



HAL
open science

Rôle des tiges-boucles TAR et cTAR sur le premier transfert de brin du VIH-1

Sandra da Silva Amaral

► **To cite this version:**

Sandra da Silva Amaral. Rôle des tiges-boucles TAR et cTAR sur le premier transfert de brin du VIH-1. Biologie moléculaire. Université Paris-Saclay, 2023. Français. NNT : 2023UPASL030 . tel-04129753

HAL Id: tel-04129753

<https://theses.hal.science/tel-04129753>

Submitted on 15 Jun 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Rôle des tiges-boucles TAR et cTAR sur le premier transfert de brin du VIH-1

Role of TAR and cTAR stem-loops on HIV-1 first strand transfer

Thèse de doctorat de l'université Paris-Saclay

École doctorale n° 577 – Structure et Dynamique des Systèmes Vivants (SDSV)
Spécialité de doctorat : Biologie moléculaire et cellulaire
Graduate School : Live Sciences and Health. Référent : ENS Paris Saclay

Thèse préparée au LBPA (Université Paris-Saclay, ENS Paris-Saclay, CNRS),
sous la direction de **Philippe Fossé**, directeur de recherche CNRS

Thèse soutenue à Paris-Saclay, le 30 mars 2023, par

Sandra DA SILVA AMARAL

Composition du Jury

Membres du jury avec voix délibérative

Carine TISNÉ

Directrice de recherche

CNRS, Université Paris-Cité

Présidente

Jean-Christophe PAILLART

Directeur de recherche

CNRS, Université de Strasbourg

Rapporteur et examinateur

Hugues DE ROCQUIGNY

Chargé de recherche, HDR

INSERM, Université de Tours

Rapporteur et examinateur

Nelly MORELLET

Ingénieure de recherche

CNRS, Université Paris-Saclay

Examinatrice

Remerciements

J'aimerais exprimer ma gratitude à toutes les personnes qui ont contribué à la réalisation de mes travaux de thèse ainsi qu'à rendre cette expérience plus qu'agréable.

Tout d'abord, je tiens à remercier mon directeur de thèse, le Dr Philippe Fossé, pour la confiance continue qu'il a su m'accorder tout au long de ces années de thèse. Mais également pour ses enseignements, sa disponibilité, sa bienveillance et enfin son soutien.

Merci au Dr Éric Deprez de m'avoir accueillie dans son laboratoire ainsi que pour nos échanges riches en leçons et conseils.

Je suis également grandement reconnaissante envers le Dr Olivier Delelis qui m'a été d'un grand soutien scientifique. Ses éclaircissements et mises en lumière ont été aussi stimulants qu'encourageants pour ce travail de thèse.

Un remerciement particulier au Dr Frédéric Subra qui a su être patient dans la transmission de son savoir en matière de pratiques de culture cellulaire et de virologie. Ta bonne humeur a pour moi été aussi importante que tes enseignements.

Je tiens aussi à remercier le Dr Marylène Mougel et le Dr Bruno Sargueil d'avoir fait partie de mon comité de suivi de thèse. Les échanges d'un regard extérieur ont apporté des questionnements pertinents.

Je suis également reconnaissante envers les Dr Carine Tisné, Nelly Morellet, Hugues de Rocquigny et Jean-Christophe Paillart d'avoir accepté d'être membres de mon jury de thèse. Je tiens d'ailleurs à remercier mes rapporteurs Hugues de Rocquigny et Jean-Christophe Paillart pour leurs remarques constructives améliorant ainsi mon manuscrit.

Un chaleureux merci au Dr Olivier Mauffret, responsable de l'équipe, pour nos quelques intéressantes discussions ainsi qu'aux membres de l'équipe.

Un remerciement tout particulier au Dr Brahim Heddi dont sa présence, son soutien, son aide quotidienne et ses encouragements ont été indispensables tant pour mes travaux de thèse que pour moi-même d'un point de vue personnel.

Aussi, un grand et tendre merci à Maryline Havard pour sa gentillesse, sa fraîcheur et son aide. Ainsi qu'au Dr Clémence Richetta pour ses conseils scientifiques et nos nombreux sympathiques échanges.

Merci à tous les membres du laboratoire sans qui ces années de thèse n'auraient pas la même saveur et plus spécifiquement au Dr Yann Ferrandez et au Dr Marc Mirande qui ont toujours su répondre présent pour m'aider.

Il est important pour moi de remercier les étudiants rencontrés avec qui une amitié est née ainsi qu'à Samira Zouhir, Stéphanie Lejalle, et à l'équipe de la start-up Kimialys qui ont rendu le quotidien de ces années moins stressant et plus plaisant. Merci à Vincent Léguillier, Xavier Télot, Déborah Lécuyer, Nassima Teibi, Candice Sartre, Pavlina Dubois, Constanza Avendaño Avila et Nhat Quang Tu. Ainsi qu'aux étudiants qui ont contribué aux travaux de la thèse, Mehdi Ben Sassi et Imène Keraim.

Enfin, j'aimerais remercier mes parents et mes sœurs qui ont été d'un important appui depuis le début de mes études et les principaux moteurs de ma réussite scolaire. Ainsi que mon mari pour sa confiance et son soutien continu lors de cette aventure.

TABLE DES MATIERES

ABRÉVIATIONS.....	1
INTRODUCTION.....	2
1. VIH : AGENT PATHOGENE RESPONSABLE DU SIDA.....	2
1.1 LES RETROVIRUS ET LEUR CLASSIFICATION.....	2
1.2 DECOUVERTE SCIENTIFIQUE DU VIH.....	4
1.3 ORIGINE ET VARIABILITE GENETIQUE.....	5
1.4 SITUATION EPIDEMIOLOGIQUE.....	8
1.5 PATHOGENESE DU VIH.....	10
2. VIH : ANATOMIE DU VIRUS, GENOME ET CYCLE DE VIE.....	12
2.1 STRUCTURE DE LA PARTICULE VIRALE.....	12
2.2 ORGANISATION GENETIQUE.....	13
2.3 CYCLE DE REPLICATION.....	16
2.3.1 Entrée du virus.....	16
2.3.2 Transcription inverse de l'ARN et décapsidation.....	18
2.3.3 Intégration du provirus.....	24
2.3.4 Synthèse, maturation et export des ARN viraux.....	27
2.3.5 Synthèse des protéines virales.....	29
2.3.6 Encapsidation de l'ARN génomique dans la particule virale.....	43
2.3.7 Assemblage, bourgeonnement et maturation de la particule virale.....	45
3. TRANSCRIPTION INVERSE.....	47
3.1 FIXATION DE L'AMORCE ET FORMATION DE L'HYBRIDE ARNG-ARN ^{T_{Lys3}}	48
3.2 INITIATION DE LA TRANSCRIPTION INVERSE.....	52
3.3 SYNTHESE DE L'ADN STRONG-STOP.....	54
3.4 LE PREMIER TRANSFERT DE BRIN.....	56
3.5 SYNTHESE DU BRIN (-) D'ADN VIRAL.....	56
3.6 INITIATION DE LA SYNTHESE DU BRIN (+) D'ADN VIRAL.....	56
3.7 LE SECOND TRANSFERT DE BRIN.....	58
3.7.1 Acides nucléiques intervenant dans le second transfert de brin.....	58
3.7.2 Mécanismes intervenant dans le second transfert de brin.....	59
3.8 SYNTHESE COMPLETE DU BRIN (+) ET DES DEUX LTR.....	60
3.9 RECOMBINAISONS CHEZ LE VIH-1.....	62
3.9.1 La recombinaison dépend du transfert de brin interne.....	62
3.9.2 Modèles de recombinaison par transfert de brin interne.....	64
3.9.3 Mécanismes du transfert de brin interne.....	65
3.10 ROLE DE LA RNA HELICASE A DANS LA TRANSCRIPTION INVERSE.....	67
3.11 INHIBITION DE LA TRANSCRIPTION INVERSE PAR LES AGENTS PHARMACOLOGIQUES.....	68
3.11.1 Inhibiteurs ciblant la RT.....	68
3.11.2 Inhibiteurs ciblant la NC.....	69
3.11.3 Inhibiteurs ciblant les acides nucléiques viraux.....	70
4. LE PREMIER TRANSFERT DE BRIN.....	71
4.1 STRUCTURES ET INTERACTIONS INTERVENANT DANS LE PREMIER TRANSFERT DE BRIN.....	71
4.1.1 Structures ARN.....	71
4.1.2 Structures de l'ADNss.....	74
4.1.3 Interactions ADN-ARN régissant le premier transfert de brin.....	75
4.1.4 Rôle de l'activité RNase H de la RT dans le premier transfert de brin.....	77
4.1.5 Rôle de la NC dans le premier transfert de brin.....	77
OBJECTIF.....	81

MATÉRIEL ET MÉTHODES	84
1. CONSTRUCTION DES PLASMIDES ET VECTEURS	84
1.1 OLIGONUCLEOTIDES	84
1.2 PLASMIDES	84
1.3 CONSTRUCTION DES PLASMIDES pLF1 ET PEP2 MUTES.....	85
1.4 CONSTRUCTION DES VECTEURS pLeVD5 ET pLeVD3	86
2. PRODUCTION DES PARTICULES VIRALES	87
2.1 TRANSFECTION DES CELLULES HEK293T	87
2.2 QUANTIFICATION DE LA P24	87
2.3 QUANTIFICATION DES ARN VIRAUX.....	88
3. ANALYSE DE L'EXPRESSION VIRALE	89
3.1 TRANSDUCTION DES CELLULES MT4.....	89
3.2 MESURE DE L'EXPRESSION DE LA GFP	90
3.3 QUANTIFICATION PAR QPCR DES PRODUITS DE LA TRANSCRIPTION INVERSE	90
RÉSULTATS ET DISCUSSION	93
1. DESCRIPTION DU SYSTEME EXPERIMENTAL.....	93
2. CONSTRUCTION DES PLASMIDES ET VECTEURS	94
2.1 CONSTRUCTION DES PLASMIDES pLF1 ET PEP2 MUTES	95
2.2 CONSTRUCTION DES VECTEURS MUTES pLeVD5 ET pLeVD3.....	98
3. BASES RATIONNELLES DE LA CONCEPTION DES MUTANTS	102
4. ANALYSE DES EFFETS PRODUITS PAR LES MUTATIONS DANS LA SEQUENCE 5' TAR.....	103
4.1 STRUCTURES SECONDAIRES DES MUTANTS 5'TAR ET cTAR.....	103
4.2 RESULTATS DES TRANSFECTIONS PAR LES MUTANTS 5'	105
4.3 RESULTATS DES TRANSDUCTIONS PAR LES MUTANTS 5'	106
4.4 RESULTATS DE L'ANALYSE DE L'ENCAPSIDATION DE L'ARNg DES MUTANTS 5'.....	112
4.5 QPCR DE L'ADN VIRAL	113
5. ANALYSE DES EFFETS PRODUITS PAR LES MUTATIONS DANS LA SEQUENCE 3' TAR.....	118
5.1 STRUCTURES SECONDAIRES DES MUTANTS 3'TAR.....	118
5.2 RESULTATS DES TRANSFECTIONS PAR LES MUTANTS 3'	119
5.3 RESULTATS DES TRANSDUCTIONS PAR LES MUTANTS 3'	120
5.4 RESULTATS DE L'ANALYSE DE L'ENCAPSIDATION DE L'ARNg DES MUTANTS 3'.....	124
5.5 QPCR ADN VIRAL.....	125
6. ANALYSE DES EFFETS PRODUITS PAR LES MUTATIONS EN 5'TAR ET 3'TAR.....	127
6.1 BUT D'ETUDIER LES DOUBLES MUTANTS	127
6.2 RESULTATS DES TRANSFECTIONS PAR LES MUTANTS 5'-3'	128
6.3 RESULTATS DES TRANSDUCTIONS PAR LES MUTANTS 5'-3'	129
6.4 RESULTATS ENCAPSIDATION ARNg DES MUTANTS 5'-3'	132
6.5 QPCR ADN VIRAL.....	133
CONCLUSIONS ET PERSPECTIVES	134
REFERENCES BIBLIOGRAPHIQUES	138

ABRÉVIATIONS

ADN : acide désoxyribonucléique	NHEJ : <i>Non-Homologous End-Joining</i>
ADNt : ADN tardif	NXF1 : <i>nuclear RNA export factor 1</i>
ADNss : ADN <i>strong-stop</i>	PAS : <i>primer activation signal</i>
AIDS : <i>Acquired immunodeficiency syndrome</i>	pb : paire de bases
ARN : acide ribonucléique	PBS : <i>primer binding site</i>
ARNg : ARN génomique	PE : Ψ <i>element</i>
ARNt : ARN de transfert	PIC : <i>Pre-integration complex</i>
ASP : <i>anti-sens protein</i>	PPT : <i>Poly purine tract</i>
ARV : <i>AIDS-associated retrovirus</i>	Psi : <i>packaging signal</i>
BET : bromure d'éthidium	PTEFb : <i>positive transcription elongation factor</i>
BICD2 : <i>bicaudal D2 protein</i>	Rev : <i>Regulator of virion expression</i>
BST2 : <i>bone marrow stromal antigen 2</i>	RRE : <i>Rev responsive element</i>
Cdk9 : <i>Cyclin-dependent kinase 9</i>	RT : <i>reverse transcriptase</i>
CRFs : <i>circulating recombinant forms</i>	RTC : <i>reverse transcription complex</i>
CRS : <i>cis-acting repressive sequences</i>	SAMHD1 : <i>SAM and HD domain-containing protein 1</i>
CMH II : complexe majeur d'histocompatibilité de classe II	SIDA : syndrome de l'immunodéficience acquise
CPSF6 : <i>Cleavage and Polyadenylation Specific Factor 6</i>	SIV : <i>Simian immunodeficiency virus</i>
CypA : cyclophiline A	SLIP : <i>Slippery site</i>
DTG : dolutégravir	snRNP : <i>small nuclear ribonucleoprotein</i>
EFV : efavirenz	STLV-III : <i>Simian T-Lymphotropic Virus III</i>
ERK : <i>extracellular signal-regulated kinases</i>	TAP :
FEZ1 : <i>fasciculation and elongation protein zeta 1</i>	TAR : <i>Transactivator response element</i>
FP : <i>fusion protein</i>	Tat : <i>Transactivator of transcription</i>
HBR : <i>Highly Basic Region</i>	TLE : <i>tRNA-like element</i>
hnRNP : <i>heterogenous nuclear ribonucleoproteins</i>	URFs : <i>unique recombinant forms</i>
HR : <i>heptad protein</i>	USE : <i>upstream sequence element</i>
HTLV-III : <i>Human T-Lymphotropic Virus</i>	Vif : <i>Viral infectivity factor</i>
ICTV : <i>International Committee on Taxonomy of Viruses</i>	VIH-1 : Virus de l'Immunodéficience Humaine de type 1
INS : <i>inhibitory/instability RNA sequences</i>	VIH-2 : Virus de l'Immunodéficience Humaine de type 2
IP6 : <i>inositol hexakisphosphate</i>	Vpr : <i>Viral protein R</i>
kDa : kilodalton	Vpu : <i>Viral protein U</i>
LAV : <i>lymphadenopathy associated-virus</i>	
LT CD4 ⁺ : Lymphocytes T CD4 ⁺	
LEDGF/p75 : <i>lens epithelium-derived growth factor</i>	
LTR : <i>Long Terminal Repeat</i>	
MPER : <i>membrane-proximal external region</i>	
Nef : <i>Negative regulatory factor</i>	
NES : <i>nuclear export sequence</i>	

INTRODUCTION

1. VIH : agent pathogène responsable du SIDA

1.1 Les rétrovirus et leur classification

Les rétrovirus sont une famille de virus caractérisés par un génome diploïde d'ARN monocaténaire de polarité positive (1). Leur matériel génétique est présent dans la capsidie de la particule virale contenant aussi des enzymes comme la transcriptase inverse et l'intégrase. Cette capsidie est entourée par une membrane lipidique, l'enveloppe virale (Figure 1) (2).

Les rétrovirus se répliquent au sein de cellules infectées qui fournissent les systèmes de synthèse et les sources d'énergie nécessaires. Ils sont caractérisés par leur capacité à réaliser la transcription inverse. Au cours de ce processus, l'ARN simple-brin est rétro-transcrit en une molécule d'ADN double-brin intégrée dans le génome cellulaire (l'ADN proviral) afin d'exprimer de nouveaux génomes viraux et les différentes protéines virales qui s'assembleront ensuite pour former de nouveaux virions. Le génome de tous les rétrovirus infectieux possède deux *Long Terminal Repeats* (LTR) qui encadrent au minimum trois principaux gènes : *gag* (codant pour les protéines structurales), *pol* (codant pour les enzymes virales), *env* (codant pour les glycoprotéines d'enveloppe) (Figure 1) (2).

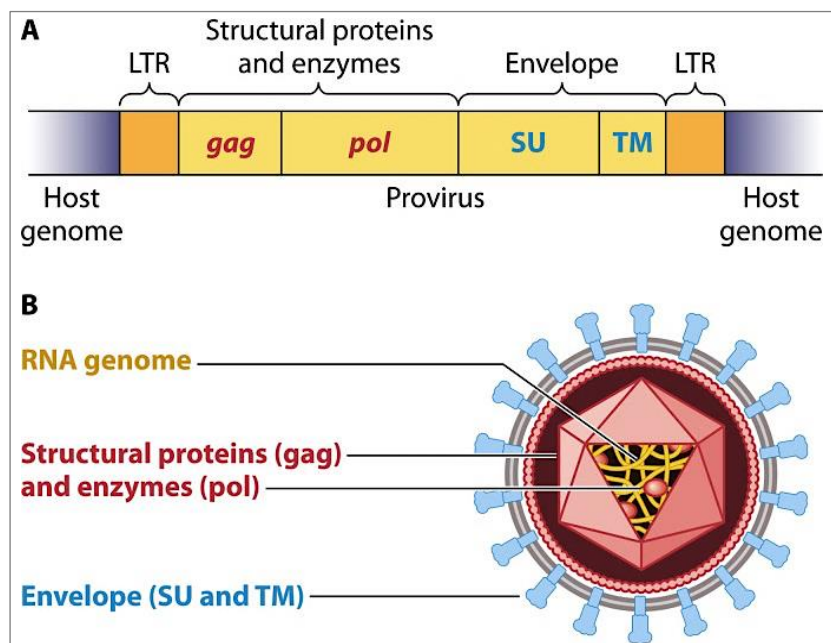


Figure 1. Caractéristiques d'un rétrovirus (2). (A) Organisation génétique. (B) Structure de la particule virale.

Le comité international de taxonomie des virus (ICTV) est chargé de la classification des virus. Celle-ci est basée principalement sur la nature du génome viral et son organisation, les propriétés de réplication ainsi que la structure de la particule virale (géométrie de la capsid, existence ou non d'une enveloppe par exemple). D'autres critères comme la pathogénicité, le spectre d'hôte « naturel » ou encore les propriétés physico-chimiques du virion peuvent être considérés (3).

L'ICTV a regroupé cinq familles de virus nécessitant une transcription inverse, dont celle des rétrovirus (*Retroviridae*), en un ordre nommé *Ortervirales*, de par leurs caractéristiques communes (notamment une monophylie des transcriptases inverses, une architecture de la polymérase, une amorce ARNt, la présence de capsid) (4).

En se basant sur le potentiel pathogénique, la famille *Retroviridae* possède onze genres, dont cinq représentent des rétrovirus avec un potentiel oncogénique. Les spumavirus n'ont actuellement pas été associés à des pathologies. En revanche, les lentivirus sont cytopathogènes et responsables de maladie à évolution lente chez les vertébrés. (Tableau 1).

Famille	Sous-famille	Genre	Nombre d'espèces
<i>Retroviridae</i>	<i>Orthoretrovirinae</i>	<i>Alpharetrovirus</i>	9
		<i>Betaretrovirus</i>	5
		<i>Deltaretrovirus</i>	4
		<i>Epsilonretrovirus</i>	3
		<i>Gammaretrovirus</i>	18
		<i>Lentivirus</i>	10 dont les VIH-1 et VIH-2
	<i>Spumaretrovirinae</i>	<i>Bovispumavirus</i>	1
		<i>Equispumavirus</i>	1
		<i>Felispumavirus</i>	1
		<i>Prosimiispumavirus</i>	1
		<i>Simiispumavirus</i>	15
<i>Metaviridae</i>	-	2 genres	34
<i>Caulimoviridae</i>	-	11 genres	94
<i>Belpaoviridae</i>	-	1 genre	11
<i>Pseudoviridae</i>	-	3 genres	34

Tableau 1. Classification des *Ortervirales*. (D'après la version 2022 de la taxonomie des virus, <https://talk.ictvonline.org/taxonomy/>).

1.2 Découverte scientifique du VIH

Les rétrovirus sont responsables de plusieurs affections chez diverses espèces animales telles que des tumeurs, des dégénérescences du système nerveux ou encore des pathologies pulmonaires et articulaires. Le virus de l'immunodéficience humaine (VIH) est l'agent pathogène causant le syndrome de l'immunodéficience humaine acquise (SIDA). Celui-ci entraîne la disparition progressive de ses principales cellules cibles, les lymphocytes T CD4⁺ (LT CD4⁺), essentielles au bon fonctionnement du système immunitaire et permettant notamment l'élimination des cellules infectées. Avec l'évolution de la maladie, le système immunitaire n'est plus en mesure de protéger suffisamment l'organisme laissant alors des maladies opportunistes se développer.

Ce syndrome a été décrit pour la première fois au début des années 1980, aux États-Unis dans la communauté homosexuelle. Il correspondait à une augmentation importante de cas d'infections opportunistes telles que la pneumonie causée par *Pneumocystis carinii* ou encore l'apparition d'une pathologie rare comme le sarcome de Kaposi. À ces maladies était associé une diminution importante du taux de lymphocytes T CD4⁺. Le syndrome d'abord baptisé « *Gay syndrome* » a rapidement été remplacé par « *Acquired immunodeficiency syndrome* » (AIDS ou SIDA en français) suite à l'apparition de cas chez les toxicomanes (prise de drogues par voie intraveineuse), les hémophiles recevant des fractions plasmatiques, des patients transfusés, des partenaires sexuels de patients atteints (hommes et femmes) et d'enfants nés de mères atteintes du syndrome. En Europe, des cas identiques ont rapidement été décrits chez ces différents groupes. Un foyer majeur a ensuite été identifié en Afrique Centrale et en Afrique de l'Est, notamment à Kinshasa, principalement chez les hétérosexuels suggérant l'existence d'un agent transmissible par voie sexuelle et sanguine (5). Cela a pu être mis en évidence en 1983 par Françoise Barré-Sinoussi et Luc Montagnier à l'Institut Pasteur (6), par la découverte du premier isolat responsable du SIDA, appelé *lymphadenopathy associated-virus* (LAV) chez un patient atteint d'une lymphadénopathie cervicale.

L'année suivante, l'équipe de Robert Gallo et Jay A. Levy ont identifié deux rétrovirus décrits comme responsable du SIDA nommés respectivement *Human T-Lymphotropic Virus* (HTLV-III) (7) et *AIDS-associated retrovirus* (ARV) (8). En 1985, la communauté scientifique a conclu que les trois rétrovirus étaient identiques. L'année suivante, un nouveau virus a été isolé de deux malades d'Afrique de l'Ouest ayant également pour cible les lymphocytes T CD4⁺, mais se distinguant du LAV-1 par sa séquence génomique et ses protéines, il fut nommé LAV-2 (9). La même année, LAV-1 et LAV-2 ont été respectivement renommés VIH-1 et VIH-2.

Le même nom a été attribué à ces virus et ils font partis de la même classe d'espèces de par leur similarité en termes d'organisation génétique, de cycles infectieux et de symptômes de la maladie mais leurs origines phylogénétiques sont bien distinctes.

1.3 Origine et variabilité génétique

En 1987, il a été établi un lien de parenté antigénique entre un virus simien provoquant une immunodéficience chez le macaque rhésus et le VIH-2 (10). Très vite, des virus, nommés *Simian immunodeficiency virus* (SIV), affectant divers primates tels que les singes verts africains, les mangabeys fuligineux ou encore les chimpanzés ont été identifiés (11). Des virus simiens affectant les chimpanzés (12, 13) et les mangabeys fuligineux (14) ont été caractérisés respectivement comme des proches parents du VIH-1 et du VIH-2. Ces observations ont pu fournir un argument fort que le SIDA chez l'Homme était le résultat de transferts zoonotiques de lentivirus de différentes espèces de primates (15). Le passage des SIVs à l'homme semble s'être fait par la chasse et la consommation de la viande de brousse, ou encore la domestication des singes. La classification phylogénique des lentivirus a pu mettre en évidence l'origine et l'évolution différente des deux VIH (Figures 2 et 3) (11).

Le VIH-1 est divisé en quatre groupes, M (pour major), O (pour outlier), N (pour non-M/non-O), et P selon sa diversité génétique. Le groupe M est classé en 9 sous-types et 125 formes recombinantes circulantes (CRFs) (16). Les CRFs sont des virus recombinants générés à partir d'une co-infection de deux (ou plusieurs) sous-types différents et possèdent une séquence génomique distincte des virus parents ne pouvant alors être classés dans aucun des 9 sous-types existant. Les CRFs sont nommés par ordre de découverte et identifiés par leur sous-type d'origine s'ils ont été retrouvés chez au moins trois patients épidémiologiquement non liés. Il arrive qu'une forme recombinante ne se retrouve que chez un seul patient ou des patients épidémiologiquement liés, on l'appelle alors forme recombinante unique (URF). Le groupe O fut identifié en 1990 chez deux patients d'Afrique centre-ouest et est endémique dans cette même région (environ 100 000 personnes sont infectées) (17). Le groupe N découvert chez une femme camerounaise est responsable d'une centaine d'infections (18). Le dernier identifié en 2009 est le groupe P (19). Les groupes N, O et P ne se limitent qu'aux résidents du Cameroun (et du Gabon pour le groupe N) ou aux voyageurs en provenance du Cameroun (mis à part pour un cas) (20, 21) et ne représentent qu'entre 1 et 2% des infections au VIH-1 (20).

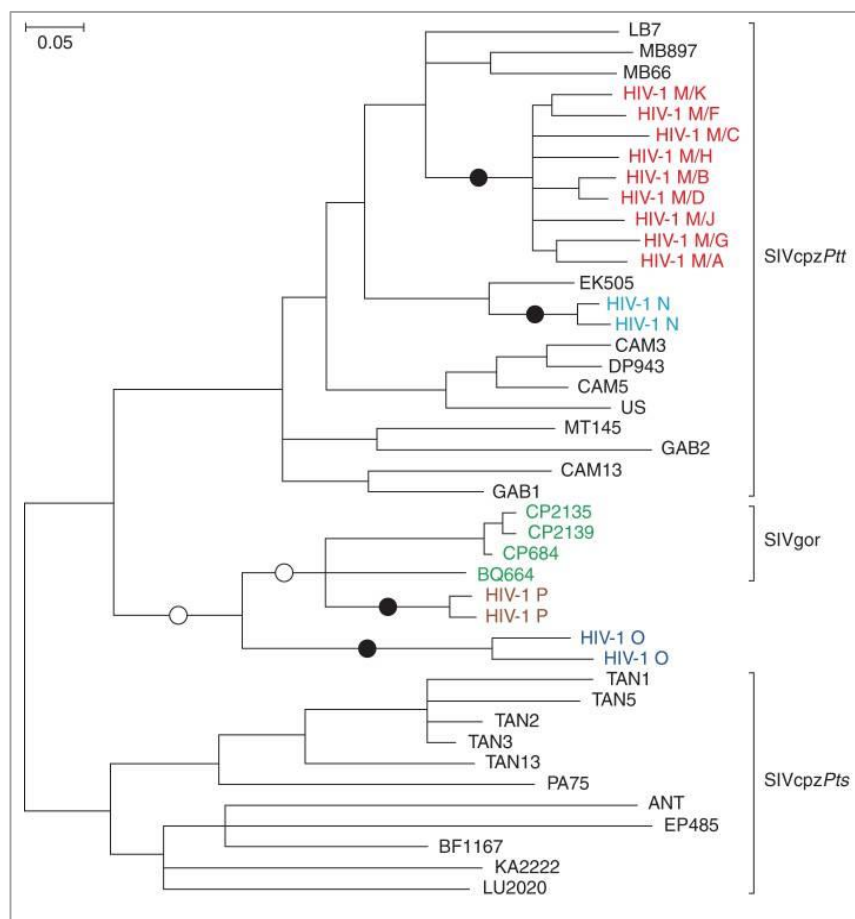


Figure 2. Arbres phylogénétiques du VIH-1 (11). Relations phylogénétiques représentées pour une région du gène viral *pol*. Les quatre groupes du VIH-1 représentent une transmission croisée indépendante. Les cercles noirs indiquent les quatre branches où la transmission inter-espèces à l'homme s'est produite. Les cercles blancs indiquent deux autres branches possibles sur lesquelles la transmission du chimpanzé au gorille s'est produite. Les crochets à droite indiquent respectivement SIVcpz de *P. troglodytes* (SIVcpz *Ptt*) et *P. t. schweinfurthii* (SIVcpz *Pts*).

Les groupes M et N ont pour origine le SIV du chimpanzé commun (SIVcpz de *Pan troglodytes troglodytes*) en Afrique centrale. Les deux groupes se sont formés à la suite de deux transmissions distinctes du virus de l'animal à l'homme. Des études phylogénétiques ont mis en évidence que le groupe M serait responsable du premier foyer endémique autour des années 1920 à Kinshasa (Léopoldville à l'époque) se disséminant en Afrique Centrale puis vers Haïti autour des années 60 se retrouvant rapidement à New York au début des années 70 (22). Cela a entraîné la propagation du virus dans le reste du continent et du monde avec plusieurs autres évènements indépendants comme en Angleterre en 1980 (23) (Figure 4) (24).

Le SIV du gorille du Cameroun (SIVgor) est issu de la transmission par le chimpanzé et semble être un réservoir potentiel d'infection pour les groupes O (25) et P (26). Les deux groupes semblent être eux aussi issus de deux évènements de transmission indépendants de l'espèce *Gorilla gorilla* vers *Homo sapiens*.

Le VIH-2 est divisé en huit groupes nommés de A à H (27), les groupes A et B étant les plus représentés. Ce second type du VIH est surtout répandu en Afrique occidentale avec des taux de prévalence plus élevés en Guinée-Bissau et au Sénégal (28). Aujourd'hui ces taux diminuent (29). Entre 1 et 2 million des personnes sont infectés par le VIH-2 d'après les dernières données de 2018 (30). Par comparaison avec le VIH-1, les personnes séropositives au VIH-2 présentent une période de latence clinique plus longue, une progression de la pathologie plus lente et une plus faible charge virale au cours de la phase asymptomatique (31). En 1989, l'origine mangabey fuligineuse (SIV_{smm}) du VIH-2 a été proposée et a été confirmée par plusieurs équipes (32). Les différentes lignées semblent être un transfert zoonotique indépendant.

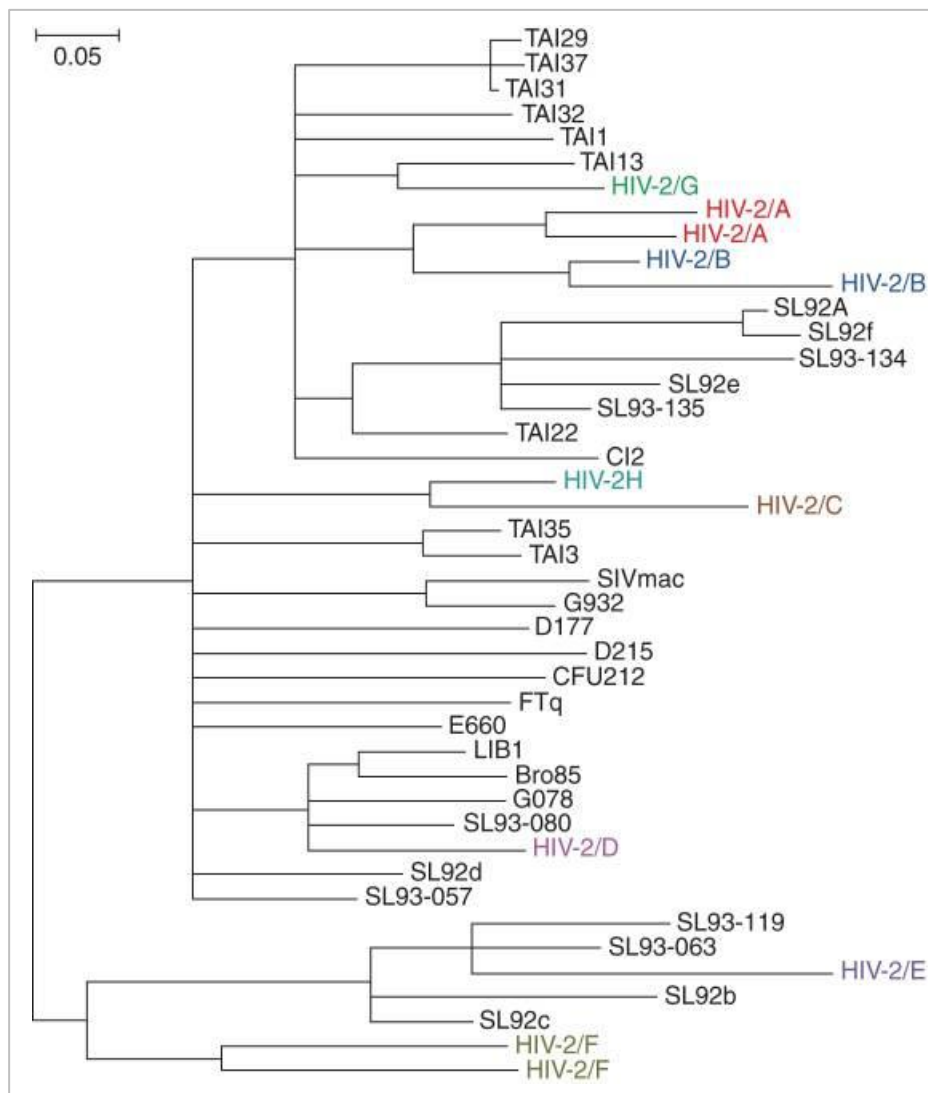


Figure 3. Arbres phylogénétiques du VIH-2 (11). Relations phylogénétiques représentées pour une région du gène viral *gag*. Les différents groupes du VIH-2 représentent une transmission croisée indépendante. Les virus non notés HIV-2 et SIV_{smm} sont des SIV_{smm}.

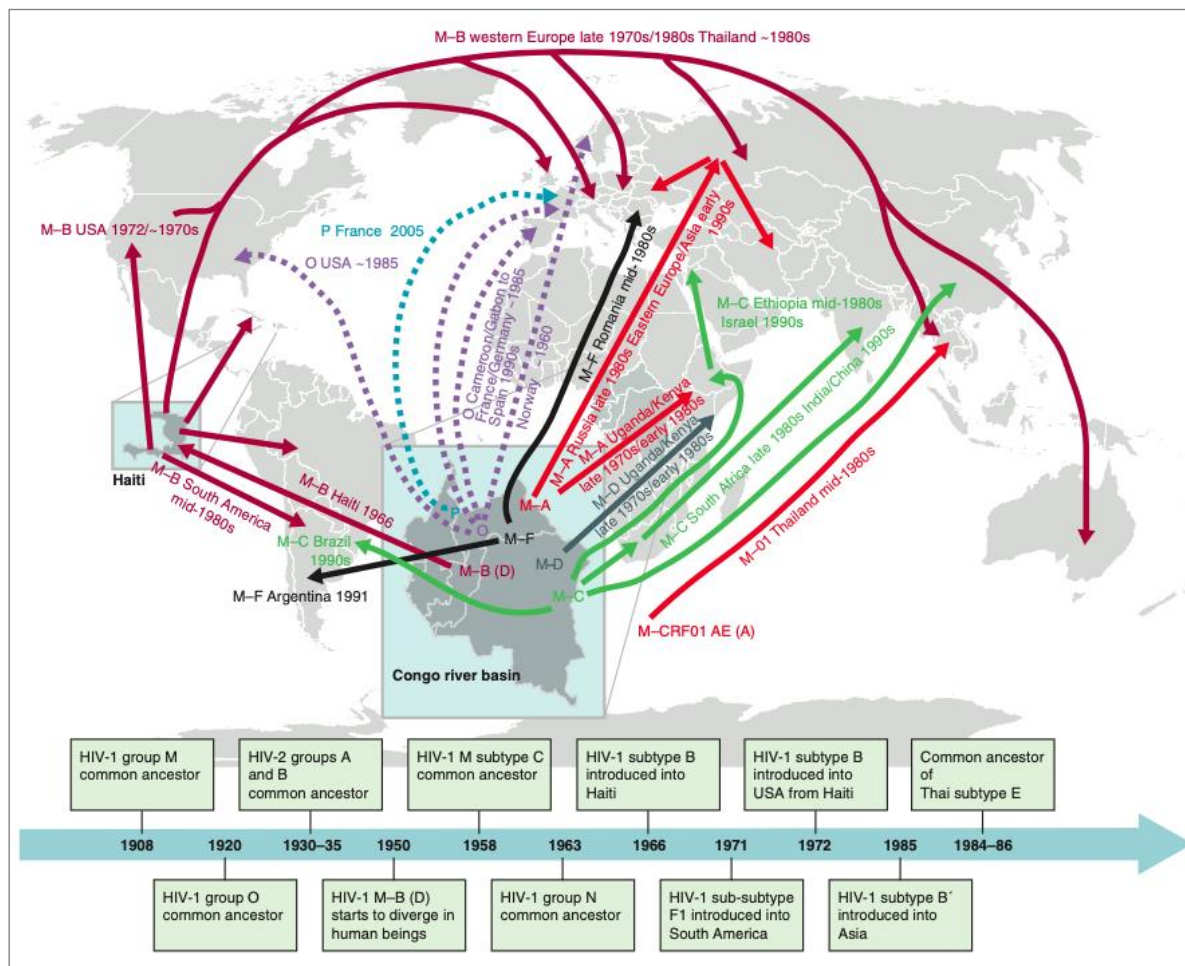


Figure 4. Propagation mondiale supposée à partir du premier cœur endémique au Congo (24). Sur la carte sont indiquées par des flèches les voies ainsi que les dates estimées d'introduction des sous-types M, N, O et P du VIH-1 et des CRF du bassin du fleuve du Congo vers d'autres régions du monde. La flèche bleue indique la chronologie estimée des ancêtres communs des VIH-1 et 2 ainsi que l'introduction estimée de certains sous-types.

L'étude en laboratoire du VIH-1 se fait principalement à partir des isolats du groupe M, ce sont HXB2, NL4-3 et LAI du sous-type B et MAL qui est un recombinant des sous-types A1, D et K.

1.4 Situation épidémiologique

Fin 2021, il a été estimé que 38,4 millions de personnes vivaient avec le VIH dont les deux tiers se trouvent en Afrique avec 25,6 millions de personnes. C'est le continent dont la prévalence est la plus élevée (Figure 5). Il a été estimé qu'un million et demi de personnes ont été infectées en 2021, soit deux fois moins qu'à la fin des années 90 (Figure 6). Cela représente un peu plus de 4000 nouvelles infections quotidiennes dont 400 sont des enfants. À l'échelle mondiale, plus de 0,5% de la population mondiale est infectée par le VIH et le taux de mortalité est de 11 décès pour 100 000 habitants. Depuis 2010, l'accès aux traitements antirétroviraux est passé de 25% à 76% pour les adultes mais seulement la moitié des enfants vivants avec le VIH (52%) y ont accès.

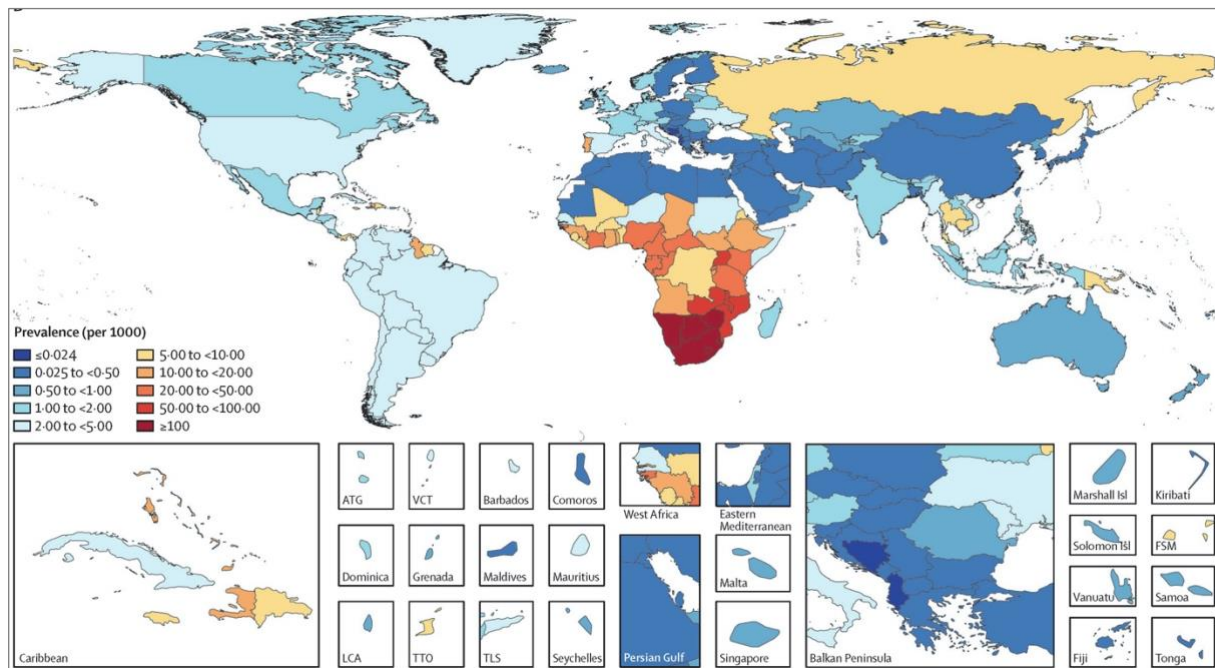


Figure 5. Prévalence du VIH-1 en 2017. (Source : World Health Organization)

En 2014, un programme mondial de lutte ayant pour objectif de mettre fin au VIH/SIDA a été rédigé par l'ONUSIDA. L'objectif était appelé « 90-90-90 » et devait être atteint avant 2020.

Le but :

- 90% des personnes atteintes doivent connaître leur séropositivité,
- 90% de toutes les personnes infectées par le VIH et dépistées doivent avoir un traitement anti-rétroviral durable
- Et 90% des personnes recevant un traitement antirétroviral doivent avoir une charge virale détectable durablement supprimée.

Cela a été le cas pour une quinzaine de pays mais la pandémie de la COVID-19 a ralenti considérablement les dépistages, la mise en place des traitements et la prévention dans plusieurs régions. À la fin 2019, les résultats étaient de 81% des personnes vivant avec le VIH connaissaient leur statut, parmi elles, 67% avaient accès à un traitement et 59% des personnes traitées avaient une charge virale indétectable. Les chiffres à atteindre sont maintenant de 95% d'ici 2030.

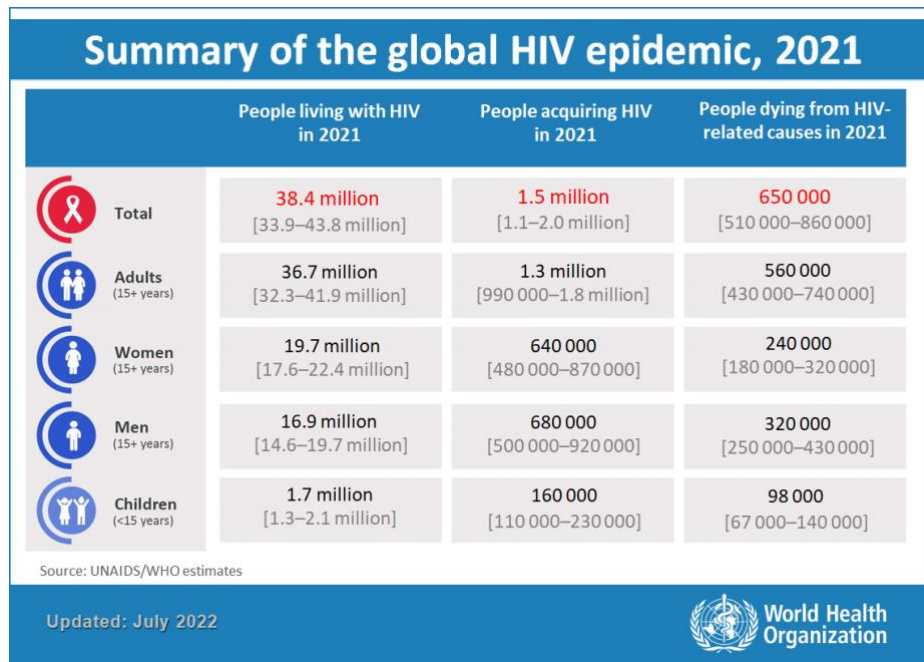


Figure 6. Principaux chiffres de l'épidémie du VIH en 2021. (Source : World Health Organization)

1.5 Pathogénèse du VIH

Le VIH se transmet par l'échange de fluide corporel par voie sexuelle, sanguine, parentérale et de la mère à l'enfant (allaitement ou accouchement) (33). La principale voie de transmission du virus est la voie sexuelle lors de rapports non protégés. Cette transmission peut en plus être facilitée lors d'une inflammation du tractus génital, d'un déséquilibre de la flore bactérienne ou des conséquences de maladies sexuellement transmissibles tels que les ulcères génitaux.

Le VIH a un tropisme pour les lymphocytes T CD4⁺. Le virus infecte également d'autres cellules comme les macrophages et les cellules dendritiques qui assurent la mise en place d'une réponse immunitaire. Les cellules cibles du virus ont une longue durée de vie, sont circulantes et résident dans les tissus lymphoïdes (ganglions lymphatiques, thymus, moelle osseuse, rate et tissus lymphoïdes associés aux muqueuses tels que les intestins). Ces derniers sont les principaux réservoirs anatomiques du virus avec le cerveau et le sang périphérique. Ce réservoir anatomique est caractéristique de chaque patient et s'établit lors de la primo-infection. Les cellules infectées quiescentes dans lesquelles se trouvent des ADN viraux représentent un réservoir cellulaire. Elles échappent aux traitements antirétroviraux et au système immunitaire pendant plusieurs dizaines d'années. En cas d'arrêt de traitement, ces cellules peuvent être activées et le virus se multiplier massivement. L'infection par le VIH suit différentes phases : aigüe, chronique et le stade du SIDA (Figure 7) (34).

Lors de la primo-infection, le virus traverse dans un premier temps la barrière mucoale sous sa forme libre ou via des cellules infectées du « donneur » ou en infectant les premières cellules rencontrées telles que les LT CD4⁺ épithéliales, les macrophages ou encore les cellules dendritiques dont les cellules de Langerhans. Les dix premiers jours, le virus n'atteint que les cellules cibles des muqueuses du site de transmission et les vaisseaux sanguins proches (33). Durant cette courte phase, le virus est indétectable. Rapidement, celui-ci parvient à coloniser les tissus lymphoïdes. Ceci se traduit par une charge virale élevée dans le sang au bout de trois à quatre semaines, un risque considérable de transmission et une diminution importante, progressive et transitoire des LT CD4⁺ (Figure 7). Cette diminution est la résultante de trois mécanismes majeurs :

- i) le rétrovirus détruit les cellules infectées,
- ii) des cellules infectées entrent en apoptose,
- iii) les lymphocytes cytotoxiques CD8⁺ (LT CD8⁺) reconnaissent les lymphocytes T CD4⁺ infectés et les détruisent à la suite de la mise en place de la réponse immunitaire.

Ces derniers provoquent une diminution de la charge virale et une augmentation des LT CD4⁺ périphériques (35). Cette phase aigüe dure entre trois et douze semaines après l'installation de l'infection. La phase de latence clinique, dite chronique, est une phase asymptomatique d'une durée de trois à dix ans selon les patients. Certains patients infectés par le VIH sont capables de contrôler l'infection sans avoir recours à un traitement antirétroviral grâce à une réponse immunitaire suffisamment efficace qui conduit à une charge virale indétectable (36). Chez les patients dont la progression suit un schéma classique la charge virale reste stable et un réservoir cellulaire se met en place. Les provirus intégrés au génome de cellules immunitaires restent sous forme latente mais cette intégration permet une propagation du virus grâce à une communication constante entre les cellules immunitaires circulantes. C'est durant cette phase que le taux de LT CD4⁺ diminue progressivement jusqu'à atteindre des niveaux critiques (< à 500 LT CD4⁺/mm³ dans le sang). La dernière phase correspond au stade du SIDA. La réplication virale épuise le système immunitaire. Le déclin du nombre de lymphocytes T CD4⁺ en dessous du seuil critique entraîne une perte d'immunité (< 200 LT CD4⁺/mm³). Progressivement, l'organisme devient une cible pour les infections opportunistes et les complications tumorales. Le décès du patient survient en moyenne neuf mois après l'entrée dans le stade du SIDA.

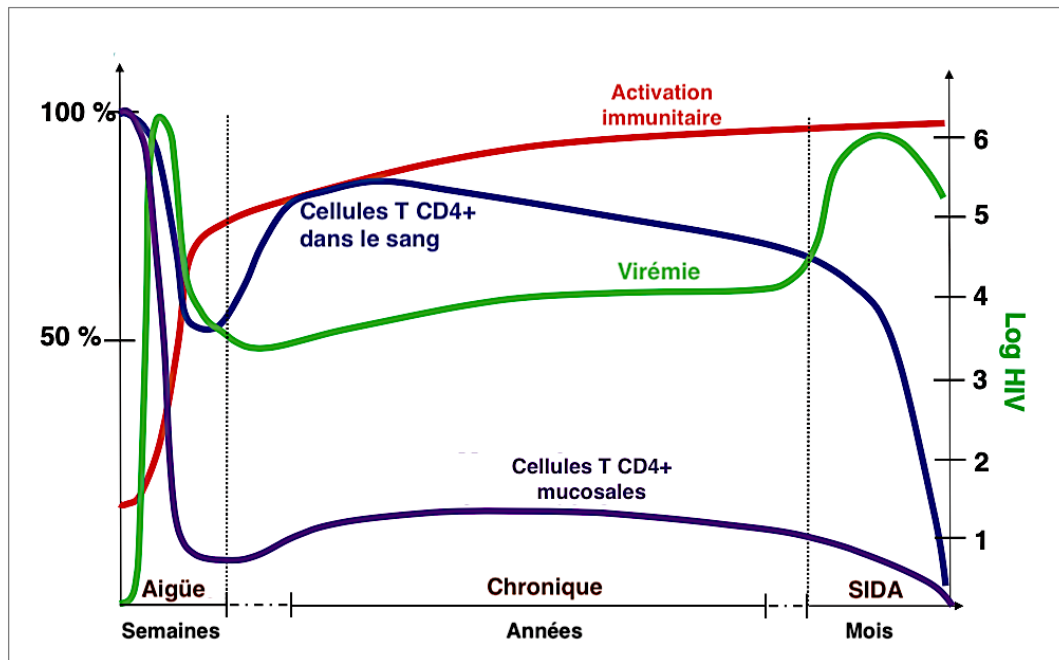


Figure 7. Évolution de l'infection naturelle du VIH-1 chez une personne infectée sans traitement antirétroviral (34). L'infection par le VIH-1 décrit trois phases caractérisées par une évolution au cours du temps de la virémie (courbe verte), du taux de LT CD4+ dans les muqueuses (courbe violette) et dans le sang (courbe bleue) au cours du temps. Une réponse immunitaire stable est rapidement observée, au bout de quelques semaines après la primo-infection jusqu'au stade SIDA (courbe rouge).

2. VIH : Anatomie du virus, génome et cycle de vie

2.1 Structure de la particule virale

La particule virale mature et infectieuse (virion) du VIH-1 est de forme sphérique et a un diamètre d'environ 145 nm (37). Sa structure est caractérisée par une enveloppe virale, dérivée de la membrane cellulaire de la cellule hôte. Elle est constituée d'une bicouche lipidique enrichie en protéines d'enveloppe, qui sont trois glycoprotéines gp120 formant la coiffe et trois glycoprotéines gp41 ancrées dans la membrane virale. Le cœur de la particule mature est composé de la capsidie présentant une forme conique. Celle-ci contient le matériel génétique du virus, deux molécules d'ARN génomique (ARNg), associées à des protéines de nucléocapside et des enzymes virales telles que la protéase, l'intégrase et la transcriptase inverse. La matrice composée de protéines de matrice est située entre l'enveloppe virale et la capsidie (Figure 8) (38).

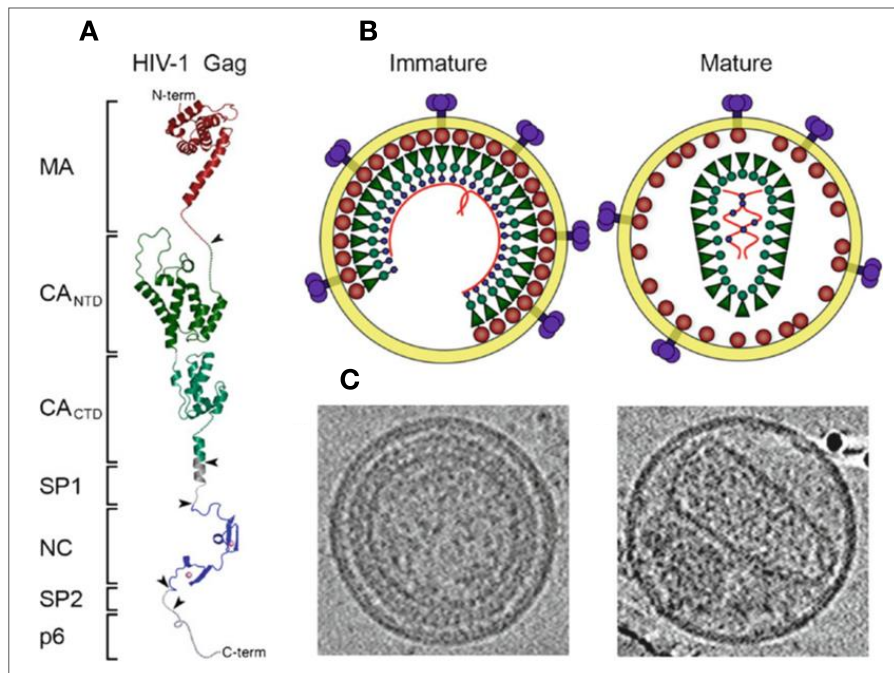


Figure 8. Structure des particules virales immature et mature du VIH-1 (38). (A) Structure tertiaire de Gag. (B) Organisation schématique de la structure des particules virales immature et mature. En mauve : gp120, en jaune : bicouche lipidique, en rouge : matrice (MA), en vert : capsid (CA), en bleu : protéine de nucléocapside (NC), filament rouge : ARNg. (C) Observation des particules virales immature et mature par cryotomographie électronique.

2.2 Organisation génétique

L'ARN viral de presque dix kilo bases possède sept marqueurs structuraux (UTR, TAR, RRE, PE, SLIP, CRS et INS) et neuf gènes (*gag*, *pol*, *env*, *tat*, *rev*, *nef*, *vif*, *vpr*, *vpu*) codant pour quinze protéines canoniques (39) (dix-neuf en incluant les précurseurs et la protéine anti-sens Asp (40)) grâce à des cadres de lectures chevauchants (Figure 9). Les gènes structuraux *gag* et *env* permettent la synthèse des protéines jouant un rôle crucial dans l'architecture de la particule virale. Le gène *pol* code pour les enzymes virales réalisant les principales étapes de la réplication telles que la transcription inverse, l'intégration et la protéolyse des protéines virales. Les gènes *tat* et *rev* sont des gènes régulateurs tandis que les gènes *nef*, *vif*, *vpr* et *vpu* sont des gènes accessoires. Ces deux derniers types de gènes permettent aux virus d'infecter des cellules, de réguler les gènes viraux, d'assembler de nouvelles particules virales et sont responsables des propriétés cytopathogènes du virus.

L'existence d'un dixième gène, hautement conservé parmi 12 souches du groupe M, présent sur le brin anti-sens de l'ADN viral au niveau du gène *env*, a été mis en évidence *in vitro* (41) et *in vivo* (42, 43). Ce gène code pour la protéine ASP (*Anti-Sens Protein*) de 189 acides aminés dont le transcrit aiderait à induire et/ou à maintenir la latence virale (44, 45) de façon épigénétique. La fonction de la protéine n'est pas encore clairement établie.

Des hypothèses d'un avantage sélectif du virus (46) ou une corrélation entre l'origine de la pandémie et le sous-type M, chez qui ASP est uniquement présente, ont été émises (45).

Il existe plusieurs protéines de fusion non canoniques dans les cellules infectées qui sont codées par exemple par le locus *tat-env-rev* dont le produit est la protéine Tev. Celles-ci ont été retrouvées uniquement chez certaines souches et peuvent être issues d'événements d'épissages alternatifs, d'autres sont produites lors d'un changement de cadre de lecture du ribosome. Elles peuvent réguler l'activité de leurs protéines parentales ou avoir de nouvelles fonctions mais *in vivo* il n'a pas été démontré de fonctions pertinentes de ces protéines. Elles ne semblent pas indispensables à la réplication du virus (47).

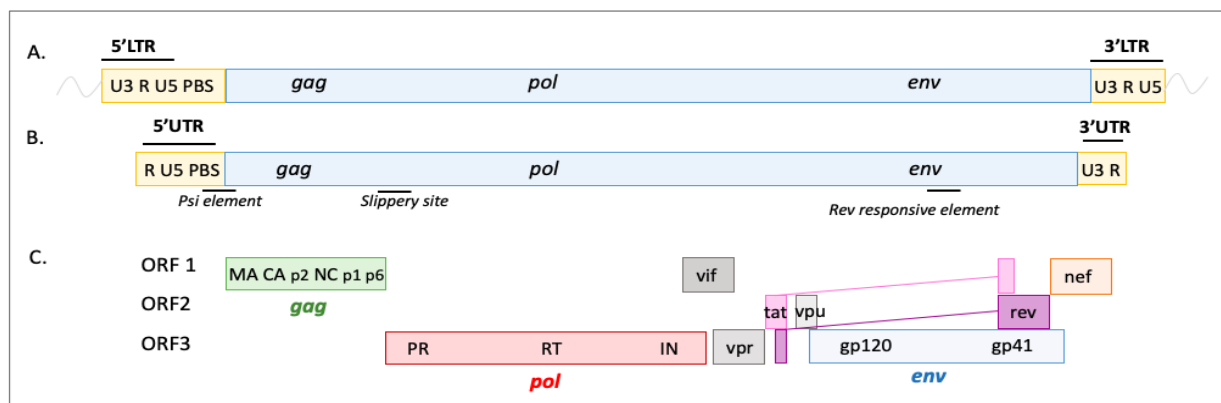


Figure 9. Organisation génétique du VIH-1. (A) ADN proviral. (B) ARN génomique. (C) Protéines du VIH-1 traduites selon les différents cadres de lecture ouverts (ORF). U5 : *unique 5' region* ; PBS : *primer binding site* ; U3 : *unique 3' region* ; R : *repeated region* ; MA : *matrix* ; CA : *capsid* ; NC : *nucleocapsid protein* ; p1 : *spacer domain 2* ; p2 : *spacer domain 1* ; PR : *protease* ; RT : *reverse transcriptase* ; IN : *integrase* ; vif : *viral infectivity factor* ; vpu : *viral protein U* ; vpr : *viral protein R* ; gp : *glycoprotein* ; nef : *negative factor*.

Les LTR (*Long Terminal Repeat*) ont une longueur d'environ 640 nucléotides segmentés en trois régions, U3, R et U5 et se trouvent aux extrémités 5' et 3' de l'ADN proviral. Ils possèdent de nombreux éléments *cis* régulateurs notamment ceux permettant la régulation des gènes viraux (48) ou encore l'initiation de la transcription (49) et de la polyadénylation (50). La région U3 est subdivisée en trois parties fonctionnelles : modulatrice (plusieurs sites de fixation de facteurs de transcriptions), activatrice (qui contient des sites de liaison pour le facteur de transcription NF- κ B par exemple) et promotrice (contenant la TATA box) (51, 52).

Le TAR (*Transactivator response element*) constitué de 54 à 57 nucléotides selon les isolats forme une tige-boucle située dans la partie R présente aux deux extrémités de l'ARNg. Le TAR en 5' fixe des protéines cellulaires et la protéine virale Tat pour former un complexe ribonucléoprotéique activant fortement la transcription du génome du VIH-1 (53–56).

Le RRE (*Rev responsive element*) est un domaine ARN d'environ 350 nucléotides situé dans la région *env*, plus précisément dans la partie englobant respectivement les extrémités 3' et 5' des domaines gp120 et gp41 (57). Le RRE adopte une structure secondaire complexe en plusieurs tiges-boucles (58). Cet élément est nécessaire à la fonction de la protéine Rev qui s'y fixe permettant ainsi l'export des ARNm (ARNm) partiellement épissés et non épissés (59). Cet élément structural est notamment très conservé à travers les sous-types du groupe M du VIH-1 (57).

Le SLIP (*Slippery site*), qui est présent entre les gènes *gag* et *pol*, est un domaine contenant le motif UUUUUUA suivi d'une structure en tige-boucle. Ce marqueur structural régule le décalage du cadre de lecture ribosomal en -1 hors du cadre de lecture du gène *gag* et dans le cadre de lecture du gène *pol*. Ce décalage de cadre de lecture permet la traduction du précurseur polypeptidique Gag-Pol, à partir duquel les enzymes virales sont obtenues, et permet de réguler un rapport entre les enzymes et les protéines structurales nécessaire à l'infectivité (60–64).

Le PE (Ψ *element*), présent dans la région 5'UTR de l'ARNg jusqu'à l'extrémité 5' du gène *gag*, est constitué de cinq structures en tige-boucle (PBS et SL1 à SL4). Le Ψ *element* est retrouvé uniquement dans les transcrits non épissés. SL1 possède une séquence palindromique de six nucléotides décrit comme le site d'initiation de la dimérisation (DIS) de l'ARNg (65, 66) et est importante dans la sélection du génome viral par le précurseur polypeptidique Gag (67, 68). SL2 contient le principal site donneur d'épissage (69). SL3 est la région la plus conservée et jouerait un rôle dans l'encapsidation de l'ARNg dans la particule virale en interagissant avec la protéine Gag (70) mais cette structure ne semble pas indispensable à ce processus puisque SL1 intervient également et serait plus importante (69, 71). SL4 contient le codon d'initiation AUG du gène *gag* mais sa structure est instable (72).

Les CRS (*cis-acting repressive sequences*) sont des séquences inhibitrices de l'expression des protéines structurales en absence de la protéine Rev. Un site a notamment été cartographié dans les séquences du gènes *pol* (73, 74) La fonction exacte des CRS est encore méconnue et certains sites d'épissage agiraient comme des séquences CRS (62).

Les INS (*inhibitory/instability RNA sequences*) sont dans les gènes de structure de certains rétrovirus complexes dont le VIH-1. Ils semblent agir comme des éléments inhibant l'expression génique de manière post-transcriptionnelle en affectant la stabilité de l'ARN ou l'exportation nucléaire. L'effet inhibiteur a pu être annulé par des mutations. Les éléments INS existent dans le génome du VIH-1, par exemple dans le gène *gag* (62, 74).

2.3 Cycle de réplication

Le cycle répliatif du VIH-1 (Figure 10) se divise en deux phases : précoce et tardive. La phase précoce comprend les étapes d'attachement à la cellule hôte jusqu'à l'intégration de l'ADN viral bicaténaire. La phase tardive inclue les étapes suivantes à partir de la transcription de l'ADN proviral jusqu'au bourgeonnement des nouvelles particules virales et leur maturation.

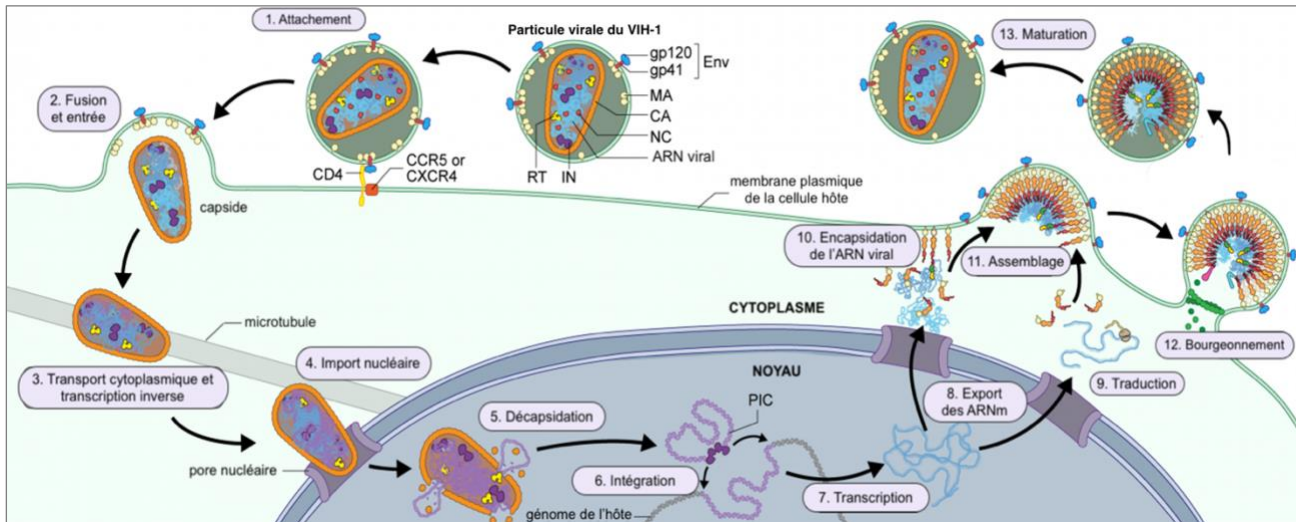


Figure 10. Cycle de réplication du VIH-1. Adapté d'un schéma du site scienceofhiv.org

2.3.1 Entrée du virus

a. Les protéines d'enveloppe du VIH

Au niveau de l'enveloppe virale trois sous unités de gp120 ancrées à la surface du virion sont liées de manière non-covalente à trois sous-unités transmembranaires de gp41 (75). Le complexe trimérique gp120-gp41 initie l'entrée dans la cellule hôte en interagissant avec l'un de ses récepteurs CD4. La gp120 contient cinq régions hautement variables (V1-V5), formant des boucles au sommet du complexe trimérique, et cinq régions conservées (C1-C5) (76). La gp41 possède une queue cytoplasmique, un domaine transmembranaire et un ectodomaine contenant plusieurs régions caractéristiques qui sont le *fusion peptide* (FP), deux séquences *heptad repeat* (HR1 et HR2), une boucle et la *membrane-proximal external region* (MPER) (Figure 11) (77).

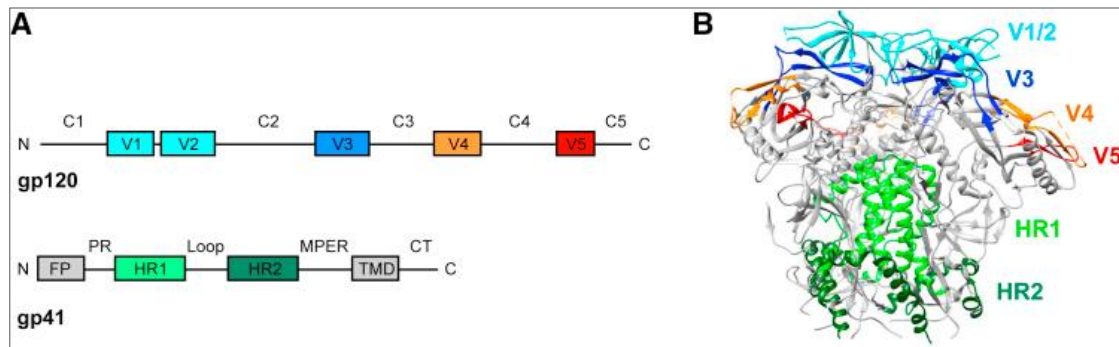


Figure 11. Protéines de l'enveloppe virale. (A) Représentation schématique de l'organisation structurale de gp120 et gp41 (77). Sont représentées les régions variables en couleurs (V1 à V5) et les régions constantes de gp120 C1 à C5). Les différents acronymes indiquent : queue cytoplasmique (CT), domaine transmembranaire (TMD), *fusion peptide* (FP), *fusion peptide proximal region* (PR), *N-terminal heptad repeat-1* (HR1), *C-terminal heptad repeat-2* (HR2), *membrane proximal external region* (MPER). (B) Structure de la protéine Env (PDB : 5FUU).

b. Les récepteurs et corécepteurs cellulaires

Les macromolécules nécessaires à l'entrée du virus sont le récepteur CD4 et les corécepteurs CXCR4 et CCR5. Le récepteur CD4 est une glycoprotéine exprimée à la surface des cellules T, des monocytes, des macrophages et certaines cellules dendritiques. Celui-ci interagit avec le complexe majeur d'histocompatibilité de classe II (CMH II) présent à la surface de cellules présentatrices d'antigènes permettant notamment l'activation des LT CD4⁺. Les récepteurs de chimiokines CXCR4 et CCR5 appartiennent à la famille des récepteurs couplés aux protéines G et possèdent sept domaines transmembranaires. Le récepteur CXCR4 est exprimé par les cellules sanguines (globules rouges et blancs notamment) et joue un rôle dans le recrutement tissulaire des cellules immunitaires. Le récepteur CCR5 est principalement exprimé par les LT CD4⁺ et les macrophages et est impliqué dans le recrutement des cellules immunitaires sur les sites inflammatoires. Néanmoins, ce récepteur ne semble pas être indispensable à l'Homme et les individus nés avec une mutation du gène CCR5 présenteraient une résistance naturelle au VIH (78).

c. Le mécanisme d'entrée du virus

Le processus peut être divisé en trois temps : la liaison de gp120 au récepteur CD4, la liaison de gp120 au corécepteur (CXCR4 ou CCR5) et la fusion membranaire déclenchée par gp41. La première étape induit des changements structuraux du complexe trimérique adoptant alors un état ouvert et exposant la boucle V3 de gp120 et la tige centrale de gp41 (Figure 12 A et B). Cela permet la liaison au corécepteur et l'amorçage de la fusion membranaire par gp41. Précisément, la liaison CXCR4 ou CCR5 à la gp120 étend les séquences HR de gp41, libérant le *fusion peptide* qui s'insère dans la membrane cellulaire (Figure 12 C et D).

Les HR en interagissant provoque le repliement sur lui-même de gp41 pour former une structure en boucle (ou « épingle à cheveux ») (Figure 12 E). Les membranes cellulaires et virales sont ainsi à une très faible distance conduisant à leur fusion et libérant le contenu de la particule virale, dont la capside, dans le cytoplasme de la cellule (Figure 12 F) (79).

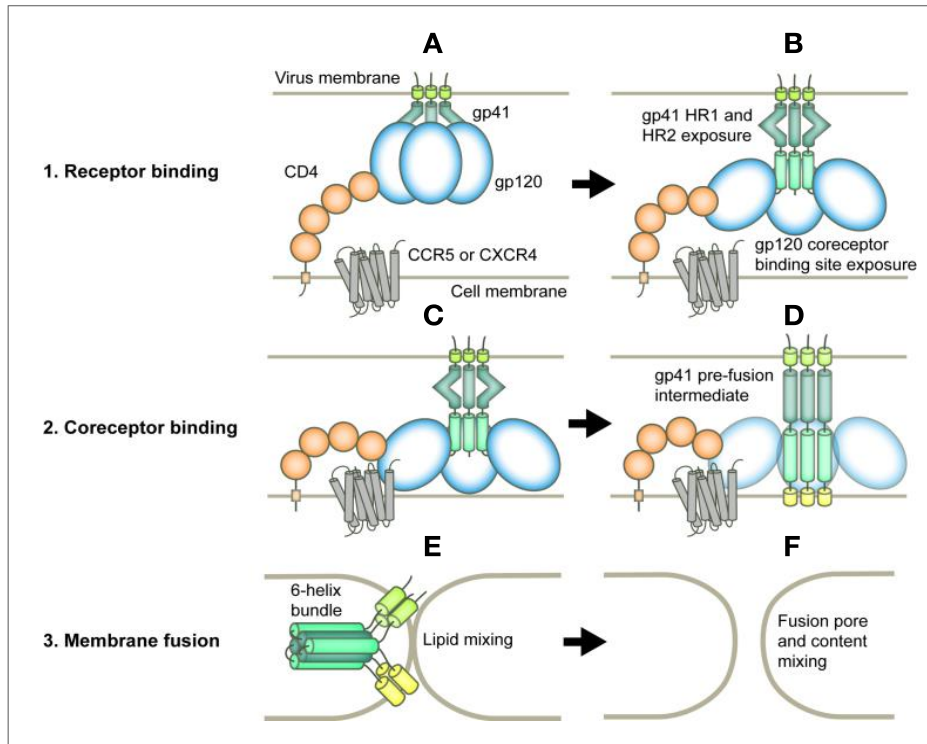


Figure 12. Entrée du VIH-1 dans la cellule (79). (1) Liaison au récepteur CD4. (2) Liaison au corécepteur. (3) Fusion membranaire.

2.3.2 Transcription inverse de l'ARN et décapsidation

La transcription inverse permet la conversion de l'ARNg en ADN double-brin pouvant intégrer le génome de la cellule hôte. Pour cela, le virus possède une protéine adéquate : la transcriptase inverse. L'étape de transcription inverse débute après l'entrée de la capside dans la cellule. Cette étape du cycle de réplication est détaillée dans le chapitre suivant.

