

UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTÉ

UMR CNRS 7213, Faculté de Pharmacie, ILLKIRCH

THÈSE présentée par

Manuel PIRES

Soutenue le : 17 mars 2017

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

Discipline/ Spécialité : Science du vivant/virologie

**Caractérisation cellulaire de nouveaux
traitements anti-VIH-1 ciblant la
nucléocapside et observation d'une
forme compacte de Gag *in cellulo*.**

THÈSE dirigée par :

Pr MELY Yves
Dr De ROCQUIGNY Hugues

Professeur, Université de Strasbourg
Chargé de Recherches, INSERM

RAPPORTEURS :

Dr BOUAZIZ Serge
Dr CIMARELLI Andrea

Directeur de Recherches, CNRS
Directeur de Recherches, CNRS

AUTRES MEMBRES DU JURY :

Dr BERNACCHI Serena
Dr DELELIS Olivier

Chargé de Recherches, CNRS
Chargé de Recherches, CNRS

INVITEE

Dr REAL Eleonore

Maitre de conférences, Université de Strasbourg

Remerciements

Tout d'abord, je souhaite remercier le Dr BERNACCHI Serena, le Dr CIMARELLI Andrea, le Dr DELELIS Olivier et le Dr BOUAZIZ Serge pour avoir accepté d'analyser et de critiquer mon travail de thèse.

Je remercie mon directeur de thèse, le Pr MELY Yves, pour l'accueil dans son laboratoire et la direction de mes travaux. Merci pour la découverte du salon VIP de l'aéroport d'Amsterdam lors de notre retour d'Italie.

Merci à mon co-directeur, le Dr Hugues de ROCQUIGNY, dont l'expertise aussi bien scientifique que bricolagique m'a permis d'avancer pendant ces trois ans de thèse.

Mes plus sincères remerciements vont au Dr REAL Eléonore. Pendant ces trois dernières années tu as été ma maman scientifique en me suivant d'abord rigoureusement, puis en me surveillant du coin de l'œil pour finalement me faire confiance et me laisser indépendant. Pendant cette thèse tu m'as écouté, recadré, mis sur la bonne voie, défendu et surtout soutenu dans mes choix scientifiques, professionnels et personnels. MERCI pour tout.

Je remercie également le Dr BOUTANT Emmanuel pour les discussions scientifiques mais aussi et surtout pour les rigolades, le soutien psychologique et toutes les discussions des réveille-tôt.

Un merci tout particulier au Dr ZGHEIB Sarwat, ma collègue, qui a le plus supporté mes sautes d'humeurs, subi mes colères sans se laisser marcher dessus. On a quand même bien rigolé et ma dernière année sans toi n'était pas pareille. Je te souhaite pleins de bonnes choses et je souhaite du courage à Hassan et Daniel quand même ;-).

Merci à mes compagnons de lunch et de café : Avisek, Mayeul, Doriane et bien sûr mon ami Redouane. Grâce à vous les repas ont été de vraies pauses avec nos discussions Top-chef, politiques, révolutions, TV, X et j'en passe et des meilleures. Merci.

Un grand merci aux biologistes du labo : Nedal, Hassan, Waseem, Tanveer, Liliana, Faisal, Salah, Halah pour avoir subi aussi ma maniaquerie et mon caractère. Pour les rigolades et les bons moments passés au labo, thank you all.

Lesia, comment ne pas te faire un remerciement spécial. Tout d'abord comme compagnon de projet, puis pour les bons moments passés dans les différents meetings, ainsi

que pour la préparation des molécules et l'écriture des rapports et bien entendu pour ta bonne humeur.

Merci à tous les autres membres du laboratoire encore présents ou déjà partis : Halina, Fred, Nicolas, Pascal, Youri, Guy, Andreas, Andrey, Kamal, Rajhans, Noémie, Marianna, Vasyl, Katya, Iryna, Sasha, Alex ... Je m'excuse pour les oublis mais ce laboratoire est trop grand.

Je remercie aussi le Dr Julien Godet pour son aide et sa participation dans les statistiques de ce manuscrit.

Je tiens aussi à remercier les secrétaires (Ingrid et Marlyse), toujours efficaces quel que soit le problème, mais aussi les membres des autres équipes de l'UMR : Tania, Emilie, Anne-Flo, Marion, Guillaume, Jérôme, Maxime...

J'adresse également mes remerciements à Romain Vauchelles, Ludovic Richert et Denis Dujardin pour l'aide et l'accompagnement en microscopie.

Merci à mes amis Jéromine et Pierre pour les bonnes soirées et les bons moments passés ensemble à Strasbourg.

Merci aux IPI 11 (Thibaut, Xavi, Cyrille, Thanie, Alice-Anne, Georges, Nico, Mihaela, Patricia, Marie-Charlotte, Ghislain et Daphné) pour votre soutien, votre humour et tous les bons comme les mauvais moments des dernières semaines.

Merci spécial aux pompiers de Bucey (Alex, Damien, Stéphane, Lilian, David, Marc, Nico, Steeve et les anciens Gaby, Guy, Daniel) mais aussi à l'équipe 1 d'Erstein (Roby, Thibault, Jerem, Mathieu le franc comtois, Mathieu O, Jo, Johan, Pascal, Hervé, Gino, Choupi, Jean-Yves ...). Sans le savoir ou sans le vouloir, votre soutien a été indispensable pour ma survie psychologique.

Merci à mes potes de toujours sur qui je peux compter pour passer de bons moments : Kevin, Deborah (et Kélyan), Julie, Vincent, Cynthia, Maxime, Coralie, Nourin, Romain L, Romain le grain et Elodie et sans oublier François, Quentin et Johan.

Merci à mes potes de fac pour les meilleurs moments de mes 26 années de scolarité : Besançon : Dondon, Billod, Grands Cheveux Laurent, les jumelles ; Toulouse : Montda, Nath, Nicho, Phung, Pipin, Marianne, Flo ; Poitiers : Marie K/C, Marion, Anne-Liseuh et Benjamin, Julie, Claire et Baptiste, Georgy, David, Marie P, Marie L ...

Merci à mes amis et amies : Léa, Aurélien et la petite puce Adèle, Le Grand, Ludo, Céline et les loulous Paul et Lisa, Gaëtan, Simon, Nanard, Elise, Louliana et ma filleule Lizéa,

Chago et Laure, Fab, Philou, Anne-so et la petite Louise, Michèle, Alex, Rose et mon filleul Jules, Sam et Oph sur qui je peux compter à tout moment pour m'aider à surmonter les difficultés.

Aux personnes qui m'ont soutenu et qui ont cru en moi : Mr et Mme Duède, Geogeo et Manue, Paulette et Daniel, Nelly, Julie et Camille.

Comment ne pas remercier l'ensemble de ma famille pour votre amour et votre confiance en moi : Tonton René, Liliane, Seb, Angéline, Camille, Louise, Thierry (Ourson), Tata Evelyne, Dominique, Vanessa, Christophe, Maelis, Laetitia, Bastien, Fred, Laure, Mathis, Nolan, Tata Valérie, Jérôme, Matthieu, Antoine. En citant la famille je ne peux bien sûr pas oublier mon père, ma mère et Pat qui je l'espère sont fiers de moi.

Un remerciement plus particulier pour ma Tata Sylvie et mon Tonton Didier adorés qui sont derrière moi pour m'aimer, me conseiller, me soutenir, me faire voyager, me supporter sans jamais broncher. Je ne peux pas oublier mes chiasses Paul et Manon, qui sont comme mes frères et sœurs.

Enfin, je pourrais écrire une thèse pour décrire tout ce que tu as fait et que tu fais encore pour moi. Ton soutien est à toute épreuve. Ta fierté, ton sourire, tes papouilles, ton écoute, tes conseils, ton amour sont mon plus gros soutien. Maman pour tout, MERCI, MERCI, MERCI !

« C'est bien mais tu peux encore mieux faire ! »

(Maurice PIREs)

Maman, cette thèse est pour toi !

Sommaire

Table des illustrations
Liste des abréviations	1
Chapitre 1 : Introduction.....	5
I. Comprendre le VIH = Virus de l'Immunodéficience Humaine	5
A. Classification : à quelle famille appartient le VIH ?	5
B. Origine, Évolution Phylogénétique et Historique : d'où vient ce virus découvert il y a 33 ans ?..	5
C. Transmission et Infection du virus : comprendre pour ne pas infecter ou être infecté	9
1. Mode de transmission	9
2. Infection : de la contamination au SIDA	10
D. Épidémiologie : pourquoi l'infection par le VIH est toujours bel et bien une épidémie ?.....	12
II. Comment est structuré le VIH-1 ?	15
A. Photos d'identités de la particule virale immature et mature	15
B. Génome viral : un plan simple pour une structure complexe	16
1. L'ARN viral :	16
2. L'ADN proviral :	21
C. Protéines virales : les matériaux de construction pour faire le virus	22
1. Protéines de structure.....	22
2. Protéines enzymatiques	29
3. Protéines d'enveloppe.....	32
4. Protéines régulatrices	34
5. Protéines auxiliaires	36
III. Comment le VIH se multiplie-t-il ?	39
A. Le cycle réplicatif : se multiplier pour mieux infecter	39
B. Contact entre le virus et la cellule : le baiser mortel	40
C. Entrée et décapsidation : attaquer le « rempart » et rentrer dans la « cour » cellulaire.....	43
D. Transcription inverse : quand l'ARN devient ADN.....	47
E. Import nucléaire : rentrer dans le donjon pour prendre les commandes	50
F. Intégration : insertion du génome viral dans le génome cellulaire	51
G. Expression du génome viral, épissage et export nucléaire : synthétiser et sortir les ARNs du noyau	53
H. Traduction des ARNm viraux : à partir des ARNs, il faut produire les protéines.....	56
I. Assemblage et bourgeonnement : constituer une armée pour accentuer l'invasion.....	57
1. Éléments indispensables à l'assemblage	58
2. Bourgeonnement.....	61

J.	Maturation : s'organiser pour mieux infecter.....	62
IV.	Forme compacte de Gag	64
V.	Comment soigne-t-on le VIH ?	70
A.	Avant de soigner, il faut dépister.....	70
B.	Les thérapies antirétrovirales : les outils pour lutter contre le virus.....	71
1.	Inhibiteurs nucléosidiques de la transcriptase inverse (NRTIs)	72
2.	Inhibiteurs non-nucléosidiques de la transcriptase inverse (NNRTIs)	72
3.	Inhibiteurs de la maturation	72
4.	Inhibiteur de fusion.....	73
5.	Inhibiteur d'entrée.....	73
6.	Inhibiteurs de l'intégrase (INIs)	73
7.	Combinaisons thérapeutiques.....	74
C.	Innovation, quels seront les prochains traitements ?.....	74
1.	À court terme : quelles stratégies pour quelles cibles ?.....	74
2.	À long terme : vers la suppression du réservoir viral et la guérison.....	75
VI.	Pourquoi la NCp7 est une bonne cible thérapeutique ?.....	79
A.	Structure de la nucléocapside.....	79
B.	Fonctions de la NucléoCapside : petit mais costaud.....	81
1.	Liaison des acides nucléiques	81
2.	Activité chaperonne	83
C.	Rôles de la nucléocapside dans le cycle viral : le petit plus qui change tout	84
1.	Dans les phases précoces.....	84
2.	Dans les phases tardives	87
D.	Interactions de la NucléoCapside avec les protéines cellulaires : qui sont les partenaires de cette protéine ?.....	88
E.	Existe-il des molécules antivirales ciblant la NucléoCapside ?.....	92
1.	Éjecteurs de zinc	92
2.	Inhibiteurs non covalents	95
3.	Oligonucléotides, aptamères et peptido-mimétiques	96
F.	THINPAD : un projet européen pour développer une nouvelle molécule antivirale ciblant la NucléoCapside.....	97
1.	Les membres du projet.....	98
2.	Organisation du projet.....	100
	Chapitre 2 : Objectifs de la thèse.....	108
	Chapitre 3 : Matériels et Méthodes.....	112
I.	Matériels	112

A.	Cellules.....	112
B.	Plasmides.....	113
C.	Anticorps primaires et secondaires.....	118
D.	Oligonucléotides.....	118
E.	Enzymes de restrictions.....	120
II.	Biologie cellulaire.....	121
A.	Culture cellulaire.....	121
B.	Transfection cellulaire.....	122
C.	Fixation cellulaire.....	122
D.	Marquages post-fixation.....	123
1.	Marquage nucléaire par du Hoescht 33342.....	123
2.	Marquage du motif tétracystéine par un dérivé de la résorufine (ReAsH).....	123
E.	Préparation des échantillons pour la cytométrie en flux.....	124
III.	Infection par des vecteurs lentiviraux.....	125
A.	Production des vecteurs lentiviraux.....	125
B.	Dosage de l'infection.....	125
1.	Par mesure de l'activité luciférase.....	125
2.	Par Cytométrie en flux.....	126
C.	Test d'infectivité et de cytotoxicité.....	126
1.	Test d'infectivité.....	126
2.	Test de cytotoxicité.....	127
D.	Test d'ajout des molécules à divers temps post-infection.....	128
IV.	Imagerie cellulaire.....	130
A.	Microscopie confocale.....	130
B.	Transfert d'énergie entre molécules fluorescentes (FRET : Florescence Resonance Energy Transfer) par mesure du temps de vie de fluorescence (FLIM : Fluorescence Lifetime Imaging Microscopy).....	131
C.	Micro-injection et suivie de la fluorescence par vidéo-microscopie.....	132
D.	Microscopie à haute résolution.....	133
V.	Biochimie.....	134
A.	Lyse des cellules et dosage protéique par Bradford.....	134
B.	Fractionnement cellulaire.....	134
C.	Immunoprécipitation (IP).....	135
D.	Cross-link.....	136
E.	SDS-PAGE et gel natif.....	137
1.	SDS-PAGE.....	137
2.	Gel natif.....	138

F.	Western blot	139
G.	Marquage des gels au bleu de coomassie	139
VI.	Biologie moléculaire.....	140
A.	Transformation de bactéries compétentes et purification de l'ADN plasmidique.....	140
B.	Mutagenèse.....	141
C.	Digestion par enzyme de restriction	141
VII.	Purification de la protéine Gag	144
VIII.	Spectrométrie de masse (MS).....	146
Chapitre 4 : Résultats et discussions		148
I.	Criblage de molécules antivirales.....	148
A.	Projet européen THINPAD	148
1.	Développement et validation du test PSEUDO	148
2.	Criblage des molécules THINPAD	154
3.	Mécanismes d'action des hits.....	159
B.	Molécules antivirales testées dans le test PSEUDO hors projet THINPAD	166
1.	Effet antiviral de duplex bactéricides/antiviraux.....	166
2.	Effet antiviral des Calixarènes	167
C.	Conclusions et perspectives	169
II.	Forme compacte de Gag	171
A.	Études in cellulo.....	172
1.	Utilisation du FRET-FLIM entre eGFP et ReAsH	172
2.	Approche et utilisation de la split-GFP	176
3.	Observations sur différents mutants.....	187
4.	Approche biochimique de la forme compacte	196
B.	Analyse en spectrométrie de masse par mobilisation ionique	209
C.	Tentatives d'observations de la forme compacte avec d'autres techniques de microscopie....	212
1.	Observation de l'évolution de la fluorescence par vidéo-microscopie.....	212
2.	Étude de la forme compacte et détection de l'ARN spécifique.....	214
3.	Création d'un split-DRONPA : observation de la forme compacte par détection en molécule unique.....	218
D.	Conclusions et perspectives	222
Liste de mes publications et communications.....		228
Références.....		230

Table des illustrations

Figure 1 : Arbre phylogénétique du VIH-1 montrant la grande diversité génétique des VIH-1	7
Figure 2 : Principales dates à retenir sur les découvertes du VIH-1	9
Figure 3 : Évolution naturelle de l'infection par le VIH-1 sur la base de l'évolution de deux marqueurs clés de l'infection.....	12
Figure 4 : Prévalence du VIH-1 chez l'adulte (15-49 ans) dans différentes région du monde en 2014	13
Figure 5 : Cascade de la prise en charge des personnes vivant avec le VIH en comparaison avec l'objectif horizon-2020 90-90-90	14
Figure 6 : Particules virales immature (A et B) et mature (C et D) observées en microscopie électronique.....	16
Figure 7 : Organisation génomique du VIH-1 sous forme d'ARN	17
Figure 8 : Structure secondaire possible de la région 5'-UTR (U5) de l'ARN du VIH-1	18
Figure 9 : La transcription inverse permet d'obtenir l'ADN proviral double brin à partir de l'ARN viral génomique.....	21
Figure 10 : Modèle des structures linéaires et tridimensionnelles hypothétique de la polyprotéine Gag	23
Figure 11 : Organisation membranaire de plusieurs Matrice de HIV-1.....	24
Figure 12 : Chevauchement des sites de liaison préférentiels à l'ARN et au PI(4,5)P2 de la partie N-terminal de la Matrice.....	26
Figure 13 : Représentation schématique de la forme conique empruntée par la capsid virale mature.....	27
Figure 14 : Structure linéaire de la Nucléocapside	28
Figure 15 : Comparaison des précurseurs Gag et Gag-Pol	30
Figure 16 : Cycle réplcatif du VIH-1 dans lequel on peut distinguer les étapes précoces et les étapes tardives	40
Figure 17 : Schéma d'entrée du VIH-1 dans la cellule	42
Figure 18 : Les différents modèles de décapsidation actuellement proposés	46
Figure 19 : Différentes étapes de la transcription inverse.....	49
Figure 20 : Représentation schématique de l'intégration de l'ADN pro viral dans le génome cellulaire	52
Figure 21 : Expression et export des ARNm plus ou moins épissés.....	56
Figure 22 : Modèle de bourgeonnement d'une particule virale de VIH-1.....	62
Figure 23 : Distribution du FRET entre les fluorophores couplés en position N-terminale.....	66
Figure 24 : Modèle dynamique de liaison de Gag à l'ARN et à la membrane cellulaire.....	67
Figure 25 : Modèle possible de traduction et d'assemblage du génome viral non épissé.....	69
Figure 26 : Séquence et structure de la Nucléocapside.....	80
Figure 27 : La séquence de la NCp7 est hautement conservée.....	81
Figure 28 : Différentes étapes du cycle viral dans lesquelles la NC est impliquée et dont l'activité dépend du ratio NC/AN	83
Figure 29 : Modèle d'initiation de la transcription.....	84
Figure 30 : Rôle de la NCp7 dans le second transfert de brin.....	86
Figure 31 : Mécanisme de dimérisation de l'ARN du VIH-1 lors de l'assemblage.....	88
Figure 32 : Localisation et appartenance des différents membres du projet européen THINPAD.....	100

<i>Figure 33 : Diagramme de PERT du projet THINPAD montrant la relation entre les différents blocs de tâches (WP).....</i>	<i>101</i>
<i>Figure 34 : Structures tridimensionnelles servant au criblage in silico de molécules potentiellement inhibitrices.....</i>	<i>102</i>
<i>Figure 35 : Test des molécules sur l'activité de déstabilisation de la NCp7 vis-à-vis de cTAR ADN.....</i>	<i>103</i>
<i>Figure 36 : Test des molécules sur la capacité de NCp7 à se lier aux ANs.....</i>	<i>103</i>
<i>Figure 37 : Comparaison des différents tests cellulaires utilisés pour évaluer l'activité antivirale des inhibiteurs de nucléocapside (NCIs).....</i>	<i>105</i>
<i>Figure 38 : principe du test MTT.....</i>	<i>128</i>
<i>Figure 39 : Plan de plaque pour le test d'ajout des molécules (TOAA).....</i>	<i>129</i>
<i>Figure 40 : Comparaison des différents systèmes de microscopie utilisés dans cette thèse.....</i>	<i>130</i>
<i>Figure 41 : Principe de la co-immunoprécipitation.....</i>	<i>136</i>
<i>Figure 42 : Utilisation de cross-linkers BS3-d0 et d4 pour figer la structure de la protéine.....</i>	<i>137</i>
<i>Figure 43 : Production de pseudo-particules virales (PPs) et utilisation dans le test PSEUDO.....</i>	<i>150</i>
<i>Figure 44 : Infectivité des PPs eGFP et luciférase.....</i>	<i>151</i>
<i>Figure 45 : Détermination du nombre de cellules à déposer par puits pour un test PSEUDO de 48H.....</i>	<i>152</i>
<i>Figure 46 : Influence du DMSO sur la viabilité cellulaire.....</i>	<i>153</i>
<i>Figure 47 : Validation du test PSEUDO par détermination de l'effet antiviral et de la toxicité du WDO-217.....</i>	<i>154</i>
<i>Figure 48 : Application du test PSEUDO à deux molécules issues du criblage.....</i>	<i>155</i>
<i>Figure 49 : Application du test PSEUDO à une molécule interagissant avec la rhodamine.....</i>	<i>156</i>
<i>Figure 50 : Relation structure-activité.....</i>	<i>157</i>
<i>Figure 51 : Bilan des molécules testées lors de la phase d'optimisation de la structure.....</i>	<i>158</i>
<i>Figure 52 : Description et validation du test d'ajout des molécules à différents temps post-infection (TOAA).....</i>	<i>161</i>
<i>Figure 53 : Exemples de détermination du mécanisme d'action de deux molécules par TOAA.....</i>	<i>162</i>
<i>Figure 54 : Suivi de l'assemblage de Gag et GagΔNC par FRET-FLIM.....</i>	<i>164</i>
<i>Figure 55 : Effet inhibiteur de l'assemblage de la molécule ROUGE suivi par FRET-FLIM.....</i>	<i>165</i>
<i>Figure 56 : Analyse du mécanisme d'action des calixarènes dans le test d'ajout des molécules à différents temps post-infection (TOAA).....</i>	<i>169</i>
<i>Figure 57 : Outils utilisés pour l'étude de la forme compacte de Gag en FRET-FLIM.....</i>	<i>173</i>
<i>Figure 58 : Observation au microscope confocal de cellules HeLa exprimant Gag_eGFP_TC ou Gag_eGFP et Gag_TC.....</i>	<i>174</i>
<i>Figure 59 : Etude de la forme compacte de Gag par FRET-FLIM.....</i>	<i>176</i>
<i>Figure 60 : Fragments de la split-GFP en deux ou en trois parties.....</i>	<i>177</i>
<i>Figure 61 : Utilisation de la split-eGFP bipartite pour montrer l'existence d'une forme compacte de Gag in cellulo.....</i>	<i>178</i>
<i>Figure 62 : Insertion de deux « multi-cloning sites » (MCS) entre la MA et la CA pour MCS1 et en C-terminal de Gag pour MCS2.....</i>	<i>179</i>
<i>Figure 63 : Validation du système split-GFP pour l'observation de la forme compacte de Gag.....</i>	<i>181</i>

Figure 64 : Observation par microscopie confocale de la fluorescence GFP	182
Figure 65 : Explication du Plugin ImageJ utilisé pour quantifier la fluorescence membranaire et cytoplasmique	183
Figure 66 : Recombinaison intra- et intermoléculaire de la split-GFP portée par Gag.....	184
Figure 67 : Influence des fragments de la split-GFP sur l'activité de la luciférase. (A) Structure théorique de la luciférase portant les fragments de la split-GFP aux extrémités N- et C-terminales.	186
Figure 68 : Utilisation du système split-GFP tripartite pour l'observation par microscopie confocale de la forme compacte de Gag. Des cellules HeLa ont été transfectées.....	187
Figure 69 : Localisation cellulaire de mutants de Gag-eGFP défectifs pour l'interaction avec l'ARN.....	189
Figure 70 : Répartition membrane vs cytoplasme de la fluorescence obtenue par reconstitution de la split-GFP dans différents mutants de Gag.	191
Figure 71 : Etude du mutant de multimérisation (GagWM).....	193
Figure 72 : Etude de mutants de la région liant les domaines CAN et CAC de CA (Gag Δ link CAC-CAN, GagPPPP) et de la région Sp1 (Gag Δ Sp1).	196
Figure 73 : Analyse par gels natif et dénaturant des formes observées pour diverses protéines Gag.....	201
Figure 74 : Principales étapes de l'étude de la forme compacte de Gag par pontage moléculaire couplé à la spectrométrie de masse.....	203
Figure 75 : Pontage moléculaire de lysats cellulaires exprimant diverses protéines Gag.	206
Figure 76 : Identification des peptides pontés obtenus après traitement au BS3 et immunoprécipitation anti-eGFP	208
Figure 77 : Purification de la protéine Gag et vérification de son activité.....	210
Figure 78 : Observation de protéines Gag purifiées avec des tampons à quantité réduite en NaCl.....	211
Figure 79 : Suivi de la fluorescence lors de la microinjection de plasmides codant Gag_GFP11_GFP1-10 ou un mix Gag_GFP11_MCS2 + Gag_MCS1_GFP1-10	214
Figure 80 : Principe d'observation de l'ARN viral grâce à la reconnaissance de tige-boucles MS2 par la protéine de capsid du bactériophage MS2 fusionnée à une protéine fluorescente.....	215
Figure 81 : Observation simultanée de la forme compacte de Gag et de l'ARN viral.....	217
Figure 82 : Comparaison de la microscopie conventionnelle et de la microscopie par localisation de fluorophore photo-activable (PALM).....	218
Figure 83 : Le chromophore de la DRONPA est capable de « switcher » d'un état éteint (conformation trans) à un état allumé (conformation cis)	219
Figure 84 : Séquences des protéines ffDRONPA et eGFP sont très différentes (31 % d'homologie) mais leurs structures tridimensionnelles sont équivalentes (tonneaux β composés de 11 brins β).....	220
Figure 85 : Observation en microscopie à haute résolution (PALM) de la split-DRONPA	221
Figure 86 : Observation en microscopie à haute résolution (PALM) de la split-DRONPA	222

<i>Tableau 1 : Données épidémiologiques du VIH-1 et du SIDA dans le monde en 2015.....</i>	<i>13</i>
<i>Tableau 2 : Partenaires cellulaires et viraux connus de la protéine NCp7 ou du domaine NC de Gag.....</i>	<i>91</i>
<i>Tableau 3 : Structure, activité antivirale et cytotoxicité des principales familles d'éjecteurs de zinc de la NC.94</i>	
<i>Tableau 4 : Mutations observées sur deux souches résistantes à chacun des différents traitements.....</i>	<i>106</i>
<i>Tableau 5 : liste des plasmides utilisés au cours de cette thèse.....</i>	<i>117</i>
<i>Tableau 6 : liste des anticorps primaires et secondaires utilisés lors de cette thèse.....</i>	<i>118</i>
<i>Tableau 7 : liste des oligonucléotides nécessaires pour les mutagénèses.....</i>	<i>120</i>
<i>Tableau 8 : enzymes de restriction utilisées pour les différents clonages.....</i>	<i>120</i>
<i>Tableau 9 : Gel de poly-acrylamide SDS-PAGE.....</i>	<i>138</i>
<i>Tableau 10 : Mélange réalisé lors de la PCR pour une mutagénèse, une insertion ou une délétion.....</i>	<i>141</i>
<i>Tableau 11 : Activités antibactérienne, antirétrovirale et cytotoxicité des principaux inhibiteurs de RNAP et de RT.....</i>	<i>167</i>
<i>Tableau 12 : Structures et activités antivirales de molécules de la famille des calixarènes.....</i>	<i>168</i>
<i>Tableau 13 : Points isoélectriques et masses moléculaires des différentes protéines étudiées.....</i>	<i>198</i>
<i>Équation 1 : calcul du pourcentage d'inhibition de l'infection et de cytotoxicité.....</i>	<i>127</i>
<i>Équation 2 : détermination de la concentration inhibitrice à 50% (CI₅₀).....</i>	<i>127</i>
<i>Équation 3 : Analyse à deux composantes du temps de vie de Gag-eGFP en présence de Gag-mCherry.....</i>	<i>132</i>
<i>Équation 4 : Détermination du volume d'enzyme de restriction nécessaire pour la digestion.....</i>	<i>142</i>

Liste des abréviations

A

Å = Angstrom
 ADME = Administration, Distribution, Métabolisme et Excrétion
 ADN = Acide DésoxyriboNucléique
 AIE = Anémie Infectieuse des Equidés
 AIP1 = ALIX = Apoptosis-Linked gene 2-Interacting protein X
 Ala = Alanine
 ALIX = ALG2-interacting protein X
 AMP = Adénosine Monophosphate
 AMPS = AmMonium PerSulfate
 ANs = Acides Nucléiques
 ANRS = Agence Nationale de la Recherche sur le SIDA
 AP-1/2 = Adaptor Protein 1 ou 2
 APOBEC3G/F = A3G = APOLipoprotein B mRNA editing Enzyme, Catalytic polypeptide-like 3G ou 3F
 ARN = Acide RiboNucléique
 ARNm = ARN messenger
 ARNt^{Lys3} = ARN de transfert
 ATCC = American Type Culture Collection
 ATF/CREB = Activating Transcription Factors / cAMP Response Element Binding protein
 ATP = Adénosine TriPhosphate
 AZT = AZidoThymidine

B

BAF = Barrier to Auto-integration Factor
 BAL = British Anti-Lewisite = 2,3-dimercaptopropanol
 BiFC = Bimolecular Fluorescence Complementation
 BSA = Bovine Serum Albumine

C

CA = Capside
 Ca²⁺ = Calcium
 CA-CTD = Domaine C-Terminal de CA
 CA-NTD = Domaine N-Terminal de CA
 CBP = cAMP responsive Binding Protein
 CC₅₀ = Concentration Cytotoxique à 50 %
 CCR5 = Récepteur à la Chimiokine de type 5
 CD4 = Cluster de Différenciation 4
 CDC = Center for Disease Control

Cf = Concentration finale
 CI₅₀ = Concentration Inhibitrice à 50 %
 CMH-I ou -II = Complexe Majeur d'Histocompatibilité de classe 1 ou 2
 CMV = CytoMegalovirus
 Co-IP = co-ImmunoPrécipitation
 CPP = Cell Penetrating Peptide
 cPPT = central PolyPurine Tract
 CPSF = Facteurs de Clivages Spécifiques à la Polyadénylation
 CRISPR-Cas9 = Clustered Regularly Interspaced Short Palindromic Repeats - CRISPR associated protein 9
 CRM1 = Chromosome Maintenance Region 1
 C-ter = partie C-terminale
 CTI = Complexe de Transcription Inverse
 CXCR4 = C-X-C Chemokine Receptor type 4
 cycT-1 = cycline T1
 CypA = Cyclophilin A
 Cys = Cystéine

D

DC-SIGN : Dendritic Cell Specific Intercellular adhesion molecular-3-Grabbing Non-integrin
 DDX9 = RHA = RNA Hélicase A
 DIS = Dimer Initiation Site = site d'initiation de la dimérisation
 DMEM = Dubelcco's Modified Eagle Medium
 DMSO = DiMéthylSulfOxyde
 dNTPs = désoxyribonucléotides
 DPBS = Dubelcco's Phosphate Buffered Saline
 DTT = DiThioThreitol

E

ECL = Enhanced Chemiluminescence
 EDTA = Ethylene Diamine Tetraacetic Acid
 EED = Embryonic Ectoderm Development
 EF1α = Elongation Factor 1 – α
 eIF4E = Eukaryotic translation Initiation Factor 4E
 ELISA = Enzyme-Linked Immunosorbent Assay
 EMCCD = Electron-Multiplying Charge-Coupled Device
 Env = Enveloppe
 eRF1 = Eukaryotic Release Factors 1

ESCRT = Endosomal Sorting Complex
Required for Transport

F

FACS = Fluorescence-Activated Cell Sorting
FDA = Food and Drug Administration
ffDRONPA = fast folding DRONPA
FLIM = Fluorescence-Lifetime Imaging
Microscopy
FRET = Fluorescence Resonance Energy
Transfert

G

Gag = Group antigen
GFP = Green Fluorescent Protein
Gln = Glutamine
Gly = Glycine
GM3 = Monosialodihexosyl Ganglioside
GMP = Guanosine MonoPhosphate
Gp120 = Enveloppe Glycoprotein 120
Gp160 = Enveloppe Glycoprotein 160
Gp41 = Enveloppe glycoprotein 41
GRPE = Genomic RNA Packaging Enhancer

H

HAART = Highly Active AntiRetroviral
Treatment
HBR = Highly Basique Region
HBSS = Hank's Balanced Salt Solution
HCD = Higher-energy Collisional Dissociation
hCG1 = Nucleoporin-like protein NLP1
HDACs = Histone DeACetylases
HF = High Fidelity
hMGA1 = High Mobility Group protein HMG-
I/HMG-Y
HP1 γ = Heterochromatin Protein 1 γ
HR = Heptad-Repeat
HRP = Horse Radish Peroxydase
HSP70 = Heat Shock Protein 70

I

IAA = IodoAcetic Acid
ICAM-1 = Intercellular Adhesion Molecule 1
Ile = Isoleucine
IMP1 = Insulin-like growth factor II mRNA
binding Protein 1
IN = Intégrase
Inh = Inhibiteur
INI = Inhibiteurs de l'Intégrase
INNRT = Inhibiteurs Non Nucléosidiques de
Transcriptase Inverse

INTER = INTERmoléculaire
INTRA = INTRAmoléculaire
IP = Immuno Précipitation
IP5 = Inositol Pentaphosphate
IP6 = Inositols-Phosphates = Acide Phytique
IPHC = Institut Pluridisciplinaire Hubert
Curien
IPTG = IsoPropyl β -D-1-
ThiolGalactopyranoside
IRES = Internal Ribosome Entry Site

K

Kb = Kilobase
kDa = Kilo Dalton

L

LANL = Los Alamos National Lab
LAV = Lymphadenopathy Associated Virus
LB = Lysogeny Broth
LDL = Low Density Lipoproteins
LEDGF = Lens Epithelium-Derived Growth
Factor
LFA-1 = Lymphocyte Function-associated
Antigen 1
LSMBO = Laboratoire de Spectrométrie de
Masse Bio-Organique
LT CD4⁺ = Lymphocyte T auxiliaire
LT CD8⁺ = Lymphocyte T cytotoxique
LTR = Long Terminal Repeat
Lucif = Luciférase

M

MA = Matrice
MCS 1 et 2 = Multi Cloning Site 1 et 2
Met = Methionine
Mg²⁺ = Magnésium
MHR = Major Homology Region
MI = Inhibiteurs de Maturation
miARN = micro-ARN
MMP9 = MétalloProtéinase Matricielle 9
MMTV = Mouse Mammary Tumour Virus
Mn²⁺ = Manganèse
MS = Mass Spectrometry
MTT = bromure de 3-(4,5-dimethylthiazol-2-
yl)-2,5-diphenyl tetrazolium
MW = Molecular Weight
Myr = Myristate

N

n = Nombre d'échantillon analysé

N = Nombre d'expérience indépendante
 NC = NCp7 = NucléoCapside
 NCIs = NCinh = Inhibiteurs de la Nucléocapside
 Nef = Negative Regulatory Factor
 NFAT = Nuclear Factor of Activated T-cells
 NFX1 = Nuclear transcription Factor, X-box binding 1
 NF- κ B = Nuclear Factor-kappa B
 NLS = Nuclear Localisation Site
 nm = Nanometre
 NMD = Nonsense-Mediated Decay
 NMT = N-MyristoylTransférase
 NR = Neutron Reflectivity
 NRTI = Nucleoside analog Reverse-Transcriptase Inhibitors
 Nt = Nucléotides
 N-ter = Partie N-terminale
 NtRTIs = Nucleotide analog Reverse-Transcriptase Inhibitors
 NUPs = NUcléoPorines

O

OMS = Organisation Mondiale de la Santé
 ONUSIDA = Programme commun des nations unies sur le VIH/SIDA
 Opti-MEM = Optimisation du milieu Eagle's Minimum Essentials Media

P

P(T/S)AP = Motif Proline (Thréonine/Sérine) Alanine Proline
 p.i. = Post infection
 PACS2 = Phosphofurin Acidic Cluster Sorting protein 2
 PAGE = Poly Acrylamide Gel Electrophoresis
 PALM = Photo-Activated Localization Microscopy
 PARP = Poly(ADP-ribose) Polymerase
 PAS = Primer Activation Signal
 PATEs = PyridinioAlkanoyl ThioEster)
 PBS = Primer Binding Site
 PCR = Polymérase Chain Reaction
 PDB = Protein Data Base
 PEI = PolyEthylène Imine
 PFA = ParaFormAldéhyde
 Phe = Phenylalanine
 PI = Inhibiteurs de Protéase
 pI = Point Isoélectrique

PI_(4,5)P2 = Phosphatidyl-sérine et PhosphatidyInositol-4,5-bisphosphate
 PIC = Complexe de Pré-Intégration
 PIQ = Plateforme d'Imagerie Quantitative
 PMT = Photomultiplicateur
 Pol = Polymérase
 polyA = PolyAdenylation
 PPs = Pseudo-Particules
 PPT = Poly Purine Tract
 PPTc = Poly Purine Tract central
 PR = Protéase
 PrEP = Pre-Exposition Prophylaxie
 PSF = Point Spread Function
 PTD = Domaine de Transduction
 P-TEFb = Positive Transcription Elongation Factor b
 PVDF = PolyVinylidène Fluoride

R

Ran = RAs-related Nuclear protein
 RanBP1 = Ran-specific binding protein 1
 RanGAP = Ran GTPase-activating protein
 Ran-GTP = GTP-binding nuclear protein Ran
 Rh6G = Rhodamine
 R_{m/c} = Ratio entre la fluorescence moyenne des pixels membranaires et celle des pixels cytoplasmiques
 RNAP = ARN Polymérase
 RNaseH = RiboNucléase de type H
 RNP = RiboNucléoProtéine
 Rpm = Rotation par minute
 RRE = Rev Response Element
 RT = Transcriptase Inverse
 RT-PCR = Real Time Polymerase Chain Reaction

S

SAMTs = S-Acyl-2-Mercaptobenzamide Thioester
 SANS = Small Angle Neutron Scattering
 SAR = Relation Structure-Activité
 SD = Splice Donor
 SDS = Sodium Dodecyl Sulfate
 SELEX = Systematic Evolution of Ligands by Exponential Enrichment
 Ser = Serine
 SIDA = Syndrome d'Immunodéficience Acquise

SIVcpzPtt = Virus de l'Immunodéficience
Simienne du chimpanzé
SIVsm = Virus de l'Immunodéficience
Simienne du mangabey
SNARE = Soluble N-ethylmaleimide-sensitive
factor Attachment protein REceptor
SOCS1 = Suppressor of Cytokine Signaling
protein 1
SUV39H = Histone-lysine N-methyltransferase
protein
SV40 = Simian Virus 40
SVF = Sérum de Veau Fœtal

T

TALENs = Nucléases des Effecteurs Mimant
l'Activation de Transcription
TAR = TransActivation Response element
TBS-T = Tris Buffer Saline-Tween
TC = Tétracystéine
TEMED = Tétraméthyléthylènediamine
THINPAD = Targeting the HIV-1 Nucleocapsid
Protein to fight Antiretroviral Drug Resistance
TI₅₀ = Temps nécessaire pour perdre 50 % de
l'effet d'une molécule (TOAA)
TNPO3 = Transportin 3
TOAA = Time Of Addition Assay
Tpm = Tour par minute
TREX1 = Three-prime Repair Exonucléase
Trp = Tryptophane
Tsg101 = Tumor susceptibility gene 101

U

U. A. = Unité Arbitraire
UNG2 = Uracil DNA glycosylase
UTR = UnTranslated Region

V

Val = Valine
vDNA = Viral DNA
VIB = Virus de l'Immunodéficience Bovine
Vif = Virus de l'Immunodéficience Féline
VIH = Virus de l'Immunodéficience Humaine
VIS = Virus de l'Immunodéficience Simienne
Vpr = Viral Protein R
VPS4 = Vacuolar Protein Sorting 4
Vpu = Viral Protein U
VSV-G = Glycoprotéine du Virus de la
Stomatite Vésiculaire

W

WB = Western Blot

WP = Working package

WT = Wild-Type

Y

YPxL = motif Lysine Proline (nucléotide au
hasard) Lysine

Z

ZFN = nucléases à doigt de zinc

Symbole

$\Delta R_{inter-intra} = R_{m/c inter} - R_{m/c intra}$

%inh = pourcentage d'inhibition

Chapitre 1 : Introduction

I. Comprendre le VIH = Virus de l'Immunodéficience Humaine

A. Classification : à quelle famille appartient le VIH ?

Le Virus de l'Immunodéficience Humaine (VIH) est un virus :

- de la famille des **Retroviridae** : regroupe tous les virus enveloppés à ARN monocaténaire diploïde de polarité positive. Tous les rétrovirus possèdent une organisation génomique semblable. Sont ainsi retrouvés un gène **gag** (group antigen), un gène **pol** (codant la **polymérase**, la transcriptase inverse et l'intégrase) et un gène **env** (codant les protéines d'enveloppe).
- de la sous-famille des **Orthoretrovirinae** : la caractéristique principale de ces virus est l'étape de transcription inverse qui permet de rétrotranscrire l'ARN génomique viral en ADN viral double brin qui sera par la suite intégré dans le génome de la cellule infectée.
- et du genre **Lentivirus** : caractérise les virus qui ont une période d'incubation longue (peut durer de 1 à 15 ans sans signe ou symptôme apparent de maladie) et qui sont cytopathogènes (toxiques pour la cellule que le virus infecte) (Campbell and Robinson, 1998). Plusieurs espèces de mammifères sont touchées par cette famille de virus présentant des caractéristiques espèce-dépendantes : bovins (VIB = Virus de l'Immunodéficience Bovine), félins (VIF = Virus de l'Immunodéficience Féline), équidés (AIE = Anémie Infectieuse des Équidés), les primates (VIS = Virus de l'Immunodéficience Simienne) et l'Homme (VIH).

B. Origine, Évolution Phylogénétique et Historique : d'où vient ce virus découvert il y a 33 ans ?

Le VIH-1 (ou à cette époque LAV : Lymphadenopathy-Associated Virus) a été découvert en 1983 par l'équipe de recherche de l'Institut Pasteur du Professeur Montagnier où travaillaient

les Drs Barré-Sinoussi et Chermann (Barré-Sinoussi et al., 1983)(Prix Nobel médecine 2008). Ce virus semble être apparu au début du XXème siècle en Afrique centrale de l'ouest (Korber et al., 2000; Worobey et al., 2008) et plus précisément dans les années 1920 dans la région de Kinshasa (aujourd'hui en République Démocratique du Congo) (Faria et al., 2014). Il s'agirait d'une transmission inter-espèce où l'Homme aurait été en contact avec des tissus contaminés de primates (chasse de viande de brousse, primates comme animal de compagnie...). La datation approximative de l'origine de transmission du virus à l'Homme a été déterminée en étudiant l'évolution génétique du virus à partir du plus ancien échantillon humain positif pour le VIH (biopsie d'un nœud lymphatique prélevé en 1960 d'une femme de Kinshasa, République Démocratique du Congo).

L'étude phylogénétique (**Figure 1**) de différents rétrovirus rappelle la grande diversité génétique du VIH-1 constituant un des obstacles majeurs de l'éradication de ce virus (Smyth et al., 2012). Elle permet également de montrer que le VIH-1 est plus proche du Virus de l'Immunodéficience Simienne du chimpanzé (SIVcpzPtt) (Keele et al., 2006) pour les groupes M et N ou du gorille (SIVgor) pour les groupes P et O (D'arc et al., 2015). Le VIH-2 quant à lui, serait apparu suite à une transmission inter-espèces avec des mangabeys (SIVsm) (Reeves and Doms, 2002). Le VIH-1 est classé en 4 groupes (**Figure 1**) (Mourez et al., 2013) :

- **Groupe M** (pour Main = Majoritaire) : regroupe les virus à l'origine de l'épidémie du SIDA puisqu'ils sont responsables de 99 % des infections soit plus de 40 millions de personnes à travers le monde et qui est aussi la lignée de virus la plus ancienne. Ce groupe est également divisé en 9 sous-types (A, B, C, D, F, G, H, J, K), tous distincts génétiquement. On distingue également dans les sous-groupes A et F des sous-sous-groupes (A1, A2, A3, A4, F1 et F2). Lorsqu'un patient est co-infecté par plusieurs sous-types, une forme recombinante apparaît (actuellement une quarantaine observée) accentuant alors la diversité génétique du virus (appelé CRF : Circulating Recombinant Forms) (Hemelaar, 2012).
- **Groupe N** (pour New = Nouveau, ou Non-N Non-O) : est plus rare et retrouvé uniquement chez des individus Camerounais.
- **Groupe O** (pour Outlier = aberrant) : est également rare (environ 100000 personnes infectées) et a été décrit suite à l'isolement d'une nouvelle souche présentant une divergence importante avec le groupe M (De Leys et al., 1990). Ce

groupe est retrouvé uniquement en Afrique centrale de l'ouest (Cameroun, Gabon, Nigeria).

- **Groupe P** (lettre choisie à la suite des groupes M, N et O) : qui est le groupe découvert le plus récemment (Plantier et al., 2009), mais également celui dont la transmission inter-espèces est la plus récente.

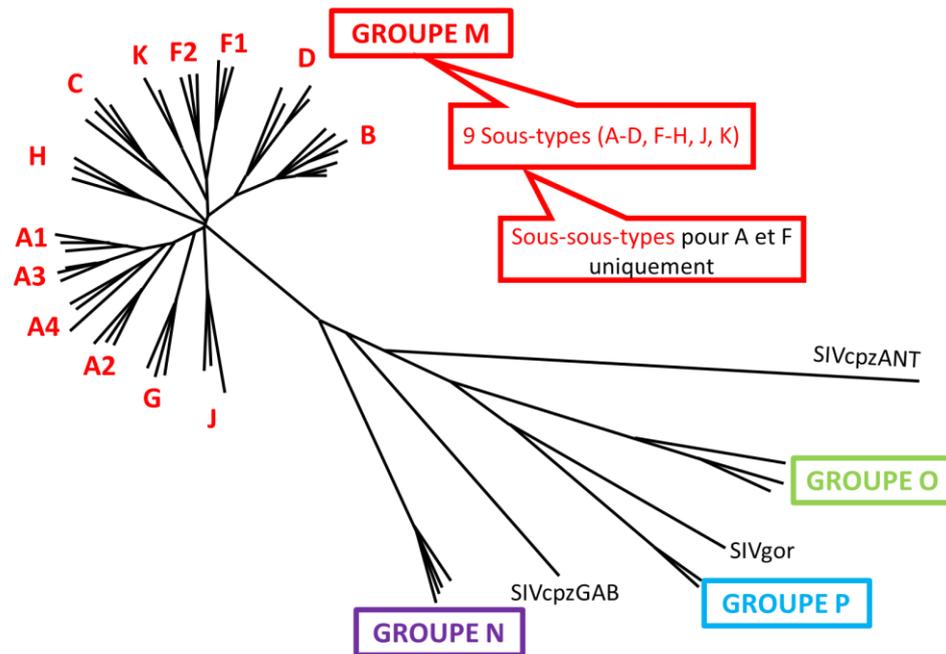


Figure 1 : Arbre phylogénétique du VIH-1 montrant la grande diversité génétique des VIH-1 (simplifié de Tebit and Arts, 2011)

Jusqu'au début des années 1980, ce virus a donc pu se développer sans être inquiété par une quelconque action humaine. En 1981, le doute s'installe dans la communauté médicale et scientifique suite à l'apparition simultanée de plusieurs cas d'infections pulmonaires rares et de cancers agressifs (Sarcomes de Kaposi). Ces maladies, couplées à une sévère immunodéficience, sont retrouvées dans la communauté gay de plusieurs villes états-uniennes. Après avoir été rapporté chez les drogués par injection, les hémophiles, les haïtiens et finalement les hétérosexuels, le virus à l'origine du **S**ndrome d'**I**mmuno**D**éficience **A**cquis (SIDA) est décrit par des chercheurs français sous le nom de LAV (Lymphadénopathie-Associated Virus) (Barré-Sinoussi et al., 1983). Si le prix Nobel et la découverte de ce virus ont été attribués à l'équipe française, une équipe américaine a identifié que la particule virale découverte par les français est celle qui est à l'origine du SIDA (Gallo et al., 1983). À partir de cette date, les connaissances sur le virus n'ont cessé d'augmenter, permettant le

développement de traitements de plus en plus efficaces. Plusieurs dates importantes sont à retenir sur la découverte du VIH (*Figure 2*) :

- *Début XX^{ème}* : origine possible du VIH dans l'Afrique centrale de l'ouest (Worobey et al., 2008)
- **1982** : première description du Syndrome d'Immuno-Déficience Acquise par le centre de contrôle des maladies états-uniens (CDC = Center for Disease Control).
- **1983 : isolation et découverte du Virus de l'Immunodéficience Humaine** (Barré-Sinoussi et al., 1983)
- **1984** : description du récepteur CD4 comme étant indispensable pour l'infection par le VIH-1 (Dalglish et al., 1984)
- **1985** : séquençage du génome viral (Ratner et al., 1985)
- **1986** : agrément par l'agence du médicament états-uniennes (FDA, Food and Drug Administration) de la zidovudine (AZT) développée à l'origine en 1964 comme anti-cancer (Horwitz et al., 1964)
- **1996** : découverte des corécepteurs : CCR5 (Alkhatib et al., 1996; Choe et al., 1996; Deng et al., 1996) et CXCR4 (Feng et al., 1996a)
 - + Première tri-thérapie antirétrovirale (HAART = Highly active Antiretroviral treatment) comprenant un inhibiteur de protéase (Saquinavir) et deux inhibiteurs nucléosidiques de transcriptase inverse (zidovudine, AZT et zalcitabine) (Collier et al., 1996).
- **2005** : description des patients dit « contrôleurs » du VIH-1, c'est-à-dire qu'ils sont infectés mais capable de contrôler la réplication virale (Lambotte et al., 2005).
- **2008** : patient Berlin qui a reçu une transplantation de cellules souches hématopoïétiques d'un donneur portant une mutation sur le gène du récepteur CCR5. Ce patient, positif au VIH-1 depuis 1995, a stoppé la trithérapie depuis cette transplantation et la charge virale est restée indétectable. Il est ainsi considéré comme soigné du VIH-1 (Hütter et al., 2009).
- **2013** : contrôleurs post-traitements. Ces patients, traités précocement après infection par le VIH-1, arrivent à maîtriser le développement du virus naturellement (Sáez-Cirión et al., 2013).

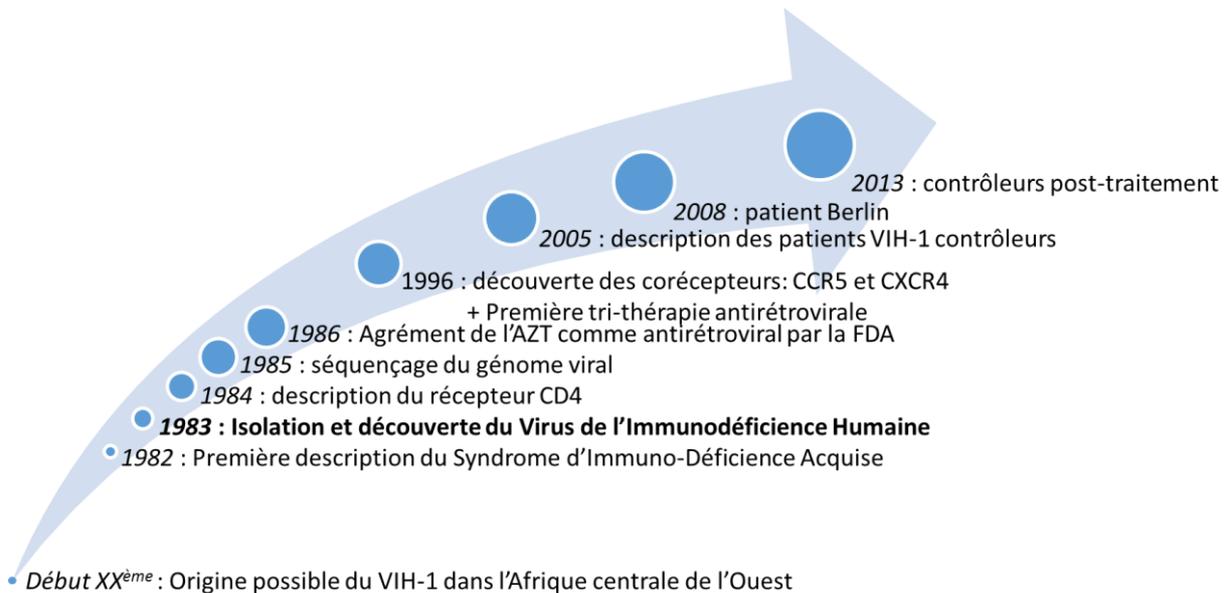


Figure 2 : Principales dates à retenir sur les découvertes du VIH-1 (adapté de Barré-Sinoussi et al., 2013)

C. Transmission et Infection du virus : comprendre pour ne pas infecter ou être infecté

1. Mode de transmission

Le VIH-1 étant présent dans la plupart des liquides biologiques de l'organisme mais à des concentrations différentes, la transmission du virus d'un individu à un autre se fait donc par l'échange de ces liquides (sang, sperme, liquide séminal, sécrétions vaginales et lait maternel).

Il existe trois modes de transmission :

- Voie sexuelle : c'est la voie majoritaire de transmission. Il y a pénétration du virus au travers des muqueuses vaginales, buccales ou anales. Si des lésions sont présentes au niveau de ces muqueuses, ou s'il y a présence d'autres maladies sexuellement transmissibles, une forte augmentation du risque de contamination est observée. Le meilleur moyen de limiter l'infection par ce type de transmission reste l'utilisation du préservatif. Récemment, un traitement prophylactique a été autorisé à la commercialisation en France (PrEP = Prophylaxie Pré-Exposition) afin de cibler les populations clés pouvant être réfractaires à l'utilisation du préservatif (essai ANRS Ipergay).

- Voie sanguine : rapidement après la découverte du virus, la voie sanguine a été démontrée comme étant une voie de transmission du VIH-1. En effet, de nombreuses personnes se sont retrouvées infectées suite à une transfusion sanguine ou une transplantation (risque presque nul en France de nos jours mais existant encore dans les pays où le sang n'est pas systématiquement testé lors du don). Aujourd'hui, la transmission par cette voie persiste principalement chez les toxicomanes (injection de drogues par voie intraveineuse) ou chez les personnels de santé ou de recherche (blessure par un objet souillé).
- Mère – enfant : le risque de transmission verticale peut exister pendant la grossesse, pendant l'accouchement (principal risque de transmission verticale) ou encore pendant l'allaitement maternel. Ce risque reste très faible en France grâce à un accompagnement adapté.

Il est important de noter que la transmission se fait essentiellement par le virus de type CCR5 qui reste prédominant dans les phases initiales de l'infection (Keele et al., 2006; Rowland-Jones, 2003). Puis les virus capables d'utiliser CXCR4 ou ayant un double tropisme sont détectés au fur et à mesure des années et cela de manière concomitante à la déplétion des lymphocytes et l'apparition des signes cliniques. Or rapidement après l'infection, la réplication virale étant exponentielle, l'homogénéité de la population de virus circulants devient très importante. Il existe donc plusieurs systèmes « d'entonnoir » qui permettent la transmission d'une seule souche virale alors que la diversité génétique intra-individu est importante (Joseph et al., 2015).

2. Infection : de la contamination au SIDA

Sans traitement, l'infection par VIH-1 évolue en 3 étapes aboutissant à la disparition du système immunitaire et la dégradation de l'état général de la personne infectée conduisant à sa mort (**Figure 3**). Ces trois étapes sont :

- Phase aigüe ou primo-infection : cette phase correspond aux quelques semaines qui suivent l'exposition au VIH-1. Elle est caractérisée par une forte augmentation du nombre de virus dans la circulation sanguine qui résulte en une forte diminution du nombre de lymphocytes T CD4⁺. Parfois, à cause de la dégradation du système immunitaire, cette phase est accompagnée de symptômes mimant ceux d'une grippe ou d'une mononucléose (fièvre, éruption cutanée...). D'un point de vue du

dépistage, cette phase est caractérisée par un test ELISA (Enzyme-Linked Immunosorbent Assay) négatif (anticorps anti-VIH-1 pas encore synthétisés), alors qu'il est possible de détecter l'ARN viral et/ou l'antigène p24 permettant le diagnostic. C'est pendant cette étape que le réservoir viral est constitué (Simon and Ho, 2003). Ce réservoir est ainsi constitué par une partie du virus qui reste à l'état latent (génome viral intégré mais non transcrit) dans certaines cellules (cellules dendritiques, macrophages, cellules microgliales) rendant ainsi très compliqué le développement d'un traitement anti-VIH-1 curatif.

- Phase chronique asymptomatique : l'apparition des anticorps anti-VIH-1 dans le sang marque le début de cette phase. On observe alors une nouvelle augmentation de la quantité de lymphocytes T CD4⁺ permettant dans un premier temps une diminution du nombre de copies d'ARN viral qui se stabilise à un faible niveau tout au long de cette période. Cette phase de latence, pouvant durer plusieurs dizaines d'années si le patient est sous traitements antirétroviraux, est caractérisée par une absence de symptômes mais, si la charge virale est forte, la transmission reste toujours possible à un autre individu.
- Phase SIDA (Syndrome d'ImmunoDéficience Acquise) : caractérisée par une chute drastique du nombre de lymphocytes T CD4⁺ (inférieur à 200 par mm³ de sang au lieu de 500 à 1600 LT CD4⁺/mm³) facilitant l'apparition d'une maladie opportuniste grave conduisant à la mort du patient (cancer invasif, sarcome de Kaposi, lymphome, candidose...). Sans traitement, cette phase dure rarement plus de 3 ans.

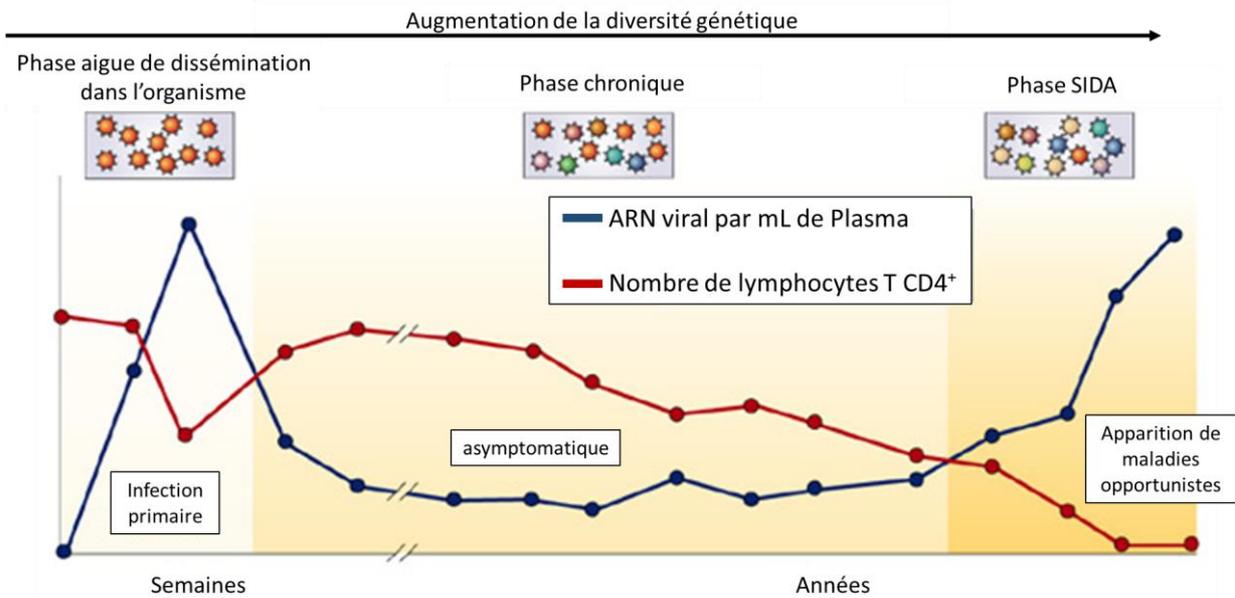


Figure 3 : Évolution naturelle de l'infection par le VIH-1 sur la base de l'évolution de deux marqueurs clés de l'infection : nombre de lymphocytes T CD4⁺ qui diminue à l'inverse de la charge virale (quantité d'ARN viral dans le plasma sanguin) qui augmente. Il est également rappelé ici que la diversité génétique du virus au sein du même individu augmente avec l'évolution de l'infection (Simon and Ho, 2003).

D. Épidémiologie : pourquoi l'infection par le VIH est toujours bel et bien une épidémie ?

En 2015, on décompte encore plus de 2 millions de nouvelles infections par le VIH-1 et 1,1 millions de décès dus au SIDA (**Tableau 1**). Lorsque l'on s'intéresse à la prévalence (nombre de cas de la maladie par rapport à la population totale), il est possible d'observer que dans le monde environ une personne sur 100 est infectée par le VIH (0,8 %) et qu'en Afrique presque une personne sur 20 est infectée (4,8 %) (**Figure 4**).

	Total	36,7 millions [34,0 – 39,8 millions]
Nombre de personnes vivant avec le VIH-1 en 2015	Adultes	34,9 millions [32,4 – 37,9 millions]
	Femmes (+ de 15 ans)	17,8 millions [16,4 – 19,4 millions]
	Enfants (- de 15 ans)	1,8 millions [1,5 – 2,0 millions]
Nombre de personnes nouvellement infectées par le VIH-1 en 2015	Total	2,1 millions [1,8 – 2,4 millions]
	Adultes	1,9 millions [1,7 – 2,2 millions]
	Enfants (- de 15 ans)	150 000 [110 000 – 190 000]
Nombre de morts dus au SIDA en 2015	Total	1,1 millions [940 000 – 1,3 millions]
	Adultes	1,0 millions [840 000 – 1,2 millions]
	Enfants (- de 15 ans)	110 000 [84 000 – 130 000]

Tableau 1 : Données épidémiologiques du VIH-1 et du SIDA dans le monde en 2015 (Traduit d'un document publié par l'OMS)

En France (mais aussi dans d'autres pays d'Europe de l'ouest), l'épidémie n'est pas totalement contrôlée et ce particulièrement chez les hommes ayant des relations sexuelles avec des hommes (HSH) mais aussi dans d'autres populations clés (professions du sexe, toxicomanes par injection). En effet, malgré une prévalence globale en France de moins de 0,2 %, elle est de 17 % chez les HSH.

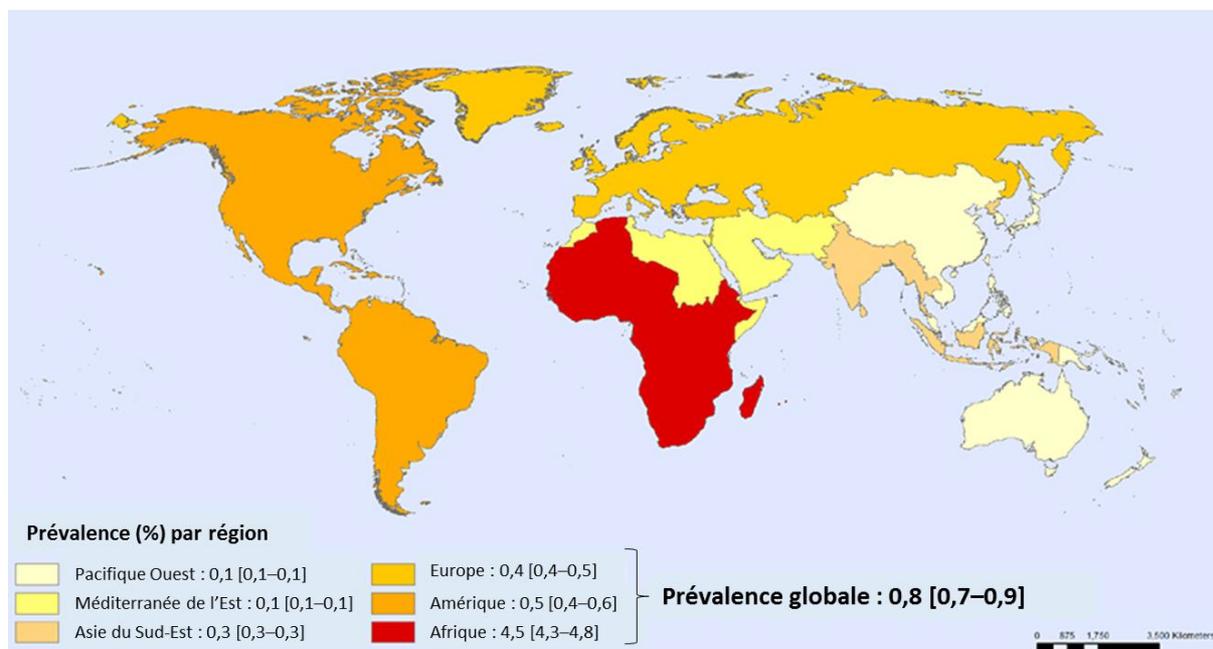


Figure 4 : Prévalence du VIH-1 chez l'adulte (15 - 49 ans) dans différentes régions du monde en 2014 (Traduit d'un document publié par l'OMS)

Pour mettre fin à l'épidémie du SIDA à l'horizon 2020, une cible ambitieuse a été mise en place par l'ONUSIDA en 2014 : 90 % des personnes vivant avec le VIH devront être dépistées, 90 % de celles-ci devront recevoir un traitement antirétroviral durablement et 90 % de ces dernières devront présenter une charge virale indétectable (**Figure 5**) (ONUSIDA, 2014). Cet objectif permettra de multiplier par trois le nombre de personnes vivant avec le VIH qui auront une charge virale indétectable (on atteindra 73 %) donc potentiellement non infectieuses. La modélisation de l'épidémie si cette cible est atteinte prédit une fin de l'épidémie du SIDA d'ici à 2030 (ONUSIDA, 2014). Si pour la France cet objectif semble réalisable, il paraît beaucoup plus difficile à atteindre dans d'autres régions du monde comme l'Afrique Sub-saharienne où le principal effort devra être fait au niveau du dépistage (seulement 45 % des personnes vivant avec le VIH savent qu'elles sont infectées). L'apparition de nouveaux matériels, plus simples d'utilisation, portatifs et peu onéreux devrait faciliter ce dépistage (Ritchie et al., 2014). L'idée est donc de fixer des objectifs difficilement atteignables pour concentrer les efforts et les ressources sur celui-ci.

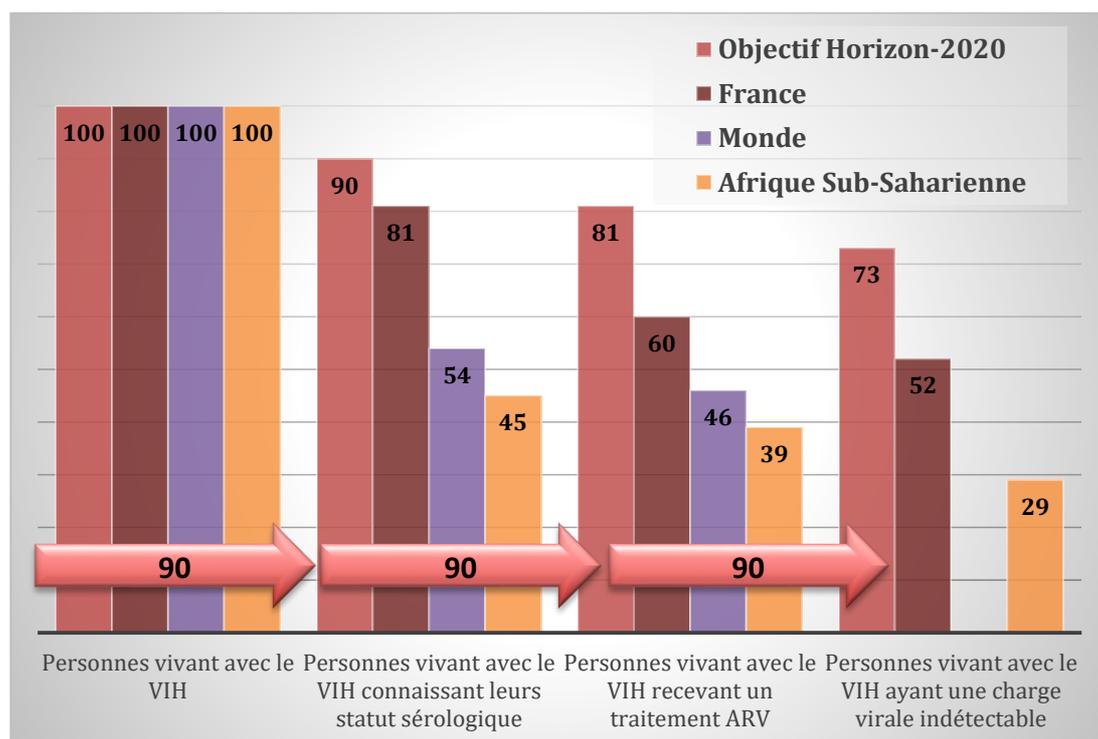


Figure 5 : Cascade de la prise en charge des personnes vivant avec le VIH en comparaison avec l'objectif horizon-2020 90-90-90 (apparaît ici en rouge). Les pourcentages donnés pour chaque catégorie sont des estimations en prenant comme 100 % le nombre de personnes vivant avec la VIH-1 dans chaque région. Issus du site de l'Organisation Mondiale de la Santé pour les données Monde et de (ONUSIDA, 2014) pour la France et l'Afrique Sub-saharienne. Aucune donnée n'est disponible concernant le nombre de personnes vivant avec le VIH dans le monde avec une charge virale indétectable.

II. Comment est structuré le VIH-1 ?

A. Photos d'identités de la particule virale immature et mature

La particule virale contient le génome viral ainsi que l'ensemble des protéines indispensables pour l'infection et la multiplication du virus. Une particule virale est sphérique et d'un diamètre moyen de 130 nm pour une particule mature ou immature (Briggs et al., 2003). Ces deux particules présentent des morphologies fortement différentes, la phase de maturation étant absolument nécessaire pour obtenir des particules infectieuses.

La particule immature est composée de l'alignement des précurseurs Gag et Gag-Pol dont l'extrémité N-terminale est ancrée dans la bicouche lipidique et dont l'extrémité C-terminale interagit avec les deux copies d'ARN viral (Wilk et al., 2001) (*Figure 6 A et B*). La particule est entourée d'une enveloppe constituée d'une bicouche lipidique empruntée à la cellule hôte lors du bourgeonnement de la particule. Finalement, à la surface de cette membrane, on retrouve les glycoprotéines virales reconnues par les récepteurs des lymphocytes T CD4⁺ indispensables pour l'infection.

La particule mature est le résultat d'une modification structurelle des protéines de structure (Mattei et al., 2016) et possède une forme caractéristique avec, en son centre, une capside de forme conique formée par l'assemblage d'hexamères et de pentamères de protéine de Capside (CA, *Figure 6 C et D*) (Briggs et al., 2003). Cette dernière renferme le génome viral recouvert de nombreuses protéines de Nucléocapside (NC). Cette ribonucléoprotéine comporte également les protéines Transcriptase inverse (RT), Intégrase (IN) et Protéase (PR). L'enveloppe de cette particule est en revanche identique à celle observée au niveau de la particule immature.

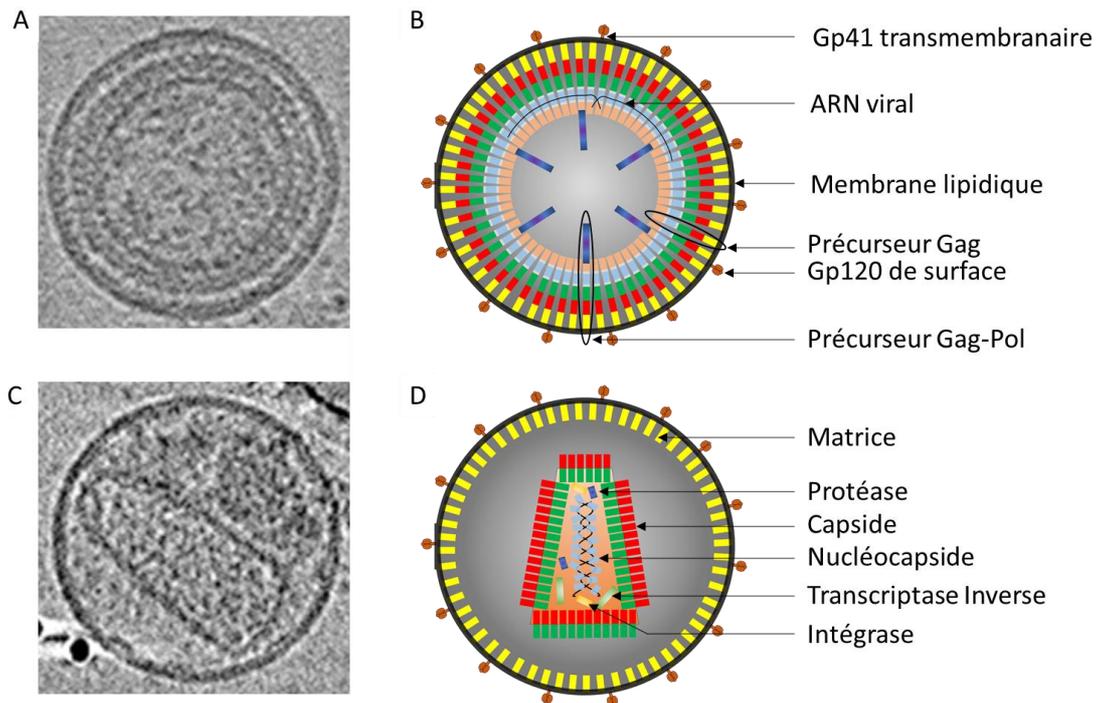


Figure 6 : Particules virales immatures (A et B) et matures (C et D) observées en microscopie électronique (A et C, particules d'environ 130 nm de diamètre) et en représentations schématiques (B et D). (Adapté et traduit de Balasubramaniam and Freed, 2011; Ganser-Pornillos et al., 2008).

B. Génome viral : un plan simple pour une structure complexe

1. L'ARN viral :

Une particule virale du VIH-1 contient deux copies d'ARN génomique simple brin de polarité positive, coiffé à l'extrémité 5' et poly-adénylé à son extrémité 3' ce qui signifie qu'il est susceptible, comme les ARNs messagers cellulaires, d'être traduits en protéines. Les séquences de ces deux copies peuvent être différentes augmentant ainsi la diversité génétique du virus.

a) Région codante de l'ARN

Si la séquence de l'ARN viral est connue de longue date (Ratner et al., 1985), la structure secondaire de l'intégralité de l'ARN n'a été proposée que récemment (Sükösd et al., 2015). Cette structure (et notamment les extrémités 5'-UTR et 3'-UTR) est riche en appariement sous forme de tige-boucle et permet l'activation de la transcription, l'initiation de la transcription inverse, facilite la dimérisation de 2 ARNs viraux ainsi que l'assemblage, modifie le cadre de

lecture, interagit avec des protéines virales et cellulaires (Watts et al., 2009). Mais il est important de se souvenir que la fonction principale de l'ARN du VIH-1 (d'une taille de 9,2 kb) est de coder le génome viral nécessaire pour la production de l'ensemble des protéines virales. Celui-ci contient neuf cadres de lecture ouverts qui permettent la production de 15 protéines (**Figure 7**). Trois de ces cadres de lecture encodent les polyprotéines Gag, Pol et Env qui seront par la suite protéolysés pour donner l'ensemble des protéines virales. Ainsi, le précurseur Gag est clivé en plusieurs protéines, MA (Matrice), CA (Capside) et NC (Nucléocapside) qui sont les composants indispensables pour la structure de la particule virale. Le précurseur Env est lui clivé en une protéine de surface (gp120) et une protéine transmembranaire (gp41) indispensable pour la reconnaissance du virus par les cellules lymphocytaires de l'organisme infecté par le VIH-1. Les trois protéines issues de Pol, PR (Protéase), RT (Transcriptase inverse) et IN (Intégrase) portent les activités enzymatiques essentielles à l'infection et sont encapsulées dans la particule virale (Frankel and Young, 1998). L'ARN code également six protéines accessoires dont certaines sont encapsidées dans la particule virale (Vif, Vpr, Nef), d'autres participent à la régulation de la transcription génique (Tat et Rev) et une dernière facilite l'assemblage de la particule (Vpu). Les protéines Tat et Rev sont indispensables pour la transcription virale, alors que les protéines auxiliaires (Vif, Vpr, Nef et Vpu) sont nécessaires *in vivo* mais pas *in vitro*.

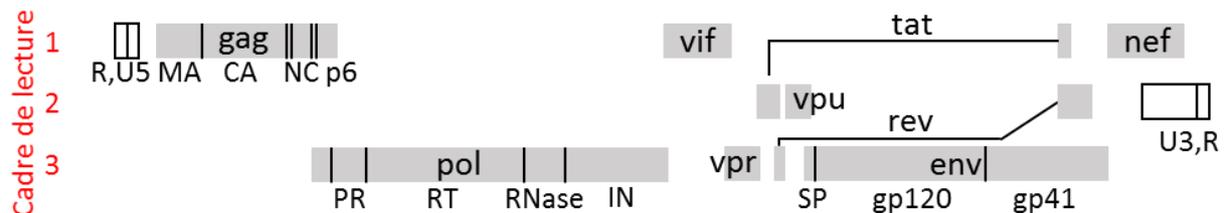


Figure 7 : Organisation génomique du VIH-1 sous forme d'ARN. Les régions codant les protéines sont en gris, les domaines de jonctions des protéines au sein des polyprotéines (gag, pol et env) sont indiqués par un trait vertical noir (Watts et al., 2009).

b) Région non codante de l'ARN

Les régions codant toutes ces protéines virales sont encadrées par des séquences non codantes appelées UTR (UnTranslated Region) : 5'-UTR et 3'-UTR situées aux deux extrémités du génome. Ces deux séquences sont composées d'une région identique aux deux extrémités et d'une région propre à chaque extrémité (U5 et PBS pour l'extrémité 5' et U3 et PPT pour l'extrémité 3').

(1) Partie identique aux extrémités 5' et 3' : R

Cette partie, appelée la région **R** (en rose dans la **Figure 8**), est composée de deux structures en tige-boucle :

- **TAR** (TransActivation Response element, Nucléotides 1 à 57) : est le site de fixation de la protéine Tat indispensable pour l'activation transcriptionnelle du VIH-1. Le rôle de cette protéine sera expliqué dans la partie [II.4.a)] de ce chapitre. Cette région, hautement conservée et très structurée, joue un rôle important dans la transcription inverse (Beerens and Kjems, 2010; Harrich et al., 1997) ainsi que lors de l'assemblage en facilitant l'empaquetage (Helga-Maria et al., 1999) et la dimérisation de l'ARN viral (Andersen et al., 2004; Jalalirad et al., 2012).
- **polyA** (PolyAdénylation, nucléotides 58 à 104) : contient le signal de polyadénylation (AAUAAA). Malgré la présence de cette zone aux deux extrémités de l'ARN viral, le signal en 5' semble être séquestré dans une boucle, alors que celui en 3' semble alterner entre conformation où la polyadénylation peut avoir lieu et séquestration dans une boucle, permettant la régulation de la transcription virale (Zarudnaya et al., 2013).

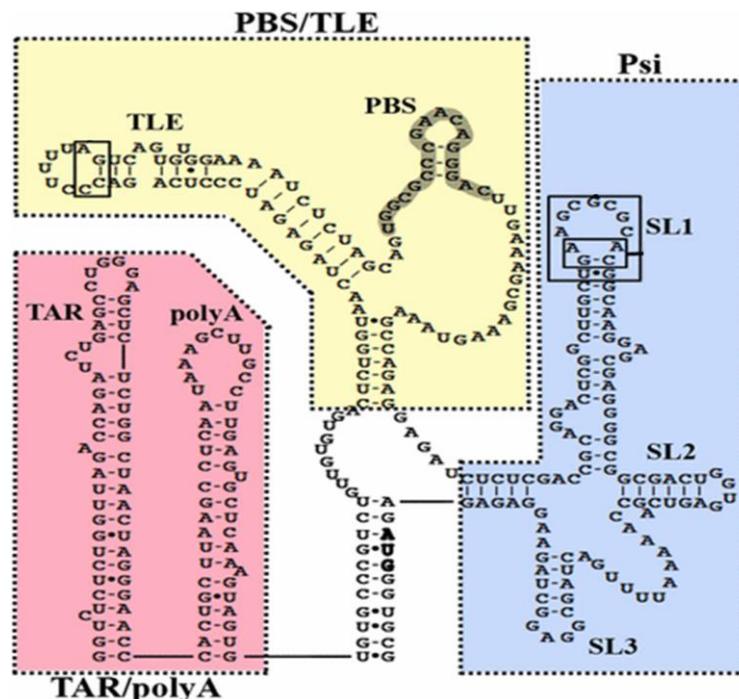


Figure 8 : Structure secondaire possible de la région 5'-UTR (U5) de l'ARN du VIH-1 (Jones et al., 2014). Les lignes pointillées englobent les différentes parties de cette région : TAR/polyA en rose qui est identique en 5' et en 3' de l'ARN viral et des séquences uniques à l'extrémité 5' : PBS/TLE en jaune et Psi en bleu.

(2) Région unique de l'extrémité 5' : U5

En 5' de l'ARN messager (Heng et al., 2012), on retrouve une région unique (**U5**) à cette extrémité comprenant :

- **PBS** (Primer Binding Site)/TLE (tRNA-Like Element) (en jaune dans la **Figure 8**) : cette zone a un rôle majeur dans la réplication car c'est la zone de fixation de l'initiateur de transcription inverse, l'ARN_t^{Lys3} (Beerens and Berkhout, 2002). Elle est, de plus, importante dans l'assemblage des particules virales en facilitant l'accès de la protéine Gag pour son interaction avec la boucle SL3 (Clever et al., 2002).
- **Psi** (en bleue dans la **Figure 8**) : région d'environ 120 nucléotides nécessaire mais non suffisante pour l'encapsidation du génome viral grâce à son interaction avec Gag (Lu et al., 2011). Cette zone est composée de quatre structures tige-boucle :
 - SL1 ou DIS (Dimer Initiation Site) : permet l'initiation de la dimérisation du génome avec le même site d'un autre ARN via une interaction boucle-boucle (kissing-loop) (Skripkin et al., 1994). Cette structure SL1 semble être un site à haute affinité pour la polyprotéine Gag (Abd El-Wahab et al., 2014) et la dimérisation facilitée au niveau de SL1 permet l'exposition du signal d'encapsidation Ψ . C'est pourquoi SL1 est impliquée dans l'assemblage des particules (Lawrence et al., 2003). Cependant sa mutation a un effet drastique sur l'encapsidation des ARNs viraux encapsidés ce qui conditionne le pouvoir infectieux (Houzet et al., 2007).
 - SL2 ou SD (Splice Donor) : est le site principal d'épissage de l'ARN viral. Tous les ARNs viraux, qui sont épissés, le sont au niveau de ce site pour former différents ARNm viraux et donc certaines protéines virales. De plus, cette boucle est une zone de haute-affinité pour la fixation de la NC (Amarasinghe et al., 2000).
 - SL3 : contient le signal d'encapsidation (Ψ) permettant l'encapsidation préférentielle du génome viral lors de l'assemblage (Lever et al., 1989; Russell et al., 2003) grâce à la très forte affinité qu'elle a avec la NC (Athavale et al., 2010; Guzman et al., 1998).
 - SL4 : contient le codon d'initiation (AUG) pour la traduction des précurseurs Gag et Gag-Pol.

(3) Région unique de l'extrémité 3' : U3

Cette zone possède un rôle moins étudié que la région U5. Elle est ainsi décrite comme site d'interactions préférentielles de micro-ARN (miARN) qui régulent le cycle répliatif (Sun et al., 2012) et facilitent la mise en place du réservoir du VIH-1 (Huang et al., 2007).

(4) Autres régions non codantes

En dehors de ces régions, d'autres séquences, non codantes mais situées dans des régions codantes, sont retrouvées dans la séquence de l'ARN viral :

- PPT (Poly Purine Tract) : riche en purines, est retrouvée juste avant la séquence unique U3 en 3' de l'ARN génomique viral au niveau de la séquence codant Nef. Cette région, très conservée dans tous les rétrovirus, est indispensable pour la synthèse du brin (+) de l'ADN viral lors la transcription inverse. Son rôle lui est conféré par sa résistance à l'activité RNase H de la transcriptase inverse. Alors que l'ARN viral est dégradé, ce fragment d'ARN reste apparié au brin (+) d'ADN viral nouvellement synthétisé, formant un hybride ARN/ADN stable qui devient ainsi l'amorce pour la synthèse du brin (-) de l'ADN viral (Rausch and Le Grice, 2004). Si cette région est mutée, la réplication virale est limitée voire supprimée conduisant à une diminution drastique de l'infectivité du virus (Miles et al., 2005).
- PPTc (Poly Purine Tract central) : située en plein centre du génome viral, dans la région codant l'intégrase, est structurellement similaire au 3'-PPT. Cette région participe à la synthèse discontinue du brin (+) d'ADN en formant un 2^{ème} site d'initiation de la transcription (Rausch and Le Grice, 2004). En effet, lors de la transcription inverse, au niveau de la séquence cPPT est retrouvée une structure à trois brins d'ADN appelée « ADN-flap » (Charneau and Clavel, 1991; Charneau et al., 1992). Cette structure semble également être très importante dans la formation et l'import nucléaire du complexe de pré-intégration (PIC) (Arhel et al., 2007; Rivière et al., 2010), mais aussi comme bloqueur de l'action des facteurs de restrictions APOBEC3G/F (Hu et al., 2010).
- RRE (Rev Response Element) : cette séquence, située sur le gène env, est uniquement présente au niveau des ARNm non épissés ou mono-épissés (Malim et al., 1989) afin de permettre la fixation de plusieurs protéines Rev (DiMattia et al.,

2010), favorisant le recrutement de facteurs d'exports nucléaires (Crm1 et Ran-GTP). Alors que l'épissage total est normalement nécessaire pour que les ARNm sortent du noyau, le complexe formé autour de RRE, permet la régulation de la sortie des ARNm non- ou mono-épissés. De par sa capacité à adopter des conformations différentes cette structure permet également le contrôle de la réplication virale (Blissenbach et al., 2010; Grewe et al., 2012; Sherpa et al., 2015).

2. L'ADN proviral :

Une des étapes clés du cycle viral du VIH-1 est la transcription inverse. Elle permet ainsi la synthèse d'ADN proviral à partir de l'ARN viral génomique (**Figure 9**).

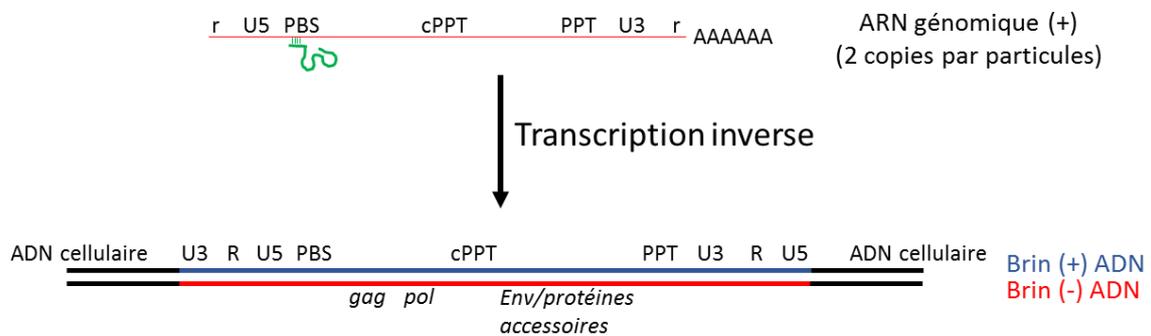


Figure 9 : La transcription inverse permet d'obtenir l'ADN proviral double brin à partir de l'ARN viral génomique. Les étapes de cette transcription inverse seront décrites dans la section [III.D] de ce chapitre (Sarafianos et al., 2009).

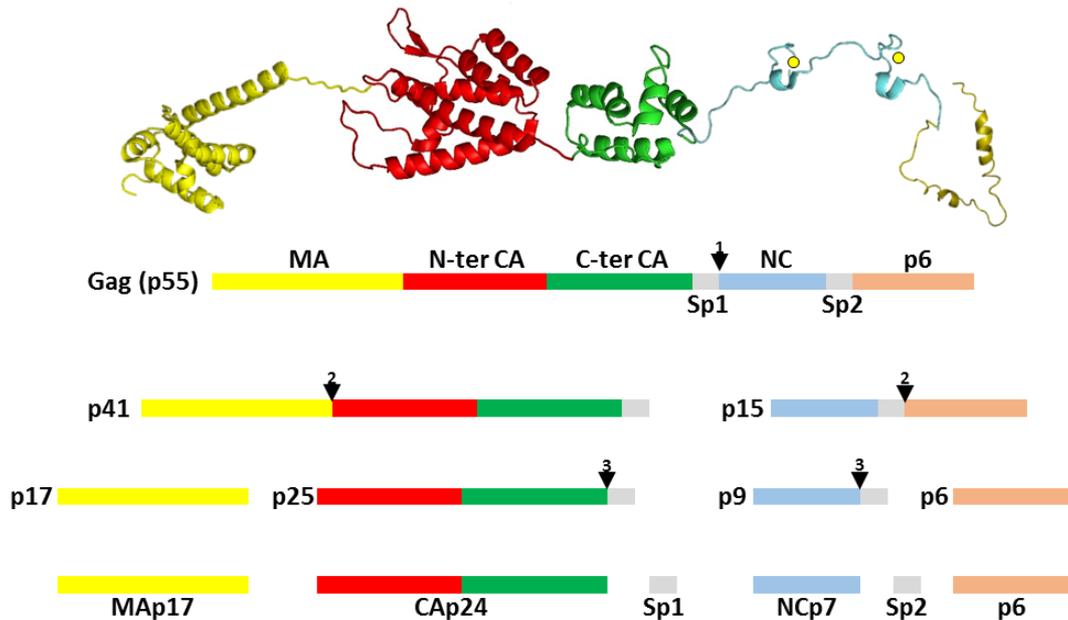
Les régions codantes sont identiques à celles retrouvées dans l'ARN et les régions UTR (5'-UTR et 3'-UTR) sont dupliquées pour former les LTR (Long Terminal Repeat). Ces LTR sont identiques aux deux extrémités de l'ADN proviral et sont composées des régions U3, R et U5. Si cette structure est très importante pour l'intégration du génome viral dans le génome cellulaire, elle contient également tous les signaux indispensables à l'expression des gènes : le promoteur, des éléments activateurs et régulateurs et les sites d'initiation (coiffe) et de terminaison (polyadénylation) de la transcription.

C. Protéines virales : les matériaux de construction pour faire le virus

1. Protéines de structure

a) Polyprotéine Gag

Gag, ou **group specific antigen**, est la polyprotéine structurale majeure du VIH-1 qui est nécessaire et suffisante pour la formation de particules virales sphériques (**Figure 6**). Cette protéine de 55 kDa est composée de quatre domaines qui formeront, après maturation, des protéines aux rôles bien définis (**Figure 10**). Seuls les trois premiers domaines sont indispensables à l'assemblage de 1500 à 3000 polyprotéines Gag sur l'ARN viral nécessaire à la formation d'une particule : la Matrice (MA) permet notamment l'interaction avec les lipides membranaires, la Capside (CA) possède des acides aminés indispensables aux interactions Gag-Gag et la Nucléocapside (NC) permet l'interaction avec l'ARN viral génomique. Le domaine p6 quant à lui, permet le recrutement de partenaires cellulaires qui sont nécessaires au bourgeonnement et à la libération des particules virales (Lingappa et al., 2014). Le rôle de la polyprotéine Gag dans l'assemblage sera étudié dans la partie [III-I] de ce chapitre et les rôles des différents domaines et protéines seront décrites dans cette section. En revanche, la structure tridimensionnelle de la protéine Gag complète reste non élucidée certainement à cause de l'abondance de régions non structurées et flexibles empêchant la cristallisation de Gag.



b) Matrice (MA)

Cette protéine de 17 kDa (132 acides aminés) est composée de cinq hélices α et de trois feuillets β et possède un myristate (modification post-traductionnelle) sur la glycine en 2^{ème} position côté N-terminal. Lors de l'assemblage à la membrane, ce domaine s'assemble sous forme d'un hexamère de trimères autour d'un axe central formant un creux où viendra s'ancrer la protéine transmembranaire d'enveloppe (gp41) (Figure 11) (Alfadhli et al., 2016; Hill et al., 1996; Tedbury et al., 2016). Dans la particule virale mature, cette protéine forme une « seconde membrane » directement située sous la bicouche lipidique (Figure 6).

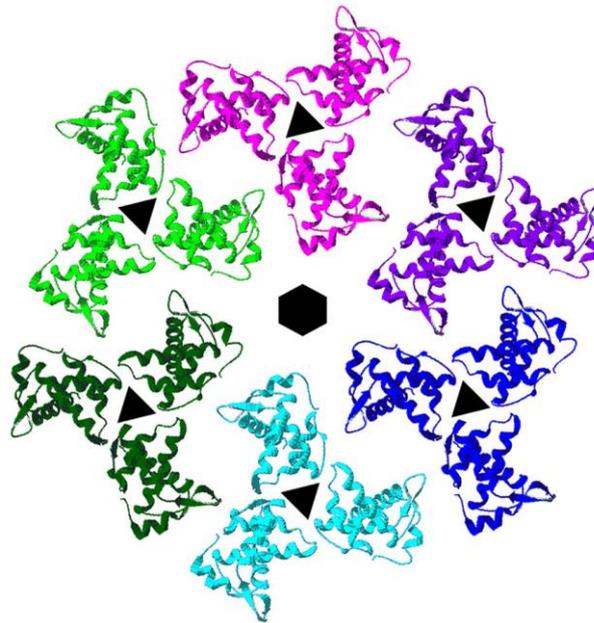


Figure 11 : Organisation membranaire de plusieurs Matrice de HIV-1 (PDB 1HIW). Cette organisation montre la formation d'un hexamère de trimères formant un creux (hexagone noir au centre) dans lequel vient se placer l'ancrage transmembranaire de l'enveloppe (gp41). Les triangles noirs indiquent l'espace autour duquel se forment les trimères de CA (Alfadhli and Barklis, 2014).

Dans les étapes précoces du cycle viral, immédiatement après l'entrée du virus dans le cytoplasme de la cellule infectée, la MA semble intervenir dans le transport vers le noyau du complexe de pré-intégration (PIC) permettant l'infection par le VIH-1 des cellules quiescentes (Kiernan et al., 1998). Ce transport actif du PIC est permis grâce à la présence de deux signaux de localisation nucléaire (NLS) sur la MA (Bukrinsky et al., 1993; Haffar et al., 2000; von Schwedler et al., 1994). Mais l'implication directe de la MA dans le transport du PIC est sujet à controverse et n'est à ce jour, pas totalement élucidée (Fouchier et al., 1997; Freed et al., 1995). Il a également été démontré l'importance de la phosphorylation de résidus sérine de la MA dans les étapes qui suivent immédiatement l'entrée de la CA dans le cytoplasme pour les sérines 9, 67, 72, 77 (Kaushik and Ratner, 2004).

Dans les étapes tardives, le principal rôle décrit du domaine MA de Gag, outre l'incorporation de la protéine d'enveloppe, est son implication dans la liaison avec les lipides membranaires. Plusieurs régions (Lingappa et al., 2014) sont importantes pour que cette interaction soit correcte :

- Acides aminés 2 à 5 : indispensables pour la reconnaissance par la N-myristoyltransférase (NMT) du site de myristoylation. Cette modification post-traductionnelle permet la fixation d'un acide gras saturé (myristate) sur la glycine

en position 2 de la MA lorsque la méthionine N-terminale a préalablement été déléetée. La mutation G2A (remplacement de la Glycine en position 2 par une Alanine) prouve que la myristoylation est indispensable pour la production de particules virales immatures en réduisant sévèrement la liaison de Gag à la membrane et inhibe la production de particules virales (Bryant and Ratner, 1990; Freed et al., 1994; Göttlinger et al., 1989; Zhou et al., 1994). Lorsqu'il est lié à la MA, le myristate existe sous 2 conformations. La première, lorsque la MA est sous forme monomérique, où il est séquestré au sein d'une poche constituée par les hélices-alpha de la MA. La seconde est observée lors de la trimérisation de la MA, où un changement structural permet l'exposition de la chaîne carbonée du myristate, facilitant l'ancrage de Gag dans les lipides membranaires (Chukkapalli and Ono, 2011).

- Acides aminés 7 et 8 : non nécessaires à la myristoylation. Leur mutation entraîne d'importants défauts d'assemblage et de bourgeonnement en diminuant la liaison à la membrane plasmique et inhibe l'exposition du myristate en présence de phospholipides en bloquant la trimérisation de MA (Ono and Freed, 1999; Paillart and Göttlinger, 1999; Saad et al., 2007).
- HBR (Highly Basic Region) : qui, avec l'ancrage par le myristate, est le second signal qui stabilise l'interaction avec les lipides membranaires. Cette zone, composée des acides aminés 15 à 32, permet la formation d'interactions électrostatiques avec la tête polaire des phospholipides membranaires $PI_{(4,5)}P2$, stabilisant la liaison de Gag à la membrane plasmique (Saad et al., 2006). Il apparaîtrait ainsi que la liaison de Gag spécifique au $PI_{(4,5)}P2$ soit médiée par trois mécanismes : (i) exposition du myristate et stabilisation (ii) par le $PI_{(4,5)}P2$ et (iii) par les acides aminés basiques (HBR).
- Acides aminés 85 à 89 : qui assurent avec le domaine HBR le ciblage de Gag spécifiquement à la membrane plasmique qui est le principal site membranaire d'assemblage de Gag pour la formation de particules virales immatures. Ainsi la mutation de ces acides aminés entraîne la relocalisation de l'assemblage sur les membranes intracellulaires (Golgi, réticulum endoplasmique, ...) (Ono et al., 2000a).

Toujours dans les étapes tardives, la liaison de la MA à l'ARN est également importante. En revanche, une compétition entre acides nucléiques et lipides membranaires semble avoir lieu au niveau de la MA car les acides aminés impliqués dans les différentes liaisons se retrouvent tous localisés dans la partie N-terminale de la MA. Ainsi, certains acides aminés se retrouvent impliqués d'une part dans la liaison à l'ARN et d'autre part dans la liaison lipidique (**Figure 12**). La MA semble plutôt lier une séquence C-terminale correspondant au gène *Pol* (Purohit et al., 2001), alors que la nucléocapside permet la liaison de la séquence Ψ située en N-terminale de l'ARN viral génomique. Cette double liaison à l'ARN peut être simultanée, le précurseur Gag empruntant alors une forme compacte dite en U (Datta et al., 2007a; Munro et al., 2014). Cette forme compacte caractéristique pourrait alors être impliquée dans la sélection des lipides membranaires spécifiques ou encore dans la différenciation entre ARNs cellulaires et ARNs viraux. La caractérisation de cette forme compacte étant un de mes projets de thèse, le rôle et la structure de celle-ci seront détaillés dans la partie [IV] de ce chapitre.

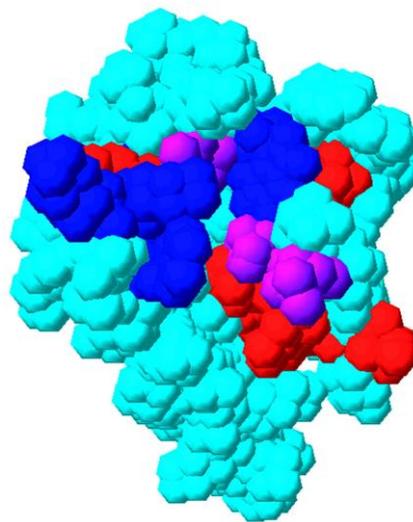


Figure 12 : Chevauchement des sites de liaison préférentiels à l'ARN et au PI(4,5)P2 de la partie N-terminal de la Matrice. Cette illustration de la partie N-ter de la matrice (PDB 1UPH) montre les sites impliqués préférentiellement dans la liaison à l'ARN (en rouge), ou aux phospholipides PI(4,5)P2 (en bleu), ou impliqués dans la liaison avec l'ARN et les phospholipides membranaires (en violet) (Alfadhli and Barklis, 2014).

c) Capside (CA)

La CA est la protéine indispensable à l'oligomérisation du précurseur Gag et permet la formation de la capsid virale conique contenant le génome viral caractéristique des particules virales matures. La CA, protéine de 24 kDa, est sous divisée en deux sous-

domaines reliés par une portion flexible : N-terminal ou CA-NTD (acides aminés 1-145) et C-terminal ou CA-CTD (acides aminés 146-231) (Berthet-Colominas et al., 1999; Gamble et al., 1997; Gitti et al., 1996; Lingappa et al., 2014b). Lors de l'assemblage de CA sous forme d'hexamères (ou de pentamères favorisant les courbes de la base et du sommet de la capsid), le CA-NTD est situé à la surface extérieure de la capsid, alors que le CA-CTD est dirigé vers l'intérieur contenant le génome viral (**Figure 13**). Dans le CA-CTD est retrouvée une zone appelée MHR (Major Homology Region), composée d'une vingtaine d'acides aminés très conservés (de D152 à L172) qui permet la dimérisation de Gag (Doménech et al., 2011). De plus, des protéines Gag portant la double mutation W184A et M185A dans CA-CTD présentent un fort défaut de multimérisation limitant alors la production de particules virales (Burniston et al., 1999; von Schwedler et al., 2003).

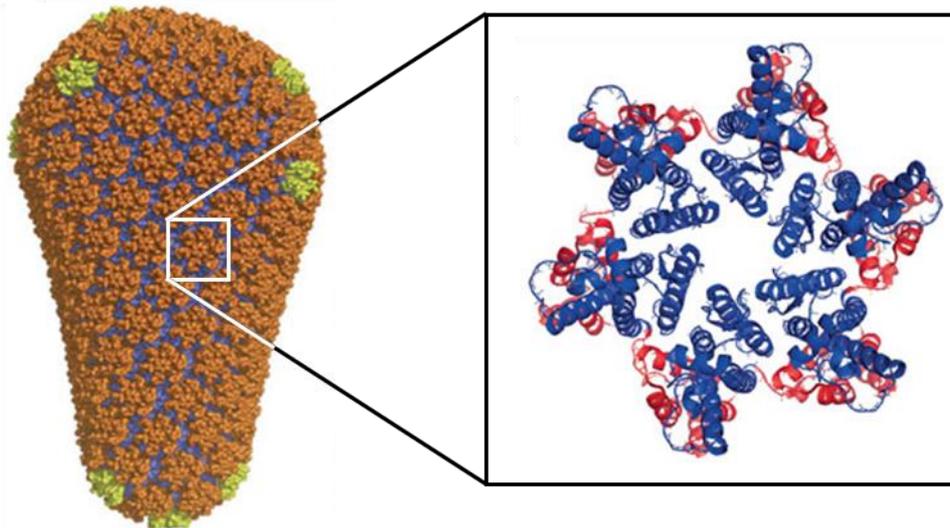


Figure 13 : Représentation schématique de la forme conique empruntée par la capsid virale mature et observation de l'assemblage sous forme hexamérique (existe aussi des formes pentamériques) des protéines de Capsid (adapté de Campbell and Hope, 2015)

La capsid virale a deux rôles majeurs : le premier est un rôle de protection des acides nucléiques contre l'activité de facteurs cellulaires capables de reconnaître des acides nucléiques étrangers (la synthase cyclique des enzymes GMP-AMP (Lahaye et al., 2013) ou TREX1 (three-prime Repair Exonucléase, Yan et al., 2010) ; le second permet de maintenir dans un espace restreint l'ensemble des éléments nécessaires à la transcription inverse (ARN viral, transcriptase inverse...).

d) Nucléocapside (NC)

La NC, protéine de 7 kDa, est une petite protéine dont la structure et la séquence sont très conservées dans les différents sous-types de VIH-1. Dans ce manuscrit, une nomenclature particulière sera utilisée afin de faciliter la compréhension. Ainsi, la NC désigne le domaine NC retrouvé dans le précurseur Gag, alors que la NCp7 permettra de décrire la protéine après maturation. Sa structure caractérisée par deux doigts de zinc de type CX₄CX₂HX₄C permet son activité de liaison spécifique à la boucle SL3 de l'ARN viral (**Figure 14** ; Guzman et al., 1998). Cette protéine étant la cible du premier de mes projets de thèse, sa structure, ses fonctions ainsi que son intérêt comme nouvelle cible thérapeutique seront détaillées dans la partie [VI] de ce chapitre.

