



Intégration du VIH-1 : Contrôle et régulation de l'interaction fonctionnelle entre l'intégrase et la chromatine

Julien Matysiak

► To cite this version:

Julien Matysiak. Intégration du VIH-1 : Contrôle et régulation de l'interaction fonctionnelle entre l'intégrase et la chromatine. Biologie cellulaire. Université de Bordeaux, 2016. Français. NNT : 2016BORD0284 . tel-01515700

HAL Id: tel-01515700

<https://theses.hal.science/tel-01515700>

Submitted on 28 Apr 2017

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

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

THÈSE PRÉSENTÉE
POUR OBTENIR LE GRADE DE
DOCTEUR DE
L'UNIVERSITÉ DE BORDEAUX

ÉCOLE DOCTORALE
SPÉCIALITÉ IMMUNOLOGIE-MICROBIOLOGIE

Par Julien MATYSIAK

**Intégration du VIH-1 : Contrôle et régulation de l'interaction
fonctionnelle entre l'intégrase et la chromatine**

Sous la direction de : Vincent PARISSI

Soutenue le 15 décembre 2016

Membres du jury :

M. GARBAY Bertrand, Professeur Bordeaux	Président
Mme LESAGE Pascale, Directrice de recherche CNRS	Rapporteur
M. EMILIANI Stéphane, Directeur de recherche Institut Cochin Paris	Rapporteur
M. DELELIS Olivier, Directeur de recherche CNRS	Examineur
M. TEICHMANN Martin, Professeur Bordeaux	Examineur
M. TAOUJI Saïd, Responsable Plateforme de criblage BMYScreen	Invité

Titre : Intégration du VIH-1 : Contrôle et régulation de l'interaction fonctionnelle entre l'intégrase et la chromatine

Résumé : L'intégrase (IN) du VIH-1 est une enzyme clé du cycle viral catalysant l'insertion stable de l'ADN viral dans le génome de la cellule infectée. L'IN participe également à de nombreuses étapes du cycle viral telles que la transcription inverse ou la maturation virale. Ainsi, la compréhension des mécanismes régulant l'intégration du VIH-1 représente un enjeu majeur dans le cadre notamment d'approches thérapeutiques. En effet, les études montrent que ces mécanismes sont finement régulés dans la cellule par des facteurs viraux et cellulaires agissant à différentes étapes du cycle viral. C'est donc dans ce contexte que nous avons étudié les facteurs à la fois viraux et cellulaires régulant ce processus. Dans un premier temps, les déterminants viraux modulant l'intégration dans la chromatine ont été analysés dans le cas de plusieurs modèles rétroviraux. Puis, dans un second temps, nous avons étudié l'impact de facteurs cellulaires, identifiés au laboratoire, sur les mécanismes d'insertion de l'ADN viral dans le génome cellulaire.

Mon travail de thèse s'est ainsi articulé en trois axes majeurs aboutissant à :

- La démonstration de la régulation de l'intégration rétrovirale par la structure chromatinienne de l'hôte
- L'identification de nouveaux cofacteurs cellulaires participant à la régulation de l'intégration dans la chromatine dont le complexe de remodelage FACT « Facilitates Chromatin Transcription »
- L'identification d'une nouvelle interaction fonctionnelle entre l'IN du VIH-1 et la queue amino-terminale de l'histone humaine H4 et de son rôle dans la sélectivité de l'intégration

Mots clés : VIH-1, Intégration, Intégrase, IN, Chromatine, Sélectivité, Cofacteurs cellulaires, FACT, MLV, PFV, ASV, H4K20me1

Title: HIV-1 Integration: Control and regulation of the functional interaction between integrase and chromatin

Abstract: HIV-1 integrase (IN) is a key enzyme of the viral cycle that catalyzes the stable insertion of viral DNA into the genome of the infected cell. IN also participates in many stages of the viral cycle such as reverse transcription or viral maturation. Thus, an understanding of the mechanisms regulating the integration of HIV-1 is a major challenge, particularly in the context of therapeutic approaches. Indeed, studies show that these mechanisms are finely regulated in the cell by viral and cellular factors acting at different stages of the viral cycle. It is in this context that we studied both viral and cellular factors regulating this process. Initially, the viral determinants modulating the integration in chromatin were analyzed in the case of several retroviral models. Then, we studied the impact of cellular factors, identified in the laboratory, on the mechanisms of insertion of the viral DNA in the cellular genome.

My thesis work has thus been articulated in three major axes leading to:

- The demonstration of the regulation of retroviral integration by the chromatin structure of the host
- The identification of new cellular cofactors participating in the regulation of chromatin integration, including the FACT remodeling complex "Facilitates Chromatin Transcription"
- The identification of a new functional interaction between the HIV-1 IN and the amino-terminal tail of human H4 histone and its role in the selectivity of integration

Keywords : HIV-1, Integration, Integrase, IN, Selectivity, Chromatin, Cellular Cofactors, FACT, ASV, MLV, PFV, H4K20me1

Laboratoire de Microbiologie fondamentale et de Pathogénicité

MFP 5234 Equipe: Replication and mobility of viral and bacterial genomes
Adresse: 146 rue Léo saignat BAT 3A 33076 Bordeaux Cedex

Sommaire

Sommaire	3
Abréviations	15
Synthèse bibliographique	25
Introduction.....	27
1. Le VIH-1 et pathologie associée : le SIDA	27
1.1. Le syndrome d'immunodéficience acquise « SIDA ».....	27
a) Historique et épidémiologie.....	27
b) Transmission du virus.....	29
c) Diversité génétique.....	29
d) Pathologie.....	29
1.2. Le Virus de l'immunodéficience humaine « VIH ».....	31
a) Classification.....	31
b) Structure du virion.....	33
c) Génome du VIH-1.....	35
d) Protéines virales.....	35
i- Glycoprotéines d'enveloppe.....	35
ii- Protéines structurales.....	37
iii- Protéines non structurales.....	39
iv- Protéines régulatrices.....	43
e) Cycle viral.....	43
i- Reconnaissance spécifique, entrée du virus et intégration	45
ii- Production et libération de nouveaux virions	45

1.3. Moyens de luttres antirétrovirales.....47

- a) Prévention à l'infection à VIH-1.....47
- b) Stratégie thérapeutique.....49
- c) Résistance aux antirétroviraux.....51
- d) Les essais vaccinaux.....53
- e) Un cas de guérison : « Le patient de Berlin ».....53

2. L'intégration et intégrase du VIH-1.....55

2.1. Intégrase virale.....55

- a) Domaine N-terminal (1-50)55
- b) Le core catalytique (51-212)55
- c) Domaine C-terminal (213 -288).....57
- d) L'intasome du VIH-1.....57

2.2. Intégration virale : mécanisme et fonction.....59

- a) Introduction.....59
- b) 3'end proccesing.....61
- c) Transfert de brins.....61
- d) Réparation post-intégration.....61
- e) Les produits non intégrés.....63

2.3. Activités de l'intégrase *in vitro*.....63

- a) Test de 3'end processing et transfert de brin.....65
- b) L'intégration concertée.....65
- c) La désintégration67

2.4. Facteurs cellulaires modulateurs de l'intégration.....67

a) Integrase interactor 1 “INI1”.....	69
b) Barrier to auto integration factor “BAF”.....	69
c) High mobility group chromosomal protein A1 “HMGA1”.....	71
d) Human heat-shock protein “HSP60”.....	71
e) RAD51.....	73
f) Lens epithelium-derived growth factor “LEDGF/p75.....	73
 2.5. Sélectivité de l'intégration.....	75
a) Introduction.....	75
b) Import nucléaire et intégration.....	77
c) Facteurs cellulaires participant à la sélectivité de l'intégration.....	79
d) Nucléosome, substrat de l'intégration.....	79
e) Conclusion.....	81
 3. Chromatine	83
 3.1. Chromatine cellulaire.....	83
a) Structure de la chromatine.....	83
b) Les histones de cœur.....	83
c) Les variants d'histones.....	85
d) Histones de liaisons.....	85
 3.2. Modifications post-traductionnelles des histones et domaine de reconnaissance.....	87
a) Acétylation.....	87
i- Bromodomaine.....	87
b) Méthylation.....	89
i- Chromodomaine.....	89
ii- Tudor domaine.....	89
iii- Domaine MBT.....	91

iv- Chromo barrel domaine.....	91
v- Domaine PWWP.....	93
c) Mono-ubiquitination.....	93
d) Phosphorylation.....	93
3.3. Dynamique de la chromatine	95
a) Complexes de remodelage ATP-dépendants.....	95
b) Enzymes responsables des modifications post-traductionnelles.....	97
c) Histones chaperones.....	99
4. Chromatine et intégration	99
4.1. Influence de la structure de la chromatine sur l'intégration.....	101
4.2. Impact de LEDGF/p75 sur l'intégration dans la chromatine	101
4.3. Impact des facteurs de remodelage ATP-dépendants.....	101
5. Objectifs de l'étude	103
Matériels et Méthodes	107
1. Matériels	108
1. 1. Culture cellulaire	108
a) Cellules de mammifères	108
1. 2. Vecteurs lentiviraux pseudotypés VSV-G	108
1. 3. Plasmides	109

a) Plasmide d'expression bactérien	109
b) Plasmide d'expression pour la production des vecteurs lentiviraux pseudotypés VSV-G	109
1. 4. Protéines commerciales	110
1. 5. Les anticorps	110
1. 6. Les oligonucléotides	111
1. 7. siRNA	111
1. 8. Peptides	112
1. 9. Nucléosomes	113
2. Méthodes	113
2. 1. Cultures bactériennes	113
a) Culture	113
b) Transformation	113
c) Extraction et purification de l'ADN plasmidique	114
2. 2. Cultures de cellules mammifères	114
2. 3. Infection des cellules par les vecteurs lentiviraux	114
a) Production des vecteurs lentiviraux pseudotypés VSV-G	114
b) Titration et dosage	115
c) Traitement DNase	115
d) Transduction	115
2. 4. Double transfection par siRNA-SSRP1	116

a) Extraction des protéines cellulaires	116
2. 5. Manipulation des acides nucléiques	116
a) Extraction des acides nucléiques de cellules infectées	116
b) Mutagenèses dirigées	116
2. 6. Préparation des cellules pour le séquençage des sites d'intégrations	117
2. 7. Manipulation des protéines	117
a) Production de l'IN recombinante	117
2. 7. Tests <i>in vitro</i>	118
a) Test d'intégration concertée	118
b) Histone peptide array	119
c) FAIRE "Formaldehyde-Assisted Isolation of Regulatory Elements"	119
d) Expériences de copréciptation	120
e) Microscale thermophoresis	121
f) Far dot blot	121
g) Sélection des cofacteurs cellulaires du complexe IN-ADNv.....	122
2. 8. Analyse <i>in silico</i>	122
a) Docking moléculaire	122
Résultats	123

Partie 1 : Impact de la structure chromatinienne sur l'intégration rétrovirale.....125

1. Objectifs125

2. Effet de la chromatine sur l'intégration de différents modèles rétroviraux127

2. 1. Intégration *in vitro* de différents systèmes rétroviraux.....127

- a) Analyse *in silico* du pBSK-Zeo 5S-G5E4 chromatinisé129
- b) Analyse biochimique du pBSK-Zeo 5S-G5E4 chromatinisé129
- c) Impact de la structure chromatinienne sur l'intégration de différents modèles rétroviraux131
- d) Impact de la structure chromatinienne sur la sélectivité de différents systèmes rétroviraux131
- e) Impact des structures intasomales sur la sélectivité d'intégration du VIH-1.133

2. 2. Rôle du « N-terminal extension domain » (NED) de PFV sur l'intégration dans la chromatine135

- a) Impact du domaine NED sur l'intégration *in vitro*135
- b) Impact du domaine NED sur la sélectivité de l'intégration *in vitro*137

2. 3. Étude comparative des sites d'intégration rétroviraux et du positionnement des nucléosomes dans la cellule infectée137

3. Conclusion et discussion139

Partie 2 : Identification de nouveaux cofacteurs cellulaires de l'intégration : rôle du complexe protéique FACT « Facilitates Chromatin Transcription »	145
1. Objectifs	145
2. Identification et validation des cofacteurs cellulaires de l'intasome du VIH-1	147
2. 1. Stratégie d'identification de cofacteurs cellulaires	147
2. 2. Sélection des cofacteurs putatifs de l'intasome du VIH-1	149
2. 3. Obtention de l'interactome de l'intasome du VIH-1	149
3. Rôle du complexe protéique FACT sur l'intégration du VIH-1	151
3. 1. Introduction	151
3. 2. Mode d'interaction du FACT avec l'intasome du VIH-1	153
3. 3. Rôle du complexe FACT sur les phases précoces de la réplication <i>in cellulo</i>	155
3. 4. Rôle du complexe FACT sur l'intégration concertée <i>in vitro</i>	157
4. Étude du mécanisme d'action impliqué dans la restauration de l'intégration médié par le complexe FACT	157

4. 1. Étude de l'activité de remodelage du complexe FACT	157
4. 2. Étude de l'intégration sur nucléosome tétramérique assemblé <i>in vitro</i>	159

5. Conclusion et discussion161

Partie 3 : Interaction fonctionnelle entre l'IN du VIH-1 et le peptide d'histone H4K20me1165

1. Objectifs165

2. Interaction entre l'IN du VIH-1 et la chromatine167

2. 1. Interaction entre l'IN du VIH-1 et les queues d'histones	167
2. 2. Identification du domaine de l'IN du VIH-1 impliqué dans l'interaction	169
2. 3. Analyse structurale et modélisation <i>in silico</i> de l'interaction	169
2. 4. Identification des acides aminés impliqués dans l'interaction avec la queue H4	173
2. 5. Impact des mutations impliquées dans l'interaction sur la réplication et l'intégration <i>in vivo</i>	175

3. Étude fonctionnelle de l'interaction IN-queue H4.....177

3. 1. Impact de l'interaction avec la queue d'histone H4 sur l'intégration <i>in vitro</i>	177
---	-----

3. 2. Impact des mutants d'interaction pour la queue H4 sur l'interaction avec le mononucléosome	179
3. 3. Impact des mutants d'interaction pour la queue H4 sur l'intégration dans le mononucléosome	181

4. Conclusion et discussion	183
--	------------

Partie 4 : Conclusion générale et perspectives.....189

1. 1 Contrôle de l'intégration par la structure de la chromatine et de l'intasome	191
--	------------

1. 2 Le FACT : un nouveau cofacteur de l'intégration du VIH-1	193
--	------------

1. 3 Un chromodomaine dans l'IN du VIH-1	195
---	------------

Bibliographie	205
----------------------------	------------

Résumé	223
---------------------	------------

Abréviations

0-9

3TC 2',3'-didéoxy-3'-thiacytidine

A

ABL Abelson

ACF ATP-utilizing chromatin assembly and remodeling factor

Ad5 Adenovirus type 5

ADN Acide désoxyribonucléique

AIDS Acquired immune deficiency syndrome

ALV Avian Leukosis Virus

ARNv Acide ribonucléique viral

ARNt Acide ribonucléique de transfert

AT AT hook like domain

ATF-3 Activating Transcription Factor 3

ATP Adénosine-5'-triphosphate

ATR ATM and Rad3-related

ATV Atazanavir

AZT Azidothymidine

B

BAF Barrier to auto integration factor

BCR Breakpoint Cluster Region

BET Bromodomain and Extra-Terminal motif

BRCA1 Breast Cancer 1

BRCA2 Breast Cancer 2

BRD2 Bromodomain-containing protein 2

BRD3 Bromodomain-containing protein 3

BRD4 Bromodomain-containing protein 4

BSA Bovine serum albumine

C

CA Capside

CBX1 Chromobox protein homolog 1

CCD Catalytic core domain

CCR5 C-C chemokine receptor type 5
CD4 Cluster of differentiation 4
CD8 Cluster of differentiation 8
CD28 Cluster of differentiation 28
CDC Center for Disease Control
CDK9 Cyclin-dependent kinase 9
CHAPS 3-[(3-cholamidopropyl)diméthylammonio]-1-propanesulfonate
CHD Chromodomain Helicase DNA-Binding
CHRAC Chromatin accessibility complex
CMH-I Complexe majeur d'histocompatibilité de type 1
CMH-II Complexe majeur d'histocompatibilité de type 2
CMV Cytomegalovirus
CoA Coenzyme A
CPI Complexe de pré-intégration
CR Charged region
CRF Circulating recombinant form
CTD C-terminal domain
CXCR4 CXC chemokine receptor 4
CYCT1 Cyclin-T1

D

DMEM Dulbecco's modified minimal essential meduim
DNA-PK DNA-dependent protein kinase
DO Densité optique
DRV Darunavir
DTG Dolutégravir
DSS Disuccinimidyl subérate
DTT Dithiotréitol

E

ENCODE Encyclopedia of DNA Elements
EDTA Ethylenediaminetetraacetic acid
EFV Efavirenz

EGFP Enhanced Green Fluorescent Protein
ELISA Enzyme-linked immunosorbent assay
Env Enveloppe
ETS E26 transformation-specific
EVG Elvitégravir

F

FACT Facilitates chromatin transcription
FAIRE Formaldehyde-Assisted Isolation of Regulatory Elements
FEN-1 Flap endonuclease 1
FIs Fusion inhibitors
FSI Full site integration

G

GFP Green fluorescent protein
gp41 glycoprotein 41
gp120 glycoprotein 120

H

HAART Highly Active Antiretroviral Therapy
HAT Histone acétyltransférase
HDAC Histone désacétylase
HEK Human embryonic kidney
HeLa Henrietta lacks
HFV Human Foamy Virus
HTLV Human T-Lymphotropic Virus
HIV Human immunodeficiency virus
HMGA1 High mobility group chromosomal protein A1
hMSL3 Human male-specific lethal-3 protein
HP1 heterochromatin protein 1
HR1 Heptad repeat 1
HR2 Heptad repeat 2
HR3 Homology region 3
HR Homologous recombination

HRP2 Hepatoma-derived growth factor related protein 2

HSI Half site integration

HSP60 Human Heat-Shock protein 60

HTH Hélice-tour-hélice

I

IBD Integrase binding domain

ICTV International Committee on Taxonomy of Viruses

IN Integrase

ING2 Inhibitor of growth protein 2

INI1 Integrase interactor 1

INO80 Inositol-requiring 80

InR séquence initiatrice

INSTIs Integrase strand transfer inhibitors

IPTG Isopropyl β -D-1-thiogalactopyranoside

IRES Internal ribosome entry sequence

ISWI Imitation Switch

J

JMJ JUMONJI

K

KO Knockout

L

L3MBTL1 Lethal(3)malignant brain tumor-like protein

LAP2 Lamina-associated protein 2

LAV Lymphadenopathy Associated Virus

LB Luria-Bertani

LEDGF/p75 Lens epithelium-derived growth factor

LTR Long terminal repeat

LT Lymphocyte T

M

MA Matrice
MAN1 LEM domain-containing protein 3
MBT Malignant brain tumor
MLV Murine leukemia virus
MMTV Mouse mammary tumor virus
MN Mononucléosome
MOI Multiplicity of infection
MSK1/2 Mitogen- and stress-activated protein kinase 1 and 2
MVC Maraviroc
MSL3 Male-specific lethal-3 protein-like
MST Microscale thermophoresis

N

NC Nucléocapside
NED NTD extension domain
Nef Negative factor
NFAT Nuclear factor of activated T-cells
NFkB Nuclear factor kappa B
NHEJ Non-Homologous End-Joining
NLS Nuclear localisation signal
NNRTIs Non nucleoside reverse transcriptase inhibitors
NPC Nuclear pore complex
NRTIs Nucleoside reverse transcriptase inhibitors
N-TEF Negative transcription elongation factor
NTD N-terminal domain
Nup96/98/153/358 Nucleoporin
NURF Nucleosome remodeling factor
NVP Nevirapine

O

ODN Oligodésoxyribonucléotide
OMS Organisation mondiale de la santé

P

PARP1 Poly(ADP-ribose) polymérase 1
Pb Paire de bases
PBS Primer binding site
PEG Polyéthylène glycol
PFV Prototype foamy virus
PHF1 PHD finger protein 1
PHF19 PHD Finger Protein 19
PIs Protease inhibitors
PMSF Fluorure de phénylméthylsulfonyl
PN Polynucléosomal
PN H4 Polynucléosomal tétramérique
PN H8 Polynucléosomal octamérique
PPE Prophylaxie post-exposition
PPT Polypurine tract
PR Protease
PrEP Prophylaxie pré-exposition
P-TEFb Positive transcription elongation factor b
PWWP Pro-Trp-Trp-Pro motif

R

RAL Raltegravir
REA Restriction enzyme accessibility assay
Rev Regulator of virion
RMN Résonance magnétique nucléaire
RMSD Root Mean square deviation
RNAPII RNA polymerase II
RRE Rev response element
RSC Remodeling the structure of chromatin
RSF Remodeling and spacing factor
RSK2 Ribosomal S6 Kinase 2
RSV Rous sarcoma virus
RT Reverse transcriptase
RTC Reverse transcription complex

S

SAM S-adénosyl méthionine

SANT switching-defective protein 3 (Swi3), adaptor 2 (Ada2), nuclear receptor co-repressor (N-CoR), transcription factor (TF)IIIB

SDS Sodium dodécyl sulfate

SH3 Sarc homology 3 domain

shRNA short hairpin RNA

SIDA Syndrome d'immunodéficience acquise

siRNA Small interfering RNA

SIV Simian immunodeficiency virus

SLIDE SANT-like ISWI domain

SMIPPIs Small-molecule inhibitor of protein-protein interaction

SNF2H Sucrose nonfermenting protein 2 homolog

SP1 Specific protein 1

Spt16 Suppressor of Ty16

SRD Supercoiled-DNA recognition region

SSRP1 Structure-specific recognition protein 1

SVF Sérum de veau fœtal

SWI/SNF Switch/sucrose nonfermenting

SWR1 SWI/SNF-related protein 1

T

T-20 enfuvirtide

TAF TBP-associated factors

Tat Transactivateur de la transcription

TAR Transactivating-responsive region

TBP TATA-binding protein

TCA Trichloroacetic acid

TNPO3 Transportine 3

TRN-SR2 Transportine-SR2

U

UV Ultraviolet

V

Vif Viral infectivity factor

VIH-1 Virus de l'immunodéficience humaine de type 1

VIH-2 Virus de l'immunodéficience humaine de type 2

Vpr Viral protein R

Vpu Viral protein unique

VSV Virus de la stomatite vésiculaire

W

WDSV Walleye Dermal Sarcoma Virus

WT Wild-type

Synthèse bibliographique

a



b

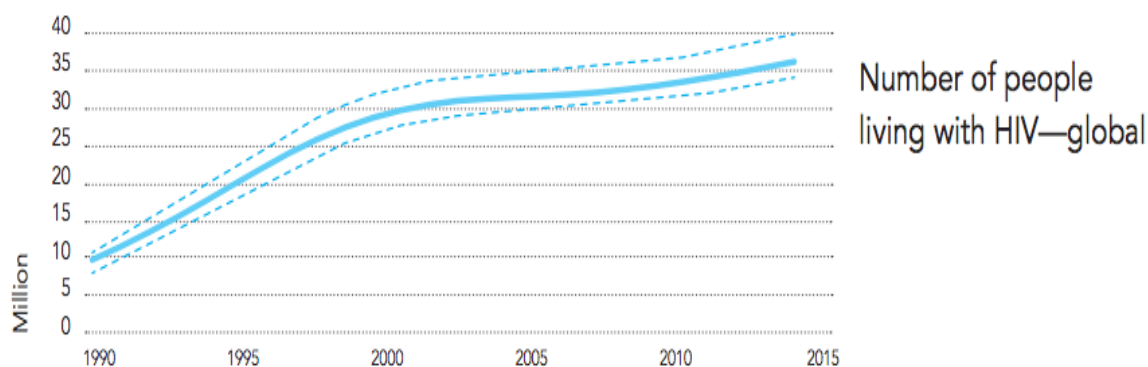


Figure 1 : Répartition des personnes vivant avec le VIH en 2012 (a) et évolution du nombre de personnes vivant avec le VIH-1 de 1990 à 2015 (b) Plus de 36,7 millions de personnes vivent avec le SIDA. Le continent le plus durement touché est l'Afrique avec près de 25 millions de personnes infectées (66%) et une fréquence d'infection en hausse en Afrique du Nord (a) Le taux annuel de nouvelles infections a chuté de 21% entre 1997 et 2010 cependant, le nombre de nouvelles infections en 2015 est de 2,1 millions (b) (Maartens *et al*, 2014) (UNAIDS - AIDS by the numbers - 2016).

Introduction

1. Le VIH-1 et pathologie associée : le SIDA

1.1. Le syndrome d'immunodéficience acquise « SIDA »

Le SIDA (Syndrome de l'immunodéficience humaine acquise) est caractérisé par une infection conduisant à une élimination progressive du système immunitaire de l'hôte infecté, le rendant vulnérable à de multiples infections opportunistes. Le virus responsable de ce syndrome est le VIH (Virus de l'immunodéficience humaine). Ce rétrovirus infecte l'homme et touche aujourd'hui plus de 36,7 millions de personnes dans le monde (UNAIDS- AIDS by the numbers - 2016). **(Figure 1)**

a) Historique et épidémiologie

L'épidémie de SIDA commence en juin 1981 quand le CDC (Center for Disease Control) repère une recrudescence d'infections opportunistes telles que des pneumonies à *Pneumocystis carinii* et des sarcomes de Kaposi chez les jeunes homosexuels américains. Il est observé chez ces patients une baisse anormale du taux de lymphocytes T CD4+ jouant un rôle essentiel dans le système immunitaire.