La capside doit se désassembler, processus appelé décapsidation, pour permettre l'achèvement de la transcription inverse ou du moins permettre l'intégration de l'ADN viral au sein des chromosomes de l'hôte. A ce jour, les études ne sont pas parvenues à un consensus sur le moment, le lieu exact, la vitesse et les mécanismes régissant le processus de décapsidation. En effet, l'étude de la protéine composant la capside peut s'avérer complexe du fait de plusieurs facteurs :

- de nombreuses interactions observées avec les protéines cellulaires (80–85),
- une fragilité génétique significative et notoire de la capside (86),

- une minorité des particules virales entrantes entraîne une infection productive, les nombreux événements d'infection abortifs peuvent créer un bruit de fond lors de l'expérience compromettant l'interprétation des résultats,
- une distinction entre les capsides entières et les capsides partiellement désassemblées est difficile à mettre en évidence,
- des capsides entières pouvant être issues de particules non infectieuses (87, 88),
- une utilisation de marquage direct de la protéine pouvant engendrer des modifications de la morphologie de la capside,
- des résultats dépendant des techniques ou des lignées cellulaires utilisées.

Actuellement, plusieurs groupes développent des stratégies d'imageries sophistiquées afin de comprendre le processus de décapsulation *in situ* au niveau d'une seule particule ayant permis une infection productive (89–93).

a. Protéines responsables du désassemblage

La protéine de capsid (CA) est la composante formant la capsid conique du VIH (Figure 13A) (94). Avec 186 hexamères et 12 pentamères (soit 1176 molécules de CA), la capsid a une longueur de 100 à 200 nm et une largeur de 45 à 50 nm (95) ce qui lui offre une large surface d'interactions avec des protéines cellulaires. Ces dernières peuvent agir comme des facteurs antiviraux ou des cofacteurs viraux. Certaines interactions vont notamment favoriser la stabilité de la capsid, participer à l'import nucléaire de la capsid et/ou promouvoir la décapsulation. Parmi ces cofacteurs cellulaires (96), nous pouvons citer :

- La cyclophiline A (CypA), connue pour interagir avec les lentivirus (97), stabilise la capsid selon le type cellulaire étudié (97–99) et serait impliquée dans l'import nucléaire (100). Elle protège également le virus du facteur de restriction TRIM5 α (97, 101).
- BICD2 (*bicaudal D2 protein*) et FEZ1 (*fasciculation and elongation protein zeta 1*) : deux protéines agissant comme des adaptateurs de transport de la dynéine et de la kinésine lors du trafic du virus à travers la cellule qui lorsqu'elles sont altérées conduisent à des défauts de trafic cytoplasmique, de décapsulation et d'import nucléaire de la capsid (102–104).
- Les nucléoporines, constituant des pores nucléaires, Nup153 et Nup358 sont impliquées dans l'import nucléaire de la capsid tout comme le facteur CPSF6 (97) (*Cleavage and Polyadenylation Specific Factor 6*).

Plus précisément, Nup358 favoriserait la décapsidation (97, 105) tandis que Nup153 et CPSF6 influenceraient le choix du site d'intégration (97, 106, 107).

- Pin1 (108) (*peptidyl-prolyl isomérase*), ERK (109) (*extracellular signal-regulated kinase 2*), MELK (110) (*maternal embryonic leucine zipper kinase*) et TRN-1 (111) (*β -karyopherin transportin 1*) sont des protéines cellulaires qui interagissent avec la capsidite et ont un effet sur sa décapsidation.

b. Mécanismes de décapsidation

Comme dit précédemment, le déroulement temporel et à l'échelle moléculaire de la décapsidation n'est que partiellement élucidé. Néanmoins, nous savons que la décapsidation ne peut avoir lieu prématurément et de manière complète dans le cytoplasme. En effet, le contenu de la capsidite se diluerait dans l'environnement cytoplasmique et pourrait alors être exposé à une réponse antivirale. La capsidite ne peut cependant rester longtemps intacte dans la cellule hôte car cela aurait pour conséquence de retarder l'intégration de l'ADN viral et la poursuite du cycle répliatif. Plusieurs modèles sont décrits actuellement :

- **une décapsidation cytoplasmique** (Figure 13B première colonne) : plusieurs études ont montré la présence de protéines de capsidite et/ou de capsidites partiellement intactes dans des cellules infectées dont certaines évoquent un couplage avec la transcription inverse (84, 112–116). Des études récentes (83, 97, 117) suggèrent que le désassemblage de la capsidite ne se fait que partiellement dans le cytoplasme, révélant le caractère protecteur de la capsidite pour son contenu.
- **une décapsidation aux portes du noyau cellulaire** (Figure 13B deuxième colonne) : la capsidite intacte ne pourrait être importée dans le noyau car sa taille excèderait celle du diamètre du pore nucléaire. La capsidite, après avoir utilisé les microtubules comme voie de transport intracellulaires (118), serait amarrée au pore afin de libérer le *reverse transcription complex/pre-integration complex* (RTC/PIC) au cours de son désassemblage (100, 119–124).
- **une décapsidation à l'intérieur du noyau cellulaire** (Figure 13B troisième et quatrième colonne) : le dernier modèle correspond à la traversée du pore nucléaire par la capsidite intacte ou partiellement intacte qui se désassemble à l'intérieur du nucléoplasme (92, 93, 120, 125–128).

Plusieurs travaux récents sont en faveur de ce modèle. Premièrement, la capsid intervient dans les étapes post-entrée intranucléaire (129–132). Deuxièmement, les facteurs d'interactions nécessaires à l'import nucléaire de la capsid se fixent aux hexamères de CA ou aux poches hydrophobes formés par les hexamères de CA et non aux monomères. La capsid doit alors être en partie intacte ou une partie des capsides qui sont entrées dans le noyau doit l'être (133, 134). Troisièmement, un remodelage des protéines de capsid, en collier de perles autour de l'ADN viral, a été mis en évidence par le groupe de F. Di Nunzio (93). Il a aussi été montré, par des études de cryo-microscopie électronique, que le linker des domaines N-terminal et C-terminal des monomères sont flexibles permettant aux hexamères et donc à la capsid de se courber (87, 135). Enfin, des données cristallographiques démontrent que les liaisons hydrogènes médiées par l'eau au niveau des chaînes latérales de certains résidus, important notamment pour les interactions CA-CA, rendent le réseau hexamérique de la capsid flexible (136). De plus, une flexibilité structurelle de la taille des pores nucléaires a été décrite suggérant que le passage d'une capsid assemblée ou partiellement assemblée est possible (92, 136–140). Celle-ci serait modulée par des signaux allostériques tels que le processus de transcription inverse ou la fixation à la capsid de co-facteurs (95, 141). À noter, que les dNTPs cellulaires seraient capables d'entrer via le canal des hexamères de la protéine de capsid permettant ainsi la transcription inverse tout en excluant les molécules de tailles supérieures au diamètre du canal d'hexamères (142, 143).

Le groupe de Pathak grâce à un marquage de la CA avec la GFP et une détection par imagerie des cellules vivantes (*live-cell imaging*) a publié (132) que la capsid intacte ou quasi intacte pénètre dans le noyau et que la décapsidation a lieu près du site d'intégration (144). L'équipe de Campbell a également publié un article dans ce sens, et a de plus montré que la transcription inverse se termine dans le noyau (145). Finalement, il y a de plus en plus d'études publiées sur la décapsidation nucléaire. L'entrée nucléaire de la capsid est permise par une décapsidation partielle ou une plasticité du pore nucléaire ou les deux à la fois. Les modèles de décapsidation dans le cytoplasme, au niveau du pore nucléaire ou dans le noyau ne sont pas forcément contradictoires mais peuvent être complémentaires selon le type cellulaire considéré.

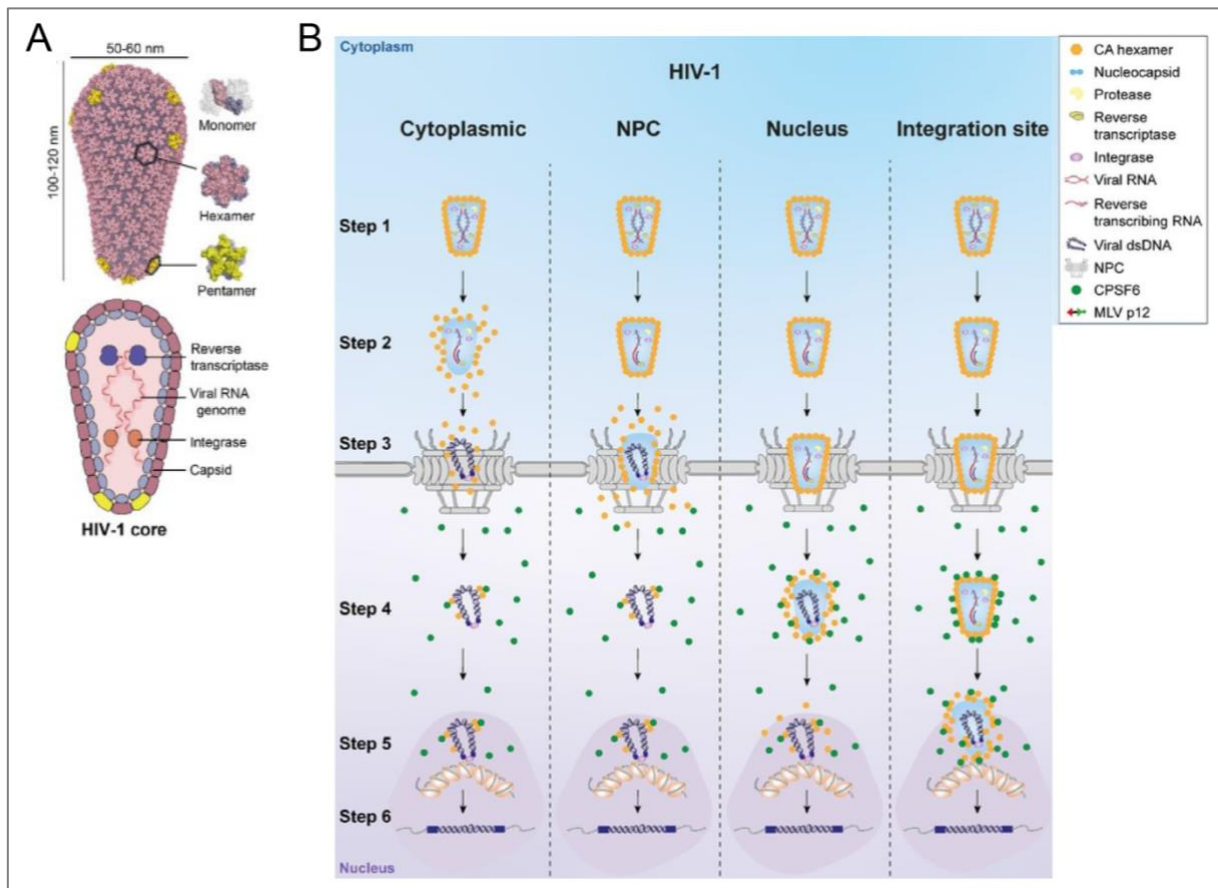


Figure 13. Schéma d'une capsid et des quatre modèles de décapsidation. (A) Capsid du VIH-1 assemblée à partir de hexamères et de pentamères de protéines de capsid. L'ARN viral, l'intégrase et la transcriptase inverse y sont logés (94). **(B)** Sont schématisés : la capsid (en orange) du VIH-1 contenant l'ARN viral (en rouge) recouvert de protéine de NC (en bleu clair), des enzymes virales (transcriptase inverse en vert, protéase en jaune, intégrase en violet) (étape 1). La transcription inverse commence et la capsid se déplace vers le noyau grâce aux microtubules de la cellule infectée (étape 2). L'ADN proviral peut traverser le pore nucléaire (étape 3), seul dans le modèle de décapsidation cytoplasmique (première colonne) ou dans la capsid partiellement décapsidée dans le modèle de décapsidation NPC (seconde colonne), ou dans la capsid intacte dans les modèles de décapsidation dans le noyau (troisième colonne) ou près du site d'intégration (quatrième colonne). L'ADN en association avec CPSF6 (en vert) se déplace vers les sites d'intégration (étape 4), s'intègre dans la chromatine (étape 5) pour former le provirus (étape 6) (146).

c. Lien entre transcription inverse et décapsidation

La capsid est une coque protectrice du matériel génétique du virus et du processus de la transcription inverse. Plusieurs études suggèrent que la décapsidation est initiée par la transcription inverse de l'ARNg (84, 90, 124, 147, 148). La transcription inverse commencerait dans le cytoplasme, dans des capsides intactes ou quasiment intactes (82, 127, 128, 132, 146), et se terminerait dans le noyau (122, 125, 127, 128, 132, 145, 149–151). Depuis de nombreuses années, des études rapportent que la transcription inverse est liée à la stabilité de la capsid.

Lorsque la stabilité de la capsid est fortement augmentée, la transcription inverse ne se produit pas (152). De plus, lorsque la transcription inverse est bloquée par des inhibiteurs de la transcriptase inverse ou des mutations de cette enzyme, la décapsidation est retardée (84, 115) suggérant qu'un certain degré de transcription inverse est nécessaire pour réaliser la décapsidation. Finalement, la décapsidation semble être fortement dépendante de la transcription inverse (84, 90, 99, 115, 116, 152–154).

Les premières études par vitesse de sédimentation avaient montré la présence de protéine de CA dans le RTC quelques heures après l'infection et une diminution significative plus de 16 h après. Plus récemment, des études par imageries ont révélé la présence de protéines de CA dans les RTC pendant la transcription inverse et des capsides intactes près de l'enveloppe nucléaire (Références cités dans la revue (85)). Un consensus a été établi définissant le RTC comme une capsid en cours de désassemblage composée des protéines de CA, de la RT, de l'intégrase, de Vpr et d'ARNg sur lequel est fixé des NC (83). Des expériences biochimiques ont montré que des capsides isolées effectuent une transcription inverse plus efficacement en présence d'IP₆, incorporé dans les virions, et se désassemblent avec une cinétique plus lente que les capsides en absence d'IP₆ (155, 156). Ces résultats sont en accord avec l'idée qu'une capsid intacte facilite la transcription inverse. *In vitro*, il a cependant été montré que la transcription inverse en présence d'IP₆ génère différentes morphologies de noyau de la capsid, allant d'intacte à 50% désassemblée (82). Ces résultats permettent de conclure que la décapsidation n'est pas un processus de « tout ou rien », mais peut se produire partiellement de façon progressive.

Une décapsidation partielle permettrait l'afflux de dNTPs et par conséquent la réalisation de la transcription inverse qui induirait une modification de la structure de la capsid telle qu'elle a été observée lors des études par vitesse de sédimentation et par imagerie. Cette hypothèse est étayée par une étude récente de Rankovic et al. qui suggère que certaines étapes de la transcription inverse induisent des changements mécaniques dans la capsid, la remodelant progressivement, par étapes, afin d'induire son désassemblage (141). Une décapsidation progressive permettrait également d'empêcher la dégradation des acides nucléiques viraux par les nucléases cellulaires bien qu'ils existent d'autres mécanismes comme l'action de SAMHD1. D'autres éléments vont dans ce sens comme une transcription inverse capable de se poursuivre dans le noyau (127, 150) et plusieurs moyens d'échappement au système immunitaire de la cellule, préservant l'intégrité des acides nucléiques viraux (85).

2.3.3 Intégration du provirus

L'ADN bicaténaire linéaire néosynthétisé est intégré dans le génome de la cellule infectée afin d'exprimer les gènes viraux et produire de nouvelles particules virales. L'intégration de l'ADN double-brin, caractéristique des rétrovirus, permet d'avoir une copie du génome viral dans chaque cellule fille lors de divisions cellulaires. Il peut aussi permettre au virus de persister dans l'organisme en état de dormance. L'ADN du VIH s'intègre préférentiellement dans des gènes transcriptionnellement actifs (157).

L'intégration est catalysée par l'intégrase (IN) virale et nécessite trois réactions : la maturation des extrémités 3' (*3'-end processing*), le transfert de brin réalisé par l'intégrase et la réparation du site d'intégration. L'intégrase est présente aux deux extrémités de l'ADN viral correspondant aux LTR. Ensemble ils forment l'intasome, sous-structure du complexe de pré-intégration (PIC) (158, 159). Ce dernier est défini comme un complexe nucléoprotéique d'ADN viral associé à des protéines virales telles que la RT, l'IN, la MA, la CA, la NC et Vpr ainsi que des facteurs cellulaires comme LEDGF, INI 1 ou HMG I (160–162). Il est formé après la rétro-transcription du génome viral afin de réaliser le processus d'intégration. Le génome viral intégré est appelé « provirus ».

L'intégrase possède deux activités : la maturation des extrémités 3' et le transfert de brin. La première étape du processus d'intégration, celle de la maturation des extrémités 3' consiste à hydrolyser la liaison phosphodiester à chaque extrémités 3' de l'ADN viral éliminant ainsi le dinucléotide 5'-GT-3' et exposant le groupe hydroxyle du dinucléotide conservé 5'-CA-3'(163) (Figure 14). Puis la seconde étape est le transfert de brin. L'intégrase catalyse l'insertion covalente des extrémités 3' de l'ADN viral dans l'ADN chromosomique.

Plus précisément, l'intégrase réalise une attaque nucléophile des liaisons phosphodiester pour cliver simultanément les deux brins de l'ADN cellulaire en utilisant le groupement 3' hydroxyle de l'ADN viral comme nucléophile (164, 165). Cela permet la formation d'une liaison covalente entre les extrémités 3' de l'ADN viral et l'ADN cible (162) (Figure 14). La dernière étape est la réparation par des enzymes cellulaires des deux sites encadrant l'ADN intégré.

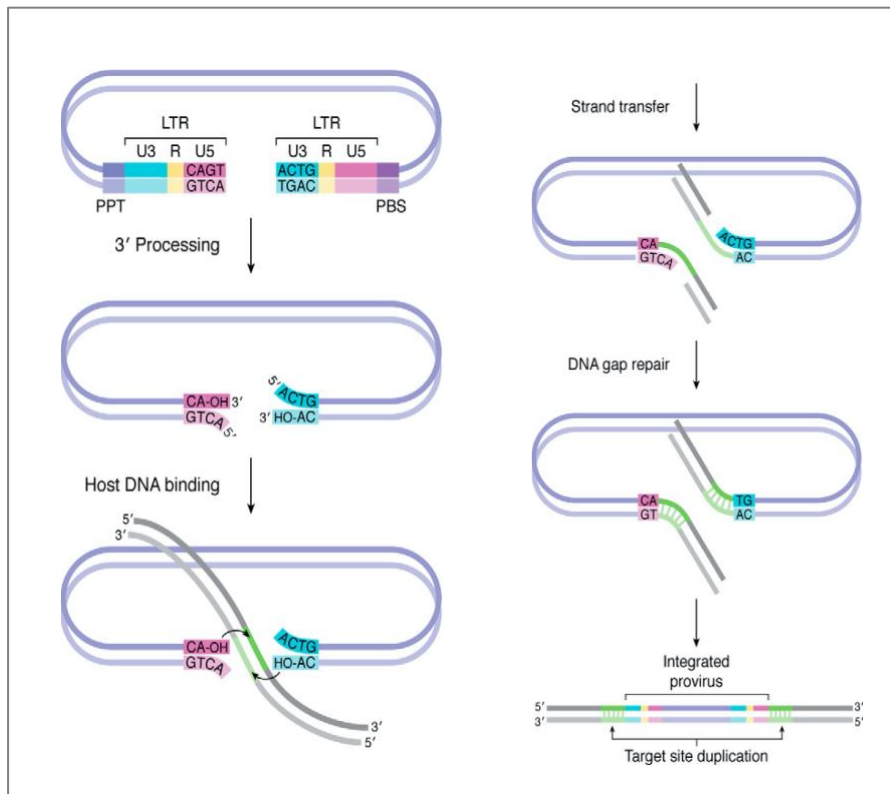


Figure 14. Processus d'intégration (166). Le processus d'intégration débute avec la maturation des extrémités 3' puis l'insertion de l'ADN viral dans l'ADN cellulaire et se termine avec la réparation des sites encadrant l'ADN viral.

Lors d'une infection naturelle au VIH-1, l'ADN linéaire connaît plusieurs destins. Il peut être intégré ou dégradé si celui-ci n'est pas rapidement intégré. Des formes non intégrées d'ADN viral sont également formées pendant la transcription inverse (167–169) donnant naissance à des formes circulaires (Figure 15) (167) (170). Parmi elles, des cercles d'ADN ne contenant pas la longueur totale de l'ADN d'origine ou des cercles d'ADN réarrangés issus d'une auto-intégration catalysée par l'intégrase (167).

Il y a aussi la formation d'ADN circulaire à 1-LTR et 2-LTR. Les circulaires à 1-LTR sont principalement formés dans le noyau bien qu'une faible proportion, environ 10%, puisse être formée dans le cytoplasme (171). Ils peuvent être générés par des recombinaisons homologues d'ADN linéaire au niveau des LTR. Les facteurs cellulaires de l'hôte peuvent également conduire à la formation de ces formes circulaires à 1-LTR ainsi qu'aux formes à 2-LTR grâce à la voie de réparation de l'ADN *non-homologous end joining* (NHEJ) qui est une réponse de la cellule hôte à la présence d'ADN double brin dans le noyau (171, 172).

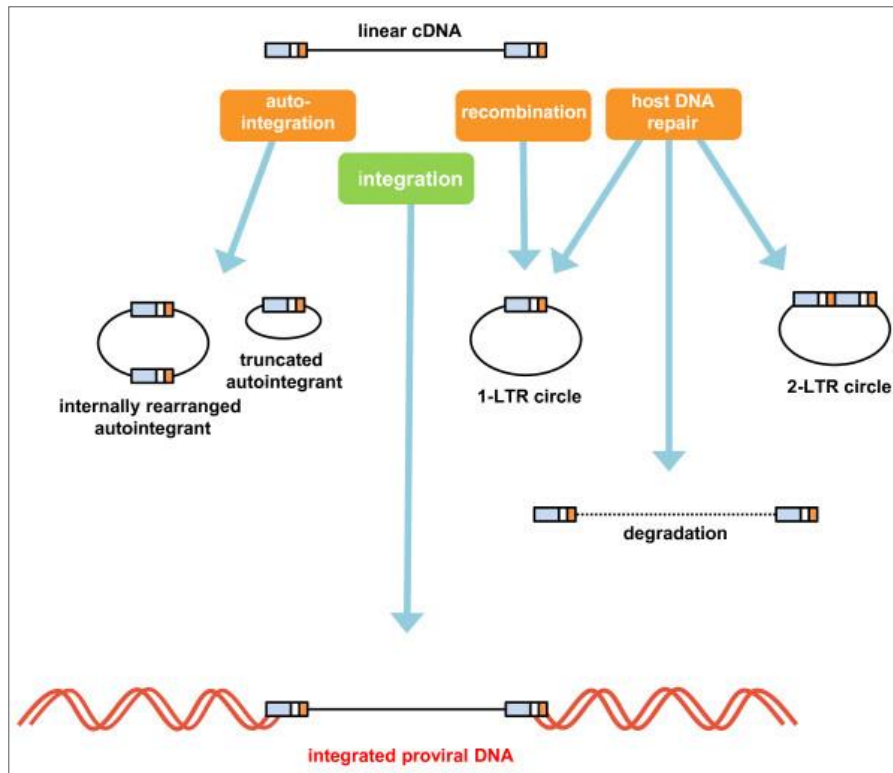


Figure 15. Différentes formes non intégrées d'ADN viral (167). L'ADN viral linéaire est la principale forme susceptible d'être intégrée. L'auto-intégration de cette forme linéaire peut conduire à la formation d'une forme circulaire tronquée non intégrée ou réarrangée. Par la voie de *non-homologous end joining*, les formes circulaires à 1-LTR et 2-LTR sont formées. La recombinaison homologue des LTR peut aboutir à la formation d'une forme circulaire à 1-LTR.

Le rôle de ces formes circulaires n'est actuellement pas établi. Des études suggèrent que les formes à 1-LTR et 2-LTR seraient responsables d'une faible transcription des gènes accessoires tels que Nef (167). Deux équipes ont montré que la production virale pouvait être détectée à partir de ces formes dans des LT CD4+ quiescents après leur infection, suggérant un rôle dans le processus de réplication des 1 et 2-LTR (173, 174). Il a été montré que l'ADN à 2-LTR peut être utilisé comme une réserve de génomes pour l'intégration provirale (173, 175, 176), suggérant que l'intégrase est capable de cliver la jonction entre les 2 LTR. Cette jonction possède d'ailleurs une séquence palindromique capable d'être clivée *in vitro* par des intégrases recombinantes du VIH-1 conduisant à la formation d'ADN linéaire mûré aux extrémités 3' (175, 177).

Plus récemment, l'équipe d'O. Delelis a mis en évidence dans un contexte virologique que l'intégrase catalyse l'intégration d'une forme circulaire à 2-LTR dans l'ADN génomique (178). L'ADN circulaire à 2-LTR est donc probablement une autre source de génome rétroviral pouvant être intégré (178).

2.3.4 Synthèse, maturation et export des ARN viraux

La phase tardive du cycle répliatif du VIH-1 débute avec la transcription du provirus en ARNm par l'ADN polymérase II de la cellule-hôte. Les transcrits subissent des modifications post-transcriptionnelles similaires aux pré-ARNm de la cellule infectée comprenant l'ajout de la coiffe en 5', le clivage avec la polyadénylation de l'extrémité 3' et l'épissage par les machineries cellulaires.

Le signal de polyadénylation est présent dans la séquence R située aux deux extrémités du génome (Figure 9). Le signal poly(A) correspond à un motif hautement conservé « AAUAAA », une région riche en GU, un site de clivage ainsi que des séquences activatrices ou inhibitrices (USE : *upstream sequence element*) (179). Les rôles de la coiffe 5' et la queue poly(A) sont multiples. Ils permettent notamment l'export des ARNm du noyau vers le cytoplasme et stabilisent les ARNm en évitant leur dégradation par des ribonucléases (180).

Afin de produire ses différentes protéines virales tout en ayant une taille de génome pouvant être contenu dans une capsid, la majorité des souches du VIH-1 présente cinq sites donneurs d'épissage et huit sites accepteurs. Ces derniers peuvent produire plus d'une centaine d'isoformes d'ARN dans les cellules infectées (Figure 16) (181) (182).

Différentes tailles de transcrits sont générées :

- Les ARNm d'environ 1kb complètement épissés et identifiés en 2012 par l'équipe de F. Bushman (183). Ces ARN pourraient coder pour des protéines de fusion ou des ARNs régulateurs (183, 184).
- Les ARNm faisant entre 1,8 et 2 kb qui codent les protéines virales Tat, Rev et Nef selon le site accepteur d'épissage (185) (Figure 16). Ces ARNm sont constitutivement exportés vers le cytoplasme via la voie TAP/NXF1 et traduits de la même façon que les ARNm cellulaires (59, 186). Ces transcrits n'utilisent pas le système d'export dépendant de la protéine virale Rev.
- Les ARNm de 4 kb produisant les protéines Vif, Vpr, une forme à un exon de Tat et l'ARNm bi-cistronique Env-Vpu non épissés ou partiellement (181).
- Les ARNm de 9 kb non épissés servant d'ARN génomique pour les nouvelles particules virales mais codant également pour Gag et Gag-Pol (181).

Les deux dernières classes d'ARNm (non épissés de 9 kb et partiellement de 4 kb), obtenues lors d'une phase plus tardive que les ARNm entièrement épissés, sont exportées vers le cytoplasme par la protéine virale Rev interagissant avec l'élément structural RRE présent sur ces ARNm et des facteurs cellulaires (187).

2.3.5 Synthèse des protéines virales

L'expression des gènes viraux peut être séparée en deux phases : la phase précoce (indépendante de Rev) et la phase tardive (dépendante de Rev). La phase précoce correspond à l'expression des protéines issues des transcrits de 2 kb (complètement épissés et exportés de façon constitutive vers le cytosol pour y être traduit) codant pour Tat, Rev et Nef. Le taux de transcription virale est fortement augmenté en présence de la protéine trans-activatrice Tat qui se fixe sur l'élément structural TAR de la région R en 5' influençant la processivité de l'ARN polymérase II cellulaire (189, 190). Grâce à cette interaction et au recrutement de facteurs cellulaires, Tat joue un rôle important dans l'expression des gènes de la phase tardive qui sont les transcrits non épissés ou partiellement. Le virus utilise donc l'épissage alternatif pour exprimer tous les gènes qui lui sont nécessaires tout en limitant la taille de son génome. Rev permet l'export des ARNm non épissés ou partiellement épissés. Nef joue un rôle dans l'échappement à la réponse immunitaire en « masquant » la cellule infectée. Une expression en deux temps des protéines virales peut s'expliquer par le fait que les protéines structurales sont plus immunogènes. Il est donc plus intéressant pour le virus de les produire après l'échappement au système immunitaire grâce à l'action de Nef. S'en suit une phase tardive correspondant à l'expression des autres protéines virales (plus immunogènes) produites à partir des transcrits non épissés (9 kb) ou partiellement épissés (4 kb) dont le transport vers le cytoplasme dépend de Rev.

a. La glycoprotéine de l'enveloppe

Le gène *env* code pour un précurseur polypeptidique glycosylé de 160 kDa (gp160). Le clivage de ce dernier par la furine, une protéase cellulaire (191), produit les glycoprotéines gp120 et gp41 (192). Celles-ci forment la protéine d'enveloppe (Env) constituée d'un trimère de gp120 formant le spicule présent à la surface du virion et d'un trimère de gp41 formant la tige transmembranaire (Figures 11 et 12). Les trois sous-unités gp120 sont liées par des interactions non covalentes aux trois sous-unités de gp41. La protéine d'enveloppe est essentielle à l'infection via la fusion membranaire (191). Elle représente donc la première et principale cible de la réponse humorale car c'est le seul antigène à la surface de la particule virale.

b. Les protéines de régulation

Tat (*Transactivator of transcription*) est une protéine de 86 à 104 acides aminés selon la souche (193) et est codée par deux exons dans le génome viral. Elle est synthétisée à partir d'ARNm partiellement ou complètement épissés (Figure 16). Elle est localisée principalement dans le noyau mais peut aussi être libérée dans l'espace extracellulaire par les cellules infectées (194). Cette protéine virale est l'un des deux facteurs de régulation viraux qui sont essentiels à l'expression du virus. Cette protéine est impliquée dans plusieurs mécanismes moléculaires ayant pour but d'augmenter le taux de transcription de l'ADN proviral. Elle a notamment un effet sur le remodelage de la chromatine (195), l'initiation et l'élongation de la transcription (196, 197) et la phosphorylation de l'ARN polymérase II (198). Plus précisément, Tat possède plusieurs domaines (Figure 17) (193) dont un domaine basique grâce auquel elle peut interagir avec l'élément structural TAR présent dans la région R située en 5' de l'ARNg.

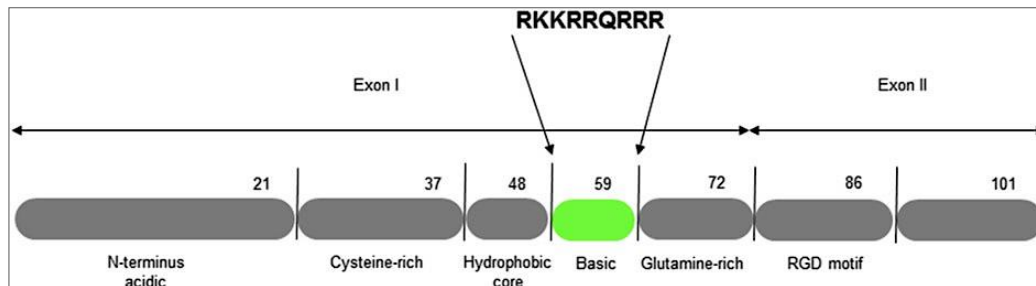


Figure 17. Domaines de la protéine Tat. Tat contient les domaines suivants : domaine N-terminal (1-21), domaine riche en cystéine (22-37), domaine noyau hydrophobe (38-47), domaine basique (48-59), domaine riche en glutamine (60-72) et domaines codés par le second exon.

Cette interaction facilite la transcription virale (189, 199, 200) par le recrutement de la forme active du *positive transcription elongation factor b* (P-TEFb) composé de la cycline T1 et la *cyclin-dependent kinase 9* (Cdk9). Tat se lie à la cycline T1 et à la tige-boucle TAR formant un complexe ternaire TAR:Tat:complexe P-TEFb sur les ARN naissants. Les changements de conformation du complexe P-TEFb induit par Tat permet à Cdk9 de phosphoryler l'ARN polymérase II (201–203) (Figure 18) (204). Cela rend ainsi l'enzyme plus stable, plus processive et capable de transcrire l'intégralité de l'ADN proviral (204).

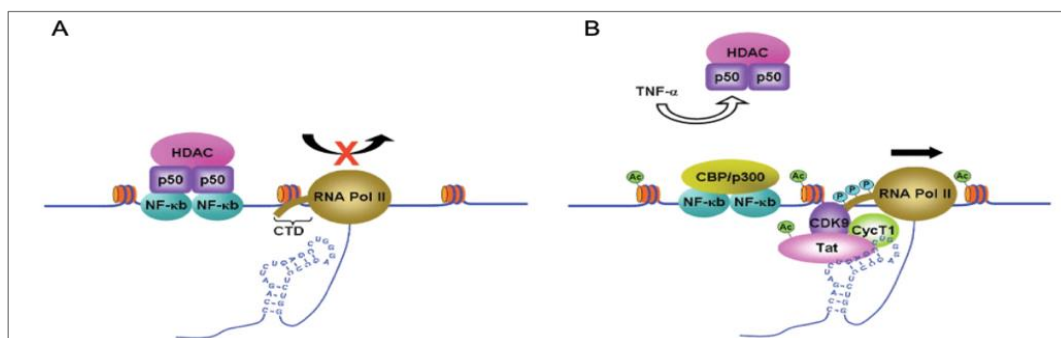


Figure 18. Transactivation par Tat de l'expression virale. (A) Lorsque le complexe NF-κB:p50:p50:HDAC est lié au LTR du virus latent, une déacétylation des histones et des changements de la structure de la chromatine ont lieu (205). En absence de la protéine Tat, l'ARN polymérase II initie la synthèse des transcrits viraux sans parvenir dans la majorité des cas à réaliser l'étape d'élongation nécessaire à la synthèse de l'ARN viral complet. (B) En présence d'une stimulation comme le TNF- α , les complexes induisant une latence du virus sont éliminés du LTR viral. Le complexe CBP/p300 se lie au promoteur et acétyle les histones environnantes ainsi que Tat (206, 207). Tat interagit avec le complexe P-TEFb et se fixe sur la tige-boucle TAR de l'ARN viral. S'ensuit une phosphorylation de l'ARN polymérase II à son domaine C-terminal afin de la rendre plus efficace et capable de synthétiser intégralement l'ARN viral.

Rev (*Regulator of virion expression*) est une phosphoprotéine de 116 acides aminés (19 kDa) qui est synthétisée à partir d'ARN viraux entièrement épissés (Figure 16). Rev est le second facteur de régulation essentiel pour l'expression du génome viral. Cette protéine se fixe spécifiquement sur l'élément structural RRE (*Rev responsive element*) situé dans le gène *env* (Figure 9). Plusieurs molécules de Rev se fixent sur RRE en formant un complexe homo-oligomérique (208). Ce complexe interagit avec la machinerie d'exportation de la cellule-hôte et permet ainsi le transport des ARN viraux non épissés ou partiellement épissés du noyau vers le cytoplasme (63, 188).

c. Les protéines accessoires

Nef (*Negative regulatory factor*) est l'une des quatre protéines accessoires du VIH-1. Elle est constituée de 207 acides aminés et est myristoylée à son extrémité N-terminale (209). Elle est produite à partir d'ARN viraux entièrement épissés (Figure 16). Nef est principalement localisée dans le cytoplasme et la membrane plasmique de la cellule infectée. Cette protéine possède un noyau structuré, une boucle interne et des régions flexibles aux extrémités C et N-terminales. Ces régions peu structurées sont capables de subir des changements conformationnels permettant à Nef d'avoir de nombreuses interactions avec les protéines cellulaires (210, 211). Nef a plusieurs activités dont l'une est de réguler négativement l'expression du récepteur CD4 (212) ainsi que des molécules du CMH de classe I des cellules

infectées (213). Nef est également capable d'inhiber l'incorporation des protéines transmembranaires SERINCs dans les virions (214, 215). Les facteurs de restriction SERINCs influenceraient la taille du pore de fusion membranaire entre le virion et la cellule, empêchant la capsidation d'entrer. Enfin, nef est décrite comme un facteur de pathogénicité requis pour une réplication optimale du virus (216).

Vpr (*Viral protein R*) est une protéine de 96 acides aminés produite à partir d'ARN viraux partiellement ou complètement épissés (Figure 16). Cette protéine est incorporée dans les particules virales grâce à une interaction spécifique avec le domaine p6 du précurseur Gag (217). Vpr est présente dans les noyaux des cellules infectées et est impliquée dans l'importation nucléaire du PIC. Cette protéine accessoire interagit avec des protéines de la cellule hôte afin d'induire leur dégradation ou pour moduler leurs activités (218). Les molécules cibles de Vpr jouent notamment un rôle dans le cycle cellulaire, la différenciation cellulaire, l'apoptose ou encore la réparation de l'ADN cellulaire. Vpr se trouve être aussi impliqué dans la transactivation des gènes viraux et cellulaires, et dans la régulation de l'activité de $\text{Nf-}\kappa\text{B}$ (216, 217). Vpr peut également induire la dégradation du facteur de restriction SAMHD1, régulateur majeur de l'homéostasie des nucléotides (220).

Vpu (*Viral protein U*) est une protéine accessoire de 81 acides aminés propre au VIH-1 et à certains SIV dont le SIVcpz (le plus proche du VIH-1). Il n'existe pas une protéine similaire chez le VIH-2. Elle est générée à partir de transcrits partiellement épissés (Figure 16). Cette protéine transmembranaire possède deux principales fonctions biologiques. Une première fonction qui est la régulation négative de l'expression des récepteurs CD4 en modulant la dégradation des CD4 nouvellement synthétisés par le protéasome dans le réticulum endoplasmique (221). Et une seconde fonction qui est la stimulation du relargage des virions à partir de la membrane plasmique des cellules infectées par le VIH-1 en contrant l'effet du facteur de restriction BST2. Cette protéine cellulaire est capable de retenir physiquement les nouveaux virus formés à la surface des cellules infectées en s'incorporant dans la membrane des particules virales (222, 223).

Vif (*Viral infectivity factor*) est une protéine accessoire basique constituée de 220 acides aminés. Elle est produite à partir d'ARN viraux partiellement épissés (Figure 16) et est localisée dans le cytoplasme des cellules infectées ou à l'intérieur des particules virales (224, 225).

Vif favorise l'infectiosité. Son rôle principal est de contrer l'effet antiviral de la protéine cellulaire APOBEC3G (225, 226). Cette dernière est capable d'induire la conversion de cytosine en uracile dans l'ADN simple-brin négatif, induisant ainsi des mutations de guanine en adénine dans l'ADN viral double-brin. De façon très simplifiée, vif empêche l'incorporation d'APOBEC3G dans les virions en phase de bourgeonnement en déclenchant sa dégradation par le protéasome via la voie d'ubiquitination (227, 228). Vif peut également inhiber la transcription d'APOBEG3G par la fixation du facteur de transcription CBF- β (*core binding factor* β), sa traduction en se liant aux UTR de ses ARNm, son encapsidation dans les particules virales en cours de bourgeonnement dont le mécanisme n'est pas encore clairement établi. (225, 229–232) Enfin, d'autres études suggèrent que Vif inhibe l'activité désaminase dans les virions (225, 233–235).

d. Les enzymes virales

- La protéase

La protéase (PR) est générée à partir du précurseur Gag-Pol (Figure 9) qui est synthétisé à partir de l'ARNm 9 kb non épissé (Figure 16). La dimérisation du précurseur Gag-Pol induit l'activation de la protéase située dans le domaine Pol et permet ainsi la libération de la protéase mature (236–238). Cette dernière est active sous la forme d'un dimère dont chaque sous-unité est constituée de 99 acides aminés (239). Elle fait partie de la famille des protéases à acide aspartique (acide aminé présent dans le site actif avec deux autres acides aminés formant la triade catalytique). La protéase en réalisant les clivages protéolytiques des précurseurs polypeptidiques Gag et Gag-Pol joue un rôle essentiel dans la maturation des particules virales (224, 228, 229). Pour cette raison, les inhibiteurs qui ciblent la protéase sont utilisés depuis 1995 comme médicaments dans le traitement des personnes infectées par le VIH (229, 230).

- L'intégrase

L'intégrase est une enzyme essentielle pour l'intégration de l'ADN viral dans le génome cellulaire en réalisant la maturation des extrémités 3' et le transfert de brin. L'intégrase est une polynucléotidyle transférase de 288 acides aminés possédant trois domaines structurellement indépendants : le domaine amino-terminal (NTD) arborant un motif conservé HisHisCysCys liant le zinc, le domaine du cœur catalytique (CCD), comprenant la triade catalytique Asp-Asp35-Glu spécifique des intégrases rétrovirales et rétrotransposons, et le domaine carboxy-terminal (CTD) (243).

L'intégrase est en équilibre entre les états monomérique, dimérique, tétramérique et d'autres états oligomériques (166). Elle se lie à l'ADN uniquement sous sa forme dimérique qui représente la forme active de l'enzyme (244). Deux protéines cellulaires ont été caractérisées comme intervenant dans l'import du PIC. Le complexe IN-ADN est transporté vers le noyau par la transporine-SR2 (TR2) (245). LEDGF/p75 est un autre cofacteur cellulaire de l'intégrase du VIH-1 qui interagit avec elle par son *IN binding domain* (IBD) et fixe le PIC à la chromatine de la cellule hôte (246, 247) (Figure 19).

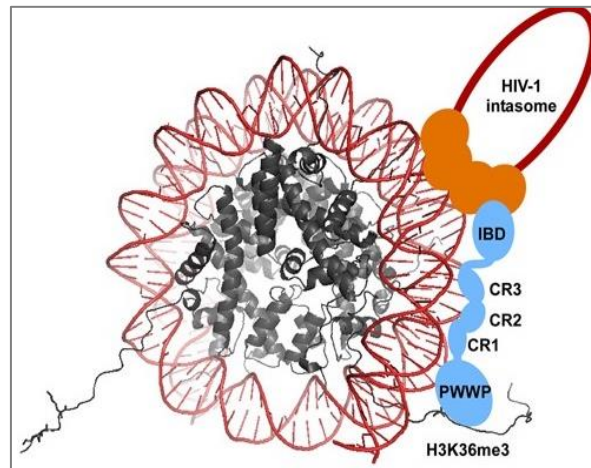


Figure 19. Interactions entre LEDGF/p75, l'intasome et l'ADN cellulaire (246). L'ADN cellulaire (trait fin rouge) est enroulé autour des histones (en gris) formant un mononucléosome. LEDGF/p75 (en bleu) se fixe sur ce dernier grâce à une liaison entre son domaine PWWP et la queue de l'histone H3 et ses trois régions chargées (CR1-3) avec l'ADN cellulaire. Le *IN binding domain* (IBD) permet au LEDGF/p75 de se lier à l'intégrase (tétramère en orange), elle-même liée à l'ADN viral (trait épais rouge).

- *La transcriptase inverse*

La transcriptase inverse (RT) (Figure 20) (248, 249) est un hétérodimère généré à partir du clivage du précurseur Gag-Pol par la protéase virale. Elle possède deux sous-unités : une grande, p66 et une plus petite, p51 qui est un produit protéolytique de la p66 (250).

La sous-unité p66 possède deux domaines :

- le domaine N-terminal contenant l'activité ADN polymérase ADN-ARN dépendante responsable de la conversion de l'ARNg viral simple-brin en ADN double-brin. Sa structure tridimensionnelle est comparée à une main droite composée de sous-domaines nommés : *Fingers*, *Palm* et *Thumb* (227, 239). Le sous-domaine *Palm* possède l'activité polymérase avec les trois acides aspartiques formant le site actif (D110, D185 et D186). Les dNTP se fixent entre les sous-domaines *Palm* et *Fingers* et d'importants contacts se produisent entre les riboses et les bases azotées avec les résidus L74, Y115, M184 et Q151. La poche catalytique est formée par le sous-domaine *Fingers* qui se rabat sur le sous-domaine *Palm* (252).

- le domaine C-terminal qui contient l'activité RNase H qui dégrade le brin d'ARN dans un hybride ADN-ARN. Il est constitué des 120 derniers acides aminés de la sous-unité p66, correspondant au fragment p15 généré par le clivage de la sous-unité p66 par la protéase virale. Le domaine RNase H est à proximité du domaine catalytique lié par le domaine de connexion au sous-domaine *Thumb*. Le clivage de p66 par la protéase virale génère la sous-unité p51 de la RT et p15. La p51, à l'exception du domaine RNaseH, contient les mêmes sous domaines que p66. Néanmoins, elle ne présente pas d'activité catalytique comme p66, présentant une structure ouverte et flexible, et jouerait davantage un rôle structural malgré une séquence d'acides aminés similaire (253).

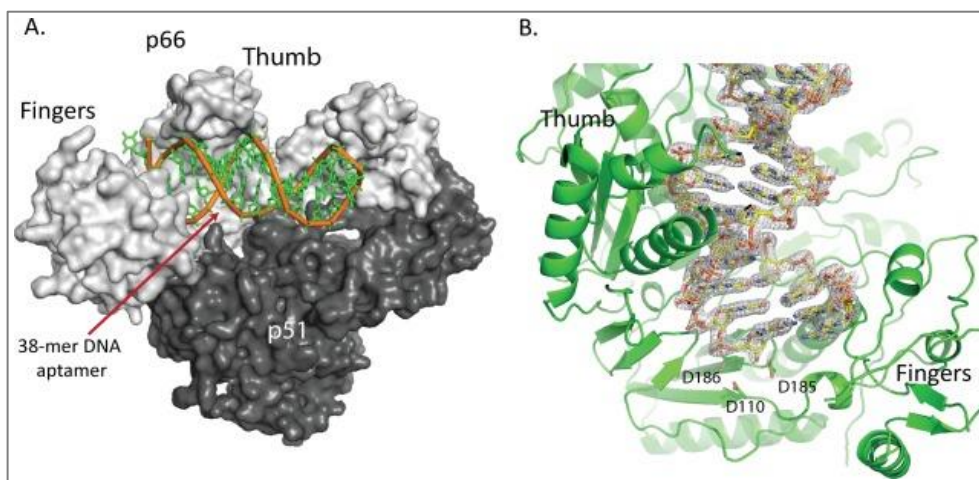


Figure 20. Structure de la transcriptase inverse du VIH-1 liée à un aptamère d'ADN (248). (A) Hétérodimère p66-p51 (gris) et un aptamère d'ADN (vert). La première paire de bases est située au niveau du site actif de l'activité polymérase et la boucle apicale de l'aptamère est située près du site actif de l'activité RNase H. (B) Représentation de la densité électronique de l'aptamère d'ADN au centre du site actif de l'activité polymérase.

Propriétés de la transcriptase inverse

Les deux activités de la RT (activité ADN polymérase et activité RNase H) sont nécessaires et coopèrent pour convertir le génome viral d'ARN simple-brin en ADN double-brin. La RT ne possède pas d'activité de relecture exonucléotidique permettant à l'enzyme de vérifier que le transcrit est l'exact copie de la matrice ADN ou ARN. Son taux d'erreur a été mesuré en culture cellulaire et est d'environ $1,4$ à $2,4 \cdot 10^{-5}$ par base et par cycle de réplication (254, 255). Cette estimation peut ne pas refléter le réel taux de mutations du VIH-1 *in vivo* car plusieurs facteurs cellulaires entrent en compte chez les patients atteints, notamment la concentration de dNTP, le taux de co-infection cellulaire avec des génomes viraux distincts ou la taille de la population virale, qui peut alors influencer le type de mutation et la fréquence.

Cuevas *et al* (256) ont mesuré *in vivo* le taux de mutation en analysant la fréquence des codons stop trouvés dans les séquences virales des patients. Le taux est de $4,1 \pm 1,7 \cdot 10^{-3}$ mutations par base et par cellule dans les séquences d'ADN des cellules mononucléaires du sang périphériques de patients. Le virus est également soumis aux facteurs cellulaires antiviraux tel que APOBEC3. Cette même étude a révélé que 98% des mutations identifiées ont été générées par les protéines APOBEC3 et non par la RT. Néanmoins, l'hypermutation induite par APOBEC3 est généralement létale pour le virus (86, 257, 258).

L'activité ADN polymérase :

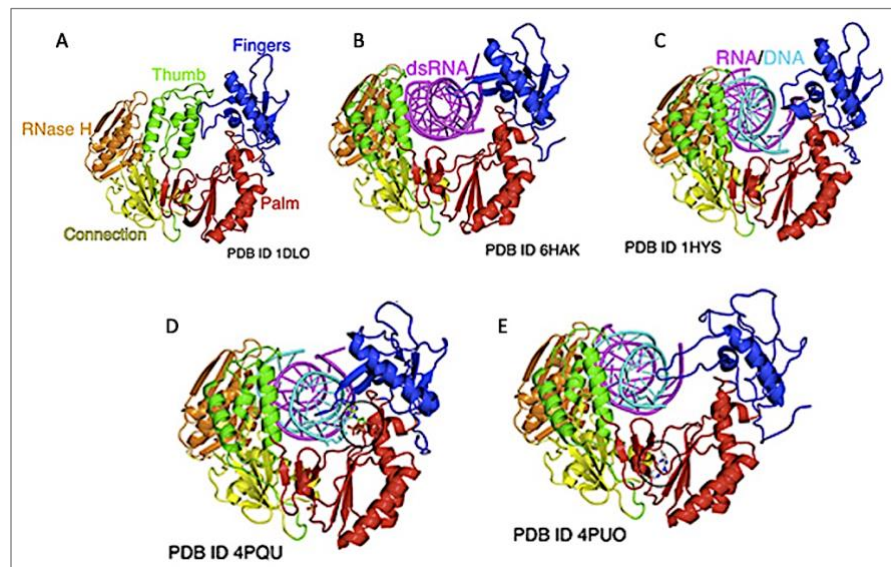


Figure 21. États conformationnels de la transcriptase inverse du VIH-1 (259). (A) Forme Apo de la RT où le domaine *Thumb* est fermé et le domaine *Fingers* ouvert. (B) Forme complexée de la RT avec l'ARN viral et l'amorce ARNt. (C) RT complexée avec l'ARN viral et l'ADN viral néosynthétisé. Sous ces deux conformations, la RT peut incorporer un dNTP où le domaine *Thumb* est maintenant ouvert. (D) Un dNTP s'est fixé au site N de la RT sur lequel le domaine *Fingers* s'est rabattu. (E) La liaison phosphodiester se forme entre l'amorce et le nucléoside, le domaine *Fingers* s'ouvre de nouveau.

La fixation de la RT sur l'ARNg viral induit un changement conformationnel du sous-domaine *Thumb* de la sous-unité p66, d'une conformation fermée à une conformation ouverte (Figure 21 A, B et C). Comme d'autres ADN polymérases, la RT a besoin à la fois d'une amorce et d'une matrice pour démarrer la polymérisation. L'enzyme se fixe préférentiellement sur les acides nucléiques double-brin. Ainsi, l'extrémité 3' de l'amorce appariée à la matrice interagit avec le site P (*Priming site*) de la RT qui est adjacent au site actif de l'activité polymérase. L'étape initiale d'incorporation nucléotidique est la fixation des dNTP entrants au niveau du site N (*Nucleotide binding site*) de l'enzyme formant alors un complexe ternaire (Figure 21 D et E).

La limitation de la cinétique de la réaction de polymérisation est due à un changement conformationnel du sous-domaine *Fingers* qui se rabat sur le dNTP entrant et le site actif de polymérisation (Figure 21 D), permettant d'aligner le 3'-OH de l'amorce, le phosphate α du dNTP et le site actif de la polymérase. A lieu ensuite la formation d'un pont phosphodiester entre le nucléoside entrant et l'amorce, générant alors un pyrophosphate. Le sous-domaine *Fingers* s'ouvre (Figure 21 E) et conduit le pyrophosphate hors du site actif. À mesure de la synthèse du brin d'ADN, la matrice est transloquée de façon à libérer le site de liaison nucléotidique pour que l'enzyme fixe le prochain dNTP entrant (Figure 21 C). La polymérisation nécessite la présence de deux ions Mg^{2+} , ils coordonnent les oxygènes des trois phosphates d'un dNTP entrant et la chaîne latérale des résidus aspartate catalytiques (D110, D185 et D186). Cette étape de coordination au niveau du site polymérase permet au 3'OH d'attaquer le phosphate α du nucléotide entrant. Les deux ions métalliques contribuent également à la stabilisation de la charge des intermédiaires de la réaction formant la liaison phosphodiester.

L'activité RNase H :

Le site actif de la RNase H est situé dans le domaine C-terminal de la RT et est séparé du site de polymérisation par dix-huit paires de bases lorsque la RT est fixée sur un hétéroduplex ARN-ADN (250, 251). De cette façon, la RT peut cliver l'ARN au cours de la synthèse de l'ADN, ce processus est appelé *polymerization-dependent RNase H activity* (262). Néanmoins, le taux de polymérisation est supérieur au taux d'hydrolyse de l'ARN. Des coupures RNase H indépendantes de la polymérisation sont aussi réalisées par la RT car nécessaire pour la dégradation complète de l'ARN copié (262). Si des virus sont déficients en RNase H, ils sont non-infectieux (263–266). La dégradation de l'ARN est normalement indépendante de la séquence mais la RNase H du VIH-1 peut effectuer certains clivages spécifiques. Similairement aux autres polymérases associées à des nucléases, l'activité RNase H, produisant des extrémités 5'P et 3'OH au niveau du site de clivage, est dépendante des ions Mg^{2+} (267).

Il a été suggéré que la spécificité de l'activité de la RNase H aux hybrides ARN-ADN soit liée à la largeur du petit sillon d'environ 9 à 10 Å (268, 269).

La RNase H utilise trois modes de clivage distincts pour dégrader l'ARN dans l'hybride ADN-ARN (Figure 22) (270) :

- Clivage dirigé par l'extrémité 3' de l'ADN et situé à 15-20 nucléotides de celle-ci. Ce clivage est couplé à la synthèse d'ADN (clivage dépendant de l'activité polymérase).

- Clivage dirigé par l'extrémité 5' de l'ARN et situé à 13-19 nucléotides de celle-ci, dans lequel l'extrémité 5' de l'ARN est en retrait sur un long fragment d'ADN,
- Clivage interne.

Les deux derniers modes de clivages sont indépendants de l'activité polymérase de la RT (271) mais ceux-ci fonctionnent simultanément (272). L'orientation de la RT lorsqu'elle est fixée sur des duplexes ARN-ADN semble déterminer son action, soit la synthèse d'ADN ou l'hydrolyse du brin ARN hybridé à l'ADN (273).

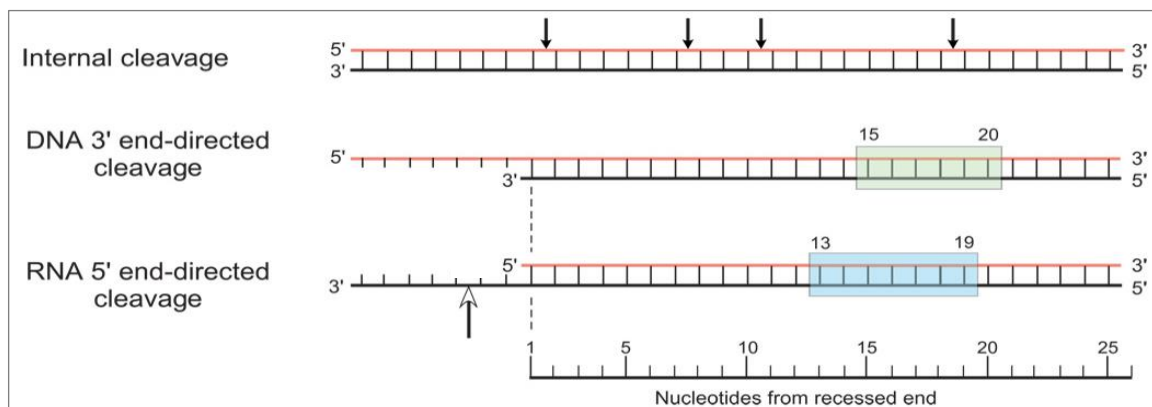


Figure 22. Modes de clivage de la RNase H du VIH-1 (38). En noir l'ADN et en rouge l'ARN forme l'hybride ADN/ARN. Les flèches du clivage interne correspondent aux sites de clivage par la RNase H déterminé par la séquence. La région de clivage pour le mode de clivage dirigé par l'extrémité 3' de l'ADN (15 à 20 nucléotides à partir de l'extrémité d'ADN en amont) est surlignée en vert. La région de clivage pour le clivage dirigé par l'extrémité 5' de l'ARN (13 à 19 nucléotides à partir de l'extrémité) est surlignée en bleu.

e. Les protéines de structure

Les protéines de structure sont codées par le gène *gag* et issues du précurseur polypeptidique Gag de 55 kDa (Pr55^{Gag}) (Figure 9). Les protéines matures obtenues après clivage par la protéase virale sont la protéine de matrice (MA, p17), la protéine de capsid (CA, p24) et la protéine de nucléocapsid (NC, p7) (Figure 23). Ces trois protéines constituent les éléments fondamentaux de la structure du virion et ont des rôles divers.

Notamment, les protéines, MA assurant la liaison aux membranes et tapissent la membrane interne de la particule virale. Les protéines CA formant la capsid contenant l'ARNg dimérique ainsi que les protéines de nucléocapsid dont la forme mature (NCp7 nommée communément NC) recouvre l'ARNg dimérique dans le virion. Les domaines protéiques du précurseur Gag sont reliés entre eux par des peptides espaceurs, SP1 et SP2. La région C-terminale de Gag correspond au peptide p6 impliqué dans le bourgeonnement des particules virales naissantes (274) et dans l'encapsidation de Vpr en le fixant (275, 276).

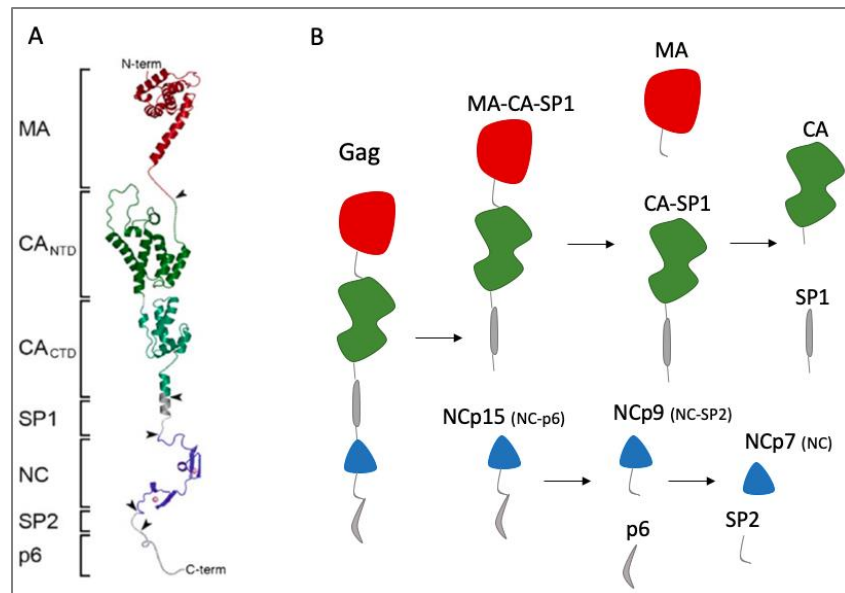


Figure 23. Protéolyse séquentielle de Gag. (A) Modèle structural de la protéine Gag (38). Les points de flèches correspondent aux clivages effectués par la protéase virale. (B) Clivage séquentiel du précurseur Gag permettant d’obtenir plusieurs protéines et peptides (277).

- *La protéine de matrice*

La protéine de matrice (MA) représente le domaine N-terminal du précurseur Pr55^{Gag} (Figure 23). Cette protéine basique possède environ 130 acides aminés et est principalement constituée de cinq hélices α majeures et trois feuilletts β . Environ 30 résidus du domaine C-terminal forment une hélice α servant à relier le domaine MA au domaine CA du précurseur Gag (268, 269) (Figures 8 et 24).

MA est impliquée dans l’interaction de Gag avec la membrane plasmique, l’incorporation de la glycoprotéine Env dans l’enveloppe des particules virales naissantes et dans des événements de la phase précoce suivant l’entrée du virion (278).

Dans la particule virale mature du VIH-1, la protéine de matrice est localisée sous l’enveloppe virale dérivant de la membrane de la cellule infectée. MA possède un double signal avec le groupe myristoyl à l’extrémité N-terminale et la région hautement basique (HBR).

La liaison à la membrane plasmique se fait par ancrage du groupement myristoyl dans la bicouche lipidique et par le domaine N-terminal basique qui lie les phospholipides acides, particulièrement le phosphatidylinositol-4,5-biphosphate PI(4,5)P2 qui est fortement présent dans la membrane plasmique (278).

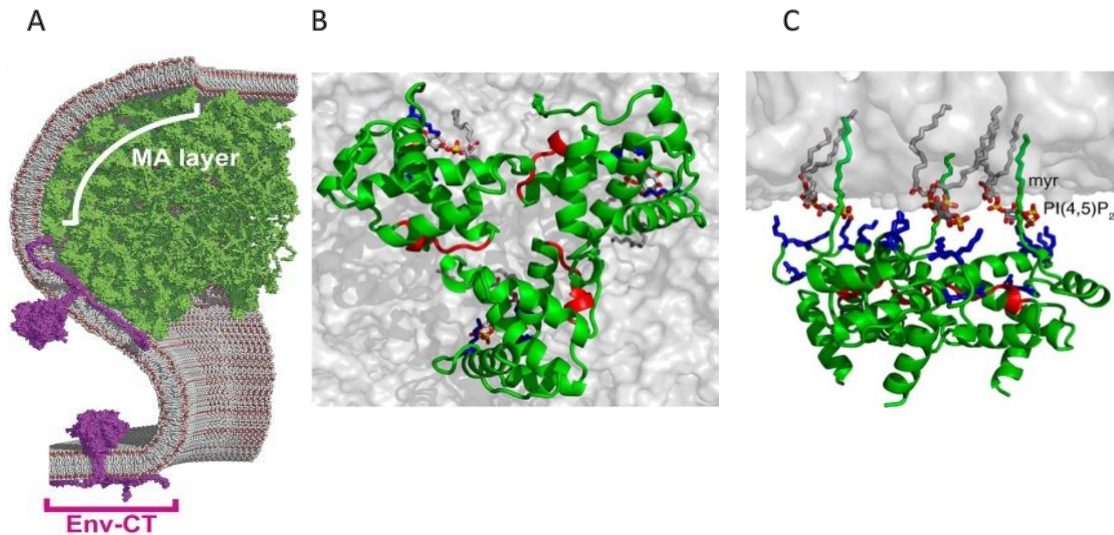


Figure 24. Modèle de reconnaissance entre MA et la membrane plasmique (280, 281). (A) Modèle en cartoon du précurseur Gag (en vert) avec la MA (identifiée par la ligne blanche) ancrée dans une bicouche membranaire. (B) Trimère de MA vu du dessous. (C) Trimère de MA lié la membrane par ses domaines HBR (en bleu), ses groupements myr et la tête polaire de PI(4,5)P₂. Les résidus en rouge sont localisés à l'interface trimérique.

- ***La protéine de capside***

La protéine de capside est une protéine d'environ 230 acides aminés composée de deux domaines liés par une région flexible : la région N-terminale dite « le domaine central » intervenant dans la maturation du virion et l'incorporation de la Cyclophiline A (CypA, protéine favorisant la réplication du VIH) et une région C-terminale dite « domaine de dimérisation » impliquée dans les interactions Gag-Gag (282).

Le domaine central est composé de sept hélices α , deux feuillets β et une boucle fortement exposée servant de site de liaison canonique pour la CypA (283). Le domaine CA du précurseur Gag joue un rôle important dans l'assemblage et la maturation du virion. Son extrémité N-terminale n'est d'ailleurs pas structurée et est liée au domaine MA lui-même relié à la membrane plasmique. Cette extrémité est libérée lors de la protéolyse du précurseur, permettant à la région N-terminale de se replier en épingle à cheveux β (284).

- ***La protéine de nucléocapside***

La protéine de nucléocapside mature (NCp7 communément nommée NC) est codée par le gène *gag* et est présente entre 1000 à 1500 copies par virion (285, 286). L'ARNg dimérique est couvert de NC dans le virion. Le rapport NC:ARNg est de 1 molécule de NC pour 12 à 18 nucléotides. Cette protéine, sous forme mature ou comme domaine du précurseur Gag, est impliquée dans plusieurs étapes du cycle réplcatif, notamment la transcription inverse (détaillé

dans le chapitre suivant), l'intégration de l'ADN double-brin dans le génome cellulaire, l'assemblage de la particule virale ainsi que la dimérisation et l'encapsidation de l'ARNg.

La NC est une petite protéine de 55 acides aminés générée par le clivage protéolytique du précurseur Gag qui produit séquentiellement trois protéines contenant le domaine NC:NCp15, NCp9 et finalement NCp7 (Figure 23). La NC se fixe sur les acides nucléiques grâce à son caractère basique et à ses deux doigts de zinc (ZF1 et ZF2). Ces derniers contenant le motif conservé CCHC fixent le zinc avec une très forte affinité (Figure 25). Les extrémités N et C-terminales ainsi que la région charnière entre ZF1 et ZF2 sont flexibles et faiblement structurées (287, 288)

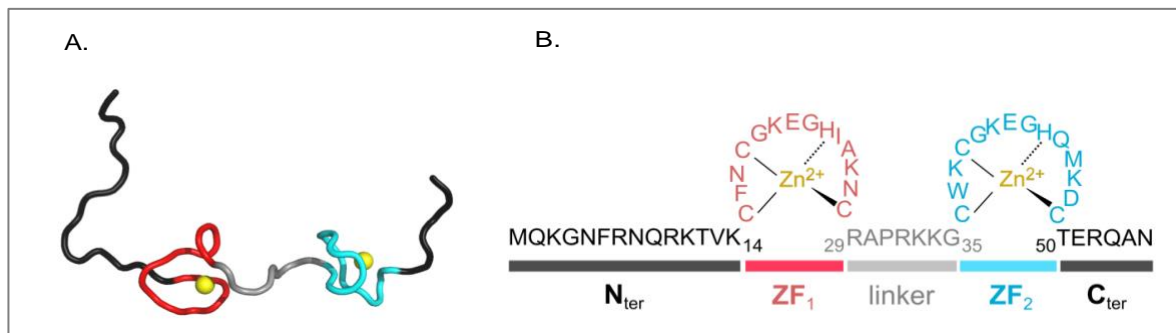


Figure 25. Structure 3D (A) et primaire (B) de la NC (289). Les deux domaines en doigts de zinc sont en rouge (ZF₁) et bleu (ZF₂), la région charnière en gris clair et les extrémités N- et C-terminales en gris foncé.

Propriétés de la protéine de nucléocapside

Interaction avec les acides nucléiques :

La NC se fixe sur l'ADN et l'ARN préférentiellement au niveau des régions simple-brin. La constante d'association de cette protéine à un acide nucléique dépend de la nature, la séquence et le repliement de la molécule cible, c'est pourquoi elle peut varier de plusieurs ordres de grandeur (de quelques nM à 1400 nM (290)). Cela lui permet d'exercer différentes fonctions selon la nature des séquences et les concentrations des acides nucléiques.

Cinq à huit nucléotides contenant généralement au moins une guanine non appariée constituent le site d'interaction de la NC (291). De plus, les motifs TG, GNG et TNG constituent des sites de fixation préférentielle pour la NC (292–295). Ces motifs sont notamment présents dans les tiges-boucles : TAR de l'ARNg, cTAR et PBS de l'ADN (-) viral (284, 287). Les interactions spécifiques ont un rôle crucial dans la reconnaissance du signal d'encapsidation de l'ARNg viral par le domaine NC du précurseur Gag, impliquant la reconnaissance et la sélection spécifique parmi de nombreux ARN cellulaires lors de l'assemblage de la particule virale (298, 299).

Une autre caractéristique de la NC est qu'elle présente une polarité de fixation opposée selon la nature de l'acide nucléique (300, 301). La NC se fixe sur l'ARN avec une orientation antiparallèle, c'est-à-dire que son extrémité N-terminale est orientée vers l'extrémité 3' de l'ARN. Inversement, la NC se fixe sur l'ADN de façon parallèle avec son extrémité N-terminale orientée vers l'extrémité 5' de l'ADN (Figure 26) (292, 293).

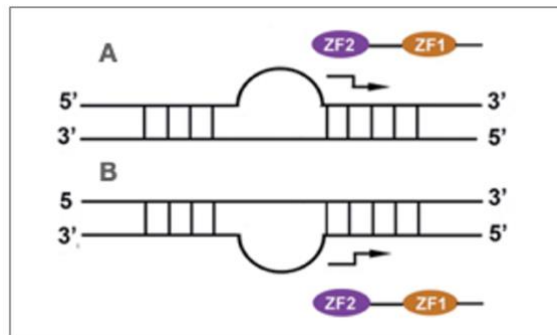


Figure 26. Polarité de fixation de la NC aux acides nucléiques (301). (A) La NC se fixe sur l'ARN avec une orientation antiparallèle. (B) La NC se fixe sur l'ADN avec une orientation parallèle. Dans les deux cas, ZF2 est à proximité de la région simple-brin et induit une déstabilisation de la tige dans le sens indiqué par la flèche

Activité chaperon des acides nucléiques :

La NC possède une activité chaperon des acides nucléiques lui permettant de réarranger l'ADN et l'ARN dans la structure secondaire la plus stable thermodynamiquement contenant généralement le maximum de paires de bases. La NC a un rôle prépondérant dans l'appariement des acides nucléiques complémentaires qui forment des structures secondaires en tige-boucle. Cette activité chaperon repose sur trois activités (302) :

1) **La déstabilisation transitoire des structures secondaires des acides nucléiques qui est principalement due aux deux doigts de zinc.** Les motifs tels que les *bulges*, les boucles apicales et internes, les mis-appariements et les paires de bases terminales constituent des éléments qui facilitent l'action déstabilisatrice de la NC (303, 304).

Celle-ci est souvent initiée dans les régions double-brin qui sont proches de nucléotides non appariés. La déstabilisation se propage ensuite le long de l'acide nucléique cible permettant le glissement de la NC et/ou la fixation de nouvelles molécules de NC (298, 299). La déstabilisation des structures secondaires favorise les processus d'hybridation inter ou intramoléculaires durant le cycle répliatif.

2) **L'agrégation des acides nucléiques due aux résidus basiques.** Cette activité facilite le contact et par conséquent l'appariement des acides nucléiques complémentaires.

3) **Un temps de résidence court sur les acides nucléiques.**

2.3.6 Encapsidation de l'ARN génomique dans la particule virale

Au cours de la phase tardive, deux copies d'ARNg non épissé contenant une coiffe en 5' et une queue poly(A) en 3' sont sélectionnées. Elles sont recrutées près de la membrane plasmique au niveau du site d'assemblage de la particule virale pour y être encapsidées. On retrouve dans le cytoplasme proche de la zone d'assemblage plusieurs types d'ARN pouvant être cellulaires ou viraux. Ces derniers sont alors soit épissés, codant pour les protéines virales accessoires, soit non épissés, codant pour les protéines virales Gag et Gag-Pol et servant d'ARNg à encapsider.

Plusieurs études ont montrés que l'encapsidation de l'ARNg est liée à sa dimérisation (299, 306, 307). L'ARNg dimérique est préférentiellement incorporé par rapport aux autres ARNm viraux et cellulaires (308, 309). L'interaction entre les deux copies d'ARNg se réalise au niveau du signal d'encapsidation Ψ qui débute dans la région non codante 5' et s'étend jusqu'au premiers nucléotides du gène *gag*. L'interaction ARNg-ARNg est initiée plus précisément par la tige-boucle SL1 (310). Cette dernière possède une séquence palindromique dans sa boucle apicale, appelée site d'initiation de dimérisation (DIS) permettant une interaction boucle-boucle entre deux molécules d'ARNg (310, 311). Pour rendre possible cette interaction intermoléculaire, onze nucléotides de la séquence U5 doivent être appariés avec la séquence contenant le codon d'initiation AUG du gène *gag* (Figure 27) (312) laissant exposé le DIS. Deux autres conformation non compatible avec la dimérisation de l'ARNg existe où le DIS est apparié avec les nucléotides de la séquence U5 (Figure 27) (312) ou avec le poly(A) (313).

Les études *in vitro* suggèrent fortement que la NC joue un rôle important dans la formation du dimère d'ARNg (312, 314–317). Le domaine NC de Gag interagit spécifiquement avec l'ARNg grâce à ses deux doigts de zinc au niveau du signal d'encapsidation Ψ constitué des tiges-boucles PBS, SL1 à SL4 situé dans la région 5'UTR (67, 70, 308, 310, 318). Le PBS ne contribue pas à la dimérisation ou l'encapsidation de l'ARNg (319). La tige-boucle SL3 est un élément important de ce signal (320) mais plusieurs études de mutagénèse dirigée suggèrent que la tige-boucle SL1 serait un élément structural plus essentiel à l'encapsidation de l'ARNg que la tige-boucle SL3 (67, 69, 71, 321). La région Ψ de l'ARNg contient sept sites contenant des guanines non appariées qui peuvent être reconnues par le domaine NC du précurseur Gag.

La substitution de ces guanines par d'autres nucléotides a montré qu'elles agissaient de façon synergique suggérant une fixation d'un précurseur Gag sur chacun des sites pour conduire à une encapsidation optimale de l'ARNg dimérique (322). En effet, l'encapsidation était peu diminuée lorsqu'il n'y avait que deux sites putatifs mutés mais l'était d'un facteur trente-quatre quand six des sept sites étaient mutés (322).