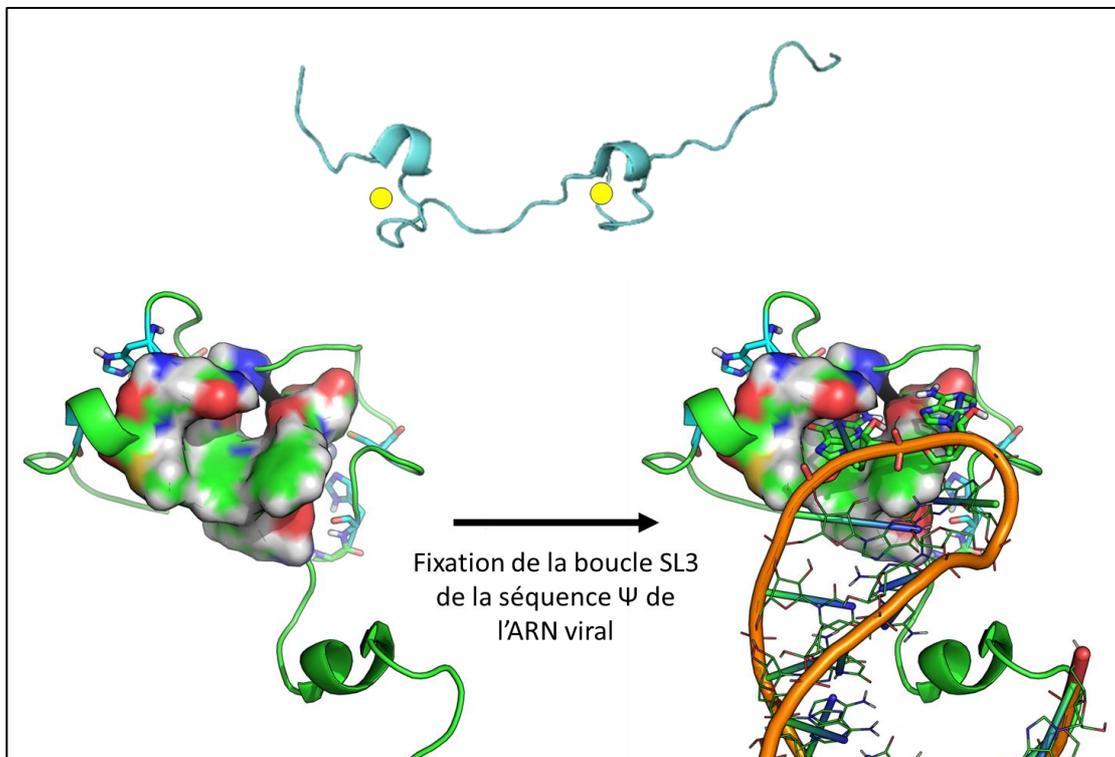


Figure 14 : Structure linéaire de la Nucléocapside (PDB 1AAF) et représentation de la poche hydrophobe formée par les doigts de zinc où vient se fixer la boucle SL3 de la séquence Ψ de l'ARN viral (adapté par M. Mori de Guzman et al., 1998).

e) Protéine p6

Aucun rôle après la maturation n'a été identifié pour ce domaine, d'une masse de 6 kDa, de l'extrémité C-terminale du précurseur Gag. En revanche, en tant que domaine de Gag, p6 permet la fixation de nombreux facteurs cellulaires et viraux indispensables au bourgeonnement de particules virales infectieuses. Parmi eux, sont retrouvés des membres

de la machinerie cellulaire, à savoir les facteurs cellulaires ESCRT (Endosomal sorting Complex Required for Transport) impliqués dans le remodelage et la scission membranaire en particulier dans la formation des corps multivésiculaires. Ces facteurs permettent de réaliser la coupure entre particules virales et membrane cellulaire à la fin du bourgeonnement (environ vingt protéines participent à ce processus). Ces protéines sont principalement recrutées au niveau de deux motifs très conservés de p6 appelés « late domain » (Votteler and Sundquist, 2013) :

- **domaine P(T/S)AP** (code à une lettre de l'acide aminé dans la séquence ; P = proline, T = thréonine ou S = sérine, A = alanine) : connu pour fixer la protéine Tsg-101, membre du complexe ESCRT-I (Garrus et al., 2001; Martin-Serrano et al., 2001; VerPlank et al., 2001) qui permet le recrutement et la fixation des autres protéines du complexe ESCRT. Ce domaine est le plus important des deux « late domain ».
- **domaine YPxL** (code à une lettre également ; Y = tyrosine, x = longueur et nombre d'acides aminés variable, L = leucine) : interagit avec la protéine ALIX qui recrute et active le complexe ESCRT-III (Strack et al., 2003).
- Enfin, son interaction avec la protéine auxiliaire Vpr permet l'embarquement de cette dernière dans la particule virale (Solbak et al., 2013). Son rôle ainsi que celui des partenaires ESCRT sera discuté dans la partie [III.I] de ce chapitre.

2. Protéines enzymatiques

a) Polyprotéine Gag-Pol

Au même titre que le précurseur Gag mais en plus faible quantité, cette polyprotéine de 160 kDa (Pr160) est encapsulée lors de l'assemblage des particules. En revanche, cette protéine est environ vingt fois moins abondante que Gag (Pettit et al., 2005). En effet, la traduction de Gag est obtenue par un décalage du cadre de lecture à l'extrémité 3' du gène *gag* rendant silencieux le codon stop de fin de traduction de Gag (*Figure 15*; Coffin et al., 1997). Sa plus faible concentration ne la rend pourtant pas moins importante, le précurseur Gag-Pol permet, d'abord par clivage auto-protéolytique, l'apport de la protéase virale, puis grâce au clivage réalisé par cette dernière, l'apport des autres enzymes indispensables du VIH-1, la transcriptase inverse et l'intégrase (Pettit et al., 2004). Le rôle de ces différentes protéines sera décrit dans les parties suivantes et les différentes étapes qui découlent de l'activité

enzymatique, à savoir la maturation [Partie III.J], la transcription inverse [Partie III.D] et l'intégration [Partie III.F], seront étudiées plus tard.

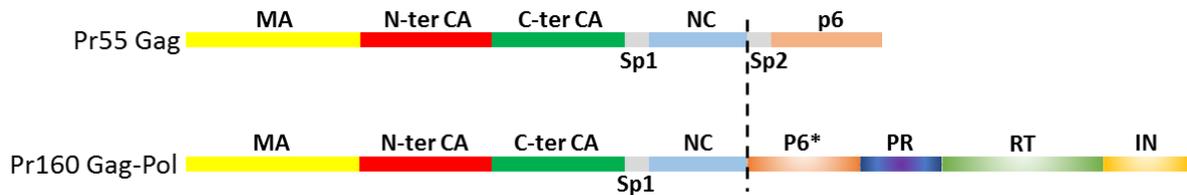


Figure 15 : Comparaison des précurseurs Gag et Gag-Pol. La ligne discontinue indique le décalage du cadre de lecture qui aboutit à la synthèse du précurseur Gag-Pol dans lequel le domaine p6 (indiqué ici p6*) est directement associé au domaine NC sans domaine Sp2 (d'après Leisher et al., 2009; Louis et al., 2013).

b) Protéase (PR)

Cette protéine virale est indispensable à la maturation de la particule virale, la rendant alors infectieuse. Son action de clivage du précurseur Gag-Pol permet alors la libération des protéines de structure et enzymatique. Comme les autres protéases à acide aspartique, la protéase du VIH-1 présente un site actif avec un motif Asp-Thr-Asp dit « fireman's grip » (Cooper, 2002). Toutefois, elle a la particularité, comme celles des autres rétrovirus, de n'être active que sous forme d'homodimères (Wlodawer et al., 1989). Les acides aminés Asp-Thr-Gly permettent l'activation d'une molécule d'eau nucléophile qui clivera alors la liaison peptidique de la protéine cible (Sundquist and Krausslich, 2012). Synthétisée en tant que domaine du précurseur Gag-Pol, la première action de cette protéine est de réaliser un clivage auto-protéolytique selon un mécanisme encore mal connu (Tang et al., 2008). Cette protéine étant indispensable à la formation de particules virales infectieuses, l'inhibition de son activité par des molécules antivirales constitue un axe de recherche important qui a donné naissance à plusieurs molécules actuellement sur le marché (Voir partie [V-B-V.3]) (Kunze et al., 2014).

c) Transcriptase Inverse (RT)

Cette protéine possède deux activités enzymatiques indispensables à la synthèse d'ADN double-brin à partir de l'ARN viral, à savoir les activités polymérase (synthèse de l'ADN proviral) et RNase H (dégradation de l'ARN matrice). Cette étape qui a lieu dans le cytoplasme est appelée transcription inverse et sera discutée dans la partie [III.D] de ce chapitre. La RT est un dimère asymétrique composé des protéines p66 (560 acides aminés, contient les sites actifs des deux activités enzymatique) et p51 (440 acides aminés, permet le maintien de la

structure de la RT) toutes deux issues du clivage de Gag-Pol (Lowe et al., 1988; di Marzo Veronese et al., 1986; Sarafianos et al., 2009). Grâce à la caractérisation de la structure tridimensionnelle de la protéine, il a été décrit que les deux sites actifs de la RT sont situés dans deux zones distinctes (Jacob-Molina et al., 1993; Kohlstaedt et al., 1992). La réplication rapide et l'absence de système de correction sont très importantes pour le virus et pour expliquer les difficultés rencontrées dans la lutte contre l'infection. En effet, ces propriétés contribuent à l'importante diversité génétique du virus rendant, jusqu'à aujourd'hui, le développement de tout vaccin impossible. De plus, ce mécanisme est également à l'origine des résistances observées chez les patients sous traitement rendant ce dernier inefficace à très long terme. Le rôle de cette enzyme étant rapidement devenu un des mieux documentés dans l'historique de la compréhension du VIH-1, il a été la première cible d'inhibiteurs antirétroviraux (description dans la partie [V-B-V.1 et V-B-V.2] de ce manuscrit).

d) Intégrase (IN)

Cette protéine permet de réaliser l'étape d'intégration de l'ADN viral nouvellement synthétisé par la RT dans l'ADN de l'hôte cellulaire. Cette étape sera décrite dans la partie [III.F] de ce chapitre. Cette protéine de 32 kDa (288 acides aminés) est retrouvée sous forme de monomères, mais également de dimères, de tétramères ou de dimères de tétramères (Deprez et al., 2000). La protéine monomérique est composée de trois domaines (Andrake and Skalka, 2015; Bushman et al., 1993; Engelman et al., 1993) :

- Domaine N-terminal (résidus 1 - 49) : contient un motif en doigt de zinc liant deux zinc qui est nécessaire à l'oligomérisation et qui pourrait être impliqué dans l'amélioration de la transcription inverse (Cai et al., 1997),
- Domaine central catalytique (résidus 50 - 212) : très conservé, lie les régions LTR (Esposito and Craigie, 1998) et contient le site catalytique (Aspartates 64 et 116 et Glutamine 152) qui permet ainsi de réaliser les étapes de préparation de l'ADN viral (3'-processing) et de transfert de brin. Pour être actif ce site doit lier deux cations divalents (Mg^{2+} ou Mn^{2+}) (Dyda et al., 1994; Kulkosky et al., 1992),
- Domaine C-terminal (résidus 213 - 288) : qui est le domaine le moins conservé de cette protéine (Eijkelenboom et al., 1995; Lodi et al., 1995) mais permet une liaison non spécifique de l'ADN (Engelman et al., 1994), son import dans le noyau (Gallay et al., 1997) et son oligomérisation par des domaines de type SH3. Ce domaine

permet également la liaison à la RT qui améliore l'activité de transfert de brin de l'Intégrase (Hehl et al., 2004).

Comme les deux autres enzymes précédemment décrites, celle-ci est également la cible de molécules antirétrovirales (Voir partie [V-B-V.6]).

3. Protéines d'enveloppe

a) Polyprotéine Env (gp160)

Les protéines d'enveloppe sont indispensables à l'infection virale car elles permettent la reconnaissance des cellules cibles et la fusion membranaire entre la particule virale et la cellule. Le précurseur polyprotéique (gp160) synthétisé à partir d'un ARN mono-épissé est clivé par les protéases cellulaires (la furine, au niveau de l'appareil de Golgi) pour donner une protéine transmembranaire gp41 et une autre de surface appelée gp120 (Hallenberger et al., 1992; Vollenweider et al., 1996) bien que la furine ne semble pas indispensable à ce clivage (Gu et al., 1995). Ces protéines s'organisent sous forme de dimères qui s'assemblent eux-mêmes sous forme de trimères pour finalement se retrouver à la surface membranaire des particules virales. Le précurseur Env possède un signal N-terminal pour être dirigé dans le réticulum endoplasmique où ce signal sera éliminé. Dans ce dernier, gp160 est alors glycosylée et oligomérisée (souvent sous forme de trimères) facilitant son transport vers l'appareil de Golgi où le précurseur sera protéolysé. Les liaisons non covalentes entre gp41 et gp160 sont en revanche conservées permettant la formation du complexe gp120/gp41 (Checkley et al., 2011). Ce complexe permet l'interaction du virus avec les récepteurs et corécepteurs à la surface des cellules cibles. Le faible nombre de complexes gp41/gp120 à la surface des particules virales (environ 14) contribue à la faible réponse immunitaire induite par l'infection au VIH-1 (Schiller and Chackerian, 2014; Zhu et al., 2003, 2006).

b) Glycoprotéine de surface (gp120)

La glycoprotéine gp120, fortement glycosylée, est composée d'une succession de cinq domaines variables (V1 à V5) et de cinq domaines constants (C1 à C5 ; Starcich et al., 1986; Willey et al., 1986). Un grand nombre de résidus cystéines est retrouvé favorisant la mise en place de ponts disulfures indispensables à la stabilisation de la structure tridimensionnelle du complexe gp41/gp120 qui sera reconnu par le récepteur cellulaire CD4 (Leonard et al., 1990).

V1 et V2 sont stabilisées par un pont disulfure et associées entre elles par un autre pont permettant de former la boucle V1/V2. Cette boucle protège le domaine variable 3 (V3) qui est très important pour la fusion membranaire et dans le tropisme viral (Freed et al., 1990; Hwang et al., 1994; Leonard et al., 1990). Ainsi, c'est sa séquence de 35 résidus et la variation du nombre de charges de ce domaine qui permettent l'interaction avec l'un ou l'autre des co-récepteurs CXCR4 ou CCR5 (Pollakis et al., 2004; Shioda et al., 1994). Concernant les domaines constants, C1 et C5 constituent le site de liaison à la gp41 alors que C2 - C4 sont dans le core de la protéine et nécessaires à la liaison avec CD4). La liaison entre V3 et CCR5 est plus particulièrement ciblée par un inhibiteur d'entrée actuellement sur le marché (Maraviroc® ; (Dragic et al., 2000) ; voir partie [V-B-V.5]). La grande variabilité de cette glycoprotéine, sa capacité de modification de structure et son important degré de glycosylation (Montefiori et al., 1988) lui permettent de ne pas induire de forte réponse immunitaire en évitant d'être reconnue par des anticorps spécifiques (Checkley et al., 2011; Wyatt et al., 1998).

c) Glycoprotéine transmembranaire (gp41)

Cette glycoprotéine permet l'ancrage des spicules d'Env dans le même hôte, l'incorporation d'Env au sein des particules virales et la fusion entre les membranes de la cellule cible et de la particule virale. Elle est composée de trois domaines majeurs (Checkley et al., 2011; Lu et al., 1995; Wyatt et al., 1998) :

- **Domaine extracellulaire ou ectodomaine** (Weissenhorn et al., 1997) : présente une région hydrophobe N-terminale connue comme étant le peptide de fusion (Bosch et al., 1989) suivie d'une région polaire et de deux régions hydrophobes appelées «heptad-repeat , HR» puis d'un domaine riche en tryptophane nommé MPER (Membrane Proximal External Region, aussi appelé HR2) impliqués dans le processus de fusion (Dubay et al., 1992; Muñoz-Barroso et al., 1999; Salzwedel et al., 1999). C'est le changement structurel engendré par la fixation de gp120 au récepteurs CD4 et aux co-récepteurs qui permet l'exposition du domaine de fusion qui pénètre dans la membrane de la cellule hôte, déstabilisant celle-ci formant alors un pore par lequel passera la capsid virale (Brasseur et al., 1990). Un peptide dérivé du domaine HR2, l'enfuvirtide ou fuzeon, est actuellement sur le marché (voir partie [V-B-V.4]).

- **Domaine transmembranaire** : constitué de 40 acides aminés environ dont 25 sont très conservés, permet au complexe gp120/gp41 de s'insérer dans la bicouche lipidique membranaire.
- **Domaine C-terminal cytoplasmique** : ce domaine est composé de plusieurs motifs présentant des activités structurales et fonctionnelles :
 - Séquence YxxL qui interagit avec le complexe adaptateur pour la clathrine (AP-2, Adaptor Protein 2) permettant l'endocytose clathrine-dépendante du virion, le bourgeonnement viral dans les cellules épithéliales (Lodge et al., 1997), le passage de cellule T à cellule T (Deschambeault et al., 1999) et l'infectivité virale (Day et al., 2004).
 - Motif di-leucine qui interagit avec la protéine AP-1, autre complexe permettant l'endocytose clathrine-dépendante indépendamment du motif YxxL (Byland et al., 2007; Wyss et al., 2001).
 - Trois hélices amphipathiques (LLP-1, LLP-2 et LLP-3), souvent citées comme peptides lytiques, structurellement très conservées et impliquées dans la stabilité d'Env, sa capacité de fusion (Kalia et al., 2003), sa multimérisation (Lee, 2000), son expression membranaire (Bültmann et al., 2001) et son incorporation dans la particule virale (Piller et al., 2000).

4. Protéines régulatrices

a) Protéine Tat (*Trans-activator of Transcription*)

La protéine Tat, encodée par deux exons différents (rapprochés par un double épissage), est une petite protéine dont la masse varie de 14 à 16 kDa (suivant la souche virale considérée) jouant un rôle crucial dans l'activation de la transcription du VIH-1. C'est une des premières protéines exprimées lors de l'infection qui reconnaît la séquence TAR à l'extrémité 5' de l'ARN viral en cours de transcription (**Figure 8**). Tat recrute alors des protéines (Cyclin dépendent kinase 9 et Cyclin T1) du complexe P-TEFb qui se retrouvent proches de l'ARN-polymérase-II qui, alors débloquée de sa pause, peut reprendre la synthèse de l'ARN (Bagashev and Sawaya, 2013; Mancebo et al., 1997; Zhu et al., 1997). Tat est composée de trois domaines fonctionnels dont le :

- Domaine de transduction (PTD) : permet l'internalisation de Tat dans les cellules (domaine minimal a.a. 47 - 57). Ce motif a été utilisé comme CPP (cell penetrating peptide) pour importer dans des cellules de nombreux peptides, protéines ou molécules chimiques (Mäe and Langel, 2006).
- Domaine de transactivation : situé en position N-terminale, riche en prolines et cystéines, interagit avec de nombreuses protéines cellulaires et est essentiel à l'activité de transactivation de la transcription, à la liaison de zinc et à la dimérisation de Tat (Frankel and Pabo, 1988; Johri et al., 2011).
- Domaine riche en arginine : permet l'interaction avec la séquence TAR de l'ARN viral (Berkhout et al., 1989).

Tat est également impliquée dans l'augmentation de l'expression de protéines cellulaires dans les cellules infectées comme plusieurs cytokines (IL-6, 8 et 10 ; Buonaguro et al., 1992; Ott et al., 1998), le co-récepteur CCR5 (Huang et al., 1998a), ou le récepteur à l'interleukine-2 (CD25 ; Mayol et al., 2007). À l'inverse, Tat permet la diminution de l'expression d'un grand nombre de protéines cellulaires dont des molécules du complexe majeur d'histocompatibilité de classe I (CMH-I) indispensables pour une réponse immunitaire correcte (De La Fuente et al., 2002; Howcroft et al., 1993). De plus, comme Tat est libérée dans l'espace intercellulaire par les cellules infectées, elle a également montré une capacité à induire l'angiogenèse (Albini et al., 1996), la prolifération des cellules de Kaposi (Ensoli et al., 1990) ou l'apoptose de cellules T non infectées (Li et al., 1995; Westendorp et al., 1995) en agissant sur des récepteurs extracellulaires ou après entrée dans ces dernières. Il a également été montré que Tat augmente l'expression de la protéine MMP9 connue pour dégrader la matrice extracellulaire de la barrière cérébrale entraînant un désordre neurologique causant des démences dans de plus de 50 % des patients touchés par le SIDA (Ju et al., 2009; Romani et al., 2010).

Comme cette protéine joue un rôle critique dans la transcription du génome viral, elle constitue une cible intéressante pour le développement de nouveaux inhibiteurs ou même pour le développement d'une nouvelle stratégie vaccinale (Programme de recherche TAT/ BIOSANTECH SA®, Loret et al., 2016).

b) Protéine Rev (Regulatory of Viral Expression)

Cette phosphoprotéine de 116 acides aminés fait la navette entre le noyau et le cytoplasme de la cellule (Pollard and Malim, 1998). Son extrémité N-terminale se fixe sur la séquence RRE

(Rev Response Element) des ARNs viraux permettant l'export nucléaire des ARNm viraux non ou mono-épissés (Böhnlein et al., 1991; Daly et al., 1989). En effet, en condition normale, les cellules eucaryotes régulent l'expression et contrôlent l'export nucléaire des ARNm correctement épissés. Le système d'export NFX1 est utilisé par les ARNm viraux totalement épissés (Cullen, 2003). Toutefois, afin d'exporter de grandes quantités d'ARNm non totalement épissés, comme le génome viral, le VIH-1 utilise le système Rev/RRE pour outrepasser les différents points de régulations cellulaires. Le complexe Rev/RRE est ainsi reconnu au niveau du signal d'export nucléaire riche en leucine de Rev par les protéines cellulaires CRM1 (Chromosome Maintenance 1) lié à Ran-GTP formant un complexe d'export ribonucléoprotéique (RNP) (Fischer et al., 1995; Fornerod et al., 1997; Williamson, 2015). Finalement, il a récemment été démontré que la structure secondaire adoptée par la séquence RRE permet de contrôler l'activation de la réplication (Sherpa et al., 2015).

5. Protéines auxiliaires

a) Protéine Nef (*Negative Regulatory Factor*)

Cette protéine de 27 kDa, myristoylée en N-terminal (Harris, 1995), est, conjointement aux protéines Rev et Tat, une des premières protéines synthétisées lors de l'infection par VIH-1 (Klotman et al., 1991). Nef est retrouvée dans le cytoplasme ou associée à la membrane cellulaire. Elle permet la régulation de plusieurs molécules de surfaces favorisant ainsi la dissémination virale. Parmi ces molécules sont retrouvés (Kirchhoff et al., 2008) :

- Le récepteur CD4, dont l'inhibition est également obtenue par l'action d'autres protéines virales (Vpu et Env). Nef favoriserait dès le début de l'infection, en recrutant les protéines AP-2 et ALIX (Amorim et al., 2014; Greenberg et al., 1997), l'endocytose des récepteurs CD4 membranaires déjà exprimés et entrainerait leur dégradation par la voie lysosomale (Anderson et al., 1994; Rhee and Marsh, 1994).
- Les complexes majeurs d'histocompatibilité de classe I et II (CMH-I et -II) : ces complexes présentent des antigènes étrangers à la surface des cellules présentatrices d'antigènes qui sont reconnus par les lymphocytes T cytotoxiques (LT CD8⁺) et provoquent la mort cellulaire des cellules infectées (Schwartz et al., 1996). Nef agirait sur deux fronts : d'abord elle favoriserait l'endocytose des complexes du CMH-I retrouvés à la surface cellulaire selon un processus dépendant

- du recrutement des protéines AP-1, PACS-2 et kinases de la famille Src (Atkins et al., 2008; Hung et al., 2007). Puis, Nef bloquerait le transport vers la membrane des CMH-I nouvellement synthétisés (Kasper et al., 2005; Pawlak and Dikeakos, 2015).
- Différents récepteurs de lymphocytes T comme CD28 (Bell et al., 2001; Swigut et al., 2001) ou CXCR4 (Hrecka et al., 2005) inhibant ainsi la réponse antigénique spécifique des lymphocytes T.

b) Protéine Vif (Viral Infectivity Factor)

Cette protéine basique et phosphorylée de 23 kDa est essentielle pour la réplication virale dans les cellules telles que les lymphocytes et les macrophages mais dispensable dans d'autres lignées cellulaires dites permissives (HeLa, 293T, Cos7, ...). Moins évident est le fait que la nature de la cellule (permissive ou non) produisant le virus est plus importante pour l'action de Vif que la cellule infectée elle-même. En effet, un virus produit en cellule permissive est capable d'infecter une cellule non permissive. Cette particularité est due à l'expression dans les cellules non permissives du facteur de restriction cellulaire APOBEC3G (effet similaire pour APOBEC3F) (Sheehy et al., 2002). APOBEC3G est une cytidine déaminase qui mute les cytosines en uraciles dans l'ADN produit lors de la transcription inverse entraînant des mis-appariements qui sont détectés et détruits par des systèmes de réparation de l'ADN inhibant la synthèse du provirus (Mangeat et al., 2003; Zhang et al., 2003). Vif interagit avec APOBEC ce qui entraîne la dégradation protéasomale d'APOBEC3G (Rose et al., 2004) limitant son incorporation dans les virions permettant alors de contrer l'action d'APOBEC3G (Henriet et al., 2009; Mariani et al., 2003; Marin et al., 2003). Vif parvient également à inhiber la traduction de APOBEC3G et 3F.

c) Protéine Vpr (Viral Protein R)

Cette protéine de 14 kDa, exprimée tardivement au cours de l'infection semble jouer des rôles importants dans les étapes précoces de l'infection virale. En effet, cette protéine auxiliaire est incorporée dans les particules virales via les domaines p6 et NC de Gag (Bachand et al., 1999; Fritz et al., 2008, 2010; de Rocquigny et al., 1997; Salgado et al., 2009). Cette protéine a plusieurs fonctions (Guenzel et al., 2014; Morellet et al., 2009) :

- Induction de l'arrêt du cycle cellulaire en phase G2 bloquant ainsi la prolifération des lymphocytes T (He et al., 1995; Jowett et al., 1995). En effet, cet arrêt

favoriserait la transcription virale car les LTR du VIH-1 sont plus actifs en phase G2 (Goh et al., 1998) et, pendant cette phase, la traduction est augmentée stimulant la production virale.

- Induction de l'apoptose des lymphocytes T infectés (pas des macrophages) conduisant à la diminution drastique de leur nombre dans les phases précoces de l'infection.
- Amélioration de la fidélité de la transcription inverse. Ainsi, en l'absence de Vpr, l'incorporation de mutations est quatre fois plus élevée lors de l'étape de transcription inverse qu'en présence de Vpr (Mansky, 1996). Vpr interagit avec l'uracil DNA glycosylase (UNG2) une protéine d'un système de réparation de l'ADN permettant le remplacement d'uraciles (normalement retrouvés uniquement dans l'ARN) incorporés par erreur dans l'ADN lors de la transcription inverse ou obtenus par déamination de cytosines (Mansky et al., 2000).
- Import nucléaire de l'ADN viral en étant membre à part entière du complexe de pré-intégration (Jenkins et al., 1998; Lyonnais et al., 2013; McDonald et al., 2002) et notamment dans le contact de ce complexe avec les pores nucléaires (Fouchier et al., 1998; Piller, 2001).

d) Protéine Vpu (Viral protein U)

Cette protéine de 16 kDa est composée d'un domaine N-terminal transmembranaire et d'un domaine intra-cytoplasmique dont une partie, située entre deux hélices α , est hautement conservée et phosphorylée. Plusieurs activités sont attribuées à Vpu (Bour and Strebel, 2003; González, 2015) :

- Dégradation des récepteurs CD4 par un processus impliquant le protéasome. Cette dégradation cible les récepteurs CD4 bloqués dans le réticulum endoplasmique par leur association avec la protéine d'enveloppe gp160. Ainsi, la dégradation des CD4 par Vpu permet à Env de poursuivre son trafic vers la membrane cellulaire (Willey et al., 1992a, 1992b).
- Stimulation de la libération de particules virales. Vpu est capable par un mécanisme mal défini à ce jour mais impliquant le domaine transmembranaire d'inhiber la protéine tétherine, protéine transmembranaire capable de former une attache pour les virus à la membrane cellulaire (Neil et al., 2008; Van Damme et al., 2008).

III. Comment le VIH se multiplie-t-il ?

A. Le cycle répliatif : se multiplier pour mieux infecter

L'objectif principal du virus étant de se multiplier, il a mis en place de nombreuses stratégies lui permettant d'utiliser les outils cellulaires qui lui sont indispensables pour atteindre cet objectif. Il se sert du système de défense immunitaire pour infecter les cellules qui le composent, puis contourne les systèmes de contrôle que ces cellules utilisent pour se protéger et finalement laisse périr cette cellule tant aimée pour en infecter d'autres.

On peut diviser le cycle en deux grandes étapes (*Figure 16*) :

- **Étapes précoces** : comprenant les phases de contact et d'entrée dans la cellule [B], de décapsidation [C], de transcription inverse [D], d'import nucléaire [E] et d'intégration [F]. Ces étapes permettent ainsi d'entrer dans la cellule pour insérer le génome viral dans l'ADN cellulaire.
- **Étapes tardives** : comprenant les étapes qui suivent l'intégration du génome viral, à savoir, la transcription du génome viral en divers ARNs, leur épissage et transport du noyau vers le cytoplasme [G], la traduction des protéines virales [H], l'assemblage permettant le bourgeonnement de nouvelles particules [I] et finalement la maturation [J]. Ces étapes permettent ainsi l'utilisation de la machinerie cellulaire pour la production de nouvelles particules virales capables d'infecter d'autres cellules.

Les mécanismes de chacune des étapes citées précédemment seront étudiés dans les sections suivantes.

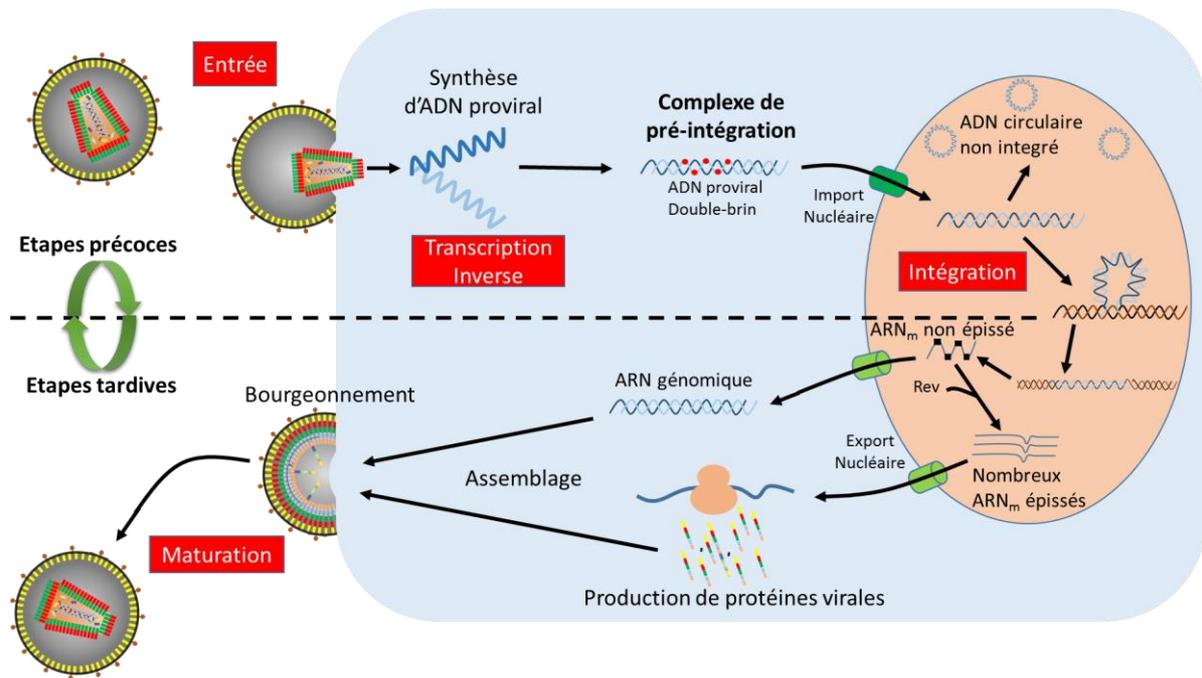


Figure 16 : Cycle réplcatif du VIH-1 dans lequel on peut distinguer les étapes précoces et les étapes tardives. Afin de rendre ce cycle facilement compréhensible, certaines étapes sont ici simplifiées. Le détail de chacune de ces étapes sera étudié dans les parties et figures suivantes (Adapté de Coiras et al., 2009).

B. Contact entre le virus et la cellule : le baiser mortel

Le contact entre la particule virale et la cellule cible est une étape cruciale pour l'infection par le VIH-1. Il débute par un premier contact souvent peu spécifique entre Env et des protéoglycanes (sur les macrophages ; Saphire et al., 2001) ou plus spécifique avec des intégrines ou d'autres molécules comme DC-SIGN (Dendritic Cell Specific Intercellular adhesion molecular-3-Grabbing Non-integrin ; Geijtenbeek et al., 2000). Des contacts entre des protéines cellulaires incluses dans la membrane virale et des composants de la cellule peuvent aussi avoir lieu comme entre ICAM-1 et LFA-1 par exemple (Fortin et al., 1998). Après ce premier contact, les trimères de dimères gp41/gp120 à la surface des particules virales sont reconnus, via la glycoprotéine gp120, par le récepteur CD4 (Maddon et al., 1986; McDougal et al., 1986, dont le rôle physiologique majeur est de participer au processus de stimulation des lymphocytes T auxiliaires) des cellules cibles : lymphocytes T CD4⁺, macrophages, microglie ou cellules dendritiques. Cette interaction induit un changement conformationnel de gp120 exposant alors les chaînes variables qui la composent et en particulier le domaine V3 qui est reconnu par des récepteurs aux chimiokines (CCR5 et CXCR4, récepteurs couplés

aux protéines G caractérisés par sept domaines transmembranaires) considérés comme des co-récepteurs du VIH (Huang et al., 2005; Kwong et al., 1998 ; **Figure 17 A**). L'utilisation de l'un ou l'autre des co-récepteurs permet de définir le tropisme du virus. Il existe ainsi des virus R5 qui utilisent principalement le récepteur à la chimiokine CCR5 et des virus X4 qui eux se fixent plutôt sur le co-récepteur CXCR4. Finalement les souches virales R5X4 peuvent utiliser les deux types de co-récepteurs. Généralement, en début d'infection virale, on retrouve principalement des VIH-1 à tropisme R5 (malgré la forte expression de CXCR4 dans les cellules cibles du VIH), évoluant vers un double tropisme, pour finalement observer majoritairement des virus à tropisme X4 dans les étapes tardives de la maladie (Regoes and Bonhoeffer, 2005). Après fixation des récepteurs et co-récepteurs, le virus migrerait le long de la membrane cellulaire pour atteindre les sites de fusion (Lehmann et al., 2005) où Env subit une nouvelle modification structurale qui induit l'exposition du peptide hydrophobe de fusion de gp41 (**Figure 17 B**). L'exposition du peptide de fusion entraîne le rapprochement de l'hélice- α N-terminale d'une gp41 et de l'hélice- α C-terminale d'une autre gp41 du trimère formant alors au sein de chaque trimère un faisceau composé de 6 hélice- α . Les hélices en N-terminal étant proches de la membrane cellulaire et les hélices en C-terminal de la membrane virale, la formation du faisceau entraîne d'abord le rapprochement des deux membranes, puis la formation du pore membranaire et finalement l'entrée de la capsid virale dans le cytoplasme cellulaire (Eckert and Kim, 2001; Melikyan, 2008).

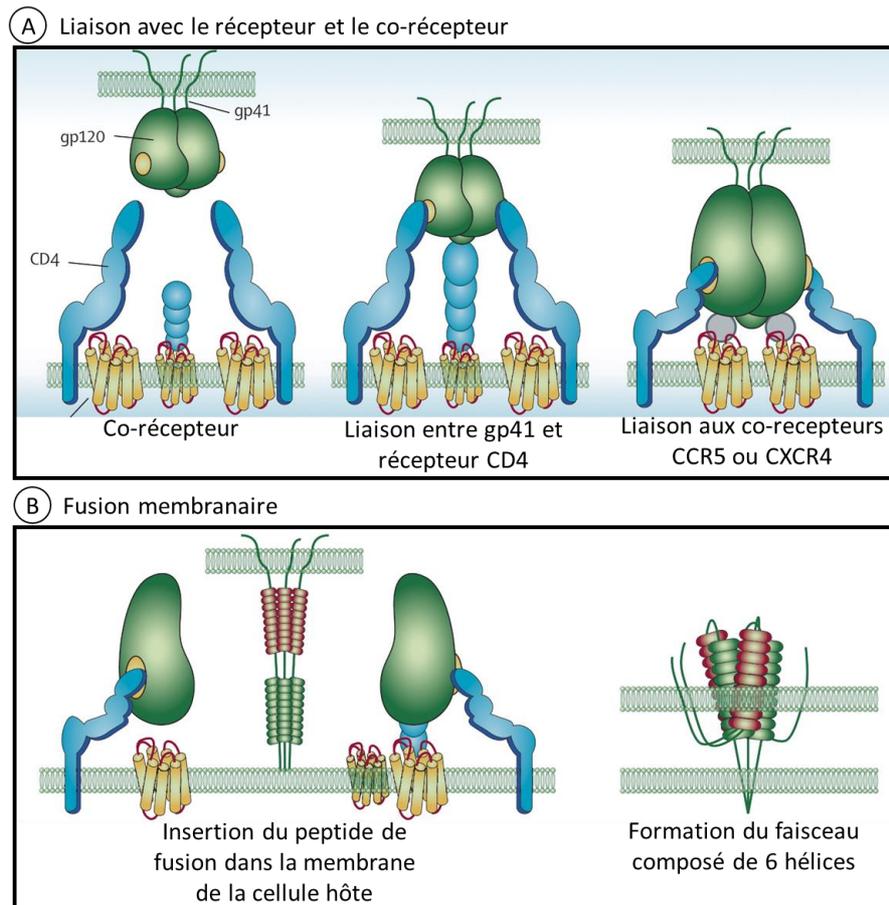


Figure 17 : Entrée du VIH-1 dans la cellule. On peut ainsi distinguer deux étapes : (A) liaison avec le récepteur CD4 suivie de celle avec le co-récepteur aux chimiokines CCR5 ou CXCR4 et (B) fusion membranaire par insertion du peptide de fusion, porté par la gp41, dans la membrane cellulaire et création d'un pore membranaire par formation d'un faisceau composé de six hélices- α provenant du trimère de gp41 (traduit de Esté and Telenti, 2007).

Ce mécanisme dépendant des récepteurs et corécepteurs n'a pas le monopole de l'infection. Ainsi, malgré les controverses, il est évident qu'il existe une transmission de cellules à cellules au travers d'une structure appelée « synapse virologique » entre cellules immunitaires (Jolly et al., 2004, 2011; McDonald et al., 2003). Il apparaît également que cette transmission de cellules à cellules, qui implique un phénomène d'endocytose dynamine dépendante (Sloan et al., 2013) reste malgré tout sensible à l'action des inhibiteurs d'entrée et des anticorps neutralisants (Martin et al., 2010). Cette transmission cellule-cellule est prédominante (en comparaison avec la transmission classique) dans les tissus lymphatiques où les virus sont présents en grande quantité.

L'entrée du virus dans la cellule étant la première étape du cycle viral, elle constitue évidemment une cible thérapeutique très intéressante. Ainsi, plusieurs stratégies ont été développées afin d'inhiber l'entrée du VIH-1 dans la cellule cible :

- Cibler la **gp41**, comme par exemple avec l'Enfuvirtide®, voir la partie [V-BV.4],
- Cibler le **CCR5**, comme par exemple par le Maraviroc®, voir la partie [V-B-V.5],
- Cibler le **CXCR4**, aucune molécule n'est actuellement sur le marché pour cibler la liaison entre gp120 et ce co-récepteur. Plusieurs laboratoires se penchent sur la question mais la toxicité cardiaque et le faible effet clinique de l'AMD-3100 ont stoppé les développements des molécules de cette classe (Hendrix et al., 2004).
- Cibler le **CD4**, récepteur permettant la fixation du virus à la membrane de la cellule, a rapidement constitué une cible de recherche avec par exemple le développement d'une forme soluble de CD4 qui pourrait se fixer préférentiellement à gp120 et entrer en compétition avec le CD4 à la surface de la cellule. Or cette stratégie s'est avérée complètement inefficace chez l'Homme (Daar et al., 1990). En revanche, récemment, un inhibiteur (BMS-663068, basé sur la structure de la gp120 lié au CD4) a été développé et l'étude clinique NCT01009814 est toujours en cours (Nettles et al., 2012). Cette voie est toujours le sujet d'importants travaux comme le montre la description d'un nouvel inhibiteur de l'interaction entre gp120 et CD4 qui bloque l'entrée du virus (Heredia et al., 2015).

C. Entrée et décapsidation : attaquer le « rempart » et rentrer dans la « cour » cellulaire

L'entrée du virus dans la cellule hôte correspond au passage de la capsid virale contenant le génome viral, à travers le pore membranaire. La décapsidation correspond littéralement à l'ouverture de la capsid virale libérant alors le génome viral qui est prêt à subir la transcription inverse. Une des caractéristiques du VIH-1 (et des autres lentivirus) est leur capacité à infecter des cellules qui ne sont pas en division. La difficulté est ainsi de faire entrer le génome viral dans le noyau pour qu'il soit intégré dans le génome cellulaire.

La capsid constituée, après clivage protéolytique de Gag par la protéase, d'environ 1500 monomères de CA est caractérisée par une forme conique d'environ 100 nm de long et de 52 nm de diamètre à sa plus grande base (Welker et al., 2000). Les 1500 protéines de CA sont, selon un modèle présenté récemment, assemblées en 216 hexamères et 12 pentamères (Zhao et al., 2013). Sa forme conique caractéristique contient deux copies de l'ARN viral recouvertes de protéines de NCp7, associées aux enzymes virales (transcriptase inverse et intégrase) et à

des protéines accessoires. Cette capsidite permet ainsi de protéger le génome viral des facteurs cellulaires potentiellement néfastes pour lui et permet également de maintenir dans un environnement proche l'ARN et la transcriptase inverse qui pourra alors produire l'ADN viral. Le dogme voulait que la décapsidation précède la transcription inverse. Or plusieurs rôles impliquant la CA dans l'import nucléaire (Matreyek and Engelman, 2011) et l'intégration du génome viral dans le génome cellulaire (Koh et al., 2013) montrent l'implication de la CA dans plusieurs étapes de la phase précoce du cycle viral. La décapsidation semble donc être incomplète tant que le génome viral n'est pas entré dans le noyau. En effet, la taille des éléments qui passent au travers d'un pore nucléaire n'excède pas 39 nm (Panté and Kann, 2002) alors que la taille de la capsidite mesure 52 nm de diamètre (au plus large), montrant qu'une capsidite complète ne peut pas rentrer dans le noyau. Plusieurs modèles de décapsidation ont donc été proposés (Campbell and Hope, 2015 ; *Figure 18*) :

- **Décapsidation immédiate** (*Figure 18 A*) qui pencherait en faveur d'une décapsidation totale immédiatement après l'entrée de la capsidite dans le cytoplasme cellulaire. Ce modèle décrit depuis plusieurs années est basé sur l'absence d'étude montrant la présence de protéine de CA dans le complexe de transcription inverse (Karageorgos et al., 1993). Cette absence est probablement causée par un artefact d'expérience dans les procédures de purification qui supprimerait la majeure partie des protéines (CA) composant la capsidite. De plus, plusieurs observations remettent en question ce modèle, comme le rôle de la capsidite dans la protection du génome viral (empêche la reconnaissance par les senseurs d'ADN qui peuvent déclencher une réponse de l'immunité innée ; Lahaye et al., 2013) ou son implication dans l'infection de cellules quiescentes (Yamashita and Emerman, 2004; Yamashita et al., 2007).
- **Décapsidation cytoplasmique** (*Figure 18 B*) où une perte progressive de nombreuses CA se fait lors du trafic vers le noyau cellulaire. Plusieurs techniques expérimentales permettent d'avancer ce modèle (McDonald et al., 2002; Perez-Caballero et al., 2005; Xu et al., 2013). Puisque l'intégrité de la capsidite diminue avec la perte continue de CA, ce modèle est contradictoire avec le rôle de la capsidite dans la protection du génome viral. Il est en revanche possible que cette

décapsidation progressive couplée au recrutement de différents facteurs cellulaires permette de prolonger le rôle protecteur de la capsid.

- **Décapsidation dans le pore nucléaire (Figure 18 C)** où l'intégrité de la capsid reste totale tout au long du trafic vers le noyau où il y a libération du génome viral dans le noyau au travers du pore nucléaire. Ce modèle est celui qui correspond le plus au rôle protecteur du génome viral joué par la capsid (les différents facteurs cellulaires sont incapables d'atteindre le génome viral). La présence de capsid intacte au niveau du pore nucléaire penche ainsi en faveur d'une décapsidation au niveau du pore nucléaire favorisant alors l'entrée de l'ADN viral dans le noyau cellulaire (Arhel et al., 2007). Enfin, l'interaction de la CA avec les protéines du pore nucléaire (NUP358 et NUP153) montrent également la pertinence de ce modèle.

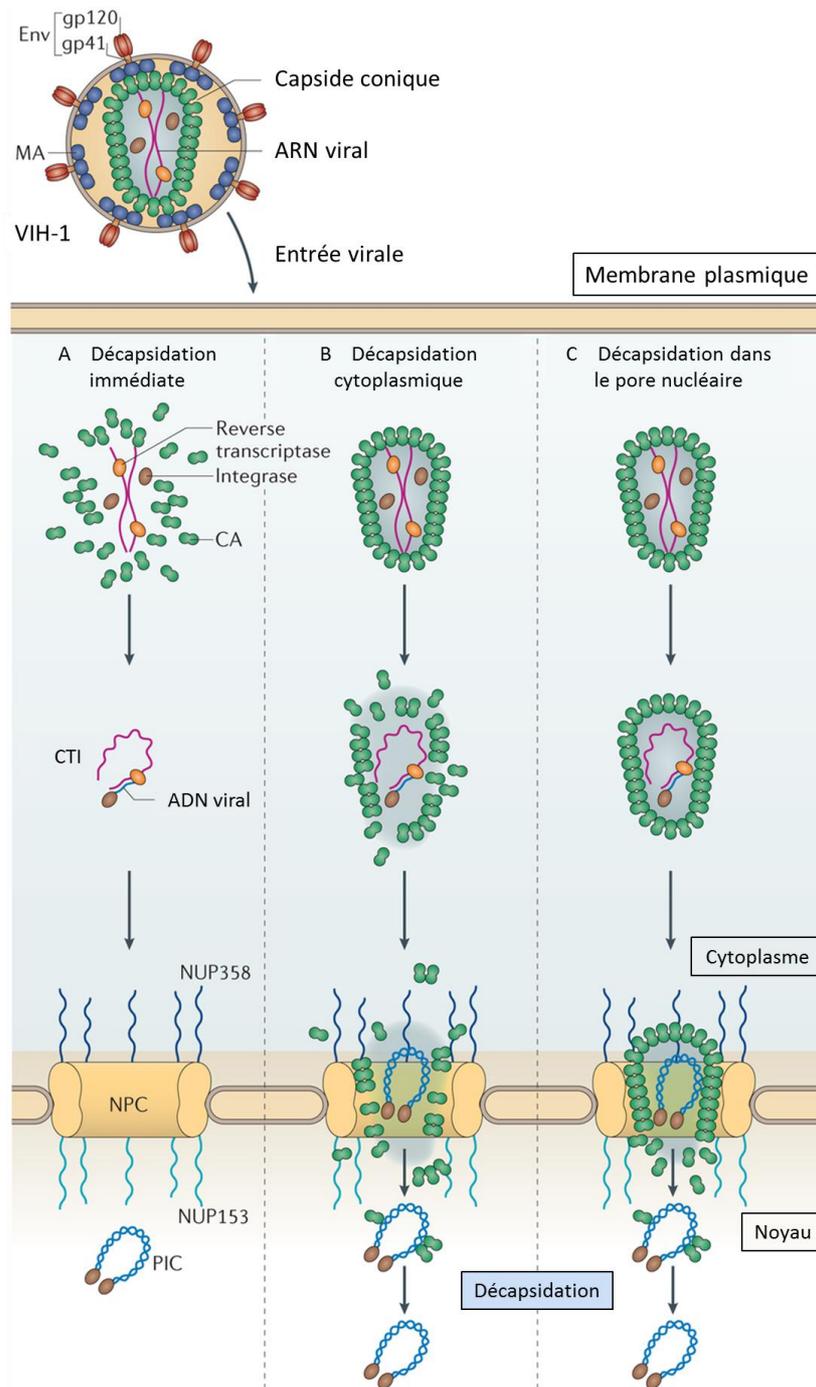


Figure 18 : Différents modèles de décapsidation actuellement proposés. Ces différents modèles ne sont pas exclusifs et leurs complémentarités peuvent contribuer à expliquer les différences observées entre les divers résultats expérimentaux visant à prouver chaque modèle indépendamment (Traduit de Campbell and Hope, 2015).

Ces trois modèles, plus complémentaires qu'exclusifs, montrent que le phénomène de décapsidation n'est pas encore complètement compris. Malgré cela, la protéine CA est une cible pour le développement d'inhibiteurs antiviraux. Ainsi, malgré des mécanismes d'actions différents, PF74 (Shi et al., 2011) ou BI-2 (Lamorte et al., 2013) permettent la déstabilisation

de la capsid virale. Les différentes implications de la capsid dans les étapes précoces du cycle viral montrent que c'est une cible thérapeutique très intéressante. En revanche, pour développer une molécule efficace, il est indispensable de comprendre les différents mécanismes de recrutement de partenaires et de décapsidation qui ne sont pour l'instant que des spéculations même si des études récentes montrent de plus en plus de précisions sur ce mécanisme (Ma et al., 2016).

D. Transcription inverse : quand l'ARN devient ADN

La transcription inverse est l'étape indispensable qui permet la production d'ADN double brin viral à partir de l'ARN (+) viral retrouvé dans les particules infectieuses. Cette étape est réalisée par la transcriptase inverse (RT pour Reverse Transcriptase en anglais) qui possède à la fois une activité ADN polymérase (synthèse d'ADN) et une activité RNase H (dégradation de l'ARN). La RT agit au sein du complexe de transcription inverse composé de deux copies d'ARN viral, d'ARNt^{Lys3}, des protéines virales MA, CA, NC, RT, IN, Vpr et de protéines cellulaires dont l'Uracil DNA glycosylase (UNG2), la Cyclophiline A (CypA) et l'Heat Shock Protein 70 (HSP70) (Iordanskiy and Bukrinsky, 2007). Lors de la transcription inverse le complexe évolue alors en un complexe appelé complexe de pré-intégration (PIC, description dans la partie [III-E] de ce manuscrit). La transcription inverse peut être divisée en plusieurs étapes (**Figure 19**) (Cimarelli and Darlix, 2014; Onafuwa-Nuga and Telesnitsky, 2009; Sarafianos et al., 2009) :

- Liaison de l'ARNt^{Lys3} : cette amorce, empruntée à la cellule productrice, est retrouvée dans les particules virales. Elle se lie de façon complémentaire à la séquence PBS située sur l'ARN viral (**Figure 19 A**) ;
- Synthèse du brin (-) d'ADN « strong-stop » (région U5 et R) par l'activité polymérase 3'→5' de la RT qui a reconnu le complexe ARN/ARNt^{Lys3} et dégradation de la partie correspondante de l'ARN matrice par l'activité RNaseH de la RT (**Figure 19 B**) ;
- Premier saut de brin de l'ADN « strong-stop » et de l'amorce qui se repositionne au niveau de l'extrémité 3' terminale de l'ARN par appariement à la séquence complémentaire R. Il y a alors synthèse complète du brin d'ADN (-) à partir de l'ARN viral et dégradation de celui-ci par la RT (**Figure 19 C**) ;

- Synthèse du brin (+) d'ADN en utilisant comme amorces des régions ARNs résistantes à l'action de la RNase H (cPPT et PPT). Cette étape permet notamment la synthèse de la séquence U3-R-U5 de l'extrémité 3' du brin d'ADN (+) (*Figure 19 D*);
- Dégradation de l'ARNt^{Lys3} entraînant le second saut de brin et permettant l'hybridation des séquences PBS complémentaires des deux brins d'ADN, provoquant par ailleurs la circularisation du brin d'ADN (-) (*Figure 19 E*);
- Finalement la synthèse d'ADN viral se termine. Il y a alors synthèse des portions manquantes par complémentarité aboutissant à l'ADN viral linéaire possédant à ses deux extrémités les séquences LTR (Long Terminal Repeat) composés des séquences U3-R-U5. Lors de cette synthèse, une structure particulière est retrouvée au niveau du cPPT qui correspond à un fragment simple brin riche en purine. Il est important de noter que cette structure, dite ADN flap semble être à l'origine du déclenchement de la décapsidation (dans le modèle de décapsidation cytoplasmique progressive, Arhel et al., 2007). Il peut également se produire des phénomènes de réparation de l'ADN et de ligation pour donner un ADN circulaire possédant une ou deux séquences LTR unique ou double (*Figure 19 F*).

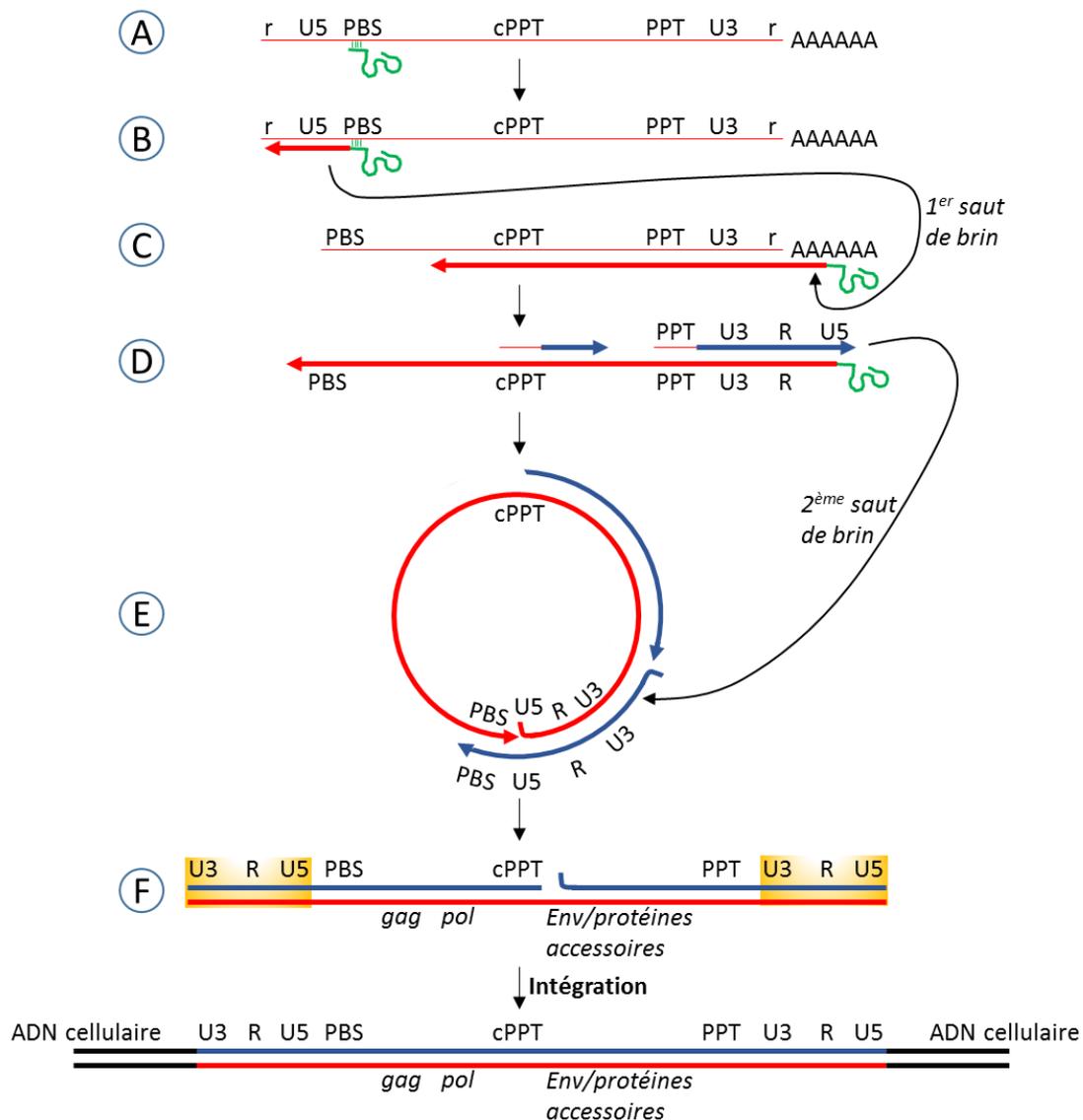


Figure 19 : Différentes étapes de la transcription inverse. L'ARN viral est représenté par une ligne fine rouge, le brin ADN (-) par une ligne épaisse rouge, le brin ADN (+) par une ligne épaisse bleue, l'ARN_t^{Lys3} est en vert, les LTR sont surlignés en jaune et enfin l'ADN cellulaire est représenté par des lignes épaisses noires. L'ADN circulaire avec un LTR unique ou double résultant de la réparation et de la ligation d'un ADN circulaire n'est pas représenté ici (Adapté de Onafuwa-Nuga and Telesnitsky, 2009).

La transcription inverse, indispensable pour l'infection virale, est la cible de molécules antivirales (description dans la partie [V-B-V.1V.2] de ce manuscrit). Ce fût la première étape ciblée par un traitement antirétroviral dans les années 1980.

L'activité polymérase de la transcriptase inverse n'étant pas fidèle à 100 %, cette étape est ainsi à l'origine de la variabilité génétique du VIH-1 et est également la cause des résistances aux traitements actuels. En effet, certaines mutations modifient positivement le virus le rendant ainsi insensible aux inhibiteurs de RT utilisés pour limiter son développement. Ces

mutations peuvent ainsi réduire l'incorporation de NRTI par rapport aux nucléotides classiques, ou peuvent faciliter l'excision de la base modifiée incorporée, bloquant préférentiellement la synthèse d'ADN viral (Hu and Hughes, 2012).

Outre l'insertion de mutations, il existe également un système de recombinaison génétique qui participe à la variabilité génétique du VIH-1. Ce système permet ainsi l'apport d'une mutation favorable, la délétion de mutants défavorables, ou encore la combinaison des deux. Cet échange de portion d'ARN (par exemple entre les 2 copies d'ARNs encapsidées dans la particule virale) est exceptionnellement fréquent dans le VIH-1 favorisant la variabilité génétique et la sélection de « super-virus » (Onafuwa-Nuga and Telesnitsky, 2009).

E. Import nucléaire : rentrer dans le donjon pour prendre les commandes

Une des caractéristiques des lentivirus est leur capacité à infecter des cellules qui ne sont pas en division. Ainsi, ils doivent passer la membrane nucléaire qui protège le génome cellulaire en passant au travers du pore nucléaire. Le génome viral va alors pouvoir s'intégrer dans le génome cellulaire (voir partie [F]). Le pore nucléaire est une structure stable qui permet le passage de molécules entre le noyau et le cytoplasme. Il est composé d'une trentaine de protéines différentes environ appelées les nucléoporines. Celles-ci forment un pore capable de laisser diffuser passivement les molécules de moins de 9 nm de diamètre, de faciliter le passage actif (par un transporteur des molécules jusqu'à 39 nm) et finalement d'empêcher l'entrée des molécules plus grosses (Hoelz et al., 2011). Le PIC (complexe de pré-intégration) est la continuité du complexe de transcription inverse et, par convention, le PIC commence quand la transcription inverse est terminée et que l'Intégrase a réalisé son activité de 3'processing (clivage des deux derniers nucléotides en 3' de chacun des deux brins d'ADN). Le PIC permet le passage de l'ADN viral néo-synthétisé et de protéines indispensables à l'intégration au travers du pore nucléaire (comme le PIC est plus gros que la capacité de passage du pore nucléaire, l'ensemble du PIC ne rentre pas dans le noyau). L'adressage du PIC vers le noyau se fait grâce aux signaux de localisation nucléaire (NLS) présents sur la MA, l'IN et Vpr. Ces différents signaux peuvent alors interagir avec des adaptateurs facilitant le transport nucléaire (Rch1, KPNA4, KPN β 1, TNPO3...) ou directement avec des protéines du pore nucléaire (NUP153, hCG1) (Matreyek and Engelman, 2013). Ces signaux ne semblent en

revanche pas indispensables pour l'import nucléaire, alors que l'ADN flap à la fin de la transcription inverse serait le seul signal nécessaire à l'import nucléaire (Rivière et al., 2010). Les protéines CA semblent également jouer un rôle important dans l'import nucléaire du PIC (Dismuke and Aiken, 2006). Une décapsidation lente pendant le transport nucléaire du complexe de transcription inverse puis du PIC pourrait contribuer au démasquage des NLS des protéines du PIC, tout en laissant l'espace pour le recrutement de protéines (BAF, hMGA1) favorisant l'entrée nucléaire. La CA interagit également directement avec des protéines favorisant l'import nucléaire (Cyclophilin A, TNPO3, NUPs, KPNs...) (Matreyek and Engelman, 2013). Finalement, il apparaîtrait que certaines protéines du PIC qui rentrent dans le noyau avec l'ADN pourraient participer à la sélection du site spécifique d'intégration (TNPO3 par exemple).

Malgré le nombre important d'études, les mécanismes de décapsidation et d'import nucléaires sont peu connus. La compréhension de ces mécanismes pourrait permettre le développement de nouveaux inhibiteurs empêchant ainsi l'infection de cellules quiescentes.

F. Intégration : insertion du génome viral dans le génome cellulaire

Après l'import nucléaire, l'ADN proviral est intégré dans le génome de la cellule hôte (Craigie and Bushman, 2012). L'IN est la protéine virale responsable suffisante pour catalyser la réaction d'intégration. En outre, la localisation de cette intégration n'est pas aléatoire mais cible préférentiellement les sites de transcription active principalement dans des régions très proches des pores nucléaires (Berry et al., 2006; Marini et al., 2015; Schröder et al., 2002).

L'intégration du génome a lieu en plusieurs étapes (Suzuki et al., 2012) :

- Reconnaissance par un dimère d'IN des deux extrémités de l'ADN proviral au niveau de sites d'attachement à l'extrémité des LTRs (**Figure 20 A**),
- Préparation de l'ADN par clivage de deux nucléotides très conservés aux deux extrémités 3' (3'-processing). Ceci fait ainsi apparaître des extrémités –OH en 3' de chaque brin d'ADN (**Figure 20 B**),
- Reconnaissance et clivage de l'ADN cible par attaque nucléophile de l'IN via les groupes hydroxyles précédemment créés (**Figure 20 C**),

- Liaison simultanée entre les extrémités 3' de l'ADN viral et les extrémités 5' de l'ADN cible par création d'une liaison phosphodiester (**Figure 20 D**),
- Excision des fragments ne s'appariant plus en 5' de l'ADN viral (car suppression de 2 nucléotides lors du 3'-processing) et liaison avec les extrémités de l'ADN cellulaire encore libre (**Figure 20 E**).

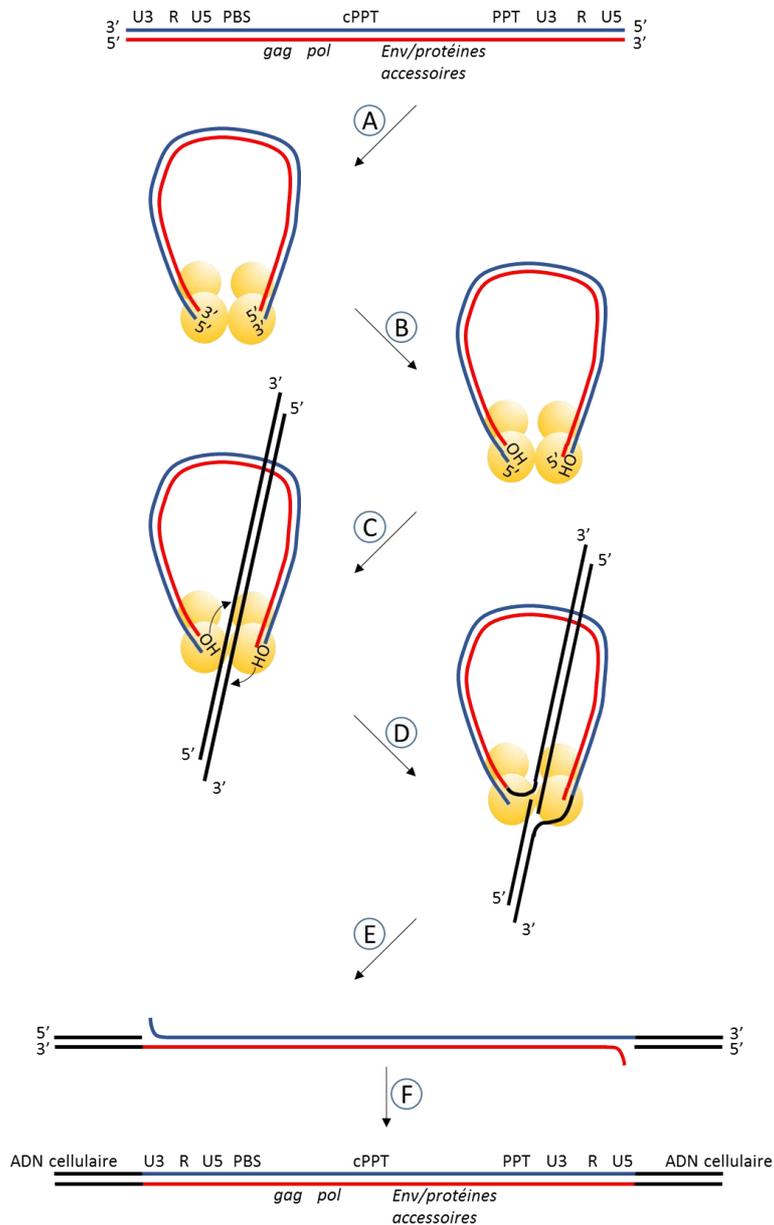


Figure 20 : Représentation schématique de l'intégration de l'ADN proviral dans le génome cellulaire. L'ADN proviral est représenté en bleu et rouge, l'ADN cellulaire en noir et l'Intégrase en jaune. Les lettres correspondent aux différentes étapes décrites dans le texte (Adapté de Suzuki et al., 2012).

L'IN agit au sein du complexe d'intégration qui implique également des facteurs cellulaires permettant l'optimisation de cette réaction. Parmi ces derniers, on retrouve LEDGF (Lens

Epithelium-Derived Growth Factor) qui stimule l'activité de l'IN (Cherepanov et al., 2004) en rapprochant cette protéine et l'ADN cible. L'intervention d'autres protéines est également suggérée comme celle de BAF (barrier to auto-integration factor, (Chen and Engelman, 1998)), de protéines liant l'ADN de la famille des « High Mobility Group » (Farnet and Bushman, 1997), de protéines de la chromatine (INI-1, EED, SUV39H ou HP1 γ) mais aussi la protéine NC du virus (Carteau et al., 1997 ; voir description dans la partie [VI-C.a]).

Mais, la proportion d'ADN viral intégré est faible par rapport à la quantité d'ADN produit lors de la transcription inverse. La majorité de cet ADN non-intégré existe sous forme linéaire (rapidement dégradée) ou sous forme circulaire contenant une ou deux copies de LTR (Butler et al., 2002; Munir et al., 2013). Le rôle de l'ADN viral non intégré est de plus en plus étudié puisqu'il a été démontré qu'il pouvait avoir un rôle dans la faible expression de protéines virales (Gelderblom et al., 2008), qu'il constitue une « réserve » qui peut être intégrée ultérieurement (Thierry et al., 2015), ou encore qu'il semble être à l'origine de la production de virus infectieux (Sloan and Wainberg, 2011). C'est pourquoi, l'ADN viral non-intégré est considéré comme un réservoir viral qui peut être impliqué dans la persistance du virus et la résistance aux traitements. En effet, il a récemment été démontré que l'ADN viral non intégré est transcrit lorsque la transcription de l'ADN intégré médiée par NF- κ B est non fonctionnelle (Thierry et al., 2016).

Cette étape étant indispensable pour que le virus se multiplie, elle a récemment été la cible pour le développement de nouvelles molécules antirétrovirales (Delelis et al., 2008). Ainsi, le Raltegravir® inhibe le transfert de brin en interagissant avec le complexe IN-vDNA après le 3' processing du vDNA. Il fût le premier inhibiteur de l'intégrase à avoir été mis sur le marché (Summa et al., 2008) et est toujours considéré comme référence pour le développement d'autres inhibiteurs de l'IN comme le Elvitegravir ou Dolutégravir mis sur le marché en 2012 et 2013 respectivement aux Etats-Unis.

G. Expression du génome viral, épissage et export nucléaire : synthétiser et sortir les ARNs du noyau

La première étape des éléments tardifs du cycle viral consiste en la transcription du génome viral sous forme d'ARNm qui seront plus ou moins épissés avant l'export vers le cytoplasme où ils seront soit encapsidés spécifiquement dans de nouvelles particules virales, soit traduits

sous forme de protéines. La cellule possède un système de contrôle très développé pour contrôler la production et l'export d'ARNm totalement épissés. Pour outrepasser ces contrôles, le VIH-1 utilise les protéines régulatrices apportées dans la cellule hôte lors de l'infection.

L'initiation de la transcription, par l'ARN polymérase II cellulaire (RNAPII), du génome viral se fait grâce à un promoteur unique situé dans la séquence 5'-LTR de l'ADN proviral intégré. Le LTR contient également en amont du promoteur de nombreux sites de régulation. Le cœur du promoteur est situé dans la région U3 qui contient une boîte TATA et plusieurs sites de liaison du facteur de transcription Sp1. En amont, se trouve une région enhancer portant deux sites de liaison à la protéine NK-κB. Encore plus en amont, se trouve la région modulatrice qui porte des sites de liaison à divers facteurs comme NFAT, CBP, AP-1 ou ATF/CREB. D'autres séquences liant des facteurs de régulation sont retrouvées dans la région U5 du LTR (Kilareski et al., 2009). L'expression du génome est également régulée par des modifications épigénétiques qui participent à la mise en place de la latence virale dans laquelle il a été montré que la transcription est réprimée par la présence de fortes concentrations d'HDACs (Histone deacetylases), d'histones méthylées ou désacétylées et d'ADN méthylé (Margolis, 2010).

Peu de temps après le début de la transcription, la polymérase s'arrête. La protéine Tat stimule alors la reprise de l'élongation en se liant sur l'ARN en cours de transcription au niveau d'un site de liaison de la région R appelé TAR (réponse à la transactivation). Une fois liée à TAR, Tat recrute la cycline T1 (cycT-1) membre du complexe P-TEFb (positive transcription elongation factor). Cdk9 (autre membre du complexe P-TEFb) phosphoryle l'extrémité C-terminale de la RNAPII ce qui stimule l'élongation (Tahirov et al., 2010). Tat recrute également de nombreux facteurs permettant le remodelage de la chromatine comme une histone acyltransférase (Tip60) ou les co-activateurs p300 et CBP (cAMP responsive binding protein) (Caputi, 2011).

Avant le transport vers le cytoplasme, les ARNm doivent subir deux modifications co-transcriptionnelles :

- insertion d'une coiffe en 5' : a lieu pendant la transition entre l'initiation de la transcription et l'élongation, lorsque l'ARN fait environ 20 à 40 nucléotides ;
- et insertion d'une queue poly(A) en 3' au niveau du site AAUAAA dans la région R. Le site étant présent aux extrémités 5' et 3', il est indispensable que le site en 5'

soit réprimé, alors que celui en 3' doit être activé. Dans un premier temps, il y a clivage des nucléotides situés entre l'extrémité 3' et le site AAUAAA (modifiés par les facteurs de clivages spécifiques à la polyadénylation (CPSF), dont l'expression est favorisée par Tat) puis insertion de la queue poly(A) (médiée par la poly(A) polymérase dont l'activité est activée par Vpr).

Le VIH-1 utilise de multiples sites d'épissage afin de générer de nombreux ARNm différents. L'épissage de l'ARN viral est contrôlé par des facteurs cellulaires qui se lient à de courtes séquences flanquant les sites d'épissage de l'ARN du VIH-1 (Karn and Stoltzfus, 2012) mais aussi par des facteurs viraux comme la protéine Tat et la structure de l'ARN lui-même (Jablonski et al., 2008, 2010). Ces ARNs messagers peuvent être multi-épissés ayant alors une taille de l'ordre de 1,8 kb ou mono-épissés (taille d'environ 4 kb).

Au début de la transcription (**Figure 21 A**), seuls les ARNm viraux totalement épissés codant les protéines Tat, Rev et Nef peuvent sortir du noyau en utilisant le système cellulaire. La protéine Rev alors synthétisée va entrer dans le noyau. La concentration en protéine Rev dans le noyau augmente ainsi progressivement permettant finalement à cette dernière de se lier à la séquence RRE pour permettre l'export actif vers le cytoplasme des ARNm non ou partiellement épissés grâce à son interaction avec le facteur d'export cellulaire CRM1 (Chromosome Maintenance Region 1) et Ran-GTP (**Figure 21 B**). Une fois dans le cytoplasme, le complexe ARN/Rev/CRM1/Ran-GTP est déstabilisé par RanGAP et RanBP1 libérant l'ARNm et la protéine Rev qui retourne dans le noyau par son interaction avec le facteur d'import nucléaire Importin- β (Henderson and Percipalle, 1997).

Ces différents ARNs serviront alors de matrice pour la synthèse de l'ensemble des protéines virales et pourront également, dans le cas de l'ARN complet non épissé de 9 kb, être encapsidés lors de l'assemblage de nouvelles particules (Kharytonchik et al., 2016).

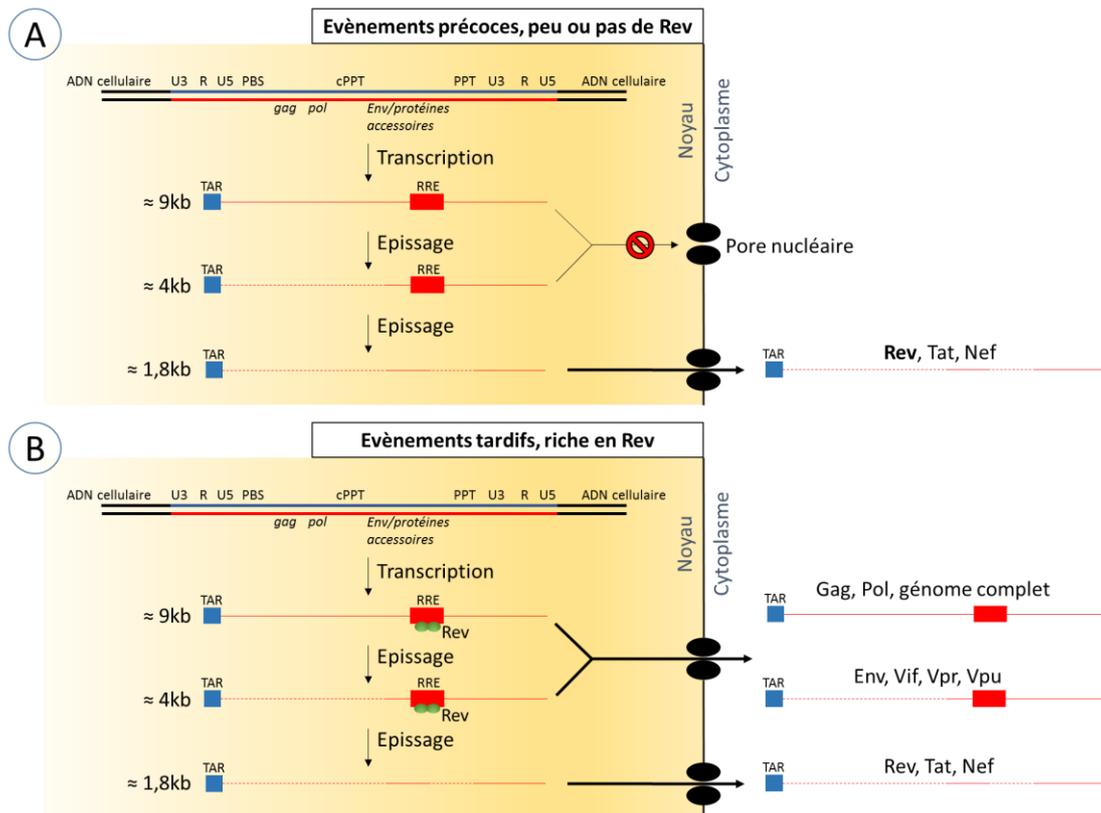


Figure 21 : Expression et export des ARNm plus ou moins épissés. Sont retrouvés l'ARNm complet (9 kb), partiellement épissé (4 kb) ou totalement épissé (1,8 kb) dans les évènements précoces et tardifs de l'infection caractérisée ici par une modification de la concentration de Rev dans le noyau (D'après Karn and Stoltzfus, 2012).

H. Traduction des ARNm viraux : à partir des ARNs, il faut produire les protéines

Les ARNs messagers viraux peuvent être traduits par la traduction coiffe-dépendante classique de la cellule qui dépend de la reconnaissance de la coiffe en 5' de l'ARN par le facteur d'initiation cellulaire eIF4E permettant l'assemblage du ribosome 40S.

Mais, comme représenté dans la **Figure 8**, la région 5'UTR est très structurée ce qui lui permet de recruter les ribosomes indépendamment de la coiffe. En effet, la traduction des ARNm du VIH-1 est également initiée au niveau de deux structures appelées IRES (Internal Ribosome Entry Site ; Chamond et al., 2010). Le premier, dans la région 5' UTR, permet la synthèse de Gag spécialement lors du blocage de la cellule en phase G2 au cours de l'infection et le second, dans le cadre de lecture de Gag, permet la synthèse de divers isoformes de Gag de petites tailles (Buck et al., 2001). Les deux types de traduction ont été montrés *in vivo* et *in vitro* (Brasey et al., 2003; Miele and Lever, 1996; Ricci et al., 2008). Il semble donc que les ARNs de

VIH-1 puissent passer de l'un à l'autre type de traduction en fonction de la nature et de l'état de la cellule infectée (Balvay et al., 2007).

Nous ne rentrerons pas ici dans les détails de la traduction de l'ensemble des ARNs messagers de VIH-1 mais il est important de rappeler qu'un décalage du cadre de lecture est observé lors de la traduction de Gag permettant alors au ribosome de ne pas « lire » le codon STOP à la fin de la synthèse de Gag ce qui résulte en la synthèse de Gag-Pol. Cet événement apparaît environ 5 % du temps donnant 1 Gag-Pol pour 20 précurseurs Gag synthétisés (Bolinger and Boris-Lawrie, 2009). Ceci pourrait ainsi participer au choix dans le dilemme proposé à l'ARN présent dans le cytoplasme entre traduction des protéines virales ou encapsidation sous forme de dimère dans les particules virales (Butsch and Boris-Lawrie, 2002). Plusieurs scénarios d'utilisation de l'ARN dans le cytoplasme sont alors possibles mais l'impossibilité de distinguer les deux types d'ARNs rend compliqué l'étude de ce mécanisme. Une des hypothèses les plus récentes est que l'ARN servant à la synthèse du précurseur Gag-Pol soit celui qui est reconnu par Gag pour être transporté à la membrane et encapsidé dans des particules virales (Chamanian et al., 2013).

I. Assemblage et bourgeonnement : constituer une armée pour accentuer l'invasion

L'assemblage et le bourgeonnement permettent l'empaquetage de tous les composants nécessaires à l'infection virale sous forme de particules virales immatures. Pour former ces particules, la polyprotéine Gag est nécessaire et suffisante. L'assemblage est rapide et seulement 8 à 9 minutes sont nécessaires pour former 90 % d'une particule virale et 13 à 35 minutes après le début de l'assemblage les premières particules sont libérées dans le milieu extracellulaire (Ivanchenko et al., 2009; Jouvenet et al., 2008). En effet, grâce à ces différents domaines, le précurseur Gag est capable d'interagir avec les acides nucléiques (ARNs), les lipides membranaires (principalement $PI_{(4,5)}P2$: Phosphatidylinositol-4,5-bisphosphate) et l'ensemble des protéines virales et cellulaires nécessaires à l'infection.

1. Éléments indispensables à l'assemblage

a) Interaction Gag-ARN

Pour former une particule virale infectieuse, il est nécessaire que celle-ci contienne deux copies du génome viral sous forme d'ARN (+). En revanche, le mécanisme de sélection de seulement deux copies d'ARNs viraux alors que 1500 à 3000 polyprotéines Gag se multimérisent pour former une particule reste mal connu (Kuzembayeva et al., 2014). La dimérisation des ARNs, indispensable pour la production de virus infectieux, se fait par l'intermédiaire du site d'initiation de la dimérisation (DIS) présent au niveau de la boucle SL1 du signal d'encapsidation Ψ dans le 5'-UTR (**Figure 8**; Moore et al., 2007). Deux séquences DIS portées par deux copies d'ARNs forment alors un « baiser entre les boucles » (kissing-loop) (Moore and Hu, 2009). Cette dimérisation, Gag-dépendante, a lieu dans le cytoplasme tout en se faisant dans un environnement proche de la membrane (Chen et al., 2016). Les deux copies d'ARNs liées chacune à une protéine Gag dans le cytoplasme sont ensuite transportées vers la membrane, où la proximité entre les signaux DIS permet la dimérisation Gag-dépendante. Alors qu'environ 10 % des particules virales du VIH-1 contiennent des ARNs cellulaires ou des ARNs viraux épissés, la plupart des particules virales (> 90 %) contient l'ARN génomique viral complet (Chen et al., 2009). La discrimination entre ARNs viraux complets, ARNs viraux épissés et les nombreux ARNs cellulaires est donc très efficace (Carlson et al., 2016) et sélective de la région Ψ en U5 (Abd El-Wahab et al., 2014).

La polyprotéine Gag possède plusieurs sites de liaison à l'ARN : au niveau de la NC et au niveau de la MA. L'encapsidation spécifique de l'ARN viral est réalisée par la liaison entre la Nucléocapside et le signal d'encapsidation appelé région Ψ située dans la région 5' UTR juste en amont du codon d'initiation de la traduction de Gag (**Figure 8**; Lu et al., 2011). Une ou plusieurs mutations dans les doigts de zinc (structure caractéristique de la NC) entraînent une production de virus contenant une faible quantité d'ARN viral génomique aboutissant à une forte diminution de l'infection, montrant que ces structures sont indispensables pour la sélection de l'ARN génomique viral (Gorelick et al., 1990; Grigorov et al., 2007). De plus, il a été montré que chaque doigt de zinc ne possède pas un rôle équivalent dans la sélection de l'ARN génomique viral (Gorelick et al., 1993). La délétion du signal d'encapsidation Ψ n'empêche pas la production de particules virales mais celles-ci contiennent une proportion plus importante d'ARNs cellulaires, les rendant non infectieuses (Houzet et al., 2007; Rulli et

al., 2007). Récemment, il a d'ailleurs été démontré que cette interaction entre Gag et le signal d'encapsidation a lieu dans le cytoplasme (Hendrix et al., 2015; Kutluay et al., 2014).

En revanche, l'échange, dans Gag du VIH-1, du domaine NC avec celui d'un autre rétrovirus (MMTV : Mouse Mammary Tumour virus) ne bloque pas l'incorporation spécifique de l'ARN viral du VIH-1 par un autre domaine de la protéine Gag ou par la présence des deux doigts de zinc de la NC de MMTV (Poon et al., 1998). Par contre, le remplacement par une NC ne portant qu'un seul doigt de zinc (Virus Moloney) entraîne la perte de l'incorporation préférentielle du génome de VIH-1. Ces observations suggèrent que, dans la sélection de l'ARN spécifique, la présence de deux doigts de zinc est la propriété la plus importante et qu'un autre domaine de Gag est impliqué. Le domaine MA de Gag possède la capacité de lier non spécifiquement les ARNs et les ADNs (Alfadhli et al., 2009; Chukkapalli et al., 2010). Plus particulièrement, une région très basique en N-terminal lierait avec une forte affinité une région de l'ARN située dans la séquence codante pour Pol (Purohit et al., 2001). La mutation de cette région basique de MA résulte en des particules présentant une réplication ralentie. L'interaction avec les acides nucléiques au niveau du domaine NC et du domaine MA suggère un repliement de la protéine Gag sous forme compacte. L'étude de cette forme étant un de mes projets de thèse, une description plus détaillée de ce modèle est présentée dans la partie [IV] de ce manuscrit. Cette région basique est également impliquée dans la liaison à la membrane. La compétition entre ARNs et lipides contribuerait également à la sélection des lipides spécifiques au niveau de la membrane (Alfadhli and Barklis, 2014; Chukkapalli et al., 2010).

Si le rôle du domaine NC pour la fixation de l'ARN a été largement développé, il intervient également dans le trafic membranaire de Gag. Ainsi, que ce soit lors de la délétion complète de NC, des résidus basiques ou des doigts de zinc uniquement, l'assemblage est fortement perturbé. Il a été montré que la mutation des quinze résidus basiques de la NC en alanine suffit à accélérer la diffusion de Gag dans la cellule alors que des mutations affectant la multimérisation (W184A-M185A) ou l'ancrage à la membrane (G2A donnant une Gag dépourvue de myristate) n'ont pas d'effet sur celle-ci (Hendrix et al., 2015). Cette différence de mobilité est donc attribuée à l'interaction de Gag avec l'ARN via le domaine NC. Si Hendrix et ses collaborateurs n'attribuent pas le ralentissement de la diffusion de Gag liée à l'ARN à la forme compacte de Gag, c'est que cette forme devrait plutôt accélérer la diffusion du fait de sa plus faible surface de contact. Un autre laboratoire s'est quant à lui concentré sur la

visualisation de la diffusion des ARNs génomiques dans le cytoplasme (Chen et al., 2014). Les auteurs concluent que le complexe Gag-ARN ne joue pas un rôle crucial dans l'assemblage du virus. En effet, que ce soit en présence ou en absence de Gag, la diffusion de l'ARN reste identique. Ce dernier atteindrait donc la membrane par simple diffusion. Les auteurs suggèrent ainsi que seule une faible proportion de Gag est liée à l'ARN, confirmant ainsi les résultats préalablement publiés (Jouvenet et al., 2009).

b) Interaction Gag-membrane

La membrane lipidique des particules virales est enrichie en certains types de lipides qui sont retrouvés dans des micro-domaines de la membrane cellulaire appelés rafts : lipides à chaînes saturées, sphingolipides, cholestérol, plasményléthanolamine, phosphatidyl-sérine et phosphatidylinositol-4,5-bisphosphate (PI_(4,5)P2). Ceci suggère que les virus bourgeonnent préférentiellement au niveau de ces domaines (Brügger et al., 2006). Cette composition est notamment le résultat de la forte affinité de Gag, via la MA, pour ces lipides. L'interaction Gag-membrane est de plus stabilisée par la présence d'un résidu myristate sur la glycine en position N-terminale. Cette modification post-traductionnelle permet ainsi l'ancrage de Gag dans la membrane et plus particulièrement au niveau du PI_(4,5)P2 (Ono et al., 2004; Saad et al., 2006). Cette interaction est doublement stabilisée par la chaîne insaturée du PI_(4,5)P2 qui s'insère dans la poche hydrophobe de la MA qui permettait la séquestration du myristate avant l'insertion de ce dernier dans la membrane (Chukkapalli and Ono, 2011). Très récemment, Barros et ses collaborateurs ont montré que la liaison de MA avec la membrane n'est pas plus forte uniquement en la présence de PI_(4,5)P2 seul. Il semble plutôt que la présence de cholestérol, augmente l'affinité de Gag pour la membrane et les PI_(4,5)P2 (Barros et al., 2016). Cette stabilisation pourrait également être favorisée par un pH acide (Fledderman et al., 2010) et par la liaison Ca²⁺-dépendante de la calmoduline à Gag (Ghanam et al., 2010).

c) Multimérisation de Gag

Finalement, la multimérisation du précurseur Gag (et à moindre mesure Gag-Pol) est un élément majeur de l'assemblage. L'expression de la protéine Gag est ainsi nécessaire et suffisante pour la production de particules virales (celles-ci étant incomplètes, elles sont bien entendu non infectieuses). Si la multimérisation de Gag est prédominante au niveau de la

membrane (Derdowski et al., 2004; Ono et al., 2000b), il semble qu'elle soit initiée dans le cytoplasme (El Meshri et al., 2015).

Deux régions de Gag ont été identifiées comme impliquées dans sa multimérisation, le domaine C-terminal de CA (CA-CTD) et la NC. Dans la CA, une région, appelée MHR (Major Homology Region) est cruciale pour la multimérisation de Gag. En effet l'introduction de la double mutation W184A et M185A dans cette région limite la multimérisation de Gag, limitant alors la production de particule virale (Datta et al., 2007b). Il apparaît également que la délétion du domaine NC de Gag limite l'oligomérisation membranaire (El Meshri et al., 2015). En effet, l'interaction de plusieurs protéines Gag (via leur domaine NC) avec un ARN permet le rapprochement des protéines, ce qui favorise l'initiation de la multimérisation (Burniston et al., 1999; Coren et al., 2007). Le caractère basique de nombreux résidus du domaine NC de Gag favorisant l'interaction avec les acides nucléiques, leurs délétions entraînent une diminution de la multimérisation de Gag (Cimarelli and Luban, 2000; Cimarelli et al., 2000).

2. Bourgeonnement

Cette étape constitue la dernière phase de l'assemblage et elle permet la séparation entre la cellule productrice et la particule virale (Freed, 2015). Elle est principalement médiée par le domaine p6 de Gag comme le montre l'accumulation à la membrane cellulaire de particules maintenues par un fin faisceau membranaire lorsque ce domaine est supprimé (Göttlinger et al., 1991). Plus précisément, le motif PTAP de p6 (Huang et al., 1995) est responsable du recrutement de la protéine TSG101 (Tumour Susceptibility Gene 101) du complexe ESCRT-I (Endosomal Sorting Complex Required for Transport) qui permet un bourgeonnement correct (Demirov et al., 2002; Garrus et al., 2001). Il apparaît également que le recrutement de la protéine ALIX (ALG2-interacting protein X) via le domaine YPXL de p6 (Fujii et al., 2009) ou par la NC (Dussupt et al., 2009) permette la libération optimale des particules virales. Il est important de rappeler que le système cellulaire ESCRT est détourné par le virus car il sert habituellement à la libération des vésicules excrétées ou encore à la séparation des membranes lors de la division cellulaire (MVB).

Le bourgeonnement se déroule en plusieurs étapes (*Figure 22*) :

- Recrutement par le domaine p6 de Gag de TSG101 du complexe ESCRT-I et d'ALIX (qui n'appartient à aucun complexe ESCRT),

- Ce premier recrutement va permettre la liaison des protéines du complexe ESCRT-III lié à la membrane qui s'assemblent sous forme de structures en spirale qui resserrent la membrane cellulaire.
- Finalement, la protéine VPS4 (Vacuolar Protein Sorting 4) est recrutée et catalyse de manière ATP-dépendante la dissociation de la particule de la cellule (Cashikar et al., 2014).

Le mécanisme exact du bourgeonnement n'est encore pas complètement compris et plusieurs modèles sont actuellement proposés qui diffèrent notamment sur l'étape impliquant le complexe ESCRT III (Cashikar et al., 2014; Freed, 2015; Hurley and Hanson, 2010; Prescher et al., 2015; Van Engelenburg et al., 2014).

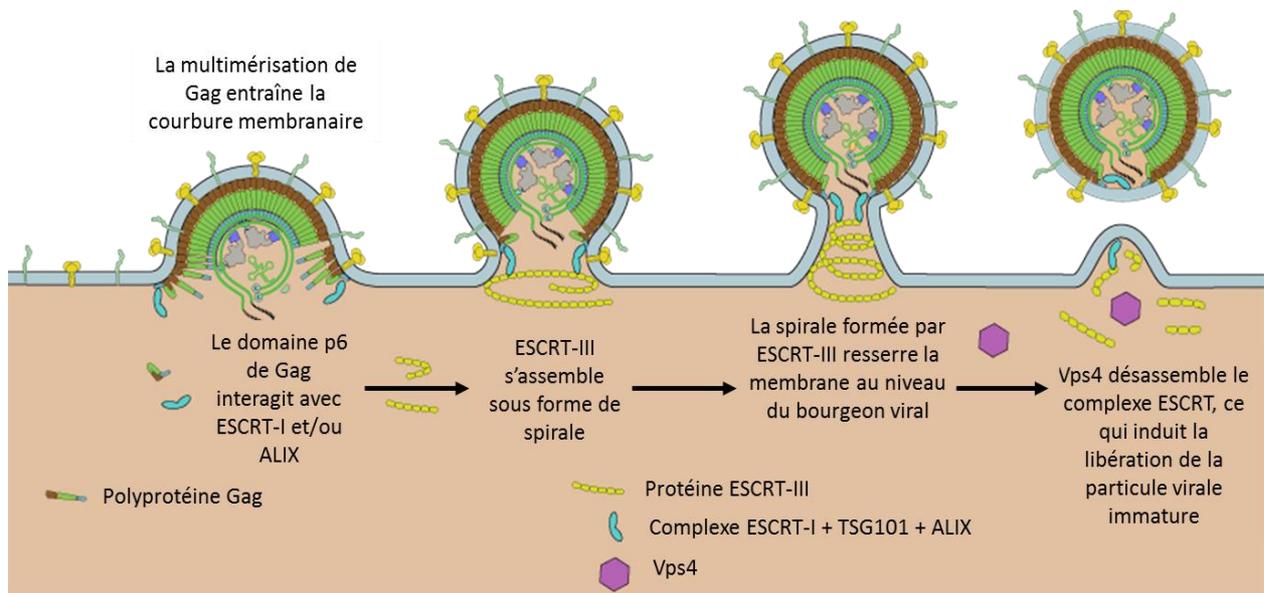


Figure 22 : Modèle de bourgeonnement d'une particule virale de VIH-1 grâce aux complexes ESCRT-I, ALIX, ESCRT-III et Vps4 (traduit de la figure du site http://viralzone.expasy.org/all_by_species/1536.html visité le 22/07/2016).

J. Maturation : s'organiser pour mieux infecter

Dernière étape du cycle viral, la maturation n'en est pas moins indispensable. En effet, une particule virale immature reste non infectieuse. Elle débute en parallèle du bourgeonnement ou immédiatement après la scission. La maturation est réalisée par la Protéase virale (PR), protéase à aspartate qui permet le clivage de tous les sites protéolytiques retrouvés au sein de Gag et Gag-Pol. Grâce à la libération de différents protéines (MA, CA, NC, p6, PR, RT et IN), un réarrangement très important a lieu au sein de la protéine virale pour aboutir à sa structure

caractéristique contenant une capsidie conique (**Figure 6**). Cette protéine étant présente en tant que domaine du précurseur Gag-Pol dans la particule immature, la première étape de la maturation consiste, au sein d'un dimère de précurseur, en un clivage auto-protéolytique de la protéine Gag-Pol. Ceci libère alors un dimère de PR (seul forme active de PR) qui clive alors l'ensemble des sites protéolytiques des précurseurs (voir **Figure 10** pour les sites de clivages de Gag).

Cette étape étant indispensable pour la production de particules infectieuses, elle est évidemment la cible d'inhibiteurs (description dans la partie [V-B-V.3] de ce manuscrit).

IV. Forme compacte de Gag

Les mécanismes actuellement proposés pour la sélection de l'ARN viral et l'assemblage de la particule comportent encore un grand nombre de suppositions. Notre hypothèse est qu'une forme compacte de Gag, peu fréquente *in cellulo*, permette la fixation de l'ARN viral spécifiquement pour la transporter vers la membrane.

Basé sur les capacités d'assemblage de la protéine Gag en solution, ce modèle a d'abord été décrit en 2007 par l'équipe d'Alan Rein (Datta et al., 2007a, 2007b). L'utilisation de la diffusion des neutrons aux petits angles (SANS, Small Angle Neutron Scattering) suggère ainsi un rapprochement entre les extrémités N-terminale (domaine MA) et C-terminale (Domaine NC) de Gag. Cette même équipe a également montré que ces deux domaines permettent l'interaction avec les inositols-phosphates (IP6, permettant de mimer les phosphatidylinositol 4,5-bisphosphate retrouvés en forte concentration dans les particules virales, (Datta et al., 2007b)). En solution, la présence d'acides nucléiques et l'absence de phospholipides entraînent la formation de particules de 30 nm de diamètre alors qu'une particule virale immature fait un diamètre de 100 à 150 nm (Campbell et al., 2001). Basés sur ces résultats, les auteurs suggèrent un repliement de la protéine Gag dans le cytoplasme autour d'acides nucléiques, puis un retour dans une structure étendue lors de l'association du domaine MA avec les phospholipides membranaires. En effet, les résidus basiques de la MA possèdent une affinité plus importante pour les lipides que pour les acides nucléiques. Cette transition conformationnelle permet par ailleurs de favoriser la multimérisation des protéines Gag qui s'assemblent sous la membrane pour former une particule virale.

Le même groupe a par la suite validé ce modèle en utilisant la technique de réflectivité de neutron (NR, Neutron Reflectivity, (Datta et al., 2011a)). Cette technique permet de mesurer la capacité de réflexion de neutron au travers d'un film fin mimant la structure membranaire. Les auteurs ont ainsi été capables de différencier une forme de Gag plutôt compacte liée à la membrane par ses extrémités N- et C-terminales, d'une forme linéaire liée par son extrémité N-terminale (MA) à la membrane et à son extrémité C-terminale (NC) aux acides nucléiques (Datta et al., 2011a). La même année, les auteurs ont d'ailleurs proposé que le domaine Sp1 (situé entre CA et NC) soit impliqué dans le passage d'une forme compacte de Gag à une forme

plus allongée (Datta et al., 2011b). Sur la base de ces résultats, ce groupe de recherche a proposé le modèle de forme compacte de Gag sur lequel mon travail de thèse s'est appuyé. De récents travaux ont contribué à prouver l'existence de cette forme *in vitro* (Munro et al., 2014), en utilisant une protéine Gag ne portant ni de domaine C-terminal p6, ni de myristate. La protéine a été modifiée par l'ajout de peptides de 12 acides aminés permettant l'arrimage de fluorophores par réaction enzymatique. Ainsi, le peptide S6 permettant l'accrochage du fluorophore Alexa-594 a été fusionné en position C-terminale de Gag Δ myr Δ p6 et le peptide A1 permettant l'accrochage du fluorophore Alexa-488 a été fusionné soit en position N-terminale de Gag Δ myr Δ p6, soit entre les domaines MA et CA. Si ces deux fluorophores se retrouvent suffisamment proches, il est possible de mesurer le transfert d'énergie (FRET) entre eux. Lorsque Gag est sous forme linéaire, les fluorophores sont éloignés, l'efficacité de FRET est proche de zéro. En revanche, lorsque Gag adopte la forme compacte, il y a rapprochement des fluorophores induisant un FRET avec une efficacité proche de « un » (**Figure 23 A**). En présence d'ADN (polyA), la protéine adopte aussi bien une conformation linéaire que compacte (**Figure 23 B**). Après l'ajout d'IP6 plus aucun FRET n'est observé, Gag empruntant alors uniquement la forme linéaire (**Figure 23 C**). Dans cette expérience, la position du fluorophore donneur (A488) est importante. En effet, s'il est placé entre MA et CA, l'efficacité de FRET maximale n'est pas atteinte suggérant une distance plus élevée entre les fluorophores que s'il est placé en position N-terminale de Gag. Les auteurs suggèrent ainsi que la flexibilité du peptide situé entre les domaines MA et CA soit nécessaire pour induire le rapprochement des extrémités N- et C-terminales de Gag.

Ces résultats confirment donc l'existence d'une forme compacte de Gag permettant le rapprochement des extrémités N- et C-terminales (Datta et al., 2007a, 2007b) validant le modèle proposé quelques années plus tôt (Datta et al., 2011a). En revanche, le domaine Gag facilitant le repliement de la protéine n'est pas décrit avec précision puisque Datta *et al.* proposaient en 2011 que le domaine Sp1 facilitait le repliement de la protéine, alors que Munro *et al.* en 2014 penchait en faveur de l'implication des résidus situés entre MA et CA.

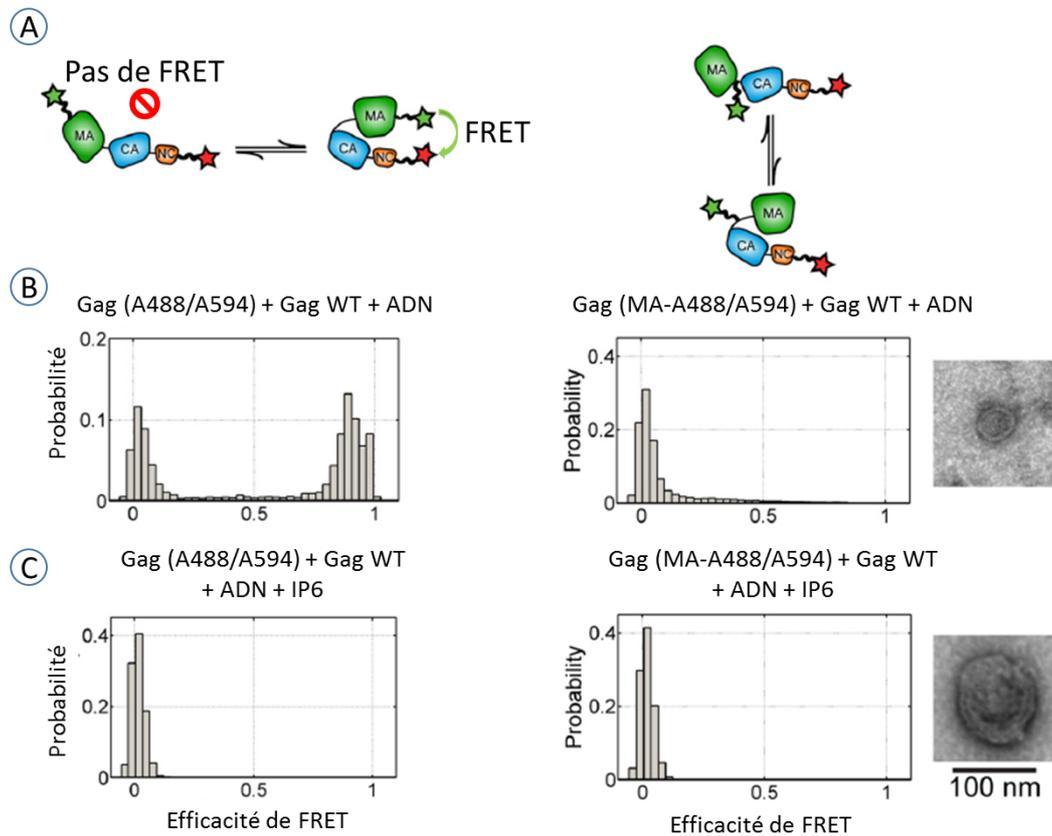


Figure 23 : Distribution du FRET entre les fluorophores couplés en position N-terminale ou entre MA et CA (étoile verte, A-488) et en position C-terminale (étoile rouge, A-594) de Gag. (A) Représentation schématique de Gag sous forme linéaire (efficacité de FRET égale à 0) ou sous forme compacte (efficacité de FRET maximale proche de 1) dans le cas où les fluorophores sont aux extrémités N- et C terminales de Gag (Panel de gauche). Pas ou peu de FRET observé dans le panel de droite où le donneur (A-488) est inséré entre MA et CA. (B) En présence d'ADN et d'une forte concentration de Gag non marquée, la protéine adopte aussi bien une conformation linéaire que compacte (panel de gauche). (C) Après l'ajout d'IP6, qui mime la membrane lipidique, plus aucun FRET n'est retrouvé, Gag emprunte uniquement la forme linéaire (panel de gauche) (Extrait de Munro et al., 2014).

Le modèle proposé est donc que la forme compacte de Gag permette la liaison de l'ARN génomique viral dans le cytoplasme (**Figure 24 A**). Dans cette configuration, la MA est dite tendue ([T]) et la protéine Gag apparaîtrait ainsi sous forme de monomères ou de dimères.