Deux équipes de chercheurs l'une à l'Institut Pasteur à Paris et l'autre au National Cancer Institut aux États-Unis travaillent sur l'identification de l'agent pathogène et c'est en démontrant la présence d'une activité de transcription inverse que l'origine rétrovirale de l'épidémie est déterminée. Plusieurs virus sont alors mis en cause comme le HLTV (Human T-Cell Lymphotropic Virus) en raison de fortes similitudes d'action (Popovic *et al*, 1984). Il faudra cependant attendre 1983, pour que l'équipe de Luc Montagnier identifie l'agent causal du SIDA. Ce virus sera préalablement appelé *Lymphadenopathy Associated Virus* (LAV) puis VIH-1 (Virus de l'immunodéficience humaine de type 1) en 1986 (Barré-Sinoussi *et al*, 1983; Coffin *et al*, 1986).

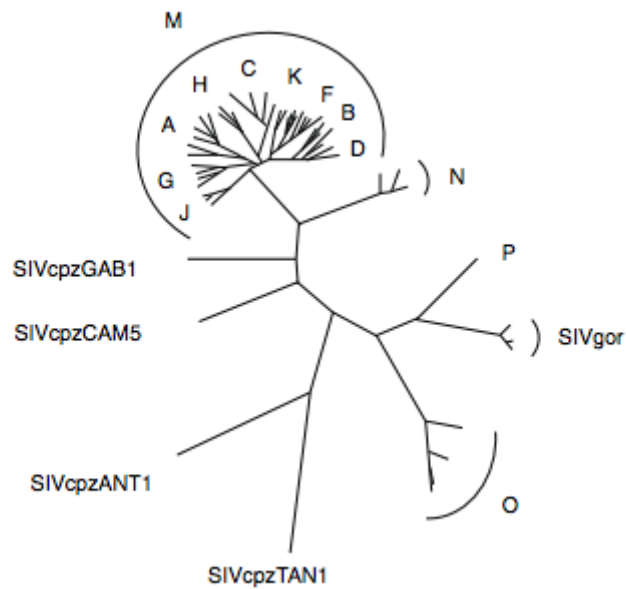


Figure 2 : Arbre phylogénétique du VIH-1 et du SIV (Hemelaar, 2012). Le VIH se décompose en 4 groupes majoritaires (M, N, O, P). Le groupe M comprend 9 sous-types (A à D, F à H, et J à K). SIVcpz (Simian immunodeficiency virus) correspond aux virus identifiés chez le chimpanzé et le SIVgor correspond à celui retrouvé chez le gorille.

b) Transmission du virus

Le VIH se transmet à l'homme par trois voies majoritaires.

- Par le sang : concerne les usagers de drogues injectables, les transfusés et les professionnels de santé.
- Par contact sexuel : se fait par contact entre les sécrétions sexuelles et les muqueuses génitales.
- Par voie verticale : *in-utero* : via une transmission transplacentaire, *Intra-partum* : durant l'accouchement via un contact sanguin et les sécrétions vaginales, *Post-partum* : suite à l'accouchement via l'allaitement maternel.

c) Diversité génétique

La variabilité du virus de l'immunodéficience humaine de type 1 (VIH-1) est très élevée. Les virus sont classés en deux types : Le VIH-1, type majoritaire et le VIH-2 type endémique présente 50 % d'homologie avec le VIH-1 au niveau des gènes *gag* et *pol* et apparaît comme moins infectieux. Le VIH-1 se décompose en plusieurs groupes (**Figure 2**) dont le groupe majoritaire dénommé M (Majeur), le groupe *outlier* dénommé O, les très rares variants des groupes N (New ou Non M, Non O) et un quatrième groupe appelé P (Plantier *et al*, 2009). C'est le VIH-1 du groupe M qui est responsable de la pandémie de SIDA. À l'intérieur du groupe M, il existe 9 sous-types (de A à D, F à H et J à K) dont le sous-type majoritaire est le sous-type C qui représente près de 50 % des infections dans le monde. Il existe également des variations au sein des sous-types tels que le sous-type F qui peut être divisé en F1 et F2 et le sous-type A en A1, A2, A3 et A4. S'ajoutent ensuite les CRF (circulating recombinant form) issues de mélanges dans le génome viral entre plusieurs sous-types lors de co-infections (Hemelaar, 2012). En France, la présence de sous-type B prédomine, mais la circulation d'autres sous-types ou de recombinants augmente tel que le CRF02_AG.

d) Pathologie

Les principales cibles sont les cellules présentant les récepteurs CD4 à leur surface. Ces cellules comprennent les lymphocytes T CD4+, les monocytes et macrophages, ainsi

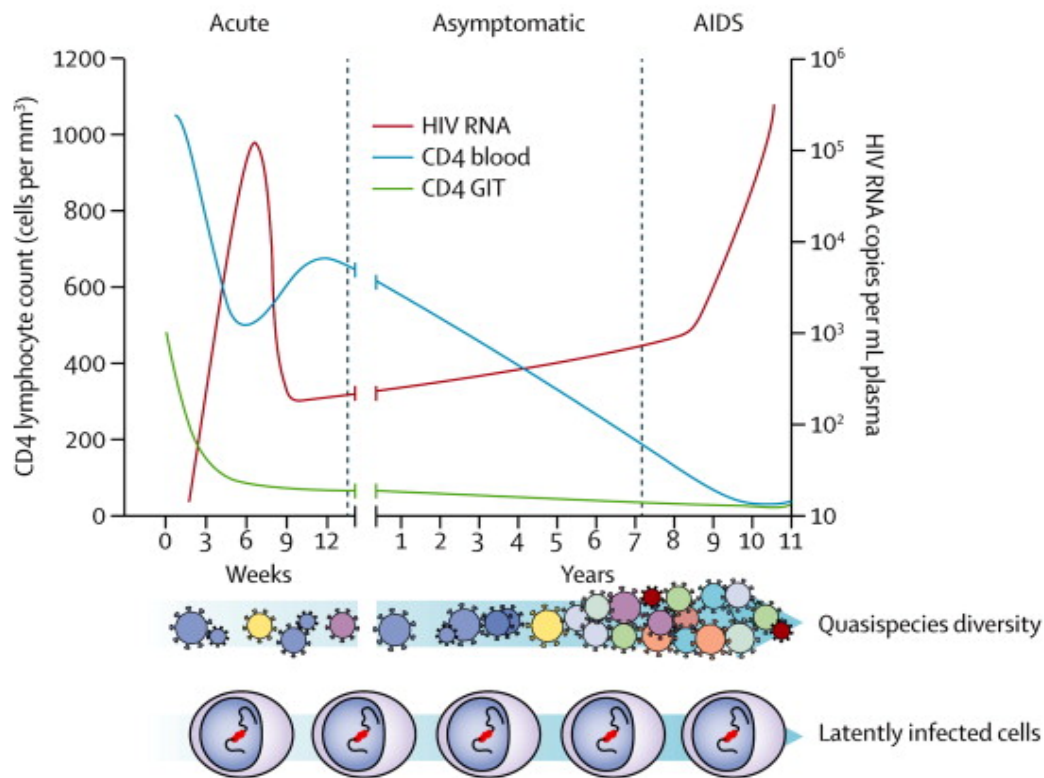


Figure 3 : Phases caractéristiques de l'infection du VIH-1 en l'absence de traitement antirétroviral (Maartens *et al*, 2014). La primo-infection caractérisée par une répllication importante du virus entraînant la destruction des cellules immunitaires visualisée par une augmentation des copies d'ARN du VIH-1 et une diminution du nombre de lymphocytes CD4 sanguins et des lymphocytes CD4 du tractus gastro-intestinal permettant de contenir l'infection. Durant la phase asymptomatique, le virus se multiplie progressivement entraînant la chute lente des lymphocytes CD4. La phase SIDA est caractérisée par l'épuisement total des lymphocytes CD4 sanguins. L'individu est alors immunodéprimé et de multiples pathologies opportunistes ne peuvent plus être contrôlées par le système immunitaire.

que les cellules dendritiques. L'infection par VIH-1 évolue en plusieurs étapes (**Figure 3**) se succédant dans le temps et conduisant à une destruction du système immunitaire de l'hôte. On distingue trois phases caractéristiques à l'infection du VIH-1 (Maartens *et al*, 2014).

- La primo-infection

La primo-infection commence après la transmission du virus, elle est caractérisée par une réplication importante du virus entraînant la destruction des lymphocytes CD4+. Cette étape est suivie par une réponse inflammatoire importante caractérisée par l'apparition d'anticorps anti-VIH par stimulation des lymphocytes B et une réponse cytotoxique des lymphocytes CD8+ permettant de contenir l'infection. C'est cette production d'anticorps anti-VIH qui rend le patient séropositif.

- La phase asymptomatique

Durant la phase asymptomatique qui est de durée variable (quelques mois à plusieurs années), le virus se multiplie activement diminuant progressivement le taux de lymphocytes CD4+. Le système immunitaire contrôle l'infection jusqu'à l'épuisement des CD4+.

- La phase symptomatique ou phase SIDA

Cette phase est caractérisée par l'apparition des premiers signes cliniques de l'infection. Le sujet sans traitement est immunodéprimé et des pathologies infectieuses ne peuvent plus être contrôlées par le système immunitaire préalablement détruit qui conduit alors au décès du patient.

1.2. Le Virus de l'immunodéficience humaine « VIH »

a) Classification

Le VIH-1 est un rétrovirus appartenant à la famille des *Retroviridae* appartenant au genre *Lentivirus*. D'après l'*International Committee on Taxonomy of Viruses* (ICTV), on distinguera 7 genres de rétrovirus (**Figure 4**) selon divers critères tels que la pathogénicité, la morphologie, la structure du génome, les propriétés antigéniques parmi lesquelles :

Retroviridae

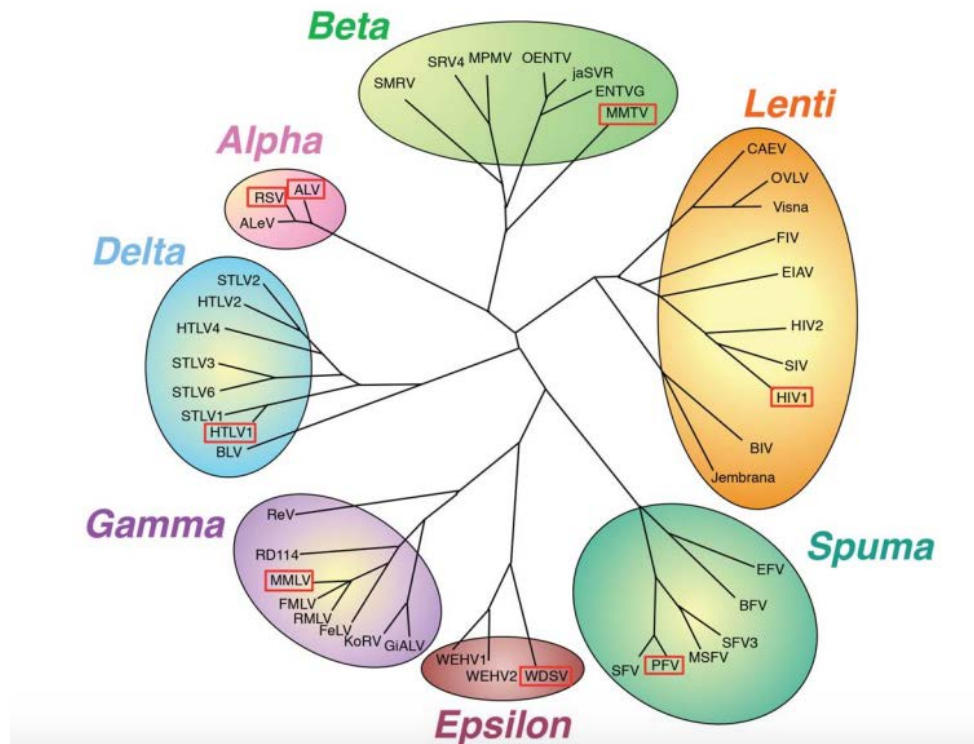


Figure 4 : Arbre phylogénétique des rétrovirus basé sur l'alignement des séquences de la RT et de l'IN (Skala, 2014).

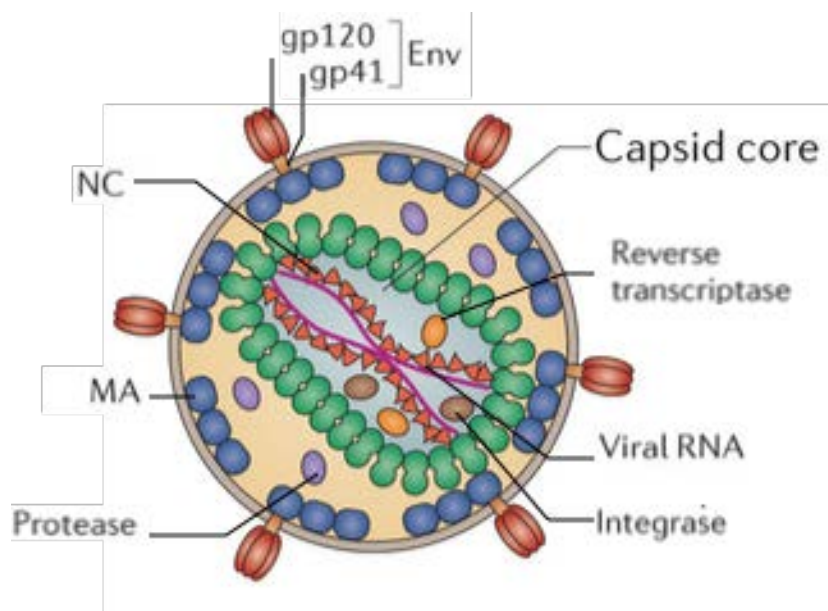


Figure 5 : Structure schématique d'une particule du VIH-1 mature. Adapté de (Campbell & Hope, 2015) (Freed, 2015) MA : Matrice, NC : Nucléocapside, Capsid core : Capside. Le virus du VIH-1 est un rétrovirus d'environ 100 nm de diamètre. Ce virus enveloppé possède un génome à ARN de polarité positive monocaténaire diploïde
Des rétrovirus simples dont le génome code uniquement pour des protéines Gag, Pol et Env :

- les Alpharétrovirus (ALV : Avian Leukosis Virus)
- les Betarétrovirus (MMTV : Mouse Mammary Tumor Virus)
- les Gammarétrovirus (MLV : Murine Leukemia Virus)

Des rétrovirus complexes codants en plus des protéines Gag, Pol, Env possèdent dans leur génome des séquences codantes pour des protéines régulatrices :

- les Deltarétrovirus (HLTV : Human T-Lymphotropic Virus)
- les Epsilonrétrovirus (WDSV : Walleye Dermal Sarcoma Virus)
- les Spumavirus (HFV : Human Foamy Virus)
- les Lentivirus (HIV : Human Immunodeficiency Virus) sont des virus responsables d'infections virales à développement lent.

b) Structure du virion

La particule virale (**Figure 5**) dont le diamètre est de 80 à 120 nm comporte une enveloppe externe en bicouche lipidique provenant de la cellule productrice dans laquelle sont ancrées des glycoprotéines de surface gp120 et gp41. Gp120 et gp41 s'organisent en trimères formant une structure en spicule responsable de l'attachement des virions sur les récepteurs CD4 et corécepteurs CCR5 ou CXCR4 des cellules cibles. L'intérieur de la particule virale comprend une membrane interne, constituée de protéines de matrice (MAp17) fixée à la bicouche lipidique par son extrémité N-terminale. Le génome ARN du virus est protégé par une nucléocapside protéique (NCp7), incluse dans une capsidie formée d'hexamères de p24 (CAp24) de forme conique.

L'autocatalyse de la polyprotéine Gag-Pol génère la protéase (PRp10) active sous forme d'homodimère. La protéase permet la maturation des différents précurseurs Gag et Gag-Pol en libérant premièrement les protéines de structure puis dans un second temps les enzymes virales RT et IN. La transcriptase inverse (p66/p51) est une enzyme hétérodimérique assurant la synthèse d'ADN proviral à partir de l'ARN génomique viral. Cette enzyme catalyse deux activités principales : une activité ADN polymérase, et une activité RNase H ARN/ADN dépendant incapable d'édition qui génère une forte variabilité observée chez le VIH-1. L'intégrase (INp32) est responsable de l'insertion stable de l'ADN proviral dans l'ADN génomique de la cellule infectée.

En plus de toutes ces protéines, le VIH-1 possède des protéines dites accessoires telles que Tat (Trans-Activator for viral Transcription), Vpu (Viral protein U), Nef (Negative Factor), Rev (Regulator of Expression of Viral proteins), Vif (Viral infectivity Factor) et Vpr

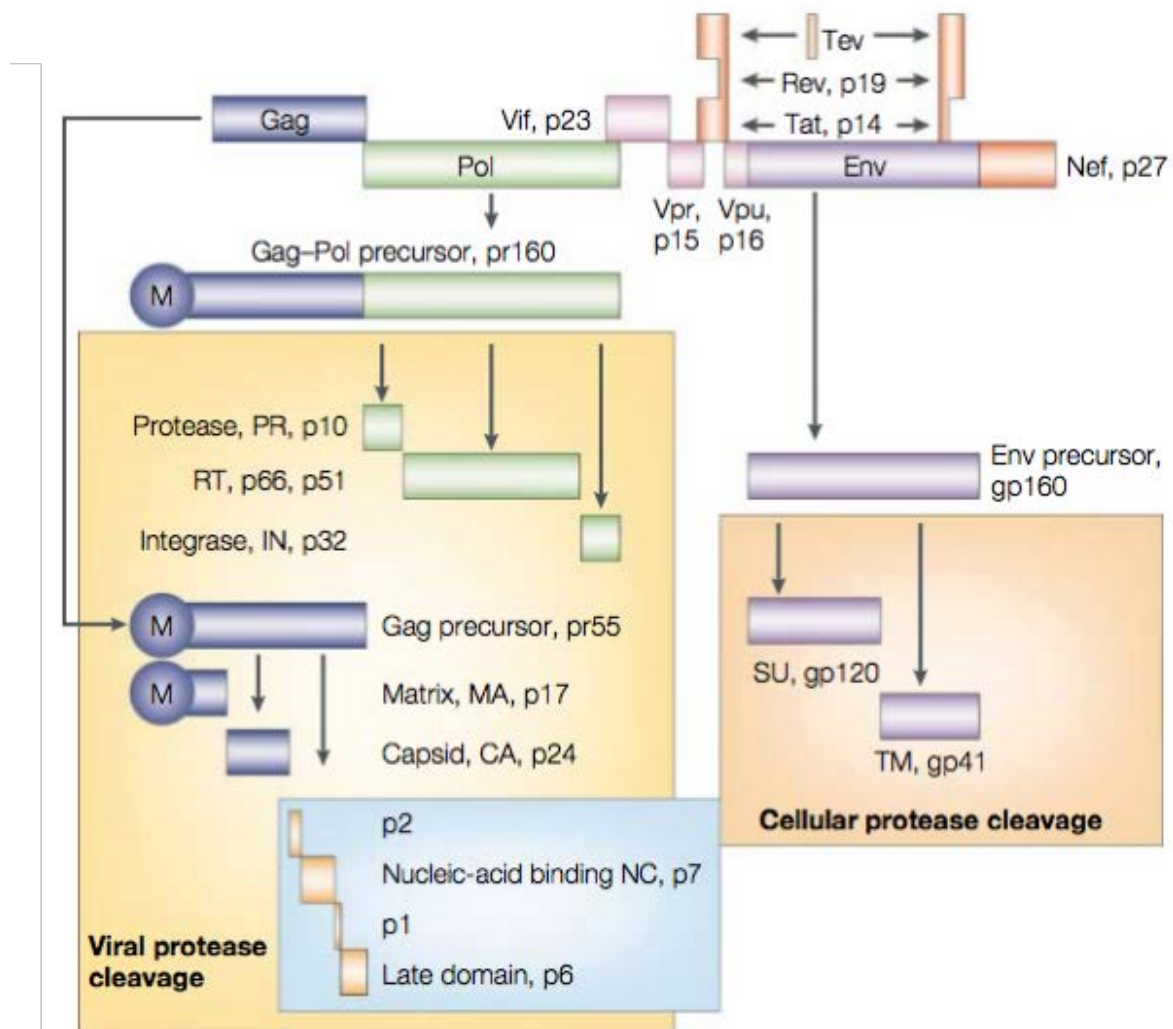


Figure 6 : Organisation génomique du VIH-1 et représentation des précurseurs protéiques et des protéines issues de leur maturation (Peterlin & Trono, 2003). L'ORF *gag* code pour un précurseur pr55^{gag} clivé par la protéase virale en 4 protéines : MAp17, CAp24, NCp7 et le peptide C-terminal P6 ainsi que deux spacers : le peptide p2 et le peptide p1. Le précurseur pr160^{gag-pol} est clivé par la protéase virale en trois enzymes : la protéase PRp10, l'intégrase INp32 et la transcriptase inverse RTp66/p51. Le gène *env* code pour le précurseur gp160 qui est clivé par une protéase cellulaire en deux protéines, la glycoprotéine de surface (SU) gp120 et la glycoprotéine transmembranaire (TM) gp41. En plus des protéines Gag, Pol et Env, le génome du VIH-1 code pour des protéines régulatrices Tat et Rev, mais aussi des protéines accessoires Vpu, Vpr, Vif et Nef.

(Viral protein R) jouant des rôles régulateurs au niveau de plusieurs étapes du cycle viral. Ces gènes sont situés dans différents cadres ouverts de lecture. Seules Nef, Vif et Vpr sont retrouvées dans les virions matures.

c) Génome du VIH-1

Le génome rétroviral est constitué de deux copies d'ARN simple brin de polarité positive de 9181 nucléotides. L'ARN viral est encadré de séquences répétées « R », de la région U5 en 5' et U3 en 3' qui sera lors de la transcription inverse convertie en ADN permettant la création aux extrémités de séquences répétées non codantes appelées LTR (Long Terminal Repeat) composé des régions U3, R et U5 jouant un rôle crucial dans la réplication virale. Le VIH-1 comporte plusieurs gènes communs à l'ensemble des rétrovirus codants pour les protéines de structure et les enzymes de réplication ainsi que les gènes accessoires (**Figure 6**). On trouve le gène *gag* (group antigen) codant la polyprotéine structurale du virus Pr55gag maturée par la protéase virale en protéine de matrice MA (p17), de capsid (p24), en nucléocapsid NC (p7) et en peptide SP1, SP2 et p6. Le gène *pol* code pour les enzymes virales, protéase, transcriptase inverse et l'intégrase. Le gène *env* code quant à lui pour la polyprotéine gp160 précurseur des glycoprotéines de surface gp120 et gp41. On retrouve ensuite les gènes codants pour *tat*, *rev*, *vpr* et les autres régulateurs *vif*, *nef* et *vpu* (Pluta & Kacprzak, 2009).

d) Les protéines virales

i- Glycoprotéines d'enveloppe

Les glycoprotéines du VIH-1 sont initialement exprimées à partir du précurseur gp160 qui est clivé en deux protéines associées de façon non covalente en gp120 et gp41. La glycoprotéine virale gp120 est responsable de l'interaction du virus avec le récepteur CD4 et le co-récepteur CCR5 ou CXCR4, la glycoprotéine gp41 est responsable de la fusion de l'enveloppe virale avec la membrane cellulaire (Blumenthal *et al*, 2012).

Les glycoprotéines gp120 et gp41 s'associent pour former un complexe gp120/gp41 sous forme de trimère contenant trois unités de gp120 et trois unités de gp41. La glycoprotéine gp120 est organisée en cinq régions constantes (C1-C5) et cinq régions variables appelées boucles variables (V1-V5). La glycoprotéine gp41 contient quatre domaines fonctionnels : le peptide fusion, le domaine heptad repeat1 (HR1), le domaine

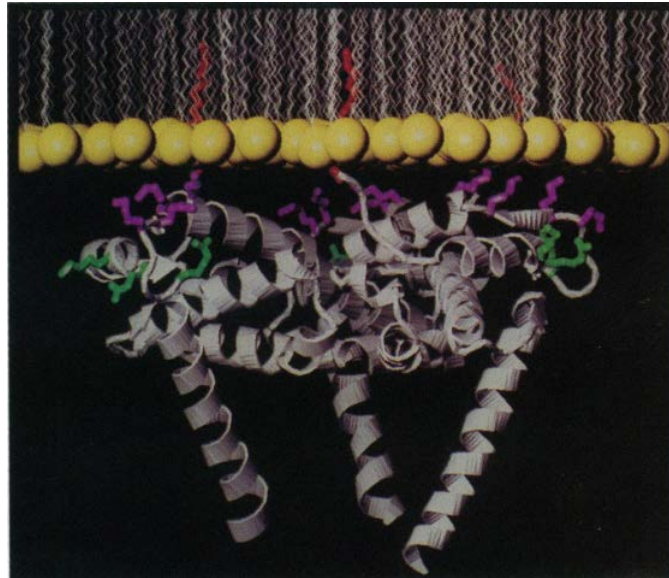


Figure 7 : Modèle schématisant les interactions entre un trimère de MA et la membrane (Hill *et al*, 1996).