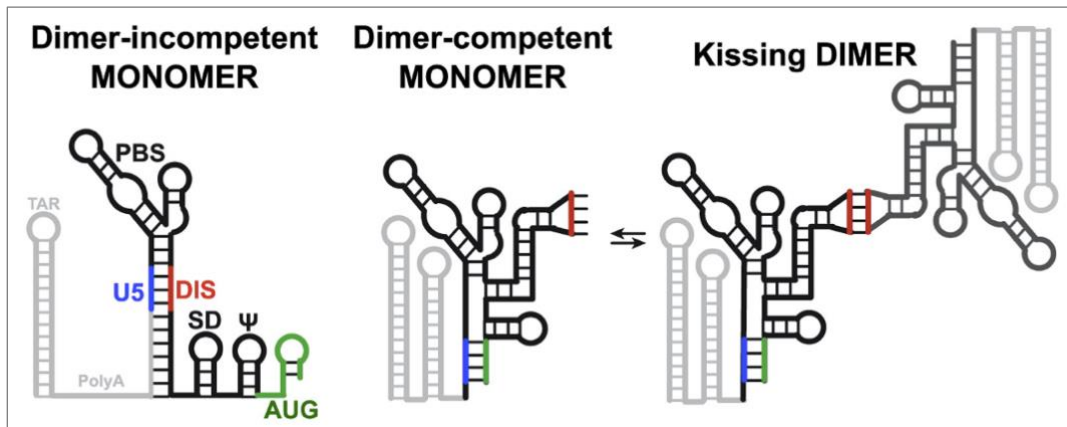


Figure 27. Conformations adoptées par la région 5'UTR de l'ARNg (312). Dans la conformation monomère *dimer-incompetent* la séquence du DIS (en rouge) est impliquée dans des appariements avec une partie de la séquence U5 (en bleu) empêchant la formation d'une interaction intermoléculaire entre deux molécules d'ARNg. Dans la conformation monomère *dimer-competent* le codon AUG (en vert) du gène *gag* est apparié avec une partie de la séquence U5 libérant ainsi le DIS qui est capable d'interagir avec le DIS d'une autre molécule d'ARNg pour former le complexe boucle-boucle (*kissing dimer*).

L'interaction ARNg-ARNg agit comme un point de nucléation afin de permettre l'oligomérisation des unités Gag et Gag-Pol par des interactions Gag-Gag et Gag-ARN contribuant à l'assemblage de la particule virale immature. La délétion du domaine NC du précurseur Gag entraîne l'accumulation de ce dernier au niveau de la membrane plasmide *in cellula* (323). Des mutations au niveau des doigts de zinc de la NC ont un impact négatif sur le bourgeonnement et le clivage du précurseur Gag (314, 315). L'assemblage du virion est ainsi médié par l'encapsidation de l'ARNg dimérique et les interactions Gag-Gag où le précurseur Gag est l'acteur protéique principal du processus d'encapsidation (323, 324).

Il a été montré que les tiges-boucles TAR et poly(A) de la région 5', qui n'est pas comprise dans le signal Ψ , jouent un rôle dans l'encapsidation de l'ARNg sans une quelconque intervention dans l'interaction du domaine NC avec le génome viral. L'ouverture de la tige-boucle TAR provoque une dimérisation aberrante et une encapsidation réduite de l'ARNg (325). Le motif 5'poly(A) a été identifié comme nécessaire à l'encapsidation de l'ARNg bien que généralement réprimé lors de la réplication virale afin d'éviter la génération de courts transcrits. Le signal 5'poly(A) pourrait faciliter l'encapsidation par une interaction directe ou indirecte du

précurseur Gag avec la machinerie cellulaire de polyadénylation (326). CPSF6 semble être un possible candidat puisqu'il intervient dans la répression du signal 5'poly(A) mais également dans le 3'-end processing de l'ARN (326). Néanmoins, des études RMN ont montré que les tiges-boucles TAR et poly(A) forment un empilement coaxial où la coiffe 5' de l'ARNg est « prise en sandwich » entre les deux tiges-boucles (327). Cette séquestration de la coiffe 5' permettrait aux ARNg de ne pas être reconnu par le facteur d'initiation de la traduction eIF4E, d'éviter leur dégradation par la machinerie cellulaire et donc d'être disponibles pour interagir avec les polypeptides Gag (327). Elle ferait donc partie, avec le reste de la région 5' non codante (signal d'encapsidation Ψ et de dimérisation DIS), des éléments prérequis pour une encapsidation optimale de l'ARNg (328).

2.3.7 Assemblage, bourgeonnement et maturation de la particule virale

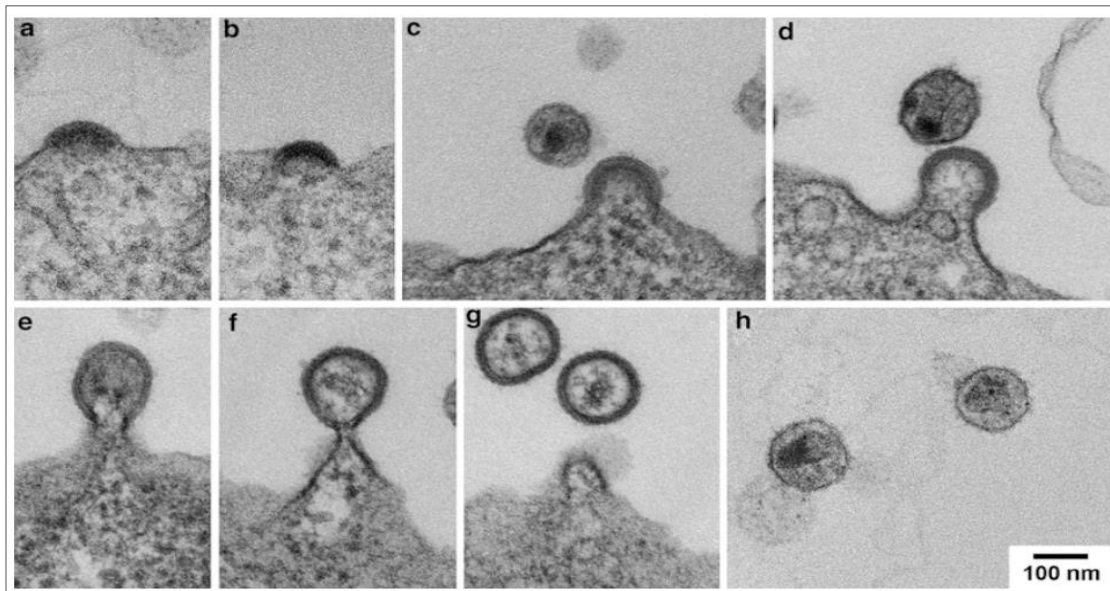
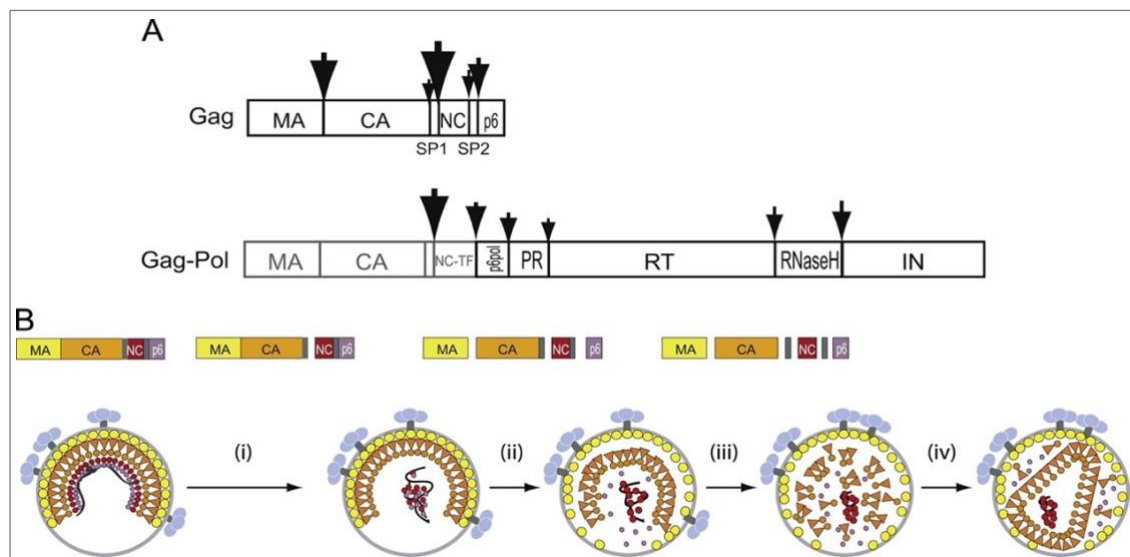


Figure 28. Étape de bourgeonnement du VIH-1 observée par microscopie électronique dans des cellules T humaines (329). (a, b) Lors des premières étapes du bourgeonnement, on observe une accumulation de protéine Gag proche de la membrane plasmique. (c, d) Formation de la coque sphérique immature de la particule conduisant à la courbure de la membrane. (e, f, g) Le lien membranaire s'affine entre le virion naissant et la cellule infectée conduisant à sa fission. (h) Gag est clivé par la protéase virale, la capsidie virale est ainsi formée.

L'assemblage de la particule virale a lieu au niveau de la membrane plasmique de la cellule infectée. Les différents composants (ARNg, Gag, Env, Gag-Pol, protéines accessoires) de la particule virale doivent être acheminés simultanément vers la membrane cellulaire depuis leur lieu de synthèse. Tout d'abord, le précurseur polypeptidique Env (gp160) est traduit près du réticulum endoplasmique puis est inséré dans la membrane de ce dernier afin d'être glycosylé.

Il est ensuite transporté vers la membrane plasmique via l'appareil de Golgi où il est clivé par la furine, une protéase cellulaire, afin de générer les deux glycoprotéines, gp120 et gp41. Par la suite, ces dernières sont transportées vers la membrane plasmique et gp41 ancre gp120 à la membrane des cellules infectées. Les précurseurs polypeptidiques Gag (Pr^{55}) et Gag-Pol (Pr^{160}) sont exprimés à partir de l'ARNg non épissé. Ils se fixent à la membrane plasmique par leur extrémité N-terminale myristoylée (Figure 24) (330). Gag oligomérisse via son domaine CA et interagit avec l'ARNg via son domaine NC permettant ainsi l'assemblage de la particule virale (331, 332). Le domaine p6 du précurseur Gag contient une région riche en proline et un domaine di-leucine qui sont les cibles des protéines cellulaires Tsg101 et ALIX. Ces dernières sont impliquées dans le processus d'export cellulaire par les protéines de classe E et la machinerie nécessaire au bourgeonnement de la particule virale immature (Figure 28) (172, 333).



Après le bourgeonnement, la particule virale n'est pas infectieuse et nécessite une maturation. Celle-ci repose sur le clivage des précurseurs Gag et Gag-Pol par la protéase virale, permettant ainsi d'obtenir les protéines fonctionnelles du VIH-1. Environ 1000 à 1500 copies du précurseur Gag sont nécessaires à la formation d'un virion infectieux (285) (286). Gag présente cinq sites de clivages protéolytiques ciblés par la protéase virale (Figure 29) (241). Le clivage entre SP1

et NC est le premier à se produire, tandis que les clivages entre SP2-p6 et MA-CA se produisent en second. Les deux sites restants CA-SP1 et NC-SP2 sont les derniers à être clivés. Le clivage entre CA et SP1 permet de déclencher le réarrangement structurel de la capsid de la particule virale en forme conique (334). Le clivage entre SP1 et NC conduit à séparer le domaine MA-CA du domaine NC associé à l'ARNg. Lors de la protéolyse, le domaine NC existe sous deux formes intermédiaires : NCp15 (produit de clivage partiel comprenant les domaines NC-SP2-p6) et NCp9 (produit de clivage partiel contenant les domaines NC-SP2). La forme finale NCp7, communément nommée NC, résulte du clivage total (38).

3. Transcription inverse

Lorsque qu'un virion infecte une cellule, les interactions entre les glycoprotéines virales, les récepteurs et corécepteurs cellulaires entraînent une fusion des membranes. Cette fusion conduit à la libération du contenu de la particule virale dans le cytoplasme cellulaire (Figure 10) et permet ainsi à la transcription inverse d'être initiée. La transcription inverse est un processus complexe (Figure 30) (335) permettant de convertir l'ARNg simple-brin en ADN double-brin linéaire qui est nommé ADN proviral lorsqu'il est intégré dans le génome cellulaire. Un seul ADN proviral est produit à partir d'un dimère d'ARNg (336). Deux protéines virales sont utilisées lors de la transcription inverse : la RT avec ses activités ADN polymérase et RNase H, ainsi que la NC possédant une activité chaperon des acides nucléiques (291, 302, 337). Des études ont montré que des mutations de la NC peuvent affecter le déroulement de la transcription inverse (337–341).

Il est établi que la majorité de l'ADN proviral est synthétisé dans la cellule infectée (342). Néanmoins, plusieurs études ont montré qu'une faible proportion de dNTPs de la précédente cellule infectée peuvent être incorporé lors du bourgeonnement, permettant une initiation de la transcription inverse dans les particules avant l'entrée dans une nouvelle cellule (343, 344).

Plusieurs équipes ont développé des systèmes acellulaires afin d'identifier les mécanismes impliqués dans les étapes clés de la transcription inverse. La transcription inverse endogène naturelle (NERT) permet de synthétiser dans des virions purifiés les différents intermédiaires et produits de la transcription inverse (345, 346). Ce processus peut être amplifiée dans des virions purifiés en la présence de désoxyribonucléotides et de détergents doux, alors appelé transcription inverse endogène (ERT) (347, 348). Toutefois, le rendement du produit complet est extrêmement bas (346).

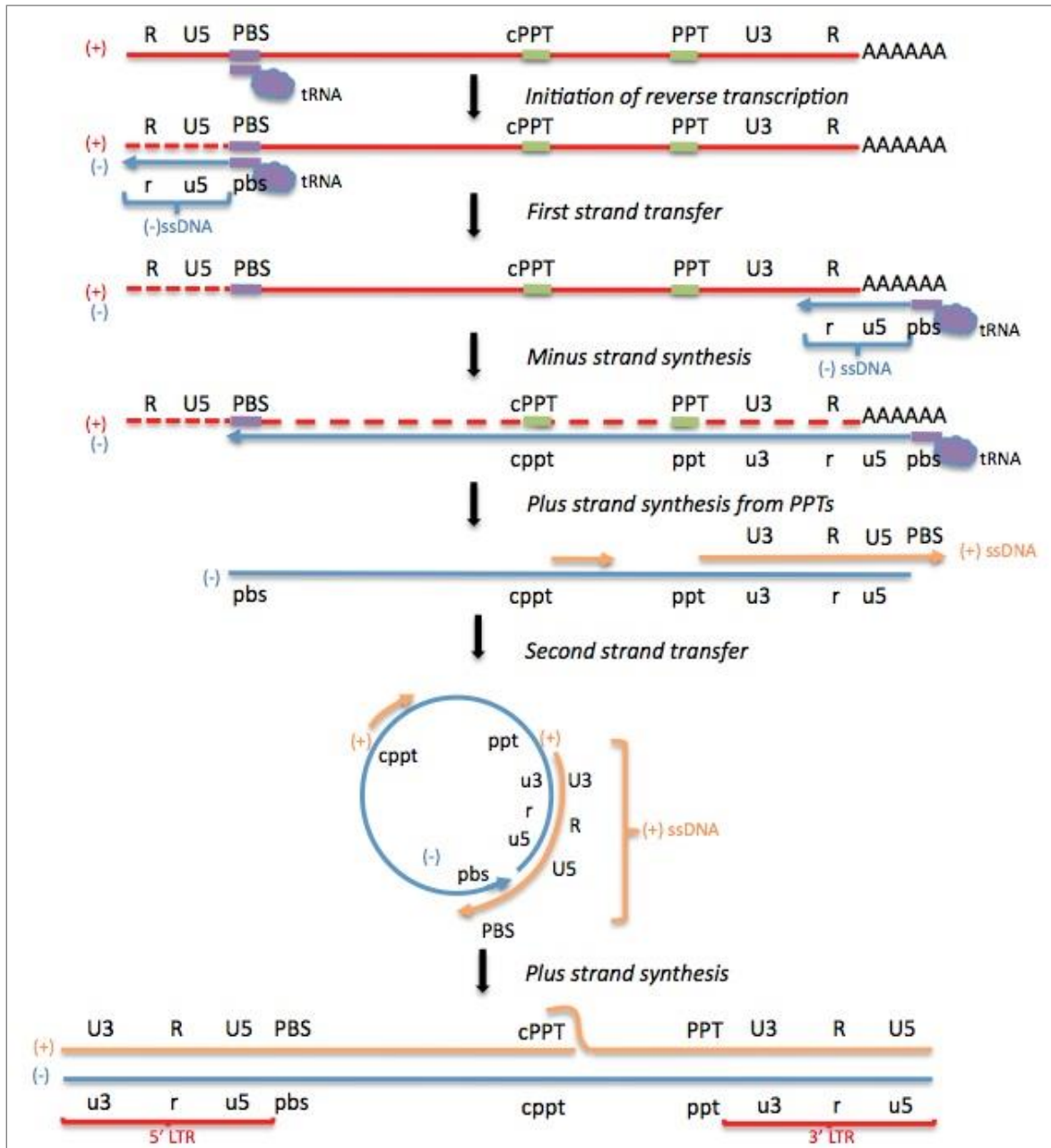


Figure 30. Schéma de la transcription inverse chez le VIH-1, adapté de (349). L'ARN viral est en rouge, l'ADN viral en bleu (-) et jaune (+), l'ARNt (trait et nuage) est en violet.

3.1 Fixation de l'amorce et formation de l'hybride ARNg-ARNt^{Lys3}

Lors de l'assemblage des particules virales, les isoaccepteurs cellulaires ARNt lysine 1, 2 et 3 sont incorporés sélectivement (350). Cette incorporation sélective implique les précurseurs polypeptidiques Gag et Gag-Pol, et les Lysyl-ARNt synthétases de la cellule hôte (351, 352). Les Lysyl-ARNt synthétases servent d'intermédiaires pour le recrutement des ARNt (353). Seul l'ARNt^{Lys3} sert d'amorce pour la transcriptase inverse. Les deux autres ARNt auraient un rôle dans l'importation du PIC jusqu'au noyau cellulaire (354).

La séquence sur laquelle se fixe l'amorce est nommée *Primer binding site* (PBS), elle est située au niveau du 5'UTR du génome viral (Figure 30) (355, 356). Elle est complémentaire des dix-huit derniers nucléotides de l'extrémité 3' de l'amorce ARNt^{Lys3} (357). L'hybridation entre l'amorce ARNt et le PBS permet l'initiation de la transcription inverse. Plusieurs études *in vitro* suggèrent que par son activité chaperon des acides nucléiques, la NC est l'acteur principal de cette hybridation (337, 340, 358). L'ARN hélicase A humaine pourrait être un acteur secondaire qui faciliterait l'hybridation en induisant un changement conformationnel dans la région 5'-UTR de l'ARNg (359, 360) mais ceci a été contredit par l'équipe de X. Heng (361). L'ARN hélicase A n'agirait que sur la processivité de la RT durant la phase d'élongation (361). Gag peut aussi faciliter l'hybridation ARNt^{Lys3}-PBS dans des virions dépourvus de protéase virale fonctionnelle (362–365).

La détermination de la structure secondaire de la région PBS dans les particules immatures et matures a permis de proposer un processus d'hybridation en deux étapes : une initiation de l'hybridation par le domaine NC de Gag suivi d'un remodelage par la NC mature du complexe ARNt^{Lys3}-PBS vers l'état thermodynamiquement le plus stable (356). La NC agirait tout au long du processus de transcription inverse en déstabilisant les structures secondaires comme les tiges-boucles afin d'éviter les pauses de la RT et sa dissociation de la matrice ARN (291).

Les régions du génome viral sur lesquelles agit la NC ne sont pas réparties de manière uniforme sur l'ARNg. En effet, six sites double-brin situés entre l'extrémité 5' de l'ARNg et la séquence PBS (les 185 premiers nt) sont préférentiellement déstabilisés par la NC (366). Ces sites sont situés justement dans la région contenant la séquence PBS ainsi que la séquence copiée par la RT lors de la synthèse de l'ADN *strong-stop* (ADNss) (Figure 31) (366).

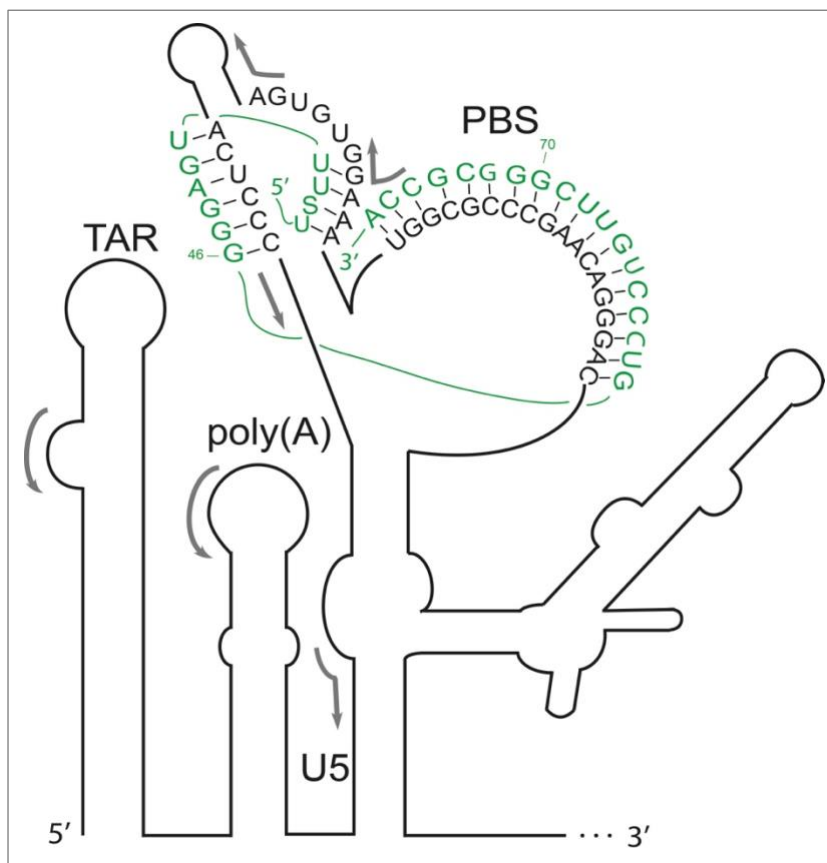


Figure 31. Schéma de la structure secondaire de l'extrémité 5' de l'ARNg. En vert, l'ARNt engagé dans trois interactions avec l'ARNg viral. Les flèches grises indiquent les sites destabilisés par la NC.

Deux interactions supplémentaires entre le 5'UTR et l'amorce ARNt, formées *in vitro* en l'absence de NC à haute température (70 °C), ont été identifiées (367). Elles impliquent deux séquences de l'ARNg en amont du PBS : une boucle riche en A et une séquence riche en C engagée dans une tige (Figure 31) (338, 347). Ces deux interactions supplémentaires sont supportées par la détermination de la structure secondaire de l'ARNg dans le virion (366). Une étude récente rapporte que l'interaction entre la boucle riche en A du PBS et l'anticodon l'ARNt^{Lys3} est conservé entre différentes souches du VIH-1(348). Lorsque cette interaction est perturbée par des mutations, une diminution de l'infectiosité, une diminution de la synthèse des produits de transcription inverse et une augmentation de la dégradation de de l'ARNg ont lieu (348). Ces résultats suggèrent que l'interaction entre la boucle riche en A du PBS et l'anticodon l'ARNt^{Lys3} protège l'ARNg de la dégradation par l'activité RNase H de la RT en l'absence de dNTPs lorsque le virion n'est pas entré dans la cellule hôte (348). La NC semble jouer un rôle important dans la formation du duplex ARNt^{Lys3}-ARNg car elle permet sa formation *in vitro* à température physiologique (368). L'étude de Heng et al. parvient à la même conclusion (365).

Les structures secondaires de la région PBS au sein des particules virales matures et immatures ont été déterminées au moyen de la méthode SHAPE (356, 366). Ces structures ont permis de proposer la formation du duplex ARN^{t^{Lys3}}-ARNg en deux étapes (Figure 32) (356) :

- i) un duplex instable qui est facilité par le domaine NC du précurseur Gag avant la maturation de la particule virale ;
- ii) un duplex stable, impliquant les deux interactions supplémentaires, qui est facilité par la NC présente dans la particule virale mature.

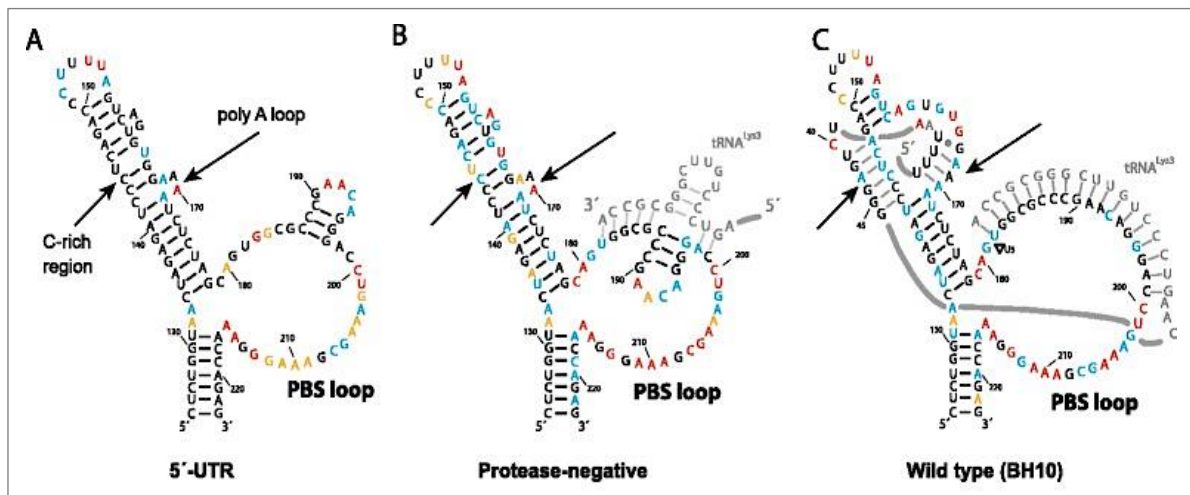


Figure 32. Modèle de liaison de l'ARN^{t^{Lys3}} à l'ARNg (356). (A) Structure secondaire de la région PBS synthétisée *in vitro* en l'absence de l'ARN^{t^{Lys3}}. (B) Interaction initiale entre l'ARN^{t^{Lys3}} et le PBS générée par l'action du polypeptide Gag. (C) Interactions principales et secondaires entre l'ARN^{t^{Lys3}} et l'ARNg générées par l'action de la NC.

Les études *in vitro* du groupe de Berkhout suggèrent un autre type d'interaction supplémentaire entre la séquence 48-55 de l'ARN^{t^{Lys3}} et une séquence conservée de huit nucléotides en amont du PBS, nommée *primer activation signal* (PAS) (369, 370). Toutefois, cette interaction supplémentaire n'est pas supportée par les études structurales de l'ARNg incorporé dans le virion.

La résolution par RMN de la structure du PBS décrit trois tiges-boucles distinctes riches en adénosine à partir d'une jonction trois voies : la tige *tRNA-like element* (TLE) (nt 135 à 177), la tige PAS (nt 125 à 223) et la tige fixant l'amorce ARNt (Figure 33) (371). Des mutations affectant la structure altèrent la transcription inverse dans les cellules infectées (371).

La fixation de l'ARNt amorce au PBS ne dépendrait pas que de la NC mais serait aussi facilitée par la lysyl-ARNt synthétase et l'ARN hélicase A qui se lie aux tiges TLE et PAS (353, 371–373). L'implication des enzymes cellulaires permettrait une fixation optimale de l'ARN^{t^{Lys3}} et un contrôle temporel de l'initiation de la transcription inverse afin d'assurer son efficacité.

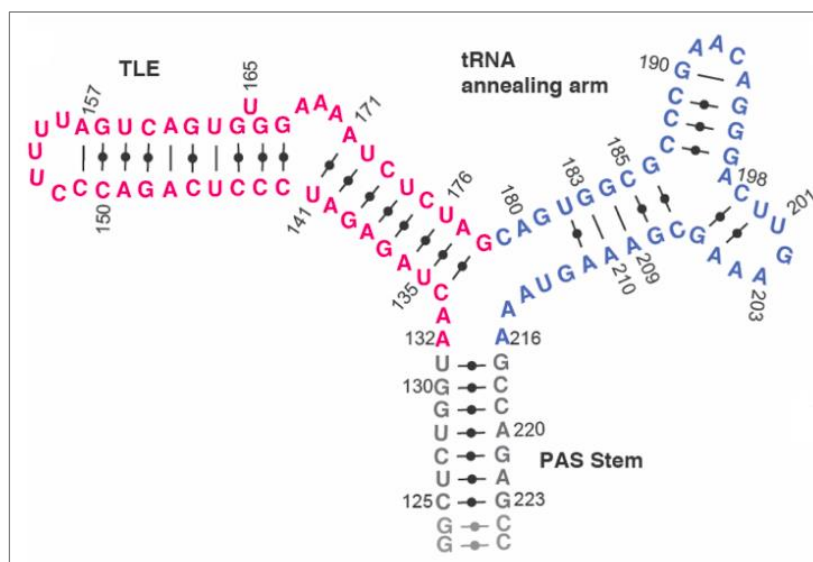


Figure 33. Structure de la région PBS déterminée par RMN (371). Les tiges TLE, PAS et celle fixant l'ARNt sont respectivement en rouge, gris et bleu.

3.2 Initiation de la transcription inverse

La transcriptase inverse a une affinité beaucoup plus faible pour l'ARN double-brin que pour l'ADN double-brin ou les duplex ADN-ARN expliquant la difficulté d'une caractérisation structurale du complexe amorce ARN^{t_{Lys3}}-ARNg lié à la transcriptase inverse (355). L'équipe de Puglisi a été la première à publier une structure tridimensionnelle du complexe ternaire (RT-ARN^{t_{Lys3}}-portion d'ARNg) d'initiation de la transcription inverse, grâce à la cryo-microscopie électronique avec une résolution de la région duplex de 4,5 Å (PDB ID 6B19) (374). L'année suivante l'équipe d'Arnold a publié la structure cristalline avec une résolution de 3,95 Å d'un complexe ternaire (RT-matrice ARN-amorce ARN, PDB ID code 6HAK) qui mime en partie le complexe naturel (375). Les deux structures apportent des informations complémentaires (Figure 34) (376). La structure cristalline représente le cœur du complexe d'initiation de la transcription inverse formé par un duplex ARN-ARN en partie double-brin (23 mer pour la matrice et 17 mer pour l'amorce) afin de mimer la région PBS de l'ARNg appariée avec l'extrémité 3' de l'ARN^{t_{Lys3}}. Cette structure décrit le complexe prêt à incorporer un nucléotide (site P : position de l'amorce 0). La structure déterminée par cryo-microscopie électronique décrit le complexe d'initiation de la transcription inverse formé par la RT et une portion d'ARNg (101 pb) hybridée avec l'ARN^{t_{Lys3}} entier. Ce complexe représente l'étape juste après l'incorporation d'un nucléotide et avant la translocation (site N : position de l'amorce +1).

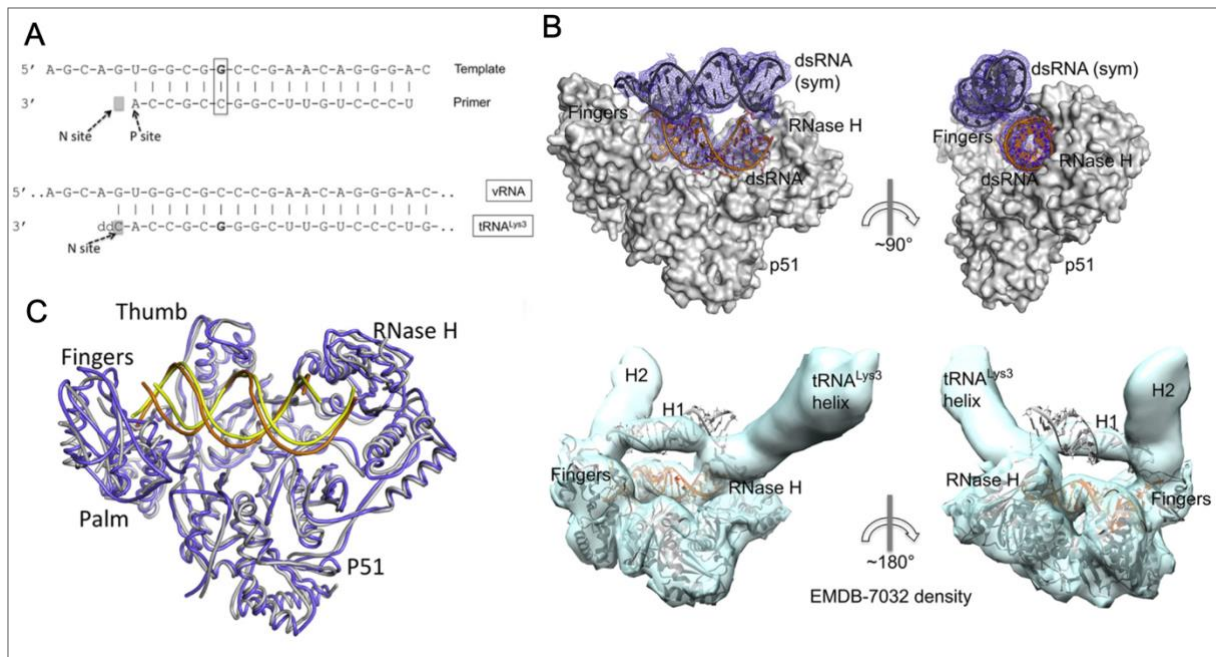


Figure 34. Structures des complexes d'initiation de la RT obtenus par cryomicroscopie électronique et cristallographie (376). (A) Séquences PBS appariées utilisées pour la cristallographie (en haut) et la cryomicroscopie électronique (en bas). (B) Cartes de densité du complexe d'initiation de la transcription inverse obtenues par la cristallographie (en haut, PDB ID 6B19) et par cryo-ME (en bas, PDB ID 6B19). (C) Superposition des complexes montrant les fortes ressemblances structurales.

L'addition des premiers désoxyribonucléotides est lente et non-processive puis une fois les cinq à six premiers nucléotides ajoutés à l'extrémité 3' de l'amorce ARNt la synthèse de l'ADN viral s'accélère (377). À partir des deux structures, il a été observé un déplacement important de l'extrémité 3' de l'amorce du site actif de la RT (259). Ce qui pourrait expliquer le faible taux d'ajout des premiers désoxyribonucléotides (376). De plus, l'équipe de Zhuang (378) a montré que la RT adopte deux orientations : une favorisant l'activité polymérase et une orientation inversée dans laquelle le site actif de la polymérase se place à l'opposé de l'extrémité 3' de l'ARNg. L'équilibre entre ces deux orientations joue un rôle dans la régulation de l'activité de la RT au cours de l'initiation de la transcription inverse. En effet, les structures secondaires de l'ARNg au niveau de la région PBS orientent la RT vers la position non favorable à la polymérisation conduisant à une pause de la RT. L'ajout des premiers nucléotides à l'ARNt^{Lys3} conduit à un changement d'orientation de la RT vers la position favorisant la polymérisation (Figure 35) (378).

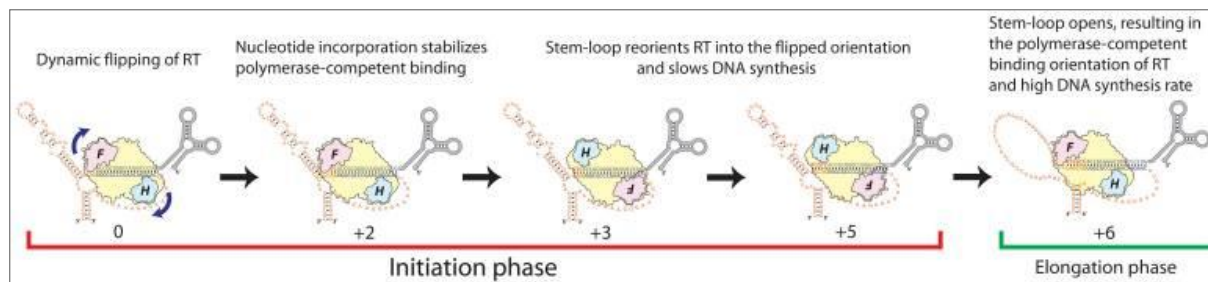


Figure 35. Orientation de la RT au cours de l'initiation de la transcription inverse (378). La RT adopte deux orientations lorsqu'elle se fixe sur le duplex ARN^{t^{Lys3}}-ARNg (stade 0 schématisé par les deux flèches incurvées) : une favorable à l'activité polymérase et une défavorable à l'activité polymérase. L'orientation non favorable à la polymérisation (domaine RNase H de la RT du côté 3' de l'amorce) est adoptée par l'enzyme après ajout des 3^{ème}, 4^{ème} et 5^{ème} nucléotide. Le taux de synthèse diminue considérablement à la position +3 car la RT rencontre les structures secondaires de l'ARNg induisant l'enzyme à se lier de manière inversée au duplex matrice-amorce, augmentant la probabilité de pause de la RT. La synthèse par déplacement de brin jusqu'en position +6 conduit à la déstabilisation de la structure secondaire permettant à la RT de se réorienter dans le mode de liaison compétent pour la polymérisation et entrer dans la phase d'élongation processive de la synthèse d'ADN.

Les interactions supplémentaires décrites dans la partie 3.1 faciliteraient la fixation de la RT sur le duplex matrice-amorce en évitant des encombrements stériques (355, 379). Il a été montré *in vitro* qu'en l'absence de NC, l'interaction de l'ARN^{t^{Lys3}} avec la boucle riche en A est plus importante chez l'isolat MAL que chez l'isolat HXB2 pour l'initiation de la transcription inverse (380–382). Pour les isolats HXB2 et LAI, c'est l'interaction entre la séquence 48-55 de l'ARN^{t^{Lys3}} et le PAS qui favorise l'initiation *in vitro* (369). Des mutations dans la boucle riche en A entraînent une diminution de la synthèse de l'ADN viral et de l'infectivité pour les souches MAL et NL4-3 (365). La tige PAS de la région PBS (Figure 33) serait importante la fixation et le recrutement l'ARN hélicase A, enzyme augmentant la processivité de la RT (361, 371). La suppression de la boucle riche en A dans l'isolat HB2X altère probablement l'initiation de la transcription inverse car elle conduit à une diminution de la synthèse d'ADN viral dans les tests *in vitro* et *ex vivo* (365).

3.3 Synthèse de l'ADN *strong-stop*

L'élongation de l'ARNt amorce par la RT conduit à la synthèse de l'ADNss, un court ADN qui est la copie de la région R-U5 de l'ARNg (Figure 30). Cet ADN a été découvert lors de la synthèse *in vitro* de l'ADN du virus du sarcome de Rous par la RT endogène du virus (383). Un fragment d'ADN de 101 nucléotides a été le principal produit généré par la synthèse incomplète de l'ADN viral, il a pour cette raison été nommé ADN *strong-stop* (ADNss). Chez le VIH-1, l'ADNss a une longueur d'environ 180 nucléotides qui peut légèrement varier selon

les isolats (384). Une étude récente montre qu’au sein d’un même isolat la longueur de l’ADNss peut être allongée de 1 à 2 nucléotides car il résulte de la copie de l’ARNg encapsidé qui peut posséder une, deux ou trois guanines à son extrémité 5’ (385). L’ARNg qui est préférentiellement encapsidé contient une seule guanine à son extrémité 5’ et par conséquent la forme prépondérante d’ADNss est celle qui contient une seule cytosine à son extrémité 3’ (385). Des études phylogénétiques prédisent que la séquence R de l’ARN forme les tiges-boucles TAR et poly(A) (386, 387).

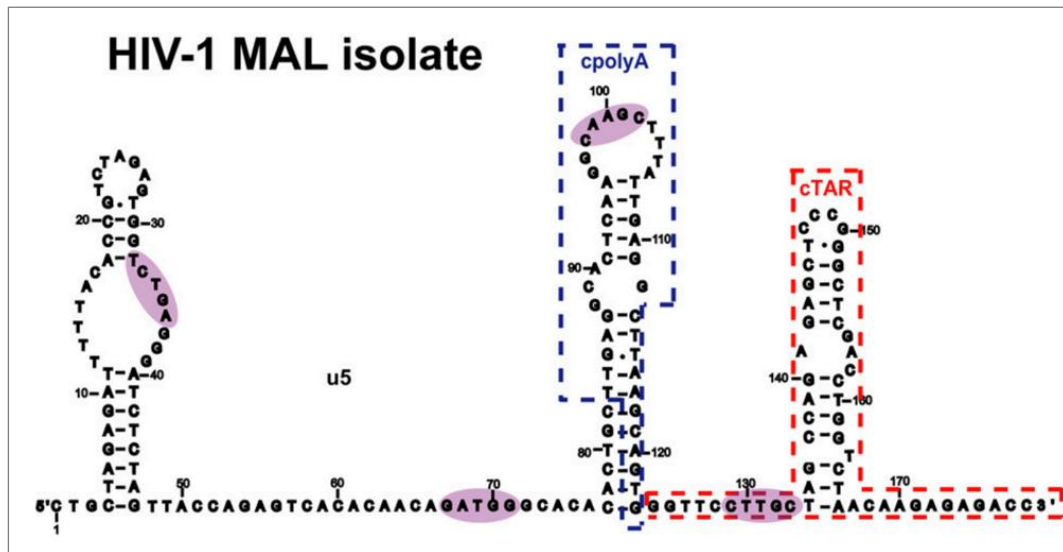


Figure 36. Structure secondaire prédite pour l’ADNss de l’isolat MAL en présence de NC (384). Les ellipses mauves correspondent aux quatre sites forts de la NC. Les domaines cpoly(A) et cTAR sont encadrés.

Des études *in vitro* et *in virio* ont montré que la séquence R dans les 5’ et 3’UTR forme ces tiges-boucles (388, 389), suggérant que la copie de R présente dans l’ADNss forme aussi deux structures en tige-boucle nommées cpoly(A) et cTAR, en référence aux tiges-boucles analogues présentes dans l’ARN viral (390, 391). Notre équipe a été la première à montrer *in vitro* que l’ADNss de l’isolat MAL adopte une structure secondaire constituée de trois tiges-boucles dont l’analyse phylogénétique suggère qu’elle est conservée chez les différents isolats (384). Une tige-boucle est dans la région u5 et deux sont dans la région r (cpoly(A) et cTAR) (Figure 36). Avant l’entrée du virus dans la cellule-hôte, il y a très peu d’ADN viral synthétisé dans le virion environ une molécule d’ADNss pour 60 à 1000 molécules d’ARNg selon le type cellulaire produisant les virions (392–394). La transcription inverse dans le virion produit 100 fois moins de molécules d’ADN simple-brin de polarité négative (Figure 30, étape *Minus strand synthesis*) que de molécules d’ADNss (345).

Lors de la synthèse de l'ADNss par l'activité polymérase de la RT, les régions copiées R et U5 sont dégradées par son activité RNase H. L'ADNss peut donc se désengager de sa matrice ARN et se replier avant de s'apparier en partie via sa séquence r avec la séquence R complémentaire située à l'extrémité 3' de l'ARNg. Notre équipe a identifié *in vitro* quatre sites forts de fixation de la NC sur l'ADNss structuré, confortant l'idée qu'elle déstabiliserait les tiges-boucles cTAR et cpoly(A) afin de permettre la formation de l'hétéroduplex ADN-ARN de 95 pb qui est nécessaire au premier transfert de brin (130).

3.4 Le premier transfert de brin

Le premier transfert de brin correspond au passage de l'ADNss de l'extrémité 5' à l'extrémité 3' de l'ARNg. Il peut se produire sur la même molécule d'ARNg ou sur la seconde incorporée dans le virion et permet la synthèse du brin (-) de l'ADN viral. Il est basé sur la complémentarité entre les séquences r de l'ADNss et R du 3'UTR de l'ARNg (Figure 30). Cette partie sera détaillée dans le chapitre 4.

3.5 Synthèse du brin (-) d'ADN viral

Une fois le premier transfert de brin réalisé, la RT catalyse la polymérisation de l'ADNss tout en dégradant par son activité RNase H la matrice d'ARNg qui a été copiée (Figure 30, étape *Minus strand synthesis*). La synthèse du brin (-) d'ADN viral peut-être ralentie lorsque la RT rencontre des structures secondaires formées par l'ARNg. Les études *in vitro* montrent que la NC peut réduire les pauses de la RT en déstabilisant ces régions structurées (395, 396).

3.6 Initiation de la synthèse du brin (+) d'ADN viral

Chez tous les rétrovirus, l'initiation du brin (+) d'ADN nécessite au minimum une séquence constituée de purines située juste en amont de la séquence U3. Au sein de l'ARNg du VIH-1, il existe deux séquences identiques constituées uniquement de purines (5'AAAAGAAAAGGGGGG-3) nommées *poly purine tracts* (PPTs) (Tableau 2) (397).

3'PPT	CUUUUAAAAGAAAAGGGGGGACUGGAAG
cPPT	AAUUUAAAAGAAAAGGGGGGAUUGGGGG

Tableau 2 : Séquences PPT et cPPT du HIV-1 NL4-3 (GenBank accession number : [AF324493](#)).

Le groupe de Weeks a déterminé au moyen de la méthode SHAPE la structure secondaire de l'ARNg du VIH-1 dans le virion (389). Suite à une optimisation de la méthode SHAPE en 2013, il a apporté des modifications sur le repliement de certains éléments dont les PPT (Figure 37) (398).

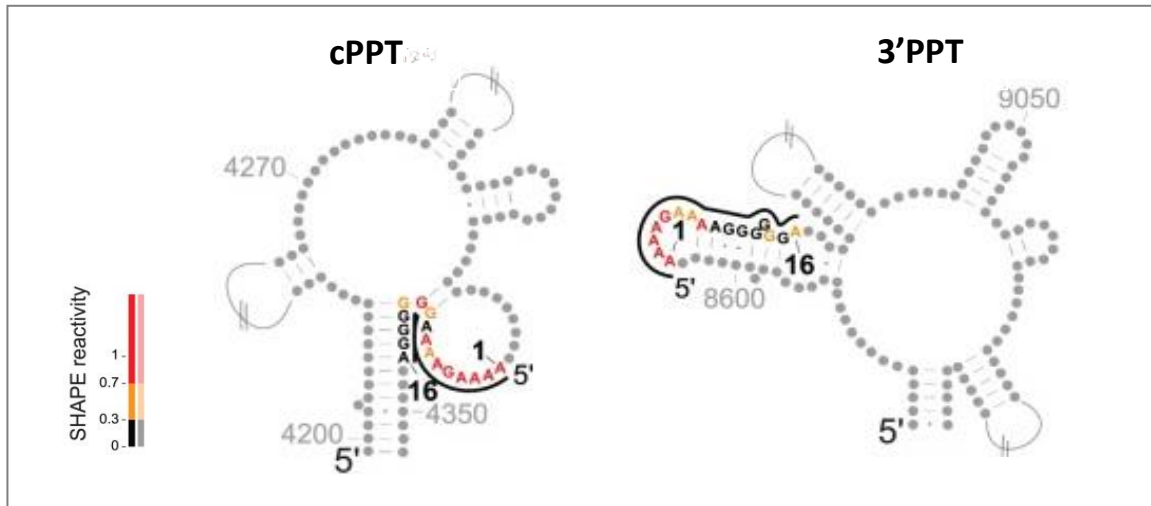


Figure 37. Structure secondaire du cPPT et du 3'PPT (398). En couleur sont représentées les séquences PPT selon leur réactivité au SHAPE.

Les séquences 3'PPT et cPPT sont situées respectivement à l'extrémité 5' du U3 et au centre du génome dans le gène de l'intégrase virale (Figure 30). Ces séquences sont utilisées comme amorces par la RT pour la synthèse du brin (+) d'ADN viral. Pendant la synthèse du brin (-) d'ADN viral, les motifs PPT de l'ARNg sont résistants au clivage par la RNase H tandis que le reste est dégradé. Un clivage précis de la jonction 3'PPT-U3 est important car il définit l'extrémité 3'LTR servant de substrat à l'intégrase virale (399). Les amorces PPT présentant des clivages aberrants ou initiant mal la synthèse du brin (+) d'ADN empêchent un déroulement efficace de la transcription inverse.

La répétition des quatre adénines contribuerait à la reconnaissance des PPT et au positionnement du site actif de l'activité polymérase de la RT (400). Le duplex formé par l'amorce PPT d'ARN et l'ADN (-) peut être fixé par la RT dans les deux orientations qu'elle peut adopter, lui conférant soit son activité RNase H soit son activité de polymérase. Le PPT n'est pas intrinsèquement résistant au clivage puisqu'il peut être hydrolysé par la RNase H1 cellulaire et l'activité RNase H de la RT du VIH-1 (401). Les propriétés conférant la résistance au clivage lors du processus de transcription inverse n'ont pas encore été clairement identifiées. Une équipe a montré dernièrement que la séquence PPT poly(rA/dT) adoptait une conformation

d'une rigidité importante inadéquate pour le clivage par l'activité RNase H à n'importe quelle position (402). Pour générer l'amorce PPT par le domaine RNase de la RT, le PPT doit subir un changement de conformation sans distorsion pour permettre l'hydrolyse (400, 402).

La synthèse du brin (+) d'ADN à partir du 3'PPT a lieu après la synthèse du brin (-) d'ADN et la dégradation de la matrice d'ARNg (+). La RT allonge l'amorce 3'PPT en utilisant comme matrice le brin (-) d'ADN, encore lié à ARNt^{Lys3} (Figure 38) (401). Le produit de cette extension d'ADN est constitué des régions U3-R-U5-PBS et est nommé ADN *strong-stop* (+) (ADNss(+)). La RT ne copie pas l'intégralité de l'ARNt^{Lys3} car la polymérisation est stoppée par la présence d'une adénine méthylée en position 58, mais reconstitue la séquence du PBS (335, 403, 404). L'élimination de l'ARNt^{Lys3} nécessite un clivage primaire par la RNase H qui ne se produit pas à la jonction ARNt-ADN mais au niveau de l'extrémité 3' de l'ARNt^{Lys3} entre l'adénine et la cytosine adjacente. Cela conduit à la présence d'un ribonucléotide (rA) et non d'un désoxyribonucléotide à l'extrémité 5' de l'ADN (-) et la formation d'un duplex transitoire ADN-ARN de dix-sept paires de bases entre l'extrémité 3' de l'ARNt et la séquence PBS de l'ADNss (+) (291).

3.7 Le second transfert de brin

3.7.1 Acides nucléiques intervenant dans le second transfert de brin

Le second transfert de brin est nécessaire à la synthèse complète du brin (+) de l'ADN viral (Figure 30, étape *second strand transfer*). C'est un processus correspondant au passage de l'ADNss (+) de l'extrémité 5' à l'extrémité 3' du brin d'ADN(-). Il repose sur l'appariement entre la séquence PBS de l'extrémité 3' de l'ADNss (+), appelée PBS (+), et la séquence complémentaire située à l'extrémité 3' de l'ADN(-), appelée pbs(-) ou cPBS (Figure 38). Comme il existe deux copies d'ARNg dans chaque virion, les transferts de brin peuvent être intra ou intermoléculaire. Il était jusqu'à présent considéré que le second transfert de brin est majoritairement intramoléculaire (405) mais une étude récente basée sur l'analyse des séquences de virus recombinants suggère que ce transfert peut être intra ou intermoléculaire (406).

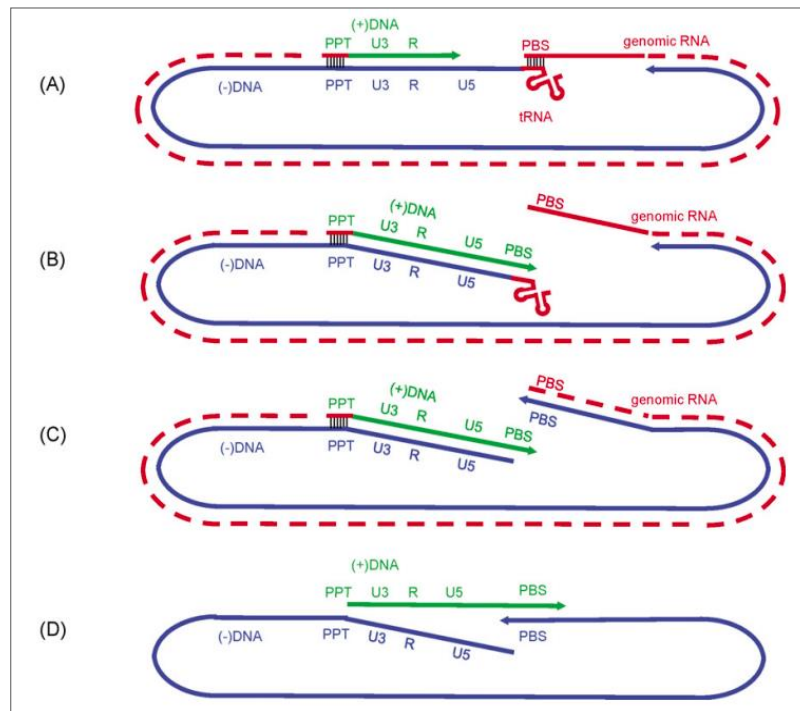


Figure 38. Étapes conduisant au second transfert de brin chez tous les rétrovirus (349). (A) Initiation de la synthèse du brin (+) d'ADN (en vert) grâce à l'amorce ARN 3'PPT. En parallèle, la dégradation de l'ARNg (en rouge) par l'activité RNase H de la RT se poursuit. (B) L'ADN (+) en cours de synthèse copie le pbs de l'ARNt. (C) L'ARNt est dégradé par l'activité RNase H de la RT. (D) La séquence PBS de l'ADN (+) s'apparie avec la séquence pbs de l'ADN (-) : cette étape constitue le second transfert de brin. Le cPPT du VIH-1 n'a pas été représenté.

3.7.2 Mécanismes intervenant dans le second transfert de brin

L'activité RNase H de la RT et l'appariement de la séquence PBS de l'ADN (+) avec la séquence pbs de l'ADN (-) sont indispensables à la réalisation du second transfert de brin (Figure 38). En effet, après avoir copié la séquence pbs de l'ARNt^{Lys3}, la RT via son activité RNase H dégrade l'ARNt libérant ainsi l'extrémité 5' de l'ADN (-) (Figure 38). L'activité chaperon des acides nucléiques de la NC jouerait un rôle dans l'élimination de l'extrémité 3' de l'amorce par dissociation des courts duplex ARN-ADN générés par l'activité RNase H (407, 408). Plusieurs études *in vitro* ont montré que l'ajout de la NC dans des tests de transcription inverse augmente de deux à trois fois le second transfert de brin (409, 410). La NC déstabilise la structure en tige-boucle qu'adoptent les 18 nt du pbs à l'extrémité 3' de l'ADN (-). Cette déstabilisation facilite l'appariement du pbs avec le PBS de ADNss (+) (411). Une autre étude (297) a montré que la fixation de la NC sur le pbs ouvre sa boucle apicale favorisant ainsi la formation d'un complexe boucle-boucle avec la tige-boucle PBS de l'ADNss (+). Cette étude suggère que la NC déstabilise via son doigt de zinc N-terminal la paire de bases adjacente à la boucle apicale. D'après les travaux du groupe de Mély, les tiges-boucles pbs et PBS possèdent trois sites sur lesquels se fixe la NC (412).

Une étude par fluorescence a montré que la NC multiplie par 10 la cinétique d'hybridation des deux tiges-boucles (413). La même étude a montré qu'une forme tronquée de NC multiplie par 6 la cinétique d'hybridation des tiges-boucles complémentaires en favorisant l'interaction boucle-boucle.

3.8 Synthèse complète du brin (+) et des deux LTR

Le second transfert de brin conduit à la formation d'une forme circulaire intermédiaire permettant à la RT d'allonger les brins d'ADN viral (Figure 39) et notamment de synthétiser les deux séquences LTR. La synthèse discontinue du brin (+) d'ADN conduit à la formation d'une structure triple nommée *central flap* avec deux segments d'ADN double-brin et un segment du brin (+) d'ADN non apparié d'une centaine de nucléotides (Figure 39) (406). Le segment non apparié dans ce triplex est issu du cPPT tandis que l'ADN double-brin en amont a été initié à partir du 3'PPT.

Des enzymes cellulaires seraient impliquées dans l'élimination du *central flap* et la réparation des coupures dans l'ADN double-brin mais cela n'a pas été démontré *in cellula* (414). Une étude suggère que le *central flap* aurait un rôle dans la décapsidation au niveau des pores nucléaires (124). Deux études indépendantes ont présenté des résultats indiquant que le cPPT protège le VIH-1 de l'activité des facteurs de restriction APOBEC3 humains (415, 416). La transcription inverse s'achève par un déplacement de brin, autrement dit, une synthèse d'un brin d'ADN à travers une région d'ADN double-brin. Cette étape est facilitée *in vitro* par la NC (417–419). Un échec dans la finalisation de la synthèse par déplacement de brin entraînerait la formation d'extrémités LTR incomplètes créant un ADN double-brin ne pouvant être intégré dans le génome cellulaire.

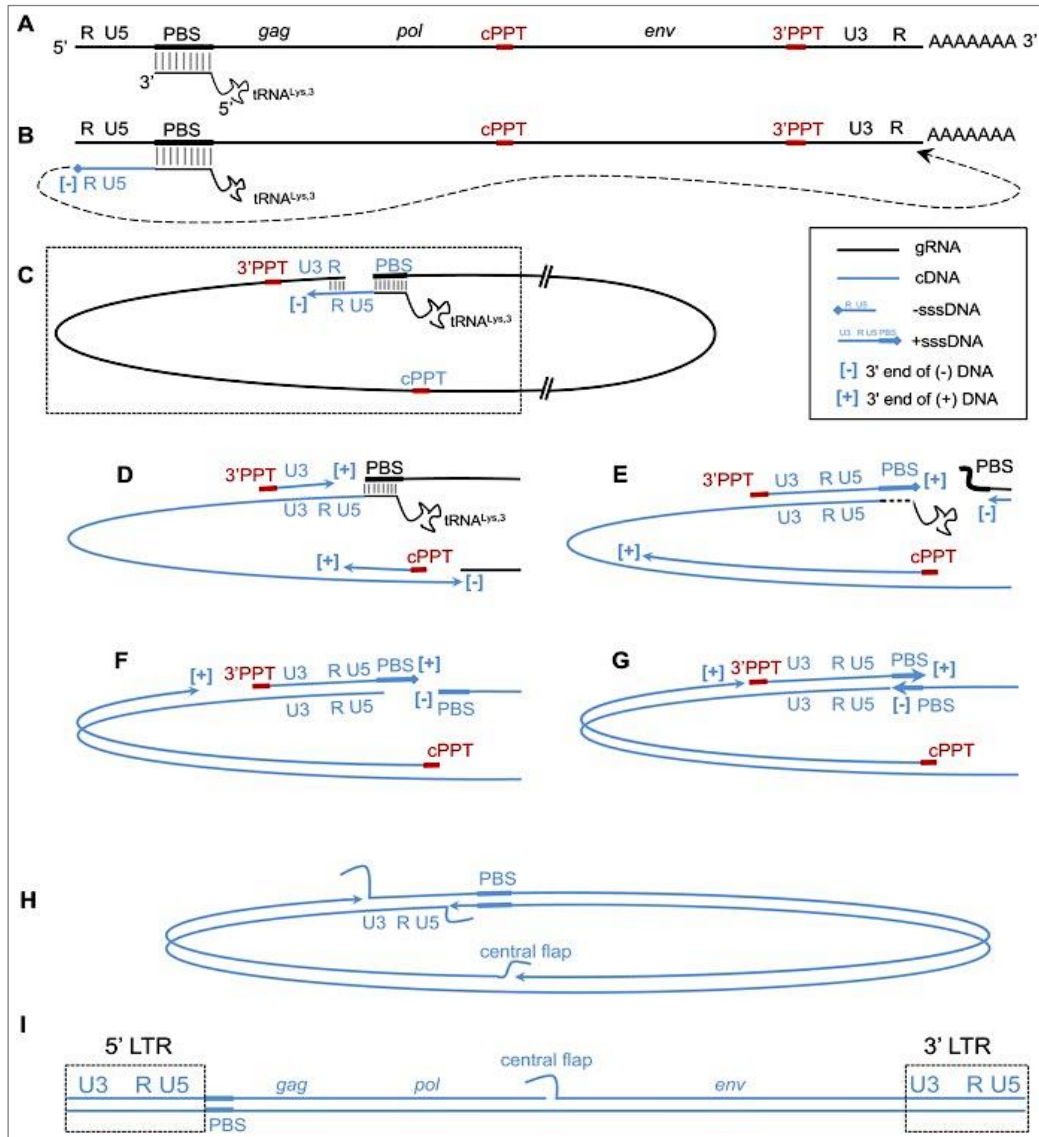


Figure 39. Transcription inverse chez le VIH-1 Adapté de (406). **(A)** Brin d'ARNg (en noir) avec les trois principaux gènes, les régions non traduites R, U5 et U3, le 3'PPT et le cPPT (en rouge) et l'ARNt servant d'amorce hybridé au PBS de l'ARNg. **(B)** Synthèse de l'ADNss (-) (en bleu). Le premier transfert de brin est illustré par la flèche pointillée. **(C)** Le premier transfert de brin conduit à la circularisation de l'ARNg permettant la poursuite de la synthèse de l'ADN (-). Le contenu de l'encadrement correspond aux étapes D à G. **(D)** La synthèse du brin (+) d'ADN initiée au niveau du 3'PPT se produit de l'extrémité 5' à l'extrémité 3' de la matrice ADN (-). La synthèse du brin (+) d'ADN est aussi initiée au niveau du cPPT. **(E)** La synthèse du brin d'ADN (+) commencé à partir du 3'PPT se prolonge jusqu'à la copie des 18 nucléotides de l'extrémité 3' de l'ARNt qui nécessite le déplacement du PBS de l'ARNg ; cela génère le fragment d'ADNss (+). **(F)** Le PBS de l'ARNg est dégradé par la RT après avoir été copié. **(G)** L'ADNss (+) se déplace de l'extrémité 5' vers l'extrémité 3' du brin d'ADN (-) grâce à la complémentarité des séquences PBS. Cette étape correspond au second transfert de brin. **(H)** La synthèse d'ADN (+) initiée à partir du 3'PPT atteint le brin d'ADN (+) débutée au cPPT et ne le déplace que partiellement avant de s'arrêter à la *central terminal sequence* et générant ainsi le *central flap* ; la synthèse du brin d'ADN (-) se poursuit en déplaçant le brin (+) pour générer le 5' LTR ; la synthèse d'ADN (+) initiée à partir du cPPT se poursuit en déplaçant le brin d'ADN (+) pour générer le 3' LTR. **(I)** ; ADN pré-proviral complet contenant les deux LTR et le *central flap*.

3.9 Recombinaisons chez le VIH-1

3.9.1 La recombinaison dépend du transfert de brin interne

Les rétrovirus recombinent à une fréquence élevée. C'est notamment le cas pour le VIH-1 qui présente à travers le monde de nombreux isolats générés par recombinaison (258, 420, 421). Le génome viral du VIH-1 étant constitué de deux copies d'ARNg, les transferts de brin peuvent se produire de manière intra ou intermoléculaire (422). La recombinaison génétique résulte de la présence dans un virion de deux molécules d'ARNg qui ne sont pas identiques. En effet, ce type de virion peut être produit par une cellule infectée par deux virus ayant des séquences très proches mais non identiques (Figure 40) (423). La diversité génétique générée par recombinaison peut conférer certains avantages aux virus comme un mécanisme de résistance aux facteurs antirétroviraux de l'hôte et aux agents pharmacologiques (396, 398–400).

En plus des deux transferts de brins terminaux pour générer un ADN viral double-brin (Figure 30), la RT change fréquemment de matrice au niveau des régions internes du génome pendant la synthèse du brin d'ADN (-). Ces transferts de brin internes sont principalement responsables de la recombinaison rétrovirale (Figure 41) (427). *In vitro*, en l'absence de NC, l'efficacité des transferts est plus importante avec des matrices ARN faiblement structurées qu'avec des matrices hautement structurées (428). La NC augmente légèrement le transfert lorsque les matrices sont faiblement structurées tandis qu'elle l'augmente significativement dans le cas de matrices très structurées (429–432). L'activité déstabilisatrice de la NC via ses doigts de zinc ainsi que la structure de l'ARN accepteur jouent des rôles essentiels dans le processus de transfert de brin (433–435).

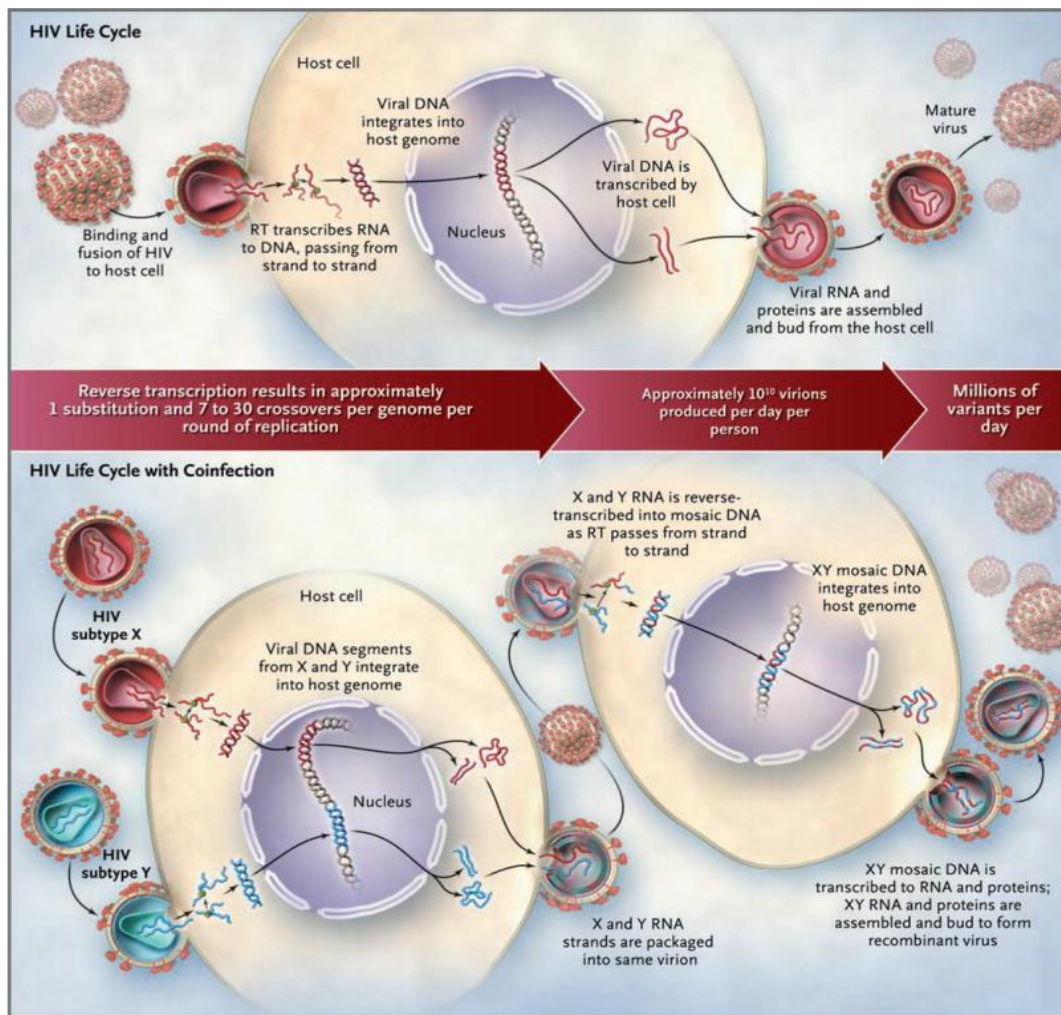


Figure 40. Évolution de la diversité génétique chez le VIH-1 au cours du cycle réplcatif viral et création de formes recombinantes uniques dans le cas d'une co-infection par deux sous-types (423).

Il y a environ 5 à 14 événements de recombinaison par génome et par cycle de transcription inverse *in vivo* (436). Lorsque les deux ARNg sont génétiquement divergents, un transfert de brin interne peut conduire à la production d'un virus recombinant (Figure 40) (423). Les événements de transfert de brin intermoléculaire conduisent à une recombinaison majoritairement homologue et à de la recombinaison non homologue qui est 100 à 1000 fois moins fréquente que la recombinaison homologue (437–439). Les événements de transfert de brin intramoléculaires peuvent générer des mutations (duplications, suppressions, suppressions avec insertions) (438). De plus, la faible fidélité de la RT crée des mutations pouvant être propagées rapidement par les transferts de brin internes (Figure 41) (427).

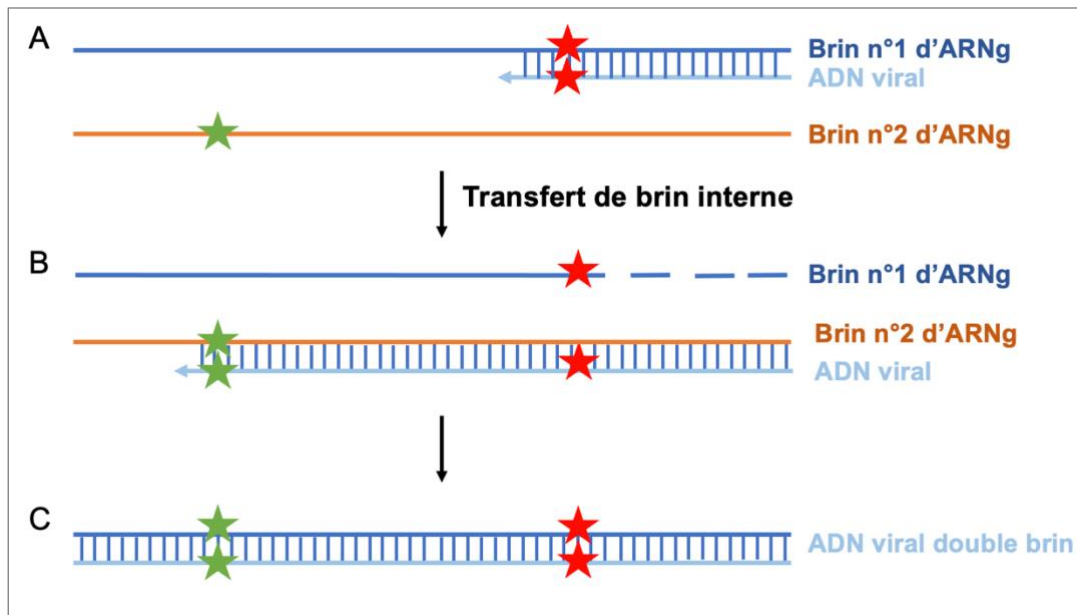


Figure 41. Présentation schématique de la genèse de la diversité du VIH-1. (A) Mutation « orange » présente sur le brin n°1 d'ARNg copiée par la RT lors de la synthèse du brin (-) d'ADN viral. (B) Transfert de brin intermoléculaire et interne conduisant à la copie de la mutation « verte », qui est présente sur le brin n°2 d'ARNg ; ce transfert permet la synthèse du brin (-) d'ADN viral contenant deux mutations. (C) Obtention de l'ADN viral double-brin et doublement muté après synthèse du brin (+).

3.9.2 Modèles de recombinaison par transfert de brin interne

Trois modèles de recombinaison par transfert de brin interne sur la base des homologies de séquence sont actuellement décrits dans la littérature (439) :

1. Le *forced copy choice* (440) propose qu'en raison de coupures présentes dans le brin matrice de l'ARNg la RT stoppe sa progression au niveau des coupures avant de poursuivre la synthèse du brin (-) d'ADN sur le second brin d'ARNg encapsidé. L'ADN contiendra donc les informations génétiques des deux copies du génome viral.
2. Le *copy choice* ne nécessite pas la présence de coupure dans l'ARNg et repose sur la possibilité de changement de matrice pendant la synthèse du brin (-) d'ADN dans des régions non altérées du génome viral (441).
3. Le *dynamic copy choice* (442) propose que le changement de matrice n'est pas forcément causé par des cassures de l'ARNg mais dépendrait d'avantage de la dynamique de la RT. D'après ce modèle, le changement de matrice est influencé par la balance entre les activités polymérase et RNase H de la RT (443). Plus précisément, pendant la synthèse du brin (-) d'ADN, l'activité polymérase synthétise l'ADN tandis que l'ARNg de l'hybride ARN-ADN est dégradé par l'activité RNase H, exposant alors

l'ADN néosynthétisé. Celui-ci serait plus susceptible de s'hybrider avec le second brin d'ARNg encapsidé et de l'utiliser pour la suite de la synthèse de l'ADN viral.

Ce modèle découle d'études démontrant que la recombinaison est favorisée par une diminution de l'activité polymérase ou une augmentation de l'activité RNase H (443). Cette étude suggère que le complexe RT-matrice devient moins stable lorsque la synthèse d'ADN ralentit ou que l'activité RNase H dégrade plus rapidement la matrice d'ARNg, provoquant ainsi la dissociation de la RT du brin copié, lui permettant alors de se fixer sur la seconde matrice.

3.9.3 Mécanismes du transfert de brin interne

Le transfert de brin interne est reconstitué *in vitro* pour déterminer les mécanismes moléculaires qui régissent ce processus. Pour cela, un ARN donneur et un ARN accepteur sont utilisés en tant que matrices. La synthèse d'ADN en présence de RT et des quatre dNTP est initiée à partir d'une amorce hybridée à l'extrémité 3' de l'ARN donneur (435). L'extension de l'amorce s'effectue le long de cet ARN, homologue à l'extrémité 3' de l'ARN accepteur. Au cours de cette extension, l'ADN complémentaire néosynthétisé (ADNc) est transféré vers l'ARN accepteur servant de matrice pour la synthèse d'ADN (435).

Les différentes études *in vitro* suggèrent l'existence de trois mécanismes de transfert de brin interne :

a. Transfert de brin interne par invasion de brin

Le transfert de brin interne par invasion est un mécanisme caractérisé par une première interaction entre l'ADNc et l'ARN accepteur se produisant en amont de l'extrémité 3' de l'ADNc (349, 444). Le site en question est nommé « site d'invasion » et correspond à une région d'ADNc libérée par l'activité RNase H de la RT. Ce modèle (Figure 42) (445) repose sur le fait que l'ARN donneur pendant la synthèse d'ADN est clivé par l'activité RNase H de la RT, libérant ainsi l'ADN néosynthétisé qui peut s'apparier avec l'ARN accepteur. La synthèse d'ADN se poursuit en copiant le brin accepteur dès que la longueur du duplex ARN accepteur-ADN est suffisante pour permettre un appariement stable. *In vitro*, la NC facilite le transfert de brin interne en déstabilisant les structures de l'ADN et de l'ARN accepteur et permettre ainsi leur appariement (444).