Si le transport des protéines de l'enveloppe à la membrane se fait par la voie d'excrétion (via des vésicules transitant par le réticulum endoplasmique et l'appareil de Golgi), le transport vers la membrane du complexe Gag-ARN, du monomère linéaire Gag, ou même de différents multimères reste en revanche à élucider (Freed, 2015). À la membrane, l'affinité de la MA pour les lipides étant plus importante que pour l'ARN et, considérant que les acides aminés impliqués dans les deux types de liaisons sont identiques (Alfadhli and Barklis, 2014), la liaison de la MA se fait alors préférentiellement sur les lipides membranaires (**Figure 24 C**) (Chukkapalli

et al., 2010). Il apparaît également que la NC facilite le recrutement de lipides chargés négativement contribuant à une première attache de la forme compacte de Gag à la membrane (Kempf et al., 2014). Finalement, la membrane étant le lieu privilégié de multimérisation de Gag, la forme linéaire est préférentiellement adoptée, favorisant la multimérisation qui, par modification conformationnelle va entraîner le passage de la MA sous forme relâchée ([R]), provoquant l'ancrage du myristate dans la membrane (Saad et al., 2006; Tang et al., 2004). Cet ancrage sera alors doublement stabilisé par le remplacement dans la poche hydrophobe de la MA du myristate par une chaîne carbonée insaturée du PI_{(4,5)P2} (Figure 24 D). Que ce soit par le mouvement des protéines Gag en cours d'assemblage, ou par le recrutement de ces protéines, la membrane des particules virales est fortement enrichie en lipides saturés principalement retrouvés au niveau des rafts (PI_{(4,5)P2}, Phosphatidylsérine, cholestérol, céramides, GM3 et sphingolipides (Brügger et al., 2006; Chan et al., 2008).

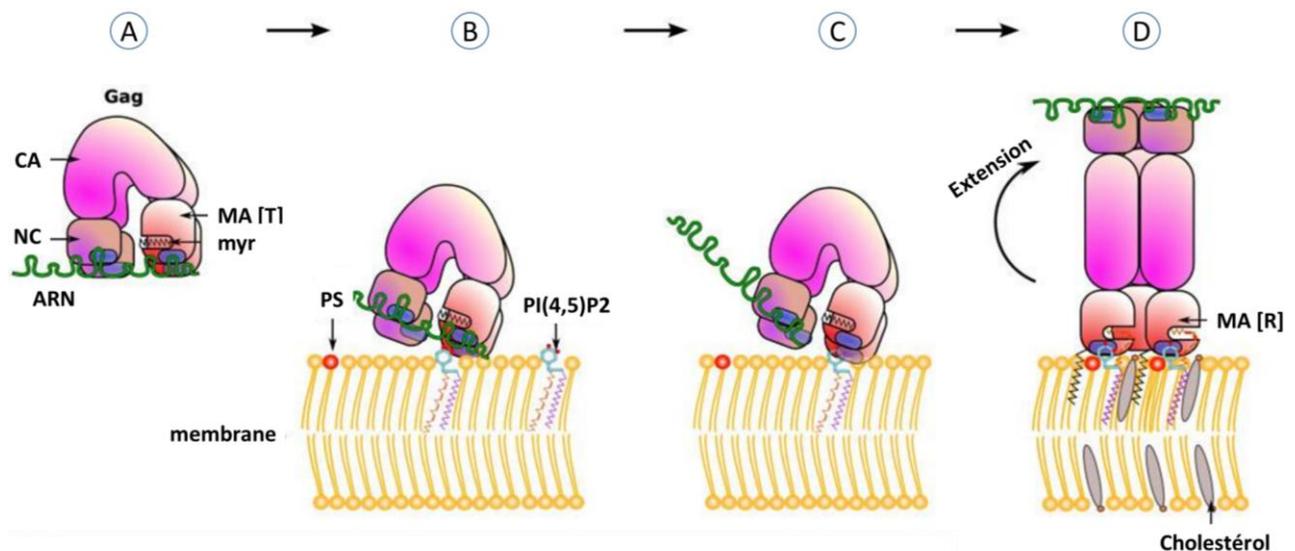


Figure 24 : Modèle dynamique de liaison de Gag à l'ARN et à la membrane cellulaire. Les lettres correspondent à l'explication de chaque étape retrouvée dans le texte (Adapté de Hamard-Peron and Muriaux, 2011).

Finalement, un modèle propose que l'ARN viral complet subissant le décalage du cadre de lecture pour la synthèse du précurseur Gag-Pol soit encapsidé de façon spécifique (Chamanian et al., 2013). Si ce modèle théorique n'est pas validé expérimentalement, il permet de lier plusieurs observations. Ainsi, comme nous l'avons vu précédemment, la traduction est majoritaire pour la polyprotéine Gag (95 % du temps) et minoritaire pour celle de Gag-Pol (5 % du temps) alors que l'ARN servant de matrice est identique. Chamanian et ses collaborateurs ont ainsi mis à jour l'importance d'une région de l'ARN viral appelé GRPE

(Genomic RNA Packaging Enhancer) chevauchant la région du décalage de cadre de lecture pour la traduction de Gag-Pol. Dans la majorité des cas, ce site semble être recouvert par des protéines cellulaires (comme eRF1) qui favorisent l'élimination de l'ARN par la voie de dégradation NMD (nonsense-mediated decay) une fois que la synthèse de la protéine Gag est terminée (Hogg and Goff, 2010). Plus rarement, il y a décalage du cadre de lecture afin de permettre au ribosome d'éviter le codon STOP arrêtant habituellement la synthèse (pour produire la protéine Gag) afin de traduire le précurseur Gag-Pol. Cette poursuite de la synthèse libère alors le site GRPE de ses interactions avec différents facteurs cellulaires. Ainsi, ce site pourrait interagir avec la séquence Ψ située à l'extrémité 3' de l'ARN viral afin d'augmenter la liaison de cette région par le domaine NC de Gag. Notre hypothèse est que la structure secondaire et/ou tertiaire empruntée par l'ARN favorise la liaison de la NC au niveau du site GRPE/ Ψ , rapprochant alors le site de liaison de la MA à l'ARN situé au niveau du gène *Pol*. La forme compacte de Gag interviendrait donc dans la liaison et le transport membranaire de l'ARN génomique viral complet. Finalement, la multimérisation de protéines Gag et Gag-Pol à la membrane couplée à l'apport de l'ARN spécifique par la forme compacte de Gag permet la libération de particules virales immatures contenant spécifiquement le génome viral.

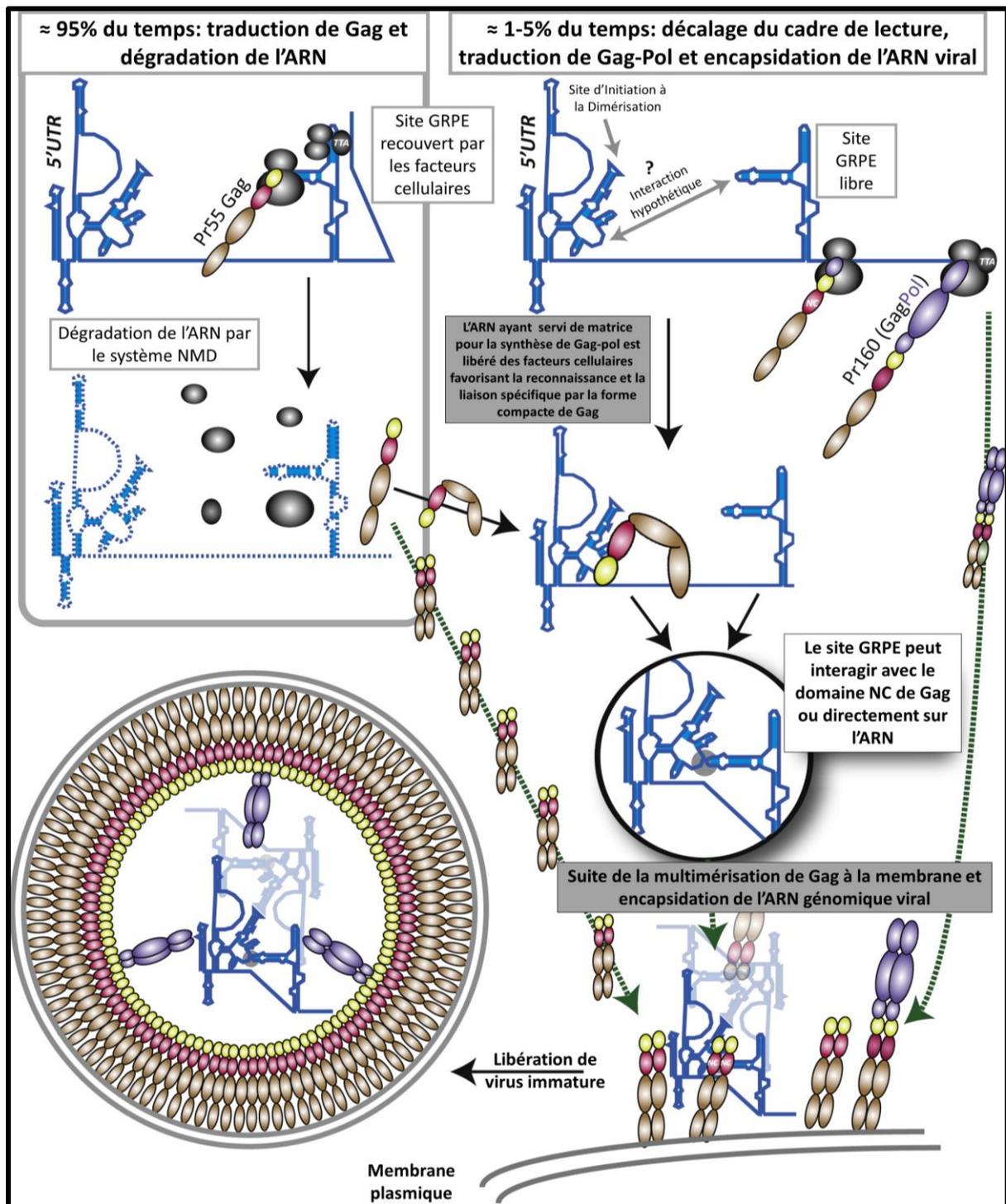


Figure 25 : Modèle possible de traduction et d'assemblage du génome viral non épissé. L'ARN sert dans la majorité du temps (95 %) à la synthèse de la polyprotéine Gag suivie d'une dégradation de l'ARN par le système NMD. Dans 5 % du temps, l'ARN sert de matrice pour la synthèse du précurseur Gag-Pol grâce à un décalage du cadre de lecture. Cet ARN, libéré des facteurs cellulaires adopte une structure secondaire favorisant le recrutement de la protéine Gag compacte et le transfert membranaire pour poursuivre l'assemblage (Adapté de Chamanian et al., 2013).

V. Comment soigne-t-on le VIH ?

A. Avant de soigner, il faut dépister

Comme il est décrit dans la partie épidémiologie de ce manuscrit [I-D], une cible ambitieuse pour 2030 a été formulée par l'ONUSIDA : 90-90-90 (*Figure 5*). Afin d'atteindre cet objectif, un effort important doit être fait au niveau du dépistage de la maladie. En effet, dans le monde, on estime que seulement 54 % des personnes infectées par le VIH-1 sont au courant de leur séropositivité. Ce pourcentage passe même à 45 % en Afrique Sub-Saharienne où l'épidémie reste un véritable fléau (ONUSIDA, 2014). Dans ces pays, l'accès aux différents tests de dépistage reste limité premièrement par leur coût mais également par le manque de matériels nécessaires pour l'analyse. En France, le dépistage est obligatoire uniquement lors de dons de sang, d'organes ou de tissus. Il existe trois catégories de tests : (<http://www.cdc.gov/hiv/basics/testing.html>, site visité le 25/07/2016) :

- **Détection des anticorps** : la plupart des tests actuellement utilisés entrent dans cette catégorie. Ils permettent la détection des anticorps produits par le système immunitaire en réponse à l'infection virale. Ce type de test ne peut donc pas être utilisé immédiatement après l'exposition à risque. En effet, plusieurs semaines (de 3 à 12 en fonction des personnes) sont nécessaires à l'organisme pour produire une quantité d'anticorps détectable. Ces tests permettent de diagnostiquer la séropositivité d'une personne en quelques minutes par un prélèvement de sang ou de salive. C'est généralement cette méthode qui est utilisée par les autotests (mis sur le marché français en mars 2015) permettant de fournir un résultat fiable (> 99 % pour le test Mylan® commercialisé en France) sur le statut sérologique d'une personne. En cas de résultats positifs, ces tests sont confirmés par des tests couplant la détection des anticorps et des antigènes viraux.
- **Détection couplée des anticorps et des antigènes** : ces tests, appelés aussi tests ELISA combinés ou tests de 4^{ème} génération, sont généralement ceux utilisés pour la validation de la séropositivité. Ils permettent la détection couplée des anticorps anti-VIH-1 et de l'antigène p24 (correspond au domaine capsidique de Gag ou à la protéine de Capsid après maturation). Un résultat négatif de ce test six semaines après l'exposition supposée est considéré comme signant l'absence d'infection par

le VIH (selon un rapport de la Haute Autorité de santé publié en octobre 2008 : « Dépistage de l'infection par le VIH en France »). En cas de résultat positif, ou de doute sur le résultat (négatif alors qu'il y a eu exposition à risque), un second test de 4^{ème} génération sera réalisé quelques semaines plus tard pour confirmer le diagnostic, ou un test de détection des acides nucléiques sera effectué.

- **Détection des acides nucléiques** : ce test permet de mesurer directement la quantité de virus présente dans le sang. Il permet la quantification de l'ARN viral par utilisation de la RT-PCR (Real Time Polymerase Chain Reaction) quantitative. Ce test, assez coûteux, ne peut pas être utilisé pour le dépistage systématique. Il permet en revanche la détection des personnes en primo-infection car quelques jours seulement sont nécessaires pour que la quantité de virus soit supérieure au seuil de détection du test.

Plusieurs stratégies de prévention sont mises en place pour éviter de contaminer ou de se faire contaminer (utilisation du préservatif, dépistage régulier du VIH et des autres infections sexuellement transmissibles, traitement prophylactique, ou encore recours au traitement comme outils de prévention). En France par exemple, afin de pallier l'absence d'utilisation du préservatif dans les populations clés (hommes ayant des relations sexuelles avec des hommes, travailleurs du sexe, usagers de drogues, personnes originaires de région à forte prévalence, personnes ayant des partenaires multiples...), le Truvada® (association d'emtricitabine et de ténofovir, respectivement inhibiteurs nucléosidiques et nucléotidiques de la transcriptase inverse) a récemment été autorisé à titre prophylactique.

B. Les thérapies antirétrovirales : les outils pour lutter contre le virus

L'utilisation de médicaments antirétroviraux permet de supprimer la réplication virale jusqu'à un niveau indétectable (< 50 copies d'ARNs / mL de sang) permettant une reconstitution presque normale du système immunitaire (Arts and Hazuda, 2012). Ces médicaments ne permettent malheureusement pas de supprimer le réservoir viral, obligeant une prise de traitement consciencieuse et durant toute la vie du patient. L'apparition des trithérapies en 1996 a révolutionné la prise en charge des personnes infectées bien que plusieurs comprimés

devaient alors être pris à heures régulières, rendant l'observance compliquée. Depuis lors, de nombreux efforts ont été réalisés afin de faciliter l'adhésion thérapeutique.

Selon l'agence du médicament états-unienne, il existe six classes de médicaments actuellement sur le marché.

1. Inhibiteurs nucléosidiques de la transcriptase inverse (NRTIs)

Dans cette classe, sont retrouvés des analogues structurels des substrats naturels de l'activité ADN polymérase de la RT. Ainsi, toutes ces molécules n'ont pas d'extrémité 3'-OH empêchant la liaison d'un autre nucléotide sur la chaîne néo-synthétisée agissant ainsi comme un terminateur de chaîne. Souvent les NRTIs sont des pro-drogues nécessitant une étape de phosphorylation par les enzymes cellulaires pour être actives, ce qui peut être limitant pour l'activité de la molécule. Pour éviter cette étape, les inhibiteurs nucléotidiques (NtRTIs) dont le mécanisme d'action est identique aux NRTIs sont apparus.

Cette classe est la plus ancienne. Elle inclut la première molécule anti-VIH-1 mise sur le marché, l'AZT sous le nom de Retrovir® en 1987 (Chaisson RE et al., 1988). Cette classe est également aujourd'hui celle qui contient le plus grand nombre de molécules mises sur le marché (13) dont les dernières ont été approuvées en 2004-2005 (Tenofovir + Emtricitabine sous le nom de Truvada® et Abacavir + lamivudine sous le nom d'Epzicom®),

2. Inhibiteurs non-nucléosidiques de la transcriptase inverse (NNRTIs)

Ces inhibiteurs sont non compétitifs n'interférant pas directement avec la liaison des nucléotides ou du substrat de la RT (chaîne monobrin d'acide nucléique couplée à une amorce). Ces molécules se placent ainsi dans une poche hydrophobe voisine du site actif de l'activité polymérase de la RT apparaissant suite au réarrangement structural de la RT après la fixation des molécules. C'est également une classe très importante car un représentant de cette classe est toujours présent dans les combinaisons thérapeutiques. La première molécule de cette classe est apparue en 1997 (Delavirdine sous le nom commercial de Rescriptor®) et la dernière en 2012 (Rilpivirine sous le nom d'Edurant®),

3. Inhibiteurs de la maturation

Cette classe peut être divisée en deux groupes (Wang et al., 2015) :

- Inhibiteurs de protéase (PI) : ciblent et inhibent directement l'activité enzymatique de la PR. Approuvée en 1995, la première molécule de cette classe (Saquinavir sous le nom d'Invirase®) a permis d'initier le développement de nouvelles stratégies thérapeutiques. La dernière molécule de cette famille a été approuvée en 2006 (darunavir sous le nom de Prezista®);
- Inhibiteurs de maturation (MI) : interfèrent avec l'action de la PR en interagissant avec les sites de clivages sur les précurseurs. Il n'y a actuellement aucune molécule de cette classe sur le marché, le développement du chef de file de cette classe, le Bevirimat®, a été arrêté en cours d'essai clinique car le virus a rapidement développé des résistances à cette molécule (Li et al., 2003). D'autres molécules dérivées du Bevirimat® ont été synthétisées afin d'améliorer la solubilité, l'activité antivirale et diminuer la cytotoxicité (Coric et al., 2013).

4. Inhibiteur de fusion

On retrouve une seule molécule dans cette famille sous le nom d'enfuvirtide (Fuzeon®) approuvée en 2003 mais dont l'utilisation est rare de par son administration par injection et son fort coût de production. C'est un peptide synthétique de 36 acides aminés qui permet d'éviter le changement conformationnel de la gp41 indispensable à la fusion membranaire (Fletcher, 2003; Matthews et al., 2004).

5. Inhibiteur d'entrée

Une seule molécule, approuvée en 2007, permet l'inhibition de l'entrée des virus à tropisme R5 exclusivement (maraviroc sous le nom de Selzentry®). Cette petite molécule, dont l'administration se fait par voie orale, empêche la liaison entre gp120 et CCR5 (Dorr et al., 2005). Actuellement, c'est la seule molécule sur le marché qui cible les corécepteurs car les autres molécules en développement étaient hépatotoxiques ou d'effet thérapeutique limité (Aplaviroc, Vicriviroc). L'inconvénient de cette molécule est qu'elle cible uniquement les virus de tropisme R5, n'affectant pas ceux de tropisme X4 ou de double tropisme.

6. Inhibiteurs de l'intégrase (INIs)

Deux molécules sur le marché font actuellement partie de cette classe : le raltegravir (Isentress®) et le dolutegravir (Tivicay®). Ainsi, le Raltegravir® fût le premier inhibiteur de

l'Intégrase (inhibe la coupure de l'ADN hôte) à avoir été mis sur le marché (Summa et al., 2008) et est toujours considéré comme la référence pour le développement d'autres inhibiteurs de l'IN.

7. Combinaisons thérapeutiques

Aujourd'hui, la plupart de ces molécules ne sont pas administrées individuellement. Il existe des médicaments qui combinent :

- des molécules appartenant à plusieurs classes comme l'Atripla® (Efavirenz (NNRTI) + Emtricitabine (NtRTI) + Ténofovir (NRTI)) ou le Complera® (Emtricitabine (NtRTI) + rilpivirine (NRTI) + tenofovir (NRTI)),
- un PI (atazanavir ou darunavir) et un potentialisateur pharmacocinétique (cobicistat) qui permettent d'améliorer l'effet de la molécule utilisée (Evotaz® ou Prezcobix®),
- ou encore deux NRTIs (ténofovir et emtricitabine), un INI (elvitégravir) accompagnés du potentialisateur cobicistat (Stribild®).

C. Innovation, quels seront les prochains traitements ?

L'apparition de résistances aux traitements, l'échec thérapeutique causé par la faible compliance, ou encore le coût élevé de certains traitements obligent à penser à des stratégies innovantes pour permettre l'accès au traitement à un maximum de personnes VIH-1 positives. Je vais ici citer et expliquer quelques-unes des approches actuellement étudiées. La liste n'est bien sûr pas exhaustive tant le nombre d'axes de recherche dans ce domaine est important.

1. À court terme : quelles stratégies pour quelles cibles ?

À court terme, les principales innovations sont :

- **Nouvelles combinaisons thérapeutiques** : cinq sont actuellement en phase d'essai clinique II ou III (Gubernick et al., 2016) ;
- **Prises espacées dans le temps** d'antirétroviraux comme le MK-8591 (Merck), un NRTI, actuellement en phase I qui permettrait la prise du traitement une fois par semaine par voie orale ou deux fois par an par injection (Grobler et al., 2016). Une combinaison de cabotegravir (laboratoire ViiV) et de Rilpivirine (laboratoire Janssen, Edurant) est actuellement en phase II d'essai clinique pour un traitement

par injection mensuelle ou bimestrielle et en phase III pour une prise par voie orale quotidienne ;

- **Amélioration du traitement prophylactique (PrEP)** : Cabotegravir, actuellement en phase II serait efficace avec une injection tous les deux mois en intramusculaire;
- **Développement d'anticorps monoclonaux** : comme par exemple le VRC01 qui va se fixer au niveau de l'enveloppe virale (sur gp120) afin de mimer la liaison entre ce site et le récepteur CD4. Après avoir prouvé son efficacité neutralisante sur différentes souches virales et anti infectieuse sur plusieurs modèles animaux, cet anticorps monoclonal est actuellement en phase III pour une administration en PrEP, ou en phase I comme traitement antiviral (Lynch et al., 2015) ;
- **Nouvelles classes d'antirétroviraux** :
 - Inhibiteurs de maturation : BMS-955176 qui, en bloquant la dernière étape de clivage lors de la maturation de Gag, produit des particules virales immatures non infectieuses (Bristol-Myers Squibb Company, 2015) ;
 - Inhibiteurs de contact : BMS-663068, actuellement en phase III, inhibe la liaison entre le virus et la cellule cible (Lalezari et al., 2015). Cette molécule permettrait de compléter cette classe d'inhibiteurs qui ne comporte actuellement que le Maraviroc inhibant seulement l'infection des virus à tropisme R5 uniquement ;
 - Inhibiteurs de Nucléocapside : un de mes projets de thèse étant le développement de cette classe, les différentes molécules ciblant cette protéine seront décrites dans la partie [VI] de ce chapitre.

2. À long terme : vers la suppression du réservoir viral et la guérison

À long terme, l'objectif est de supprimer le réservoir viral. En effet, l'ensemble des thérapies actuellement sur le marché ou celles citées ci-dessus ne permettent ou ne permettront pas de guérir du VIH-1 (Barré-Sinoussi et al., 2013). De plus, ces traitements sont coûteux, souvent toxiques pour l'organisme (surtout à long terme) et l'adhésion thérapeutique des patients est compliquée sur une longue période.

Le réservoir viral (aussi appelé latence thérapeutique) correspond à l'intégration persistante d'ADN viral (qui n'est pas activement transcrit et reste invisible pour les traitements antirétroviraux), dans le génome des lymphocytes T CD4⁺ mémoires et naïfs principalement.

D'autres types cellulaires semblent également y jouer un rôle (Churchill et al., 2016). Elle constitue une barrière majeure pour soigner le VIH-1. L'implication de plusieurs mécanismes cellulaires a été démontrée dans la mise en place du réservoir :

- régulation négative des facteurs de transcription cellulaire qui participent à la transcription virale (NF- κ B, PTEF-b) (Tyagi et al., 2010),
- inactivation de la transcription par action d'histones déacétylases (HDAC) recrutées au niveau des LTRs à proximité du site d'initiation de la transcription (Coull et al., 2000). Ceci est en accord avec le fait que l'inactivation de ces enzymes (par le Vorinostat® par exemple) entraîne un rebond de la quantité d'ARNs dans le sang et semble perturber le réservoir viral (Archin et al., 2012). Un essai clinique a montré que cette réactivation du réservoir viral semble efficace chez l'Homme (Søgaard et al., 2015).

Le mécanisme exact de la mise en place et du maintien de la latence virale reste encore un mystère. Afin de développer de nouvelles stratégies thérapeutiques, il est indispensable de comprendre les mécanismes impliqués dans la mise en place de la latence virale : restriction de l'initiation de la transcription ; l'épissage, l'export nucléaire et la traduction des ARNm viraux ; l'expression et l'exposition des antigènes viraux (Deeks et al., 2016).

Plusieurs stratégies sont étudiées afin de développer un traitement curatif du VIH-1 :

- **Action moléculaire :**
 - Activateurs de latence virale : au même titre que les inhibiteurs de HDAC, cette classe de molécules permettrait de rendre « visible » la latence virale en rendant la production et la libération de virus de nouveau effective. Les cellules contenant le virus latent seraient alors démasquées et supprimées par le système immunitaire : c'est la stratégie « shock and kill ». L'administration d'activateurs de latence virale à des patients sous traitements antirétroviraux a montré une augmentation de la quantité d'ADN viral circulant (Elliott et al., 2015; Rasmussen et al., 2014). Il est donc indispensable de développer une seconde génération de molécules de cette classe afin d'améliorer la spécificité et de diminuer la toxicité.
 - Bloqueurs permanents du réservoir : à l'inverse de la stratégie précédente, l'idée ici est de bloquer de manière irréversible et spécifiquement la

transcription virale en bloquant les protéines nécessaires à l'activation de la transcription (par exemple CycT-1 ou Cdk9 membres du complexe P-TEFb; Mousseau et al., 2015).

- **Action sur le système immunitaire :**

- Développement de vaccins induisant la production de lymphocytes T cytotoxiques (lymphocytes T CD8⁺) dirigés contre de nouveaux épitopes, permettant ainsi l'augmentation du contrôle immunitaire (Casazza et al., 2013). D'autres stratégies ciblant les lymphocytes T sont également en cours de développement (pour revue Leibman and Riley, 2015) ;
- Développement d'anticorps monoclonaux comme le VRC01 par exemple (Lynch et al., 2015). Ces anticorps pourraient être utilisés pour supprimer la réplication du virus, maintenir la suppression virale induite par les thérapies antirétrovirales, ou encore réduire le réservoir viral ;
- Diminution de l'inflammation chronique observée chez les personnes contaminées par le VIH-1 qui contribue à la persistance du virus lors du traitement antirétroviral (Micci et al., 2015).

- **Thérapie génique et cellulaire** (Dey and Pillai, 2015) :

- Le premier type de thérapie est basé sur l'observation d'un patient positif au VIH-1 qui a été guéri suite à une transplantation de cellules souches hématopoïétiques d'un donneur portant une mutation sur le gène du récepteur CCR5 (CCR5Δ32 qui permet aux caucasiens homozygotes porteurs de ce gène de ne pas être sensibles à l'infection). Ce patient dit « patient Berlin », Timothy Brown, est ainsi devenu résistant à l'infection au VIH-1 (Hütter et al., 2009). L'approche ici correspondrait à l'utilisation de cellules souches hématopoïétiques du patient à traiter dans lesquelles la délétion de l'expression du récepteur CCR5 serait réalisée par l'utilisation de nucléases à doigt de zinc (ZFNs, permet de supprimer une séquence d'ADN spécifique) avant réintroduction chez le patient (Cannon and June, 2011; Perez et al., 2008),
- Suppression des gènes codant des protéines indispensables à l'infection par le VIH-1 par des méthodes d'ingénierie génomique :

- limiter l'expression de CCR5 à la surface des cellules ciblées par le VIH par utilisation de nucléases des effecteurs mimant l'activation de transcription (TALENs ; Mussolino et al., 2011) bloquant la capacité d'infection des virus à tropisme R5,
- supprimer l'ADN viral complet intégré grâce au système CRISPR-Cas9 (Kaminski et al., 2016). Cette dernière stratégie doit maintenant être démontrée *in vivo*.

Aucune des stratégies citées ci-dessus n'est considérée comme étant **LA** méthode curative du VIH-1. Ainsi, une combinaison de différentes stratégies semble être la solution la plus probable pour aboutir au développement d'un traitement efficace, peu ou pas cytotoxique, sans résistance développée par le virus et peu coûteux.

VI. Pourquoi la NCp7 est une bonne cible thérapeutique ?

A. Structure de la nucléocapside

La Nucléocapside possède de nombreux rôles et existe sous deux formes au cours de l'infection : la première comme domaine de Gag (nommé NC dans la suite de ce manuscrit) et la seconde comme protéine à part entière (nommée NCp7 dans la suite de ce manuscrit) une fois que la maturation de Gag est effectuée. Comme il a été montré dans la **Figure 10**, la NCp7 apparaît lors du clivage séquentiel du précurseur Gag par la protéase virale au sein du virion (Pettit et al., 1998). Après un premier clivage entre les domaines Sp1 et NC, un intermédiaire, appelé NCp15 composé des domaines NC, Sp2 et p6 est généré. Ce premier clivage est indispensable pour la formation et la stabilisation du dimère d'ARNs (Shehu-Xhilaga et al., 2001). Un deuxième clivage libère la protéine p6 de ce précurseur, formant un second intermédiaire, NCp9 composé de NC et Sp2. Finalement, même si l'infection virale n'est pas impactée par celui-ci (Coren et al., 2007), un troisième et dernier clivage libère la NCp7 qui est la forme la plus efficace en termes d'activité chaperonne.

NCp7 est une petite protéine basique composée de 55 acides aminés et possède deux structures en doigts de zinc très conservées flanquées par de petits domaines basiques (**Figure 26**). Les doigts de zinc permettent la fixation de deux ions zinc via les acides aminés cystéines et histidines du motif (Cys-X₂-Cys-X₄-His-X₄-Cys, ou X représente un acide aminé quelconque) avec une forte affinité ($K_d = 10^{13}$ - 10^{14} M^{-1}) (Mély et al., 1996). La fixation des ions zinc permet une modification conformationnelle de la protéine rapprochant les deux doigts de zinc et exposant les plateaux hydrophobes proximal (Val13, Phe16, Ile24 et Ala25) et distal (Trp37, Gln45 et Met46).

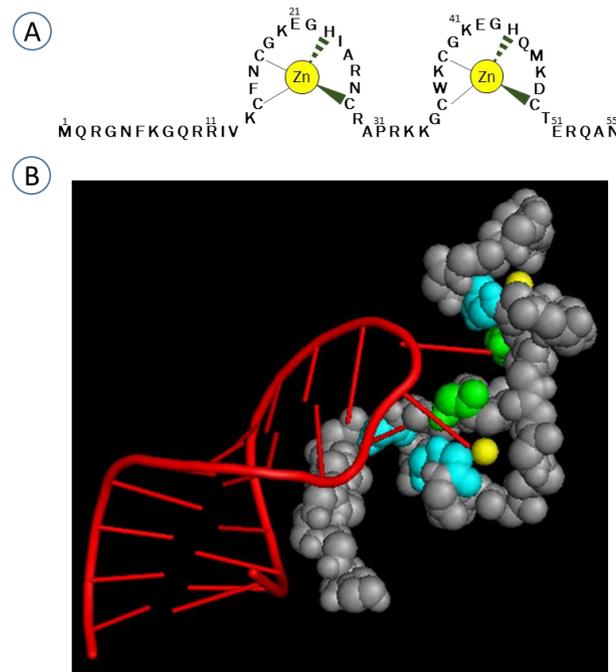
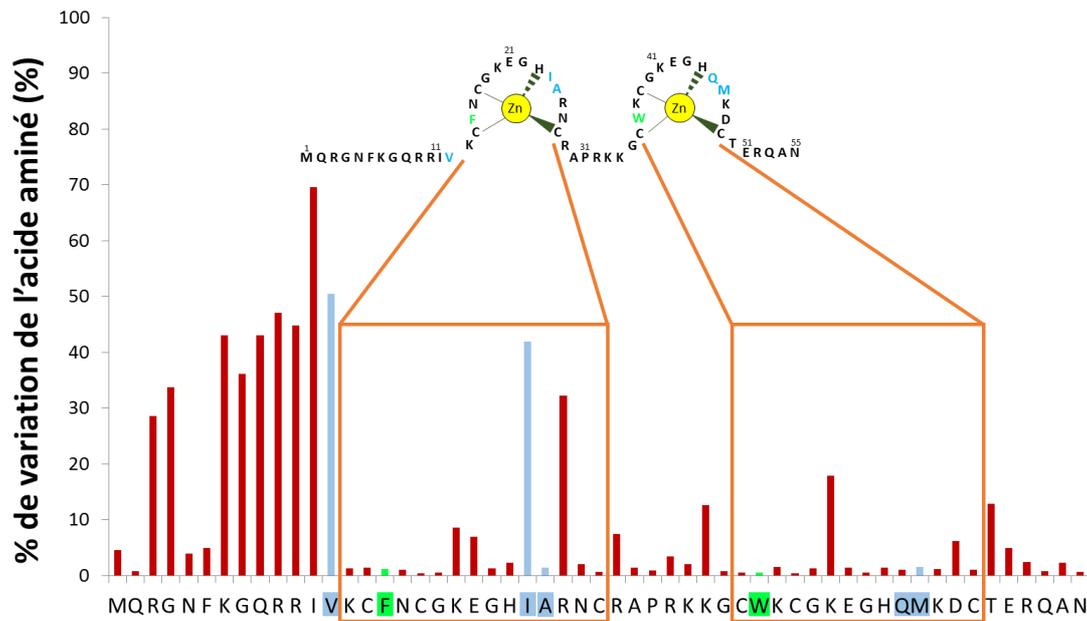


Figure 26 : Séquence et structure de la Nucléocapside. (A) Séquence d'acides aminés de la NCp7 montrant les deux doigts de zinc. (B) Structure tridimensionnelle de l'interaction entre NCp7 et la boucle SL3 de l'ARN viral (selon PDB 1A1T). Les sphères jaunes représentent les ions zinc. Les résidus de NCp7 impliqués dans la formation du plateau hydrophobe (Val13, Ile24, Ala25, Gln45, Met46) sont représentés en bleu, ceux retrouvés dans la liaison avec les acides nucléiques (Phe16, Trp37) en vert et les autres en gris.

La NCp7 possède une séquence très conservée dans les différentes souches du virus (**Figure 27**). La partie N-terminale est très variable, alors que celle des doigts de zinc est très conservée. Il a été démontré qu'une mutation non conservative dans le plateau hydrophobe de la NCp7 entraîne la production de particules virales non infectieuses (Dorfman et al., 1993; Wu et al., 2013). De la même manière, une mutation dans les résidus de liaison du zinc rend l'infection virale non efficace (Déméné et al., 1994; Dorfman et al., 1993), sans modifier la capacité de fixation du zinc (Bombarda et al., 2005), mais en empêchant le repliement correct de la protéine. Habituellement, ce repliement permet de favoriser la liaison avec les acides nucléiques et plus particulièrement avec les guanines via le Trp37, un des acides aminés les plus importants et conservés de la NCp7.

Enfin, cette petite protéine est très flexible ce qui lui permet de lier de nombreuses séquences d'acides nucléiques pour lesquelles elle possède une propriété chaperonne (Darlix et al., 2011; Godet et al., 2013; Levin et al., 2005). Toutefois cela rend le développement d'inhibiteurs anti-VIH-1 ciblant NCp7 difficile.



Fait par Francesco Saladini, 2014

Figure 27 : La séquence de la NCp7 est hautement conservée. Cet histogramme montrant la fréquence de variation de chaque acide aminé constituant la NCp7 a été réalisé par l'alignement de 10493 séquences de patients issues de la LANL HIV Database (réalisé par Francesco Saladini, Université de Sienne, collaborateur du projet THINPAD, en 2014). Le bleu permet de mettre en évidence les résidus impliqués dans la formation du plateau hydrophobe (Val13, Ile24, Ala25, Gln45 et Met46) et en vert les acides aminés aromatiques impliqués dans la liaison avec les acides nucléiques (Phe16 et Trp37).

B. Fonctions de la NucléoCapside : petit mais costaud

1. Liaison des acides nucléiques

Une des premières fonctions décrites de la NCp7 a été la liaison aux acides nucléiques grâce à ses domaines en doigts de zinc. Cette protéine est ainsi capable de se lier de manière non spécifique à n'importe quelle séquence de 5 à 8 nucléotides (nt) mais généralement avec une faible affinité (Darlix et al., 2011). Ainsi, que ce soit en tant que domaine de Gag ou en tant que protéine à part entière, la liaison de la NCp7 avec les acides nucléiques (AN) est évidente mais le rôle joué par la protéine dépendrait de son taux d'occupation des séquences d'AN (Rapport molaire NC/AN ; **Figure 28**) :

- À faible ratio NC/AN (1/100) :
 - Phases tardives, la protéine au sein de Gag, fixe l'ARN viral avec une préférence pour les motifs GGAG ou GGUG retrouvés sur les boucles SL1 (Abd El-Wahab et al., 2014), SL2 (Amarasinghe et al., 2000), SL3 (Guzman

et al., 1998) ou au niveau des séquences PBS (motif TG, Bourbigot et al., 2008) ou cTAR (Bazzi et al., 2011) du génome viral.

- Phases précoces, l'affinité étant moins forte pour l'ADN double brin, la NCp7 se détache au fur et à mesure lors de la transcription inverse. Ainsi quelques molécules de NCp7 se retrouveraient dans le noyau où elles peuvent stimuler l'intégration.
- Un ratio NC/AN intermédiaire (de 1/15 à 1/7) permettrait l'agrégation des acides nucléiques :
 - Phases tardives, lors de l'assemblage où le domaine NC de Gag favoriserait alors la dimérisation des deux copies d'ARNs viral (van Bel et al., 2014)
 - Phases précoces, immédiatement après l'entrée de la particule virale lors de la décapsidation progressive et de la transcription inverse où la NCp7 jouerait un rôle dans le transfert de brin (Bampi et al., 2006).
- À fort ratio (de 1/5nt à 1/2nt) où la protéine recouvre alors complètement les acides nucléiques. Cette situation est retrouvée au sein de la nucléocapside virale où les nombreuses NCp7 recouvrent les deux brins d'ARNs viraux les protégeant d'une dégradation en formant un complexe nucléoprotéique compact.

La NCp7 lie préférentiellement l'ARN, mais elle peut également lier l'ADN. La liaison de l'un ou l'autre de ces acides nucléiques par la NCp7 se fait par une orientation différente (Bazzi et al., 2011) :

- lors de la fixation d'un ribose (ARN), la NCp7 est positionnée de façon parallèle à la chaîne nucléotidique (5' vers 3') pour favoriser l'élongation de l'ARN impliqué dans l'empaquetage puis la dimérisation ;
- lors de la fixation d'un désoxyribose (ADN), la NCp7 se positionne plutôt de manière anti-parallèle (3' vers 5') pour déstabiliser l'ADN (permet les deux transferts de brins de la transcription inverse).

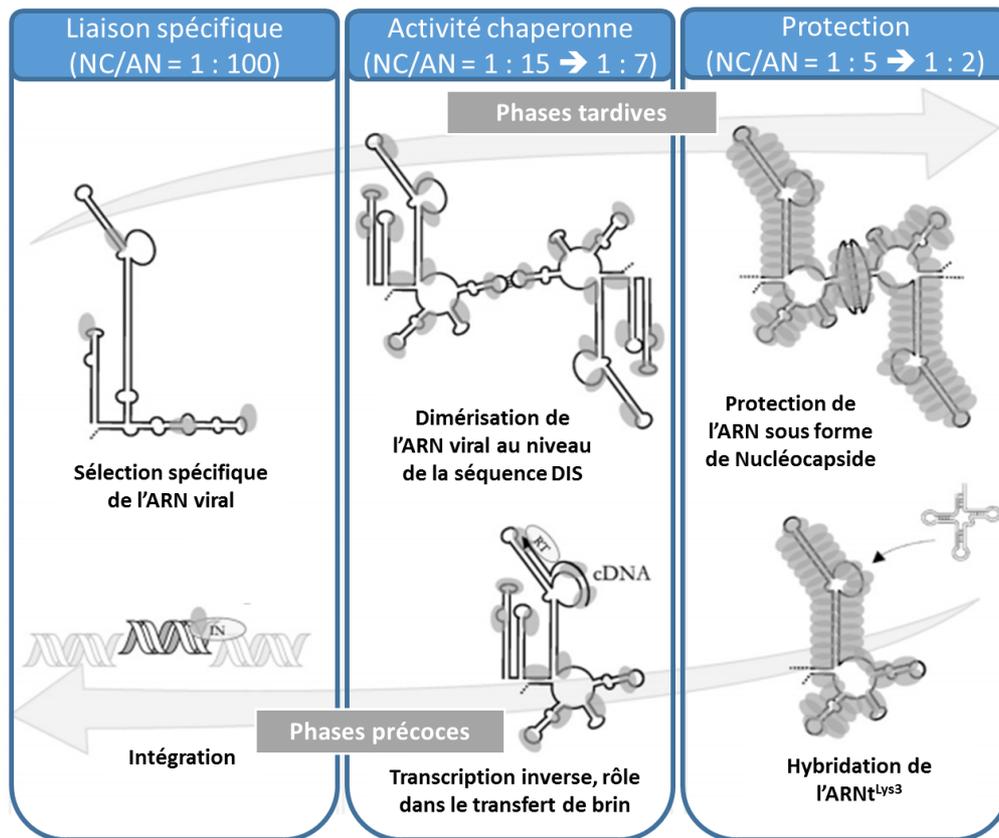


Figure 28 : Différentes étapes du cycle viral dans lesquelles la NC est impliquée et dont l'activité dépend du ratio NC/AN. Le taux d'occupation de la NCp7 : nucléotides (mol : mol) est compris entre 1 : 100 à gauche et 1 : 2 à droite et définit la fenêtre d'activité de la protéine (adaptée de Darlix et al., 2011).

2. Activité chaperonne

En plus de sa capacité de liaison aux acides nucléiques, la NC possède une activité chaperonne pour ces derniers. Celle-ci intervient lorsque la protéine ne recouvre que partiellement les ANs (fenêtre centrale dans la **Figure 28**) que ce soit lors des étapes tardives (assemblage) ou précoces (transcription inverse). La caractéristique d'une protéine chaperonne des acides nucléiques (ARN et/ou ADN) est d'assister les ANs pour qu'ils atteignent plus facilement leur conformation la plus stable. L'activité chaperonne de la NCp7 repose ainsi sur trois propriétés :

- Capacité d'agréger les acides nucléiques,
- Capacité à se lier et se séparer rapidement des acides nucléiques,
- Déstabilisation partielle de la structure des acides nucléiques.

Il est important de noter que les formes intermédiaires de la nucléocapside (Gag, NCp15 et NCp9) ont également une activité chaperonne qui augmente avec le degré de maturation (Gag < NCp15 < NCp9 < NCp7) (Cruceanu et al., 2006) .

C. Rôles de la nucléocapside dans le cycle viral : le petit plus qui change tout

1. Dans les phases précoces

La NCp7 intervient au niveau de plusieurs étapes des phases précoces du cycle viral du VIH-1. Ces différents rôles montrent ainsi l'importance de cette petite protéine.

a. Au niveau de la transcription inverse

Cette étape du cycle viral permet la synthèse de l'ADN proviral à partir de l'ARN génomique viral apporté lors de l'infection (voir partie [III-D] de ce manuscrit). Afin d'améliorer l'efficacité et la rapidité de cette étape, la NCp7 intervient au niveau de plusieurs étapes :

- **Initiation** : la NCp7 (Tisé et al., 2001) ou le domaine NC de Gag (Cen et al., 1999) permet de stimuler la formation du complexe entre la séquence PBS de l'ARN viral et l'ARNt^{Lys3} indispensable à l'initiation de la transcription inverse. La structure obtenue, thermodynamiquement stable, permet le recrutement de la transcriptase inverse qui pourra débiter son action (**Figure 29**). L'hybridation des séquences complémentaires PBS (sur l'ARN viral) et anti-PBS (sur l'ARNt^{Lys3}) se fait par plusieurs réarrangements facilités par la NCp7 qui accélère fortement la vitesse d'hybridation (Hargittai et al., 2001; Tisé et al., 2001).

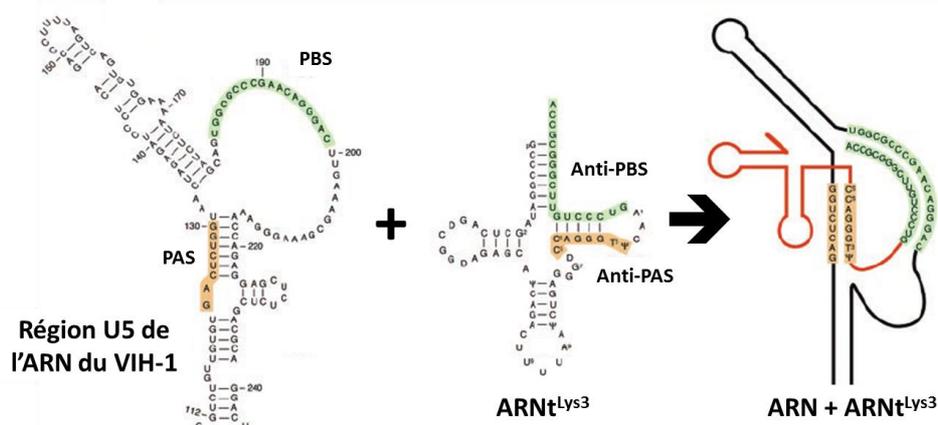


Figure 29 : Modèle d'initiation de la transcription par liaison de l'ARNt^{Lys3} sur la séquence PBS présente dans la région U5 de l'ARN viral. PAS = Primer Activation Signal ; PBS = Primer Binding Site. Dans le complexe à droite, on retrouve la structure de l'ARN viral en noire et celle de l'ARNt^{Lys3} en rouge (adaptée de (Abbink et al., 2004)).

- **Synthèse du brin ADN (-)** : il a été démontré *in vitro* que l'ADN généré par transcription inverse est fréquemment incomplet. Ceci résulte des « pauses » de la RT qui apparaissent souvent au niveau des structures en tige-boucle de l'ARN viral. La présence de NCp7 permet la déstabilisation de ces structures ce qui facilite le rôle de la RT (Rein et al., 1998; Sharma et al., 2016) ;
- **Premier saut de brin** : dans la transcription inverse, le premier saut de brin est requis lors de la synthèse du brin d'ADN (-) (**Figure 19 C**). Le fragment d'ADN, dit « strong-stop » est synthétisé par la RT à partir de l'amorce ARN^{t^{Lys3}}. Le premier transfert de brin consiste en l'hybridation de la séquence R de l'ADN (-) « strong-stop » nouvellement synthétisé avec la séquence complémentaire R en 3' de l'ARN génomique viral. La NCp7 en déstabilisant les structures en tige de la région R (TAR pour l'ARN et cTAR pour l'ADN) (Azoulay et al., 2003; Bernacchi et al., 2002) accélère leur hybridation (Godet et al., 2006; Vo et al., 2006). La NCp7 semble également stimuler l'activité RNase H de la RT accélérant alors la dégradation de l'ARN matrice (Peliska et al., 1994) ;
- **Second saut de brin** : il est nécessaire à la synthèse du brin d'ADN (+) dans la transcription inverse (**Figure 19 E**). Ce saut de brin implique l'hybridation de la séquence PBS(-) (située sur le brin ADN (-) précédemment synthétisé) avec la séquence PBS(+) (située sur le brin ADN(+)) qui sont toutes les deux structurées sous forme de tige-boucle. En absence de NCp7, la formation du duplex PBS (+)/(-) passe par une interaction entre les extrémités protrudentes (bases en rouge en haut de la **Figure 30**; Ramalanjaona et al., 2007). En présence de NCp7, il y a extension des boucles PBS et exposition des bases de la boucle qui peuvent alors participer à des interactions boucle/boucle (bases bleues en bas de la **Figure 30**; Godet et al., 2011).

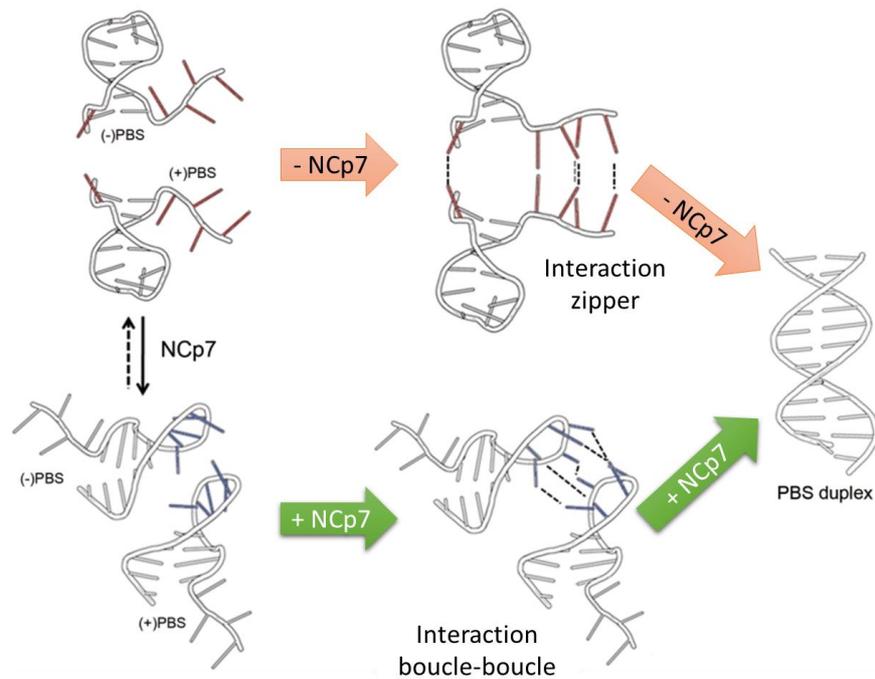


Figure 30 : Rôle de la NCp7 dans le second transfert de brin. L'activité chaperonne de la NCp7 favorise la formation du duplex PBS (+)/(-) en permettant la formation d'un intermédiaire boucle-boucle, alors qu'en absence de cette protéine, la formation du duplex plus lente et moins efficace passe par un intermédiaire de type zipper (Godet et al., 2011).

b. Dans le complexe de pré-intégration

La NCp7 a été retrouvée au niveau des pores nucléaires (Arhel et al., 2007). Comme nous l'avons décrite précédemment, la NCp7 interagit probablement avec le site permettant le flap de l'ADN lors de la dernière étape de la transcription inverse (Mirambeau et al., 2007). Cette dernière étape constituerait l'élément déclencheur de l'import nucléaire de l'ADN viral. La NCp7 serait également présente dans le PIC, mais son rôle reste encore à prouver (Lyonnais et al., 2013).

c. Au niveau de l'intégration

La protéine NCp7 se localise dans le noyau (Zhang and Crumpacker, 2002) et se concentre dans les nucléoles (Anton et al., 2015; Yu et al., 2016) après surexpression. NCp7 a également été décrite dans le noyau huit heures après infection. Diverses études réalisées *in vitro* montrent que la NCp7 stimule le transfert de brin du processus d'intégration en présence d'ions Mg^{2+} . Cette activité est dépendante des doigts de zinc et de la région basique de la NCp7 et est observée même à faible concentration de NCp7 (Carteau et al., 1997, 1999; Gao et al., 2003; Poljak et al., 2003). Ces résultats *in vitro* sont soutenus par des résultats en cellules

montrant que des mutations induites dans la nucléocapside au niveau des doigts de zinc perturbent l'intégration de l'ADN viral dans l'ADN cellulaire (Buckman et al., 2003; Thomas et al., 2006). Cependant, le mécanisme exact de cette stimulation de l'intégration par la NCp7 reste mal connu (Thomas and Gorelick, 2008).

2. Dans les phases tardives

a) Sélection spécifique de l'ARN viral

La nucléocapside est capable de se lier de manière spécifique à des ANs. L'affinité peut varier en fonction de la nature (ARN ou ADN), de la séquence et de la structure de l'AN (Vuilleumier et al., 1999). La reconnaissance spécifique par le domaine NC de Gag, de la séquence d'encapsidation Ψ , permet la sélection du génome viral perdu comme une aiguille dans une botte de foin au sein d'un large excès d'ARNs cellulaires (Cimarelli and Darlix, 2002; Muriaux and Darlix, 2010). Il semble d'ailleurs que cette interaction Gag-ARN via le domaine NC soit indispensable à l'initiation de l'assemblage des particules. En effet, la délétion du domaine NC retarde la production des particules virales au niveau de l'initiation de l'assemblage, mais, une fois démarrée, cette production se déroule à une vitesse similaire à celle du virus sauvage (Ott et al., 2009). En effet, la délétion du domaine NC diminue la capacité de multimérisation de Gag, entraînant la diminution de l'efficacité de l'assemblage (El Meshri et al., 2015).

b) Dimérisation des ARNs viraux

La dimérisation du génome viral est indispensable pour que l'infection et notamment la transcription inverse soit efficace. Au cours de l'assemblage, cette dimérisation a lieu par association de séquences palindromiques complémentaires appelées DIS, ou site d'initiation à la dimérisation. Cette séquence serait indisponible lorsque l'ARN est sous forme linéaire (**Figure 31 A**). La nucléocapside permet la modification de la structure de l'ARN en formant une structure en « épingle à cheveux » (**Figure 31 B**) (Feng et al., 1996b; Mundigala et al., 2014). Cette dernière expose la séquence DIS qui pourra être associée sous forme de « kissing complex » à la structure DIS d'un autre ARN (**Figure 31 C**). Finalement, la nucléocapside permet la formation et l'extension des duplex d'ARNs sous forme de dimère (**Figure 31 D**; Rist and Marino, 2002).

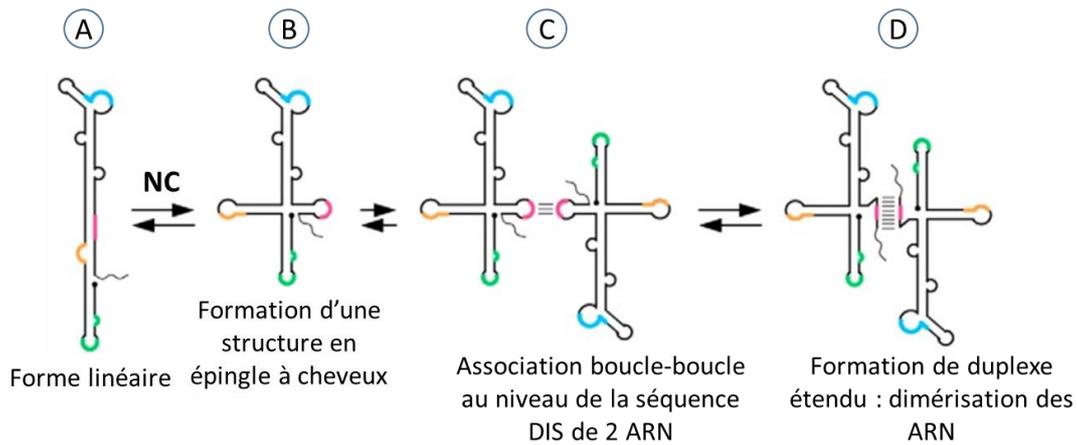


Figure 31 : Mécanisme de dimérisation de l'ARN du VIH-1 lors de l'assemblage. La séquence d'initiation de la dimérisation DIS apparaît en rose. Elle reste non accessible lorsque l'ARN est sous forme linéaire (A) mais est exposée lorsque celui-ci adopte une structure en épingle à cheveux (B). Il y a alors formation d'un « kissing complex » (interaction boucle-boucle) au niveau de la séquence palindromique DIS de deux ARNs viraux (C), puis extension de l'hybridation (D) pour former alors un dimère d'ARNs (d'après van Bel et al., 2014).

D. Interactions de la NucléoCapside avec les protéines cellulaires : qui sont les partenaires de cette protéine ?

Que ce soit en tant que domaine de Gag ou après maturation, la nucléocapside interagit avec de nombreuses protéines cellulaires et virales. Le **Tableau 2** récapitule l'ensemble des partenaires ainsi que leurs rôles physiologiques et dans le cycle viral. D'autres études ont été menées afin d'identifier de nouveaux partenaires cellulaires de la protéine Gag. Plusieurs centaines d'interactions ont ainsi été révélées (Brass et al., 2008; Jäger et al., 2012). La validation et la caractérisation de ces interactions pourraient permettre le développement de molécules ciblant de nouvelles cibles thérapeutiques.

Partenaire d'interaction	Rôle de la protéine	Domaine(s) d'interaction de Gag	Rôle dans le cycle viral	Publication(s)
Actine	Microfilaments d'actine : constituant essentiel du cytosquelette	NC	Interaction Gag/actine pourrait jouer un rôle dans l'assemblage et/ou d'autres étapes du cycle de réplication	(Liu et al., 1999; Rey et al., 1996; Wilk et al., 1999)
AIP1 = ALIX (Apoptosis-Linked gene 2-Interacting protein X)	Protéine adaptatrice impliquée dans plusieurs processus intracellulaires (apoptose, endocytose, trafic membranaire...)	NC et p6	Nécessaire pour le bourgeonnement du virus	(Dussupt et al., 2009; Popov et al., 2008, 2009; Strack et al., 2003; Usami et al., 2007)
A3G (APOBEC3G)	Superfamille des cytidines déaminases, enzyme ayant un rôle dans l'immunité innée	NC (région entre les deux doigts de zinc)	A3G incorporé dans les virions, diminue la transcription inverse et l'infectivité du VIH-1	(Alce and Popik, 2004; Cen et al., 2004; Douaisi et al., 2004; Friew et al., 2009)
DDX6 et DDX9 = RHA (RNA Hélicase A)		NC	DDX6 facilite l'assemblage de la capsid ; DDX9 stimule la réplication du virus aux étapes de traduction et de transcription inverse et l'hybridation de l'ARNt ^{Lys,3} à l'ARN du VIH-1	(Reed et al., 2012; Roy et al., 2006; Xing et al., 2011)
EF1α (Elongation Factor 1 - α)	Composant de la machinerie traductionnelle, délivrant les aminoacyl-ARNt aux ribosomes	MA et NC	EF1α permettrait l'incorporation de l'ARNt dans les virions. Dans la cellule cible, ce partenaire fait partie du complexe de transcription inverse où il améliore cette étape du cycle viral.	(Cimarelli and Luban, 1999; Warren et al., 2012)

IMP1 (Insulin-like growth factor II mRNA binding Protein 1)	Régule le transport de l'ARN et la traduction locale	NC (second doigt de zinc)	Empêche la formation des particules virales et diminue l'infectivité des virions	(Zhou et al., 2008)
IP5 (Inositol Pentaphosphate) et IP6 (acide phytique)	Rôle dans la croissance cellulaire, la différenciation, la migration, l'endocytose ou encore l'apoptose	MA et NC	Essentielle pour l'assemblage des particules virales	(Campbell et al., 2001; Datta et al., 2007b)
Lyric (HIV-inducible gene)	Impliquée dans plusieurs voies de signalisation, aurait des effets anti-apoptotiques et serait impliquée dans la tumorigénèse	MA et NC	Lyric pourrait réguler l'infectivité virale	(Engeland et al., 2011)
SOCS1 (Suppressor of Cytokine Signaling protein 1)	Réduit l'inflammation	MA et NC	SOCS1 est un facteur cellulaire inductible durant l'infection par le VIH-1, capable de réguler les stades tardifs du cycle de réplication	(Nishi et al., 2009; Ryo et al., 2008)
SNARE (Soluble N-ethylmaleimide-sensitive factor Attachment protein REceptor)	Famille de protéines : essentielles pour la fusion des membranes. Elles sont nécessaires pour chaque étape de l'exocytose ainsi que les voies d'endocytose	NC	SNARE affecte l'assemblage et le relargage du VIH-1, en empêchant probablement le trafic et l'association de Gag à la membrane plasmique	(Joshi et al., 2011)

Staufen-1 (double-stranded RNA-binding protein)	Impliquée dans de nombreux procédés liés à l'ARN (localisation...)	NC	Staufen-1 influence la multimérisation de Gag	(Chatel-Chaix et al., 2008; Milev et al., 2010)
Tsg101 (Tumor susceptibility gene 101)	Famille des VPs (Vacuolar Protein sorting)	p6 et NC	Permet le bourgeonnement du virus	(Dussupt et al., 2011; Garrus et al., 2001)
RPL7 (ribosomale protein L7)	Protéine de la grande sous-unité ribosomale	NC	Permet d'améliorer l'activité chaperonne de Gag	(El Mekdad et al., 2016)
Gag	Protéine structurale du VIH-1	CA et NC	La multimérisation de Gag est un événement clé dans l'assemblage du VIH-1	(Hogue et al., 2009; Kutluay and Bieniasz, 2010; Maldonado et al., 2014)
Vif	Protéine régulatrice du VIH-1	NC et un domaine entre MA et l'extrémité N-terminal de CA	Interaction Vif/Pr55 ^{Gag} : rôle dans la réplication	(Bouyac et al., 1997)
Vpr	Protéine régulatrice du VIH-1	p6 et NC	Vpr est encapsidé dans le virion par son interaction avec p6 et cela permet à Vpr d'exercer ses fonctions lors de l'infection	(Paxton et al., 1993; de Rocquigny et al., 1997; Zhu et al., 2004)

Tableau 2 : Partenaires cellulaires et viraux connus de la protéine NCp7 ou du domaine NC de Gag (extrait de Klingler, 2015)

E. Existe-il des molécules antivirales ciblant la Nucléocapside ?

La nucléocapside en intervenant dans de nombreuses étapes du cycle viral est une cible thérapeutique intéressante. De plus, cette protéine est fortement conservée dans les différents sous-types du VIH-1 et les souches virales résistantes. Les inhibiteurs de NC peuvent être classés en trois classes : les éjecteurs de zinc, les inhibiteurs non covalents et les oligonucléotides, aptamères et autres peptido-mimétiques.

1. Éjecteurs de zinc

Les doigts de zinc étant très conservés dans les différentes souches de VIH-1 (*Figure 27*) et impliqués dans la plupart des étapes du cycle viral, ces structures sont la cible pour le développement de nouveaux inhibiteurs. La liaison du zinc est indispensable pour donner la structure caractéristique de la NC. C'est pourquoi, les premières molécules ciblant la nucléocapside ont été des éjecteurs de zinc. L'activité de ces éjecteurs est basée sur leurs capacités à altérer l'interaction de la NCp7 avec ses acides nucléiques cibles (Huang et al., 1998b; Jenkins et al., 2005; Tummino et al., 1996). Ainsi, en agissant sur la NCp7, les éjecteurs de zinc ont montré qu'ils pouvaient perturber la transcription inverse (Morcock et al., 2005; Pannecouque et al., 2010; Rice and Turpin, 1996; Sharmeen et al., 2001), l'activité chaperonne zinc-dépendante de la NCp7 (Pannecouque et al., 2010; Vercruysse et al., 2012), ou encore l'assemblage des particules virales (Miller Jenkins et al., 2010; Turpin et al., 1999).

Il est possible de les classer en 3 groupes (Mori et al., 2015) :

- **Attaque électrophile des doigts de zinc** accompagnée par la formation de pont disulfure intra- ou inter-moléculaire entre les résidus cystéines des doigts de zinc. Cette classe contient les molécules des familles NOBA (3-nitrobenzamide) et DIBA (disulfide benzamide), ou encore le WDO-217 (3-phenyl-2H-[1,2,4] thiadiazol-5-ylideneamine) (*Tableau 3* :). L'attaque électrophile peut également être accompagnée d'acétylation pour les familles PATEs (pyridinioalkanoyl thioester) et SAMTs (S-acyl-2-mercaptobenzamide thioester).
- **Chélation de zinc** par interaction avec la liaison ester du NV38 (N,N-bis(4-ethoxycarbonyl-1,2,3-thiadiazol-5-yl)benzene-1,2-diamine) (*Tableau 3*).

- **Liaison covalente des cystéines** impliquées dans la liaison au zinc par le platine (Pt). Il y a d'abord une reconnaissance du Trp37 de la NCp7, puis formation d'une liaison covalente entre le Pt et les cystéines. Parmi ces molécules, on trouve par exemple le [SP-4-2]-[PtCl(NH₃)(quin)(9-EtGH)]

Comme le montre leur CI₅₀ (concentration inhibant 50 % de l'activité virale), la plupart de ces éjecteurs de zinc ont une activité antivirale à faible concentration sur des cellules infectées par le VIH (voir colonne CI₅₀ dans le [Tableau 3](#)). Certaines d'entre elles se sont également avérées efficaces sur des souches résistantes aux traitements actuels de VIH-1 et sur le VIH-2 ou le SIV. Enfin, l'absence d'apparition de souches virales résistantes à ces molécules confirme le fort potentiel de ces inhibiteurs.

En revanche, la toxicité de ces molécules, mesurée par la concentration nécessaire pour tuer 50 % des cellules (CC₅₀), est souvent un frein important pour le développement d'éjecteurs de zinc comme inhibiteurs de NCp7. En effet, un certain nombre de protéines cellulaires possèdent des doigts de zinc (PARP, ou poly(ADP-ribose) polymérase ; GATA-1 par exemple) expliquant ainsi la forte cytotoxicité de ces molécules. C'est pourquoi seulement deux molécules ont été testées cliniquement chez l'homme, montrant peu d'efficacité antivirale et/ou une forte toxicité. La famille des SAMT, testée dans un modèle primate, n'a pas montré d'efficacité sur la charge virale mais est apparue comme étant non cytotoxique. Cette caractéristique, couplée au fait que ces molécules présentent un effet antiviral *in cellulo*, ont permis d'envisager le développement de virucides à application locale qui utilisent ces molécules. Ceux-ci se sont avérés efficaces dans l'inhibition de la transmission virale *in vivo* (Wallace et al., 2009).

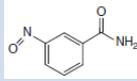
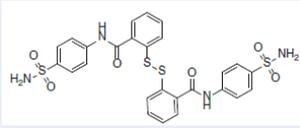
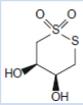
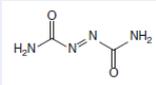
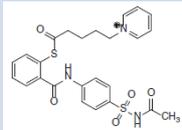
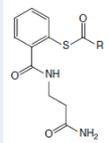
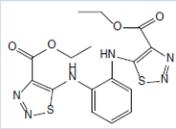
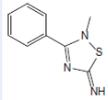
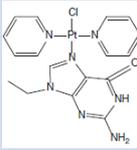
Molécule	Cl ₅₀ (μM)	CC ₅₀ (μM)	IT	Structure	Références
NOBA	ND	10,6	-		(Huang et al., 1998b)
DIBA-1	2,3	>200	>87		(Huang et al., 1998b; Rice et al., 1995)
Dithiane	6,6	184	30		(Huang et al., 1998b)
ADA	38	>200	>5		(Huang et al., 1998b; Rice et al., 1997a, 1997b)
PATE-45	6,2	>316	>51		(Turpin et al., 1999)
SAMT-19	2,9	461	160		(Srivastava et al., 2004)
NV038	17	>300	>17		(Pannecouque et al., 2010)
WDO-217	7,9	72	9		(Vercruyse et al., 2012)
[SP-4-2]- [PtCl(NH ₃) (quin)(9-EtGH)]	41,9	>200	>4,75		(Sartori et al., 2000)

Tableau 3 : Structure, activité antivirale et cytotoxicité des principales familles d'éjecteurs de zinc de la NC. Cl₅₀ représente la concentration d'inhibiteur nécessaire pour inhiber 50 % de la réplication virale ; CC₅₀ correspond à la dose de cette molécule pour tuer 50 % des cellules ; IT = Index thérapeutique est égale à CC₅₀/Cl₅₀ ; ND = non déterminé. (Mori et al., 2015)

2. Inhibiteurs non covalents

Actuellement, la stratégie privilégiée est le développement d'inhibiteurs non covalents de la NCp7. Contrairement aux éjecteurs de zinc, ces molécules semblent être beaucoup plus spécifiques pour la NCp7 et donc moins toxiques. On peut séparer ces molécules en deux classes : celles qui se lient à la NCp7 et celles liées aux ANs.

a) Ligands de la Nucléocapside

L'objectif est d'élaborer une molécule qui se lie spécifiquement à la NCp7 sans en éjecter le zinc. Le plateau hydrophobe de la NCp7 qui est le domaine de liaison aux ANs est la cible privilégiée pour le développement d'un nouvel inhibiteur. Celui-ci pourra alors bloquer par un mécanisme de compétition l'interaction entre les ANs et la NCp7. Un des moyens utilisés pour identifier ces molécules est le criblage à haut débit de banques de molécules. Pour cela, plusieurs laboratoires ont développé des tests basés sur la capacité de la NCp7 à déstabiliser la structure secondaire des ANs (Shvadchak et al., 2009) ou sur l'interaction entre la NCp7 et les ANs (Breuer et al., 2012). Chacun de ces criblages a permis d'identifier plusieurs molécules dont le squelette principal est similaire. Ces molécules présentent une activité *in vitro* et *in cellulo* intéressante suggérant qu'elles peuvent être utilisées comme point de départ pour une amélioration en chimie médicinale. La description de la structure tridimensionnelle de NCp7 couplée à un de ces inhibiteurs a permis de mieux comprendre leur mécanisme d'action (Goudreau et al., 2013). Dans la structure résolue, deux molécules d'inhibiteurs se fixent au niveau de la protéine permettant de créer une liaison entre la poche hydrophobe et la région N-terminale de la protéine. La structure de ce complexe NCp7/inhibiteurs, associée aux structures connues de la protéine couplée aux ANs (Bourbigot et al., 2008; Guzman et al., 1998) a permis le développement d'un criblage virtuel (*in silico*) permettant de cribler l'interaction potentielle de NCp7 avec plusieurs millions de molécules (Mori et al., 2010, 2011a, 2011b, 2014).

Un autre inhibiteur découvert récemment (A1752) est particulièrement intéressant puisqu'il semble être actif en agissant sur les étapes tardives du cycle viral et en particulier l'assemblage de Gag en se fixant de manière spécifique à la nucléocapside. Il est également capable d'agir sur les processus de décapsidation et de transcription inverse dans les cellules infectées par des virus produits en présence de l'inhibiteur. En absence d'information concernant sa capacité à éjecter le zinc, A1752 reste toutefois un éjecteur de zinc potentiel (Kim et al., 2015).

Actuellement, malgré leur effet antiviral *in vitro* et *in cellulo*, aucun des inhibiteurs de nucléocapside de cette classe n'a réussi à atteindre le stade des essais cliniques. Le développement d'un inhibiteur de NCp7 est l'objectif du projet THINPAD dans lequel j'ai été impliqué (voir partie [VI-F] de ce chapitre) et le chapitre résultats [I-A].

b) Ligands des acides nucléiques partenaires de la Nucléocapside

En parallèle au développement d'inhibiteurs liant la NCp7 pour empêcher sa liaison avec les ANs, certains laboratoires ont travaillé sur la stratégie inverse, à savoir cibler spécifiquement l'ARN partenaire spécifique de la NCp7 pour inhiber leur interaction. Suite à la preuve de concept montrant l'intérêt de cibler les ANs pour empêcher la fixation et l'activité de la NC (Turner et al., 2009), plusieurs équipes ont travaillé sur des molécules pouvant interagir avec la boucle SL1 (Chung et al., 2008, 2010) ou la boucle SL3 (Warui and Baranger, 2009, 2012). Aucune de ces molécules n'a en revanche été testée *in cellulo* pour évaluer un potentiel effet antiviral. Selon le même principe, une autre équipe a travaillé sur le développement d'inhibiteurs ciblant spécifiquement la région TAR (Sosic et al., 2013). *In vitro*, ces derniers semblent d'ailleurs contribuer simultanément à l'inhibition des étapes de la transcription inverse stimulées par la NCp7 et de l'initiation de la transcription médiée par Tat (Sosic et al., 2016). En revanche, aucune activité antivirale n'a été observée *in cellulo*, certainement de par la faible biodisponibilité intracellulaire de ces molécules fortement chargées.

Ces molécules doivent donc subir des modifications structurales pour augmenter leur biodisponibilité et diminuer leur toxicité afin d'être réellement considérées comme une nouvelle classe d'inhibiteurs.

3. Oligonucléotides, aptamères et peptido-mimétiques

Une stratégie alternative est le développement d'oligonucléotides pouvant interagir directement avec la NCp7 et entrer en compétition avec les ANs partenaires naturels de la NCp7. Comme pour les inhibiteurs non-covalents, l'idée est de développer une séquence oligonucléotidique capable de se lier avec le plateau hydrophobe de la NCp7 (Druillennec and Roques, 2000; Druillennec et al., 1999). Sur la base de ce concept, des oligoribonucléotides méthylés mimant la séquence LTR de l'ADN proviral ont été synthétisés et testés *in vitro* et *in cellulo* (Avilov et al., 2012; Grigorov et al., 2011).

Des aptamères d'ARNs d'environ 40 nucléotides capables d'inhiber par compétition l'activité de la NCp7 ont été identifiés par la méthode SELEX (Mely et al., 2008). La NCp7 présente une affinité plus forte pour les aptamères générés que pour la séquence Ψ de l'ARN viral (Berglund et al., 1997). Lorsqu'un de ces aptamères est exprimé dans une cellule infectée, l'empaquetage de l'ARN viral lors de l'assemblage est partiellement inhibé (Lochrie et al., 1997). Le fait que cette diminution ne soit que partielle pourrait s'expliquer par la présence d'un second domaine de liaison aux ANs suffisant pour permettre l'empaquetage de l'ARN lors de l'assemblage (au niveau de la MA, voir [II-CII.1II.1.b])).

En parallèle, le développement de petits peptides ciblant spécifiquement la séquence d'encapsidation Ψ a également été mis en place (Pustowka et al., 2003). Le meilleur des peptides issus de ce groupe (séquence HWWPWW) semble cibler plus spécifiquement la boucle SL3 et inhiber la production de nouvelles particules virales (Dietz et al., 2008).

F. THINPAD : un projet européen pour développer une nouvelle molécule antivirale ciblant la NucléoCapside

Comme il a été décrit précédemment, il existe actuellement six classes de médicaments qui permettent de lutter contre le VIH-1. Celles-ci ciblent principalement les enzymes virales (transcriptase inverse, intégrase et protéase) mais aussi l'entrée du virus, ou encore la fusion entre la particule virale et la cellule cible. Près d'une trentaine de molécules ont été approuvées aux États-Unis et en Europe mais seules quelques-unes sont aujourd'hui couramment utilisées en clinique. La plupart du temps administrées sous forme de combinaisons thérapeutiques, elles peuvent devenir chez certains patients inefficaces de par l'apparition de virus résistants où l'on observe des mutations dans les gènes codant les protéines cibles des traitements. Les dernières avancées ont permis le développement de molécules plus efficaces, moins toxiques et engendrant moins de résistances mais un certain nombre de patients deviennent malgré tout insensibles à ces médicaments entraînant l'évolution de la maladie vers le stade SIDA et la mort (Imaz et al., 2011). Il est donc nécessaire de développer une stratégie innovante permettant de pallier à ce problème de résistance. Une des stratégies est de cibler une protéine très conservée dans l'ensemble des souches virales qui ne pourra donc pas évoluer en réponse à la pression de sélection de ce nouvel inhibiteur. C'est pourquoi la **nucléocapside** est une cible particulièrement intéressante.

Comme il a été précédemment expliqué, cette petite protéine intervient dans plusieurs étapes importantes du cycle viral (transcription inverse, intégration et assemblage) et une mutation ponctuelle de celle-ci diminue très fortement l'infection virale. Elle ouvre en outre la possibilité de cibler plusieurs étapes du cycle viral par l'utilisation d'une seule molécule. C'est dans ce contexte que le projet **THINPAD** (Targeting the **HIV-1 Nucleocapsid Protein** to fight **Antiretroviral Drug Resistance**) a été développé. Les objectifs de ce projet sont donc :

- Identifier et développer un inhibiteur de la nucléocapside (NCI) capable d'inhiber la réplication virale de souches sauvages et résistantes.
- Vérifier l'absence d'émergence de souches résistantes à cette nouvelle classe de molécules.
- Étudier la pharmacocinétique et le profil ADME (Administration, Distribution, Métabolisme et Excrétion du principe actif et de ses métabolites) des molécules développées pendant le projet afin d'apporter le maximum d'informations pour ce nouveau candidat médicament.

1. Les membres du projet

Les partenaires du projet sont issus des milieux académique et industriel, tous choisis pour leurs expertises dans la recherche sur la nucléocapside et dans le développement d'inhibiteurs antiviraux. Il y a ainsi cinq organisations qui participent au projet (**Figure 32**) :

- **Partenaire 1** : Université de Sienne. Ce partenaire possède une plateforme capable de réaliser différents types de modélisation informatique. Le laboratoire est également capable de synthétiser et caractériser de nouvelles molécules organiques, ainsi que de travailler sur des souches de VIH-1 issues de patients (accès à un laboratoire de confinement de type L3 indispensable pour l'étude de ce type d'échantillon). La coordination et la dissémination du projet THINPAD sont réalisées par ce partenaire sous la direction du Pr Maurizio Botta. Son équipe de recherche participe à la première étape du projet, à savoir la sélection de molécules par criblage virtuel (*in silico*) basé sur plusieurs structures tridimensionnelles de la protéine. L'équipe du Dr Maurizio Zazzi est quant à elle impliquée dans l'étude de l'activité antivirale des molécules sur des souches sauvages de virus mais également sur des souches résistantes. Cette équipe a également étudié l'induction de souches résistantes aux NCIs.

- **Partenaire 2** : IDIBAPS (Consorti Institut D'Investigacions Biomèdiques August Pi I Sunyer) est un laboratoire translationnel qui combine des expertises dans différentes spécialités (clinique, virologie, biologie cellulaire, biochimie, ...). Sous la direction des Pr Jose M Gatell et Dr Gilles Mirambeau, ce partenaire est impliqué dans la validation de l'interaction entre les molécules et la nucléocapside. Par l'utilisation d'un test cellulaire, cette équipe teste également les effets antiviraux des NCIs sur une souche virale sauvage (génomme virale complet intégré dans des cellules latentes).
- **Partenaire 3** : Université de Strasbourg, qui au sein de l'UMR biophotonique et pharmacologie regroupe de nombreux spécialistes de la protéine NCp7. Sous la direction du Pr Yves Mely, ce partenaire utilise un test *in vitro* qui permet de vérifier l'effet des NCIs sur l'activité chaperonne de la NCp7. Ce criblage à moyen débit est la pierre angulaire du projet car il est en première ligne pour sélectionner les NCIs potentiels. Ce même partenaire est également impliqué dans un troisième et dernier test *in cellulo* de l'activité antivirale des NCIs (c'est dans cette partie que mon travail de thèse s'est principalement inscrit). Finalement, conjointement aux deux premiers partenaires, le partenaire a étudié le mécanisme d'action des molécules les plus actives à l'aide de différents tests.
- **Partenaire 4** : IRBM Science Park à l'origine de la découverte du premier inhibiteur d'intégrase mis sur le marché (Isentress™). IRBM possède une expérience importante dans l'optimisation de molécules pour leur développement comme candidats médicaments. De plus l'accès à un laboratoire d'expérimentation animale leur permet d'étudier le profil ADME des meilleures molécules. Enfin, ce partenaire a mis à disposition de l'université de Sienne, une banque de molécules exclusive pour le screening *in silico*. Ce partenaire est impliqué dans la synthèse et l'optimisation des meilleures molécules (conjointement au partenaire 1). Finalement ce partenaire industriel étudie les caractéristiques physico-chimiques des touches afin de pouvoir initier l'étude préclinique sur un modèle de souris humanisées.
- **Partenaire 5** : Virostatics. Ce deuxième partenaire industriel est impliqué dans l'évaluation de la toxicité des diverses molécules sélectionnées.

L'expertise, le matériel et les techniques nécessaires au développement d'un nouvel inhibiteur sont donc réunis dans ce consortium, chaque partenaire apportant une pierre à l'édifice.

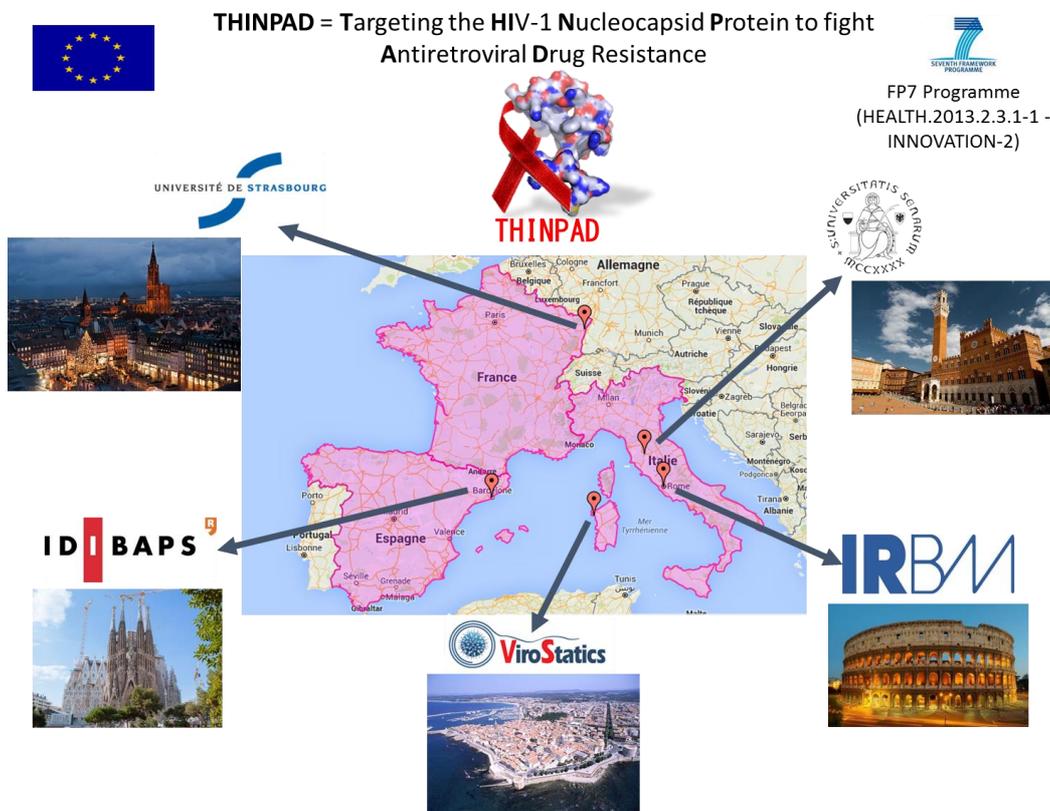


Figure 32 : Localisation et appartenance des différents membres du projet européen THINPAD.

2. Organisation du projet

Afin de mettre en relation les différents partenaires et leurs expertises dans la recherche d'inhibiteurs ciblant la NCp7, un plan détaillé a été mis en place afin d'optimiser les partenariats entre les différents laboratoires. Ainsi, THINPAD s'est organisé autour de 7 blocs de tâches (WP ou Working package, **Figure 33**). Les WP 2 à 5 sont continuellement interconnectés afin d'aboutir le plus rapidement possible à une molécule intéressante.

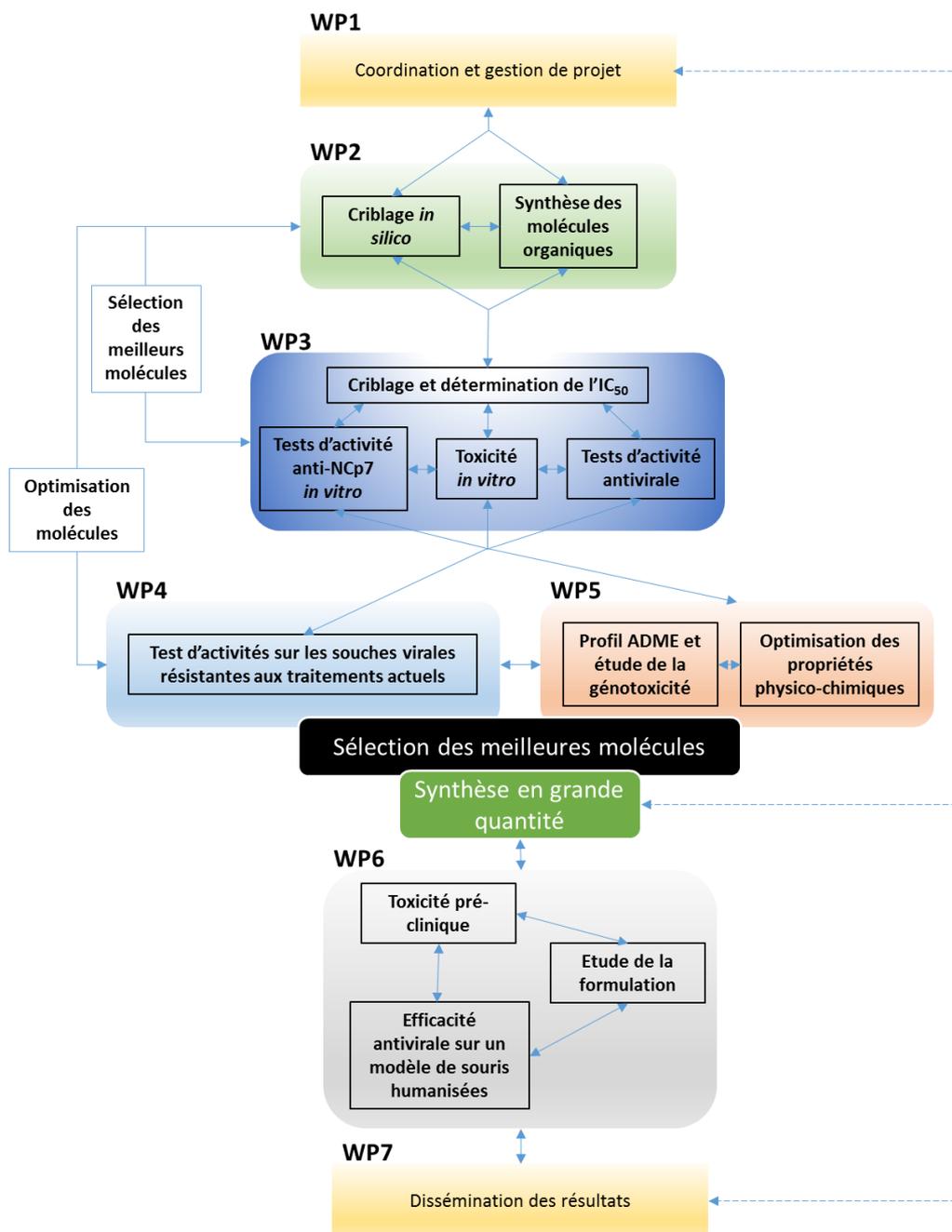


Figure 33 : Diagramme de PERT du projet THINPAD montrant la relation entre les différents blocs de tâches (WP).

a) WP1 : coordination et gestion de projet

Ce WP dirigé par le partenaire 1 mais impliquant l'ensemble des membres du projet permet de réaliser un suivi du projet en rappelant aux groupes concernés les étapes intermédiaires à atteindre pour rester dans la planification initiale ainsi que l'écriture des différents rapports à remettre au contrôleur européen à des dates prédéfinies. Ce WP inclut notamment l'organisation de réunions semestrielles entre les membres, la correspondance avec le conseil scientifique européen et la mise en place d'un site internet (<http://www.thinpad.unisi.it/>).

b) WP2 : sélection *in silico* et optimisation des molécules

C'est ce WP qui a alimenté le projet par la pré-sélection *in silico* de molécules en se basant sur deux banques de données (la première étant une collection privée du partenaire 4 contenant environ 65 000 molécules et la deuxième est une collection de 11 290 000 molécules commerciales). La modélisation informatique se fait grâce à trois structures tridimensionnelles de la NCp7 en interaction avec la boucle SL3 (Guzman et al., 1998), la séquence PBS (Bourbigot et al., 2008) et un inhibiteur non covalent de nucléocapside déjà décrit (Goudreau et al., 2013) (**Figure 34**). Ce criblage informatique réalisé par le partenaire 1 est complété par la synthèse et la mise à disposition pour les partenaires 2 et 3 des molécules par le partenaire 4. Les nombreux échanges entre l'université de Sienne et IRBM, ainsi que les différents résultats du criblage ont permis d'optimiser la structure des molécules afin d'améliorer l'effet ou diminuer la cytotoxicité.

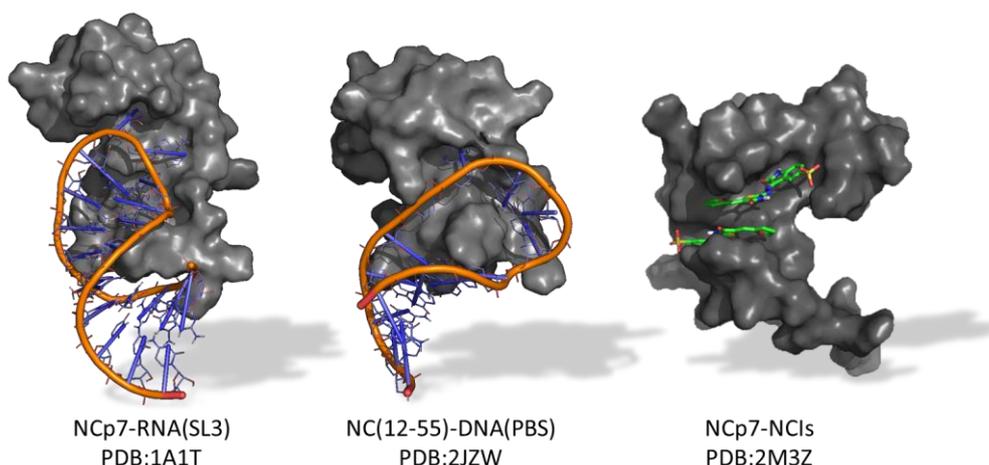


Figure 34 : Structures tridimensionnelles servant au criblage *in silico* de molécules potentiellement inhibitrices.

c) WP3 : criblage *in vitro* et *in cellulo*, caractérisation du mécanisme d'action

Les molécules issues du WP2 sont criblées dans deux tests *in vitro* différents :

(i) Test d'inhibition de l'activité chaperonne des ANs de NCp7 (**NCinh**). Ce test, réalisé dans notre laboratoire (Bernacchi and Mély, 2001; Bernacchi et al., 2002), est basé sur la capacité de la NCp7 à déstabiliser la structure en tige-boucle de la séquence d'ADN cTAR (séquence complémentaire à la séquence TAR de l'ARN viral, impliquée dans le premier transfert de brin lors de la transcription inverse). En plaçant aux extrémités de cTAR un fluorophore (rhodamine 6G, abrégé sous le nom Rh6G) et un inhibiteur de fluorescence

(Dabcyl) capable d'éteindre l'émission du fluorophore lorsqu'ils sont proches, il est possible de suivre la déstabilisation de l'ADN par la NCp7 par mesure de l'intensité de fluorescence (**Figure 35**).

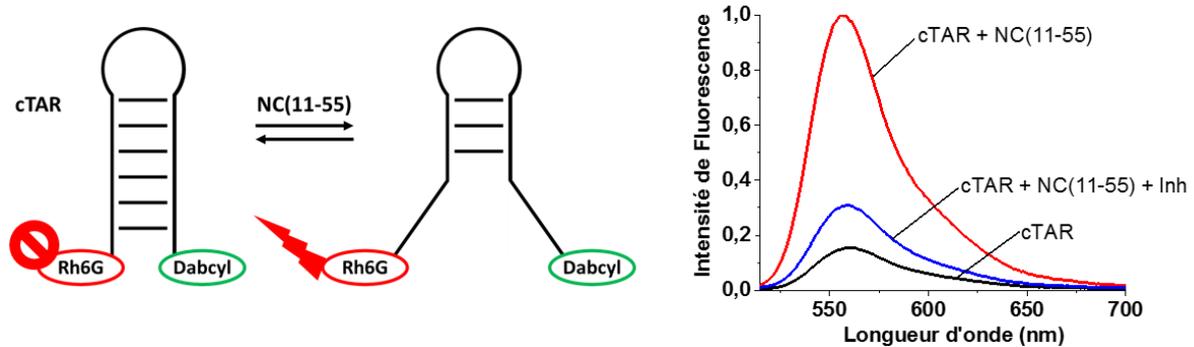


Figure 35 : Test des molécules sur l'activité de déstabilisation de la NCp7 vis-à-vis de cTAR ADN. L'organisation en tige-boucle de l'ADN rapproche la Rh6G et le Dabcyl, diminuant fortement l'intensité de fluorescence de la Rh6G (courbe noire). En présence de NCp7, il y a déstabilisation de la structure de l'ADN, qui éloigne le fluorophore de son inhibiteur, induisant ainsi une augmentation de l'intensité de fluorescence (courbe rouge). Enfin avec un inhibiteur, l'activité chaperonne de la NCp7 est inhibée, ce qui diminue à nouveau la fluorescence de la Rh6G (courbe bleue).

(ii) Test d'inhibition de la liaison de la NCp7 aux ANs (**EMSA**). Ce test, mis au point par le partenaire 2 permet de confirmer l'activité des molécules qui ciblent la NCp7. La présence de NCp7 permet de compacter l'ADN circulaire utilisé dans ce test entraînant une migration sur gel d'agarose différente. En présence d'inhibiteurs, une fraction importante de l'ADN migrera au même niveau que l'ADN libre. L'utilisation de concentrations croissantes d'inhibiteurs permet de déterminer une CI_{50} (**Figure 36**).

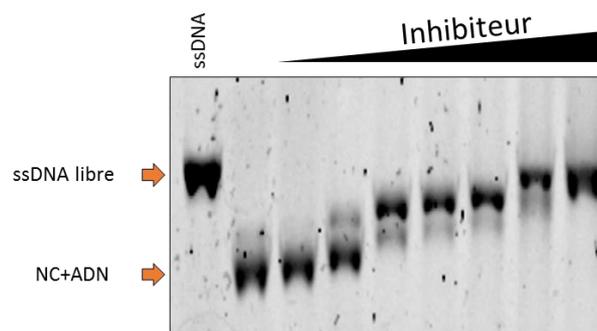


Figure 36 : Test des molécules sur la capacité de NCp7 à se lier aux ANs. NCp7 permet une réorganisation de l'ADN circulaire utilisé dans ce test entraînant une migration différente. La présence d'inhibiteurs permet un retour à la normale prouvant l'inhibition de la liaison de la NCp7 aux ANs.

Quand une molécule est positive dans les deux tests de liaison/déstabilisation, elle est ensuite testée dans trois autres tests *in cellulo* (**Figure 37**) :

- Test d'effet antiviral sur des pseudo-particules (**PSEUDO**) : ce test central pour un de mes projets de thèse sera décrit dans le chapitre résultat [I-A-I.1] de ce manuscrit.
- Test d'effet antiviral sur des cellules infectées de manière chronique et latente (**RevCycle**). Ces cellules possèdent une seule copie du génome viral dont la transcription peut être réactivée pharmacologiquement (Vorinostat, inhibiteur de HDAC utilisé dans la stratégie « shock and kill »). Les virus produits en présence des molécules à tester sont utilisés pour infecter un autre type de cellules possédant le gène de la luciférase intégrée à leur génome et dont l'expression est sous contrôle des LTR viraux (TZM-bl). Il est alors possible de mesurer la transactivation de ces LTRs par les protéines Tat apportées par les particules virales. Dans ce test, on est donc capable de mesurer la capacité des NCIs à inhiber les étapes tardives du cycle viral, mais également dans une moindre mesure les étapes précoces. Le partenaire 2 réalise ce test.
- Test d'effet antiviral (**Mono-** et **BiCycle**). Le partenaire 1 utilise également les cellules TZM-bl afin de quantifier le niveau d'infection par mesure de l'activité luciférase. L'infection en présence de molécules se fait ici avec des virus sauvages soit sur les cellules TZM-bl (MonoCycle), soit pour l'infection d'une autre lignée cellulaire (MT-2). Les particules virales produites en différentes quantités en fonction de l'efficacité de la molécule sont ensuite utilisées pour infecter les cellules TZM-bl (BiCycle).

Quand une molécule est positive dans deux tests cellulaires sur trois, plusieurs tests de toxicité cellulaire sont ensuite effectués sur cette molécule par le partenaire 5.

Finalement, lorsqu'une molécule présente un effet intéressant et une absence de cytotoxicité (ou faible), une batterie de tests est réalisée afin d'élucider son mécanisme d'action. Une partie de ces tests est effectuée dans notre laboratoire et au cours de ma thèse, j'ai pu mettre au point et utiliser plusieurs tests (test d'ajout des molécules à différents temps post-infection, tests d'assemblage).

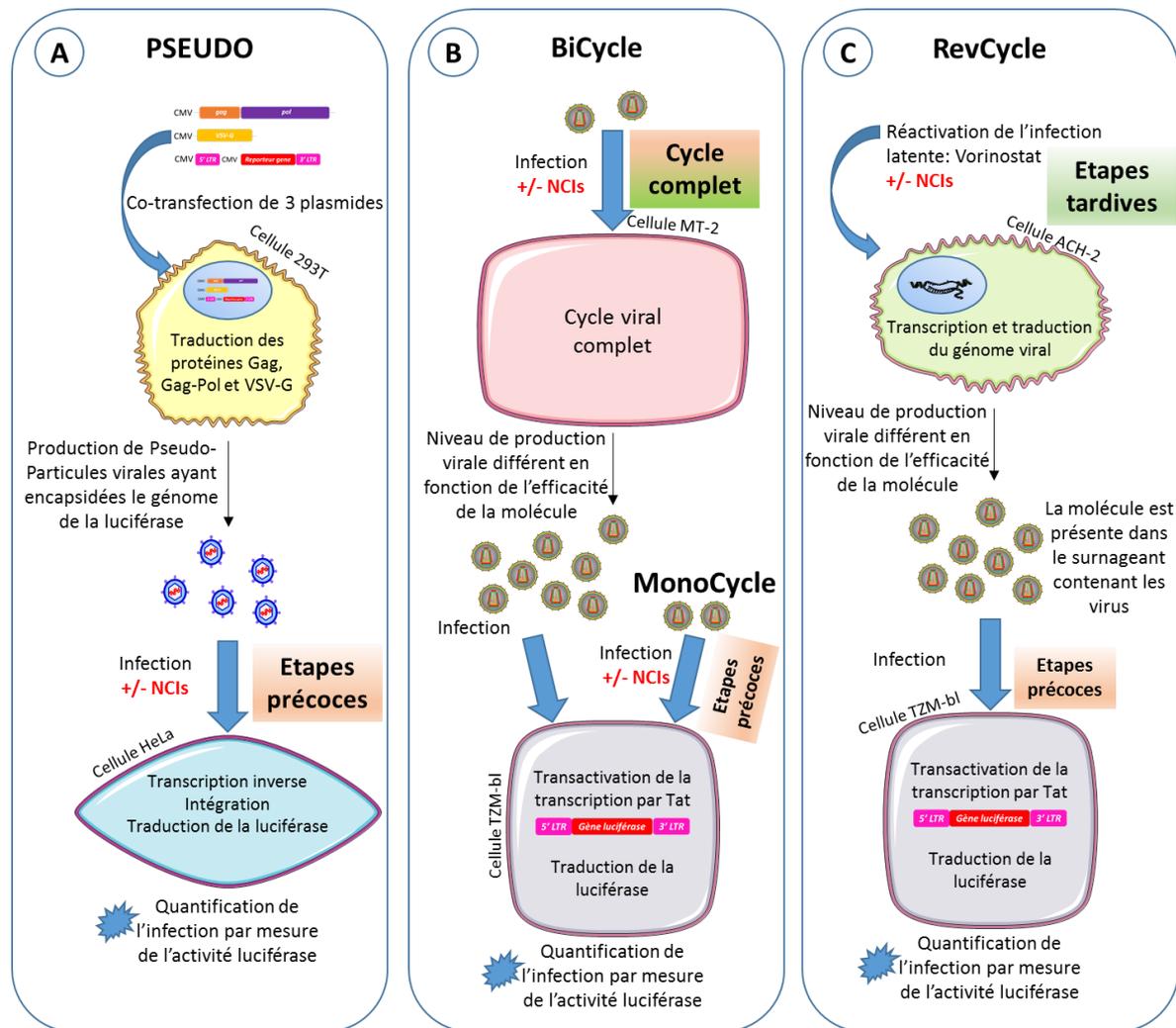


Figure 37 : Comparaison des différents tests cellulaires utilisés pour évaluer l'activité antivirale des inhibiteurs de nucléocapside (NCIs). Le test PSEUDO (A), basé sur la production de pseudo particules par co-transfection de trois plasmides permet l'étude de l'effet des NCIs sur les phases précoces uniquement. Le MonoCycle (B) cible également uniquement les phases précoces alors que l'utilisation du BiCycle (B) permet l'étude de l'effet des molécules sur un cycle viral complet. Finalement le RevCycle (C) utilise des cellules infectées de manière latente au VIH-1 et permet l'étude de l'effet des NCIs principalement sur les phases tardives mais aussi sur les phases précoces.

d) WP4 : souches virales résistantes

Ce WP peut être divisé en deux parties :

- Tout d'abord, un test d'efficacité des molécules sur des souches résistantes aux traitements actuels. Le test utilisé est le même que le BiCycle en utilisant des souches virales résistantes aux NRTIs, aux NNRTIs, aux INIs, ou encore aux PIs (**Tableau 4**).

Classe des molécules inactives sur ces mutants	Mutations par rapport à la souche WT
Inhibiteurs de protéase (PIs)	V32I, I54V, I84V, L90M
	M46I, I54V, V82L, L90M
Inhibiteurs nucléosidique de la transcription inverse (NRTIs)	M41L, D67N, T69D, M184V, L210W, T215Y
	M41L, D67N, M184V, L210W, T215Y
Inhibiteurs non-nucléosidique de la transcription inverse (NNRTIs)	K103N, V179F, Y181C
	L100I, K103N, H221Y
Inhibiteurs d'intégrase (INIs)	G140S, Q148H
	E92Q, N155H

Tableau 4 : Mutations observées sur deux souches résistantes à chacun des différents traitements.

Par la suite, dans ce WP, le partenaire 1 a également essayé de créer une souche virale résistante aux NCIs en soumettant une pression de sélection au virus par ajout de concentrations progressives de la molécule.

e) WP5 : optimisation pharmaceutique

L'étude pharmacocinétique et l'établissement du profil ADME (Absorption, Distribution, Métabolisme et Excretion) sont rapidement pris en compte pour l'optimisation des molécules. Le but de ce WP est d'identifier des NCIs comme candidat médicament ayant des profils ADME et pharmacocinétique optimaux, sans génotoxicité afin de pouvoir rapidement être testés dans un modèle de souris humanisées. Ces études sont réalisées par les partenaires 4 et 5.

f) WP6 : études précliniques

Ce WP, principalement effectué par le partenaire 4 permet l'évaluation de la tolérance des souris humanisées à cette nouvelle molécule dans un premier temps et de l'effet antiviral dans un deuxième temps. Ces souris sont dites humanisées car, ne possédant pas de système immunitaire murin, elles subissent une greffe de cellules progénitrices hématopoïétiques CD34⁺ mimant le système immunitaire humain. Les souris sont ensuite contaminées par différentes souches virales afin de tester l'efficacité des NCIs *in vivo* (Legrand et al., 2009; Traggiai et al., 2004).

g) WP7 : exploitation et dissémination des résultats

Ce WP permet le recensement des différents résultats pour qu'ils soient publiés dans des revues scientifiques et/ou protégés par un dépôt de brevet international afin de permettre une exploitation commerciale.

Chapitre 2 : Objectifs de la thèse

Mon travail de thèse s'est articulé autour de deux projets :

1) Caractérisation cellulaire de nouveaux traitements anti-VIH-1 ciblant la NucléoCapside :

Le SIDA (Syndrome d'Immuno Déficience Acquise), maladie provoquée par le VIH (Virus d'Immunodéficience Humaine) est une des pandémies parmi les plus sévères des dernières années. Les thérapies antirétrovirales, qui permettent d'empêcher l'évolution de la maladie au stade SIDA, sont basées sur l'utilisation de molécules inhibitrices qui ciblent les étapes clés du cycle viral : entrée, transcription inverse (RT), intégration (IN) et maturation. Chacune de ces molécules antivirales induit une pression de sélection à l'origine de l'émergence de virus résistants. Les trithérapies combinent plusieurs molécules ciblant différentes étapes du cycle viral afin d'améliorer l'efficacité, la sélectivité et limiter les résistances aux traitements. Toutefois, ces traitements demeurent imparfaits et il devient donc urgent de développer de nouvelles stratégies thérapeutiques limitant l'apparition de virus résistants. Notre laboratoire participe à un projet européen (THINPAD) ayant pour objectif de développer un traitement anti-VIH-1 ciblant la NucléoCapside (NCp7). Du fait de son activité chaperonne des acides nucléiques (AN), la NCp7 est impliquée dans de nombreuses étapes du cycle viral, notamment lors de la transcription inverse et de l'assemblage de la particule virale (rôle du domaine NC au sein du précurseur Gag). De plus, elle stimulerait la phase d'intégration du génome viral à celui de la cellule hôte. Ces activités et le fait que la séquence de la NC est fortement conservée dans les différents sous-types du VIH-1 font de cette protéine une cible thérapeutique intéressante pour le développement de nouvelles molécules anti-VIH.