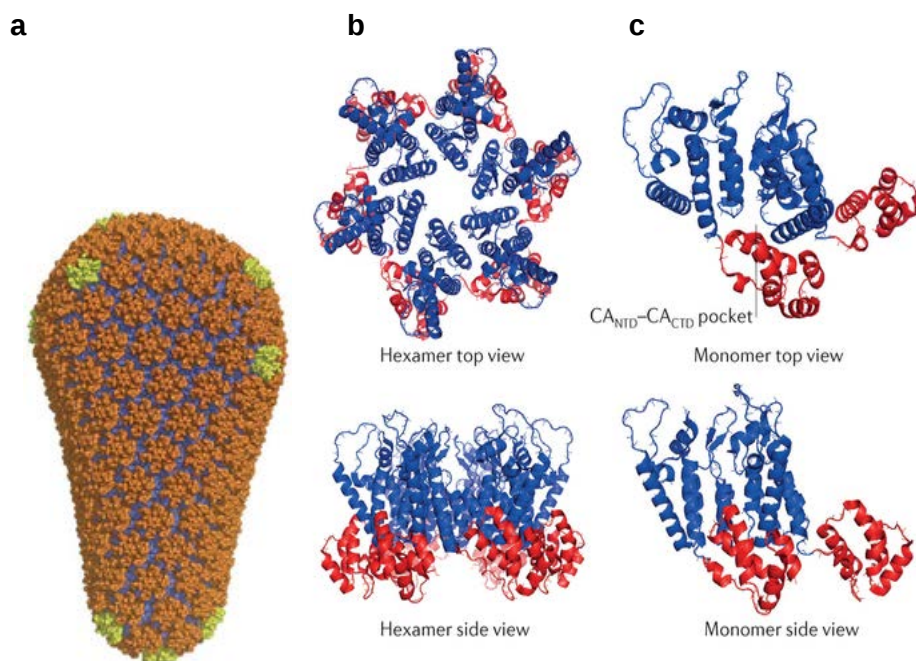


Figure 8 : Structure de la capsid de VIH-1, d'un dimère et un hexamère de capsid (Campbell & Hope, 2015; Pornillos *et al*, 2011). **(a)** Modèle d'assemblage de la capsid conique contenant des hexamères de CA (orange) et des pentamères de CA (jaune). **(b)** Structure d'un hexamère de CA (PDB : 3GV2). **(c)** Structure de deux monomères de CA de l'hexamère illustrant la poche formée par le CA_{NTD} et le CA_{CTD} permettant l'interaction avec des protéines cellulaires.

heptad repeat 2 (HR2) et le domaine transmembranaire qui permet l'ancrage de la gp41 (Buzon *et al*, 2010).

ii- Les protéines structurales

Les protéines structurales sont issues du précurseur Gag Pr55 (Pr55gag) maturé par la protéase virale.

- La matrice (p17, MA)

La matrice est le domaine N-terminal du précurseur polyprotéique Pr55gag. la protéine p17 est une protéine de 132 acides aminés qui recouvre la surface interne de l'enveloppe virale (**Figure 7**). La structure tridimensionnelle de la protéine p17 a pu être déterminée par résonance magnétique de surface (RMN) et par cristallographie aux rayons X (Massiah *et al*, 1994, 1996). La protéine p17 s'assemble en trimère pour former des hexamères de trimère. Cette protéine possède une séquence signal de localisation nucléaire, une séquence d'export nucléaire et d'une séquence de transport à la membrane (Bukrinskaya, 2007). La protéine p17 joue un rôle clé dans plusieurs étapes précoces et tardives de la réplication virale.

- La capsid (p24, CA)

Pendant la maturation du virion, la capsid p24 est libérée de la polyprotéine Pr55gag par la protéase virale, les monomères de capsid s'assemblent pour former des structures hexamériques et pentamériques. Cette protéine possède deux domaines structurés : le domaine N-terminal (CA-NTD) et le domaine C-terminal (CA-CTD) connecté par un linker flexible (**Figure 8**). Les interactions entre les oligomères de capsid permettent la formation de la structure conique (Briggs *et al*, 2009; Pornillos *et al*, 2011; Fassati, 2012; Campbell & Hope, 2015).

- La nucléocapsid (p7, NC)

La nucléocapsid est une protéine de 55 acides aminés caractérisée par la présence de deux domaines à doigts de zinc CCHC (Mirambeau *et al*, 2010). Ces domaines sont



Figure 9 : Structure d'un homodimère de protéase en complexe avec TL-3 (Brik & Wong, 2003). (PDB : 3TLH). TL-3 est un inhibiteur de protéase.

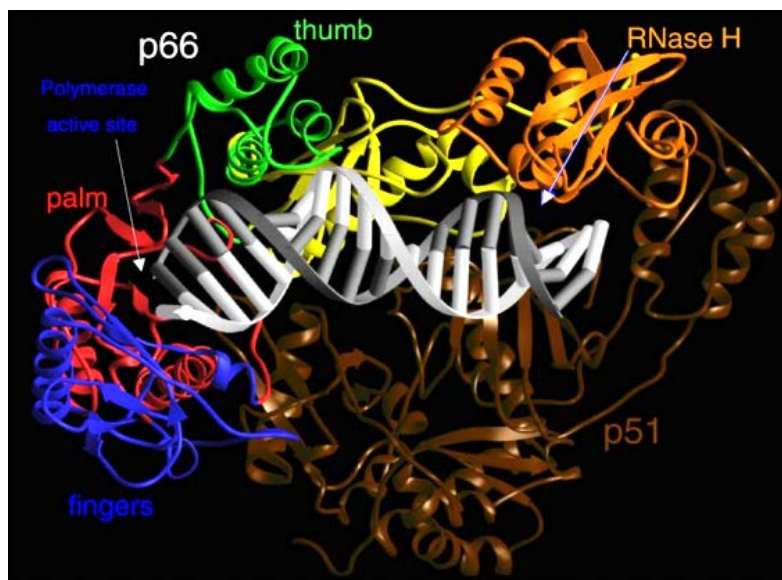


Figure 10 : Représentation du complexe de la RT avec un acide nucléique (Sarafianos *et al*, 2009). Le monomère p66 constitué des doigts (bleu), du pouce (vert), la paume (rouge) et du domaine RNase H (orange). Le monomère p51 est représenté en marron.

nécessaires à la réplication virale et participent directement à la reconnaissance du génome et à son encapsidation. Cette protéine possède une activité chaperonne et permet la protection de l'ARN viral (Darlix *et al*, 1995). Plus récemment, il a été montré l'implication de la NC lors de la transcription inverse (Thomas & Gorelick, 2008/6). La structure de la nucléocapside a été résolue par RMN (De Guzman *et al*, 1998).

- La protéine p6

Cette protéine de 52 acides aminés est codée par la région la plus polymorphique du gène *gag*. Cette protéine subit de nombreuses modifications post-traductionnelles telles que des sumoylation, ubiquitination et phosphorylation. Elle est décrite comme permettant l'incorporation de Vpr lors du bourgeonnement du virion (Friedrich *et al*, 2016).

iii- Les protéines non structurales

- La protéase (p10, PR)

La protéase virale est une aspartate protéase qui est active sous sa forme dimérique (**Figure 9**). Le site actif de l'homodimère est localisé à l'interface des dimères et est constitué d'une triade catalytique Asp25-Thr26-Gly27. Dans la structure obtenue par diffraction des rayons X, chaque monomère est constitué de 99 acides aminés (Navia *et al*, 1989). La protéase virale est essentielle à la maturation des précurseurs polyprotéiques Pr55gag.

- La transcriptase inverse (p66/p51, RT)

La RT est active sous forme hétérodimérique composée de deux sous-unités, la p66 de 560 acides aminés et la p51 de 440 acides aminés (**Figure 10**). La sous-unité p51 est issue d'un clivage par la PR réalisé dans le virion du dimère p66/p66 entre la phénylalanine F440 et la tyrosine Y441 qui est réalisé dans le virion. La RT comporte plusieurs activités enzymatiques : une activité ADN polymérase ADN ou ARN dépendante et une activité RNase H qui dégrade le brin d'ARN dans un hybride ARN-ADN. La RT nécessite la présence de cations divalents Mg²⁺ pour ses deux activités. C'est la sous-unité p66 qui comporte ces activités au sein de l'hétérodimère. Le domaine N-terminal porte l'activité ADN

polymérase et le domaine C-terminal contient l'activité RNaseH. La structure tridimensionnelle de la RT a permis de déterminer quatre sous domaines au domaine polymérase : les doigts (1-85 et 118-155), la paume (86-117 et 156-236), le pouce (237-318) et le domaine de connexion (319-426) (Jacobo-Molina *et al*, 1993; Kohlstaedt *et al*, 1992).

- L'intégrase (p32, IN)

L'IN est une protéine de la famille des poly-nucléotidyl-transférases. Cette enzyme se compose de trois domaines et catalyse l'étape d'intégration virale. L'intégrase étant l'objet majeur de ma thèse, elle sera décrite plus précisément dans la partie 2.

iv- Les protéines régulatrices

- Tat

Tat est une protéine régulatrice de taille variable entre 84 et 104 acides aminés en fonction des isolats. Cette protéine contient 6 différents domaines. La région I (1-21) est un domaine riche en proline et possède un tryptophane en position 11 conservée. La région II (22-37) possède six cystéines conservées. La région III (38-48) contient une séquence conservée ⁴³LGISYG⁴⁸. La région IV (49-59) riche en résidus basiques possède une séquence conservée ⁴⁹RKKRRQRRAP⁵⁹. La région V (60-72) riche en glutamine qui est le domaine le moins conservé et la région VI qui constitue le domaine C-terminal de Tat. Tat est impliquée dans la régulation de la transcription des gènes via l'interaction avec une structure ARN de 59 nucléotides appelée TAR (transactivation-responsive region) (Romani *et al*, 2010; Musinova *et al*, 2016). De plus, Tat peut être libérée par la cellule infectée et être internalisée par des cellules non infectées.

- Rev

Rev est une protéine de 116 acides aminés contenant un domaine NLS (Nuclear localization signal) situé dans le domaine N-terminal de la protéine permettant son importation au noyau, le domaine C-terminal contient un signal d'export nucléaire.

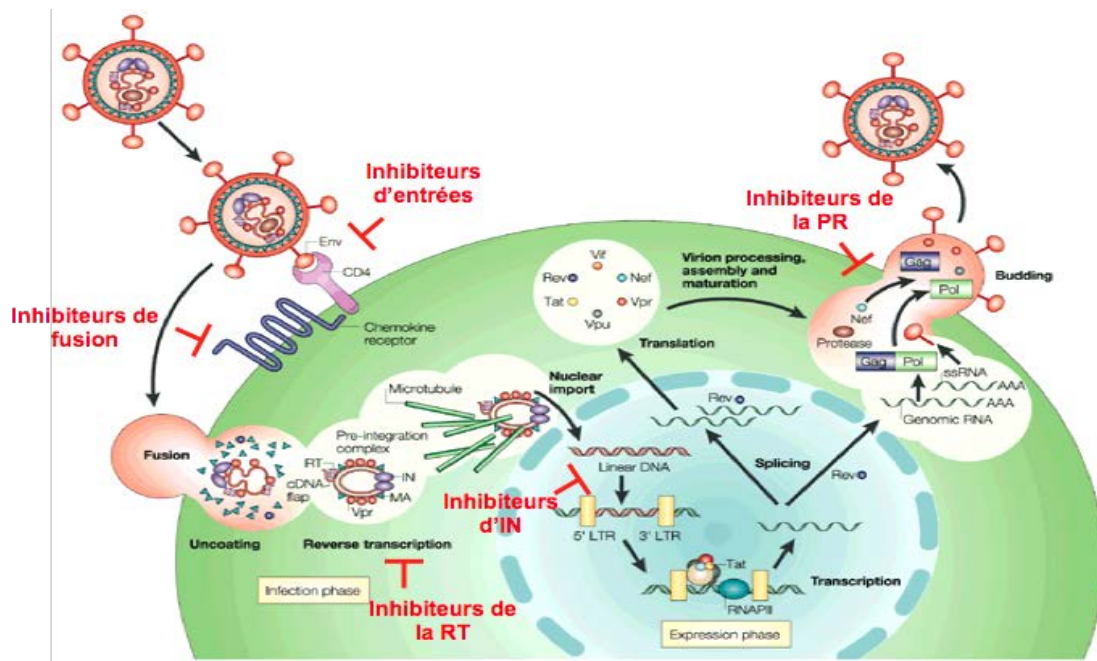


Figure 11 : Représentation du cycle réplcatif du VIH-1 et cibles thérapeutiques actuelles. Adapté de (Peterlin & Trono, 2003). Lors des étapes précoces du cycle réplcatif viral, il y a entrée du virion par fusion des bicouches lipidiques. C'est au sein du complexe de pré-intégration que l'étape de rétrotranscription va permettre la synthèse de l'ADN viral. Ce complexe est ensuite transporté au noyau afin de permettre l'intégration de l'ADN viral dans l'ADN de la cellule. Il y a ensuite transcription et traduction des protéines virales permettant la production et la libération de virions. L'ensemble des cibles thérapeutiques utilisé aujourd'hui en thérapie touche les différentes étapes du cycle réplcatif.

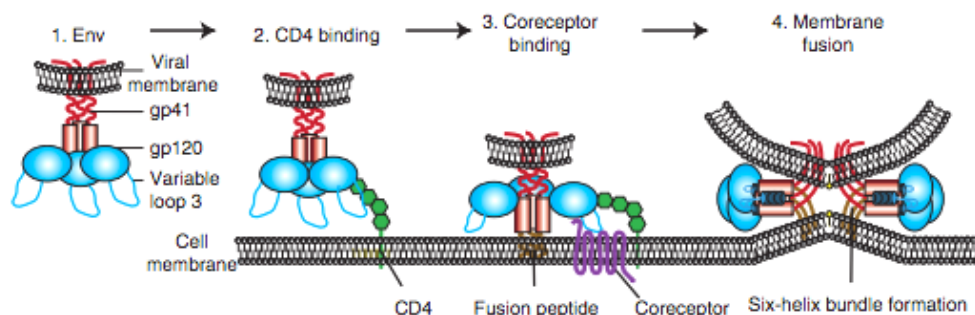


Figure 12 : Représentation schématique des différentes étapes d'entrée du VIH-1 dans la cellule cible (Wilen *et al*, 2012). (1) L'entrée nécessite les protéines Env comprenant les sous-unités gp120 (bleu) et gp41 (rouge). (2) Interaction entre la gp120 et le récepteur cellulaire CD4. (3) Cette interaction va permettre un changement de conformation de Env permettant l'interaction de gp41 avec le corecepteur CCR5 ou CXCR4. (4) Cette interaction induit un changement de conformation et permet l'insertion du peptide fusion de gp41 dans la membrane de la cellule cible induisant la fusion des deux membranes.

Rev en se fixant sur les ARNm viraux non ou mono-épissés au niveau de la séquence RRE (Rev response element) permet leurs transports vers le cytoplasme (Groom *et al*, 2009).

v- Les protéines accessoires

- Vpu

Vpu est une protéine membranaire de 81 acides aminés. Il est décrit que Vpu induit la dégradation par le protéasome du CD4 dans le reticulum endoplasmique par le recrutement d'ubiquitin ligase (Margottin *et al*, 1998/3). Vpu permet aussi la neutralisation de facteurs cellulaires améliorant alors la libération des virions.

- Vpr

Vpr est une protéine de 96 acides aminés. Cette protéine est décrite comme jouant un rôle important dans l'import nucléaire du CPI (Complexe de pré-intégration) dans les macrophages (Connor *et al*, 1995). Vpr est aussi impliqué dans l'arrêt du cycle cellulaire en phase G2, l'apoptose et la prolifération cellulaire, cependant le rôle exact de ce processus reste encore mal connu.

- Vif

Vif est une protéine de 23 kDa, elle induit la dégradation par le protéasome de la protéine APOBEC3G qui est un facteur cellulaire de restriction à l'infection virale (Mangeat *et al*, 2003).

- Nef

Nef est une protéine de 27 à 35 kDa, elle est décrite comme permettant l'inhibition d'expression des protéines cellulaires : CD4, CD8, CD28, CMH-I et le CMH-II (Li *et al*, 2005).

e) Cycle viral

Le cycle viral se subdivise en deux phases majeures : la reconnaissance spécifique et l'entrée dans la cellule cible ainsi que le transport de l'intasome viral vers le noyau et

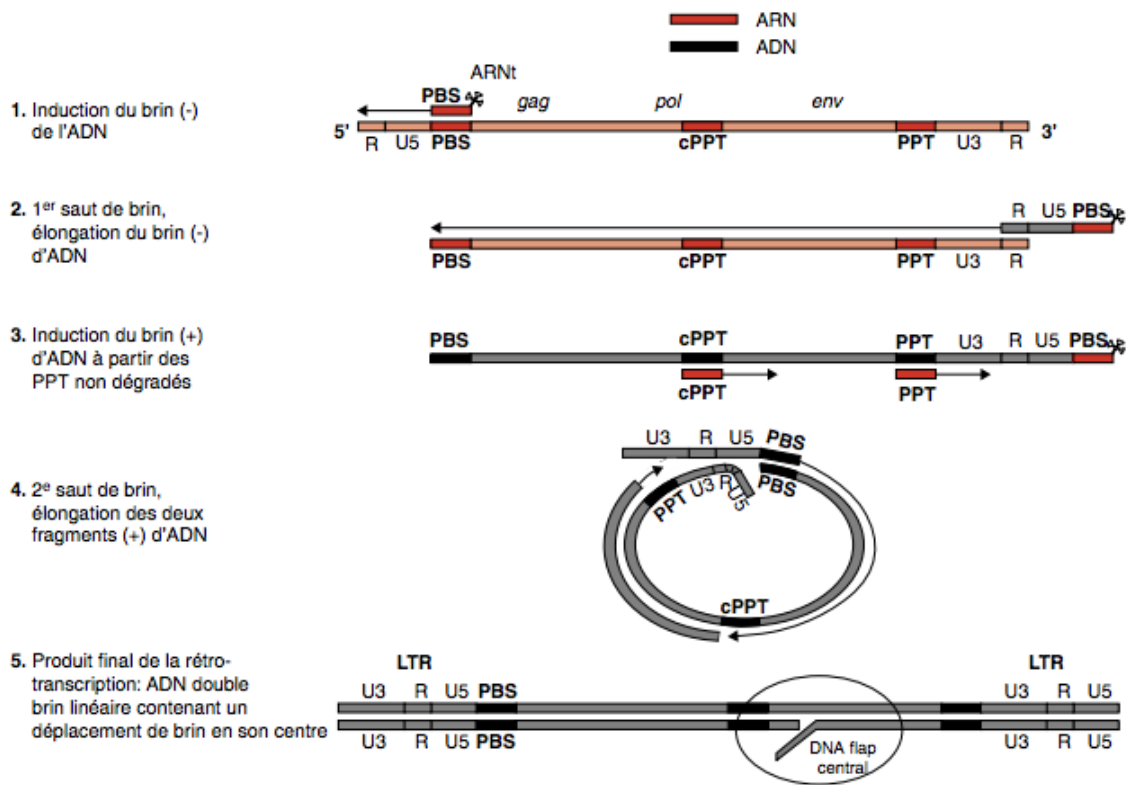


Figure 13 : Étapes de la transcription inverse du VIH-1 (Jacquot & Benichou, 2006). (1) Pour initier la transcription inverse, la RT du VIH-1 a besoin d'une matrice et d'une amorce qui est l'ARN^{lys3}. Cet ARN^{lys3} va se fixer au niveau de la séquence PBS (primer binding site) située en 5' de l'ARN viral et permet la synthèse de l'ADN (-) jusqu'à l'extrémité 5' de l'ARN viral. Un duplex ADN/ARN néosynthétisé est alors généré. (2) L'activité RNaseH de la RT va dégrader l'extrémité 5' de l'ARN viral permettant le 1^{er} saut de brin. L'ADN (-) est alors transféré à l'extrémité 3' de l'ARN grâce à des séquences R répétées. La synthèse de l'ADN (-) se poursuit alors jusqu'à l'extrémité 5'. (3) La RNase H va ensuite dégrader le reste de l'ARN en excluant les régions PPT (Polypurine tract). Ces régions vont servir d'amorces à la synthèse d'une portion de l'ADN (+). (4) Un deuxième saut de brin a lieu permettant l'hybridation des séquences PBS des deux brins et la synthèse complète de l'ADN (+). (5) L'ensemble de ces étapes aboutit à un ADN double brin linéaire comportant à chaque extrémité les LTR viraux composés des séquences U3-R-U5. La génération du « DNA flap central » durant la synthèse de l'ADN (+) joue un rôle dans l'import nucléaire de l'ADN viral.

l'intégration de l'ADN viral dans le génome de la cellule. Puis, dans des étapes plus tardives, il y a transcription et traduction des protéines virales permettant la production ainsi que libération par bourgeonnement des nouveaux virions. L'ensemble de ces étapes fait intervenir de nombreux facteurs cellulaires (**Figure 11**).

i- Reconnaissance spécifique, entrée du virus et intégration

Le cycle commence par l'interaction spécifique entre la glycoprotéine gp120 du virus avec les récepteurs membranaires CD4 (cluster of differentiation 4) à la surface de la cellule cible (**Figure 12**) (Maddon *et al*, 1986). Cette interaction provoque une modification conformationnelle de la gp120. Le corecepteur est soit le CCR5 (C-C chemokine receptor 5) soit le CXCR4 (C-X-C chemokine receptor 4). La liaison aux corécepteurs CCR5 ou CXCR4 va ensuite provoquer un changement conformationnel de la glycoprotéine gp41 permettant la fusion des deux bicouches lipidiques et la libération de la capside virale dans le cytoplasme de la cellule (Wilén *et al*, 2012). Puis, il y a désassemblage de la capside pour former le complexe de pré-intégration (CPI) quand l'ARN viral est converti en ADN double brin grâce à la transcriptase inverse (RT) du virus, activité caractéristique des rétrovirus (**Figure 13**) (Hu & Hughes, 2012). Ce complexe est ensuite activement transporté au noyau via le réseau microtubulaire afin de permettre son import dans le noyau via un transport actif par les pores nucléaires. L'intégration de l'ADN viral sera ensuite catalysée par l'intégrase virale (IN) ciblant majoritairement des régions transcrites du génome permettant la production de multiples virions.

ii- Production et libération de nouveaux virions

Une fois intégré, le virus est alors sous sa forme de provirus. Le LTR en 5' est un promoteur viral. En effet, il y a recrutement au niveau des séquences régulatrices de plusieurs facteurs favorisant la transcription via l'ARN polymérase de type II au niveau du LTR viral (**Figure 14**). En amont du site d'initiation de la transcription (InR), une TATA box ainsi que trois sites de liaison pour SP1 (Specific protein 1) permettent le recrutement et le positionnement correct de l'ARN polymérase de type II au site d'initiation de la transcription. De plus, des régions d'amplification vont permettre la fixation de NF- κ B (nuclear factor- κ B), NFAT (nuclear factor of activated T cells) et des facteurs de la famille ETS. La première étape de transcription est dite « Tat indépendante » et permet la production de faible

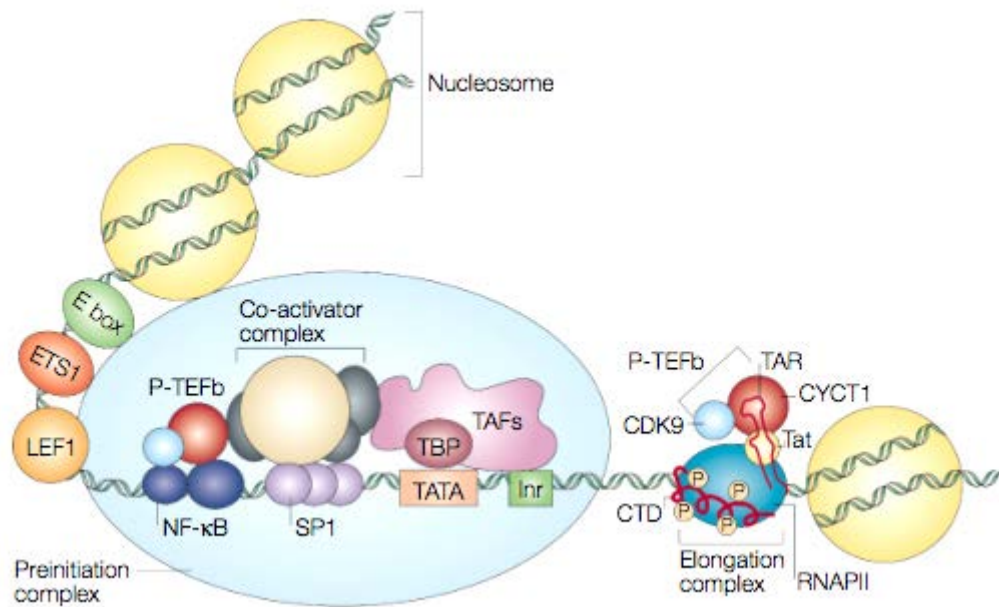


Figure 14 : Représentation du LTR du VIH-1 (Peterlin & Trono, 2003). Le promoteur viral contient la séquence initiatrice (Inr) et la « TATA box » (TATA). Ces domaines sont reconnus par les « TATA-binding protein » (TBP) et les « TBP-associated factors » (TAFs). Ces domaines sont flanqués en amont par 3 sites de fixation à SP1 et en aval par une séquence TAR. Le complexe de co-activation est capable de se fixer au domaine SP1 afin de recruter l'ARN polymérase de type 2 (RNAPII). Le facteur P-TEFb (Positive transcription elongation factor b) contenant CYCT1 (cyclin T1), CDK9 (cyclin-dependant kinase 9) et le transactivateur Tat vont se fixer au niveau de la séquence TAR. Suite à la transcription, ce recrutement va permettre une phosphorylation de l'ARN polymérase de type II et du facteur N-TEF (Negative transcription elongation factor) afin de convertir le complexe de transcription en complexe d'élongation de la transcription. En absence de Tat, NF-κB est capable de recruter P-TEFb afin d'assurer l'élongation de la transcription.

quantité d'ARN messagers viraux appelés ARNs précoces de régulation qui vont être à l'origine de trois protéines de régulation : Tat, Nef et Rev. La fixation de Tat et de P-TEFb au niveau de la séquence TAR (transactivating-responsive region) situé en 5' de l'ARN viral transcrit va ensuite permettre le recrutement de facteurs afin d'augmenter l'activité de l'ARN polymérase de type II dite « Tat dépendante ». Les ARN multiépissés sont exportés vers le cytoplasme de manière indépendante de Rev via le système d'export des ARN messagers cellulaires, ce qui permet la production des protéines Rev, Nef et Tat. Le taux de Rev va inhiber la synthèse des ARNm multiépissés et activer la synthèse d'ARNm monoépissés et non épissés. L'assemblage de nouvelles particules virales débute avec l'encapsidation de l'ARN génomique et des autres facteurs viraux tel que les protéines accessoires Vif, Vpr et Nef. La libération des nouvelles particules virales se fait par bourgeonnement au niveau de la membrane cytoplasmique de la cellule infectée. Il y a finalement maturation par la protéase virale des polyprotéines gag et gag-pol résultant en des changements de morphologies importantes du virion incluant la génération de la capside conique et permettant la formation de virions matures infectieux (Peterlin & Trono, 2003).