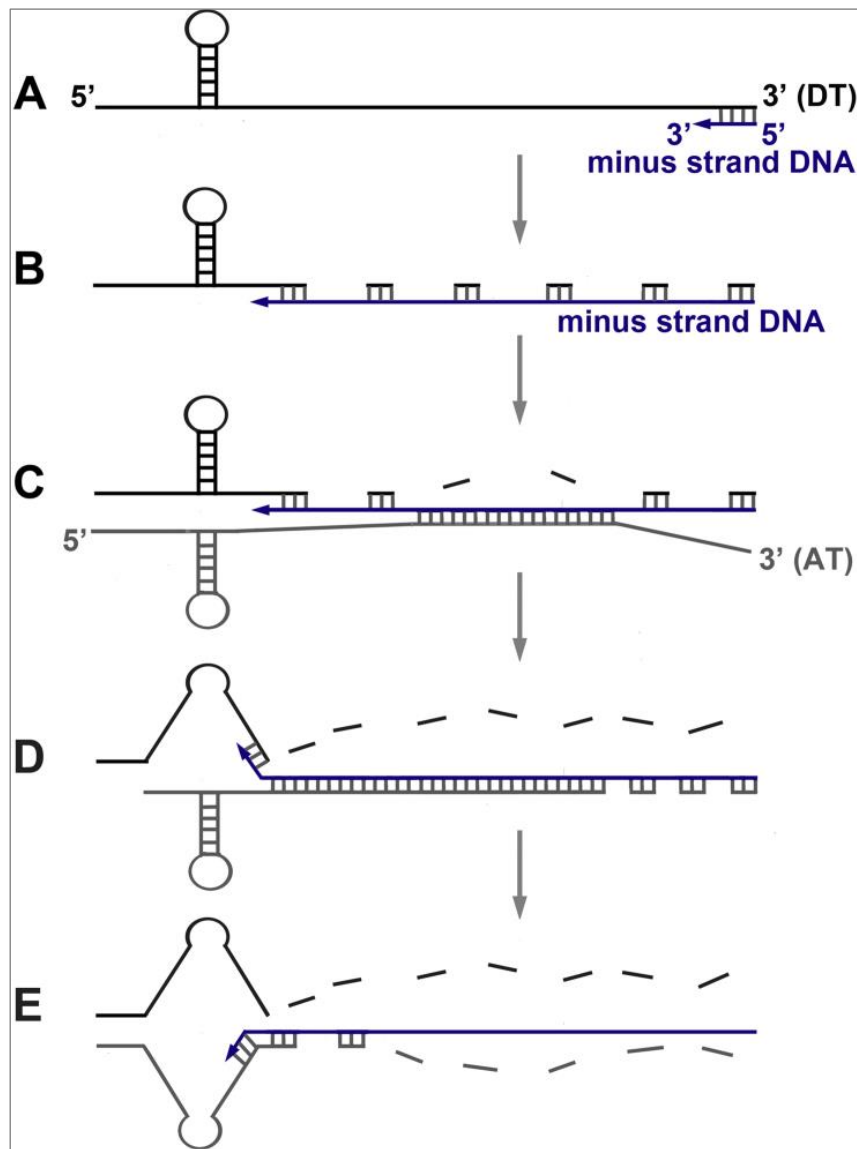


Figure 42. Mécanisme de transfert de brin par invasion (445). (A) Le brin d'ADN (-) (en bleu) est hybridé à la matrice d'ARN donneur (*donor template*, DT en noir). (B) La RT synthétise le brin d'ADN et clive en parallèle par son activité RNase H l'ARN donneur. (C) À la suite des coupures, une partie de l'ADN n'est plus appariée à l'ARN donneur et peut donc s'hybrider en partie avec l'ARN accepteur (*acceptor template*, AT en gris). Cette étape est appelée « invasion » ou « docking ». (D) L'appariement de l'ADN avec l'ARN accepteur se propage et déplace les fragments d'ARN clivés. (E) La synthèse d'ADN se poursuit jusqu'à l'extrémité 5' de l'ARN accepteur

b. Transfert de brin interne pause-dépendant

Lors de la synthèse d'ADN, les structures secondaires de l'ARNg tels que les tiges-boucles et les tétrades de G peuvent ralentir la RT et ainsi favoriser le clivage de l'ARN donneur par l'activité RNase H (378, 379, 446–448). Cela facilite ainsi l'interaction entre l'ADN et l'ARN accepteur (435, 449).

c. Transfert de brin interne pause-indépendant

En l'absence d'une forte pause de la RT au niveau de l'ARN donneur une interaction entre l'ARN accepteur et l'ADNc en cours de synthèse est possible dans certains cas et peut permettre par conséquent un transfert de brin interne efficace (428). Par exemple, *in vitro* en l'absence de NC, lorsqu'une matrice ARN est faiblement structurées comme c'est le cas de la région pol-vif, l'efficacité du transfert est significativement plus élevée qu'avec une matrice ARN très structurées comme la région gag-pol (428).

Des études *in vitro* ont montré que des mutations ou une délétion de la tige-boucle SL1 (contenant le DIS, responsable de la dimérisation de l'ARNg) peut conduire à une diminution d'environ quatre fois le taux de transfert de brin interne *in vitro* et par conséquent la recombinaison entre ces deux ARN encapsidés dans le même virion (450, 451). De plus, deux études montrent que le taux de recombinaison augmente lorsque les deux ARNg encapsidés présentent chacun un SL1 complémentaire qui en rapprochant les deux molécules d'ARNg faciliterait les transferts de brin interne (309, 451).

3.10 Rôle de la RNA hélicase A dans la transcription inverse

Lors du bourgeonnement de la nouvelle particule virale, de nombreuses molécules de la cellule hôte infectée vont y être incorporées. C'est notamment le cas des isoaccepteurs cellulaires 1, 2 et 3 de l'ARNt lysine, de nucléotides ou de protéines (452, 453). La RNA hélicase A humaine (RHA) est incorporée dans les virions selon une stœchiométrie proportionnelle à l'ARNg (490, 491). Cette enzyme est capable de dérouler à la fois l'ARN double-brin et l'ADN double-brin dans le sens 3' vers 5' (454, 455). Elle possède deux domaines de liaison à l'ARN double-brin dans sa partie N-terminale, un domaine central lui conférant son activité d'hydrolyse des nucléotides triphosphates et sa partie C-terminale possédant la séquence RGG répétée qui a une affinité plus élevée pour l'ARN simple-brin que pour l'ARN double-brin (455). La RHA intervient dans les étapes du métabolisme des ARN cellulaires comme la transcription et la traduction (359) mais elle est aussi impliquée dans des réponses immunitaires liées à la présence de pathogènes dont les virus (359, 456). L'infectiosité d'une particule virale diminue lorsqu'elle ne contient pas de RHA (457) car cette hélicase intervient dans plusieurs étapes du cycle de réplication du VIH-1 tels que la transcription inverse (458, 459), la transcription (460), la traduction (457) et l'export des ARNm viraux (461). La RHA favorise la transcription inverse du VIH-1 en facilitant l'hybridation de l'ARNt^{Lys3} avec la séquence PBS (458). Il a été montré que la RHA se fixe sur une structure ARN formant une jonction à trois hélices et contenant la

séquence PBS (371, 462). La RHA joue aussi un rôle positif dans la transcription inverse en augmentant la processivité de la RT pendant la phase d'élongation de la synthèse de l'ADN (459).

3.11 Inhibition de la transcription inverse par les agents pharmacologiques

Il existe aujourd'hui plusieurs types d'agents pharmacologiques pour traiter les personnes infectées par le VIH. En 2019, on compte une trentaine de médicaments antirétroviraux approuvés par l'agence américaine des médicaments (FDA) (463). La prise en charge des patients le plus tôt possible permet de faire de l'infection par le VIH une maladie chronique mais gérable. Les traitements actuels sont à prendre tout au long de la vie du patient pouvant ainsi contribuer à l'apparition de résistances. Ces traitements ciblent les différentes étapes du cycle de réplication du virus. Ainsi, sont aujourd'hui commercialisés les inhibiteurs d'entrée du virus, de RT, de protéase et d'intégrase.

3.11.1 Inhibiteurs ciblant la RT

Deux types d'inhibiteurs de RT sont actuellement commercialisés pour empêcher la progression de la transcription inverse. Ce sont les inhibiteurs nucléosidiques/nucléotidiques (INTI) et les inhibiteurs non nucléosidiques (INNTI). Les INTI sont des analogues nucléosidiques/nucléotidiques qui sont dépourvus du 3'OH du ribose nécessaire à l'ajout du prochain nucléotide ; ils bloquent ainsi la synthèse de l'ADN par la RT (464). Les INNTI ciblent directement la transcriptase inverse en se fixant près du site actif de polymérisation et empêchent les mouvements de l'enzyme (465). Un nouvel inhibiteur nucléosidique prometteur a été identifié et est actuellement en phase d'essai clinique, c'est un inhibiteur de la translocation des nucléosides de la RT (466–469). Une nouvelle classe d'agents pharmacologiques agissant sur la fonction polymérase de la RT est aussi développée (376, 470, 471).

L'inhibition de l'activité RNase H de la RT conduirait à l'inhibition du premier transfert de brin et par conséquent à l'inhibition de la transcription inverse. Parmi les antirétroviraux utilisés par les patients infectés par le VIH, il n'en existe pas à ce jour qui ciblent l'activité RNase H de la RT. En effet, plusieurs obstacles subsistent pour développer des inhibiteurs ciblant l'activité RNase H de la RT. Les activités RNase H et polymérase de la RT sont étroitement liées dans le temps et dans l'espace. Le substrat naturel, l'hybride ADN-ARN, interagit plus fortement avec le domaine polymérase qu'avec le domaine RNase H (261). De plus, l'hybride ADN-ARN empêcherait les petites molécules d'atteindre le site actif du domaine RNase H (472).

Il faut également souligner que le domaine RNase H de la RT appartient à la superfamille des intégrases rétrovirales, il possède un haut degré de similitudes structurales et mécanistiques avec l'intégrase virale (dépendance à deux ions Mg^{2+} et motif du site actif) (473). Actuellement, plusieurs composés ont montré une efficacité en terme d'inhibition de la RNase H mais ils avaient une faible activité antivirale *ex vivo* et/ou avait une toxicité importante en culture cellulaire (472). Des études rationnelles avec une approche basée sur les pharmacophores chélateurs de métaux et la structure du domaine RNase H sont en cours et pourraient conduire à la conception d'inhibiteurs spécifiques de l'activité RNase H de la RT (376, 472). D'autres inhibiteurs potentiels dérivés de la coumarine sont également étudiés (474).

3.11.2 Inhibiteurs ciblant la NC

Grâce à ses fonctions pléiotropes et sa haute conservation dans plusieurs sous-types du VIH-1, la NC est une cible attractive pour la conception de médicaments antirétroviraux. Les études *in vitro* suggèrent que la NC joue un rôle important lors du premier et du second transfert de brin qui sont des étapes cruciales de la transcription inverse (291, 358). De plus, la production de virions non infectieux lors de mutations ponctuelles au niveau des doigts de zinc de la NC ne s'explique pas uniquement par une encapsidation diminuée de l'ARNg dans la particule virale (475, 476). L'inhibition de la NC devrait donc empêcher ou réduire fortement le premier et le second transfert de brin (290, 358, 379, 477). Des inhibiteurs dits éjecteurs de zinc (inhibiteurs covalents), des inhibiteurs non covalents se fixant à la NC, ou encore des inhibiteurs non covalents compétiteurs de la NC ont été conçus ces dernières années (478–483).

Les éjecteurs de zinc diminuent l'affinité de la NC pour les acides nucléiques et provoquent un dépliement de la NC induisant alors une activité antivirale importante mais aussi des effets cytotoxiques (481). Ceci peut s'expliquer par la faible sélectivité de ces inhibiteurs qui ciblent également les protéines cellulaires possédant des doigts de zinc (480, 482). Les éjecteurs de zinc sont plutôt envisagés pour une application microbicide dans le but de prévenir la transmission du VIH (480, 484–486). Des molécules de types inhibiteurs non covalents sont étudiées ces dernières années mais n'ont pas encore atteint les phases pré-cliniques. La caractérisation des interactions et les informations structurales suggèrent que la NC a le potentiel pour être une cible thérapeutique des inhibiteurs non covalents. Leur mode d'action serait plus spécifique que les inhibiteurs éjecteurs de zinc car ils ciblent l'activité chaperon de la NC soit en entrant en compétition avec l'acide nucléique viral en se fixant sur la NC, soit en se fixant sur l'acide nucléique complexé avec la NC afin d'en perturber l'interaction, soit en se fixant sur l'acide nucléique qui n'est pas encore lié à la NC pour empêcher sa fixation (478, 480,

481, 487–489). Une étude sur une classe d'inhibiteurs non covalents a identifié un dérivé pyrimidique, le 5,6-dihydroxypyrimidine, avec des propriétés antivirales prometteuses notamment un bon profil pharmacocinétique chez la souris (490, 491). Une optimisation est tout de même nécessaire. Des études préliminaires avec d'autres molécules sont actuellement en cours (489, 490, 492, 493).

En 2015, un nouveau composé inhibant directement la fonction chaperon de la NC a été identifié, le A1752 (481, 488). En présence de ce dernier, les particules virales produites sont non infectieuses et présentent des défauts de transcription inverse et de décapsidation. Récemment, des calixarènes ont montré une inhibition de l'activité chaperon de la NC par fixation à sa région N-terminale basique (494). Ces composés entrant en compétition avec les acides nucléiques diminuent l'infectiosité des souches sauvages et des souches résistantes aux antirétroviraux et présentent une faible cytotoxicité. Ils peuvent également inhiber la polymérisation et le changement d'orientation de la RT. Cependant, l'étude de Humbert *et al.* (2020) sur un modèle animal n'a pas montré un effet significatif. Une clairance importante et une faible distribution dans les tissus en serait responsable. Ces composés à base de calixarène sont fortement polarisés ce qui peut jouer sur leur perméabilité. Des modifications chimiques sont envisagées pour améliorer la pharmacocinétique de ces composés.

3.11.3 Inhibiteurs ciblant les acides nucléiques viraux

Afin de cibler plus spécifiquement les étapes de la réplication du VIH qui dépendent probablement de l'action de la NC, des inhibiteurs se fixant spécifiquement sur des structures des acides nucléiques viraux sont développés (482). Les tiges-boucles TAR et SL de la région Ψ de l'ARN sur lesquelles se fixe la NC sont les principales cibles. Cette approche est prometteuse du fait d'une spécificité plus importante de ce type d'inhibiteurs et d'une faible cytotoxicité observée pour certains composés. Néanmoins, les composés testés jusqu'à ce jour ne présentent pas un effet antiviral suffisamment important pour conduire à des applications cliniques. Plusieurs composés sont actuellement étudiés par le groupe de Gatto, notamment des dérivées d'antraquinones (495, 496) ciblant la tige-boucle TAR de l'ARNg qui est un domaine de la région R nécessaire au premier transfert de brin. Ces composés sont conçus pour se fixer sur TAR et empêcher ainsi la fixation de la NC. D'autres composés, les bis-chloropiperidines ont été conçus pour stabiliser la tige-boucle TAR par alkylation et empêcher ainsi son ouverture par la NC, ouverture qui est nécessaire au premier transfert de brin (497, 498). Ces composés ciblant les acides nucléiques viraux nécessitent d'autres études pour envisager d'éventuels tests précliniques.

4. Le premier transfert de brin

Le premier transfert de brin a lieu de l'extrémité 5'UTR vers l'extrémité 3'UTR de l'ARNg viral après la synthèse de l'ADNss par l'activité polymérase de la RT. Pendant la synthèse de l'ADNss, la partie R-U5 de l'ARNg est dégradée par l'activité RNase H de la RT et libère ainsi l'ADNss (Figure 30). Ce dernier s'apparie via sa région r avec la région R complémentaire de l'extrémité 3'UTR. L'appariement de ces séquences complémentaires est indispensable à la poursuite de la transcription inverse et à la réplication virale. Les particules virales possèdent deux copies d'ARNg, le premier transfert de brin peut se produire de manière intra- ou intermoléculaire (499, 500). Les études traitant du mécanisme du premier transfert de brin ont été principalement réalisées au moyen de systèmes *in vitro*. On observe d'ailleurs que les transferts de brin sont peu efficaces *in vitro* (347). L'isolat NL4-3 du VIH-1 est l'un des isolats les plus utilisés pour les études *ex vivo* et *in vitro*. L'ADNss du NL4-3 (180 nt) possède les régions r (95 nt) et u5 (85 nt) (Figure 36) qui sont respectivement complémentaires des régions R et U5 du 3'UTR. L'alignement des séquences 5' et 3'LTR du NL4-3 montrent quelques différences mais les régions R en 5' et 3' sont presque identiques.

4.1 Structures et interactions intervenant dans le premier transfert de brin

4.1.1 Structures ARN

a. Repliement de la séquence R

La longueur de la séquence R varie considérablement chez les rétrovirus, de 16 nt chez le MMTV (*Mouse mammary tumor virus*), de 95 nt chez l'isolat NL4-3 du VIH-1, jusqu'à 247 nt chez le HLTV-2 (501, 502). Dans la majorité des cas, le premier transfert de brin chez le VIH-1 se produit *ex vivo* après la copie complète de la séquence R située en 5' (503). La réplication du VIH-1 est cependant possible avec des séquences R tronquées (504). Deux études *in vitro* suggèrent que la structure adoptée par R joue un rôle important dans le premier transfert de brin chez le VIH-1 (505, 506).

Les études *in vitro* et *ex vivo* ont montré que la séquence R peut se replier en formant les tiges-boucles TAR et poly(A) en 5' et en 3' de l'ARN viral (388, 389). Au niveau de l'extrémité 5' de l'ARNg, la tige inférieure de la tige-boucle poly(A) est allongée grâce à un appariement avec une partie de la région U5 (366, 389). Au niveau du 3'UTR, la tige-boucle TAR a été caractérisée dans l'ARNg extrait du virion tandis que la tige-boucle poly(A) n'a pas pu l'être (389). Toutefois, elle a été caractérisée *in vitro* (388). L'étude *in vitro* d'un ARN de 104 nucléotides au moyen de la diffusion des rayons X aux petits angles (SAXS) suggère qu'au

niveau tridimensionnel les tiges-boucles TAR et poly(A) adoptent une structure coaxiale (Figure 43) (507). Les éléments de la séquence R jouent d'autres rôles importants pour la réplication du virus, comme c'est le cas de la tige-boucle TAR qui est indispensable à l'activation de la transcription.

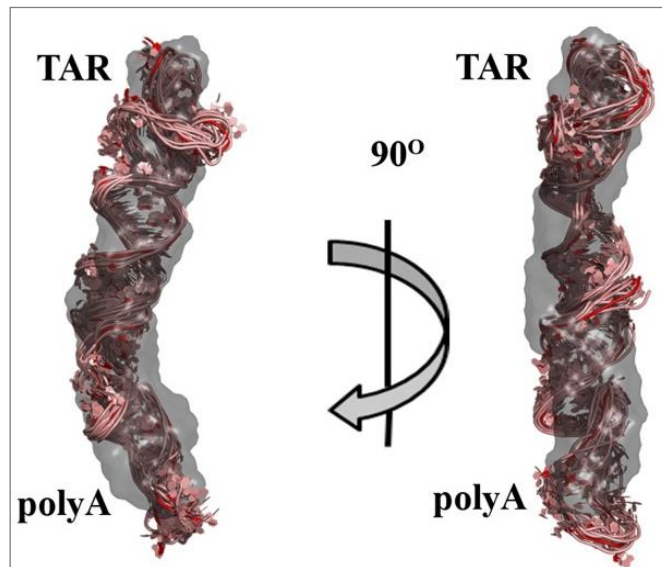


Figure 43. Modèle de structure coaxiale adoptés par les éléments TAR et poly(A) (507).

b. Circularisation du génome par l'interaction TAR-TAR

La conformation tertiaire de l'ARNg viral pourrait faciliter certaines interactions entre des régions éloignées. La tige-boucle TAR, présente aux deux extrémités du génome, pourrait via les deux séquences palindromiques interagir pour conduire à la circularisation du génome (Figure 44) (508) et favoriser les étapes de la transcription inverse tels que les transferts de brin. Une étude *in vitro* n'est cependant pas en faveur de cette hypothèse. En effet, les mutations introduites dans TAR n'ont pas affecté la formation d'un duplex entre deux ARN représentant les extrémités 5' et 3' de l'ARNg (509).

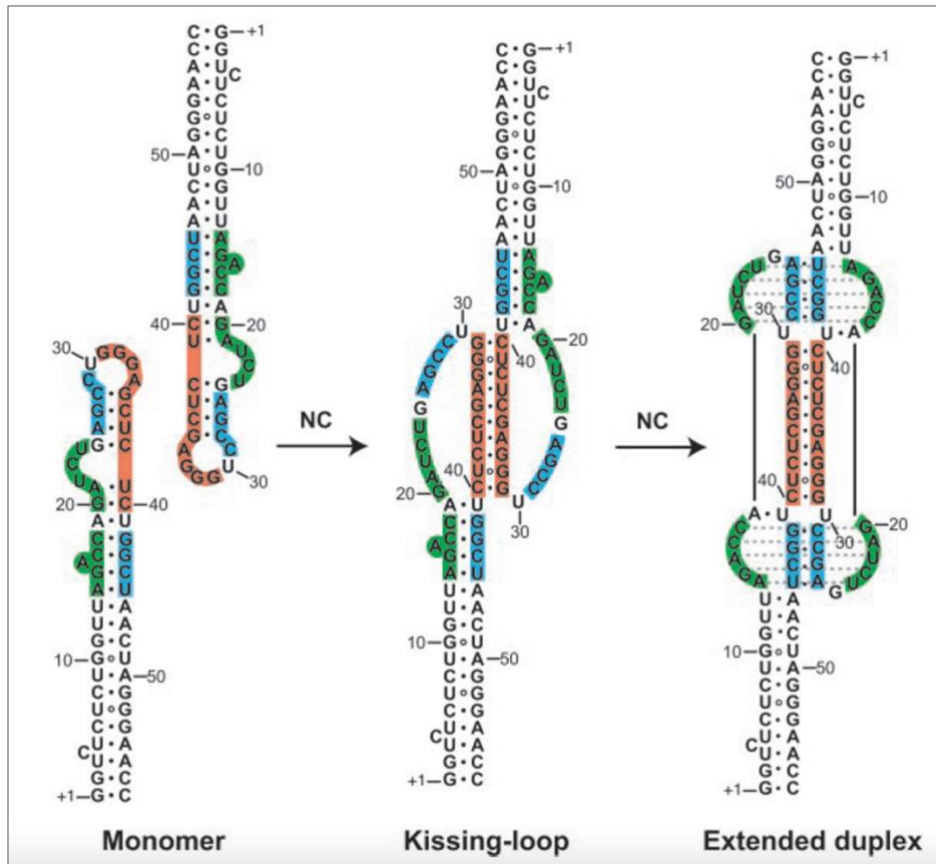


Figure 44. Modèle de dimérisation via la tige-boucle TAR (508). La NC facilite la formation d'un dimère TAR-TAR dont les sous-unités sont liées par une interaction boucle-boucle (*kissing complex*) ou un duplex étendu (*extended duplex*). Les séquences impliquées dans l'interaction intermoléculaire sont représentées par les couleurs rose, bleu et vert.

c. Circularisation du génome par l'interaction U3-ARNt^{Lys3}

La séquence U3 présente à l'extrémité 3' de l'ARNg possède une séquence de 9 nt hautement conservée. Celle-ci se trouve être complémentaire de la tige anticodon et de la boucle variable de l'ARNt^{Lys3} (510). Il a été suggéré que cet ARNt servant d'amorce pour l'initiation de la transcription inverse formerait aussi un duplex de 9 pb avec une partie du U3 pour relier les extrémités du génome viral (Figure 45) (511). La séquence de 9 nt et le segment 1, qui est voisin de celle-ci (Figure 45), favorisent le premier transfert de brin *in vitro* (511, 512). Des mutations dans ces deux éléments réduisent de 9 à 26% le premier transfert de brin dans les cellules infectées (512). Il a également été suggéré que l'interaction de l'extrémité 3' de l'ARNt^{Lys3} avec le segment 1 pourrait faciliter le déplacement de l'ARNt^{Lys3} du PBS et ainsi faciliter le second transfert de brin. Finalement, l'interaction U3-ARNt^{Lys3} pourrait faciliter deux étapes de la transcription inverse.

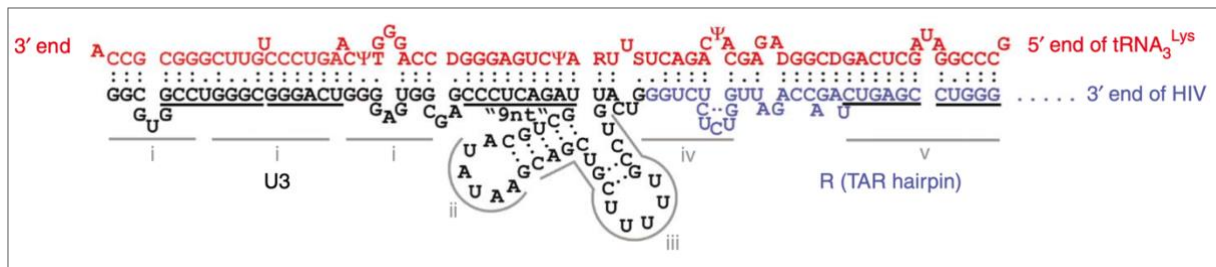


Figure 45. Complémentarité de séquence entre U3 R et l'ARN^{t^{Lys3}} (511).

d. Circularisation du génome par l'interaction gag-U3 et poly(A)

Une analyse des hétérodimères formés *in vitro* entre les extrémités 5' et 3' de l'ARNg a montré que des séquences situées dans les régions poly(A) et U3 peuvent s'apparier avec des séquences situées dans le gène gag (369). La suppression de cet appariement par mutagenèse dirigée réduit *in vitro* le premier transfert de brin induit par la NC (509). La circularisation de l'ARNg *in vivo* n'a pas été démontrée mais les études phylogénétiques des sous-types du groupe M suggère que l'interaction gag-U3 poly(A) a été conservé au cours de l'évolution.

4.1.2 Structures de l'ADNss

La région r de l'ADNss est entièrement complémentaire à la région 5' R de l'ARNg. Dans r, les séquences complémentaires des éléments TAR et poly(A) du 5'R sont respectivement nommées cTAR et cpoly(A) (390, 391). Les séquences TAR, poly(A), cTAR et cpoly(A) peuvent adopter chacune une structure en tige-boucle. Il a été montré que la séquence cTAR isolée forme une structure en tige-boucle dynamique impliquant un équilibre entre la conformation fermée et la conformation partiellement ouverte dite « Y » (305, 391, 513, 514). Notre équipe a été la première à déterminer *in vitro* la structure secondaire de l'ADNss de l'isolat MAL du VIH-1 en l'absence et présence de NC (384). Elle a aussi montré par analyse phylogénétique que le repliement de l'ADNss en trois tiges-boucles est conservé parmi les différents groupes du VIH-1(384) (Figure 46).

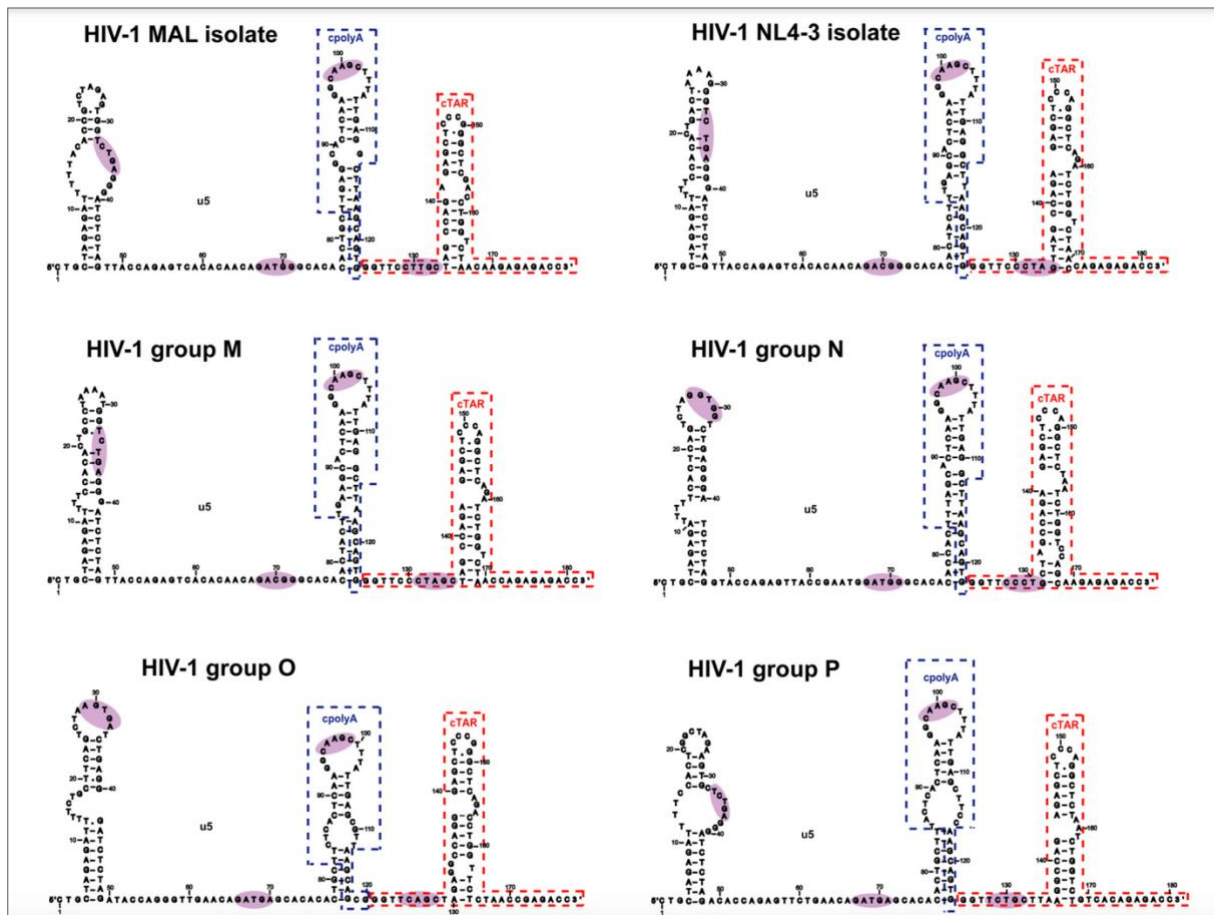


Figure 46. Structure secondaire de l'ADNss chez différents groupes et isolats du VIH-1 (384).

4.1.3 Interactions ADN-ARN régissant le premier transfert de brin

a. Interaction poly(A)-cpoly(A)

L'équipe de Berkhout a été la première à proposer un modèle dans lequel le premier transfert de brin serait facilité par des interactions boucle-boucle impliquant les tiges-boucles TAR et poly(A) de l'ARN accepteur et les tiges-boucles cTAR et cpoly(A) de l'ADNss (390). Le modèle du premier transfert de brin par invasion de l'ARN accepteur proposé à partir d'études *in vitro* suggère un rôle important de la séquence poly(A) (512, 515). Dans ce modèle, au cours de la synthèse de l'ADNss, la RT fait une pause juste en amont de la tige-boucle 5'TAR de l'ARN donneur, initiant la dégradation de la séquence 5'poly(A) par l'activité RNase H de la RT, permettant ainsi l'appariement de la séquence 3'poly(A) de l'ARN accepteur avec la séquence cpoly(A) de l'ADNss (Figure 47) (516). De plus, le transfert de brin est inhibé lorsque les tiges-boucles poly(A) en 5' et 3' sont stabilisées par des mutations (390).

Les transferts de brin précoce avec un ADNss tronqué (504), qui est dépourvu de la séquence TAR, sont des événements peu fréquents qui montrent que l'interaction poly(A)-cpoly(A) peut permettre le premier transfert de brin.

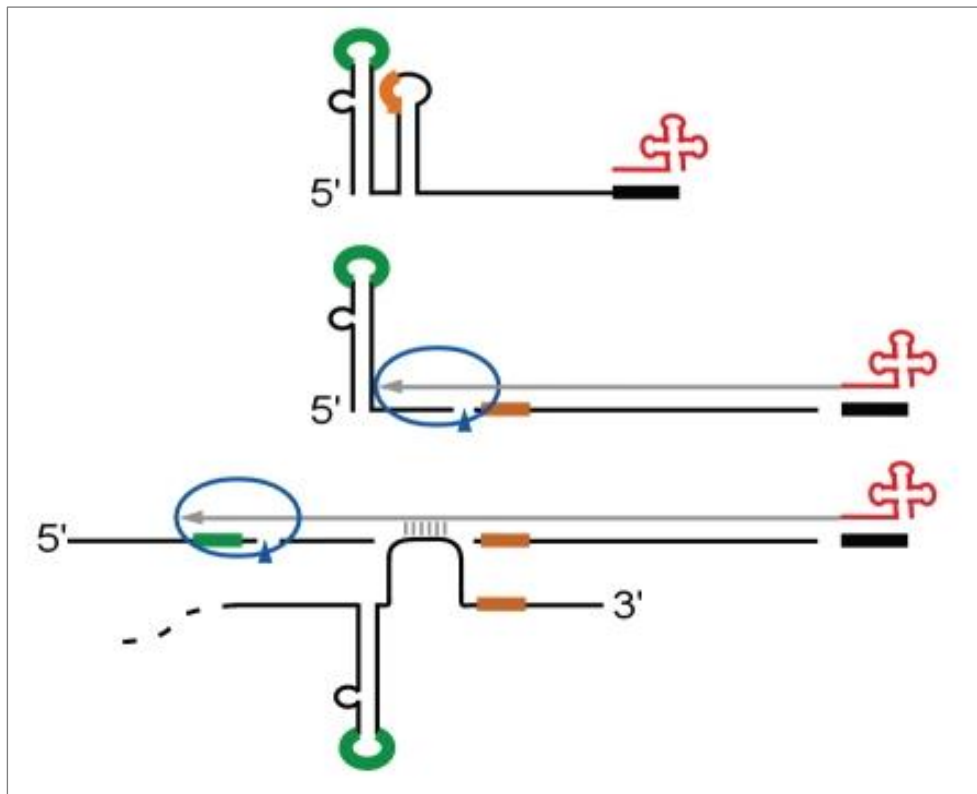


Figure 47. Modèle du premier transfert de brin par invasion (516). La tige-boucle poly(A) (boucle apicale partiellement en orange) est déstabilisée lorsqu'elle est copiée par la RT (ovale bleu). La tige-boucle TAR (boucle apicale en vert) déclenche la pause de la RT qui par son activité RNase H induit des clivages (triangle bleu) dans la séquence poly(A). Ceci permet l'invasion de la séquence cpoly(A) de l'ADNss naissant par la séquence poly(A) de l'extrémité 3' de l'ARNg. Le premier transfert de brin a lieu lorsque l'ADNss naissant est transféré au niveau de la séquence 3' TAR (non montré dans ce schéma).

b. Interaction TAR-cTAR

L'analyse *ex vivo* du premier transfert de brin par mutagenèse dirigée a montré que le premier transfert de brin se produit dans 90 à 99% des cas après la synthèse complète de l'ADNss (503). Cette observation suggère que l'hybridation TAR-cTAR est nécessaire à un transfert de brin efficace. La plupart des études ayant pour but de caractériser les mécanismes régissant l'appariement des tiges-boucles TAR et cTAR ont été réalisées *in vitro* à partir de courts acides nucléiques formant l'intégralité des tiges-boucles ou uniquement leur partie apicale. Deux modèles d'appariement entre les tiges-boucles TAR et cTAR ont été proposés (Figure 48) :

i) le modèle *kissing complex* (296, 390) correspond à une interaction initiale entre les boucles apicales TAR et cTAR ; ii) le modèle *zipper intermediate* (517, 518) repose sur une interaction initiale impliquant les extrémités des tiges inférieures.

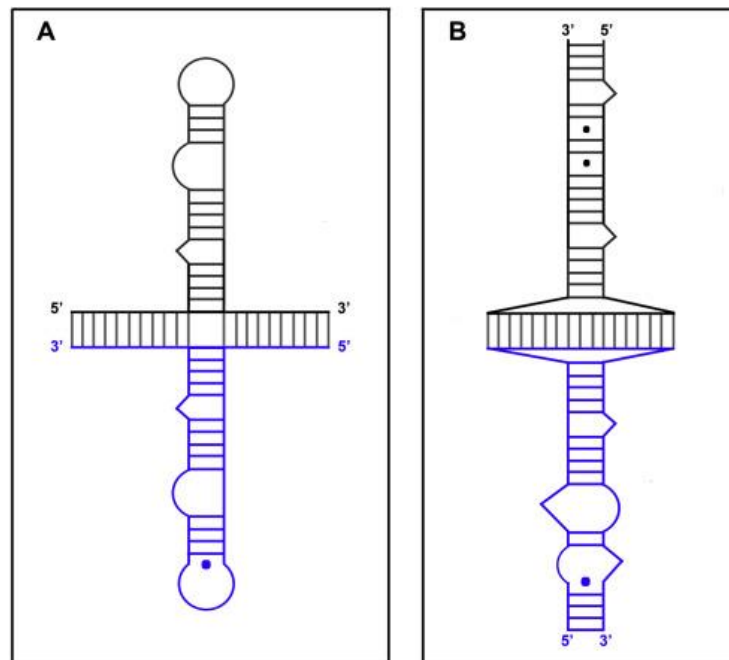


Figure 48. Modèles d'appariement entre les tiges-boucles TAR et cTAR (292). (A) Modèle *zipper intermediate* impliquant les extrémités des tiges inférieures. (B) Modèle *kissing complex* impliquant les boucles apicales.

4.1.4 Rôle de l'activité RNase H de la RT dans le premier transfert de brin

L'activité RNase H de la RT est essentielle pour la réplication du VIH. En effet, des études *in vitro et ex vivo* ont montré que les RT dépourvues d'activité RNase H ne permettent pas les réactions de transferts de brin (444, 519–521). Il a également été rapporté que des virus possédant des RT avec une activité RNase H réduite peuvent se répliquer mais présentent une fréquence de transfert de brin plus faible (116, 522, 523). La dégradation de la matrice grâce aux clivages de l'ARN par la RNase H favorise les transferts de brin en rendant l'ADN néosynthétisé accessible à l'appariement avec la région complémentaire ou en créant un site d'invasion (524, 525). Des études *ex vivo* utilisant des virus mutants dépourvus d'activité RNase H vont dans ce sens (520).

4.1.5 Rôle de la NC dans le premier transfert de brin

Plusieurs études ont montré que la NC facilite *in vitro* les transferts de brin grâce à son activité chaperon des acides nucléiques (526). Elle joue un rôle majeur dans le premier transfert de brin, notamment en facilitant par différentes actions décrites ci-après l'appariement entre les régions complémentaires de l'ADNss et l'extrémité 3'UTR (384, 430, 527).

a. La NC empêche l'auto-amorçage de l'ADNss

Guo et al. ont démontré *in vitro* que la transcription inverse après production de l'ADNss peut générer des produits d'auto-amorçage en l'absence de NC. Ceci est possible car l'ADNss contient la tige-boucle cTAR qui sert d'amorce lors du premier transfert de brin mais peut aussi servir de matrice en l'absence d'ARN accepteur (430) (Figure 49).

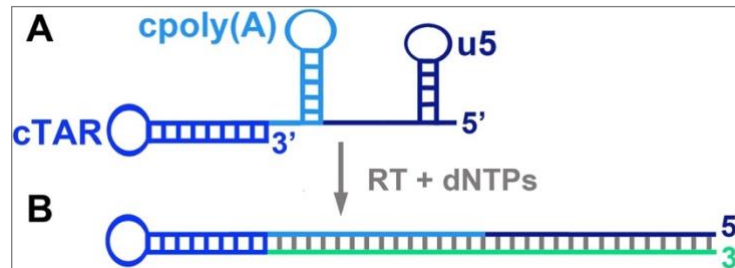


Figure 49. Auto-amorçage de l'ADNss (292). (A) Repliement de l'ADNss en tiges-boucles avec dans le sens 5' vers 3' u5, cpoly(A) et cTAR. La région r est composée des domaines cpoly(A) et cTAR. (B) Allongement de l'ADNss (-) en présence de RT et de dNTPs. La ligne verte indique le brin néosynthétisé.

Les produits d'auto-amorçage ne peuvent cependant pas servir pour la suite de la transcription inverse car leurs extrémités 3' ne sont pas complémentaires de l'extrémité 5' de la séquence 3' TAR. La NC réduit significativement l'auto-amorçage de l'ADNss en l'absence d'un ARN accepteur lorsque la transcription inverse est effectuée avec un ARN donneur servant de matrice et contenant la totalité de la région R-U5-PBS (528). La NC empêcherait l'auto-amorçage de en favorisant un repliement partiel de la tige-boucle cTAR dans lesquelles 14-18 nucléotides de son extrémité 3' resteraient appariés à un fragment d'ARN donneur qui n'aurait pas été clivé par l'activité RNase H de la RT. La NC favoriserait donc la formation de ce duplex ARN-ADN mais pas la formation complète de la tige-boucle cTAR nécessaire à l'auto-amorçage. En accord avec cette hypothèse, la NC a peu ou pas d'effet sur l'auto-amorçage d'ADNss synthétiques tronqués en l'absence d'ARN accepteur et donneur (528). En revanche, la NC réduit fortement l'auto-amorçage d'ADNss synthétiques tronqués en présence d'un ARN accepteur car la NC facilite la formation du duplex ARN-ADN contenant plus de paires de bases que les structures en tiges-boucles des ADNss.

b. La NC augmente le clivage de l'ARN donneur par l'activité RNase H

La NC augmente *in vitro* l'efficacité du premier transfert de brin en augmentant la dégradation par la RNase H de la matrice d'ARN donneur lors de la synthèse de l'ADNss (519), notamment les clivages secondaires des duplex à extrémités franches (529). De plus, la NC et le domaine RNase H de la RT peuvent interagir via les doigts de zinc (530, 531).

Néanmoins, il n'a pas été rapporté que cette interaction soit responsable de l'augmentation de l'activité RNase H. Purohit et al (532) ont montré que l'augmentation de la dégradation de l'ARN par l'activité RNase H de la RT. Celle-ci améliore les hybridations entre ARN et ADN, la RT étend et clive ainsi d'avantage de duplex. Il est donc probable que l'augmentation du premier transfert de brin en présence de NC ne résulte pas directement d'une interaction entre cette protéine et le domaine RNase H de la RT (525). Enfin, il a été montré que la NC, en présence d'un ARN accepteur, est nécessaire pour éliminer les fragments de 14-18 nucléotides de l'ARN donneur qui restent appariés à l'extrémité 3' de l'ADNss. Autrement dit, la NC favorise le déplacement des ARN courts par l'ARN accepteur plus long formant ainsi un duplex thermodynamiquement plus stable avec l'ADNss (-) (477).

c. La NC facilite l'appariement des tiges-boucles complémentaires

La réaction d'hybridation, nécessaire au premier transfert de brin, est accélérée *in vitro* en présence de NC suivant une cinétique de pseudo premier ordre (533). De nombreuses études *in vitro* supportent l'hypothèse que la NC est l'acteur principal qui facilite le premier transfert de brin en déstabilisant les structures secondaires des deux séquences r/R complémentaires (291, 305, 526, 534, 535). Le premier doigt de zinc de la NC joue le rôle le plus important dans la déstabilisation des structures secondaires formées par les acides nucléiques. Des contacts entre des acides aminés hydrophobes appartenant au premier doigt de zinc et des paires de bases ont été mis en évidence dans les complexes NC-acides nucléiques analysés par RMN (540, 541). Les travaux de notre équipe suggèrent que la NC adopte une position parallèle à la direction 5'-3' de la chaîne d'acide nucléique dans le cas de l'ADN et une position antiparallèle dans le cas de l'ARN (300, 301).

Le dépliement des tiges-boucles TAR et poly(A) dans l'ARN et cTAR et cpoly(A) dans l'ADNss sont nécessaires pour former l'hétéroduplex ADN-ARN de 93 paires de base dans le cas de l'isolat NL4-3 (391). L'étude de Clever et al. suggère fortement qu'*ex vivo* l'hybridation ADN cTAR – ARN TAR est requise pour un transfert de brin efficace. En outre, des études *in vitro* suggèrent que les tiges-boucles TAR et cTAR sont beaucoup plus importantes que les tiges-boucles poly(A) et cpoly(A) dans le premier transfert de brin. Des mutations dans la boucle apicale de TAR diminuent *in vitro* ce transfert ou l'hybridation TAR-cTAR, suggérant que l'hybridation est initiée via un complexe boucle-boucle formé par les boucles apicales de TAR et cTAR (388, 390). Une molécule de NC grâce à ses doigts de zinc peut se fixer avec une haute affinité au niveau de la boucle apicale de l'ARN TAR (293, 542). La boucle possède une guanine flexible qui joue probablement un rôle dans cette interaction (543). La première étude

in vitro dédiée à l'action de la NC sur les tiges-boucles TAR et cTAR suggère que cette protéine ne déstabilise qu'une paire de bases par molécule d'ARN TAR tandis qu'elle déstabilise 7 à 8 paires de bases par molécule d'ADN cTAR (391). Une étude plus récente réalisée à l'échelle de la molécule unique suggère que la NC déstabilise la tige-boucle ARN TAR en ciblant 4 guanines impliquées dans des appariements CG/GC qui sont adjacents à des régions instables de la tige (paires GU, boucles ou *bulges*). Ces guanines se trouvent dans la partie supérieure de la tige-boucle (paire GU, boucles ou *bulges*) (304). Puisque ces guanines se trouvent dans la partie supérieure de la tige-boucle, il est probable que cette partie joue un rôle primordial dans l'initiation de la déstabilisation de l'élément TAR qui est requise pour l'appariement TAR-cTAR (304). Une des guanines a été identifiée par RMN comme un site clé pour l'initiation de la déstabilisation (301). Les boucles apicale et interne de la tige-boucle cTAR isolée sont respectivement des sites faible et fort de fixation de la NC (296). Pour des rapports NC:nucléotides \leq à 1:8, les tiges-boucles TAR et cTAR isolées s'apparient en utilisant l'une ou l'autre des deux voies qui ont été décrites dans la partie 4.1.3 (Figure 48) (518). En revanche, en présence de concentrations élevées de NC, c'est la voie *zipper intermediate* qui est principalement utilisée (517, 518).

Notre équipe a récemment mis en évidence *in vitro* que l'hybridation médiée par la NC entre l'ADNss entier et un ARN représentant une partie de l'extrémité 3' de l'ARNg ne repose pas uniquement sur une seule des deux voies et peut n'impliquer ni les boucles apicales ni les extrémités des tiges inférieures de TAR et cTAR (384). En effet, des mutations supprimant soit la formation du *kissing complex* soit du *zipper intermediate* n'empêchent pas l'appariement des deux acides nucléiques (384). Il est donc possible que l'appariement de l'ADNss avec l'ARN 3'UTR puisse être aussi initié via une troisième voie dépendant de l'interaction poly(A)-cpoly(A) qui est décrite dans le modèle par invasion (Figure 47) (515, 516). Un modèle d'interaction de la NC avec l'ADNss a été proposé par notre équipe. Dans ce modèle, la première étape est la fixation de la NC au niveau de quatre sites qui déclenche l'ouverture de la structure tridimensionnelle de l'ADNss et favorise ainsi l'accessibilité de sa région r. La seconde étape étant la déstabilisation par la NC des structures secondaires et la troisième étape étant l'appariement par la NC des régions complémentaires r et R (292, 384).

OBJECTIF

Comme je l'ai déjà mentionné dans l'introduction, le premier transfert de brin implique l'appariement de la séquence r de l'ADNss avec la séquence R située en 3' de l'ARNg. Chacune de ces séquences forme deux tiges-boucles (cTAR et cpoly(A) dans l'ADNss, 3'TAR et 3'poly(A) dans l'ARNg) (Figure 50). Les tiges-boucles cTAR et 3'TAR joueraient le rôle le plus important dans le premier transfert de brin. Ces deux tiges-boucles doivent s'ouvrir pour pouvoir s'apparier. Plusieurs études dont la plupart *in vitro* suggèrent que la NC est principalement responsable de l'ouverture de ces structures et de leur appariement.

A partir de l'étude *in vitro* de l'interaction de la NC avec l'ADNss de l'isolat MAL, notre équipe a proposé pour l'ADNss de l'isolat NL4-3 (384) que la NC déstabilise la tige-boucle cTAR en se fixant préférentiellement au niveau de la guanine 134 (Figure 50). Une autre étude *in vitro* réalisée avec la tige-boucle cTAR isolée suggère aussi que l'interaction NC-G134 joue un rôle important dans l'ouverture de la tige-boucle (544). Cette étude suggère également que l'interaction NC-G174 est impliquée dans la déstabilisation de la tige-boucle. Toutefois, cette interaction n'a pas été observée avec l'ADNss entier (384). Deux études *in vitro* suggèrent que l'ouverture de la tige-boucle 3'TAR est déclenchée par la déstabilisation de quatre paires C-G/G-C (Figure 50) résultant de l'interaction de la NC avec les guanines impliquées dans ces paires de bases (303, 304). La déstabilisation de la tige-boucle cTAR par la NC est fortement diminuée *in vitro* lorsque des mutations aux extrémités 5' et 3' permettent la formation d'une tige composée d'au moins onze paires de bases consécutives (545, 546). L'équipe a montré que la NC n'est pas capable de convertir deux tiges-boucles identiques d'ARN en un homodimère lorsque la tige contient dix paires de bases consécutives (547). L'activité déstabilisatrice de la NC est donc limitée par le nombre de paires de bases consécutives ainsi que de la reconnaissance de guanines non appariées ou engagées dans des appariements qui sont adjacents à une région simple-brin, une boucle, un bulge ou un mésappariement.

Deux études ont comparé *ex vivo* et *in vitro* l'effet de mutations introduites dans la séquence de la NC (548, 549). Les résultats de ces études suggèrent qu'*in cellula* la transcription inverse de l'ARNg nécessite l'activité chaperon des acides nucléiques de la NC. Elles ne le démontrent cependant pas directement car les mutations, en plus d'un effet sur l'hybridation TAR-cTAR *in vitro* et l'encapsidation de l'ARNg *ex vivo*, pourraient avoir un impact négatif sur d'autres étapes de la réplication du virus (dimérisation de l'ARNg, assemblage de la particule virale). A ce jour, il n'existe pas d'études *ex vivo* montrant que la transcription inverse est altérée par des

L'objectif majeur de la thèse est de déterminer si *ex vivo* le premier transfert de brin dépend principalement de la déstabilisation par la NC des tiges-boucles ADN cTAR et ARN TAR. Pour atteindre cet objectif nous déterminons si la réplication du virus et plusieurs étapes de la transcription inverse sont altérées par des mutations dans les tiges-boucles TAR et cTAR qui ont été conçues pour que ces dernières soient peu ou pas déstabilisées par la NC.

MATÉRIEL ET MÉTHODES

1. Construction des plasmides et vecteurs

1.1 Oligonucléotides

Les oligodésoxyribonucléotides achetés chez Eurogentec et utilisés comme amorces pour les mutagénèses dirigées par PCR sont présentés dans le Tableau 3.

Noms	Séquences des amorces 5'-3'	Matrice	But
Owt(+)-fwd Om1-rev	CGGCACGAGCTCTCTGGCTAACTAGG GACGAGCTCCCAGGCTCAGATCTGGCTAACCAGAGAGAC	pLF1 pEP2	pLF1-m1 pEP2-m1
Om2-fwd Om1-rev	GCAGAGCTCTCTGGCTAACCAGGGAACCCACTGC GACGAGCTCCCAGGCTCAGATCTGGCTAACCAGAGAGAC	pLF1 pEP2	pLF1-m2 pEP2-m2
Om3-fwd Owt(-)-rev	GCAGAGCTCTCTGGCTAACTAGAGAGACCACTGCTTAAGCCTC GCCTGGGAGCTCCCAGGCTCAG	pLF1 pEP2	pLF1-m3 pEP2-m3
Om4-fwd Om4-rev	CGGCACGAGCTCTCTGGCTAACCAGGGAACCCACTG GAGCTGAGCTCCCAGGCTCAGATCTGGTCTAACCAGGGAGACCCAGTACAG	pLF1 pEP2	pLF1-m4 pEP2-m4
Om6-fwd Om6-rev	CGTCACGAGCTCTCTGGTCTAACTAGGGAGACCACTGCTTAAGCCTC CCGCTCGTCGTGGGAGCTCCCAGGCTCAG	pLF1 pEP2	pLF1-m6 pEP2-m6
OmS-1fwd OmS-1rev	CGGCACGAGCTCTCTGGCTAACTACTCTCACCCTGCTTAAGCCTC CATCGTGAGCTCCCAGGCTCAGATCTGGTCTAACCACCCCTACCCAGTACAGGCCAAAAAGC	pLF1 pEP2	pLF1-mS1 pEP2-mS1
Ofwd-m7 Owt(-)-rev	GTCACGAGCTCAGATCTGGCTAACTAGGGAACCC GCCTGGGAGCTCCCAGGCTCAG	pEP2	pEP2-m7
OmΔR-fwd OmU3-rev	CGCAGCGAATTCTTCGGAGAATAAACGTCCTCTTCTCTCGAACAAGTAGTGTGTGCCCG GCCGAGGCGCGACGGAATTCAGTACAGGCCAAAAAGCAGC	pEP2	pEP2-ΔR
OmG-fwd Om3G-rev	CGGCACGAGCTCTCTGATTAATTAGGGAACCCACTGCTTAAG CATAGTGAGCTCCCAGGCTCAGATCTGATTTAATCAGAGAGACCCAGTACAG	pEP2	pEP2-m3G
OmG-fwd Om4G-rev	CGGCACGAGCTCTCTGATTAATTAGGGAACCCACTGCTTAAG CATAGTGAGCTCCCAGGCTTAGATCTGATTTAATCAGAGAGACCCAGTACAG	pEP2	pEP2-m4G

Tableau 3. Amorces utilisées pour la mutagénèse dirigé par PCR des plasmides pLF1 et pEP2.

1.2 Plasmides

Le vecteur NLENG1-ES-IRES (176) nommé pLévy dans le laboratoire a été généreusement donné par David Lévy (NY, USA). Il est depuis plusieurs années utilisé dans notre laboratoire par l'équipe d'Olivier Delelis (551).

Le plasmide pLF1 (5267 bp) a été construit dans l'équipe par une étudiante de M1 qui était sous la direction de Brigitte René (CRHC CNRS). Il résulte du clonage dans le plasmide pCG44 (552) d'un insert généré par PCR à partir du vecteur pLévy.

Le plasmide pEP2 (5940 bp) a été construit dans l'équipe par un étudiant de M1 qui était sous la direction de Philippe Fossé. Il résulte d'une PCR amplifiant qu'une partie du vecteur pLévy.

1.3 Construction des plasmides pLF1 et PEP2 mutés

Les mutations dans le 5'R introduites dans le pLF1 et celles dans le 3'R introduites dans le pEP2 ont été réalisées grâce à la mutagenèse dirigée par PCR. Les amorces utilisées pour les différentes mutations sont présentées dans le tableau 3. Celles-ci ont été utilisées à des concentrations de 10 µM. Les essais PCR ont été réalisés avec 1 ng et 0,3 ng du plasmide pLF1 ou pEP2. Le volume de la réaction pour un essai était de 50 µl avec :

- 2,5 µl de chacune des deux amorces à 10 µM, la *forward* et la *reverse*
- 1 µl du plasmide à 1 ng/µl ou 0,3 ng/µl
- 1 µl de dNTPs à 10 mM
- 10 µl du tampon de réaction Q5 5X de chez *New England Biolabs* (NEB)
- 32,5 µl d'eau
- 0,5 µl Q5 High Fidelity DNA polymerase (2 U/µl) de chez NEB

Les conditions de la PCR sont les suivantes et la seule différence entre les couples d'amorces est la température des hybridations pouvant varier de quelques degrés mais qui était généralement comprise entre 51 et 55°C pour les 6 cycles et 71 à 76°C pour les 25 cycles :

1^{er} cycle de 1 min 30 s à 98°C, 40 s à 54°C, 50 s à 72°C,

6 cycles de 13 s à 98°C, 40 s à 54°C, 50 s à 72°C,

25 cycles de 13 s à 98°C, 40 s à 71°C, 50 s à 72°C,

Un cycle final de 7 min à 72°C.

Un contrôle sans amorce et un contrôle sans plasmide sont réalisés dans les mêmes conditions. Les produits de PCR ont été précipités à l'éthanol afin de récupérer l'ADN amplifié. Puis les produits de PCR ont été digérés par *SacI* et *DpnI*. Après une nouvelle précipitation, dans le but de réduire le volume, et une purification sur gel d'agarose 0,8 % TAE 1X-BET (0,3 µg/mL), les produits de PCR ont pu être liés par la ligase (*Instant Sticky-End Ligase Master Mix* de chez NEB) afin de produire les plasmides mutants pLF1 (pLF1-m1, pLF1-m2, pLF1-m3, pLF1-m4, pLF1-m6, et pLF1-mS1) et pEP2 (pEP2-m1, pEP2-m2, pEP2-m3, pEP2-m4, pEP2-m6, pEP2-m7, pEP2-m3G, pEP2-m4G et pEP2-mS1). Un mutant pEP2-ΔR a été conçu ne contenant pas la partie R du 3'LTR utilisé comme contrôle pour nos expériences *ex vivo*. Celui-ci conserve le site de polyadénylation. Pour sa construction, la stratégie est la même que pour les autres mutants pEP2 mais un site *EcoRI* a été ajouté plutôt qu'un site *SacI*.

Les plasmides ont tous été produits par clonage moléculaire en utilisant les bactéries DH5α. Les bactéries ont été transformées par choc thermique et cultivées à 37°C. Pour les préparations d'ADN plasmidique des cultures dans du LB-Ampicilline (100 µg/ml) à 37°C ont été réalisées.

Les kits de préparation d'ADN plasmidique (*Monarch® Plasmid Miniprep Kit et QIAfilter Midi Plasmid Purification* de Qiagen) repose sur la purification par colonne de silice avec quatre étapes : une lyse des bactéries, une fixation de l'ADN à la colonne de silice, des lavages de colonnes pour éliminer les composants biologiques telles que protéines, les sels, etc, et une dernière étape d'élution dans du tampon Tris-HCl 10 mM pH 7. Les aliquotes des plasmides purifiés étaient digérés avec les mêmes enzymes de restriction ayant servis à leur construction et analysées par électrophorèse sur gel d'agarose 0,6% TAE 1X-BET (0,3 µg/mL). Les régions LTR ont ensuite été séquencées par Eurofins Genomics avec la méthode de Sanger.

1.4 Construction des vecteurs pLevD5 et pLevD3

Les fragments AatII-BssHIII des plasmides pLF1 et les fragments XhoI-NgoMIV des plasmides pEP2 ont ensuite été isolés (digestion et purification sur gel des fragments d'intérêt) afin de produire les plasmides mutants pLevD5 (pLevD5-m1, pLevD5-m2, pLevD5-m3, pLevD5-m4, pLevD5-m6 et pLevD5-mS1) et pLevD3 (pLevD3-m1, pLevD3-m2, pLevD3-m3, pLevD3-m4, pLevD3-m6, pLevD3-m7, pLevD3-ΔR, pLevD3-m3G, pLevD3-m4G et pLevD3-mS1). Les plasmides mutants pLF1 et le pLévy ont d'abord été digérés par les enzymes, AatII et BssHIII. La partie 5'R du pLF1 a ensuite été liée par la ligase (*Instant Sticky-End Ligase Master Mix* de chez NEB) au pLévy dépourvu de la région 5'R sauvage. De la même façon, les plasmides mutants pEP2 et le pLévy ont été digérés par les enzymes XhoI et NgoMIV. La partie 3'R du pEP2 a ensuite été liée au pLévy dépourvu de la région 3'R sauvage. Pour la construction des doubles mutants, les vecteurs mutants pLevD5 ont été digérés par les enzymes XhoI et NgoMIV afin d'éliminer la région 3'R sauvage et la remplacer par la région 3'R des pEP2 mutés.

Les vecteurs pLevD5 et pLevD3 ont tous été produits par clonage moléculaire en utilisant les bactéries *Stable Competent E.coli* de NEB. Les bactéries ont été transformées par choc thermique et cultivés à 30°C dans leur milieu *Stable Outgrowth* délivré par le fournisseur. Pour les préparations d'ADN plasmidiques, des cultures dans du LB-Ampicilline (100 µg/ml) à 30°C ont été réalisées. Les ADN plasmidiques ont été préparés et analysés comme décrit dans la partie 1.3 du Matériel et Méthodes.

2. Production des particules virales

2.1 Transfection des cellules HEK293T

Les cellules HEK293T sont une lignée cellulaire dérivée de rein embryonnaire humain exprimant l'antigène T du *simian virus 40* (SV40). Cette lignée est sensible à la transfection et est donc utilisée pour la production rétrovirale. Pour produire les virus, nous utilisons la méthode chimique de transfection par phosphate de calcium. Les virus sont pseudotypés avec la glycoprotéine du *vesicular stomatitis virus* (VSV-G). Pour une flasque de 75 cm² contenant les cellules HEK293T à une confluence comprise entre 60 et 80 % le jour de la production des virus, il faut étaler entre un et deux millions de cellules deux jours avant dans 12 ml de milieu DMEM supplémenté avec 10% de sérum fœtal bovin et 1% de pénicilline/streptomycine (100 units/mL). Le jour de la production, il faut mélanger : 1,35 ml d'eau, 150 µl de CaCl₂ 3M, 3 µg de plasmide pMD.G codant pour l'enveloppe VSV-G, 20 µg d'ADN plasmidique. Pour former les précipités ADN-phosphate de calcium, il faut ajouter goutte à goutte et sous agitation ce mélange à 1,5 ml de tampon HEPES salin 2X- pH 7 (HBS) contenant des ions phosphate. Après avoir attendu environ 15 min que les précipités se forment, le mélange plasmides-HBS-calcium est ajouté aux cellules. Le précipité formé entre dans les cellules par endocytose. Le lendemain, il est nécessaire de retirer le milieu, de rincer les cellules avec 3 ml de PBS, puis d'ajouter 12 ml de milieu DMEM réchauffé à 37°C 48 h après la transfection, il faut centrifuger une première fois à 528 g pendant 5 min à 12°C puis à 2113 g pendant 20 min à 12°C avant de récupérer le surnageant de culture qui contient les particules virales. Celui-ci est ensuite aliquoté et conservé à -80°C.

2.2 Quantification de la p24

Pour transduire les cellules MT4 avec la même proportion de virus, une quantification de la protéine de 24 kDa (p24) caractéristique du VIH-1 est réalisée à partir d'une aliquote du surnageant mentionné ci-dessus. Le test se fait à l'aide du kit HIV p24 ELISA de chez Perkin Elmer.

Des puits sont recouverts d'anticorps monoclonaux de lapin spécifique de p24 du VIH-1. 20 µl de Triton X-100 5% sont déposés avant 200 µl de dilution (10⁵ et 10⁴) de surnageant de particules virales. La plaque est incubée 2 h à 37°C. Après 5 lavages avec 200 µl de tampon de lavage 1X, 100 µl de solution contenant un anticorps secondaire polyclonal biotinylé sont déposés avant une incubation d'1 h à 37°C. Les puits sont de nouveau lavés et 100 µl de solution Streptavidine-HRP (peroxydase de raifort) dilué au 100^{ème} dans son tampon fourni sont

déposés. La plaque est ensuite incubée 30 min à température ambiante. La Streptavidine-HRP est un oxydoréducteur qui va se fixer à la biotine de l'anticorps secondaire. De nouveau, les puits sont lavés et 100 µl d'orthophénylènediamine (cachet d'OPD dilué dans son tampon citrate contenant 0,03% de peroxyde d'hydrogène), substrat hydrosoluble pour la peroxydase HRP, sont déposés. Le temps de la coloration de l'OPD en jaune-orangé prend environ 15 à 20 min à l'abri de la lumière puis la réaction est arrêtée avec 100 µl de solution stop contenant de l'acide sulfurique 4N. La coloration jaune-orangé est détectable à 492 nm et est proportionnelle à la quantité de p24 quantifiée par mesure de la densité optique (DO).

2.3 Quantification des ARN viraux

La quantification des ARN viraux se fait pour un volume donné d'aliquot de surnageant viral. Pour extraire les ARN le *QIAshredder homogenizer* et le kit *RNeasy Plus Mini* de chez Qiagen sont utilisés. 10 µl de surnageant viral sont déposés à un culot de $5 \cdot 10^5$ cellules obtenu après 5 min de centrifugation à 300 g. Toutes les centrifugations de ce protocole sont réalisées à température ambiante. Les cellules servent ici de transporteur.

350 µl de tampon RLT Plus- β -mercaptoethanol (10 µl de β -mercaptoethanol 14,3 M pour 1 ml de RLT Plus) sont utilisés pour re-suspendre et lyser le tout. Le β -mercaptoethanol et le thiocyanate de guanidinium présent dans le tampon RLT Plus sont utilisés comme agent réducteur des RNases pouvant dégrader les ARN libérés pendant la lyse. L'ensemble est déposé sur une colonne de *QIAshredder* et centrifugé 2 min à pleine vitesse pour homogénéiser le lysat obtenu. Ce surnageant est déposé sur une colonne de *gDNA eliminator* et centrifugé 30 s à 8000 g. Cette étape permet l'élimination des ADN génomiques. La suite des étapes correspond à une purification sur colonne de silice avec une fixation des ARN, des lavages et une élution. 350 µl d'éthanol 70 % sont ajoutés au surnageant de la colonne *gDNA eliminator*, déposés sur la colonne de silice après homogénéisation par pipetage et centrifugé 15 s à 8000 g. Une première étape de lavage est appliquée à la colonne avec 700 µl de tampon RW1 et une centrifugation de 15 s à 8000 g éliminant les molécules telles que les glucides, les protéines, les acides gras, etc. provenant des cellules et des virions ne s'étant pas fixés à la colonne. Une seconde étape de lavage permettant l'élimination des sels des tampons précédents est réalisée avec 500 µl de tampon doux RPE et une centrifugation de 15 s à 8000 g. Une dernière étape de lavage avec 500 µl de tampon RPE et une centrifugation de 2 min à 8000 g est appliqué. Une plus longue centrifugation avec ce même tampon permet l'élimination de l'éthanol restant qui pourrait être élué avec les ARN. Enfin, les ARN sont élués dans 50 µl d'eau RNase-free par une centrifugation d'une min à 8000 g. La concentration des ARN est mesurée par la mesure de la

DO à 260 nm au spectrophotomètre (une unité de DO correspond à 40 µg/ml d'ARN). Le ratio DO_{260}/DO_{280} doit être compris entre 1,9 et 2,1 pour indiquer des ARN d'une pureté satisfaisante.

Les ARN non-épissés sont ensuite quantifiés par RT-qPCR à l'aide des amorces (Figure 51) : La 9 : 5'-GACGCTCTCGCACCCATCTC-3' et La 8.1 : 5'-CTGAAGCGCGCACGGCAA-3' et de la sonde de Taqman TM : 5'-FAM-TAGCCTCCGCTAGTCAAATTTTTGGCGTXT-Ph-3' avec FAM = 6-carboxyfluoroceïne ; XT = 6-TAMRA (6-carboxytetramethylrhodamine) ; Ph=phosphorylé. Les quantités pour une réaction de RT-PCR sont les suivantes :

- 10,58 µl d'eau
- 0,06 µl de chacune des deux amorces à 100 µM
- 0,2 de la sonde TM à 20 µM
- 4 µl de *Reaction Mix 5X* contenant le tampon de réaction, l'AptaTaq polymérase, des dNTPs, et du MgCl₂.
- 0,1 µl d'enzyme RT

Les mix de réactions sont faits à partir du *LightCycler Multiplex RNA Virus Master* de chez Roche. Un pré-mélange est réalisée pour le nombre d'essais à faire, 15 µl sont déposés au sommet des capillaires et 5 µl d'ARN avec une concentration comprise entre 50 et 200 ng/µl ensuite. Les capillaires sont centrifugés à 735 g pendant 15 s avant d'être placés dans l'appareil. Le programme de RT-qPCR est réalisé avec le *Light Cycler instrument (Roche Life Science)* et est le suivant : une transcription inverse à 65°C pendant 10 min, puis une dénaturation à 95°C pendant 5 min et une amplification avec 50 cycles à 95°C pendant 10 s et 60°C pendant 40 s.

3. Analyse de l'expression virale

3.1 Transduction des cellules MT4

Les cellules MT4 sont des lymphocytes CD4⁺ porteuses du *Human T Lymphotropic virus* (HTLV-1) et permissives à la réplication du VIH-1. 5.10⁵ cellules MT4 sont transduites avec 20 ng de p24_{gag} après traitement à la benzonase (Sigma) des virions pendant 30 min à 37 °C (100 µl de surnageant viral pour 0,5 µl de benzonase à 10 U/µl) dans leur milieu RPMI supplémenté avec 10% de sérum fœtal bovin et 1% de penicilline/streptomycine (100 units/ml). Cette étape permet l'élimination d'ADN plasmidique utilisé lors de la transfection qui aurait pu se fixer aux virions produits. Des contrôles d'inhibition de la transcription inverse, avec l'efavirenz, et d'intégration, avec le dolutégravir, sont utilisés à 1 µM pour 5.10⁵ cellules.

3.2 Mesure de l'expression de la GFP

La fluorescence des cellules MT4 est ensuite mesurée toutes les 24 h jusque trois ou quatre jours post-infection. Pour cela, un volume de 400 μ l de cellules est prélevé et ajouté à 2 ml de PBS, centrifugé 5 min à 528 g. Le surnageant est éliminé et les cellules sont remises en suspension dans 100 μ l de PBS pour être analysées au cytomètre (FACS Celesta BD). Le laser bleu (488 nm) du cytomètre est utilisé pour l'excitation de la GFP dont un de ses maxima d'excitation est 475 nm.

3.3 Quantification par qPCR des produits de la transcription inverse

Après avoir infecté 5.10^5 cellules MT4/ml avec 20 ng de p24_{gag}, 1 ml de cellules est prélevé à chaque temps souhaité, centrifugé 5 min à 300 g et 1 ml de PBS est ajouté au culot pour lavage. Une nouvelle centrifugation de 5 min à 300 g est réalisée. Le surnageant est éliminé et les culots sont congelés à -80°C pour une extraction d'ADN ultérieure.

L'extraction d'ADN se fait à l'aide du kit *QIAamp DNA blood mini* de chez Qiagen sur des culots secs. La purification des ADN se fait à l'aide d'une colonne de silice en trois étapes : une fixation des ADN, des lavages et une élution. Toutes les centrifugations sont réalisées à température ambiante. 20 μ l de protéase, 200 μ l de PBS et 200 μ l de tampon AL sont ajoutés au culot pour lyser les cellules. Le tout est vortexé 15 s puis incubé 10 min à 56 °C. 200 μ l d'éthanol absolu sont ajoutés et le tout est vortexé 15 s. Le mélange est déposé sur une colonne de silice et centrifugé une min à 6000 g. 500 μ l de tampon astringent AW1, éliminant les molécules telles que les glucides, les protéines, les acides gras, etc., sont déposés au sommet de la colonne pour lavage et centrifugés une min à 6000 g. Une seconde étape de lavage est réalisée avec 500 μ l de tampon doux AW2 éliminant les sels et une centrifugation de 3 min à 20000 g. Pour éliminer l'éthanol restant, les colonnes sont centrifugées une min à 20000 g. Enfin, les ADN sont élués avec 100 μ l d'eau après une min d'incubation et une centrifugation d'une min à 6000 g. Les concentrations des ADN sont déterminées au moyen d'un spectrophotomètre en mesurant la DO à 260 nm. Le ratio DO_{260}/DO_{280} doit être compris entre 1,7 et 1,9 pour indiquer des ADN d'une pureté satisfaisante. Une fois extrait, les ADN des différents produits de la transcription inverse sont quantifiés par qPCR à l'aide d'une gamme étalon de 1000 à 10.10^8 copies du pLévy. Deux sondes fluorescentes sont utilisées afin d'avoir une meilleure spécificité. La quantification se base sur le principe de transfert d'énergie (FRET). Lors de l'hybridation, chaque sonde se fixe à son complémentaire amplifié à une distance inférieure à 10 nm. La sonde donneuse possède une fluorescéine verte à son extrémité 3' excitée par la source de lumière de l'appareil. Celle-ci excite à son tour le fluorochrome LC640-Red présent à l'extrémité 5' de la

sonde acceptrice émettant alors une lumière fluorescente à une longueur d'onde différente de la sonde donneuse et mesurée pendant la PCR. Pendant l'étape de polymérisation, les deux sondes se retrouvent en solution pour se fixer de nouveau aux amplicons du cycle suivant. La quantité de fluorescence émise est proportionnelle à la quantité d'ADN synthétisé pendant la PCR. Pour quantifier l'ADNss avant le premier transfert de brin, les amorces utilisées sont :

- FwSS (-) : 5'-ATCTGAGCCTGGGAGCTCTCT-3'
- AA55 : 5'-CTGCTAGAGATTTTCCCACTGAC-3'

Les séquences des sondes sont :

- SS/AA/FL : 5'-TCAAGGCAAGCTTTATTGAGGCT-FL-3'
 - SS/AA/LC : 5'-LC640-GCAGTGGGTCCCTAGTTAGCCA-Ph-3'
- FL = fluoresceine ; LC = fluorochrome LC640-Red ; Ph=phosphorylé

Pour quantifier l'ADN simple-brin de polarité négative synthétisé après le premier transfert de brin, les amorces utilisées sont :

- U3NL9496 : 5'-GCTGCATATAAGCAGCTGCTTTTTGCCT-3'
- AA55 : 5'-CTGCTAGAGATTTTCCCACTGAC-3'

et les sondes sont les mêmes que celles utilisées pour quantifier l'ADNss.