L'objectif du consortium européen THINPAD est d'identifier puis d'améliorer des inhibiteurs de la NCp7 en appliquant successivement un crible *in silico*, une sélection *in vitro* des molécules capables d'inhiber l'activité chaperonne des AN de la NCp7 et des cribles *in cellulo* de l'activité antivirale des molécules anti-NCp7. Pour les molécules les plus prometteuses, les mécanismes d'action seront analysés avant de passer à des analyses d'activité et de toxicité *in vivo*. Au sein de ce consortium, l'**objectif** de mon premier projet de thèse a consisté à développer un test d'infection utilisant des pseudo-particules virales mimant la phase précoce de l'infection virale (post-entrée jusqu'à l'intégration). L'effet des molécules ciblant la nucléocapside est mesuré par quantification de l'infection de ces pseudo-

particules. En effet, le génome intégré dans les particules porte un gène rapporteur dont l'expression est sous contrôle du promoteur CMV (Luciférase ou GFP). Les résultats obtenus sont alors analysés en regard des résultats des autres tests réalisés en parallèle par les partenaires du consortium THINPAD afin de sélectionner les molécules potentiellement actives.

Comme la NCp7 est impliquée dans plusieurs étapes du cycle viral, il est important d'identifier celle (ou celles) qui est (ou sont) inhibée(s) par les diverses molécules. C'est pourquoi, en utilisant l'infection par les pseudo-particules, nous avons appliqué un test permettant de déterminer jusqu'à quel temps post-infection (p.i.) une molécule antivirale peut être ajoutée tout en conservant son effet. Ce test permet ainsi de vérifier si les molécules inhibent la transcription inverse ou l'intégration. Afin d'observer l'effet des molécules sur l'assemblage des particules virales, un nouveau test a également été développé. Pour cela, des protéines fluorescentes eGFP ou mCherry ont été insérées dans la polyprotéine Gag dont la multimérisation à la membrane cellulaire est nécessaire et suffisante pour le bourgeonnement de nouvelles particules virales. La multimérisation de Gag et donc l'assemblage peut être quantifiée par des mesures de transfert d'énergie par résonance de type Förster (FRET) entre un fluorophore donneur (eGFP) et un accepteur (mCherry) portés par deux protéines Gag.

2) Observation d'une forme compacte de Gag *in cellulo* :

Les précurseurs Gag recrutent les deux copies d'ARNs viral et se multimérisent au cours de leur transport vers la membrane cellulaire et au niveau de cette dernière. Ces protéines Gag sont nécessaires et suffisantes pour permettre le bourgeonnement de la particule virale. Il a été montré, en solution, par l'utilisation d'une protéine Gag tronquée de son domaine p6 et non myristoylée, que l'interaction de Gag avec l'ARN par l'intermédiaire à la fois du domaine NC mais également du domaine Matrice (MA), entraîne la formation d'une forme repliée de Gag. Cette forme compacte serait une forme transitoire de Gag observée à distance de la membrane cellulaire. *In vitro*, l'addition de lipides, mimant une membrane lipidique, entraîne la perte de la forme compacte en faveur d'une forme étendue de Gag similaire à celle observée dans les virions immatures. En effet, la MA étant plus affine pour les lipides que pour l'ARN, il y aurait perte d'interaction entre le domaine MA et ce dernier alors que le domaine

NC resterait lié à l'ARN. Dans ce contexte, notre **objectif** est de montrer l'existence de la forme repliée de Gag *in cellulo*.

Pour atteindre cet objectif, nous avons utilisé une stratégie basée sur l'utilisation de Split-eGFP. Dans cette stratégie, l'eGFP est clivée en deux parties (GFP1-10 et GFP11). Les deux parties, correspondant aux extrémités N- et C-terminales de l'eGFP, exprimées seules sont non fluorescentes mais sont capables de reconstituer une eGFP entière fluorescente lorsqu'elles se rapprochent. Les parties C- (GFP11, constituée d'un seul des brins beta constituant l'eGFP) et N-terminales (GFP1-10, N-ter formée des brins 1 à 10 du tonneau beta de l'eGFP) ont été insérées entre les domaines Matrice et Capside de Gag et à l'extrémité C-terminale de Gag respectivement au sein d'une protéine Gag complète. Ainsi, lorsque Gag prend sa conformation repliée, il y a reconstitution du fluorophore par rapprochement des deux parties (intramoléculaire) et donc émission de fluorescence. Étant donné le caractère multimérique de Gag une eGFP fluorescente peut également être reconstituée lors de l'interaction de deux protéines Gag portant chacune des fragments complémentaires de l'eGFP (intermoléculaire). Plusieurs techniques d'imagerie, de biologie moléculaire et de biochimie ont ainsi été utilisées pour montrer l'existence de cette forme compacte de Gag.

Chapitre 3 : Matériels et Méthodes

I. Matériels

A. Cellules

- **Cellules HeLa** (ATCC®CCL-2™) : ces cellules épithéliales dérivées d'un cancer du col de l'utérus (infectées par le virus du papillome humain ; prélèvement sur **Henrietta Lacks**) constituent la première lignée cellulaire humaine immortalisée. Elles nous permettent principalement de réaliser des transfections transitoires pour l'observation de protéines fluorescentes en microscopie, ou pour l'extraction protéique nécessaire pour les techniques biochimiques.
- **Cellules 293T** (ATCC®CRL-3216™) : ces cellules épithéliales dérivées de cellules embryonnaires rénales humaines (contiennent Adénovirus) permettent la réplication épisomale de plasmide sous la dépendance d'un promoteur SV40 (*Simian Virus 40*). Dans notre cas, elles sont utilisées pour la production de pseudo-particules virales par co-transfection de trois plasmides bactériens.
- Bactéries compétentes :
 - **DH5α** : cette lignée d'*Escherichia coli* est celle que nous utilisons au laboratoire pour les clonages et l'amplification des différents plasmides. La forte efficacité de transformation de ces bactéries nous permet ainsi d'obtenir des transformations efficaces malgré la faible quantité d'ADN utilisée. Elles sont cultivées à 37 °C dans un milieu stérile de type Lysogeny Broth (LB, 10 g.L⁻¹ tryptone, 5 g.L⁻¹ extrait de levures, 5 g.L⁻¹ NaCl) contenant l'antibiotique dont le gène de résistance est retrouvé sur le plasmide à transformer.
 - **BL21(DE3)pLysS** : ces bactéries (*E. coli*) compétentes permettent la production de protéines. Elles expriment en effet l'ARN polymérase du bactériophage T7 dont la synthèse est sous contrôle du promoteur *lacUV5*, donc induite par l'IPTG (Isopropyl β-D-1-thiogalactopyranoside). Afin de réduire l'expression basale du gène d'intérêt souvent toxique pour les

cellules, ces bactéries produisent le lysozyme T7 (pLysS) qui inhibe l'activité de l'ARN polymérase de T7 (Zhang and Studier, 1997). Ces bactéries sont cultivées dans un milieu LB contenant du chloramphénicol ($35 \mu\text{g.mL}^{-1}$; le plasmide pLysS portant un gène de résistance à cet antibiotique) et l'antibiotique dont le gène de résistance est retrouvé sur le plasmide à transformer.

B. Plasmides

L'ensemble des plasmides utilisés au cours de cette thèse est présenté dans le **Tableau 5**. Afin d'être certain de la séquence nucléotidique du gène codant pour la protéine d'intérêt, l'ensemble des séquences a été vérifié grâce à un séquençage effectué par la société GATC biotech (<https://www.gatc-biotech.com/fr/index.html>).

Nom	Gène de résistance	Tag	Position du Tag ou de l'insertion	Promoteur	Vecteur
gag-eGFP	Ampiciline	eGFP	Entre les domaines MA et CA de Gag	CMV	pcDNA3.1
Gag_mCherry	Ampiciline	mCherry	Entre les domaines MA et CA de Gag	CMV	pcDNA3.1
gag_eGFP_MCS2	Ampiciline	eGFP/MCS2	Insertion d'un site de clonage multiple en C-terminal de Gag	CMV	pcDNA3.1
gag_MCS1_MCS2	Ampiciline	MCS1/MCS2	Insertion d'un site de clonage multiple entre MA et CA de Gag et en C-ter de Gag	CMV	pcDNA3.1
eGFP	Kanamycine			CMV	
Gag (pNL4.3)	Ampiciline	ϕ		CMV	pcDNA3.1
pcDNA3.1	Ampiciline			CMV	pcDNA3.1
pGL3	Ampiciline		Luciférase	SV40	pGL-3
Plasmides nécessaires pour l'observation de la forme compacte par utilisation du FRET-FLIM					
gagTC	Ampiciline	TC	Insertion du motif TC en C-terminale de Gag	CMV	pcDNA3.1
gag_eGFP_TC	Ampiciline	eGFP/TC	Insertion de l'eGFP entre MA et CA de gag et du motif TC en C-terminale de Gag	CMV	pcDNA3.1
Plasmides codant les différents fragments de la split-GFP tri-partite dans la protéine Gag WT					
Zipper_GFP10 (a)	Ampiciline	GFP10	Motif Leucine Zipper associé au brin β #10 de la split-GFP tri-partite	CMV	pcDNA3
Zipper_GFP11 (a)	Ampiciline	GFP11	Motif Leucine Zipper associé au brin β #11 de la split-GFP tri-partite	CMV	pcDNA3
GFP10_Zipper_GFP11 (a)	Ampiciline	GFP10/GFP11	Motif Leucine zipper encadré par les brins β #10 et #11 de la split-GFP tri-partite	CMV	pcDNA3
GFP1-9 (a)	Kanamycine	GFP1-9	Brinβ #1 à 9 constituant le principal fragment de la split-GFP tri-partite	T7	pET
gag_GFP1-9_MCS2	Ampiciline	GFP1-9	MCS1	CMV	pcDNA3.1
gag_MCS1-GFP1-9	Ampiciline	GFP1-9	MCS2	CMV	pcDNA3.1
gag_GFP10_MCS2	Ampiciline	GFP10	MCS1	CMV	pcDNA3.1
gag_GFP10-GFP11	Ampiciline	GFP10/GFP11	MCS1 pour GFP10 et MCS2 pour GFP11	CMV	pcDNA3.1
gag_MCS1-GFP10	Ampiciline	GFP10	MCS1	CMV	pcDNA3.1

Plasmides codant les différents fragments de la split-GFP bi-partite dans la protéine Gag WT					
Zipper_GFP1-10 (a)	Ampiciline		Motif Leucine Zipper associé aux brins β #1 à 10 de la split-GFP bi-partite	CMV	pcDNA3.1
gag_GFP1-10_MCS2	Ampiciline	GFP1-10	MCS1	CMV	pcDNA3.1
gag_GFP1-10_GFP11	Ampiciline	GFP1-10/GFP11	GFP1-10 inséré dans MCS1 et GFP11 dans MCS2	CMV	pcDNA3.1
gag_MCS1_GFP11	Ampiciline	GFP11	MCS2	CMV	pcDNA3.1
gag_MCS1_GFP1-10	Ampiciline	GFP1-10	MCS2	CMV	pcDNA3.1
gag_GFP11_MCS2	Ampiciline	GFP11	MCS1	CMV	pcDNA3.1
gag_GFP11_GFP1-10	Ampiciline	GFP11/GFP1-10	GFP11 inséré dans MCS1 et GFP1-10 dans MCS2	CMV	pcDNA3.1
Insertion des fragments de la split-GFP bi-partite aux extrémités N- et C-terminales de la luciférase					
MCS1_lucif_MCS2	Ampiciline	MCS1/MCS2	Insertion des sites de clonages multiples aux extrémités N- et C-terminales de la luciférase	CMV	pcDNA3.1
GFP11_lucif_MCS2	Ampiciline	GFP11	MCS1	CMV	pcDNA3.1
GFP11_lucif_GFP1-10	Ampiciline	GFP11/GFP1-10	GFP11 inséré dans MCS1 et GFP1-10 dans MCS2	CMV	pcDNA3.1
MCS1_lucif_GFP1-10	Ampiciline	GFP1-10	MCS2	CMV	pcDNA3.1
Plasmides codant les différents fragments de la split-GFP bi-partite dans la protéine Gag délétée du linker entre les deux domaines de la Capside					
gag Δ linkCAC-CAN_GFP11-GFP1-10	Ampiciline	GFP11/GFP1-10	GFP11 inséré dans MCS1 et GFP1-10 dans MCS2	CMV	pcDNA3.1
gag Δ linkCAC-CAN_GFP11-MCS2	Ampiciline	GFP11	MCS1	CMV	pcDNA3.1
gag Δ linkCAC-CAN_MCS1-GFP1-10	Ampiciline	GFP1-10	MCS2	CMV	pcDNA3.1
gag Δ linkCAC-CAN_eGFP_MCS2	Ampiciline	eGFP/MCS2	Insertion d'un site de clonage multiple en C-terminale de Gag	CMV	pcDNA3.1
Plasmides codant les différents fragments de la split-GFP bi-partite dans la protéine Gag où le linker entre les deux domaines de la Capside est remplacé par plusieurs prolines					
Gag_link-PPPP_GFP11-GFP1-10	Ampiciline	GFP11/GFP1-10	MCS1 pour GFP11 et MCS2 pour GFP1-10	CMV	pcDNA3.1
Gag_link-PPPP_GFP11-MCS2	Ampiciline	GFP11	MCS1	CMV	pcDNA3.1
Gag_link-PPPP_MCS1-GFP1-10	Ampiciline	GFP1-10	MCS2	CMV	pcDNA3.1
Gag_link-PPPP_eGFP_MCS2	Ampiciline	eGFP/MCS2	Insertion d'un site de clonage multiple en C-terminal de Gag	CMV	pcDNA3.1

Plasmides codant les différents fragments de la split-GFP bi-partite dans la protéine Gag déletée du domaine Sp1 (entre CA et NC)					
gagΔSp1_GFP11-GFP1-10	Ampiciline	GFP11/GFP1-10	GFP11 inséré dans MCS1 et GFP1-10 dans MCS2	CMV	pcDNA3.1
gagΔSp1_GFP11-MCS2	Ampiciline	GFP11	MCS1	CMV	pcDNA3.1
gagΔSp1_MCS1-GFP1-10	Ampiciline	GFP1-10	MCS2	CMV	pcDNA3.1
gagΔSp1_eGFP-MCS2	Ampiciline	eGFP/MCS2	Insertion d'un site de clonage multiple en C-terminal de Gag	CMV	pcDNA3.1
Plasmides codant les différents fragments de la split-GFP bi-partite dans la protéine Gag dont la multimérisation est limitée par la mutation W318A-M319A					
gagWM_GFP11_GFP1-10	Ampiciline	GFP11/GFP1-10	GFP11 inséré dans MCS1 et GFP1-10 dans MCS2	CMV	pcDNA3.1
gagWM_GFP11-MCS2	Ampiciline	GFP11	MCS1	CMV	pcDNA3.1
gagWM_MCS1-GFP1-10	Ampiciline	GFP1-10	MCS2	CMV	pcDNA3.1
gagWM_eGFP_MCS2	Ampiciline	eGFP/MCS2	Insertion d'un site de clonage multiple en C-terminal de Gag	CMV	pcDNA3.1
Plasmides codant les différents fragments de la split-GFP bi-partite dans la protéine Gag déletée du domaine NC					
gagΔNC_GFP11-GFP1-10	Ampiciline	GFP11/GFP1-10	GFP11 inséré dans MCS1 et GFP1-10 dans MCS2	CMV	pcDNA3.1
gagΔNC_GFP11-MCS2	Ampiciline	GFP11	MCS1	CMV	pcDNA3.1
gagΔNC_MCS1-GFP1-10	Ampiciline	GFP1-10	MCS2	CMV	pcDNA3.1
gagΔNC_eGFP-MCS2	Ampiciline	eGFP/MCS2	Insertion d'un site de clonage multiple en C-terminal de Gag	CMV	pcDNA3.1
Plasmides codant les différents fragments de la split-GFP bi-partite dans la protéine Gag dont la liaison à l'ARN par la MA est limitée par le mutant KKK-25,26,27-TTT					
gagKK-TT_GFP11_MCS2	Ampiciline	GFP11/GFP1-10	GFP11 inséré dans MCS1 et GFP1-10 dans MCS2	CMV	pcDNA3.1
gagKK-TT_MCS1_GFP1-10	Ampiciline	GFP11	MCS1	CMV	pcDNA3.1
gagKK-TT_eGFP_MCS2	Ampiciline	GFP1-10	MCS2	CMV	pcDNA3.1
gagKK-TT_GFP11_GFP1-10	Ampiciline	eGFP/MCS2	Insertion d'un site de clonage multiple en C-terminal de Gag	CMV	pcDNA3.1
Plasmides codant les différents fragments de la split-GFP bi-partite dans la protéine Gag dont la délétion du domaine NC est associé au mutant KK-TT					
gagΔNC_KK-TT_GFP11-GFP1-10	Ampiciline	GFP11/GFP1-10	GFP11 inséré dans MCS1 et GFP1-10 dans MCS2	CMV	pcDNA3.1
gagΔNC_KK-TT_GFP11-MCS2	Ampiciline	GFP11	MCS1	CMV	pcDNA3.1
gagΔNC_KK-TT_MCS1-GFP1-10	Ampiciline	GFP1-10	MCS2	CMV	pcDNA3.1
gagΔNC_KK-TT_eGFP-MCS2	Ampiciline	eGFP/MCS2	Insertion d'un site de clonage multiple en C-terminal de Gag	CMV	pcDNA3.1

Plasmides codant les différents fragments de la split-DRONPA bi-partite dans la protéine Gag (b)					
ff DRONPA	Ampiciline			CMV	pcDNA3.1
Nter-DRONPA-LZ	Ampiciline		Motif Leucine Zipper associé aux brins β #1 à 10 de la split-DRONPA bi-partite	CMV	pcDNA3.1
LZ-c-ter DRONPA	Ampiciline		Motif Leucine Zipper associé au brin β #11 de la split-DRONPA bi-partite	CMV	pcDNA3.1
Plasmides utilisées la purification : insertion de Gag-6H et des différentes insertions de split-GFP bi-partite dans Gag WT					
pET28A_gag (pNL4.3)	Kanamycine			T7	pET28A
pET28A_gagGFP11_GFP1-10	Kanamycine	GFP11/GFP1-10	GFP11 inséré dans MCS1 et GFP1-10 dans MCS2	T7	pET28A
pET28A_gagGFP11_MCS2	Kanamycine	GFP11	GFP11 dans MCS1	T7	pET28A
pET28A_eGFP	Kanamycine	eGFP	Entre les domaines MA et CA de Gag	T7	pET28A
pET28A_gageGFP_MCS2	Kanamycine	eGFP/MCS2	eGFP dans MCS1	T7	pET28A
pET28A_gagMCS1_MCS2	Kanamycine	MCS1/MCS2	Insertion d'un site de clonage multiple entre MA et CA de Gag et en C-ter de Gag	T7	pET28A
Plasmides utilisées la production de pseudo-particules virales (c)					
Ψ cor eGFP	Ampiciline		SIN-LTR-Psi-cPPT-LoxP-CMV-eGFP	CMV	pSicor
Ψ cor Lucif	Ampiciline		SIN-LTR-Psi-cPPT-LoxP-CMV-luciferase	CMV	pSicor
Dr8.91 (gag-pol)	Ampiciline		Gag-Pol-cPPT-RRE (souche pNL4.3)	CMV	pCMV
VSV-G	Ampiciline		VSV-G	CMV	pMD2.G
Plasmides utilisées la localisation spécifique de l'ARN viral (d)					
mCherry_MS2		mCherry	Protéine de Capside du bactériophage MS2	CMV	
pRSV-Rev			Protéine Rev sous contrôle du promoteur RSV	RSV-U3	pRSV-Rev
pCMV-intro			ARN viral modifié, possédant les boucles MS2	CMV	pcDNA3.1

Tableau 5 : liste des plasmides utilisés au cours de cette thèse. (a) ces plasmides ont été généreusement fournis par le Dr Stephanie Cabantous (Cancer Research Center, Toulouse) ; (b) plasmides envoyés par le Dr Peter Dedecker (KU LEUVEN) ; (c) ces plasmides ont été obtenus du laboratoire de Trono (Addgene) ; (d) Plasmides fournis par le Dr Nolwenn Jouvenet (Institut Pasteur, Paris).

C. Anticorps primaires et secondaires

Les différents anticorps primaires et secondaires utilisés dans l'analyse de la forme compacte de Gag *in cellulo* par analyses biochimiques sont présentés dans le **Tableau 6**.

Nom	Espèce	Fournisseur	Référence		Application/dilution
Anticorps primaires					
α-p24 Gag	Souris	NIH reagent program	6521	Monoclonal	WB 1/10000
α-eGFP	Souris	Life technologies	A11120	Monoclonal	IP 1 µg/mg protéine
α-GAPDH	Souris	Millipore	MAB374	Monoclonal	WB 1/5000
Anticorps secondaires					
α-souris_HRP	Chèvre	Promega	W402B	Couplé HRP	WB 1/10000

Tableau 6 : liste des anticorps primaires et secondaires utilisés lors de cette thèse.

D. Oligonucléotides

Les oligonucléotides synthétiques ont été commandés chez Sigma Aldrich et leurs séquences sont présentées dans le **Tableau 7**.

#	Nom	Séquence 5'→3'
Insertion des multi cloning sites entra MA et CA pour MCS1 et en C-terminale de Gag pour MCS2		
1	Fwd MCS1	GATCCGGTGGAGGTGGATCATTAAATTTAAAAGGCGCGCCGGTGGAGG TGGATCAG
2	Rev MCS1	AATTCTGATCCACCTCCACCGGCGCGCCTTTTAATTAATGATCCACCT CCACCG
3	Rev MCS2	Phos-GGTCGTTGCCGAACAGGCTG
4	Fwd MCS2	Phos-CCAGCAGCCAGACCGGTTAAATCGATTCTCGAGTCTAGAG
Insertion des split-eGFP bi- et tri-partites dans MCS1 et/ou MCS2		
5	Fwd GFP11 MCS2 avec STOP	CCGGTGGTGGAGGTGGATCAGGGGCCGGCCTAGAGAAGCGTGACC ACATGGTCCTTCTAGAGTACGTAACAGCTGCTGGGATTACAGACGCA AGTTAAAT
6	Rev GFP11 MCS2	CGATTTAACTTGCCTGTGTAATCCCAGCAGCTGTTACGTACTCTAGAA GGACCATGTGGTCACGCTTCTCTAGGCCGGCCCTGATCCACCTCCA CCA
7	Fwd GFP11 MCS1	TAACGGTGGAGGTGGATCAGGGGCCGGCCTAGAGAAGCGTGACCA CATGGTCCTTCTAGAGTACGTAACAGCTGCTGGGATTACAGACGCAA GTGGG
8	Rev GFP11 MCS1	CGCGCCCACTTGCCTGTGTAATCCCAGCAGCTGTTACGTACTCTAGAA GGACCATGTGGTCACGCTTCTCTAGGCCGGCCCTGATCCACCTCCA CCGTTAAT
9	Fwd GFP1-10 MCS1	CTCTTAATTAACGTTTCGAAAGGCGAGGAG
10	Rev GFP1-10 MCS1	ACACTAATCGATTCAATTTCTCGTTTGGGTCTTTG
11	Fwd GFP1-10 MCS2	CTCACCGGTGGTGGAGGTGGATCAGGGGCCGGCCTAGTTTCGAAAG GCGAGGAG

Manuel PIRES

12	Rev GFP1-10 MCS2 avec STOP	ACACTAATCGATTCAATTTCTCGTTTGGGTCTTTG
13	Fwd GFP10 MCS1	TAACGGTGGAGGTGGATCAGGGGCCGCTAATGGGCGACCTGCC GACGACCACTACCTGTCCACCCAGACCATCCTGAGCAAGGACCTGAA CGGG
14	Rev GFP10 MCS1	CGCGCCCGTTCAGGTCCTTGCTCAGGATGGTCTGGGTGGACAGGTA GTGGTCGTCGGGCAGGTCGCCATTAGGCCGGCCCCTGATCCACCTC CACCGTTAAT
15	Fwd GFP10 MCS2	CCGGTGGTGGAGGTGGATCAGGGGCCGCTAATGGGCGACCTGC CCGACGACCACTACCTGTCCACCCAGACCATCCTGAGCAAGGACCTG AACAT
16	Rev GFP10 MCS2	CGATGTTCAAGTCCTTGCTCAGGATGGTCTGGGTGGACAGGTAGTG GTCGTCGGGCAGGTCGCCATTAGGCCGGCCCCTGATCCACCTCCAC CA
Insertion des fragments de split-DRONPA bi-partite dans MCS1 et MCS2 de Gag ou avec le motif leucine zipper		
17	Fwd Nter-DRONPA_Lzip	CTAGCTAGCTAGACCATGAGTGTGATTAACAGACATGAAG
18	Rev Nter-DRONPA_Lzip	GCCATCGATGGCCTTCTTAGCTTTATAAGTAGTTTTAACGTCACATC
19	Fwd Lzipper_Cter-DRONPA	GCCATCGATGGTTGGCGGTGGCTCTGGAGGTGGTGGGTCCGTTGTC CAGTTGCCAGACTATCAC
20	Rev Lzipper_Cter-DRONPA	GCTCTAGAGCGTTACTTGGCCTGCCTCGG
21	Fwd Luciferase MCS1	CGGGATCCCGCCTTAATTAAGGCGGCGCCATGGAAGACGCCAAA AACATAAAG
22	Rev Luciferase MCS2	GGAATTCCACCGTCTTACCATCGATGGCCACGGCGATCTTTCCG C
23	Fwd N-Ter DRONPA MCS1	CTCTTAATTAACATGAGTGTGATTAACAGACATGAAG
24	Rev N-ter DRONPA MCS1	ATAGGCGCGCCCCTTCTTAGCTTTATAAGTAGTTTTAACGTCACATC
25	Fwd N-Ter DRONPA MCS2	CTCACCGGTGGTGGAGGTGGATCAGGGGCCGCTAATGAGTGTGA TTAAACAGACATGAAG
26	Rev N-ter DRONPA MCS2	ACAACCATCGATGGCTTACTTCTTAGCTTTATAAGTAGTTTTAACGTC ACATC
27	Fwd C-Ter DRONPA MCS1	CTCTTAATTAACGTTGTCCAGTTGCCAGACTATCAC
28	Rev C-ter DRONPA MCS1	ATAGGCGCGCCCCTTGGCCTGCCTCGG
29	Fwd C-Ter DRONPA MCS2	CTCACCGGTGGTGGAGGTGGATCAAGGGGCCGCTAGTTGTCCAG TTGCCAGACTATCAC
30	Rev C-ter DRONPA MCS2	ACAACCATCGATGGCTTACTTGGCCTGCCTCGG
Création des différents mutants de Gag		
31	Fwd gag WM	Phos-GACGTGAAGAACGCCGGCGACCCAGACCCTG
32	Rev gag WM	Phos-CTGGCTGGCCTGCTCAGC
33	Fwd gag Del-link-CAC-CAN	Phos-GACATCCGCCAGGGCCCC
34	Rev gag Del-link-CAC-CAN	Phos-CATCCGCACGATCTTGTTCAAGG
35	Fwd Gag_link-PPPP	Phos-GTACAGCCCCCGCCACCTCCGGACATCCGCC
36	Rev Gag_link-PPPP	Phos-ATCCGCACGATCTTGTTCAAGG
37	Fwd gag Del Sp1	Phos-ATGCAGCGCGCAACTTCC
38	Rev gag Del Sp1	Phos-CAGCACGCGGGCCTTGTG
39	Fwd DeltaNC	Phos-CTGGGCAAGATCTGGCCAG
40	Rev DeltaNC	Phos-GAACATGATGGTCGCCGGTTCG
41	Fwd MA KK26-27TT	Phos-CCCGGCGCACGACGACGTACAAGCTGAAG
42	Rev MA KK26-27TT	Phos-GCGCAGGCGGATCTTCTCCAC

Création des plasmides nécessaires pour la purification de protéines		
45	Fwd gag pet28a (digestion Nco I)	ATCTACCATGGGCGCCCGC
46	Rev gag_..._GFP1-10 pet28a (digestion Xho I)	GTGGTGCTCGAGTTTCTCGTTTGGGTCTTTGCTCAGCAC
47	Rev gag_..._MCS2 pet28a (digestion Xho I)	GTGGTGCTCGAGACCGGTCTGGCTGCTGGGGT
48	Fwd eGFP pet28a (digestion Nco I)	CTACCATGGTGAGCAAGGGCGAGGAGCTG
49	Rev eGFP pet28a (digestion Xho I)	GTGGTGCTCGAGCTTGTACAGCTCGTCCATGCCGAGAGTGAT
50	Age1_GFP1-10 Fwd	CACCGGTGGTGGAGGTGGATCAATGGTTTCGAAAGGGCGAGGAGC
51	GFP1-10-without stop Xho1 Rev	ACCACCCTCGAGTTTCTCGTTTGGGTCTTTGCTCAGC
52	Pac1_eGFP Fwd	CTTAATTAACATGGTGAGCAAGGGCGAGGAG
53	eGFP_without Stop_Asc1 Rev	CGGCGCGCCCTTGTACAGCTCGTCCATGCCG
Oligonucléotide utilisé pour l'étude de la forme compacte de Gag par spectrométrie de masse à mobilisation ionique		
54	polyA(25)	AAAAAAAAAAAAAAAAAAAAAAAAAAAA

Tableau 7 : liste des oligonucléotides nécessaires pour les mutagénèses. Phos- indique une phosphorylation des amorces en 5'.

E. Enzymes de restrictions

Les différentes enzymes de restriction utilisées pour la digestion de plasmides et de séquences d'intérêt amplifiées par PCR sont présentées dans le **Tableau 8**.

Nom	concentration (U/μl)	Nombre de site		site de clivage 5'-3'	Réf. New England BioLabs®
		vecteur lambda (48,5 Kb)	pNEB198 (2,7 Kb)		
Age1-HF	20	13		A/CCGGT	R3552S
Cla1	10	15		AT/CGT	R0197S
Asc1	10	2		GG/CGCGCC	R0558S
Pac1	10		1	TTAAT/TAA	R0547S
BamH1-HF	20	5		G/GATCC	R3136S
EcoR1-HF	20	5		G/AATTC	R3101S
Nhe1-HF	20	1		G/CTAGC	R3131S
Nco1-HF	20	4		C/CATGG	R3193S
Xho1	20	1		C/TCGAG	R0146S

Tableau 8 : enzymes de restriction utilisées pour les différents clonages. HF = High Fidelity®NEB

II. Biologie cellulaire

A. Culture cellulaire

Les cellules HeLa sont cultivées dans des flasques de 75 cm² (T75), dans un milieu de culture DMEM 1 g.L⁻¹ glucose (Dubelcco's Modified Eagle Medium ; Gibco®Life technologie, réf. 21885) dit complet car supplémenté par 10 % (V/V) de Sérum de Veau Fœtal (SVF ; Gibco® Thermo Fisher Scientific, réf. 10270-106) et par des antibiotiques (Pénicilline et Streptomycine 100 UI/mL final). Les cellules 293T sont quant à elles cultivées dans un milieu DMEM complet riche en glucose (4,5 g.L⁻¹ ; Gibco®Life technologie, réf. 31966).

Afin d'être maintenues dans des conditions physiologiques, les cellules sont placées dans un incubateur thermostaté (37 °C) dont l'atmosphère est composée de 5 % de CO₂ et 95 % d'humidité environ. Un passage des cellules bi-hebdomadaires permet de faciliter la prolifération cellulaire en évitant que le tapis cellulaire atteigne la confluence. Cette méthode nous permet ainsi de maintenir les cellules dans le temps (maximum 20 passages pour les HeLa et 10 passages pour les 293T). Pour effectuer le passage, les cellules sont dans premier temps lavées avec du DPBS 1X (Dubelcco's Phosphate Buffered Saline ; Lonza, BE17-512F), puis détachées par action de la trypsine (Lonza, BE02-007E). Un volume de DMEM complet est ensuite ajouté pour stopper l'action de la trypsine et une centrifugation (1300 rpm, 5 min) permet de supprimer l'enzyme encore présente dans le milieu. Le culot cellulaire est alors repris dans 10 mL de DMEM complet et la concentration cellulaire est déterminée à l'aide d'une cellule de Neubauer. Une quantité déterminée de cellules est remise en culture dans une nouvelle flasque contenant 20 mL de DMEM complet. Ce passage nous permet également d'ensemencer d'autres contenants (Ibidi 4 puits, plaque 6 puits, flasques, ...) utilisés pour les différentes expériences.

Les cellules Hela (10⁶ par tube) sont conservées (azote liquide) dans un milieu composé de 10 % (V/V) DMSO, 50 % (V/V) Serum de veau Fœtal (SVF) et 40 % (V/V) de milieu de culture de type DMEM (1 g.L⁻¹ glucose, Dulbecco's Modified Eagle Medium).

Les cellules 293T sont conservées dans un milieu contenant 10 % (V/V) DMSO et 90 % (V/V) SVF dans de l'azote liquide. Elles ne sont utilisées qu'à des passages inférieurs à dix.

B. Transfection cellulaire

La transfection cellulaire est un processus de transfert de gènes qui permet l'introduction d'ADN exogène (ici plasmide bactérien) dans des cellules eucaryotes. Le matériel génétique introduit dans la cellule ne s'intègre pas dans le génome de la cellule hôte (on parle de transfection transitoire). Plusieurs techniques de transfection existent mais nous avons utilisé exclusivement le kit commercial JetPEI (Polyplus transfection™). Ce kit utilise le PEI (Polyéthylène Imine) comme polymère cationique qui se fixe aux phosphates négatifs de l'ADN pour l'englober entièrement. Les complexes formés, globalement positifs vont se fixer aux polysaccharides de la membrane plasmique chargés négativement. Ils sont ensuite internalisés par endocytose et l'ADN libéré peut alors être transporté dans le noyau. Le PEI, en inactivant les hydrolases acides en tamponnant le pH cytoplasmique protège l'ADN d'une dégradation (Polyplus-transfection, Illkirch-Graffenstaden, France).

Vingt-quatre heures avant la transfection, les cellules sont déposées :

- Pour l'imagerie dans une Ibidi 4 puits, $4 \cdot 10^4$ cellules par puits dans un volume de 800 μ L
- Pour un western blot dans une plaque 6 puits, $2 \cdot 10^5$ cellules par puits dans un volume de 3 mL
- Pour une migration en conditions natives ou un cross-link dans une boîte de pétri à 10^6 cellules dans un volume final de 10 mL

Pour chaque transfection, deux tubes sont préparés : (i) l'un contenant l'ADN à transférer dans une solution de 150 mM NaCl stérile et (ii) l'autre contenant le PEI (stock à 7,5 mM, 2 μ l de PEI stock pour 1 μ g d'ADN) également dans une solution de 150 mM NaCl stérile. Pour former des complexes ADN-PEI, le tube contenant le PEI est ensuite ajouté à celui contenant l'ADN, puis vortexé pendant 15 s et laissé à température ambiante pendant 30 min. Le mélange est ensuite déposé sur les cellules dont le milieu de culture a été fraîchement remplacé et celles-ci sont replacées dans l'incubateur pendant 12 ou 24 h.

C. Fixation cellulaire

La fixation permet de figer les membranes et les constituants cellulaires afin de conserver les structures cellulaires dans un état proche de celui du vivant. Dans notre cas, l'utilisation du ParaFormaldéhyde (PFA) permet la formation de ponts méthylènes entre les acides aminés

Manuel PIRES

des protéines cellulaires. Ce procédé nous permet de conserver les cellules pendant plusieurs jours avant d'être observées au microscope confocal (ni thermostaté, ni maintenu avec 5 % de CO₂). Après plusieurs lavages par du DPBS 1X pour éliminer le maximum de complexes PEI-ADN et de débris cellulaires, les cellules sont incubées pendant 15 min à température ambiante et sous agitation douce avec une solution de PFA 4 % (stock à 16 % dilué dans du DPBS 1X ; Electron Microscopy Science, réf. 15710). Après une autre série de lavages, les échantillons peuvent être conservés dans du DPBS à 4 °C.

D. Marquages post-fixation

1. Marquage nucléaire par du Hoescht 33342

Après fixation et pour pouvoir observer les noyaux cellulaires, un traitement par du Hoechst (ThermoFisher-33342) sans perméabilisation est réalisé. Ce marqueur émet une fluorescence bleue (excitation à 350 nm et émission à 461 nm) lors de sa liaison à l'ADN double brin. Après les lavages post-fixation, les cellules sont incubées pendant 20 min à température ambiante sous agitation avec une solution d'H-33342 à 1 µg.mL⁻¹. Afin d'éliminer l'excédent de colorant, trois lavages avec du DPBS 1X sont ensuite réalisés.

2. Marquage du motif tétracystéine par un dérivé de la résorufine (ReAsH)

Cette technique est basée sur la forte affinité de l'arsenic (As) pour les groupements thiols (SH) retrouvés notamment sur les acides aminés cystéines. L'insertion par mutagénèse d'un petit motif constitué de 6 résidus (dont quatre cystéines) est suffisante pour permettre le marquage et le suivi d'une protéine d'intérêt. En effet, la reconnaissance du motif et la formation de liaisons covalentes permettent de rendre fluorescents les dérivés de fluorescéines tels que le FIAsh ou résorufine comme le ReAsH (**Figure 57B**). Ce système permet de limiter la modification structurelle de la protéine marquée puisque le motif TC est de petite taille (1 kDa). Après trois lavages par de l'Opti-MEM (Optimisation du milieu Eagle's Minimum Essentials Media) sans sérum, les cellules sont incubées pendant 1 h à 37 °C avec une solution d'Opti-MEM contenant 0,8 µM de ReAsH-EDT2 (EDT2 = 1,2-ethanedithiol qui permet de stabiliser et solubiliser le ReAsH ; Thermo Fisher Scientific, T34562). Les cellules sont ensuite lavées plusieurs fois par de l'Opti-MEM et incubées 10 min avec le réactif BAL (2,3-

Manuel PIRES

dimercaptopropanol, ou British Anti-Lewisite) afin d'éliminer le maximum de liaison non spécifique, avant un autre lavage par de l'Opti-MEM. Il est alors possible d'observer directement les cellules si le microscope est équipé d'une enceinte thermostatée ou la concentration en CO₂ est maintenue à 5 %. Sinon, les cellules peuvent être fixées à l'aide de PFA par exemple.

E. Préparation des échantillons pour la cytométrie en flux

Afin d'observer les caractéristiques individuelles de plusieurs milliers de cellules et notamment la fluorescence (ici pour mesurer la capacité d'infection des pseudo-particules virales), il est possible de les analyser par cytométrie en flux (FACS). Pour se faire, les cellules sont détachées de leurs supports par trypsination comme lors d'un passage cellulaire. Le culot cellulaire est alors lavé une fois dans du DPBS 1X (centrifugation à 1300 rpm pendant 5 min pour éliminer le surnageant), avant d'être fixé dans une solution de PFA 2 % (dans du DPBS 1X) pendant 15 min. Les cellules sont lavées 2 fois dans du DPBS 1X et finalement resuspendues à une concentration de 10⁶ cellules/mL dans une solution de DPBS 1X contenant 0,5 % (V/V) de BSA (Bovine Serum Albumine ; euromedex, réf. 04-100-811-C) et 2 mM EDTA (Ethylene Diamine Tetraacetic Acid ; Euromedex, Réf. EU0007). Les échantillons sont ensuite analysés sur un BDFACS Aria II à l'IGBMC. Pour chaque cellule, une évaluation de la granularité et de la taille des cellules permettent d'éliminer notamment les débris cellulaires. Finalement, une excitation par un laser à 488 nm couplée à l'utilisation d'un filtre « long pass » 505 nm (conserve uniquement la fluorescence supérieure à 505 nm) et d'un filtre « band pass » 525/50 (conserve uniquement la fluorescence entre 475 et 575 nm) permet la détection des cellules exprimant l'eGFP.

III. Infection par des vecteurs lentiviraux

A. Production des vecteurs lentiviraux

Les vecteurs lentiviraux, aussi appelés pseudo-particules (PPs) sont produits par transfection transitoire de trois plasmides (d'emballage, d'enveloppe et de transfert) dans des cellules 293T (**Figure 43**). Un total de $4 \cdot 10^6$ cellules 293T estensemencé dans des boîtes de pétri de 10 cm de diamètre à J0. La co-transfection est réalisée 24 h après l'ensemencement (J1) et sans changement de milieu par (i) 3 μg de plasmide d'enveloppe (VSV-G), (ii) 6 μg de plasmide d'emballage (dR8.91(gag-pol)) et (iii) 12 μg de plasmide de transfert ($\Psi\text{cor-lucif}$ ou $\Psi\text{cor-eGFP}$). Le milieu de culture est changé 24 h après la transfection (J2) afin d'éliminer les complexes de transfection non internalisés et pour booster l'activité cellulaire. Finalement, 48 h après transfection (J3), le surnageant cellulaire contenant les PPs est récolté et filtré pour supprimer les débris cellulaires sans fixer les PPs par passage sur un filtre d'acétate de sodium 0,45 μM (Millipore réf. SLHV033RS), puis concentré deux fois par centrifugation sur des colonnes Vivaspin qui permettent d'éliminer les éléments d'un poids moléculaire inférieur à 50 kDa (Sartorius Intec, réf. VS2031). Finalement, les PPs sont conservées à -80°C par aliquot de 500 μL et la concentration est évaluée grâce à la quantification du contenu en antigène p24 (correspond à la CA de Gag) par un test ELISA selon le protocole du fabricant (Innogenetics, Innostest[®] HIV Antigen mAb, réf. 80563).

B. Dosage de l'infection

1. Par mesure de l'activité luciférase

Cette méthode permet d'évaluer l'infectivité des PPs-luciférase sur les cellules HeLa. Un total de $3 \cdot 10^5$ cellules HeLa/puits dans du milieu DMEM complet est déposé 24 h avant l'infection en présence de 8 $\mu\text{g/mL}$ de polybrène par différentes concentrations de PPs-luciférase (0, 3, 10, 30, 100 et 300 ng équivalent p24). Après 24 h d'infection, les cellules subissent un lavage au DPBS 1X puis sont lysées directement dans le puit par 500 μL de tampon de lyse (cell culture lysis buffer, Promega réf. E153A) supplémenté de 0,5 % (V/V) de Triton-X100 pendant 30 min à température ambiante et sous agitation. L'activité luciférase est alors mesurée en utilisant un luminomètre (Berthold, Tristar LB941). Pour cela, 20 μL d'échantillon sont placés dans une

Manuel PIRES

plaque 96 puits blanche et 50 μ l du substrat de la luciférase (25 mM Tricine Buffer pH = 7,8 ; 0,5 mM EDTA pH = 7,9 ; 5 mM MgSO₄ ; 5 mM DTT ; 0,5 mM ATP ; 1,65 mM D-luciferin sel de sodium et 0,325 mM Coenzyme A hydrate de sel de sodium) sont injectés et la luminescence est mesurée en intégrant le signal pendant 10 s. L'activité luciférase mesurée est proportionnelle à l'expression de la luciférase et donc à l'infectivité des PPs.

2. Par Cytométrie en flux

Cette méthode est utilisée pour déterminer la capacité d'infection des PPs virales exprimant l'eGFP. En effet, la quantité de cellules infectées exprimant l'eGFP est proportionnelle à l'infectivité et à la quantité de PPs (en ng d'équivalent p24). Des cellules HeLa sontensemencées dans des plaques 6 puits à 3.10⁵ cellules/puits 24h dans du DMEM complet avant l'infection par des concentrations croissantes de PPs (0, 3, 10, 30, 100 et 300 ng équivalent p24) en présence de 8 μ g/mL de polybrène. 48h après infection, les cellules sont récoltées, fixées et analysées selon le protocole de cytométrie en flux pour déterminer le pourcentage d'infection.

C. Test d'infectivité et de cytotoxicité

1. Test d'infectivité

Afin de tester un nombre important de molécules, le test d'infection a été adapté dans des plaques 96 puits. Pour cela, 5.10³ cellules HeLa (100 μ l) sontensemencées par puits dans une plaque 96 puits blanche (Greiner Bio-One, réf. 655 075). Le jour de l'infection (24 h après dépôt des cellules), le milieu est remplacé par 50 μ l d'une solution de DMEM contenant la molécule d'intérêt à la concentration 2X et par 50 μ l de DMEM contenant des PPs (1,7 ng μ l⁻¹ p24) en présence de polybrène (8 μ g/mL). La concentration des PPs a été déterminée pour obtenir 50 % d'infection (voir résultats). Afin de lisser la variabilité observée dans les puits, chaque concentration est testée en sextuplicate. Chaque plaque contient une série de contrôles positifs où l'inhibition est maximum (AZT, AZidoThymidine 0,5 μ M, inhibiteur de la transcription inverse) et négatifs (DMEM contenant un % de DMSO identique à celui utilisé pour les molécules testées). Les cellules sont lysées directement dans le puit 24 h après infection et traitement par 30 μ l de tampon de lyse (cell culture lysis buffer, Promega réf. E153A) supplémenté par 0,5 % (V/V) de Triton-X100 pendant 30 min à température ambiante

Manuel PIRES

et sous agitation. L'activité luciférase est alors mesurée par un luminomètre (Berthold, Tristar LB941) par injection de 50 µl de substrat de la luciférase (25 mM Tricine Buffer pH = 7,8 ; 0,5 mM EDTA pH = 7,9 ; 5 mM MgSO₄ ; 5 mM DTT ; 0,5 mM ATP ; 1,65 mM D-luciferin sel de sodium et 0,325 mM Coenzyme A hydrate de sel de sodium) et la luminescence est accumulée pendant 10 s (valeur molécule). Un contrôle est également réalisé en absence de molécule mais à pourcentage de DMSO équivalent (valeur DMSO)

Il est alors possible de calculer pour chaque concentration le pourcentage d'inhibition de l'infection (%inh) d'après l'équation suivante :

$$\%inh[molécule] = \frac{(valeur\ DMSO) - (valeur\ molécule)}{(valeur\ DMSO)} \times 100$$

Équation 1 : Calcul du pourcentage d'inhibition de l'infection et de cytotoxicité.

Le pourcentage d'inhibition de l'infection est rapporté en fonction de la concentration de la molécule (exemple sur la **Figure 47**). La courbe obtenue peut être affinée par l'équation suivante :

$$y = \frac{A_1 + (A_2 - A_1)}{1 + 10^{(\log(CI_{50}) - \log(CI)) \times p}}$$

Équation 2 : Détermination de la concentration inhibitrice à 50 % (CI₅₀).

A1 et A2 représentent respectivement le pourcentage d'inhibition en absence de molécule (DMSO seul) et en présence d'une concentration saturante d'inhibiteur ; CI représente la concentration en inhibiteur ; CI₅₀ représente la concentration inhibitrice à 50 % de l'effet maximal ; et p correspond au coefficient de Hill.

La concentration inhibitrice à 50 % de l'effet maximal (CI₅₀) servira de valeur de référence permettant de comparer les molécules entre elles.

2. Test de cytotoxicité

En parallèle de la plaque permettant de déterminer la capacité des molécules à inhiber l'infection par des PPs, une seconde plaque 96 puits est établie afin de tester la cytotoxicité des molécules. Ce test permet de vérifier que l'inhibition de l'infection n'est pas la conséquence de la cytotoxicité de la molécule. Pour cela, une plaque 96 puits transparente est préparée dans les mêmes conditions que la plaque du test d'infectivité. 24 h après l'infection, le milieu est remplacé par 110 µL d'un mélange composé de 100 µL de DMEM

Manuel PIRES

complet et de 10 μL d'une solution à 12 mM de MTT (bromure de 3-(4,5-diméthylthiazol-2-yl)-2,5-diphényl tétrazolium). Après une incubation de 4 h à 37 °C, 80 μL du milieu est remplacé par 50 μL de DMSO permettant de dissoudre les cristaux de formazan. En effet, le test MTT permet d'évaluer la viabilité cellulaire car le MTT est réduit par la succinate deshydrogénase mitochondriale des cellules vivantes actives en formazan (*Figure 38*). L'absorbance, mesurée à 540 nm, est révélatrice de la viabilité cellulaire. Il est possible, en utilisant l'équation 1 de calculer le [% de cytotoxicité] et le [% de toxicité] (concentration nécessaire pour tuer 50 % des cellules (CC_{50})).

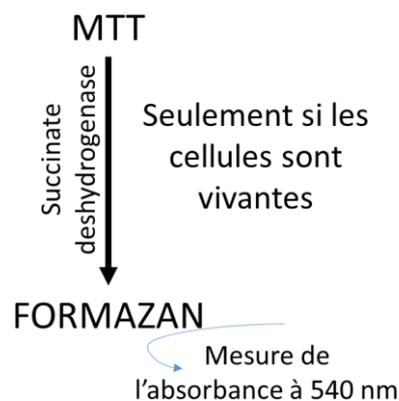


Figure 38 : principe du test MTT (bromure de 3-(4,5-diméthylthiazol-2-yl)-2,5-diphényl tétrazolium)

D. Test d'ajout des molécules à divers temps post-infection

Ce test permet de déterminer sur quelle étape précoce du cycle viral la molécule effectue son action antivirale. Pour cela, deux plaques blanches 96 puits sontensemencées avec $5 \cdot 10^3$ cellules par puits (100 μL). Vingt-quatre heures plus tard, les surnageants sont remplacés dans la première colonne (t_0) par 100 μL d'une solution de DMEM contenant soit la molécule d'intérêt à une concentration correspondant à vingt fois la CI_{50} , soit de l'AZT à 1 μM (inhibiteur de transcription inverse), soit du Raltegravir à 2 μM (inhibiteur d'intégration) ou du DMSO (*Figure 39*). L'ensemble des molécules y compris l'AZT et le Raltegravir ont été dissoutes dans du DMSO dont la quantité est normalisée pour toutes les molécules testées à 0,4 % final maximum. Les PP (100 μL de DMEM complet contenant 1,7 ng d'éq. p24 en présence de polybrène 8 $\mu\text{g}/\text{mL}$) sont ensuite ajoutées dans tous les puits de la colonne t_0 ou introduits en remplacement des surnageants, dans les autres puits de la plaque (t_1 à t_{12}). Puis toutes les heures post-infection, 100 μL de la dilution de la molécule d'intérêt ou des contrôles (DMSO, AZT et Raltegravir) sont ajoutées dans une colonne différente. Finalement, 24 h après

Manuel PIREs

l'infection, l'activité luciférase est mesurée de la même manière que pour le test d'infectivité (aucune molécule n'est ajoutée dans la colonne t24, correspondant alors au maximum d'infection). Chaque condition est testée en quadruplicate. Chaque molécule est testée à une concentration finale correspondant à dix fois la CI_{50} déterminée au préalable dans le test d'infectivité. Pour chaque molécule et chaque temps d'addition, le pourcentage d'inhibition de l'infection est déterminé grâce à l'*équation 2*.

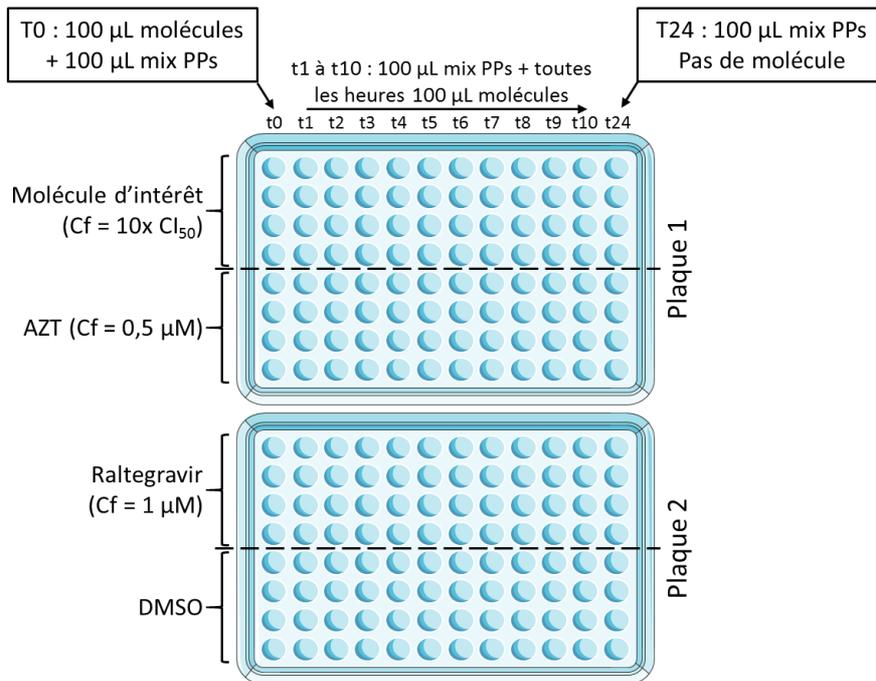


Figure 39 : Plan de plaque pour le test d'ajout des molécules (TOAA). Cf = Concentration finale

IV. Imagerie cellulaire

A. Microscopie confocale

La plupart des images utilisées pendant cette thèse a été réalisée avec ce type de microscope (**Figure 40 A**), qui permet la visualisation de protéines fluorescentes dans un plan focal. Si les fluorophores excités par un faisceau laser, émettent sur toute l'épaisseur de l'échantillon, la présence d'un diaphragme à diamètre variable (pinhole) permet de sélectionner la fluorescence émise dans un seul plan focal en éliminant la fluorescence provenant des régions en dehors de ce plan. Les photons émis sont détectés par un photomultiplicateur (PMT) qui permet l'observation de la fluorescence d'un plan focal par recombinaison d'une image où chaque pixel possède une valeur d'intensité qui lui est propre.

Le microscope confocal Leica TC SPE-II de la Plateforme d'Imagerie Quantitative (PIQ) excite l'échantillon grâce à un balayage laser, avec un objectif à immersion à huile G63X (HCX PL APO CS). Des longueurs d'onde de 405 nm (DAPI), 488 nm (eGFP) et 561 nm (mCherry) ont été utilisées pour exciter les différents échantillons. La fluorescence émise a été recueillie avec des filtres d'émission de bandes passantes 500-555 nm (eGFP) et 570-625 nm (mCherry).

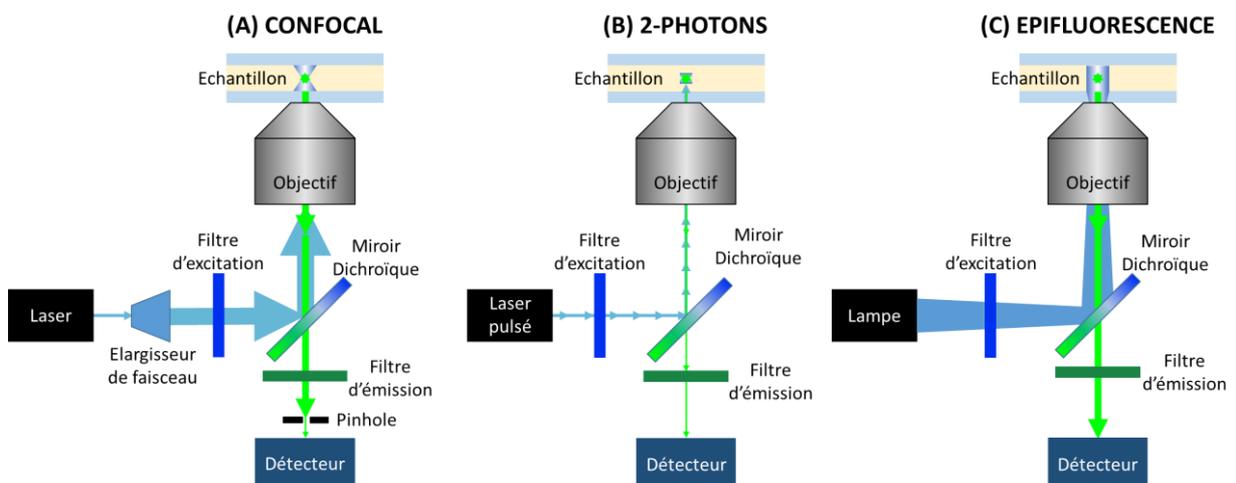


Figure 40 : Comparaison des différents systèmes de microscopie utilisés dans cette thèse.

B. Transfert d'énergie entre molécules fluorescentes (FRET : Florescence Resonance Energy Transfer) par mesure du temps de vie de fluorescence (FLIM : Fluorescence Lifetime Imaging Microscopy)

Un principe utilisé à plusieurs reprises dans cette thèse est le FRET, qui consiste en la mesure du transfert non radiatif de l'énergie d'un fluorophore à l'état excité (donneur) à un second fluorophore (accepteur). Pour que ce transfert d'énergie ait lieu, il est indispensable que la distance entre les fluorophores n'excède pas 10 nm. Cette très courte distance permet de conclure que les protéines qui portent les fluorophores sont en interaction. Le spectre d'émission du fluorophore donneur et le spectre d'absorption de l'accepteur doivent se chevaucher pour pouvoir donner un transfert d'énergie. Plusieurs techniques d'observation du FRET existent mais au laboratoire, il est possible de le quantifier par mesure du temps de vie de la fluorescence émise par le donneur (FLIM). Ainsi, s'il y a transfert d'énergie, il y a diminution du temps de vie du donneur. À l'inverse, en l'absence de FRET, le temps de vie du donneur, en présence de l'accepteur, reste identique à celui du fluorophore donneur seul (2,4 ns en moyenne pour l'eGFP). L'avantage majeur est que la mesure de transfert d'énergie par temps de vie est directe et ne requiert pas de corrections complexes. Le microscope est couplé à un système d'excitation bi-photons qui permet d'exciter uniquement les fluorophores situés au niveau du plan focal (**Figure 40 B**). Contrairement au système confocal, où toute l'épaisseur de l'échantillon est excitée mais seule la fluorescence du plan focal est récupérée, l'excitation bi-photonique permet de n'exciter que les fluorophores du plan focal et donc de récupérer spécifiquement l'émission de fluorescence du plan focal. Dans le montage à deux photons, un microscope Olympus IX70 est utilisé avec un objectif 60X à eau (UPLSAPO 60XW) et la fluorescence de l'échantillon est détectée par une photodiode à avalanche permettant une détection de signaux extrêmement faibles. Dans notre cas, la technique de FRET-FLIM a été utilisée pour :

- tester l'effet de molécules antivirales sur l'assemblage en plaçant les fluorophores donneurs et accepteurs entre les domaines MA et CA de deux protéines Gag. La multimérisation de Gag peut alors être quantifiée par FRET-FLIM.

Manuel PIRES

- vérifier l'existence d'une forme compacte de Gag en utilisant le transfert d'énergie entre l'eGFP (insérée entre les domaines MA et CA de Gag) et le ReAsH (marquage du motif TC placé en C-terminal de Gag).

L'analyse permet ensuite de déterminer le pourcentage de FRET entre les fluorophores (dans notre cas le donneur est toujours l'eGFP et l'accepteur est soit mCherry, soit le système TC/ReAsH). Une analyse à deux composantes (**Équation 3**) permet de définir plus précisément la population de donneur seul (non engagée dans un FRET) et la population de donneur en interaction avec un accepteur (qui donne un FRET) et le pourcentage de FRET de cette population.

$$F = \alpha_1 e^{-t/\tau_1} + \alpha_2 e^{-t/\tau_2}$$

Équation 3 : Analyse à deux composantes du temps de vie de Gag-eGFP en présence de Gag-mCherry.

F représente la décroissance de fluorescence, τ_2 représente le temps de vie long (ici $\approx 2,4$ ns correspondant au temps de vie de l'eGFP dans la protéine Gag-eGFP en absence d'accepteur ou monomérique), α_2 correspond à la population de Gag-eGFP sous forme monomérique, τ_1 est le temps de vie moyen de la population de Gag qui transfère de l'énergie à Gag-mCherry et α_1 montre la population de Gag multimérique.

C. Micro-injection et suivi de la fluorescence par vidéo-microscopie

La technique de micro-injection permet de micro-injecter de faibles quantités de substances liquides dans des compartiments microscopiques. Nous avons injecté de l'ADN plasmidique directement dans le noyau cellulaire. Cet ADN est alors immédiatement transcrit sous forme d'ARN et la fluorescence de la protéine d'intérêt apparaît en quelques minutes. Nous pouvons suivre l'apparition de la fluorescence et sa localisation dans le temps grâce à la vidéo-microscopie.

Techniquement, les cellules sontensemencées 24 h avant l'expérience dans des plaques 6 puits contenant une lamelle (préalablement stérilisée par de l'éthanol 70 % et lavée trois fois avec du DPBS 1X) à raison de $2 \cdot 10^5$ cellules par puits dans 2 mL de DMEM complet. Le jour de l'expérience, les plasmides sont dilués dans de l'eau milliQ à une concentration de $100 \text{ ng} \cdot \mu\text{L}^{-1}$ puis centrifugés 20 min à 13000 rpm afin d'éliminer les gros éléments contenus dans la solution. L'ADN contenu dans le surnageant est re-suspendu à une concentration de $0,1 \text{ } \mu\text{g}/\mu\text{L}$

Manuel PIRES

dans une solution de rhodamine-dextran (0,5 $\mu\text{g}/\mu\text{L}$) qui permet de vérifier que les cellules ont bien été micro-injectées. Les lamelles sont ensuite montées dans une chambre de Ludin (Life Imaging Services, Bâle, Suisse) contenant du milieu DMEM sans rouge phénol puis placées sur le microscope Leica DMIRE2 (microscope à épifluorescence, **Figure 40 C**) où la température et le CO_2 sont maintenus dans des conditions physiologiques (37 °C et 5 % CO_2). Les cellules à visualiser sont repérées et leurs coordonnées mémorisées ce qui permet de micro-injecter plusieurs cellules. La micro injection se fait à l'aide d'un micro-capillaire par un système eppendorf (Femtojet/injectMan N12 microinjector). L'acquisition des images avec un objectif à huile 100X HCX PL APO se fait toutes les 10 min grâce à une caméra (Leica DC350FX CCD). Il est ainsi possible d'observer séquentiellement l'évolution de la localisation de la fluorescence de plusieurs champs d'observation.

D. Microscopie à haute résolution

Un total de 10^5 cellulesensemencées sur lamelle de verre stérile dans une plaque 12 puits est transfecté 24 h avant fixation par du PFA 4 % avec un maximum de 2 μg d'ADN plasmidique. Les lamelles sont ensuite montées dans une chambre de Ludin contenant du D-PBS. Toutes les mesures utilisant Dronpa et Split-Dronpa pour la microscopie à haute résolution de type PALM (Photo-Activated Localization Microscopy) ont été réalisées sur un microscope inversé Nikon Eclipse Ti équipé d'un objectif à huile APO TIRF 100X avec une ouverture numérique (NA) de 1,49.

Les protéines fluorescentes sont excitées avec un laser à 488 nm (Oxxius) et réactivées par irradiation continue avec un laser à diode à 405 nm (Spectra-Physics). Les intensités maximales étaient de 6,5 kW/cm^2 et 170 W/cm^2 à 488 nm et 405 nm respectivement. Les faisceaux laser ont été réfléchis par un miroir dichroïque multibande Di-405/488/561/635. Des filtres Notch (488 nm) et Long Pass (488 nm) (Semrock) ont été utilisés pour filtrer la fluorescence recueillie. Les images ont été acquises en utilisant une caméra de type EMCCD (electron-multiplying charge-coupled Device ; Hamamatsu) avec un temps d'exposition de 50 ms.

V. Biochimie

A. Lyse des cellules et dosage protéique par Bradford

Les cellules sont récoltées par raclement du tapis cellulaire directement dans le milieu de culture puis centrifugées pendant 5 min à 1500 rpm. Une seconde centrifugation, identique à la première, est réalisée suite à un lavage avec du DPBS 1X. Les culots cellulaires sont alors remis en suspension dans 500 µL de tampon de lyse froid (Tris-HCl 10 mM pH = 7,5 ; NaCl 150 mM ; EDTA 1 mM ; NP40 1 % ; SDS 0,05 %) dans lequel une tablette d'inhibiteurs de protéases a été ajoutée (complete Mini, EDTA free Protase Inhibitor Cocktail Tablets, Roche, Réf. 04693159001). Après 20 min d'incubation sur glace, une centrifugation à 14000 rpm pendant 20 min et à 4 °C permet d'éliminer les débris cellulaires. Les protéines retrouvées dans le surnageant sont quantifiées par un dosage de Bradford.

Ce dosage colorimétrique utilise la modification de l'absorbance à 525 nm provoquée par la liaison du bleu de Coomassie (constituant du réactif de Bradford, kit Bio-rad Protein Assay II, Réf. 5000002) aux acides aminés des protéines présentes dans la solution. Une gamme étalon réalisée à partir de BSA (Bovine Serum Albumine) permet d'établir une droite d'étalonnage. Il est alors possible de déterminer la concentration protéique des échantillons grâce à cette droite.

B. Fractionnement cellulaire

Comme pour la lyse, les cellules sont directement récoltées par raclement de la boîte (un million de cellulesensemencées dans une boîte de pétri de 10 cm de diamètre 8 h avant la transfection), puis lavées par du DPBS 1X. Le culot est en revanche re-suspendu et incubé sur glace pendant 20 min dans 200 µL de tampon A (10 mM HEPES pH = 7,4 ; 10 mM KCl ; 1 mM EDTA ; 1 mM EGTA ; 250 mM sucrose ; 1,5 mM MgCl₂ ; tablette d'inhibiteurs de protéases). Après homogénéisation à l'aide d'une aiguille 22G (environ vingt passages), une centrifugation à 850 g pendant 5 min à 4 °C est réalisée. Le culot est suspendu dans 50 µL de tampon C (20 mM HEPES pH = 7,4 ; 2,5 % (V/V) glycérol ; 420 mM NaCl ; 1 mM EDTA ; 1 mM EGTA ; 1,5 mM MgCl₂ ; tablette d'inhibiteurs de protéases) correspond à la **fraction nucléaire**. Le surnageant est quant à lui ultracentrifugé à 55000 rpm pendant 45 min à 4 °C. Le culot correspondant à la **fraction membranaire** est re-suspendu dans 50 µL de tampon de lyse froid (Tris-HCl 10 mM

Manuel PIRES

pH = 7,5 ; NaCl 150 mM ; EDTA 1 mM ; NP40 1 % ; SDS 0,05 % ; tablette d'inhibiteurs de protéases) alors que le surnageant correspondant à la **fraction cytosolique** est conservé. Chaque fraction est alors quantifiée par dosage de Bradford.

C. Immunoprécipitation (IP)

L'immunoprécipitation a été utilisée afin de concentrer les protéines d'intérêts pour être analysées en spectrométrie de masse. Une fois que les protéines sont extraites et quantifiées (voir partie [Lyse des cellules et dosage protéique par Bradford]), l'échantillon est séparé en deux, une première partie ne subissant pas de cross-link et une seconde subissant un cross-link (voir partie suivante [cross-link]). Pour chaque échantillon, un dépôt « input » permet de vérifier que les protéines sont bien exprimées. Afin d'être déposé dans un gel de polyacrylamide, un aliquot de lysat correspondant à 40 µg de protéines totales est prélevé, dans lequel 5 mM de DTT (Dithiothreitol, réduction des ponts disulfures) et du bleu de Laemmli 4X (Tris-HCl 277,8 mM, glycérol 44,4 % (V/V), SDS 4,4 % (V/V) et bleu de bromophénol 0,02 % (m/V)) sont ajoutés. Le volume est complété à 40 µL (concentration à 1 µg/µL). Le mélange est chauffé pour dénaturer les protéines 5 min à 95 °C puis conservé sur glace en attendant le dépôt sur gel. Pour les fractions non cross-linkées et cross-linkées destinées à l'IP, un volume équivalent à 1 mg de protéines est incubé pendant 2 h à 4 °C sous agitation en présence de 1 µg d'anticorps anti-eGFP (**Figure 41 A**). L'anticorps va ainsi se fixer spécifiquement aux protéines eGFP et au grand fragment de la split-GFP (GFP1-10) présentes dans l'échantillon (**Figure 41 B**). Après avoir été équilibré dans du tampon de lyse, un volume de 50 µL de suspension de billes magnétiques recouvertes de protéines A (Dynabeads® Protein A for IP, ThermoFisher scientific, Réf. 10001D) est ajouté et incubé 1 h à 4 °C sur roue (**Figure 41 C**). Les protéines A vont se lier à la partie constante des anticorps anti-eGFP. Grâce aux billes magnétiques, les protéines d'intérêt vont être purifiées et l'ensemble des protéines qui n'est ni lié aux protéines d'intérêt (interaction protéique, co-IP), ni reconnu par l'anticorps (IP), ni lié directement aux protéines A (bruit de fond) va être éliminé lors des différents lavages (trois lavages avec 1 mL de tampon de lyse). Les billes sont finalement resuspendues dans 50 µL d'un mélange contenant 5 mM de DTT, du bleu de Laemmli 4X et de l'eau. Une dénaturation pendant 5 min à 95 °C permet de supprimer les liaisons existantes entre les protéines d'intérêt, les anticorps (supprime également les ponts disulfures entre les chaînes lourdes et

légères des anticorps) et les protéines A (les billes magnétiques sont également séparées). L'utilisation du rack magnétique lors du dépôt sur gel va permettre de ne déposer que les protéines dénaturées.

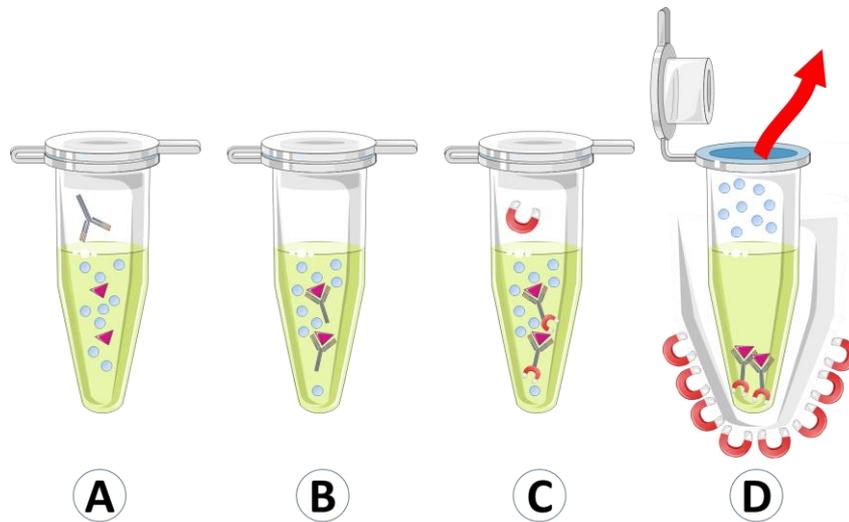


Figure 41 : Principe de la co-immunoprécipitation. (A) Dans le lysat cellulaire est ajouté l'anticorps anti-eGFP. (B) Après 2 h d'incubation sous agitation, l'anticorps reconnaît spécifiquement la protéine d'intérêt, ici la protéine GFP couplée à Gag (triangle rouge). (C) Des billes magnétiques (ici représentées par un aimant) recouvertes de protéine A reconnaissent la partie constante de l'anticorps, ce qui permet (D) de faire précipiter les billes et la protéine d'intérêt qui y sont attachées. Après avoir été décrochées des billes et dénaturées, les protéines sont analysées par migration par un gel d'acrylamide SDS-PAGE.

D. Cross-link

Afin de vérifier par spectrométrie de masse, si la reformation de la split-GFP se fait au sein de la forme compacte de Gag ou par multimérisation de cette dernière, nous devons au préalable figer les structures tridimensionnelles des protéines dans une conformation proche de leur état natif. Pour cela, l'utilisation de cross-linkers permet de former des liaisons covalentes entre résidus lysines et/ou arginines. Le choix d'un cross-linker BS3 lourd/léger (BS3-d4/d0 ; Thermo Fisher Scientific, Réf.21595 et 21590 ; **Figure 42 A et B**) composé d'un mélange de BS3 (hydrosoluble, imperméable aux cellules, les groupes réactifs sont des esters d'amines primaires, bras espaceur non clivable de 11,4 Å) dont quatre des hydrogènes ont été remplacés par des deutériums (d4) et de BS3 non modifié (d0), permet, lors de l'analyse en spectrométrie de masse, de repérer facilement les fragments cross-linkés qui apparaissent alors comme un doublet. Dans le lysat protéique (équivalent à 1 mg de protéine totale), les deux cross-linkers sont ajoutés à une concentration finale de 2,5 mM pendant 30 min à

Manuel PIRES

température ambiante permettant la formation d'un lien covalent entre des lysines et/ou des arginines distantes de 11,4 Å ou moins (**Figure 42 C**). Afin d'arrêter la réaction, le lysat est par la suite incubé pendant 15 min à température ambiante avec une solution de Tris pH = 7,5 à une concentration finale de 20 mM. Les protéines cross-linkées sont ensuite préparées pour être dénaturées (ajout des protéines dans un mélange contenant 5 mM de DTT, du bleu de Laemmli 4X et de l'eau) et déposées sur gel de polyacrylamide.

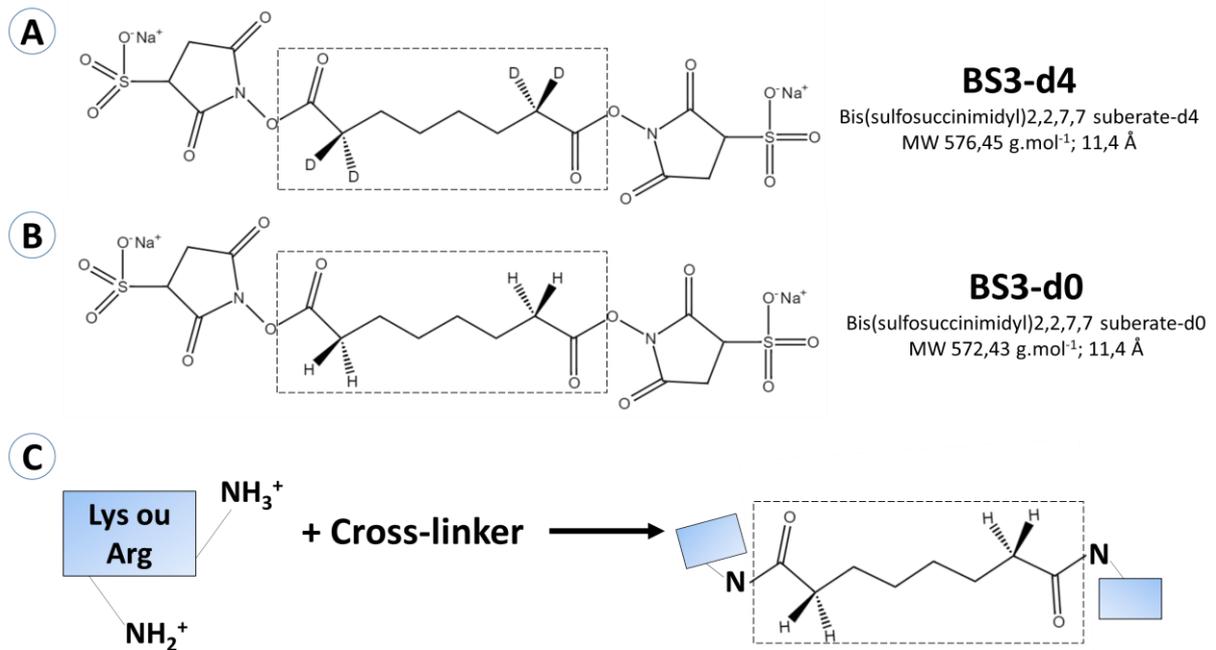


Figure 42 : Utilisation de cross-linkers BS3-d0 et d4 pour figer la structure de la protéine. Cross-linkers (A) BS3-d4 possédant quatre isotopes de l'hydrogène dits lourds (D pour Deutérium) et (B) BS3-d0 possédant quatre hydrogènes dits légers (H). (C) Création d'une liaison covalente entre les cross-linkers et les fonctions amines de la chaîne latérale des résidus lysine et arginine d'une protéine.

E. SDS-PAGE et gel natif

1. SDS-PAGE

Les échantillons ayant subi une dénaturation sont ensuite analysés sur un gel de polyacrylamide (PAGE) préalablement préparé (**Tableau 9**). Le SDS en se liant à un ratio approximativement constant (une molécule de SDS pour deux acides aminés) charge négativement les protéines permettant alors leur séparation sur la base de leurs poids moléculaires. Les migrations ont lieu dans du tampon de migration (Tris 25 mM ; Glycine 192 mM ; SDS 0,1 % (V/V), pH = 8) à 120 V pendant un temps variant de 1 h 30 à 2 h. Un marqueur

de taille est également déposé (Page Ruler Plus, Thermo Fisher scientific, Réf. 26619) afin de déterminer expérimentalement la taille des bandes apparaissant lors de la révélation.

	Gel de séparation		Gel de concentration
	10 %	12 %	4 %
H₂O	4,8 mL	4,3 mL	5,1 mL
Acrylamide:bisacrylamide 29:1 (40 %)	2,5 mL	3 mL	0,8 mL
Tris	2,5 mL (1,5 M; pH = 8,8)		2 mL (0,5 M; pH = 6,8)
SDS 20 %	50 µL		40 µL
AMPS 10 %	100 µL		40 µL
TEMED	10 µL		10 µL

Tableau 9 : Gel de poly-acrylamide SDS-PAGE. SDS = Sodium Dodecyl Sulfate ; AMPS = AmMonium PerSulfate; TEMED = TÉtraMÉthyléthylèneDiamine

2. Gel natif

La migration en conditions natives permet de conserver les structures secondaires, tertiaires et quaternaires des protéines grâce à une migration principalement dépendante du point isoélectrique (pI) de la protéine.

Le protocole présenté ici a été adapté de (David E; Garfin, 2003). Pour cela, l'équivalent de 15 µg de protéines (lyse totale ou fractionnement) ajouté à 15 µL de tampon de dépôt de McLellan (10 % (V/V) glycérol ; 0,01 % (m/V) de bleu de bromophénol ; 0,5 % (V/V) de tampon 5X de McLellan (pH = 10,2; 37 mM Ammoniaque et 20 mM CAPS : 3-(cyclohexylamino)-1-propanesulfonate)) est analysé par électrophorèse sur un gel de polyacrylamide 6 % continu (pas de gel de concentration ; 20 % (V/V) de tampon de gel 5X de McLellan pH = 10,2 ; 20 % (V/V) d'Acrylamide/bisacrylamide 30 % (37,5:1) ; 0,5 % AmMonium PerSulfate (AMPS) ; et 0,1 % TEMED). Cette migration se fait dans du tampon de McLellan (pH = 10,2) pendant 2 h 30 sur glace afin d'éviter la dénaturation des protéines. Avant de réaliser un transfert sur une membrane de PVDF (protocole de transfert identique à celui des gels dénaturants) puis de faire un western blot, la fluorescence de l'eGFP est directement observée sur gel en utilisant une excitation à 460 nm et en observant la fluorescence émise par un lecteur LAS4000 (GE Healthcare Life Sciences) muni d'un filtre de bande-passe 510DF10.

F. Western blot

Les protéines séparées soit sur gel SDS-PAGE, soit sur gel natif sont ensuite transférées (Tampon de transfert : Tris 25 mM ; Glycine 192 mM ; Ethanol 20 % (V/V) ; pH = 6,8) sur une membrane de PVDF (PolyVinylidène Fluoride ; Amersham Hybond P Réf. 10600023) préalablement « activée » dans de l'éthanol 100 % pendant 10 min puis équilibrée dans du tampon de transfert pendant 20 min minimum. Afin de limiter l'augmentation de la température, le transfert se fait à 4 °C (glace et agitation du tampon) pendant 2 h 30 à 110 V. Pour bloquer les sites non spécifiques des membranes, celles-ci sont par la suite incubées pendant une heure sous agitation avec une solution de caséine 3 % (m/V ; Blotting-Grade Blocker, Bio-Rad Réf.1706404) dans du tampon TBS-T (Tris Buffer Saline-Tween ; Tris 50 mM ; NaCl 138 mM ; KCl 2,7 mM ; Tween 0,2 % (V/V) ; Sigma Réf. T9039). Les membranes sont ensuite incubées pendant 1 h à température ambiante ou sur la nuit à 4 °C sous agitation avec une solution d'anticorps primaire dilué dans de la caséine 3 % (m/V). Après trois lavages par du tampon TBS-T, les membranes sont incubées 1 h à température ambiante et sous agitation avec une solution de caséine 3 % (m/V) contenant l'anticorps secondaire couplé à la HRP (Horse Radish Peroxydase, qui permet la production de photon de luminescence en réagissant avec le luminol). Après trois lavages par du tampon TBS-T, l'activité de l'enzyme HRP est alors révélée par ajout d'un substrat de peroxydase ECL (Enhanced Chemiluminescence ; Pierce™ ECL Western Blotting Substrate, ThermoFisher Scientific, Réf. 32106) et la luminescence émise est quantifiée par l'appareil Image Quant LAS4000 (GE Healthcare Life Sciences).

G. Marquage des gels au bleu de coomassie

Au cours de la purification de protéines ou de la préparation de protéines cross-linkées pour une étude en spectrométrie de masse, il est parfois nécessaire de visualiser les protéines dans le gel de polyacrylamide qui est alors coloré au bleu de coomassie. Le bleu de Coomassie se lie aux acides aminés basiques et aromatiques des protéines présentes dans le gel. Pour ce faire, après migration, le gel est incubé sous agitation pendant 30 min dans une solution de Quick Coomassie Stain (CliniSciences, Réf. GEN-QC-STAIN-1L), puis lavé pendant 30 min sous agitation dans de l'eau distillée. Une image du gel est par la suite réalisée sur un appareil GeneGenius (B10 Imaging System, Syngène).

VI. Biologie moléculaire

A. Transformation de bactéries compétentes et purification de l'ADN plasmidique

Les bactéries chimiquement compétentes (50 μ L) DH5 α sont décongelées lentement dans la glace puis mises en contact pendant 30 min sur glace avec le plasmide d'intérêt (ou le produit de ligation). Un choc thermique de 30 s à 42 °C permet de faire entrer le plasmide. Le tube est ensuite incubé sur glace pendant 2 min puis les bactéries sont mises en suspension dans 500 μ L de milieu LB (10 g.L⁻¹ tryptone, 10 g.L⁻¹ NaCl, 5 g.L⁻¹ extrait de levure) sans antibiotiques et sous agitation (200 tpm) pendant 1 h à 37 °C. Cette étape permet la reconstitution de la membrane bactérienne et l'expression du gène de résistance à l'antibiotique. Les bactéries sont étalées sur des boîtes de pétri contenant du milieu LB Agar (15 g.L⁻¹ Agar, 10 g.L⁻¹ tryptone, 5 g.L⁻¹ NaCl, 5 g.L⁻¹ extrait de levure) supplémenté par l'antibiotique dont le plasmide d'intérêt possède le gène de résistance (ampiciline 100 μ g/mL ; kanamycine 50 μ g/mL) et placées à 37 °C pendant la nuit. Le lendemain, une (dans le cas d'une amplification de plasmide, préculture) ou plusieurs (dans le cas d'une mutagenèse, miniculture) colonie(s) est/sontensemencée(s) dans 5 mL de milieu LB (chacune) avec antibiotique et incubées à 37 °C sous agitation à 220 tpm. Cette préculture ou miniculture peut alors être utilisée pour :

- après 8 h d'incubation, ensemenecer 200 mL de milieu LB avec antibiotique qui seront incubés la nuit à 37 °C sous agitation à 220 tpm. La purification de l'ADN plasmidique est alors réalisée sur des colonnes échangeuses d'anions grâce au kit Xtra Midi Plus® (Macherey-Nagel, Réf. 740410).
- après 16h d'incubation, extraire l'ADN plasmidique de 3 mL de culture sur une membrane de silice grâce au kit NucleoSpin® Plasmid (Macherey-Nagel, Réf.740588). Cet ADN est ensuite envoyé à la compagnie GATC Biotech pour séquençage et vérification de la séquence et de la correcte insertion de l'insert ou de la présence de la mutation attendue.

Les concentrations d'ADN sont ensuite déterminées par lecture de l'absorbance à 280 nm grâce à un spectrophotomètre de type NanoDrop (Thermo Fisher Scientifique).

B. Mutagénèse

L'insertion du multi-cloning site 2 (MCS2) ou la création des différents mutants de Gag ont été réalisées grâce au kit « Site Directed Mutagenesis » (Thermo Fisher Scientific, Réf. F-541) suivant un protocole en trois étapes :

- réaction de PCR (Polymérase Chain Reaction, **Tableau 10**) : qui permet l'amplification de l'intégralité du plasmide à muter par l'utilisation d'oligonucléotides phosphorylées en 5'. Les différents oligonucléotides utilisés sont présentés dans le **Tableau 7**. Certaines PCR étant plus difficiles que d'autres par la présence de régions riches en GC dans la séquence à amplifier, nous avons parfois changé le tampon « HF » pour du tampon « GC rich » ou ajouté dans les cas les plus difficiles 2 % de DMSO afin de faciliter la polymérisation.

Composant	Volume (µL)	Concentration finale
H ₂ O	Qsp 50 µL	
Tampon 5X Phusion HF	10	1X
10 mM dNTPs mix	1	250 µM chacun
Amorce Fwd	2,5	10 µM
Amorce Rev	2,5	10 µM
ADN à muter	X	1 à 10 ng
ADN polymérase Phusion Hot Start (2 U/µL)	0,5	0,02 U/µl

Tableau 10 : Mélange réalisé lors de la PCR pour une mutagénèse, une insertion ou une délétion

- ligation qui se fait par la T4 ADN quick ligase, permet la circularisation du plasmide muté grâce aux phosphates portés par les oligonucléotides.
- transformation de l'ADN circularisé dans des bactéries compétentes DH5α puis vérification par séquençage que l'insertion, la mutation ou la délétion est correcte.

C. Digestion par enzyme de restriction

Après avoir inséré le MCS2 par mutagénèse en 3' de la région codante de Gag_eGFP (aboutissant au plasmide gag_eGFP_MCS2), le MCS1 est quant à lui inséré entre les sites de restriction BamH1 et EcoR1 à la place de l'eGFP. Pour cela, les oligonucléotides sens (oligo #1 dans le **Tableau 7**) et anti-sens (oligo #2 dans le **Tableau 7**) correspondant au MCS1 sont dans un premier temps appariés : chaque oligonucléotide est resuspendu à une concentration de

Manuel PIRES

5 μM dans un 5 μL de tampon composé de 8 mM TrisHCl pH = 7,5, 5 mM NaCl, 8 mM MgCl_2 . Dans un second temps, le mélange est chauffé à 95 °C pendant 5 min, puis pour permettre l'appariement des oligonucléotides, la température de l'échantillon est progressivement diminuée jusqu'à température ambiante. Après digestion du vecteur codant pour Gag_eGFP_MCS2 par les enzymes de restriction (BamH1 et EcoR1), les oligonucléotides appariés sont insérés par ligation pour finalement obtenir le plasmide **gag_MCS1_MCS2**. Un séquençage de l'ADN plasmidique est par la suite réalisé par la société GATC biotech.

Ces MCS sont ensuite utilisés pour insérer les régions codantes des fragments de split-GFP grâce à la présence des sites de restriction. Il est d'abord nécessaire d'insérer par PCR les sites des enzymes de restriction aux extrémités 3' et 5' du gène d'intérêt. Il faut ensuite réaliser une digestion enzymatique de ces sites ainsi que du vecteur plasmidique dans lequel nous souhaitons insérer le fragment (dans MCS1 ou MCS2 de gag_MCS1_MCS2). La quantité d'enzyme de restriction est dépendante du type d'enzyme et le calcul se fait selon l'**Équation 4**. Un mélange (qsp 100 μL) contenant 20 μL de tampon CutSmart 5X, l'ADN à digérer et la/les enzyme(s) de restriction(s) est incubé sur la nuit à 37 °C. Pour dénaturer l'enzyme de restriction, un chauffage à 65 °C pendant 20 min est réalisé.

$$\text{quantité d'enzyme}(\mu\text{L}) = \frac{\text{Kb vecteur réf}}{\text{Kb ADN}} \times \frac{\text{site ADN}}{\text{site référence}} \times \frac{\text{qtt ADN à digérer}(\mu\text{g})}{\text{concentration de l'enzyme}(\frac{\text{U}}{\mu\text{L}})}$$

Équation 4 : Détermination du volume d'enzymes de restriction nécessaire pour la digestion.

Kb vecteur réf correspond à la taille (en kilobase) du vecteur d'étude de digestion de l'enzyme (en général Lambda (48,5 Kb) ; *Kb ADN* correspond à la taille (en kilobase) du plasmide à digérer ou de l'ADN préalablement amplifié par PCR ; *site ADN* correspond au nombre de sites sur l'ADN ou le plasmide à digérer (souvent 1) ; *site référence* correspond au nombre de sites de restriction pour l'enzyme présent sur le vecteur référence. *qtt* dans l'équation correspond à quantité.

Avant de réaliser une ligation, le vecteur (plasmide digéré par les enzymes de restriction) est d'abord déphosphorylé (10 % (V/V) Tampon Antarctic phosphatase, 2 % (V/V) d'enzyme de déphosphorylation antarctic phosphatase, New England Biolabs, Réf. M0289S) afin d'empêcher une « auto-ligation ». La ligation de l'insert et du vecteur se fait en utilisant la Quick T4 DNA ligase du kit « Quick Ligation » (New England Biolabs, Réf. M2200). Les quantités d'insert et de vecteur sont déterminées en utilisant le NEBIOCalculator® (<http://nebiocalculator.neb.com/#!/ligation>).

Manuel PIRES

L'ADN circulaire finalement obtenu est ensuite transformé selon le protocole détaillé dans la partie [Transformation de bactéries compétentes et purification de l'ADN plasmidique].

Si nécessaire, une migration sur gel d'agarose (1,5 % (V/V) dans du tampon TAE, Tris Acetate, EDTA) permet de vérifier la bonne amplification et la taille de l'ADN obtenu. Si besoin, les résultats de PCR ou de digestion sont purifiés sur gel ou sur colonne pour ôter les oligonucléotides ou les fragments de digestion (Nucleospin® Gel and PCR Clean-up, Macherey-Nagel, Réf.740609).

VII. Purification de la protéine Gag

La purification de la protéine Gag a été adaptée du protocole établi dans par l'équipe de Johnson Mak (McKinstry et al., 2014).

Des bactéries compétentes *E. Coli* BL21(DE3)pLysS transformées par un plasmide pET28A codant la protéine Gag fusionnée en C-terminal à une étiquette contenant six résidus histidines (Gag_6H) ont été cultivées dans du milieu LB contenant 100 µg/mL d'ampicilline et 25 µg/mL de chloramphenicol. Une pré-culture (20 mL) de ces bactéries (12 h d'incubation à 37 °C) sert à inoculer 1 L de milieu LB contenant les antibiotiques (Ampicilline et chloramphenicol). Cette culture a été menée à 37 °C sous agitation à 200 tpm jusqu'à l'obtention d'une DO₆₀₀ de 0,4. La température de l'incubateur est alors réduite à 18 °C jusqu'à obtention d'une DO₆₀₀ de 0,6. L'expression de la protéine a été induite par ajout de 0,5 mM d'IPTG (Isopropyl β-D-1-thigalactopyranoside) pendant 18 h à 18 °C sous agitation. Le culot bactérien obtenu après une centrifugation de 30 min à 3500 rpm à 4 °C est alors congelé rapidement dans de l'azote liquide et conservé à -80 °C. Afin de vérifier l'induction, 1 mL de milieu LB contenant les bactéries avant et après l'induction sont prélevés puis centrifugés à 13000 rpm pendant 2 min. Les culots sont resuspendus dans 70 µL d'H₂O puis chauffés pendant 5 min à 95°C. Après une seconde centrifugation à 13000 rpm pendant 2 min, 35 µL de surnageant sont mélangés avec 15 µL de Laemmli 4X et 10 µL de DTT. Avant d'être déposés sur gel de polyacrylamide 12 %, les échantillons sont à nouveau chauffés 5 min à 95 °C. La migration se fait à 150 V pendant 1 h et le gel est ensuite révélé en bleu de Coomassie pour vérifier si l'induction a été efficace.

Le jour de la purification, le culot bactérien est décongelé sur glace dans 10 mL de tampon A (Tris 50 mM, NaCl 1 M, DTT 2,5 mM, ZnCl₂ 1 µM, à pH = 8) supplémenté d'inhibiteur de protéases (protease inhibitor cocktail, Roche) et de 200 µL de Tween 20. Pour améliorer la destruction des membranes bactériennes, les bactéries sont incubées 30 min à 4 °C sous agitation rotative après ajout de lysozyme (1 mg/mL). La lyse des bactéries est finalisée par une série de huit sonications (IKASONIC U50) de 20 s intercalées par un refroidissement du lysat bactérien sur glace. Le lysat est ensuite clarifié par centrifugation à 15000 rpm pendant 1 h à 4 °C puis chargé sur une colonne de nickel (25 mL, Ni-NTA agarose column, GE Healthcare) préalablement équilibrée avec 100 mL de tampon A. Le lysat cellulaire est ensuite

Manuel PIRES

chargé sur la colonne à une vitesse de 1 mL/min pour que les protéines se fixent sur le nickel au niveau de l'étiquette histidine. Après un lavage de 2 h à 1 mL/min par un mélange composé de tampon A et de 6 % de tampon B (Tampon A + 500 mM Imidazole), les protéines sont éluées par un gradient linéaire d'imidazole de 5 à 500 mM (élution de la protéine Gag_{6H} par compétition d'affinité pour le nickel) obtenu en mélangeant du tampon A et du tampon B à un débit de 0,5 mL/min. Les fractions éluées (2 mL par fraction) sont analysées sur gel SDS-PAGE 12 % et le gel est coloré au bleu de Coomassie afin de sélectionner les fractions contenant majoritairement la protéine Gag_{6H}. Ces fractions sont alors regroupées et dialysées pendant la nuit contre le tampon A. La protéine est ensuite chargée sur une colonne d'exclusion de taille (100 mL, Superdex 200 10/300, GE Healthcare Life Sciences) préalablement équilibrée par 200 mL de tampon A à 0,5 mL/min. Les fractions de 2 ml issues d'une élution à 0,5 mL/min sont également analysées par SDS-PAGE coloré au bleu de Coomassie. Les fractions contenant un rapport protéine d'intérêt sur contaminant élevé sont rassemblées et concentrées sur des filters à centrifuger (Centricon 30 kDa filters Merck Millipore Réf. UFC903024) afin d'éliminer toutes les protéines dont la masse moléculaire est inférieure à 30 kDa. Pour déterminer la concentration de notre protéine, nous mesurons l'absorbance à 280 nm par spectrométrie. La concentration finale en protéine a été déterminée en utilisant la loi de Beer-Lambert (coefficient d'extinction molaire $62\,000\text{ M}^{-1}\cdot\text{cm}^{-1}$) en tenant compte de l'ajout de glycérol (5 % (V/V)) pour la conservation. La protéine est aliquotée par 50 μL et rapidement congelée dans de l'azote liquide avant d'être conservée à $-80\text{ }^{\circ}\text{C}$.

La qualité de la purification est évaluée par mesure de la capacité de Gag à déstabiliser les ANs. Pour cela, un test utilisant la boucle cTAR couplée à un fluorophore et un quencher à ses deux extrémités et mis en présence de Gag (*Figure 35*). La déstabilisation de la structure en tige boucle par l'activité chaperonne de Gag_{6H} permet d'observer une augmentation de la fluorescence.

VIII. Spectrométrie de masse (MS)

Après cross-link, immuno-précipitation et migration sur gel de polyacrylamide, les bandes de gel contenant les protéines d'intérêts sont découpées et envoyées au Laboratoire de Spectrométrie de masse bio-organique (LSMBO ; Maxime Bourguet, Dr Sarah Sanglier-Cianferani).

Les échantillons sont dans un premier temps réduits (DTT, 1,4-Dithiothreitol), puis alkylés (IAA, IodoAcetic Acid), pour finalement être déshydratés à l'acétonitrile. Entre chacune de ces étapes, plusieurs lavages sont réalisés au bicarbonate d'ammonium (NH_4HCO_3). Les protéines sont digérées par la trypsine (37 °C toute la nuit), puis, les peptides sont extraits, concentrés sous vide et repris dans un mélange $\text{H}_2\text{O}/\text{Acétonitrile}/\text{HCOOH}$ (acide formique).

Les échantillons sont ensuite injectés et analysés par un système nanoLC-Orbitrap (Thermo Scientific, Waltham, USA) qui permet une séparation chromatographique des protéines. Dans ce système sont intégrées deux colonnes : une colonne ACQUITY UPLC® BEH130 C18 column (250 mm x 75 μm avec une taille de particules de 1,7 μm ; Waters Corporation, Milford, USA) et une pré-colonne Symmetry C18 precolumn (20 mm x 18 μm avec une taille de particules de 5 μm Waters Corporation, Milford, USA). La séparation chromatographique se fait en deux étapes :

- Étape d'enrichissement : les échantillons sont injectés et retenus sur la pré-colonne par 99 % de solvant A (H_2O + 0,1 % acide formique) et de 1 % de solvant B (acétonitrile + 0,1 % acide formique) ;
- Étape d'éluion : les peptides sont élués à l'aide d'un gradient de solvant B de 1 % à 35 %.

Les dix peptides les plus abondants sont sélectionnés et isolés en MS pour ensuite subir une fragmentation de type HCD (Higher-energy Collisional Dissociation) et les fragments ainsi générés sont également analysés en MS (analyse MS/MS). Une recherche Mascot® est effectuée pour comparer la séquence théorique des protéines cibles et les données MS collectées pendant l'analyse. Cette comparaison permet d'évaluer la couverture de séquence des protéines cibles ainsi que le nombre de peptides identifiés cross-linkés ou non. Parallèlement, deux logiciels de traitement sont utilisés afin d'identifier les peptides cross-linkés : pLink® et Mass Spec Studio®. On relève le rapport m/z du peptide et le temps auquel

Manuel PIRES

le peptide est élué au cours de l'analyse (temps de rétention) afin d'obtenir le spectre de masse sur une fenêtre de temps de l'analyse. En diminuant progressivement les fenêtres de temps, on cherche à identifier plus finement sur quelle fenêtre de temps sortent les peptides cross-linkés lourd (BS3-d4) et léger (BS3-d0). Une fois identifié, on combine le spectre sur une fenêtre de temps incluant l'identification des deux peptides et on vérifie qu'il y a bien un ratio de 50/50 en termes d'intensité pour les deux peptides en question. Si c'est le cas, le peptide est alors validé, permettant de montrer l'existence d'un cross-link intramoléculaire ou intermoléculaire. En revanche, si le ratio n'est pas de 50/50, le peptide n'est généralement pas retenu.

Chapitre 4 : Résultats et discussions

Mon travail de thèse s'est articulé autour de deux projets, un premier dans lequel j'ai testé plusieurs molécules grâce à un test développé et mis au point par mes soins au laboratoire utilisant des pseudo-particules virales qui miment les phases précoces du cycle viral du VIH-1. Ce travail faisait partie d'un projet européen important THINPAD et a été complété par le développement de tests en vue d'identifier les mécanismes d'action des molécules les plus intéressantes. Mon second projet s'est quant à lui axé sur l'étude par différentes techniques de la forme compacte de Gag *in cellulo*.

I. Criblage de molécules antivirales

A. Projet européen THINPAD

1. Développement et validation du test PSEUDO

Notre laboratoire n'est pas équipé d'un laboratoire de sécurité microbiologique de type L3 nous empêchant de travailler avec une souche virale de VIH-1 sauvage. Comme nous travaillons plus particulièrement sur la protéine NC qui exerce une partie de ses rôles dans les phases précoces du cycle viral nous avons choisi de travailler avec des particules infectieuses mais non répliquatives. Les activités antivirales des molécules ciblant la NC ont été testées dans un test d'infectivité appelé PSEUDO qui permet de valider leur effet sur les phases précoces de l'infection uniquement.

Dans ce test, nous utilisons des pseudo-particules virales (PPs) de deuxième génération (Naldini et al., 1996). Ce système présente plusieurs avantages. D'abord, il nous permet de travailler en sécurité car ces particules sont infectieuses mais non répliquatives (génomme viral incomplet et absence de gènes accessoires) ne permettant donc qu'un seul cycle d'infection. Ensuite, ces particules sont pseudo-typées par la glycoprotéine membranaire du virus de la stomatite vésiculaire (VSV-G) permettant l'infection de cellules n'exprimant pas les récepteurs naturels utilisés par le VIH-1. Enfin, le génome encapsidé dans les particules, n'est pas un génome viral complet mais porte un gène rapporteur (Luciférase ou GFP) sous contrôle du promoteur CMV (Cytomegalovirus, reconnu par l'ARN polymérase des cellules de mammifère)

permettant de quantifier l'infection. L'incorporation de trois plasmides dans les cellules productrices (293T) est nécessaire pour la production de ces particules (**Figure 43**) :

- **Plasmide d'emballage** dit « Gag-Pol » : permet l'expression sous la dépendance du promoteur CMV des protéines du VIH-1. Parmi celles-ci, on retrouve les protéines de structure et enzymatique (Gag-Pol) et les protéines régulatrices (Tat et Rev). Les gènes accessoires, d'enveloppe et les LTR aux extrémités 3' et 5', ont été supprimés empêchant l'encapsidation et la transcription inverse LTR dépendantes de l'ARNm produit à partir de ce plasmide.
- **Plasmide d'enveloppe** dit « VSV-G » : qui permet l'expression à la surface des PP de la Glycoprotéine du Virus de la Stomatite Vésiculaire (VSV-G) remplaçant ainsi les glycoprotéines du VIH-1 (gp120 et gp41). Ce pseudotypage par la VSV-G permet d'infecter un large panel de cellules car cette glycoprotéine est connue pour interagir avec un récepteur cellulaire (récepteur aux Low Density Lipoproteins, LDL) présent sur un grand nombre de types cellulaires (Amirache et al., 2014; Finkelshtein et al., 2013). Il est important de noter que l'entrée du pseudovirus est réalisée par endocytose alors que dans le cas du VIH-1 sauvage, c'est une fusion des membranes virale et cellulaire récepteurs (CD4) et corécepteurs (CCR5 ou CXCR4) dépendante qui est observée.
- **Plasmide de transfert** dit « Ψ » : code sous la dépendance d'un promoteur CMV pour un pseudo ARN viral qui porte une séquence codante pour un gène rapporteur (luciférase ou eGFP). L'ARN viral produit présente à ses extrémités des séquences UTR modifiées. En effet, la séquence U3 en 5' est absente et on retrouve également une délétion partielle de la séquence U3 en 3' (boîte TATA et sites de liaison aux facteurs de transcriptions Sp1 et NF- κ B indispensables à la transcription du VIH-1) (Zufferey et al., 1998). Cette délétion permet, après transfert de cette région U3 mutée en 5' de l'ADN proviral lors de la transcription inverse, d'obtenir un ADN qui ne peut pas être transcrit par le promoteur viral. La région U5 est en revanche intacte, permettant l'encapsidation de cet ARN grâce à la présence de la séquence d'encapsidation Ψ . Enfin, l'ADN intégré présente un gène rapporteur entre les deux LTR placés sous la dépendance d'un promoteur CMV permettant l'expression de l'eGFP ou de la luciférase une fois l'ADN proviral intégré. Ce rapporteur rend possible la quantification et la visualisation de l'infection virale.