Les différentes cibles actuelles des thérapies antirétrovirales découlent directement du cycle de réplication du VIH-1.

1.3. Moyens de lutttes antirétrovirales

a) Prévention de l'infection à HIV-1

Depuis 2000, les derniers résultats de l'OMS montrent que la propagation du VIH-1 s'est enrayée montrant sur une même période une baisse de 35 % de nouvelles infections. Cependant, la prévention de l'infection à VIH-1 joue un rôle prépondérant dans un objectif de limitation de la pandémie d'ici 2030. La prévention à l'infection de VIH-1 consiste au port du préservatif, à la circoncision des hommes et à l'initiation rapide d'une thérapie antirétrovirale chez l'individu néo-infecté, mais aussi au développement de programmes visant à réduire les comportements à risques dans les pays en voie de développement.

Très récemment, un nouvel outil de prévention de l'infection à VIH-1 est la prophylaxie pré-exposition (PrEP) reposant sur une prévention médicamenteuse. La stratégie majeure est de permettre à une personne séronégative qui court le risque de contracter le VIH de réduire son risque d'infection (Molina *et al*, 2015; Guideline on When to Start Antiretroviral Therapy and on Pre-Exposure Prophylaxis for HIV, 2015). La PrEP est officiellement autorisée en France depuis fin novembre 2015 par la prise de Truvada®. Un deuxième outil de prévention est la prophylaxie post-exposition (PPE) permettant de réduire les risques d'infections suite à une potentielle exposition au virus (Whitlock *et al*, 2015).

b) Stratégie thérapeutique

Il existe aujourd'hui de nombreux traitements contre le SIDA, cependant aucune thérapie ne permet de guérir de cette maladie. La stratégie actuelle consiste à l'utilisation d'inhibiteurs ciblant l'entrée du virus en empêchant l'interaction avec la cellule cible ou la fusion des deux bicouches lipidiques, mais aussi des inhibiteurs des enzymes virales comme la RT, la PR et l'IN. La thérapie anti-VIH actuelle tourne autour d'une combinaison de molécules HAART (Highly Active Antiretroviral Therapy) ciblant différentes étapes du cycle viral. L'objectif majeur étant de réduire le risque de développement de virus mutant résistant dû au taux d'erreur important généré par la RT du virus.

La première classe regroupe les inhibiteurs nucléosidiques ou NRTIs (nucleoside/nucleotide reverse transcriptase inhibitor). C'est historiquement la première classe de drogues antivirales développée contre le VIH-1. Ces inhibiteurs sont des analogues de nucléosides entrant en compétition avec les nucléosides naturels permettant le blocage de la RT (Tressler & Godfrey, 2012) (Zidovudine AZT, Lamivudine 3TC).

La deuxième classe regroupe les inhibiteurs non nucléosidiques ou NNRTIs (non nucleoside reverse transcriptase inhibitor). Ces inhibiteurs se fixent à la RT du virus empêchant un changement conformationnel de la structure tridimensionnelle affectant l'activité de cette dernière (de Béthune, 2010) (nevirapine NVP, efavirenz EFV).

Une troisième classe d'inhibiteur regroupe les inhibiteurs de la protéase ou PIs. Ce sont des inhibiteurs peptidomimétiques mimant le substrat précurseur polypeptidique entrant en compétition avec le substrat viral. Ces peptides non clivables par la PR affectent son activité et sont alors utilisés comme inhibiteurs (Lv *et al*, 2015) (ataznavir ATV, darunavir DRV).

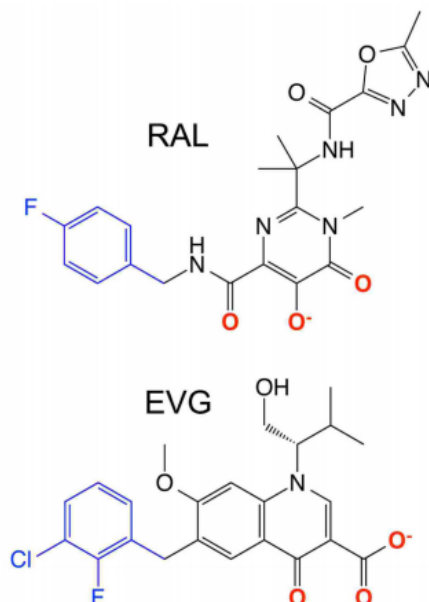


Figure 15 : Structure chimique de deux molécules anti-intégrase. RAL : Raltegravir, EVG : Elvitegravir.

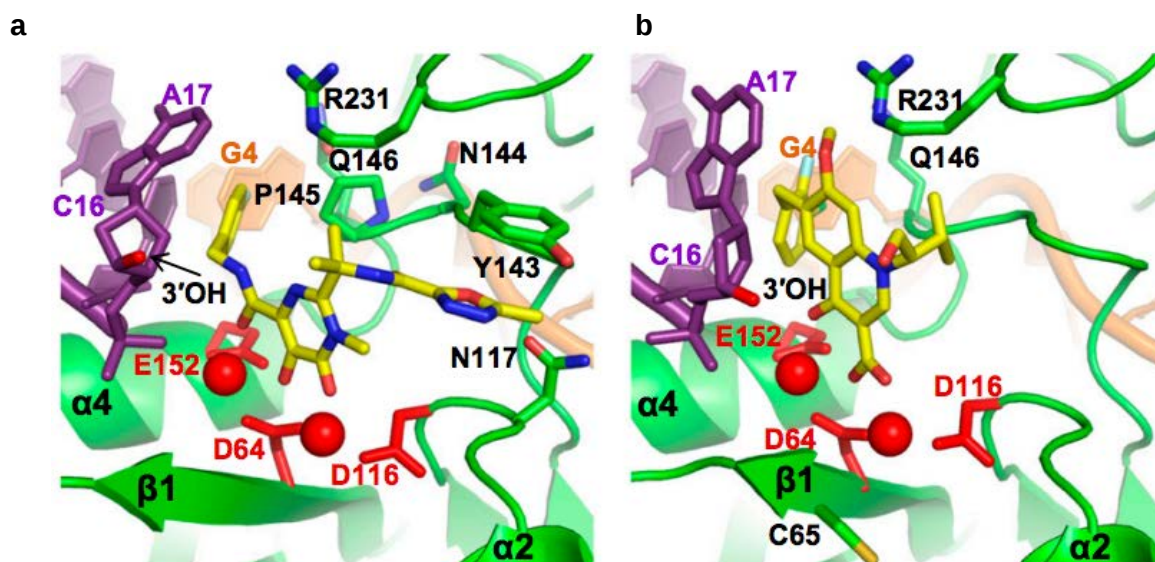


Figure 16 : Modèle de fixation de deux molécules anti-intégrase : Raltégravir (a) et Elvitegravir (b) dans le site actif de l'intégrase (Krishnan *et al*, 2010). Les différents atomes sont représentés en jaune (carbone), bleu (azote), rouge (oxygène), gris (Fluor) et vert (Chlore) ainsi que les sphères rouges (Mg^{2+}).

Une quatrième classe d'inhibiteurs comprend les inhibiteurs de fusion (FIs). Ce sont des inhibiteurs peptidomimétiques du domaine HR2 de la glycoprotéine gp41 et agissent par fixation sur le domaine HR1 afin d'inhiber la fusion des bicouches lipidiques (Zhang *et al*, 2015) (enfuvirtide T-20).

Une cinquième classe d'inhibiteurs regroupe les inhibiteurs d'entrées. Ces inhibiteurs bloquent l'interaction de la glycoprotéine virale gp120 à son corécepteur et sont des molécules antagonistes du CCR5 ou du CXCR4 (Maraviroc MVC).

Une dernière classe d'inhibiteurs regroupe les inhibiteurs de l'activité de transfert de brin de l'intégrase (IN) (INSTIs). Ces inhibiteurs viennent se fixer au niveau du complexe ADNviral/IN empêchant l'intégration dans l'ADN cellulaire (Raltegravir RAL, Dolutégravir DTG, Elvitégravir EVG) (**Figure 15 et 16**).

Plus récemment, de nouveaux inhibiteurs ont été décrits. Il s'agit des LEDGIN (Christ *et al*, 2012). Ces inhibiteurs allostériques vont se fixer à l'IN permettant d'inhiber l'interaction entre l'intégrase (IN) et son cofacteur cellulaire LEDGF/p75. Le mécanisme d'action de ces molécules est décrit comme affectant l'oligomérisation de l'IN et l'assemblage du virion.

c) Résistance aux antirétroviraux

La résistance aux antirétroviraux va permettre aux virus de se répliquer en présence d'une molécule antivirale à des concentrations qui inhibent la réplication d'un virus sensible. Dans le cas du VIH-1, la sélection des mutations concerne les gènes codant les protéines cibles des antirétroviraux tels que la RT, IN, PR, Gp41. Ainsi, les protéines portant les mutations deviennent insensibles aux antirétroviraux concernés. La génération des différentes mutations implique principalement la RT du virus et la recombinaison génomique (Hu & Temin, 1990). En effet, la fidélité de l'activité polymérase de cette protéine va permettre l'introduction d'erreurs au hasard dans le génome viral (Drosopoulos *et al*, 1998). L'émergence de variants résistants occupe aujourd'hui une place centrale dans les problématiques de lutte antirétrovirale. Ainsi, le développement de nouvelles cibles thérapeutiques comme le ciblage des interactions entre facteurs cellulaires et viraux permet une diminution du processus de résistance virale.

d) Les essais vaccinaux

Bien que la recherche d'un vaccin anti-VIH ait commencé dès 1980, et après plus de 250 essais cliniques, aucun vaccin n'est disponible. L'ensemble des échecs s'explique par la grande variabilité virale ainsi que l'apparition de souches mutantes échappant au système immunitaire. Un essai clinique STEP développé par Merck aussi connu sous le nom MerckV520 Protocol 023 a eu pour objectif la vaccination s'appuyant sur l'utilisation de trois vecteurs viraux basés sur l'Adénovirus de type 5 (Ad5) un pour gag, pol, et nef. Les premiers résultats ont montré un taux plus élevé d'infections chez les individus vaccinés versus le groupe contrôle poussant alors à l'abandon de l'étude STEP pour cause d'inefficacité du concept (Sekaly, 2008).

Depuis ce jour, seul un essai clinique (RV144) a pu montrer un effet modeste d'efficacité contre l'infection à VIH-1 (31,2 % après 42 mois de suivi). En effet les vaccinés ont pu recevoir quatre injections d'ALVAC-HIV (vecteur viral permettant l'expression de six gènes viraux modifiés de sous types B : *env*, *gag*, *pol* développé par Sanofi Pasteur. Ces injections sont ensuite suivies de deux injections d'AIDSVAX (mélange de glycoprotéine gp120 de sous-type B et E) (Rerks-Ngarm *et al*, 2009). Basé sur ces résultats encourageants, un nouvel essai clinique a été développé en 2015 sous le nom de HVTN 100 en Afrique du Sud.

f) Un cas clinique de guérison : « Le patient de Berlin »

Un seul cas de guérison a pour le moment été décrit, il s'agit de l'américain Timothy Brown qui ne montre aucun signe d'infection depuis 2007. Chez ce patient infecté par le VIH-1 depuis 1995 dont une leucémie myéloïde aiguë a été diagnostiquée en 2006. Il reçoit alors afin de soigner sa leucémie, une radiothérapie puis une allogreffe de moelle osseuse issue d'un donneur homozygote pour l'allèle *ccr5* portant une rare mutation au niveau du récepteur CCR5 (CCR5 Δ 32) (Hütter *et al*, 2009). Le patient infecté par un virus à tropisme CCR5 a alors contrôlé spontanément la réplication virale. Fin 2010, le VIH restait indétectable dans le sang circulant et dans les différents tissus (moelle osseuse, cerveau, colon) sans traitement antirétroviral (Yukl *et al*, 2013). Cette première guérison est d'autant plus questionnable que l'homozygotie CCR5 Δ 32 est associé à une forte, mais pas complète

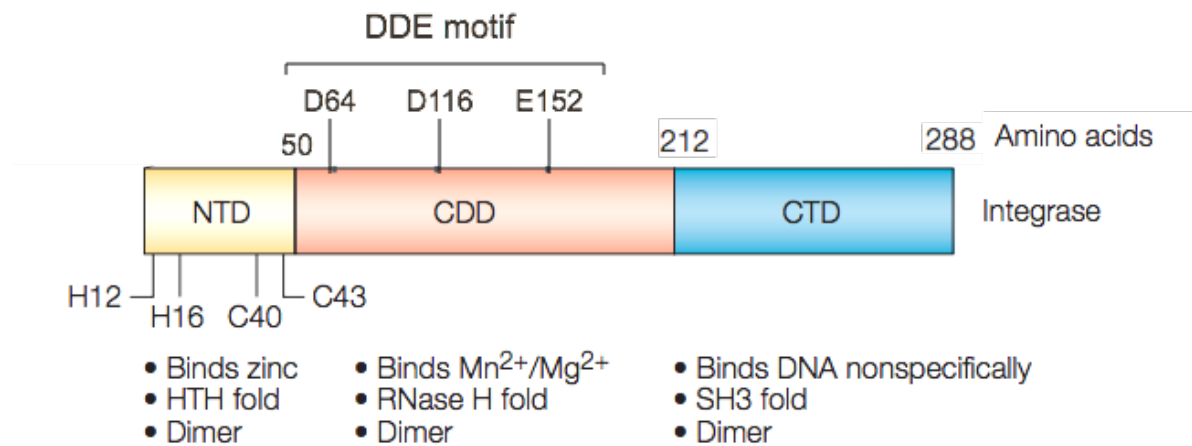


Figure 17 : Organisation schématique de l'intégrase du VIH-1. Adapté de (Pommier *et al*, 2005)

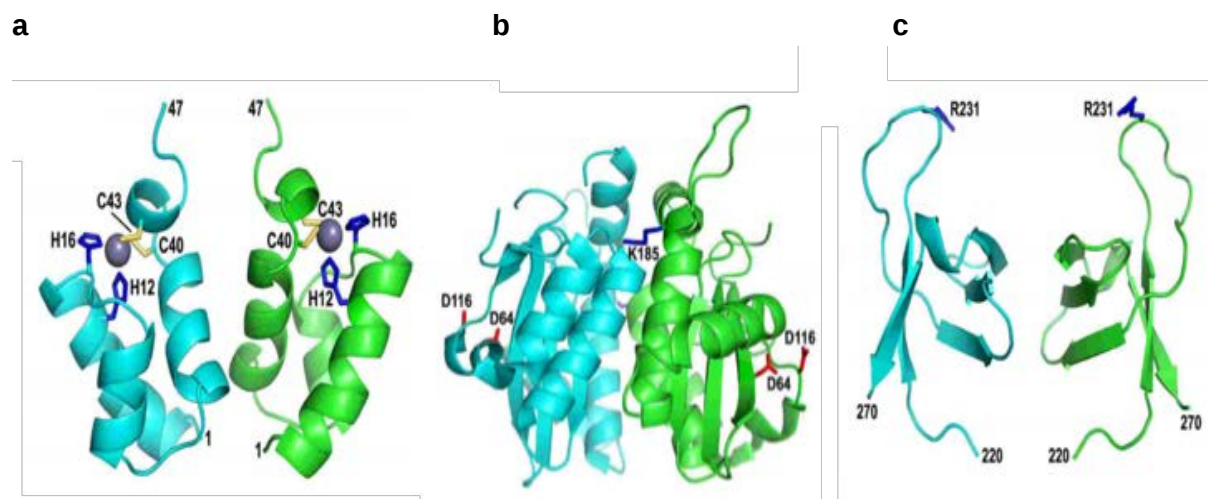


Figure 18 : Structure des domaines isolés de l'intégrase du VIH-1. (a) Structure d'un dimère du domaine N-terminale (NTD) (résidus 1 à 47), la sphère représente un atome de zinc (Cai *et al*, 1997) (b) Structure d'un dimère du CCD, les résidus de la triade catalytique sont représentés en rouge (Dyda *et al*, 1994). (c) Structure d'un dimère du domaine C-terminale (CTD) de l'intégrase du VIH-1 (résidus 220 à 270) (Lodi *et al*, 1995; Eijkelenboom *et al*, 1995).

résistance au VIH-1. De plus, l'analyse par séquençage a révélé la présence de variants à tropisme CXCR4 et qu'il existe des macrophages exprimant le CCR5 qui constituent potentiellement des réservoirs dans la muqueuse intestinale.

2. L'intégration et intégrase du VIH-1

2.1. Intégrase virale

L'intégrase du VIH-1 est une protéine de 32kDa composé de 288 acides aminés, cette protéine est codée par le gène *pol* et est issue de la maturation du précurseur Gag-Pol. Elle se compose de trois domaines qui ont pu être identifiés par protéolyse partielle et par étude fonctionnelle (Engelman *et al*, 1993; van Gent *et al*, 1993). Ces trois domaines qui sont requis pour réaliser l'intégration de l'ADN viral (**Figure 17**) consistent en un domaine N-terminal (1-50), un domaine central ou core catalytique (50-212), et un domaine C-terminal (212-288). À ce jour, aucune structure complète de l'IN du VIH-1 n'a pu être résolue. Cependant des structures de domaines isolés ainsi qu'une structure complète de l'IN associée à ses substrats ADN complexé avec son cofacteur LEDGF/p75 a pu être déterminée par cryo-microscopie (Dyda *et al*, 1994; Eijkelenboom *et al*, 1995; Lodi *et al*, 1995; Cai *et al*, 1997; Chen *et al*, 2000; Wang *et al*, 2001; Michel *et al*, 2009).

a) Le domaine N-terminal (1-50)

Le domaine N-terminal contient un motif conservé HHCC (His-12, His-16, Cys-40, Cys43). Ce motif permet la liaison du zinc impliqué dans l'oligomérisation de la protéine et dans la fixation à l'ADN viral (van den Ent *et al*, 1999; Zhao *et al*, 2008). Les données de RMN ont pu identifier la structure isolée de ce domaine composé de trois hélices alpha formant un motif hélice-tour-hélice (HTH) (**Figure 18a**) (Cai *et al*, 1997; Eijkelenboom *et al*, 1997).

b) Le core catalytique (51-212)

Le core catalytique (CCD) comporte le site actif de l'enzyme composé par la triade catalytique D₆₄D₁₁₆D₁₅₂ conservée dans la famille des poly-nucléotidyl-transférases à laquelle appartient l'IN, transposases et recombinaisons (Engelman & Craigie, 1992; Kulkosky *et al*, 1992).

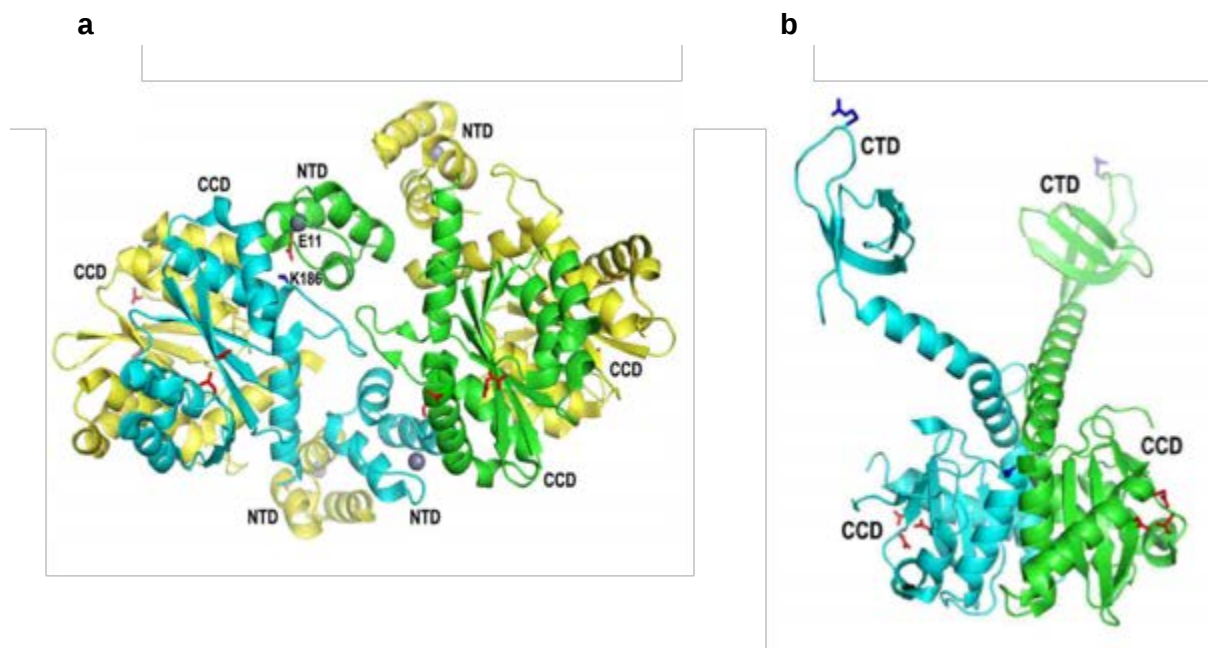


Figure 19 : Structure du double domaine NTD-CCD et CCD-CTD de l'intégrase du VIH-1. (a) Structure d'un dimère du double domaine NTD-CCD (PDB : 1K6Y) (Wang *et al*, 2001). (b) Structure d'un dimère du double domaine CCD-CTD (PDB : 1EX4) (Chen *et al*, 2000).

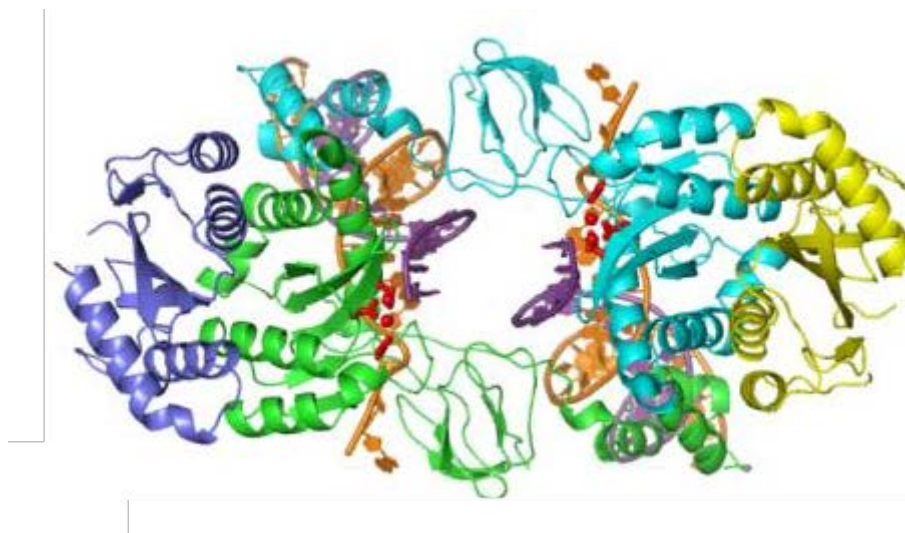


Figure 20 : Modèle de l'intasome du VIH-1 (Krishnan *et al*, 2010). Les sous-unités internes engagées avec l'ADN viral (résidus 1 à 270) sont représentées en vert et en cyan. Les CCD externes (résidus 56 à 202) sont représentés en bleu et jaune. L'ADN est en orange et magenta.

La triade catalytique est capable de coordonner deux ions divalents Mg^{2+} qui sont requis pour l'activité catalytique de l'IN. La structure cristallographique de ce domaine a pu être déterminée par l'introduction d'une mutation F185K permettant d'augmenter sa solubilité tout en conservant partiellement l'activité de l'IN (**Figure 18 b**) (Dyda *et al*, 1994). La structure du monomère montre qu'il contient 6 hélices α et 5 feuillets β . Ce domaine est aussi impliqué dans l'interaction avec LEDGF/p75, cofacteur cellulaire de la protéine.

c) Le domaine C-terminal (213-288)

Le domaine C-terminal (CTD) est le domaine qui est le moins conservé des trois domaines de l'intégrase. La structure de ce domaine a été caractérisée par RMN et se compose de 5 feuillets β (résidus 222-229, 232-245, 248-253, 256-262, 266-270) ayant un repliement analogue aux domaines de types SH3 (sarc homology 3 domain) impliqués dans les interactions protéines-protéines (**Figure 18c**) (Lodi *et al*, 1995; Eijkelenboom *et al*, 1995). De plus, ce domaine possède des régions chargées positivement permettant une fixation à l'ADN de façon non spécifique. Ce domaine est également impliqué dans l'oligomérisation de l'IN (Jenkins *et al*, 1996; Lutzke & Plasterk, 1998).

d) L'intasome du VIH-1

Jusqu'à ce jour, aucune structure complète du monomère d'intégrase ou de l'oligomère n'est disponible dû en majeure partie à la faible stabilité et solubilité de la protéine. Seul le double domaine NTD/CCD portant les mutations F185K, W131D et F139D ainsi que le double domaine CCD/CTD portant les mutations C56S, W131D, F139D, F185K et C286S sont disponibles (**Figure 19**) (Chen *et al*, 2000; Wang *et al*, 2001).