Pour quantifier l'ADN tardif, les amorces utilisées sont :

- MH531 : 5'-TGTGTGCCCGTCTGTTGTGT-3'
- MH532 : 5'-GAGTCCTGCGTCGAGAGATC- 3'

et les sondes sont :

- MH FL : 5'-CCCTCAGACCCTTTTAGTCAGTGTGGAA-FL-3'
- MH LC : 5'-LC-TCTCTAGCAGTGGCGCCCGAACAG-Ph-3'.

Les temps post-infection prélevés pour quantifier les ADN sont 0, 2, 4, 8 et 24 h. Les quantités nécessaires pour une réaction de qPCR sont les suivantes :

- 11,08 µl d'eau
- 2,4 µl de MgCl₂ 25 mM
- 0,06 µl de chaque amorce à 100 µM (Cf = 0,3 µM)
- 0,2 µl de chaque sonde à 20 µM (Cf = 0,2 µM)
- 2 µl d'ADN polymérase FastStart Taq préalablement mélangé au *Reaction Mix 10X* contenant les dNTPs, le tampon de réaction et 10 mM de MgCl₂ (*LightCycler FastStart DNA Master HybProbe* de chez Roche).

Un pré-mélange est réalisé pour le nombre d'essais à faire et 16 µl sont déposés au sommet d'un capillaire, 4 µl d'ADN (autour de 500 ng) est ensuite déposé. Les capillaires sont

centrifugés à 735 g pendant 15 s avant d'être placés dans l'appareil. Les programmes de PCR sont réalisés avec le *Light Cycler instrument (Roche Life Science)* et sont les suivants :

- Pour l'ADN tardif : une dénaturation à 95°C pendant 8 min, une amplification avec 50 cycles de 95 °C pendant 10 s ; 60 °C pendant 10 s ; 72 °C pendant 8 s.

Une dénaturation à 95 °C pendant 8 min permet également à l'ADN polymérase Taq de s'activer, d'après les recommandations du fournisseur.

- Pour l'ADNss quantifié à l'aide des amorces FwSS (-) et AA55 : une dénaturation à 95 °C pendant 8 min, une amplification avec 50 cycles de 95 °C pendant 10 s ; 60 °C pendant 10 s ; 72 °C pendant 6 s.

RÉSULTATS ET DISCUSSION

1. Description du système expérimental

Pour tester les mutations introduites dans la région R des LTR 5' et 3', le vecteur NLENG1-ES-IRES (176) (16150 bp) construit à partir du génome NL4-3 du VIH-1 a été utilisé (Figure 51) (176). Il a été nommé pLévy pour le WT, pLevD5-m pour les mutations introduites dans la région 5'R et pLevD3-m pour les mutations introduites dans la région 3'R. Pour empêcher la production de virus répliatifs, dont les effets sont observables sur plusieurs cycles de réplication, les vecteurs possèdent tous deux codons stop contigus au début du gène codant pour les protéines d'enveloppe.

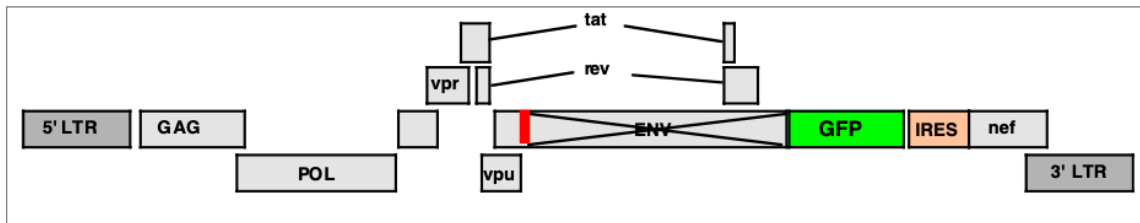


Figure 51. Génome modifié du VIH-1 dans le vecteur pLévy (176). Le trait rouge indique les deux codons stops insérés dans le gène des protéines d'enveloppe. Le gène de la GFP est inséré en aval de ENV. L'IRES est utilisé pour la traduction du gène *nef*.

Les particules virales ont été produites à partir de cellules HEK293T qui ont été transfectées par le pLévy et le plasmide pMD.G qui permet de synthétiser la protéine d'enveloppe du VSV-G. Ces particules virales ne peuvent conduire qu'à un seul cycle de réplication lors de la transduction des cellules MT4 (Figure 52) et permettre ainsi une meilleure interprétation des résultats. Les virus possèdent un génome contenant le gène de la GFP permettant de suivre l'évolution au cours du temps de l'expression virale des cellules transduites. La fluorescence a été détectée par FACS (*Fluorescent Activated Cell Sorter*) à la suite de la traduction de l'ARNm de la GFP. Une expression relativement élevée de la GFP nécessite que l'ADN viral soit intégré dans le génome de la cellule sous la forme d'ADN proviral (176, 178, 553). L'ADN proviral est le produit d'une transcription inverse complète.

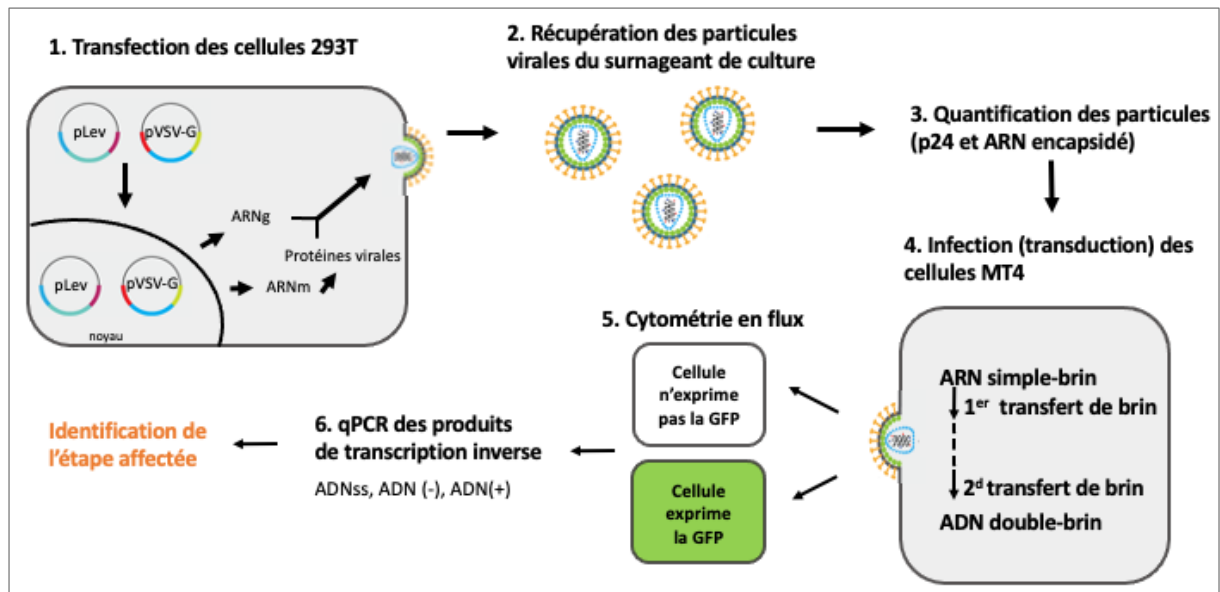


Figure 52. Stratégie expérimentale pour analyser l'étape de la transcription inverse affectée par les mutations dans les tiges-boucles ARN TAR et ADN cTAR.

2. Construction des plasmides et vecteurs

La construction des vecteurs mutés a été basée sur l'utilisation d'un vecteur rétroviral, de deux plasmides et de la méthode de mutagenèse dirigée par PCR suivie du clonage moléculaire.

- Le vecteur rétroviral pLévy, possédant les deux LTR, a servi à construire les vecteurs mutés pLevD5 et pLevD3. Il contient le génome modifié du VIH-1 (Figure 51) et a été utilisé pour la suite des expériences *ex vivo* (Figure 52). A cause des deux LTR, ce vecteur ne peut pas être utilisé directement pour muter un seul LTR et il est enclin à recombiner lorsqu'il se réplique dans des bactéries qui n'ont pas le génotype adéquat. De plus, à cause de sa grande taille (16150 pb), la quantité de vecteur produit dans les bactéries adéquates (par exemple la souche *NEB*[®] *stable competent E. coli*) n'est pas élevée. Pour ces raisons, il a été nécessaire d'utiliser les plasmides décrits ci-dessous pour construire les vecteurs mutés.
- Le plasmide pLF1, possédant un seul LTR, permet la construction des plasmides mutés dans la séquence R du 5'LTR.
- Le plasmide pEP2, possédant un seul LTR, permet la construction des plasmides mutés dans la séquence R du 3'LTR.

2.1 Construction des plasmides pLF1 et pEP2 mutés

Seize plasmides mutés ont été construits dont dix par moi. Les plasmides pLF1 (5266 pb) et pEP2 (5951 pb) contiennent respectivement les extrémités 5' et 3' du génome de type sauvage du VIH-1 (isolat NL4-3). Ils ont servi à produire par mutagenèse dirigée au moyen de la PCR les plasmides pLF1 et pEP2 mutés respectivement dans les séquences 5'R et 3'R (Figure 53). Les mutations dans la séquence R ciblent uniquement l'élément TAR à l'exception du mutant pEP2- Δ R qui est délété de l'élément TAR et d'une partie de l'élément poly(A).

Les couples d'amorces utilisées pour les PCR ont été conçues de sorte à contenir les mutations 5' ou 3' et le site de restriction Sac I (GAGCTC) (Tableau 3). Ce site unique présent dans la séquence TAR permet directement l'obtention des plasmides mutés après digestion par Sac I et ligation des produits de PCR.

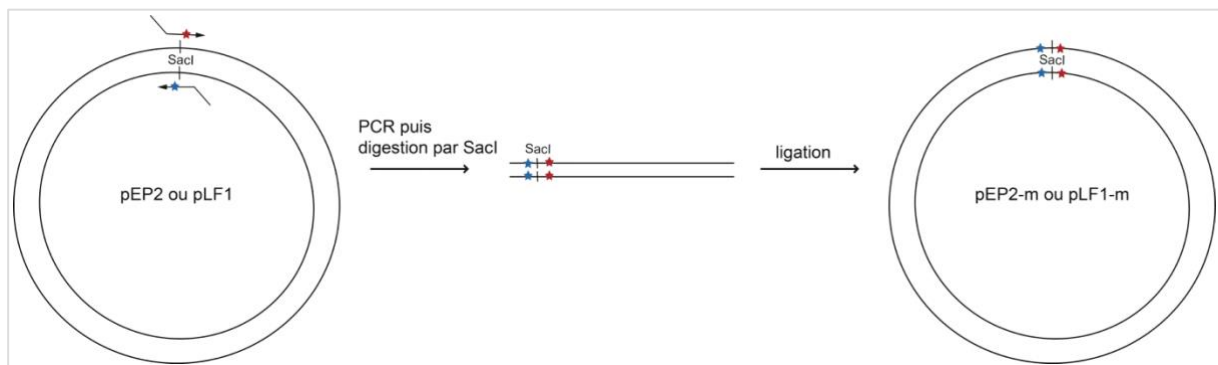


Figure 53. Stratégie de mutagenèse dirigée (exemple du mutant m4). La première étape est l'amplification par PCR d'un plasmide dont une mutation a été introduite dans chacune des amorces servant à la PCR. Puis les produits de PCR ont été digérés par SacI et liés pour former le plasmide muté.

Après amplification par PCR (exemples de la construction des plasmides pEP2-m6 et pLF1-m6, Figure 54), les produits de PCR ont été précipités à l'éthanol afin de purifier uniquement les ADN. Ces derniers ont été digérés par Sac I afin de générer les extrémités cohésives et par DpnI pour digérer le plasmide d'origine (pLF1 ou pEP2) qui était méthylé et avait servi de matrice à la PCR. L'élimination du plasmide sauvage d'origine par la digestion par DpnI a permis de purifier uniquement le plasmide muté et d'éviter une transformation bactérienne par le plasmide sauvage.

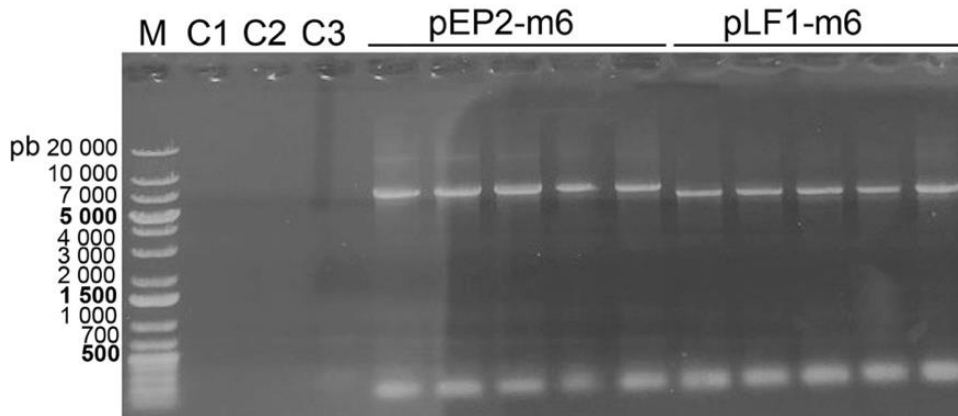


Figure 54. Analyse des produits de PCR générant les plasmides pEP2-m6 et pLF1-m6. L'analyse a été effectuée par électrophorèse sur gel d'agarose 0,6% TAE1X BET 0,3 µg/ml . M : marqueur de taille, C1 : contrôle PCR sans amorces avec pLF1, C2 : contrôle PCR sans amorces avec pEP2, C3 : contrôle sans plasmide avec les amorces Om6 fwd et Om6 rev. Une bande principale migrant entre 5000 pb et 7000 pb pour chacun des cinq essais réalisés avec 1 ng/µl de matrice est observée. Le pLF1-m6 migre un peu plus loin que le pEP2-m6 du fait de sa plus petite taille. Les bandes observées en bas du gel correspondent aux amorces également présente dans la piste C3 mais en quantité beaucoup moins importante.

Une purification sur gel d'agarose des produits PCR digérés par Sac I et DpnI a été réalisée pour isoler le plasmide muté et linéaire. La purification a été efficace comme le montre l'analyse par électrophorèse sur gel d'agarose (Figure 55). Le plasmide purifié a ensuite été circularisé par ligation avant de transformer les bactéries DH5α qui ont servi à l'amplifier.

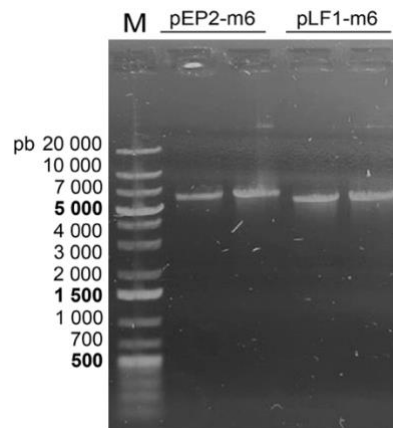


Figure 55. Analyse des plasmides pEP2-m6 et pLF1-m6 digérés par SacI et DpnI après purification sur gel d'agarose. L'analyse a été effectuée par électrophorèse sur gel d'agarose 0,6% TAE1X BET 0,3 µg/ml. M : marqueur de taille.

Après une transformation bactérienne, un repiquage de plusieurs colonies, une préparation d'ADN plasmidique et une digestion analytique ont été réalisées (Figure 56). Puis les plasmides des clones sélectionnés correspondant au bon profil de digestion ont été envoyés à une société de séquençage pour vérifier que la partie VIH-1 de leur séquence contienne uniquement la mutation choisie.

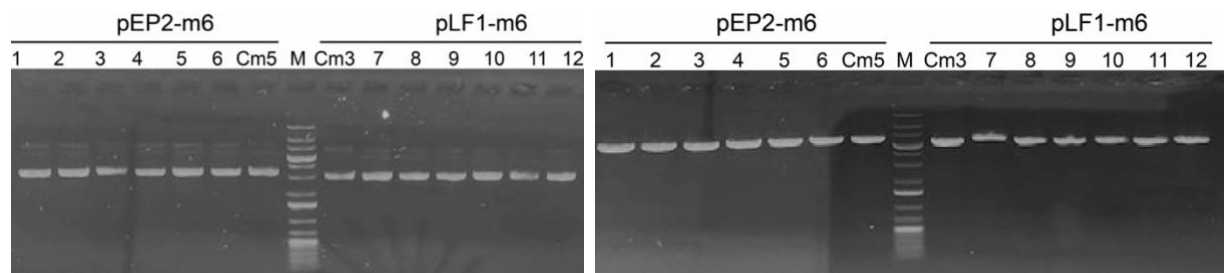


Figure 56. Analyse des plasmides pEP2-m6 et pLF1-m6 avant (panneau de gauche) et après digestion par SacI (panneau de droite). L'analyse a été effectuée par électrophorèse sur gel d'agarose 0,8% TAE1X BET 0,3 µg/ml. Dans le panneau de gauche, migration des plasmides pEP2-m6 pour les clones 1 à 6 et pLF1-m6 pour les clones 7 à 12 avec un pEP2-m5 (Cm5) et pLF1-m3 (Cm3) nous servant de contrôles de taille. Les plasmides surenroulés migrent au-dessus de la bande à 3000 pb pour les pEP2 et juste au niveau pour les pLF1. On peut observer une ou deux bandes selon les clones au-dessus des plasmides surenroulés correspondant probablement à des plasmides circulaires relâchés et des multimères. Dans le panneau de droite, les mêmes clones digérés par SacI. Les clones 1 à 6 de pEP2-m6 ainsi que les clones 8 à 12 de pLF1-m6 sont linéarisés comme attendu. Le clone 7 de pLF1-m6 migre un peu moins loin que les autres. Il n'a donc pas été considéré pour la suite des expériences. M : marqueur de taille (le même que celui utilisé dans les Figures 54 et 55).

Les résultats ne sont montrés que pour deux plasmides sur les seize construits, mais la même démarche a été appliquée aux autres mutants. Cependant, la construction du mutant pEP2- Δ R diffère par l'insertion d'un site EcoRI plutôt que d'un site SacI à partir des amorces de PCR. La digestion par EcoRI génère les extrémités cohésives qui permettent la circularisation du mutant pEP2- Δ R après ligation. Pour rappel, ce mutant est délété d'une grande partie de sa région R mais conserve cependant le site de polyadénylation. Il sert de contrôle négatif dans les tests *ex vivo*.

Le séquençage de plusieurs des mutants a montré des insertions au niveau du site SacI des plasmides pLF1 et pEP2 mutés. Nous n'avons pas réellement identifié la cause de ces insertions, mais on suppose que les mutations présentes dans l'amorce lui font adopter un repliement particulier conduisant à une duplication complète ou partielle du site SacI lors de la PCR. Pour certains des plasmides mutés contenant des insertions, il a pu être possible de les corriger. Les insertions contenaient une dizaine à une quinzaine de nucléotides. Chez certains plasmides l'insertion était encadrée par deux sites SacI. Une digestion par SacI a permis l'élimination de cette insertion (Figure 57). Puis une purification sur gel, une ligation-

transformation, des repiquages de bactéries et des séquençages ont dû de nouveau être réalisés pour s'assurer de la correction. Pour d'autres plasmides contenant également des insertions, il n'a pas été possible de les corriger de cette façon, il a donc fallu reprendre les constructions depuis la PCR.

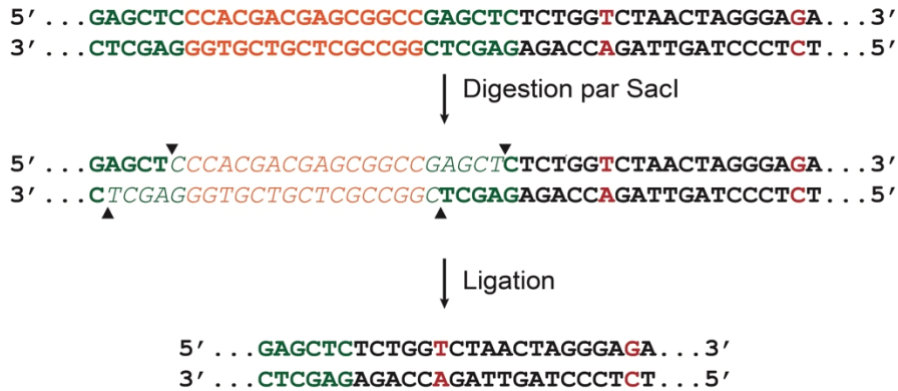


Figure 57. Exemple de correction des insertions grâce au site SacI du plasmide pEP2-m6. Après digestion par SacI dont le site a été dupliqué, le plasmide est de nouveau lié et amplifié par clonage moléculaire. En vert, les sites SacI, en orange l'insertion de 15 nucléotides en plus du site SacI dupliqué, en rouge les deux insertions T et G correspondant au mutant m6. Les flèches noires au niveau des sites SacI correspondent aux coupures réalisées par l'enzyme.

2.2 Construction des vecteurs mutés pLevD5 et pLevD3

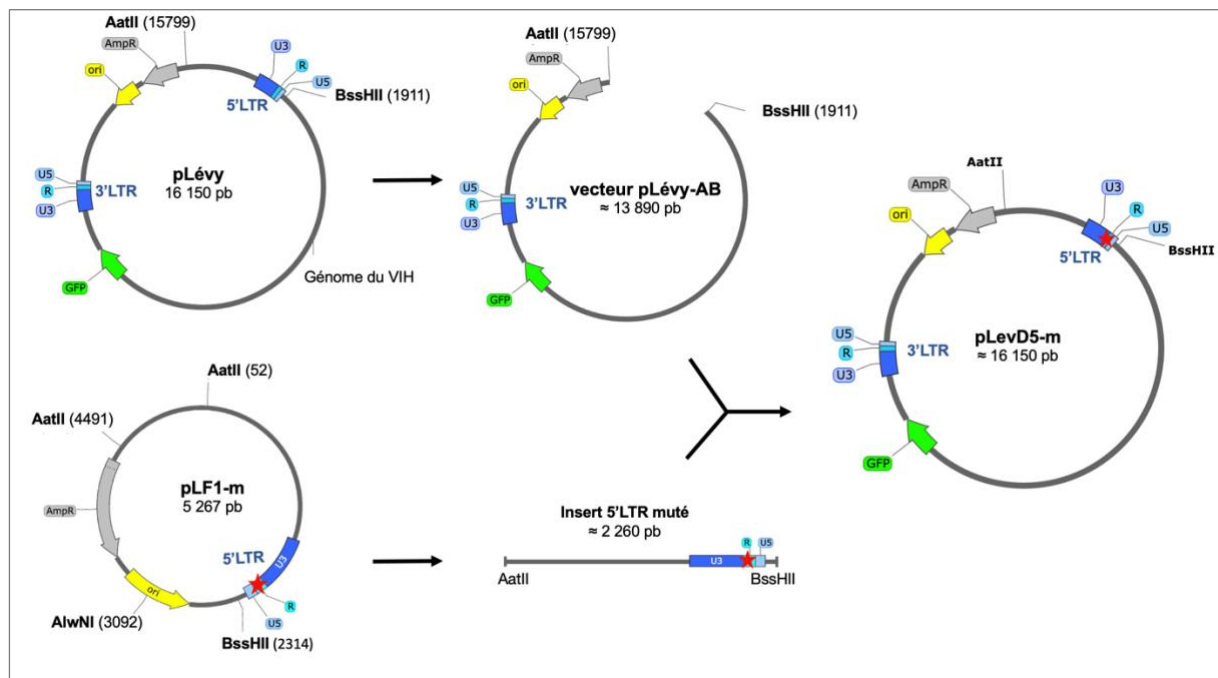


Figure 58. Stratégie de construction d'un vecteurs muté pLevD5.

J'ai construit vingt-et-un vecteurs mutés à partir des inserts des régions 5'LTR et 3'LTR des plasmides pLF1 et pEP2 mutés et du vecteur pLévy dont la région 5'LTR ou la région 3'LTR de type sauvage avait été éliminée (Figures 58 et 59). Les inserts et les vecteurs pLévy-AB et pLévy-XN ont été produits au moyen des enzymes de restriction AatII, BssHIII, NgoMIV et XhoI.

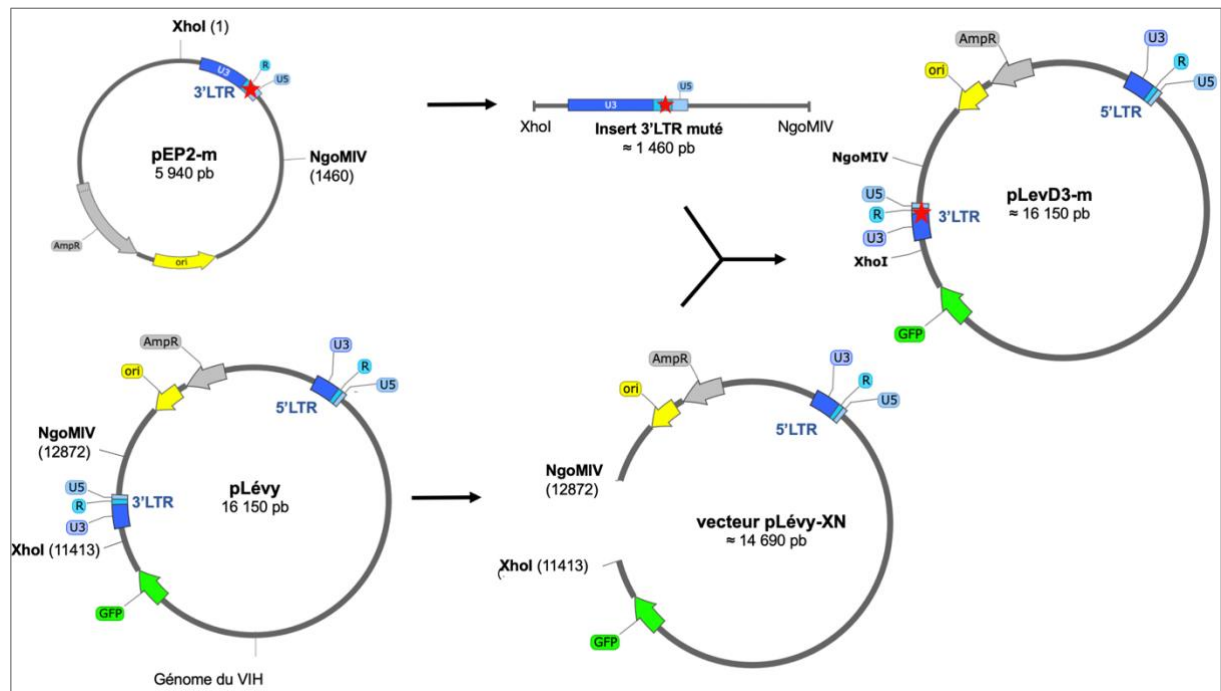


Figure 59. Stratégie de construction d'un vecteurs muté pLevD3.

Dans un premier temps, le plasmide pLévy, contenant le génome du VIH-1 dans lequel est inséré le gène de la GFP et un IRES (Figure 51), a été digéré soit par les enzymes AatII et BssHIII, soit par les enzymes XhoI et NgoMIV afin de générer respectivement les vecteurs pLévy-AB et pLévy-XN (Figures 58 et 59). Le vecteur pLévy-AB est délété du 5'LTR tandis que le vecteur pLévy-XN est délété du 3' LTR. Après digestion, ils ont été purifiés sur gel d'agarose et comme le montre la Figure 60, la purification était efficace.

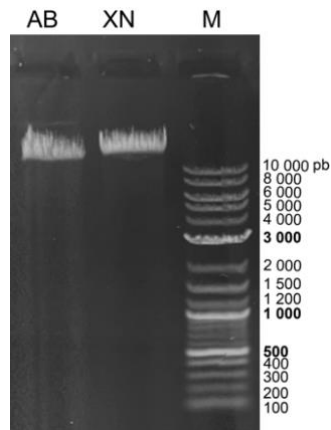


Figure 60. Analyse des vecteurs pLévy-AB et pLévy-XN après purification sur gel. L'analyse a été effectuée par électrophorèse sur gel d'agarose 0,6% TAE1X BET 0,3 µg/ml. M : marqueur de taille ; AB : vecteur pLévy-AB ; XN : vecteur pLévy-XN.

Les plasmides pLF1 mutés et pEP2 mutés ont été digérés par les mêmes enzymes, AatII et BssHII pour les inserts-5'-AB et XhoI et NgoMIV pour les inserts-3'-XN. Les inserts ont ensuite été purifiés sur gel d'agarose. Comme précédemment, les résultats ne sont présentés que pour deux mutants, pLevD5-m6 et pLevD3-m6, mais la construction des autres mutants suit le même protocole. Comme le montre la Figure 61, la purification des inserts mutés m6 était efficace.

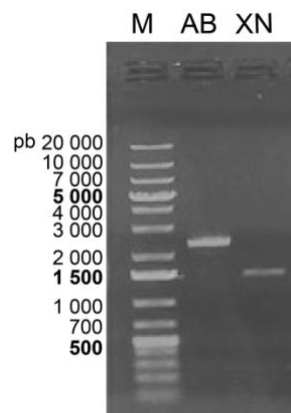


Figure 61. Analyse des inserts 5' et 3'LTR-m6 après purification sur gel d'agarose. L'analyse a été effectuée par électrophorèse sur gel d'agarose 0,8% TAE1X BET 0,3 µg/ml. Piste AB : insert 5'LTR-m6 (environ 2260 pb) généré au moyen des enzymes AatII et BssHII ; piste XN : insert 3'LTR-m6 a (environ 1460 pb) généré au moyen des enzymes XhoI et NgoMIV. M : marqueur de taille.

Les produits de ligation des mélanges vecteurs pLévy-AB/insert 5'LTR-m6 et pLévy-XN/insert 3'LTR-m6 ont servi à transformer les bactéries *NEB® stable competent E. coli* qui sont préconisées dans le cas de clonage de plasmides ayant des séquences répétées telles que les LTR des vecteurs rétroviraux. Après un peu moins de 24 h à 30°C pour une meilleure stabilité des clones plasmidiques, un repiquage et une préparation d'ADN plasmidique d'au moins six clones ont été réalisés. Les plasmides ont été analysés par une digestion avec les enzymes de

restriction ayant servi à les construire, XhoI et NgoMIV pour les pLevy-3' muté (Figure 61) et AatII et BssHII pour les pLevy-5' muté (Figure 62).

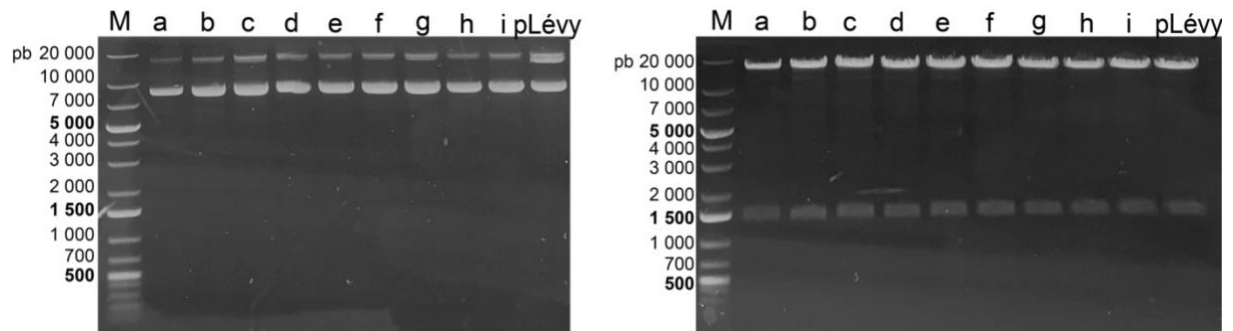


Figure 62. Analyse de neuf clones plasmidiques pLevD3-m6 avant (panneau de gauche) et après digestion par XhoI et NgoMIV (panneau de droite). L'analyse a été effectuée par électrophorèse sur gel d'agarose 0,8% TAE1X BET 0,3 µg/ml. L'insert 3'LTR a une taille d'environ 1460 pb. La migration du pLévy nous sert de contrôle de taille et de contrôle de digestion. M : marqueur de taille.

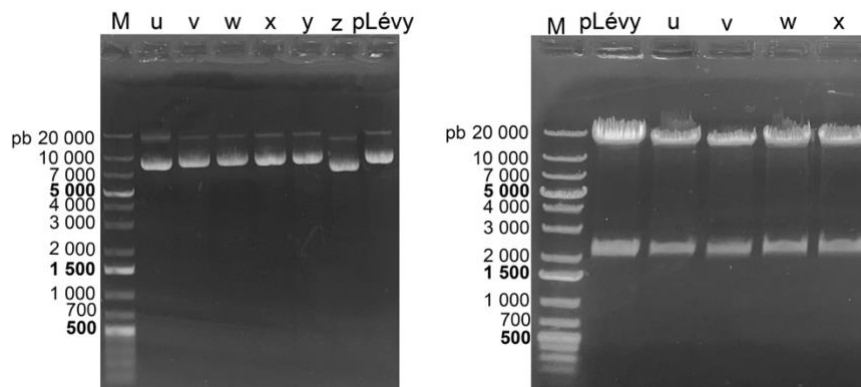


Figure 63. Analyse de clones plasmidiques pLevD5-m6 avant (panneau de gauche) et après digestion par AatII et BssHII (panneau de droite). L'analyse a été effectuée par électrophorèse sur gel d'agarose 0,8% TAE1X BET 0,3 µg/ml. L'insert 5'LTR a une taille d'environ 2260 pb. M : marqueur de taille.

Les mutations introduites peuvent produire des mésappariements entre les régions r de l'ADNss et 3'R de l'ARNg. Selon leur localisation, certains de ces mésappariements peuvent empêcher le premier transfert de brin et par conséquent la poursuite de la transcription inverse (554). C'est pourquoi, pour certaines mutations dans le 5'R ou le 3'R produisant un effet négatif sur la réplication virale, des doubles mutants supprimant les mésappariements ont été construits pour déterminer si la réplication virale était restaurée. Cinq doubles mutants (vecteurs pLevD5-3-m2, pLevD5-3-m3, pLevD5-3-m4, pLevD5-3-m6 et pLevD5-3-mS1) ont été construits. Pour cela, les régions 3' des plasmides pLevD5-m2, pLevD5-m3, pLevD5-m4, pLevD5-m6, et pLevD5-mS1 ont été éliminées par digestion par les enzymes XhoI et NgoMIV. Elles ont été remplacées par les fragments XhoI-NgoMIV des plasmides pEP2-m2, pEP2-m3, pEP2-m4, pEP2-m6 et pEP2-mS1 contenant la région 3'R mutée.

3. Bases rationnelles de la conception des mutants

La conception des mutants a reposé sur les observations et résultats ci-après.

- Dans l'ADNss, la tige-boucle cTAR ne contient pas plus de cinq paires de bases consécutives (Figure 50 A).
- La tige-boucle ARN TAR contient onze paires de bases consécutives mais deux sont des paires G-U et U-G (Figure 50 B).
- *In vitro* la déstabilisation de la tige-boucle ADN cTAR par la NC est fortement réduite par des mutations qui génèrent une tige constituée d'au moins onze paires de bases Watson-Crick consécutives (545, 546).
- La NC n'est pas capable de convertir deux tiges-boucles identiques d'ARN en un homodimère lorsque la tige contient dix paires de bases Watson-Crick consécutives (547).
- La NC déstabilise les tiges-boucles en interagissant préférentiellement avec des guanines non appariées ou des guanines appariées mais adjacentes à une paire G-U ou à des régions non appariées (304, 555, 556).

Les mutations choisies correspondent à des délétions, additions et/ou des substitutions dans la région R de l'ARNg et dans la région r de l'ADNss qui ont pour but de stabiliser les tiges-boucles ADN cTAR et ARN TAR ou d'altérer leur interaction avec la NC. Elles suppriment par exemple des *bulges*, allongent des tiges, remplacent des paires de bases G-U par des paires de base de type Watson-Crick ou remplacent des paires de bases G-C par des paires de bases A-U. Pour cibler uniquement l'éventuel rôle *ex vivo* de la NC sur la déstabilisation des tiges-boucles TAR et cTAR, le choix des mutations a été guidé par quatre critères :

- La partie supérieure de la tige-boucle 5'TAR ARN, c'est-à-dire entre les nucléotides 18 à 44, est préservée (Figure 64). En effet, cette région est indispensable à l'activation de la transcription de l'ADN proviral (voir la sous-partie 2.3.5.b de l'introduction).
- La partie double-brin inférieure de la tige-boucle 5'TAR ARN est préservée pour ne pas impacter la dimérisation et l'encapsidation de l'ARNg (voir la sous-partie 2.3.6 de l'introduction).
- La partie double-brin de la tige-boucle 3'TAR ARN est préservée car son ouverture peut impacter la polyadénylation des ARNm (557).
- A l'exception de deux mutants, aucune mutation n'est introduite entre les nucléotides 1 et 10 des tiges-boucles 5'TAR et 3'TAR (Figures 64 et 71) car les mutations dans cette région conduisent à des mésappariements entre l'ADNss et la région 3'R de l'ARNg qui réduisent significativement la transcription inverse (554).

4. Analyse des effets produits par les mutations dans la séquence 5' TAR

4.1 Structures secondaires des mutants 5'TAR et cTAR

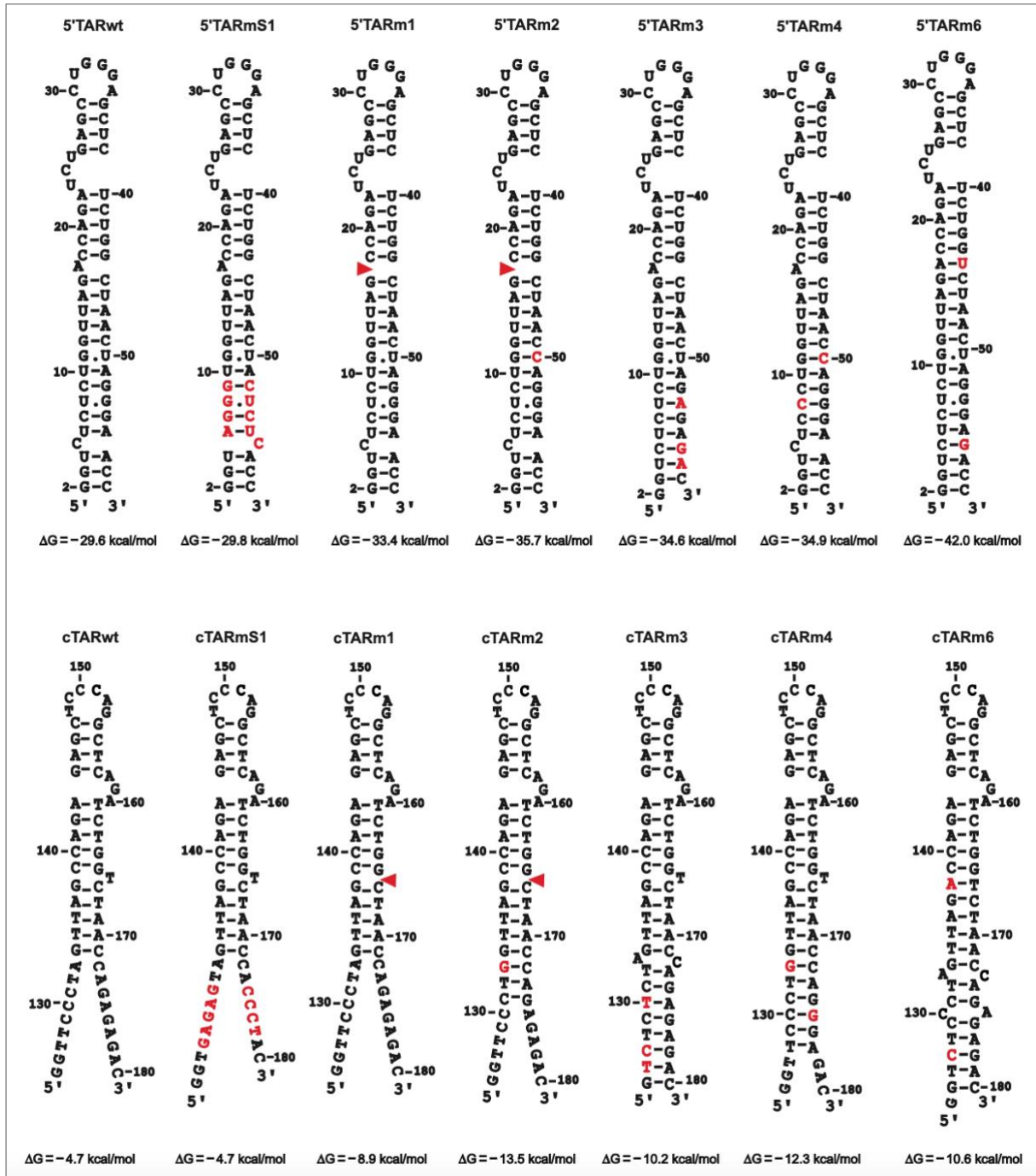


Figure 64 : Structures secondaires prédites pour les 5'TAR et cTAR mutés. Les mutations sont en rouge. Un triangle rouge correspond à une délétion d'un nucléotide. Les repliements prédictifs ont été réalisés à l'aide du logiciel *mfold* (558). Pour l'ADN cTARwt, le modèle de base est l'élément cTAR dans l'ADNss (559). La numérotation du 5'TARwt est en fonction des sites alternatifs de la transcription (327). Le G2 dans le 5' TARwt correspond à la coiffe en 5' (7-méthylguanosine triphosphate) (327). La numérotation de l'ADN cTAR est en fonction de celle de l'ADNss (559). La tige-boucle cTAR ne possède pas deux cytosines à son extrémité 3' car la RT du VIH-1 ne copie probablement pas la coiffe en 5' de l'ARNg (560).

Les mutations dans la tige-boucle ARN 5'TAR ont été conçues pour en modifier finement la structure mais aussi celle de l'ADN cTAR si la RT peut copier la tige-boucle TAR stabilisée par ces mutations (Figure 64). Comme le montre les ΔG calculés par mfold, tous les mutants à l'exception des mS1 sont prédits posséder une tige-boucle plus stable que celle du sauvage. Le mutant mS1, qui sert de contrôle négatif, est identique dans la séquence 5'TAR à celui publié par Ohi et Clever (554). En effet, le mésappariement induit par les mutations du mS1 empêche le premier transfert de brin et réduit fortement la réplication du virus.

Dans le 5'TARm1, la suppression du *bulge* A17 génère une tige de seize paires de bases consécutives dont deux sont G-U et U-G. La RT ne devrait pas pouvoir copier la tige-boucle 5'TARm1 si l'ouverture de cette dernière exige l'action de la NC. La mutation m1 correspond à la délétion du *bulge* T166 et à la formation d'une tige constituée de dix paires de bases consécutives dans l'ADN cTAR. La NC devrait donc être capable de déstabiliser cette tige-boucle et de se fixer efficacement sur celle-ci car son site fort putatif (CCTAG) (Figure 50 A) est conservé.

La tige-boucle 5'TARm2 diffère de la tige-boucle 5'TARm1 par le remplacement de la paire G-U par la paire G-C qui augmente légèrement la stabilité de la structure. En revanche, la tige-boucle cTARm2 est significativement différente de la tige-boucle cTARm1 car le remplacement de l'adénine par la guanine à la position 133 permet la formation d'une tige constituée de treize paires de bases Watson-Crick consécutives. Le site fort putatif de la NC est modifié au niveau d'une base (CCTGG) et les deux guanines sont appariées et non adjacentes à une région simple-brin. La NC ne devrait donc pas pouvoir faciliter l'appariement cTARm2-3'TAR en déstabilisant la tige-boucle ADN.

Dans le 5'TARm3, trois substitutions permettent de supprimer le *bulge* en bas de la tige, de remplacer la paire U-G par la paire U-A et de former une tige inférieure composée de quatorze paires de bases consécutives dont une paire G-U. La RT ne devrait pas pouvoir copier la tige-boucle 5'TARm3 si l'ouverture de cette dernière exige l'action de la NC. La structure de la tige-boucle cTARm3 diffère significativement de celle du sauvage car les extrémités 5' et 3' sont appariées sur une longueur de huit paires de bases consécutives. Le site fort putatif de la NC est modifié au niveau d'une base (TCTAG) et la guanine est appariée mais adjacente d'une boucle interne. La NC devrait donc être capable de déstabiliser la tige-boucle cTARm3 et de se fixer efficacement sur celle-ci.

Dans le 5'TARm4 les remplacements de U8 et U50 par C8 et C50 augmentent la stabilité de la tige-boucle en remplaçant respectivement les paires G-U et U-G par les paires G-C et C-G mais ne changent pas le nombre de paires de bases consécutives par rapport au sauvage.

En revanche, les remplacements de A133 et A175 par G133 et G175 dans l'ADN cTARm4 modifient significativement la tige-boucle par rapport au sauvage. En effet, la tige inférieure est allongée de six paires de bases et est donc constituée de onze paires de base Watson-Crick consécutives. De plus, le site fort putatif de la NC est modifié au niveau d'une base (CCTGG) et les deux guanines sont appariées et non adjacentes à une région simple-brin. La NC ne devrait donc pas pouvoir faciliter l'appariement cTARm4-TAR en déstabilisant la tige-boucle ADN. Dans le 5'TARm6, deux insertions (U et G) suppriment deux *bulges* et conduisent à la formation d'une tige constituée de vingt-et-une paires de bases dont deux sont G-U et U-G. La RT ne devrait pas pouvoir copier la tige-boucle 5'TARm6 si l'ouverture de cette dernière exige l'action de la NC. La structure de la tige-boucle cTARm6 diffère significativement de celle du sauvage car les extrémités 5' et 3' sont appariées et la tige intermédiaire est constituée de onze paires de bases Watson-Crick consécutives. Le site fort putatif de la NC (CCTAG) est partiellement apparié mais la guanine appariée est adjacente d'une boucle interne. La NC devrait donc être capable de se fixer sur la tige-boucle cTARm6 tout en ne la déstabilisant pas ou peu à cause des onze paires de bases Watson-Crick consécutives.

4.2 Résultats des transfections par les mutants 5'

Une fois les différents vecteurs pLevD5-m construits, les particules virales ont été obtenues grâce à des transfections des cellules HEK293T comme décrit dans la partie Matériel et Méthodes. Les virions ont été quantifiés par un test ELISA permettant une quantification de la p24, protéine constituant la capsid. Les résultats des quantifications p24 sont présentés dans le Tableau 4 et représentent une moyenne d'au moins deux quantifications par type de virus.

WT	5m1	5m2	5m3	5m4	5m6	5mS1
2,39 ng/μl	2,17 ng/μl	1,68 ng/μl	2,58 ng/μl	2,20 ng/μl	1,75 ng/μl	1,62 ng/μl

Tableau 4 : Quantités de p24 mesurées pour les sept productions de particules virales.

Les six productions de particules virales des mutants présentent des valeurs de p24 proches de celle du WT. En effet, la diminution la plus importante observée avec le 5mS1 n'est que de 32% par rapport au sauvage. Ces résultats suggèrent que l'expression des virus dans les cellules HEK293T n'est pas altérée par les mutations que nous avons conçues. Les surnageants contenant les virions ont été traités à la benzonase (DNase) comme mentionné dans la partie Matériel et Méthodes afin d'éliminer l'ADN plasmidique qui a été utilisé pour produire les particules virales et qui pourrait être amplifié lors des quantifications par qPCR des ADN viraux.

4.3 Résultats des transductions par les mutants 5'

Les transductions des cellules MT4 par les virions ont été réalisées en utilisant la même quantité de p24 que le WT, c'est-à-dire 20 ng de p24 pour 500 000 cellules, ce qui correspond à une multiplicité d'infection (MOI) de 0,2. Les virus réalisent leur unique cycle de réplication en entrant dans la cellule par endocytose grâce à l'enveloppe du VSV-G, entrée qui est suivie séquentiellement par : 1) la transcription inverse du génome viral ; 2) l'intégration de l'ADN viral ; 3) la production des différentes protéines virales et de la GFP qui permet de suivre l'expression du génome viral.

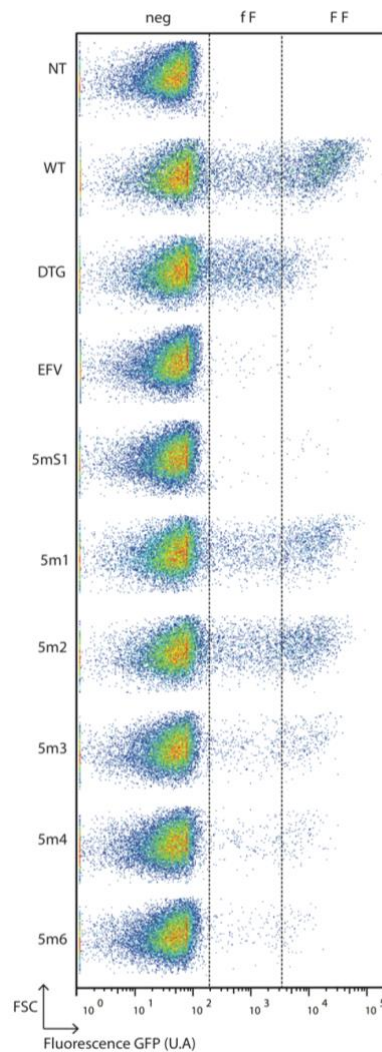


Figure 65. Fluorescence des cellules MT4 déterminée par cytométrie en flux deux jours après les transductions. Trois niveaux d'intensité de fluorescence sont détectés, un négatif (neg), un de faible intensité (f F) et un de forte intensité (F F). Le f F correspond aux formes ADN viraux circulaires à 1 et 2-LTR exprimant faiblement la GFP. Le F F correspond à une forte expression de la GFP à partir de l'ADN proviral (171, 176, 178). FSC : *Foward scatter* (paramètre mesuré permettant de discriminer les cellules par leur taille). NT, cellules MT4 non transduites ; WT, cellules MT4 transduites par le virus sauvage ; DTG, cellules MT4 transduites par le virus sauvage en présence de DTG ; EFV, cellules MT4 transduites par le virus sauvage en présence de EFV ; 5m1 à 5m6, cellules MT4 transduites par les virus mutants 5m1 à 5m6.

Les mesures de la fluorescence de la GFP grâce au cytomètre en flux ont été effectuées un à quatre jours après l'infection. Les expériences ont été réalisées en triplicat biologique. Les paramètres du cytomètre ont permis l'excitation des protéines GFP à 488 nm et une mesure de son absorbance à 510 nm. Les cellules en doublet, en amas et les débris cellulaires ont été éliminés de l'analyse par le logiciel Flowjo. Les analyses ont été réalisées sur une acquisition de 20 000 cellules. De plus, un pourcentage de cellules vivantes similaire entre les cellules transduites par le virus sauvage et les cellules transduites par les virus mutants a été mesuré chaque jour. À titre d'exemple, les résultats d'une expérience obtenus par cytométrie en flux deux jours après les transductions sont présentés dans la Figure 65. Comme attendu, les cellules non transduites (NT) ne sont pas fluorescentes. Comme cela a déjà été décrit dans d'autres études pour les cellules transduites par le virus sauvage (176, 551), on observe deux types de fluorescence, une fluorescence faible (f F) et une fluorescence forte (F F) que l'on sépare par les deux traits verticaux en pointillé.

Un contrôle négatif correspondant à la transduction des cellules MT4 par le virus sauvage en présence de dolutégravir (DTG) a été effectué (Figure 65). On observe peu de cellules présentant une forte fluorescence car le DTG inhibe l'intégrase et ne permet donc pas la formation de l'ADN proviral qui est nécessaire à une expression importante du gène de la GFP. On observe cependant beaucoup de cellules présentant une faible fluorescence en présence de DTG. Cette fluorescence correspond à la détection de la GFP faiblement exprimée par les formes circulaires à 1 et 2-LTR (voir Introduction partie 2.3.3). Les histogrammes de la Figure 66 permettent de visualiser dans la fluorescence totale les proportions de cellules présentant une faible (f F) et une forte (F F) fluorescence un à quatre jours après la transduction. En présence de DTG, le pourcentage de cellules à forte fluorescence a fortement diminué tandis que celui des cellules présentant une faible fluorescence n'a pas diminué (Figure 66). Le même effet du DTG a été observé dans une étude publiée (551). Le contrôle DTG permet de confirmer que la forte fluorescence résulte de l'expression de la GFP à partir de l'ADN viral intégré et que la faible fluorescence résulte de l'expression de la GFP à partir des formes circulaires à 1 et 2-LTR de l'ADN viral non intégré. Pour les cellules transduites avec les virus mutés, on observe aussi que le pourcentage de cellules présentant une forte fluorescence est fortement diminué en présence de DTG (Figure 66). Ces résultats suggèrent que le processus d'intégration de l'ADN viral dans le génome de la cellule hôte n'est pas affecté par les mutations (Figure 66).

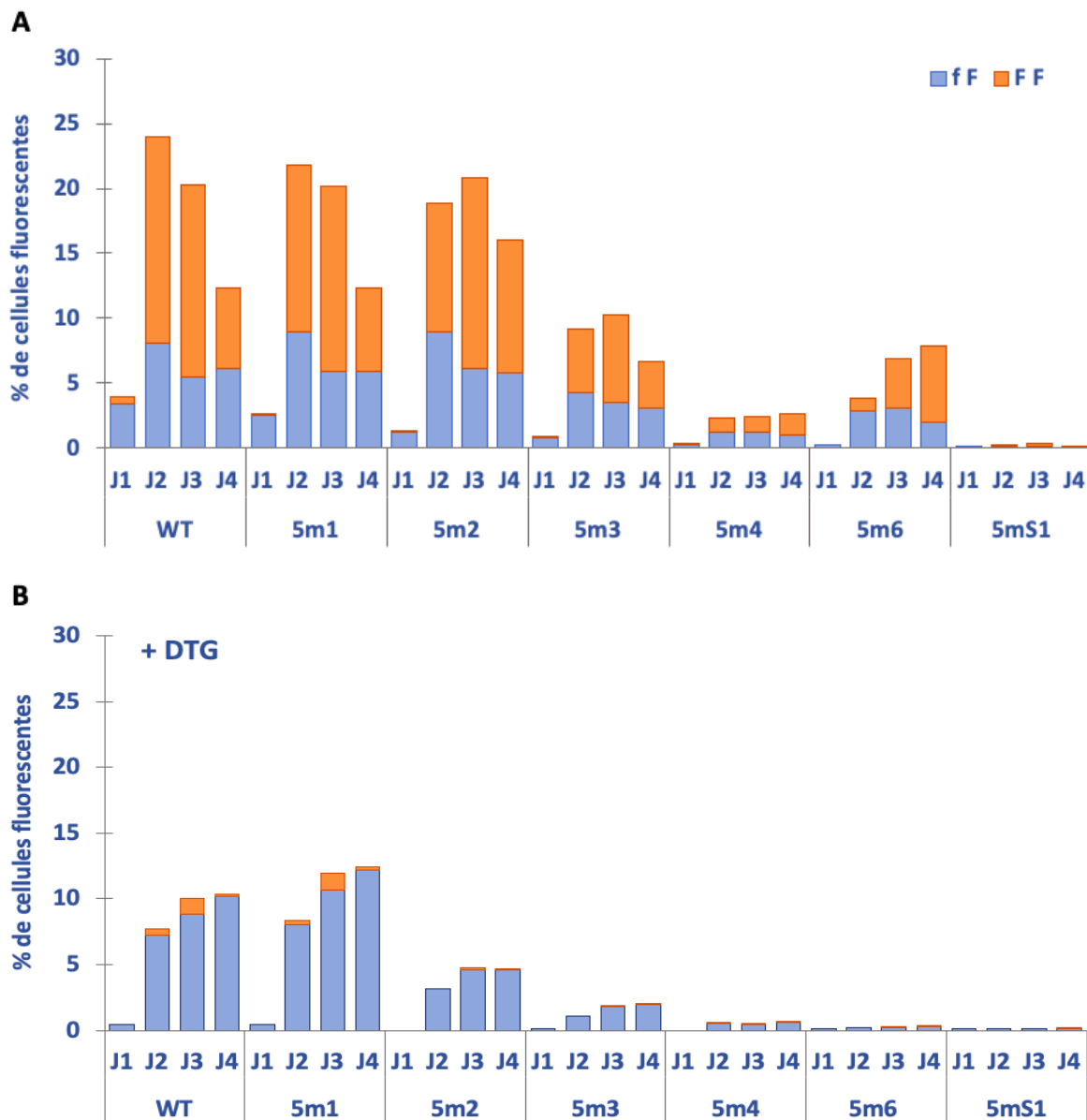


Figure 66. Pourcentages de cellules fluorescentes un à quatre jours après les transductions par le virus sauvage (WT) et les mutants 5' sans (A) ou avec DTG (B). Les pourcentages de cellules fluorescentes et les proportions des faibles (en bleu) et fortes (en orange) fluorescences ont été déterminés par cytométrie en flux à l'aide du logiciel FlowJo.

Un autre contrôle négatif correspondant à la transduction des cellules MT4 par le virus sauvage en présence d'éfavirenz (EFV) a été réalisé. Très peu de fluorescence est détectée en présence d'EFV (Figure 65) qui est un inhibiteur de la RT. Comme le montre la Figure 67, qui présente les résultats de trois expériences indépendantes, le pourcentage de cellules fluorescentes en présence d'EFV est proche de zéro (Figure 67). Ce contrôle négatif permet de vérifier que presque toutes les cellules transduites ne peuvent exprimer la GFP lorsque la transcription inverse du génome viral est inhibée. En effet, la synthèse d'ADN viral intégré et non intégré

est quasiment nulle lorsque la transcription inverse est inhibée par l'EFV. Il y avait également comme attendu très peu de cellules fluorescentes transduites par les virus mutants en présence d'EFV (résultats non montrés).

Comme attendu, la transduction des cellules MT4 par le virus mutant 5mS1, qui nous sert de contrôle négatif pour le premier transfert de brin, produit très peu de cellules fluorescentes (Figures 65-67). À J2, le pourcentage de cellules fluorescentes transduites par le 5mS1 (0,28%) est 74 fois plus petit que celui observé avec les cellules fluorescentes transduites par le virus sauvage (Figure 67A). L'expression de la GFP est donc très fortement réduite avec le 5mS1 car il ne peut réaliser efficacement la transcription inverse et par conséquent produire de l'ADN viral intégré et non intégré. Il est intéressant de noter que l'infectivité de ce mutant était réduite d'un facteur 100 par rapport au sauvage dans une étude antérieure (561). Les résultats obtenus avec ce mutant confirment la validité du système expérimental que nous utilisons pour étudier la réplication virale.

Nous n'observons pas une différence significative entre le pourcentage des cellules transduites par le virus sauvage exprimant la GFP et celui des cellules transduites par le mutant 5m1 ou 5m2 (Figure 67). Ces résultats montrent que les mutations 1 et 2 dans la séquence 5'TAR n'ont pas un effet notable sur l'expression de la GFP à partir de l'ADN viral intégré et non intégré et par conséquent sur la transcription inverse. La tige-boucle 5'TAR des deux mutants possède une tige intermédiaire constituée de seize paires de bases consécutives (Figure 64). De plus, celle du mutant 5m2 contient treize paires de bases Watson-Crick consécutives et une seule paire U-G. Cette longue tige est cependant copiée par la RT pendant la synthèse de l'ADNss. Il est intéressant de noter que parmi les quatre paires C-G/G-C déstabilisées *in vitro* par la NC (Figure 50), trois ne sont pas adjacentes à un *bulge* ou à une paire G-U chez le mutant 5m2. Les résultats suggèrent que la NC n'est pas requise *ex vivo* pour permettre la déstabilisation de la tige-boucle 5'TAR qui se produit lorsque celle-ci est copiée par la RT. Au niveau de l'ADNss, la mutation m1 n'est pas sensée empêcher le premier transfert de brin car la tige-boucle cTARm1 est constituée de dix paires de bases consécutives et le site fort putatif (CCTAG) de la NC est conservé (Figure 64). Comme mentionné dans la partie 4.1, la NC ne devrait pas pouvoir faciliter l'appariement cTARm2-3'TAR car elle ne devrait pas pouvoir déstabiliser la tige-boucle ADN cTARm2. Puisque le mutant 5m2 n'est pas significativement affecté pour la transcription inverse, la NC n'est probablement pas requise pour l'appariement cTARm2-3'TAR.

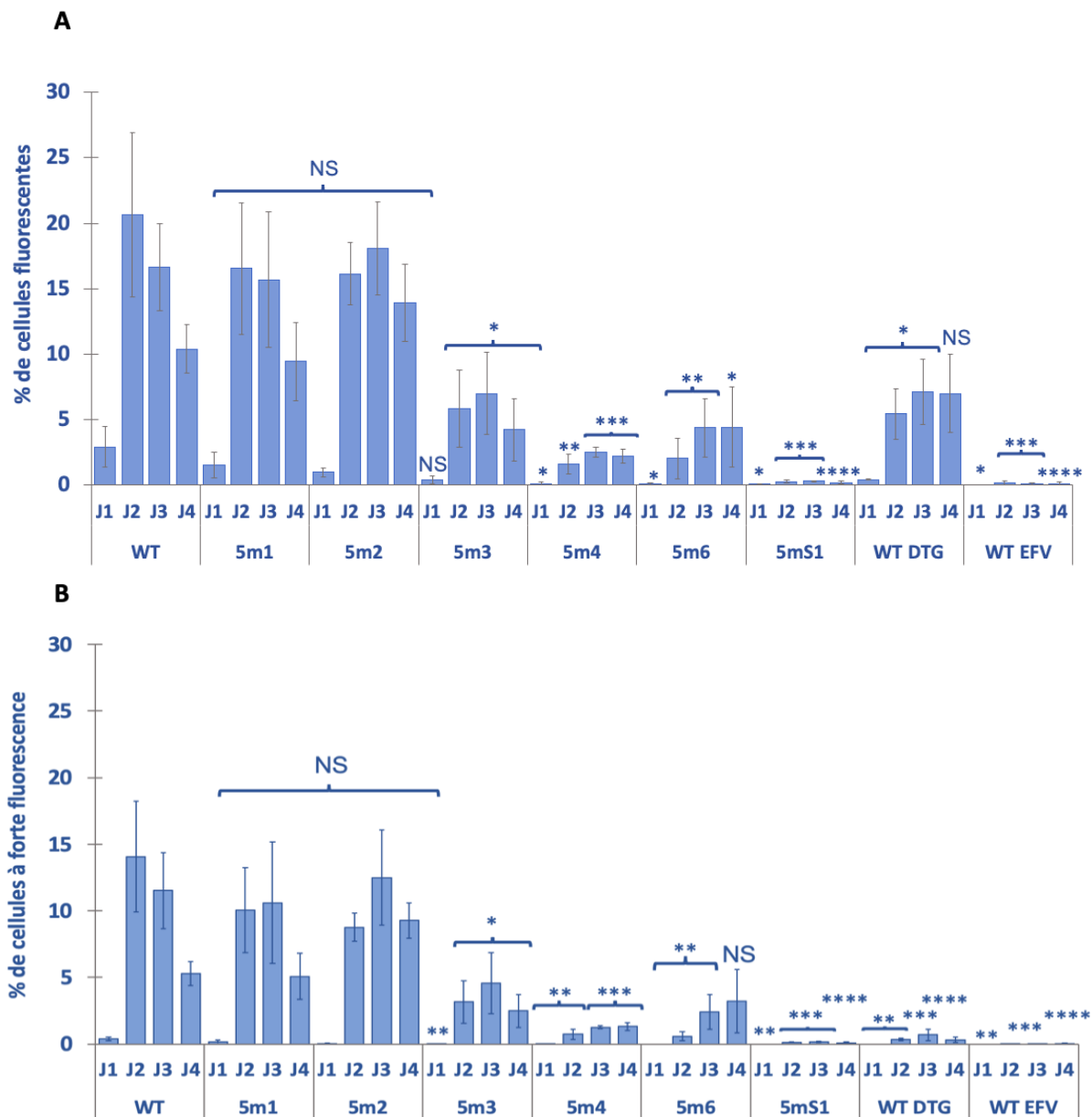


Figure 67. Pourcentages des cellules fluorescentes un à quatre jours après les transductions. Les transductions des cellules MT4 par les virus sauvage et mutants ont été effectuées trois fois. Les pourcentages de cellules fluorescentes ont été déterminés par cytométrie en flux. **(A)** Pourcentages de cellules fluorescentes (f F + F F). **(B)** Pourcentages de cellules à forte fluorescence (F F). Le WT EFV correspond à une transduction avec le virus sauvage en présence de 1 μ M EFV. Le WT DTG correspond à une transduction avec le virus sauvage en présence de 1 μ M DTG. Les étoiles représentent le p du test student obtenu lors de la comparaison des moyennes du sauvage avec celles des mutants, * : $p < 0,05$; ** : $p < 0,01$; *** : $p < 0,005$; **** : $p < 0,001$; NS : différence non significative.

Par rapport aux cellules transduites par le virus sauvage, les cellules transduites par les mutants 5m3, 5m4 et 5m6 présentent une diminution significative des pourcentages de fluorescence totale (f F + F F) (Figure 67 A). Cette diminution à J2 est 3,6 fois moins pour le 5m3, 12,8 fois moins pour le 5m4 et 10,2 fois moins pour 5m6. Si nous comparons les pourcentages de cellules

présentant une forte fluorescence (Figure 67 B), c'est-à-dire l'expression de la GFP par l'ADN proviral, nous obtenons à J2, 4,5 fois moins pour le 5m3, 19,3 fois moins pour le 5m4 et 24,2 fois moins pour le 5m6. Ces résultats sont compatibles avec une diminution importante de la transcription inverse chez ces mutants. On remarque que le pourcentage de fluorescence totale (f F + F F) ou de forte fluorescence (F F) atteint un maximum deux jours après la transduction des cellules par le virus sauvage et diminue progressivement de J3 à J4 (Figure 67). Il est à noter que les cellules transduites meurent plus rapidement que les cellules non transduites. Le pourcentage de fluorescence totale (f F + F F) atteint un maximum trois jours après la transduction des cellules par les mutants 5m3, 5m4 et 5m6 (Figure 67 A). Le retard de l'expression de la GFP peut s'expliquer par un retard de la transcription inverse qui serait affectée par les mutations présentes dans l'élément 5'TAR. Comme mentionné dans la partie 4.1, les mutations dans les mutants 5m3, 5m4 et 5m6 devraient empêcher ou diminuer significativement la transcription inverse si la NC joue un rôle essentiel ou important dans la synthèse de l'ADNss ou le premier transfert de brin. La quantification par qPCR des ADN viraux, qui sont présents dans les cellules transduites, est nécessaire pour déterminer si la synthèse de l'ADNss et/ou le premier transfert de brin sont affectés par les mutations 5m3, 5m4 et 5m6.

Finalement, nous observons chez trois mutants, parmi les cinq que nous avons conçus, une diminution de l'expression de la GFP qui traduit probablement une diminution de la transcription inverse. Plusieurs hypothèses peuvent expliquer cette diminution :

- La première, **il existe un mésappariement entre l'ADNss muté synthétisé et la région 3'TAR ARN lors du premier transfert de brin**. La mutation n'aurait pas d'effet sur la région 5'TAR ARN, l'ADNss pourrait être synthétisé et les structures secondaires déstabilisées par la NC, si celle-ci agit. Mais les mésappariements interrompent la transcription inverse. Ce scénario est identique à celui décrit par Ohi et Clever et serait vraisemblable dans le cas du mutant 5m4 (554).
- La seconde hypothèse est que **la tige-boucle 5'TAR ARN est stabilisée et empêche la synthèse de l'ADNss**. Cette hypothèse est vraisemblable pour le mutant 5m6 dont la tige-boucle 5'TAR est la plus stable (Figure 64). En revanche, la stabilité de la tige-boucle 5'TAR n'affecte pas la transcription inverse des mutants 5m1 et 5m2 et probablement pas celle du mutant 5m3 dont la stabilité est inférieure à celle du 5m2.

- La troisième hypothèse est que **la tige-boucle 5'TAR ARN stabilisée est cependant copiée par la RT et l'ADNss peut donc être synthétisé mais n'est pas déstabilisé par la NC**. Cette hypothèse peut expliquer les résultats obtenus pour les mutants 4 et 6 où les mutations dans l'ADNss peuvent impacter la fixation de la NC et la déstabilisation des structures secondaires par celle-ci.
- La dernière hypothèse est **un effet des mutations sur le processus d'encapsidation** de l'ARNg réduisant le nombre de particules virales infectieuses. Néanmoins, les mutants ont été conçus pour ne pas altérer le processus d'encapsidation de l'ARNg dans la particule virale. Pour exclure totalement l'hypothèse que la réplication réduite des mutants soit due à une diminution d'encapsidation, la quantification de l'encapsidation de l'ARNg chez les différents mutants a été déterminée.

4.4 Résultats de l'analyse de l'encapsidation de l'ARNg des mutants 5'

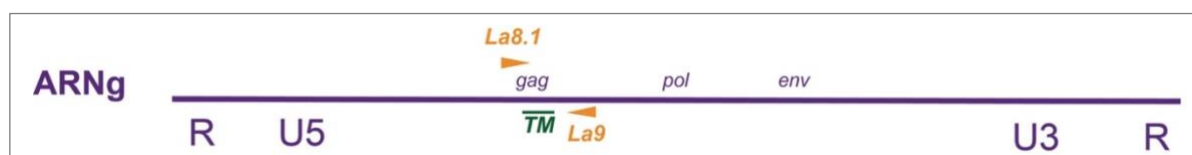


Figure 68. Amorces et sonde utilisées pour quantifier par RT-qPCR l'ARNg encapsidé dans les particules virales. L'ARNg viral qui est non épissé est quantifié grâce à l'amorce La9 s'hybridant à la séquence du gène *gag* qu'il contient mais qui est absente dans les ARN viraux épissés

Une quantification par RT-qPCR de l'ARNg présent dans les virions, qui ont servi aux transductions, a été réalisée afin de s'assurer que la même quantité d'ARNg est encapsidée dans les particules virales sauvages et mutantes. Les amorces utilisées permettent de ne quantifier que l'ARN non épissé (Figure 68). Les résultats de la quantification par RT-qPCR de l'ARNg par volume sont présentés en fonction du ratio ARNg/p24 normalisé par rapport au ratio ARNg/p24 du sauvage (Figure 69). La quantification a été réalisée au moins deux fois pour chaque type de virus. Les ratios ARN/p24 traduisent la quantité d'ARNg par rapport à la quantité de p24 mesurée en test ELISA qui reflète le nombre de particules virales utilisées. En accord avec ce qui avait été montré dans une étude antérieure (561), nous observons que l'encapsidation de l'ARNg du mutant 5mS1 est équivalente à celle du virus sauvage. L'encapsidation de l'ARNg chez les mutants 5m4 et 5m6 est équivalente ou légèrement supérieure à celle du virus sauvage. La diminution de l'expression de la GFP chez ces deux mutants n'est donc pas due à une réduction de l'encapsidation.

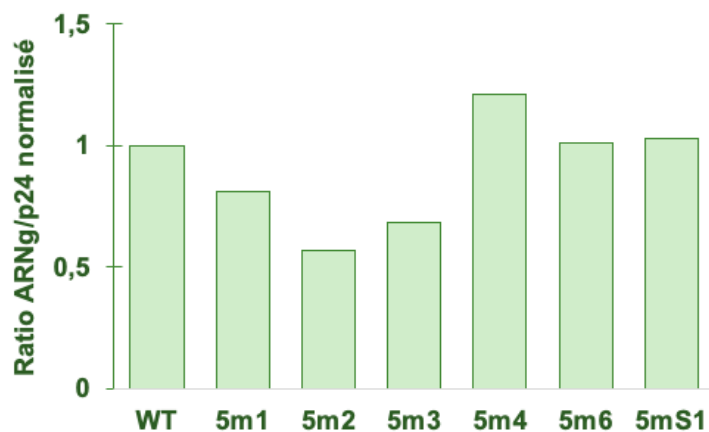


Figure 69. Histogrammes représentant les ratios ARNg/p24 des mutants 5'. Les ratios sont issus d'au moins deux quantifications d'ARNg et deux quantifications p24. Les quantités d'ARNg sont cohérentes avec les quantités de p24 mesurées pour les virus mutants par rapport au virus sauvage.

L'encapsulation de l'ARNg chez les mutants 5m1, 5m2 et 5m3 est diminuée respectivement de 19%, 43% et 32% par rapport à celle du sauvage. Ces diminutions d'encapsulation sont faibles et ne peuvent pas avoir un effet significatif sur la réplication de ces trois mutants. Cette assertion est étayée par l'observation que le mutant 5m2 n'est pas altéré dans l'expression de la GFP bien que l'encapsulation soit diminuée de 43% par rapport à celle du sauvage. Par conséquent et comme attendu, nos mutations dans l'élément 5'TAR n'ont pas un effet important sur l'encapsulation de l'ARNg dans la particule virale.

4.5 qPCR de l'ADN viral

Afin de déterminer si la diminution de la transcription inverse est due à un effet des mutations sur la synthèse de l'ADNss et/ou sur l'hybridation cTAR-3'TAR, une quantification des ADN viraux a été réalisée pour certains mutants. Les ADN ont été extraits comme décrit dans la partie Matériel et Méthodes à différents temps après la transduction. Le couple d'amorce MH531/MH532 permet de quantifier l'ADN tardif (ADNt) viral obtenu après le second transfert de brin (562, 563). Le couple d'amorce FwSS/AA55 a été conçu pour quantifier l'ADNss avant le premier transfert de brin (554, 564). Le dernier couple U3NL9496/AA55 a été conçu (564) pour mesurer la quantité d'ADNss allongé après le premier transfert de brin (Figure 70). Une quantification de l'ADNt viral à partir des cellules transduites par les particules virales sauvages traitées à la benzonase et mises en présence d'EFV, inhibiteur de transcriptase inverse, a été réalisée afin de vérifier qu'il ne reste pas d'ADN plasmidique non digéré qui serait amplifié. Aucune amplification par qPCR d'ADNt n'a été observée aux temps 4 h et 24 h post-transduction (résultats non présentés).