Manuel PIRES

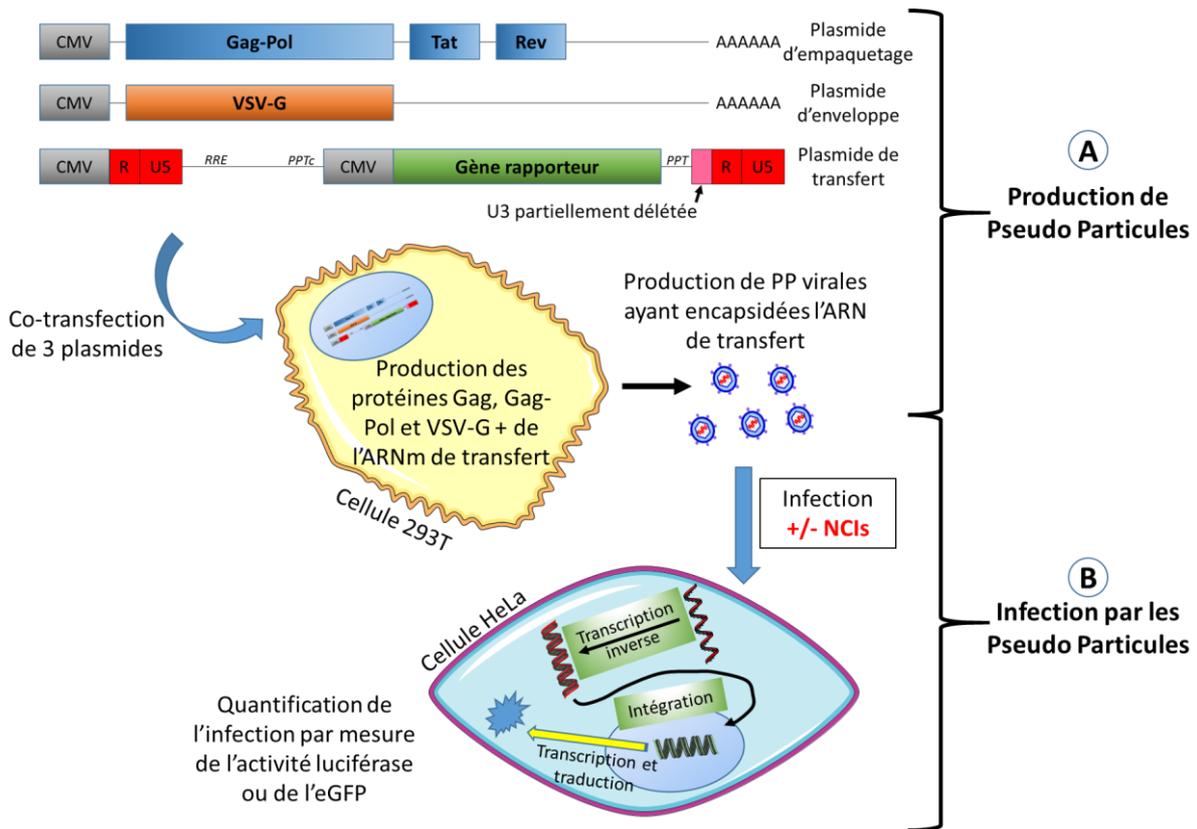


Figure 43 : Production de pseudo-particules virales (PPs) et utilisation dans le test PSEUDO. (A) Plasmides nécessaires pour la production de PPs : plasmide d'empaquetage permet l'apport des protéines de structure et les enzymes du VIH-1 (Gag-Pol) et deux protéines régulatrices indispensables (Tat et Rev) ; plasmide d'enveloppe code pour la protéine d'enveloppe du Virus de la Stomatite Vésiculaire ; le plasmide de transfert est transcrit sous forme d'ARN qui sera encapsidé dans les PPs, rétrotranscrit puis intégré dans le génome de la cellule cible où le gène rapporteur (luciférase ou eGFP) permettra la quantification de l'infection. (B) Ces PP sont alors utilisées pour infecter des cellules HeLa en présence d'inhibiteurs de la nucléocapside (NCIs) permettant de tester l'effet de ces derniers sur les phases précoces de l'infection virale.

Détermination de la quantité de PPs à utiliser

La première étape du développement du test d'activité antivirale des molécules sélectionnées par le projet THINPAD a consisté à mettre au point les conditions expérimentales pour produire des PPs et infecter des cellules HeLa par ces PPs d'une façon reproductible, analysable et sans saturer les cellules lors de l'infection. Les premières particules produites permettaient l'expression de l'eGFP dans les cellules infectées. Après production, la concentration en p24 a été déterminée par un test ELISA (Innotest® HIV Antigen mAb, InnoGenetics). Différentes quantités de virus (exprimées en ng d'équivalent p24) ont été utilisées pour infecter des cellules HeLa dans des plaques 6 puits. Après 48 h d'infection, les cellules ont été fixées et comptées par cytométrie en flux afin de déterminer le nombre de cellules infectées en fonction de la quantité de virus déposée (Figure 44). Grâce à cette

Manuel PIRES

expérience, nous avons pu conclure que 100 ng d'équivalent p24 permettaient d'infecter la moitié des 300 000 cellules HeLa contenues dans nos puits. Ce pourcentage d'infection est une valeur intéressante car elle permet de ne pas saturer les cellules infectées par les virus. Nous voulions pouvoir quantifier facilement l'infection sans recourir à la cytométrie en flux. Nous avons donc produit et testé des PPs portant la luciférase comme gène rapporteur de l'infection. L'infektivité de ces PPs est similaire à celle des PPs-eGFP. De plus, 100 ng d'équivalent p24 de PPs-luciférase permettent d'obtenir une activité luciférase importante mais non saturante pour le luminomètre. Nous avons donc choisi d'utiliser 100 ng d'équivalents p24 pour infecter 300 000 cellules. Par la suite, toutes les productions de PPs ont été testées selon ce protocole pour vérifier que celles-ci peuvent être utilisées pour tester l'activité antivirale des molécules du projet THINPAD. Cette démarche nous a permis de comparer l'activité antivirale des molécules testées avec plusieurs séries de PPs.

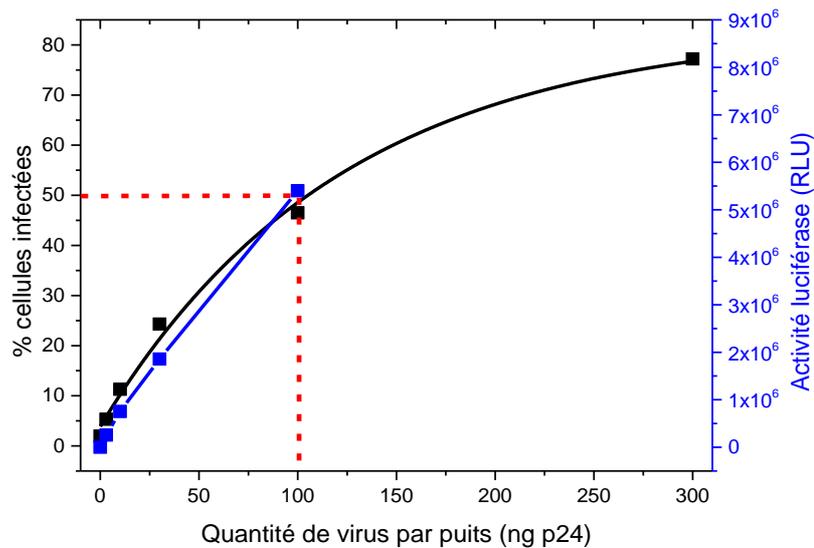
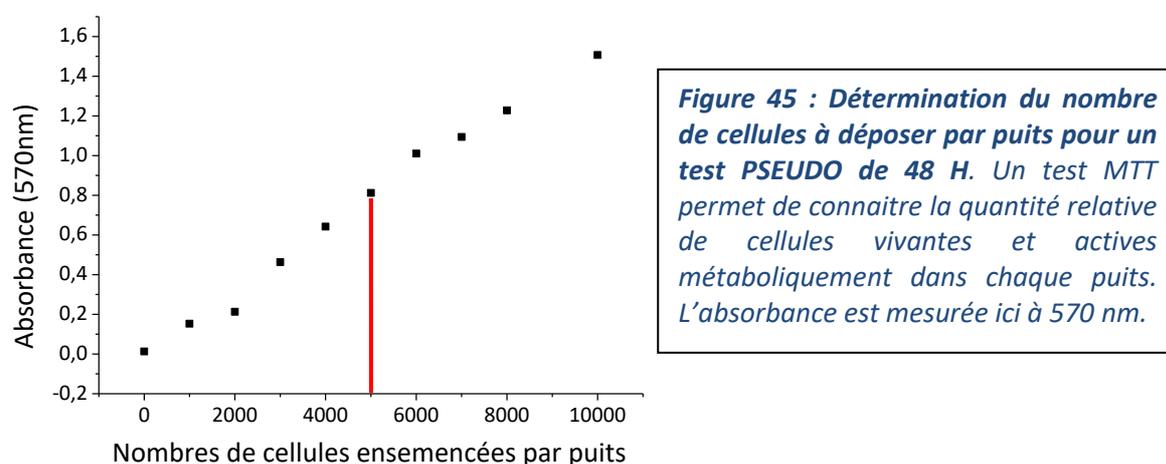


Figure 44 : Infectivité des PPs eGFP et luciférase. 300 000 cellules HeLa par puits d'une plaque 6 puits sont infectées pendant 48h par différentes quantités de PPs-eGFP (noir) ou pendant 24h par des PPs-luciférase (bleu). Le nombre de cellules infectées par PPs-eGFP est quantifié par cytométrie en flux en comptant le nombre de cellules exprimant l'eGFP (noir). La capacité d'infection des PPs-luciférase est mesurée par luminométrie (bleu). Les pointillés rouges indiquent la quantité de PPs nécessaire pour infecter 50 % des cellules dans le puits.

Manuel PIRES

Adaptation en plaque 96 puits

Afin de tester un grand nombre de molécules, nous avons décidé de travailler en plaque 96 puits. Le nombre de cellules à utiliser dans ces plaques a été déterminé grâce à un test MTT, qui permet de mesurer la viabilité cellulaire (les cellules vivantes sont capables de réduire le sel de tétrazolium (MTT) en bleu de formazan grâce à la succinate déshydrogénase mitochondriale). Nous avons donc placé des quantités variables de cellules HeLa dans chaque puits d'une plaque 96 puits et réalisé le test MTT 48 h post étalement. Nous avons alors établi que le nombre optimal de cellules permettant une lecture non saturée de l'absorbance est de 5000 cellules par puits (*Figure 45*). Par comparaison avec les résultats obtenus en plaque 6 puits (100 ng et 300000 cellules), il faut donc 1,7 ng d'équivalent p24 pour infecter 50 % des 5000 cellules contenues dans un puits d'une plaque 96 puits.



Détermination de la concentration maximale en DMSO

La plupart des molécules testées dans le projet THINPAD est dissoute dans du Diméthylsulfoxyde (DMSO). À faible concentration, ce solvant n'est pas toxique pour les cellules. À forte concentration, le DMSO est toxique pour les cellules mais permet leur cryoprotection lors de la conservation de cellules à -80 °C ou dans l'azote liquide. Lorsque sa concentration ou le temps de contact avec les cellules augmente, il devient toxique pour la cellule. Nous avons vérifié la toxicité du DMSO par un test de cytotoxicité au MTT sur des cellules infectées ou non par des PPs. Pour les cellules infectées, la mesure de l'activité luciférase nous permet également d'évaluer l'influence du DMSO sur l'infectivité des PPs. Nous avons observé qu'au-delà de 0,4 % (Volume de DMSO/Volume total du puits), le DMSO réduit de façon notable la capacité d'infection des PPs (courbe noire, *Figure 46*, diminution de l'infectivité de 20 % par rapport à l'absence de DMSO). Par ailleurs, la cytotoxicité du DMSO n'apparaît notable qu'à une concentration supérieure à 0,5 % (courbe bleue, *Figure 46*). Nous

Manuel PIRES

avons ainsi décidé de limiter la concentration de DMSO à 0,4 % dans l'ensemble de nos tests afin de garder une capacité d'infection maximale de nos PPs dans des conditions non toxiques pour nos cellules.

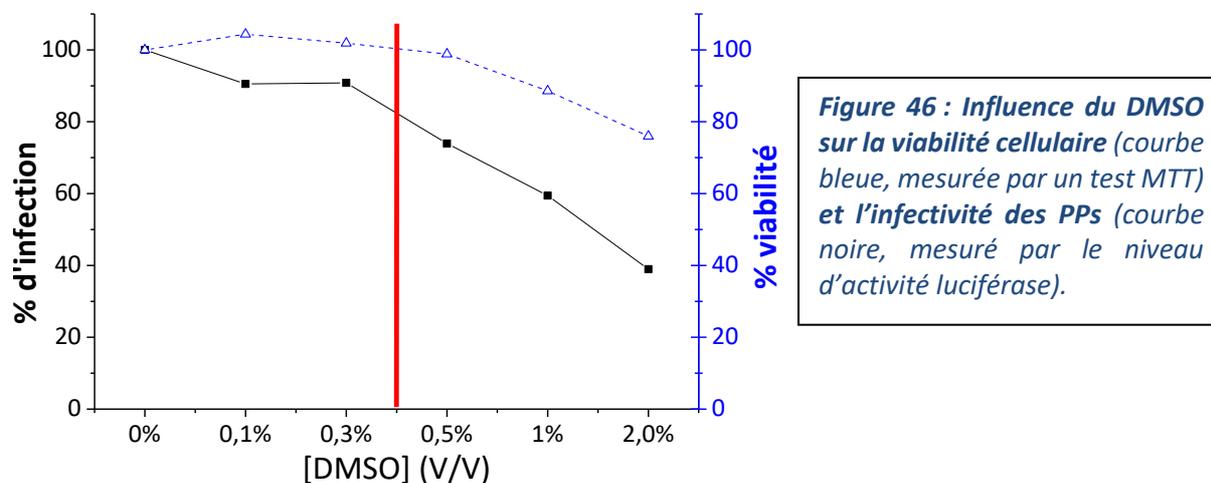


Figure 46 : Influence du DMSO sur la viabilité cellulaire (courbe bleue, mesurée par un test MTT) et l'infektivité des PPs (courbe noire, mesurée par le niveau d'activité luciférase).

Validation du test par une molécule inhibitrice témoin

Pour valider le test, nous avons utilisé un inhibiteur spécifique de la nucléocapside, le WDO-217 (Vercruyse et al., 2012) qui éjecte le zinc.

Le test a été réalisé en utilisant 5 000 cellules par puits dans 100 μ L étalées 24 h avant d'être traitées par des concentrations croissantes de WDO-217 pendant 30 min en limitant la concentration en DMSO à 0,4 % (V/V DMSO). Puis, des PPs-luciférase (1,7 ng eq. p24 par puits) ont été ajoutées pour 24 h. Une autre plaque en parallèle permet de tester la toxicité de la molécule (test MTT). Pour chaque concentration de molécule, les pourcentages d'inhibition et de cytotoxicité sont alors calculés et représentés pour déterminer la concentration inhibitrice à 50 % de l'effet maximal (CI_{50}) et la concentration nécessaire pour tuer 50 % des cellules (CC_{50}).

Pour WDO-217, l' $IC_{50} = 8,5 \pm 0,2 \mu$ M (Figure 47), indiquant qu'une concentration de 8,5 μ M de WDO-217 permet de réduire de moitié de l'infection par les PPs. Comme on le voit dans la Figure 47, cette molécule ne présente aucune cytotoxicité dans les concentrations testées (maximum 150 μ M).

La CI_{50} expérimentale est du même ordre de grandeur que celle obtenue par les auteurs de l'étude sur cette molécule (CI_{50} variait de 0,9 à 1,7 μ M en fonction de la souche virale étudiée (Vercruyse et al., 2012)). Ces résultats valident donc l'utilisation du test pour le criblage de l'ensemble des molécules envoyées par les partenaires du projet THINPAD. Il est important de

Manuel PIRES

noter qu'un contrôle positif utilisant 0,5 μM d'AZT (AZidoThymidine, inhibiteur nucléosidique de la transcription inverse dont la CI_{50} varie de 0,002 à 2 μM en fonction de la souche virale) sera utilisé pour vérifier l'efficacité du test pour chaque plaque.

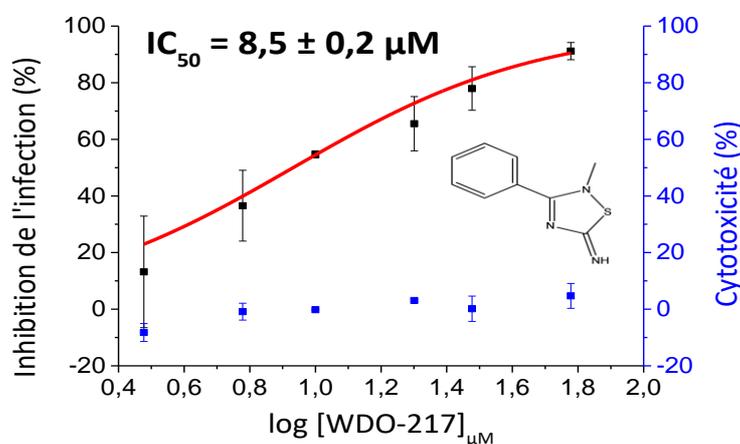


Figure 47 : Validation du test PSEUDO par détermination de l'effet antiviral et de la toxicité du WDO-217. L'inhibition de l'infection est calculée en fonction de la diminution de l'activité luciférase en présence de la molécule après infection par des PPs-lucif (test PSEUDO). La cytotoxicité est mesurée grâce au test MTT. L'ajustement des données par l'Équation 2 permettant la détermination de la concentration inhibitrice à 50 % (CI_{50}) est représenté par la ligne rouge. La structure de la molécule testée est représentée en insert. Les valeurs représentent la moyenne de trois expériences ($N = 3$) dans lesquelles chaque concentration a été testée six fois ($n = 6$).

2. Criblage des molécules THINPAD

Pour des raisons de confidentialité, aucune structure des molécules du projet européen THINPAD ne sera divulguée dans ce manuscrit. À chaque molécule d'intérêt sera attribuée une couleur afin de les nommer. Un brevet est en cours de dépôt afin de protéger les découvertes faites dans ce projet. Une publication est également en cours de rédaction.

Le criblage s'est déroulé en plusieurs étapes, comme indiqué précédemment. Dans un premier temps, les composés sélectionnés *in silico* comme potentiels compétiteurs des ANs pour la liaison à NCp7 ont été testés pour leur capacité à inhiber *in vitro* l'activité chaperonne de la NCp7 (test NCinh) ou son interaction avec les ANs (test EMSA). Les molécules ainsi sélectionnées ont, dans un second temps, été testées dans notre test PSEUDO. Les résultats obtenus dans ces tests ont permis aux partenaires 1 et 4 du projet d'optimiser la structure de ces molécules afin d'améliorer leurs activités, leurs solubilités, ou de limiter leurs toxicités (démarche « hit to lead »).

Manuel PIRES

a) Validation de l'effet antiviral de molécules issues du criblage *in silico* et *in vitro*

La première étape de la sélection a consisté en un criblage *in silico* de deux banques de molécules. Au total, plus de 11 millions de molécules ont été criblées pour leur aptitude à interagir avec la poche hydrophobe de NCp7. Parmi elles, 245 ont été retenues pour être testées *in vitro* dans les tests NCinh et EMSA (Figure 35). La majorité des molécules (226) n'ont présenté aucun effet ou un effet en de ça du seuil fixé (75% d'inhibition à 100 μ M) et leur parcours s'est arrêté là. Par conséquent, seules 19 ont été retenues pour passer le test PSEUDO. À ces 19 molécules, ont été rajoutés les 34 faux positifs du test NCinh (20 molécules fluorescentes + 10 molécules interagissant avec la rhodamine). Parmi ces 53 molécules, seules quatre ont montré un effet antiviral. Ainsi, à titre d'exemple, la molécule VERTE présente une CI_{50} de 4,9 μ M (Figure 48) et semble non-cytotoxique jusqu'à 30 μ M. L'effet observé est donc bien dû à son effet antiviral et non à sa cytotoxicité. À l'inverse, la molécule ORANGE atteint son maximum d'efficacité à 60 μ M alors qu'à cette concentration la toxicité est déjà de 40 %. La CI_{50} de 14,5 μ M est donc surévaluée. Pour cette molécule, la CC_{50} est de 85 μ M (Figure 48).

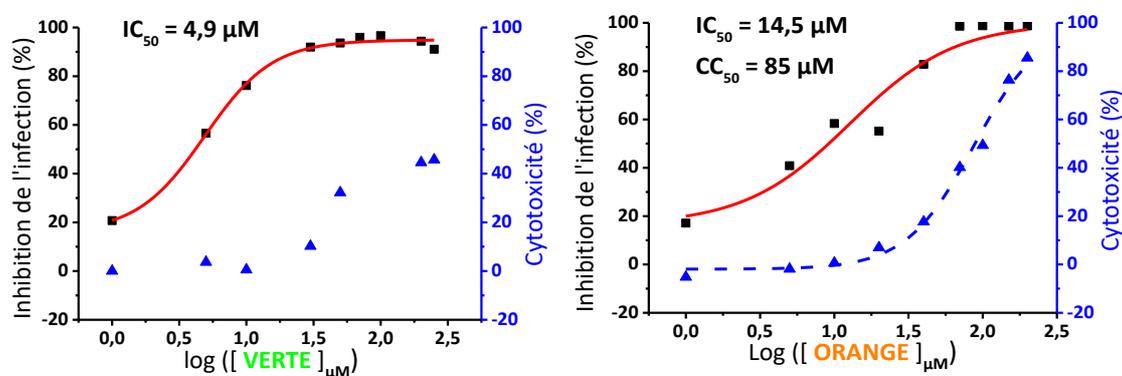


Figure 48 : Application du test PSEUDO à deux molécules issues du criblage. Pour chaque molécule, la toxicité (bleu) et le pourcentage d'inhibition de l'infection (rouge) ont été mesurés. La concentration inhibitrice à 50 % (CI_{50}) a été déterminée comme indiqué dans la figure 47. Les analyses ont été réalisées en sextuplicats ($n = 6$).

La plupart des molécules interagissant avec la rhodamine, qui appartiennent toutes à la même famille, comme par exemple la molécule ROUGE, présentent un effet antiviral à faible concentration. Ainsi, ROUGE possède une CI_{50} de 540 nM avec une toxicité de seulement 40 % à 100 mM (Figure 49). Cette famille de molécules est très intéressante puisque la toxicité n'apparaît qu'à des concentrations plus de cent fois supérieures à la CI_{50} .

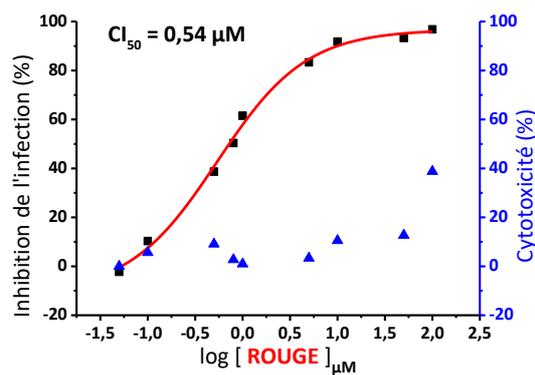


Figure 49 : Application du test PSEUDO à une molécule interagissant avec la rhodamine. La toxicité (bleu) et le pourcentage d'inhibition de l'infection (rouge) ont été mesurés. La concentration inhibitrice à 50 % (CI_{50}) a été déterminée comme indiqué dans la figure 47. Les analyses ont été réalisées en sextuplicats ($n = 6$).

Au total, douze molécules se sont avérées positives dans le test PSEUDO, montrant une capacité à inhiber les phases précoces du cycle viral du VIH-1 avec une CI_{50} inférieure à 50 μM . Les résultats du test PSEUDO, conjointement aux résultats obtenus par les autres partenaires, ont permis de distinguer cinq familles de molécules qui ont été modifiées structurellement afin d'améliorer leurs structures.

b) Optimisation des molécules antivirales : « Hit to lead ».

Après avoir sélectionné des molécules caractérisées par cinq squelettes structuraux différents, l'objectif a été d'améliorer les propriétés de ces molécules en insérant ou modifiant les substituants ou les groupes fonctionnels de ces molécules. Ce travail a été réalisé par les partenaires 1 (Université de Sienna) et 4 (IRBM), qui sans toucher au squelette de la molécule, ont modifié les structures latérales droites et gauches des diverses molécules (**Figure 50**). Cette étape, appelée relation structure-activité (SAR) a permis d'atteindre plusieurs objectifs :

- Améliorer **l'activité antivirale** de la molécule : il a été possible, au sein d'une des familles de molécules, de réduire la CI_{50} de 170 μM à 1,5 μM .
- Améliorer **la solubilité** de la molécule : la plupart des molécules n'étant soluble que dans le DMSO, il est important d'optimiser leur structure pour les rendre solubles dans un solvant mieux toléré par l'organisme et idéalement dans l'eau.
- Améliorer **la stabilité** de la molécule : afin de conserver une activité antivirale importante et stable au cours du temps, la molécule ne doit pas être métabolisée trop rapidement par l'organisme.

- Améliorer la **sélectivité** : afin de cibler uniquement la protéine de Nucléocapside et donc limiter les effets secondaires.
- Comprendre le **mécanisme d'action** de la molécule : si en changeant une partie de la molécule, l'activité antivirale est totalement ou partiellement perdue, ceci implique que la structure initiale est importante pour le mécanisme d'action.

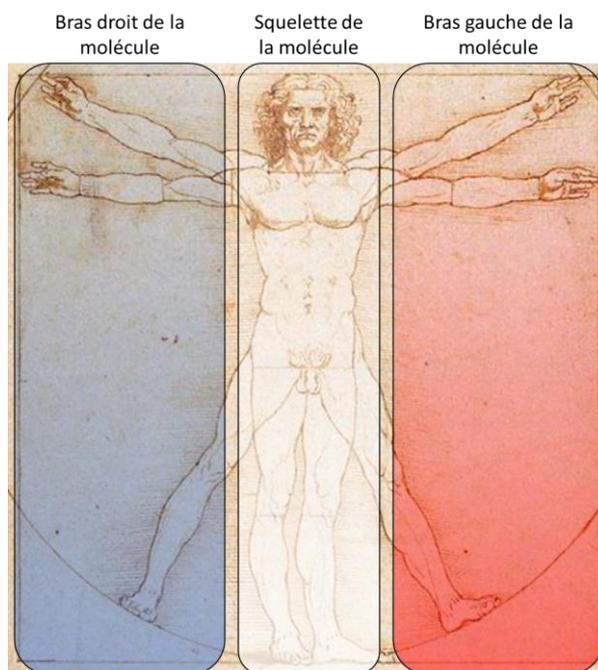


Figure 50 : Relation structure-activité. Sans toucher au squelette de la molécule (ici en blanc), les partenaires 1 et 4 de THINPAD ont modifié les structures latérales droite (en bleu) ou gauche (en rouge) afin d'améliorer l'activité antivirale de la molécule.

Au total, plus de 400 molécules ont été créées en modifiant la structure des cinq familles de molécules sélectionnées dans le premier criblage. Comme précédemment, ces molécules ont été testées dans le test NCI_{inh} puis, pour les plus intéressantes, dans le test PSEUDO (**Figure 51 A**). Dans ce graphique, on remarque que les investigations se sont principalement concentrées sur les familles II et III. Étant dans une phase d'amélioration des molécules, le seuil de sélection est cette fois une CI₅₀ inférieure à 100 µM dans le test NCI_{inh} et une CI₅₀ inférieure à 10 µM dans le test PSEUDO et les autres tests cellulaires. Concernant les familles I, IV et V, les modifications pouvant être apportées étant limitées, peu d'amélioration a été observé. De plus, dans ces familles, aucune molécule n'a présenté d'effet antiviral dans le test PSEUDO (**Figure 51 B**). Comme la plupart des modifications apportées dans la famille II se sont avérées délétères pour l'activité dans les tests *in vitro* (NCI_{inh} et EMSA), peu de molécules ont été testées *in cellulo*. En revanche, pour la famille III, les modifications ont permis d'obtenir un grand nombre de molécules actives dans le test NCI_{inh} qui dans le test PSEUDO (**Figure 51 B**)

Manuel PIRES

ont montré une amélioration notable de leur activité antivirale. Plus de 30 molécules ont donné une CI_{50} inférieure à $10 \mu M$ ($n = 2$).

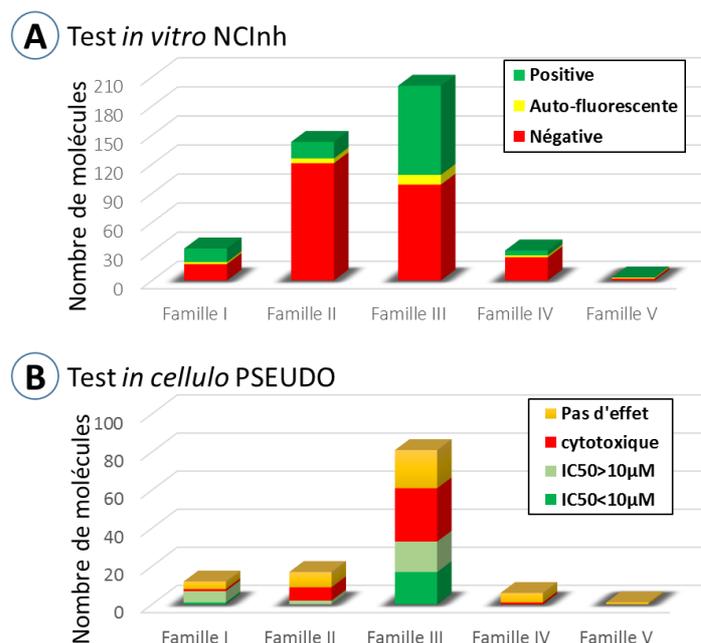


Figure 51 : Bilan des molécules testées lors de la phase d'optimisation de la structure. Les molécules sont classées par famille et par activité. (A) Plus de 400 molécules ont été testées dans le test NCIH et parmi elles (B) plus de 100 l'ont été dans les tests cellulaires et notamment le test PSEUDO.

Conjointement au test PSEUDO, deux autres tests d'activité antivirale ainsi que des tests de toxicité des molécules sur les cellules cibles du virus (lymphocytes, macrophages...) ont été réalisés par les autres partenaires. La sélection des meilleures molécules (leads) a été basée sur la capacité des molécules à inhiber l'infection virale, mais également à ne pas être cytotoxique, à être soluble et à présenter une structure chimique facilement synthétisable et modifiable. Dans certains cas, des molécules ont été sélectionnées malgré l'absence d'activité dans les tests PSEUDO ou MonoCycle. Ces molécules présentent une activité antivirale importante dans les tests BiCycle et RevCycle (CI_{50} de l'ordre du nanomolaire pour les plus intéressantes) suggérant qu'elles agissent plutôt sur les phases tardives du cycle viral. Afin d'étudier plus précisément les mécanismes d'action des molécules finales sélectionnées, plusieurs tests ont été mis en place.

3. Mécanismes d'action des hits

Notre laboratoire (ainsi que le partenaire 1) a utilisé plusieurs tests *in vitro* et *in cellulo* pour caractériser les mécanismes d'action des leads :

- **Éjection de zinc** : l'absence de l'éjection des zincs par les inhibiteurs de nucléocapside sélectionnés par le THINPAD est une condition indispensable pour le développement d'une nouvelle molécule antivirale. Nous avons, en collaboration avec le Dr Sarah Sanglier-Cianferani, testé par spectrométrie de masse l'aptitude des leads à éjecter le zinc (Vercruysse et al., 2012). Seules les molécules n'éjectant pas le zinc sont sélectionnées pour d'autres tests complémentaires.
- **Effet sur la transcription inverse** : un test *in vitro* développé dans notre laboratoire permet de vérifier l'activité d'inhibiteurs sur cette étape du cycle viral. La NCp7 favorisant cette étape, il est également possible de vérifier l'effet des molécules ciblant la NCp7 sur la transcription inverse (Sharma et al., 2016).
- **Effet sur l'intégration** : l'utilisation d'un kit commercial (HIV-1 Integrase assay kit, Xpress Bio Co, Réf. EZ-1700) permet d'évaluer la capacité des molécules à inhiber directement cette étape du cycle viral (pas de NCp7 dans ce test).
- **Quantification des produits 2-LTR** : ce test cellulaire, permet de quantifier par PCR quantitative les produits linéaires de transcription inverse, l'ADN intégré et enfin l'ADN circulaire non-intégré. La proportion de chacune de ces formes varie lorsqu'un défaut est observé dans les étapes de transcription inverse ou d'intégration. La comparaison du niveau relatif de chaque produit permet alors de déterminer quelle étape du cycle viral est ciblée par les molécules (Butler et al., 2001).

En parallèle de ces tests développés par nos collaborateurs, j'ai également mis au point deux tests additionnels. Il s'agit d'un test cellulaire d'ajout de molécules à différents temps post infection (TOAA) et d'un test d'assemblage *in cellulo* visant à déterminer si les molécules agissent également sur les phases tardives de l'infection.

a) Test d'ajout des molécules à différents temps post-infection (TOAA)

Ce test permet de déterminer quelle étape du cycle viral est ciblée par les molécules testées. Sachant que la fusion membranaire a lieu à H0, la transcription inverse vers H3-4, l'intégration

Manuel PIRES

vers H6, l'assemblage et le bourgeonnement à partir de H16, on peut déterminer quelle étape est ciblée en déterminant jusqu'à quel temps après infection, l'ajout de la molécule permet de visualiser son activité antivirale (**Figure 52 A**). L'utilisation des PPs nous limite aux étapes précoces du cycle (post-entrée jusqu'à l'intégration). Le test a été validé en utilisant des molécules de cible connue (AZT ciblant la transcription inverse qui agit jusqu'à 3 h post infection (p.i.) et Raltegravir ciblant l'intégration qui agit jusqu'à 6 h p.i.). Un des facteurs limitants de ce test est la concentration à utiliser. En effet, afin d'observer l'étape ciblée par la molécule, le test doit être effectué avec une concentration qui inhibe 100 % de l'infection virale. Pour cela, nous avons fixé la concentration à dix fois la CI_{50} qui, pour la plupart des molécules testées, correspond à une concentration à laquelle leur cytotoxicité empêche l'utilisation de ce test.

Du fait de l'absence de synchronisation des cellules et de la variabilité de la population virale, tous les virus d'un puits n'infectent pas les cellules au même instant, entraînant une perte progressive de l'activité d'une molécule testée. Pour l'AZT, la perte d'effet est observée dès 2 h p.i. et jusqu'à 10 h p.i. Si on regarde par exemple, entre 2 h et 3 h p.i. on passe de 100 % d'inhibition de l'infection à 90 %. Cette perte d'effet indique que 10 % des infections ont déjà passé l'étape de transcription inverse après 3 h d'infection et ne sont donc plus sensibles à l'AZT. On peut ainsi déterminer le temps nécessaire pour perdre 50 % de l'effet d'une molécule (TI_{50}). Pour l'AZT, il est de 7,7 h p.i. Pour le Raltegravir, inhibiteur d'intégration, cette perte d'effet apparaît à partir de 6 h p.i. avec une TI_{50} de 12,2 h p.i.

Nous avons ensuite testé le WDO-217, un inhibiteur de NC éjecteur de zinc (Vercruysse et al., 2012). L'infection commence à perdre sa sensibilité au WDO-217, à des temps aussi courts qu'une heure post-infection (**Figure 52**). WDO-217 agit donc avant même l'AZT ($TI_{50} = 2,6$ h), en accord avec son action virucide sur les particules virales avant leur entrée dans la cellule et sur la transcription inverse (Vercruysse et al., 2012). L'utilisation de ces différentes molécules a permis de valider notre test et de montrer qu'il est capable de discriminer des molécules ciblant différentes étapes des phases précoces du cycle viral. Nous avons essayé de diminuer le temps d'infection en ne mettant les PPs en contact avec les cellules que pendant un temps limité (2 heures), espérant rendre la chute de l'activité des molécules moins étalée dans le temps mais les résultats ont montré une plus grande variabilité nous poussant à conserver le protocole initial.

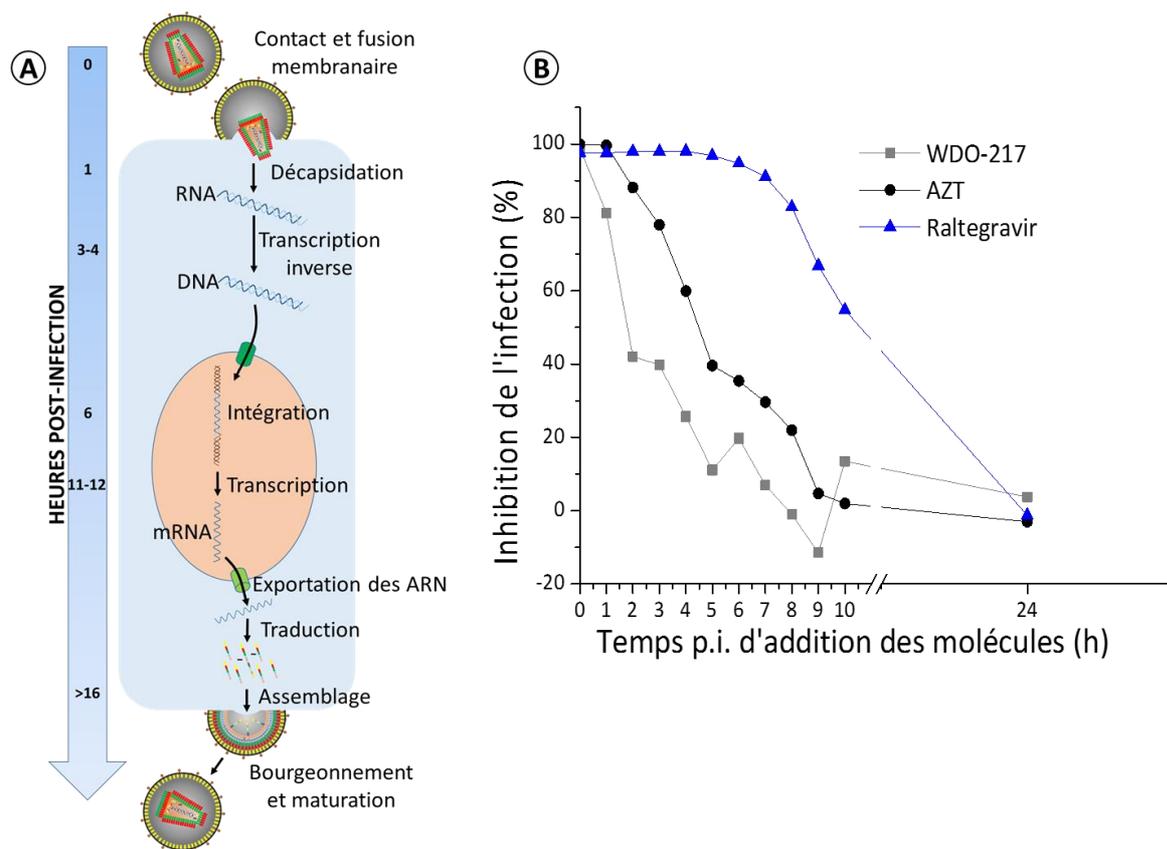


Figure 52 : Description et validation du test d'ajout des molécules à différents temps post-infection (TOAA). (A). Le cycle viral présente une chronologie bien définie. La perte d'efficacité d'un inhibiteur signifie qu'une partie de l'infection est devenue insensible à celui-ci. (B) Les résultats de TOAA montrent que l'AZT ciblant la transcription inverse perd de son activité dès 3 h post infection (p.i.) montrant qu'une partie des PPs a déjà terminé leur transcription inverse 3 h p.i. Le Raltegravir cible l'intégration et agit jusqu'à 6 h p.i. Pour le WDO-217, son activité antivirale est affectée dès la première heure de décalage de l'infection. Les analyses ont été réalisées en quadruplicats ($n = 4$).

Plusieurs molécules présentant une CI_{50} de l'ordre du micro-molaire (μM) ont été soumises à ce test. J'ai choisi de présenter seulement les molécules **VERTE** et **ROUGE**, dont la détermination de l' IC_{50} a été présentée en **Figure 48** et **Figure 49**. La molécule **ROUGE**, testée à $25 \mu M$, présente un TI_{50} de 12,2 h et une cinétique de perte d'effet similaire à celle du Raltegravir, indiquant que cette molécule inhibe l'étape d'intégration (**Figure 53**). Pour la molécule **VERTE**, l'identification de l'étape ciblée est plus complexe. En effet, comme pour l'AZT, la molécule **VERTE** commence à perdre son efficacité 3 h p.i. mais cette perte est par la suite plus lente que pour l'AZT avec un TI_{50} de 10,7 h. Ainsi, la molécule semble impacter la transcription inverse à une étape plus tardive que l'AZT comme par exemple lors des sauts de brins. On ne peut toutefois pas exclure que cette molécule agisse sur la structure du PIC et son entrée dans le noyau (**Figure 53**).

Manuel PIRES

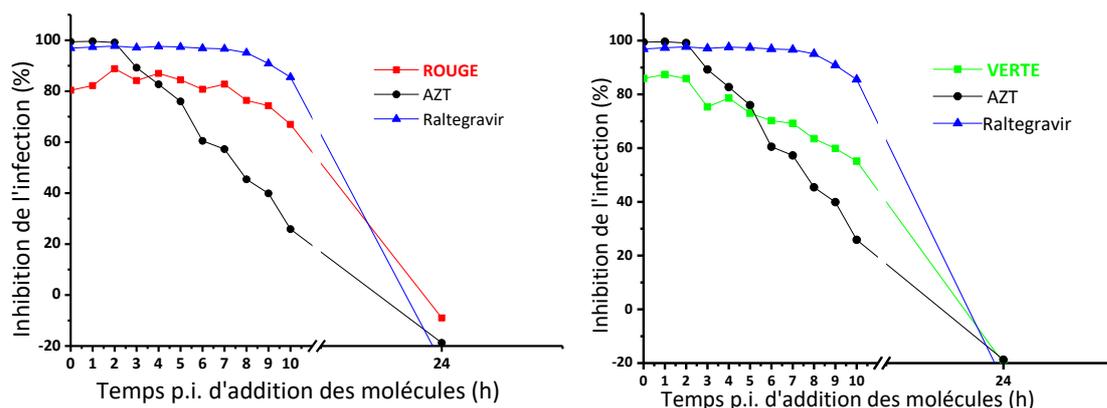


Figure 53 : Exemples de détermination du mécanisme d'action de deux molécules par TOAA. Deux molécules issues de la sélection du THINPAD ont été testées dans le test TOAA et les cinétiques de décroissance de l'effet antiviral ont été comparées à celles de l'AZT et du Raltegravir. La molécule **ROUGE** semble être un inhibiteur de l'intégration (A) alors que la molécule **VERTE** semble inhiber la transcription inverse (B). Les analyses ont été réalisées en quadruplicats ($n = 4$).

L'inconvénient principal de ce test est que pour des raisons de sensibilité, les molécules doivent être utilisées à très forte concentration (10 fois la CI_{50}). Cette limitation nous a empêchés de tester bon nombre de molécules, car les molécules sont très souvent toxiques à cette concentration.

Comme mentionné plus haut, certaines molécules sont inactives dans le test PSEUDO, mais présentent une forte activité dans les BiCycle et RevCycle suggérant une action dans les phases tardives. Nous avons donc développé des tests de détermination du mécanisme d'action ciblant les phases tardives dont l'assemblage.

Le test PSEUDO permet d'étudier l'effet des inhibiteurs ciblant la nucléocapside sur les étapes précoces du cycle viral uniquement. Or NC, en tant que domaine de la polyprotéine Gag, est impliquée dans les étapes tardives du cycle viral du VIH-1 et notamment l'assemblage. C'est pourquoi, afin de tester l'effet des inhibiteurs sur cette phase tardive, nous avons développé deux tests utilisant soit la technique d'imagerie FRET-FLIM, soit le Western blot. L'absence d'inhibiteur d'assemblage sur le marché a rendu le développement de ces tests difficile. Nous avons cherché à nous procurer des molécules récemment décrites comme inhibitrices de l'assemblage mais sans succès (Lemke et al., 2012).

b) Tests d'assemblage par FRET-FLIM

Des travaux récents au laboratoire ont montré que la multimérisation de la protéine Gag lors de l'assemblage peut être suivie par FRET-FLIM en surexprimant des polyprotéines Gag marquées par des protéines fluorescentes (eGFP ou mCherry) insérées entre les domaines MA

Manuel PIRES

et CA (El Meshri et al., 2015). Un transfert d'énergie par résonance de type Förster (FRET) a lieu entre le fluorophore donneur (eGFP) et l'accepteur (mCherry) lorsque ces derniers sont proches (moins de 10 nm), ce qui est le cas lors de la multimérisation de Gag. Ce transfert d'énergie est mesuré par la diminution du temps de vie du fluorophore donneur (FLIM). La multimérisation de protéines Gag marquées par eGFP et mCherry est donc observable par la diminution du temps de vie de l'eGFP. Des études réalisées au laboratoire par Salah El Meshri dans des cellules HeLa exprimant, après transfection transitoire, les protéines Gag recombinantes en présence d'une certaine proportion de Gag non marquée (protéine Gag sauvage) ont montré que l'assemblage et le bourgeonnement de particules virales étaient peu affectés par la présence de ces protéines fluorescentes. De plus, cette technique peut également être utilisée pour observer des défauts d'assemblage comme c'est le cas pour la protéine Gag délétée du domaine NC (**Figure 54**, El Meshri et al., 2015). En fait, les déclins de Gag-eGFP à chaque pixel de l'image sont analysés en deux composantes, selon l'**équation 3**. La première composante de temps de vie court (τ_1) et de population (α_1) décrit la population de Gag-eGFP multimérique où Gag-eGFP transfère son énergie à Gag-mCherry. La seconde population (α_2) est sous forme monomérique avec un temps de vie (τ_2) inchangé par rapport à celui de Gag-eGFP exprimée seule.

L'étude d'El Meshri et al, 2015 a montré une augmentation du FRET et donc de la multimérisation de Gag au fur et à mesure de l'accumulation de cette dernière à la membrane (comparer GagWT entre 12 h et 24 h, **Figure 54 E et F**). Ce transfert de fluorescence est fortement diminué avec le mutant Gag Δ NC montrant un défaut de multimérisation qui n'est pas récupéré par une expression plus longue de Gag Δ NC (comparer 12h et 24h pour Gag Δ NC, **Figure 54**). Ces résultats montrent que la suppression du domaine NC de Gag perturbe fortement l'assemblage de Gag.

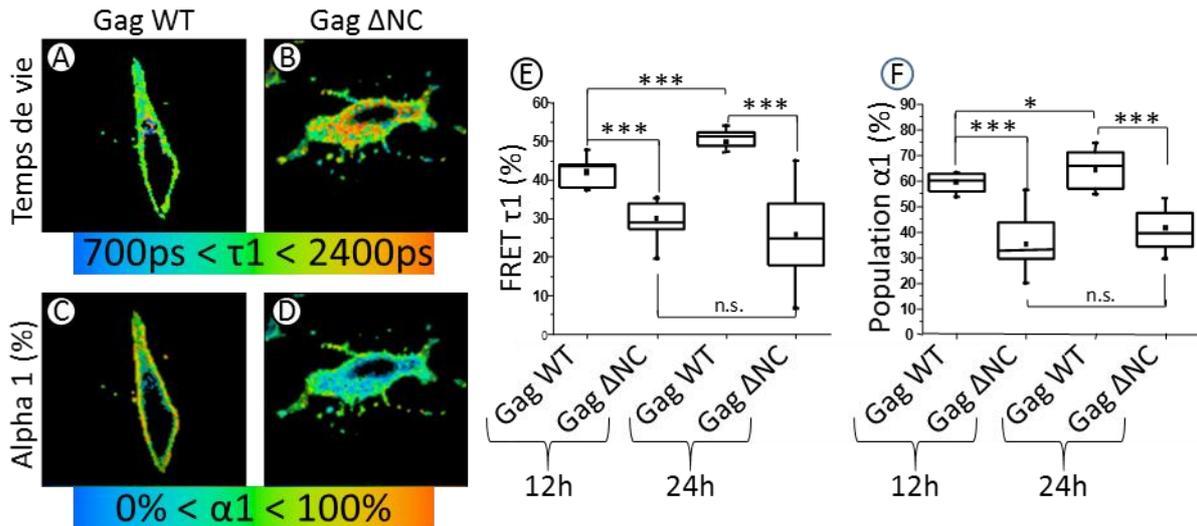


Figure 54 : Suivi de l'assemblage de Gag et GagΔNC par FRET-FLIM (extrait de El Meshri et al., 2015). Des cellules HeLa ont été transfectées par un mélange de plasmides codant pour (i) Gag/Gag-eGFP/Gag-mCherry ou (ii) GagΔNC/GagΔNC-eGFP/GagΔNC-mCherry avec un ratio de 0,7/0,1/0,2. (A) et (B) montrent la distribution du temps de vie court (τ_1) de Gag-eGFP après 12 h de transfection ; (C) et (D) montrent le pourcentage de la population multimérique de Gag-eGFP (α_1) après 12 h de transfection. (E) Pourcentage de transfert d'énergie (FRET) des populations de Gag-eGFP et GagΔNC-eGFP multimériques (τ_1). (F) Pourcentage de ces populations (α_1). Ces graphiques ont été obtenus à partir de plus de 150 cellules issues de trois expériences indépendantes. Test-t de Student * : $p < 0,005$ et *** : $p < 0,001$. (El Meshri et al., 2015).

Ces expériences nous ont servi de base pour notre test. L'idée de ce test est que l'inhibition de la multimérisation par une molécule anti-NC devrait entraîner une diminution de FRET comme lors de la délétion du domaine NC de Gag.

Plusieurs molécules ont été testées en prenant GagΔNC comme témoin positif d'un défaut d'assemblage. Malheureusement, de nombreuses molécules présentaient une auto-fluorescence avec un temps de vie très court (en général inférieur à 0,5 ns) empêchant l'étude de leur effet sur l'assemblage. Seule la molécule **ROUGE** a pu être testée. Elle ne semble pas induire d'inhibition de l'assemblage (Figure 55), puisqu'aucune diminution du FRET ou de la population de Gag-eGFP multimérique (α_1) n'a été observée. Il est à noter que les résultats obtenus pour GagΔNC lors de mes manipulations, sont plus nuancés que dans ceux de El Meshri et al (comparaison de τ_1 et α_1 entre la Figure 54 et la Figure 55). Ceci peut être lié au fait que mes expériences portaient sur un faible nombre de cellules. De plus, l'absence d'inhibiteur d'assemblage de référence nous a empêché de valider le test.

Manuel PIRES

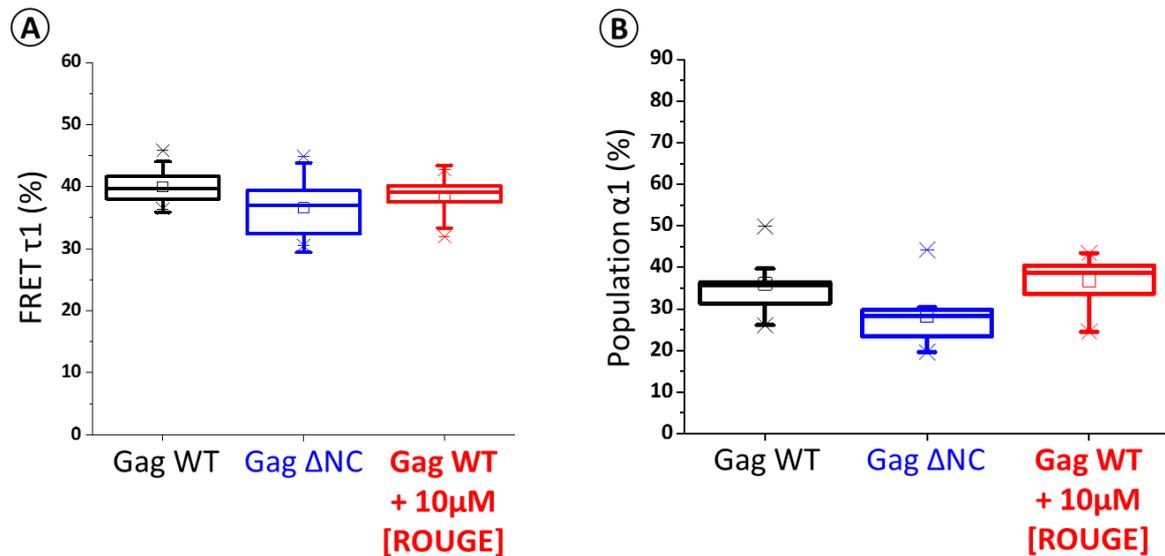


Figure 55 : Effet inhibiteur de l'assemblage de la molécule ROUGE suivi par FRET-FLIM. Des cellules HeLa ont été transfectées par un mélange de plasmides codant pour Gag/Gag-eGFP/Gag-mCherry en absence (Noir) ou en présence de molécule ROUGE (10 μM, environ 10 fois la CI_{50}) (rouge) ou codant GagΔNC/GagΔNC-eGFP/GagΔNC-mCherry (Bleu) avec un ratio 0,7/0,1/0,2. La quantité de DMSO a été normalisée pour tous les échantillons. Après 12 h de transfection, les cellules préalablement fixées par du PFA 4 % ont été observées en FRET-FLIM. (A). Représentation du temps de vie court (τ_1) de l'eGFP. (B) Pourcentage de la population multimérique de Gag (α_1). Ces graphiques ont été réalisés à partir des données de dix cellules issues d'une expérience.

Nous avons utilisé un second test pour étudier l'effet des molécules sur l'assemblage (von Schwedler et al., 2003). Dans ce test, 12 h ou 24 h après la transfection d'un plasmide codant pour la protéine Gag en présence de différentes concentrations de molécules, nous avons analysé sur gel de polyacrylamide couplé à un Western blot anti-p24 les contenus protéiques des particules virales récoltées dans le surnageant cellulaire et le lysat cellulaire. En théorie, en présence d'un inhibiteur d'assemblage, la quantité de Gag dans le cytoplasme devrait rester inchangée ou être augmentée alors que celle dans le surnageant cellulaire devrait diminuer. Les premiers résultats pour un des inhibiteurs de NC montrent l'absence d'effet sur l'assemblage. Pour valider ce test, il sera à nouveau indispensable d'utiliser un inhibiteur de référence.

À terme, nous pourrions utiliser ce test par western blot en première intention pour déterminer si les molécules candidates ont la capacité d'inhiber la production de PPs pour, dans un second temps, pour les positives, tester leur capacité à inhiber la multimérisation de Gag dans le test de FRET-FLIM.

B. Molécules antivirales testées dans le test PSEUDO hors projet THINPAD

1. Effet antiviral de duplex bactéricides/antiviraux

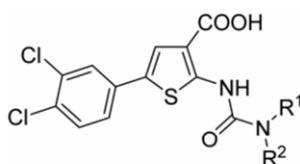
En collaboration avec l'Université de Saarland à Saarbrücken en Allemagne, nous avons également testé l'activité antivirale et la toxicité de molécules développées pour inhiber de concert la réplication du VIH-1 et une infection bactérienne. Le développement de ces molécules a été motivé par le fait que les patients contaminés par le VIH-1 contractent dix-huit fois plus fréquemment que la population générale une infection par des bactéries (*Staphylococcus aureus*) résistantes à un traitement par la méthiciline (MRSA) (Crum-Cianflone et al., 2007). De plus, des interactions médicamenteuses entre les molécules anti-VIH-1 et celles permettant le traitement des infections opportunistes peuvent apparaître chez les personnes infectées par le VIH-1 et diminuer l'effet des deux traitements (Piscitelli and Gallicano, 2001).

L'enzyme la plus fréquemment ciblée par les traitements anti-VIH-1 est la transcriptase inverse (RT). Les agents antibactériens quant à eux ciblent souvent l'ARN polymérase (RNAP), responsable de la transcription des ARNs messagers bactériens. Des similitudes fonctionnelles et structurales entre ces deux enzymes ont été démontrées (Szilvay et al., 2000). Ainsi, un site de liaison de la RNAP (switch region) ayant des similitudes avec le site de liaison des INNRT de la RT a été identifié (Mukhopadhyay et al., 2008). Sur la base de ces similitudes, l'idée de ce travail publié en 2016 (Elgaher et al., 2016) a été de concevoir une molécule capable d'inhiber à la fois l'infection par le VIH-1 en ciblant la RT et l'infection par les MRSA en ciblant la « switch region » de RNAP.

Dans un premier temps, les molécules testées dans le test PSEUDO n'ont pas montré d'activité. Le caractère fortement lipophile de ces molécules favorise probablement l'interaction avec des protéines du milieu de culture, ce qui diminue leur internalisation cellulaire (faible biodisponibilité). Nous avons donc adapté le protocole du test PSEUDO afin d'analyser leurs activités antivirales. Le milieu de culture DMEM supplémenté par 10 % de Sérum de Veau Fœtal (SVF) a été remplacé par du tampon HBSS (Hank's Balanced Salt Solution), pour éviter la présence de fortes concentrations de protéines lors de l'ajout de molécules sur les cellules tout en maintenant un pH et une balance osmotique optimaux pour les cellules. Les PPs, habituellement diluées dans un milieu de culture complet sont ici

Manuel PIRES

mélangées dans un milieu DMEM sans sérum. Les solutions obtenues en mélangeant volume à volume les solutions de PPs et de molécules ont été mises au contact des cellules pendant 2 h, un temps jugé suffisant pour permettre leur entrée dans la cellule. Deux heures après infection et traitement, les cellules sont replacées dans un milieu DMEM complet contenant du sérum afin de minimiser la toxicité. Comme pour le test PSEUDO, un contrôle en absence de molécule est réalisé en maintenant la concentration de DMSO constante dans tous les essais. Grâce aux adaptations de protocole, des CI_{50} ont pu être extrapolées (**Tableau 11**). Les CI_{50} variant de 20 à 35 μM sont largement supérieures à celles obtenues *in vitro* (0,1 à 1,2 μM), probablement du fait de la faible biodisponibilité de ces molécules.



#	Molécules		Inhibition de l'activité de la RT <i>in vitro</i> (CI_{50} , μM)	CI_{50} dans le test PSEUDO (μM)	% d'inhibition de la viabilité cellulaire des cellules HeLa à 25 μM (MTT)
	R1	R2			
4			0,9	27	23
11		H	1,2	20	0
12		H	0,7	Pas d'inhibition	0
13		H	0,8	35	12
17		H	0,3	30	6
18		H	0,1	30	0
19		H	0,9	Non mesuré	Non mesuré
20		H	0,6	Non mesuré	Non mesuré

Tableau 11 : Activités antibactérienne, antirétrovirale et cytotoxicité des principaux inhibiteurs de RNAP et de RT. Les molécules diffèrent principalement par leur chaîne latérale R1 (seule la molécule 4 diffère par la chaîne latérale R2). L'inhibition *in vitro* de la RT est mesurée par un test de FRET développé au laboratoire (Sharma et al., 2016). Les mesures de cytotoxicité (test MTT) et d'activité antivirales dans le test PSEUDO ont été réalisées en parallèle. Les résultats et les structures sont extraits de Elgaher et al., 2016.

2. Effet antiviral des Calixarènes

En collaboration avec le Pr Kalchenko de l'académie nationale des sciences d'Ukraine, nous avons également testé l'effet anti-VIH-1 de molécules de la famille des calixarènes. L'activité

Manuel PIRES

antivirale des calixarènes, sur des virus à enveloppes, a été mentionnée pour la première fois en 1994 (Hwang et al., 1994). Plus récemment, le rythme des recherches sur les calixarènes en tant qu'inhibiteurs du VIH-1 s'est accéléré (Luo et al., 2013, 2015, 2014). En se basant sur ces études, nous avons développé des calixarènes capables de se lier à la nucléocapside.

Le test PSEUDO a été utilisé pour tester l'activité antivirale de quatre calixarènes sur les étapes précoces du cycle viral. Comme les résultats du [tableau 12](#) l'indiquent, les molécules CIP200 et 201 ($CI_{50} = 10,3 \pm 0,3 \mu\text{M}$ et $21 \pm 3 \mu\text{M}$ respectivement) ont une activité antivirale supérieure à celles de C933 et 934 ($CI_{50} = 28 \pm 3 \mu\text{M}$ et $74 \pm 7 \mu\text{M}$ respectivement). Pour toutes les molécules, la toxicité n'est observée qu'à des concentrations très supérieures à la CI_{50} , donnant de bons index de sélectivité (SI), qui rendent possible l'étude de leurs mécanismes d'action dans le test d'ajout des molécules (TOAA).

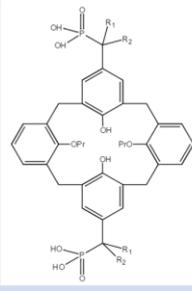
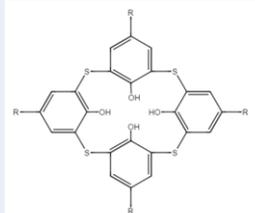
	#	Structure	CI_{50} (μM)	CC_{50} (μM)	SI
	C933	R1: H R2: PO(OH) ₂	28 ± 3	800 ± 50	28
	C934	R1: OH R2: PO(OH) ₂	74 ± 7	750 ± 70	10
	CIP200	R: SO ₃ Na	10,3 ± 0,3	> 50	> 4,8
	CIP201	R: CH ₂ ⁻ (NaHPO ₃)	21 ± 3	> 100	> 4,8

Tableau 12 : Structures et activités antivirales de molécules de la famille des calixarènes. Les concentrations inhibitrices à 50 % (CI_{50}) et les concentrations pour tuer 50 % des cellules (CC_{50}) sont déterminées à partir de la moyenne de trois expériences indépendantes et par l'utilisation des Équation 1 et Équation 2. L'index de sélectivité (SI) est calculé par le rapport CC_{50}/CI_{50} .

Il est habituellement nécessaire d'utiliser une concentration équivalente à dix fois la CI_{50} pour le TOAA. Comme C933 et C934 sont toxiques à ces concentrations, nous avons choisi de diminuer la concentration utilisée dans le TOAA en utilisant sept fois la CI_{50} pour C933 (200 μM) et 4 fois pour C934 (300 μM). Comme on le voit dans la [Figure 56](#), les deux calixarènes (C933 et C934) présentent des cinétiques de perte d'activité identiques suggérant un mécanisme d'action identique. La perte d'effet antiviral observé dès la première heure de décalage du traitement suggère que comme WDO-217, ([Figure 52](#)) ces molécules agissent

Manuel PIRES

directement sur les virus. Le fait que la perte d'activité soit progressive (entre 1 h et 12 h p.i.) suggère que ces molécules pourraient avoir une action sur la transcription inverse.

Les investigations sur ces molécules sont poursuivies afin de préciser leurs mécanismes d'actions. Le pouvoir infectieux de PPs prétraitées par C933 et C934 sera mesuré pour évaluer l'activité virucide de ces molécules. Par ailleurs, ces molécules seront aussi testées pour déterminer leur aptitude à inhiber la transcriptase inverse (Sharma et al, 2016). L'ensemble de cette étude sera soumise à publication prochainement.

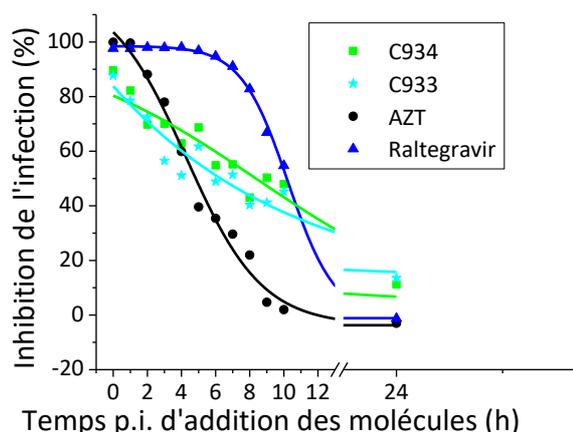


Figure 56 : Analyse du mécanisme d'action des calixarènes dans le test d'ajout des molécules à différents temps post-infection (TOAA). L'AZT (0,5 μ M) cible la transcription inverse et agit jusqu'à 3h post infection (p.i.), le Raltegravir (1 μ M) cible l'intégration et agit jusqu'à 6 h p.i. L'activité antivirale des calixarènes (200 μ M pour le C933 et 300 μ M pour le C934) est affectée dès la première heure montrant des similitudes avec l'éjecteur de zinc WDO-217.

C. Conclusions et perspectives

Le VIH-1, découvert en 1983 mais présent chez l'Homme depuis le début des années 1900 a causé 1,1 million de morts en 2015. Les thérapies actuelles pour lutter contre l'infection du VIH-1 sont puissantes et sélectives. Depuis la commercialisation de l'AZT en 1986, premier inhibiteur nucléosidique de la transcription inverse, de nombreuses recherches ont permis d'allonger la durée de vie d'un patient infecté par le VIH-1. L'apparition de la trithérapie en 1996, puis le développement de thérapies spécifiques de nouvelles cibles (Intégrase du VIH-1, entrée dans la cellule cible, maturation, ...) et enfin les innovations galéniques permettent aujourd'hui aux patients de vivre plusieurs dizaines d'années avec le VIH-1 sans entrer dans le stade SIDA tout en ayant un quotidien de moins en moins affecté par les effets secondaires des médicaments. Mais une des caractéristiques majeures du VIH-1 est qu'il peut s'adapter génétiquement à ces différents traitements, conduisant à des souches virales résistantes. De

Manuel PIRES

plus, à cause du coût de ces médicaments, leur accès reste difficile dans les pays les plus pauvres, où l'infection par le VIH-1 reste mal contrôlée. Il est donc important de développer des molécules dirigées contre des cibles thérapeutiques, générant peu ou pas de résistance. La protéine virale NCp7 est ainsi une candidate très intéressante car elle est impliquée dans plusieurs étapes du cycle viral et sa séquence est très conservée dans les différentes souches virales. Le projet européen THINPAD, dans lequel notre laboratoire est impliqué avait pour objectif de développer une molécule ciblant spécifiquement la protéine de NCp7 pour inhiber ses différentes activités dans le cycle viral.

Dans le projet THINPAD, j'ai participé au criblage des molécules en testant l'activité antivirale de plus de 100 molécules grâce à un test développé au laboratoire, utilisant des pseudo-particules virales (PSEUDO). Afin d'étudier le mécanisme d'action des leads, j'ai utilisé plusieurs tests pour caractériser leur effet sur les phases précoces (TOAA) ou tardives (FRET-FLIM ou western blot) du cycle.

Après trois ans d'efforts, une nouvelle famille d'inhibiteurs a été développée dans le projet THINPAD. Les molécules présentant les plus grandes activités anti-NC *in vitro* et antivirales *in cellulo*, associées à une cytotoxicité cellulaire modérée ont également montré un effet antiviral équivalent sur les souches sauvages et résistantes aux traitements actuels. Les molécules les plus prometteuses sont testées *in vivo*, dans un modèle murin d'infection au VIH-1.

II. Forme compacte de Gag

Malgré la présence de plus de 2000 protéines Gag au sein de chaque particule virale, toutes capables d'interagir avec le génome, seules deux copies de ce dernier sont retrouvées au sein du virion. Le mécanisme responsable de cette sélection des ANs reste mal connu. Un autre aspect de cette sélection est la préférence de Gag pour les ARNs génomiques viraux encapsidés au dépend des ARNs cellulaires.

En effet, Gag posséderait une plus forte affinité pour l'ARN viral que pour les ARNs cellulaires pourtant présents en abondance dans le cytoplasme de la cellule infectée (Abd El-Wahab et al., 2014; Carlson et al., 2016). Cette forte affinité serait liée à la présence, sur Gag, de deux sites de liaison ayant une affinité particulière pour deux motifs retrouvés sur l'ARN viral. Le site préférentiel est constitué par le domaine NC qui fixe la séquence Ψ retrouvée dans la région 5'-UTR de l'ARN viral (Darlix et al., 2011). Il existe un second site dans la partie N-terminale de Gag (acides aminés basiques de l'extrémité N-terminale de la matrice) qui semble avoir une forte affinité pour une région de l'ARN retrouvée dans le gène *Pol* (Purohit et al., 2001). Ces deux sites ont également la capacité de se lier non spécifiquement aux ARNs. Une hypothèse avancée pour expliquer la sélection spécifique de seulement deux ARNs viraux est le repliement de Gag sous une forme compacte dite en U où les extrémités N- et C-terminales sont rapprochées par la fixation de l'ARN au niveau des domaines MA (partie N-terminale de Gag) et NC (partie C-terminale de Gag) (Datta et al., 2007a; Munro et al., 2014). Cette forme observée *in vitro* avec des formes tronquées de Gag (Gag non myristoylée et sans domaine p6), n'a jamais été décrite *in vivo*. L'objectif de cette partie de la thèse est donc de prouver l'existence de cette forme compacte dans l'environnement cellulaire mais aussi *in vitro* avec une Gag pleine longueur.

Le travail a donc été divisé en deux parties. La première partie a consisté à développer des outils afin d'observer *in cellulo* la forme compacte de Gag en microscopie de fluorescence mais aussi par des études biochimiques. Dans la seconde partie, nous avons souhaité prouver l'existence de la forme en U par des études de spectrométrie de masse à mobilisation ionique réalisées sur une protéine Gag complète purifiée au laboratoire. Enfin, dans une dernière partie, différentes techniques de microscopies pour visualiser cette forme compacte seront décrites.

A. Études in cellulo

1. Utilisation du FRET-FLIM entre eGFP et ReAsH

La forme compacte de Gag induit théoriquement le rapprochement des extrémités N- et C-terminales de la protéine lors de l'interaction avec l'ARN viral par les deux domaines MA et NC. Nous avons donc imaginé utiliser le transfert d'énergie entre deux fluorophores (FRET) situés à proximité du domaine MA pour l'un et de l'extrémité C-terminale pour l'autre. Les fluorophores les plus classiquement utilisés en fusion à des protéines exprimées par transfection transitoire dans des cellules sont l'eGFP et la mCherry.

L'inconvénient de cette stratégie est que nous augmentons très fortement la taille de la protéine étudiée. En effet, la taille de la protéine Gag (55 kDa) serait doublée si nous lui greffions l'eGFP (27 kDa) et la mCherry (28,8 kDa). Cette augmentation de taille n'est certainement pas sans conséquence d'un point de vue fonctionnel. Afin de contourner ce problème, nous avons choisi d'utiliser le marquage TC-ReAsH (Adams et al., 2002; Griffin et al., 1998). Cette approche est basée sur la forte affinité de l'arsenic (As) pour les groupements thiols retrouvés notamment sur les cystéines. Ainsi, l'insertion d'un petit motif de 6 résidus (CCPGCC dit TétraCystéines (TC)) à l'extrémité C-terminale de Gag permet le marquage de la protéine. En effet, la reconnaissance du motif par un réactif arsénié dérivé de la fluorescéine (FIAsH) ou de résosfurine (ReAsH) permet la formation de liaisons covalentes qui rendent fluorescente la protéine portant le motif TC (**Figure 57 B**). Cette approche permet de limiter les modifications structurelles de la protéine marquée puisque le motif TC est de petite taille (1 kDa). De plus, le marquage par le ReAsH ne nécessite pas de perméabilisation des cellules (contrairement aux anticorps) autorisant ainsi l'observation de cellules non fixées. Enfin, le spectre d'absorption du ReAsH se chevauchant avec le spectre d'émission de l'eGFP, il est possible d'envisager un transfert d'énergie entre les deux fluorophores (comme l'indique le losange noir dans la **Figure 57 C**). En revanche, ce marquage reste faiblement spécifique principalement à cause des nombreuses protéines contenant des cystéines qui interagissent avec le ReAsH.

Nous avons donc inséré la protéine eGFP entre les domaines MA et CA de Gag car il a été décrit que ce site d'insertion est le moins délétère pour l'activité de la protéine Gag (El Meshri et al., 2015; Hübner et al., 2007). Le motif TC, quant à lui, est fusionné en C-terminal du domaine p6 de Gag (**Figure 57 A**). La protéine qui porte à la fois l'eGFP et le motif TC

Manuel PIRES

(Gag_eGFP_TC) servira ainsi à prouver l'existence d'une forme compacte de Gag par FRET intramoléculaire entre les deux fluorophores. Les résultats de FRET obtenus avec cette construction seront comparés à ceux obtenus (FRET intermoléculaire) en co-exprimant une Gag fusionnée à l'eGFP (Gag_eGFP) et une Gag fusionnée au motif TC (Gag_TC). En effet, la protéine Gag_eGFP_TC étant capable de se multimériser, un signal de FRET intermoléculaire est également possible en plus de l'intramoléculaire.

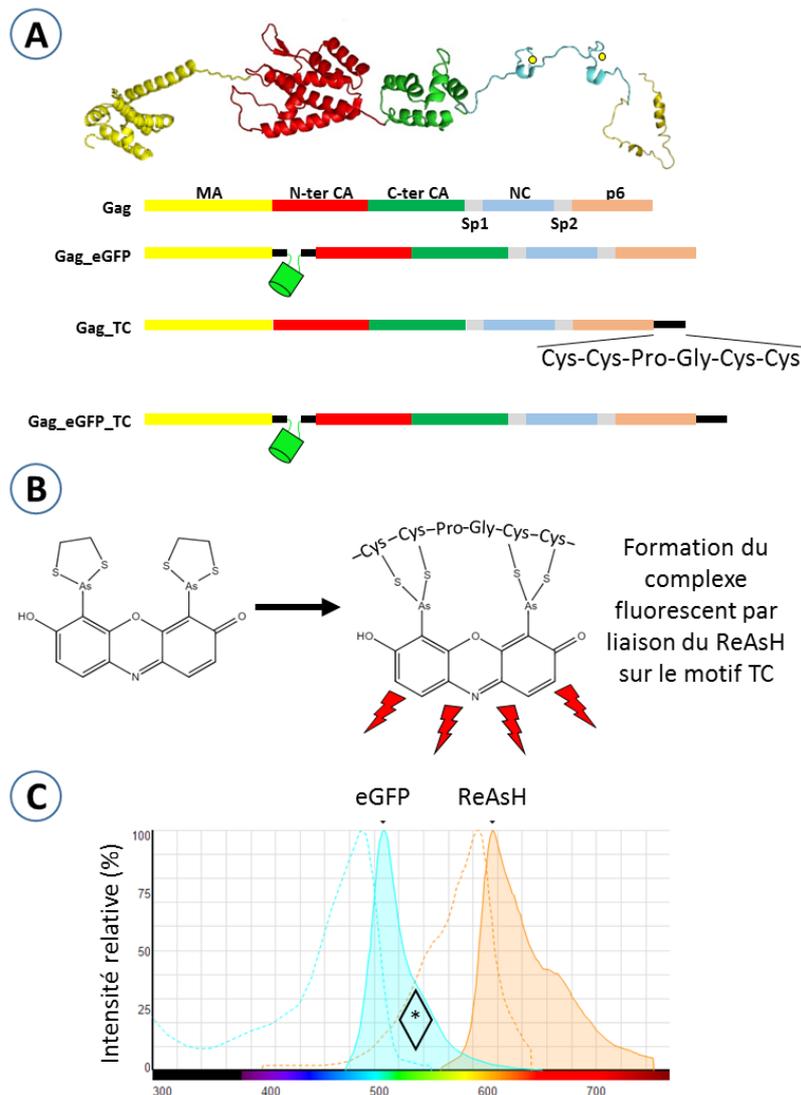


Figure 57 : Outils utilisés pour l'étude de la forme compacte de Gag en FRET-FLIM. (A) Structure tridimensionnelle des différents domaines de Gag sauvage. Représentation linéaire des différentes protéines Gag utilisées dans cette étude (jaune pour MA, rouge pour la partie N-terminale de CA, vert pour la partie C-terminale de CA, bleu pour la NC et orange pour p6). L'eGFP, représentée par un cylindre vert, est insérée entre les domaines MA et CA. Le motif tétracystéine (TC, CCPGCC) est inséré en C-terminal de la protéine. (B) Fonctionnement du marquage par le ReAsH qui est non-fluorescent lorsqu'il n'est pas lié puis devient fluorescent lorsqu'il forme une liaison covalente avec le motif TC. (C) Spectre d'absorption (en pointillé) et d'émission (spectre plein) de l'eGFP et du ReAsH montrant qu'un transfert d'énergie est possible entre ces deux fluorophores (losange noir avec l'étoile).

Manuel PIRES

Après avoir réalisé et vérifié la séquence des différentes constructions, nous avons observé la caractéristique des marquages par microscopie confocale (**Figure 58**). Sans incubation par le ReAsH aucune fluorescence n'est observée (**Figure 58 C et D**). En présence de ReAsH, un marquage non spécifique apparait dans les cellules non transfectées (résultat non montré) Ce marquage non spécifique est également observable dans les marquages des cellules transfectées (**Figure 58 A et B**) où nous observons un marquage ReAsH dans l'ensemble du tapis cellulaire alors que seules les cellules présentant une fluorescence verte (eGFP) sont positivement transfectées. Dans le cas de cellules exprimant les protéines Gag_eGFP_TC ou Gag_TC, on note la présence d'un marquage ReAsH plus intense à la membrane cellulaire reflétant l'accumulation des protéines Gag dans cet espace. L'intensité du marquage cytoplasmique reste quant à elle identique dans les cellules transfectées et non transfectées montrant la faible spécificité et sensibilité de ce type de marquage. Au niveau membranaire, il est de plus possible d'observer une co-localisation de la fluorescente verte (GFP) insérée entre les domaines MA et CA de Gag et du marquage ReAsH (**Figure 58 A et C**) validant l'aspect spécifique du marquage au ReAsH à la membrane rendue possible par l'accumulation à la membrane de la protéine Gag lors de l'assemblage. Le fort bruit de fond cytoplasmique rend impossible toute étude de co-localisation et donc toute étude de FRET dans le cytoplasme qui est pourtant le compartiment dans lequel la forme compacte est attendue étant donné qu'*in vitro* Gag ré-adapte une forme linéaire en présence de membranes.

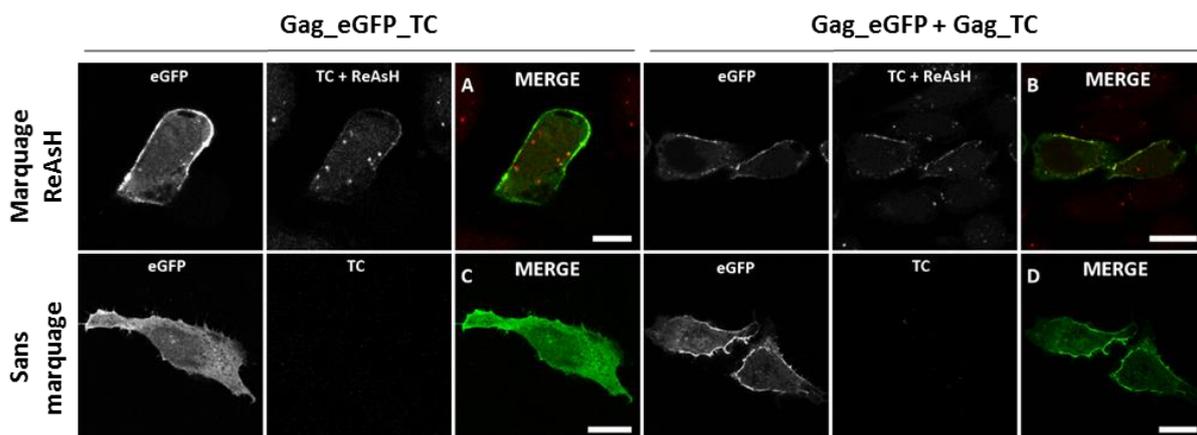


Figure 58 : Observation au microscope confocal de cellules HeLa exprimant Gag_eGFP_TC ou Gag_eGFP et Gag_TC. Des cellules HeLa ont été observées par microscopie confocale 6 h post transfection par des constructions codant Gag_eGFP_TC (1 μ g) ou co-exprimant Gag_eGFP (0,5 μ g) et Gag_TC (0,5 μ g). En présence de ReAsH (ligne supérieure), un marquage diffus apparait (TC + ReAsH) alors qu'il n'y a aucune fluorescence du motif TC sans marquage (ligne inférieure). La fluorescence de l'eGFP présente une localisation membranaire et cytoplasmique. La superposition des images (MERGE, eGFP en vert et ReAsH en rouge) montre une co-localisation membranaire de l'eGFP et de TC-ReAsH.

Manuel PIRES

Pour quantifier le FRET au niveau membranaire, nous avons utilisé le FLIM à deux photons pour mesurer le temps de vie de l'eGFP (FLIM). Comme on le voit dans la **Figure 59**, en absence de marquage ReAsH, le temps de vie de l'eGFP dans Gag_eGFP_TC (2,39 ns, carré noir, **Figure 59 C et E**) est presque identique à celui du contrôle Gag_eGFP (2,42 ns, triangle noir, **Figure 59 A et E**). L'insertion du motif TC dans la partie C-terminale de Gag n'influence donc pas le temps de vie de fluorescence de l'eGFP. Après marquage par le ReAsH, lorsque les marqueurs fluorescents sont portés par deux protéines différentes (Gag_eGFP + Gag_TC), le temps de vie de l'eGFP passe à 2,38 ns (triangle noir vs triangle rouge) indiquant une absence de FRET intermoléculaire. Lorsque les deux fluorophores sont dans la même protéine (Gag_eGFP_TC), le décalage du temps de vie est plus important (de 2,39 ns à 2,31 ns), indiquant un transfert d'énergie de 3,3 %. Cette valeur de FRET reste toutefois en dessous du seuil de 5 % requis pour valider un transfert d'énergie entre deux fluorophores en FLIM.

Malgré plusieurs essais, nous n'avons pas été en mesure d'obtenir un transfert d'énergie supérieur à 5 %. La raison principale est liée au fort bruit de fond induit par le marquage ReAsH, qui nous limite à observer les zones où le marquage est fortement concentré, c'est-à-dire au niveau des membranes. Or, d'après notre hypothèse, la forme compacte est principalement cytoplasmique. Il est donc peu probable que de fortes proportions de formes en U de Gag se retrouvent au niveau membranaire, expliquant la faible valeur de FRET. Pour limiter ce bruit de fond important lié au marquage ReAsH, nous avons décidé d'utiliser l'approche split-GFP.

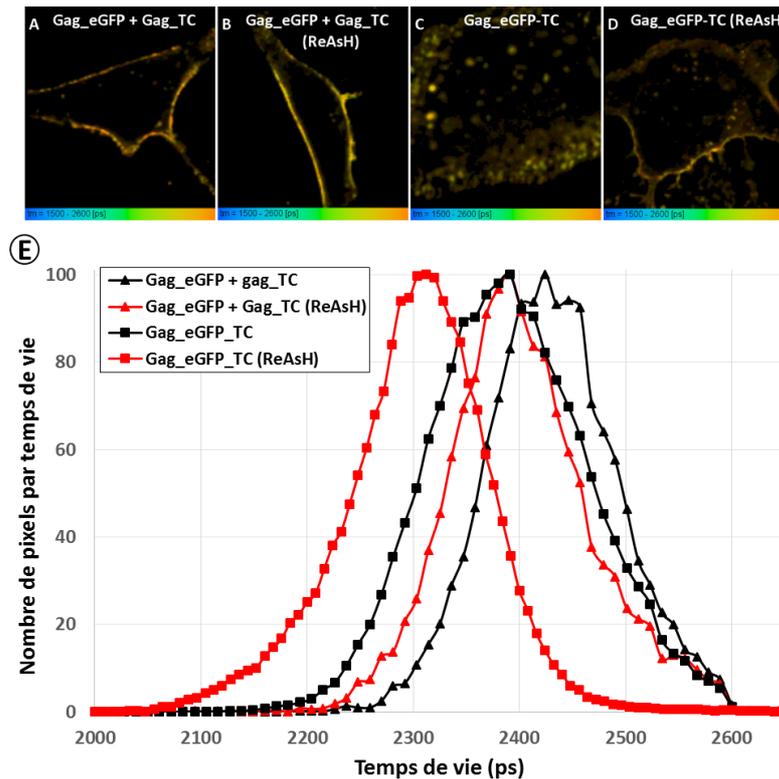


Figure 59 : Etude de la forme compacte de Gag par FRET-FLIM. Des cellules HeLa ont été transfectées pendant 6 h par des plasmides exprimant les différentes protéines indiquées puis marquées ou non par le ReAsH. Le temps de vie de l'eGFP est déterminé par l'analyse de la décroissance de fluorescence pour chaque pixel (images (A), (B), (C) et (D)). (E) Distribution normalisée des temps de vie (moyenne représentant le résultat de cinq cellules issues d'une expérience, $n = 5$). Le rouge représente les échantillons marqués par le ReAsH, alors que le noir décrit les échantillons sans marquage. L'eGFP et le motif TC sont soit dans la même protéine (carrés), soit dans deux protéines différentes (triangles).

2. Approche et utilisation de la split-GFP

Afin de s'affranchir des problèmes de marquage non spécifiques, nous avons choisi d'utiliser la technique du BiFC (Bimolecular Fluorescence Complementation), basée sur l'association de deux fragments non fluorescents qui reforment une protéine fluorescente. Nous avons ainsi choisi d'utiliser la GFP coupée en deux ou trois fragments (Cabantous et al., 2005, 2013). Le Dr Stephanie Cabantous nous a généreusement fourni les plasmides codant pour les différents fragments de split-GFP bi-partite et tri-partite. Les fragments ont été optimisés afin de perturber le moins possible la solubilité de la protéine sur laquelle ils sont insérés, mais également afin de s'auto-assembler lorsque les fragments sont suffisamment proches. La GFP (Green fluorescent Protein) est une protéine composée de 230 résidus (environ 27 kDa) formant une structure tertiaire dite en tonneau bêta. Ce tonneau est lui-même composé de 11 brins bêta qui s'associent pour former puis maintenir le chromophore de la protéine (composé de trois acides aminés : Serine 65, Tyrosine 66 et Glycine 67 portés par une hélice

Manuel PIRES

alpha qui traverse le tonneau bêta) (Heim et al., 1994). La split-GFP a été obtenue à partir d'une protéine GFP modifiée pour accélérer son repliement (super-folder GFP, S30R, Y39N, N105T, Y145F, I171V et A206V). La split-GFP bi-partite consiste en la séparation en deux fragments de la GFP, un premier conservant la majeure partie de la structure de la protéine ([GFP1-10], brins bêta 1 à 10, acides aminés 1 à 214) alors que le second beaucoup plus petit est composé uniquement du 11^{ème} et dernier brin bêta de la GFP ([GFP11], résidus 215 à 230 ; **Figure 60 A**). En ce qui concerne la split-GFP tri-partite, la partie majoritaire est composée des brins 1 à 9 ([GFP1-9], acides aminés 1-193) alors que chacun des petits fragments correspond au 10^{ème} brin ([GFP10], résidus 194-214) et 11^{ème} brin (identique à la split-bi-partite), respectivement (**Figure 60 B**).

L'utilisation de ces systèmes solutionne le fait que la forme compacte de Gag est rare et transitoire. En effet, l'utilisation de la split-GFP permet de bloquer la structure de Gag sous forme compacte car la reconstitution du fluorophore à partir des différents fragments est quasi irréversible. Grâce à cela, on augmente artificiellement la quantité cellulaire de la forme compacte facilitant ainsi son observation. Les inconvénients de ce système sont ainsi tournés à notre avantage afin d'observer la forme compacte de Gag *in cellulo*.

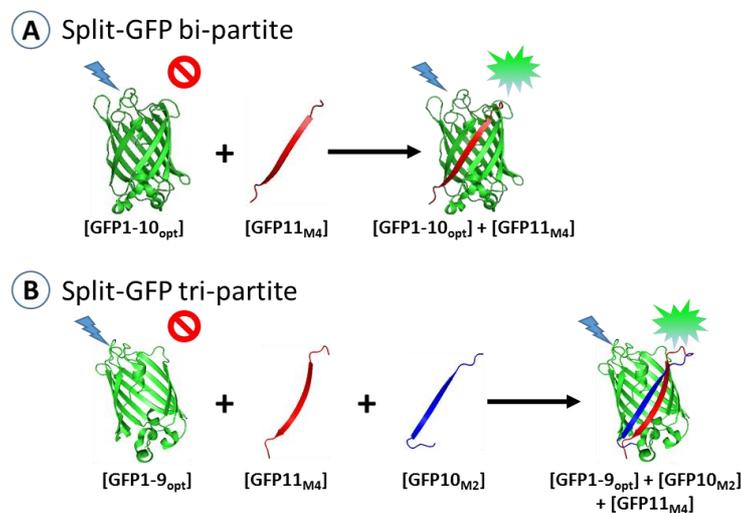


Figure 60 : Fragments de la split-GFP en deux ou en trois parties. Les différents fragments ont été optimisés pour ne pas modifier la solubilité de la protéine suivie par ce système et pour s'auto-assembler lorsqu'ils sont suffisamment proches (Cabantous et al., 2005, 2013). Les différentes structures ont été obtenues à partir de la séquence PDB 2Y0G

Nous avons donc inséré le petit fragment de la split-GFP entre les domaines MA et CA de Gag et le grand fragment à l'extrémité C-terminale de Gag. L'hypothèse est que lorsque Gag adopte une forme en U, le rapprochement des deux fragments permet une recombinaison

Manuel PIRES

INTRAmoléculaire de l'eGFP (*Figure 61*). Toutefois, ce système n'élimine pas une possible recombinaison INTERmoléculaire de la fluorescence (*Figure 61*), suite à la multimérisation des protéines Gag. L'utilisation de contrôles adaptés consistant en la co-expression de deux Gag portant chacune un des fragments de la GFP nous permettra de mesurer le signal INTER.

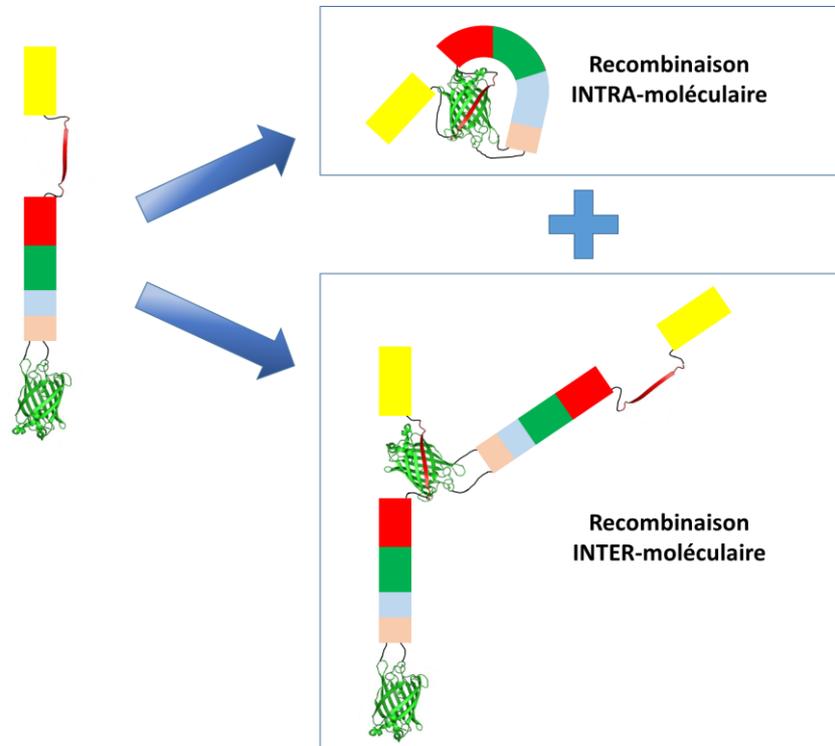


Figure 61 : Utilisation de la split-eGFP bipartite pour montrer l'existence d'une forme compacte de Gag in cellulo. En insérant le petit fragment (GFP11) entre les domaines MA et CA de Gag et le grand fragment (GFP1-10) en C-terminal de Gag, il est possible de reformer une GFP fluorescente par recombinaison INTRAmoléculaire (forme en U) ou INTERmoléculaire (multimérisation). La recombinaison INTERmoléculaire peut également être obtenue en exprimant deux protéines Gag portant chacune un seul des fragments de la split-GFP.

Réalisation des constructions plasmidiques

Afin de réaliser les constructions nécessaires pour ce projet à partir d'un même squelette protéique de Gag, nous avons inséré deux sites de clonage multiples (MCS) contenant plusieurs sites de restriction. Le premier (MCS1) est situé au niveau du site classique d'insertion de la GFP dans Gag (entre MA et CA) et le second (MCS2) à l'extrémité C-terminale de Gag. Ces MCS nous permettent d'insérer les différents fragments du système split GFP et de remplacer ou de supprimer des portions de Gag afin de générer des mutants (*Figure 62*). Le MCS1 porte les sites de restriction PacI et AclI dans lesquels peuvent être insérés les fragments de split GFP ou l'eGFP et le MCS2 porte les sites AgeI et ClaI afin d'insérer ces mêmes types de fragments en C-terminal. Le clonage de MCS1 au sein de Gag permet également d'introduire deux peptides de liaison de part et d'autre de l'insert (linkers 1 et 2,

Manuel PIRES

Gly-Gly-Gly-Gly-Ser) pour apporter un peu de flexibilité. La présence d'autres sites comme BamHI et EcoRI nous permet également de transférer d'une construction à une autre des parties de Gag afin de faciliter l'obtention de tous les plasmides nécessaires pour l'étude d'un mutant (Gag portant une eGFP pleine longueur, GFP11_MCS2, MCS1_GFP1-10 ou GFP11_GFP1-10). Il est important de noter que dans notre nomenclature, lorsqu'un fragment est inséré dans un MCS, seul le nom du fragment apparaît alors que lorsqu'il n'y a aucune insertion, le MCS conserve sa place dans la nomenclature du plasmide ou de la protéine. Par exemple, le plasmide gag Δ NC_GFP11_MCS2 code pour la protéine Gag délétée du domaine NC avec le 11^{ème} brin bêta de la GFP inséré en MCS1 et aucune insertion en MCS2.

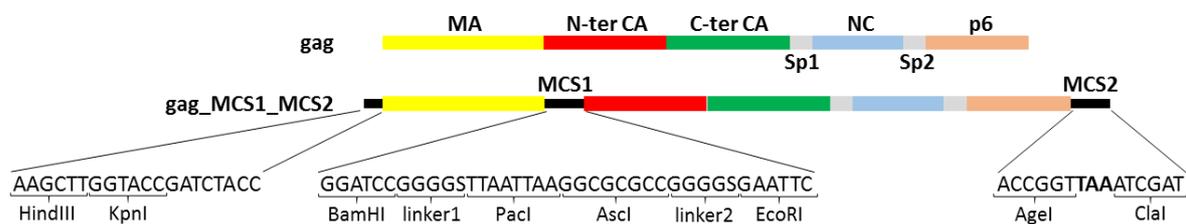


Figure 62 : Insertion de deux « multi-cloning sites » (MCS) entre la MA et la CA pour MCS1 et en C-terminal de Gag pour MCS2.

Validation des contrôles et du principe

Nous avons par la suite validé les différentes constructions et les contrôles. Pour ce faire, des cellules HeLa ont été transfectées pendant 12 h avec les différents plasmides, puis fixées à l'aide de PFA 4 %. Tout d'abord, nous avons vérifié que, comme attendu d'après les résultats de Cabantous et collègues, les fragments de la GFP au sein de Gag soient totalement non fluorescents (**Figure 63 A, B**). Contrairement au marquage ReAsH, nous pouvons ainsi nous affranchir complètement du bruit de fond qui aurait pu être formé par la fluorescence résiduelle de chaque fragment. Lorsque les deux fragments de la GFP sont co-exprimés dans une cellule, au sein de la même protéine ou de deux protéines différentes, leur rapprochement entraîne l'apparition d'un signal de fluorescence (**Figure 63 C et D**). Lorsque les fragments sont portés par la même protéine Gag (**Figure 63 C**, Gag_GFP11_GFP1-10), nous observons une forte fluorescence cytoplasmique sans accumulation spécifique à la membrane (intensités de fluorescence équivalentes) alors que lorsqu'ils sont portés par deux protéines Gag différentes, la fluorescence globale est moins intense et elle est surtout observée à la membrane (site de forte multimérisation de Gag) (**Figure 63 D**, Gag_GFP11_MCS2 + Gag_MCS1_GFP1-10). Dans le cas de Gag_GFP11_GFP1-10, la fluorescence peut résulter à la fois de la reformation d'une GFP par des fragments en intramoléculaire ou en intermoléculaire

Manuel PIRES

mais le contrôle montre que le signal INTER est majoritairement membranaire. Le signal cytoplasmique observé en INTRA est donc spécifique de Gag_GFP11_GFP1-10 et probablement lié à la forme compacte de la protéine Gag.

Ces premières observations nous ont permis de valider notre approche. Dans l'idéal, nous aurions souhaité une absence totale de fluorescence lors de l'expression des fragments de la GFP portés par des Gag différentes (pas de recombinaison intermoléculaire). Ce n'est pas le cas, ce qui suggère que Gag est capable de se multimériser sous des formes variables et éloignées structurellement de l'alignement parallèle de forme allongées souvent représenté pour la Gag multimérique.

Toutefois, les signaux de fluorescence observés dans les cas INTRA et INTER sont très différents autant qualitativement (cytoplasmique et membranaire dans la recombinaison INTRA et membranaire dans l'INTER) que quantitativement (fluorescence beaucoup plus intense dans l'intramoléculaire). Cette approche semble donc adaptée pour l'observation de la forme compacte de Gag *in cellulo* et nos résultats suggèrent que le signal cytoplasmique est probablement lié à la forme compacte de Gag. La confirmation de cette hypothèse constitue une grande partie du travail présenté par la suite.

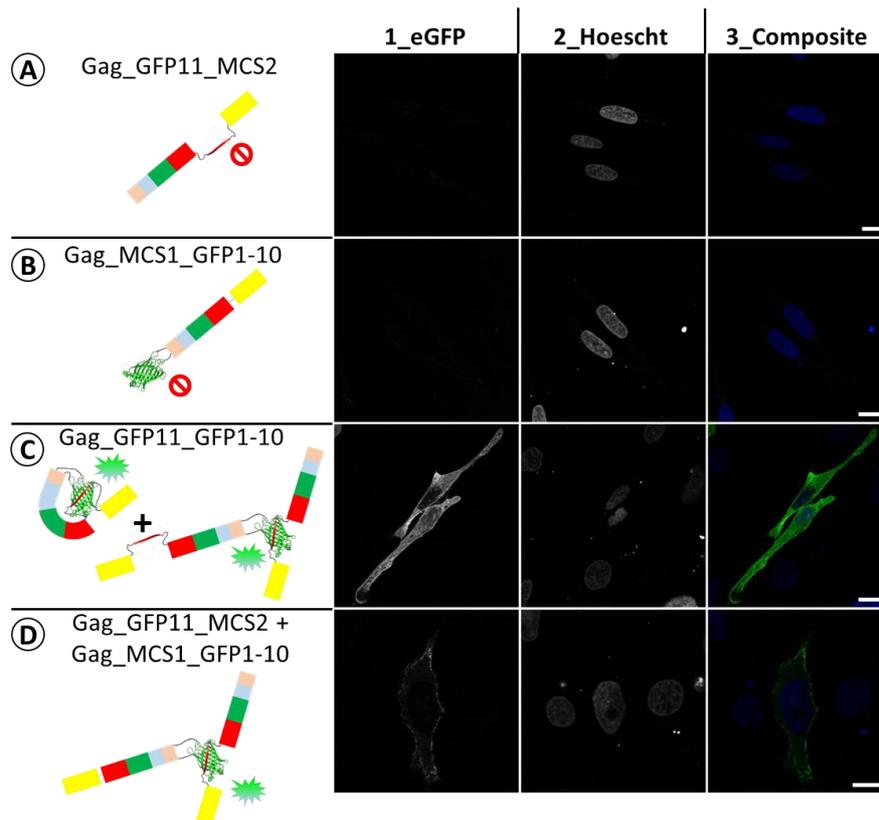


Figure 63 : Validation du système split-GFP pour l'observation de la forme compacte de Gag. Observation par microscopie confocale de la fluorescence GFP (colonne 1) de cellules HeLa 12 h post transfection par des plasmides codant Gag_GFP11_MCS2 (A, 1 µg), Gag_MCS1_GFP1-10 (B, 1 µg), Gag_GFP11_GFP1-10 (C, 1 µg) ou co-transfectées par gag_GFP11_MCS2 (0,5 µg) et gag_MCS1_GFP1-10 (D, 0,5 µg). Les cellules ont été fixées au PFA 4 % et traitées au Hoechst (colonne 2). Les images composites des signaux GFP et Hoechst sont présentées dans la colonne 3. Le schéma des protéines exprimées par chaque construction est donné à gauche de chaque série d'images. La barre d'échelle indique une longueur de 20 µm.

Résultats obtenus en échangeant les positions des fragments de la split-GFP

Nous avons également inversé le placement des fragments, à savoir le petit fragment (GFP11) en C-terminal et le grand fragment (GFP1-10) entre MA et CA (**Figure 64**). Aucune différence n'a été observée au niveau de la morphologie cellulaire. Si la recombinaison intermoléculaire est inchangée et reste principalement localisée à la membrane (comparaison de la **Figure 63 D** et de la **Figure 64 B**), la fluorescence intramoléculaire est toujours principalement cytoplasmique, mais son intensité est plus faible (comparaison de la **Figure 63 C** et de la **Figure 64 A**). C'est pourquoi, nous avons décidé d'utiliser pour la suite des études les constructions où GFP11 est placée en MCS1 et GFP1-10 en MCS2. Ces résultats confirment la présence d'une forme compacte cytoplasmique.

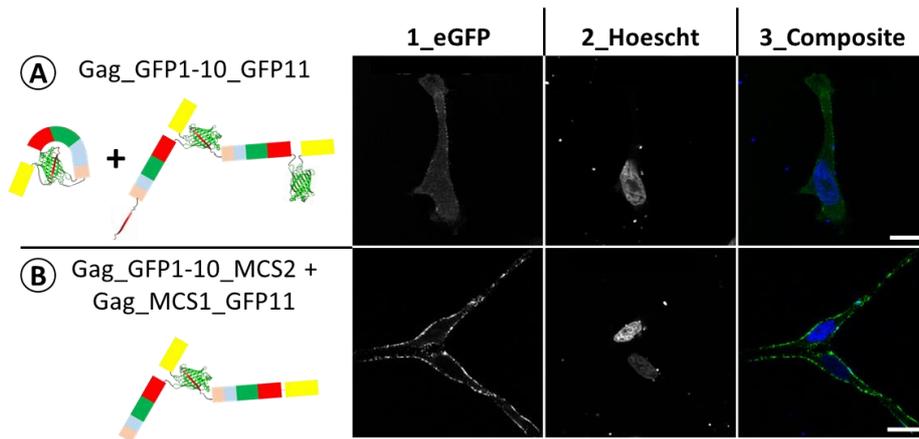


Figure 64 : Observation par microscopie confocale de la fluorescence GFP de cellules HeLa 12 h après transfection par un plasmide codant Gag_GFP1-10_GFP11 (A, 1 μ g) ou cotransfectées par des plasmides codant Gag_MCS1_GFP11 (0,5 μ g) et Gag_GFP1-10_MCS2 (B, 0,5 μ g). Les cellules ont été fixées au PFA 4 % et traitées au Hoechst (A2 et B2). Les images composites des signaux GFP et Hoechst sont données en A3 et B3. La barre d'échelle correspond à 20 μ m.

Quantification de la différence des fluorescences membranaires et cytoplasmiques

Nous voulions par la suite pouvoir quantifier de manière reproductible la différence de fluorescence entre recombinaisons intramoléculaire et intermoléculaire en observant un grand nombre de cellules pour chaque type de transfection. Pour cela, un plugin ImageJ a été développé spécifiquement par l'ingénieur de la Plateforme d'Imagerie Quantitative (PIQ, Romain Vauchelles). Cette analyse semi-automatique a pour but d'obtenir rapidement l'intensité moyenne par pixel de l'espace dit « membranaire » et de celui dit « cytoplasmique ». La première étape consiste en une délimitation approximative de la bordure cellulaire par ajustement du seuil de détection (dit threshold, apparaît en rouge dans la **Figure 65 A2 et B2**). Cette délimitation approximative permet d'établir le contour cellulaire (ligne jaune dans la **Figure 65 A3 et B3**), puis l'espace membranaire est défini à partir de ce contour par un algorithme qui décrit cet espace comme une bande de largeur régulière (13 pixels, détermination arbitraire) contournant la cellule vers l'intérieur du contour. L'espace cytoplasmique contient tout ce qui se trouve à l'intérieur de cet espace, excepté la zone nucléaire (détectée par sa coloration au Hoechst et soustraite de la zone précédemment définie). Une fois ces zones définies, les valeurs de fluorescence des pixels qui les composent permettent de définir une fluorescence moyenne par pixel de chaque zone. Les différentes zones sont schématisées dans le panel du haut de la **Figure 65**.