Même si l'état oligomérique fonctionnel de l'IN *in vivo* n'est encore pas encore connu, de multiples études ont pu montrer que l'activité d'intégration était réalisée par un oligomère d'intégrase. En effet, un dimère d'intégrase est suffisant à la réaction de 3' processing et qu'un tétramère est requis pour une activité d'intégration efficace *in vitro*. De plus, un tétramère apparaît comme nécessaire pour l'activité de transfert de brin. Ces résultats sont confortés par des images de microscopie à force atomique montrant un tétramère d'IN en complexe avec l'ADN substrat (Kotova *et al*, 2010).

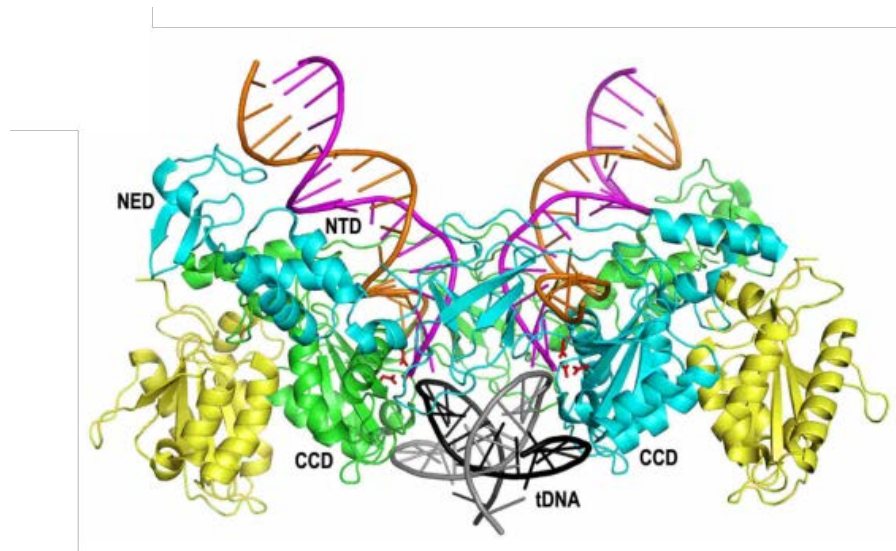


Figure 21 : Structure du complexe de transfert de brin de PFV (Hare *et al*, 2010). Les monomères internes du tétramère en contact avec l'ADN viral sont représentés en vert et en cyan. Les monomères externes sont représentés en jaune. Le brin d'ADN transféré contenant l'extrémité CA 3'OH est en magenta. Le brin complémentaire non transféré est en orange. (PDB : 3OS1) NED : NTD extension domain.

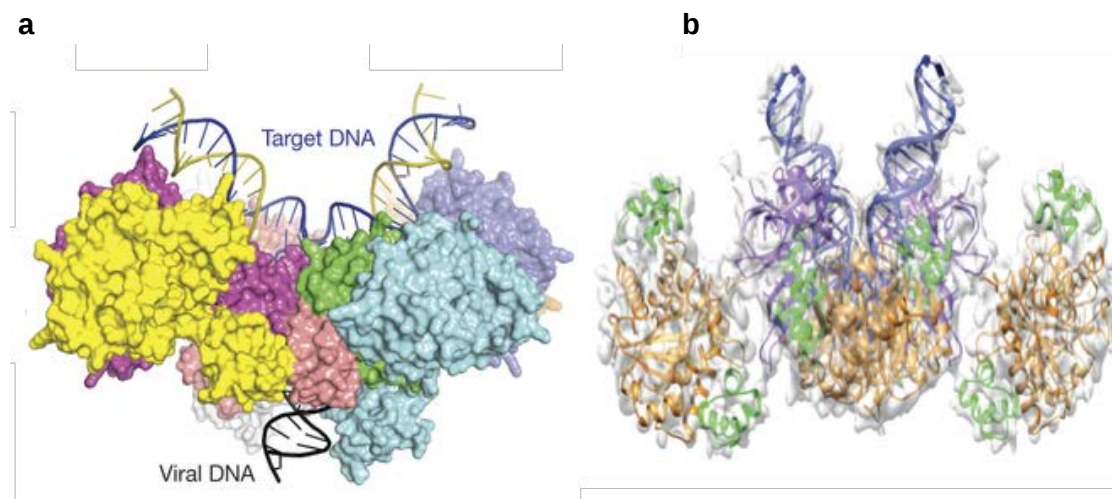


Figure 22 : Structure de l'intasome de RSV et de MMTV (Ballandras-Colas *et al*, 2016; Yin *et al*, 2016). (a) Structure de l'intasome de RSV. Les monomères des dimères externes sont représentés en jaune, magenta, orange et violet. Les monomères des dimères internes sont représentés en blanc, rose, vert et azur. Les monomères nécessaires à l'activité catalytique sont en vert et rose. (b) Structure de l'intasome de MMTV. Dans chaque monomère, le domaine NTD est représenté en vert, le domaine CCD en orange et le CTD en violet. L'ADN est représenté en bleu.

Basés sur ces résultats, plusieurs modèles ont été proposés pour représenter ce nucléoprotéo-complexe qu'est l'intasome du VIH-1 (**Figure 20**) (Podtelezhnikov *et al*, 2003;

Wielens *et al*, 2005; Krishnan *et al*, 2010). Une avancée majeure dans la compréhension de l'état oligomérique de l'IN et de ses interactions avec l'ADN a été permise grâce à la structure entière de l'intégrase du Prototype Foamy Virus (PFV) complexé avec son ADN substrat (**Figure 21**). L'intasome de PFV est composé de deux dimères d'IN formant un tétramère où seul un site actif de chaque dimère du CCD participe à l'activité enzymatique (Maertens *et al*, 2010).

Plus récemment, les structures des intasomes d'autres rétrovirus, le virus du sarcome de Rous (RSV) et le virus de la tumeur mammaire de la souris (MMTV) ont pu être obtenus par cryo-électromicroscopie et par diffraction aux rayons X (**Figure 22**) (Ballandras-Colas *et al*, 2016; Yin *et al*, 2016). L'intasome de RSV et de MMTV possède une architecture différente de PFV, elle se compose d'un octamère d'intégrase. Ces différentes structures d'intasome mettent en relief la diversité d'architecture intasomale des rétrovirus et suggèrent ainsi des différences fonctionnelles.

2.2. Intégration virale : mécanismes et fonctions

a) Introduction

Après l'entrée du virus dans la cellule, il y a désassemblage du virion qui va permettre la formation du complexe de rétro-transcription (RTC) qui inclut les deux brins d'ARN viraux, la RT, l'IN et d'autres facteurs viraux et cellulaires. L'étape de rétro-transcription se déroule au sein du RTC qui va permettre la formation d'un nouveau nucléocomplexe comprenant l'ADN néo-synthétisé appelé le complexe de pré-intégration (CPI). Le CPI est formé dans le cytoplasme de la cellule et va ensuite être importé dans le noyau de la cellule où l'ADN viral est intégré dans l'ADN chromosomique. L'organisation du CPI est encore mal connue dû à la faible quantité de ces complexes dans les extraits protéiques de cellules infectées. L'intégration de l'ADN viral est une étape clé du cycle viral et nécessite trois réactions consécutives : le 3'-end processing, le transfert de brin et la réparation du site d'intégration (**Figure 23**).

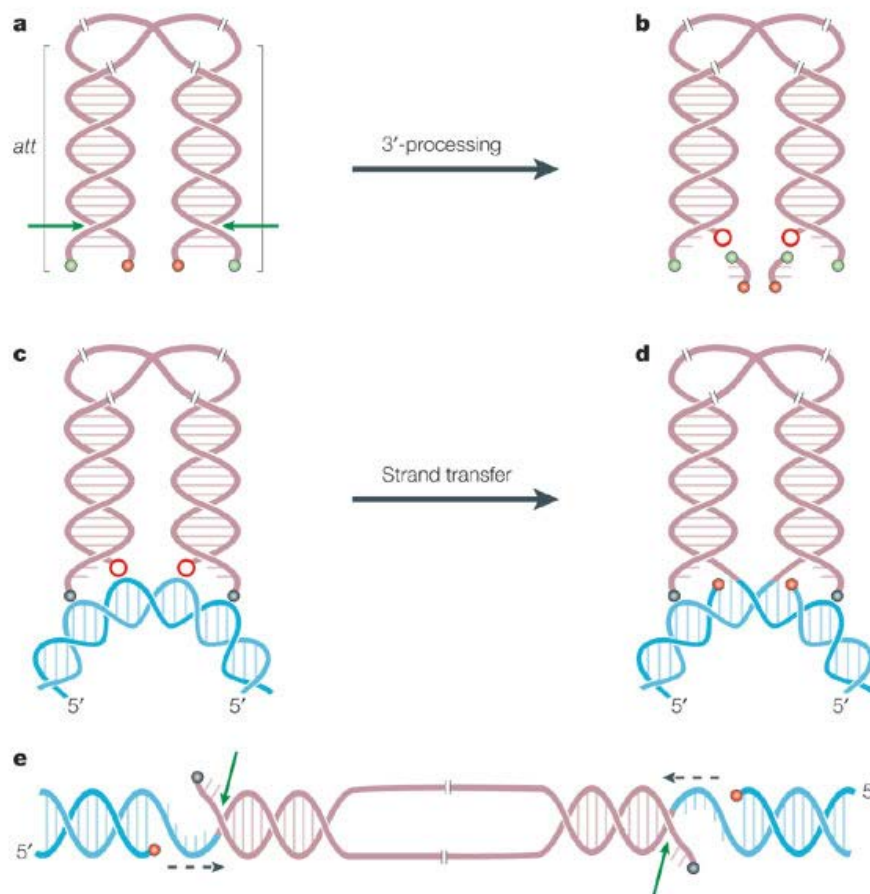


Figure 23 : Schéma des différentes étapes de l'intégration (Pommier *et al*, 2005). (a) Suite à la transcription inverse, l'activité de 3'-processing a lieu dans le cytoplasme de la cellule. Cette activité est un clivage endonucléotidique de l'ADN viral au niveau des extrémités 3' (b) Ce clivage d'un dinucléotide GT à chaque extrémité va générer deux extrémités CA 3' OH. L'extrémité 3' OH est représentée par un point rouge et le groupement 5' P par un point vert. (c) L'oligomère d'IN complexé avec l'ADN viral est transloqué dans le noyau afin de réaliser la réaction de transfert de brin. (d) Le groupement 3'OH attaque la liaison phosphodiester de l'ADN cible avec un décalage de 5bp permettant la jonction de l'ADN viral avec l'ADN cellulaire (e) Puis, il y a jonction de l'extrémité 5' de l'ADN viral avec l'extrémité 3' de l'ADN cible par des facteurs de réparation cellulaire.

b) 3'-end processing

En présence de Mg^{2+} ou de Mn^{2+} et d'un ADN substrat, l'intégrase virale est capable de catalyser le clivage d'un dinucléotide GT aux extrémités 3' d'U3 et d'U5 des Long Terminal Repeat (LTR) de l'ADN. Cette activité va permettre de libérer des extrémités CA 3'OH rentrantes marquant alors la jonction entre l'ADN viral et cellulaire du produit d'intégration (Sherman & Fyfe, 1990; Bushman & Craigie, 1991).

c) Transfert de brins

L'intégrase du VIH-1 est capable de catalyser la seconde étape du processus d'intégration qui est l'insertion covalente des extrémités virales dans l'ADN cellulaire. Cette étape est dépendante de la présence de Mg^{2+} ou de Mn^{2+} *in vitro*. Dans cette réaction, le groupement hydroxyle OH en 3' de l'adénosine attaque la liaison phosphodiester de l'ADN cible. Cette réaction résulte en un clivage simultané des deux brins de l'ADN cible permettant la liaison de l'extrémité 3' de l'ADN viral à l'ADN cible au niveau du site de clivage par une réaction de substitution nucléophile. Cette attaque nucléophile s'effectue en deux sites localisés à chaque extrémité de l'ADN viral avec un décalage de 5 paires de bases chez les lentivirus (Bushman & Craigie, 1991).

d) La réparation post-intégration

L'étape de réparation du site d'intégration consiste en l'élimination des deux dinucléotides des extrémités 5' de l'ADN viral, la ligation de ces extrémités avec les extrémités 3' de l'ADN chromosomique et le remplissage des gaps. L'ensemble de ces étapes est réalisé par des facteurs cellulaires tels que l'endonucléase FEN-1 ainsi que des ligases qui vont permettre de terminer le processus d'intégration (Brin *et al*, 2000). Cette étape reste néanmoins encore peu étudiée, mais plusieurs protéines cellulaires sont susceptibles de jouer un rôle dans la réparation du site d'intégration telles que :

- La voie NHEJ (Non Homologous End-Joining) composée par la DNA-PK et l'hétérodimère Ku70/80 représentant le mécanisme majeur de réparation des cassures doubles brins de l'ADN.

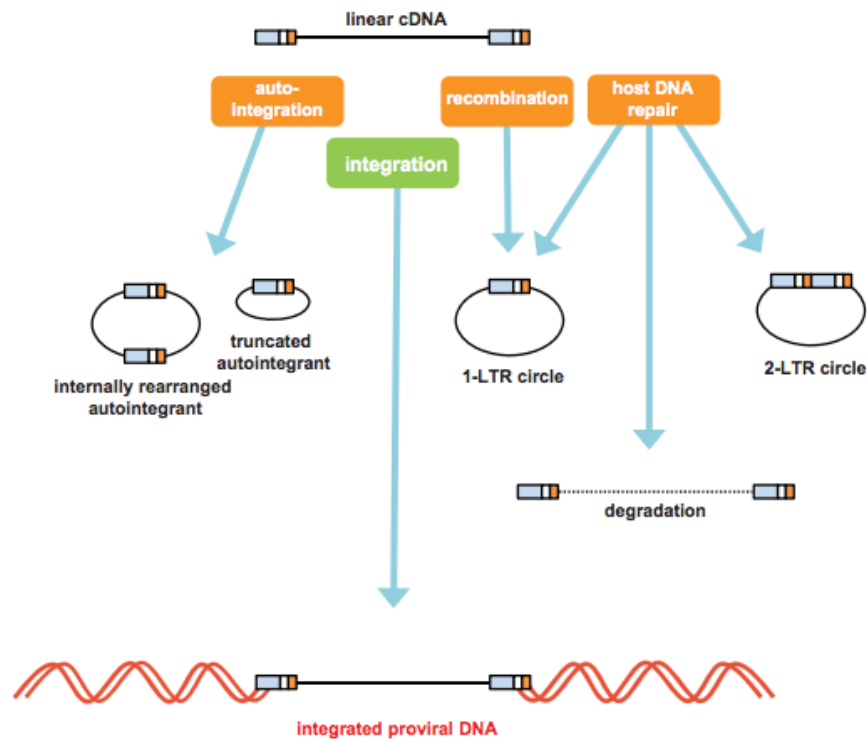


Figure 24 : Représentation des différentes formes d'ADN non intégrées (Sloan & Wainberg, 2011).

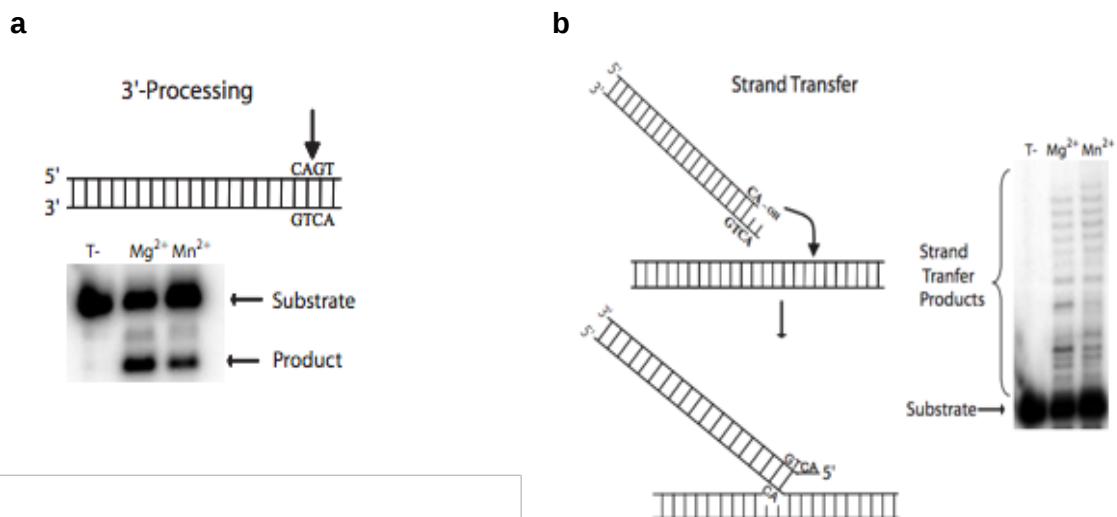


Figure 25 : Représentation des activités de 3'end processing (a) et de transfert de brin (b) réalisé par l'IN *in vitro* (Delelis et al, 2008).

- La Poly (ADP-ribose) polymérase 1 (PARP1) qui est un facteur cellulaire potentiellement impliqué lors de son activation par des cassures simple et double brins de l'ADN (Dantzer *et al*, 2006).
- La kinase ATR (ATM and Rad3-related) senseur des dommages à l'ADN semble être impliquée (Daniel *et al*, 2003).
- RAD51, facteur cellulaire jouant un rôle crucial dans la recombinaison homologue (HR) pourrait également jouer un rôle dans la régulation des étapes de la réparation de l'ADN (Cosnefroy *et al*, 2012).

e) Les produits non intégrés

Plusieurs formes d'ADN viraux sont retrouvées dans le noyau de la cellule infectée sous formes non intégratives. Ces différentes formes d'ADN non-intégrés sont issues de la rétro-transcription de l'ARN viral sous forme d'ADN viral complémentaire linéaire, les autres formes retrouvées sont des formes d'ADN circulaire (**Figure 24**). Certaines formes d'ADN circulaires sont issues de l'auto-intégration, processus dans lequel l'extrémité 3' de l'ADN viral transcrit s'insère dans l'ADN viral aboutissant à la formation de cercles tronqués ou réarrangés. Les autres formes circulaires sont les cercles à 1-LTR formés par recombinaison homologue au niveau des LTR de l'ADN linéaire. D'autres études ont montrées que ces cercles à 1-LTR peuvent aussi être formés par ligation d'intermédiaires de la rétro-transcription. Quant aux cercles à 2-LTR, leur présence est liée aux événements de réparation non homologue de la voie NHEJ (non homologue end joining). Étant donné que les cercles à 2-LTR sont exclusivement retrouvés dans le noyau, ils peuvent être utilisés comme marqueur de l'import nucléaire viral (Sloan & Wainberg, 2011).

2.3. Activités de l'intégrase *in vitro*

Plusieurs tests *in vitro* ont été développés afin de mieux caractériser les activités de l'intégrase. Les tests de 3'end processing et de transfert de brin permettent de reproduire ces deux activités de façon isolée alors que le test d'intégration concertée reproduit l'intégration totale telle qu'elle peut se faire dans la cellule afin de se rapprocher d'une situation physiologique. De plus, l'IN est capable de catalyser la réaction inverse du transfert de brin *in vitro* : la désintégration.

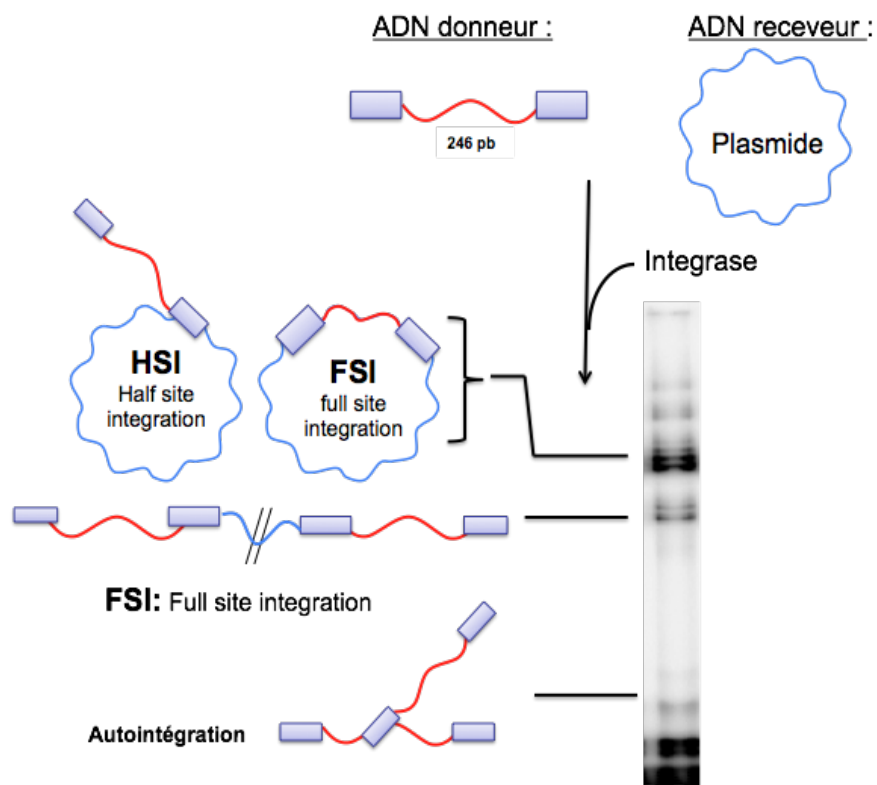


Figure 26 : Représentation de l'activité d'intégration concertée.

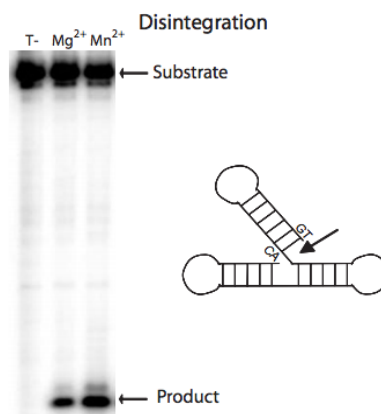


Figure 27 : Représentation de l'activité catalytique de désintégration de l'IN (Delelis *et al*, 2008).

a) Test de 3'end processing et transfert de brin

Le test de 3'end processing reproduit la première étape de l'intégration. Ainsi, un substrat oligonucléotidique (ODN) de 21 nucléotides correspondant à l'extrémité U5 du LTR viral préalablement radiomarké en 5' va être mis en présence de l'intégrase virale recombinante (**Figure 25a**) (Leh *et al*, 2000). De cette réaction aboutit un produit de 19 nucléotides visualisable après électrophorèse sur un gel de polyacrylamide. De plus, l'IN catalyse ensuite le transfert de brin de cet ODN de 19 nucléotides dans un autre ODN (**Figure 25b**). L'insertion se faisant de manière aléatoire, les produits obtenus sont de différentes tailles. Les différents produits de la réaction sont visualisés après électrophorèse sur gel de polyacrylamide. Cette activité de transfert de brin peut également être spécifiquement testée en utilisant un ODN pré-processé mimant ainsi une extrémité virale ayant déjà subi le 3'end processing.

b) L'intégration concertée

Ce test fait intervenir les partenaires minimaux permettant de mimer la réaction d'intégration de deux extrémités de l'ADN viral dans l'ADN cellulaire. Pour effectuer cette réaction, nous utilisons un ADN donneur radiomarké de 246 pb comportant les 21 derniers nucléotides des extrémités U3 et U5 mimant l'ADN viral et un ADN receveur mimant l'ADN cellulaire. Ainsi en présence d'intégrase, plusieurs produits d'intégration sont visualisables après électrophorèse sur gel d'agarose (**Figure 26**). On distinguera premièrement des produits d'autointégration d'un donneur dans un autre donneur, et une famille de produit d'intégration d'un donneur dans un ADN receveur (hétérointégration). Parmi ces produits, on distinguera des produits d'intégration d'une seule extrémité d'un donneur dans l'ADN receveur aboutissant à un produit d'intégration partielle ou « half site » (HSI). Les produits d'intégration concertée ou « full site » (FSI) sont la résultante de l'intégration de deux extrémités virales, d'une ou de deux molécules différentes d'ADN donneur, dans l'ADN receveur de façon concertée au même site d'intégration. Ainsi, on distinguera les produits d'intégration concertée de deux extrémités d'un même donneur, et d'une intégration concertée d'une extrémité de deux donneurs différents permettant la linéarisation de l'ADN receveur.

c) La désintégration

Une troisième activité peut être caractérisée seulement *in vitro* : la désintégration (**Figure 27**). Cette réaction correspond à la réaction inverse de la réaction de transfert de brin (Chow *et al*, 1992). Dans cette réaction, un substrat ADN correspondant à un ADN replié en Y mime le produit d'intégration d'une extrémité virale dans un brin d'ADN cible. Ainsi, en présence d'IN recombinante, l'enzyme purifiée est capable de cliver un oligonucléotide. Cette réaction peut être catalysée par des IN tronquées NTD-CCD (IN¹⁻²¹⁰) et CCD-CTD (IN⁵⁵⁻²⁸⁸) ainsi que par le domaine catalytique seul (IN⁵⁵⁻²⁰⁹) (Leh *et al*, 2000). Cette activité comme les autres activités de l'IN *in vitro*, nécessite la présence d'un cation divalent Mg²⁺ ou Mn²⁺.

2. 4. Facteurs cellulaires modulateurs de l'intégration

Bien que l'IN soit suffisante à catalyser la réaction de 3'end processing et de transfert de brin *in vitro* en l'absence de cofacteurs. De nombreuses études montrent la participation de plusieurs cofacteurs cellulaires lors de l'étape de rétro-transcription de l'ARN viral et de l'intégration de l'ADN viral dans la cellule. Les protéines cellulaires qui facilitent l'intégration du provirus sont appelées cofacteurs de l'intégration. De nombreux cofacteurs de l'intégration ont été décrits comme jouant un rôle sur la catalyse, l'import nucléaire, le choix du site d'intégration et la réparation du site d'intégration. De façon intéressante, l'étude de ces cofacteurs de l'intégration constitue autant de choix possibles dans le développement d'inhibiteurs antirétroviraux.