Les courbes d'étalonnage utilisées pour la quantification des ADN viraux sont présentées dans la Figure 71. Elles montrent que nous avons défini les conditions optimales pour quantifier de faibles quantités d'ADN viraux. Les mutants 5m1 et 5m2 présentent des pourcentages de cellules fluorescentes similaires au virus sauvage. C'est pourquoi nous avons réalisé uniquement une quantification de l'ADNt chez les cellules transduites par le 5m2 afin de confirmer nos observations précédentes.

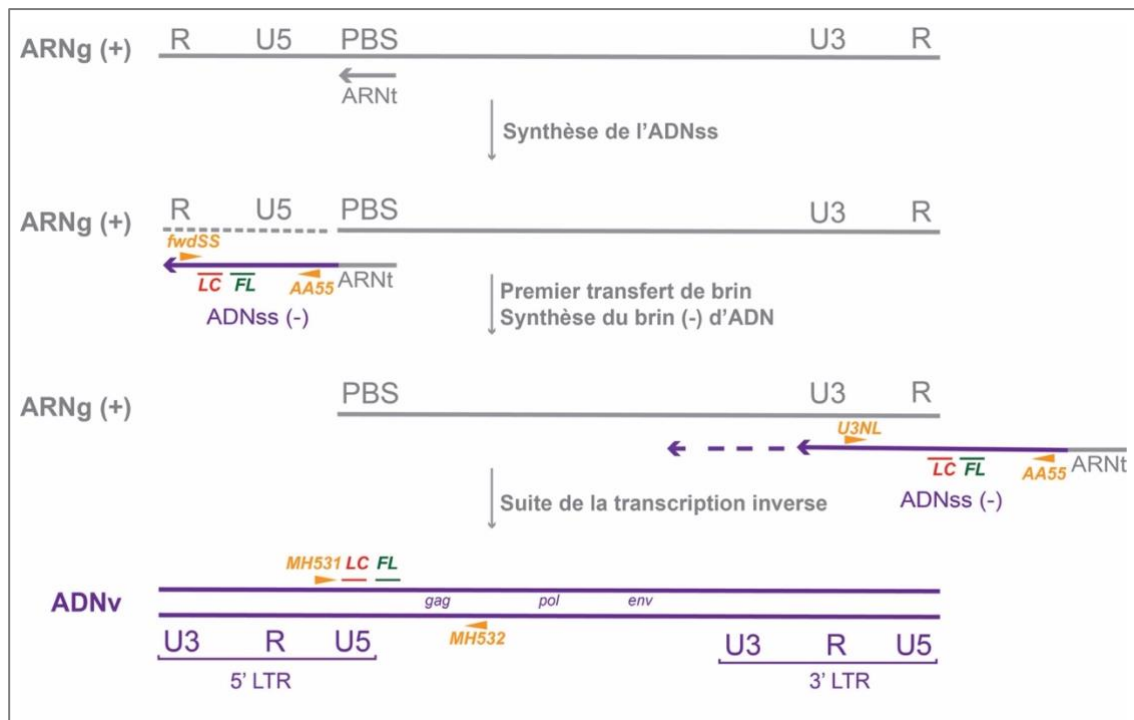


Figure 70. Schéma simplifié de la transcription inverse avec les amorces et sondes utilisées en qPCR. En gris l'ARNg et l'ARNt. En violet, l'ADN viral. En orange, les amorces utilisées en qPCR pour quantifier soit l'ADNss, soit le brin (-) d'ADN en cours de synthèse après le premier transfert de brin, soit l'ADN viral tardif après le second transfert de brin. En rouge et vert, les sondes. Ce sont les mêmes pour les couples d'amorces FwSS(-)/AA55 et U3NL9496/AA55.

Le nombre de copies d'ADNt à 8 h et 24 h post-transduction dans les cellules transduites par le 5m2 est similaire à celui obtenu avec le sauvage (Figure 72 A). Ces résultats sont cohérents avec ceux de l'expression de la GFP et confirment que le mutant 5m2 ne présente pas une transcription inverse altérée. Les cellules transduites par les mutants 5m3 et 5mS1 contiennent très peu d'ADNt à 8 h et 24 h post-transduction (Figure 72 B). Ces résultats sont cohérents avec la faible expression de la GFP chez ces mutants. Pour le mutant 5mS1, cette observation est en cohérence avec les résultats de Ohi et Clever (554).

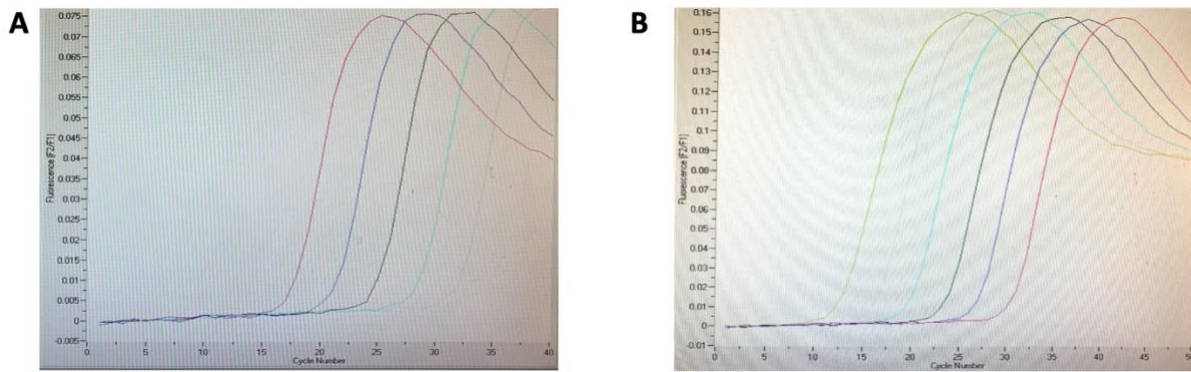


Figure 71. Courbes d'étalonnage de l'amplification par qPCR de l'ADNt (A) et l'ADNss (B). La gamme de la quantification de l'ADNt correspond à une dilution de 10^7 copies à 10^3 copies de pLevy. La gamme de la quantification de l'ADNss correspond à une dilution de 10^8 copies à 10^3 copies de pLevy.

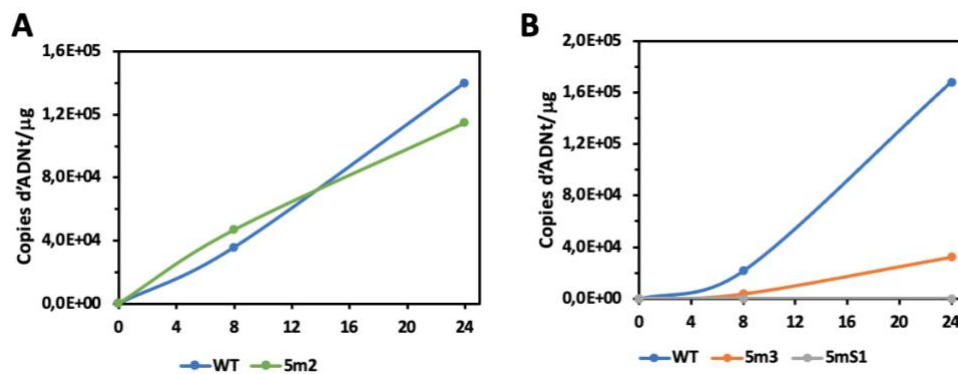


Figure 72. Expériences représentatives de la quantification par qPCR de l'ADNt. (A) Quantification de l'ADNt des cellules transduites par le virus sauvage et le mutant 5m2. (B) Quantification de l'ADNt des cellules transduites par le virus sauvage et les mutants 5m3 et 5mS1. L'ADNt a été quantifié à 0, 8 h et 24 h post-transduction. Le nombre de copies d'ADNt est déterminé en fonction de la quantité en μg d'ADN total (ADN viral plus ADN cellulaire).

Les résultats de plusieurs quantifications de l'ADNt et de l'ADNss sont présentés dans la Figure 73. La quantification de l'ADNss dans les cellules transduites par le mutant 5mS1 confirme les résultats de Ohi et Clever (554) qui avaient montré une très forte réduction de la quantité d'ADNss dans les cellules infectées par ce mutant. L'hypothèse la plus vraisemblable qui avait été proposée est que le mésappariement de l'extrémité 3' de la séquence cTARmS1 avec l'extrémité 5' de la séquence 3'TAR sauvage conduit à une dégradation rapide de l'ADNss. Pour le mutant 5m3, nous observons respectivement aux temps 8 h et 24 h post-transduction une diminution de 8 fois et 6 fois de la quantité d'ADNt par rapport au sauvage. La quantification de l'ADNss pour ce mutant montre une diminution de 3,6 fois du nombre de copies par rapport à celui du sauvage à 2 h et jusqu'à 150 fois 4 h après la transduction (Figure 73 B). La synthèse de l'ADNss pourrait être affectée pour le mutant 5m3 dès le début de la transcription inverse. La partie inférieure de la tige-boucle 5'TARm3 qui est constituée de quatorze paires de bases consécutives serait trop stable pour pouvoir être déstabilisée par la NC

et donc être copiée par la RT. L'hypothèse d'une dégradation rapide de l'ADNss à cause de mésappariements entre l'extrémité 5' de l'ADN cTARm3 et l'extrémité 3' de l'ARN 3'TAR sauvage est défavorisée car il avait été montré que des mésappariements à l'extérieur de la région 1-10 de la séquence TAR n'ont pas d'effet sur la transcription inverse (554). Nous ne pouvons cependant pas totalement exclure l'hypothèse que la structure de la tige-boucle cTARm3 déclenche la dégradation de l'ADNss.

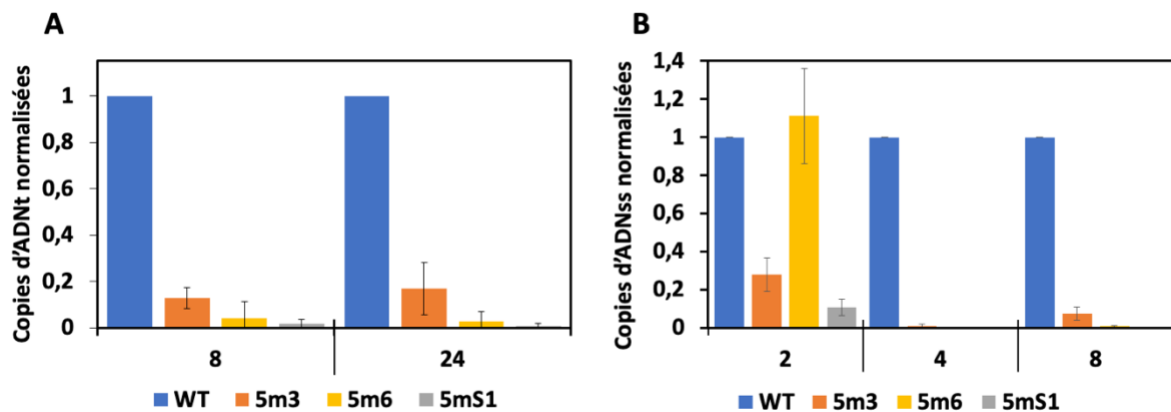


Figure 73. Copies d'ADNt et ADNss normalisées par rapport au WT. (A) Copies d'ADNt normalisées par rapport au : WT (bleu), 5m3 (orange), 5m6 (jaune) et 5mS1 (gris) aux temps 8 h et 24 h post-transduction. (B) Copies d'ADNss normalisées par rapport au : WT (bleu), 5m3 (orange), 5m6 (jaune) et 5mS1 (gris) aux temps 2 h, 4 h et 8 h post-transduction.

Pour le mutant 5m6, nous observons respectivement aux temps 8 h et 24 h post-transduction une diminution de 24 fois et 39 fois de la quantité d'ADNt par rapport au sauvage (Figure 73 A). La quantification de l'ADNss pour ce mutant montre un nombre de copies similaire à celui du sauvage 2 h après la transduction (Figure 73 B). La RT serait donc capable de copier la tige-boucle 5'TARm6 possédant vingt-et-une paires de bases consécutives dont deux sont des U-G/G-U. En revanche, quasiment aucune copie n'est détectée à 4 h et 8 h. Ces résultats suggèrent que la structure de la tige-boucle cTARm6 déclencherait la dégradation de l'ADNss. Ceci pourrait être dû à ce que la NC soit peu efficace pour déstabiliser cette tige-boucle (voir partie 4.1) et donc permettre le premier transfert de brin. Toutefois, quelques molécules d'ADNss produites à 2 h pourrait être déstabilisées, donc servir au premier transfert de brin et ainsi conduire à la synthèse de quelques molécules d'ADNt (Figure 73 A). Une autre hypothèse pour expliquer la présence d'ADNt bien que l'on ne détecte quasiment pas d'ADNss à 4 h et 8 h résulte des observations ci-après. Dans la majorité des cas, le premier transfert de brin chez le VIH-1 se produit *ex vivo* après la copie complète de la séquence R située en 5' (554). La réplication du VIH-1 est cependant possible avec des séquences R tronquées (502). Plus précisément, un ADNss tronqué dont la séquence cpoly(A) constitue son extrémité 3' peut

s'apparier avec la séquence poly(A) située à l'extrémité 3' de l'ARNg et donc réaliser le premier transfert de brin (Figure 74).

Les résultats des quantifications de l'ADNt et de l'ADNss confirment que la faible expression de la GFP chez les cellules transduites par les mutants 5m3 et 5m6 est due à une altération de la transcription inverse. La dégradation de l'ADNss chez le mutant 5m6 est fortement suggérée par nos résultats tandis que cette hypothèse semble moins probable pour le mutant 5m3. En effet, il est difficile d'expliquer que la RT ne pourrait pas copier la tige-boucle 5'TARm3 tandis qu'elle en serait capable pour la tige-boucle 5'TARm6 qui contient plus de paires de bases consécutives et est plus stable (Figure 64).

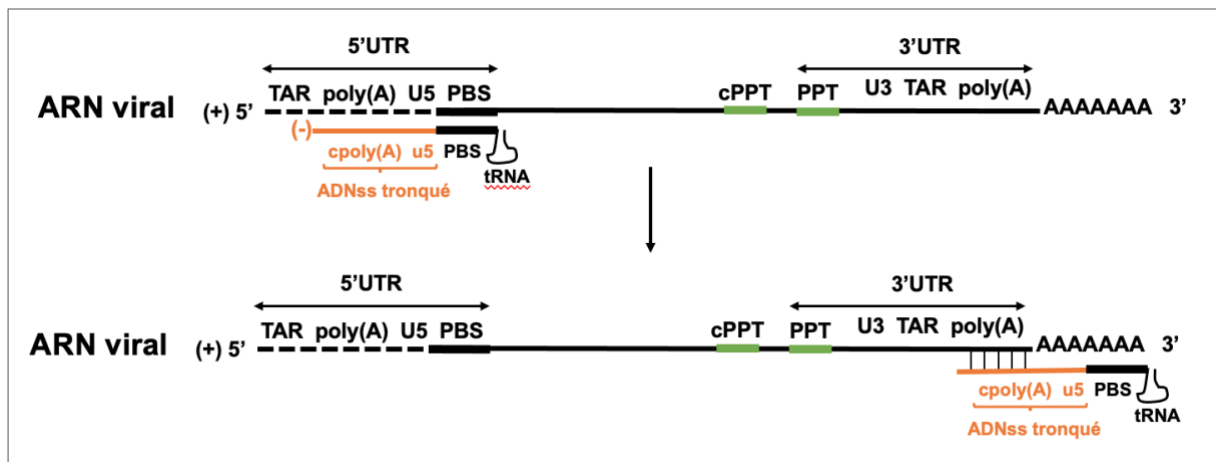


Figure 74. Premier transfert de brin avec un ADNss tronqué.

5. Analyse des effets produits par les mutations dans la séquence 3' TAR

5.1 Structures secondaires des mutants 3'TAR

Chez le virus sauvage, la tige boucle 3'TAR (Figure 75) diffère légèrement de la tige-boucle 5'TAR (Figure 64 A) car la tige inférieure est allongée de deux paires de bases (550). Toutefois, la tige intermédiaire est de même longueur (onze paires de bases consécutives) chez les tiges-boucles 5'TAR et 3'TAR.

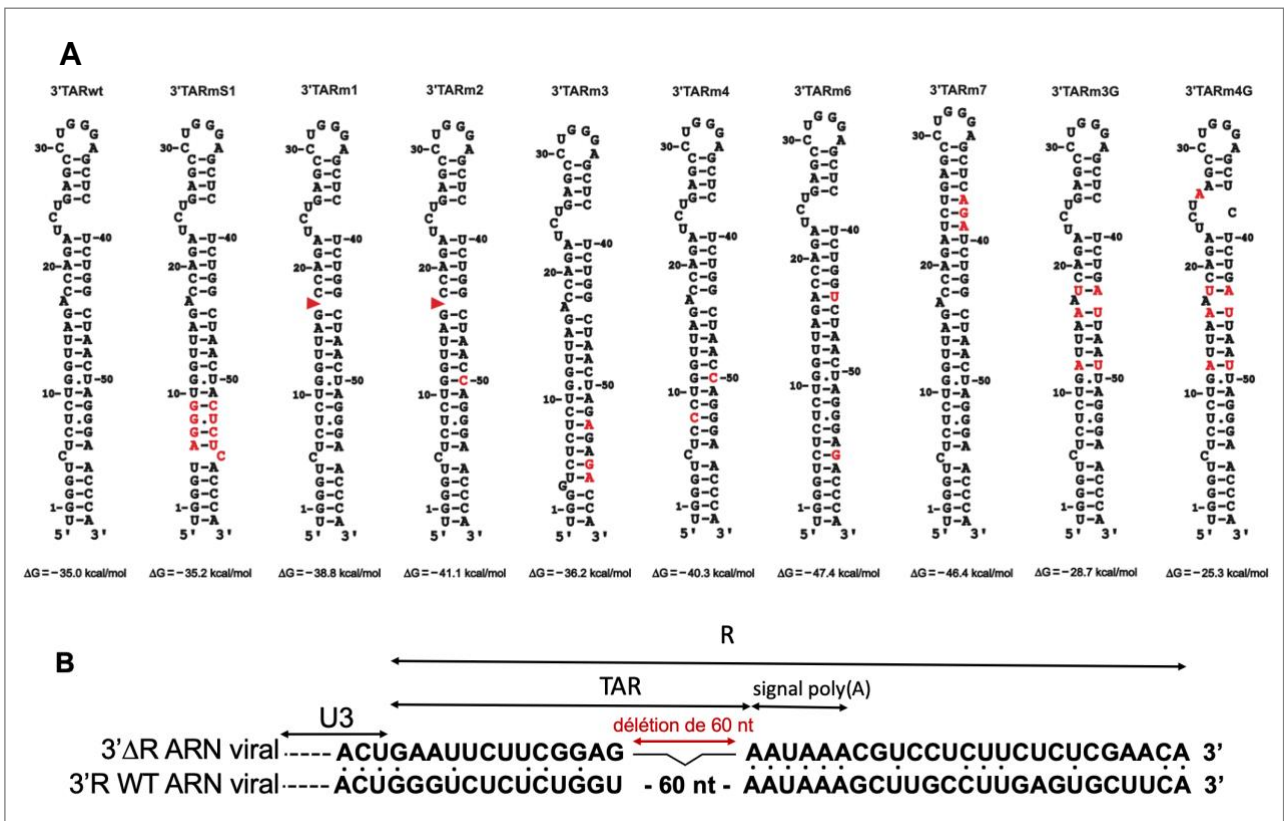


Figure 75. Structures secondaires prédites pour les 3'TAR mutés (A) et schéma des modifications du mutant 3'ΔR. Le modèle de base est l'élément 3'TAR dans la structure secondaire du génome du VIH-1 (550). Les mutations sont en rouge. Les triangles correspondent à des délétions. Les repliements prédictifs ont été réalisés à l'aide du logiciel *mfold* (558). **(B)** Blast d'une partie de la région 3'UTR du mutant 3'DR et le sauvage. 60 nucléotides du R ont été supprimés, le signal de polyadénylation a été conservé, et des mutations dans région R ont été ajoutées afin de garder dans l'ARN pré-messager la même distance entre le signal de polyadénylation et le site de clivage.

Trois mutants supplémentaires (m7, m3G et m4G) ont été conçus. Le mutant m7 a été construit uniquement pour le 3'TAR car la suppression du *bulge* à trois nucléotides en amont de la tige apicale du 5'TAR abolirait l'activation par Tat de la transcription de l'ADN proviral. La partie supérieure de la tige-boucle est composée de douze paires de bases Watson-Crick consécutives et la paire G26-C39 n'est plus adjacente à une région simple-brin mais est complètement intégrée dans une tige. De plus, le ΔG calculé par *mfold* de la tige-boucle 3'TAR est fortement

augmenté par rapport à celui du sauvage. Par conséquent, la déstabilisation de la tige-boucle 3'TAR_{m7} par la NC ne devrait pas être possible ou fortement réduite. Comme le montre les ΔG calculés par mfold, les mutants m3G et m4G sont prédits posséder une tige-boucle moins stable que celle du sauvage. Chez ces deux mutants, les mutations ne sont pas conçues pour augmenter la stabilité de la tige-boucle 3'TAR, elles ont pour objectif de diminuer l'interaction NC-TAR dépendant de guanines (303, 304). Plus précisément, dans le mutant 3'TAR_{m3G}, les trois paires C-G/G-C qui sont déstabilisées par la NC ont été remplacées par les paires U-A/A-U. Le mutant 3'TAR_{m4G} possède une substitution supplémentaire par rapport au 3'TAR_{m3G} qui supprime une paire G-C de la tige apicale. La suppression des sites préférentiellement déstabilisés par la NC devrait empêcher ou fortement diminuer la déstabilisation des tiges-boucles 3'TAR_{m3G} et 3'TAR_{m4G}.

Un autre contrôle négatif, le mutant 3'TAR- ΔR ne possédant pas de structure 3'TAR, a été conçu. Dans ce mutant, la séquence R est presque complètement déléetée, il ne subsiste que le signal de polyadénylation. Le premier transfert de brin ne devrait donc pas se produire et la transcription inverse devrait être stoppée à cette étape.

5.2 Résultats des transfections par les mutants 3'

Une fois les différents vecteurs pLevD3-m construits, les particules virales ont été obtenues grâce à des transfections des cellules HEK293T comme décrit dans la partie Matériel et Méthodes. Les virions ont été quantifiés par un test ELISA permettant une quantification de la p24, protéine constituant la capsid. Les résultats des quantifications p24 sont présentés dans le Tableau 5 et représentent une moyenne d'au moins deux quantifications par type de virus. Neuf des dix productions de particules virales des mutants présentent des valeurs de p24 inférieures à celle du WT, excepté le 3m7. Cette diminution de p24 par rapport au sauvage peut être légère (inférieure ou égale à 30% pour les mutants 3m1, 3m2 et 3m6), modérée (de 40% à 60% pour les mutants 3m3, 3m4 et 3mS1) ou assez élevée (de 69% à 76% pour les mutants 3m3G, 3m4G et 3 ΔR). Ces résultats suggèrent que l'expression des virus 3m1, 3m2 et 3m6 dans les cellules HEK293T n'est pas significativement altérée. En revanche, la production de particules virales par les cellules HEK293T transfectées par les mutants 3m3, 3m4, 3mS1, 3m3G, 3m4G et 3 ΔR est réduite par rapport au sauvage d'un facteur d'environ 1,7 à 4. La traduction pourrait être diminuée dans ces six mutants à cause des mutations présentes dans l'extrémité 3' de l'ARNg. Ces mutations pourraient provoquer une dégradation plus rapide des ARNm viraux et donc diminuer leur traduction. En revanche, les particules virales produites devraient contenir de l'ARNg car les mutations ne devraient pas avoir un effet sur

l'encapsidation. Les surnageants contenant les virions ont été traités à la benzonase (DNase) comme mentionné dans la partie Matériel et Méthodes afin d'éliminer l'ADN plasmidique qui a été utilisé pour produire les particules virales et qui pourrait être amplifié lors des quantifications par qPCR des ADN viraux.

WT	3m1	3m2	3m3	3m4	3m6	3m7	3mS1	3m3G	3m4G	3ΔR
2,39 ng/μl	1,9 ng/μl	1,67 ng/μl	1,43 ng/μl	1,27 ng/μl	1,76 ng/μl	2,59 ng/μl	0,96 ng/μl	0,64 ng/μl	0,57 ng/μl	0,75 ng/μl

Tableau 5 : Quantités de p24 mesurées pour les onze productions de particules virales.

5.3 Résultats des transductions par les mutants 3'

Les transductions des cellules MT4 par les mutants 3' ont été réalisées dans les mêmes conditions que pour les mutants 5', c'est-à-dire en triplicat biologique et avec la même quantité de p24 correspondant à priori à la même quantité de particules virales. À titre d'exemple, les résultats d'une expérience obtenus par cytométrie en flux deux jours après les transductions sont présentés dans la Figure 76. On observe que par rapport aux cellules transduites par le virus sauvage, le nombre de cellules fluorescentes diminue fortement lorsque les cellules sont transduites par les mutants 3'. Pour les cellules transduites par les mutants 3m1 à 3m6 en présence de DTG, on observe comme pour le virus sauvage que le pourcentage de cellules présentant une forte fluorescence est fortement diminué par rapport à celui des cellules présentant une faible fluorescence (Figure 77). Ces résultats suggèrent que le processus d'intégration de l'ADN viral dans le génome de la cellule hôte n'est pas affecté par les mutations en 3'.

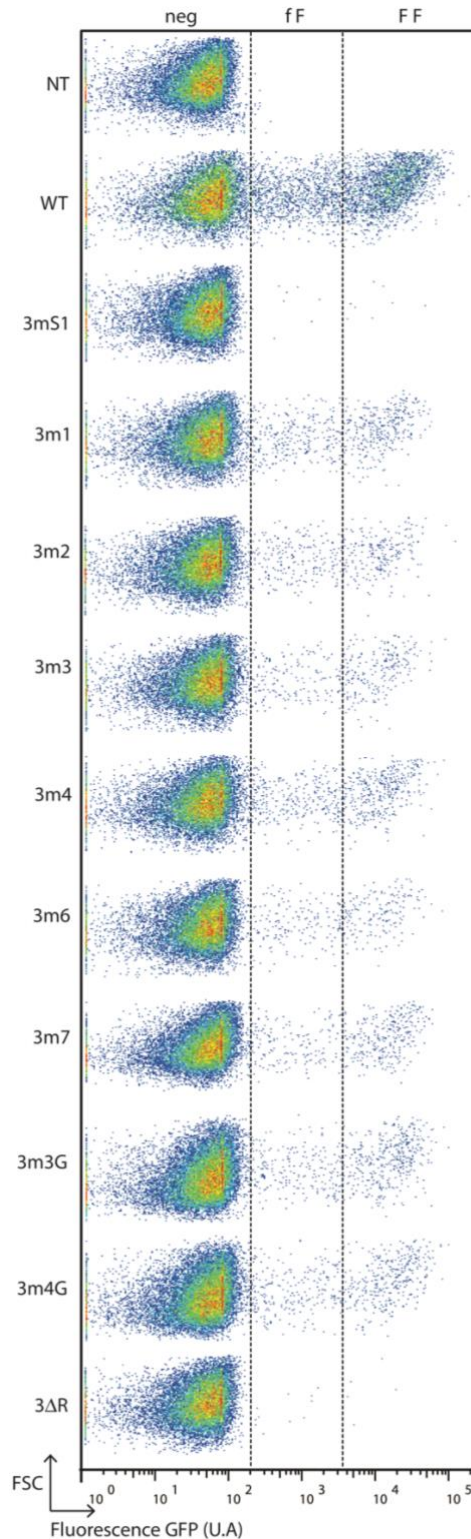


Figure 76. Fluorescence des cellules MT4 déterminée par cytométrie en flux deux jours après les transductions. Trois niveaux d'intensité de fluorescence sont détectés, un négatif (neg), un de faible intensité (f F) et un de forte intensité (F F). Le f F correspond aux formes ADN viraux circulaires à 1 et 2-LTR exprimant faiblement la GFP. Le F F correspond à une forte expression de la GFP à partir de l'ADN proviral (171, 176, 178). FSC : *Foward scatter* (paramètre mesuré permettant de discriminer les cellules par leur taille). NT, cellules MT4 non transduites ; WT, cellules MT4 transduites par le virus sauvage. 3m1 à 3mΔR, cellules MT4 transduites par les virus mutants 3m1 à 3mΔR.

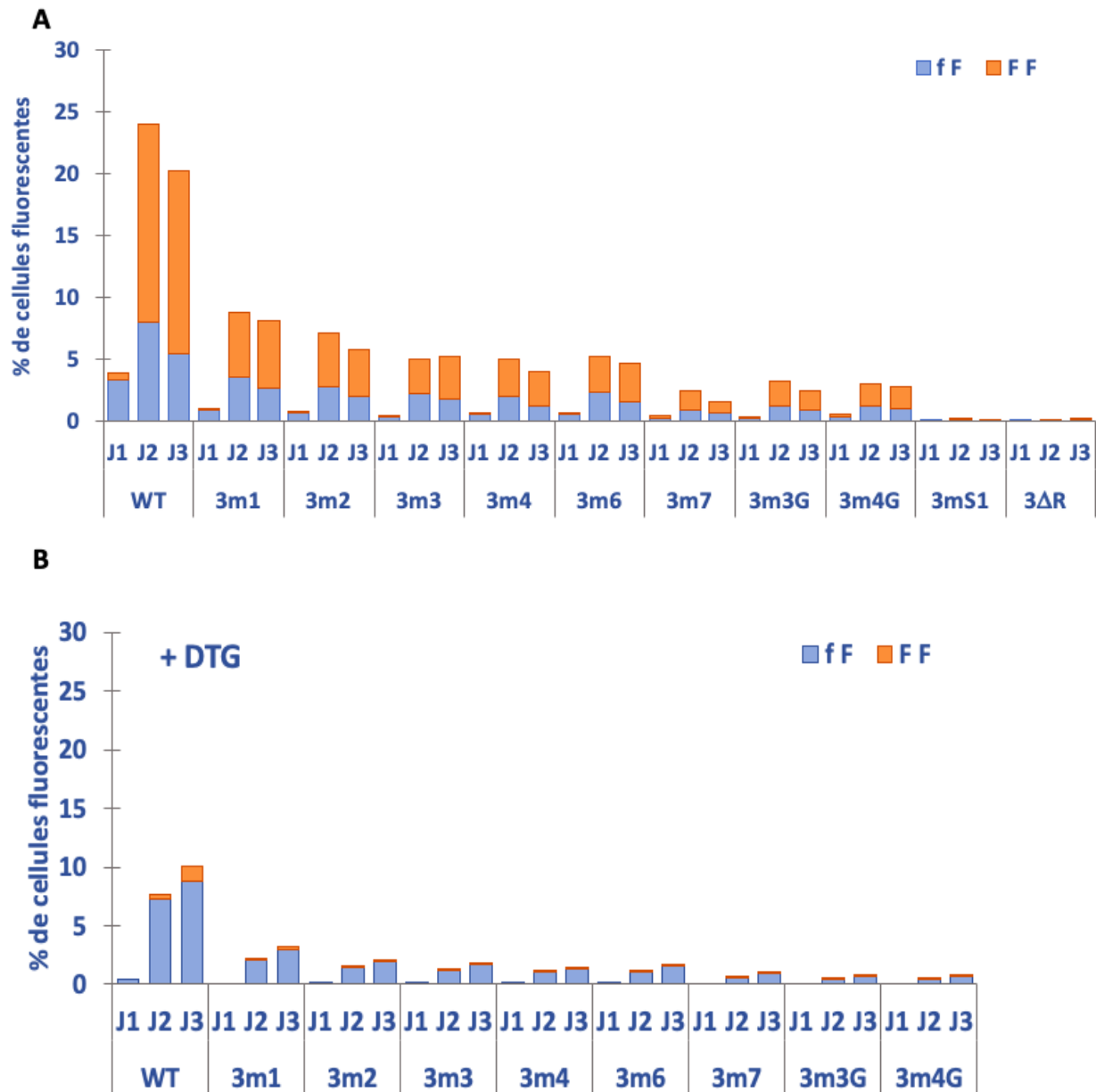


Figure 77. Pourcentages de cellules fluorescentes un à trois jours après les transductions par le virus sauvage (WT) et les mutants 3' sans (A) ou avec DTG (B). Les pourcentages de cellules fluorescentes et les proportions des faibles (en bleu) et fortes (en orange) fluorescences ont été déterminés par cytométrie en flux à l'aide du logiciel FlowJo.

Comme attendu, avec le mutant 3mS1 de Ohi et Cleaver (554) qui nous sert de contrôle négatif, nous n'observons quasiment pas de cellules fluorescentes (0,18% en moyenne de cellules exprimant la GFP à J2) (Figure 78 A) ; le pourcentage de cellules fluorescentes transduites par le 3mS1 est 115 fois plus petit que celui observé avec les cellules fluorescentes transduites par le virus sauvage. Il est intéressant de noter que l'infectivité de ce mutant était réduite d'un facteur 100 par rapport au sauvage dans une étude antérieure (554). Le premier transfert de brin est prédit ne pas se produire chez ce mutant à cause du mésappariement entre l'extrémité 3' de la tige-boucle sauvage cTAR et l'extrémité 5' de la tige-boucle 3'TARmS1. Pour l'autre contrôle négatif, le mutant 3ΔR dont une grande partie du R est délété, nous observons une quasi-absence de cellules fluorescences. En effet, à J2 le pourcentage de cellules fluorescentes transduites par le 3ΔR est 148 fois plus petit que celui observé avec les cellules fluorescentes transduites par le virus sauvage. Cela confirme que la séquence R en 3' est indispensable au premier transfert de brin en s'appariant avec la séquence r de l'ADNss. Autrement dit, si les tiges-boucles 3'TAR mutées sont trop stables pour être ouvertes et ne peuvent donc s'apparier avec la tige-boucle cTAR de l'ADNss, le premier transfert de brin ne devrait pas avoir lieu ou être moins efficace et par conséquent l'expression de la GFP à partir de l'ADN proviral ne devrait pas se produire ou être fortement réduite. À J2 et par rapport aux cellules transduites par le virus sauvage, nous observons une diminution importante et significative du pourcentage de cellules fluorescentes lorsqu'elles ont été transduites par les mutants 3' (Figure 78 A) : 3m1 (4,4 fois moins), 3m2 (4,8 fois moins), 3m3 (6,4 fois moins), 3m4 (5,6 fois moins), 3m6 (7,4 fois moins), 3m7 (6 fois moins), 3m3G (10,2 fois moins) et 3m4G (8,8 fois moins). Finalement, nous observons chez les huit mutants 3', qui ont été conçus pour ne pas permettre à la NC d'apparier les tiges-boucles TAR et cTAR, une diminution significative de l'expression de la GFP. Cette diminution est compatible avec une diminution de la transcription inverse qui serait causée par la perte d'efficacité du premier transfert de brin. La quantification par qPCR des ADN viraux, qui sont présents dans les cellules transduites, est nécessaire pour confirmer que le premier transfert de brin et la transcription inverse sont affectés par les mutations. Les mutants ont été conçus pour ne pas altérer le processus d'encapsidation de l'ARNg dans la particule virale. Pour exclure totalement l'hypothèse que la réplication réduite des mutants soit due à une diminution d'encapsidation, la quantification de l'encapsidation de l'ARNg chez les différents mutants 3' a été déterminée.

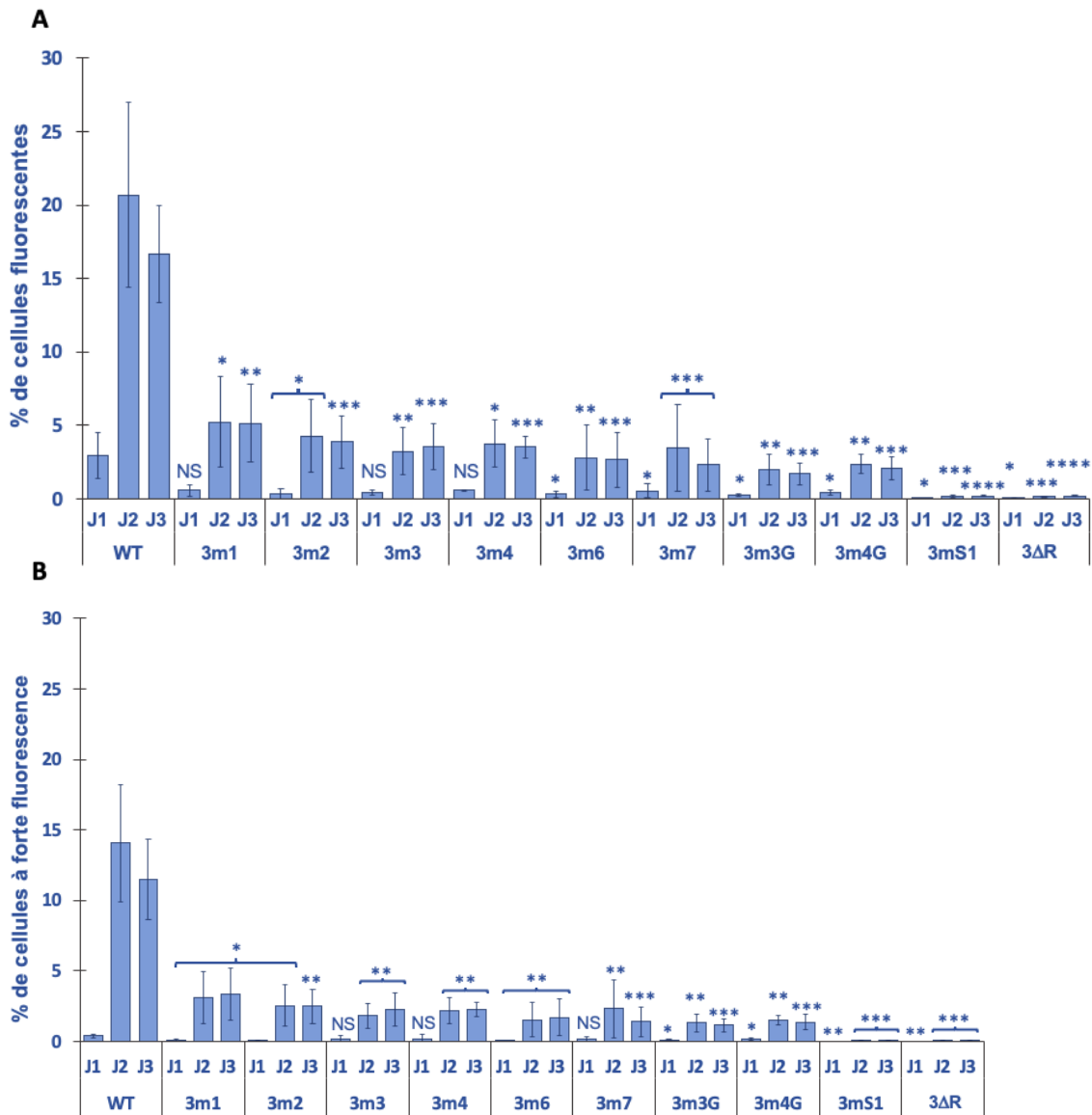


Figure 78. Pourcentages des cellules fluorescentes un à trois jours après les transductions. Les transductions des cellules MT4 par les virus sauvage et mutants ont été effectuées trois fois. Les pourcentages de cellules fluorescentes ont été déterminés par cytométrie en flux. **(A)** Pourcentages de cellules fluorescentes (f F + F F). **(B)** Pourcentages de cellules à forte fluorescence (F F). Les étoiles représentent le p du test student obtenu lors de la comparaison des moyennes du sauvage avec celles des mutants, * : $p < 0,05$; ** : $p < 0,01$; *** : $p < 0,005$; **** : $p < 0,001$; NS : différence non significative.

5.4 Résultats de l'analyse de l'encapsidation de l'ARNg des mutants 3'

La quantification a été réalisée au moins deux fois pour chaque type de mutant 3'. Les ratios ARNg/p24 traduisent la quantité d'ARNg par rapport à la quantité de p24 mesurée en test ELISA qui reflète le nombre de particules virales utilisées. Les résultats ont été normalisés par rapport au sauvage (Figure 79). L'encapsidation de l'ARNg chez les mutants 3m3G, 3m4G et 3ΔR est équivalente ou légèrement supérieure à celle du virus sauvage. La forte diminution de

l'expression de la GFP chez ces trois mutants n'est donc pas due à une réduction de l'encapsidation. L'encapsidation de l'ARNg chez les mutants 3m1, 3m2 et 3m7 est diminuée respectivement de 11%, 25% et 17% par rapport à celle du sauvage. Ces diminutions d'encapsidation sont faibles et ne peuvent pas rendre compte de la forte diminution de l'expression de la GFP observée avec ces trois mutants. L'encapsidation de l'ARNg chez les mutants 3mS1, 3m3 et 3m4 est diminuée respectivement de 41%, 45% et 41% par rapport à celle du sauvage. Ces diminutions d'encapsidation sont modérées et ne pourraient rendre compte que partiellement de la forte diminution de l'expression de la GFP observée avec ces trois mutants. En revanche, l'encapsidation de l'ARNg chez le 3m6 est réduite de 72%, c'est-à-dire d'un facteur 3,6 qui explique en partie la diminution d'environ un facteur 7,4 de l'expression de la GFP.

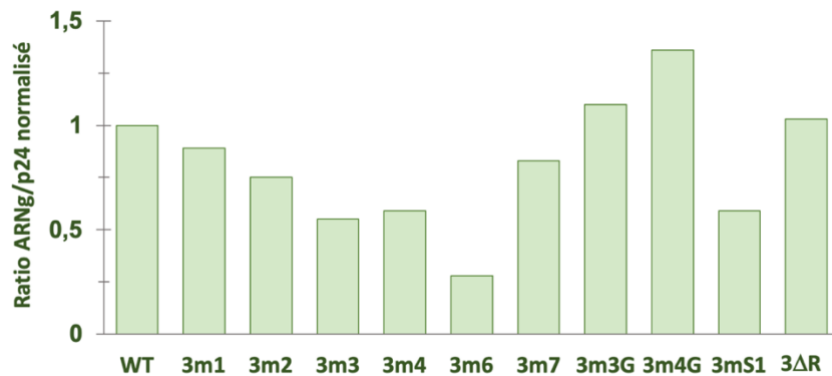


Figure 79. Histogrammes représentant les ratios ARNg/p24 des mutants 3'. Les ratios sont issus d'au moins deux quantifications d'ARNg et deux quantifications p24. Les quantités d'ARNg sont cohérentes avec les quantités de p24 mesurées pour les virus mutants par rapport au virus sauvage.

5.5 qPCR ADN viral

Comme pour les mutants 5', une quantification des ADN viraux est réalisée à différents temps après transduction avec les mêmes conditions de transductions que pour les résultats de fluorescence. Les quantifications ont été réalisées avec les mutants 3m2, 3m3, 3m7, 3m3G et 3m4G qui nous semblent les plus intéressants d'après les résultats décrits précédemment. Les quantifications avec les mutants 3m3G et 3m4G ont été effectuées plusieurs fois mais les résultats obtenus étaient non reproductibles et sont pour cette raison non présentés. Le mutant 3m2 présente des niveaux faiblement diminués d'ADNt à 8 h et 24 h de 1,3 et 1,7 fois par rapport au sauvage (Figures 80 A et 81 A). Les résultats d'ADNt ne semblent expliquer qu'en partie la diminution de 4,8 fois de l'expression de la GFP à J2 (Figure 78). En ce qui concerne l'ADNss, nous observons des diminutions de 2,5 à 3 fois entre le 3m2 et le sauvage à 2 h et 4 h (Figures 80 B et 81 B). Les résultats d'ADNss pour le 3m2 montrent que celui-ci est bien

synthétisé, comme c'est attendu, mais à un niveau significativement diminué par rapport au sauvage. Ce résultat peut s'interpréter comme une synthèse diminuée de l'ADNss ou une dégradation de ce dernier résultant de son incapacité à s'apparier via la tige-boucle cTAR sauvage avec la tige-boucle 3'TARm2, cette dernière étant trop stable pour être ouverte par la NC. Il est intéressant de souligner que trois des quatre guanines reconnues par la NC ne sont plus à proximité d'un bulge ou d'une paire U-G dans la tige-boucle 3'TARm2 (Figures 50 et 75). Nous favorisons l'hypothèse de la synthèse diminuée par rapport à l'hypothèse de la dégradation car on observe une augmentation de l'ADNss au cours du temps (Figure 80 B).

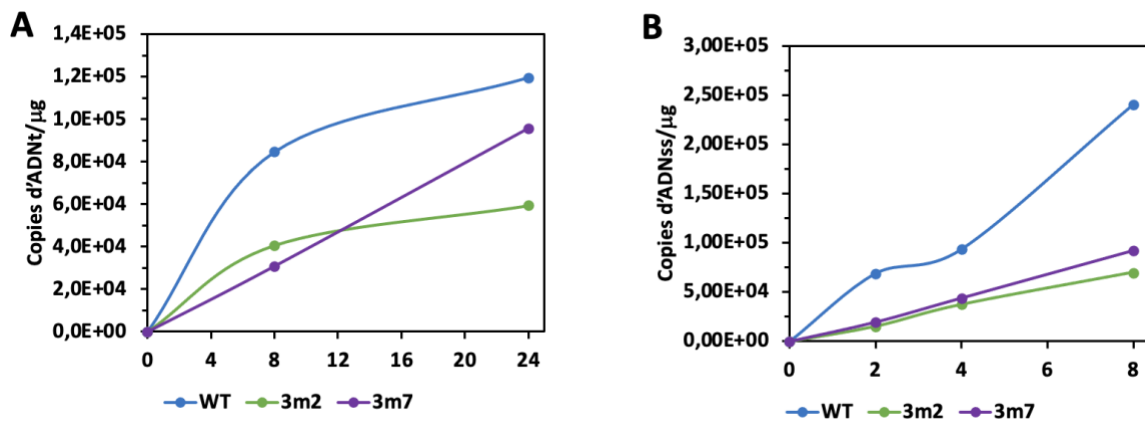


Figure 80. Expériences représentatives de la quantification par qPCR des ADNt et ADNss. (A) Quantification de l'ADNt des cellules transduites par le virus sauvage et les mutants 3m2 et 3m7. L'ADNt a été quantifié à 0, 8 h et 24 h post-transduction. (B) Quantification de l'ADNss des cellules transduites par le virus sauvage et les mutants 3m3 et 3m7. L'ADNss a été quantifié à 0, 2 h, 4 h et 8 h post-transduction. Le nombre de copies d'ADNt et ADNss est déterminé en fonction de la quantité en μg d'ADN total (ADN viral plus ADN cellulaire).

Pour le mutant 3m3, nous ne détectons pas d'ADN tardif à 8 h mais uniquement à 24 h dont la quantité est 3 fois inférieure à celle du sauvage (Figure 81 A). La diminution d'ADNt explique au moins en partie la diminution de 6,4 fois de l'expression de la GFP (Figure 78). Pour l'ADNss, une diminution de 11 fois à 2 h et de 2,5 fois à 4 h par rapport au sauvage a été mesurée (Figure 81 B). L'augmentation d'ADNss entre 2 h et 4 h est plutôt en faveur d'une synthèse diminuée de l'ADNss qu'à une dégradation de ce dernier. La diminution d'ADNt est cohérente avec la diminution d'ADNss et peut donc expliquer une diminution de la transcription inverse conduisant à moins d'ADN proviral et donc à une diminution de l'expression de la GFP. Pour le mutant 3m7, les quantités d'ADNss sont 5 et 2,5 fois inférieures à celles du sauvage pour les temps 2 h et 4 h (Figures 80 B et 81 B). Néanmoins, il présente un niveau d'ADNt

proche du sauvage à 24 h (Figures 80 A et 81 A). La faible expression de la GFP chez ce mutant pourrait s'expliquer par une intégration non efficace de l'ADN double-brin viral.

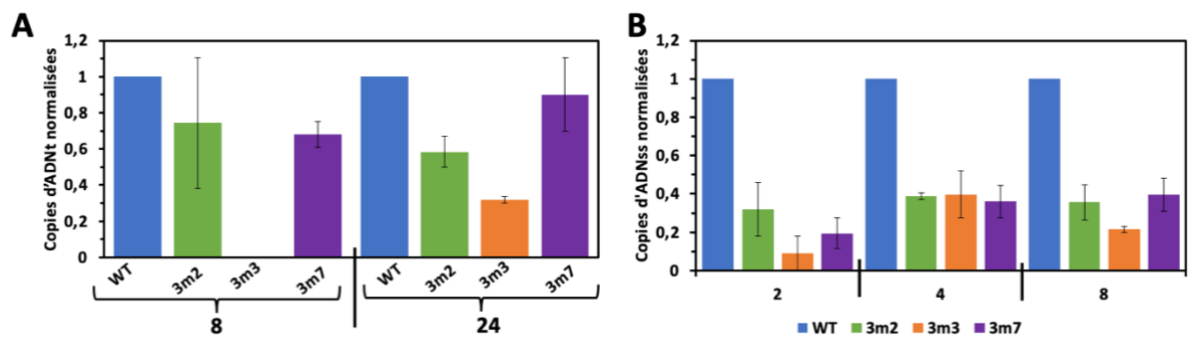


Figure 81. Copies d'ADNT et ADNss normalisées par rapport au WT. (A) Copies d'ADNT normalisées par rapport au : WT (bleu), 3m2 (vert), 3m3 (orange), 3m7 (violet) aux temps 8 h et 24 h post-transduction. (B) Copies d'ADNss normalisées par rapport au : WT (bleu), 3m2 (vert), 3m3 (orange), 3m7 (violet) aux temps 2 h, 4 h et 8 h post-transduction.

Il est intéressant de souligner que pour les trois mutants étudiés (3m2, 3m3 et 3m7), la quantité d'ADNss est significativement diminuée par rapport au sauvage. Ces résultats peuvent s'interpréter comme une dégradation de l'ADNss lorsqu'il ne peut effectuer rapidement le premier transfert de brin à cause d'une tige-boucle 3'TAR trop stable mais nous favorisons l'hypothèse d'une synthèse diminuée de l'ADNss car pour les trois mutants on observe une augmentation de l'ADNss au cours du temps.

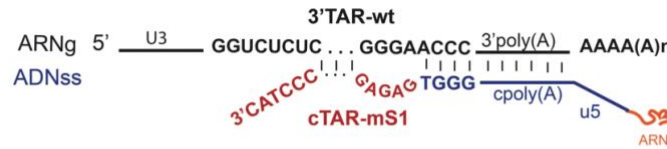
6. Analyse des effets produits par les mutations en 5'TAR et 3'TAR

6.1 But d'étudier les doubles mutants

Nos mutants ont été conçus pour ne pas pouvoir se répliquer à cause des tiges-boucles TAR ou cTAR qui ne pourraient pas s'ouvrir sous l'action de la NC et donc s'apparier pour conduire au premier transfert de brin. A l'exception des mutants mS1 et m4, les mésappariements entre l'ADNss et la région 3'R de l'ARNg introduits par les mutations ne devraient pas avoir un effet significatif sur la transcription inverse car ils n'impliquent pas la séquence 1-10 des tiges-boucles 5'TAR et 3'TAR (554) (Figures 64 et 75). Lorsque le même type de mutation en 5'TAR ou 3'TAR empêche la répllication virale, elle est testée dans les deux éléments TAR (double mutant) afin de vérifier que l'effet négatif n'est pas dû au mésappariement. En effet, la répllication virale ne doit pas être restaurée chez le double mutant si le mésappariement n'est pas la cause d'une forte réduction de la transcription inverse chez le simple mutant. Le double mutant 5'-3'TARmS1, qui nous sert de contrôle positif et a été utilisé dans une étude antérieure (554), peut se répliquer car l'appariement TAR-cTAR a été complètement restauré (Figure 82).

C'est pourquoi nous avons construit et testé le contrôle positif 5-3mS1 et quatre doubles mutants (5-3m2, 5-3m3, 5-4m4 et 5-3m6), contenant les mêmes mutations en 5'TAR (et ADNss donc) et en 3'TAR.

cTARmS1/3'TARwt : 1^{er} transfert de brin ne se produit pas



cTARwt/3'TARmS1 : 1^{er} transfert de brin ne se produit pas



cTARmS1/3'TARmS1 : 1^{er} transfert de brin se produit



Figure 82. Schématisation de l'effet des mutations mS1 sur l'appariement ADNss-3'R. Les deux premiers schémas représentent le mésappariement en 3' de l'ADN cTAR qui empêche le premier transfert de brin. Ce mésappariement est supprimé par la présence des mutations dans les deux régions 5'R et 3'R qui restaurent le premier transfert de brin.

6.2 Résultats des transfections par les mutants 5'-3'

De la même façon que précédemment, les particules virales produites par les cellules HEK293T ont été quantifiées par test ELISA p24 et les résultats sont présentés dans le Tableau 6 et représentent une moyenne d'au moins deux quantifications par type de virus. Trois productions de particules virales (mutants 5-3m2, 5-3m4 et 5-3mS1) présentent des valeurs de p24 supérieures de 20% à 51% à celle du WT. La diminution de p24 par rapport au sauvage est légère pour le 5-3m3 (31%) et assez élevée pour le 5-3m6 (70%), c'est-à-dire d'un facteur 3,3. Ces résultats suggèrent que l'expression des virus 5-3m2, 5-3m4 et 5-3mS1 dans les cellules HEK293T n'est pas altérée et très peu dans le cas du 5-3m3. En revanche, la production de particules virales par les cellules HEK293T transfectées par le mutant 5-3m6 est réduite par rapport au sauvage d'un facteur 3,3. La traduction pourrait être diminuée chez ce mutant. Les surnageants contenant les virions ont été traités à la benzonase (DNase) comme mentionné dans

la partie Matériel et Méthodes afin d'éliminer l'ADN plasmidique qui a été utilisé pour produire les particules virales et qui pourrait être amplifié lors des quantifications par qPCR des ADN viraux.

WT	5-3m2	5-3m3	5-3-m4	5-3m6	5-3mS1
2,39 ng/μl	2,88 ng/μl	1,65 ng/μl	3,24 ng/μl	0,72 ng/μl	3,62 ng/μl

Tableau 6 : Quantités de p24 mesurées pour les six productions virales.

6.3 Résultats des transductions par les mutants 5'-3'

Les transductions des cellules MT4 par les mutants 5'-3' ont été réalisées dans les mêmes conditions que précédemment. A titre d'exemple, les résultats d'une expérience obtenus par cytométrie en flux 2 jours après les transductions sont présentés dans la Figure 83. Comme attendu, on observe que par rapport aux cellules transduites par les virus 5mS1 et 3mS1, le nombre de cellules fluorescentes augmente lorsque les cellules sont transduites par le mutant 5-3mS1. En revanche, on observe que par rapport aux cellules transduites par le mutant 3m2, le nombre de cellules fluorescentes n'augmente pas lorsque les cellules sont transduites par le mutant 5-3m2. Nous n'observons pas non plus une augmentation du nombre de cellules fluorescentes lorsque les cellules sont transduites par les doubles mutants 5-3m3 et 5-3m6 par rapport à celles transduites par les simples mutants m3 et m6. En revanche, le nombre de cellules fluorescentes augmente légèrement lorsque les cellules sont transduites par le double mutant 5-3m4.

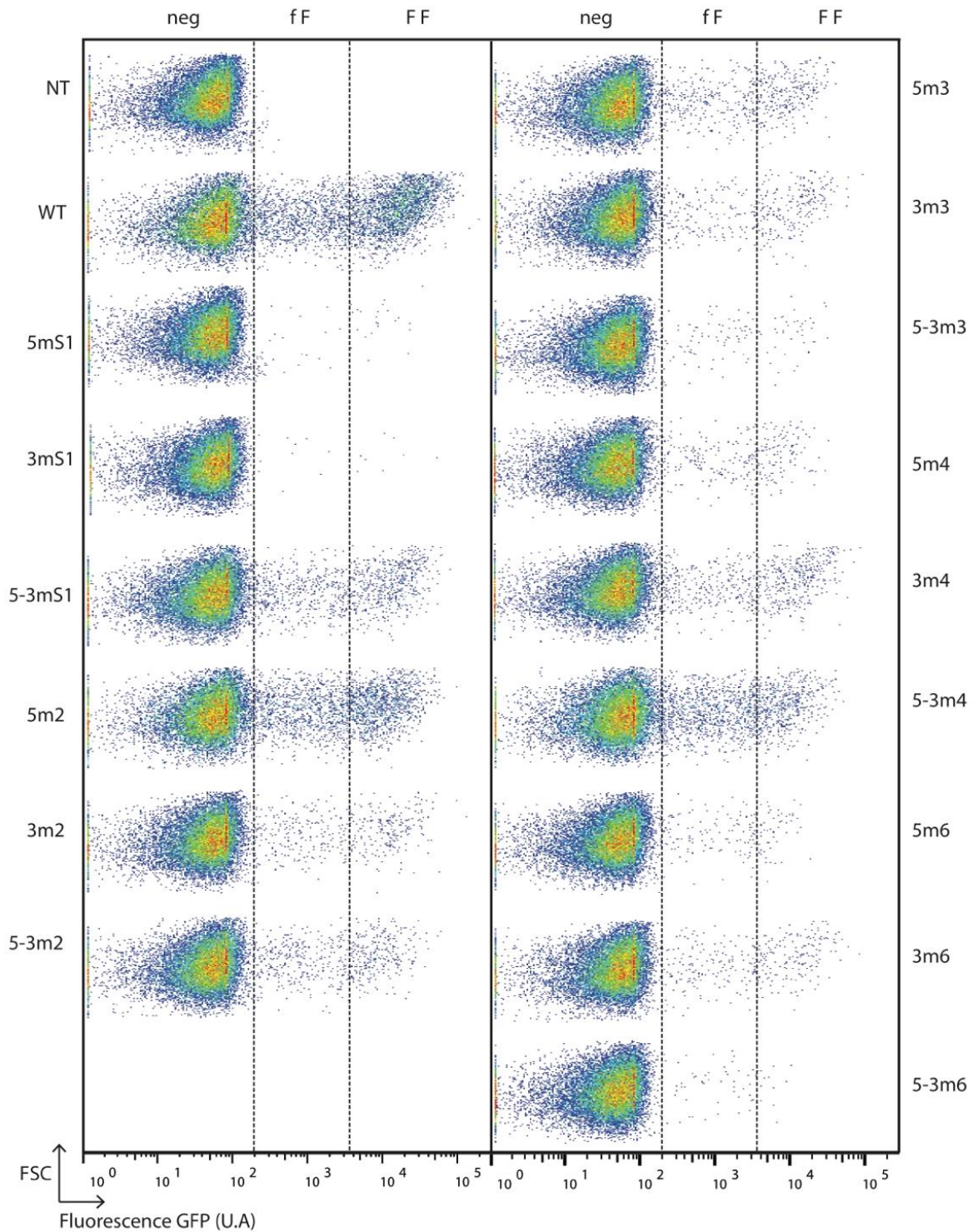


Figure 83. Fluorescence des cellules MT4 déterminée par cytométrie en flux deux jours après les transductions. Trois niveaux d'intensité de fluorescence sont détectés, un négatif (neg), un de faible intensité (f F) et un de forte intensité (F F). Le f F correspond aux formes ADN viraux circulaires à 1 et 2-LTR exprimant faiblement la GFP. Le F F correspond à une forte expression de la GFP à partir de l'ADN proviral (171, 176, 178). FSC : *Foward scatter* (paramètre mesuré permettant de discriminer les cellules par leur taille). NT, cellules MT4 non transduites ; WT, cellules MT4 transduites par le virus sauvage. 5mS1 à 5-3m6, cellules MT4 transduites par les virus simples et doubles mutants.

Le pourcentage de cellules fluorescentes transduites par le double mutant 5-3mS1 est respectivement augmenté de 32 et 49 fois par rapport aux cellules fluorescentes transduites par les simples mutants 5mS1 et 3mS1 (Figure 84). Nous retrouvons bien les résultats de Ohi et Clever qui ont montré que l'infectivité était restaurée chez le double mutant 5-3mS1 et que la très faible infectivité des mutants 5mS1 et 3mS1 résultait d'une quasi-incapacité à réaliser le premier transfert de brin (554). Le mésappariement a été supprimé par la présence des mutations au niveau des séquences 5'TAR et 3'TAR (Figure 81).

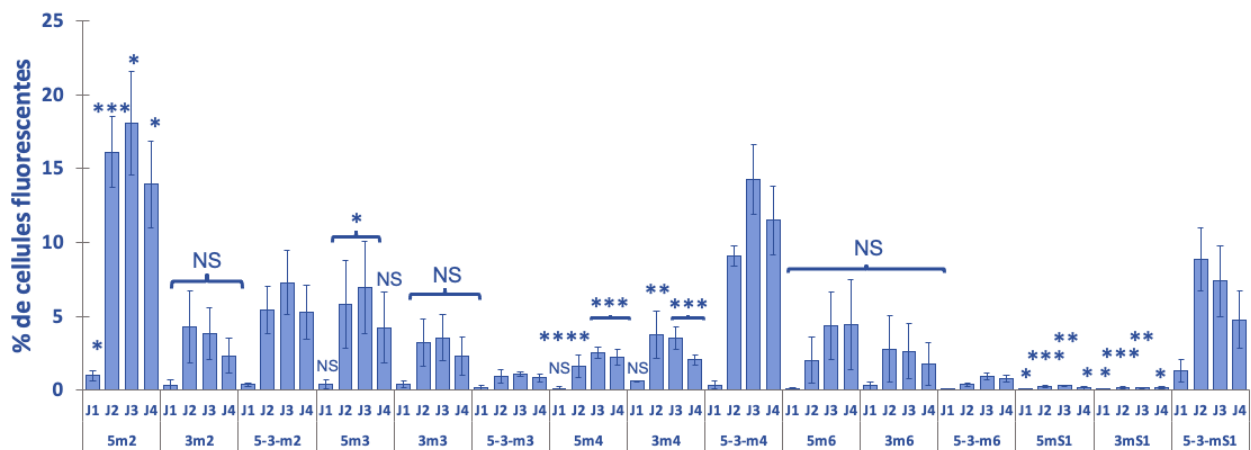


Figure 84. Pourcentage de cellules fluorescentes pour les simples mutants et les doubles mutants un à quatre jours après les transductions. Les étoiles représentent le p du test student avec * : $p < 0,05$; ** : $p < 0,01$; *** : $p < 0,005$; **** : $p < 0,001$ entre les simples mutants et leur double mutant, 5m2 par rapport au 5-3m2 par exemple.

Le mutant 5m2 n'a pas d'effet sur l'expression de la GFP contrairement au mutant 3m2 (Figures 67 et 77). Pour rappel, à l'exception des mutants m4, les mutations testées ne sont pas prévues produire des mésappariements qui auraient un effet négatif sur le premier transfert de brin car elles ne sont pas dans la région 1-10 des séquences 5' et 3'TAR (554). L'expression de la GFP est également faible chez les cellules transduites par le double mutant 5-3m2 (Figure 83). Ces résultats montrent que ce n'est pas un mésappariement entre la tige-boucle 3'TARm2 et la tige-boucle cTAR de type sauvage qui est responsable de la faible expression de la GFP chez le mutant 3m2. Ils suggèrent aussi que c'est la stabilité de la tige-boucle 3'TARm2 qui est responsable du phénotype du mutant 3m2. L'expression de la GFP est également faible chez les cellules transduites par les doubles mutants 5-3m3 et 5-3m6 (Figure 84). Elle est même diminuée d'un facteur 3 à 7 par rapport aux simples mutants, ce qui suggère un effet additif des mutations en 5' et 3'. Par conséquent, l'effet négatif des mutations m3 et m6 chez les simples mutants n'est pas dû comme attendu aux mésappariements qu'elles génèrent. Les observations

pour les mutants m4 sont différentes. L'expression de la GFP est faible chez les cellules transduites par les simples mutants tandis qu'elle est significativement restaurée chez le double mutant. Ce résultat n'est pas inattendu car une mutation (remplacement d'une uridine par une cytosine) est située dans la région 1-10 des séquences 5' et 3'TAR (Figures 64 et 75). Toutefois, il montre qu'un seul mésappariement suffit pour avoir un effet significatif sur l'expression de la GFP et donc probablement sur le premier transfert de brin. En effet, Ohi et Clever n'avaient pas étudié un seul mésappariement mais au minimum trois (554).

6.4 Résultats encapsidation ARNg des mutants 5'-3'

La quantification a été réalisée au moins deux fois pour chaque type de mutant. Les ratios ARNg/p24 traduisent la quantité d'ARNg par rapport à la quantité de p24 mesurée en test ELISA qui reflète le nombre de particules virales utilisées. Les résultats ont été normalisés par rapport au sauvage (Figure 85).

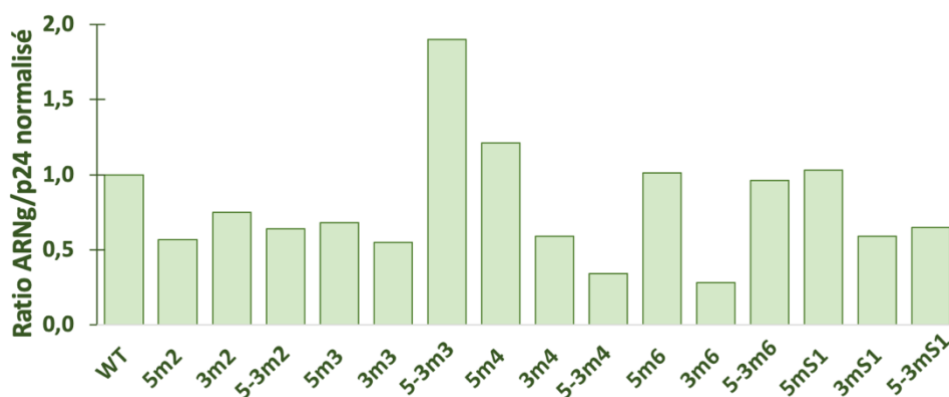


Figure 85. Histogrammes représentant les rapports ARNg/p24 normalisé par rapport au sauvage. Les ratios sont issus d'au moins deux quantifications d'ARNg et deux quantifications p24. Les quantités d'ARNg sont cohérentes avec les quantités de p24 mesurées pour les virus mutants par rapport au virus sauvage.

L'encapsidation de l'ARNg pour les doubles mutants m2 et mS1 n'est pas significativement augmentée ou diminuée par rapport à au moins l'un des simples mutants, c'est-à-dire 5m2 et 3m2 par rapport au 5-3m2 et 5mS1 et 3mS1 par rapport au 5-3mS1. En revanche et de façon surprenante, l'encapsidation de l'ARNg pour le double mutant 5-3m3 est fortement augmentée par rapport aux simples mutants et est presque le double de ce qui est mesuré pour le virus sauvage. Nous n'avons pas d'explications logiques pour expliquer ce résultat. Toutefois, cela montre que la faible expression de la GFP par les cellules transduites par le double mutant 5-3m3 n'est pas due à une diminution de l'encapsidation. Nous observons aussi que l'encapsidation de l'ARNg pour le double mutant 5-3m4 est diminuée de 42% à 72% par

rapport aux simples mutants. L'expression de la GFP par les cellules transduites par les simples mutants m4 est cependant significativement plus faible que celles transduites par le double mutant 5-3m4 (Figure 84). Ceci confirme que le phénotype du mutant 3m4 n'est pas dû à une diminution de 41% de l'encapsidation de l'ARNg par rapport au sauvage.

6.5 qPCR ADN viral

Les quantifications de l'ADNt ont été réalisées dans les mêmes conditions de transduction que pour les résultats de fluorescence et uniquement pour les mutants 5-3m3 et 5-3mS1. Les résultats des quantifications d'ADNt sont présentés dans la Figure 86. Cette dernière présente les ratios par rapport au sauvage des ADNt entre les simples et doubles mutants. Comme cela a été mentionné précédemment, les cellules transduites par le mutant contrôle 5mS1 n'expriment quasiment pas la GFP car elles contiennent très peu d'ADN proviral à cause d'une très faible quantité d'ADNt. L'expression de la GFP qui est augmentée de 32 fois chez le double mutant 5-3mS1 par rapport au simple mutant 5mS1 (Figure 84) se traduit par une augmentation d'ADNt de 120 fois à 8 h et 600 fois à 24 h post-transduction. La transcription inverse est donc restaurée chez le double mutant mS1 même si celle-ci n'est pas aussi efficace que chez le sauvage. Cette restauration avait aussi été observée par Ohi et Clever (554).

Pour les mutants m3, nous observons des quantités similaires entre les simples et le double mutant. La présence de la mutation dans les deux régions 5'TAR et 3'TAR ne permet pas à la transcription inverse d'être restaurée et conforte les résultats précédemment obtenus en fluorescence. Les mutations de type m3 ont bien un effet sur la transcription inverse et ne sont pas dues à des mésappariements.

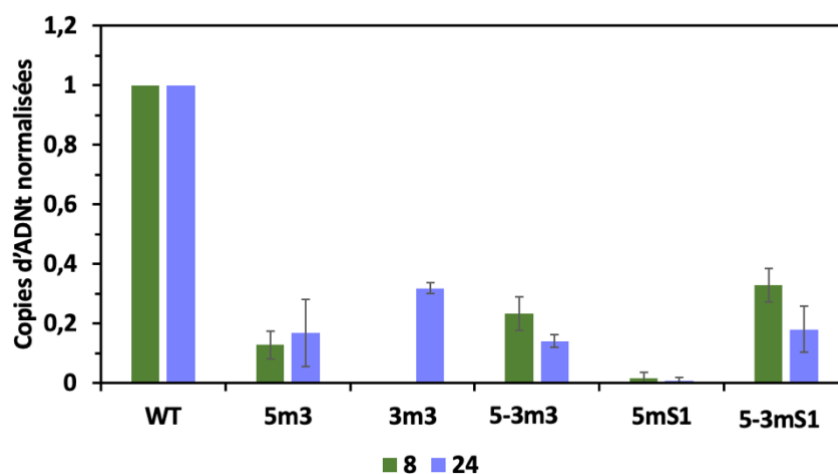


Figure 86. Ratio des quantifications de l'ADNt des simples et doubles mutants mS1 et m3 à 8 h et 24 h post-transduction.

CONCLUSIONS ET PERSPECTIVES

À notre connaissance, notre étude est la première à avoir pour objectif de déterminer si *ex vivo* le premier transfert de brin chez le VIH-1 dépend principalement de la déstabilisation par la NC des tiges-boucles ADN cTAR et ARN TAR. Pour atteindre cet objectif, nous avons conçu et analysé une batterie de virus mutants pour leurs capacités à se répliquer en suivant l'expression de la GFP et à effectuer la transcription inverse en mesurant par qPCR des intermédiaires de ce processus. Les mutations choisies ont été introduites dans les séquences 5'TAR et 3'TAR d'un vecteur rétroviral contenant l'intégralité du génome du VIH-1 (isolat NL4-3) mais déficient au niveau de l'expression du gène *env* et possédant le gène de la GFP.

Les mutations dans la séquence 5'TAR ont pour but de stabiliser les tiges-boucles 5'TAR et cTAR et donc de diminuer ou d'empêcher leur déstabilisation par la NC. D'après une étude *in vitro* récente de notre équipe (547), la NC ne devrait pas être capable de déstabiliser les tiges-boucles 5'TARm1, 5'TARm2, 5'TARm3 et 5'TARm6 car elles contiennent de quatorze à vingt-et-une paires de bases consécutives (Figure 64). De plus, trois des quatre guanines reconnues par la NC ne sont plus à proximité d'un bulge ou d'une paire U-G dans la tige-boucle 5'TARm2 (Figures 50 et 64). Si la tige-boucle 5'TAR ne peut pas être ouverte, elle ne peut pas être copiée par la RT, ce qui se traduit par l'obtention d'un ADNss tronqué conduisant à une transcription inverse diminuée car il est peu efficace pour le premier transfert de brin (502, 554).

Nous avons montré que l'expression de la GFP n'est pas significativement diminuée dans les cellules transduites par les mutants 5m1 et 5m2 tandis qu'elle l'est dans les cellules transduites par les mutants 5m3, 5m4 et 5m6. La restauration de l'expression de la GFP dans les cellules transduites par le double mutant 5-3m4 suggère fortement que la diminution observée avec le mutant 5m4 est due à une diminution de la transcription inverse provoquée par un mésappariement entre l'extrémité 3' de l'ADN cTARm4 et l'extrémité 5' de l'ARN 3'TAR sauvage. Ce résultat est attendu car le mésappariement est dans la région 1-10 de la séquence TAR qui joue un rôle essentiel dans le premier transfert de brin (554). Les résultats d'expression de la GFP obtenus avec les mutants 5m1 et 5m2 suggèrent fortement que la transcription inverse et par conséquent la synthèse de l'ADNss chez ces deux mutants ne sont pas diminuées. L'analyse par qPCR de l'ADN viral tardif (ADNt) dans les cellules transduites par le mutant 5m2 confirme que celles-ci en contiennent autant que les cellules transduites par le virus

sauvage. La transcription inverse n'est donc pas altérée bien que la NC ne devrait pas pouvoir déstabiliser la tige-boucle cTARm2 et permettre le premier transfert de brin (voir partie 4.1). Les résultats obtenus avec le 5m2 suggèrent donc que la NC n'est pas nécessaire à la déstabilisation de la tige-boucle cTARm2 qui possède treize paires de bases Watson-Crick consécutives et un site fort NC dans une configuration non-optimale pour interagir avec la protéine. Une autre protéine telle qu'une hélicase pourrait être impliquée dans la déstabilisation de la tige-boucle cTARm2. Mon équipe d'accueil vérifiera cependant que la NC n'est pas capable *in vitro* d'apparier la tige-boucle ADN cTARm2 avec la tige-boucle ARN 3'TAR. L'analyse par qPCR de l'ADN viral strong-stop (ADNss) dans les cellules transduites par le mutant 5m6 montre que celles-ci en contiennent autant que les cellules transduites par le virus sauvage 2 h après la transduction (Figure 73 B). Les résultats obtenus avec les mutants 5m1, 5m2 et 5m6 suggèrent que la NC n'est pas nécessaire à l'ouverture de la tige-boucle 5'TAR lorsque celle-ci est copiée par la RT pour générer l'extrémité 3' de l'ADNss. L'ARN hélicase A humaine, qui agit sur la processivité de la RT durant la phase d'élongation (361), est un bon candidat pour l'ouverture de la tige-boucle 5'TAR lors de la synthèse de l'ADNss. Mon équipe d'accueil vérifiera que la NC ne peut pas faciliter *in vitro* la copie par la RT des tiges-boucles ARN TARm2 et TARm6.

