Tong-Sevinc H, Stone E, Brisson JR & MacKenzie CR (2004) Pentamerization of single-domain antibodies from phage libraries: a novel strategy for the rapid generation of high-avidity antibody reagents. *J Mol Biol* 335: 49–56
- Zhao N, Sebastiano V, Moshkina N, Mena N, Hultquist J, Jimenez-Morales D, Ma Y, Rialdi A, Albrecht R, Fenouil R, *et al* (2018) Influenza virus infection causes global RNAPII termination defects. *Nat Struct Mol Biol* 25: 885–893
- Zhirnov OP (2020) Unique bipolar gene architecture in the RNA genome of Influenza A virus. *Biochemistry (Mosc)* 85: 387–392
- Zhirnov OP, Akulich KA, Lipatova AV & Usachev EV (2017) Negative-sense virion RNA of segment 8 (NS) of influenza A virus is able to translate in vitro a new viral protein. *Dokl Biochem Biophys* 473: 122–127
- Zhirnov OP & Isaeva EI (2019) NSP protein encoded in negative NS RNA strand of Influenza A virus induces cellular immune response in infected animals. *Dokl Biochem Biophys* 486: 201–205
- Zhirnov OP & Klenk HD (2010) [Integration of influenza A virus NSP gene into baculovirus genome and its expression in insect cells]. *Vopr Virusol* 55: 4–8
- Zhirnov OP, Poyarkov SV, Vorob'eva IV, Safonova OA, Malyshev NA & Klenk HD (2007) Segment NS of influenza A virus contains an additional gene NSP in positive-sense orientation. *Dokl Biochem Biophys* 414: 127–133
- Zhong W, Reche PA, Lai C-C, Reinhold B & Reinherz EL (2003) Genome-wide characterization of a viral cytotoxic T lymphocyte epitope repertoire. *J Biol Chem* 278: 45135–45144
- Zhu J, Adli M, Zou JY, Verstappen G, Coyne M, Zhang X, Durham T, Miri M, Deshpande V, De Jager PL, *et al* (2013) Genome-wide chromatin state transitions associated with developmental and environmental cues. *Cell* 152: 642–654
- Zhu M, Gong X, Hu Y, Ou W & Wan Y (2014a) Streptavidin-biotin-based directional double Nanobody sandwich ELISA for clinical rapid and sensitive detection of influenza H5N1. *J Transl Med* 12: 352

Zhu M, Hu Y, Li G, Ou W, Mao P, Xin S & Wan Y (2014b) Combining magnetic nanoparticle with biotinylated nanobodies for rapid and sensitive detection of influenza H3N2. *Nanoscale Res Lett* 9: 528

Zürcher T, Marión RM & Ortín J (2000) Protein synthesis shut-off induced by influenza virus infection is independent of PKR activity. *J Virol* 74: 8781–8784