Les cofacteurs de l'intégration ont été identifiés par différentes méthodes : par reconstitution d'activité *in vitro* de CPI déprotéinisé (Lee & Craigie, 1994), double-hybride (Kalpana *et al*, 1994; Rain *et al*, 2009/4) et par co-immunoprécipitation (Cherepanov *et al*, 2003). Les différents cofacteurs de l'intégration détectés par ces méthodes doivent ensuite être validés pour leur effet sur l'intégration *in vitro* en utilisant l'IN recombinant. De plus, un rôle clair sur l'intégration *in vivo* doit être démontré pour valider ces cofacteurs putatifs de l'intégration. Seul un cofacteur de l'intégration a pu être déterminé à ce jour, ce cofacteur a été identifié par co-immunoprécipitation et par double hybride et validé en culture cellulaire : Lens epithelium-derived growth factor/p75 (LEDGF/p75). D'autres cofacteurs ont pu être identifiés, cependant un rôle direct sur l'intégration reste à être démontré : Barrier to autointegration factor (BAF), integrase interactor 1 (INI1), HMGA1, HSP60 et RAD51.

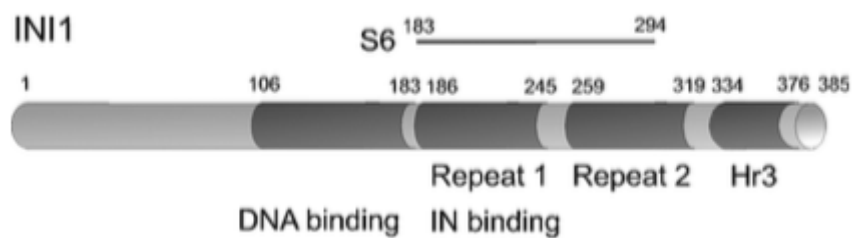


Figure 28 : Représentation des différents domaines de INI1 (Neamati & Wang, 2011).

a) Integrase Interactor 1 (INI1)

INI1 est le premier cofacteur de l'intégrase identifié par double-hybride (Kalpana *et al*, 1994). INI1 est une protéine de 385 acides aminés appartenant au complexe de remodelage de la chromatine SWI/SNF, complexe décrit comme permettant d'activer la transcription par remodelage de la chromatine ATP dépendant (**Figure 28**) (Wang *et al*, 1996). Cette protéine est constituée de trois domaines conservés Rpt1, Rpt2, et un domaine HR3 (homology region 3). La région Rpt1 est décrite comme nécessaire et suffisante à l'interaction avec l'IN du VIH-1 (Morozov *et al*, 1998). Le domaine d'interaction minimal est nommé S6 (aa183-294) et agit comme inhibiteur transdominant spécifique des phases tardives de la réplication du VIH-1 (Yung *et al*, 2001). Il a été montré que la présence du fragment S6 dans les cellules productrices réduit la libération des virions d'un facteur 10000 à 100 000 et que INI1 était retrouvé à l'intérieur du virion. Les études cellulaires ont permis de montrer que INI1 se colocalise avec le CPI (Turelli *et al*, 2001). D'autres études ont montré que la mutation K71R de l'intégrase du VIH-1 diminue l'interaction avec INI1 de 55 %. Il est observé chez ce mutant une réplication deux fois plus importante suggérant qu'INI1 pourrait agir comme facteur de restriction dans les phases précoces de la réplication (Maroun *et al*, 2006). Au contraire, un Knock down de INI1 par shRNA n'a pas montré d'influence sur la transduction de vecteur VIH pseudotypé VSV-G-GFP alors qu'une baisse de la réplication du VIH-1 de deux fois a été observée. Ces données suggèrent un rôle de INI1 dans la transactivation induite par Tat (Ariumi *et al*, 2006). L'ensemble de ces études montre qu'INI1 joue un rôle dans les phases pré et post-intégratives cependant son rôle dans l'intégration du VIH-1 reste à être clairement défini.

b) Barrier to autointegration factor (BAF)

BAF est une protéine conservée de 89 acides aminés sous forme dimérique en solution (Cai *et al*, 1998). Chaque monomère contient deux domaines HhH (Hélice-hairpin-Hélice) de fixation non spécifique à l'ADN. BAF a été mise en évidence par restauration de l'activité *in vitro* de CPI déprotéinisé (Lee & Craigie, 1994; Chen & Engelman, 1998) et par co-immunoprécipitation (Lin & Engelman, 2003). BAF est décrite comme permettant d'inhiber l'activité d'auto-intégration de l'intégrase au sein du CPI. Cette inhibition de l'activité d'auto-intégration est une caractéristique importante chez les rétrovirus, car l'auto-intégration résulte à la destruction du génome viral.

De plus, il a été décrit que BAF joue un rôle sur la structure de la chromatine grâce à ses interactions avec LAP2, emerin, MAN1, core histone H3 et l'histone H1 (Montes de Oca *et al*, 2005).

c) High mobility group chromosomal protein A1 (HMGA1)

HMGA1 est une protéine non-histone (Farnet & Bushman, 1997) appartenant à la famille HMGA. Cette protéine est capable de se lier à l'ADN et à d'autres protéines leur permettant de participer à de nombreux processus cellulaires comme la régulation de la transcription (pour une revue (Reeves, 2001)). Il a été montré que la protéine HMGA1 recombinante est capable de restaurer l'activité *in vitro* du CPI déprotéinisé par une forte concentration en sels et à stimuler l'intégration concertée dans un test d'intégration *in vitro* par stabilisation de l'IN dans une conformation active (Hindmarsh *et al*, 1999). Il a été montré que HMGA1 est capable de reconnaître et d'altérer l'ADN et des facteurs de transcription suggérant un rôle important de HMGA1 lors de phases de transcription du VIH-1 (Henderson *et al*, 2000). C'est en analysant l'interaction de HMGA1 avec les LTR viraux qu'il a été montré que la protéine était capable de se lier au niveau des sites de fixation des facteurs de transcription. Cette liaison va permettre le recrutement du complexe de remodelage de la chromatine SWI/SNF via ATF-3 (Henderson *et al*, 2004).

d) Human Heat-Shock protein 60 (HSP60)

HSP60 est une protéine chaperonne jouant un rôle essentiel dans la structuration de protéines néo-synthétisées (Fink, 1999). HSP60 recombinante a été montrée comme interagissant avec l'IN du VIH-1 *in vitro*, le CCD de l'IN semble nécessaire et suffisant à cette interaction. De plus, il a été montré que l'ajout de hHSP60 permet de stimuler l'activité de 3'end processing et de transfert de brin de l'IN et protège l'IN d'une dénaturation thermique (Parissi *et al*, 2001). Un rôle *in vivo* reste cependant à être déterminé.

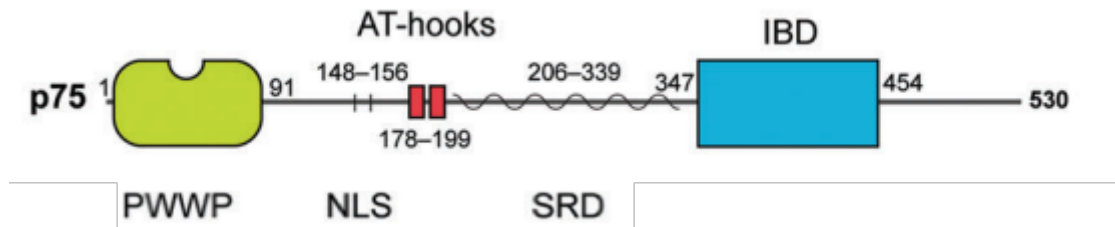


Figure 29 : Représentation des différents domaines de LEDGF/p75 (Demeulemeester *et al*, 2015). SRD : supercoiled-DNA recognition region, NLS : nuclear localization signal, AT : AT hook like domain, IBD : integrase binding domain, PWWP : Pro-Trp-Trp-Pro domain.

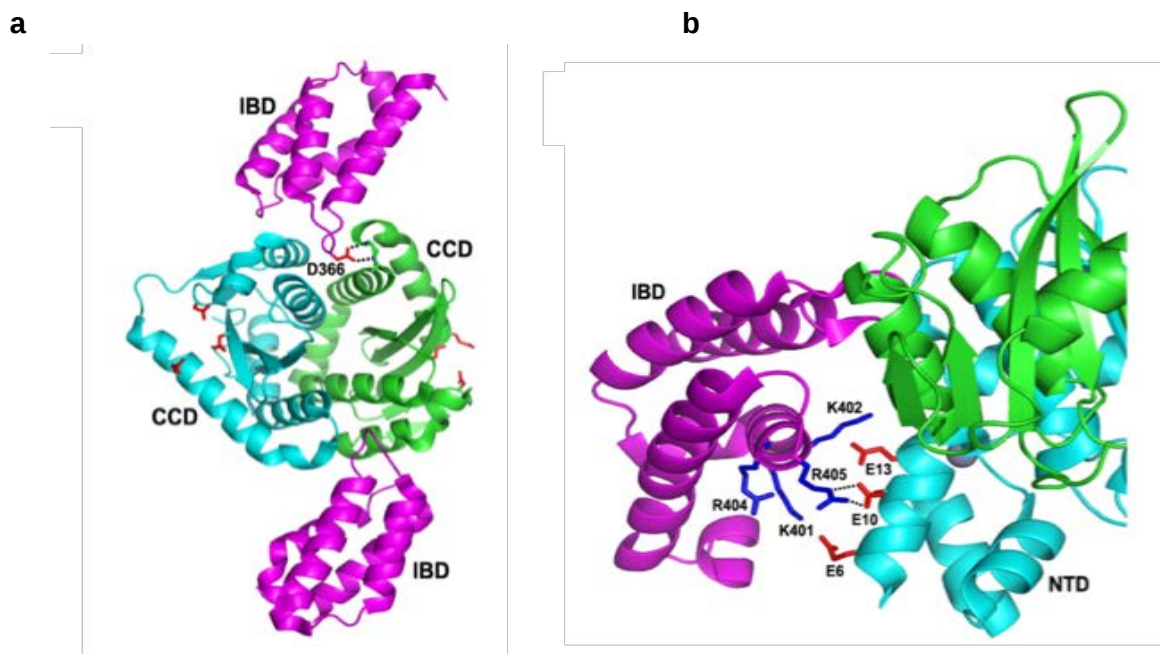


Figure 30 : Structure de complexe IN-LEDGF/p75 du VIH-1. (a) Complexe IBD-CCD/F185K du VIH-1 : les domaines IBD de LEDGF/p75 sont représentés en magenta, le dimère de CCD de l'IN du VIH-1 sont représentés en vert et en azur (PDB : 2B4J). **(b) Complexe IBD-NTD/CTD du VIH-2** : les domaines IBD de LEDGF/p75 sont représentés en magenta, le domaine NTD en azur et le CCD en vert. (PDB : 3F9K). (Cherepanov *et al*, 2005; Hare *et al*, 2009a).

e) RAD51

RAD51 est une recombinase de 339 acides aminés, jouant un rôle essentiel dans la recombinaison homologue (HR) lors de cassure double brin de l'ADN. RAD51 est capable de se polymériser afin de former un nucléofilament sur l'ADN. Cette protéine est capable d'interagir avec de multiples protéines dont BRCA1 et BRCA2 afin de réguler son activité. Plus récemment, il a été montré une interaction entre l'IN du VIH-1 et RAD51 par test ELISA et par pull-down (Desfarges *et al*, 2006). RAD51 apparaît comme jouant un rôle important sur différentes étapes de la réplication virale. En effet, il a été montré qu'une stimulation de RAD51 permet une l'inhibition de l'intégration, du VIH-1 conduisant à une baisse de la réplication virale. De plus, les données montrent qu'une inhibition précoce de RAD51 conduit à une stimulation de l'intégration et de la réplication virale (Thierry *et al*, 2015).

f) Lens Epithelium-Derived Growth Factor (LEDGF/p75)

LEDGF/p75 est le cofacteur de l'intégrase le plus documenté à ce jour. LEDGF/p75 est une protéine de 530 acides aminés codé par le gène *psip1* et constitué de plusieurs domaines fonctionnels (**Figure 29**) : le domaine PWWP situé dans le domaine N-terminal de la protéine qui est responsable de l'interaction avec la modification épigénétique H3k36me3 (Eidahl *et al*, 2013), le domaine NLS (₁₄₈GRKRKAQ₁₅₆) (Maertens *et al*, 2004), un domaine comportant deux motifs AT-hook et un domaine IBD (intégrase binding domain) conservé capable d'interagir avec l'intégrase du VIH-1 de 80 acides aminés (347-429) au niveau du C-terminal de la protéine (Cherepanov *et al*, 2004). La structure d'un dimère de CCD de l'intégrase interagissant avec deux monomères d'IBD (Cherepanov *et al*, 2005) ainsi que la structure du NTD-CCD de l'intégrase du VIH-2 en complexe avec l'IBD de LEDGF/p75 a pu être résolue (**Figure 30**) (Hare *et al*, 2009 b). La stœchiométrie du complexe IN/LEDGF/p75 n'est pas élucidée, mais des modèles moléculaires suggèrent un complexe de 400 kDa constitué d'un tétramère d'IN avec deux LEDGF/p75. Des expériences de double-hybride ont permis de montrer que le résidu Q168 de l'IN est essentiel à l'interaction avec LEDGF/p75 (Emiliani *et al*, 2005). LEDGF/p75 a été montré par co-immunoprécipitation et double hybride comme interagissant avec l'intégrase du VIH-1. De plus, l'ajout de LEDGF/p75 au test de transfert de brin augmente l'activité de l'IN suggérant la contribution de LEDGF/p75 durant le processus d'intégration *in vitro* (Cherepanov *et al*, 2003). Les expériences cellulaires montrent que la déplétion de LEDGF/p75 abolit la localisation nucléaire de l'IN et son association avec la chromatine (Maertens *et al*, 2003).

Genre	Unités de transcription	+/- 2kb promoteurs	+/- 2kb ilots CpG
Lentivirus	++	0	0
Alpharétrovirus	+	0	+
Betarétrovirus	0	0	0
Gammarétrovirus	+	++	++
Deltarétrovirus	+	+	+
Spumavirus	0	++	++
Epsilonrétrovirus	ND	ND	ND

Figure 31 : Sélectivité d'intégration des différents rétrovirus. Adapté de (Desfarges & Ciuffi, 2010). La préférence de sélectivité d'intégration en fonction des rétrovirus est représentée par 0 (nul), + (faible), ++ (importante) et ND (non déterminé). Un intervalle de 2 kb autour des sites d'intégration a été pris en compte.

Il a ensuite été démontré que la déplétion de LEDGF/p75 permet la réduction de la réplication d'un facteur deux à quatre (Vandekerckhove *et al*, 2006). Le séquençage des sites d'intégration dans des cellules infectés KO LEDGF montre une redistribution des sites d'intégration (Shun *et al*, 2007; Hombrouck *et al*, 2007). L'ensemble de ces données suggère un rôle de ciblage de LEDGF/p75 vers les unités de transcriptions actives de la chromatine permettant ainsi d'expliquer la sélectivité de l'intégration, c'est-à-dire le choix du site d'intégration, du VIH-1, dans l'ADN cellulaire.

2.5. Sélectivité du VIH-1

a) Introduction

Afin d'assurer leur réplication, les rétrovirus doivent insérer leur copie d'ADN rétrotranscrit à partir de leur génome ARN dans la chromatine de la cellule infectée. Le choix du site d'intégration dans la chromatine n'est pas un processus aléatoire et apparaît comme spécifique du rétrovirus (**Figure 31**). En effet, pour les lentivirus, l'intégration est favorisée dans les unités de transcription active (Schröder *et al*, 2002) alors que les sites d'initiation de la transcription et les îlots CpG sont favorisés dans le cas des gamma-rétrovirus. En revanche, les alpha-rétrovirus ne montrent pas de préférence de sélectivité marquée (Mitchell *et al*, 2004). La sélectivité de l'intégration est un processus déterminant. En effet, l'expression des gènes viraux va varier en fonction du site d'intégration. L'intégration dans une unité de transcription active va permettre une forte expression des gènes viraux aboutissant à une forte infectivité virale. Au contraire, l'intégration dans des régions silencieuses hétérochromatiniennes va conduire à une faible transcription conduisant à un potentiel mécanisme de latence viral (Lewinski *et al*, 2005).

De nombreux mécanismes peuvent influencer la sélectivité de l'intégration : la structure de la chromatine, en fonction de sa condensation peut influencer la sélectivité d'intégration, il a été montré qu'elle peut favoriser ou être inhibitrice à l'accès du CPI. L'intégration rétrovirale lors de certaines phases du cycle cellulaire quand la chromatine est remodelée pour la réplication de l'ADN peut favoriser l'intégration. De plus, les facteurs cellulaires capables

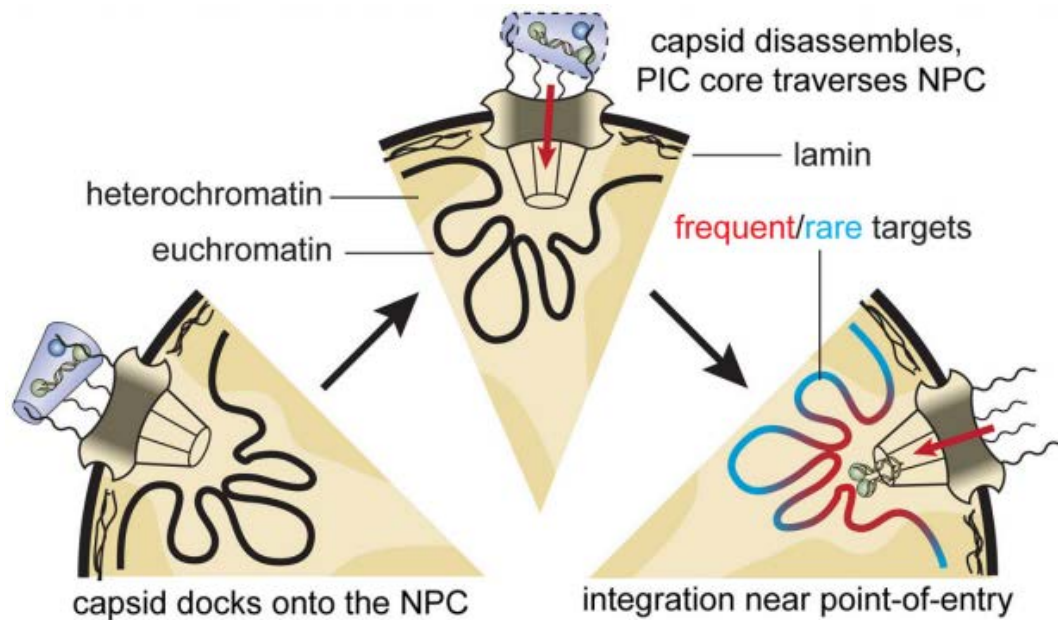


Figure 32 : Représentation schématique de l'entrée de virus dans le noyau et de son intégration (Demeulemeester *et al*, 2015). PIC : Pre integration complex, NPC : Nuclear pore complex.

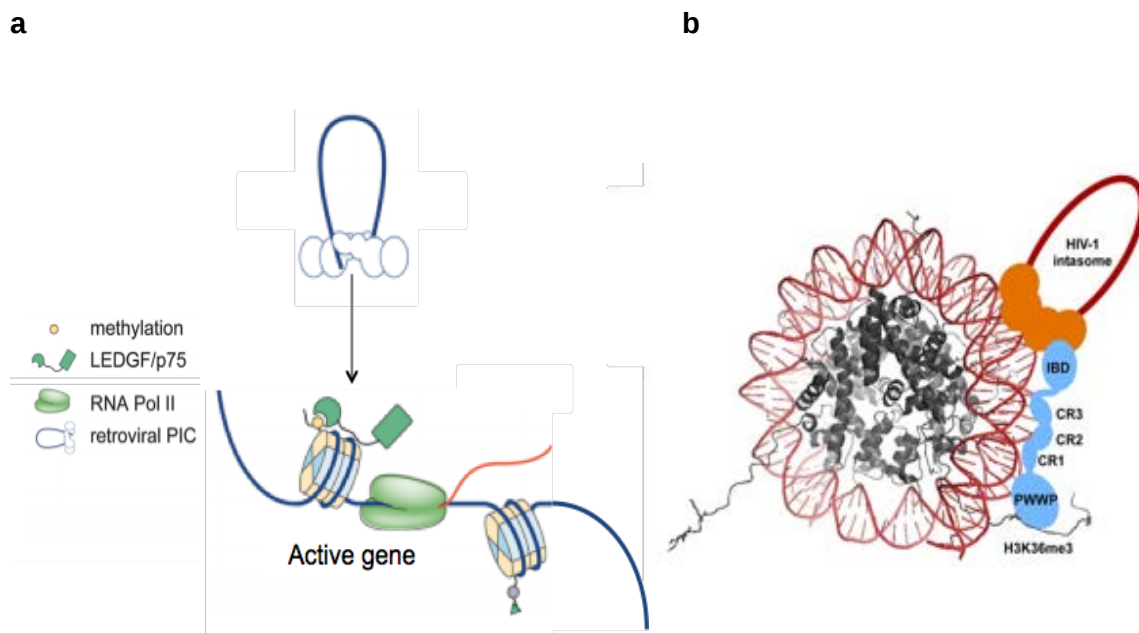


Figure 33 : Modèle du ciblage de l'intégration via LEDGF/p75 (a) et modèle d'interaction du CPI avec le nucléosome (Kvaratskhelia *et al*, 2014; Demeulemeester *et al*, 2015). (a) Le CPI du VIH-1 est ciblé par LEDGF/p75 vers les régions transcriptionnellement actives du génome où les marques H3K36me3 sont présentes. (b) LEDGF/p75 (bleu) est capable d'interagir spécifiquement avec le mononucléosome via d'une part l'interaction entre le domaine PWWP et la marque H3K36me3 et d'autre part les différentes CR (charged regions) avec l'ADN nucléosomal (rouge) afin de cibler l'intasome du VIH-1 (représenté par un tétramère d'IN en orange et l'ADN viral en rouge).

d'interagir avec la chromatine et avec le CPI permettent de cibler l'intégration vers des régions spécifiques de la chromatine. De plus, le séquençage massif de 40 569 sites d'intégrations du VIH-1 a pu montrer que l'intégration du VIH-1 était favorisée à l'extérieur de sillon majeur de l'ADN nucléosomale et était associée à des modifications épigénétiques liées à une transcription active (Wang *et al*, 2007). La sélectivité de l'intégration constitue un enjeu majeur. En effet, l'intégration rétrovirale peut aussi affecter l'expression de gènes cellulaires dû à l'insertion de promoteurs viraux proche de gènes cellulaires aboutissant à l'apparition de processus oncogéniques notamment par l'utilisation de vecteurs rétroviraux lors de thérapie génique (Hacein-Bey-Abina *et al*, 2003).

b) Import nucléaire et intégration

Les lentivirus se caractérisent par leur capacité à s'intégrer dans des cellules qui ne se divisent pas. Ainsi, l'entrée du CPI dans le noyau nécessite un processus actif mettant en jeu les pores nucléaires, des nucléoporines tels que Nup153 et Nup98-Nup96 (Di Nunzio *et al*, 2013) et des facteurs d'imports nucléaires tels que Transportine-SR2 (TRN-SR2, TNPO3) (Christ *et al*, 2008). Suite à l'import nucléaire, le CPI doit accéder à la chromatine de la cellule infectée. Ce processus reste encore mal connu, mais des études d'imageries cellulaires suggèrent que l'intégration se localise dans les régions euchromatiques à la périphérie du noyau proche du pore nucléaire (Albanese *et al*, 2008; Di Primio *et al*, 2013). Ces données basées sur l'architecture nucléaire contribuent à un premier biais dans la sélectivité de l'intégration vers des gènes actifs enrichis en modifications d'histone liées à une transcription active. Ces données suggèrent un lien étroit entre l'import nucléaire et le choix du site d'intégration (**Figure 32**). En effet la déplétion des facteurs d'imports nucléaires tels que TRN-SR2, Nup358/RanBP2, Nup153, Nup98-Nup96 et Tpr altèrent le choix du site d'intégration suggérant un rôle important de ces protéines dans ce processus de sélectivité (Ocwieja *et al*, 2011; Lelek *et al*, 2015). De plus, des études de la CA montrent qu'elle joue un rôle dans le choix du site d'intégration (Schaller *et al*, 2011). En effet, lors du remplacement du gène *gag* du VIH-1 par son homologue MLV, le virus hybride montre une intégration diminuée vers des régions denses en gènes (Ocwieja *et al*, 2011). De plus, des mutants de la CA qui perturbent son interaction avec Nup358, Nup153 et cyclophiline A affectent le choix du site d'intégration.