Nous avons montré que par rapport aux cellules transduites par le virus sauvage, l'expression de la GFP est significativement diminuée dans les cellules transduites par le mutant 5m3 et est associée à une diminution des quantités d'ADNt et d'ADNss (Figure 73). La diminution de la quantité d'ADNss observée 2 h après la transduction (Figure 73 B) pourrait être due à une diminution de la synthèse d'ADNss causée par la stabilité de la tige-boucle 5'TARm3. Toutefois, cette hypothèse n'est pas étayée par les résultats obtenus avec les mutants 5m2 et 5m6. En effet, il ne nous semble pas logique que la RT puisse copier les tiges-boucles 5'TARm2 et 5'TARm6 possédant respectivement seize et vingt-et-une paires de bases consécutives mais pas la tige-boucle 5'TARm3 qui possède quatorze paires de bases consécutives (Figure 64). Une dégradation rapide de l'ADNss du mutant 5m3 pourrait être déclenchée par la structure de la tige-boucle ADN cTARm3. Nos résultats suggèrent aussi une dégradation de l'ADNss du mutant 5m6 qui serait déclenchée par la structure de la tige-boucle ADN cTARm6. Il est intéressant de souligner que la partie inférieure des tiges-boucles cTARm3 et cTARm6 est appariée et non simple-brin comme chez le sauvage. L'extrémité double-brin de la tige-boucle cTAR pourrait servir de signal à la dégradation de l'ADNss.

Les mutations présentes dans les mutants 3m1 à 3m6 (Figure 75) ont pour but de stabiliser la tige-boucle ARN 3'TAR afin d'empêcher son appariement avec la tige-boucle cTAR de l'ADNss. Les mutations m7, m3G et m4G ciblent le *bulge* et les nucléotides ayant été rapportés comme important pour l'activité *in vitro* de la NC dans la déstabilisation de la tige-boucle 3'TAR (303, 304). Pour tous les mutants, l'expression de la GFP est significativement diminuée, ce qui suggère une transcription inverse altérée. Dans le cas du mutant 3m6, cette diminution est au moins due en partie à un défaut d'encapsidation de l'ARNg. Il pourrait être intéressant de comprendre pourquoi la tige-boucle 3'TARm6 a un effet négatif sur l'encapsidation bien qu'elle soit éloignée du signal d'encapsidation qui est situé dans la partie 5' de l'ARNg. Pour les mutants 3m2, 3m3 et 3m7, nous avons mesuré des quantités d'ADNss diminuées d'environ deux fois par rapport au sauvage. L'hypothèse d'une dégradation de l'ADNss nous semble peu probable car la quantité d'ADNss ne diminue pas au cours du temps (Figure 81). Nous favorisons l'hypothèse d'une diminution de la synthèse de l'ADNss chez ces trois mutants bien que cela soit surprenant car la copie par la RT de la tige-boucle 5'TAR ne devrait pas dépendre de la tige-boucle 3'TAR qui est à l'autre extrémité du génome. Une circularisation de l'ARNg favorisant la synthèse de l'ADNss et dépendant de plusieurs interactions intramoléculaires dont l'interaction 5'TAR - 3'TAR (565) pourrait expliquer l'effet à distance de mutations dans la tige-boucle 3'TAR. Des expériences *in vitro* sont envisagées par mon équipe d'accueil pour vérifier que la NC ne facilite pas l'appariement de l'ADN cTAR sauvage avec les tiges-boucles ARN 3'TARm2, 3'TARm3 et 3'TARm7.

Le mutant 3m7 montre une diminution de la production d'ADNss mais un niveau d'ADNt similaire à celui du sauvage. Nous n'avons pas une explication logique pour ce résultat. L'ADNt du mutant 3m7 conduit à peu d'ADN viral intégré car l'expression de la GFP est significativement diminuée. Ce résultat est surprenant car des mutations dans la séquence 3'TAR ne devraient pas avoir un effet sur le processus d'intégration. En effet, à la fin de la transcription inverse de l'ARNg du mutant 3m7, l'ADN viral double-brin produit devrait contenir deux LTR de type sauvage. Une quantification par qPCR des formes à 1 et 2-LTR pourrait nous apporter des informations sur une altération au niveau de l'intégration.

L'étude des cellules transduites par les doubles mutants 5-3m2, 5-3m3 et 5-3m6 montre que l'expression de la GFP n'est pas augmentée par rapport aux cellules transduites par les simples mutants m2, m3 et m6. La quantification de l'ADNt dans les cellules transduites par le double mutant 5-3m3 confirme que la transcription inverse n'est pas augmentée chez ce dernier. Ces résultats suggèrent fortement que les mésappariements entre l'ADN cTAR et l'ARN 3'TAR ne

sont pas responsables du phénotype des simples mutants m2, m3 et m6. Ces résultats étaient attendus car les mutations n'affectent pas les nucléotides 1 à 10 de la séquence TAR.

Finalement, l'étude avec le dernier couple d'amorce permettant de quantifier par qPCR l'ADN viral généré par le premier transfert de brin n'a pas été réalisée mais devra l'être. Une optimisation de cette qPCR sera nécessaire pour nous apporter des informations complémentaires sur les mutants ayant un effet sur la synthèse de l'ADNt.

Nos travaux contribuent à mieux comprendre le processus de transcription inverse et plus spécifiquement deux étapes qui sont la synthèse de l'ADNss et le premier transfert de brin. Toutefois, des études complémentaires sont nécessaires pour déterminer au niveau moléculaire le rôle de l'élément TAR dupliqué dans ces deux étapes.

REFERENCES BIBLIOGRAPHIQUES

1. Varmus, H. (1988) Retroviruses. *Science*. **240**, 1427–1435
2. Greenwood, A. D., O'Brien, S. P., Roca, A. L., and Eiden, M. V. (2017) Transmission, Evolution, and Endogenization: Lessons Learned from Recent Retroviral Invasions. *Microbiol. Mol. Biol. Rev. MMBR*. 10.1128/MMBR.00044-17
3. Murphy, F. A., Fauquet, C. M., Bishop, D. H. L., Ghabrial, S. A., Jarvis, A. W., Martelli, G. P., Mayo, M. A., and Summers, M. D. (2012) *Virus Taxonomy: Classification and Nomenclature of Viruses*, Springer Science & Business Media
4. Krupovic, M., Blomberg, J., Coffin, J. M., Dasgupta, I., Fan, H., Geering, A. D., Gifford, R., Harrach, B., Hull, R., Johnson, W., Kreuze, J. F., Lindemann, D., Llorens, C., Lockhart, B., Mayer, J., Muller, E., Olszewski, N. E., Pappu, H. R., Pooggin, M. M., Richert-Pöggeler, K. R., Sabanadzovic, S., Sanfaçon, H., Schoelz, J. E., Seal, S., Stavelone, L., Stoye, J. P., Teycheney, P.-Y., Tristem, M., Koonin, E. V., and Kuhn, J. H. (2018) Ortervirales: New Virus Order Unifying Five Families of Reverse-Transcribing Viruses. *J. Virol*. 10.1128/JVI.00515-18
5. Infection par le VIH et travail. 25 ans après le début de la pandémie - Article de revue - INRS [online] <http://www.inrs.fr/media.html?refINRS=TC%20103> (Accessed April 14, 2020)
6. Barré-Sinoussi, F., Chermann, J. C., Rey, F., Nugeyre, M. T., Chamaret, S., Gruest, J., Dautet, C., Axler-Blin, C., Vézinet-Brun, F., Rouzioux, C., Rozenbaum, W., and Montagnier, L. (1983) Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science*. **220**, 868–871
7. Popovic, M., Sarngadharan, M. G., Read, E., and Gallo, R. C. (1984) Detection, isolation, and continuous production of cytopathic retroviruses (HTLV-III) from patients with AIDS and pre-AIDS. *Science*. **224**, 497–500
8. Levy, J. A., Hoffman, A. D., Kramer, S. M., Landis, J. A., Shimabukuro, J. M., and Oshiro, L. S. (1984) Isolation of lymphocytopathic retroviruses from San Francisco patients with AIDS. *Science*. **225**, 840–842
9. Clavel, F., Guétard, D., Brun-Vézinet, F., Chamaret, S., Rey, M. A., Santos-Ferreira, M. O., Laurent, A. G., Dautet, C., Katlama, C., and Rouzioux, C. (1986) Isolation of a new human retrovirus from West African patients with AIDS. *Science*. **233**, 343–346
10. Chakrabarti, L., Guyader, M., Alizon, M., Daniel, M. D., Desrosiers, R. C., Tiollais, P., and Sonigo, P. (1987) Sequence of simian immunodeficiency virus from macaque and its relationship to other human and simian retroviruses. *Nature*. **328**, 543–547
11. Sharp, P. M., and Hahn, B. H. (2011) Origins of HIV and the AIDS Pandemic. *Cold Spring Harb. Perspect. Med*. 10.1101/cshperspect.a006841
12. Gao, F., Bailes, E., Robertson, D. L., Chen, Y., Rodenburg, C. M., Michael, S. F., Cummins, L. B., Arthur, L. O., Peeters, M., Shaw, G. M., Sharp, P. M., and Hahn, B. H. (1999) Origin of HIV-1 in the chimpanzee *Pan troglodytes*. *Nature*. **397**, 436–441
13. Huet, T., Cheynier, R., Meyerhans, A., Roelants, G., and Wain-Hobson, S. (1990) Genetic organization of a chimpanzee lentivirus related to HIV-1. *Nature*. **345**, 356–359
14. Hirsch, V. M., Olmsted, R. A., Murphey-Corb, M., Purcell, R. H., and Johnson, P. R. (1989) An African primate lentivirus (SIVsm) closely related to HIV-2. *Nature*. **339**, 389–392
15. Hahn, B. H., Shaw, G. M., De Cock, K. M., and Sharp, P. M. (2000) AIDS as a zoonosis: scientific and public health implications. *Science*. **287**, 607–614
16. HIV Circulating Recombinant Forms (CRFs) [online] <https://www.hiv.lanl.gov/content/sequence/HIV/CRFs/CRFs.html> (Accessed April 14, 2020)
17. De Leys, R., Vanderborght, B., Vanden Haesevelde, M., 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
18. Simon, F., Maucière, P., Roques, P., Loussert-Ajaka, I., Müller-Trutwin, M. C., Saragosti, S., Georges-Courbot, M. C., Barré-Sinoussi, F., and Brun-Vézinet, F. (1998) Identification of a new human immunodeficiency virus type 1 distinct from group M and group O. *Nat. Med.* **4**, 1032–1037
 19. 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**, 871–872
 20. Vallari, A., Holzmayer, V., Harris, B., Yamaguchi, J., Ngansop, C., Makamche, F., Mbanya, D., Kaptué, L., Ndembi, N., Gürtler, L., Devare, S., and Brennan, C. A. (2011) Confirmation of putative HIV-1 group P in Cameroon. *J. Virol.* **85**, 1403–1407
 21. Delauger, C., De Oliveira, F., Lascoux-Combe, C., Plantier, J.-C., and Simon, F. (2011) HIV-1 group N: travelling beyond Cameroon. *Lancet Lond. Engl.* **378**, 1894
 22. 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., Gilbert, M. T. P., and Wolinsky, S. M. (2008) Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960. *Nature.* **455**, 661–664
 23. Hué, S., Pillay, D., Clewley, J. P., and Pybus, O. G. (2005) Genetic analysis reveals the complex structure of HIV-1 transmission within defined risk groups. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 4425–4429
 24. 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
 25. Van Heuverswyn, F., Li, Y., Neel, C., Bailes, E., Keele, B. F., Liu, W., Loul, S., Butel, C., Liegeois, F., Bienvenue, Y., Ngolle, E. M., Sharp, P. M., Shaw, G. M., Delaporte, E., Hahn, B. H., and Peeters, M. (2006) Human immunodeficiency viruses: SIV infection in wild gorillas. *Nature.* **444**, 164
 26. Vallari, A., Holzmayer, V., Harris, B., Yamaguchi, J., Ngansop, C., Makamche, F., Mbanya, D., Kaptué, L., Ndembi, N., Gürtler, L., Devare, S., and Brennan, C. A. (2011) Confirmation of putative HIV-1 group P in Cameroon. *J. Virol.* **85**, 1403–1407
 27. Kannangai, R., David, S., and Sridharan, G. (2012) Human immunodeficiency virus type-2-A milder, kinder virus: an update. *Indian J. Med. Microbiol.* **30**, 6–15
 28. de Silva, T. I., Cotten, M., and Rowland-Jones, S. L. (2008) HIV-2: the forgotten AIDS virus. *Trends Microbiol.* **16**, 588–595
 29. Olesen, J. S., Jespersen, S., da Silva, Z. J., Rodrigues, A., Erikstrup, C., Aaby, P., Wejse, C., and Hønge, B. L. (2018) HIV-2 continues to decrease, whereas HIV-1 is stabilizing in Guinea-Bissau. *AIDS Lond. Engl.* **32**, 1193–1198
 30. Gottlieb, G. S., Raugi, D. N., and Smith, R. A. (2018) 90-90-90 for HIV-2? Ending the HIV-2 epidemic by enhancing care and clinical management of patients infected with HIV-2. *Lancet HIV.* **5**, e390–e399
 31. MacNeil, A., Sarr, A. D., Sankalé, J.-L., Meloni, S. T., Mboup, S., and Kanki, P. (2007) Direct Evidence of Lower Viral Replication Rates In Vivo in Human Immunodeficiency Virus Type 2 (HIV-2) Infection than in HIV-1 Infection. *J. Virol.* **81**, 5325–5330
 32. Chen, Z., Telfier, P., Gettie, A., Reed, P., Zhang, L., Ho, D. D., and Marx, P. A. (1996) Genetic characterization of new West African simian immunodeficiency virus SIVsm: geographic clustering of household-derived SIV strains with human immunodeficiency virus type 2 subtypes and genetically diverse viruses from a single feral sooty mangabey troop. *J. Virol.* **70**, 3617–3627
 33. Shaw, G. M., and Hunter, E. (2012) HIV transmission. *Cold Spring Harb. Perspect.*

Med. **2**, a006965

34. Grossman, Z., Meier-Schellersheim, M., Paul, W. E., and Picker, L. J. (2006) Pathogenesis of HIV infection: what the virus spares is as important as what it destroys. *Nat. Med.* **12**, 289–295
35. Cohen, M. S., Shaw, G. M., McMichael, A. J., and Haynes, B. F. (2011) Acute HIV-1 Infection. *N. Engl. J. Med.* **364**, 1943–1954
36. Lambotte, O., Boufassa, F., Madec, Y., Nguyen, A., Goujard, C., Meyer, L., Rouzioux, C., Venet, A., Delfraissy, J.-F., and Group, T. 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
37. 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
38. Ganser-Pornillos, B. K., Yeager, M., and Pornillos, O. (2012) Assembly and Architecture of HIV. *Adv. Exp. Med. Biol.* **726**, 441–465
39. Powell, C. (2019) *Fungi, Bacteria and Viruses*, Scientific e-Resources
40. van Dijk, D., Ertaylan, G., Boucher, C. A., and Sloot, P. M. (2010) Identifying potential survival strategies of HIV-1 through virus-host protein interaction networks. *BMC Syst. Biol.* **4**, 96
41. Miller, R. H. (1988) Human immunodeficiency virus may encode a novel protein on the genomic DNA plus strand. *Science.* **239**, 1420–1422
42. Bet, A., Maze, E. A., Bansal, A., Sterrett, S., Gross, A., Graff-Dubois, S., Samri, A., Guihot, A., Katlama, C., Theodorou, I., Mesnard, J.-M., Moris, A., Goepfert, P. A., and Cardinaud, S. (2015) The HIV-1 Antisense Protein (ASP) induces CD8 T cell responses during chronic infection. *Retrovirology.* **12**, 15
43. Berger, C. T., Llano, A., Carlson, J. M., Brumme, Z. L., Brockman, M. A., Cedeño, S., Harrigan, P. R., Kaufmann, D. E., Heckerman, D., Meyerhans, A., and Brander, C. (2015) Immune Screening Identifies Novel T Cell Targets Encoded by Antisense Reading Frames of HIV-1. *J. Virol.* **89**, 4015–4019
44. Toyoda, K., and Matsuoka, M. (2022) Functional and Pathogenic Roles of Retroviral Antisense Transcripts. *Front. Immunol.* **13**, 875211
45. Cassan, E., Arigon-Chifolleau, A.-M., Mesnard, J.-M., Gross, A., and Gascuel, O. (2016) Concomitant emergence of the antisense protein gene of HIV-1 and of the pandemic. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 11537–11542
46. Pavesi, A., and Romerio, F. (2022) Extending the Coding Potential of Viral Genomes with Overlapping Antisense ORFs: A Case for the De Novo Creation of the Gene Encoding the Antisense Protein ASP of HIV-1. *Viruses.* **14**, 146
47. Langer, S., and Sauter, D. (2016) Unusual Fusion Proteins of HIV-1. *Front. Microbiol.* **7**, 2152
48. Burnett, J. C., Miller-Jensen, K., Shah, P. S., Arkin, A. P., and Schaffer, D. V. (2009) Control of Stochastic Gene Expression by Host Factors at the HIV Promoter. *PLoS Pathog.* [10.1371/journal.ppat.1000260](https://doi.org/10.1371/journal.ppat.1000260)
49. Garcia, J. A., Harrich, D., Soultanakis, E., Wu, F., Mitsuyasu, R., and Gaynor, R. B. (1989) Human immunodeficiency virus type 1 LTR TATA and TAR region sequences required for transcriptional regulation. *EMBO J.* **8**, 765–778
50. Guntaka, R. V. (1993) Transcription termination and polyadenylation in retroviruses. *Microbiol. Rev.* **57**, 511–521
51. Gaynor, R. (1992) Cellular transcription factors involved in the regulation of HIV-1 gene expression. *AIDS Lond. Engl.* **6**, 347–363
52. Mbondji-wonje, C., Dong, M., Wang, X., Zhao, J., Ragupathy, V., Sanchez, A. M.,

- Denny, T. N., and Hewlett, I. (2018) Distinctive variation in the U3R region of the 5' Long Terminal Repeat from diverse HIV-1 strains. *PLoS ONE*. 10.1371/journal.pone.0195661
53. Bannwarth, S., and Gatignol, A. (2005) HIV-1 TAR RNA: the target of molecular interactions between the virus and its host. *Curr. HIV Res.* **3**, 61–71
54. Schulze-Gahmen, U., Echeverria, I., Stjepanovic, G., Bai, Y., Lu, H., Schneidman-Duhovny, D., Doudna, J. A., Zhou, Q., Sali, A., and Hurley, J. H. (2016) Insights into HIV-1 proviral transcription from integrative structure and dynamics of the Tat:AFF4:P-TEFb:TAR complex. *eLife*. **5**, e15910
55. Schulze-Gahmen, U., and Hurley, J. H. (2018) Structural mechanism for HIV-1 TAR loop recognition by Tat and the super elongation complex. *Proc. Natl. Acad. Sci.* **115**, 12973–12978
56. Dorin, D., Bonnet, M. C., Bannwarth, S., Gatignol, A., Meurs, E. F., and Vaquero, C. (2003) The TAR RNA-binding protein, TRBP, stimulates the expression of TAR-containing RNAs in vitro and in vivo independently of its ability to inhibit the dsRNA-dependent kinase PKR. *J. Biol. Chem.* **278**, 4440–4448
57. Fernandes, J., Jayaraman, B., and Frankel, A. (2012) The HIV-1 Rev response element. *RNA Biol.* **9**, 6–11
58. Malim, M. H., Hauber, J., Le, S.-Y., Maizel, J. V., and Cullen, B. R. (1989) The HIV-1 rev trans-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. *Nature*. **338**, 254–257
59. Rausch, J. W., and Le Grice, S. F. J. (2015) HIV Rev Assembly on the Rev Response Element (RRE): A Structural Perspective. *Viruses*. **7**, 3053–3075
60. Jacks, T., Power, M. D., Masiarz, F. R., Luciw, P. A., Barr, P. J., and Varmus, H. E. (1988) Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. *Nature*. **331**, 280–283
61. Dulude, D., Baril, M., and Brakier-Gingras, L. (2002) Characterization of the frameshift stimulatory signal controlling a programmed –1 ribosomal frameshift in the human immunodeficiency virus type 1. *Nucleic Acids Res.* **30**, 5094–5102
62. Foley, B., Leitner, T., Apetrei, C., Hahn, B., Mizrahi, I., Mullins, J., Rambaut, A., Wolinsky, S., Korber, B., Singh, A., Abfalterer, W., Giori, E. E., Fischer, W., Hraber, P., Macke, J., Szinger, J. J., Wagh, K., and Yoon, H. (2018) HIV Sequence Compendium 2018
63. Biswas, P., Jiang, X., Pacchia, A. L., Dougherty, J. P., and Peltz, S. W. (2004) The Human Immunodeficiency Virus Type 1 Ribosomal Frameshifting Site Is an Invariant Sequence Determinant and an Important Target for Antiviral Therapy. *J. Virol.* **78**, 2082–2087
64. Shehu-Xhilaga, M., Crowe, S. M., and Mak, J. (2001) Maintenance of the Gag/Gag-Pol ratio is important for human immunodeficiency virus type 1 RNA dimerization and viral infectivity. *J. Virol.* **75**, 1834–1841
65. 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
66. 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. U. S. A.* **91**, 4945–4949
67. 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**, 1–13
68. Durand, S., Seigneuret, F., Burlaud-Gaillard, J., Lemoine, R., Tassi, M.-F., Moreau, A., Mougel, M., Roingard, P., Tauber, C., and de Rocquigny, H. (2021) Quantitative

- analysis of the formation of nucleoprotein complexes between HIV-1 Gag protein and genomic RNA using transmission electron microscopy. *J. Biol. Chem.* **298**, 101500
69. 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
70. Bell, N. M., L’Hernault, A., Murat, P., Richards, J. E., Lever, A. M. L., and Balasubramanian, S. (2013) Targeting RNA–Protein Interactions within the Human Immunodeficiency Virus Type 1 Lifecycle. *Biochemistry.* **52**, 9269–9274
71. Mailler, E., Bernacchi, S., Marquet, R., Paillart, J.-C., Vivet-Boudou, V., and Smyth, R. P. (2016) The Life-Cycle of the HIV-1 Gag–RNA Complex. *Viruses.* 10.3390/v8090248
72. 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
73. Cochrane, A. W., Jones, K. S., Beidas, S., Dillon, P. J., Skalka, A. M., and Rosen, C. A. (1991) Identification and characterization of intragenic sequences which repress human immunodeficiency virus structural gene expression. *J. Virol.* **65**, 5305–5313
74. Ostermann, P. N., Ritchie, A., Ptok, J., and Schaal, H. Let It Go: HIV-1 cis-Acting Repressive Sequences. *J. Virol.* **95**, e00342-21
75. Pancera, M., Zhou, T., Druz, A., Georgiev, I. S., Soto, C., Gorman, J., Huang, J., Acharya, P., Chuang, G.-Y., Ofek, G., Stewart-Jones, G. B. E., Stuckey, J., Bailer, R. T., Joyce, M. G., Louder, M. K., Tumba, N., Yang, Y., Zhang, B., Cohen, M. S., Haynes, B. F., Mascola, J. R., Morris, L., Munro, J. B., Blanchard, S. C., Mothes, W., Connors, M., and Kwong, P. D. (2014) Structure and immune recognition of trimeric prefusion HIV-1 Env. *Nature.* **514**, 455–461
76. 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
77. Lee, J. H., Ozorowski, G., and Ward, A. B. (2016) CryoEM structure of a native, fully glycosylated and cleaved HIV-1 envelope trimer. *Science.* **351**, 1043–1048
78. Samson, M., Libert, F., Doranz, B. J., Rucker, J., Liesnard, C., Farber, C. M., Saragosti, S., Lapoumeroulie, C., Cognaux, J., Forceille, C., Muyldermans, G., Verhofstede, C., Burtonboy, G., Georges, M., Imai, T., Rana, S., Yi, Y., Smyth, R. J., Collman, R. G., Doms, R. W., Vassart, G., and Parmentier, M. (1996) Resistance to HIV-1 infection in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. *Nature.* **382**, 722–725
79. Falkenhagen, A., and Joshi, S. (2018) HIV Entry and Its Inhibition by Bifunctional Antiviral Proteins. *Mol. Ther. Nucleic Acids.* **13**, 347–364
80. Yamashita, M., and Engelman, A. N. (2017) Capsid-dependent host factors in HIV-1 infection. *Trends Microbiol.* **25**, 741–755
81. Wilbourne, M., and Zhang, P. (2021) Visualizing HIV-1 Capsid and Its Interactions with Antivirals and Host Factors. *Viruses.* **13**, 246
82. Christensen, D. E., Ganser-Pornillos, B. K., Johnson, J. S., Pornillos, O., and Sundquist, W. I. (2020) Reconstitution and visualization of HIV-1 capsid-dependent replication and integration in vitro. *Science.* **370**, eabc8420
83. Novikova, M., Zhang, Y., Freed, E. O., and Peng, K. (2019) Multiple Roles of HIV-1 Capsid during the Virus Replication Cycle. *Virol. Sin.* **34**, 119–134
84. Hulme, A. E., Perez, O., and Hope, T. J. (2011) Complementary assays reveal a relationship between HIV-1 uncoating and reverse transcription. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 9975–9980

85. AlBurtamani, N., Paul, A., and Fassati, A. (2021) The Role of Capsid in the Early Steps of HIV-1 Infection: New Insights into the Core of the Matter. *Viruses*. **13**, 1161
86. Rihn, S. J., Wilson, S. J., Loman, N. J., Alim, M., Bakker, S. E., Bhella, D., Gifford, R. J., Rixon, F. J., and Bieniasz, P. D. (2013) Extreme genetic fragility of the HIV-1 capsid. *PLoS Pathog.* **9**, e1003461
87. Mattei, S., Glass, B., Hagen, W. J. H., Kräusslich, H.-G., and Briggs, J. A. G. (2016) The structure and flexibility of conical HIV-1 capsids determined within intact virions. *Science*. **354**, 1434–1437
88. Thomas, J. A., Ott, D. E., and Gorelick, R. J. (2007) Efficiency of human immunodeficiency virus type 1 postentry infection processes: evidence against disproportionate numbers of defective virions. *J. Virol.* **81**, 4367–4370
89. Mukherjee, S., Boutant, E., Réal, E., Mély, Y., and Anton, H. (2021) Imaging Viral Infection by Fluorescence Microscopy: Focus on HIV-1 Early Stage. *Viruses*. **13**, 213
90. Francis, A. C., Marin, M., Shi, J., Aiken, C., and Melikyan, G. B. (2016) Time-Resolved Imaging of Single HIV-1 Uncoating In Vitro and in Living Cells. *PLoS Pathog.* **10**.1371/journal.ppat.1005709
91. Müller, T. G., Zila, V., Peters, K., Schifferdecker, S., Stanic, M., Lucic, B., Laketa, V., Lusic, M., Müller, B., and Kräusslich, H.-G. (2021) HIV-1 uncoating by release of viral cDNA from capsid-like structures in the nucleus of infected cells. *eLife*. **10**, e64776
92. Zila, V., Margiotta, E., Turoňová, B., Müller, T. G., Zimmerli, C. E., Mattei, S., Allegretti, M., Börner, K., Rada, J., Müller, B., Lusic, M., Kräusslich, H.-G., and Beck, M. (2021) Cone-shaped HIV-1 capsids are transported through intact nuclear pores. *Cell*. **184**, 1032-1046.e18
93. Blanco-Rodriguez, G., Gazi, A., Monel, B., Frabetti, S., Scoca, V., Mueller, F., Schwartz, O., Krijnse-Locker, J., Charneau, P., and Di Nunzio, F. (2020) Remodeling of the Core Leads HIV-1 Preintegration Complex into the Nucleus of Human Lymphocytes. *J. Virol.* **94**, e00135-20
94. Shen, Q., Wu, C., Freniere, C., Tripler, T. N., and Xiong, Y. (2021) Nuclear Import of HIV-1. *Viruses*. **13**, 2242
95. Perilla, J. R., and Schulten, K. (2017) Physical properties of the HIV-1 capsid from all-atom molecular dynamics simulations. *Nat. Commun.* **8**, 15959
96. McFadden, W. M., Snyder, A. A., Kirby, K. A., Tedbury, P. R., Raj, M., Wang, Z., and Sarafianos, S. G. (2021) Rotten to the core: antivirals targeting the HIV-1 capsid core. *Retrovirology*. **18**, 41
97. Toccafondi, E., Lener, D., and Negroni, M. (2021) HIV-1 Capsid Core: A Bullet to the Heart of the Target Cell. *Front. Microbiol.* **12**, 652486
98. Li, Y., Kar, A. K., and Sodroski, J. (2009) Target cell type-dependent modulation of human immunodeficiency virus type 1 capsid disassembly by cyclophilin A. *J. Virol.* **83**, 10951–10962
99. Maillet, S., Fernandez, J., Decourcelle, M., El Koulali, K., Blanchet, F. P., Arhel, N. J., Maarifi, G., and Nisole, S. (2020) Daxx Inhibits HIV-1 Reverse Transcription and Uncoating in a SUMO-Dependent Manner. *Viruses*. **12**, 636
100. Schaller, T., Ocwieja, K. E., Rasaiyaah, J., Price, A. J., Brady, T. L., Roth, S. L., Hué, S., Fletcher, A. J., Lee, K., KewalRamani, V. N., Noursadeghi, M., Jenner, R. G., James, L. C., Bushman, F. D., and Towers, G. J. (2011) HIV-1 Capsid-Cyclophilin Interactions Determine Nuclear Import Pathway, Integration Targeting and Replication Efficiency. *PLoS Pathog.* **7**, e1002439
101. Yu, A., Skorupka, K. A., Pak, A. J., Ganser-Pornillos, B. K., Pornillos, O., and Voth, G. A. (2020) TRIM5 α self-assembly and compartmentalization of the HIV-1 viral capsid. *Nat. Commun.* **11**, 1307

102. Dharan, A., Opp, S., Abdel-Rahim, O., Keceli, S. K., Imam, S., Diaz-Griffero, F., and Campbell, E. M. (2017) Bicaudal D2 facilitates the cytoplasmic trafficking and nuclear import of HIV-1 genomes during infection. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E10707–E10716
103. Huang, P.-T., Summers, B. J., Xu, C., Perilla, J. R., Malikov, V., Naghavi, M. H., and Xiong, Y. (2019) FEZ1 Is Recruited to a Conserved Cofactor Site on Capsid to Promote HIV-1 Trafficking. *Cell Rep.* **28**, 2373–2385.e7
104. Naghavi, M. H. (2021) HIV-1 capsid exploitation of the host microtubule cytoskeleton during early infection. *Retrovirology.* **18**, 19
105. Bichel, K., Price, A. J., Schaller, T., Towers, G. J., Freund, S. M., and James, L. C. (2013) HIV-1 capsid undergoes coupled binding and isomerization by the nuclear pore protein NUP358. *Retrovirology.* **10**, 81
106. 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
107. Chin, C. R., Perreira, J. M., Savidis, G., Portmann, J. M., Aker, A. M., Feeley, E. M., Smith, M. C., and Brass, A. L. (2015) Direct Visualization of HIV-1 Replication Intermediates Shows that Capsid and CPSF6 Modulate HIV-1 Intra-nuclear Invasion and Integration. *Cell Rep.* **13**, 1717–1731
108. Misumi, S., Inoue, M., Dochi, T., Kishimoto, N., Hasegawa, N., Takamune, N., and Shoji, S. (2010) Uncoating of Human Immunodeficiency Virus Type 1 Requires Prolyl Isomerase Pin1. *J. Biol. Chem.* **285**, 25185–25195
109. Dochi, T., Nakano, T., Inoue, M., Takamune, N., Shoji, S., Sano, K., and Misumi, S. (2014) Phosphorylation of human immunodeficiency virus type 1 capsid protein at serine 16, required for peptidyl-prolyl isomerase-dependent uncoating, is mediated by virion-incorporated extracellular signal-regulated kinase 2. *J. Gen. Virol.* **95**, 1156–1166
110. Takeuchi, H., Saito, H., Noda, T., Miyamoto, T., Yoshinaga, T., Terahara, K., Ishii, H., Tsunetsugu-Yokota, Y., and Yamaoka, S. (2017) Phosphorylation of the HIV-1 capsid by MELK triggers uncoating to promote viral cDNA synthesis. *PLoS Pathog.* **13**, e1006441
111. Fernandez, J., Machado, A. K., Lyonais, S., Chamontin, C., Gärtner, K., Léger, T., Henriquet, C., Garcia, C., Portilho, D. M., Pugnère, M., Chaloin, L., Muriaux, D., Yamauchi, Y., Blaise, M., Nisole, S., and Arhel, N. J. (2019) Transportin-1 binds to the HIV-1 capsid via a nuclear localization signal and triggers uncoating. *Nat. Microbiol.* **4**, 1840–1850
112. 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
113. Stremlau, M., Perron, M., Lee, M., Li, Y., Song, B., Javanbakht, H., Diaz-Griffero, F., Anderson, D. J., Sundquist, W. I., and Sodroski, J. (2006) Specific recognition and accelerated uncoating of retroviral capsids by the TRIM5 α restriction factor. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 5514–5519
114. Kutluay, S. B., Perez-Caballero, D., and Bieniasz, P. D. (2013) Fates of retroviral core components during unrestricted and TRIM5-restricted infection. *PLoS Pathog.* **9**, e1003214
115. Yang, Y., Fricke, T., and Diaz-Griffero, F. (2013) Inhibition of reverse transcriptase activity increases stability of the HIV-1 core. *J. Virol.* **87**, 683–687
116. Cosnefroy, O., Murray, P. J., and Bishop, K. N. (2016) HIV-1 capsid uncoating initiates after the first strand transfer of reverse transcription. *Retrovirology.* **13**, 58
117. Guedán, A., Caroe, E. R., Barr, G. C. R., and Bishop, K. N. (2021) The Role of Capsid in HIV-1 Nuclear Entry. *Viruses.* **13**, 1425

118. Walsh, D., and Naghavi, M. H. (2019) Exploitation of cytoskeletal networks during early viral infection. *Trends Microbiol.* **27**, 39–50
119. Francis, A. C., and Melikyan, G. B. (2018) Single HIV-1 Imaging Reveals Progression of Infection through CA-Dependent Steps of Docking at the Nuclear Pore, Uncoating, and Nuclear Transport. *Cell Host Microbe.* **23**, 536-548.e6
120. Burdick, R. C., Delviks-Frankenberry, K. A., Chen, J., Janaka, S. K., Sastri, J., Hu, W.-S., and Pathak, V. K. (2017) Dynamics and regulation of nuclear import and nuclear movements of HIV-1 complexes. *PLoS Pathog.* **13**, e1006570
121. Zurnic Bönisch, I., Dirix, L., Lemmens, V., Borrenberghs, D., De Wit, F., Vernailen, F., Rocha, S., Christ, F., Hendrix, J., Hofkens, J., and Debysse, Z. (2020) Capsid-Labelled HIV To Investigate the Role of Capsid during Nuclear Import and Integration. *J. Virol.* **94**, e01024-19
122. Francis, A. C., Marin, M., Prellberg, M. J., Palermino-Rowland, K., and Melikyan, G. B. (2020) HIV-1 Uncoating and Nuclear Import Precede the Completion of Reverse Transcription in Cell Lines and in Primary Macrophages. *Viruses.* **12**, E1234
123. 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
124. 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
125. Selyutina, A., Persaud, M., Lee, K., KewalRamani, V., and Diaz-Griffero, F. (2020) Nuclear Import of the HIV-1 Core Precedes Reverse Transcription and Uncoating. [10.1101/2020.03.31.018747](https://doi.org/10.1101/2020.03.31.018747)
126. Otto, G. (2021) HIV-1 capsid gains nuclear entry. *Nat. Rev. Microbiol.* **19**, 223
127. Li, C., Burdick, R. C., Nagashima, K., Hu, W.-S., and Pathak, V. K. (2021) HIV-1 cores retain their integrity until minutes before uncoating in the nucleus. *Proc. Natl. Acad. Sci. U. S. A.* **118**, e2019467118
128. Müller, T. G., Zila, V., Peters, K., Schifferdecker, S., Stanic, M., Lucic, B., Laketa, V., Lusic, M., Müller, B., and Kräusslich, H.-G. HIV-1 uncoating by release of viral cDNA from capsid-like structures in the nucleus of infected cells. *eLife.* **10**, e64776
129. Bhattacharya, A., Alam, S. L., Fricke, T., Zadrozny, K., Sedzicki, J., Taylor, A. B., Demeler, B., Pornillos, O., Ganser-Pornillos, B. K., Diaz-Griffero, F., Ivanov, D. N., and Yeager, M. (2014) Structural basis of HIV-1 capsid recognition by PF74 and CPSF6. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 18625–18630
130. Chen, N.-Y., Zhou, L., Gane, P. J., Opp, S., Ball, N. J., Nicastro, G., Zufferey, M., Buffone, C., Luban, J., Selwood, D., Diaz-Griffero, F., Taylor, I., and Fassati, A. (2016) HIV-1 capsid is involved in post-nuclear entry steps. *Retrovirology.* **13**, 28
131. Balasubramaniam, M., Zhou, J., Addai, A., Martinez, P., Pandhare, J., Aiken, C., and Dash, C. (2019) PF74 Inhibits HIV-1 Integration by Altering the Composition of the Preintegration Complex. *J. Virol.* **93**, e01741-18
132. Burdick, R. C., Li, C., Munshi, M., Rawson, J. M. O., Nagashima, K., Hu, W.-S., and Pathak, V. K. (2020) HIV-1 uncoats in the nucleus near sites of integration. *Proc. Natl. Acad. Sci.* **117**, 5486–5493
133. Price, A. J., Jacques, D. A., McEwan, W. A., Fletcher, A. J., Essig, S., Chin, J. W., Halambage, U. D., Aiken, C., and James, L. C. (2014) Host cofactors and pharmacologic ligands share an essential interface in HIV-1 capsid that is lost upon disassembly. *PLoS Pathog.* **10**, e1004459
134. Buffone, C., Martinez-Lopez, A., Fricke, T., Opp, S., Severgnini, M., Cifola, I., Petiti,

- L., Frabetti, S., Skorupka, K., Zadrozny, K. K., Ganser-Pornillos, B. K., Pornillos, O., Di Nunzio, F., and Diaz-Griffero, F. (2018) Nup153 Unlocks the Nuclear Pore Complex for HIV-1 Nuclear Translocation in Nondividing Cells. *J. Virol.* **92**, e00648-18
135. Ni, T., Gerard, S., Zhao, G., Dent, K., Ning, J., Zhou, J., Shi, J., Anderson-Daniels, J., Li, W., Jang, S., Engelman, A. N., Aiken, C., and Zhang, P. (2020) Intrinsic curvature of HIV-1 CA hexamer underlies capsid topology and interaction with cyclophilin A. *Nat. Struct. Mol. Biol.* **27**, 855–862
136. Gres, A. T., Kirby, K. A., KewalRamani, V. N., Tanner, J. J., Pornillos, O., and Sarafianos, S. G. (2015) X-Ray Structures of Native HIV-1 Capsid Protein Reveal Conformational Variability. *Science.* **349**, 99–103
137. Stanley, G. J., Fassati, A., and Hoogenboom, B. W. (2018) Atomic force microscopy reveals structural variability amongst nuclear pore complexes. *Life Sci. Alliance.* **1**, e201800142
138. Lin, D. H., and Hoelz, A. (2019) The Structure of the Nuclear Pore Complex (An Update). *Annu. Rev. Biochem.* **88**, 725–783
139. Knockenhauer, K. E., and Schwartz, T. U. (2016) The Nuclear Pore Complex as a Flexible and Dynamic Gate. *Cell.* **164**, 1162–1171
140. Mahamid, J., Pfeffer, S., Schaffer, M., Villa, E., Danev, R., Cuellar, L. K., Förster, F., Hyman, A. A., Plitzko, J. M., and Baumeister, W. (2016) Visualizing the molecular sociology at the HeLa cell nuclear periphery. *Science.* **351**, 969–972
141. Rankovic, S., Deshpande, A., Harel, S., Aiken, C., and Rousso, I. (2021) HIV-1 uncoating occurs via a series of rapid biomechanical changes in the core related to individual stages of reverse transcription. *J. Virol.* 10.1128/JVI.00166-21
142. Jacques, D. A., McEwan, W. A., Hilditch, L., Price, A. J., Towers, G. J., and James, L. C. (2016) HIV-1 uses dynamic capsid pores to import nucleotides and fuel encapsidated DNA synthesis. *Nature.* **536**, 349–353
143. Xu, C., Fischer, D. K., Rankovic, S., Li, W., Dick, R. A., Runge, B., Zadorozhnyi, R., Ahn, J., Aiken, C., Polenova, T., Engelman, A. N., Ambrose, Z., Rousso, I., and Perilla, J. R. (2020) Permeability of the HIV-1 capsid to metabolites modulates viral DNA synthesis. *PLoS Biol.* **18**, e3001015
144. Burdick, R. C., Li, C., Munshi, M., Rawson, J. M. O., Nagashima, K., Hu, W.-S., and Pathak, V. K. (2020) HIV-1 uncoats in the nucleus near sites of integration. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 5486–5493
145. Dharan, A., Bachmann, N., Talley, S., Zwickelmaier, V., and Campbell, E. M. (2020) Nuclear pore blockade reveals that HIV-1 completes reverse transcription and uncoating in the nucleus. *Nat. Microbiol.* **5**, 1088–1095
146. Guedán, A., Donaldson, C. D., Cosnefroy, O., Taylor, I. A., and Bishop, K. N. (2021) HIV-1 requires capsid remodelling at the nuclear pore for nuclear entry and integration, 10.1101/2021.03.18.436028
147. Yang, Y., Fricke, T., and Diaz-Griffero, F. (2013) Inhibition of Reverse Transcriptase Activity Increases Stability of the HIV-1 Core. *J. Virol.* **87**, 683–687
148. Rankovic, S., Varadarajan, J., Ramalho, R., Aiken, C., and Rousso, I. (2017) Reverse Transcription Mechanically Initiates HIV-1 Capsid Disassembly. *J. Virol.* 10.1128/JVI.00289-17
149. Francis, A. C., Marin, M., Singh, P. K., Achuthan, V., Prellberg, M. J., Palermino-Rowland, K., Lan, S., Tedbury, P. R., Sarafianos, S. G., Engelman, A. N., and Melikyan, G. B. (2020) HIV-1 replication complexes accumulate in nuclear speckles and integrate into speckle-associated genomic domains. *Nat. Commun.* **11**, 3505
150. Rensen, E., Mueller, F., Scoca, V., Parmar, J. J., Souque, P., Zimmer, C., and Di Nunzio, F. (2021) Clustering and reverse transcription of HIV-1 genomes in nuclear niches of

macrophages. *EMBO J.* **40**, e105247

151. Scoca, V., Louveaux, M., Morin, R., Ershov, D., Tinevez, J.-Y., and Nunzio, F. D. (2021) HIV-induced membraneless organelles orchestrate post-nuclear entry steps. [10.1101/2020.11.17.385567](https://doi.org/10.1101/2020.11.17.385567)
152. Forshey, B. M., von Schwedler, U., Sundquist, W. I., and Aiken, C. (2002) Formation of a Human Immunodeficiency Virus Type 1 Core of Optimal Stability Is Crucial for Viral Replication. *J. Virol.* **76**, 5667–5677
153. Rankovic, S., Varadarajan, J., Ramalho, R., Aiken, C., and Rousso, I. (2017) Reverse Transcription Mechanically Initiates HIV-1 Capsid Disassembly. *J. Virol.* **91**, e00289-17
154. Mamede, J. I., Cianci, G. C., Anderson, M. R., and Hope, T. J. (2017) Early cytoplasmic uncoating is associated with infectivity of HIV-1. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E7169–E7178
155. Mallery, D. L., Márquez, C. L., McEwan, W. A., Dickson, C. F., Jacques, D. A., Anandapadamanaban, M., Bichel, K., Towers, G. J., Saiardi, A., Böcking, T., and James, L. C. IP6 is an HIV pocket factor that prevents capsid collapse and promotes DNA synthesis. *eLife.* **7**, e35335
156. Mallery, D. L., Faysal, K. M. R., Kleinpeter, A., Wilson, M. S. C., Vaysburd, M., Fletcher, A. J., Novikova, M., Böcking, T., Freed, E. O., Saiardi, A., and James, L. C. (2019) Cellular IP6 Levels Limit HIV Production while Viruses that Cannot Efficiently Package IP6 Are Attenuated for Infection and Replication. *Cell Rep.* **29**, 3983-3996.e4
157. 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
158. Li, M., Mizuuchi, M., Burke Jr, T. R., and Craigie, R. (2006) Retroviral DNA integration: reaction pathway and critical intermediates. *EMBO J.* **25**, 1295–1304
159. Craigie, R. (2018) Nucleoprotein intermediates in HIV-1 DNA integration: structure and function of HIV-1 intasomes. *Subcell. Biochem.* **88**, 189–210
160. Schweitzer, C. J., Jagadish, T., Haverland, N., Ciborowski, P., and Belshan, M. (2013) Proteomic Analysis of Early HIV-1 Nucleoprotein Complexes. *J. Proteome Res.* **12**, 559–572
161. Chin, C. R., Perreira, J. M., Savidis, G., Portmann, J. M., Aker, A. M., Feeley, E. M., Smith, M. C., and Brass, A. L. (2015) Direct Visualization of HIV-1 Replication Intermediates Shows that Capsid and CPSF6 Modulate HIV-1 Intra-nuclear Invasion and Integration. *Cell Rep.* **13**, 1717–1731
162. Craigie, R. (2018) Nucleoprotein intermediates in HIV-1 DNA integration: structure and function of HIV-1 intasomes. *Subcell. Biochem.* **88**, 189–210
163. Sherman, P. A., and Fyfe, J. A. (1990) Human immunodeficiency virus integration protein expressed in Escherichia coli possesses selective DNA cleaving activity. *Proc. Natl. Acad. Sci. U. S. A.* **87**, 5119–5123
164. Bushman, F. D., and Craigie, R. (1991) Activities of human immunodeficiency virus (HIV) integration protein in vitro: specific cleavage and integration of HIV DNA. *Proc. Natl. Acad. Sci. U. S. A.* **88**, 1339–1343
165. Engelman, A., Mizuuchi, K., and Craigie, R. (1991) HIV-1 DNA integration: Mechanism of viral DNA cleavage and DNA strand transfer. *Cell.* **67**, 1211–1221
166. Engelman, A. N. (2019) Multifaceted HIV integrase functionalities and therapeutic strategies for their inhibition. *J. Biol. Chem.* **294**, 15137–15157
167. Sloan, R. D., and Wainberg, M. A. (2011) The role of unintegrated DNA in HIV infection. *Retrovirology.* **8**, 52
168. Pang, S., Koyanagi, Y., Miles, S., Wiley, C., Vinters, H. V., and Chen, I. S. (1990) High levels of unintegrated HIV-1 DNA in brain tissue of AIDS dementia patients. *Nature.* **343**, 85–89

169. Teo, I., Veryard, C., Barnes, H., An, S. F., Jones, M., Lantos, P. L., Luthert, P., and Shaunak, S. (1997) Circular forms of unintegrated human immunodeficiency virus type 1 DNA and high levels of viral protein expression: association with dementia and multinucleated giant cells in the brains of patients with AIDS. *J. Virol.* **71**, 2928–2933
170. Chun, T. W., Carruth, L., Finzi, D., Shen, X., DiGiuseppe, J. A., Taylor, H., Hermankova, M., Chadwick, K., Margolick, J., Quinn, T. C., Kuo, Y. H., Brookmeyer, R., Zeiger, M. A., Barditch-Crovo, P., and Siliciano, R. F. (1997) Quantification of latent tissue reservoirs and total body viral load in HIV-1 infection. *Nature.* **387**, 183–188
171. 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
172. Kilzer, J. M., Stracker, T., Beitzel, B., Meek, K., Weitzman, M., and Bushman, F. D. (2003) Roles of host cell factors in circularization of retroviral dna. *Virology.* **314**, 460–467
173. Chan, C. N., Trinité, B., Lee, C. S., Mahajan, S., Anand, A., Wodarz, D., Sabbaj, S., Bansal, A., Goepfert, P. A., and Levy, D. N. (2016) HIV-1 latency and virus production from unintegrated genomes following direct infection of resting CD4 T cells. *Retrovirology.* **13**, 1
174. Delelis, O., Parissi, V., Leh, H., Mbemba, G., Petit, C., Sonigo, P., Deprez, E., and Mouscadet, J.-F. (2007) Efficient and Specific Internal Cleavage of a Retroviral Palindromic DNA Sequence by Tetrameric HIV-1 Integrase. *PLoS ONE.* **2**, e608
175. Thierry, S., Munir, S., Thierry, E., Subra, F., Leh, H., Zamborlini, A., Saenz, D., Levy, D. N., Lesbats, P., Saïb, A., Parissi, V., Poeschla, E., Deprez, E., and Delelis, O. (2015) Integrase inhibitor reversal dynamics indicate unintegrated HIV-1 dna initiate de novo integration. *Retrovirology.* **12**, 24
176. 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
177. Zhang, D., He, H., and Guo, S. (2014) Hairpin DNA probe-based fluorescence assay for detecting palindrome cleavage activity of HIV-1 integrase. *Anal. Biochem.* **460**, 36–38
178. Richetta, C., Thierry, S., Thierry, E., Lesbats, P., Lapaillerie, D., Munir, S., Subra, F., Leh, H., Deprez, E., Parissi, V., and Delelis, O. (2019) Two-long terminal repeat (LTR) DNA circles are a substrate for HIV-1 integrase. *J. Biol. Chem.* **294**, 8286–8295
179. Colgan, D. F., and Manley, J. L. (1997) Mechanism and regulation of mRNA polyadenylation. *Genes Dev.* **11**, 2755–2766
180. Parent, A., and Bisailon, M. (2006) [Synergy between transcription and mRNA processing events]. *Med. Sci. MS.* **22**, 626–632
181. Sertznig, H., Hillebrand, F., Erkelenz, S., Schaal, H., and Widera, M. (2018) Behind the scenes of HIV-1 replication: Alternative splicing as the dependency factor on the quiet. *Virology.* **516**, 176–188
182. Nguyen Quang, N., Goudey, S., Ségéral, E., Mohammad, A., Lemoine, S., Blugeon, C., Versapuech, M., Paillart, J.-C., Berlioz-Torrent, C., Emiliani, S., and Gallois-Montbrun, S. (2020) Dynamic nanopore long-read sequencing analysis of HIV-1 splicing events during the early steps of infection. *Retrovirology.* **17**, 25
183. Ocwieja, K. E., Sherrill-Mix, S., Mukherjee, R., Custers-Allen, R., David, P., Brown, M., Wang, S., Link, D. R., Olson, J., Travers, K., Schadt, E., and Bushman, F. D. (2012) Dynamic regulation of HIV-1 mRNA populations analyzed by single-molecule enrichment and long-read sequencing. *Nucleic Acids Res.* **40**, 10345–10355
184. Langer, S., and Sauter, D. (2016) Unusual Fusion Proteins of HIV-1. *Front. Microbiol.* **7**, 2152
185. Purcell, D. F., and Martin, M. A. (1993) Alternative splicing of human immunodeficiency virus type 1 mRNA modulates viral protein expression, replication, and

infectivity. *J. Virol.* **67**, 6365–6378

186. Taniguchi, I., Mabuchi, N., and Ohno, M. (2014) HIV-1 Rev protein specifies the viral RNA export pathway by suppressing TAP/NXF1 recruitment. *Nucleic Acids Res.* **42**, 6645–6658

187. Li, G., and De Clercq, E. (2016) HIV Genome-Wide Protein Associations: a Review of 30 Years of Research. *Microbiol. Mol. Biol. Rev. MMBR.* **80**, 679–731

188. Pollard, V. W., and Malim, M. H. (1998) The Hiv-1 Rev Protein. *Annu. Rev. Microbiol.* **52**, 491–532

189. Weeks, K. M., and Crothers, D. M. (1991) RNA recognition by Tat-derived peptides: Interaction in the major groove? *Cell.* **66**, 577–588

190. Obayashi, C. M., Shinohara, Y., Masuda, T., and Kawai, G. (2021) Influence of the 5'-terminal sequences on the 5'-UTR structure of HIV-1 genomic RNA. *Sci. Rep.* **11**, 10920

191. Blumenthal, R., Durell, S., and Viard, M. (2012) HIV Entry and Envelope Glycoprotein-mediated Fusion. *J. Biol. Chem.* **287**, 40841–40849

192. McCune, J. M., Rabin, L. B., Feinberg, M. B., Lieberman, M., Kosek, J. C., Reyes, G. R., and Weissman, I. L. (1988) Endoproteolytic cleavage of gp160 is required for the activation of human immunodeficiency virus. *Cell.* **53**, 55–67

193. Kurnaeva, M. A., Sheval, E. V., Musinova, Y. R., and Vassetzky, Y. S. (2019) Tat basic domain: A “Swiss army knife” of HIV-1 Tat? *Rev. Med. Virol.* **29**, e2031

194. De Marco, A., Dans, P. D., Knezevich, A., Maiuri, P., Pantano, S., and Marcello, A. (2010) Subcellular localization of the interaction between the human immunodeficiency virus transactivator Tat and the nucleosome assembly protein 1. *Amino Acids.* **38**, 1583–1593

195. Mahmoudi, T., Parra, M., Vries, R. G. J., Kauder, S. E., Verrijzer, C. P., Ott, M., and Verdin, E. (2006) The SWI/SNF chromatin-remodeling complex is a cofactor for Tat transactivation of the HIV promoter. *J. Biol. Chem.* **281**, 19960–19968

196. Veschambre, P., Simard, P., and Jalinet, P. (1995) Evidence for functional interaction between the HIV-1 Tat transactivator and the TATA box binding protein in vivo. *J. Mol. Biol.* **250**, 169–180

197. Zhou, Q., Chen, D., Pierstorff, E., and Luo, K. (1998) Transcription elongation factor P-TEFb mediates Tat activation of HIV-1 transcription at multiple stages. *EMBO J.* **17**, 3681–3691

198. Zhou, Q., Chen, D., Pierstorff, E., and Luo, K. (1998) Transcription elongation factor P-TEFb mediates Tat activation of HIV-1 transcription at multiple stages. *EMBO J.* **17**, 3681–3691

199. Delling, U., Roy, S., Sumner-Smith, M., Barnett, R., Reid, L., Rosen, C. A., and Sonenberg, N. (1991) The number of positively charged amino acids in the basic domain of Tat is critical for trans-activation and complex formation with TAR RNA. *Proc. Natl. Acad. Sci. U. S. A.* **88**, 6234–6238

200. Calnan, B. J., Biancalana, S., Hudson, D., and Frankel, A. D. (1991) Analysis of arginine-rich peptides from the HIV Tat protein reveals unusual features of RNA-protein recognition. *Genes Dev.* **5**, 201–210

201. Wei, P., Garber, M. E., Fang, S. M., Fischer, W. H., and Jones, K. A. (1998) A novel CDK9-associated C-type cyclin interacts directly with HIV-1 Tat and mediates its high-affinity, loop-specific binding to TAR RNA. *Cell.* **92**, 451–462

202. 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

203. Rustanti, L., Jin, H., Lor, M., Lin, M. H., Rawle, D. J., and Harrich, D. (2017) A mutant Tat protein inhibits infection of human cells by strains from diverse HIV-1 subtypes. *Virol. J.* **14**, 52

204. 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
205. Williams, S. A., Chen, L.-F., Kwon, H., Ruiz-Jarabo, C. M., Verdin, E., and Greene, W. C. (2006) NF-kappaB p50 promotes HIV latency through HDAC recruitment and repression of transcriptional initiation. *EMBO J.* **25**, 139–149
206. Deng, L., de la Fuente, C., Fu, P., Wang, L., Donnelly, R., Wade, J. D., Lambert, P., Li, H., Lee, C. G., and Kashanchi, F. (2000) Acetylation of HIV-1 Tat by CBP/P300 increases transcription of integrated HIV-1 genome and enhances binding to core histones. *Virology.* **277**, 278–295
207. Marzio, G., Tyagi, M., Gutierrez, M. I., and Giacca, M. (1998) HIV-1 Tat transactivator recruits p300 and CREB-binding protein histone acetyltransferases to the viral promoter. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 13519–13524
208. Malim, M. H., and Cullen, B. R. (1991) HIV-1 structural gene expression requires the binding of multiple Rev monomers to the viral RRE: implications for HIV-1 latency. *Cell.* **65**, 241–248
209. Fackler, O. T., Kienzle, N., Kremmer, E., Boese, A., Schramm, B., Klimkait, T., Kücherer, C., and Mueller-Lantzsch, N. (1997) Association of human immunodeficiency virus Nef protein with actin is myristoylation dependent and influences its subcellular localization. *Eur. J. Biochem.* **247**, 843–851
210. Jäger, S., Cimermancic, P., Gulbahce, N., Johnson, J. R., McGovern, K. E., Clarke, S. C., Shales, M., Mercenne, G., Pache, L., Li, K., Hernandez, H., Jang, G. M., Roth, S. L., Akiva, E., Marlett, J., Stephens, M., D’Orso, I., Fernandes, J., Fahey, M., Mahon, C., O’Donoghue, A. J., Todorovic, A., Morris, J. H., Maltby, D. A., Alber, T., Cagney, G., Bushman, F. D., Young, J. A., Chanda, S. K., Sundquist, W. I., Kortemme, T., Hernandez, R. D., Craik, C. S., Burlingame, A., Sali, A., Frankel, A. D., and Krogan, N. J. (2012) Global landscape of HIV–human protein complexes. *Nature.* **481**, 365–370
211. Pereira, E. A., and daSilva, L. L. P. (2016) HIV-1 Nef: Taking Control of Protein Trafficking. *Traffic.* **17**, 976–996
212. Garcia, J. V., and Miller, A. D. (1991) Serine phosphorylation-independent downregulation of cell-surface CD4 by nef. *Nature.* **350**, 508–511
213. Schwartz, O., Maréchal, V., Le Gall, S., 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
214. Usami, Y., Wu, Y., and Göttlinger, H. G. (2015) SERINC3 and SERINC5 restrict HIV-1 infectivity and are counteracted by Nef. *Nature.* **526**, 218–223
215. Wu, Y., Olety, B., Weiss, E. R., Popova, E., Yamanaka, H., and Göttlinger, H. (2019) Potent Enhancement of HIV-1 Replication by Nef in the Absence of SERINC3 and SERINC5. *mBio.* **10**, e01071-19
216. Mwimanzi, P., Markle, T. J., Ogata, Y., Martin, E., Tokunaga, M., Mahiti, M., Kuang, X. T., Walker, B. D., Brockman, M. A., Brumme, Z. L., and Ueno, T. (2013) Dynamic range of Nef functions in chronic HIV-1 infection. *Virology.* **439**, 74–80
217. Bachand, F., Yao, X. J., Hrimech, 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
218. Desai, T. M., Marin, M., Sood, C., Shi, J., Nawaz, F., Aiken, C., and Melikyan, G. B. (2015) Fluorescent protein-tagged Vpr dissociates from HIV-1 core after viral fusion and rapidly enters the cell nucleus. *Retrovirology.* **12**, 88
219. Fabryova, H., and Strebel, K. (2019) Vpr and Its Cellular Interaction Partners: R We There Yet? *Cells.* 10.3390/cells8111310
220. Baldauf, H.-M., Pan, X., Erikson, E., Schmidt, S., Daddacha, W., Burggraf, M.,

- Schenkova, K., Ambiel, I., Wabnitz, G., Gramberg, T., Panitz, S., Flory, E., Landau, N. R., Sertel, S., Rutsch, F., Lasitschka, F., Kim, B., König, R., Fackler, O. T., and Keppler, O. T. (2012) SAMHD1 restricts HIV-1 infection in resting CD4(+) T cells. *Nat. Med.* **18**, 1682–1687
221. Magadán, J. G., Pérez-Victoria, F. J., Sougrat, R., Ye, Y., Strebel, K., and Bonifacino, J. S. (2010) Multilayered mechanism of CD4 downregulation by HIV-1 Vpu involving distinct ER retention and ERAD targeting steps. *PLoS Pathog.* **6**, e1000869
222. 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
223. Faust, T. B., Binning, J. M., Gross, J. D., and Frankel, A. D. (2017) Making Sense of Multifunctional Proteins: Human Immunodeficiency Virus Type 1 Accessory and Regulatory Proteins and Connections to Transcription. *Annu. Rev. Virol.* **4**, 241–260
224. Liu, H., Wu, X., Newman, M., Shaw, G. M., Hahn, B. H., and Kappes, J. C. (1995) The Vif protein of human and simian immunodeficiency viruses is packaged into virions and associates with viral core structures. *J. Virol.* **69**, 7630–7638
225. Stupfler, B., Verriez, C., Gallois-Montbrun, S., Marquet, R., and Paillart, J.-C. (2021) Degradation-Independent Inhibition of APOBEC3G by the HIV-1 Vif Protein. *Viruses.* **13**, 617
226. Sheehy, A. M., Gaddis, N. C., and Malim, M. H. (2003) The antiretroviral enzyme APOBEC3G is degraded by the proteasome in response to HIV-1 Vif. *Nat. Med.* **9**, 1404–1407
227. Faust, T. B., Binning, J. M., Gross, J. D., and Frankel, A. D. (2017) Making Sense of Multifunctional Proteins: Human Immunodeficiency Virus Type 1 Accessory and Regulatory Proteins and Connections to Transcription. *Annu. Rev. Virol.* **4**, 241–260
228. Sheehy, A. M., Gaddis, N. C., and Malim, M. H. (2003) The antiretroviral enzyme APOBEC3G is degraded by the proteasome in response to HIV-1 Vif. *Nat. Med.* **9**, 1404–1407
229. Anderson, B. D., and Harris, R. S. (2015) Transcriptional regulation of APOBEC3 antiviral immunity through the CBF- β /RUNX axis. *Sci. Adv.* **1**, e1500296
230. Guerrero, S., Libre, C., Batisse, J., Mercenne, G., Richer, D., Laumond, G., Decoville, T., Moog, C., Marquet, R., and Paillart, J.-C. (2016) Translational regulation of APOBEC3G mRNA by Vif requires its 5'UTR and contributes to restoring HIV-1 infectivity. *Sci. Rep.* **6**, 39507
231. Mercenne, G., Bernacchi, S., Richer, D., Bec, G., Henriët, S., Paillart, J.-C., and Marquet, R. (2010) HIV-1 Vif binds to APOBEC3G mRNA and inhibits its translation. *Nucleic Acids Res.* **38**, 633–646
232. 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
233. Santa-Marta, M., Silva, F. A. da, Fonseca, A. M., and Goncalves, J. (2005) HIV-1 Vif Can Directly Inhibit Apolipoprotein B mRNA-editing Enzyme Catalytic Polypeptide-like 3G-mediated Cytidine Deamination by Using a Single Amino Acid Interaction and Without Protein Degradation *. *J. Biol. Chem.* **280**, 8765–8775
234. Feng, Y., Love, R. P., and Chelico, L. (2013) HIV-1 Viral Infectivity Factor (Vif) Alters Processive Single-stranded DNA Scanning of the Retroviral Restriction Factor APOBEC3G*[S]. *J. Biol. Chem.* **288**, 6083–6094
235. Bernacchi, S., Henriët, S., Dumas, P., Paillart, J.-C., and Marquet, R. (2007) RNA and DNA Binding Properties of HIV-1 Vif Protein: A FLUORESCENCE STUDY *. *J. Biol.*

Chem. **282**, 26361–26368

236. Louis, J. M., Wondrak, E. M., Kimmel, A. R., Wingfield, P. T., and Nashed, N. T. (1999) Proteolytic processing of HIV-1 protease precursor, kinetics and mechanism. *J. Biol. Chem.* **274**, 23437–23442
237. 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
238. Agniswamy, J., Sayer, J. M., Weber, I. T., and Louis, J. M. (2012) Terminal interface conformations modulate dimer stability prior to amino terminal autoprocessing of HIV-1 protease. *Biochemistry*. **51**, 1041–1050
239. 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
240. Potempa, M., Nalivaika, E., Ragland, D., Lee, S.-K., Schiffer, C. A., and Swanstrom, R. (2015) A Direct Interaction With RNA Dramatically Enhances the Catalytic Activity of the HIV-1 Protease In Vitro. *J. Mol. Biol.* **427**, 2360–2378
241. Konvalinka, J., Kräusslich, H.-G., and Müller, B. (2015) Retroviral proteases and their roles in virion maturation. *Virology*. **479–480**, 403–417
242. Ghosh, A. K., Osswald, H. L., and Prato, G. (2016) Recent Progress in the Development of HIV-1 Protease Inhibitors for the Treatment of HIV/AIDS. *J. Med. Chem.* **59**, 5172–5208
243. 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
244. Guiot, E., Carayon, K., Delelis, O., Simon, F., Tauc, P., Zubin, E., Gottikh, M., Mouscadet, J.-F., Brochon, J.-C., and Deprez, E. (2006) Relationship between the Oligomeric Status of HIV-1 Integrase on DNA and Enzymatic Activity. *J. Biol. Chem.* **281**, 22707–22719
245. Christ, F., Thys, W., De Rijck, J., Gijssbers, R., Albanese, A., Arosio, D., Emiliani, S., Rain, J.-C., Benarous, R., Cereseto, A., and Debysier, Z. (2008) Transportin-SR2 imports HIV into the nucleus. *Curr. Biol. CB*. **18**, 1192–1202
246. Kvaratskhelia, M., Sharma, A., Larue, R. C., Serrao, E., and Engelman, A. (2014) Molecular mechanisms of retroviral integration site selection. *Nucleic Acids Res.* **42**, 10209–10225
247. Engelman, A. N., and Singh, P. K. (2018) Cellular and molecular mechanisms of HIV-1 integration targeting. *Cell. Mol. Life Sci. CMLS*. **75**, 2491–2507
248. Miller, M. T., Tuske, S., Das, K., DeStefano, J. J., and Arnold, E. (2016) Structure of HIV-1 reverse transcriptase bound to a novel 38-mer hairpin template-primer DNA aptamer. *Protein Sci. Publ. Protein Soc.* **25**, 46–55
249. Ruiz, F. X., and Arnold, E. (2020) Evolving understanding of HIV-1 reverse transcriptase structure, function, inhibition, and resistance. *Curr. Opin. Struct. Biol.* **61**, 113–123
250. Lightfoote, M. M., Coligan, J. E., Folks, T. M., Fauci, A. S., Martin, M. A., and Venkatesan, S. (1986) Structural characterization of reverse transcriptase and endonuclease polypeptides of the acquired immunodeficiency syndrome retrovirus. *J. Virol.* **60**, 771–775
251. Steitz, T. A. (1999) DNA Polymerases: Structural Diversity and Common Mechanisms. *J. Biol. Chem.* **274**, 17395–17398
252. Huang, H., Chopra, R., Verdine, G. L., and Harrison, S. C. (1998) Structure of a Covalently Trapped Catalytic Complex of HIV-1 Reverse Transcriptase: Implications for Drug Resistance. *Science*. **282**, 1669–1675
253. Hu, W.-S., and Hughes, S. H. (2012) HIV-1 Reverse Transcription. *Cold Spring Harb.*

Perspect. Med. 10.1101/cshperspect.a006882

254. Abram, M. E., Ferris, A. L., Shao, W., Alvord, W. G., and Hughes, S. H. (2010) Nature, Position, and Frequency of Mutations Made in a Single Cycle of HIV-1 Replication. *J. Virol.* **84**, 9864–9878
255. Sanjuán, R., Nebot, M. R., Chirico, N., Mansky, L. M., and Belshaw, R. (2010) Viral Mutation Rates. *J. Virol.* **84**, 9733–9748
256. Cuevas, J. M., Geller, R., Garijo, R., López-Aldeguer, J., and Sanjuán, R. (2015) Extremely High Mutation Rate of HIV-1 In Vivo. *PLoS Biol.* 10.1371/journal.pbio.1002251
257. Ho, Y.-C., Shan, L., Hosmane, N. N., Wang, J., Laskey, S. B., Rosenbloom, D. I. S., Lai, J., Blankson, J. N., Siliciano, J. D., and Siliciano, R. F. (2013) Replication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure. *Cell.* **155**, 540–551
258. Rihn, S. J., Hughes, J., Wilson, S. J., and Bieniasz, P. D. (2015) Uneven genetic robustness of HIV-1 integrase. *J. Virol.* **89**, 552–567
259. Ruiz, F. X., and Arnold, E. (2020) Evolving understanding of HIV-1 reverse transcriptase structure, function, inhibition, and resistance. *Curr. Opin. Struct. Biol.* **61**, 113–123
260. Corona, A., Meleddu, R., Esposito, F., Distinto, S., Bianco, G., Masaoka, T., Maccioni, E., Menéndez-Arias, L., Alcaro, S., Le Grice, S. F. J., and Tramontano, E. (2016) Ribonuclease H/DNA Polymerase HIV-1 Reverse Transcriptase Dual Inhibitor: Mechanistic Studies on the Allosteric Mode of Action of Isatin-Based Compound RMNC6. *PLOS ONE.* **11**, e0147225
261. Furfine, E. S., and Reardon, J. E. (1991) Reverse transcriptase.RNase H from the human immunodeficiency virus. Relationship of the DNA polymerase and RNA hydrolysis activities. *J. Biol. Chem.* **266**, 406–412
262. Figiel, M., Krepl, M., Poznański, J., Gołąb, A., Šponer, J., and Nowotny, M. (2017) Coordination between the polymerase and RNase H activity of HIV-1 reverse transcriptase. *Nucleic Acids Res.* **45**, 3341–3352
263. Tisdale, M., Schulze, T., Larder, B. A., and Moelling, K. (1991) Mutations within the RNase H domain of human immunodeficiency virus type 1 reverse transcriptase abolish virus infectivity. *J. Gen. Virol.* **72** (Pt 1), 59–66
264. Volkmann, S., Wöhrl, B. M., Tisdale, M., and Moelling, K. (1993) Enzymatic analysis of two HIV-1 reverse transcriptase mutants with mutations in carboxyl-terminal amino acid residues conserved among retroviral ribonucleases H. *J. Biol. Chem.* **268**, 2674–2683
265. Tanese, N., and Goff, S. P. (1988) Domain structure of the Moloney murine leukemia virus reverse transcriptase: mutational analysis and separate expression of the DNA polymerase and RNase H activities. *Proc. Natl. Acad. Sci. U. S. A.* **85**, 1777–1781
266. Schatz, O., Cromme, F. V., Grüninger-Leitch, F., and Le Grice, S. F. (1989) Point mutations in conserved amino acid residues within the C-terminal domain of HIV-1 reverse transcriptase specifically repress RNase H function. *FEBS Lett.* **257**, 311–314
267. Krug, M. S., and Berger, S. L. (1989) Ribonuclease H activities associated with viral reverse transcriptases are endonucleases. *Proc. Natl. Acad. Sci. U. S. A.* **86**, 3539–3543
268. Fedoroff OYu, null, Salazar, M., and Reid, B. R. (1993) Structure of a DNA:RNA hybrid duplex. Why RNase H does not cleave pure RNA. *J. Mol. Biol.* **233**, 509–523
269. Beilhartz, G. L., Wendeler, M., Baichoo, N., Rausch, J., Le Grice, S., and Götte, M. (2009) HIV-1 reverse transcriptase can simultaneously engage its DNA/RNA substrate at both DNA polymerase and RNase H active sites: implications for RNase H inhibition. *J. Mol. Biol.* **388**, 462–474
270. Champoux, J. J., and Schultz, S. J. (2009) Ribonuclease H: Properties, Substrate Specificity, and Roles in Retroviral Reverse Transcription. *FEBS J.* **276**, 1506–1516

271. Champoux, J. J., and Schultz, S. J. (2009) Ribonuclease H: Properties, Substrate Specificity, and Roles in Retroviral Reverse Transcription. *FEBS J.* **276**, 1506–1516
272. Li, A., Li, J., and Johnson, K. A. (2016) HIV-1 Reverse Transcriptase Polymerase and RNase H (Ribonuclease H) Active Sites Work Simultaneously and Independently. *J. Biol. Chem.* **291**, 26566–26585
273. Abbondanzieri, E. A., Bokinsky, G., Rausch, J. W., Zhang, J. X., Le Grice, S. F. J., and Zhuang, X. (2008) Dynamic binding orientations direct activity of HIV reverse transcriptase. *Nature.* **453**, 184–189
274. Friedrich, M., Setz, C., Hahn, F., Matthaei, A., Fraedrich, K., Rauch, P., Henklein, P., Traxdorf, M., Fossen, T., and Schubert, U. (2016) Glutamic Acid Residues in HIV-1 p6 Regulate Virus Budding and Membrane Association of Gag. *Viruses.* 10.3390/v8040117
275. Salgado, G. F., Marquant, R., Vogel, A., Alves, I. D., Feller, S. E., Morellet, N., and Bouaziz, S. (2009) Structural studies of HIV-1 Gag p6ct and its interaction with Vpr determined by solution nuclear magnetic resonance. *Biochemistry.* **48**, 2355–2367
276. Kondo, E., Mammano, F., Cohen, E. A., and Göttlinger, H. G. (1995) The p6gag domain of human immunodeficiency virus type 1 is sufficient for the incorporation of Vpr into heterologous viral particles. *J. Virol.* **69**, 2759–2764
277. Pornillos, O., and Ganser-Pornillos, B. K. (2019) Maturation of retroviruses. *Curr. Opin. Virol.* **36**, 47–55
278. Mercredi, P. Y., Bucca, N., Loeliger, B., Gaines, C. R., Mehta, M., Bhargava, P., Tedbury, P. R., Charlier, L., Floquet, N., Muriaux, D., Favard, C., Sanders, C. R., Freed, E. O., Marchant, J., and Summers, M. F. (2016) Structural and Molecular Determinants of Membrane Binding by the HIV-1 Matrix Protein. *J. Mol. Biol.* **428**, 1637–1655
279. 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. U. S. A.* **93**, 3099–3104
280. Pezeshkian, N., Groves, N. S., and van Engelenburg, S. B. (2019) Single-molecule imaging of HIV-1 envelope glycoprotein dynamics and Gag lattice association exposes determinants responsible for virus incorporation. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 25269–25277
281. Murphy, R. E., Samal, A. B., Vlach, J., Mas, V., Prevelige, P. E., and Saad, J. S. (2019) Structural and biophysical characterizations of HIV-1 matrix trimer binding to lipid nanodiscs shed light on virus assembly. *J. Biol. Chem.* **294**, 18600–18612
282. Perilla, J. R., and Gronenborn, A. M. (2016) Molecular Architecture of the Retroviral Capsid. *Trends Biochem. Sci.* **41**, 410–420
283. Gamble, T. R., Vajdos, F. F., Yoo, S., Worthylake, D. K., Houseweart, M., Sundquist, W. I., and Hill, C. P. (1996) Crystal structure of human cyclophilin A bound to the amino-terminal domain of HIV-1 capsid. *Cell.* **87**, 1285–1294
284. von Schwedler, U. K., Stemmler, T. L., Klishko, V. Y., Li, S., Albertine, K. H., Davis, D. R., and Sundquist, W. I. (1998) Proteolytic refolding of the HIV-1 capsid protein amino-terminus facilitates viral core assembly. *EMBO J.* **17**, 1555–1568
285. Briggs, J. A. G., Simon, M. N., Gross, I., Kräusslich, H.-G., Fuller, S. D., Vogt, V. M., and Johnson, M. C. (2004) The stoichiometry of Gag protein in HIV-1. *Nat. Struct. Mol. Biol.* **11**, 672–675
286. Chertova, E., Chertov, O., Coren, L. V., Roser, J. D., Trubey, C. M., Bess, J. W., Sowder, R. C., Barsov, E., Hood, B. L., Fisher, R. J., Nagashima, K., Conrads, T. P., Veenstra, T. D., Lifson, J. D., and Ott, D. E. (2006) Proteomic and biochemical analysis of purified human immunodeficiency virus type 1 produced from infected monocyte-derived macrophages. *J. Virol.* **80**, 9039–9052