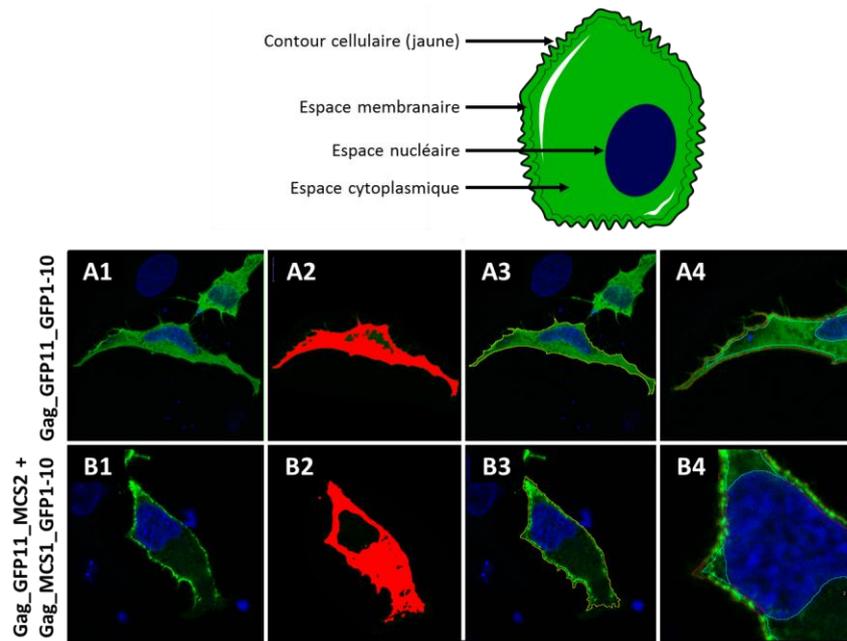


Figure 65 : Explication du Plugin ImageJ utilisé pour quantifier la fluorescence membranaire et cytoplasmique. Un exemple d'analyse est donné pour Gag_GFP11_GFP1-10 (A1 à A4) et pour Gag_GFP11_MCS2 + Gag_MCS1_GFP1-10 (B1 à B4). En A2 et B2, l'image montre l'ajustement du seuil de détection permettant de déterminer le contour cellulaire (trait jaune en A3 et B3 et zoom en A4 et B4). L'espace membranaire est alors déterminé par soustraction à partir de ce contour d'une épaisseur de 13 pixels. Enfin, l'espace nucléaire est soustrait de la fluorescence correspondant à l'espace cytoplasmique.

Malgré des réglages identiques du microscope dans les différentes expériences et bien sûr entre les différents échantillons (puissance du laser, gain, exposition, ...), l'intensité de fluorescence apparaît assez variable d'une cellule à une autre. Afin de s'affranchir de cette variabilité, nous avons décidé d'exprimer les intensités de fluorescence sous forme de ratio ($R_{m/c}$) entre la fluorescence moyenne des pixels membranaires et celle des pixels cytoplasmiques. Il nous est alors possible de comparer les ratios obtenus pour les différentes constructions (**Figure 66 A**). Ainsi, dans le cas où les fragments sont insérés dans la même protéine (Gag_GFP11_GFP1-10), on remarque une fluorescence moyenne cytoplasmique de 1279 ± 630 U.A. par pixel et membranaire de 1429 ± 788 U.A. par pixel ce qui donne un ratio $R_{m/c}$ de $1,09 \pm 0,2$ (correspond à la moyenne des ratios obtenus pour chaque cellule). Dans le cas de la recombinaison intermoléculaire (Gag_GFP11_MCS2 + Gag_MCS1_GFP1-10), la fluorescence membranaire est inférieure à celle observée en INTRA (moyenne de 1076 ± 671 U.A. par pixel). La différence la plus flagrante est la très forte diminution de l'intensité de fluorescence dans le cytoplasme (moyenne de 524 ± 385 U.A. par pixel), résultant en une forte augmentation du $R_{m/c}$ à $2,09 \pm 0,4$, qui devient significativement différent de celui obtenu pour l'étude en INTRA.

Manuel PIRES

Grâce à ces analyses qui confirment les observations préalables, il apparaît que la forme compacte de Gag permette une reconstitution intramoléculaire cytoplasmique de la fluorescence à partir des deux fragments de GFP. La fluorescence membranaire de ces échantillons est plus probablement engendrée par la recombinaison intermoléculaire causée par le rapprochement des fragments lors de la multimérisation de Gag.

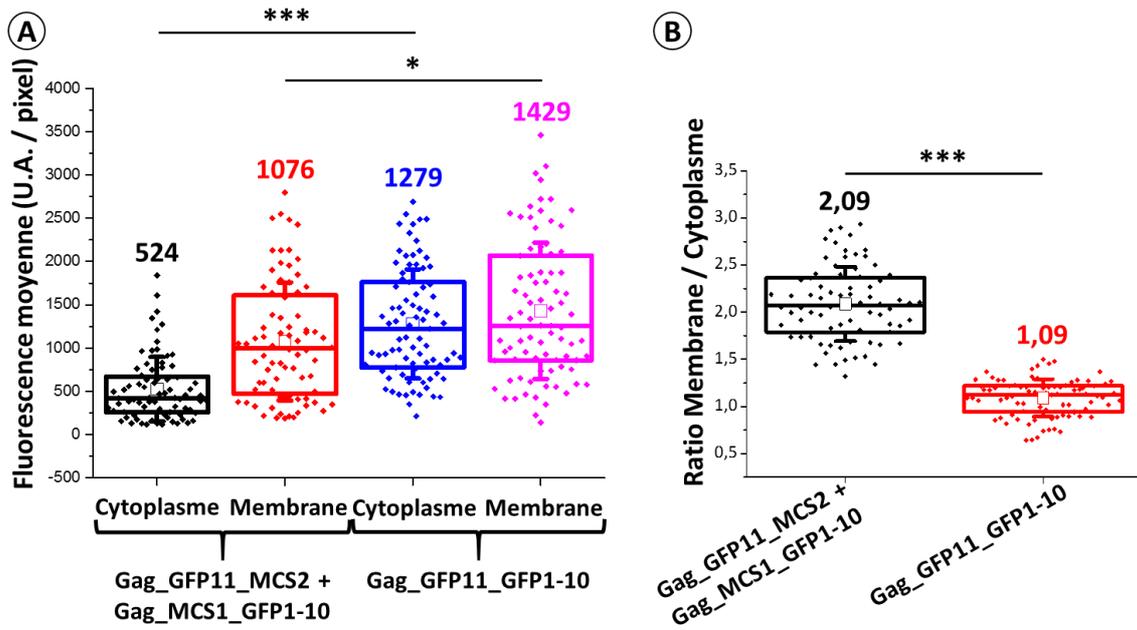


Figure 66 : Recombinaison intra- et intermoléculaire de la split-GFP portée par Gag. Les intensités de fluorescence moyennes par pixel des zones membranaire et cytoplasmique et les Rm/c (ratio membrane/cytoplasme) sont représentés en A et B respectivement. Chaque point représente une cellule (intra moléculaire $N = 6$ $n = 83$, intermoléculaire $N = 6$ $n = 76$, où N représente le nombre d'expérience indépendante et n le nombre total d'échantillons analysés). La boîte représente la médiane (trait horizontal au centre), 75 % des échantillons (trait horizontal supérieur) et 25 % des échantillons (trait horizontal inférieur). La moyenne est indiquée par un carré blanc et les barres d'erreurs montrent l'écart type à la moyenne. Les valeurs p obtenues dans un test HSD de Tukey sont données avec *** $p < 0,001$; ** $p < 0,01$ et * $p < 0,5$.

Un contrôle important pour valider nos conclusions est de vérifier que le repliement de Gag est bien à l'origine de la reconstitution de la GFP. En effet, comme mentionné précédemment, les fragments de la split-GFP présentent une forte affinité l'un pour l'autre et pourraient donc conduire la recombinaison de la GFP. Pour vérifier cette hypothèse, nous avons choisi de placer les fragments de la split-GFP aux extrémités N- et C-terminales de la firefly-luciférase, afin de suivre les signaux de fluorescence INTER et INTRA moléculaires obtenus lors de l'expression en cellules HeLa et de les corrélés avec l'activité enzymatique des luciférases en supposant qu'une luciférase dont les deux extrémités seraient rapprochées est non active.

Manuel PIRES

La distance théorique entre les deux extrémités de la luciférase étant de 52,4 Å (5,24 nm, mesurée par PyMOL sur la structure PDB 2D1S), les deux extrémités seraient donc artificiellement rapprochées dans le cas d'une reconstitution intra moléculaire liée à l'affinité de GFP11 pour GFP1-10. Cette distance est largement inférieure à la longueur théorique de la protéine Gag (\approx 280 Å, MA = 68,6 Å (PDB 1UPH), CA = 58,8 Å (PDB 2M8N), Sp1 = 61,7 Å (PDB 1U57), NC = 68,3 Å (PDB 1AAF) et p6 = 21,4 Å (PDB 2C55)). Il est donc raisonnable de penser que si l'affinité des fragments de la GFP provoquait la reconstitution d'une protéine complète cet événement serait plus probable pour la luciférase que pour Gag (même si cette hypothèse ne tient pas compte de la flexibilité de la protéine). De plus, la luciférase étant une protéine majoritairement monomérique (Conti et al., 1996), le signal INTER avec cette protéine devrait donc être uniquement lié à l'interaction entre les fragments de la split-GFP.

Lors de l'observation en microscopie des cellules transfectées par des constructions codant pour la protéine luciférase fusionnée à un ou aux deux fragments de la split-GFP, nous n'avons observé qu'un très faible nombre de cellules fluorescentes et uniquement dans le cas [GFP11_Lucif_MCS2 + MCS1_Lucif_GFP1-10] correspondant à un signal intermoléculaire. La protéine luciférase issue de *Photinus pyralis*, en général trouvée sous forme monomérique, est capable de se dimériser à forte concentration ce qui peut expliquer ce résultat (Denburg and McElroy, 1970). Afin de vérifier si les fragments de split-GFP perturbent la structure de la luciférase, nous avons mesuré l'activité des différentes fusions. MCS1_Lucif_GFP1-10 apparait plus active que GFP11_Lucif_MCS2 (**Figure 67**), suggérant que la GFP11 en N-terminal est délétère pour l'activité de la protéine. Toutefois, nous n'avons pas contrôlé le niveau d'expression de ces constructions, il est donc possible que les différences observées soient liées à des différences de niveaux d'expression. Ce qui est le plus important est que la co-expression de MCS1_Lucif_GFP1-10 et GFP11_Lucif_MCS2 n'entraîne pas de diminution de l'activité. De même lorsque la luciférase porte les deux fragments de la GFP il n'y a pas perte d'activité en comparaison des contrôles. Ces résultats montrent que les fragments de la GFP ne dirigent pas le repliement de la luciférase et suggèrent que la fluorescence observée pour Gag dans le cas d'une recombinaison intramoléculaire (Gag_GFP11_GFP1-10) est provoquée par un repliement de la protéine Gag sous forme compacte plutôt que par l'affinité des fragments entre eux.

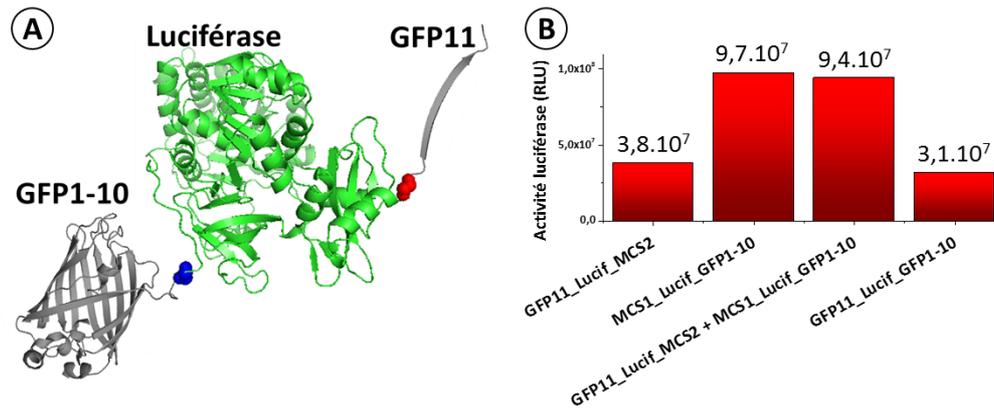


Figure 67 : Influence des fragments de la split-GFP sur l'activité de la luciférase. (A) Structure théorique de la luciférase portant les fragments de la split-GFP aux extrémités N- et C-terminales. Structure tridimensionnelle adaptée de PDB 2D1S (luciférase) et de 2Y0G (eGFP). (B) Activités enzymatiques des divers mutants de la luciférase exprimés en cellules HeLa. Les cellules HeLa sont transfectées pendant 24 h par les différents plasmides (1 μ g d'ADN pour GFP11_lucif_MCS2, MCS1_lucif_GFP1-10 et GFP11_lucif_GFP1-10 ; ou 0,5 μ g de chaque plasmide pour GFP11_lucif_MCS2 + MCS1_lucif_GFP1-10). Les cellules sont ensuite lysées et l'activité luciférase est mesurée sur un échantillon de 20 μ L dans un lecteur de plaque après ajout de 50 μ L de réactif luciférine.

Utilisation du système split-GFP tri-partite

Nous disposons également de la split tri-partite dans laquelle la GFP est divisée en trois parties : une grande constituée des brins β 1 à 9 (GFP1-9) et deux petites, le 10^{ème} feuillet β seul (GFP10) et le 11^{ème} feuillet β seul (GFP11) placées chacune dans une MCS (GFP10 dans MCS1 et GFP11 dans MCS2). La GFP1-9 exprimée dans les cellules est non fluorescente (Figure 68 A), comme évidemment les deux brins beta indépendants (résultats non présentés). La présence de seulement deux des trois fragments de la GFP n'est pas suffisante pour observer une fluorescence (Figure 68 B et C). Lorsque nous nous plaçons dans la configuration de recombinaison intermoléculaire (Gag_GFP10_MCS2 + Gag_MCS1_GFP11 + GFP1-9), aucune fluorescence n'est observée, alors que lors de l'observation de la recombinaison intramoléculaire, quelques cellules fluorescentes sont visibles (Figure 68 E), indiquant un rapprochement des deux brins beta, lié au repliement de Gag en forme compacte.

Dans ces expériences, la GFP1-9 était exprimée libre dans le cytoplasme des cellules, ce qui est vraisemblablement en grande partie responsable des faibles signaux observés dans cette approche. Il est possible d'imaginer de la fusionner à la protéine MCP du bactériophage MS2 qui est elle-même capable de reconnaître des tiges-boucles d'ARN dites MS2 (Bertrand et al., 1998; Buxbaum et al., 2015). L'expression simultanée de Gag_GFP10_GFP11 et MCP-GFP1-9 permettrait de savoir si la forme compacte de Gag est associée spécifiquement à l'ARN génomique.

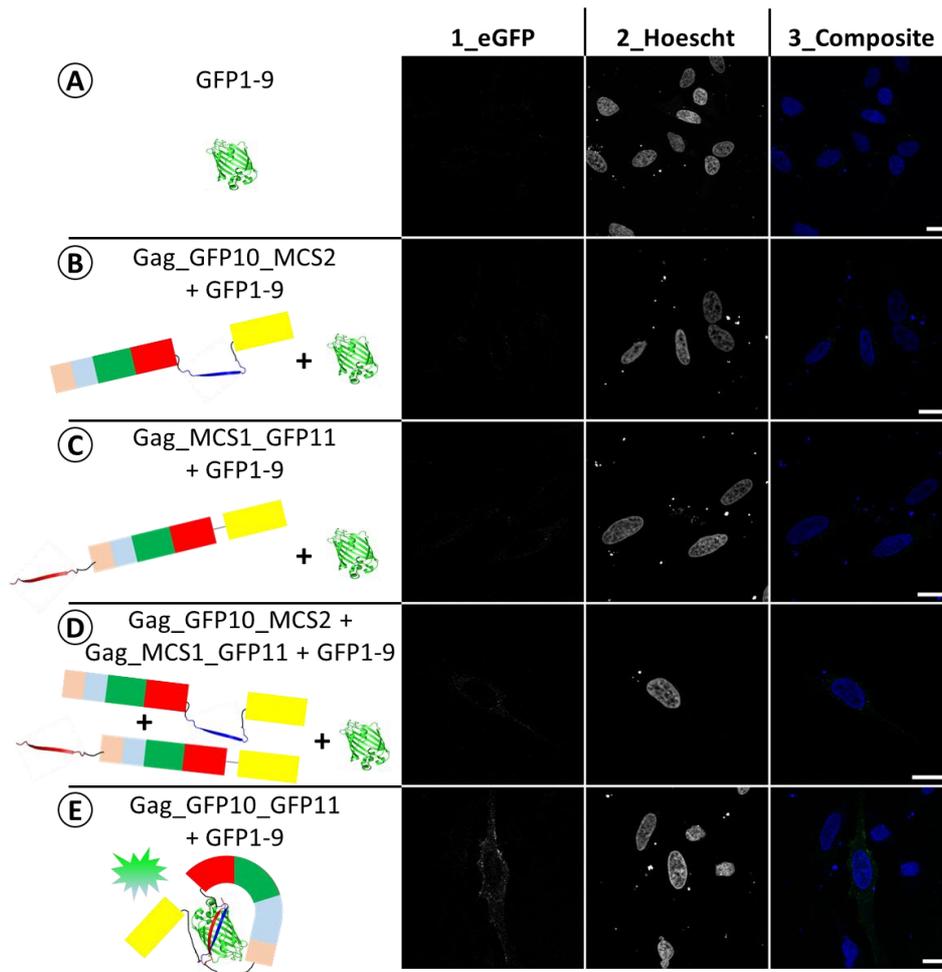


Figure 68 : Utilisation du système split-GFP tripartite pour l'observation par microscopie confocale de la forme compacte de Gag. Des cellules HeLa ont été transfectées par des plasmides codant GFP1-9 (A, 1 μ g) ; GFP1-9 (0,5 μ g) + Gag_GFP10_MCS2 (B, 1 μ g) ; GFP1-9 (0,5 μ g) + Gag_MCS1_GFP11 (C, 1 μ g) ; GFP1-9 (0,5 μ g) + Gag_GFP10_MCS2 (0,5 μ g) + Gag_MCS1_GFP11 (D, 0,5 μ g) ; GFP1-9 (0,5 μ g) + Gag_GFP10_GFP11 (E, 1 μ g) et fixées 12 h plus tard au PFA 4 % et traitées au Hoechst (colonne 2). Les images composites des signaux GFP (colonne 1) et Hoechst sont présentées dans la colonne 3. Le schéma des protéines exprimées par chaque construction est donné à gauche de chaque série d'images. La barre d'échelle indique une longueur de 20 μ m.

3. Observations sur différents mutants

Afin de vérifier la pertinence de la forme compacte *in cellulo*, nous avons par la suite construit différents mutants pour vérifier l'importance des deux sites de la liaison à l'ARN, de la multimérisation ou de certains domaines de Gag dans la formation de la Gag compacte.

a) Perte d'interaction avec l'ARN

Selon les études *in vitro*, la forme compacte de Gag lie les ARNs par deux sites différents :

- **Nucléocapside (NC)** permet la sélection du génome viral parmi un large excès d'ARNs cellulaires (Cimarelli and Darlix, 2002; Muriaux and Darlix, 2010). Comme

la suppression de ce domaine modifie fortement l'affinité de la protéine Gag pour les ANs, nous avons choisi de déléter ce domaine afin de supprimer une des interactions de Gag aux ARNs (mutant **Delta-NC** ou **Δ NC**).

- **Domaine hyper-basique** (HBR, Highly Basic Region) : ce domaine MA de Gag permet l'interaction avec la région de l'ARN qui code pour Pol (Purohit et al., 2001). Comme les lysines en positions 26, 27 et 28 seraient impliquées dans la liaison avec les ANs (Chukkapalli et al., 2010; Ott et al., 2005), nous avons choisi d'introduire des mutations non conservatives en remplaçant ces trois résidus lysines par des thréonines (mutant **KKK-25,26,27-TTT**, nommé pour simplifier KK-TT).

Un troisième mutant, associant la délétion du domaine NC et les mutations KK-TT a également été réalisé afin de vérifier l'importance des deux domaines (mutant **KK-TT_ Δ NC**).

Tout d'abord, nous avons vérifié l'influence de ces mutations sur la distribution intracellulaire de la protéine Gag en marquant les mutants par l'eGFP pleine longueur entre les positions MA et CA. Des cellules HeLa ont été transfectées transitoirement et fixées au PFA 4 % à $t = 12$ h puis observées en microscopie confocale. Dans le cas de la protéine sauvage (dite WT, Wild-Type), on observe une intensité de fluorescence plus importante à la membrane (1862 ± 879 U.A. par pixel) que dans le cytoplasme (1415 ± 759 U.A. par pixel) donnant des ratios $R_{m/c}$ de $1,36 \pm 0,32$. Cette localisation préférentiellement membranaire était attendue car elle traduit l'assemblage des protéines Gag à la membrane pour la production de particules virales (**Figure 69 A**). Notons que le ratio obtenu avec Gag-eGFP est intermédiaire à ceux obtenus pour les études intermoléculaires ($R_{m/c} = 2,09 \pm 0,39$) ou intramoléculaires ($R_{m/c} = 1,09 \pm 0,2$) avec la split-GFP, ce qui suggère que la fluorescence observée avec cette dernière provient principalement d'une sous-population de Gag à la membrane dans le cas intermoléculaire et d'une sous-population cytoplasmique en intramoléculaire. Dans le cas de la délétion du domaine NC (Gag Δ NC_eGFP), l'intensité de fluorescence totale est diminuée, avec une diminution plus marquée au niveau membranaire (moyenne de 1220 ± 709 U.A. par pixel) que cytoplasmique (moyenne de 1180 ± 690 U.A. par pixel). Cette diminution de la fluorescence membranaire qui entraîne une diminution significative du ratio $R_{m/c}$ ($1,02 \pm 0,18$) est certainement due aux défauts d'assemblage déjà décrits pour ce mutant (**Figure 69 B**, El Meshri et al., 2015). La diminution de la fluorescence totale est probablement liée à une moins grande stabilité de cette protéine. Dans le cas du mutant KK-TT (**Figure 69 C**), aucune modification du ratio $R_{m/c}$ ($1,33 \pm 0,29$; moyenne des ratios) n'est observée en comparaison du WT. En effet,

Manuel PIRES

les fluorescences cytoplasmique (moyenne de 1360 ± 827 U.A. par pixel) et membranaire (moyenne de 1693 ± 850 U.A. par pixel) sont toutes les deux légèrement diminuées par rapport à celles de Gag WT suggérant que les défauts d'assemblage de ce mutant sont très faibles. Le double mutant KK-TT_ΔNC (*Figure 69 D*), n'a été observé que lors d'une seule expérience ce qui ne nous permet pas de calculer précisément les fluorescences dans les deux compartiments. Toutefois la valeur du ratio $R_{m/c}$ ($1,33 \pm 0,29$) semble être diminuée suggérant un défaut d'assemblage proche de celui observé dans le cas de la délétion du domaine NC. Toutefois, on ne retrouve pas de diminution de la fluorescence globale, suggérant une absence de déstabilisation pour ce mutant.

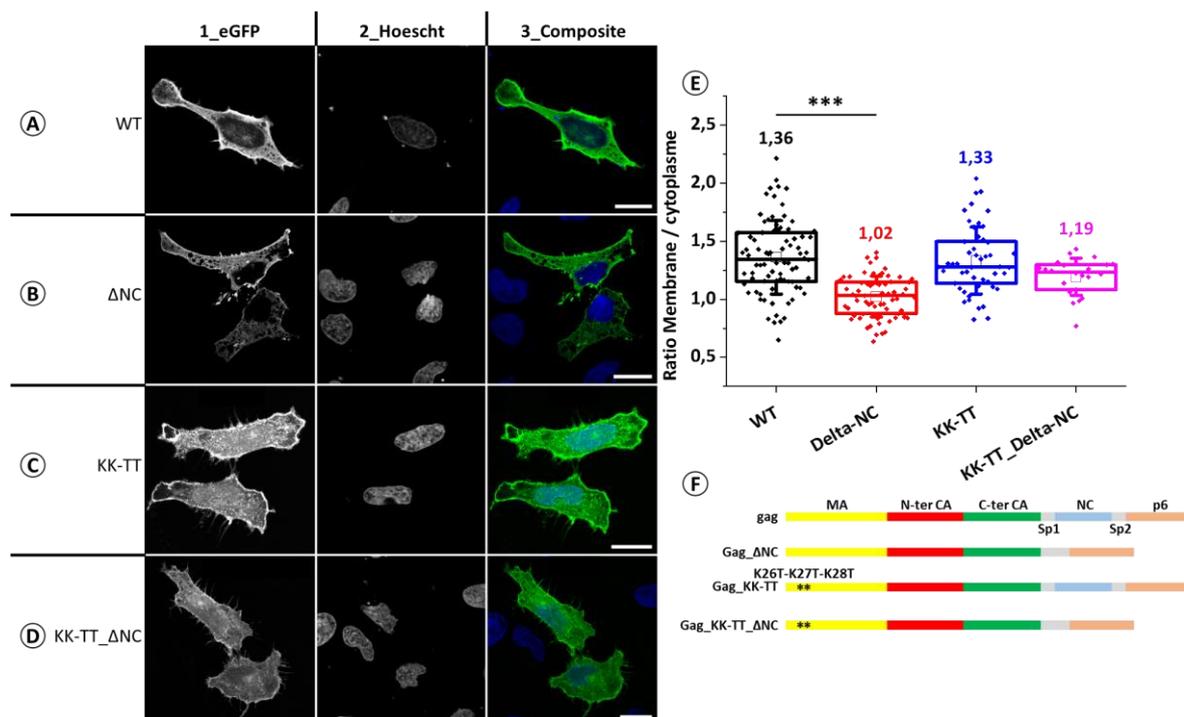


Figure 69 : Localisation cellulaire de mutants de Gag-eGFP déficients pour l'interaction avec l'ARN. Des cellules HeLa ont été transfectées par des constructions codant Gag_eGFP (A, 1 μ g) ; Gag Δ NC_eGFP (B, 1 μ g) ; GagKK-TT_eGFP (C, 1 μ g) ou Gag Δ NC_KK-TT_eGFP (D, 1 μ g), fixées au PFA 12 h post transfection et observées au microscope confocal. Pour chaque échantillon sont présentées une image de fluorescence eGFP (colonne 1), le signal du marquage Hoechst (colonne 2) et l'image composite (colonne 3). L'échelle indique une distance de 20 μ m. (E) Représentation des ratios ($R_{m/c}$) pour les différents mutants testés. Chaque point représente une cellule (Gag_eGFP N = 6, n = 82 ; Gag Δ NC_eGFP N = 5, n = 66 ; GagKK-TT_eGFP N = 4, n = 49 ; Gag Δ NC_KKTT_eGFP N = 1, n = 22). La boîte représente la médiane (trait horizontal au centre), 75 % des échantillons (trait horizontal supérieur) et 25 % des échantillons (trait horizontal inférieur). La moyenne est indiquée par un carré blanc et les barres d'erreurs représentent l'écart-type de la moyenne. (F) Schéma des différents mutants de Gag utilisés dans cette étude. Les valeurs p obtenues dans un test HSD de Tukey sont données avec *** p < 0,001.

Manuel PIRES

Cette analyse nous a permis de confirmer les défauts d'assemblage de la protéine Gag dans le cas d'une délétion du domaine NC. En revanche, la mutation connue pour diminuer l'interaction du domaine MA avec les acides nucléiques (KK-TT) ne semble pas affecter l'assemblage (Burniston et al., 1999 ; Ott et al., 2005).

Par la suite, nous avons analysé la capacité de ces différents mutants à former une forme compacte grâce au système split-GFP bi-partite. Pour chaque mutant, des constructions ont été réalisées pour observer des reconstitutions intra- ou inter-moléculaires de la GFP. Comme pour Gag WT (voir [Figure 66](#)), les ratios $R_{m/c}$ ont été calculés ([Figure 70 G](#)). Les valeurs de $R_{m/c}$ obtenues pour les études intermoléculaires montrent clairement les défauts de multimérisation des mutants de Gag. En effet, pour tous les mutants, ces ratios sont significativement inférieurs au ratio de Gag WT ($1,65 \pm 0,47$ pour Gag Δ NC, $1,74 \pm 0,47$ pour Gag KK-TT et $1,61 \pm 0,3$ pour Gag KK-TT_ Δ NC à comparer à $2,09 \pm 0,39$ pour Gag WT), indiquant que les mutants sont moins aptes à se multimériser que Gag WT et ne s'accumulent pas à la membrane. Pour les mutants Δ NC et KK-TT_ Δ NC, ces résultats sont en accord avec ce qui a été vu pour les protéines Gag marquées par l'eGFP. Pour le mutant KK-TT, il semble que ce mutant présente une accumulation à la membrane identique à celle de Gag WT (voir [Figure 66](#)) mais une multimérisation moins efficace ($R_{m/c} = 1,74 \pm 0,47$ vs $2,09 \pm 0,39$ [Figure 66](#)).

Si on regarde les $R_{m/c}$ obtenus pour les études intramoléculaires, on observe que pour tous les mutants de Gag les ratios en INTRA sont significativement plus faibles qu'en INTER ([Figure 70 G](#)). Il est important de garder en mémoire que dans l'étude INTRA nous observons à la fois la forme compacte de Gag et des formes multimériques de Gag retrouvées dans l'étude INTER dont nous ne pouvons pas nous affranchir ici. Donc, plus la diminution du ratio est forte entre INTER et INTRA, plus dans l'étude INTRA se trouve une forme fluorescente cytoplasmique, que nous attribuons à la forme compacte de Gag. En calculant $\Delta R_{inter-intra} = R_{m/c inter} - R_{m/c intra}$ pour chaque mutant, on obtient $\Delta R_{inter-intra}$ (WT) = 1,0 ; $\Delta R_{inter-intra}$ (Δ NC) = 0,44 ; $\Delta R_{inter-intra}$ (KK-TT) = 0,51 et $\Delta R_{inter-intra}$ (KK-TT_ Δ NC) = 0,23. Ces données indiquent que tous les mutants présentent une diminution de la proportion de forme compacte tout en étant capables de se replier sous cette forme. Le mutant KK-TT_ Δ NC est le plus impacté, cumulant les défauts observés pour Δ NC et KK-TT.

En ce qui concerne le mutant KK-TT nous avons décidé, dans cette étude, de muter les lysines 25, 26 et 27 de la région HBR de MA car elles étaient décrites comme impliquées dans la liaison aux ARNs (Chukkapalli et al., 2010; Hearps et al., 2008; Purohit et al., 2001). Toutefois, d'autres

Manuel PIRES

acides aminés de la HBR peuvent également participer à cette interaction (Arginine 20, Lysines 28 et 30 ; Purohit et al., 2001). Il est donc possible que le mutant KK-TT conserve la capacité d'interagir avec les ANs même si cette interaction est de plus faible affinité que celle de la Gag WT. Dans tous les cas, ces études montrent que comme attendu la formation de Gag compacte est diminuée par des mutations affectant l'interaction entre les ANs et Gag. Il semble primordial de muter plus d'acides aminés dans le domaine HBR de la matrice afin de visualiser l'impact d'une perte totale de l'interaction ANs/Gag par ce domaine.

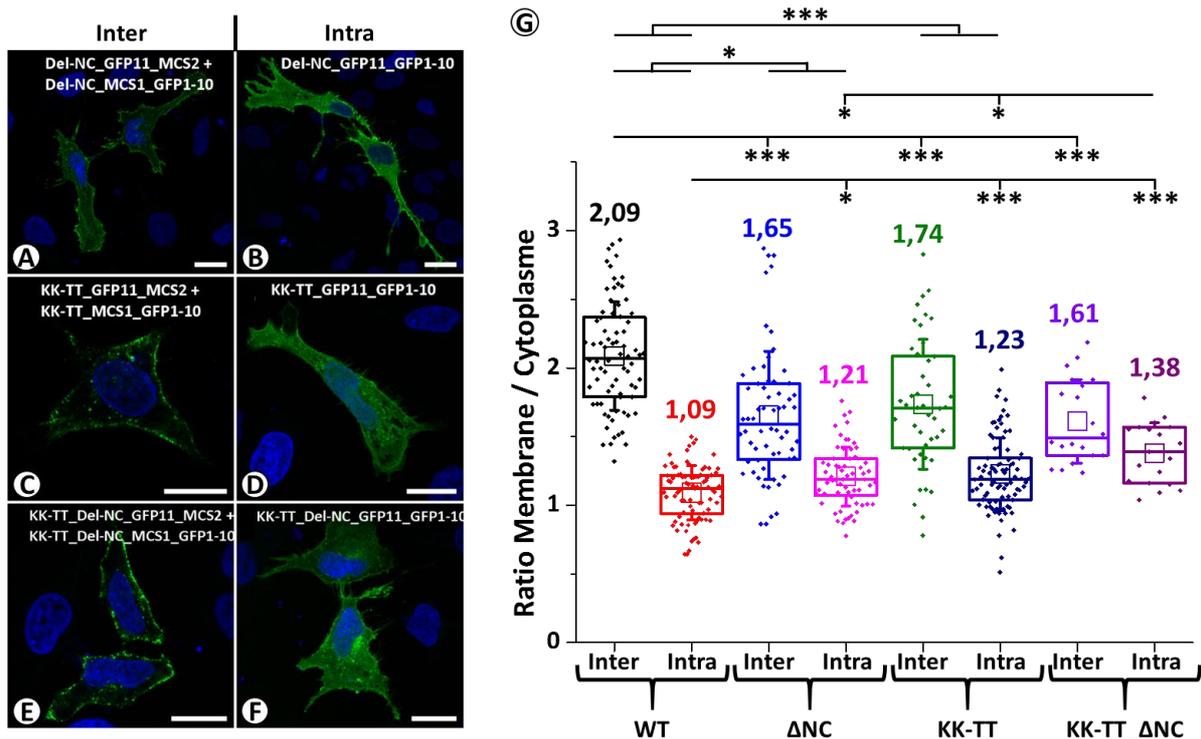


Figure 70 : Répartition membrane vs cytoplasme de la fluorescence obtenue par reconstitution de la split-GFP dans différents mutants de Gag. Images confocales (A-F) de cellules HeLa fixées au PFA 4 % à 12h post-transfection pour l'étude de la reconstitution de fluorescence pour les mutants de Gag ΔNC en intermoléculaire (A) ou intramoléculaire (B), KK-TT en intermoléculaire (C) ou intramoléculaire (D) et KK-TT_ΔNC en intermoléculaire (E) ou intramoléculaire (F). L'échelle indique une distance de 20 μm. (G) Représentation des ratios Rm/c pour la reconstitution de la split-GFP pour Gag WT (inter N = 6, n = 81 ; intra N = 6, n = 83), Gag ΔNC (inter N = 5, n = 66 ; intra N = 5, n = 64), Gag KK-TT (inter N = 3, n = 47 ; intra N = 6, n = 87) et Gag KK-TT_ΔNC (inter N = 1, n = 20 ; intra N = 1, n = 19), en inter- et en intramoléculaire. La boîte représente la médiane (trait horizontal au centre), 75 % des échantillons (trait horizontal supérieur) et 25 % des échantillons (trait horizontal inférieur). La moyenne inscrite au-dessus de chaque diagramme est indiquée par un carré blanc et les barres d'erreurs montrent l'écart type à la moyenne. INTER correspond à la transfection de 2 plasmides codant pour Gag couplée à un seul fragment de split-GFP, alors qu'INTRA correspond à une protéine Gag couplée à la fois à GFP11 et à GFP1-10. Les valeurs p obtenues dans un test HSD de Tukey sont données avec *** p < 0,001 et * p < 0,5.

Manuel PIRES

b) Étude d'un mutant portant un défaut d'assemblage

Dans les études précédentes utilisant la split-GFP, nous étions limités par la présence d'un signal intermoléculaire attribué à la capacité de Gag de se multimériser. Pour éviter cela, nous avons donc choisi d'étudier un mutant de Gag dans lequel deux résidus (Trp318 et Met319) de la région CA importants pour la multimérisation ont été mutés en Alanine (Datta et al., 2007b; Gamble et al., 1997; Ganser-Pornillos et al., 2004). En effet, des études *in vitro* ont montré une constante d'association de 5,5 μM pour Gag- Δp6 et supérieure à 20 μM pour Gag(WM-318/319-AA)- Δp6 (Kovaleski et al., 2006).

Une protéine Gag portant ces deux mutations et couplée à l'eGFP montre effectivement que l'assemblage membranaire est perturbé avec une fluorescence membranaire identique à la fluorescence cytoplasmique (**Figure 71 A**, $R_{m/c} = 1,0 \pm 0,1$ à comparer à $R_{m/c} = 1,35 \pm 0,3$ pour Gag WT, **Figure 67**). Cela indique que la faible capacité de GagWM à se multimériser limite son accumulation à la membrane. Cette inhibition de la multimérisation est également observée lors de l'étude intermoléculaire par la split-GFP de ce mutant, puisque le ratio $R_{m/c}$ passe ainsi de $2,1 \pm 0,4$ pour Gag WT à $1,75 \pm 0,6$ pour le mutant WM. Finalement, la recombinaison intramoléculaire de la split-GFP pour le mutant WM présente un ratio $R_{m/c} = 1,0 \pm 0,1$ comparable à celui de Gag WT ($1,1 \pm 0,2$) montrant que la mutation WM n'a pas d'effet sur la formation de la forme compacte de Gag. Contrairement aux mutants présentant une perte d'interaction avec les ANs, nous n'observons pas de rapprochement des valeurs de $R_{m/c}$ obtenus en inter- et intramoléculaire, qui reflétaient la perte de la forme cytoplasmique fluorescente.

Si la mutation WM diminue fortement la multimérisation et la localisation membranaire, il aurait été intéressant de l'associer à la mutation G2A. Cette mutation empêche en effet l'ancrage membranaire de la protéine Gag en supprimant le résidu impliqué dans la myristoylation de la protéine (Chukkapalli and Ono, 2011). Ainsi, en supprimant la multimérisation et l'interaction membranaire simultanément, nous devrions augmenter artificiellement la quantité de protéine Gag monomérique et donc la quantité de protéine sous forme compacte.

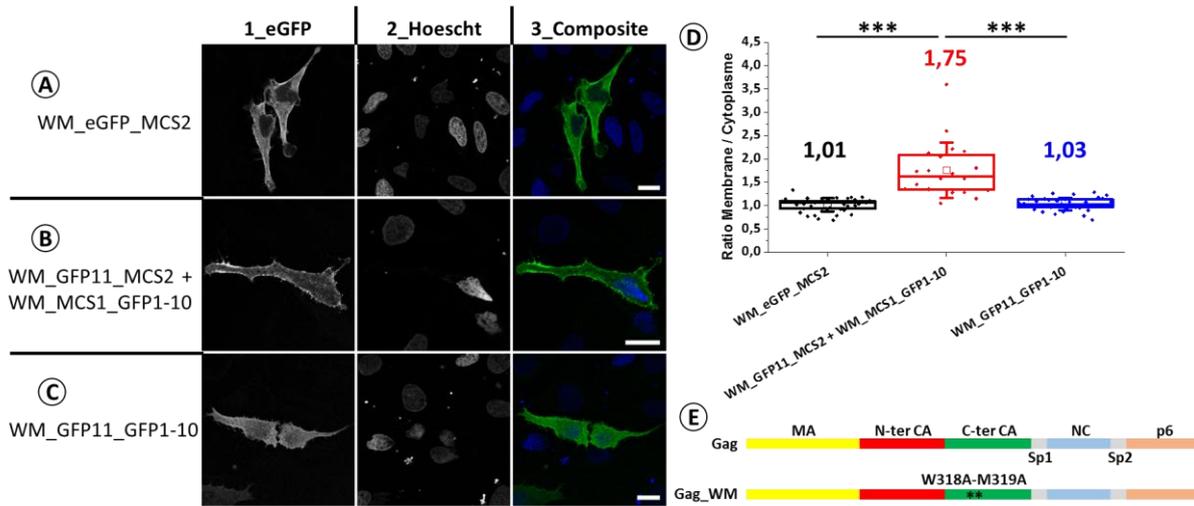


Figure 71 : Etude du mutant de multimérisation (GagWM). Images confocales (A-F) de cellules HeLa fixées au PFA 4 % à 12 h post transfection pour l'étude de la reconstitution de fluorescence pour le mutant GagWM couplé à l'eGFP (A) ou à la split-GFP pour une association intermoléculaire (B) ou intramoléculaire (C). L'échelle indique une distance de 20 μ m. (D) Représentation des ratios $R_{m/c}$ obtenus pour GagWM-eGFP (N = 3, n = 42) et la reconstitution de la split-GFP pour Gag WM (INTER N = 2, n = 20 ; INTRA N = 3, n = 39). La boîte représente la médiane (trait horizontal au centre), 75 % des échantillons (trait horizontal supérieur) et 25 % des échantillons (trait horizontal inférieur). La moyenne inscrite au-dessus de chaque diagramme est indiquée par un carré blanc et les barres d'erreurs montrent l'écart-type à la moyenne. INTER correspond à la transfection de 2 plasmides codant pour Gag couplée à un seul fragment de split-GFP, alors qu'INTRA correspond à une protéine Gag couplée à la fois à GFP11 et à GFP1-10. (E) Schéma des différentes protéines de Gag utilisées dans cette étude. Les valeurs p obtenues dans un test HSD de Tukey sont données avec *** $p < 0,001$.

c) Régions de Gag impliquées dans la formation de la forme compacte

Pour que Gag puisse adopter sa forme compacte, il existe au sein de la protéine, une ou des région(s) flexibles permettant la forme en U de la protéine. Plusieurs régions flexibles ont été identifiées :

- Lien entre MA et CA : ce domaine, peu structuré (Tang et al., 2002) a été utilisé comme site d'insertion de marqueurs fluorescents protéiques de grande taille (eGFP, mCherry) permettant de suivre la dynamique de la protéine Gag dans la cellule (Muller et al., 2004). Malgré l'insertion d'une protéine fluorescente de taille importante au niveau de ce domaine, il apparaît que les fonctions de Gag sont marginalement perturbées permettant ainsi une production de particules virales infectieuses (Muller et al., 2004). C'est dans cet espace que nous avons inséré notre premier Multi-Cloning Site (MCS1) empêchant ainsi de tester l'importance de ce domaine dans la forme compacte de Gag. Dans les différentes études *in vitro* et *in silico* publiées pour l'étude de la forme compacte de Gag, la flexibilité de ce

domaine serait importante pour permettre le rapprochement des domaines MA et CA-NTD de Gag (Datta et al., 2007a; Munro et al., 2014).

- Lien entre les parties N- et C-terminales de la CA : CA est composée de deux domaines globulaires N- et C-terminaux (CAN et CAC respectivement) reliés par une région flexible de cinq acides aminés (146 à 150 de CA) (Berthet-Colominas et al., 1999; Momany et al., 1996). Ce domaine aurait pour rôle de permettre le rapprochement des domaines CAN et CAC de CA. Les acides aminés qui le composent ont montré leur importance dans l'assemblage de particules virales. Leurs mutations entraînent la production de particules peu ou pas infectieuses présentant des défauts dans la formation de la capsid conique après maturation qui est alors moins stable qu'une capsid sauvage (Jiang et al., 2011). Les acides aminés de ce domaine sont importants, au sein de la protéine CA mature, pour permettre la formation de dimères CAC/CAC ((Byeon et al., 2009). Pour vérifier le rôle de ce domaine, nous l'avons soit délété complètement (Δ linkCAC-CAN, CA délétée des acides aminés Y279, S280, P281, T282, S283, I284 et L285), soit remplacé par une série de prolines (282-TSIL/PPPP-285, Gag_link-PPPP) afin d'inclure artificiellement un repliement dans la protéine (de par sa chaîne latérale, une proline induit un angle de 30° dans l'axe de la protéine).
- Domaine de liaison SP1 situé entre la CA et la NC est le dernier domaine de Gag à être clivé lors de la maturation. Ce domaine n'est pas indispensable aux interactions Gag/Gag mais participe à la formation de multimères de hauts poids moléculaires (Datta et al., 2011b; Guo et al., 2005; Liang et al., 2002) et est nécessaire pour la production de particules virales immatures (Kräusslich et al., 1995). Principalement non structuré dans la protéine Gag monomérique (Newman et al., 2004), ce domaine se structure en hélice lors de la multimérisation de Gag. Cette structure jouerait un rôle dans le contrôle du degré de multimérisation et stabiliserait l'assemblage de Gag sur le génome viral ((Schur et al., 2016; Wagner et al., 2016 ; Datta et al., 2011b). La flexibilité de SP1 dans la protéine Gag monomérique permettrait à la protéine Gag de se replier afin de lier l'ARN par les domaines MA et NC. Dans un second temps, une fois à la membrane, Sp1 par le switch de conformation pourrait permettre à la protéine d'adopter une forme linéaire afin de faciliter et stabiliser l'assemblage. En se basant sur cette hypothèse,

Manuel PIRES

nous avons ainsi choisi de supprimer ce domaine dans le mutant Δ Sp1, Δ A366-M379.

Ces trois mutants (Δ link, PPPP et Δ Sp1) présentent tous un défaut d'assemblage. En effet, lorsque la distribution intracellulaire de ces mutants fusionnés à l'eGFP est étudiée et comparée à celle de la forme WT de Gag (**Figure 69 A**), aucune accumulation de fluorescence à la membrane n'est observée (**Figure 72 A, D, G**) comme le révèlent les ratios $R_{m/c}$ proches de 1 ($0,91 \pm 0,15$ pour Gag_link-PPPP, $0,98 \pm 0,18$ pour Gag Δ Sp1 et $0,89 \pm 0,14$ pour Gag Δ link CAC-CAN) à comparer à $1,36 \pm 0,32$ pour GagWT (**Figure 72 J**). En ce qui concerne l'observation de la forme compacte de Gag par le système split-GFP, on remarque que les études intermoléculaires confirment partiellement les résultats obtenus avec la protéine GFP. Les mutants Sp1 et Δ link CAC-CAN présentent en effet des ratios $R_{m/c}$ de $1,73 \pm 0,38$ et $1,47 \pm 0,34$ proches de celui de Δ NC ($1,65 \pm 0,47$). Le mutant Gag Δ linkCAC-CAN est particulier. En effet, il présente un $R_{m/c}$ de $1,47 \pm 0,33$ en intermoléculaire, principalement lié au fait que la fluorescence cytoplasmique est forte (852 ± 415 comparée à celle du WT (524 ± 371)) ; **Figure 72 C**). Ceci suggère que ce mutant est capable de s'oligomériser mais pas d'aller à la membrane. Il en découle que dans l'étude intramoléculaire une forte proportion de la fluorescence cytoplasmique est probablement liée à une recombinaison intermoléculaire de la split-GFP. Seul le mutant Gag_link-PPPP a un ratio en intermoléculaire proche de celui de Gag WT ($2,03 \pm 0,64$).

Les études intra montrent que ces mutants sont capables, comme Gag sauvage, de se replier sous forme compacte (**Figure 72 E-F et H-I**). Les faibles ratios $R_{m/c}$ en intramoléculaire proche de 1 ($0,96 \pm 0,17$ pour Gag_link-PPPP et $0,97 \pm 0,25$ pour Gag Δ Sp1) sont en effet équivalents à celui de la Gag sauvage. Un faible rapport ($R_{m/c} = 0,97 \pm 0,18$) est également observé pour Gag Δ linkCAC-CAN. Ce rapport peut être sous-évalué du fait de la forte oligomérisation intermoléculaire de ce mutant dans le cytoplasme. On observe un resserrement des valeurs de $R_{m/c}$ inter et intra pour ce mutant comme pour les mutants où l'interaction avec les acides nucléiques est perturbée (Δ NC, KK-TT, KK-TT_ Δ NC, voir **Figure 70**).

Cette étude suggère ainsi que le mutant Δ Sp1 ne limite pas le repliement de Gag en forme de U. Il est possible que ce domaine soit plutôt impliqué dans le retour à une forme linéaire de Gag à la membrane. Ce retour ne pouvant pas être suivi avec le système split-GFP (la reformation de la fluorescence est irréversible), ceci ne reste actuellement qu'une hypothèse. Pour le mutant Δ linkCAC-CAN, nos données ne donnent pas de réponse claire. L'insertion dans

Manuel PIRES

ce même domaine de plusieurs résidus proline n'induit pas artificiellement la forme repliée. Ceci peut être expliqué par le fait que la mutation de quatre résidus en proline (T282P, S283P, I284P, L285P) associée à la proline281 existante apporte un angle important (150°) favorisant plutôt la forme linéaire de Gag que la forme compacte. La délétion du domaine Sp1 ne modifie pas non plus la compaction de la protéine.

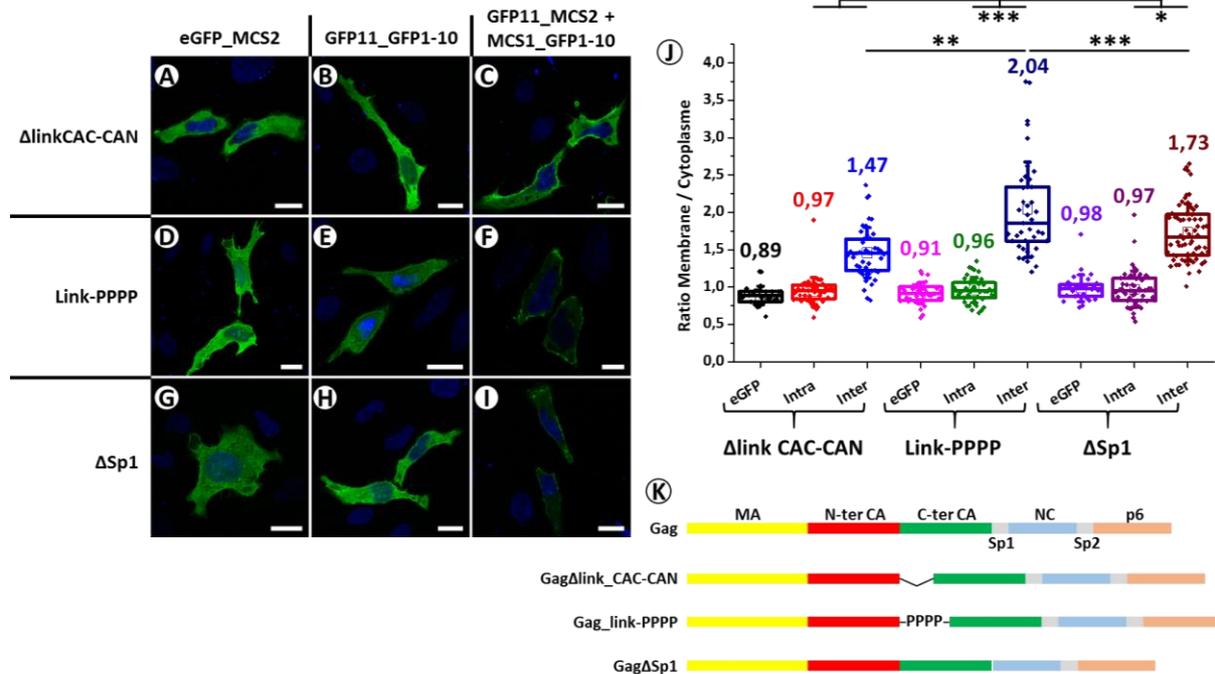


Figure 72 : Etude de mutants de la région liant les domaines CAN et CAC de CA (*GagΔlink CAC-CAN*, *GagPPPP*) et de la région Sp1 (*GagΔSp1*). Images confocales (A-I) de cellules HeLa fixées au PFA 4 % à 12 h post transfection pour l'étude de la reconstitution de fluorescence pour le mutant *GagΔlink CAC-CAN* (A à C), *GagPPPP* (D à F) et *GagΔSp1* (G à I). L'échelle indique une distance de 20 μm. (J) Représentation des ratios Rm/c obtenus avec l'eGFP ou pour la reconstitution intra- et intermoléculaire pour *GagΔlink CAC-CAN* (inter N = 3, n = 41; intra N = 3, n = 39), *GagPPPP* (inter N = 3, n = 41; intra N = 3, n = 41) et *GagΔSp1* (inter N = 4, n = 66; intra N = 4, n = 44). Chaque point représente une cellule (issue de N = 3 ou 4 expériences indépendantes). La boîte représente la médiane (trait horizontal au centre), 75 % des échantillons (trait horizontal supérieur) et 25 % des échantillons (trait horizontal inférieur). La moyenne inscrite au-dessus de chaque diagramme est indiquée par un carré blanc et les barres d'erreurs montrent l'erreur type de la moyenne. Inter correspond à la transfection de 2 plasmides codant pour Gag couplée à un seul fragment de split-GFP, alors qu'intra correspond à une protéine Gag couplée à la fois à GFP11 et à GFP1-10. (K) Schéma des différentes protéines de Gag utilisées dans cette étude. Les valeurs p obtenues dans un test HSD de Tukey sont données avec *** p < 0,001 ; ** p < 0,01 et * p < 0,5.

4. Approche biochimique de la forme compacte

Les expériences de microscopie de fluorescence ont permis de suivre la distribution des différentes formes fluorescentes exprimées dans la cellule. L'interprétation des données obtenues dans le cas de Gag fusionnée à l'eGFP pleine longueur est assez facile mais bien plus complexe dans le cas de l'utilisation du système split-GFP. Les différents contrôles nous

Manuel PIRES

permettent d'être assez confiants dans le fait que le signal cytoplasmique pour la plupart des mutants est lié à une reconstitution intramoléculaire de Gag. Toutefois, nous voulions confirmer l'existence dans les cellules de cette forme compacte par des approches biochimiques.

Nous avons donc, dans un premier temps, réalisé une analyse par électrophorèse en conditions natives, des formes de Gag exprimées dans les cellules lors de l'expression de protéines Gag dans les configurations de reconstitution d'une split-GFP en intra- et intermoléculaire. Ces analyses ont été couplées à un fractionnement cellulaire afin de séparer la fraction membranaire de la fraction cytoplasmique. Dans un second temps, des analyses par spectrométrie de masse couplée à des pontages moléculaires (« cross-link ») ont été mises en œuvre.

a) *Migration en conditions natives*

Habituellement, l'électrophorèse en gel dénaturant de polyacrylamide (SDS-PAGE) permet la séparation de protéines en fonction de leurs poids moléculaires (structure primaire uniquement). En conditions natives, la protéine conserve ses structures secondaires, tertiaires et quaternaires ce qui permet une séparation basée sur sa charge nette, sa taille et sa forme. La charge globale d'une protéine dépend de son point isoélectrique (pI). Le pI d'une protéine correspond au pH auquel sa charge globale est nulle ne lui permettant alors pas de migrer dans des conditions natives. L'utilisation d'un pH supérieur au pI d'une protéine lui donnera une charge globale négative lui permettant de migrer vers le pôle positif du système de migration. Les pIs des protéines que nous étudions étant relativement élevés (*Tableau 13*), nous avons choisi un tampon de migration avec un pH fortement basique (pH 10,2) afin que toutes les protéines soient chargées négativement et donc limiter l'impact de la différence des pIs entre protéines sur les migrations observées (David E; Garfin, 2003). La conservation des structures tridimensionnelles des protéines permet également l'observation directe de la fluorescence de la GFP dans le gel avant de réaliser le transfert des protéines sur une membrane PVDF, ce qui est très utile pour l'étude de la reconstitution de la protéine GFP.

	Point Isoélectrique (pI)	Masse moléculaire (kDa)	Charges nettes globales (pH 10,2)
Gag-eGFP_MCS2	8,28	83,5	-55,6
Gag_GFP11_GFP1-10	8,80	85,4	-49,4
Gag_GFP11_MCS2	9,10	60,8	-28,8
Gag_MCS1_GFP1-10	9,01	82,6	-44,8

Tableau 13 : Points isoélectriques et masses moléculaires des différentes protéines étudiées. Les charges nettes globales sont calculées à l'aide du logiciel en ligne développé par C. Putnam, The Scripps Research Institute (<http://protcalc.sourceforge.net/>).

Comme pour les observations en microscopie, nous nous sommes attachés à séparer les protéines membranaires et cytoplasmiques étant donné que la forme compacte de Gag est principalement attendue dans le cytoplasme alors que les formes multimériques s'accumulent à la membrane. Pour ce faire, un fractionnement cellulaire a été réalisé pour séparer les fractions membranaire, cytoplasmique et nucléaire qui ont par la suite été analysées sur un gel en conditions natives (Sever et al., 2003). Le compartiment nucléaire n'étant pas important pour cette étude des formes de la protéine Gag, seuls les résultats d'électrophorèse des fractions membranaires et cytoplasmiques sont présentés.

Nous avons donc exprimé Gag-eGFP, Gag_GFP11_GFP1-10, Gag_GFP11_MCS2 ou Gag_MCS1_GFP1-10 ou co-exprimé Gag_GFP11_MCS2 et Gag_MCS1_GFP1-10 dans des cellules HeLa pendant 12 heures avant que ces dernières soient soumises à un fractionnement cellulaire et les fractions analysées sur un gel natif de protéines (**Figure 73**) ou sur un gel dénaturant. Chacun des gels est couplé à une révélation en western blot par un anticorps anti-Gag.

Le western blot associé au gel dénaturant permet de contrôler la bonne expression et récupération des protéines Gag d'intérêt dans chaque fraction. On observe que Gag-eGFP et Gag_GFP11_GFP1-10 sont fortement exprimées et retrouvées dans les deux fractions (**Figure 73**, western blot gel dénaturant, 1, 1' et 2, 2'), Gag_GFP11_MCS2 est fortement exprimée dans les cellules mais retrouvée principalement (si exprimée seule, **Figure 73**, western blot gel dénaturant, 3 et 3') ou exclusivement (si co-exprimée avec Gag_MCS1_GFP1-10, **Figure 73**, western blot gel dénaturant, 4 et 4') dans la fraction membranaire. Par contre l'expression de Gag_MCS1_GFP1-10 est faible et elle n'est détectée que dans la fraction membranaire (**Figure 73**, western blot gel dénaturant, 5). Sur la même membrane nous avons révélé la GAPDH qui

Manuel PIRES

nous permet de vérifier la qualité du fractionnement (peu de GAPDH dans le compartiment membranaire) et la régularité des dépôts protéiques (quantités de protéines déposées identiques).

La fluorescence dans le gel natif révèle la présence de formes fluorescentes de très hauts poids moléculaires (bloquées dans les puits de dépôt) pour les fractions membranaires et cytoplasmiques des cellules exprimant Gag-eGFP (**Figure 73**, puits 1 et 1'), Gag_GFP11_GFP1-10 (**Figure 73**, puits 2 et 2') et dans une moindre mesure Gag_MCS1_GFP1-10 + Gag_GFP11_MCS2 (**Figure 73**, puits 3 et 3'). On retrouve également des bandes de faibles poids moléculaires pour les fractions cytoplasmiques et membranaires des cellules exprimant Gag-eGFP (**Figure 73**, puits 1 et 1') et Gag_GFP11_GFP1-10 (**Figure 73**, puits 2 et 2'). La présence de Gag dans ces bandes est confirmée par les résultats du western blot anti-Gag du gel natif associé (la superposition des signaux de fluorescence et du western blot est montrée en pointillés blancs). Les formes de hauts poids moléculaires sont certainement des multimères de Gag et sont logiquement observées en grande quantité dans la fraction membranaire. Les bandes de bas poids moléculaires correspondent certainement aux formes monomériques des protéines Gag. L'absence d'une telle bande dans le puits 3 confirme que la fluorescence observée par microscopie dans le cas de Gag_GFP11_GFP1-10 (**Figure 73**, puits 2 et 2') est liée à une recombinaison intramoléculaire de cette protéine qui permet de recréer une protéine GFP fluorescente. Le western blot associé au gel dénaturant montre également deux bandes de migration intermédiaire en grande quantité dans la fraction membranaire de Gag_GFP11_GFP1-10 (**Figure 73**, puits 3) dont la plus basse (**Figure 73**, pointillés bleus) est également retrouvée dans la fraction cytoplasmique. Des bandes proches des bandes hautes de Gag_GFP11_GFP1-10 (**Figure 73** puits 2, pointillés rouges) sont également retrouvées pour Gag_MCS1_GFP1-10 + Gag_GFP11_MCS2 (**Figure 73**, puits 3) et Gag_MCS1_GFP1-10 (**Figure 73**, puits 5) dans la fraction membranaire uniquement. Une hypothèse est que les formes hautes (pointillés rouges) correspondent à des dimères de Gag_GFP11_GFP1-10 (**Figure 73**, puits 2), de Gag_MCS1_GFP1-10 + Gag_GFP11_MCS2 (**Figure 73**, puits 3) ou de Gag_MCS1_GFP1-10 (puits 5) et la forme plus basse à la forme monomérique linéaire de Gag_GFP11_GFP1-10 (**Figure 73**, pointillés bleus). Le fait qu'aux bandes marquées par des pointillés rouges soient attribuées des oligomères est renforcé par le fait qu'ils sont retrouvés principalement dans la fraction membranaire. Dans cette hypothèse, la bande très basse (pointillés blancs) observée pour Gag-eGFP (**Figure 73**, puits 1 et 1') serait la forme linéaire de

Manuel PIRES

cette dernière. En effet, cette protéine a un pI de 8,28 (charge nette de -55,6 à pH 10,2) contre 8,80 pour Gag_GFP11_GFP1-10 (charge nette de -49,4 à pH 10,2), expliquant que Gag-eGFP va migrer plus rapidement que Gag_GFP11_GFP1-10. Il est possible que seule la forme linéaire de cette protéine soit visible dans ce gel. D'autres hypothèses sont possibles et il est difficile d'attribuer avec certitude une forme de Gag à chaque bande. Il est toutefois intéressant d'observer les différences importantes entre Gag_GFP11_GFP1-10 et Gag_GFP11_MCS2 + Gag_MCS1_GFP1-10 qui confortent nos conclusions en microscopie, notamment vis-à-vis de la forme compacte de Gag. Toutefois, l'impossibilité d'utiliser un marqueur de taille dans ce type de gel nous empêche d'attribuer avec certitude chacune de ces bandes. Il serait important de pouvoir analyser les protéines contenues dans les différentes bandes.

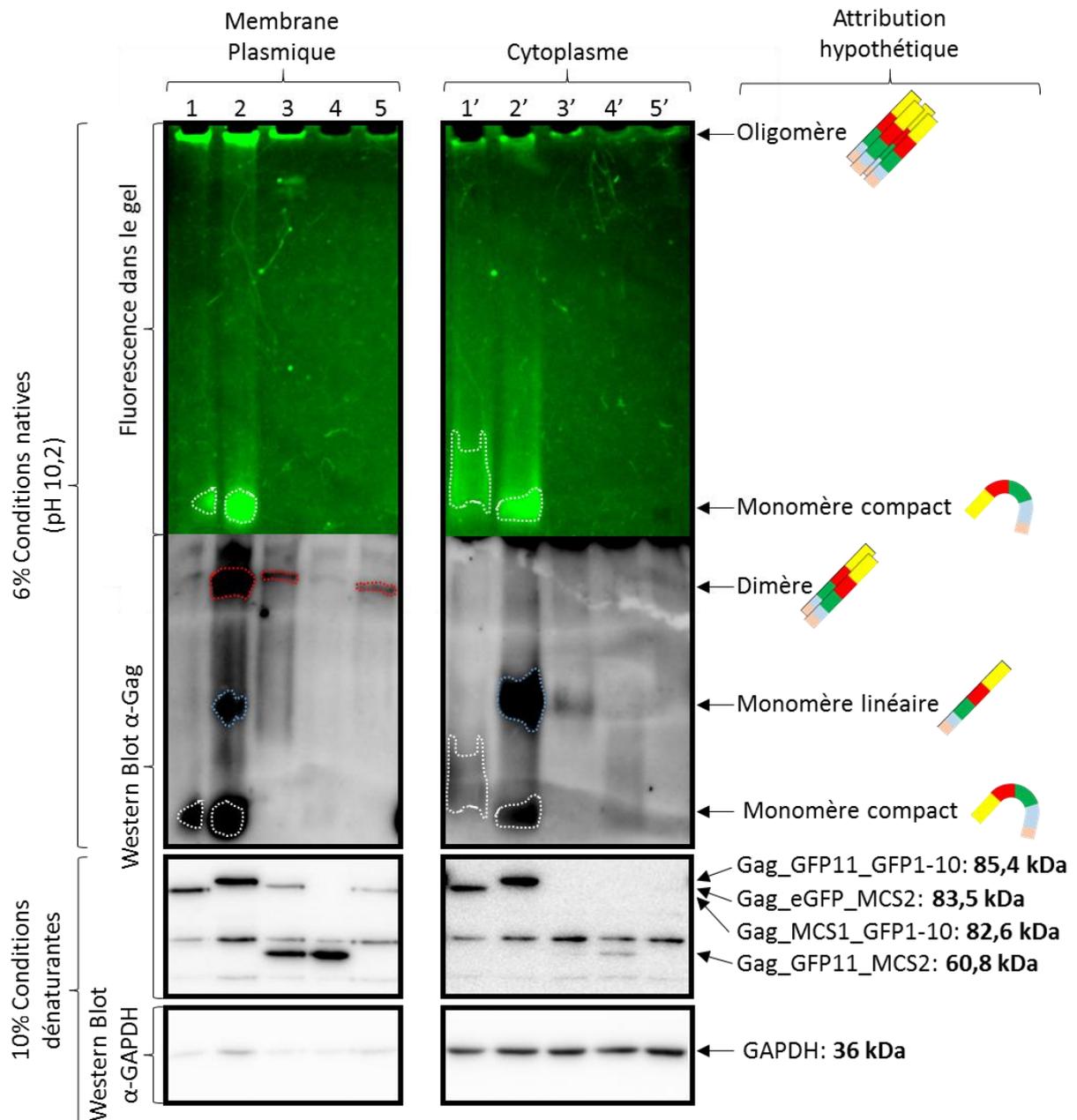


Figure 73 : Analyse par gels natif et dénaturant des formes observées pour diverses protéines Gag. Un fractionnement cellulaire à partir de cellules HeLa exprimant (1 et 1') Gag_eGFP_MCS2 ; (2 et 2') Gag_GFP11_GFP1-10 ; (3 et 3') Gag_GFP11_MCS2 + Gag_MCS1_GFP1-10 ; (4 et 4') Gag_GFP11_MCS2 ; (5 et 5') Gag_MCS1_GFP1-10 pendant 12 h a été réalisé pour séparer les fractions membranaire et cytoplasmique. Les fractions obtenues ont été séparées par un gel en conditions natives (gel 6 %, tampon de migration pH = 10,2). La fluorescence GFP a été observée (LAS4000) avant transfert sur une membrane PVDF et révélation par un western blot anti-Gag. En parallèle, une migration en conditions dénaturantes couplée à un western blot anti-Gag permet de vérifier la présence des protéines Gag dans les différentes fractions des divers échantillons. Un western blot anti-GAPDH permet de contrôler l'homogénéité des dépôts et la qualité du fractionnement.

Manuel PIRES

b) Utilisation d'agents pontant et analyse en spectrométrie de masse

L'impossibilité d'identifier avec certitude les formes de Gag présentes dans les différentes bandes du gel natif nous a dirigé vers une autre technique. Les protéines d'intérêts ont été exprimées par transfection de cellules HeLa et leurs différents états structurels bloqués par l'utilisation de molécules chimiques (cross-linkers) liant par une liaison covalente les fonctions amines des résidus lysines et arginines. Finalement, les différents fragments liés par ces pontages ont été analysés en spectrométrie de masse (technique bottom-up, (Sinz, 2006, 2014)). Les cross-linkers ne peuvent lier que des chaînes latérales de résidus lysines/arginines suffisamment proches (moins de 11,4 Å pour le BS3). Ainsi, si la reformation de la fluorescence est intramoléculaire, on peut espérer observer des peptides obtenus par le cross-link de résidus des extrémités N- et C-terminales de la même protéine rapprochés dans la forme compacte. Au contraire, si la fluorescence est reformée par recombinaison intermoléculaire, les fragments observés seront plus certainement obtenus par rapprochement des extrémités C-terminales ou N-terminales de deux protéines Gag. Une collaboration avec le Dr Sarah Sanglier-Cianferani de l'Institut Pluridisciplinaire Hubert Curien (IPHC, Laboratoire de Spectrométrie de Masse BioOrganique, Université de Strasbourg, CNRS UMR7178) a permis d'élaborer un protocole afin de réaliser cette étude (**Figure 74**).

Manuel PIRES

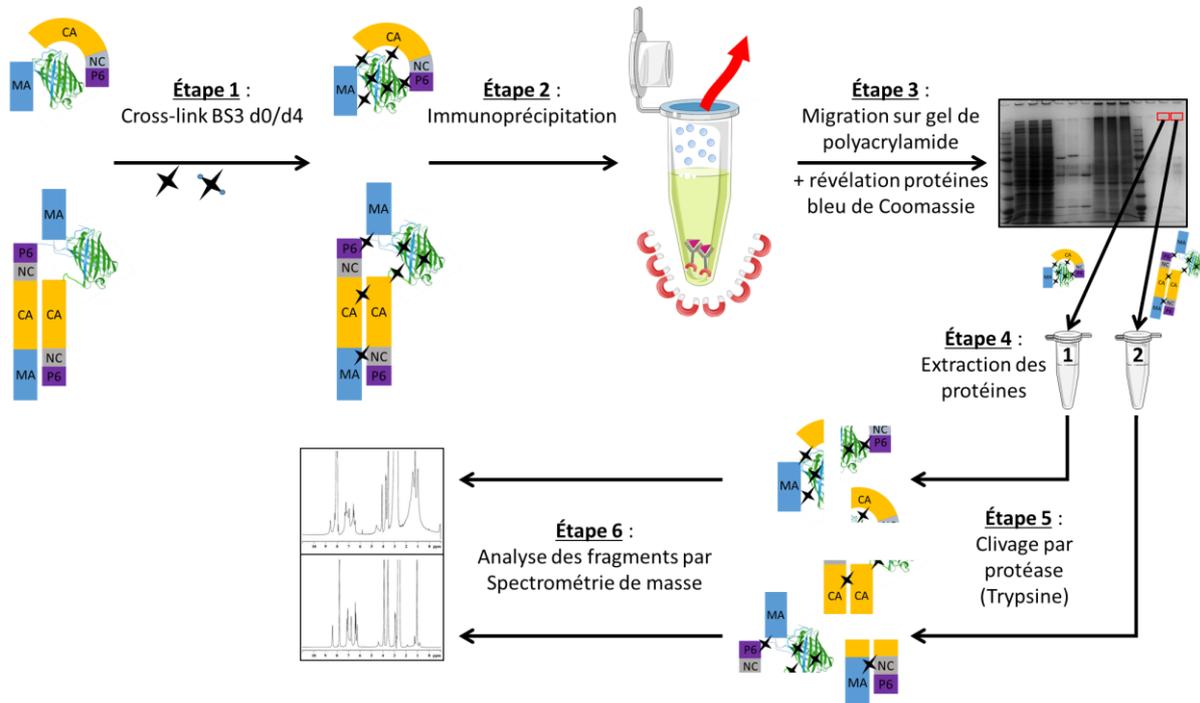


Figure 74 : Principales étapes de l'étude de la forme compacte de Gag par pontage moléculaire couplé à la spectrométrie de masse. Les différentes étapes de cette étude sont ici schématisées pour la protéine Gag-eGFP sauvage linéaire multimérique ou monomérique en forme compacte. L'étape 1 consiste à ponter les protéines retrouvées dans un lysat cellulaire par l'agent BS3. Étape 2 : les lysats sont ensuite immunoprécipités en utilisant un anticorps anti-GFP. Étape 3 : les immunoprécipitats sont séparés sur gel SDS-PAGE et les bandes d'intérêt découpées (étape 4), les protéines sont clivées par la trypsine (étape 5) et les fragments sont analysés par spectrométrie de masse (étape 6).

Les lysats protéiques ont été traités par un mélange de deux formes du même agent pontant BS3, une forme dite lourde (BS3-d₄), possédant dans sa structure quatre atomes de deutérium et une forme dite légère (BS3-d₀), possédant quatre hydrogènes. En effet, l'efficacité du pontage étant faible, la majorité des peptides obtenus après digestion protéique ne seront pas pontés. L'utilisation du mélange BS3-d₀/d₄ permet d'obtenir pour les peptides obtenus par pontage, des doublets d'intensité 1:1 séparés de 4 Da dans le spectre de masse (Müller et al., 2001; Sinz, 2014). Il est alors plus facile d'identifier ces peptides parmi la multitude de peptides non pontés et de les analyser plus en détail.

Nous avons préparé des lysats de cellules HeLa 12 h après transfection par des constructions codant Gag_eGFP_MCS2, Gag_GFP11_GFP1-10 ou Gag_GFP11_MCS2 + Gag_MCS1_GFP1-10. Les lysats ont été pontés ou non avant d'être immunoprécipités à l'aide d'un anticorps anti-GFP. Les lysats et les immunoprécipitats ont été analysés sur gel dénaturant SDS-PAGE, puis colorés au bleu de Coomassie ou transférés sur membranes pour une analyse en western-blot (WB) par un anticorps anti-Gag.

Manuel PIRES

Les pontages réalisés par le BS3 sont stables malgré les conditions dénaturantes du gel (DTT, chauffage à 95°C). Les profils de migration des protéines des lysats sans ou avec pontage sont très différents (comparer les puits 1, 2 et 3 avec les puits 7, 8 et 9 dans la **Figure 75 A**) ce qui montre l'efficacité de la réaction de pontage. On observe également dans le WB associé que Gag-eGFP et Gag_GFP11_GFP1-10 sont correctement exprimées et détectées par un anticorps anti-Gag (puits 1 et 2 du WB). Une immunoprécipitation a alors été réalisée afin de purifier spécifiquement les protéines d'intérêt par l'utilisation d'un anticorps dirigé contre un épitope conformationnel de l'eGFP (Velappan et al., 2010), plus sensible et plus spécifique qu'un anticorps anti-Gag (**Figure 75**, IP α -eGFP). Dans les puits 4 et 5, on observe, après immunoprécipitation, dans le gel coloré en bleu de Coomassie des bandes aux tailles attendues pour nos protéines d'intérêt pour les lysats non traités (encadré jaune, 83,5 kDa pour Gag_eGFP_MCS2 (puits 4) et 85,4 kDa pour Gag_GFP11_GFP1-10 (puits 5)). Dans le puits 6, aucune bande n'est observée en bleu de Coomassie pour Gag_GFP11_MCS2 (taille 60,8 kDa) et seule une bande de très faible intensité est visible pour Gag_MCS1_GFP1-10 (encadré jaune, taille 82,6 kDa) ce qui confirme la faible expression de ces protéines qui était observée dans le lysat correspondant (puits 3, WB). Dans le puits 6, après immunoprécipitation de ce lysat, on observe en WB, une légère bande correspondant à Gag_GFP11_MCS2 (encadré vert, 60,8 kDa) ainsi qu'une bande plus faible pour Gag_MCS1_GFP1-10 (82,6 kDa). Il est important de noter que l'anticorps utilisé pour l'IP est un anticorps conformationnel qui immunoprécipite une GFP correctement repliée. On peut donc penser que les protéines immunoprécipitées sont des protéines Gag_GFP11_MCS2 et Gag_MCS1_GFP1-10 associées sous formes de dimères ou d'oligomères permettant la reconstitution d'une GFP correctement repliée. En ce qui concerne les lysats traités par le BS3, on observe toujours sur le western blot les bandes correspondant à Gag-eGFP et Gag_MCS1_GFP1-10 (encadré orange) mais également des bandes de plus hauts poids moléculaires (encadré bleu puits 7 et 8). Ces dernières sont également visibles dans les lysats pontés exprimant Gag_GFP11_MCS2 + Gag_MCS1_GFP1-10 (encadré bleu, puits 9). À nouveau, les monomères ne sont pas observés, confirmant la spécificité de l'anticorps anti-GFP pour une GFP correctement repliée. Après pontage et immunoprécipitation, on observe la perte totale de signaux en coloration au bleu de Coomassie (puits 10 à 12). Toutefois, grâce au WB, on remarque que l'anticorps anti-GFP a permis de précipiter des protéines Gag qui migrent à la hauteur des formes monomériques de Gag_eGFP (encadré rose, puits 10) et Gag_GFP11_GFP1-10 (encadré rose,

Manuel PIRES

puits 11) et qui sont enrichies dans l'immunoprécipitat (comparer les puits 10 avec 7 et 11 avec 8). On observe (surtout pour Gag_eGFP) que la bande est plus large dans les IP avec pontage que sans pontage (comparer les puits 10 et 4) ce qui suggère que cette protéine est potentiellement liée à des partenaires cellulaires. Ce WB montre également la présence de formes de hauts poids moléculaires immunoprécipitées par l'anti-GFP et détectées par l'anti-Gag qui correspondent à des formes de Gag multimériques ou pontées à des partenaires de hauts poids moléculaires. Ce signal est spécifique car non retrouvé dans le puits 12 qui ne montre pas de détection spécifique de Gag_GFP11_MCS2 ou Gag_MCS1_GFP1-10 et ne présente pas ce signal diffus en western blot.

Nous souhaitons analyser en spectrométrie de masse la bande correspondant à la protéine Gag monomérique obtenue après immunoprécipitation anti-GFP et observable en western blot (*Figure 75*, encadré rose, puits 11). Toutefois, il nous fallait pour cela une quantité suffisante de protéines correspondant à une quantité détectable par le bleu de Coomassie, ce qui n'est pas le cas ici. Nous devons donc augmenter les quantités obtenues de cette forme pour pouvoir découper la bande, extraire les protéines et analyser les peptides en spectrométrie de masse à l'IPHC (réalisé par Maxime Bourguet, étudiant en thèse IPHC, Strasbourg).

Afin de mettre au point les expériences de spectrométrie de masse, nous avons réalisé les analyses sur les seules bandes visibles sur gel d'immunoprécipitation après pontage et colorées au bleu de Coomassie, à savoir les bandes correspondant aux multimères (encadré rouge *Figure 75*).

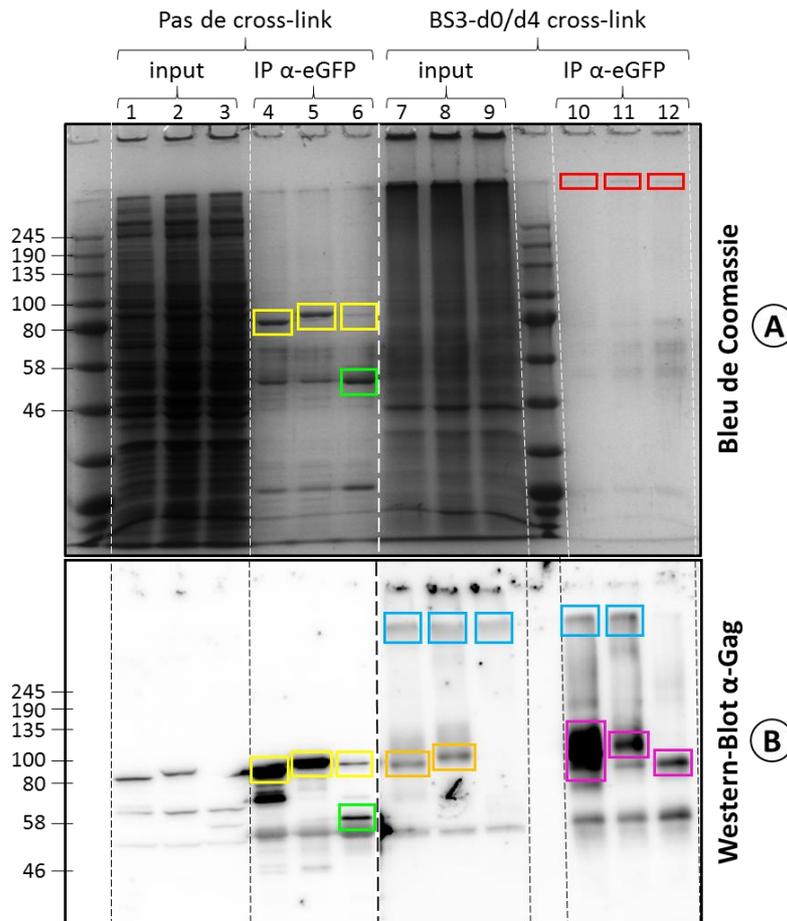


Figure 75 : Pontage moléculaire de lysats cellulaires exprimant diverses protéines Gag. Des cellules HeLa exprimant Gag_eGFP_MCS2 (1, 4 (cadre jaune), 7 et 10 (cadre rose)); Gag_GFP11_GFP1-10 (2, 5 (cadre jaune), 8 et 11 (cadre rose)) ou Gag_GFP11_MCS2 + Gag_MCS1_GFP1-10 (3, 6 (cadre vert et jaune), 9 et 12 (cadre rose)) pendant 12h ont été lysées. Les lysats ont été traités (cross-link) ou non (pas de cross link) par un mélange de formes légère et lourde de l'agent pontant BS3 (d0/d4) avant d'être immunoprecipités par un anticorps anti-eGFP (IP- α -eGFP) ou pas (input). Les échantillons obtenus ont été analysés sur deux gels dénaturants SDS-PAGE (10 % pour le gel de séparation et 4 % pour le gel de concentration) qui ont soit été colorés au bleu de Coomassie (A) soit transférés sur une membrane de PVDF pour une révélation par un western blot anti-Gag (B). Les cadres bleus indiquent les bandes de hauts poids moléculaires résultant du cross-link des protéines Gag_eGFP_MCS2 (puits 7 et 10), Gag_GFP11_GFP1-10 (puits 8 et 11) et du mix Gag_GFP11_MCS2 + Gag_MCS1_GFP1-10 (puits 9). Les cadres jaunes, vert et rose indiquent les bandes correspondant respectivement à Gag_MCS1_GFP1-10, gag_GFP11_MCS2 et Gag_GFP11_GFP1-10. Les cadres rouges indiquent la partie du gel découpée puis envoyée à l'Institut Pluridisciplinaire Hubert Curien pour analyse en spectrométrie de masse.

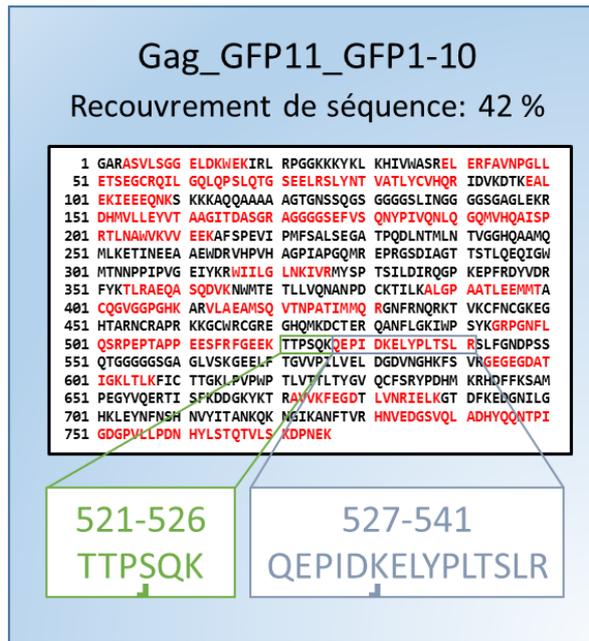
La faible quantité de protéines contenues dans ces bandes, couplée à la présence abondante de contaminants (kératine notamment) a rendu difficile la détection de peptides pontés. Un faible recouvrement de séquences a été observé : 42 % pour Gag_GFP11_GFP1-10 dans le cadre du puits 11 et seulement 17 % pour Gag_MCS1_GFP1-10 dans le cadre du puits 12, (Figure 76). Afin de valider un cross-link pour un peptide donné, il est nécessaire que les proportions du peptide ponté par l'agent lourd (BS3-d4) et par l'agent léger (BS3-d0) soient

Manuel PIRES

proches de 50/50. Dans l'échantillon du puits 11 (intramoléculaire), un seul peptide a été validé. Il résulte d'une liaison covalente entre deux lysines portées par deux fragments du domaine Sp6 de Gag (**Figure 76**, cadre du haut). Cependant, nous ne pouvons pas déterminer si cette liaison est réalisée entre des fragments de la même protéine ou deux protéines différentes. La proximité de ces deux lysines n'est toutefois pas décrite dans la structure du domaine p6 de Gag, laissant supposer une interaction intermoléculaire.

Dans l'échantillon issu de l'analyse intermoléculaire, un seul cross-link a également été validé montrant une liaison intermoléculaire entre deux lysines portées par deux peptides identiques (**Figure 76**, cadre du bas). Il est normal dans ces bandes de hauts poids moléculaires de retrouver des associations intermoléculaires. La limite de notre étude est le faible nombre de peptides pontés retrouvés dans nos échantillons qui est lié à la faible quantité de protéines dans ces échantillons. Il faudra donc augmenter les quantités de protéines afin de pouvoir réaliser la même étude sur les bandes de plus faible poids moléculaire et espérer obtenir un meilleur recouvrement de séquences et un nombre suffisant de peptides pontés pour pouvoir comparer ceux obtenus en intra- et en intermoléculaire.

Cross-link intra- ou intermoléculaire



Cross-link intermoléculaire

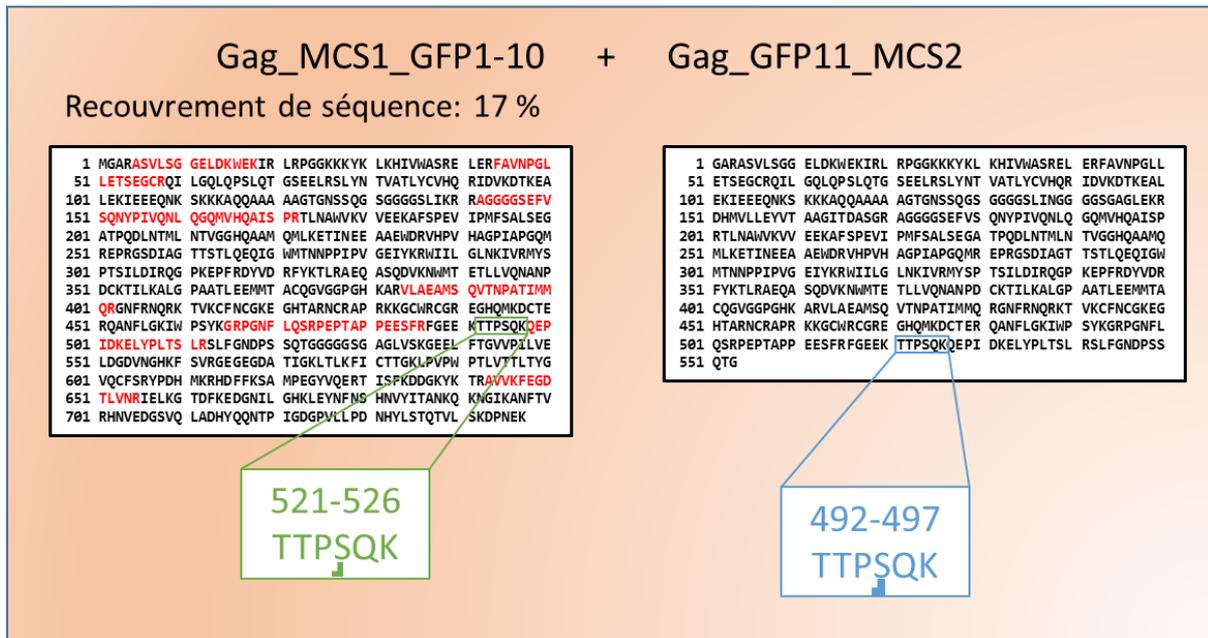


Figure 76 : Identification des peptides pontés obtenus après traitement au BS3 et immunoprécipitation anti-eGFP. Les échantillons pontés par BS3-d0/d4 ont été immunoprécipités par un anticorps anti-eGFP et séparés sur gel de polyacrylamide 10 % (voir Figure 74). Les bandes d'intérêt ont été découpées et les protéines ont été éluées et digérées par de la trypsine. Les fragments obtenus ont été analysés par LC-MS/MS (Orbitrap Q-Exacte Plus) et les fragments cross-linkés ont été validés lorsque le ratio cross-link lourd (BS3-d4) et cross-link léger (BS3-d0) est proche de 0,5 (expérience de MS réalisée par Maxime Bourquet, CNRS UMR7178).

B. Analyse en spectrométrie de masse par mobilisation ionique

L'existence d'une forme compacte de Gag a été montré *in vitro* (Datta et al., 2007a, 2011a; Munro et al., 2014). Toutefois ces études utilisaient des formes tronquées de Gag, sans domaine p6. C'est pourquoi, en parallèle de la démonstration de l'existence de la forme compacte *in cellulo* avec une protéine Gag pleine longueur, nous avons souhaité prouver que cette forme compacte existait aussi *in vitro* pour une protéine Gag pleine longueur purifiée et non modifiée. Pour cela, une seconde collaboration a été mise en place avec le Dr Sarah Sanglier-Cianferani afin d'analyser par spectrométrie à mobilisation ionique les différentes formes de la protéine Gag en solution.

La première étape a été de réaliser la purification de la protéine Gag selon le protocole de McKinstry et al., 2014 adapté au laboratoire par Kamal Sharma et Hassan Karnib. Afin de permettre sa purification, la protéine Gag a été fusionnée à une étiquette composée de six histidines (6-His-Tag, **Figure 77 A**). La transcription du cDNA correspondant est placée sous le contrôle du promoteur de l'ARN polymérase du phage T7 dont l'expression est induite dans les bactéries par l'ajout d'IPTG (isopropyl β -D-1-thiogalactopyranoside, analogue de l'allolactose) qui inhibe l'activité du répresseur de la transcription qui est physiologiquement fixé sur le promoteur de la β -galactosidase (*lacZ*), l'opéron *lac* (**Figure 77 B**). La présence de l'étiquette histidine permet de réaliser une première étape de purification par affinité sur une colonne de Nickel; les protéines étant fixées sur la colonne grâce à la forte affinité entre le Nickel et les résidus histidines (Hochuli et al., 1987, 1988). Une élution par l'application d'une concentration croissante d'imidazole permet ensuite de « décrocher » la protéine de la colonne par compétition (l'affinité du Nickel pour l'imidazole étant très forte, **Figure 77 C**). Suite à l'élimination de l'imidazole grâce à une dialyse, une seconde étape de purification par une colonne d'exclusion permet de purifier plus finement la protéine Gag (**Figure 77 D**). Cette seconde colonne permet d'éliminer les contaminants de hauts poids moléculaires (fractions 22 à 26, **Figure 77 D**) et de faibles poids moléculaires (fraction 36 à 46, **Figure 77 D**). Les fractions d'intérêts 27 à 34 sont mélangées et la protéine purifiée est concentrée. Sa capacité à déstabiliser les acides nucléiques est mesurée afin de valider la purification (**Figure 35** et **Figure 77 E**). Ce protocole nous permet d'obtenir de grandes quantités de protéine Gag sauvage pleine longueur, active pour la déstabilisation des acides nucléiques.

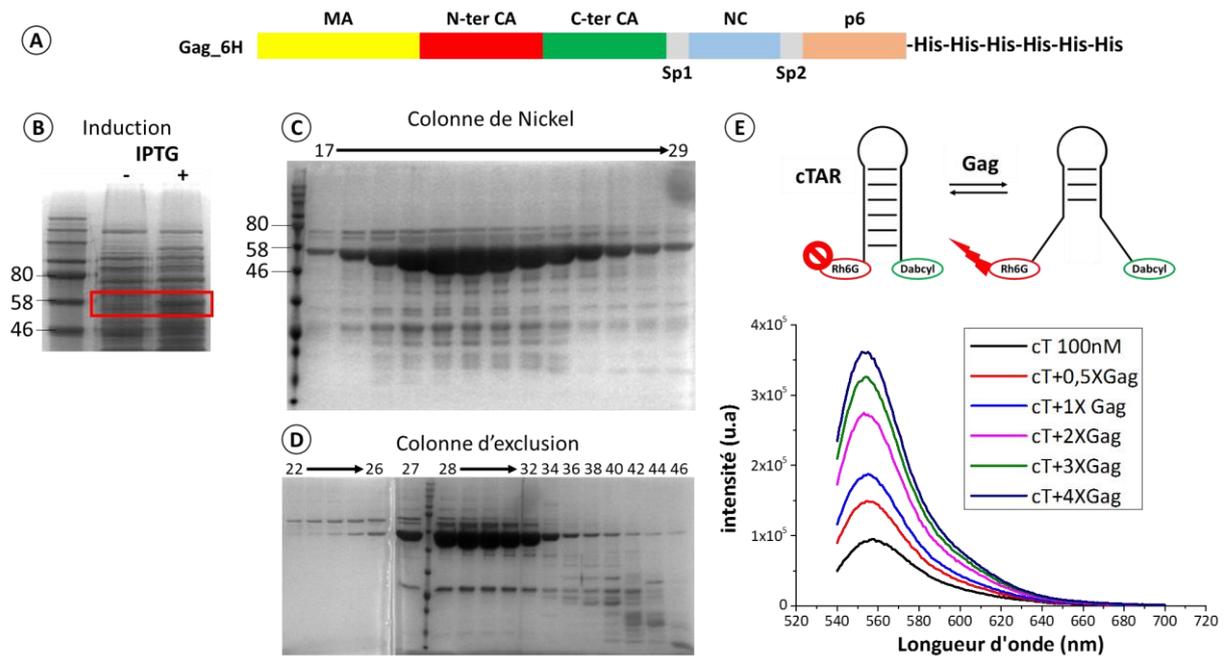


Figure 77 : Purification de la protéine Gag et vérification de son activité. (A) Schéma de la protéine Gag purifiée portant à son extrémité C-terminale une étiquette portant six résidus histidine. (B) Induction de l'expression de la protéine Gag par ajout de 0,5mM d'IPTG pendant 18h à 18°C. Les lysats bactériens avant et après induction sont analysés sur gel SDS-PAGE 12 % coloré au bleu de Coomassie. La protéine Gag_{6H} d'une masse molaire de 56,9 kDa est signalée par le rectangle rouge. (C) Analyse par gel SDS-PAGE 12 % coloré au bleu de Coomassie des fractions obtenues après élution de la colonne d'affinité au Nickel ; les fractions 17 à 29 sont mélangées et soumises à une dialyse pour éliminer l'imidazole, puis passées dans une colonne d'exclusion pour améliorer la purification. Les différentes fractions d'élution sont analysées par gel SDS-PAGE de 10 % coloré au bleu de Coomassie (D). Les fractions contenant des contaminants protéiques de poids moléculaires plus importants (fractions 22 à 26) ou plus faibles (fractions 36 à 46) que Gag (fractions 27 à 34) sont ainsi éliminées. (E) L'activité chaperonne de la protéine Gag purifiée est testée en mesurant sa capacité à déstabiliser la boucle cTAR (cT, 100nM, concentrations croissantes de Gag de 50nM (0.5XGag) à 400nM (4XGag)), le test est similaire à celui décrit dans la Figure 35 (Bernacchi and Mély, 2001; Bernacchi et al., 2002).

Comme les protéines Gag ont été purifiées dans un tampon riche en sel (1 M NaCl), les conditions ne sont pas optimales pour l'analyse en spectrométrie de masse. Une diminution de la concentration en sels dans le tampon contenant les protéines permet en général d'affiner l'analyse (Constantopoulos et al., 1999). En effet, le principe de la spectrométrie de masse est de séparer, en phase gazeuse, des molécules chargées, en fonction de leur rapport masse/charge. La concentration en sel peut modifier la charge des protéines rendant alors l'analyse plus difficile. Nous avons donc essayé de diminuer la concentration de NaCl à 0,5 M et 0,15 M lors de la purification. L'analyse sur gels de polyacrylamide colorés au bleu de Coomassie de ces purifications a révélé que la purification était moins performante qu'avec le protocole de référence (Figure 78). En effet, si la purification de référence permet de n'observer qu'une seule bande (correspondant à la protéine Gag_{6H}), de nombreuses bandes

Manuel PIRES

apparaissent lors de la diminution de la quantité de sel dans le tampon. Afin de conserver la protéine la plus pure possible pour l'analyse en spectrométrie de masse, nous avons décidé, malgré la quantité importante de sel dans le tampon, de poursuivre nos analyses avec une protéine Gag purifiée selon le premier protocole de purification contenant 1 M de NaCl.

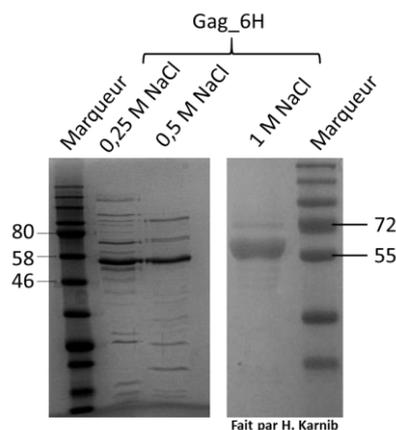


Figure 78 : Observation de protéines Gag purifiées avec des tampons à quantité réduite en NaCl. Analyse sur gel de polyacrylamide 12 % coloré au bleu de Coomassie des résultats de purification de Gag en utilisant des tampons contenant 0,25 et 0,5 M de NaCl et 1 M de NaCl (réalisé par Hassan Karnib). La migration permet d'observer l'augmentation de la quantité de contaminants protéiques lors de la diminution de la concentration de sel dans le tampon (0,5 M et 0,25 M au lieu de 1 M dans le protocole de référence).

La spectrométrie de masse en mobilité ionique permet de différencier deux protéines dont la masse est identique en fonction de leurs surfaces de collision. Après évaluation de la masse en spectrométrie de masse classique, les protéines sont redirigées dans un tube rempli de gaz (tube de dérive). La collision de la protéine avec le gaz ralentit cette dernière. Le nombre de collisions est d'autant plus important que la surface de la protéine est grande. Une protéine Gag sous forme compacte présentant une surface de collision plus faible qu'une protéine Gag linéaire devrait donc sortir du tube de dérive plus rapidement. Par ailleurs, l'ajout d'ADN (polyA de 25 nucléotides) pourrait favoriser le repliement de la protéine Gag sous forme compacte. Ces expériences ont été réalisées par Thomas Botzanowski, étudiant en thèse dans le laboratoire du Dr Sarah Sanglier-Cianferani.

Dans des expériences très préliminaires, nous avons validé la purification de protéine, en retrouvant par spectrométrie de masse classique, une masse expérimentale correspondant à la masse théorique de Gag associée à deux ions zinc. Cette vérification a de plus contribué à valider que le changement de tampon (diminution de NaCl) est néfaste pour la protéine. Nos collaborateurs ont ensuite testé la mobilité ionique de la protéine Gag en absence d'AN. Un

Manuel PIRES

seul pic est apparu montrant l'existence d'une forme unique de Gag. La prochaine étape consistera à incuber les protéines avec du polyA. Cet oligonucléotide est utilisé pour mimer les acides nucléiques autour desquels la protéine Gag pourra se replier. Nous espérons observer un second pic correspondant à la forme repliée de Gag.

Aucun résultat concernant cette étape n'est présenté ici car bien qu'encourageants ceux-ci sont trop préliminaires.

C. Tentatives d'observations de la forme compacte avec d'autres techniques de microscopie

1. Observation de l'évolution de la fluorescence par vidéo-microscopie

La vidéo-microscopie de fluorescence est une des techniques disponible au sein de la plateforme d'imagerie du laboratoire. Grâce à une micro-injection d'ADN plasmidique codant des protéines fluorescentes directement dans le noyau, il est possible de suivre l'apparition et l'évolution au cours du temps des protéines fluorescentes. Cette technique permet donc d'observer les protéines très rapidement après leur synthèse. Comme nous pensons que la forme compacte correspond à une Gag monomérique et cytoplasmique avant qu'elle se multimérise et atteigne la membrane, cette technique nous a paru particulièrement bien adapté. C'est pourquoi, nous avons suivi au cours du temps la fluorescence obtenue par reconstitution d'une GFP par des fragments de la split-GFP portés par la même protéine Gag (intramoléculaire, Gag_GFP11_GFP1-10) ou par deux protéines Gag différentes (intermoléculaire, Gag_GFP11_MCS2 et Gag_MCS1_GFP1-10). Nous remarquons que la fluorescence apparait plus rapidement dans le cas de la recombinaison intramoléculaire (20 min, **Figure 79 A**) que dans le cas de la recombinaison intermoléculaire (30 min, **Figure 79 A**). De plus, la fluorescence intermoléculaire se concentre très rapidement à la membrane (60 min, **Figure 79 A**), en accord avec le fort marquage membranaire observé en microscopie confocale. La quantification de fluorescence au niveau membranaire ou cytoplasmique n'est pas possible car dans cette expérience, l'observation se fait par un microscope à épifluorescence qui collecte la fluorescence émise sur toute l'épaisseur cellulaire (**Figure 40 C**) et non uniquement dans un plan focal comme en microscopie confocale (**Figure 40 A**). Toutefois, la quantification de la fluorescence cellulaire totale (**Figure 79 B**) révèle que la fluorescence est plus intense dans le cas intramoléculaire. En effet, après 60 min

Manuel PIRES

d'observation, la fluorescence atteint 7000 U.A. dans le cas de la recombinaison intramoléculaire, alors qu'elle n'est que de 3000 U.A. pour la recombinaison intermoléculaire. Cette technique confirme les observations en microscopie confocale. Lorsque les deux fragments de la split-GFP sont insérés dans la même protéine Gag, la fluorescence apparaît plus rapidement et est plus intense en se localisant principalement dans le cytoplasme, ce qui est en accord avec une recombinaison intramoléculaire plus précoce que la recombinaison intermoléculaire membranaire qui peut être aussi observée avec ce type de construction. Dans la recombinaison intermoléculaire, la fluorescence apparaît surtout à la membrane, confirmant que cette fluorescence est principalement causée par la multimérisation de Gag. La micro-injection couplée à la vidéo-microscopie nous permet d'observer en temps réel l'apparition, la localisation et l'intensité de la fluorescence. Cette technique pourra également être utilisée avec les mutants (Δ NC, KK-TT, KK-TT_ Δ NC, Δ linkCAC-CAN) afin de comparer les différences de cinétiques d'apparition de fluorescence.

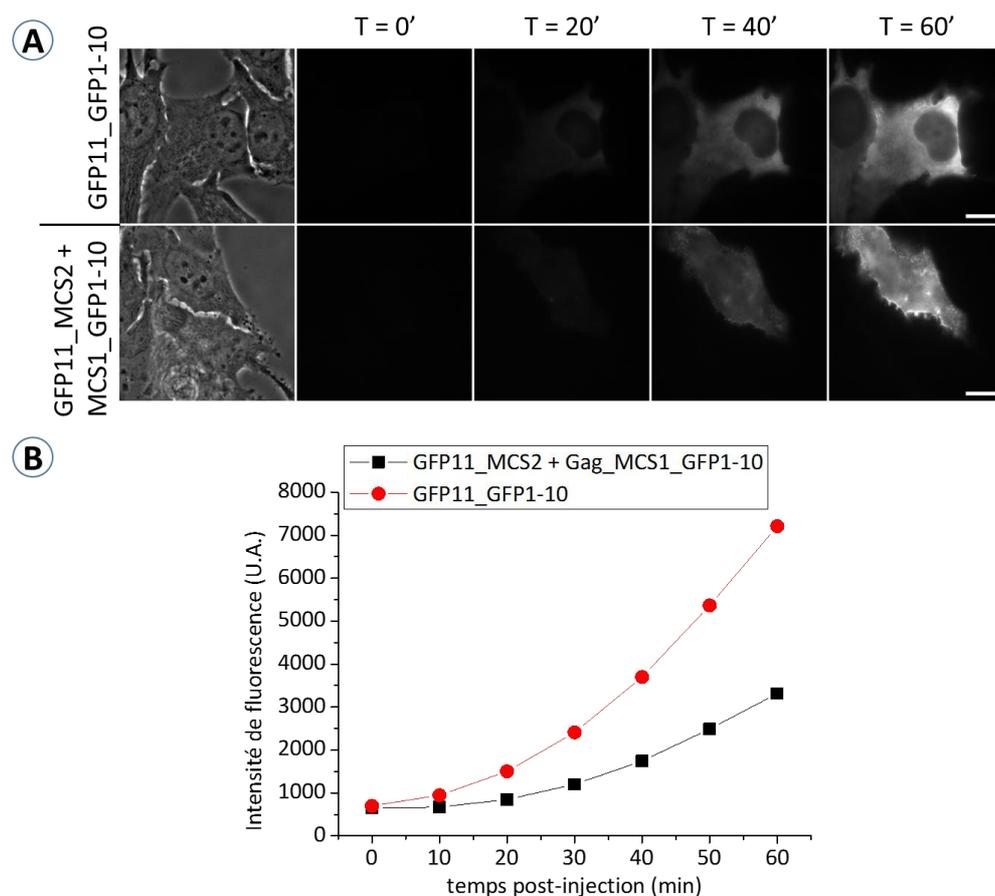


Figure 79 : Suivi de la fluorescence lors de la microinjection de plasmides codant Gag_GFP11_GFP1-10 ou un mix Gag_GFP11_MCS2 + Gag_MCS1_GFP1-10. (A) Suivi de la fluorescence en fonction du temps par vidéo-microscopie suite à la micro-injection d'ADN plasmidique (Gag_GFP11_GFP1-10 en haut et Gag_GFP11_MCS2 + Gag_MCS1_GFP1-10 en bas) dans le noyau cellulaire. La première colonne montre l'image en champ large de la cellule micro-injectée, puis les colonnes suivantes indiquent les images à 0 min, 20 min, 40 min et 60 min après la micro-injection. Les barres d'échelles correspondent à 20 μ m. (B) Mesure au cours du temps de la fluorescence cellulaire totale (Unité Arbitraire) observée sur cinq cellules différentes.