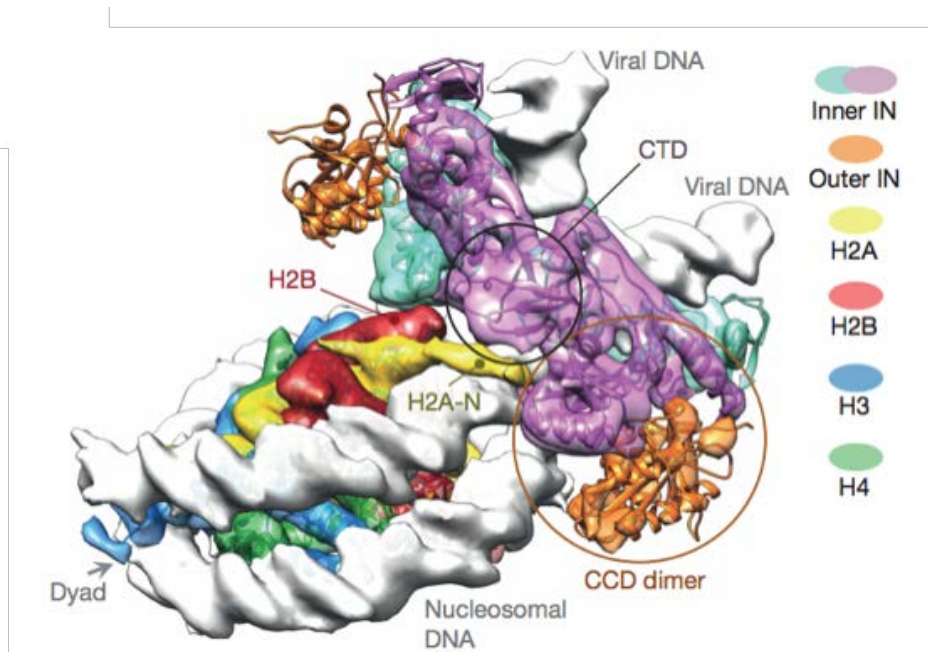


Figure 34 : Structure du complexe de l'intasome PFV avec un nucléosome (Maskell et al, 2015).

c) Facteurs cellulaires participant à la sélectivité de l'intégration

Dès l'entrée du CPI dans le noyau des cellules infectées, l'IN détourne des facteurs cellulaires à son avantage afin de cibler son intégration dans des régions spécifiques du génome. Dans le cas des lentivirus, le co-facteur de ciblage est LEDGF/p75 (Cherepanov, 2007) alors que les gamma-rétrovirus interagissent avec les protéines BET (De Rijck *et al*, 2013). L'étude de la sélectivité du VIH-1 a montré qu'une déplétion de LEDGF/p75 réduit l'intégration du VIH-1, mais aussi change le choix du site d'intégration loin des gènes actifs (Ciuffi *et al*, 2005; Marshall *et al*, 2007). De plus des protéines LEDGF/p75 chimériques dans lesquelles le domaine N-terminal de fixation à la chromatine a été remplacé par ceux d'autres protéines CBX1, HP1 et ING2 montrent une redirection de l'intégration vers les régions reconnues par ces protéines (Gijssbers *et al*, 2010; Ferris *et al*, 2010; Silvers *et al*, 2010). Par ailleurs, un rôle de la protéine HRP-2 a pu être mis en évidence dans la sélectivité de l'intégration du VIH-1. En effet, lors de la déplétion de LEDGF/p75, la protéine HRP-2 permettrait le ciblage de l'intégration vers les régions transcrites du génome (Schrijvers *et al*, 2012). Cependant, il a été montré que lors de la double déplétion LEDGF/p75 et HRP-2, une intégration résiduelle dans les régions actives est réalisable suggérant alors la participation de facteurs additionnels (Wang *et al*, 2012). L'ensemble de ces données est en faveur d'un deuxième biais de la sélectivité du VIH-1 où LEDGF/p75 et d'autres facteurs cellulaires au sein du CPI sont capables de cibler l'intégration vers les régions transcrites de la chromatine (**Figure 33**).

d) Le nucléosome substrat de l'intégration

Les récentes études de cryo-microscopie du complexe Intasome/nucléosome montrent que l'intasome de PFV est capable de catalyser son intégration sur mononucléosome (**Figure 34**) (Maskell *et al*, 2015). Dans le cas du PFV, cette intégration se réalise préférentiellement à deux positions symétriques ± 36 bp de l'axe dyade (axe passant par le centre du nucléosome et par la 74^{ème} paire de base) du nucléosome. Les données d'interaction montrent une interaction entre le domaine N-terminal de l'histone H2A et le domaine C-terminal d'une sous unité interne de l'intasome et une interaction entre le domaine C-terminal d'une sous unité extérieure avec d'autres parties du nucléosomes permettant la lecture des marques épigénétiques.

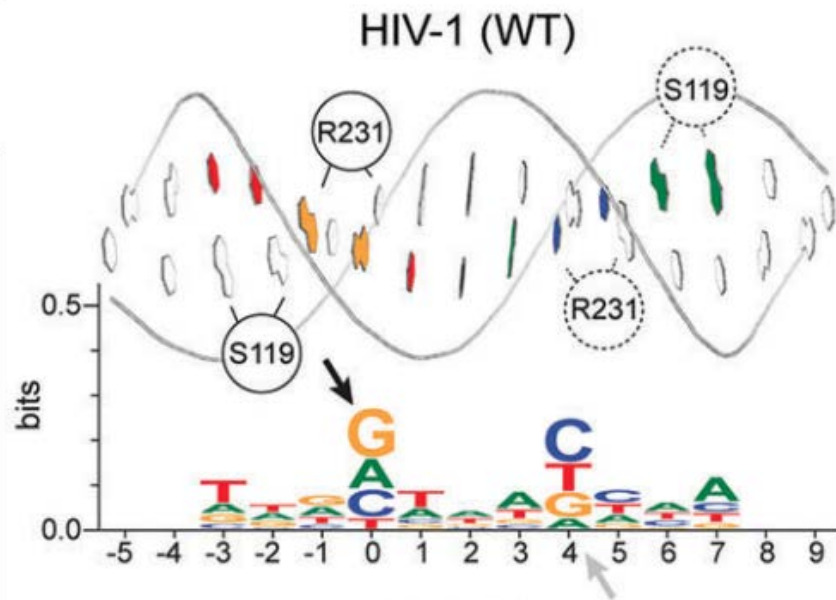


Figure 35 : Consensus de séquence des sites d'intégration du VIH-1 dans l'ADN cible. Représentation en logo de la séquence des sites d'intégration du VIH-1, les flèches indiquent les sites de transfert de brin (position 0) sur le brin plus (noir) et sur le brin moins (gris). Une représentation de l'ADN cible avec les différents contacts nucléotides/acides aminés est représentée en haut (Demeulemeester *et al*, 2015).

Dans le cas du VIH-1, les études *in vitro* sur des mononucléosomes ou sur des mini-chromosomes purifiés ont montré que l'intégration rétrovirale était favorisée sur ce type de structure par rapport à de l'ADN nu (Pryciak & Varmus, 1992; Pruss *et al*, 1994). Ces études sont supportées par les données de séquençage des sites d'intégration qui ont montré que l'ADN cible de l'intégration est enroulé à la surface du nucléosome. De plus, d'autres études sur des structures polynucléosomales montrent que l'environnement chromatinien joue un rôle prépondérant sur l'intégration *in vitro* du VIH-1 (Benleulmi *et al*, 2015). De plus, l'IN du VIH-1 possède une préférence de séquence dans l'ADN cible. Le rôle de la séquence de l'ADN cible dans la sélectivité est majoritairement local, mais une séquence consensus a pu être déterminée. Cette séquence de type (-3) TDG\G/V TWA (C/B) CHA(+7) :D: pas C, :V: pas T, :W: A ou T, :H: pas G) suggère donc un troisième biais dans la sélectivité du VIH-1 (**Figure 35**) (Serrao *et al*, 2015).

e) Conclusion

La sélectivité de l'intégration est un processus en plusieurs étapes qui commence par l'import nucléaire du CPI. Les IN lentivirales sont capables de se lier à un cofacteur cellulaire LEDGF/p75 afin de cibler l'intégration vers des régions transcrites de la chromatine, alors que les gamma-rétrovirus se lient aux protéines BET. LEDGF/p75 et les protéines BET sont qualifiés de « chromatine readers proteins » et possèdent de multiples points d'accroches dans la chromatine. Suite à son entrée du CPI dans le noyau et son ciblage vers les régions transcrites de la chromatine, le nucléosome apparaît comme substrat de l'intégration où la réaction de transfert de brin peut se réaliser permettant la fusion des deux génomes viral et cellulaire. Les interactions de l'intasome avec la séquence d'ADN cible et le nucléosome semblent donc cruciales dans le choix du site d'intégration du VIH-1.

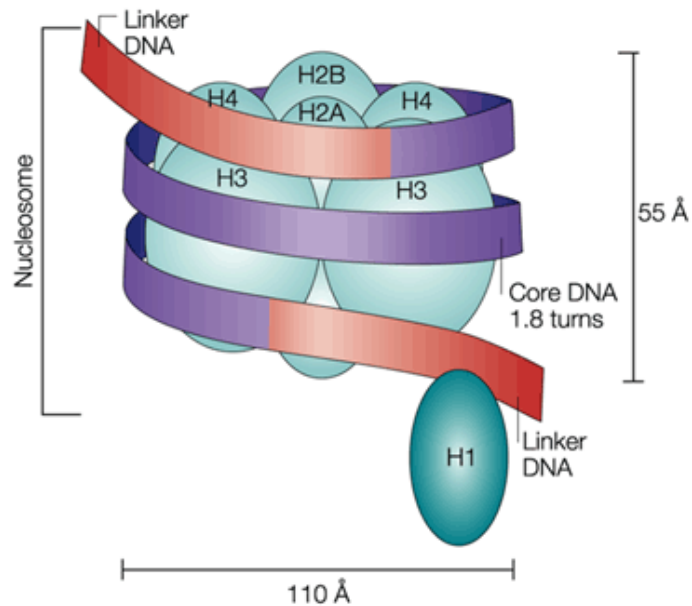


Figure 36 : Composition d'un nucléosome (Georgopoulos, 2002). Le nucléosome est constitué de deux copies de chacune des quatre histones H2A, H2B, H3 et H4 et sur lequel est enroulé un ADN de 145 à 1 7 pb. Les histones de liaison H1 stabilisent l'enroulement de l'ADN sur le nucléosome.

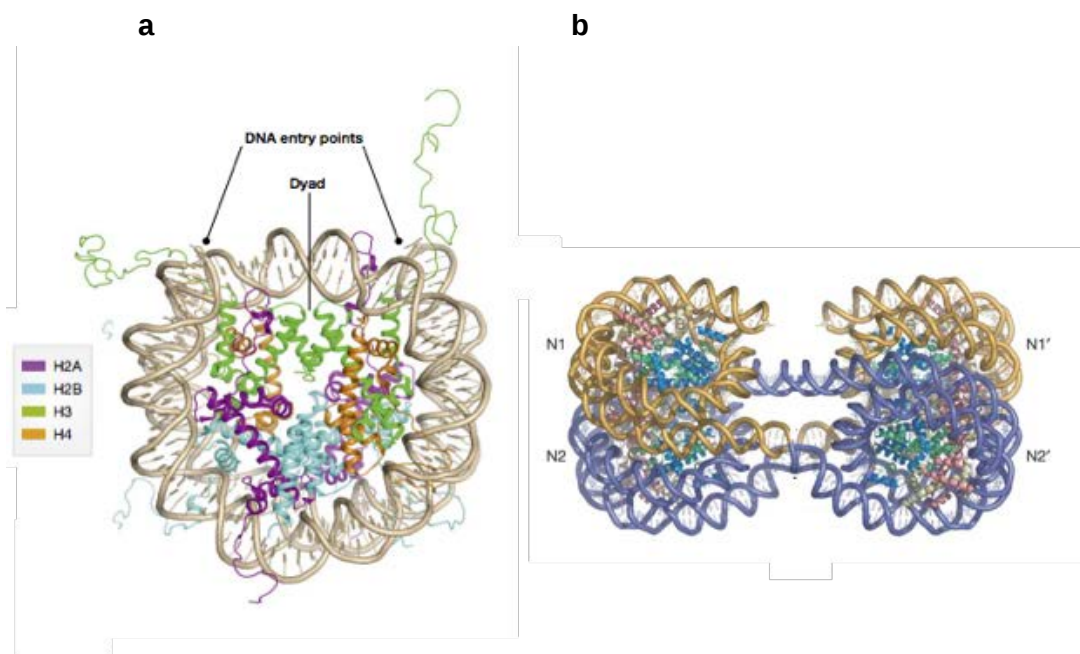


Figure 37 : Structure d'un mononucléosome et d'un tétranucléosome (Schalch *et al*, 2005; Speranzini *et al*, 2016). (b) N1, N2, N1' et N2' représentent les quatre nucléosomes.

3. Chromatine

L'intégrase au sein du CPI va être importée vers le noyau pour y intégrer son ADN viral nécessaire à l'expression du provirus. La cible de l'intégration est l'ADN chromatinisé, cet ADN possède une structure particulière et dynamique.

3.1. Chromatine cellulaire

a) Structure de la chromatine

Dans tous les systèmes eucaryotiques, l'ADN est compacté sous forme de chromatine. La condensation de la chromatine va permettre à un ADN de près de 2 mètres d'entrer dans un noyau de 5 à 20 μm de diamètre. La chromatine est constituée d'une succession de nucléosomes sous la forme de nucléofilament de 10 nm ou « collier de perles ». Ce nucléofilament est capable de se replier pour former une fibre de 30 nm de diamètre représentant un ordre supérieur de compaction (Wu *et al*, 2007). Chaque nucléosome contient un octamère d'histone (contenant deux molécules de chaque histone H2A, H2B, H3, H4) et occasionnellement d'histones de liaison H1 ou H5 sur lequel est enroulé un ADN de 145-147pb (**Figure 36**) (Luger & Hansen, 2005; Woodcock, 2006). De nombreux facteurs cellulaires ont été décrits comme s'associant avec les protéines histones comme les protéines HMGA, HMGB et HMGIC (Hock *et al*, 2007). Ces protéines non-histones sont associées à la chromatine et contribuent à son organisation. La structure cristallographique du nucléosome a été résolue en 1997 (**Figure 37**) (Czarnota *et al*, 1997; Chakravarthy *et al*, 2005) ainsi que la structure d'un tétranucléosome en 2005 (Schalch *et al*, 2005). Les queues N-terminales des histones, flexibles, ne sont pas définies dans la structure cristallographique sauf la queue H4 qui établit des interactions avec un dimère H2A/H2B d'un nucléosome voisin et se trouve ainsi stabilisée. Ces queues portent de nombreuses modifications post-traductionnelles permettant la régulation de la compaction de la chromatine, son accessibilité et le recrutement de protéines spécifiques.

b) Les histones de cœur

Les histones qui constituent l'octamère du nucléosome sont les histones de cœur. Les histones sont des protéines basiques de 11 à 16 kDa. Ce cœur protéique contient deux copies de chacune des quatre histones H2A, H2B, H3 et H4. Ces protéines partagent une

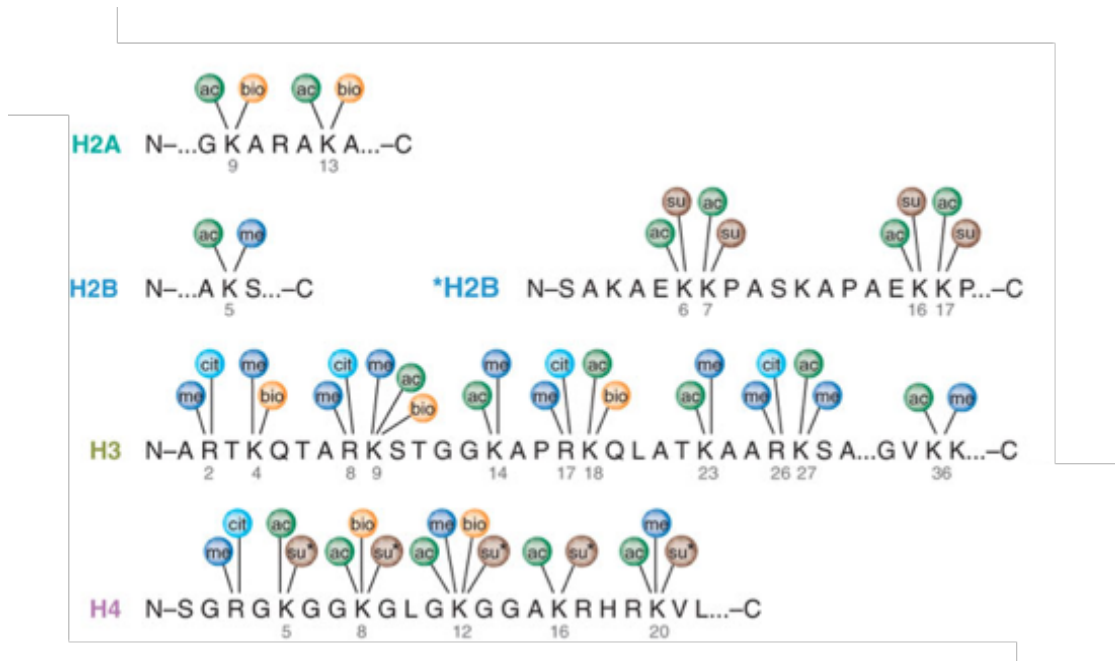


Figure 38 : Exemple de modifications post-traductionnelles des histones (Latham & Dent, 2007). ac : acétylation, me : méthylation, bio : biotinylation, su : sumoylation, cit : citrullination.

structure commune appelée « histone fold » qui comprend trois hélices α séparées par deux boucles (Arents *et al*, 1991). Ces protéines très conservées au cours de l'évolution se composent d'un domaine central globulaire et des extrémités N et C terminales non structurées appelées queues d'histones. Ces queues d'histones sont plus variables et sont dépourvues de structure secondaire. Il a été montré que ces queues riches en résidus lysine et arginine chargées positivement influencent l'accessibilité de l'ADN nucléosomal aux facteurs cellulaires. Les queues d'histones sont la cible de nombreuses modifications post-traductionnelles variées telles que l'acétylation, la méthylation, la phosphorylation, l'ubiquitination, la sumoylation, l'isomérisation et la déamination.

c) Variants d'histones

De nombreux variants d'histones ont été décrits parmi lesquels les variants H2AZ, H2AX, macroH2A, H2A.Bbd, H3.3 et CenH3 (Jin *et al*, 2005; Bernstein & Hake, 2006; Loyola & Almouzni, 2007). De plus, aucun variant de l'histone H4 n'a pu être déterminé. H2AZ est retrouvé dans des régions réprimées du génome sous sa forme non-acétylée et dans des régions actives sous sa forme acétylée. H2AX est un variant retrouvé au niveau des sites de cassures double brin d'ADN (Stucki *et al*, 2005). MacroH2A est retrouvé enrichi dans le chromosome X inactif chez le mammifère (Costanzi & Pehrson, 1998). H2ABbd est localisé dans des régions transcriptionnellement actives de la chromatine et est exclu des régions inactives. H3.3 est incorporé durant la transcription active (Hake & Allis, 2006). CenH3 ou CenpA est localisé au sein de l'hétérochromatine. Le rôle exact de ces différents variants d'histones reste à être déterminé, mais leurs fonctions semblent être reliées à leur localisation.

d) Histones de liaison

Les histones de liaison H1 se fixent à l'ADN qui lie deux nucléosomes adjacents. Ces histones ne présentent pas d'homologie structurale avec les histones de cœur formant l'octamère. L'histone H1 possède un domaine central globulaire et deux extrémités N et C terminales non structurées et modifiées post-traductionnellement (Godde & Ura, 2008). Les histones de liaisons ont été démontrées comme jouant un rôle essentiel dans la stabilisation de la fibre de 30 nm de diamètre (Fan *et al*, 2005). De plus, les sites d'interactions de l'histone H1 avec le nucléosome ont été identifiés par mutagenèse dirigée (Brown *et al*, 2006).

3.2. Modifications post-traductionnelles d'histones

Les queues d'histones et plus rarement le domaine globulaire portent de nombreux sites de modifications post-traductionnelles. Ces marques d'histones possèdent deux rôles majeurs : elles permettent de créer un site de fixation pour les protéines ayant un domaine de reconnaissance pour ces différentes marques. Elles permettent aussi de modifier la compaction de la chromatine (Kouzarides, 2007). Il existe huit types de marques covalentes : acétylation, méthylation (lysine mono, di ou triméthylée et arginine mono et diméthylée), phosphorylation, ubiquitinylation, sumoylation, adénosine diphosphate - ribosylation, déamination et isomérisation (**Figure 38**). Les enzymes responsables de ces modifications sont bien caractérisées ainsi que les enzymes responsables du retrait de ces modifications incluant des histones désacétylases, histone déméthylases et déubiquitylases.

a) Acétylation

L'acétylation est catalysée par des histones acétyl-transférases (HAT) utilisant comme donneur du groupement acétyl, l'acétyl-coenzymeA (CoA) (Peterson & Laniel, 2004). Dans cette réaction, le groupement amino NH_3^+ est remplacé par le groupement NH-CO-CH_3 non chargé. Cette neutralisation de charge affaiblit l'interaction entre les histones chargées positivement et l'ADN chargé négativement. Ces HAT ciblent de nombreuses lysines sur différentes histones. La réaction inverse de déacétylation est réalisée par des histones déacétylases (HDAC).

i- Bromodomaine

Une structure bien définie appelée bromodomaine est retrouvée dans les sous-unités de nombreuses HAT et dans les complexes protéiques de remodelage de la chromatine ATP-dépendant (Jeanmougin *et al*, 1997). Ce domaine conservé d'environ 110 acides aminés a été décrit comme reconnaissant les marques acétylées (Mujtaba *et al*, 2007). De plus, il a été montré récemment que les protéines à bromodomaine BRD2, BRD3 et BRD4 interagissent directement avec l'IN gamma rétroviral de MLV afin de permettre le ciblage du CPI vers les histones acétylées du génome (De Rijck *et al*, 2013).

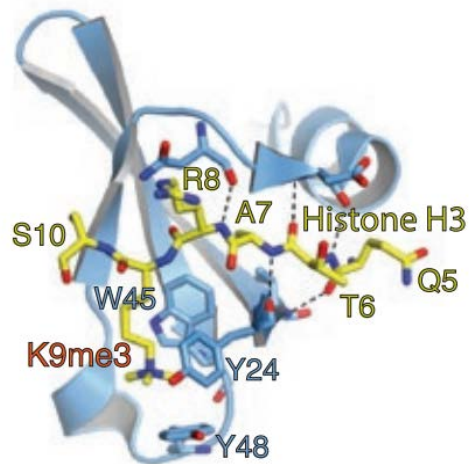


Figure 39 : Structure du chromodomaine de HP1 en complexe avec le peptide H3K9me3 (Jacobs & Khorasanizadeh, 2002). HP1 est représenté en bleu, les acides aminés formant la cage aromatique Y24, W45 et Y24 sont annotés en bleu. Le peptide H3K9me3 est représenté en jaune, et les atomes d'azotes et d'oxygènes en bleu et rouge respectivement (P B : 1KNE).

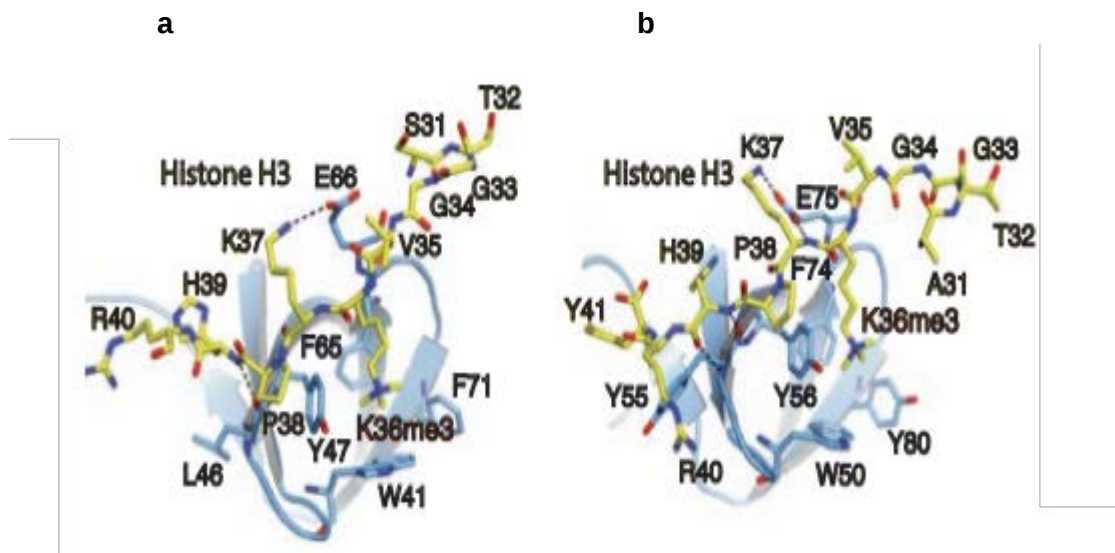


Figure 40 : Structure du Tudor domaine de PHF1 et PHF19 en complexe avec le peptide H3K36me3 (Musselman *et al*, 2012a; Ballaré *et al*, 2012). (a) PHF1 est représenté en bleu. Le peptide H3K36me3 est en jaune, et les atomes d'azotes et d'oxygènes en bleu et rouge respectivement (P B : 4HCZ) (b) Même représentation qu'en figure (a) (P B : 4BD3).

b) Méthylation

La méthylation se produit sur les résidus lysines et les arginines (Martin & Zhang, 2005). Sur les lysines, les méthyle-transférases (méthylase) catalysent le transfert d'un groupement méthyle de la S-adénosyl méthionine (SAM) sur la lysine remplaçant un, deux ou trois hydrogènes sur le groupement amino NH_3^+ . Contrairement aux acétyl-transférases, ces enzymes apparaissent comme spécifiques et ne possèdent qu'un substrat de méthylation. Les résidus arginines sont méthylés par des protéines arginine-méthylases (PRMT) permettant la méthylation des H3R2, R8, R17, R26 et H4R3. De nombreux domaines sont décrits comme permettant la reconnaissance des queues d'histones méthylées (Taverna *et al*, 2007). Ces domaines forment la « Royal family » (Maurer-Stroh *et al*, 2003) incluant les chromo, Tudor, MBT (Malignant brain tumor), chromo barrel et PWWP domaines. La différence entre ces quatre derniers domaines est faible, mais diffère majoritairement par le nombre et les différentes orientations du tonneau α . Les protéines comportant ces domaines sont capables d'interagir avec de nombreuses marques d'histones présentant une lysine méthylée ou une arginine méthylée pour les domaines Tudor. Le processus inverse de déméthylation est réalisé par deux familles de déméthylases appelés JUMONJI (JMJ) et amine oxydase.

i- Chromodomaine

Le premier chromodomaine a été identifié au sein de la protéine HP1 (heterochromatin protein 1) chez *D.melanogaster*. Il a été décrit que cette protéine ayant un rôle dans la répression des gènes était capable de reconnaître l'histone H3 méthylée sur la lysine 9 (Bannister *et al*, 2001; Lachner *et al*, 2001). Ce chromodomaine est composé de trois feuillets β antiparallèles liés à une hélice α (Ball *et al*, 1997). La structure de HP1 montre une fixation du peptide H3K9me2 et H3K9me3 (**Figure 39**) (Jacobs & Khorasanizadeh, 2002; Nielsen *et al*, 2002). Cette interaction se localise dans une cage aromatique formée par les résidus Tyr24, Tyr45 et Tyr48 permettant la fixation de la méthyle-lysine via des interactions cations- π .

ii- Tudor domaine

Le domaine Tudor le mieux caractérisé est présent chez les protéines PHF1 et PHF19 (plant homeodomain finger 1 et 19). Ces protéines reconnaissent via leur Tudor

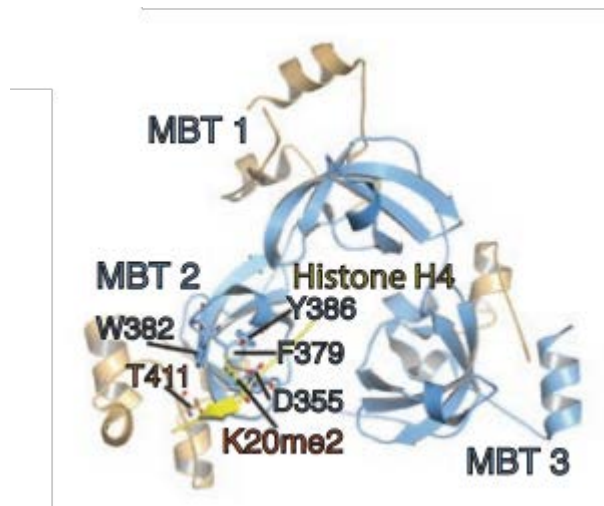


Figure 41 : Structure du domaine MBT de L3MBTL1 en complexe avec le peptide H4K20me2 (Min *et al*, 2007). Structure des trois domaines MBT (MBT1, MBT2, MBT3) de L3MBTL1 avec le peptide H4K20me2. Le peptide H4K20me2 est en jaune, et les atomes d'azotes et d'oxygènes en bleu et rouge respectivement. Les acides aminés formant la cage aromatique de MBT2 sont annotés. L3MBTL1 est représenté en bleu (P B : 2RJE).