287. Morellet, N., Jullian, N., De Rocquigny, H., Maigret, B., Darlix, J. L., and Roques, B. P. (1992) Determination of the structure of the nucleocapsid protein NCp7 from the human immunodeficiency virus type 1 by 1H NMR. *EMBO J.* **11**, 3059–3065
288. Morellet, N., de Rocquigny, H., Mély, Y., Jullian, N., Déméné, H., Ottmann, M., Gérard, D., Darlix, J. L., Fournie-Zaluski, M. C., and Roques, B. P. (1994) Conformational behaviour of the active and inactive forms of the nucleocapsid NCp7 of HIV-1 studied by 1H NMR. *J. Mol. Biol.* **235**, 287–301
289. Retureau, R., Oguey, C., Mauffret, O., and Hartmann, B. (2019) Structural Explorations of NCp7–Nucleic Acid Complexes Give Keys to Decipher the Binding Process. *J. Mol. Biol.* **431**, 1966–1980
290. 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
291. 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
292. René, B., Mauffret, O., and Fossé, P. (2018) Retroviral nucleocapsid proteins and DNA strand transfers. *Biochim. Open.* **7**, 10–25
293. Kanevsky, I., Chaminade, F., Ficheux, D., Moumen, A., Gorelick, R., Negroni, M., Darlix, J.-L., and Fossé, P. (2005) Specific interactions between HIV-1 nucleocapsid protein and the TAR element. *J. Mol. Biol.* **348**, 1059–1077
294. 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
295. Fisher, R. J., Rein, A., Fivash, M., Urbaneja, M. A., Casas-Finet, J. R., Medaglia, M., and Henderson, L. E. (1998) Sequence-specific binding of human immunodeficiency virus type 1 nucleocapsid protein to short oligonucleotides. *J. Virol.* **72**, 1902–1909
296. Kanevsky, I., Chaminade, F., Chen, Y., Godet, J., René, B., Darlix, J.-L., Mély, Y., Mauffret, O., and Fossé, P. (2011) Structural determinants of TAR RNA-DNA annealing in the absence and presence of HIV-1 nucleocapsid protein. *Nucleic Acids Res.* **39**, 8148–8162
297. 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
298. Comas-Garcia, M., Davis, S. R., and Rein, A. (2016) On the Selective Packaging of Genomic RNA by HIV-1. *Viruses*. 10.3390/v8090246
299. Olson, E. D., and Musier-Forsyth, K. (2019) Retroviral Gag protein - RNA interactions: Implications for specific genomic RNA packaging and virion assembly. *Semin. Cell Dev. Biol.* **86**, 129–139
300. 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
301. Belfetmi, A., Zargarian, L., Tisné, C., Sleiman, D., Morellet, N., Lescop, E., Maskri, O., René, B., Mély, Y., Fossé, P., and Mauffret, O. (2016) Insights into the mechanisms of RNA secondary structure destabilization by the HIV-1 nucleocapsid protein. *RNA N. Y. N.* **22**, 506–517
302. Levin, J. G., Mitra, M., Mascarenhas, A., and Musier-Forsyth, K. (2010) Role of HIV-1 nucleocapsid protein in HIV-1 reverse transcription. *RNA Biol.* **7**, 754–774
303. Belfetmi, A., Zargarian, L., Tisné, C., Sleiman, D., Morellet, N., Lescop, E., Maskri, O., René, B., Mély, Y., Fossé, P., and Mauffret, O. (2016) Insights into the mechanisms of

- RNA secondary structure destabilization by the HIV-1 nucleocapsid protein. *RNA N. Y. N.* **22**, 506–517
304. McCauley, M. J., Rouzina, I., Manthei, K. A., Gorelick, R. J., Musier-Forsyth, K., and Williams, M. C. (2015) Targeted binding of nucleocapsid protein transforms the folding landscape of HIV-1 TAR RNA. *Proc. Natl. Acad. Sci.* **112**, 13555–13560
305. Godet, J., and Mély, Y. (2010) Biophysical studies of the nucleic acid chaperone properties of the HIV-1 nucleocapsid protein. *RNA Biol.* **7**, 687–699
306. Comas-Garcia, M., Datta, S. A., Baker, L., Varma, R., Gudla, P. R., and Rein, A. (2017) Dissection of specific binding of HIV-1 Gag to the “packaging signal” in viral RNA. *eLife.* **6**, e27055
307. Rein, A. (2019) RNA Packaging in HIV. *Trends Microbiol.* **27**, 715–723
308. 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
309. 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
310. 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. U. S. A.* **91**, 4945–4949
311. Paillart, J. C., Skripkin, E., Ehresmann, B., Ehresmann, C., and Marquet, R. (1996) A loop-loop “kissing” complex is the essential part of the dimer linkage of genomic HIV-1 RNA. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 5572–5577
312. Blakemore, R. J., Burnett, C., Swanson, C., Kharytonchyk, S., Telesnitsky, A., and Munro, J. B. (2021) Stability and conformation of the dimeric HIV-1 genomic RNA 5'UTR. *Biophys. J.* **120**, 4874–4890
313. Ye, L., Gribling-Burrer, A.-S., Bohn, P., Kibe, A., Börtlein, C., Ambi, U. B., Ahmad, S., Olguin-Nava, M., Smith, M., Caliskan, N., von Kleist, M., and Smyth, R. P. (2022) Short- and long-range interactions in the HIV-1 5' UTR regulate genome dimerization and packaging. *Nat. Struct. Mol. Biol.* **29**, 306–319
314. Bell, N. M., and Lever, A. M. L. (2013) HIV Gag polyprotein: processing and early viral particle assembly. *Trends Microbiol.* **21**, 136–144
315. Kafaie, J., Song, R., Abrahamyan, L., Mouland, A. J., and Laughrea, M. (2008) Mapping of nucleocapsid residues important for HIV-1 genomic RNA dimerization and packaging. *Virology.* **375**, 592–610
316. Muriaux, D., Girard, P.-M., Bonnet-Mathonière, B., and Paoletti, J. (1995) Dimerization of HIV-1Lai RNA at Low Ionic Strength AN AUTOCOMPLEMENTARY SEQUENCE IN THE 5' LEADER REGION IS EVIDENCED BY AN ANTISENSE OLIGONUCLEOTIDE. *J. Biol. Chem.* **270**, 8209–8216
317. Keane, S. C., Van, V., Frank, H. M., Sciandra, C. A., McCowin, S., Santos, J., Heng, X., and Summers, M. F. (2016) NMR detection of intermolecular interaction sites in the dimeric 5'-leader of the HIV-1 genome. *Proc. Natl. Acad. Sci.* **113**, 13033–13038
318. 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
319. Ding, P., Kharytonchyk, S., Waller, A., Mbaekwe, U., Basappa, S., Kuo, N., Frank, H. M., Quasney, C., Kidane, A., Swanson, C., Van, V., Sarkar, M., Cannistraci, E., Chaudhary, R., Flores, H., Telesnitsky, A., and Summers, M. F. (2020) Identification of the initial nucleocapsid recognition element in the HIV-1 RNA packaging signal. *Proc. Natl. Acad. Sci.*

U. S. A. **117**, 17737–17746

320. 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
321. Didierlaurent, L., Racine, P. J., Houzet, L., Chamontin, C., Berkhout, B., and Mougel, M. (2011) Role of HIV-1 RNA and protein determinants for the selective packaging of spliced and unspliced viral RNA and host U6 and 7SL RNA in virus particles. *Nucleic Acids Res.* **39**, 8915–8927
322. Nikolaitchik, O. A., Somoulay, X., Rawson, J. M. O., Yoo, J. A., Pathak, V. K., and Hu, W.-S. (2020) Unpaired Guanosines in the 5' Untranslated Region of HIV-1 RNA Act Synergistically To Mediate Genome Packaging. *J. Virol.* **94**, e00439-20
323. 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
324. Lu, K., Heng, X., and Summers, M. F. (2011) Structural Determinants and Mechanism of HIV-1 Genome Packaging. *J. Mol. Biol.* **410**, 609–633
325. Das, A. T., Vrolijk, M. M., Harwig, A., and Berkhout, B. (2012) Opening of the TAR hairpin in the HIV-1 genome causes aberrant RNA dimerization and packaging. *Retrovirology.* **9**, 59
326. Smyth, R. P., Smith, M. R., Jousset, A.-C., Despons, L., Laumond, G., Decoville, T., Cattenoz, P., Moog, C., Jossinet, F., Mougel, M., Paillart, J.-C., von Kleist, M., and Marquet, R. (2018) In cell mutational interference mapping experiment (in cell MIME) identifies the 5' polyadenylation signal as a dual regulator of HIV-1 genomic RNA production and packaging. *Nucleic Acids Res.* **46**, e57
327. Brown, J. D., Kharytonchyk, S., Chaudry, I., Iyer, A. S., Carter, H., Becker, G., Desai, Y., Glang, L., Choi, S. H., Singh, K., Lopresti, M. W., Orellana, M., Rodriguez, T., Oboh, U., Hijji, J., Ghinger, F. G., Stewart, K., Francis, D., Edwards, B., Chen, P., Case, D. A., Telesnitsky, A., and Summers, M. F. (2020) Structural basis for transcriptional start site control of HIV-1 RNA fate. *Science.* **368**, 413–417
328. Ding, P., Kharytonchyk, S., Kuo, N., Cannistraci, E., Flores, H., Chaudhary, R., Sarkar, M., Dong, X., Telesnitsky, A., and Summers, M. F. (2021) 5'-Cap sequestration is an essential determinant of HIV-1 genome packaging. *Proc. Natl. Acad. Sci.* **118**, e2112475118
329. Hope, T. J., Richman, D., Stevenson, M., and SpringerLink (2018) *Encyclopedia of AIDS*, 1st ed. 2018., Springer New York, New York, NY, 10.1007/978-1-4939-7101-5
330. Sundquist, W. I., and Kräusslich, H.-G. (2012) HIV-1 Assembly, Budding, and Maturation. *Cold Spring Harb. Perspect. Med.* 10.1101/cshperspect.a006924
331. Sundquist, W. I., and Kräusslich, H.-G. (2012) HIV-1 Assembly, Budding, and Maturation. *Cold Spring Harb. Perspect. Med.* 10.1101/cshperspect.a006924
332. Tedbury, P. R., and Freed, E. O. (2015) HIV-1 Gag: An Emerging Target for Antiretroviral Therapy. *Curr. Top. Microbiol. Immunol.* **389**, 171–201
333. Hellmund, C., and Lever, A. M. L. (2016) Coordination of Genomic RNA Packaging with Viral Assembly in HIV-1. *Viruses.* 10.3390/v8070192
334. Wiegers, K., Rutter, G., Kottler, H., Tessmer, U., Hohenberg, H., and Kräusslich, H. G. (1998) Sequential steps in human immunodeficiency virus particle maturation revealed by alterations of individual Gag polyprotein cleavage sites. *J. Virol.* **72**, 2846–2854
335. Gilboa, E., Mitra, S. W., Goff, S., and Baltimore, D. (1979) A detailed model of reverse transcription and tests of crucial aspects. *Cell.* **18**, 93–100
336. Hu, W. S., and Temin, H. M. (1990) Genetic consequences of packaging two RNA genomes in one retroviral particle: pseudodiploidy and high rate of genetic recombination.

Proc. Natl. Acad. Sci. U. S. A. **87**, 1556–1560

337. Thomas, J. A., and Gorelick, R. J. (2008) Nucleocapsid protein function in early infection processes. *Virus Res.* **134**, 39–63
338. Didierlaurent, L., Houzet, L., Morichaud, Z., Darlix, J.-L., and Mougel, M. (2008) The conserved N-terminal basic residues and zinc-finger motifs of HIV-1 nucleocapsid restrict the viral cDNA synthesis during virus formation and maturation. *Nucleic Acids Res.* **36**, 4745–4753
339. M, M., A, C., and Ji, D. (2010) Implications of the nucleocapsid and the microenvironment in retroviral reverse transcription. *Viruses.* **2**, 939–960
340. Thomas, J. A., Bosche, W. J., Shatzer, T. L., Johnson, D. G., and Gorelick, R. J. (2008) Mutations in Human Immunodeficiency Virus Type 1 Nucleocapsid Protein Zinc Fingers Cause Premature Reverse Transcription. *J. Virol.* **82**, 9318–9328
341. Mougel, M., Houzet, L., and Darlix, J.-L. (2009) When is it time for reverse transcription to start and go? *Retrovirology.* **6**, 24
342. Whitcomb, J. M., Kumar, R., and Hughes, S. H. (1990) Sequence of the circle junction of human immunodeficiency virus type 1: implications for reverse transcription and integration. *J. Virol.* **64**, 4903–4906
343. Trono, D. (1992) Partial reverse transcripts in virions from human immunodeficiency and murine leukemia viruses. *J. Virol.* **66**, 4893–4900
344. Lori, F., di Marzo Veronese, F., de Vico, A. L., Lusso, P., Reitz, M. S., and Gallo, R. C. (1992) Viral DNA carried by human immunodeficiency virus type 1 virions. *J. Virol.* **66**, 5067–5074
345. Zhang, J., and Temin, H. M. (1993) Rate and mechanism of nonhomologous recombination during a single cycle of retroviral replication. *Science.* **259**, 234–238
346. Zhang, H., Dornadula, G., and Pomerantz, R. J. (1998) Natural endogenous reverse transcription of HIV-1. *J. Reprod. Immunol.* **41**, 255–260
347. Hooker, C. W., and Harrich, D. (2003) The first strand transfer reaction of HIV-1 reverse transcription is more efficient in infected cells than in cell-free natural endogenous reverse transcription reactions. *J. Clin. Virol.* **26**, 229–238
348. Warrilow, D., Meredith, L., Davis, A., Burrell, C., Li, P., and Harrich, D. (2008) Cell Factors Stimulate Human Immunodeficiency Virus Type 1 Reverse Transcription In Vitro. *J. Virol.* **82**, 1425–1437
349. Basu, V. P., Song, M., Gao, L., Rigby, S. T., Hanson, M. N., and Bambara, R. A. (2008) Strand transfer events during HIV-1 reverse transcription. *Virus Res.* **134**, 19–38
350. Marquet, R., Isel, C., Ehresmann, C., and Ehresmann, B. (1995) tRNAs as primer of reverse transcriptases. *Biochimie.* **77**, 113–124
351. Kaminska, M., Shalak, V., Francin, M., and Mirande, M. (2007) Viral Hijacking of Mitochondrial Lysyl-tRNA Synthetase. *J. Virol.* **81**, 68–73
352. Pavon-Eternod, M., Wei, M., Pan, T., and Kleiman, L. (2010) Profiling non-lysyl tRNAs in HIV-1. *RNA N. Y. N.* **16**, 267–273
353. Cen, S., Khorchid, A., Javanbakht, H., Gabor, J., Stello, T., Shiba, K., Musier-Forsyth, K., and Kleiman, L. (2001) Incorporation of lysyl-tRNA synthetase into human immunodeficiency virus type 1. *J. Virol.* **75**, 5043–5048
354. Zaitseva, L., Myers, R., and Fassati, A. (2006) tRNAs Promote Nuclear Import of HIV-1 Intracellular Reverse Transcription Complexes. *PLoS Biol.* [10.1371/journal.pbio.0040332](https://doi.org/10.1371/journal.pbio.0040332)
355. Isel, C., Ehresmann, C., and Marquet, R. (2010) Initiation of HIV Reverse Transcription. *Viruses.* **2**, 213–243
356. Seif, E., Niu, M., and Kleiman, L. (2015) In virio SHAPE analysis of tRNA(Lys3) annealing to HIV-1 genomic RNA in wild type and protease-deficient virus. *Retrovirology.*

357. Mak, J., and Kleiman, L. (1997) Primer tRNAs for reverse transcription. *J. Virol.* **71**, 8087–8095
358. Levin, J. G., Mitra, M., Mascarenhas, A., and Musier-Forsyth, K. (2010) Role of HIV-1 nucleocapsid protein in HIV-1 reverse transcription. *RNA Biol.* **7**, 754–774
359. Xing, L., Niu, M., and Kleiman, L. (2012) In Vitro and In Vivo Analysis of the Interaction between RNA Helicase A and HIV-1 RNA. *J. Virol.* **86**, 13272–13280
360. Xing, L., Liang, C., and Kleiman, L. (2011) Coordinate Roles of Gag and RNA Helicase A in Promoting the Annealing of $(\text{tRNA})_{3}^{\text{Lys}}$ to HIV-1 RNA. *J. Virol.* **85**, 1847–1860
361. Brady, S., Singh, G., Bolinger, C., Song, Z., Boeras, I., Weng, K., Trent, B., Brown, W. C., Singh, K., Boris-Lawrie, K., and Heng, X. (2019) Virion-associated, host-derived DHX9/RNA helicase A enhances the processivity of HIV-1 reverse transcriptase on genomic RNA. *J. Biol. Chem.* **294**, 11473–11485
362. Guo, F., Saadatmand, J., Niu, M., and Kleiman, L. (2009) Roles of Gag and NCp7 in facilitating tRNA(Lys)₃ Annealing to viral RNA in human immunodeficiency virus type 1. *J. Virol.* **83**, 8099–8107
363. Feng, Y. X., Campbell, S., Harvin, D., Ehresmann, B., Ehresmann, C., and Rein, A. (1999) The human immunodeficiency virus type 1 Gag polyprotein has nucleic acid chaperone activity: possible role in dimerization of genomic RNA and placement of tRNA on the primer binding site. *J. Virol.* **73**, 4251–4256
364. Guo, F., Saadatmand, J., Niu, M., and Kleiman, L. (2009) Roles of Gag and NCp7 in facilitating tRNA(Lys)₃ Annealing to viral RNA in human immunodeficiency virus type 1. *J. Virol.* **83**, 8099–8107
365. Gremminger, T., Song, Z., Ji, J., Foster, A., Weng, K., and Heng, X. (2020) Extended Interactions between HIV-1 Viral RNA and tRNA_{Lys3} Are Important to Maintain Viral RNA Integrity. *Int. J. Mol. Sci.* **22**, 58
366. Wilkinson, K. A., Gorelick, R. J., Vasa, S. M., Guex, N., Rein, A., Mathews, D. H., Giddings, M. C., and Weeks, K. M. (2008) High-throughput SHAPE analysis reveals structures in HIV-1 genomic RNA strongly conserved across distinct biological states. *PLoS Biol.* **6**, e96
367. Isel, C., Ehresmann, C., Keith, G., Ehresmann, B., and Marquet, R. (1995) Initiation of reverse transcription of HIV-1: secondary structure of the HIV-1 RNA/tRNA(3Lys) (template/primer). *J. Mol. Biol.* **247**, 236–250
368. Barat, C., Lullien, V., Schatz, O., Keith, G., Nugeyre, M. t., Grüniger-Leitch, F., Barré-Sinoussi, F., LeGrice, S. f., and Darlix, J. l. (1989) HIV-1 reverse transcriptase specifically interacts with the anticodon domain of its cognate primer tRNA. *EMBO J.* **8**, 3279–3285
369. Ooms, M., Cupac, D., Abbink, T. E. M., Huthoff, H., and Berkhout, B. (2007) The availability of the primer activation signal (PAS) affects the efficiency of HIV-1 reverse transcription initiation. *Nucleic Acids Res.* **35**, 1649–1659
370. Huthoff, H., Bugala, K., Barciszewski, J., and Berkhout, B. (2003) On the importance of the primer activation signal for initiation of tRNA(lys3)-primed reverse transcription of the HIV-1 RNA genome. *Nucleic Acids Res.* **31**, 5186–5194
371. Song, Z., Gremminger, T., Singh, G., Cheng, Y., Li, J., Qiu, L., Ji, J., Lange, M. J., Zuo, X., Chen, S.-J., Zou, X., Boris-Lawrie, K., and Heng, X. (2021) The three-way junction structure of the HIV-1 PBS-segment binds host enzyme important for viral infectivity. *Nucleic Acids Res.* **49**, 5925–5942
372. Jones, C. P., Saadatmand, J., Kleiman, L., and Musier-Forsyth, K. (2013) Molecular mimicry of human tRNA_{Lys} anti-codon domain by HIV-1 RNA genome facilitates tRNA

primer annealing. *RNA*. **19**, 219–229

373. Cen, S., Javanbakht, H., Niu, M., and Kleiman, L. (2004) Ability of Wild-Type and Mutant Lysyl-tRNA Synthetase To Facilitate tRNA^{Lys} Incorporation into Human Immunodeficiency Virus Type 1. *J. Virol.* **78**, 1595–1601
374. Larsen, K. P., Mathiharan, Y. K., Kappel, K., Coey, A. T., Chen, D.-H., Barrero, D., Madigan, L., Puglisi, J. D., Skinotis, G., and Puglisi, E. V. (2018) Architecture of an HIV-1 reverse transcriptase initiation complex. *Nature*. **557**, 118–122
375. Das, K., Martinez, S. E., DeStefano, J. J., and Arnold, E. (2019) Structure of HIV-1 RT/dsRNA initiation complex prior to nucleotide incorporation. *Proc. Natl. Acad. Sci.* **116**, 7308–7313
376. Xavier RUIZ, F., and Arnold, E. (2020) Evolving understanding of HIV-1 reverse transcriptase structure, function, inhibition, and resistance. *Curr. Opin. Struct. Biol.* **61**, 113–123
377. Isel, C., Lanchy, J. M., Le Grice, S. F., Ehresmann, C., Ehresmann, B., and Marquet, R. (1996) Specific initiation and switch to elongation of human immunodeficiency virus type 1 reverse transcription require the post-transcriptional modifications of primer tRNA^{3Lys}. *EMBO J.* **15**, 917–924
378. Liu, S., Harada, B. T., Miller, J. T., Le Grice, S. F. J., and Zhuang, X. (2010) Initiation complex dynamics direct the transitions between distinct phases of early HIV reverse transcription. *Nat. Struct. Mol. Biol.* **17**, 1453–1460
379. Sleiman, D., Goldschmidt, V., Barraud, P., Marquet, R., Paillart, J.-C., and Tisné, C. (2012) Initiation of HIV-1 reverse transcription and functional role of nucleocapsid-mediated tRNA/viral genome interactions. *Virus Res.* **169**, 324–339
380. Isel, C., Marquet, R., Keith, G., Ehresmann, C., and Ehresmann, B. (1993) Modified nucleotides of tRNA^(3Lys) modulate primer/template loop-loop interaction in the initiation complex of HIV-1 reverse transcription. *J. Biol. Chem.* **268**, 25269–25272
381. Goldschmidt, V., Rigourd, M., Ehresmann, C., Le Grice, S. F. J., Ehresmann, B., and Marquet, R. (2002) Direct and indirect contributions of RNA secondary structure elements to the initiation of HIV-1 reverse transcription. *J. Biol. Chem.* **277**, 43233–43242
382. Isel, C., Lanchy, J. M., Le Grice, S. F., Ehresmann, C., Ehresmann, B., and Marquet, R. (1996) Specific initiation and switch to elongation of human immunodeficiency virus type 1 reverse transcription require the post-transcriptional modifications of primer tRNA^{3Lys}. *EMBO J.* **15**, 917–924
383. Coffin, J. M., and Haseltine, W. A. (1977) Terminal redundancy and the origin of replication of Rous sarcoma virus RNA. *Proc. Natl. Acad. Sci. U. S. A.* **74**, 1908–1912
384. Chen, Y., Maskri, O., Chaminade, F., René, B., Benkaroun, J., Godet, J., Mély, Y., Mauffret, O., and Fossé, P. (2016) Structural Insights into the HIV-1 Minus-strand Strong-stop DNA. *J. Biol. Chem.* **291**, 3468–3482
385. Masuda, T., Sato, Y., Huang, Y.-L., Koi, S., Takahata, T., Hasegawa, A., Kawai, G., and Kannagi, M. (2015) Fate of HIV-1 cDNA intermediates during reverse transcription is dictated by transcription initiation site of virus genomic RNA. *Sci. Rep.* **5**, 17680
386. Berkhout, B., Klaver, B., and Das, A. T. (1995) A conserved hairpin structure predicted for the poly(A) signal of human and simian immunodeficiency viruses. *Virology*. **207**, 276–281
387. Feng, S., and Holland, E. C. (1988) HIV-1 tat trans-activation requires the loop sequence within tar. *Nature*. **334**, 165–167
388. Kanevsky, I., Chaminade, F., Ficheux, D., Moumen, A., Gorelick, R., Negroni, M., Darlix, J.-L., and Fossé, P. (2005) Specific interactions between HIV-1 nucleocapsid protein and the TAR element. *J. Mol. Biol.* **348**, 1059–1077
389. 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
390. Berkhout, B., Vastenhouw, N. L., Klasens, B. I., and Huthoff, H. (2001) Structural features in the HIV-1 repeat region facilitate strand transfer during reverse transcription. *RNA*. **7**, 1097–1114
391. 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
392. Arts, E. J., Mak, J., Kleiman, L., and Wainberg, M. A. (1994) Mature reverse transcriptase (p66/p51) is responsible for low levels of viral DNA found in human immunodeficiency virus type 1 (HIV-1). *Leukemia*. **8 Suppl 1**, S175-178
393. Arts, E. J., Mak, J., Kleiman, L., and Wainberg, M. A. (1994) DNA found in human immunodeficiency virus type 1 particles may not be required for infectivity. *J. Gen. Virol.* **75 (Pt 7)**, 1605–1613
394. Zhang, H., Zhang, Y., Spicer, T. P., Abbott, L. Z., Abbott, M., and Poiesz, B. J. (1993) Reverse transcription takes place within extracellular HIV-1 virions: potential biological significance. *AIDS Res. Hum. Retroviruses*. **9**, 1287–1296
395. Anthony, R. M., and Destefano, J. J. (2007) In vitro synthesis of long DNA products in reactions with HIV-RT and nucleocapsid protein. *J. Mol. Biol.* **365**, 310–324
396. Lanciault, C., and Champoux, J. J. (2005) Effects of unpaired nucleotides within HIV-1 genomic secondary structures on pausing and strand transfer. *J. Biol. Chem.* **280**, 2413–2423
397. Schultz, S. J., and Champoux, J. J. (2008) RNase H Activity: Structure, Specificity, and Function in Reverse Transcription. *Virus Res.* **134**, 86–103
398. Pollom, E., Dang, K. K., Potter, E. L., Gorelick, R. J., Burch, C. L., Weeks, K. M., and Swanstrom, R. (2013) Comparison of SIV and HIV-1 Genomic RNA Structures Reveals Impact of Sequence Evolution on Conserved and Non-Conserved Structural Motifs. *PLoS Pathog.* **9**, e1003294
399. Fabris, D., Marino, J. P., and Le Grice, S. F. J. (2009) Revisiting Plus-Strand DNA Synthesis in Retroviruses and Long Terminal Repeat Retrotransposons: Dynamics of Enzyme: Substrate Interactions. *Viruses*. **1**, 657–677
400. 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
401. Huber, H. E., and Richardson, C. C. (1990) Processing of the primer for plus strand DNA synthesis by human immunodeficiency virus 1 reverse transcriptase. *J. Biol. Chem.* **265**, 10565–10573
402. Figiel, M., Krepl, M., Park, S., Poznański, J., Skowronek, K., Gołąb, A., Ha, T., Šponer, J., and Nowotny, M. (2018) Mechanism of polypurine tract primer generation by HIV-1 reverse transcriptase. *J. Biol. Chem.* **293**, 191–202
403. Renda, M. J., Rosenblatt, J. D., Klimatcheva, E., Demeter, L. M., Bambara, R. A., and Planelles, V. (2001) Mutation of the Methylated tRNA\documentclass[12pt]{minimal} \usepackage{amsmath} \usepackage{wasysym} \usepackage{amsfonts} \usepackage{amssymb} \usepackage{amsbsy} \usepackage{mathrsfs} \setlength{\oddsidemargin}{-69pt} \begin{document} \begin{equation*} {\mathrm{_{3}^{\mathrm{Lys}}}} \end{equation*} \end{document} Residue A58 Disrupts Reverse Transcription and Inhibits Replication of Human Immunodeficiency Virus Type 1. *J. Virol.* **75**, 9671–9678
404. Marquet, R., Isel, C., Ehresmann, C., and Ehresmann, B. (1995) tRNAs as primer of reverse transcriptases. *Biochimie*. **77**, 113–124

405. Yu, H., Jetzt, A. E., Ron, Y., Preston, B. D., and Dougherty, J. P. (1998) The nature of human immunodeficiency virus type 1 strand transfers. *J. Biol. Chem.* **273**, 28384–28391
406. Cappy, P., Moisan, A., De Oliveira, F., Plantier, J.-C., and Negroni, M. (2017) HIV-1 sequences in the epidemic suggest an alternative pathway for the generation of the Long Terminal Repeats. *Sci. Rep.* **7**, 13715
407. Guo, J., Wu, T., Anderson, J., Kane, B. F., Johnson, D. G., Gorelick, R. J., Henderson, L. E., and Levin, J. G. (2000) Zinc finger structures in the human immunodeficiency virus type 1 nucleocapsid protein facilitate efficient minus- and plus-strand transfer. *J. Virol.* **74**, 8980–8988
408. Wu, T., Guo, J., Bess, J., Henderson, L. E., and Levin, J. G. (1999) Molecular Requirements for Human Immunodeficiency Virus Type 1 Plus-Strand Transfer: Analysis in Reconstituted and Endogenous Reverse Transcription Systems. *J. Virol.* **73**, 4794–4805
409. Wu, T., Guo, J., Bess, J., Henderson, L. E., and Levin, J. G. (1999) Molecular Requirements for Human Immunodeficiency Virus Type 1 Plus-Strand Transfer: Analysis in Reconstituted and Endogenous Reverse Transcription Systems. *J. Virol.* **73**, 4794–4805
410. Muthuswami, R., Chen, J., Burnett, B. P., Thimmig, R. L., Janjic, N., and McHenry, C. S. (2002) The HIV plus-strand transfer reaction: determination of replication-competent intermediates and identification of a novel lentiviral element, the primer over-extension sequence. *J. Mol. Biol.* **315**, 311–323
411. Johnson, P. E., Turner, R. B., Wu, Z. R., Hairston, L., Guo, J., Levin, J. G., and Summers, M. F. (2000) A mechanism for plus-strand transfer enhancement by the HIV-1 nucleocapsid protein during reverse transcription. *Biochemistry.* **39**, 9084–9091
412. Egelé, C., Schaub, E., Ramalanjaona, N., Piémont, E., Ficheux, D., Roques, B., Darlix, J.-L., and Mély, Y. (2004) HIV-1 nucleocapsid protein binds to the viral DNA initiation sequences and chaperones their kissing interactions. *J. Mol. Biol.* **342**, 453–466
413. 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
414. Lesbats, P., Engelmann, A. N., and Cherepanov, P. (2016) Retroviral DNA Integration. *Chem. Rev.* **116**, 12730–12757
415. Wurtzer, S., Goubard, A., Mammano, F., Saragosti, S., Lecossier, D., Hance, A. J., and Clavel, F. (2006) Functional central polypurine tract provides downstream protection of the human immunodeficiency virus type 1 genome from editing by APOBEC3G and APOBEC3B. *J. Virol.* **80**, 3679–3683
416. 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
417. Hameau, L., Jeusset, J., Lafosse, S., Coulaud, D., Delain, E., Unge, T., Restle, T., Le Cam, E., and Mirambeau, G. (2001) Human immunodeficiency virus type 1 central DNA flap: dynamic terminal product of plus-strand displacement dna synthesis catalyzed by reverse transcriptase assisted by nucleocapsid protein. *J. Virol.* **75**, 3301–3313
418. Zhang, X.-J., Wang, R.-R., Chen, H., Luo, R.-H., Yang, L.-M., Liu, J.-P., Sun, H.-D., Zhang, H.-B., Xiao, W.-L., and Zheng, Y.-T. (2018) SJP-L-5 inhibits HIV-1 polypurine tract primed plus-strand DNA elongation, indicating viral DNA synthesis initiation at multiple sites under drug pressure. *Sci. Rep.* **8**, 2574
419. Hsu, M., and Wainberg, M. A. (2000) Interactions between human immunodeficiency virus type 1 reverse transcriptase, tRNA primer, and nucleocapsid protein during reverse transcription. *J. Hum. Virol.* **3**, 16–26
420. Simon-Loriere, E., Rossolillo, P., and Negroni, M. (2011) RNA structures, genomic organization and selection of recombinant HIV. *RNA Biol.* **8**, 280–286

421. 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
422. Panganiban, A. T., and Fiore, D. (1988) Ordered interstrand and intrastrand DNA transfer during reverse transcription. *Science.* **241**, 1064–1069
423. Taylor, B. S., Sobieszczyk, M. E., McCutchan, F. E., and Hammer, S. M. (2008) The challenge of HIV-1 subtype diversity. *N. Engl. J. Med.* **358**, 1590–1602
424. Streeck, H., Li, B., Poon, A. F. Y., Schneidewind, A., Gladden, A. D., Power, K. A., Daskalakis, D., Bazner, S., Zuniga, R., Brander, C., Rosenberg, E. S., Frost, S. D. W., Altfeld, M., and Allen, T. M. (2008) Immune-driven recombination and loss of control after HIV superinfection. *J. Exp. Med.* **205**, 1789–1796
425. Moutouh, L., Corbeil, J., and Richman, D. D. (1996) Recombination leads to the rapid emergence of HIV-1 dually resistant mutants under selective drug pressure. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 6106–6111
426. Sakuragi, S., Shioda, T., and Sakuragi, J. (2015) Properties of human immunodeficiency virus type 1 reverse transcriptase recombination upon infection. *J. Gen. Virol.* **96**, 3382–3388
427. Malim, M. H., and Emerman, M. (2001) HIV-1 sequence variation: drift, shift, and attenuation. *Cell.* **104**, 469–472
428. Derebail, S. S., Heath, M. J., and DeStefano, J. J. (2003) Evidence for the differential effects of nucleocapsid protein on strand transfer in various regions of the HIV genome. *J. Biol. Chem.* **278**, 15702–15712
429. Darlix, J. L., Vincent, A., Gabus, C., de Rocquigny, H., and Roques, B. (1993) Trans-activation of the 5' to 3' viral DNA strand transfer by nucleocapsid protein during reverse transcription of HIV1 RNA. *C. R. Acad. Sci. III.* **316**, 763–771
430. Guo, J., Henderson, L. E., Bess, J., Kane, B., and Levin, J. G. (1997) Human immunodeficiency virus type 1 nucleocapsid protein promotes efficient strand transfer and specific viral DNA synthesis by inhibiting TAR-dependent self-priming from minus-strand strong-stop DNA. *J. Virol.* **71**, 5178–5188
431. Henriot, S., Sinck, L., Bec, G., Gorelick, R. J., Marquet, R., and Paillart, J.-C. (2007) Vif is a RNA chaperone that could temporally regulate RNA dimerization and the early steps of HIV-1 reverse transcription. *Nucleic Acids Res.* **35**, 5141–5153
432. Post, K., Olson, E. D., Naufer, M. N., Gorelick, R. J., Rouzina, I., Williams, M. C., Musier-Forsyth, K., and Levin, J. G. (2016) Mechanistic differences between HIV-1 and SIV nucleocapsid proteins and cross-species HIV-1 genomic RNA recognition. *Retrovirology.* **13**, 89
433. Negroni, M., and Buc, H. (2000) Copy-choice recombination by reverse transcriptases: Reshuffling of genetic markers mediated by RNA chaperones. *Proc. Natl. Acad. Sci.* **97**, 6385–6390
434. Moumen, A., Polomack, L., Unge, T., Véron, M., Buc, H., and Negroni, M. (2003) Evidence for a mechanism of recombination during reverse transcription dependent on the structure of the acceptor RNA. *J. Biol. Chem.* **278**, 15973–15982
435. Hanson, M. N., Balakrishnan, M., Roques, B. P., and Bambara, R. A. (2005) Effects of donor and acceptor RNA structures on the mechanism of strand transfer by HIV-1 reverse transcriptase. *J. Mol. Biol.* **353**, 772–787
436. Cromer, D., Grimm, A. J., Schlub, T. E., Mak, J., and Davenport, M. P. (2016) Estimating the in-vivo HIV template switching and recombination rate. *AIDS Lond. Engl.* **30**, 185–192
437. Zhang, J., and Temin, H. M. (1994) Retrovirus recombination depends on the length of sequence identity and is not error prone. *J. Virol.* **68**, 2409–2414

438. Pathak, V. K., and Hu, W.-S. (1997) “Might as Well Jump!” Template Switching by Retroviral Reverse Transcriptase, Defective Genome Formation, and Recombination. *Semin. Virol.* **8**, 141–150
439. Rawson, J. M. O., Nikolaitchik, O. A., Keele, B. F., Pathak, V. K., and Hu, W.-S. (2018) Recombination is required for efficient HIV-1 replication and the maintenance of viral genome integrity. *Nucleic Acids Res.* **46**, 10535–10545
440. Coffin, J. M. (1979) Structure, replication, and recombination of retrovirus genomes: some unifying hypotheses. *J. Gen. Virol.* **42**, 1–26
441. Negroni, M., Ricchetti, M., Nouvel, P., and Buc, H. (1995) Homologous recombination promoted by reverse transcriptase during copying of two distinct RNA templates. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 6971–6975
442. Svarovskaia, E. S., Delviks, K. A., Hwang, C. K., and Pathak, V. K. (2000) Structural determinants of murine leukemia virus reverse transcriptase that affect the frequency of template switching. *J. Virol.* **74**, 7171–7178
443. Hwang, C. K., Svarovskaia, E. S., and Pathak, V. K. (2001) Dynamic copy choice: Steady state between murine leukemia virus polymerase and polymerase-dependent RNase H activity determines frequency of in vivo template switching. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 12209–12214
444. Chen, Y., Balakrishnan, M., Roques, B. P., and Bambara, R. A. (2005) Acceptor RNA cleavage profile supports an invasion mechanism for HIV-1 minus strand transfer. *J. Biol. Chem.* **280**, 14443–14452
445. René, B., Mauffret, O., and Fossé, P. (2018) Retroviral nucleocapsid proteins and DNA strand transfers. *Biochim. Open.* **7**, 10–25
446. Shen, W., Gao, L., Balakrishnan, M., and Bambara, R. A. (2009) A recombination hot spot in HIV-1 contains guanosine runs that can form a G-quartet structure and promote strand transfer in vitro. *J. Biol. Chem.* **284**, 33883–33893
447. Shen, W., Gorelick, R. J., and Bambara, R. A. (2011) HIV-1 nucleocapsid protein increases strand transfer recombination by promoting dimeric G-quartet formation. *J. Biol. Chem.* **286**, 29838–29847
448. Piekna-Przybylska, D., Sharma, G., and Bambara, R. A. (2013) Mechanism of HIV-1 RNA Dimerization in the Central Region of the Genome and Significance for Viral Evolution *. *J. Biol. Chem.* **288**, 24140–24150
449. Bampi, C., Jacquenet, S., Lener, D., Décimo, D., and Darlix, J.-L. (2004) The chaperoning and assistance roles of the HIV-1 nucleocapsid protein in proviral DNA synthesis and maintenance. *Int. J. Biochem. Cell Biol.* **36**, 1668–1686
450. Andersen, E. S., Jeeninga, R. E., Damgaard, C. K., Berkhout, B., and Kjems, J. (2003) Dimerization and Template Switching in the 5' Untranslated Region between Various Subtypes of Human Immunodeficiency Virus Type 1. *J. Virol.* **77**, 3020–3030
451. Balakrishnan, M., Roques, B. P., Fay, P. J., and Bambara, R. A. (2003) Template Dimerization Promotes an Acceptor Invasion-Induced Transfer Mechanism during Human Immunodeficiency Virus Type 1 Minus-Strand Synthesis. *J. Virol.* **77**, 4710–4721
452. Santos, S., Obukhov, Y., Nekhai, S., Bukrinsky, M., and Iordanskiy, S. (2012) Virus-producing cells determine the host protein profiles of HIV-1 virion cores. *Retrovirology.* **9**, 65
453. Warren, K., Warrilow, D., Meredith, L., and Harrich, D. (2009) Reverse Transcriptase and Cellular Factors: Regulators of HIV-1 Reverse Transcription. *Viruses.* **1**, 873–894
454. Lee, C. G., and Hurwitz, J. (1992) A new RNA helicase isolated from HeLa cells that catalytically translocates in the 3' to 5' direction. *J. Biol. Chem.* **267**, 4398–4407
455. Zhang, S., and Grosse, F. (1997) Domain structure of human nuclear DNA helicase II (RNA helicase A). *J. Biol. Chem.* **272**, 11487–11494
456. Zhang, Z., Yuan, B., Lu, N., Facchinetti, V., and Liu, Y.-J. (2011) DHX9 pairs with

- IPS-1 to sense double-stranded RNA in myeloid dendritic cells. *J. Immunol. Baltim. Md* 1950. **187**, 4501–4508
457. Bolinger, C., Sharma, A., Singh, D., Yu, L., and Boris-Lawrie, K. (2010) RNA helicase A modulates translation of HIV-1 and infectivity of progeny virions. *Nucleic Acids Res.* **38**, 1686–1696
458. Xing, L., Liang, C., and Kleiman, L. (2011) Coordinate Roles of Gag and RNA Helicase A in Promoting the Annealing of $(\text{tRNA})_{3}^{\text{Lys}}$ to HIV-1 RNA. *J. Virol.* **85**, 1847–1860
459. Brady, S., Singh, G., Bolinger, C., Song, Z., Boeras, I., Weng, K., Trent, B., Brown, W. C., Singh, K., Boris-Lawrie, K., and Heng, X. (2019) Virion-associated, host-derived DHX9/RNA helicase A enhances the processivity of HIV-1 reverse transcriptase on genomic RNA. *J. Biol. Chem.* **294**, 11473–11485
460. Fujii, R., Okamoto, M., Aratani, S., Oishi, T., Ohshima, T., Taira, K., Baba, M., Fukamizu, A., and Nakajima, T. (2001) A Role of RNA Helicase A in cis-Acting Transactivation Response Element-mediated Transcriptional Regulation of Human Immunodeficiency Virus Type 1. *J. Biol. Chem.* **276**, 5445–5451
461. Li, J., Tang, H., Mullen, T. M., Westberg, C., Reddy, T. R., Rose, D. W., and Wong-Staal, F. (1999) A role for RNA helicase A in post-transcriptional regulation of HIV type 1. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 709–714
462. Boeras, I., Song, Z., Moran, A., Franklin, J., Brown, W. C., Johnson, M., Boris-Lawrie, K., and Heng, X. (2016) DHX9/RHA Binding to the PBS-Segment of the Genomic RNA during HIV-1 Assembly Bolsters Virion Infectivity. *J. Mol. Biol.* **428**, 2418–2429
463. Phanuphak, N., and Gulick, R. M. (2020) HIV treatment and prevention 2019: current standards of care. *Curr. Opin. HIV AIDS.* **15**, 4–12
464. Holec, A. D., Mandal, S., Prathipati, P. K., and Destache, C. J. (2017) Nucleotide Reverse Transcriptase Inhibitors: A Thorough Review, Present Status and Future Perspective as HIV Therapeutics. *Curr. HIV Res.* **15**, 411–421
465. Sluis-Cremer, N., and Tachedjian, G. (2008) Mechanisms of inhibition of HIV replication by nonnucleoside reverse transcriptase inhibitors. *Virus Res.* **134**, 147–156
466. Merck Sharp & Dohme Corp. (2020) *A Phase 2a, Double-Blind, Placebo-Controlled Study to Evaluate the Safety, Tolerability, and Pharmacokinetics of Oral MK-8591 Once-Monthly in Participants at Low-Risk for HIV-1 Infection*, clinicaltrials.gov, [online] <https://clinicaltrials.gov/ct2/show/NCT04003103> (Accessed August 22, 2021)
467. Merck Sharp & Dohme Corp. (2019) *A Single-Dose Clinical Trial to Study the Safety, Tolerability, Pharmacokinetics, and Anti-Retroviral Activity of MK-8591 Monotherapy in Anti-Retroviral Therapy (ART)-Naive, HIV-1 Infected Patients*, clinicaltrials.gov, [online] <https://clinicaltrials.gov/ct2/show/NCT02217904> (Accessed August 23, 2021)
468. Cambou, M. C., and Landovitz, R. J. (2020) Novel Antiretroviral Agents. *Curr. HIV/AIDS Rep.* **17**, 118–124
469. Cilento, M. E., Kirby, K. A., Tedbury, P. R., and Sarafianos, S. G. (2021) Proteins and Disease | Structural Basis of HIV Reverse Transcription, Inhibition, and Drug Resistance. in *Encyclopedia of Biological Chemistry III (Third Edition)* (Jez, J. ed), pp. 92–104, Elsevier, Oxford, 10.1016/B978-0-12-809633-8.21570-7
470. Zhang, Z., Walker, M., Xu, W., Shim, J. H., Girardet, J.-L., Hamatake, R. K., and Hong, Z. (2006) Novel nonnucleoside inhibitors that select nucleoside inhibitor resistance mutations in human immunodeficiency virus type 1 reverse transcriptase. *Antimicrob. Agents Chemother.* **50**, 2772–2781
471. Ruiz, F. X., Hoang, A., Das, K., and Arnold, E. (2019) Structural Basis of HIV-1 Inhibition by Nucleotide-Competing Reverse Transcriptase Inhibitor INDOPY-1. *J. Med. Chem.* **62**, 9996–10002

472. Wang, L., Sarafianos, S. G., and Wang, Z. (2020) Cutting into the Substrate Dominance: Pharmacophore and Structure-Based Approaches toward Inhibiting HIV Reverse Transcriptase-Associated Ribonuclease H. *Acc. Chem. Res.* **53**, 218–230
473. Nowotny, M. (2009) Retroviral integrase superfamily: the structural perspective. *EMBO Rep.* **10**, 144–151
474. Kang, D., Urhan, Ç., Wei, F., Frutos-Beltrán, E., Sun, L., Álvarez, M., Feng, D., Tao, Y., Pannecouque, C., De Clercq, E., Menéndez-Arias, L., Liu, X., and Zhan, P. (2021) Discovery, optimization, and target identification of novel coumarin derivatives as HIV-1 reverse transcriptase-associated ribonuclease H inhibitors. *Eur. J. Med. Chem.* **225**, 113769
475. 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
476. 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
477. Hergott, C. B., Mitra, M., Guo, J., Wu, T., Miller, J. T., Iwatani, Y., Gorelick, R. J., and Levin, J. G. (2013) Zinc finger function of HIV-1 nucleocapsid protein is required for removal of 5'-terminal genomic RNA fragments: A paradigm for RNA removal reactions in HIV-1 reverse transcription. *Virus Res.* **171**, 346–355
478. de Rocquigny, H., Shvadchak, V., Avilov, S., Dong, C. Z., Dietrich, U., Darlix, J.-L., and Mély, Y. (2008) Targeting the viral nucleocapsid protein in anti-HIV-1 therapy. *Mini Rev. Med. Chem.* **8**, 24–35
479. Tedbury, P. R., and Freed, E. O. (2015) HIV-1 Gag: An Emerging Target for Antiretroviral Therapy. *Curr. Top. Microbiol. Immunol.* **389**, 171–201
480. 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. *Future HIV-1 Ther.* **389**, 53–92
481. Dick, A., and Cocklin, S. (2020) Recent Advances in HIV-1 Gag Inhibitor Design and Development. *Molecules.* **25**, 1687
482. Sancineto, L., Iraci, N., Tabarrini, O., and Santi, C. (2018) NCp7: targeting a multitasking protein for next-generation anti-HIV drug development part 1: covalent inhibitors. *Drug Discov. Today.* **23**, 260–271
483. Iraci, N., Tabarrini, O., Santi, C., and Sancineto, L. (2018) NCp7: targeting a multitask protein for next-generation anti-HIV drug development part 2. Noncovalent inhibitors and nucleic acid binders. *Drug Discov. Today.* **23**, 687–695
484. 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
485. Vercruyse, T., Basta, B., Dehaen, W., Humbert, N., Balzarini, J., Debaene, F., Sanglier-Cianfèrani, S., Pannecouque, C., Mély, Y., and Daelemans, D. (2012) A phenylthiadiazolylidene-amine derivative ejects zinc from retroviral nucleocapsid zinc fingers and inactivates HIV virions. *Retrovirology.* **9**, 95
486. Helmold Hait, S., Hogge, C. J., Rahman, M. A., Ko, E.-J., Hunegnaw, R., Mushtaq, Z., Enyindah-Asonye, G., Hoang, T., Miller Jenkins, L. M., Appella, E., Appella, D. H., and Robert-Guroff, M. (2020) An SAMT-247 Microbicide Provides Potent Protection against Intravaginal Simian Immunodeficiency Virus Infection of Rhesus Macaques, whereas an Added Vaccine Component Elicits Mixed Outcomes. *J. Immunol. Baltim. Md 1950.* **204**, 3315–3328

487. Mori, M., Nucci, A., Lang, M. C. D., Humbert, N., Boudier, C., Debaene, F., Sanglier-Cianferani, S., Catala, M., Schult-Dietrich, P., Dietrich, U., Tisné, C., Mely, Y., and Botta, M. (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
488. 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
489. Wang, R., Wei, Y., Wang, M., Yan, P., Jiang, H., and Du, Z. (2021) Interaction of Natural Compounds in Licorice and Turmeric with HIV-NCp7 Zinc Finger Domain: Potential Relevance to the Mechanism of Antiviral Activity. *Molecules*. **26**, 3563
490. Mori, M., Kovalenko, L., Malancona, S., Saladini, F., De Forni, D., Pires, M., Humbert, N., Real, E., Botzanowski, T., Cianférani, S., Giannini, A., Dasso Lang, M. C., Cugia, G., Poddesu, B., Lori, F., Zazzi, M., Harper, S., Summa, V., Mely, Y., and Botta, M. (2018) Structure-Based Identification of HIV-1 Nucleocapsid Protein Inhibitors Active against Wild-Type and Drug-Resistant HIV-1 Strains. *ACS Chem. Biol.* **13**, 253–266
491. Malancona, S., Mori, M., Fezzardi, P., Santoriello, M., Basta, A., Nibbio, M., Kovalenko, L., Speziale, R., Battista, M. R., Cellucci, A., Gennari, N., Monteagudo, E., Di Marco, A., Giannini, A., Sharma, R., Pires, M., Real, E., Zazzi, M., Dasso Lang, M. C., De Forni, D., Saladini, F., Mely, Y., Summa, V., Harper, S., and Botta, M. (2020) 5,6-Dihydropyrimidine Scaffold to Target HIV-1 Nucleocapsid Protein. *ACS Med. Chem. Lett.* **11**, 766–772
492. Mori, M., Ciaco, S., Mély, Y., and Karioti, A. (2020) Inhibitory Effect of Lithospermic Acid on the HIV-1 Nucleocapsid Protein. *Molecules*. **25**, 5434
493. Mori, M., Dasso Lang, M. C., Saladini, F., Palombi, N., Kovalenko, L., De Forni, D., Poddesu, B., Friggeri, L., Giannini, A., Malancona, S., Summa, V., Zazzi, M., Mely, Y., and Botta, M. (2018) Synthesis and Evaluation of Bifunctional Aminothiazoles as Antiretrovirals Targeting the HIV-1 Nucleocapsid Protein. *ACS Med. Chem. Lett.* **10**, 463–468
494. Humbert, N., Kovalenko, L., Saladini, F., Giannini, A., Pires, M., Botzanowski, T., Cherenok, S., Boudier, C., Sharma, K. K., Real, E., Zaporozhets, O. A., Cianférani, S., Seguin-Devaux, C., Poggialini, F., Botta, M., Zazzi, M., Kalchenko, V. I., Mori, M., and Mély, Y. (2020) (Thia)calixarenephosphonic Acids as Potent Inhibitors of the Nucleic Acid Chaperone Activity of the HIV-1 Nucleocapsid Protein with a New Binding Mode and Multitarget Antiviral Activity. *ACS Infect. Dis.* **6**, 687–702
495. Gamba, E., Susic, A., Saccone, I., Magli, E., Frecentese, F., and Gatto, B. (2020) Multiple in Vitro Inhibition of HIV-1 Proteins by 2,6-Dipeptidyl-anthraquinone Conjugates Targeting the PBS RNA. *ACS Med. Chem. Lett.* **11**, 949–955
496. Susic, 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
497. Susic, A., Olivato, G., Carraro, C., Göttlich, R., Fabris, D., and Gatto, B. (2021) Bis-3-Chloropiperidines Targeting TAR RNA as A Novel Strategy to Impair the HIV-1 Nucleocapsid Protein. *Molecules*. **26**, 1874
498. Susic, A., Göttlich, R., Fabris, D., and Gatto, B. (2021) B-CePs as cross-linking probes for the investigation of RNA higher-order structure. *Nucleic Acids Res.* **49**, 6660–6672
499. Yu, H., Jetzt, A. E., Ron, Y., Preston, B. D., and Dougherty, J. P. (1998) The nature of human immunodeficiency virus type 1 strand transfers. *J. Biol. Chem.* **273**, 28384–28391
500. van Wamel, J. L., and Berkhout, B. (1998) The first strand transfer during HIV-1 reverse transcription can occur either intramolecularly or intermolecularly. *Virology*. **244**, 245–251
501. Sodroski, J., Trus, M., Perkins, D., Patarca, R., Wong-Staal, F., Gelmann, E., Gallo,

- R., and Haseltine, W. A. (1984) Repetitive structure in the long-terminal-repeat element of a type II human T-cell leukemia virus. *Proc. Natl. Acad. Sci. U. S. A.* **81**, 4617–4621
502. Klaver, B., and Berkhout, B. (1994) Premature strand transfer by the HIV-1 reverse transcriptase during strong-stop DNA synthesis. *Nucleic Acids Res.* **22**, 137–144
503. Ohi, Y., and Clever, J. L. (2000) Sequences in the 5' and 3' R elements of human immunodeficiency virus type 1 critical for efficient reverse transcription. *J. Virol.* **74**, 8324–8334
504. Klaver, B., and Berkhout, B. (1994) Premature strand transfer by the HIV-1 reverse transcriptase during strong-stop DNA synthesis. *Nucleic Acids Res.* **22**, 137–144
505. Berkhout, B., Vastenhouw, N. L., Klasens, B. I., and Huthoff, H. (2001) Structural features in the HIV-1 repeat region facilitate strand transfer during reverse transcription. *RNA.* **7**, 1097–1114
506. Moumen, A., Polomack, L., Roques, B., Buc, H., and Negroni, M. (2001) The HIV-1 repeated sequence R as a robust hot-spot for copy-choice recombination. *Nucleic Acids Res.* **29**, 3814–3821
507. 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. U. S. A.* **111**, 3395–3400
508. 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
509. Beerens, N., and Kjems, J. (2010) Circularization of the HIV-1 genome facilitates strand transfer during reverse transcription. *RNA.* **16**, 1226–1235
510. Brulé, F., Bec, G., Keith, G., Le Grice, S. F., Roques, B. P., Ehresmann, B., Ehresmann, C., and Marquet, R. (2000) In vitro evidence for the interaction of tRNA(3)(Lys) with U3 during the first strand transfer of HIV-1 reverse transcription. *Nucleic Acids Res.* **28**, 634–640
511. Piekna-Przybylska, D., DiChiacchio, L., Mathews, D. H., and Bambara, R. A. (2010) A sequence similar to tRNA 3 Lys gene is embedded in HIV-1 U3-R and promotes minus-strand transfer. *Nat. Struct. Mol. Biol.* **17**, 83–89
512. Piekna-Przybylska, D., Dykes, C., Demeter, L. M., and Bambara, R. A. (2011) Sequences in the U3 region of Human Immunodeficiency Virus 1 improve efficiency of minus strand transfer in infected cells. *Virology.* **410**, 368–374
513. 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
514. Liu, H.-W., Zeng, Y., Landes, C. F., Kim, Y. J., Zhu, Y., Ma, X., Vo, M.-N., Musier-Forsyth, K., and Barbara, P. F. (2007) Insights on the role of nucleic acid/protein interactions in chaperoned nucleic acid rearrangements of HIV-1 reverse transcription. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 5261–5267
515. Song, M., Balakrishnan, M., Chen, Y., Roques, B. P., and Bambara, R. A. (2006) Stimulation of HIV-1 minus strand strong stop DNA transfer by genomic sequences 3' of the primer binding site. *J. Biol. Chem.* **281**, 24227–24235
516. Piekna-Przybylska, D., and Bambara, R. A. (2011) Requirements for efficient minus strand strong-stop DNA transfer in human immunodeficiency virus 1. *RNA Biol.* **8**, 230–236
517. 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
518. Vo, M.-N., Barany, G., Rouzina, I., and Musier-Forsyth, K. (2009) HIV-1

- Nucleocapsid Protein Switches the Pathway of TAR RNA/DNA Annealing from Loop-Loop “Kissing” to “Zipper.” *J. Mol. Biol.* **386**, 789–801
519. Chen, Y., Balakrishnan, M., Roques, B. P., Fay, P. J., and Bambara, R. A. (2003) Mechanism of Minus Strand Strong Stop Transfer in HIV-1 Reverse Transcription *. *J. Biol. Chem.* **278**, 8006–8017
520. Tanese, N., Telesnitsky, A., and Goff, S. P. (1991) Abortive reverse transcription by mutants of Moloney murine leukemia virus deficient in the reverse transcriptase-associated RNase H function. *J. Virol.* **65**, 4387–4397
521. Blain, S. W., and Goff, S. P. (1995) Effects on DNA synthesis and translocation caused by mutations in the RNase H domain of Moloney murine leukemia virus reverse transcriptase. *J. Virol.* **69**, 4440–4452
522. Julias, J. G., McWilliams, M. J., Sarafianos, S. G., Arnold, E., and Hughes, S. H. (2002) Mutations in the RNase H domain of HIV-1 reverse transcriptase affect the initiation of DNA synthesis and the specificity of RNase H cleavage in vivo. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 9515–9520
523. Cameron, C. E., Ghosh, M., Le Grice, S. F. J., and Benkovic, S. J. (1997) Mutations in HIV reverse transcriptase which alter RNase H activity and decrease strand transfer efficiency are suppressed by HIV nucleocapsid protein. *Proc. Natl. Acad. Sci.* **94**, 6700–6705
524. Luczkowiak, J., Matamoros, T., and Menéndez-Arias, L. (2018) Template-primer binding affinity and RNase H cleavage specificity contribute to the strand transfer efficiency of HIV-1 reverse transcriptase. *J. Biol. Chem.* **293**, 13351–13363
525. Purohit, V., Balakrishnan, M., Kim, B., and Bambara, R. A. (2005) Evidence that HIV-1 reverse transcriptase employs the DNA 3’ end-directed primary/secondary RNase H cleavage mechanism during synthesis and strand transfer. *J. Biol. Chem.* **280**, 40534–40543
526. Basu, V. P., Song, M., Gao, L., Rigby, S. T., Hanson, M. N., and Bambara, R. A. (2008) Strand transfer events during HIV-1 reverse transcription. *Virus Res.* **134**, 19–38
527. Guo, J., Wu, T., Anderson, J., Kane, B. F., Johnson, D. G., Gorelick, R. J., Henderson, L. E., and Levin, J. G. (2000) Zinc finger structures in the human immunodeficiency virus type 1 nucleocapsid protein facilitate efficient minus- and plus-strand transfer. *J. Virol.* **74**, 8980–8988
528. Driscoll, M. D., and Hughes, S. H. (2000) Human Immunodeficiency Virus Type 1 Nucleocapsid Protein Can Prevent Self-Priming of Minus-Strand Strong Stop DNA by Promoting the Annealing of Short Oligonucleotides to Hairpin Sequences. *J. Virol.* **74**, 8785–8792
529. Wisniewski, M., Chen, Y., Balakrishnan, M., Palaniappan, C., Roques, B. P., Fay, P. J., and Bambara, R. A. (2002) Substrate Requirements for Secondary Cleavage by HIV-1 Reverse Transcriptase RNase H*. *J. Biol. Chem.* **277**, 28400–28410
530. Lener, D., Tanchou, V., Roques, B. P., Grice, S. F. J. L., and Darlix, J.-L. (1998) Involvement of HIV-I Nucleocapsid Protein in the Recruitment of Reverse Transcriptase into Nucleoprotein Complexes Formed in Vitro*. *J. Biol. Chem.* **273**, 33781–33786
531. Druillenec, S., Caneparo, A., de Rocquigny, H., and Roques, B. P. (1999) Evidence of interactions between the nucleocapsid protein NCp7 and the reverse transcriptase of HIV-1. *J. Biol. Chem.* **274**, 11283–11288
532. Purohit, V., Roques, B. P., Kim, B., and Bambara, R. A. (2007) Mechanisms That Prevent Template Inactivation by HIV-1 Reverse Transcriptase RNase H Cleavages *. *J. Biol. Chem.* **282**, 12598–12609
533. Guo, J., Wu, T., Bess, J., Henderson, L. E., and Levin, J. G. (1998) Actinomycin D Inhibits Human Immunodeficiency Virus Type 1 Minus-Strand Transfer in In Vitro and Endogenous Reverse Transcriptase Assays. *J. Virol.* **72**, 6716–6724
534. Darlix, J.-L., Garrido, J. L., Morellet, N., Mély, Y., and de Rocquigny, H. (2007)

- Properties, functions, and drug targeting of the multifunctional nucleocapsid protein of the human immunodeficiency virus. *Adv. Pharmacol. San Diego Calif.* **55**, 299–346
535. Post, K., Kankia, B., Gopalakrishnan, S., Yang, V., Cramer, E., Saladores, P., Gorelick, R. J., Guo, J., Musier-Forsyth, K., and Levin, J. G. (2009) Fidelity of plus-strand priming requires the nucleic acid chaperone activity of HIV-1 nucleocapsid protein. *Nucleic Acids Res.* **37**, 1755–1766
536. Guo, J., Wu, T., Kane, B. F., Johnson, D. G., Henderson, L. E., Gorelick, R. J., and Levin, J. G. (2002) Subtle Alterations of the Native Zinc Finger Structures Have Dramatic Effects on the Nucleic Acid Chaperone Activity of Human Immunodeficiency Virus Type 1 Nucleocapsid Protein. *J. Virol.* **76**, 4370–4378
537. Derebail, S. S., Heath, M. J., and DeStefano, J. J. (2003) Evidence for the differential effects of nucleocapsid protein on strand transfer in various regions of the HIV genome. *J. Biol. Chem.* **278**, 15702–15712
538. Williams, M. C., Gorelick, R. J., and Musier-Forsyth, K. (2002) Specific zinc-finger architecture required for HIV-1 nucleocapsid protein's nucleic acid chaperone function. *Proc. Natl. Acad. Sci.* **99**, 8614–8619
539. Mitra, M., Wang, W., Vo, M.-N., Rouzina, I., Barany, G., and Musier-Forsyth, K. (2013) The N-Terminal Zinc Finger and Flanking Basic Domains Represent the Minimal Region of the Human Immunodeficiency Virus Type-1 Nucleocapsid Protein for Targeting Chaperone Function. *Biochemistry.* **52**, 8226–8236
540. 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
541. Johnson, P. E., Turner, R. B., Wu, Z. R., Hairston, L., Guo, J., Levin, J. G., and Summers, M. F. (2000) A mechanism for plus-strand transfer enhancement by the HIV-1 nucleocapsid protein during reverse transcription. *Biochemistry.* **39**, 9084–9091
542. 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
543. Dethoff, E. A., Hansen, A. L., Musselman, C., Watt, E. D., Andricioaei, I., and Al-Hashimi, H. M. (2008) Characterizing complex dynamics in the transactivation response element apical loop and motional correlations with the bulge by NMR, molecular dynamics, and mutagenesis. *Biophys. J.* **95**, 3906–3915
544. 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
545. Beltz, H., Azoulay, J., Bernacchi, S., Clamme, J.-P., Ficheux, D., Roques, B., Darlix, J.-L., and Mély, Y. (2003) Impact of the Terminal Bulges of HIV-1 cTAR DNA on its Stability and the Destabilizing Activity of the Nucleocapsid Protein NCp7. *J. Mol. Biol.* **328**, 95–108
546. Cosa, G., Harbron, E. J., Zeng, Y., Liu, H.-W., O'Connor, D. B., Eta-Hosokawa, C., Musier-Forsyth, K., and Barbara, P. F. (2004) Secondary Structure and Secondary Structure Dynamics of DNA Hairpins Complexed with HIV-1 NC Protein. *Biophys. J.* **87**, 2759–2767
547. Chaminade, F., Darlix, J.-L., and Fossé, P. (2022) RNA Structural Requirements for Nucleocapsid Protein-Mediated Extended Dimer Formation. *Viruses.* **14**, 606
548. 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

549. Wu, H., Mitra, M., Naufer, M. N., McCauley, M. J., Gorelick, R. J., Rouzina, I., Musier-Forsyth, K., and Williams, M. C. (2014) Differential contribution of basic residues to HIV-1 nucleocapsid protein's nucleic acid chaperone function and retroviral replication. *Nucleic Acids Res.* **42**, 2525–2537

550. 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

551. 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**, 25678

552. Fossé, P., Motté, N., Roumier, A., Gabus, C., Muriaux, D., Darlix, J. L., and Paoletti, J. (1996) A short autocomplementary sequence plays an essential role in avian sarcoma-leukosis virus RNA dimerization. *Biochemistry.* **35**, 16601–16609

553. 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

554. Ohi, Y., and Clever, J. L. (2000) Sequences in the 5' and 3' R elements of human immunodeficiency virus type 1 critical for efficient reverse transcription. *J. Virol.* **74**, 8324–8334

555. Grohman, J. K., Gorelick, R. J., Lickwar, C. R., Lieb, J. D., Bower, B. D., Znosko, B. M., and Weeks, K. M. (2013) A Guanosine-Centric Mechanism for RNA Chaperone Function. *Science.* **340**, 190–195

556. Belfetmi, A. Les protéines de nucléocapside du VIH-1: structures, dynamiques, propriétés de fixation et de déstabilisation des acides nucléiques

557. Vrolijk, M. M., Harwig, A., Berkhout, B., and Das, A. T. (2009) Destabilization of the TAR hairpin leads to extension of the polyA hairpin and inhibition of HIV-1 polyadenylation. *Retrovirology.* **6**, 13

558. Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31**, 3406–3415

559. Chen, Y., Maskri, O., Chaminade, F., René, B., Benkaroun, J., Godet, J., Mély, Y., Mauffret, O., and Fossé, P. (2016) Structural Insights into the HIV-1 Minus-strand Strong-stop DNA. *J. Biol. Chem.* **291**, 3468–3482

560. Huang, Y.-L., Kawai, G., Hasegawa, A., Kannagi, M., and Masuda, T. (2019) Impact of 5'-end nucleotide modifications of HIV-1 genomic RNA on reverse transcription. *Biochem. Biophys. Res. Commun.* **516**, 1145–1151

561. Clever, J. L., Eckstein, D. A., and Parslow, T. G. (1999) Genetic Dissociation of the Encapsidation and Reverse Transcription Functions in the 5' R Region of Human Immunodeficiency Virus Type 1. *J. Virol.* **73**, 101–109

562. Selyutina, A., Persaud, M., Lee, K., KewalRamani, V., and Diaz-Griffero, F. (2020) Nuclear Import of the HIV-1 Core Precedes Reverse Transcription and Uncoating. *Cell Rep.* **32**, 108201

563. Brussel, A., and Sonigo, P. (2004) Evidence for gene expression by unintegrated human immunodeficiency virus type 1 DNA species. *J. Virol.* **78**, 11263–11271

564. Masuda, T., Sato, Y., Huang, Y.-L., Koi, S., Takahata, T., Hasegawa, A., Kawai, G., and Kannagi, M. (2015) Fate of HIV-1 cDNA intermediates during reverse transcription is dictated by transcription initiation site of virus genomic RNA. *Sci. Rep.* **5**, 17680

565. Piekna-Przybylska, D., Dykes, C., Demeter, L. M., and Bambara, R. A. (2011) Sequences in the U3 region of Human Immunodeficiency Virus 1 improve efficiency of minus strand transfer in infected cells. *Virology.* **410**, 368–374

Titre : Rôle des tiges-boucles TAR et cTAR sur le premier transfert de brin du VIH-1

Mots clés : transcription inverse, transfert de brin, ADN *strong-stop*, structure, protéine de nucléocapside, VIH-1

Résumé : Le cycle du VIH-1 requiert le processus de transcription inverse permettant la synthèse d'un ADN double-brin à partir du génome viral constitué de deux copies d'ARN simple-brin de polarité positive servant alors de matrice. Deux événements, appelés transferts de brin terminaux, sont nécessaires à la réalisation de la transcription inverse et donnent lieu chacun à la synthèse d'un des deux brins d'ADN. Mes travaux de thèse sont focalisés sur le premier transfert de brin. Celui-ci se produit de l'extrémité 5' vers l'extrémité 3' de l'ARN génomique viral (ARNg) au niveau des séquences R dupliquées, et ce après la synthèse de l'ADN *strong-stop* (ADNss) par la transcriptase inverse (RT). L'ADNss étant une copie des régions R et U5 de l'extrémité 5' de l'ARNg, il peut s'apparier avec la région R présente aussi en 3' de l'ARN viral et permettre la synthèse du brin d'ADN de polarité négative. Les séquences des régions R forment deux tiges-boucles (cTAR et cpoly(A) dans l'ADNss, 3'TAR et 3'poly(A) dans l'ARNg). Ces structures doivent être ouvertes pour permettre l'hybridation entre les régions R de l'ADNss et de l'ARNg.

Plusieurs études *in vitro* suggèrent que la protéine de nucléocapside du VIH-1 (NC) est en partie responsable de l'ouverture de ces tiges-boucles et de leur appariement. Cette protéine est connue pour son activité chaperon des acides nucléiques et est impliquée dans plusieurs étapes du cycle viral. Les travaux étudiant le rôle de la NC sur la transcription inverse ont montré notamment *in vitro* que cette protéine facilite le premier transfert de brin et en se fixant préférentiellement au niveau de certaines guanines déstabilise les tiges-boucles TAR et cTAR. Cette déstabilisation par la NC est dépendante du nombre de paires de bases consécutives impliquées dans la tige-boucle.

Les travaux de thèse déterminent si *ex vivo* le premier transfert de brin dépend de la déstabilisation par la NC des tiges-boucles cTAR de l'ADNss et TAR de l'ARNg. Pour cela, différents types de mutations ont été conçues et introduites par mutagénèse dirigée dans les tiges-boucles afin d'empêcher leur déstabilisation par la NC et produire ainsi un effet négatif sur la transcription inverse. La réplication des virus mutés et plus précisément leur capacité à réaliser la transcription inverse de l'ARNg sont analysées et discutées.

Title : Role of TAR and cTAR stem-loops on HIV-1 first strand transfer

Keywords : reverse transcription, strand transfert, strong-stop DNA, structure, nucleocapsid protein, HIV-1

Abstract : The HIV-1 life cycle requires the reverse transcription process. This process allows the synthesis of a double-stranded DNA from its genome, which consists of two copies of single-stranded RNA of positive polarity used as a template. Two events, called terminal strand transfers, are necessary for reverse transcription to occur and each results in the synthesis of one of the two DNA strands. My thesis work focuses on the first strand transfer. This occurs from the 5' end to the 3' end of the viral genomic RNA (gRNA) at the duplicated R sequences after the synthesis of the strong-stop DNA (ssDNA) by the reverse transcriptase (RT). As ssDNA is a copy of the R and U5 regions of the 5' end of the gRNA, it can pair with the R region also present in the 3' of the viral RNA and allow synthesis of the negative polarity DNA strand. The sequences of the R regions form two stem-loops (cTAR and cpoly(A) in ssRNA, 3'TAR and 3'poly(A) in gRNA). These structures must be open to allow hybridization between the R regions of the ssDNA and RNA.

Several *in vitro* studies suggest that the HIV-1 nucleocapsid protein (NC) is partly responsible for the opening of these stem-loops and their pairing. This protein is known for its chaperone nucleic activity and is involved in several steps of the viral cycle. The works studying the role of NC on reverse transcription showed *in vitro* that this protein facilitates the first strand transfer and destabilizes the TAR and cTAR stem-loops by binding to certain guanines. Furthermore, the destabilisation reaction by NC is dependent on the number of consecutive base pairs involved in the stem-loop.

The thesis work determines whether the first strand transfer depends on the destabilisation of the ssDNA cTAR and the gRNA TAR stem-loops by NC *ex vivo*. To this end, different types of mutations were designed and introduced in these stem-loops to prevent their destabilisation by NC and to observe an effect on reverse transcription. Replication of mutant viruses and more precisely the reverse transcription of gRNA are analysed and discussed.