2. Étude de la forme compacte et détection de l'ARN spécifique

Une hypothèse est que la forme compacte de Gag serait impliquée dans la sélection de l'ARN viral pour son empaquetage préférentiel dans la particule virale lors du bourgeonnement. De plus, l'ARN viral portant des domaines spécifiques d'interaction avec Gag serait susceptible de stabiliser la forme compacte de Gag. Par conséquent, nous avons souhaité coupler l'observation de la forme compacte de Gag à celle de l'ARN viral. Pour cela, nous avons utilisé une technique qui permet d'observer l'ARN grâce à une reconnaissance spécifique de tige-boucles MS2 par la protéine de capsid du bactériophage MS2. Dans cette technique, l'ARN viral est modifié pour porter un certain nombre de tige-boucles dites MS2 (Figure 80 A) qui sont spécifiquement reconnues par la protéine de capsid du bactériophage MS2 (Figure 80 B)

Manuel PIRES

fusionnée à une protéine fluorescente (mCherry) permettant le suivi de l'ARN viral par microscopie de fluorescence (**Figure 80 C**). La structure de l'ARN modifié (Zennou et al., 2004), présente le site de liaison spécifique du domaine NC de Gag (séquence Ψ), une forme tronquée du gène *gag*, une série de 12 tige-boucles MS2 (une tige-boucle pouvant lier un dimère de MS2, un maximum de 24 protéines MS2 couplées à mCherry pourront se fixer sur l'ARN), les séquences codant les protéines Tat, Vpu et Rev et le site viral d'export nucléaire RRE. Ce dernier permet l'export nucléaire de l'ARN lors de la reconnaissance du site spécifique RRE par la protéine Rev. Il est donc indispensable de co-exprimer cette protéine pour observer l'ARN viral dans le cytoplasme. La protéine MS2_mCherry a été fusionnée à un site de localisation nucléaire (NLS, Nuclear Localisation Site) et est retrouvée dans le noyau (Jouvenet et al., 2009) où elle reconnaît l'ARN portant les tige-boucles MS2. Il est donc nécessaire de co-transfecter en plus du plasmide codant l'ARN spécifique, le plasmide codant Rev et le plasmide codant MS2_mCherry.

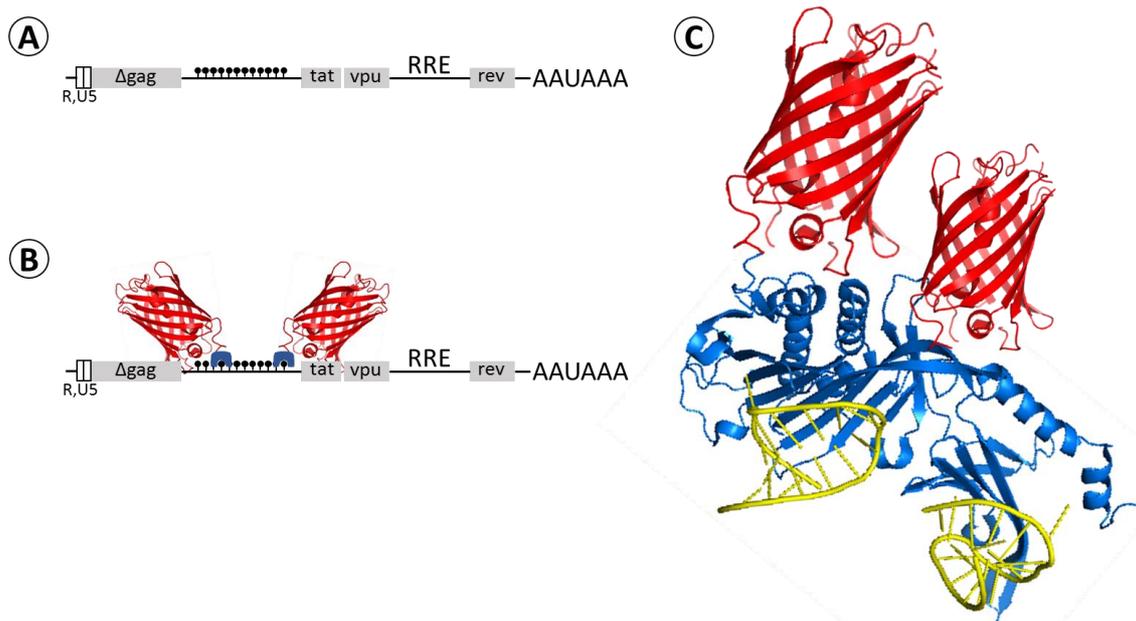


Figure 80 : Principe d'observation de l'ARN viral grâce à la reconnaissance de tige-boucles MS2 par la protéine de capsid du bactériophage MS2 fusionnée à une protéine fluorescente. (A) Séquence de l'ARN spécifique (plasmide pCMV-intro) modifié pour que sa localisation soit suivie par microscopie (Jouvenet et al., 2009; Zennou et al., 2004). Les tige-boucles MS2 sont indiquées par des cercles noirs. (B) Reconnaissance de l'ARN par liaison de deux protéines de capsid de MS2 couplées à la protéine fluorescente mCherry (chaque tige-boucle peut lier un dimère de MS2 soit un maximum de 24 MS2 par ARN). L'export nucléaire de l'ARN induit par Rev (qui se fixe sur le RRE) entraîne la sortie de la protéine MS2_mCherry ce qui permet de suivre la localisation de l'ARN. La protéine MS2 est en bleu et la mCherry en rouge. (C) Représentation tridimensionnelle de la reconnaissance de deux boucles MS2 (en jaune) par la protéine de capsid du bactériophage MS2 (en bleu, PDB 2BU1) couplée à la protéine mCherry (en rouge, PDB 2H5Q).

Manuel PIRES

Nous avons d'abord confirmé qu'en absence de la protéine Rev, la protéine mCherry_MS2 est localisée dans le noyau, que ce soit en absence (*Figure 81 B et E*) ou en présence de l'ARN viral (*Figure 81 C et F*) ou que la reconstitution de la split-GFP soit étudiée en intra- (*Figure 81 B et C*) ou en intermoléculaire (*Figure 81 E et F*). Ce n'est seulement que lorsque Rev, l'ARN viral et mCherry_MS2 sont co-exprimés que le signal MS2_mCherry est observé dans le cytoplasme, reflétant la localisation cytoplasmique de l'ARN viral et montrant la spécificité de reconnaissance de MS2 pour les tige-boucles (*Figure 81 A et D*). La localisation majoritairement cytoplasmique ou membranaire de la fluorescence issue de la recombinaison des fragments de la split-GFP en intramoléculaire (*Figure 81 A, B et C*) et intermoléculaire (*Figure 81 D, E et F*) respectivement reste conforme aux observations préalables. En revanche, l'ARN viral ne semble se concentrer à la membrane plasmique que dans le cas d'une recombinaison intramoléculaire (flèches blanches, *Figure 81 A*). Ces données suggèrent un recrutement de l'ARN à la membrane par la Gag compacte. Il est important de se rappeler qu'une Gag sauvage retrouve une forme linéaire à proximité de la membrane et ce par l'affinité plus importante du domaine MA pour les membranes que pour l'ARN. Or nous savons que les deux fragments de la GFP conservent une certaine affinité l'un pour l'autre ce qui est susceptible de « bloquer » une partie de Gag_GFP11_GFP1-10 sous forme compacte et ce même à proximité de la membrane. L'ARN viral retrouvé proche de la membrane pourrait être un ARN qui reste lié à cette forme compacte de Gag.

Afin de montrer que l'ARN viral est également associé dans le cytoplasme avec la forme compacte de Gag, nous pourrions envisager des mesures de transfert de fluorescence entre la GFP obtenue par reconstitution de la split GFP et la protéine mCherry portée par MS2. Il est important de noter que seul le site de reconnaissance du domaine NC de Gag est présent sur cet ARN. Il pourrait être intéressant d'ajouter le site de reconnaissance des acides aminés basiques de l'extrémité N-terminale de la MA (situés au niveau du gène *Pol*, Purohit et al., 2001) afin de stabiliser davantage la forme compacte. Il est également envisageable de coupler le fragment GFP1-9 de la split-GFP tripartite avec la protéine MS2 et d'utiliser dans ce cas une protéine Gag portant les deux autres fragments GFP10 et GFP11. Il serait alors peut-être possible de détecter l'ARN viral uniquement lorsqu'il est couplé à la forme compacte de Gag. Nous pourrions ainsi confirmer le rôle de la forme compacte de Gag dans la reconnaissance et le transport vers la membrane de l'ARN spécifique.

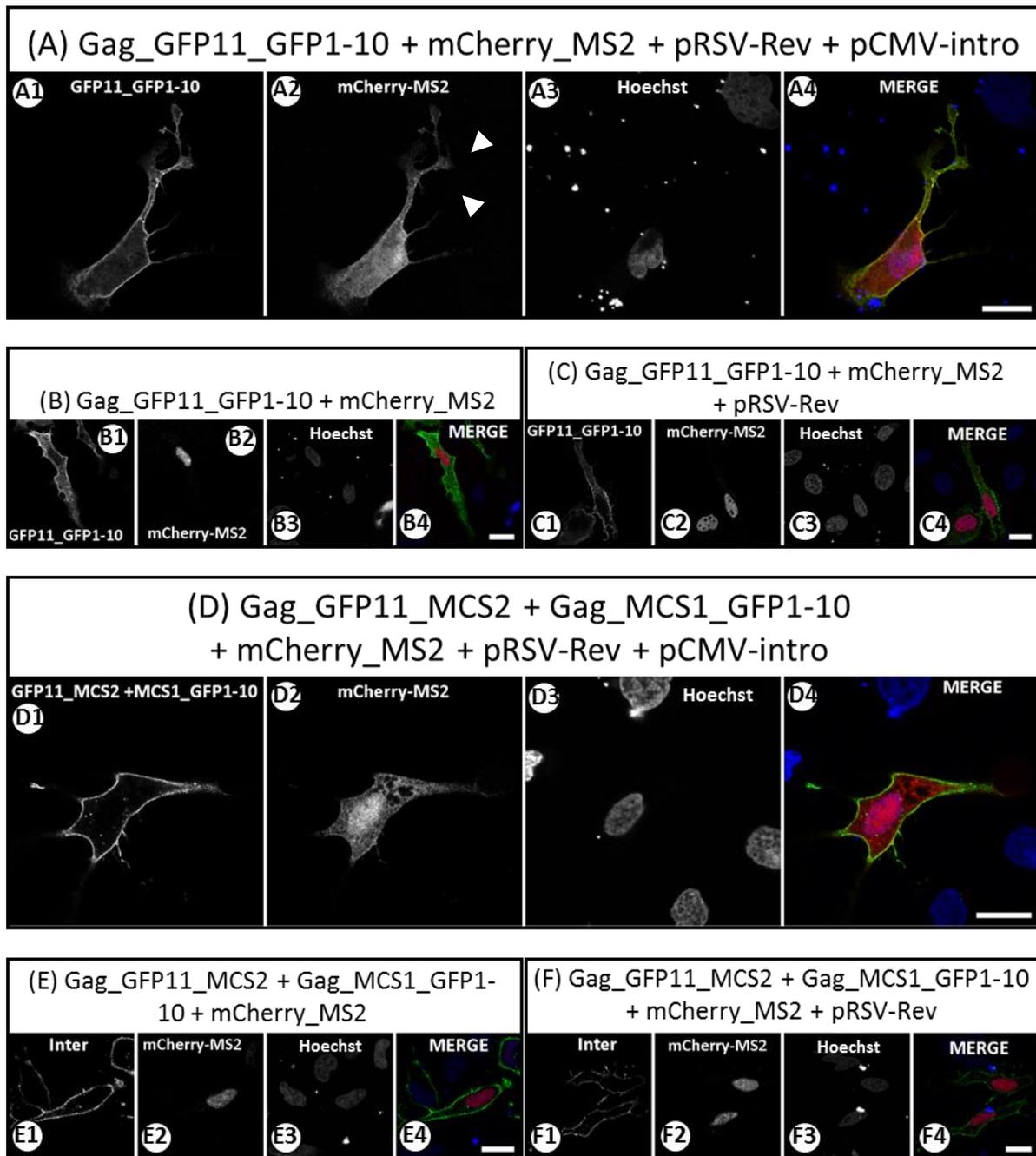


Figure 81 : Observation simultanée de la forme compacte de Gag et de l'ARN viral. Des cellules HeLa ont été transfectées par un mélange de plasmides codant mCherry_MS2 (0,5 µg) avec (A, B et C) gag_GFP11_GFP1-10 (0,5 µg) ou (D, E, F) (gag_GFP11_MCS2 (0,5µg) + gag_MCS1_GFP1-10 (0,5 µg)) en présence de plasmide codant pRSV-Rev (0,2 µg) (A et D) et/ou pCMV-Intro (0,25 µg) (A, C, D et F). Les images A4 et D4 présentent la superposition des fluorescences observées pour la GFP (A1, D1) en vert, la mCherry-MS2 (A2, D2) en rouge et Hoechst (A3, D3) en bleu. (B4) et (E4) montrent la superposition des signaux obtenus pour la GFP (vert, B1 et E1), la protéine mCherry_MS2 (B2 et E2) et Hoechst (B3 et E3) en absence de Rev et de l'ARN viral. (C4) et (F4) montrent la superposition des signaux obtenus pour la GFP (vert, B1 et E1), la protéine mCherry_MS2 (B2 et E2), le signal Hoechst (B3 et E3) en absence de Rev uniquement. Les barres d'échelles correspondent à 20 µm.

3. Création d'un split-DRONPA : observation de la forme compacte par détection en molécule unique

Afin de compléter l'étude de la forme compacte de Gag, nous avons souhaité utiliser la microscopie à haute résolution pour localiser un fluorophore unique. Alors qu'en microscopie conventionnelle, la résolution est de l'ordre de 200 nm à cause de la diffraction de la lumière émise par les fluorophores dans l'ouverture numérique (**Figure 82 A**), la microscopie de localisation de fluorophores photo-activés (PALM) permet d'observer seulement un faible nombre de fluorophores à chaque instant (**Figure 82 B**). Grâce à une analyse mathématique (courbe gaussienne bidimensionnelle tracée sur la fonction d'étalement du point (PSF = Point Spread Function)), il est possible de localiser individuellement chaque fluorophore et d'atteindre une résolution de l'ordre de 20 nm.

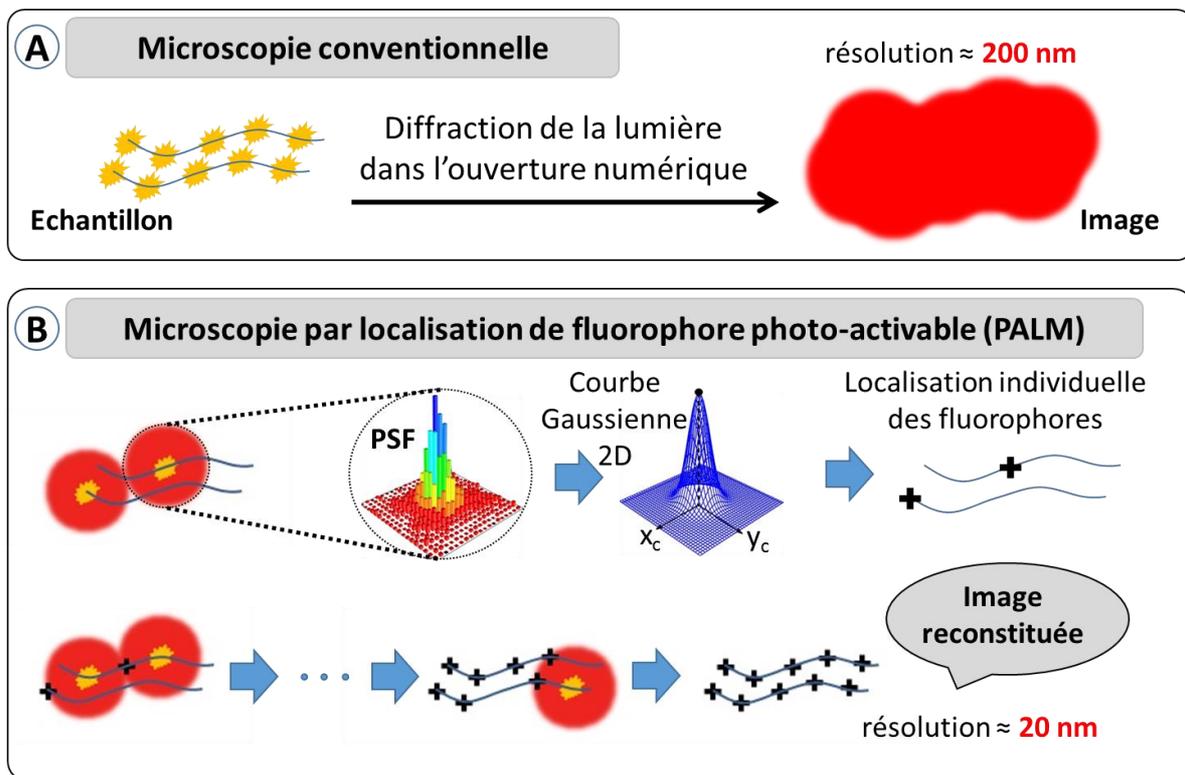


Figure 82 : Comparaison de la microscopie conventionnelle et de la microscopie par localisation de fluorophore photo-activable (PALM). PSF = Point Spread Function, pour fonction d'étalement du point (traduit d'une présentation d'Oleksandr Glushonkov, doctorant au laboratoire).

Pour mettre en place cette technique, il est nécessaire d'utiliser un fluorophore photo-activable qui passe d'un état non émissif à un état où il émet de la fluorescence pendant une courte période lors de l'illumination à la longueur d'onde d'activation. Lors de l'observation de ce type de fluorophores, un clignotement apparaît, chaque point correspondant à un fluorophore unique. Plusieurs catégories de fluorophores photo-activables existent, mais

Manuel PIRES

nous avons choisi d'utiliser la DRONPA. Cette protéine est deux fois plus brillante que l'eGFP et est issue de la modification d'une protéine trouvée chez des coraux de la famille des *pectinidae*. Le chromophore de la DRONPA (Tyr66, Gly64, Cys62) situé au cœur d'une structure tridimensionnelle en tonneau β est dit « photo-switchable » puisqu'il est possible « d'allumer » le fluorophore par une excitation à 405 nm, ou de « l'éteindre » par une excitation à 488 nm (**Figure 83**). Pour observer la fluorescence du chromophore « allumé », il est nécessaire d'exciter celui-ci à 488 nm, entraînant par le fait son extinction qui peut être de nouveau allumée par une excitation à 405 nm. Une seule molécule de DRONPA peut être allumée et éteinte plus de 100 fois. L'émission du fluorophore est donc temporaire permettant la localisation individuelle des fluorophores.

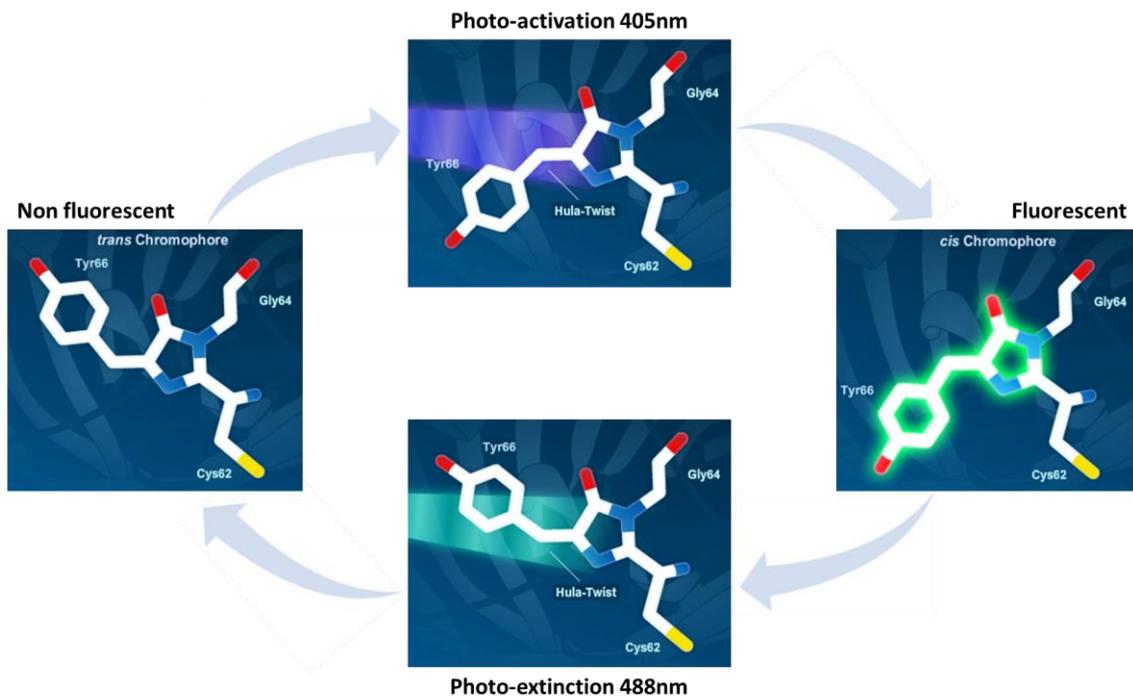


Figure 83 : Le chromophore de la DRONPA est capable de « switcher » d'un état éteint (conformation *trans*) à un état allumé (conformation *cis*) par l'utilisation de différentes longueurs d'ondes : 405 nm pour l'allumer et 488 nm pour l'éteindre.

Afin d'associer la technique d'observation en molécule unique à la méthode split, nous avons utilisé la DRONPA séparée en deux fragments généreusement envoyée par le Dr Peter Dedecker (Moeyaert et al., 2014). Comme il est visible dans la **Figure 84**, les structures primaires de la DRONPA et de l'eGFP sont peu similaires (seulement 31 % de résidus sont identiques) mais les structures secondaires et tertiaires sont semblables (structures composées de 11 brins β qui forment un tonneau β). La split-DRONPA est ainsi composée de

Manuel PIRES

deux fragments, un premier N-terminal composé des brins β 1 à 9 (Nter-DRONPA) et un second composé des brins β 10 et 11 (Cter-DRONPA).

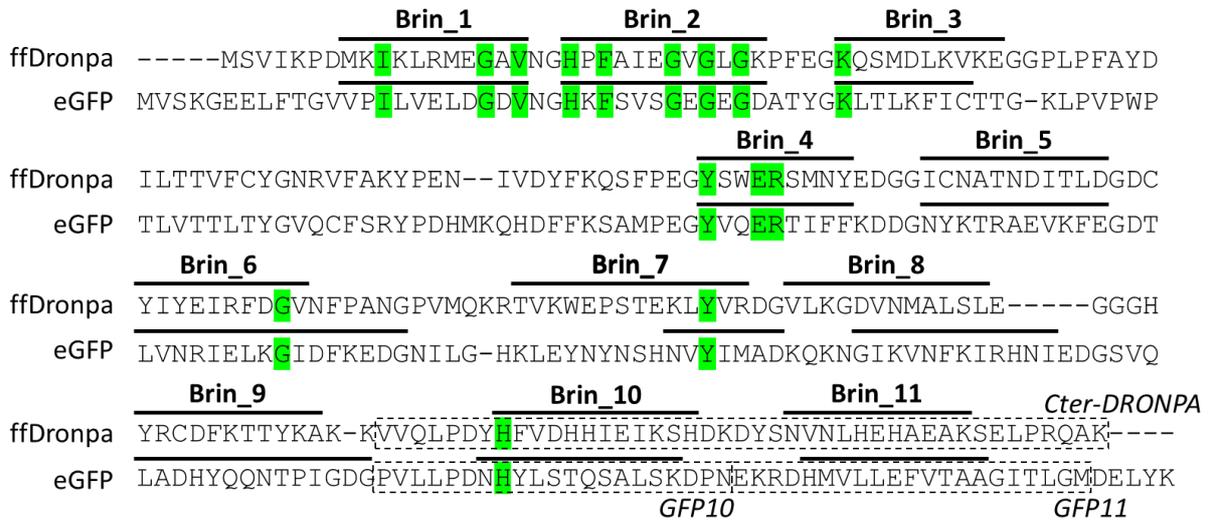


Figure 84 : Séquences des protéines ffDRONPA et eGFP sont très différentes (31 % d'homologie) mais leurs structures tridimensionnelles sont équivalentes (tonneaux β composés de 11 brins β). Les séquences protéiques ont été alignées par le logiciel ExPASy (<http://web.expasy.org/sim/>). Les brins β sont matérialisés par un trait noir au-dessus de la séquence correspondante. En vert sont indiqués les acides aminés identiques dans les brins β des deux protéines. Les cadres en pointillés indiquent les sites de clivage pour former la split GFP tri-partite et la split-DRONPA bipartite.

La première étape a consisté à vérifier que les deux fragments de DRONPA pouvaient reconstituer une protéine fluorescente lorsqu'ils étaient mis à proximité. Nous avons donc couplé chacun de ces deux fragments à un motif leucine zipper (LZ). Ce motif, connu pour favoriser la dimérisation (Landschulz et al., 1988), permet le rapprochement des deux fragments afin de forcer leurs appariements (utilisé par Cabantous et al., 2005, pour valider le split-GFP). Aucun des fragments de DRONPA, exprimés en cellules indépendamment, au même titre que ceux de la split-GFP, ne présentent de fluorescence résiduelle (**Figure 85 A et B**). En revanche, lors d'une co-transfection des plasmides codant chacun pour un fragment couplé au motif leucine zipper, une fluorescence importante apparaît dans le noyau (**Figure 85 C**). Le motif leucine-zipper étant riche en arginines et lysines (résidus basiques tous les deux) va interagir avec l'ADN, expliquant ainsi sa localisation nucléaire et nucléolaire.

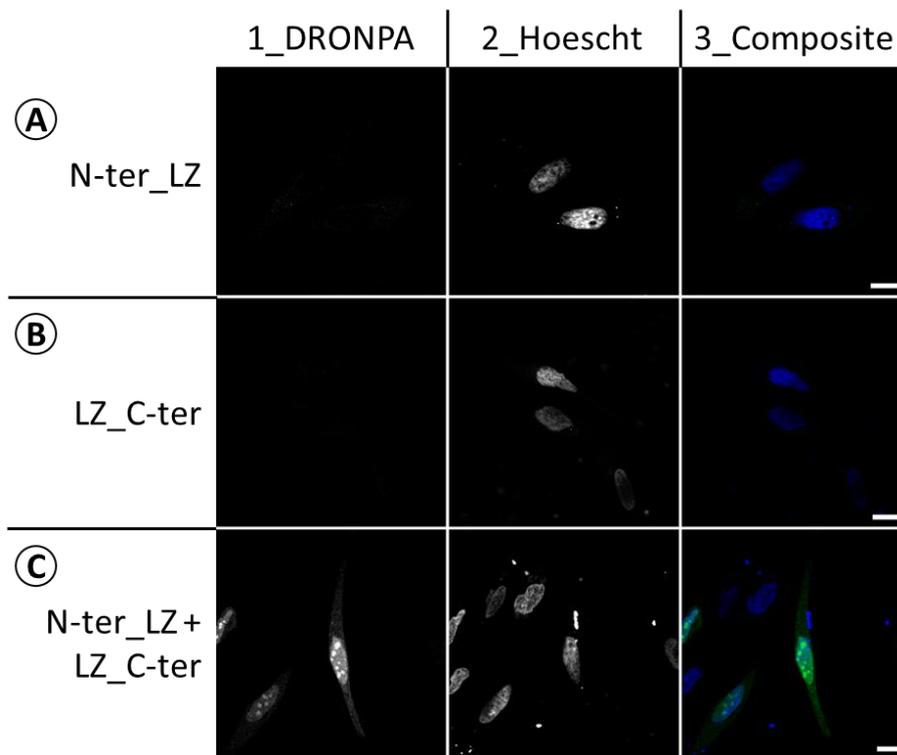


Figure 85 : Observation en microscopie à haute résolution (PALM) de la split-DRONPA. Pour chaque échantillon, l'observation de l'émission de fluorescence à 488 nm se fait par excitation permanente à 405 nm. Que ce soit pour ffDRONPA (A à E) ou pour split-DRONPA (F à J), chaque image correspond à l'observation pendant 50 ms. Les flèches blanches indiquent les photo-activations. Les barres d'échelles correspondent à 5 μm .

L'étape suivante a consisté en la vérification du photo-switching de la split-DRONPA reconstituée (possibilité de passer d'un état éteint à un état allumé et vice-versa). L'utilisation d'une DRONPA complète (ffDRONPA, [Figure 86](#)) montre l'apparition de quelques photo-activations (indiquées par des flèches blanches). Pour la split-DRONPA en revanche, malgré une localisation nucléaire similaire à l'observation en microscopie confocale, aucune photo-activation (ou très faible) n'est observée. Cette vérification en microscopie à haute résolution permet dans un premier temps de montrer que le clivage de type split de la DRONPA élimine ou limite la capacité de photo-switching de la protéine. Il est possible en effet que des ajustements des conditions de photo-switching sont probablement nécessaires avant de pouvoir utiliser cette méthode pour étudier la forme compacte de Gag en microscopie à haute résolution. De plus, dans un deuxième temps, c'est plutôt l'utilisation de la DRONPA en microscopie PALM qui est remise en question. En effet, comme on le voit dans la [Figure 86](#), la fluorescence basale avec split-DRONPA (sans excitation à 405 nm) est très intense, du fait de la présence de nombreux fluorophores au stade photo-activé. Ceci limite fortement l'observation des fluorophores après photo-activation (rapport signal/bruit de fond très

Manuel PIRES

faible). Ce type d'activation (photo-switching) semble donc moins pertinent que les activations de type photo-conversion (mEos par exemple) ou après activation, la fluorescence passe d'une émission « verte » (516 nm) à une émission « rouge » (581 nm). Sur le même principe que la split-DRONPA (clivage similaire dans la structure protéique), il pourrait être réalisé une split-Eos afin de vérifier dans un premier temps la conservation de la photo-conversion afin de valider la pertinence de l'utilisation de cet outil pour étudier la forme compacte en microscopie à haute résolution.

L'utilisation de la split-DRONPA n'est donc pas adaptée pour l'observation de la formation d'une forme compacte de Gag en microscopie PALM.

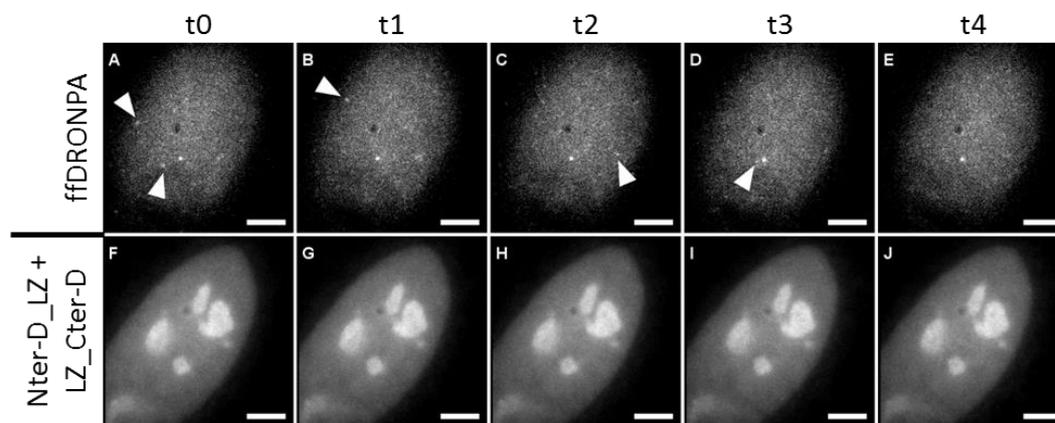


Figure 86 : Observation en microscopie à haute résolution (PALM) de la split-DRONPA. Pour chaque échantillon, l'observation de l'émission de fluorescence à 488 nm se fait par excitation permanente à 405 nm. Que ce soit pour ffDRONPA (A à E) ou pour split-DRONPA (F à J), chaque image correspond à l'observation pendant 50 ms. Les barres d'échelles correspondent à 5 μ m.

D. Conclusions et perspectives

La polyprotéine Gag est une protéine de structure du VIH-1 qui se multimérise à la membrane pour induire la formation d'une particule virale. Elle interagit via son domaine NC avec la région Ψ retrouvée à l'extrémité 5' de l'ARN viral. La présence de deux copies d'ARN viral par particule alors que plus de 2000 protéines Gag la constituent est également bien décrit dans la littérature. Cependant, le mécanisme de sélection de l'ARN viral complet non épissé parmi un large excès d'ARN cellulaires et d'ARN viraux épissés reste mal compris. Par ailleurs, l'existence d'une forme compacte transitoire de Gag a été démontrée *in vitro* (Munro et al., 2014). Cette forme compacte obtenue par le rapprochement des extrémités N- et C-terminales de Gag serait stabilisée par l'interaction simultanée des domaines MA et NC avec

Manuel PIRES

l'ARN et serait perdue, en présence de membranes, de par la forte affinité de MA pour ces dernières.

Un de mes projets a consisté à démontrer l'existence de cette forme *in cellulo*. Pour ce faire, en se basant sur la stratégie employée *in vitro*, nous avons dans un premier temps utilisé la méthode de FRET intramoléculaire. L'insertion de deux protéines fluorescentes au sein d'une même protéine Gag n'étant pas envisageable (problème d'encombrement stérique et perte d'activité), un marquage TC-ReAsH (étiquette tétra-cystéine de petite taille associée à un composé fluorescent à base d'arsenic) a été retenue. Cependant, cette technique s'est révélée sub-optimale de par le fort bruit de fond lié à ReAsH dans le cytoplasme et le caractère probablement minoritaire et transitoire de la forme compacte de Gag à la membrane (Datta et al., 2011a).

Dans une deuxième approche, nous avons utilisé une stratégie basée sur la split-eGFP. Dans cette stratégie, l'eGFP est clivée en deux parties non fluorescentes (GFP1-10 et GFP11) qui peuvent reconstituer une protéine fluorescente complète lors de leur rapprochement (Cabantous et al., 2005). Une série de constructions plasmidiques codant des protéines Gag recombinantes portant diverses combinaisons de la split-GFP a été réalisée. Pour ce faire, deux sites de clonage multiples (MCS) ont été placés entre MA et CA (MCSI) et à l'extrémité C-terminale de Gag (MCSII) respectivement. Ces sites permettent d'insérer dans les positions désirées les fragments GFP11, GFP1-10 ou l'eGFP et, si besoin, de muter certains domaines de la séquence de Gag tout en facilitant l'échange de ces mutations entre les différents plasmides nécessaires pour l'étude (Gag_MCS1_MCS2; Gag_eGFP_MCS2, Gag_GFP11_GFP1-10, gag_GFP11_MCS2 et Gag_MCS1_GFP1-10). Théoriquement, dans la forme compacte, grâce au rapprochement des fragments de GFP placés entre les domaines Matrice et Capside de Gag pour le premier et à l'extrémité C-terminale de Gag pour le second (ou vice versa), il est possible d'observer une reconstitution du fluorophore. En accord avec cette possibilité, une fluorescence cytoplasmique importante a été observée avec la protéine Gag_GFP11_GFP1-10, indiquant qu'une proportion significative de protéines Gag adopte une conformation repliée dans le cytoplasme (reconstitution INTRAmoléculaire). Ce résultat diffère de celui obtenu dans l'étude de Munro et collaborateurs où aucun FRET n'était observé dans la configuration où des fluorophores sous la forme de petits fluorophores organiques étaient insérés en C-terminal et entre MA et CA. Cette absence de FRET peut s'expliquer par le fait que contrairement au système split, les petites molécules n'interagissent pas physiquement

Manuel PIRES

ensemble et ne peuvent donc pas stabiliser la forme repliée de Gag. De plus, Munro et collaborateurs ont utilisé une protéine Gag tronquée du domaine p6 (en C-terminal), qui pourrait donner un repliement différent de la protéine Gag que nous avons utilisée. La reconstitution de fluorescence par l'utilisation de la split-eGFP apparaît donc plus adaptée que la méthode de FRET pour étudier la forme compacte de Gag *in cellulo*.

Il a ensuite fallu vérifier que la multimérisation de Gag ne contribuait pas de manière majeure à la reconstitution de la fluorescence à partir des deux fragments d'eGFP. Pour cela, les deux fragments de la split-GFP ont été insérés dans deux protéines Gag différentes (Gag_GFP11_MCS2 et Gag_MCS1_GFP1-10, reconstitution INTERmoléculaire). Dans le cas où ces deux protéines sont co-exprimées, la fluorescence est alors essentiellement observée au niveau de la membrane plasmique où la forte densité de protéines Gag favorise la reconstitution intermoléculaire et très peu dans le cytoplasme contrairement à l'observation en INTRA. Ces données confirment que la split-GFP est la technique appropriée pour observer la forme compacte de Gag *in cellulo*, qui est donc majoritairement présente dans le cytoplasme.

Afin de réaliser une mesure plus quantitative, nous avons imagé par microscopie confocale un grand nombre de cellules puis quantifié, grâce à un plugin Image J les intensités de fluorescence dans le cytoplasme et à la membrane de cellules exprimant les dérivés de Gag permettant une reconstitution INTRA ou INTER. Les analyses statistiques sur un grand nombre d'échantillons ont confirmé une différence significative entre les reconstitutions INTRA et INTER de l'eGFP. Dans le cas de la reconstitution INTRA, des fluorescences équivalentes sont observées dans le cytoplasme et à la membrane, alors que dans la reconstitution INTER, la fluorescence est beaucoup plus importante au niveau de la membrane plasmique. Cette différence confirme que la fluorescence observée à la membrane est plutôt provoquée par une reconstitution INTER liée au fort degré d'oligomérisation de Gag à la membrane (Derdowski et al., 2004; El Meshri et al., 2015; Fritz et al., 2010; Kutluay and Bieniasz, 2010). En revanche, en INTRA, la forte fluorescence dans le cytoplasme conforte l'existence d'une forme compacte de Gag dans le cytoplasme et valide les hypothèses émises dans différentes études préalables (Datta et al., 2007a, 2011a; Munro et al., 2014).

Afin d'identifier les domaines de Gag importants dans l'obtention de la forme compacte, des constructions codant pour divers mutants de Gag dont certains décrits comme limitant la localisation membranaire ou diminuant la multimérisation (diminution du signal INTER) ont

Manuel PIRES

été réalisés. Nous avons montré que le domaine de liaison (Δ linkCAC-CAN) entre les parties N- et C-terminales de la MA ainsi que le domaine Sp1 jouent un rôle limité dans le repliement de Gag. C'est pourquoi, l'investigation pourrait se prolonger sur d'autres domaines de Gag afin de déterminer le ou les régions de Gag impliquées dans le repliement sous forme compacte. Il est à noter que l'insertion d'un fragment de split eGFP entre les domaines MA et CA nous a empêché d'étudier l'importance de cette région dans le repliement. Or, la majorité des études semble en effet pencher en faveur du rôle déterminant de la région entre MA et CA dans le repliement de Gag (Datta et al., 2007a, 2011a; Munro et al., 2014). Il serait donc intéressant par la suite de placer les fragments de split-eGFP aux extrémités N- et C-terminales afin d'étudier l'impact d'une délétion ou d'une modification du domaine flexible entre MA et CA dans la formation de la forme compacte.

Comme la stabilisation de la forme compacte de Gag par les acides nucléiques nécessiterait l'interaction de ces derniers avec à la fois les domaines MA et NC (pour revue, voir Muriaux and Darlix, 2010, Chukkapalli et al., 2010; Ott et al., 2005) nous avons muté ces domaines de fixation de l'ARN. Le domaine NC a donc été supprimé (Δ NC) et/ou trois lysines de la Matrice ont été mutées en thréonines (KKK-25, 26, 27-TTT). Nos données montrent que ces mutations limitent la formation de la forme compacte, surtout lorsque ces deux modifications sont associées, confirmant le rôle des domaines MA et NC dans la stabilisation de la forme compacte de Gag (Munro et al., 2014). Afin d'améliorer ces résultats, il serait intéressant de muter un plus grand nombre de résidus impliqués dans la liaison de la MA avec les acides nucléiques. En effet, il est probable que la mutation KKK-25,26,27-TTT ne supprime pas complètement l'interaction avec les acides nucléiques.

L'utilisation d'une forme tri-partite de la split-eGFP pour montrer la forme compacte de Gag *in cellulo* s'est également montrée intéressante. En effet, l'insertion de petits fragments au sein de Gag (seulement un feuillet β entre MA et CA et un en C-terminale) limite la modification de la protéine d'intérêt. Le faible bruit de fond ainsi que la possibilité d'associer le 3^{ème} fragment à un partenaire indiquent que cet outil peut aussi être utilisé pour étudier la forme compacte de Gag. Par la suite, il serait intéressant d'utiliser la split-GFP tri-partite pour visualiser la liaison de la forme compacte et l'ARN viral grâce à la reconnaissance des boucles MS2 par la capsid du bactériophage MS2 couplé à un des fragments de la split-eGFP. Ainsi, en plaçant deux fragments de la split tri-partite au niveau de Gag (GFP10 entre MA et CA et GFP11 en C-terminale) et en couplant le reste du tonneau beta constituant l'eGFP (GF1-9) avec

Manuel PIRES

la capsid de MS2, il sera possible de valider la liaison de Gag repliée à l'ARN viral du VIH-1 auquel aura été ajouté les boucles MS2. En effet, nous pensons que la forme compacte de Gag favorise la fixation de l'ARN viral spécifiquement pour le transporter vers la membrane. C'est pourquoi, si nous sommes capables de trouver un mutant qui ne donne pas de forme compacte (autre que ΔNC_KK-TT), il serait intéressant d'utiliser ce mutant pour étudier l'impact de la forme compacte dans la sélection de l'ARN viral spécifique en comparaison de la forme sauvage.

Des études biochimiques (fractionnement cellulaire, gels de protéine natifs et dénaturants, western blot) ont également permis de confirmer la présence d'une forme repliée de Gag dans le cytoplasme. Pour compléter la caractérisation de cette forme compacte, plusieurs approches de spectrométrie de masse ont été utilisées en collaboration avec le laboratoire du Dr S. Cianferani (Institut Pluridisciplinaire Hubert CURIEN, Strasbourg) :

- Premièrement, un protocole couplant un pontage chimique des protéines et une purification par immuno-précipitation des lysats cellulaires a été développé. Les échantillons issus de cette précipitation ont été partiellement digérés par la trypsine et les peptides obtenus ont été analysés par spectrométrie de masse afin de différencier les peptides issus d'une reconstitution intramoléculaire (forme compacte de Gag) de ceux obtenus par association intermoléculaire (multimères de Gag). Les résultats préliminaires semblent valider le repliement de Gag. Les mesures devront encore être complétées et affinées afin de compléter la validation de la forme compacte de Gag *in cellulo*.
- Deuxièmement, comme à ce jour les résultats *in vitro* montrant l'existence de la forme compacte ont uniquement utilisé une forme de Gag tronquée de son domaine p6 (Gag Δ p6), une forme sauvage complète de Gag a été purifiée pour montrer sa capacité à se replier sous forme compacte. Des expériences de spectrométrie de masse à mobilité ionique ont été initiées. Pour cette technique, il est indispensable de modifier le tampon de purification de la protéine Gag car celui actuellement utilisé rend très compliqué les expériences de mobilité ionique. Nous espérons grâce à cette approche, pouvoir confirmer directement l'existence d'une forme compacte. L'utilisation de polyA25 devrait permettre d'induire le repliement de la protéine grâce à leur reconnaissance par Gag. Le passage à des études *in vitro* nous permettra aussi de rechercher plus aisément un mutant

Manuel PIRES

diminuant la forme compacte. En ajoutant à Gag purifié des ARN de séquence, de longueur ou de nature variable, nous pourrions également travailler sur la sélection de l'ARN viral.

Pour continuer ce projet, il est important de renouveler et d'améliorer l'expérience d'analyse des peptides cross-linkés par spectrométrie de masse pour confirmer les résultats obtenus mais également de continuer les investigations *in vitro* par spectrométrie de masse à mobilité ionique.

De plus, nous possédons au laboratoire de nombreuses techniques de fluorescence (anisotropie et spectroscopie de fluorescence, spectroscopie à corrélation de fluorescence par exemple) permettant l'étude de protéines *in vitro*. Plusieurs plasmides ont été préparés pour permettre la purification de protéines Gag recombinantes associées aux différents fragments de split-GFP bi-partite. Ceci permettra de visualiser *in vitro* l'apparition de la fluorescence à partir des fragments de la split-GFP placés dans Gag. Le repliement de Gag, ainsi que sa dynamique en fonction de l'environnement (ARN, membrane, pH, différents partenaires, ...) pourront être étudiés afin de déterminer les paramètres clés pour le repliement de la protéine. Grâce aux différents instruments disponibles au laboratoire, l'étude complète *in vitro* de la forme compacte de Gag par le système split-GFP pourra alors être réalisée.

Manuel PIRES

Liste de mes publications et communications

Liste de publications :

Elgaher Walid, Sharma Kamal, Hauptenthal Jörg, Saladini Francesco, **Pires Manuel**, Real Eleonore, Mély Yves, Hartmann Rolf. *Discovery and Structure-Based Optimization of 2-Ureidothiophene-3-Carboxylic Acids as Dual Bacterial RNAP and Viral RT Inhibitors*, *Journal of Medicinal Chemistry*, 06/2016.

Mori Mattia, Kovalenko Lesia, Malancona Savina, Saladini Francesco, de Forni Davide, **Pires Manuel**, Humbert Nicolas, Real Eleonore, Botzanowski Thomas, Cianferani Sarah, Giannini Alessia, Dasso Lang Chiara, Cugia Giulia, Poddesu Barbara, Lori Franco, Zazzi Maurizio, Harper Steven, Summa Vincenzo, Mely Yves, & Botta Maurizio. *Structure-based identification of HIV-1 nucleocapsid protein inhibitors active against wild-type and drug-resistant HIV-1 strains*. **Soumis à publication**.

Humbert Nicolas, Kovalenko Lesia, **Pires Manuel**, Saladini Francesco, Boudier Christian, Sharma Kamal K., Real Eleonore, Botta Maurizio, Zazzi Maurizio, Kalchenko Vitaly I., Mori Mattia, Mély Yves. *Calixarenes as potent inhibitors with a new mechanism action of the nucleic acid chaperone activity of the HIV-1 nucleocapsid protein*. **En préparation**

Pires Manuel, Botzanowski Thomas, Bourguet Maxime, Cianferani Sarah, Réal Eleonore, Mely Yves. *In cellulo visualization of the HIV-1 Gag Compact form using a split-GFP system*, **En préparation**.

Liste de communications :

Manuel Pires, Eleonore Real, Hugues de Rocquigny et Yves Mely. *VIH : vers un traitement curatif ? Doctoriales d'Alsace 2014*, Goersdrof, 09-13 juin 2014.

Manuel Pires, Lesia Kovalenko, Nicolas Humbert, Hugues de Rocquigny, Eleonore Real et Yves Mely. *Cellular and molecular characterization of new anti-HIV1 Molecules targeting the Nucleocapsid protein*. *Journée Campus Illkirch*, ESBS Illkirch, 13-14 avril 2015.

Manuel Pires, Malancona Savina, Steven Harper, Vincenzo Summa, Hugues de Rocquigny, Eleonore Real et Yves Mely. *Cellular and molecular characterization of new anti-HIV1 Molecules targeting the Nucleocapsid protein*. *Retropath2015*, Wolfsburg, Mülheim an der Ruhr, Allemagne, 24-27 Août 2015.

Manuel Pires, Malancona Savina, Steven Harper, Vincenzo Summa, Hugues de Rocquigny, Eleonore Real et Yves Mely. *Caractérisation cellulaire et moléculaire de nouveaux traitements anti-VIH1*. *Université des Jeunes Chercheurs* (UJC, Sidaction), Carry-le-Rouet, 10-16 octobre 2015.

Manuel Pires, Eleonore Real et Yves Mely. *Le VIH inscrit sur un site de rencontre pour trouver l'amour avec une nouvelle molécule*. *Finale régionale Ma thèse en 180s*, 24 mars 2016.

Manuel PIRES

Manuel Pires, Lesia Kovalenko, Nicolas Humbert, Malancona Savina, Steven Harper, Vincenzo Summa, Mattia Mori, Maurizio Botta, Eleonore Real et Yves Mely. *Screening of new anti-HIV1 Molecules targeting the Nucleocapsid protein. European Workshop in Drug Synthesis*, Sienne, 15-19 Mai 2016.

Manuel Pires, Lesia Kovalenko, Nicolas Humbert, Francesco Saladini, Maurizio Zazzi, Malancona Savina, Steven Harper, Vincenzo Summa, Eleonore Réal et Yves Mely. *Targeting the HIV-1 Nucleocapsid protein to fight antiretroviral drug resistance : in cellulo screening and characterization. The 10th International Retroviral Nucleocapsid protein and Assembly Symposium*, 18-21 septembre 2016.

Références

- Abbink, T.E.M., Beerens, N., and Berkhout, B. (2004). Forced selection of a human immunodeficiency virus type 1 variant that uses a non-self tRNA primer for reverse transcription: involvement of viral RNA sequences and the reverse transcriptase enzyme. *J. Virol.* 78, 10706–10714.
- Abd El-Wahab, E.W., Smyth, R.P., Mailler, E., Bernacchi, S., Vivet-Boudou, V., Hijnen, M., Jossinet, F., Mak, J., Paillart, J.-C., and Marquet, R. (2014). Specific recognition of the HIV-1 genomic RNA by the Gag precursor. *Nat. Commun.* 5, 4304.
- Adams, S.R., Campbell, R.E., Gross, L.A., Martin, B.R., Walkup, G.K., Yao, Y., Llopis, J., and Tsien, R.Y. (2002). New biarsenical ligands and tetracysteine motifs for protein labeling in vitro and in vivo: synthesis and biological applications. *J. Am. Chem. Soc.* 124, 6063–6076.
- Albini, A., Soldi, R., Giunciuglio, D., Giraudo, E., Benelli, R., Primo, L., Noonan, D., Salio, M., Camussi, G., Rockl, W., et al. (1996). The angiogenesis induced by HIV-1 tat protein is mediated by the Flk-1/KDR receptor on vascular endothelial cells. *Nat. Med.* 2, 1371–1375.
- Alce, T.M., and Popik, W. (2004). APOBEC3G is incorporated into virus-like particles by a direct interaction with HIV-1 Gag nucleocapsid protein. *J. Biol. Chem.* 279, 34083–34086.
- Alfadhli, A., and Barklis, E. (2014). The roles of lipids and nucleic acids in HIV-1 assembly. *Virology* 5, 253.
- Alfadhli, A., Still, A., and Barklis, E. (2009). Analysis of Human Immunodeficiency Virus Type 1 Matrix Binding to Membranes and Nucleic Acids. *J. Virol.* 83, 12196–12203.
- Alfadhli, A., Mack, A., Ritchie, C., Cylinder, I., Harper, L., Tedbury, P.R., Freed, E.O., and Barklis, E. (2016). Trimer Enhancement Mutation Effects on HIV-1 Matrix Protein Binding Activities. *J. Virol.* 90, 5657–5664.
- Alkhatib, G., Combadiere, C., Broder, C.C., Feng, Y., Kennedy, P.E., Murphy, P.M., and Berger, E.A. (1996). CC CKR5: a RANTES, MIP-1alpha, MIP-1beta receptor as a fusion cofactor for macrophage-tropic HIV-1. *Science* 272, 1955–1958.
- Amarasinghe, G.K., De Guzman, R.N., Turner, R.B., and Summers, M.F. (2000). NMR structure of stem-loop SL2 of the HIV-1 Ψ RNA packaging signal reveals a novel A-U-A base-triple platform. *J. Mol. Biol.* 299, 145–156.
- Amirache, F., Lévy, C., Costa, C., Mangeot, P.-E., Torbett, B.E., Wang, C.X., Nègre, D., Cosset, F.-L., and Verhoeven, E. (2014). Mystery solved: VSV-G-LVs do not allow efficient gene transfer into unstimulated T cells, B cells, and HSCs because they lack the LDL receptor. *Blood* 123, 1422–1424.
- Amorim, N.A., da Silva, E.M.L., de Castro, R.O., da Silva-Januário, M.E., Mendonça, L.M., Bonifacino, J.S., da Costa, L.J., and daSilva, L.L.P. (2014). Interaction of HIV-1 Nef Protein with the Host Protein Alix Promotes Lysosomal Targeting of CD4 Receptor. *J. Biol. Chem.* 289, 27744–27756.
- Andersen, E.S., Contera, S.A., Knudsen, B., Damgaard, C.K., Besenbacher, F., and Kjems, J. (2004). Role of the Trans-activation Response Element in Dimerization of HIV-1 RNA. *J. Biol. Chem.* 279, 22243–22249.
- Anderson, S.J., Lenburg, M., Landau, N.R., and Garcia, J.V. (1994). The cytoplasmic domain of CD4 is sufficient for its down-regulation from the cell surface by human immunodeficiency virus type 1 Nef. *J. Virol.* 68, 3092–3101.
- Andrake, M.D., and Skalka, A.M. (2015). RETROVIRAL INTEGRASE: THEN AND NOW.

- Annu. Rev. Virol. 2, 241–264.
- Anton, H., Taha, N., Boutant, E., Richert, L., Khatteer, H., Klaholz, B., Rondé, P., Réal, E., de Rocquigny, H., and Mély, Y. (2015). Investigating the cellular distribution and interactions of HIV-1 nucleocapsid protein by quantitative fluorescence microscopy. *PLoS One* 10, e0116921.
- Archin, N.M., Liberty, A.L., Kashuba, A.D., Choudhary, S.K., Kuruc, J.D., Crooks, A.M., Parker, D.C., Anderson, E.M., Kearney, M.F., Strain, M.C., et al. (2012). Administration of vorinostat disrupts HIV-1 latency in patients on antiretroviral therapy. *Nature* 487, 482–485.
- Arhel, N.J., Souquere-Besse, S., Munier, S., Souque, P., Guadagnini, S., Rutherford, S., Prévost, M.-C., Allen, T.D., and Charneau, P. (2007). HIV-1 DNA Flap formation promotes uncoating of the pre-integration complex at the nuclear pore. *EMBO J.* 26, 3025–3037.
- Arts, E.J., and Hazuda, D.J. (2012). HIV-1 Antiretroviral Drug Therapy. *Cold Spring Harb. Perspect. Med.* 2.
- Athavale, S.S., Ouyang, W., McPike, M.P., Hudson, B.S., and Borer, P.N. (2010). Effects of the Nature and Concentration of Salt on the Interaction of the HIV-1 Nucleocapsid Protein with SL3 RNA. *Biochemistry (Mosc.)* 49, 3525–3533.
- Atkins, K.M., Thomas, L., Youker, R.T., Harriff, M.J., Pissani, F., You, H., and Thomas, G. (2008). HIV-1 Nef binds PACS-2 to assemble a multikinase cascade that triggers major histocompatibility complex class I (MHC-I) down-regulation: analysis using short interfering RNA and knock-out mice. *J. Biol. Chem.* 283, 11772–11784.
- Avilov, S.V., Boudier, C., Gottikh, M., Darlix, J.-L., and Mély, Y. (2012). Characterization of the Inhibition Mechanism of HIV-1 Nucleocapsid Protein Chaperone Activities by Methylated Oligoribonucleotides. *Antimicrob. Agents Chemother.* 56, 1010–1018.
- Azoulay, J., Clamme, J.P., Darlix, J.L., Roques, B.P., and Mély, Y. (2003). Destabilization of the HIV-1 complementary sequence of TAR by the nucleocapsid protein through activation of conformational fluctuations. *J. Mol. Biol.* 326, 691–700.
- Bachand, F., Yao, X.J., Hrimch, M., Rougeau, N., and Cohen, E.A. (1999). Incorporation of Vpr into human immunodeficiency virus type 1 requires a direct interaction with the p6 domain of the p55 gag precursor. *J. Biol. Chem.* 274, 9083–9091.
- Bagashev, A., and Sawaya, B.E. (2013). Roles and functions of HIV-1 Tat protein in the CNS: an overview. *Virol. J.* 10, 358.
- Balasubramaniam, M., and Freed, E.O. (2011). New Insights into HIV Assembly and Trafficking. *Physiology* 26, 236–251.
- Balvay, L., Lastra, M.L., Sargueil, B., Darlix, J.-L., and Ohlmann, T. (2007). Translational control of retroviruses. *Nat. Rev. Microbiol.* 5, 128–140.
- Bampi, C., Bibillo, A., Wendeler, M., Divita, G., Gorelick, R.J., Le Grice, S.F.J., and Darlix, J.-L. (2006). Nucleotide excision repair and template-independent addition by HIV-1 reverse transcriptase in the presence of nucleocapsid protein. *J. Biol. Chem.* 281, 11736–11743.
- Barré-Sinoussi, F., Chermann, J., Rey, F., Nugeyre, M., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vezinet-Brun, F., Rouzioux, C., et al. (1983). barré-sinoussi et al, 1983, Isolation of a T-lymphotropic retrovirus from a patient at risk for AIDS.pdf. 220, 868–871.
- Barré-Sinoussi, F., Ross, A.L., and Delfraissy, J.-F. (2013). Past, present and future: 30 years of HIV research. *Nat. Rev. Microbiol.* 11, 877–883.
- Barros, M., Heinrich, F., Datta, S.A.K., Rein, A., Karageorgos, I., Nanda, H., and Lösche, M. (2016). Membrane Binding of HIV-1 Matrix Protein: Dependence on Bilayer Composition and Protein Lipidation. *J. Virol.* 90, 4544–4555.
- Bazzi, A., Zargarian, L., Chaminade, F., Boudier, C., De Rocquigny, H., René, B., Mély, Y., Fossé, P., and Mauffret, O. (2011). Structural insights into the cTAR DNA recognition by the HIV-1 nucleocapsid protein: role of sugar deoxyriboses in the binding polarity of NC. *Nucleic Acids Res.* 39, 3903–3916.
- Beerens, N., and Berkhout, B. (2002). The tRNA Primer Activation Signal in the Human

- Immunodeficiency Virus Type 1 Genome Is Important for Initiation and Processive Elongation of Reverse Transcription. *J. Virol.* 76, 2329–2339.
- Beerens, N., and Kjems, J. (2010). Circularization of the HIV-1 genome facilitates strand transfer during reverse transcription. *RNA* 16, 1226–1235.
- van Bel, N., Das, A.T., Cornelissen, M., Abbink, T.E.M., and Berkhout, B. (2014). A short sequence motif in the 5' leader of the HIV-1 genome modulates extended RNA dimer formation and virus replication. *J. Biol. Chem.* 289, 35061–35074.
- Bell, N.M., and Lever, A.M.L. (2013). HIV Gag polyprotein: processing and early viral particle assembly. *Trends Microbiol.* 21, 136–144.
- Bell, I., Schaefer, T.M., Tribble, R.P., Amedee, A., and Reinhart, T.A. (2001). Down-modulation of the costimulatory molecule, CD28, is a conserved activity of multiple SIV Nefs and is dependent on histidine 196 of Nef. *Virology* 283, 148–158.
- Berglund, J.A., Charpentier, B., and Rosbash, M. (1997). A high affinity binding site for the HIV-1 nucleocapsid protein. *Nucleic Acids Res.* 25, 1042–1049.
- Berkhout, B., Silverman, R.H., and Jeang, K.T. (1989). Tat trans-activates the human immunodeficiency virus through a nascent RNA target. *Cell* 59, 273–282.
- Bernacchi, S., and Mély, Y. (2001). Exciton interaction in molecular beacons: a sensitive sensor for short range modifications of the nucleic acid structure. *Nucleic Acids Res.* 29, e62.
- Bernacchi, S., Stoylov, S., Piémont, E., Ficheux, D., Roques, B.P., Darlix, J.L., and Mély, Y. (2002). HIV-1 nucleocapsid protein activates transient melting of least stable parts of the secondary structure of TAR and its complementary sequence. *J. Mol. Biol.* 317, 385–399.
- Berry, C., Hannenhalli, S., Leipzig, J., and Bushman, F.D. (2006). Selection of target sites for mobile DNA integration in the human genome. *PLoS Comput. Biol.* 2, e157.
- Berthet-Colominas, C., Monaco, S., Novelli, A., Sibai, G., Mallet, F., and Cusack, S. (1999). Head-to-tail dimers and interdomain flexibility revealed by the crystal structure of HIV-1 capsid protein (p24) complexed with a monoclonal antibody Fab. *EMBO J.* 18, 1124–1136.
- Bertrand, E., Chartrand, P., Schaefer, M., Shenoy, S.M., Singer, R.H., and Long, R.M. (1998). Localization of ASH1 mRNA Particles in Living Yeast. *Mol. Cell* 2, 437–445.
- Blissenbach, M., Grewe, B., Hoffmann, B., Brandt, S., and Überla, K. (2010). Nuclear RNA Export and Packaging Functions of HIV-1 Rev Revisited. *J. Virol.* 84, 6598–6604.
- Böhnlein, E., Berger, J., and Hauber, J. (1991). Functional mapping of the human immunodeficiency virus type 1 Rev RNA binding domain: new insights into the domain structure of Rev and Rex. *J. Virol.* 65, 7051–7055.
- Bolinger, C., and Boris-Lawrie, K. (2009). Mechanisms employed by retroviruses to exploit host factors for translational control of a complicated proteome. *Retrovirology* 6, 8.
- Bombarda, E., Roques, B.P., Mély, Y., and Grell, E. (2005). Mechanism of zinc coordination by point-mutated structures of the distal CCHC binding motif of the HIV-1 NCp7 protein. *Biochemistry (Mosc.)* 44, 7315–7325.
- Bosch, M.L., Earl, P.L., Fargnoli, K., Picciafuoco, S., Giombini, F., Wong-Staal, F., and Franchini, G. (1989). Identification of the fusion peptide of primate immunodeficiency viruses. *Science* 244, 694–697.
- Bour, S., and Strebel, K. (2003). The HIV-1 Vpu protein: a multifunctional enhancer of viral particle release. *Microbes Infect.* 5, 1029–1039.
- Bourbigot, S., Ramalanjaona, N., Boudier, C., Salgado, G.F.J., Roques, B.P., Mély, Y., Bouaziz, S., and Morellet, N. (2008). How the HIV-1 nucleocapsid protein binds and destabilises the (-)primer binding site during reverse transcription. *J. Mol. Biol.* 383, 1112–1128.
- Bouyac, M., Courcoul, M., Bertoia, G., Baudat, Y., Gabuzda, D., Blanc, D., Chazal, N., Boulanger, P., Sire, J., Vigne, R., et al. (1997). Human immunodeficiency virus type 1 Vif protein binds to the Pr55Gag precursor. *J. Virol.* 71, 9358–9365.

- Brasey, A., Lopez-Lastra, M., Ohlmann, T., Beerens, N., Berkhout, B., Darlix, J.-L., and Sonenberg, N. (2003). The Leader of Human Immunodeficiency Virus Type 1 Genomic RNA Harbors an Internal Ribosome Entry Segment That Is Active during the G2/M Phase of the Cell Cycle. *J. Virol.* 77, 3939–3949.
- Brass, A.L., Dykxhoorn, D.M., Benita, Y., Yan, N., Engelman, A., Xavier, R.J., Lieberman, J., and Elledge, S.J. (2008). Identification of host proteins required for HIV infection through a functional genomic screen. *Science* 319, 921–926.
- Brasseur, R., Vandenbranden, M., Cornet, B., Burny, A., and Ruyschaert, J.-M. (1990). Orientation into the lipid bilayer of an asymmetric amphipathic helical peptide located at the N-terminus of viral fusion proteins. *Biochim. Biophys. Acta BBA - Biomembr.* 1029, 267–273.
- Breuer, S., Chang, M.W., Yuan, J., and Torbett, B.E. (2012). Identification of HIV-1 inhibitors targeting the nucleocapsid protein. *J. Med. Chem.* 55, 4968–4977.
- Briggs, J.A.G., Wilk, T., Welker, R., Kräusslich, H.-G., and Fuller, S.D. (2003). Structural organization of authentic, mature HIV-1 virions and cores. *EMBO J.* 22, 1707–1715.
- Bristol-Myers Squibb Company (2015). Complete Phase 2a Study of HIV-1 Investigational Maturation Inhibitor Demonstrates Positive Results for Therapy Designed to Attack Virus Differently Than Existing Treatments | Business Wire.
- Brügger, B., Glass, B., Haberkant, P., Leibrecht, I., Wieland, F.T., and Kräusslich, H.-G. (2006). The HIV lipidome: a raft with an unusual composition. *Proc. Natl. Acad. Sci. U. S. A.* 103, 2641–2646.
- Bryant, M., and Ratner, L. (1990). Myristoylation-dependent replication and assembly of human immunodeficiency virus 1. *Proc. Natl. Acad. Sci. U. S. A.* 87, 523–527.
- Buck, C.B., Shen, X., Egan, M.A., Pierson, T.C., Walker, C.M., and Siliciano, R.F. (2001). The Human Immunodeficiency Virus Type 1 gag Gene Encodes an Internal Ribosome Entry Site. *J. Virol.* 75, 181–191.
- Buckman, J.S., Bosche, W.J., and Gorelick, R.J. (2003). Human immunodeficiency virus type 1 nucleocapsid zn(2+) fingers are required for efficient reverse transcription, initial integration processes, and protection of newly synthesized viral DNA. *J. Virol.* 77, 1469–1480.
- Bukrinsky, M.I., Haggerty, S., Dempsey, M.P., Sharova, N., Adzhubei, A., Spitz, L., Lewis, P., Goldfarb, D., Emerman, M., and Stevenson, M. (1993). A nuclear localization signal within HIV-1 matrix protein that governs infection of non-dividing cells. *Nature* 365, 666–669.
- Bültmann, A., Muranyi, W., Seed, B., and Haas, J. (2001). Identification of Two Sequences in the Cytoplasmic Tail of the Human Immunodeficiency Virus Type 1 Envelope Glycoprotein That Inhibit Cell Surface Expression. *J. Virol.* 75, 5263–5276.
- Buonaguro, L., Barillari, G., Chang, H.K., Bohan, C.A., Kao, V., Morgan, R., Gallo, R.C., and Ensoli, B. (1992). Effects of the human immunodeficiency virus type 1 Tat protein on the expression of inflammatory cytokines. *J. Virol.* 66, 7159–7167.
- Burniston, M.T., Cimarelli, A., Colgan, J., Curtis, S.P., and Luban, J. (1999). Human Immunodeficiency Virus Type 1 Gag Polyprotein Multimerization Requires the Nucleocapsid Domain and RNA and Is Promoted by the Capsid-Dimer Interface and the Basic Region of Matrix Protein. *J. Virol.* 73, 8527–8540.
- Bushman, F.D., Engelman, A., Palmer, I., Wingfield, P., and Craigie, R. (1993). Domains of the integrase protein of human immunodeficiency virus type 1 responsible for polynucleotidyl transfer and zinc binding. *Proc. Natl. Acad. Sci.* 90, 3428–3432.
- Butler, S.L., Hansen, M.S.T., and Bushman, F.D. (2001). A quantitative assay for HIV DNA integration in vivo. *Nat. Med.* 7, 631–634.
- Butler, S.L., Johnson, E.P., and Bushman, F.D. (2002). Human Immunodeficiency Virus cDNA Metabolism: Notable Stability of Two-Long Terminal Repeat Circles. *J. Virol.* 76, 3739–3747.
- Butsch, M., and Boris-Lawrie, K. (2002). Destiny of unspliced retroviral RNA: ribosome and/or virion? *J. Virol.* 76, 3089–3094.

- Buxbaum, A.R., Haimovich, G., and Singer, R.H. (2015). In the right place at the right time: visualizing and understanding mRNA localization. *Nat. Rev. Mol. Cell Biol.* *16*, 95–109.
- Byeon, I.-J.L., Meng, X., Jung, J., Zhao, G., Yang, R., Ahn, J., Shi, J., Concel, J., Aiken, C., Zhang, P., et al. (2009). Structural convergence between Cryo-EM and NMR reveals intersubunit interactions critical for HIV-1 capsid function. *Cell* *139*, 780–790.
- Byland, R., Vance, P.J., Hoxie, J.A., and Marsh, M. (2007). A Conserved Dileucine Motif Mediates Clathrin and AP-2-dependent Endocytosis of the HIV-1 Envelope Protein. *Mol. Biol. Cell* *18*, 414–425.
- Cabantous, S., Terwilliger, T.C., and Waldo, G.S. (2005). Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein. *Nat. Biotechnol.* *23*, 102–107.
- Cabantous, S., Nguyen, H.B., Pedelacq, J.-D., Koräichi, F., Chaudhary, A., Ganguly, K., Lockard, M.A., Favre, G., Terwilliger, T.C., and Waldo, G.S. (2013). A New Protein-Protein Interaction Sensor Based on Tripartite Split-GFP Association. *Sci. Rep.* *3*.
- Cai, M., Zheng, R., Caffrey, M., Craigie, R., Clore, G.M., and Gronenborn, A.M. (1997). Solution structure of the N-terminal zinc binding domain of HIV-1 integrase. *Nat. Struct. Biol.* *4*, 567–577.
- Campbell, E.M., and Hope, T.J. (2015). HIV-1 capsid: the multifaceted key player in HIV-1 infection. *Nat. Rev. Microbiol.* *13*, 471–483.
- Campbell, R.S., and Robinson, W.F. (1998). The comparative pathology of the lentiviruses. *J. Comp. Pathol.* *119*, 333–395.
- Campbell, S., Fisher, R.J., Towler, E.M., Fox, S., Issaq, H.J., Wolfe, T., Phillips, L.R., and Rein, A. (2001). Modulation of HIV-like particle assembly in vitro by inositol phosphates. *Proc. Natl. Acad. Sci. U. S. A.* *98*, 10875–10879.
- Cannon, P., and June, C. (2011). CCR5 Knockout Strategies. *Curr. Opin. HIV AIDS* *6*, 74–79.
- Caputi, M. (2011). *The Regulation of HIV-1 mRNA Biogenesis* (INTECH Open Access Publisher).
- Carlson, L.-A., Bai, Y., Keane, S.C., Doudna, J.A., and Hurley, J.H. (2016). Reconstitution of selective HIV-1 RNA packaging in vitro by membrane-bound Gag assemblies. *eLife* *5*, e14663.
- Carteau, S., Batson, S.C., Poljak, L., Mouscadet, J.F., de Rocquigny, H., Darlix, J.L., Roques, B.P., Käs, E., and Auclair, C. (1997). Human immunodeficiency virus type 1 nucleocapsid protein specifically stimulates Mg²⁺-dependent DNA integration in vitro. *J. Virol.* *71*, 6225–6229.
- Carteau, S., Gorelick, R.J., and Bushman, F.D. (1999). Coupled Integration of Human Immunodeficiency Virus Type 1 cDNA Ends by Purified Integrase In Vitro: Stimulation by the Viral Nucleocapsid Protein. *J. Virol.* *73*, 6670–6679.
- Casazza, J.P., Bowman, K.A., Adzaku, S., Smith, E.C., Enama, M.E., Bailer, R.T., Price, D.A., Gostick, E., Gordon, I.J., Ambrozak, D.R., et al. (2013). Therapeutic Vaccination Expands and Improves the Function of the HIV-Specific Memory T-Cell Repertoire. *J. Infect. Dis.* *207*, 1829–1840.
- Cashikar, A.G., Shim, S., Roth, R., Maldazys, M.R., Heuser, J.E., and Hanson, P.I. (2014). Structure of cellular ESCRT-III spirals and their relationship to HIV budding. *eLife* *3*.
- Cen, S., Huang, Y., Khorchid, A., Darlix, J.L., Wainberg, M.A., and Kleiman, L. (1999). The role of Pr55(gag) in the annealing of tRNA³Lys to human immunodeficiency virus type 1 genomic RNA. *J. Virol.* *73*, 4485–4488.
- Cen, S., Guo, F., Niu, M., Saadatmand, J., Deflassieux, J., and Kleiman, L. (2004). The interaction between HIV-1 Gag and APOBEC3G. *J. Biol. Chem.* *279*, 33177–33184.
- Chaisson RE, Leuther MD, Allain J, and et al (1988). Effect of zidovudine on serum human immunodeficiency virus core antigen levels: Results from a placebo-controlled trial. *Arch. Intern. Med.* *148*, 2151–2153.

- Chamanian, M., Purzycka, K.J., Wille, P.T., Ha, J.S., McDonald, D., Gao, Y., Le Grice, S.F.J., and Arts, E.J. (2013). A cis-Acting Element in Retroviral Genomic RNA Links Gag-Pol Ribosomal Frameshifting to Selective Viral RNA Encapsidation. *Cell Host Microbe* *13*, 181–192.
- Chamond, N., Locker, N., and Sargueil, B. (2010). The different pathways of HIV genomic RNA translation. *Biochem. Soc. Trans.* *38*, 1548–1552.
- Chan, R., Uchil, P.D., Jin, J., Shui, G., Ott, D.E., Mothes, W., and Wenk, M.R. (2008). Retroviruses human immunodeficiency virus and murine leukemia virus are enriched in phosphoinositides. *J. Virol.* *82*, 11228–11238.
- Charneau, P., and Clavel, F. (1991). A single-stranded gap in human immunodeficiency virus unintegrated linear DNA defined by a central copy of the polypurine tract. *J. Virol.* *65*, 2415–2421.
- Charneau, P., Alizon, M., and Clavel, F. (1992). A second origin of DNA plus-strand synthesis is required for optimal human immunodeficiency virus replication. *J. Virol.* *66*, 2814–2820.
- Chatel-Chaix, L., Boulay, K., Mouland, A.J., and Desgroseillers, L. (2008). The host protein Staufen1 interacts with the Pr55Gag zinc fingers and regulates HIV-1 assembly via its N-terminus. *Retrovirology* *5*, 41.
- Checkley, M.A., Luttmann, B.G., and Freed, E.O. (2011). HIV-1 Envelope Glycoprotein Biosynthesis, Trafficking, and Incorporation. *J. Mol. Biol.* *410*, 582–608.
- Chen, H., and Engelman, A. (1998). The barrier-to-autointegration protein is a host factor for HIV type 1 integration. *Proc. Natl. Acad. Sci. U. S. A.* *95*, 15270–15274.
- Chen, J., Nikolaitchik, O., Singh, J., Wright, A., Bencsics, C.E., Coffin, J.M., Ni, N., Lockett, S., Pathak, V.K., and Hu, W.-S. (2009). High efficiency of HIV-1 genomic RNA packaging and heterozygote formation revealed by single virion analysis. *Proc. Natl. Acad. Sci. U. S. A.* *106*, 13535–13540.
- Chen, J., Grunwald, D., Sardo, L., Galli, A., Plisov, S., Nikolaitchik, O.A., Chen, D., Lockett, S., Larson, D.R., Pathak, V.K., et al. (2014). Cytoplasmic HIV-1 RNA is mainly transported by diffusion in the presence or absence of Gag protein. *Proc. Natl. Acad. Sci. U. S. A.* *111*, E5205–E5213.
- Chen, J., Rahman, S.A., Nikolaitchik, O.A., Grunwald, D., Sardo, L., Burdick, R.C., Plisov, S., Liang, E., Tai, S., Pathak, V.K., et al. (2016). HIV-1 RNA genome dimerizes on the plasma membrane in the presence of Gag protein. *Proc. Natl. Acad. Sci.* *113*, E201–E208.
- Cherepanov, P., Devroe, E., Silver, P.A., and Engelman, A. (2004). Identification of an Evolutionarily Conserved Domain in Human Lens Epithelium-derived Growth Factor/Transcriptional Co-activator p75 (LEDGF/p75) That Binds HIV-1 Integrase. *J. Biol. Chem.* *279*, 48883–48892.
- Choe, H., Farzan, M., Sun, Y., Sullivan, N., Rollins, B., Ponath, P.D., Wu, L., Mackay, C.R., LaRosa, G., Newman, W., et al. (1996). The beta-chemokine receptors CCR3 and CCR5 facilitate infection by primary HIV-1 isolates. *Cell* *85*, 1135–1148.
- Chukkapalli, V., and Ono, A. (2011). Molecular determinants that regulate plasma membrane association of HIV-1 Gag. *J. Mol. Biol.* *410*, 512–524.
- Chukkapalli, V., Oh, S.J., and Ono, A. (2010). Opposing mechanisms involving RNA and lipids regulate HIV-1 Gag membrane binding through the highly basic region of the matrix domain. *Proc. Natl. Acad. Sci.* *107*, 1600–1605.
- Chung, J., Mujeeb, A., Jiang, Y., Guilbert, C., Pendke, M., Wu, Y., and James, T.L. (2008). A small molecule, Lys-Ala-7-amido-4-methylcoumarin, facilitates RNA dimer maturation of a stem-loop 1 transcript in vitro: structure-activity relationship of the activator. *Biochemistry (Mosc.)* *47*, 8148–8156.
- Chung, J., Ulyanov, N.B., Guilbert, C., Mujeeb, A., and James, T.L. (2010). Binding characteristics of small molecules that mimic nucleocapsid protein-induced maturation of stem-

- loop 1 of HIV-1 RNA. *Biochemistry (Mosc.)* 49, 6341–6351.
- Churchill, M.J., Deeks, S.G., Margolis, D.M., Siliciano, R.F., and Swanstrom, R. (2016). HIV reservoirs: what, where and how to target them. *Nat. Rev. Microbiol.* 14, 55–60.
- Cimarelli, A., and Darlix, J.L. (2002). Assembling the human immunodeficiency virus type 1. *Cell. Mol. Life Sci. CMLS* 59, 1166–1184.
- Cimarelli, A., and Darlix, J.-L. (2014). HIV-1 Reverse Transcription. In *Human Retroviruses*, E. Vicenzi, and G. Poli, eds. (Totowa, NJ: Humana Press), pp. 55–70.
- Cimarelli, A., and Luban, J. (1999). Translation elongation factor 1-alpha interacts specifically with the human immunodeficiency virus type 1 Gag polyprotein. *J. Virol.* 73, 5388–5401.
- Cimarelli, A., and Luban, J. (2000). Human Immunodeficiency Virus Type 1 Virion Density Is Not Determined by Nucleocapsid Basic Residues. *J. Virol.* 74, 6734–6740.
- Cimarelli, A., Sandin, S., Höglund, S., and Luban, J. (2000). Basic residues in human immunodeficiency virus type 1 nucleocapsid promote virion assembly via interaction with RNA. *J. Virol.* 74, 3046–3057.
- Clever, J.L., Miranda, Jr., D., and Parslow, T.G. (2002). RNA Structure and Packaging Signals in the 5' Leader Region of the Human Immunodeficiency Virus Type 1 Genome. *J. Virol.* 76, 12381–12387.
- Coffin, J.M., Hughes, S.H., and Varmus, H.E. (1997). *Synthesis of Gag and Gag-Pro-Pol Proteins* (Cold Spring Harbor Laboratory Press).
- Coiras, M., López-Huertas, M.R., Pérez-Olmeda, M., and Alcamí, J. (2009). Understanding HIV-1 latency provides clues for the eradication of long-term reservoirs. *Nat. Rev. Microbiol.* 7, 798–812.
- Collier, A.C., Coombs, R.W., Schoenfeld, D.A., Bassett, R.L., Timpone, J., Baruch, A., Jones, M., Facey, K., Whitacre, C., McAuliffe, V.J., et al. (1996). Treatment of Human Immunodeficiency Virus Infection with Saquinavir, Zidovudine, and Zalcitabine. *N. Engl. J. Med.* 334, 1011–1018.
- Constantopoulos, T.L., Jackson, G.S., and Enke, C.G. (1999). Effects of salt concentration on analyte response using electrospray ionization mass spectrometry. *J. Am. Soc. Mass Spectrom.* 10, 625–634.
- Conti, E., Franks, N.P., and Brick, P. (1996). Crystal structure of firefly luciferase throws light on a superfamily of adenylate-forming enzymes. *Structure* 4, 287–298.
- Cooper, J.B. (2002). Aspartic proteinases in disease: a structural perspective. *Curr. Drug Targets* 3, 155–173.
- Coren, L.V., Thomas, J.A., Chertova, E., Sowder, R.C., Gagliardi, T.D., Gorelick, R.J., and Ott, D.E. (2007). Mutational Analysis of the C-Terminal Gag Cleavage Sites in Human Immunodeficiency Virus Type 1. *J. Virol.* 81, 10047–10054.
- Coric, P., Turcaud, S., Souquet, F., Briant, L., Gay, B., Royer, J., Chazal, N., and Bouaziz, S. (2013). Synthesis and biological evaluation of a new derivative of bevirimat that targets the Gag CA-SP1 cleavage site. *Eur. J. Med. Chem.* 62, 453–465.
- Coull, J.J., Romerio, F., Sun, J.M., Volker, J.L., Galvin, K.M., Davie, J.R., Shi, Y., Hansen, U., and Margolis, D.M. (2000). The human factors YY1 and LSF repress the human immunodeficiency virus type 1 long terminal repeat via recruitment of histone deacetylase 1. *J. Virol.* 74, 6790–6799.
- Craigie, R., and Bushman, F.D. (2012). HIV DNA Integration. *Cold Spring Harb. Perspect. Med.* 2.
- Cruceanu, M., Urbaneja, M.A., Hixson, C.V., Johnson, D.G., Datta, S.A., Fivash, M.J., Stephen, A.G., Fisher, R.J., Gorelick, R.J., Casas-Finet, J.R., et al. (2006). Nucleic acid binding and chaperone properties of HIV-1 Gag and nucleocapsid proteins. *Nucleic Acids Res.* 34, 593–605.
- Crum-Cianflone, N.F., Burgi, A.A., and Hale, B.R. (2007). Increasing rates of community-

- acquired methicillin-resistant *Staphylococcus aureus* infections among HIV-infected persons. *Int. J. STD AIDS* 18, 521–526.
- Cullen, B.R. (2003). Nuclear RNA export. *J. Cell Sci.* 116, 587–597.
- Daar, E.S., Li, X.L., Moudgil, T., and Ho, D.D. (1990). High concentrations of recombinant soluble CD4 are required to neutralize primary human immunodeficiency virus type 1 isolates. *Proc. Natl. Acad. Sci. U. S. A.* 87, 6574–6578.
- Dalgleish, A.G., Beverley, P.C., Clapham, P.R., Crawford, D.H., Greaves, M.F., and Weiss, R.A. (1984). The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. *Nature* 312, 763–767.
- Daly, T.J., Cook, K.S., Gray, G.S., Maione, T.E., and Rusche, J.R. (1989). Specific binding of HIV-1 recombinant Rev protein to the Rev-responsive element in vitro. *Nature* 342, 816–819.
- D'arc, M., Ayouba, A., Esteban, A., Learn, G.H., Boué, V., Liegeois, F., Etienne, L., Tagg, N., Leendertz, F.H., Boesch, C., et al. (2015). Origin of the HIV-1 group O epidemic in western lowland gorillas. *Proc. Natl. Acad. Sci.* 112, E1343–E1352.
- Darlix, J.-L., Godet, J., Ivanyi-Nagy, R., Fossé, P., Mauffret, O., and Mély, Y. (2011). Flexible Nature and Specific Functions of the HIV-1 Nucleocapsid Protein. *J. Mol. Biol.* 410, 565–581.
- Datta, S.A.K., Curtis, J.E., Ratcliff, W., Clark, P.K., Crist, R.M., Lebowitz, J., Krueger, S., and Rein, A. (2007a). Conformation of the HIV-1 Gag Protein in Solution. *J. Mol. Biol.* 365, 812–824.
- Datta, S.A.K., Zhao, Z., Clark, P.K., Tarasov, S., Alexandratos, J.N., Campbell, S.J., Kvaratskhelia, M., Lebowitz, J., and Rein, A. (2007b). Interactions between HIV-1 Gag molecules in solution: an inositol phosphate-mediated switch. *J. Mol. Biol.* 365, 799–811.
- Datta, S.A.K., Heinrich, F., Raghunandan, S., Krueger, S., Curtis, J.E., Rein, A., and Nanda, H. (2011a). HIV-1 Gag extension: conformational changes require simultaneous interaction with membrane and nucleic acid. *J. Mol. Biol.* 406, 205–214.
- Datta, S.A.K., Temeselew, L.G., Crist, R.M., Soheilian, F., Kamata, A., Mirro, J., Harvin, D., Nagashima, K., Cachau, R.E., and Rein, A. (2011b). On the role of the SP1 domain in HIV-1 particle assembly: a molecular switch? *J. Virol.* 85, 4111–4121.
- David E; Garfin (2003). Chapter 7: Gel electrophoresis of proteins. In *Essential Cell Biology Vol 1: Cell Structure, A Practical Approach*, (John Davey, Mike Lord), pp. 197–268.
- Day, J.R., Munk, C., and Guatelli, J.C. (2004). The Membrane-Proximal Tyrosine-Based Sorting Signal of Human Immunodeficiency Virus Type 1 gp41 Is Required for Optimal Viral Infectivity. *J. Virol.* 78, 1069–1079.
- De La Fuente, C., Santiago, F., Deng, L., Eadie, C., Zilberman, I., Kehn, K., Maddukuri, A., Baylor, S., Wu, K., Lee, C.G., et al. (2002). Gene expression profile of HIV-1 Tat expressing cells: a close interplay between proliferative and differentiation signals. *BMC Biochem.* 3, 1.
- De Leys, R., Vanderborcht, B., Haesevelde, M.V., Heyndrickx, L., Van Geel, A., Wauters, C., Bernaerts, R., Saman, E., Nijs, P., and Willems, B. (1990). Isolation and partial characterization of an unusual human immunodeficiency retrovirus from two persons of west-central African origin. *J. Virol.* 64, 1207–1216.
- Deeks, S.G., Lewin, S.R., Ross, A.L., Ananworanich, J., Benkirane, M., Cannon, P., Chomont, N., Douek, D., Lifson, J.D., Lo, Y.-R., et al. (2016). International AIDS Society global scientific strategy: towards an HIV cure 2016. *Nat. Med. advance online publication*.
- Delelis, O., Carayon, K., Saïb, A., Deprez, E., and Mouscadet, J.-F. (2008). Integrase and integration: biochemical activities of HIV-1 integrase. *Retrovirology* 5, 114.
- Déméné, H., Dong, C.Z., Ottmann, M., Rouyez, M.C., Jullian, N., Morellet, N., Mely, Y., Darlix, J.L., Fournié-Zaluski, M.C., and Saragosti, S. (1994). 1H NMR structure and biological studies of the His23-->Cys mutant nucleocapsid protein of HIV-1 indicate that the conformation of the first zinc finger is critical for virus infectivity. *Biochemistry (Mosc.)* 33, 11707–11716.

- Demirov, D.G., Ono, A., Orenstein, J.M., and Freed, E.O. (2002). Overexpression of the N-terminal domain of TSG101 inhibits HIV-1 budding by blocking late domain function. *Proc. Natl. Acad. Sci. U. S. A.* *99*, 955–960.
- Denburg, J., and McElroy, W.D. (1970). Catalytic subunit of firefly luciferase. *Biochemistry (Mosc.)* *9*, 4619–4624.
- Deng, H., Liu, R., Ellmeier, W., Choe, S., Unutmaz, D., Burkhart, M., Di Marzio, P., Marmon, S., Sutton, R.E., Hill, C.M., et al. (1996). Identification of a major co-receptor for primary isolates of HIV-1. *Nature* *381*, 661–666.
- Deprez, E., Tauc, P., Leh, H., Mouscadet, J.F., Auclair, C., and Brochon, J.C. (2000). Oligomeric states of the HIV-1 integrase as measured by time-resolved fluorescence anisotropy. *Biochemistry (Mosc.)* *39*, 9275–9284.
- Derdowski, A., Ding, L., and Spearman, P. (2004). A Novel Fluorescence Resonance Energy Transfer Assay Demonstrates that the Human Immunodeficiency Virus Type 1 Pr55Gag I Domain Mediates Gag-Gag Interactions. *J. Virol.* *78*, 1230–1242.
- Deschambeault, J., Lalonde, J.-P., Cervantes-Acosta, G., Lodge, R., Cohen, É.A., and Lemay, G. (1999). Polarized human immunodeficiency virus budding in lymphocytes involves a tyrosine-based signal and favors cell-to-cell viral transmission. *J. Virol.* *73*, 5010–5017.
- Dey, R., and Pillai, B. (2015). Cell-based gene therapy against HIV. *Gene Ther.* *22*, 851–855.
- Dietz, J., Koch, J., Kaur, A., Raja, C., Stein, S., Grez, M., Pustowka, A., Mensch, S., Ferner, J., Möller, L., et al. (2008). Inhibition of HIV-1 by a Peptide Ligand of the Genomic RNA Packaging Signal Ψ . *ChemMedChem* *3*, 749–755.
- DiMattia, M.A., Watts, N.R., Stahl, S.J., Rader, C., Wingfield, P.T., Stuart, D.I., Steven, A.C., and Grimes, J.M. (2010). Implications of the HIV-1 Rev dimer structure at 3.2 Å resolution for multimeric binding to the Rev response element. *Proc. Natl. Acad. Sci. U. S. A.* *107*, 5810–5814.
- Dismuke, D.J., and Aiken, C. (2006). Evidence for a functional link between uncoating of the human immunodeficiency virus type 1 core and nuclear import of the viral preintegration complex. *J. Virol.* *80*, 3712–3720.
- Doménech, R., Bocanegra, R., Velázquez-Campoy, A., and Neira, J.L. (2011). The isolated major homology region of the HIV capsid protein is mainly unfolded in solution and binds to the intact protein. *Biochim. Biophys. Acta* *1814*, 1269–1278.
- Dorfman, T., Luban, J., Goff, S.P., Haseltine, W.A., and Göttlinger, H.G. (1993). Mapping of functionally important residues of a cysteine-histidine box in the human immunodeficiency virus type 1 nucleocapsid protein. *J. Virol.* *67*, 6159–6169.
- Dorr, P., Westby, M., Dobbs, S., Griffin, P., Irvine, B., Macartney, M., Mori, J., Rickett, G., Smith-Burchnell, C., Napier, C., et al. (2005). Maraviroc (UK-427,857), a Potent, Orally Bioavailable, and Selective Small-Molecule Inhibitor of Chemokine Receptor CCR5 with Broad-Spectrum Anti-Human Immunodeficiency Virus Type 1 Activity. *Antimicrob. Agents Chemother.* *49*, 4721–4732.
- Douaisi, M., Dussart, S., Courcoul, M., Bessou, G., Vigne, R., and Decroly, E. (2004). HIV-1 and MLV Gag proteins are sufficient to recruit APOBEC3G into virus-like particles. *Biochem. Biophys. Res. Commun.* *321*, 566–573.
- Dragic, T., Trkola, A., Thompson, D.A.D., Cormier, E.G., Kajumo, F.A., Maxwell, E., Lin, S.W., Ying, W., Smith, S.O., Sakmar, T.P., et al. (2000). A binding pocket for a small molecule inhibitor of HIV-1 entry within the transmembrane helices of CCR5. *Proc. Natl. Acad. Sci.* *97*, 5639–5644.
- Druillennec, S., and Roques, B.P. (2000). HIV-1 NCp7 as a target for the design of novel antiviral agents. *Drug News Perspect.* *13*, 337–349.
- Druillennec, S., Meudal, H., Roques, B.P., and Fournié-Zaluski, M.C. (1999). Nucleomimetic strategy for the inhibition of HIV-1 nucleocapsid protein NCp7 activities. *Bioorg. Med. Chem.*

Lett. 9, 627–632.

Dubay, J.W., Roberts, S.J., Brody, B., and Hunter, E. (1992). Mutations in the leucine zipper of the human immunodeficiency virus type 1 transmembrane glycoprotein affect fusion and infectivity. *J. Virol.* 66, 4748–4756.

Dussupt, V., Javid, M.P., Abou-Jaoudé, G., Jadwin, J.A., de La Cruz, J., Nagashima, K., and Bouamr, F. (2009). The nucleocapsid region of HIV-1 Gag cooperates with the PTAP and LYPXnL late domains to recruit the cellular machinery necessary for viral budding. *PLoS Pathog.* 5, e1000339.

Dussupt, V., Sette, P., Bello, N.F., Javid, M.P., Nagashima, K., and Bouamr, F. (2011). Basic residues in the nucleocapsid domain of Gag are critical for late events of HIV-1 budding. *J. Virol.* 85, 2304–2315.

Dyda, F., Hickman, A.B., Jenkins, T.M., Engelman, A., Craigie, R., and Davies, D.R. (1994). Crystal structure of the catalytic domain of HIV-1 integrase: similarity to other polynucleotidyl transferases. *Science* 266, 1981–1986.

Eckert, D.M., and Kim, P.S. (2001). Mechanisms of Viral Membrane Fusion and Its Inhibition. *Annu. Rev. Biochem.* 70, 777–810.

Eijkelenboom, A.P.A.M., Puras Lutzke, R.A., Boelens, R., Plasterk, R.H.A., Kaptein, R., and Hård, K. (1995). The DNA-binding domain of HIV-1 integrase has an SH3-like fold. *Nat. Struct. Mol. Biol.* 2, 807–810.

El Mekdad, H., Boutant, E., Karnib, H., Biedma, M.E., Sharma, K.K., Malyska, I., Laumond, G., Roy, M., Réal, E., Paillart, J.-C., et al. (2016). Characterization of the interaction between the HIV-1 Gag structural polyprotein and the cellular ribosomal protein L7 and its implication in viral nucleic acid remodeling. *Retrovirology*.

El Meshri, S.E., Dujardin, D., Godet, J., Richert, L., Boudier, C., Darlix, J.L., Didier, P., Mély, Y., and de Rocquigny, H. (2015). Role of the Nucleocapsid Domain in HIV-1 Gag Oligomerization and Trafficking to the Plasma Membrane: A Fluorescence Lifetime Imaging Microscopy Investigation. *J. Mol. Biol.* 427, 1480–1494.

Elgaher, W.A.M., Sharma, K.K., Haupenthal, J., Saladini, F., Pires, M., Real, E., Mély, Y., and Hartmann, R.W. (2016). Discovery and Structure-Based Optimization of 2-Ureidothiophene-3-carboxylic Acids as Dual Bacterial RNA Polymerase and Viral Reverse Transcriptase Inhibitors. *J. Med. Chem.*

Elliott, J.H., McMahon, J.H., Chang, C.C., Lee, S.A., Hartogensis, W., Bumpus, N., Savic, R., Roney, J., Hoh, R., Solomon, A., et al. (2015). Short-term administration of disulfiram for reversal of latent HIV infection: a phase 2 dose-escalation study. *Lancet HIV* 2, e520–e529.

Engeland, C.E., Oberwinkler, H., Schümann, M., Krause, E., Müller, G.A., and Kräusslich, H.-G. (2011). The Cellular Protein Lyric Interacts with HIV-1 Gag. *J. Virol.* 85, 13322–13332.

Engelman, A., Bushman, F.D., and Craigie, R. (1993). Identification of discrete functional domains of HIV-1 integrase and their organization within an active multimeric complex. *EMBO J.* 12, 3269–3275.

Engelman, A., Hickman, A.B., and Craigie, R. (1994). The core and carboxyl-terminal domains of the integrase protein of human immunodeficiency virus type 1 each contribute to nonspecific DNA binding. *J. Virol.* 68, 5911–5917.

Ensoli, B., Barillari, G., Salahuddin, S.Z., Gallo, R.C., and Wong-Staal, F. (1990). Tat protein of HIV-1 stimulates growth of cells derived from Kaposi's sarcoma lesions of AIDS patients. *Nature* 345, 84–86.

Esposito, D., and Craigie, R. (1998). Sequence specificity of viral end DNA binding by HIV-1 integrase reveals critical regions for protein-DNA interaction. *EMBO J.* 17, 5832–5843.

Esté, J.A., and Telenti, A. (2007). HIV entry inhibitors. *The Lancet* 370, 81–88.

Faria, N.R., Rambaut, A., Suchard, M.A., Baele, G., Bedford, T., Ward, M.J., Tatem, A.J., Sousa, J.D., Arinaminpathy, N., Pepin, J., et al. (2014). The early spread and epidemic ignition

- of HIV-1 in human populations. *Science* 346, 56–61.
- Farnet, C.M., and Bushman, F.D. (1997). HIV-1 cDNA Integration: Requirement of HMG I(Y) Protein for Function of Preintegration Complexes In Vitro. *Cell* 88, 483–492.
- Feng, Y., Broder, C.C., Kennedy, P.E., and Berger, E.A. (1996a). HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. *Science* 272, 872–877.
- Feng, Y.X., Copeland, T.D., Henderson, L.E., Gorelick, R.J., Bosche, W.J., Levin, J.G., and Rein, A. (1996b). HIV-1 nucleocapsid protein induces “maturation” of dimeric retroviral RNA in vitro. *Proc. Natl. Acad. Sci. U. S. A.* 93, 7577–7581.
- Finkelshtein, D., Werman, A., Novick, D., Barak, S., and Rubinstein, M. (2013). LDL receptor and its family members serve as the cellular receptors for vesicular stomatitis virus. *Proc. Natl. Acad. Sci. U. S. A.* 110, 7306–7311.
- Fischer, U., Huber, J., Boelens, W.C., Mattajt, L.W., and Lührmann, R. (1995). The HIV-1 Rev Activation Domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs. *Cell* 82, 475–483.
- Fledderman, E.L., Fujii, K., Ghanam, R.H., Waki, K., Prevelige, P.E., Freed, E.O., and Saad, J.S. (2010). Myristate exposure in the human immunodeficiency virus type 1 matrix protein is modulated by pH. *Biochemistry (Mosc.)* 49, 9551–9562.
- Fletcher, C.V. (2003). Enfuvirtide, a new drug for HIV infection. *The Lancet* 361, 1577–1578.
- Fornerod, M., Ohno, M., Yoshida, M., and Mattaj, I.W. (1997). CRM1 is an export receptor for leucine-rich nuclear export signals. *Cell* 90, 1051–1060.
- Fortin, J.F., Cantin, R., and Tremblay, M.J. (1998). T cells expressing activated LFA-1 are more susceptible to infection with human immunodeficiency virus type 1 particles bearing host-encoded ICAM-1. *J. Virol.* 72, 2105–2112.
- Fouchier, R.A., Meyer, B.E., Simon, J.H., Fischer, U., and Malim, M.H. (1997). HIV-1 infection of non-dividing cells: evidence that the amino-terminal basic region of the viral matrix protein is important for Gag processing but not for post-entry nuclear import. *EMBO J.* 16, 4531–4539.
- Fouchier, R.A., Meyer, B.E., Simon, J.H., Fischer, U., Albright, A.V., González-Scarano, F., and Malim, M.H. (1998). Interaction of the human immunodeficiency virus type 1 Vpr protein with the nuclear pore complex. *J. Virol.* 72, 6004–6013.
- Frankel, A.D., and Pabo, C.O. (1988). Cellular uptake of the tat protein from human immunodeficiency virus. *Cell* 55, 1189–1193.
- Frankel, A.D., and Young, J.A.T. (1998). HIV-1: Fifteen Proteins and an RNA. *Annu. Rev. Biochem.* 67, 1–25.
- Freed, E.O. (2015). HIV-1 assembly, release and maturation. *Nat. Rev. Microbiol.* 13, 484–496.
- Freed, E.O., Myers, D.J., and Risser, R. (1990). Characterization of the fusion domain of the human immunodeficiency virus type 1 envelope glycoprotein gp41. *Proc. Natl. Acad. Sci.* 87, 4650–4654.
- Freed, E.O., Orenstein, J.M., Buckler-White, A.J., and Martin, M.A. (1994). Single amino acid changes in the human immunodeficiency virus type 1 matrix protein block virus particle production. *J. Virol.* 68, 5311–5320.
- Freed, E.O., Englund, G., and Martin, M.A. (1995). Role of the basic domain of human immunodeficiency virus type 1 matrix in macrophage infection. *J. Virol.* 69, 3949–3954.
- Friew, Y.N., Boyko, V., Hu, W.-S., and Pathak, V.K. (2009). Intracellular interactions between APOBEC3G, RNA, and HIV-1 Gag: APOBEC3G multimerization is dependent on its association with RNA. *Retrovirology* 6, 56.
- Fritz, J.V., Didier, P., Clamme, J.-P., Schaub, E., Muriaux, D., Cabanne, C., Morellet, N., Bouaziz, S., Darlix, J.-L., Mély, Y., et al. (2008). Direct Vpr-Vpr interaction in cells monitored

- by two photon fluorescence correlation spectroscopy and fluorescence lifetime imaging. *Retrovirology* 5, 87.
- Fritz, J.V., Dujardin, D., Godet, J., Didier, P., De Mey, J., Darlix, J.-L., Mély, Y., and de Rocquigny, H. (2010). HIV-1 Vpr oligomerization but not that of Gag directs the interaction between Vpr and Gag. *J. Virol.* 84, 1585–1596.
- Fujii, K., Munshi, U.M., Ablan, S.D., Demirov, D.G., Soheilian, F., Nagashima, K., Stephen, A.G., Fisher, R.J., and Freed, E.O. (2009). Functional role of Alix in HIV-1 replication. *Virology* 391, 284–292.
- Gallay, P., Hope, T., Chin, D., and Trono, D. (1997). HIV-1 infection of nondividing cells through the recognition of integrase by the importin/karyopherin pathway. *Proc. Natl. Acad. Sci. U. S. A.* 94, 9825–9830.
- Gallo, R.C., Sarin, P.S., Gelmann, E.P., Robert-Guroff, M., Richardson, E., Kalyanaraman, V.S., Mann, D., Sidhu, G.D., Stahl, R.E., Zolla-Pazner, S., et al. (1983). Isolation of human T-cell leukemia virus in acquired immune deficiency syndrome (AIDS). *Science* 220, 865–867.
- Gamble, T.R., Yoo, S., Vajdos, F.F., von Schwedler, U.K., Worthylake, D.K., Wang, H., McCutcheon, J.P., Sundquist, W.I., and Hill, C.P. (1997). Structure of the carboxyl-terminal dimerization domain of the HIV-1 capsid protein. *Science* 278, 849–853.
- Ganser-Pornillos, B.K., von Schwedler, U.K., Stray, K.M., Aiken, C., and Sundquist, W.I. (2004). Assembly properties of the human immunodeficiency virus type 1 CA protein. *J. Virol.* 78, 2545–2552.
- Ganser-Pornillos, B.K., Yeager, M., and Sundquist, W.I. (2008). The structural biology of HIV assembly. *Curr. Opin. Struct. Biol.* 18, 203–217.
- Gao, K., Gorelick, R.J., Johnson, D.G., and Bushman, F. (2003). Cofactors for Human Immunodeficiency Virus Type 1 cDNA Integration In Vitro. *J. Virol.* 77, 1598–1603.
- Garrus, J.E., von Schwedler, U.K., Pornillos, O.W., Morham, S.G., Zavitz, K.H., Wang, H.E., Wettstein, D.A., Stray, K.M., Côté, M., Rich, R.L., et al. (2001). Tsg101 and the vacuolar protein sorting pathway are essential for HIV-1 budding. *Cell* 107, 55–65.
- Geijtenbeek, T.B., Kwon, D.S., Torensma, R., van Vliet, S.J., van Duijnhoven, G.C., Middel, J., Cornelissen, I.L., Nottet, H.S., KewalRamani, V.N., Littman, D.R., et al. (2000). DC-SIGN, a dendritic cell-specific HIV-1-binding protein that enhances trans-infection of T cells. *Cell* 100, 587–597.
- Gelderblom, H.C., Vatakis, D.N., Burke, S.A., Lawrie, S.D., Bristol, G.C., and Levy, D.N. (2008). Viral complementation allows HIV-1 replication without integration. *Retrovirology* 5, 60.
- Ghanam, R.H., Fernandez, T.F., Fledderman, E.L., and Saad, J.S. (2010). Binding of Calmodulin to the HIV-1 Matrix Protein Triggers Myristate Exposure. *J. Biol. Chem.* 285, 41911–41920.
- Gitti, R.K., Lee, B.M., Walker, J., Summers, M.F., Yoo, S., and Sundquist, W.I. (1996). Structure of the amino-terminal core domain of the HIV-1 capsid protein. *Science* 273, 231–235.
- Godet, J., de Rocquigny, H., Raja, C., Glasser, N., Ficheux, D., Darlix, J.-L., and Mély, Y. (2006). During the Early Phase of HIV-1 DNA Synthesis, Nucleocapsid Protein Directs Hybridization of the TAR Complementary Sequences via the Ends of their Double-stranded Stem. *J. Mol. Biol.* 356, 1180–1192.
- Godet, J., Ramalanjaona, N., Sharma, K.K., Richert, L., de Rocquigny, H., Darlix, J.-L., Duportail, G., and Mély, Y. (2011). Specific implications of the HIV-1 nucleocapsid zinc fingers in the annealing of the primer binding site complementary sequences during the obligatory plus strand transfer. *Nucleic Acids Res.* 39, 6633–6645.
- Godet, J., Kenfack, C., Przybilla, F., Richert, L., Duportail, G., and Mély, Y. (2013). Site-selective probing of cTAR destabilization highlights the necessary plasticity of the HIV-1

- nucleocapsid protein to chaperone the first strand transfer. *Nucleic Acids Res.* *41*, 5036–5048.
- Goh, W.C., Rogel, M.E., Kinsey, C.M., Michael, S.F., Fultz, P.N., Nowak, M.A., Hahn, B.H., and Emerman, M. (1998). HIV-1 Vpr increases viral expression by manipulation of the cell cycle: A mechanism for selection of Vpr in vivo. *Nat. Med.* *4*, 65–71.
- González, M.E. (2015). Vpu Protein: The Viroporin Encoded by HIV-1. *Viruses* *7*, 4352–4368.
- Gorelick, R.J., Nigida, S.M., Bess, J.W., Arthur, L.O., Henderson, L.E., and Rein, A. (1990). Noninfectious human immunodeficiency virus type 1 mutants deficient in genomic RNA. *J. Virol.* *64*, 3207–3211.
- Gorelick, R.J., Chabot, D.J., Rein, A., Henderson, L.E., and Arthur, L.O. (1993). The two zinc fingers in the human immunodeficiency virus type 1 nucleocapsid protein are not functionally equivalent. *J. Virol.* *67*, 4027–4036.
- Göttlinger, H.G., Sodroski, J.G., and Haseltine, W.A. (1989). Role of capsid precursor processing and myristoylation in morphogenesis and infectivity of human immunodeficiency virus type 1. *Proc. Natl. Acad. Sci. U. S. A.* *86*, 5781–5785.
- Göttlinger, H.G., Dorfman, T., Sodroski, J.G., and Haseltine, W.A. (1991). Effect of mutations affecting the p6 gag protein on human immunodeficiency virus particle release. *Proc. Natl. Acad. Sci. U. S. A.* *88*, 3195–3199.
- Goudreau, N., Hucke, O., Faucher, A.-M., Grand-Maître, C., Lepage, O., Bonneau, P.R., Mason, S.W., and Titolo, S. (2013). Discovery and structural characterization of a new inhibitor series of HIV-1 nucleocapsid function: NMR solution structure determination of a ternary complex involving a 2:1 inhibitor/NC stoichiometry. *J. Mol. Biol.* *425*, 1982–1998.
- Greenberg, M.E., Bronson, S., Lock, M., Neumann, M., Pavlakis, G.N., and Skowronski, J. (1997). Co-localization of HIV-1 Nef with the AP-2 adaptor protein complex correlates with Nef-induced CD4 down-regulation. *EMBO J.* *16*, 6964–6976.
- Grewe, B., Ehrhardt, K., Hoffmann, B., Blissenbach, M., Brandt, S., and Überla, K. (2012). The HIV-1 Rev Protein Enhances Encapsidation of Unspliced and Spliced, RRE-Containing Lentiviral Vector RNA. *PLOS ONE* *7*, e48688.
- Griffin, B.A., Adams, S.R., and Tsien, R.Y. (1998). Specific covalent labeling of recombinant protein molecules inside live cells. *Science* *281*, 269–272.
- Grigorov, B., Décimo, D., Smagulova, F., Péchoux, C., Mougel, M., Muriaux, D., and Darlix, J.-L. (2007). Intracellular HIV-1 Gag localization is impaired by mutations in the nucleocapsid zinc fingers. *Retrovirology* *4*, 54.
- Grigorov, B., Bocquin, A., Gabus, C., Avilov, S., Mély, Y., Agopian, A., Divita, G., Gottikh, M., Witvrouw, M., and Darlix, J.-L. (2011). Identification of a methylated oligoribonucleotide as a potent inhibitor of HIV-1 reverse transcription complex. *Nucleic Acids Res.* *39*, 5586–5596.
- Grobler, J., Friedman, E., Barrett, S.E., Wood, S.L., Ankrom, W., Fillgrove, K.L., Lai, M.-T., Gindy, M., Iwamoto, M., and Hazuda, D.J. (2016). Long-Acting Oral and Parenteral Dosing of MK-8591 for HIV Treatment or Prophylaxis | CROI Conference. p.
- Gu, M., Rappaport, J., and Leppla, S.H. (1995). Furin is important but not essential for the proteolytic maturation of gp160 of HIV-1. *FEBS Lett.* *365*, 95–97.
- Gubernick, S.I., Félix, N., Lee, D., Xu, J.J., and Hamad, B. (2016). The HIV therapy market. *Nat. Rev. Drug Discov.* *15*, 451–452.
- Guenzel, C.A., Hérate, C., and Benichou, S. (2014). HIV-1 Vpr—a still “enigmatic multitasker.” *Front. Microbiol.* *5*.
- Guo, X., Roldan, A., Hu, J., Wainberg, M.A., and Liang, C. (2005). Mutation of the SP1 sequence impairs both multimerization and membrane-binding activities of human immunodeficiency virus type 1 Gag. *J. Virol.* *79*, 1803–1812.
- Guzman, R.N.D., Wu, Z.R., Stalling, C.C., Pappalardo, L., Borer, P.N., and Summers, M.F. (1998). Structure of the HIV-1 Nucleocapsid Protein Bound to the SL3 Ψ-RNA Recognition

- Element. Science 279, 384–388.
- Haffar, O.K., Popov, S., Dubrovsky, L., Agostini, I., Tang, H., Pushkarsky, T., Nadler, S.G., and Bukrinsky, M. (2000). Two nuclear localization signals in the HIV-1 matrix protein regulate nuclear import of the HIV-1 pre-integration complex. *J. Mol. Biol.* 299, 359–368.
- Hallenberger, S., Bosch, V., Angliker, H., Shaw, E., Klenk, H.-D., and Garten, W. (1992). Inhibition of furin-mediated cleavage activation of HIV-1 glycoprotein gp160. *Nature* 360, 358–361.
- Hamard-Peron, E., and Muriaux, D. (2011). Retroviral matrix and lipids, the intimate interaction. *Retrovirology* 8, 15.
- Hargittai, M.R., Mangla, A.T., Gorelick, R.J., and Musier-Forsyth, K. (2001). HIV-1 nucleocapsid protein zinc finger structures induce tRNA(Lys,3) structural changes but are not critical for primer/template annealing. *J. Mol. Biol.* 312, 985–997.
- Harrich, D., Ulich, C., García-Martínez, L.F., and Gaynor, R.B. (1997). Tat is required for efficient HIV-1 reverse transcription. *EMBO J.* 16, 1224–1235.
- Harris, M. (1995). The role of myristoylation in the interactions between human immunodeficiency virus type I Nef and cellular proteins. *Biochem. Soc. Trans.* 23, 557–561.
- He, J., Choe, S., Walker, R., Di Marzio, P., Morgan, D.O., and Landau, N.R. (1995). Human immunodeficiency virus type 1 viral protein R (Vpr) arrests cells in the G2 phase of the cell cycle by inhibiting p34cdc2 activity. *J. Virol.* 69, 6705–6711.
- Hearps, A.C., Wagstaff, K.M., Piller, S.C., and Jans, D.A. (2008). The N-terminal basic domain of the HIV-1 matrix protein does not contain a conventional nuclear localization sequence but is required for DNA binding and protein self-association. *Biochemistry (Mosc.)* 47, 2199–2210.
- Hehl, E.A., Joshi, P., Kalpana, G.V., and Prasad, V.R. (2004). Interaction between Human Immunodeficiency Virus Type 1 Reverse Transcriptase and Integrase Proteins. *J. Virol.* 78, 5056–5067.
- Heim, R., Prasher, D.C., and Tsien, R.Y. (1994). Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *Proc. Natl. Acad. Sci.* 91, 12501–12504.
- Helga-Maria, C., Hammarskjöld, M.-L., and Rekosh, D. (1999). An Intact TAR Element and Cytoplasmic Localization Are Necessary for Efficient Packaging of Human Immunodeficiency Virus Type 1 Genomic RNA. *J. Virol.* 73, 4127–4135.
- Hemelaar, J. (2012). The origin and diversity of the HIV-1 pandemic. *Trends Mol. Med.* 18, 182–192.
- Henderson, B.R., and Percipalle, P. (1997). Interactions between HIV Rev and nuclear import and export factors: the Rev nuclear localisation signal mediates specific binding to human importin-beta. *J. Mol. Biol.* 274, 693–707.
- Hendrix, C.W., Collier, A.C., Lederman, M.M., Schols, D., Pollard, R.B., Brown, S., Jackson, J.B., Coombs, R.W., Glesby, M.J., Flexner, C.W., et al. (2004). Safety, pharmacokinetics, and antiviral activity of AMD3100, a selective CXCR4 receptor inhibitor, in HIV-1 infection. *JAIDS J. Acquir. Immune Defic. Syndr.* 37, 1253–1262.
- Hendrix, J., Baumgärtel, V., Schrimpf, W., Ivanchenko, S., Digman, M.A., Gratton, E., Kräusslich, H.-G., Müller, B., and Lamb, D.C. (2015). Live-cell observation of cytosolic HIV-1 assembly onset reveals RNA-interacting Gag oligomers. *J. Cell Biol.* 210, 629–646.
- Heng, X., Kharytonchyk, S., Garcia, E.L., Lu, K., Divakaruni, S.S., LaCotti, C., Edme, K., Telesnitsky, A., and Summers, M.F. (2012). Identification of a Minimal Region of the HIV-1 5'-Leader required for RNA Dimerization, NC Binding and Packaging. *J. Mol. Biol.* 417, 224–239.
- Henriet, S., Mercenne, G., Bernacchi, S., Paillart, J.-C., and Marquet, R. (2009). Tumultuous Relationship between the Human Immunodeficiency Virus Type 1 Viral Infectivity Factor (Vif) and the Human APOBEC-3G and APOBEC-3F Restriction Factors. *Microbiol. Mol. Biol. Rev.* MMBR 73, 211–232.

- Heredia, A., Latinovic, O.S., Barbault, F., and de Leeuw, E.P. (2015). A novel small-molecule inhibitor of HIV-1 entry. *Drug Des. Devel. Ther.* 9, 5469–5478.
- Hill, C.P., Worthylake, D., Bancroft, D.P., Christensen, A.M., and Sundquist, W.I. (1996). Crystal structures of the trimeric human immunodeficiency virus type 1 matrix protein: implications for membrane association and assembly. *Proc. Natl. Acad. Sci.* 93, 3099–3104.
- Hochuli, E., Döbeli, H., and Schacher, A. (1987). New metal chelate adsorbent selective for proteins and peptides containing neighbouring histidine residues. *J. Chromatogr.* 411, 177–184.
- Hochuli, E., Bannwarth, W., Döbeli, H., Gentz, R., and Stüber, D. (1988). Genetic Approach to Facilitate Purification of Recombinant Proteins with a Novel Metal Chelate Adsorbent. *Nat. Biotechnol.* 6, 1321–1325.
- Hoelz, A., Debler, E.W., and Blobel, G. (2011). The Structure of the Nuclear Pore Complex. *Annu. Rev. Biochem.* 80, 613–643.
- Hogg, J.R., and Goff, S.P. (2010). Upf1 senses 3'UTR length to potentiate mRNA decay. *Cell* 143, 379–389.
- Hogue, I.B., Hoppe, A., and Ono, A. (2009). Quantitative fluorescence resonance energy transfer microscopy analysis of the human immunodeficiency virus type 1 Gag-Gag interaction: relative contributions of the CA and NC domains and membrane binding. *J. Virol.* 83, 7322–7336.
- Horwitz, J.P., Chua, J., and Noel, M. (1964). Nucleosides. V. The Monomesylates of 1-(2'-Deoxy- β -D-lyxofuranosyl) thymine1, 2. *J. Org. Chem.* 29, 2076–2078.
- Houzet, L., Paillart, J.C., Smagulova, F., Maurel, S., Morichaud, Z., Marquet, R., and Mougel, M. (2007). HIV controls the selective packaging of genomic, spliced viral and cellular RNAs into virions through different mechanisms. *Nucleic Acids Res.* 35, 2695–2704.
- Howcroft, T.K., Strelbel, K., Martin, M.A., and Singer, D.S. (1993). Repression of MHC class I gene promoter activity by two-exon Tat of HIV. *Science* 260, 1320–1322.
- Hrecka, K., Swigut, T., Schindler, M., Kirchhoff, F., and Skowronski, J. (2005). Nef Proteins from Diverse Groups of Primate Lentiviruses Downmodulate CXCR4 To Inhibit Migration to the Chemokine Stromal Derived Factor 1. *J. Virol.* 79, 10650–10659.
- Hu, W.-S., and Hughes, S.H. (2012). HIV-1 Reverse Transcription. *Cold Spring Harb. Perspect. Med.* 2.
- Hu, C., Saenz, D.T., Fadel, H.J., Walker, W., Peretz, M., and Poeschla, E.M. (2010). The HIV-1 Central Polypurine Tract Functions as a Second Line of Defense against APOBEC3G/F. *J. Virol.* 84, 11981–11993.
- Huang, C., Tang, M., Zhang, M.-Y., Majeed, S., Montabana, E., Stanfield, R.L., Dimitrov, D.S., Korber, B., Sodroski, J., Wilson, I.A., et al. (2005). Structure of a V3-containing HIV-1 gp120 core. *Science* 310, 1025–1028.
- Huang, J., Wang, F., Argyris, E., Chen, K., Liang, Z., Tian, H., Huang, W., Squires, K., Verlinghieri, G., and Zhang, H. (2007). Cellular microRNAs contribute to HIV-1 latency in resting primary CD4+ T lymphocytes. *Nat. Med.* 13, 1241–1247.
- Huang, L., Bosch, I., Hofmann, W., Sodroski, J., and Pardee, A.B. (1998a). Tat protein induces human immunodeficiency virus type 1 (HIV-1) coreceptors and promotes infection with both macrophage-tropic and T-lymphotropic HIV-1 strains. *J. Virol.* 72, 8952–8960.
- Huang, M., Orenstein, J.M., Martin, M.A., and Freed, E.O. (1995). p6Gag is required for particle production from full-length human immunodeficiency virus type 1 molecular clones expressing protease. *J. Virol.* 69, 6810–6818.
- Huang, M., Maynard, A., Turpin, J.A., Graham, L., Janini, G.M., Covell, D.G., and Rice, W.G. (1998b). Anti-HIV agents that selectively target retroviral nucleocapsid protein zinc fingers without affecting cellular zinc finger proteins. *J. Med. Chem.* 41, 1371–1381.
- Hübner, W., Chen, P., Del Portillo, A., Liu, Y., Gordon, R.E., and Chen, B.K. (2007). Sequence of human immunodeficiency virus type 1 (HIV-1) Gag localization and oligomerization

- monitored with live confocal imaging of a replication-competent, fluorescently tagged HIV-1. *J. Virol.* *81*, 12596–12607.
- Hung, C.-H., Thomas, L., Ruby, C.E., Atkins, K.M., Morris, N.P., Knight, Z.A., Scholz, I., Barklis, E., Weinberg, A.D., Shokat, K.M., et al. (2007). HIV-1 Nef assembles a Src family kinase-ZAP-70/Syk-PI3K cascade to downregulate cell-surface MHC-I. *Cell Host Microbe* *1*, 121–133.
- Hurley, J.H., and Hanson, P.I. (2010). Membrane budding and scission by the ESCRT machinery: it's all in the neck. *Nat. Rev. Mol. Cell Biol.* *11*, 556–566.
- Hütter, G., Nowak, D., Mossner, M., Ganepola, S., Müssig, A., Allers, K., Schneider, T., Hofmann, J., Kücherer, C., Blau, O., et al. (2009). Long-term control of HIV by CCR5 Delta32/Delta32 stem-cell transplantation. *N. Engl. J. Med.* *360*, 692–698.
- Hwang, K.M., Qi, Y.M., Liu, S.-Y., Choy, W., and Chen, J. (1994). Inhibition and treatment of infection by enveloped virus with calix(n) arene compounds.
- Imaz, A., Falcó, V., and Ribera, E. (2011). Antiretroviral salvage therapy for multiclass drug-resistant HIV-1-infected patients: from clinical trials to daily clinical practice. *AIDS Rev.* *13*, 180–193.
- Iordanskiy, S., and Bukrinsky, M. (2007). Reverse transcription complex: the key player of the early phase of HIV replication. *Future Virol.* *2*, 49–64.
- Ivanchenko, S., Godinez, W.J., Lampe, M., Kräusslich, H.-G., Eils, R., Rohr, K., Bräuchle, C., Müller, B., and Lamb, D.C. (2009). Dynamics of HIV-1 Assembly and Release. *PLOS Pathog* *5*, e1000652.
- Jablonski, J.A., Buratti, E., Stuani, C., and Caputi, M. (2008). The secondary structure of the human immunodeficiency virus type 1 transcript modulates viral splicing and infectivity. *J. Virol.* *82*, 8038–8050.
- Jablonski, J.A., Amelio, A.L., Giacca, M., and Caputi, M. (2010). The transcriptional transactivator Tat selectively regulates viral splicing. *Nucleic Acids Res.* *38*, 1249–1260.
- Jacobo-Molina, A., Ding, J., Nanni, R.G., Clark, A.D., Lu, X., Tantillo, C., Williams, R.L., Kamer, G., Ferris, A.L., and Clark, P. (1993). Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 Å resolution shows bent DNA. *Proc. Natl. Acad. Sci.* *90*, 6320–6324.
- Jäger, S., Cimermancic, P., Gulbahce, N., Johnson, J.R., McGovern, K.E., Clarke, S.C., Shales, M., Mercenne, G., Pache, L., Li, K., et al. (2012). Global landscape of HIV-human protein complexes. *Nature* *481*, 365–370.
- Jalalirad, M., Saadatmand, J., and Laughrea, M. (2012). Dominant role of the 5' TAR bulge in dimerization of HIV-1 genomic RNA, but no evidence of TAR-TAR kissing during in vivo virus assembly. *Biochemistry (Mosc.)* *51*, 3744–3758.
- Jenkins, L.M.M., Byrd, J.C., Hara, T., Srivastava, P., Mazur, S.J., Stahl, S.J., Inman, J.K., Appella, E., Omichinski, J.G., and Legault, P. (2005). Studies on the mechanism of inactivation of the HIV-1 nucleocapsid protein NCp7 with 2-mercaptobenzamide thioesters. *J. Med. Chem.* *48*, 2847–2858.
- Jenkins, Y., McEntee, M., Weis, K., and Greene, W.C. (1998). Characterization of HIV-1 vpr nuclear import: analysis of signals and pathways. *J. Cell Biol.* *143*, 875–885.
- Jiang, J., Ablan, S.D., Derebail, S., Hercik, K., Soheilian, F., Thomas, J.A., Tang, S., Hewlett, I., Nagashima, K., Gorelick, R.J., et al. (2011). The interdomain linker region of HIV-1 capsid protein is a critical determinant of proper core assembly and stability. *Virology* *421*, 253–265.
- Johri, M.K., Mishra, R., Chhatbar, C., Unni, S.K., and Singh, S.K. (2011). Tits and bits of HIV Tat protein. *Expert Opin. Biol. Ther.* *11*, 269–283.
- Jolly, C., Kashefi, K., Hollinshead, M., and Sattentau, Q.J. (2004). HIV-1 cell to cell transfer across an Env-induced, actin-dependent synapse. *J. Exp. Med.* *199*, 283–293.
- Jolly, C., Welsch, S., Michor, S., and Sattentau, Q.J. (2011). The regulated secretory pathway

- in CD4(+) T cells contributes to human immunodeficiency virus type-1 cell-to-cell spread at the virological synapse. *PLoS Pathog.* 7, e1002226.
- Jones, C.P., Cantara, W.A., Olson, E.D., and Musier-Forsyth, K. (2014). Small-angle X-ray scattering-derived structure of the HIV-1 5' UTR reveals 3D tRNA mimicry. *Proc. Natl. Acad. Sci.* 111, 3395–3400.
- Joseph, S.B., Swanstrom, R., Kashuba, A.D.M., and Cohen, M.S. (2015). Bottlenecks in HIV-1 transmission: insights from the study of founder viruses. *Nat. Rev. Microbiol.* 13, 414–425.
- Joshi, A., Garg, H., Ablan, S.D., and Freed, E.O. (2011). Evidence of a role for soluble N-ethylmaleimide-sensitive factor attachment protein receptor (SNARE) machinery in HIV-1 assembly and release. *J. Biol. Chem.* 286, 29861–29871.
- Jouvenet, N., Bieniasz, P.D., and Simon, S.M. (2008). Imaging the biogenesis of individual HIV-1 virions in live cells. *Nature* 454, 236–240.
- Jouvenet, N., Simon, S.M., and Bieniasz, P.D. (2009). Imaging the interaction of HIV-1 genomes and Gag during assembly of individual viral particles. *Proc. Natl. Acad. Sci. U. S. A.* 106, 19114–19119.
- Jowett, J.B., Planelles, V., Poon, B., Shah, N.P., Chen, M.L., and Chen, I.S. (1995). The human immunodeficiency virus type 1 vpr gene arrests infected T cells in the G2 + M phase of the cell cycle. *J. Virol.* 69, 6304–6313.
- Ju, S.M., Song, H.Y., Lee, J.A., Lee, S.J., Choi, S.Y., and Park, J. (2009). Extracellular HIV-1 Tat up-regulates expression of matrix metalloproteinase-9 via a MAPK-NF- κ B dependent pathway in human astrocytes. *Exp. Mol. Med.* 41, 86–93.
- Kalia, V., Sarkar, S., Gupta, P., and Montelaro, R.C. (2003). Rational Site-Directed Mutations of the LLP-1 and LLP-2 Lentivirus Lytic Peptide Domains in the Intracytoplasmic Tail of Human Immunodeficiency Virus Type 1 gp41 Indicate Common Functions in Cell-Cell Fusion but Distinct Roles in Virion Envelope Incorporation. *J. Virol.* 77, 3634–3646.
- Kaminski, R., Chen, Y., Fischer, T., Tedaldi, E., Napoli, A., Zhang, Y., Karn, J., Hu, W., and Khalili, K. (2016). Elimination of HIV-1 Genomes from Human T-lymphoid Cells by CRISPR/Cas9 Gene Editing. *Sci. Rep.* 6, 22555.
- Karageorgos, L., Li, P., and Burrell, C. (1993). Characterization of HIV replication complexes early after cell-to-cell infection. *AIDS Res. Hum. Retroviruses* 9, 817–823.
- Karn, J., and Stoltzfus, C.M. (2012). Transcriptional and Posttranscriptional Regulation of HIV-1 Gene Expression. *Cold Spring Harb. Perspect. Med.* 2.
- Kasper, M.R., Roeth, J.F., Williams, M., Filzen, T.M., Fleis, R.I., and Collins, K.L. (2005). HIV-1 Nef disrupts antigen presentation early in the secretory pathway. *J. Biol. Chem.* 280, 12840–12848.
- Kaushik, R., and Ratner, L. (2004). Role of Human Immunodeficiency Virus Type 1 Matrix Phosphorylation in an Early Postentry Step of Virus Replication. *J. Virol.* 78, 2319–2326.
- Keele, B.F., Van Heuverswyn, F., Li, Y., Bailes, E., Takehisa, J., Santiago, M.L., Bibollet-Ruche, F., Chen, Y., Wain, L.V., Liegeois, F., et al. (2006). Chimpanzee reservoirs of pandemic and nonpandemic HIV-1. *Science* 313, 523–526.
- Kempf, N., Postupalenko, V., Bora, S., Didier, P., Arntz, Y., de Rocquigny, H., and Mély, Y. (2014). The HIV-1 Nucleocapsid Protein Recruits Negatively Charged Lipids To Ensure Its Optimal Binding to Lipid Membranes. *J. Virol.* 89, 1756–1767.
- Kharytonchyk, S., Monti, S., Smaldino, P.J., Van, V., Bolden, N.C., Brown, J.D., Russo, E., Swanson, C., Shuey, A., Telesnitsky, A., et al. (2016). Transcriptional start site heterogeneity modulates the structure and function of the HIV-1 genome. *Proc. Natl. Acad. Sci.* 113, 13378–13383.
- Kiernan, R.E., Ono, A., Englund, G., and Freed, E.O. (1998). Role of Matrix in an Early Postentry Step in the Human Immunodeficiency Virus Type 1 Life Cycle. *J. Virol.* 72, 4116–4126.

- Kilareski, E.M., Shah, S., Nonnemacher, M.R., and Wigdahl, B. (2009). Regulation of HIV-1 transcription in cells of the monocyte-macrophage lineage. *Retrovirology* 6, 118.
- Kim, M.-J., Kim, S.H., Park, J.A., Yu, K.L., Jang, S.I., Kim, B.S., Lee, E.S., and You, J.C. (2015). Identification and characterization of a new type of inhibitor against the human immunodeficiency virus type-1 nucleocapsid protein. *Retrovirology* 12, 90.
- Kirchhoff, F., Schindler, M., Specht, A., Arhel, N., and Münch, J. (2008). Role of Nef in primate lentiviral immunopathogenesis. *Cell. Mol. Life Sci. CMLS* 65, 2621–2636.
- Klingler, J. (2015). Caractérisation de l'interaction entre la protéine rétrovirale Gag et les protéines cellulaires et virales. Université de Strasbourg, Faculté de Pharmacie.
- Klotman, M.E., Kim, S., Buchbinder, A., DeRossi, A., Baltimore, D., and Wong-Staal, F. (1991). Kinetics of expression of multiply spliced RNA in early human immunodeficiency virus type 1 infection of lymphocytes and monocytes. *Proc. Natl. Acad. Sci.* 88, 5011–5015.
- Koh, Y., Wu, X., Ferris, A.L., Matreyek, K.A., Smith, S.J., Lee, K., KewalRamani, V.N., Hughes, S.H., and Engelman, A. (2013). Differential effects of human immunodeficiency virus type 1 capsid and cellular factors nucleoporin 153 and LEDGF/p75 on the efficiency and specificity of viral DNA integration. *J. Virol.* 87, 648–658.
- Kohlstaedt, L.A., Wang, J., Friedman, J.M., Rice, P.A., and Steitz, T.A. (1992). Crystal structure at 3.5 Å resolution of HIV-1 reverse transcriptase complexed with an inhibitor. *Science* 256, 1783–1790.
- Korber, B., Muldoon, M., Theiler, J., Gao, F., Gupta, R., Lapedes, A., Hahn, B.H., Wolinsky, S., and Bhattacharya, T. (2000). Timing the ancestor of the HIV-1 pandemic strains. *Science* 288, 1789–1796.
- Kovaleski, B.J., Kennedy, R., Hong, M.K., Datta, S.A., Kleiman, L., Rein, A., and Musier-Forsyth, K. (2006). In Vitro Characterization of the Interaction between HIV-1 Gag and Human Lysyl-tRNA Synthetase. *J. Biol. Chem.* 281, 19449–19456.
- Kräusslich, H.G., Fäcke, M., Heuser, A.M., Konvalinka, J., and Zentgraf, H. (1995). The spacer peptide between human immunodeficiency virus capsid and nucleocapsid proteins is essential for ordered assembly and viral infectivity. *J. Virol.* 69, 3407–3419.
- Kulkosky, J., Jones, K.S., Katz, R.A., Mack, J.P., and Skalka, A.M. (1992). Residues critical for retroviral integrative recombination in a region that is highly conserved among retroviral/retrotransposon integrases and bacterial insertion sequence transposases. *Mol. Cell. Biol.* 12, 2331–2338.
- Kunze, J., Todoroff, N., Schneider, P., Rodrigues, T., Geppert, T., Reisen, F., Schreuder, H., Saas, J., Hessler, G., Baringhaus, K.-H., et al. (2014). Targeting Dynamic Pockets of HIV-1 Protease by Structure-Based Computational Screening for Allosteric Inhibitors. *J. Chem. Inf. Model.* 54, 987–991.
- Kutluay, S.B., and Bieniasz, P.D. (2010). Analysis of the Initiating Events in HIV-1 Particle Assembly and Genome Packaging. *PLOS Pathog* 6, e1001200.
- Kutluay, S.B., Zang, T., Blanco-Melo, D., Powell, C., Jannain, D., Errando, M., and Bieniasz, P.D. (2014). Global changes in the RNA binding specificity of HIV-1 gag regulate virion genesis. *Cell* 159, 1096–1109.
- Kuzembayeva, M., Dilley, K., Sardo, L., and Hu, W.-S. (2014). Life of psi: How full-length HIV-1 RNAs become packaged genomes in the viral particles. *Virology* 454–455, 362–370.
- Kwong, P.D., Wyatt, R., Robinson, J., Sweet, R.W., Sodroski, J., and Hendrickson, W.A. (1998). Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature* 393, 648–659.
- Lahaye, X., Satoh, T., Gentili, M., Cerboni, S., Conrad, C., Hurbain, I., El Marjou, A., Lacabaratz, C., Lelièvre, J.-D., and Manel, N. (2013). The capsids of HIV-1 and HIV-2 determine immune detection of the viral cDNA by the innate sensor cGAS in dendritic cells. *Immunity* 39, 1132–1142.

- Lalezari, J.P., Latiff, G.H., Brinson, C., Echevarría, J., Treviño-Pérez, S., Bogner, J.R., Thompson, M., Fourie, J., Sussmann Pena, O.A., Mendo Urbina, F.C., et al. (2015). Safety and efficacy of the HIV-1 attachment inhibitor prodrug BMS-663068 in treatment-experienced individuals: 24 week results of AI438011, a phase 2b, randomised controlled trial. *Lancet HIV* 2, e427-437.
- Lambotte, O., Boufassa, F., Madec, Y., Nguyen, A., Goujard, C., Meyer, L., Rouzioux, C., Venet, A., Delfraissy, J.-F., and Group, S.-H.S. (2005). HIV Controllers: A Homogeneous Group of HIV-1—Infected Patients with Spontaneous Control of Viral Replication. *Clin. Infect. Dis.* 41, 1053–1056.
- Lamorte, L., Titolo, S., Lemke, C.T., Goudreau, N., Mercier, J.-F., Wardrop, E., Shah, V.B., von Schwedler, U.K., Langelier, C., Banik, S.S.R., et al. (2013). Discovery of novel small-molecule HIV-1 replication inhibitors that stabilize capsid complexes. *Antimicrob. Agents Chemother.* 57, 4622–4631.
- Landschulz, W.H., Johnson, P.F., and McKnight, S.L. (1988). The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. *Science* 240, 1759–1764.
- Lawrence, D.C., Stover, C.C., Noznitsky, J., Wu, Z., and Summers, M.F. (2003). Structure of the Intact Stem and Bulge of HIV-1 Ψ -RNA Stem-Loop SL1. *J. Mol. Biol.* 326, 529–542.
- Lee, S.-F. (2000). Multimerization Potential of the Cytoplasmic Domain of the Human Immunodeficiency Virus Type 1 Transmembrane Glycoprotein gp41. *J. Biol. Chem.* 275, 15809–15819.
- Legrand, N., Ploss, A., Balling, R., Becker, P.D., Borsotti, C., Brezillon, N., Debarry, J., Jong, Y. de, Deng, H., Santo, J.P.D., et al. (2009). Humanized Mice for Modeling Human Infectious Disease: Challenges, Progress, and Outlook. *Cell Host Microbe* 6, 5–9.
- Lehmann, M.J., Sherer, N.M., Marks, C.B., Pypaert, M., and Mothes, W. (2005). Actin- and myosin-driven movement of viruses along filopodia precedes their entry into cells. *J. Cell Biol.* 170, 317–325.
- Leibman, R.S., and Riley, J.L. (2015). Engineering T Cells to Functionally Cure HIV-1 Infection. *Mol. Ther.* 23, 1149–1159.
- Leihener, A., Ludwig, C., and Wagner, R. (2009). Uncoupling Human Immunodeficiency Virus Type 1 gag and pol Reading Frames: Role of the Transframe Protein p6* in Viral Replication. *J. Virol.* 83, 7210–7220.
- Lemke, C.T., Titolo, S., von Schwedler, U., Goudreau, N., Mercier, J.-F., Wardrop, E., Faucher, A.-M., Coulombe, R., Banik, S.S.R., Fader, L., et al. (2012). Distinct Effects of Two HIV-1 Capsid Assembly Inhibitor Families That Bind the Same Site within the N-Terminal Domain of the Viral CA Protein. *J. Virol.* 86, 6643–6655.
- Leonard, C.K., Spellman, M.W., Riddle, L., Harris, R.J., Thomas, J.N., and Gregory, T.J. (1990). Assignment of intrachain disulfide bonds and characterization of potential glycosylation sites of the type 1 recombinant human immunodeficiency virus envelope glycoprotein (gp120) expressed in Chinese hamster ovary cells. *J. Biol. Chem.* 265, 10373–10382.
- Lever, A., Gottlinger, H., Haseltine, W., and Sodroski, J. (1989). Identification of a sequence required for efficient packaging of human immunodeficiency virus type 1 RNA into virions. *J. Virol.* 63, 4085–4087.
- Levin, J.G., Guo, J., Rouzina, I., and Musier-Forsyth, K. (2005). Nucleic acid chaperone activity of HIV-1 nucleocapsid protein: critical role in reverse transcription and molecular mechanism. *Prog. Nucleic Acid Res. Mol. Biol.* 80, 217–286.
- Li, C.J., Friedman, D.J., Wang, C., Metelev, V., and Pardee, A.B. (1995). Induction of apoptosis in uninfected lymphocytes by HIV-1 Tat protein. *Science* 268, 429–431.
- Li, F., Goila-Gaur, R., Salzwedel, K., Kilgore, N.R., Reddick, M., Matallana, C., Castillo, A.,

- Zoumplis, D., Martin, D.E., Orenstein, J.M., et al. (2003). PA-457: A potent HIV inhibitor that disrupts core condensation by targeting a late step in Gag processing. *Proc. Natl. Acad. Sci.* *100*, 13555–13560.
- Liang, C., Hu, J., Russell, R.S., Roldan, A., Kleiman, L., and Wainberg, M.A. (2002). Characterization of a Putative α -Helix across the Capsid-SP1 Boundary That Is Critical for the Multimerization of Human Immunodeficiency Virus Type 1 Gag. *J. Virol.* *76*, 11729–11737.
- Lingappa, J.R., Reed, J.C., Tanaka, M., Chutiraka, K., and Robinson, B.A. (2014a). How HIV-1 Gag assembles in cells: Putting together pieces of the puzzle. *Virus Res.* *193*, 89–107.
- Lingappa, J.R., Reed, J.C., Tanaka, M., Chutiraka, K., and Robinson, B.A. (2014b). How HIV-1 Gag assembles in cells: Putting together pieces of the puzzle. *Virus Res.* *193*, 89–107.
- Liu, B., Dai, R., Tian, C.J., Dawson, L., Gorelick, R., and Yu, X.F. (1999). Interaction of the human immunodeficiency virus type 1 nucleocapsid with actin. *J. Virol.* *73*, 2901–2908.
- Lochrie, M.A., Waugh, S., Pratt, D.G., Clever, J., Parslow, T.G., and Polisky, B. (1997). In vitro selection of RNAs that bind to the human immunodeficiency virus type-1 gag polyprotein. *Nucleic Acids Res.* *25*, 2902–2910.
- Lodge, R., Lalonde, J.P., Lemay, G., and Cohen, E.A. (1997). The membrane-proximal intracytoplasmic tyrosine residue of HIV-1 envelope glycoprotein is critical for basolateral targeting of viral budding in MDCK cells. *EMBO J.* *16*, 695–705.
- Lodi, P.J., Ernst, J.A., Kuszewski, J., Hickman, A.B., Engelman, A., Craigie, R., Clore, G.M., and Gronenborn, A.M. (1995). Solution structure of the DNA binding domain of HIV-1 integrase. *Biochemistry (Mosc.)* *34*, 9826–9833.
- Loret, E.P., Darque, A., Jouve, E., Loret, E.A., Nicolino-Brunet, C., Morange, S., Castanier, E., Casanova, J., Caloustian, C., Bornet, C., et al. (2016). Intradermal injection of a Tat Oyi-based therapeutic HIV vaccine reduces of 1.5 log copies/mL the HIV RNA rebound median and no HIV DNA rebound following cART interruption in a phase I/II randomized controlled clinical trial. *Retrovirology* *13*, 21.
- Louis, J.M., Tözsér, J., Roche, J., Matúz, K., Aniana, A., and Sayer, J.M. (2013). Enhanced stability of monomer fold correlates with extreme drug resistance of HIV-1 protease. *Biochemistry (Mosc.)* *52*, 7678–7688.
- Lowe, D.M., Aitken, A., Bradley, C., Darby, G.K., Larder, B.A., Powell, K.L., Purifoy, D.J., Tisdale, M., and Stammers, D.K. (1988). HIV-1 reverse transcriptase: crystallization and analysis of domain structure by limited proteolysis. *Biochemistry (Mosc.)* *27*, 8884–8889.
- Lu, K., Heng, X., and Summers, M.F. (2011). Structural Determinants and Mechanism of HIV-1 Genome Packaging. *J. Mol. Biol.* *410*, 609–633.
- Lu, M., Blacklow, S.C., and Kim, P.S. (1995). A trimeric structural domain of the HIV-1 transmembrane glycoprotein. *Nat. Struct. Biol.* *2*, 1075–1082.
- Luo, Z., Xu, X., Zhang, X., and Hu, L. (2013). Development of calixarenes, cyclodextrins and fullerenes as new platforms for anti-HIV drug design: an overview. *Mini Rev. Med. Chem.* *13*, 1160–1165.
- Luo, Z., Zhao, Y., Ma, C., Li, Z., Xu, X., Hu, L., Huang, N., and He, H. (2015). Synthesis, biological evaluation and molecular docking of calix[4]arene-based β -diketo derivatives as HIV-1 integrase inhibitors. *Arch. Pharm. (Weinheim)* *348*, 206–213.
- Luo, Z.-G., Zhao, Y., Ma, C., Xu, X.-M., Zhang, X.-M., Huang, N.-Y., and He, H.-Q. (2014). Synthesis and anti-integrase evaluation of novel calix[4]arene derivatives containing the triazolyl 1,3-diketo moiety. *Chin. Chem. Lett.* *25*, 737–740.
- Lynch, R.M., Boritz, E., Coates, E.E., DeZure, A., Madden, P., Costner, P., Enama, M.E., Plummer, S., Holman, L., Hendel, C.S., et al. (2015). Virologic effects of broadly neutralizing antibody VRC01 administration during chronic HIV-1 infection. *Sci. Transl. Med.* *7*, 319ra206–319ra206.
- Lyonnais, S., Gorelick, R.J., Heniche-Boukhalfa, F., Bouaziz, S., Parissi, V., Mouscadet, J.-F.,

- Restle, T., Gatell, J.M., Le Cam, E., and Mirambeau, G. (2013). A protein ballet around the viral genome orchestrated by HIV-1 reverse transcriptase leads to an architectural switch: from nucleocapsid-condensed RNA to Vpr-bridged DNA. *Virus Res.* *171*, 287–303.
- Ma, Y., He, Z., Tan, T., Li, W., Zhang, Z., Song, S., Zhang, X., Hu, Q., Zhou, P., Wu, Y., et al. (2016). Real-Time Imaging of Single HIV-1 Disassembly with Multicolor Viral Particles. *ACS Nano* *10*, 6273–6282.
- Maddon, P.J., Dalgleish, A.G., McDougal, J.S., Clapham, P.R., Weiss, R.A., and Axel, R. (1986). The T4 gene encodes the AIDS virus receptor and is expressed in the immune system and the brain. *Cell* *47*, 333–348.
- Mäe, M., and Langel, U. (2006). Cell-penetrating peptides as vectors for peptide, protein and oligonucleotide delivery. *Curr. Opin. Pharmacol.* *6*, 509–514.
- Maldonado, J.O., Martin, J.L., Mueller, J.D., Zhang, W., and Mansky, L.M. (2014). New insights into retroviral Gag-Gag and Gag-membrane interactions. *Front. Microbiol.* *5*, 302.
- Malim, M.H., Böhnlein, S., Fenrick, R., Le, S.Y., Maizel, J.V., and Cullen, B.R. (1989). Functional comparison of the Rev trans-activators encoded by different primate immunodeficiency virus species. *Proc. Natl. Acad. Sci. U. S. A.* *86*, 8222–8226.
- Mancebo, H.S., Lee, G., Flygare, J., Tomassini, J., Luu, P., Zhu, Y., Peng, J., Blau, C., Hazuda, D., Price, D., et al. (1997). P-TEFb kinase is required for HIV Tat transcriptional activation in vivo and in vitro. *Genes Dev.* *11*, 2633–2644.
- Mangeat, B., Turelli, P., Caron, G., Friedli, M., Perrin, L., and Trono, D. (2003). Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. *Nature* *424*, 99–103.
- Mansky, L.M. (1996). The mutation rate of human immunodeficiency virus type 1 is influenced by the vpr gene. *Virology* *222*, 391–400.
- Mansky, L.M., Preveral, S., Selig, L., Benarous, R., and Benichou, S. (2000). The interaction of vpr with uracil DNA glycosylase modulates the human immunodeficiency virus type 1 In vivo mutation rate. *J. Virol.* *74*, 7039–7047.
- Margolis, D.M. (2010). Mechanisms of HIV Latency: an Emerging Picture of Complexity. *Curr. HIV/AIDS Rep.* *7*, 37–43.
- Mariani, R., Chen, D., Schröfelbauer, B., Navarro, F., König, R., Bollman, B., Münk, C., Nymark-McMahon, H., and Landau, N.R. (2003). Species-specific exclusion of APOBEC3G from HIV-1 virions by Vif. *Cell* *114*, 21–31.
- Marin, M., Rose, K.M., Kozak, S.L., and Kabat, D. (2003). HIV-1 Vif protein binds the editing enzyme APOBEC3G and induces its degradation. *Nat. Med.* *9*, 1398–1403.
- Marini, B., Kertesz-Farkas, A., Ali, H., Lucic, B., Lisek, K., Manganaro, L., Pongor, S., Luzzati, R., Recchia, A., Mavilio, F., et al. (2015). Nuclear architecture dictates HIV-1 integration site selection. *Nature* *521*, 227–231.
- Martin, N., Welsch, S., Jolly, C., Briggs, J.A.G., Vaux, D., and Sattentau, Q.J. (2010). Virological synapse-mediated spread of human immunodeficiency virus type 1 between T cells is sensitive to entry inhibition. *J. Virol.* *84*, 3516–3527.
- Martin-Serrano, J., Zang, T., and Bieniasz, P.D. (2001). HIV-1 and Ebola virus encode small peptide motifs that recruit Tsg101 to sites of particle assembly to facilitate egress. *Nat. Med.* *7*, 1313–1319.
- di Marzo Veronese, F., Copeland, T.D., DeVico, A.L., Rahman, R., Oroszlan, S., Gallo, R.C., and Sarngadharan, M.G. (1986). Characterization of highly immunogenic p66/p51 as the reverse transcriptase of HTLV-III/LAV. *Science* *231*, 1289–1291.
- Matreyek, K.A., and Engelman, A. (2011). The requirement for nucleoporin NUP153 during human immunodeficiency virus type 1 infection is determined by the viral capsid. *J. Virol.* *85*, 7818–7827.
- Matreyek, K.A., and Engelman, A. (2013). Viral and Cellular Requirements for the Nuclear

- Entry of Retroviral Preintegration Nucleoprotein Complexes. *Viruses* 5, 2483–2511.
- Mattei, S., Schur, F.K., and Briggs, J.A. (2016). Retrovirus maturation — an extraordinary structural transformation. *Curr. Opin. Virol.* 18, 27–35.
- Matthews, T., Salgo, M., Greenberg, M., Chung, J., DeMasi, R., and Bolognesi, D. (2004). Case history: Enfuvirtide: the first therapy to inhibit the entry of HIV-1 into host CD4 lymphocytes. *Nat. Rev. Drug Discov.* 3, 215–225.
- Mayol, K., Munier, S., Beck, A., Verrier, B., and Guillon, C. (2007). Design and characterization of an HIV-1 Tat mutant: Inactivation of viral and cellular functions but not antigenicity. *Vaccine* 25, 6047–6060.
- McDonald, D., Vodicka, M.A., Lucero, G., Svitkina, T.M., Borisy, G.G., Emerman, M., and Hope, T.J. (2002). Visualization of the intracellular behavior of HIV in living cells. *J. Cell Biol.* 159, 441–452.
- McDonald, D., Wu, L., Bohks, S.M., KewalRamani, V.N., Unutmaz, D., and Hope, T.J. (2003). Recruitment of HIV and its receptors to dendritic cell-T cell junctions. *Science* 300, 1295–1297.
- McDougal, J.S., Kennedy, M.S., Sligh, J.M., Cort, S.P., Mawle, A., and Nicholson, J.K. (1986). Binding of HTLV-III/LAV to T4+ T cells by a complex of the 110K viral protein and the T4 molecule. *Science* 231, 382–385.
- McKinstry, W.J., Hijnen, M., Tanwar, H.S., Sparrow, L.G., Nagarajan, S., Pham, S.T., and Mak, J. (2014). Expression and purification of soluble recombinant full length HIV-1 Pr55(Gag) protein in *Escherichia coli*. *Protein Expr. Purif.* 100, 10–18.
- Melikyan, G.B. (2008). Common principles and intermediates of viral protein-mediated fusion: the HIV-1 paradigm. *Retrovirology* 5, 111.
- Mély, Y., De Rocquigny, H., Morellet, N., Roques, B.P., and Gérard, D. (1996). Zinc Binding to the HIV-1 Nucleocapsid Protein: A Thermodynamic Investigation by Fluorescence Spectroscopy. *Biochemistry (Mosc.)* 35, 5175–5182.
- Mely, Y., Rocquigny, H. d, Shvadchak, V., Avilov, S., Dong, C.Z., Dietrich, U., and Darlix, J.-L. (2008). Targeting the viral nucleocapsid protein in anti-HIV-1 therapy. *Mini Rev. Med. Chem.* 8, 24–35.
- Micci, L., Ryan, E.S., Fromentin, R., Bosinger, S.E., Harper, J.L., He, T., Paganini, S., Easley, K.A., Chahroudi, A., Benne, C., et al. (2015). Interleukin-21 combined with ART reduces inflammation and viral reservoir in SIV-infected macaques. *J. Clin. Invest.* 125, 4497–4513.
- Miele, G., and Lever, A.M. (1996). Expression of mutant and wild-type gag proteins for gene therapy in HIV-1 infection. *Gene Ther.* 3, 357–361.
- Miles, L.R., Agresta, B.E., Khan, M.B., Tang, S., Levin, J.G., and Powell, M.D. (2005). Effect of Polypurine Tract (PPT) Mutations on Human Immunodeficiency Virus Type 1 Replication: a Virus with a Completely Randomized PPT Retains Low Infectivity. *J. Virol.* 79, 6859–6867.
- Milev, M.P., Brown, C.M., and Moulard, A.J. (2010). Live cell visualization of the interactions between HIV-1 Gag and the cellular RNA-binding protein Staufen1. *Retrovirology* 7, 41.
- Miller Jenkins, L.M., Ott, D.E., Hayashi, R., Coren, L.V., Wang, D., Xu, Q., Schito, M.L., Inman, J.K., Appella, D.H., and Appella, E. (2010). Small-molecule inactivation of HIV-1 NCp7 by repetitive intracellular acyl transfer. *Nat. Chem. Biol.* 6, 887–889.
- Mirambeau, G., Lyonais, S., Coulaud, D., Hameau, L., Lafosse, S., Jeusset, J., Borde, I., Reboud-Ravaux, M., Restle, T., Gorelick, R.J., et al. (2007). HIV-1 protease and reverse transcriptase control the architecture of their nucleocapsid partner. *PloS One* 2, e669.
- Moeyaert, B., Nguyen Bich, N., De Zitter, E., Rocha, S., Clays, K., Mizuno, H., van Meervelt, L., Hofkens, J., and Dedeker, P. (2014). Green-to-Red Photoconvertible Dronpa Mutant for Multimodal Super-resolution Fluorescence Microscopy. *ACS Nano* 8, 1664–1673.
- Momany, C., Kovari, L.C., Prongay, A.J., Keller, W., Gitti, R.K., Lee, B.M., Gorbalenya, A.E., Tong, L., McClure, J., Ehrlich, L.S., et al. (1996). Crystal structure of dimeric HIV-1 capsid

- protein. *Nat. Struct. Biol.* *3*, 763–770.
- Montefiori, D.C., Robinson, W.E., and Mitchell, W.M. (1988). Role of protein N-glycosylation in pathogenesis of human immunodeficiency virus type 1. *Proc. Natl. Acad. Sci.* *85*, 9248–9252.
- Moore, M.D., and Hu, W.-S. (2009). HIV-1 RNA dimerization: It takes two to tango. *AIDS Rev.* *11*, 91.
- Moore, M.D., Fu, W., Nikolaitchik, O., Chen, J., Ptak, R.G., and Hu, W.-S. (2007). Dimer Initiation Signal of Human Immunodeficiency Virus Type 1: Its Role in Partner Selection during RNA Copackaging and Its Effects on Recombination. *J. Virol.* *81*, 4002–4011.
- Morcock, D.R., Thomas, J.A., Gagliardi, T.D., Gorelick, R.J., Roser, J.D., Chertova, E.N., Bess, J.W., Ott, D.E., Sattentau, Q.J., Frank, I., et al. (2005). Elimination of retroviral infectivity by N-ethylmaleimide with preservation of functional envelope glycoproteins. *J. Virol.* *79*, 1533–1542.
- Morellet, N., Roques, B.P., and Bouaziz, S. (2009). Structure-function relationship of Vpr: biological implications. *Curr. HIV Res.* *7*, 184–210.
- Mori, M., Dietrich, U., Manetti, F., and Botta, M. (2010). Molecular dynamics and DFT study on HIV-1 nucleocapsid protein-7 in complex with viral genome. *J. Chem. Inf. Model.* *50*, 638–650.
- Mori, M., Manetti, F., and Botta, M. (2011a). Targeting protein-protein and protein-nucleic acid interactions for anti-HIV therapy. *Curr. Pharm. Des.* *17*, 3713–3728.
- Mori, M., Manetti, F., and Botta, M. (2011b). Predicting the binding mode of known NCp7 inhibitors to facilitate the design of novel modulators. *J. Chem. Inf. Model.* *51*, 446–454.
- Mori, M., Nucci, A., Lang, M.C.D., Humbert, N., Boudier, C., Debaene, F., Sanglier-Cianferani, S., Catala, M., Schult-Dietrich, P., Dietrich, U., et al. (2014). Functional and structural characterization of 2-amino-4-phenylthiazole inhibitors of the HIV-1 nucleocapsid protein with antiviral activity. *ACS Chem. Biol.* *9*, 1950–1955.
- Mori, M., Kovalenko, L., Lyonnais, S., Antaki, D., Torbett, B.E., Botta, M., Mirambeau, G., and Mély, Y. (2015). Nucleocapsid Protein: A Desirable Target for Future Therapies Against HIV-1. In *The Future of HIV-1 Therapeutics*, B.E. Torbett, D.S. Goodsell, and D.D. Richman, eds. (Cham: Springer International Publishing), pp. 53–92.
- Mourez, T., Simon, F., and Plantier, J.-C. (2013). Non-M Variants of Human Immunodeficiency Virus Type 1. *Clin. Microbiol. Rev.* *26*, 448–461.
- Mousseau, G., Mediouni, S., and Valente, S.T. (2015). Targeting HIV transcription: the quest for a functional cure. *Curr. Top. Microbiol. Immunol.* *389*, 121–145.
- Mukhopadhyay, J., Das, K., Ismail, S., Koppstein, D., Jang, M., Hudson, B., Sarafianos, S., Tuske, S., Patel, J., Jansen, R., et al. (2008). The RNA Polymerase “Switch Region” Is a Target for Inhibitors. *Cell* *135*, 295–307.
- Muller, B., Daecke, J., Fackler, O.T., Dittmar, M.T., Zentgraf, H., and Krausslich, H.-G. (2004). Construction and Characterization of a Fluorescently Labeled Infectious Human Immunodeficiency Virus Type 1 Derivative. *J. Virol.* *78*, 10803–10813.
- Müller, D.R., Schindler, P., Towbin, H., Wirth, U., Voshol, H., Hoving, S., and Steinmetz, M.O. (2001). Isotope-Tagged Cross-Linking Reagents. A New Tool in Mass Spectrometric Protein Interaction Analysis. *Anal. Chem.* *73*, 1927–1934.
- Mundigala, H., Michaux, J.B., Feig, A.L., Ennifar, E., and Rueda, D. (2014). HIV-1 DIS stem loop forms an obligatory bent kissing intermediate in the dimerization pathway. *Nucleic Acids Res.* *42*, 7281–7289.
- Munir, S., Thierry, S., Subra, F., Deprez, E., and Delelis, O. (2013). Quantitative analysis of the time-course of viral DNA forms during the HIV-1 life cycle. *Retrovirology* *10*, 87.
- Muñoz-Barroso, I., Salzwedel, K., Hunter, E., and Blumenthal, R. (1999). Role of the Membrane-Proximal Domain in the Initial Stages of Human Immunodeficiency Virus Type 1

- Envelope Glycoprotein-Mediated Membrane Fusion. *J. Virol.* 73, 9693–9693.
- Munro, J.B., Nath, A., Farber, M., Datta, S.A.K., Rein, A., Rhoades, E., and Mothes, W. (2014). A Conformational Transition Observed in Single HIV-1 Gag Molecules during In Vitro Assembly of Virus-Like Particles. *J. Virol.* 88, 3577–3585.
- Muriaux, D., and Darlix, J.-L. (2010). Properties and functions of the nucleocapsid protein in virus assembly. *RNA Biol.* 7, 744–753.
- Mussolino, C., Morbitzer, R., Lütge, F., Dannemann, N., Lahaye, T., and Cathomen, T. (2011). A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic Acids Res.* 39, 9283–9293.
- Naldini, L., Blömer, U., Gallay, P., Ory, D., Mulligan, R., Gage, F.H., Verma, I.M., and Trono, D. (1996). In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science* 272, 263–267.
- Neil, S.J.D., Zang, T., and Bieniasz, P.D. (2008). Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. *Nature* 451, 425–430.
- Nettles, R.E., Schürmann, D., Zhu, L., Stonier, M., Huang, S.-P., Chang, I., Chien, C., Krystal, M., Wind-Rotolo, M., Ray, N., et al. (2012). Pharmacodynamics, Safety, and Pharmacokinetics of BMS-663068, an Oral HIV-1 Attachment Inhibitor in HIV-1–Infected Subjects. *J. Infect. Dis.* 206, 1002–1011.
- Newman, J.L., Butcher, E.W., Patel, D.T., Mikhaylenko, Y., and Summers, M.F. (2004). Flexibility in the P2 domain of the HIV-1 Gag polyprotein. *Protein Sci. Publ. Protein Soc.* 13, 2101–2107.
- Nishi, M., Ryo, A., Tsurutani, N., Ohba, K., Sawasaki, T., Morishita, R., Perrem, K., Aoki, I., Morikawa, Y., and Yamamoto, N. (2009). Requirement for microtubule integrity in the SOCS1-mediated intracellular dynamics of HIV-1 Gag. *FEBS Lett.* 583, 1243–1250.
- Onafuwa-Nuga, A., and Telesnitsky, A. (2009). The remarkable frequency of human immunodeficiency virus type 1 genetic recombination. *Microbiol. Mol. Biol. Rev. MMBR* 73, 451–480, Table of Contents.
- Ono, A., and Freed, E.O. (1999). Binding of human immunodeficiency virus type 1 Gag to membrane: role of the matrix amino terminus. *J. Virol.* 73, 4136–4144.
- Ono, A., Orenstein, J.M., and Freed, E.O. (2000a). Role of the Gag Matrix Domain in Targeting Human Immunodeficiency Virus Type 1 Assembly. *J. Virol.* 74, 2855–2866.
- Ono, A., Demirov, D., and Freed, E.O. (2000b). Relationship between human immunodeficiency virus type 1 Gag multimerization and membrane binding. *J. Virol.* 74, 5142–5150.
- Ono, A., Ablan, S.D., Lockett, S.J., Nagashima, K., and Freed, E.O. (2004). Phosphatidylinositol (4, 5) bisphosphate regulates HIV-1 Gag targeting to the plasma membrane. *Proc. Natl. Acad. Sci. U. S. A.* 101, 14889–14894.
- ONUSIDA (2014). 90-90-90 Une cible ambitieuse de traitement pour aider à mettre fin à l'épidémie du sida.
- Ott, D.E., Coren, L.V., and Gagliardi, T.D. (2005). Redundant Roles for Nucleocapsid and Matrix RNA-Binding Sequences in Human Immunodeficiency Virus Type 1 Assembly. *J. Virol.* 79, 13839–13847.
- Ott, D.E., Coren, L.V., and Shatzer, T. (2009). The nucleocapsid region of human immunodeficiency virus type 1 Gag assists in the coordination of assembly and Gag processing: role for RNA-Gag binding in the early stages of assembly. *J. Virol.* 83, 7718–7727.
- Ott, M., Lovett, J.L., Mueller, L., and Verdin, E. (1998). Superinduction of IL-8 in T Cells by HIV-1 Tat Protein Is Mediated Through NF- κ B Factors. *J. Immunol.* 160, 2872–2880.
- Paillart, J.C., and Göttlinger, H.G. (1999). Opposing effects of human immunodeficiency virus type 1 matrix mutations support a myristyl switch model of gag membrane targeting. *J. Virol.* 73, 2604–2612.

- Pannecouque, C., Szafarowicz, B., Volkova, N., Bakulev, V., Dehaen, W., Mély, Y., and Daelemans, D. (2010). Inhibition of HIV-1 replication by a bis-thiadiazolbenzene-1,2-diamine that chelates zinc ions from retroviral nucleocapsid zinc fingers. *Antimicrob. Agents Chemother.* *54*, 1461–1468.
- Panté, N., and Kann, M. (2002). Nuclear pore complex is able to transport macromolecules with diameters of about 39 nm. *Mol. Biol. Cell* *13*, 425–434.
- Pawlak, E.N., and Dikeakos, J.D. (2015). HIV-1 Nef: a master manipulator of the membrane trafficking machinery mediating immune evasion. *Biochim. Biophys. Acta BBA - Gen. Subj.* *1850*, 733–741.
- Paxton, W., Connor, R.I., and Landau, N.R. (1993). Incorporation of Vpr into human immunodeficiency virus type 1 virions: requirement for the p6 region of gag and mutational analysis. *J. Virol.* *67*, 7229–7237.
- Peliska, J.A., Balasubramanian, S., Giedroc, D.P., and Benkovic, S.J. (1994). Recombinant HIV-1 nucleocapsid protein accelerates HIV-1 reverse transcriptase catalyzed DNA strand transfer reactions and modulates RNase H activity. *Biochemistry (Mosc.)* *33*, 13817–13823.
- Perez, E.E., Wang, J., Miller, J.C., Jouvenot, Y., Kim, K.A., Liu, O., Wang, N., Lee, G., Bartsevich, V.V., Lee, Y.-L., et al. (2008). Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. *Nat. Biotechnol.* *26*, 808–816.
- Perez-Caballero, D., Hatziioannou, T., Zhang, F., Cowan, S., and Bieniasz, P.D. (2005). Restriction of human immunodeficiency virus type 1 by TRIM-CypA occurs with rapid kinetics and independently of cytoplasmic bodies, ubiquitin, and proteasome activity. *J. Virol.* *79*, 15567–15572.
- Pettit, S.C., Sheng, N., Tritch, R., Erickson-Viitanen, S., and Swanstrom, R. (1998). The regulation of sequential processing of HIV-1 Gag by the viral protease. In *Aspartic Proteinases*, (Springer), pp. 15–25.
- Pettit, S.C., Everitt, L.E., Choudhury, S., Dunn, B.M., and Kaplan, A.H. (2004). Initial Cleavage of the Human Immunodeficiency Virus Type 1 GagPol Precursor by Its Activated Protease Occurs by an Intramolecular Mechanism. *J. Virol.* *78*, 8477–8485.
- Pettit, S.C., Clemente, J.C., Jeung, J.A., Dunn, B.M., and Kaplan, A.H. (2005). Ordered Processing of the Human Immunodeficiency Virus Type 1 GagPol Precursor Is Influenced by the Context of the Embedded Viral Protease. *J. Virol.* *79*, 10601–10607.
- Piller, S.C. (2001). Effect of extracellular HIV-1 Vpr protein in vitro. *J. Neurovirol.* *7*, 183–185.
- Piller, S.C., Dubay, J.W., Derdeyn, C.A., and Hunter, E. (2000). Mutational Analysis of Conserved Domains within the Cytoplasmic Tail of gp41 from Human Immunodeficiency Virus Type 1: Effects on Glycoprotein Incorporation and Infectivity. *J. Virol.* *74*, 11717–11723.
- Piscitelli, S.C., and Gallicano, K.D. (2001). Interactions among Drugs for HIV and Opportunistic Infections. *N. Engl. J. Med.* *344*, 984–996.
- Plantier, J.-C., Leoz, M., Dickerson, J.E., De Oliveira, F., Cordonnier, F., Lemée, V., Damond, F., Robertson, D.L., and Simon, F. (2009). A new human immunodeficiency virus derived from gorillas. *Nat. Med.* *15*, 871.
- Poljak, L., Batson, S.M., Ficheux, D., Roques, B.P., Darlix, J.-L., and Käs, E. (2003). Analysis of NCp7-dependent activation of HIV-1 cDNA integration and its conservation among retroviral nucleocapsid proteins. *J. Mol. Biol.* *329*, 411–421.
- Pollakis, G., Abebe, A., Kliphuis, A., Chalaby, M.I.M., Bakker, M., Mengistu, Y., Brouwer, M., Goudsmit, J., Schuitemaker, H., and Paxton, W.A. (2004). Phenotypic and Genotypic Comparisons of CCR5- and CXCR4-Tropic Human Immunodeficiency Virus Type 1 Biological Clones Isolated from Subtype C-Infected Individuals. *J. Virol.* *78*, 2841–2852.
- Pollard, V.W., and Malim, M.H. (1998). The HIV-1 Rev protein. *Annu. Rev. Microbiol.* *52*, 491–532.

- Poon, D.T., Li, G., and Aldovini, A. (1998). Nucleocapsid and matrix protein contributions to selective human immunodeficiency virus type 1 genomic RNA packaging. *J. Virol.* 72, 1983–1993.
- Popov, S., Popova, E., Inoue, M., and Göttlinger, H.G. (2008). Human immunodeficiency virus type 1 Gag engages the Bro1 domain of ALIX/AIP1 through the nucleocapsid. *J. Virol.* 82, 1389–1398.
- Popov, S., Popova, E., Inoue, M., and Göttlinger, H.G. (2009). Divergent Bro1 domains share the capacity to bind human immunodeficiency virus type 1 nucleocapsid and to enhance virus-like particle production. *J. Virol.* 83, 7185–7193.
- Prescher, J., Baumgärtel, V., Ivanchenko, S., Torrano, A.A., Bräuchle, C., Müller, B., and Lamb, D.C. (2015). Super-Resolution Imaging of ESCRT-Proteins at HIV-1 Assembly Sites. *PLoS Pathog.* 11.
- Purohit, P., Dupont, S., Stevenson, M., and Green, M.R. (2001). Sequence-specific interaction between HIV-1 matrix protein and viral genomic RNA revealed by in vitro genetic selection. *RNA* 7, 576–584.
- Pustowka, A., Dietz, J., Ferner, J., Baumann, M., Landersz, M., Königs, C., Schwalbe, H., and Dietrich, U. (2003). Identification of Peptide Ligands for Target RNA Structures Derived from the HIV-1 Packaging Signal ψ by Screening Phage-Displayed Peptide Libraries. *ChemBioChem* 4, 1093–1097.
- Ramalanjaona, N., de Rocquigny, H., Millet, A., Ficheux, D., Darlix, J.-L., and Mély, Y. (2007). Investigating the mechanism of the nucleocapsid protein chaperoning of the second strand transfer during HIV-1 DNA synthesis. *J. Mol. Biol.* 374, 1041–1053.
- Rasmussen, T.A., Tolstrup, M., Brinkmann, C.R., Olesen, R., Erikstrup, C., Solomon, A., Winckelmann, A., Palmer, S., Dinarello, C., Buzon, M., et al. (2014). Panobinostat, a histone deacetylase inhibitor, for latent-virus reactivation in HIV-infected patients on suppressive antiretroviral therapy: a phase 1/2, single group, clinical trial. *Lancet HIV* 1, e13–e21.
- Ratner, L., Haseltine, W., Patarca, R., Livak, K.J., Starcich, B., Josephs, S.F., Doran, E.R., Rafalski, J.A., Whitehorn, E.A., Baumeister, K., et al. (1985). Complete nucleotide sequence of the AIDS virus, HTLV-III. *Nature* 313, 277–284.
- Rausch, J.W., and Le Grice, S.F.J. (2004). “Binding, bending and bonding”: polypurine tract-primed initiation of plus-strand DNA synthesis in human immunodeficiency virus. *Int. J. Biochem. Cell Biol.* 36, 1752–1766.
- Reed, J.C., Molter, B., Geary, C.D., McNevin, J., McElrath, J., Giri, S., Klein, K.C., and Lingappa, J.R. (2012). HIV-1 Gag co-opts a cellular complex containing DDX6, a helicase that facilitates capsid assembly. *J. Cell Biol.* 198, 439–456.
- Reeves, J.D., and Doms, R.W. (2002). Human immunodeficiency virus type 2. *J. Gen. Virol.* 83, 1253–1265.
- Regoes, R.R., and Bonhoeffer, S. (2005). The HIV coreceptor switch: a population dynamical perspective. *Trends Microbiol.* 13, 269–277.
- Rein, A., Henderson, L.E., and Levin, J.G. (1998). Nucleic-acid-chaperone activity of retroviral nucleocapsid proteins: significance for viral replication. *Trends Biochem. Sci.* 23, 297–301.
- Rey, O., Canon, J., and Krogstad, P. (1996). HIV-1 Gag protein associates with F-actin present in microfilaments. *Virology* 220, 530–534.
- Rhee, S.S., and Marsh, J.W. (1994). Human immunodeficiency virus type 1 Nef-induced down-modulation of CD4 is due to rapid internalization and degradation of surface CD4. *J. Virol.* 68, 5156–5163.
- Ricci, E.P., Soto Rifo, R., Herbreteau, C.H., Decimo, D., and Ohlmann, T. (2008). Lentiviral RNAs can use different mechanisms for translation initiation. *Biochem. Soc. Trans.* 36, 690–693.
- Rice, null, and Turpin, null (1996). Virus-encoded Zinc Fingers as Targets for Antiviral

- Chemotherapy. *Rev. Med. Virol.* 6, 187–199.
- Rice, W.G., Supko, J.G., Malspeis, L., Buckheit, R.W., Clanton, D., Bu, M., Graham, L., Schaeffer, C.A., Turpin, J.A., Domagala, J., et al. (1995). Inhibitors of HIV nucleocapsid protein zinc fingers as candidates for the treatment of AIDS. *Science* 270, 1194–1197.
- Rice, W.G., Baker, D.C., Schaeffer, C.A., Graham, L., Bu, M., Terpening, S., Clanton, D., Schultz, R., Bader, J.P., Buckheit, R.W., et al. (1997a). Inhibition of multiple phases of human immunodeficiency virus type 1 replication by a dithiane compound that attacks the conserved zinc fingers of retroviral nucleocapsid proteins. *Antimicrob. Agents Chemother.* 41, 419–426.
- Rice, W.G., Turpin, J.A., Huang, M., Clanton, D., Buckheit, R.W., Covell, D.G., Wallqvist, A., McDonnell, N.B., DeGuzman, R.N., Summers, M.F., et al. (1997b). Azodicarbonamide inhibits HIV-1 replication by targeting the nucleocapsid protein. *Nat. Med.* 3, 341–345.
- Rist, M.J., and Marino, J.P. (2002). Mechanism of nucleocapsid protein catalyzed structural isomerization of the dimerization initiation site of HIV-1. *Biochemistry (Mosc.)* 41, 14762–14770.
- Ritchie, A.V., Ushiro-Lumb, I., Edemaga, D., Joshi, H.A., De Ruiter, A., Szumilin, E., Jendrulek, I., McGuire, M., Goel, N., Sharma, P.I., et al. (2014). SAMBA HIV semiquantitative test, a new point-of-care viral-load-monitoring assay for resource-limited settings. *J. Clin. Microbiol.* 52, 3377–3383.
- Rivière, L., Darlix, J.-L., and Cimarelli, A. (2010). Analysis of the Viral Elements Required in the Nuclear Import of HIV-1 DNA. *J. Virol.* 84, 729–739.
- de Rocquigny, H., Petitjean, P., Tanchou, V., Decimo, D., Drouot, L., Delaunay, T., Darlix, J.L., and Roques, B.P. (1997). The zinc fingers of HIV nucleocapsid protein NCp7 direct interactions with the viral regulatory protein Vpr. *J. Biol. Chem.* 272, 30753–30759.
- Romani, B., Engelbrecht, S., and Glashoff, R.H. (2010). Functions of Tat: the versatile protein of human immunodeficiency virus type 1. *J. Gen. Virol.* 91, 1–12.
- Rose, K.M., Marin, M., Kozak, S.L., and Kabat, D. (2004). The viral infectivity factor (Vif) of HIV-1 unveiled. *Trends Mol. Med.* 10, 291–297.
- Rowland-Jones, S.L. (2003). Timeline: AIDS pathogenesis: what have two decades of HIV research taught us? *Nat. Rev. Immunol.* 3, 343–348.
- Roy, B.B., Hu, J., Guo, X., Russell, R.S., Guo, F., Kleiman, L., and Liang, C. (2006). Association of RNA Helicase A with Human Immunodeficiency Virus Type 1 Particles. *J. Biol. Chem.* 281, 12625–12635.
- Rulli, S.J., Hibbert, C.S., Mirro, J., Pederson, T., Biswal, S., and Rein, A. (2007). Selective and Nonselective Packaging of Cellular RNAs in Retrovirus Particles. *J. Virol.* 81, 6623–6631.
- Russell, R.S., Roldan, A., Detorio, M., Hu, J., Wainberg, M.A., and Liang, C. (2003). Effects of a Single Amino Acid Substitution within the p2 Region of Human Immunodeficiency Virus Type 1 on Packaging of Spliced Viral RNA. *J. Virol.* 77, 12986–12995.
- Ryo, A., Tsurutani, N., Ohba, K., Kimura, R., Komano, J., Nishi, M., Soeda, H., Hattori, S., Perrem, K., Yamamoto, M., et al. (2008). SOCS1 is an inducible host factor during HIV-1 infection and regulates the intracellular trafficking and stability of HIV-1 Gag. *Proc. Natl. Acad. Sci. U. S. A.* 105, 294–299.
- Saad, J.S., Miller, J., Tai, J., Kim, A., Ghanam, R.H., and Summers, M.F. (2006). Structural basis for targeting HIV-1 Gag proteins to the plasma membrane for virus assembly. *Proc. Natl. Acad. Sci.* 103, 11364–11369.
- Saad, J.S., Loeliger, E., Luncsford, P., Liriano, M., Tai, J., Kim, A., Miller, J., Joshi, A., Freed, E.O., and Summers, M.F. (2007). Point mutations in the HIV-1 matrix protein turn off the myristyl switch. *J. Mol. Biol.* 366, 574–585.
- Sáez-Cirióñ, A., Bacchus, C., Hocqueloux, L., Avettand-Fenoel, V., Girault, I., Lecuroux, C., Potard, V., Versmisse, P., Melard, A., Prazuck, T., et al. (2013). Post-Treatment HIV-1 Controllers with a Long-Term Virological Remission after the Interruption of Early Initiated

- Antiretroviral Therapy ANRS VISCONTI Study. *PLOS Pathog* 9, e1003211.
- Salgado, G.F., Vogel, A., Marquant, R., Feller, S.E., Bouaziz, S., and Alves, I.D. (2009). The role of membranes in the organization of HIV-1 Gag p6 and Vpr: p6 shows high affinity for membrane bilayers which substantially increases the interaction between p6 and Vpr. *J. Med. Chem.* 52, 7157–7162.
- Salzwedel, K., West, J.T., and Hunter, E. (1999). A conserved tryptophan-rich motif in the membrane-proximal region of the human immunodeficiency virus type 1 gp41 ectodomain is important for Env-mediated fusion and virus infectivity. *J. Virol.* 73, 2469–2480.
- Saphire, A.C.S., Bobardt, M.D., Zhang, Z., David, G., and Gallay, P.A. (2001). Syndecans Serve as Attachment Receptors for Human Immunodeficiency Virus Type 1 on Macrophages. *J. Virol.* 75, 9187–9200.
- Sarafianos, S.G., Marchand, B., Das, K., Himmel, D.M., Parniak, M.A., Hughes, S.H., and Arnold, E. (2009). Structure and Function of HIV-1 Reverse Transcriptase: Molecular Mechanisms of Polymerization and Inhibition. *J. Mol. Biol.* 385, 693–713.
- Sartori, D.A., Miller, B., Biebach, U., and Farrell, N. (2000). Modulation of the chemical and biological properties of trans platinum complexes: monofunctional platinum complexes containing one nucleobase as potential antiviral chemotypes. *J. Biol. Inorg. Chem. JBIC Publ. Soc. Biol. Inorg. Chem.* 5, 575–583.
- Schiller, J., and Chackerian, B. (2014). Why HIV Virions Have Low Numbers of Envelope Spikes: Implications for Vaccine Development. *PLOS Pathog* 10, e1004254.
- Schröder, A.R.W., Shinn, P., Chen, H., Berry, C., Ecker, J.R., and Bushman, F. (2002). HIV-1 integration in the human genome favors active genes and local hotspots. *Cell* 110, 521–529.
- Schur, F.K.M., Obr, M., Hagen, W.J.H., Wan, W., Jakobi, A.J., Kirkpatrick, J.M., Sachse, C., Kräusslich, H.-G., and Briggs, J.A.G. (2016). An atomic model of HIV-1 capsid-SP1 reveals structures regulating assembly and maturation. *Science* 353, 506–508.
- Schwartz, O., Maréchal, V., Gall, S.L., Lemonnier, F., and Heard, J.-M. (1996). Endocytosis of major histocompatibility complex class I molecules is induced by the HIV-1 Nef protein. *Nat. Med.* 2, 338–342.
- von Schwedler, U.K., Kornbluth, R.S., and Trono, D. (1994). The nuclear localization signal of the matrix protein of human immunodeficiency virus type 1 allows the establishment of infection in macrophages and quiescent T lymphocytes. *Proc. Natl. Acad. Sci.* 91, 6992–6996.
- von Schwedler, U.K., Stray, K.M., Garrus, J.E., and Sundquist, W.I. (2003). Functional Surfaces of the Human Immunodeficiency Virus Type 1 Capsid Protein. *J. Virol.* 77, 5439–5450.
- Sever, N., Yang, T., Brown, M.S., Goldstein, J.L., and DeBose-Boyd, R.A. (2003). Accelerated degradation of HMG CoA reductase mediated by binding of insig-1 to its sterol-sensing domain. *Mol. Cell* 11, 25–33.
- Sharma, K.K., Przybilla, F., Restle, T., Godet, J., and Mély, Y. (2016). FRET-based assay to screen inhibitors of HIV-1 reverse transcriptase and nucleocapsid protein. *Nucleic Acids Res.* 44, e74.
- Sharmeen, L., McQuade, T., Heldsinger, A., Gogliotti, R., Domagala, J., and Gracheck, S. (2001). Inhibition of the early phase of HIV replication by an isothiazolone, PD 161374. *Antiviral Res.* 49, 101–114.
- Sheehy, A.M., Gaddis, N.C., Choi, J.D., and Malim, M.H. (2002). Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature* 418, 646–650.
- Shehu-Xhilaga, M., Kraeusslich, H.G., Pettit, S., Swanstrom, R., Lee, J.Y., Marshall, J.A., Crowe, S.M., and Mak, J. (2001). Proteolytic Processing of the P2/Nucleocapsid Cleavage Site Is Critical for Human Immunodeficiency Virus Type 1 RNA Dimer Maturation. *J. Virol.* 75, 9156–9164.
- Sherpa, C., Rausch, J.W., J. Le Grice, S.F., Hammarskjold, M.-L., and Rekosh, D. (2015). The

- HIV-1 Rev response element (RRE) adopts alternative conformations that promote different rates of virus replication. *Nucleic Acids Res.* gkv313.
- Shi, J., Zhou, J., Shah, V.B., Aiken, C., and Whitby, K. (2011). Small-molecule inhibition of human immunodeficiency virus type 1 infection by virus capsid destabilization. *J. Virol.* 85, 542–549.
- Shioda, T., Oka, S., Ida, S., Nokihara, K., Toriyoshi, H., Mori, S., Takebe, Y., Kimura, S., Shimada, K., and Nagai, Y. (1994). A naturally occurring single basic amino acid substitution in the V3 region of the human immunodeficiency virus type 1 env protein alters the cellular host range and antigenic structure of the virus. *J. Virol.* 68, 7689–7696.
- Shvadchak, V., Sanglier, S., Rocle, S., Villa, P., Haiech, J., Hibert, M., Van Dorsselaer, A., Mély, Y., and de Rocquigny, H. (2009). Identification by high throughput screening of small compounds inhibiting the nucleic acid destabilization activity of the HIV-1 nucleocapsid protein. *Biochimie* 91, 916–923.
- Simon, V., and Ho, D.D. (2003). HIV-1 dynamics in vivo: implications for therapy. *Nat. Rev. Microbiol.* 1, 181–190.
- Sinz, A. (2006). Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein-protein interactions. *Mass Spectrom. Rev.* 25, 663–682.
- Sinz, A. (2014). The advancement of chemical cross-linking and mass spectrometry for structural proteomics: from single proteins to protein interaction networks. *Expert Rev. Proteomics* 11, 733–743.
- Skripkin, E., Paillart, J.-C., Marquet, R., Ehresmann, B., and Ehresmann, C. (1994). Identification of the primary site of the human immunodeficiency virus type 1 RNA dimerization in vitro. *Proc. Natl. Acad. Sci.* 91, 4945–4949.
- Sloan, R.D., and Wainberg, M.A. (2011). The role of unintegrated DNA in HIV infection. *Retrovirology* 8, 52.
- Sloan, R.D., Kuhl, B.D., Mesplède, T., Münch, J., Donahue, D.A., and Wainberg, M.A. (2013). Productive Entry of HIV-1 during Cell-to-Cell Transmission via Dynamin-Dependent Endocytosis. *J. Virol.* 87, 8110–8123.
- Smyth, R.P., Davenport, M.P., and Mak, J. (2012). The origin of genetic diversity in HIV-1. *Virus Res.* 169, 415–429.
- Søgaard, O.S., Graversen, M.E., Leth, S., Olesen, R., Brinkmann, C.R., Nissen, S.K., Kjaer, A.S., Schleimann, M.H., Denton, P.W., Hey-Cunningham, W.J., et al. (2015). The Depsipeptide Romidepsin Reverses HIV-1 Latency In Vivo. *PLOS Pathog* 11, e1005142.
- Solbak, S.M.Ø., Reksten, T.R., Hahn, F., Wray, V., Henklein, P., Halskau, Ø., Schubert, U., and Fossen, T. (2013). HIV-1 p6 — a structured to flexible multifunctional membrane-interacting protein. *Biochim. Biophys. Acta BBA - Biomembr.* 1828, 816–823.
- Sosic, A., Frecentese, F., Perissutti, E., Sinigaglia, L., Santagada, V., Caliendo, G., Magli, E., Ciano, A., Zagotto, G., Parolin, C., et al. (2013). Design, synthesis and biological evaluation of TAR and cTAR binders as HIV-1 nucleocapsid inhibitors. *MedChemComm* 4, 1388–1393.
- Sosic, A., Sinigaglia, L., Cappellini, M., Carli, I., Parolin, C., Zagotto, G., Sabatino, G., Rovero, P., Fabris, D., and Gatto, B. (2016). Mechanisms of HIV-1 Nucleocapsid Protein Inhibition by Lysyl-Peptidyl-Anthraquinone Conjugates. *Bioconjug. Chem.* 27, 247–256.
- Srivastava, P., Schito, M., Fattah, R.J., Hara, T., Hartman, T., Buckheit, R.W., Turpin, J.A., Inman, J.K., and Appella, E. (2004). Optimization of unique, uncharged thioesters as inhibitors of HIV replication. *Bioorg. Med. Chem.* 12, 6437–6450.
- Starcich, B.R., Hahn, B.H., Shaw, G.M., McNeely, P.D., Modrow, S., Wolf, H., Parks, E.S., Parks, W.P., Josephs, S.F., and Gallo, R.C. (1986). Identification and characterization of conserved and variable regions in the envelope gene of HTLV-III/LAV, the retrovirus of AIDS. *Cell* 45, 637–648.
- Strack, B., Calistri, A., Craig, S., Popova, E., and Göttlinger, H.G. (2003). AIP1/ALIX Is a

- Binding Partner for HIV-1 p6 and EIAV p9 Functioning in Virus Budding. *Cell* 114, 689–699.
- Sükösd, Z., Andersen, E.S., Seemann, S.E., Jensen, M.K., Hansen, M., Gorodkin, J., and Kjems, J. (2015). Full-length RNA structure prediction of the HIV-1 genome reveals a conserved core domain. *Nucleic Acids Res.* gkv1039.
- Summa, V., Petrocchi, A., Bonelli, F., Crescenzi, B., Donghi, M., Ferrara, M., Fiore, F., Gardelli, C., Gonzalez Paz, O., Hazuda, D.J., et al. (2008). Discovery of raltegravir, a potent, selective orally bioavailable HIV-integrase inhibitor for the treatment of HIV-AIDS infection. *J. Med. Chem.* 51, 5843–5855.
- Sun, G., Li, H., Wu, X., Covarrubias, M., Scherer, L., Meinking, K., Luk, B., Chomchan, P., Alluin, J., Gombart, A.F., et al. (2012). Interplay between HIV-1 infection and host microRNAs. *Nucleic Acids Res.* 40, 2181–2196.
- Sundquist, W.I., and Krausslich, H.-G. (2012). HIV-1 Assembly, Budding, and Maturation. *Cold Spring Harb. Perspect. Med.* 2, a006924–a006924.
- Suzuki, Y., Chew, M.L., and Suzuki, Y. (2012). Role of host-encoded proteins in restriction of retroviral integration. *Front. Microbiol.* 3, 227.
- Swigut, T., Shohdy, N., and Skowronski, J. (2001). Mechanism for down-regulation of CD28 by Nef. *EMBO J.* 20, 1593–1604.
- Szilvay, A.M., Stern, B., Blichenberg, A., and Helland, D.E. (2000). Structural and functional similarities between HIV-1 reverse transcriptase and the Escherichia coli RNA polymerase β' subunit. *FEBS Lett.* 484, 43–47.
- Tahirov, T.H., Babayeva, N.D., Varzavand, K., Cooper, J.J., Sedore, S.C., and Price, D.H. (2010). Crystal structure of HIV-1 Tat complexed with human P-TEFb. *Nature* 465, 747–751.
- Tang, C., Ndassa, Y., and Summers, M.F. (2002). Structure of the N-terminal 283-residue fragment of the immature HIV-1 Gag polyprotein. *Nat. Struct. Mol. Biol.* 9, 537–543.
- Tang, C., Loeliger, E., Luncsford, P., Kinde, I., Beckett, D., and Summers, M.F. (2004). Entropic switch regulates myristate exposure in the HIV-1 matrix protein. *Proc. Natl. Acad. Sci. U. S. A.* 101, 517–522.
- Tang, C., Louis, J.M., Aniana, A., Suh, J.-Y., and Clore, G.M. (2008). Visualizing transient events in amino-terminal autoprocessing of HIV-1 protease. *Nature* 455, 693–696.
- Tebit, D.M., and Arts, E.J. (2011). Tracking a century of global expansion and evolution of HIV to drive understanding and to combat disease. *Lancet Infect. Dis.* 11, 45–56.
- Tedbury, P.R., Novikova, M., Ablan, S.D., and Freed, E.O. (2016). Biochemical evidence of a role for matrix trimerization in HIV-1 envelope glycoprotein incorporation. *Proc. Natl. Acad. Sci. U. S. A.* 113, E182-190.
- Thierry, S., Munir, S., Thierry, E., Subra, F., Leh, H., Zamborlini, A., Saenz, D., Levy, D.N., Lesbats, P., Saïb, A., et al. (2015). Integrase inhibitor reversal dynamics indicate unintegrated HIV-1 dna initiate de novo integration. *Retrovirology* 12, 24.
- Thierry, S., Thierry, E., Subra, F., Deprez, E., Leh, H., Bury-Moné, S., and Delelis, O. (2016). Opposite transcriptional regulation of integrated vs unintegrated HIV genomes by the NF- κ B pathway. *Sci. Rep.* 6.
- Thomas, J.A., and Gorelick, R.J. (2008). Nucleocapsid protein function in early infection processes. *Virus Res.* 134, 39–63.
- Thomas, J.A., Gagliardi, T.D., Alvord, W.G., Lubomirski, M., Bosche, W.J., and Gorelick, R.J. (2006). Human immunodeficiency virus type 1 nucleocapsid zinc-finger mutations cause defects in reverse transcription and integration. *Virology* 353, 41–51.
- Tisné, C., Roques, B.P., and Dardel, F. (2001). Heteronuclear NMR studies of the interaction of tRNA³Lys with HIV-1 nucleocapsid protein1. *J. Mol. Biol.* 306, 443–454.
- Traggiai, E., Chicha, L., Mazzucchelli, L., Bronz, L., Piffaretti, J.-C., Lanzavecchia, A., and Manz, M.G. (2004). Development of a Human Adaptive Immune System in Cord Blood Cell-Transplanted Mice. *Science* 304, 104–107.

- Tummino, P.J., Scholten, J.D., Harvey, P.J., Holler, T.P., Maloney, L., Gogliotti, R., Domagala, J., and Hupe, D. (1996). The in vitro ejection of zinc from human immunodeficiency virus (HIV) type 1 nucleocapsid protein by disulfide benzamides with cellular anti-HIV activity. *Proc. Natl. Acad. Sci. U. S. A.* *93*, 969–973.
- Turner, K.B., Kohlway, A.S., Hagan, N.A., and Fabris, D. (2009). Noncovalent probes for the investigation of structure and dynamics of protein-nucleic acid assemblies: the case of NC-mediated dimerization of genomic RNA in HIV-1. *Biopolymers* *91*, 283–296.
- Turpin, J.A., Song, Y., Inman, J.K., Huang, M., Wallqvist, A., Maynard, A., Covell, D.G., Rice, W.G., and Appella, E. (1999). Synthesis and biological properties of novel pyridinioalkanoyl thioesters (PATE) as anti-HIV-1 agents that target the viral nucleocapsid protein zinc fingers. *J. Med. Chem.* *42*, 67–86.
- Tyagi, M., Pearson, R.J., and Karn, J. (2010). Establishment of HIV latency in primary CD4+ cells is due to epigenetic transcriptional silencing and P-TEFb restriction. *J. Virol.* *84*, 6425–6437.
- Usami, Y., Popov, S., and Göttlinger, H.G. (2007). Potent Rescue of Human Immunodeficiency Virus Type 1 Late Domain Mutants by ALIX/AIP1 Depends on Its CHMP4 Binding Site. *J. Virol.* *81*, 6614–6622.
- Van Damme, N., Goff, D., Katsura, C., Jorgenson, R.L., Mitchell, R., Johnson, M.C., Stephens, E.B., and Guatelli, J. (2008). The interferon-induced protein BST-2 restricts HIV-1 release and is downregulated from the cell surface by the viral Vpu protein. *Cell Host Microbe* *3*, 245–252.
- Van Engelenburg, S.B., Shtengel, G., Sengupta, P., Waki, K., Jarnik, M., Ablan, S.D., Freed, E.O., Hess, H.F., and Lippincott-Schwartz, J. (2014). Distribution of ESCRT machinery at HIV assembly sites reveals virus scaffolding of ESCRT subunits. *Science* *343*, 653–656.
- Velappan, N., Fisher, H.E., Pesavento, E., Chasteen, L., D'Angelo, S., Kiss, C., Longmire, M., Pavlik, P., and Bradbury, A.R.M. (2010). A comprehensive analysis of filamentous phage display vectors for cytoplasmic proteins: an analysis with different fluorescent proteins. *Nucleic Acids Res.* *38*, e22.
- Vercruyse, T., Basta, B., Dehaen, W., Humbert, N., Balzarini, J., Debaene, F., Sanglier-Cianferani, S., Pannecouque, C., Mely, Y., and Daelemans, D. (2012). A phenylthiadiazolylidene-amine derivative ejects zinc from retroviral nucleocapsid zinc fingers and inactivates HIV virions. *Retrovirology* *9*, 95.
- VerPlank, L., Bouamr, F., LaGrassa, T.J., Agresta, B., Kikonyogo, A., Leis, J., and Carter, C.A. (2001). Tsg101, a homologue of ubiquitin-conjugating (E2) enzymes, binds the L domain in HIV type 1 Pr55Gag. *Proc. Natl. Acad. Sci.* *98*, 7724–7729.
- Vo, M.-N., Barany, G., Rouzina, I., and Musier-Forsyth, K. (2006). Mechanistic Studies of Mini-TAR RNA/DNA Annealing in the Absence and Presence of HIV-1 Nucleocapsid Protein. *J. Mol. Biol.* *363*, 244–261.
- Vollenweider, F., Benjannet, S., Decroly, E., Savaria, D., Lazure, C., Thomas, G., Chrétien, M., and Seidah, N.G. (1996). Comparative cellular processing of the human immunodeficiency virus (HIV-1) envelope glycoprotein gp160 by the mammalian subtilisin/kexin-like convertases. *Biochem. J.* *314* (Pt 2), 521–532.
- Votteler, J., and Sundquist, W.I. (2013). Virus Budding and the ESCRT Pathway. *Cell Host Microbe* *14*, 232–241.
- Vuilleumier, C., Bombarda, E., Morellet, N., Gérard, D., Roques, B.P., and Mély, Y. (1999). Nucleic acid sequence discrimination by the HIV-1 nucleocapsid protein NCp7: a fluorescence study. *Biochemistry (Mosc.)* *38*, 16816–16825.
- Wagner, J.M., Zadrozny, K.K., Chrustowicz, J., Purdy, M.D., Yeager, M., Ganser-Pornillos, B.K., and Pornillos, O. (2016). Crystal structure of an HIV assembly and maturation switch. *eLife* *5*, e17063.
- Wallace, G.S., Cheng-Mayer, C., Schito, M.L., Fletcher, P., Miller Jenkins, L.M., Hayashi, R.,

- Neurath, A.R., Appella, E., and Shattock, R.J. (2009). Human immunodeficiency virus type 1 nucleocapsid inhibitors impede trans infection in cellular and explant models and protect nonhuman primates from infection. *J. Virol.* *83*, 9175–9182.
- Wang, D., Lu, W., and Li, F. (2015). Pharmacological intervention of HIV-1 maturation. *Acta Pharm. Sin. B* *5*, 493–499.
- Warren, K., Wei, T., Li, D., Qin, F., Warrilow, D., Lin, M.-H., Sivakumaran, H., Apolloni, A., Abbott, C.M., Jones, A., et al. (2012). Eukaryotic elongation factor 1 complex subunits are critical HIV-1 reverse transcription cofactors. *Proc. Natl. Acad. Sci. U. S. A.* *109*, 9587–9592.
- Warui, D.M., and Baranger, A.M. (2009). Identification of specific small molecule ligands for stem loop 3 ribonucleic acid of the packaging signal Psi of human immunodeficiency virus-1. *J. Med. Chem.* *52*, 5462–5473.
- Warui, D.M., and Baranger, A.M. (2012). Identification of small molecule inhibitors of the HIV-1 nucleocapsid-stem-loop 3 RNA complex. *J. Med. Chem.* *55*, 4132–4141.
- Watts, J.M., Dang, K.K., Gorelick, R.J., Leonard, C.W., Bess, J.W., Swanstrom, R., Burch, C.L., and Weeks, K.M. (2009). Architecture and Secondary Structure of an Entire HIV-1 RNA Genome. *Nature* *460*, 711–716.
- Weissenhorn, W., Dessen, A., Harrison, S.C., Skehel, J.J., and Wiley, D.C. (1997). Atomic structure of the ectodomain from HIV-1 gp41. *Nature* *387*, 426–430.
- Welker, R., Hohenberg, H., Tessmer, U., Huckhagel, C., and Kräusslich, H.G. (2000). Biochemical and structural analysis of isolated mature cores of human immunodeficiency virus type 1. *J. Virol.* *74*, 1168–1177.
- Westendorp, M.O., Frank, R., Ochsenbauer, C., Stricker, K., Dhein, J., Walczak, H., Debating, K.-M., and Krammer, P.H. (1995). Sensitization of T cells to CD95-mediated apoptosis by HIV-1 Tat and gp120. *Nature* *375*, 497–500.
- Wilk, T., Gowen, B., and Fuller, S.D. (1999). Actin associates with the nucleocapsid domain of the human immunodeficiency virus Gag polyprotein. *J. Virol.* *73*, 1931–1940.
- Wilk, T., Gross, I., Gowen, B.E., Rutten, T., de Haas, F., Welker, R., Kräusslich, H.-G., Boulanger, P., and Fuller, S.D. (2001). Organization of Immature Human Immunodeficiency Virus Type 1. *J. Virol.* *75*, 759–771.
- Willey, R.L., Rutledge, R.A., Dias, S., Folks, T., Theodore, T., Buckler, C.E., and Martin, M.A. (1986). Identification of conserved and divergent domains within the envelope gene of the acquired immunodeficiency syndrome retrovirus. *Proc. Natl. Acad. Sci. U. S. A.* *83*, 5038–5042.
- Willey, R.L., Maldarelli, F., Martin, M.A., and Strebel, K. (1992a). Human immunodeficiency virus type 1 Vpu protein regulates the formation of intracellular gp160-CD4 complexes. *J. Virol.* *66*, 226–234.
- Willey, R.L., Maldarelli, F., Martin, M.A., and Strebel, K. (1992b). Human immunodeficiency virus type 1 Vpu protein induces rapid degradation of CD4. *J. Virol.* *66*, 7193–7200.
- Williamson, J.R. (2015). Really exasperating viral protein from HIV. *eLife* *4*, e05169.
- Wlodawer, A., Miller, M., Jaskólski, M., Sathyanarayana, B.K., Baldwin, E., Weber, I.T., Selk, L.M., Clawson, L., Schneider, J., and Kent, S.B. (1989). Conserved folding in retroviral proteases: crystal structure of a synthetic HIV-1 protease. *Science* *245*, 616–621.
- Worobey, M., Gemmel, M., Teuwen, D.E., Haselkorn, T., Kunstman, K., Bunce, M., Muyembe, J.-J., Kabongo, J.-M.M., Kalengayi, R.M., Van Marck, E., et al. (2008). Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960. *Nature* *455*, 661–664.
- Wu, H., Mitra, M., McCauley, M.J., Thomas, J.A., Rouzina, I., Musier-Forsyth, K., Williams, M.C., and Gorelick, R.J. (2013). Aromatic residue mutations reveal direct correlation between HIV-1 nucleocapsid protein's nucleic acid chaperone activity and retroviral replication. *Virus Res.* *171*, 263–277.
- Wyatt, R., Kwong, P.D., Desjardins, E., Sweet, R.W., Robinson, J., Hendrickson, W.A., and

- Sodroski, J.G. (1998). The antigenic structure of the HIV gp120 envelope glycoprotein. *Nature* 393, 705–711.
- Wyss, S., Berlioz-Torrent, C., Boge, M., Blot, G., Höning, S., Benarous, R., and Thali, M. (2001). The highly conserved C-terminal dileucine motif in the cytosolic domain of the human immunodeficiency virus type 1 envelope glycoprotein is critical for its association with the AP-1 clathrin adaptor [correction of adapter]. *J. Virol.* 75, 2982–2992.
- Xing, L., Liang, C., and Kleiman, L. (2011). Coordinate roles of Gag and RNA helicase A in promoting the annealing of formula to HIV-1 RNA. *J. Virol.* 85, 1847–1860.
- Xu, H., Franks, T., Gibson, G., Huber, K., Rahm, N., De Castillia, C.S., Luban, J., Aiken, C., Watkins, S., Sluis-Cremer, N., et al. (2013). Evidence for biphasic uncoating during HIV-1 infection from a novel imaging assay. *Retrovirology* 10, 70.
- Yamashita, M., and Emerman, M. (2004). Capsid is a dominant determinant of retrovirus infectivity in nondividing cells. *J. Virol.* 78, 5670–5678.
- Yamashita, M., Perez, O., Hope, T.J., and Emerman, M. (2007). Evidence for Direct Involvement of the Capsid Protein in HIV Infection of Nondividing Cells. *PLOS Pathog* 3, e156.
- Yan, N., Regalado-Magdos, A.D., Stiggelbout, B., Lee-Kirsch, M.A., and Lieberman, J. (2010). The cytosolic exonuclease TREX1 inhibits the innate immune response to human immunodeficiency virus type 1. *Nat. Immunol.* 11, 1005–1013.
- Yu, K.L., Lee, S.H., Lee, E.S., and You, J.C. (2016). HIV-1 nucleocapsid protein localizes efficiently to the nucleus and nucleolus. *Virology* 492, 204–212.
- Zarudnaya, M.I., Potyahaylo, A.L., Kolomiets, I.M., and Hovorun, D.M. (2013). Structural model of the complete poly(A) region of HIV-1 pre-mRNA. *J. Biomol. Struct. Dyn.* 31, 1044–1056.
- Zennou, V., Perez-Caballero, D., Göttlinger, H., and Bieniasz, P.D. (2004). APOBEC3G Incorporation into Human Immunodeficiency Virus Type 1 Particles. *J. Virol.* 78, 12058–12061.
- Zhang, J., and Crumpacker, C.S. (2002). Human Immunodeficiency Virus Type 1 Nucleocapsid Protein Nuclear Localization Mediates Early Viral mRNA Expression. *J. Virol.* 76, 10444–10454.
- Zhang, X., and Studier, F.W. (1997). Mechanism of inhibition of bacteriophage T7 RNA polymerase by T7 lysozyme 1. *J. Mol. Biol.* 269, 10–27.
- Zhang, H., Yang, B., Pomerantz, R.J., Zhang, C., Arunachalam, S.C., and Gao, L. (2003). The cytidine deaminase CEM15 induces hypermutation in newly synthesized HIV-1 DNA. *Nature* 424, 94–98.
- Zhao, G., Perilla, J.R., Yufenyuy, E.L., Meng, X., Chen, B., Ning, J., Ahn, J., Gronenborn, A.M., Schulten, K., Aiken, C., et al. (2013). Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics. *Nature* 497, 643–646.
- Zhou, W., Parent, L.J., Wills, J.W., and Resh, M.D. (1994). Identification of a membrane-binding domain within the amino-terminal region of human immunodeficiency virus type 1 Gag protein which interacts with acidic phospholipids. *J. Virol.* 68, 2556–2569.
- Zhou, Y., Rong, L., Lu, J., Pan, Q., and Liang, C. (2008). Insulin-like growth factor II mRNA binding protein 1 associates with Gag protein of human immunodeficiency virus type 1, and its overexpression affects virus assembly. *J. Virol.* 82, 5683–5692.
- Zhu, H., Jian, H., and Zhao, L.-J. (2004). Identification of the 15FRFG domain in HIV-1 Gag p6 essential for Vpr packaging into the virion. *Retrovirology* 1, 26.
- Zhu, P., Chertova, E., Bess, J., Lifson, J.D., Arthur, L.O., Liu, J., Taylor, K.A., and Roux, K.H. (2003). Electron tomography analysis of envelope glycoprotein trimers on HIV and simian immunodeficiency virus virions. *Proc. Natl. Acad. Sci. U. S. A.* 100, 15812–15817.
- Zhu, P., Liu, J., Bess, J., Chertova, E., Lifson, J.D., Grisé, H., Ofek, G.A., Taylor, K.A., and

Roux, K.H. (2006). Distribution and three-dimensional structure of AIDS virus envelope spikes. *Nature* 441, 847–852.

Zhu, Y., Pe'ery, T., Peng, J., Ramanathan, Y., Marshall, N., Marshall, T., Amendt, B., Mathews, M.B., and Price, D.H. (1997). Transcription elongation factor P-TEFb is required for HIV-1 tat transactivation in vitro. *Genes Dev.* 11, 2622–2632.

Zufferey, R., Dull, T., Mandel, R.J., Bukovsky, A., Quiroz, D., Naldini, L., and Trono, D. (1998). Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery. *J. Virol.* 72, 9873–9880.

ANRS - Santé publique, Sciences sociales - Actualités - L'essai ANRS Ipergay montre une diminution de 86% du risque d'infection.

Manuel PIRES

Caractérisation cellulaire de nouveaux traitements anti-VIH-1 ciblant la NucléoCapside et observation d'une forme compacte de Gag in cellulo.

Résumé

Les thérapies actuelles ciblant les étapes clés du cycle de réplication du VIH-1 sont puissantes et sélectives, mais des échecs thérapeutiques sont apparus en raison de l'émergence de résistance aux médicaments. Par conséquent, il existe un besoin urgent de développer de nouveaux médicaments qui utilisent des stratégies thérapeutiques alternatives. Cibler la protéine de nucléocapside (NCp7) est une stratégie prometteuse car NCp7 joue un rôle clé dans la réplication du VIH et sa séquence est hautement conservée dans les différentes souches du VIH. L'objectif du consortium européen **THINPAD** est de sélectionner des inhibiteurs spécifiques de la NCp7. Pour cribler les molécules, nous avons établi un test d'infectivité *in cellulo* utilisant la capacité des molécules sélectionnées à inhiber l'infection de cellules HeLa par des pseudo-particules non-réplicatives mimant la phase précoce du cycle d'infection virale du VIH-1. Afin de comprendre le mécanisme d'action des meilleures molécules sélectionnées, nous avons mis en place un test d'addition des molécules à différents temps post-infection et un test d'assemblage basé sur la technique de FRET-FLIM.

La polyprotéine Gag est un précurseur nécessaire et suffisant pour la formation de particules virales mais le mécanisme de sélection de seulement deux copies d'ARN spécifique par virion est encore mal compris. Dans Gag, il y a deux sites d'interaction pour l'ARN. L'un de ces sites est le domaine de nucléocapside (NC), qui contient deux doigts de zinc capables de recruter l'ARN génomique viral. Le second site est constitué par les résidus basiques du domaine Matrice (MA). Ces deux sites peuvent interagir avec l'ARN simultanément et stabiliser la forme compacte de Gag. Pour démontrer *in cellulo* ce repliement de Gag, nous avons utilisé un système de split-Green Fluorescent Protein (GFP) et combiné différentes techniques de microscopie et de biochimie. Afin de compléter cette étude, le repliement sous forme compacte de différents mutants de Gag qui ont perdu totalement ou partiellement leurs capacités d'interaction avec l'ARN a également été étudié.

Mots clés : VIH-1, NCp7, inhibiteurs, Gag, forme compacte, split-GFP

Abstract

Current therapies targeting key steps of the HIV replication cycle are potent and selective, but clinical failures appeared due to the emergence of drug resistance. Hence, there is an urgent need for novel drugs and alternative therapeutic strategies. Targeting the HIV NucleoCapsid protein (NCp7) is a promising strategy since NCp7 plays key roles in HIV replication and its sequence is highly conserved across HIV strains. The European consortium **THINPAD** aims at selecting specific NCp7 inhibitors. To screen molecules, we established an *in-cellulo* infectivity-assay based on the ability of the selected hits to inhibit the infection of HeLa cells by non-replicative HIV-1 pseudoparticles mimicking the early phase of the virus infection cycle. In order to decipher the mechanisms of action of the best hits, we implemented a time of addition assay and a FRET-FLIM based assembly assay.

The HIV-1 Gag polyprotein precursor is necessary and sufficient for the formation and the release of virus-like-particles but the mechanism of selection of only two copies of specific RNA per virion need to be further understood. In Gag, there are two interaction sites for RNA, namely the NucleoCapsid (NC) domain, which contains two zinc fingers able to recruit the dimeric virus genome, and the basic residues of the Matrix domain (MA). It was previously shown *in vitro* that these two sites can interact with RNA in order to stabilize a compact form of Gag. To demonstrate *in cellulo* this compact form of Gag, we used a split-Green Fluorescent Protein (GFP) system together with different microscopy and biochemical techniques. To complete this study, we monitored the folding of Gag mutants which have lost their ability to interact with RNA.

Key words: HIV-1, NCp7, inhibitors, Gag, compact shape, split-GFP