Figure 42 : Structure du Chromo barrel domaine de MSL3 en complexe avec le peptide H4K20me1 et l'ADN (Kim *et al*, 2010). MSL3 est représenté en bleu, le peptide H4K20me1 en jaune et l'ADN en vert (P B : 3OA6).

domaine l'histone H3 di ou triméthylée sur la lysine 36 (H3K36me2/3). La structure tridimensionnelle de PHF1-Tudor et PHF19-Tudor fixés au peptide H3K36me3 a pu être déterminée par résonance magnétique de surface (RMN) et par cristallographie aux rayons X (**Figure 40**) (Ballaré *et al*, 2012; Musselman *et al*, 2012a; Cai *et al*, 2013). Ces données montrent la présence d'une structure en tonneau α composée de 5 feuillets α antiparallèles. Dans cette structure, le peptide H3K36me3 se fixe dans une cavité aromatique formée par les résidus Trp41, Tyr47, Phe65 et Phe71 via des interactions cations- π avec les résidus Trp41 et Phe65.

iii- Domaine MBT

Le domaine MBT a été identifié initialement chez *D.melanogaster*. Des mutations dans le gène *mbt* conduisent à l'apparition d'un développement tumoral. La protéine humaine L3MBTL1 est la première protéine MBT qui a pu être décrite au point de vue structural (Wang *et al*, 2003). Cette protéine comprend plusieurs répétitions du motif MBT. Chaque répétition comprend un repliement en tonneau α composé de 5 feuillets α antiparallèles (α 1 - α 5) avec une hélice 3_{10} entre les feuillets α 3 et α 4 ainsi qu'une hélice α et un court feuillet α (α 6) au niveau du C-terminal. Les structures des trois domaines MBT de la protéine L3BMTL1 fixée au peptide H4K20me1 et H4K20me2 ont pu montrer que le second domaine MBT possède une cage aromatique constituée des résidus Phe379, Trp382, Tyr386 et Thr411 permettant la fixation du peptide diméthylé sur la lysine 20 (**Figure 41**). Cette interaction ($K_D = \alpha 5\mu M$) pour H4K20me1 ou H4K20me2 semble jouer de multiples rôles qui restent à être déterminés (Li *et al*, 2007; Min *et al*, 2007).

iv- Chromo barrel domaine

Le domaine chromo barrel possède une structure proche des domaines Tudor et MBT. MSL3 est une sous unité du complexe MSL (male-specific lethal) recruté au niveau de ses gènes cibles via la reconnaissance de son domaine chromo barrel en N-terminal (MSL3-CB) avec des histones méthylées (Kim *et al*, 2010; Moore *et al*, 2010). Les études *in vitro* montrent une préférence d'interaction de MSL3-CB pour les peptides H4K20me1 et H4K20me2 avec une affinité de $K_D=\alpha 30$ et $\alpha 220 \mu M$ respectivement. MSL3 adopte un repliement en tonneau α suivi d'une hélice α en C-terminal. Dans la structure, le peptide H4K20me1 se fixe dans une cage aromatique formée par les résidus Tyr31, Phe56, Trp59 et Trp63 (**Figure 42**) (Kim *et al*, 2010). De plus, MSL3 est capable d'interagir avec l'ADN

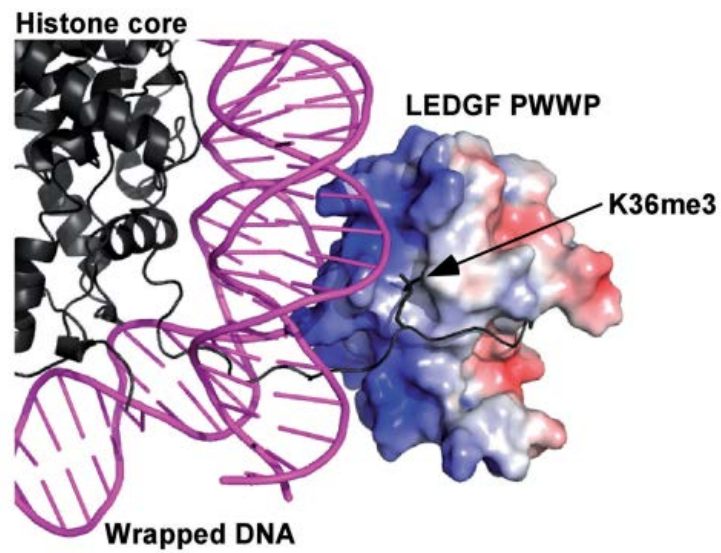


Figure 43 : Modèle d'interaction entre le domaine PWWP de LEDGF/p75 et la modification H3K36me3 (Eidahl *et al*, 2013).

proche de la cage aromatique, suggérant une interaction simultanée de MSL3 avec l'histone H4K20me1 et l'ADN.

v- Domaine PWWP

Le domaine PWWP décrit un domaine possédant un motif Pro-Trp-Trp-Pro initialement identifié chez la protéine WHSC1 (Wolf-Hirschhorn syndrome candidate 1) (Stec *et al*, 2000). LEDGF/p75 décrite précédemment (voir chapitre 2.4.) possède un domaine PWWP. La structure de ce domaine montre un repliement en tonneau α constitué de 5 feuillets α antiparallèles avec une hélice 3_{10} entre les feuillets α 4 et α 5. Récemment, il a été montré que le domaine PWWP de LEDGF/p75 est capable d'interagir avec le peptide H3K36me3 (**Figure 43**) (Eidahl *et al*, 2013). Le peptide se fixe dans une cavité hydrophobe formée par les résidus Trp21, Phe44 et Ala51. LEDGF/p75 possède également une fixation simultanée à l'ADN permettant de stabiliser l'interaction au nucléosome.

c) Mono-ubiquitination

La mono-ubiquitination des lysines est une autre modification post-traductionnelle impliquée dans la régulation des gènes (Shilatifard, 2006). Cette modification peut être associée à l'activation ou un silencing des gènes.

d) Phosphorylation

La phosphorylation des histones utilise comme substrat l'ATP, cette phosphorylation contribue à l'activation des gènes (Kouzarides, 2007). Il a été démontré que la phosphorylation de H3S10 est impliquée dans la mitose et l'activation de gènes. Les protéines responsables de la phosphorylation de H3S10 sont les kinases MSK1/2 et RSK2. De plus, H2AXS139 est phosphorylée par la kinase ATM lors de dommages à l'ADN, permettant le recrutement de facteurs cellulaires.

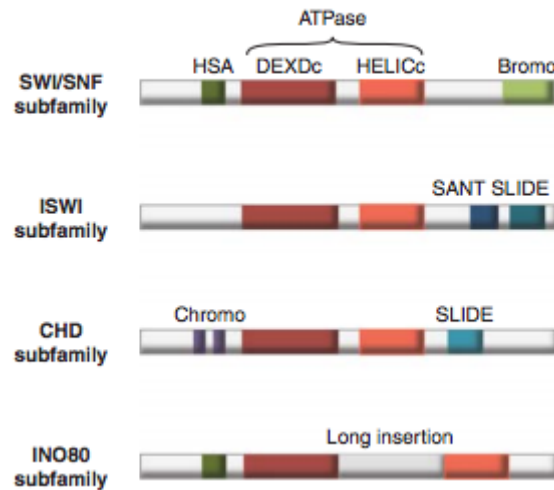


Figure 44 : Organisation schématique des différents domaines composant les complexes de remodelages de la chromatine ATP-dépendants (Becker & Workman, 2013). Le domaine caractéristique des complexes de remodelage ATP-dépendants est le domaine ATPase. La famille SWI/SNF possède un bromodomaine capable de reconnaître les histones acétylées. Les familles ISWI et CHD possèdent un domaine SANT-SLIDE, mais seule la famille CHD possède un tandem de chromodomaine. H A : Helicase SANT domain.

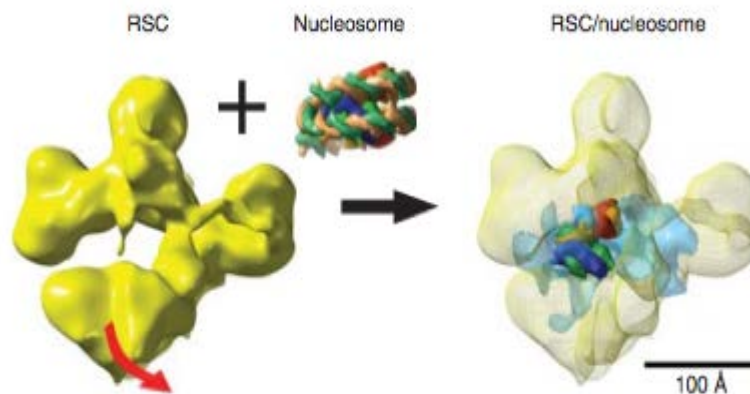


Figure 45 : Structure du complexe RSC-Nucléosome (Leschziner *et al*, 2007).

3.3. Dynamique de la chromatine

Afin de permettre l'ensemble des mécanismes impliquant l'ADN telles que l'expression des gènes, la réparation de l'ADN et la division cellulaire, la chromatine compactée représente une barrière. Ainsi, l'action orchestrée des facteurs de remodelage de la chromatine tels que les histones chaperonnes, les complexes de remodelage ATP-dépendant et les enzymes responsables de modifications post-traductionnelles des histones permettent le réarrangement de l'architecture de la chromatine au niveau du nucléosome afin de rendre l'ADN accessible à ces mécanismes (Clapier & Cairns, 2009; Becker & Workman, 2013).

a) Complexes de remodelage ATP-dépendant

Il existe à l'heure actuelle quatre grandes familles de complexes de remodelages ATP-dépendants : SWI/SNF, ISWI, CHD, et INO80. Ces complexes enzymatiques et leurs sous-unités utilisent l'ATP afin d'induire la dynamique des nucléosomes (**Figure 44**).

- La famille SWI/SNF (switch/sucrose nonfermenting) se caractérise par la présence d'une sous unité ATPase possédant un bromodomaine proche du C-terminal de la protéine (Mohrmann & Verrijzer, 2005). Ainsi, ces bromodomains vont permettre d'augmenter l'affinité de fixation du complexe SWI/SNF pour la chromatine afin d'améliorer leur activité de remodelage des nucléosomes. La structure par microscopie électronique du complexe RSC complexé au nucléosome a été résolue (**Figure 45**) (Smith *et al*, 2003; Leschziner *et al*, 2007). Cette structure montre que le nucléosome est enveloppé par le complexe RSC au sein d'une cavité centrale dans laquelle le nucléosome peut être remodelé.
- La famille ISWI (imitation switch) est caractérisée par la présence de domaine SANT-SLIDE (**Figure 44**) permettant le ciblage de ces protéines vers la chromatine (Mellor & Morillon, 2004; Saha *et al*, 2006). ISWI va s'associer avec différentes sous-unités de différents complexes tels que RSF (remodeling and spacing factor), NURF (nucleosome remodeling factor), CHRAC (chromatin accessibility complex) et ACF (ATP-utilizing chromatin assembly and remodeling factor).

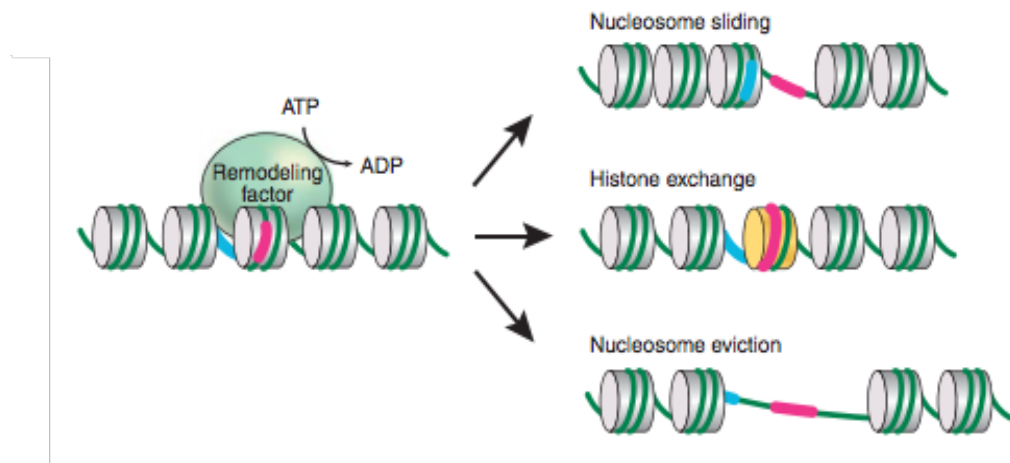


Figure 46 : Mécanismes d'actions des remodeleurs de la chromatine ATP-dépendants (Becker & Workman, 2013). Les différents modèles du remodelage de la chromatine entraînent le changement de position ou de composition du nucléosome. Le modèle du « sliding » entraîne un mouvement de translation du nucléosome afin d'exposer une région préalablement occluse. Il existe aussi le mécanisme d'échange d'une histone de cœur pour un variant d'histone et le mécanisme d'éviction du nucléosome afin de rendre l'ADN accessible.

- La famille CHD possède un tandem de chromodomaines permettant la reconnaissance des lysines méthylées des histones H3 (H3K4me3) ainsi qu'un domaine proche des domaines SLIDE, caractéristique des protéines appartenant à la famille ISWI (Hall & Georgel, 2007).
- La famille INO80 va s'associer aux protéines YY1 permettant leur ciblage vers leurs sites de fixation *in vitro* (Cai et al, 2007).

Bien qu'*in vitro*, plusieurs mécanismes d'action de ces complexes de remodelage aient été décrits, la situation cellulaire reste à être élucidée. Cependant plusieurs processus sont proposés pour ces complexes de remodelage ATP-dépendants (**Figure 46**) (Mueller-Planitz et al, 2013). Le premier mécanisme est celui du « sliding » du nucléosome, mécanisme dans lequel l'octamère d'histone se déplace le long de l'ADN sans s'en dissocier. Afin de réaliser ce processus, de nombreuses interactions entre l'ADN et l'octamère d'histones doivent être rompues puis reformées de façon coordonnée permettant de rendre l'ADN accessible. Le deuxième mécanisme est l'éviction d'un nucléosome. Dans ce mécanisme, les interactions entre l'octamère d'histone et l'ADN sont rompues permettant la dissociation de l'octamère d'histone de l'ADN. Cette éviction peut être totale ou partielle, permettant l'échange d'un core histone par des variants d'histones. En effet, le facteur de remodelage SWR1 est capable d'échanger un dimère H2A/H2B et de le remplacer par un variant d'histone H2A.Z (Ranjan et al, 2015).

b) Enzymes responsables des modifications post-traductionnelles des histones

Les enzymes responsables des modifications post-traductionnelles des histones jouent aussi un rôle important dans le remodelage de la chromatine. En effet, elles vont permettre le recrutement de facteurs de remodelage ATP-dépendant possédant des domaines de reconnaissance de type bromodomaine ou chromodomaine (SWI/SNF, CHD). De plus, elles peuvent affecter la stabilité du nucléosome en altérant les interactions au sein du nucléosome, avec le nucléosome voisin, mais aussi les interactions de l'octamère d'histone avec l'ADN permettant l'obtention d'une chromatine fermée ou ouverte. Ces modifications post-traductionnelles sont réversibles et transitoires. En effet la cellule possède des enzymes distinctes appelées « writers » et « erasers » respectivement, jouant alors sur la dynamique de la chromatine (**Figure 47**) (Musselman et al, 2012b).

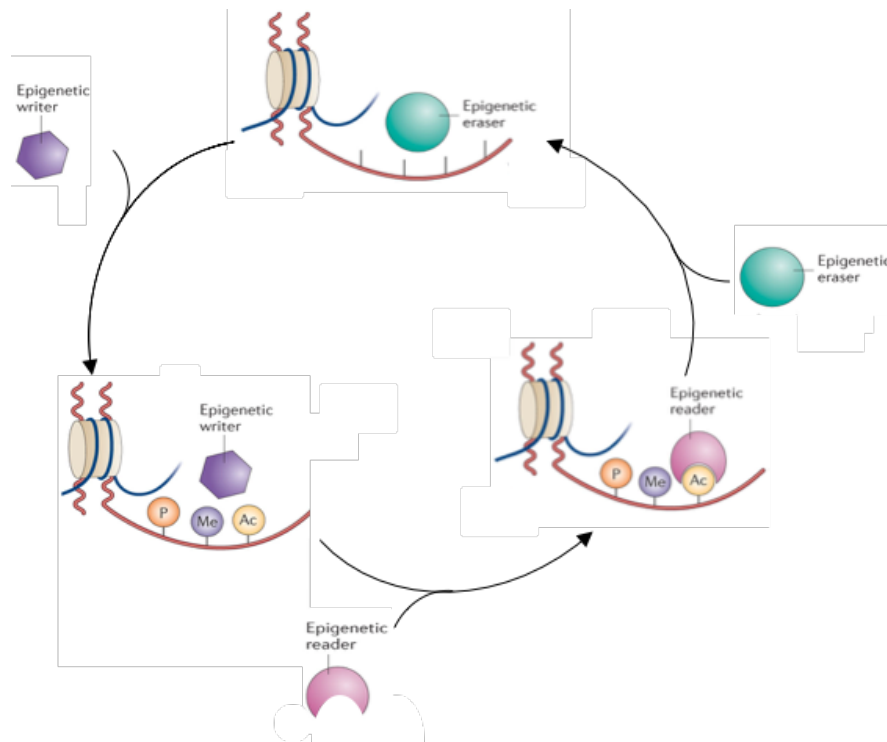


Figure 47 : Représentation schématique de la dynamique des modifications post-traductionnelles des histones. Adapté de (Falkenberg & Johnstone, 2014). Les histones « writer » comprennent les histones acétyl-transférases, histones méthyl-transférases et protéines arginine-méthylases. Les histones « readers » représentent les protéines possédant des domaines de type bromodomaine, chromodomaine et Tudor domaine. Les histones « erasers » comprennent les histones déacétylase (HDAC) et histones déméthylases.

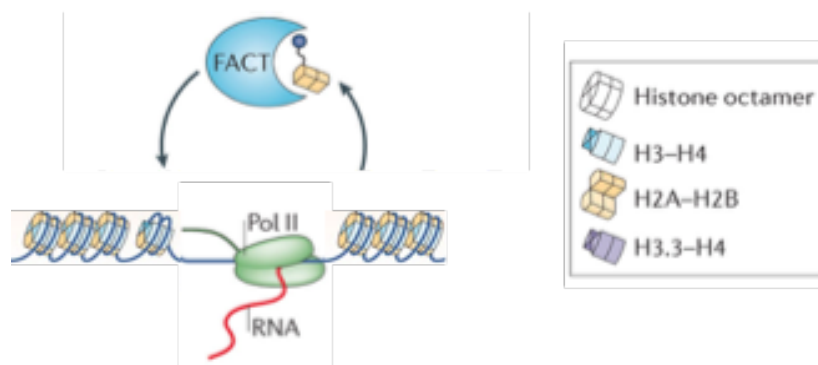


Figure 48 : Représentation schématique de l'activité du complexe de remodelage FACT. Adapté de (Venkatesh & Workman, 2015)

c) Histones chaperonnes

Le dernier acteur majeur de la régulation de la dynamique de la chromatine sont les histones chaperonnes. Ces protéines se définissent comme des protéines interagissant avec les histones et jouent un rôle dans le stockage et le transport des histones, mais aussi dans l'assemblage et le désassemblage du nucléosome. Les histones chaperonnes peuvent être classées en fonction de leur substrat de fixation H2A-H2B ou H3-H4, cependant certaines histones chaperonnes peuvent se fixer sur les deux types de dimères. De plus, certaines histones chaperonnes sont capables de reconnaître une protéine histone spécifique (Burgess & Zhang, 2013). Contrairement aux complexes de remodelages, les histones chaperonnes ne sont pas ATP-dépendantes et utilisent le mouvement de l'ADN autour de l'axe dyade pour déstabiliser le nucléosome. L'histone chaperonne FACT (Facilitates chromatin transcription) est un complexe protéique composé des protéines SSRP1 (structure-specific recognition protein 1) et Spt16 (suppressor of Ty16) (Bondarenko *et al*, 2015). Ce complexe se fixe au dimère H2A-H2B et H3-H4 via la protéine Spt16. Ce complexe joue un rôle dans la réplication, la réparation et la transcription de l'ADN grâce à son activité d'assemblage et de désassemblage du nucléosome (**Figure 48**). En effet, le complexe FACT agit en permettant l'éviction d'un dimère H2A-H2B afin d'obtenir une structure hexamérique du nucléosome. Ce complexe FACT va ensuite permettre le réassemblage du nucléosome après le passage de l'ARN polymérase de type II lors de la transcription. Très récemment, il a été montré que la sous-unité SSRP1 était impliquée dans l'interaction avec le domaine PWWP de LEDGF/p75 (Lopez *et al*, 2016). L'ensemble de ces données suggère un rôle important des facteurs régulant la dynamique de la chromatine dans l'accessibilité de la chromatine lors de l'intégration rétrovirale.

4. Chromatine et intégration

Compte tenu de la propriété intrinsèque de compaction de l'ADN génomique avec les protéines histones sous la forme de chromatine, l'étude de l'impact de cette structure sur l'intégration apparaît comme cruciale afin de mieux comprendre les déterminants du mécanisme d'intégration dans la cellule.

4.1. Influence de la structure de la chromatine sur l'intégration

L'étude de l'influence de la chromatine sur l'intégration est basée sur des tests *in vitro* reproduisant l'intégration rétrovirale dans des substrats de type mononucléosome, polynucléosome et minichromosome, afin de mimer la chromatine cellulaire. Ces études ont confirmé que la structure chromatinienne influence l'intégration. En effet, les premières études de l'influence de la structure de la chromatine ont permis de montrer que l'intégration du VIH-1 et de MLV étaient favorisées au niveau du sillon majeur de la face externe de l'ADN courbé nucléosomal par rapport à de l'ADN nu (Pryciak & Varmus, 1992; Müller & Varmus, 1994; Pruss *et al*, 1994). De plus, il a été montré que la structure de la chromatine influence différemment les rétrovirus. En effet, l'intégration du VIH-1 est fortement inhibée sur un ADN compacté avec l'histone H1 alors que l'intégration du virus du sarcome aviaire (ASV) est stimulée sur ce substrat compacté (Taganov *et al*, 2004). Plus récemment, il a été montré que la présence de nucléosomes stablement positionnés sur une matrice artificielle de type polynucléosomale circulaire était une structure réfractaire à l'intégration concertée (Lesbats *et al*, 2011).

4.2. Impact de LEDGF/p75 sur l'intégration dans la chromatine

L'étude de LEDGF/p75 sur l'intégration du VIH-1 dans la chromatine a permis de montrer un rôle important des cofacteurs cellulaires de l'IN. En effet, il a été montré que LEDGF/p75 permet de stimuler l'intégration sur une structure polynucléosomale linéaire et que le domaine PWWP de LEDGF/p75 était nécessaire à cette stimulation *in vitro* (Botbol *et al*, 2008; Naughtin *et al*, 2015).

4.3. Impact des facteurs de remodelage ATP-dépendants

Le complexe SWI/SNF est un complexe de remodelage de la chromatine ATP-dépendant dont la sous-unité INI1 est décrite comme interagissant avec l'IN du VIH-1. Il a été montré que le traitement des polynucléosomes circulaires par le complexe SWI/SNF entraîne une restauration de l'intégration concertée sur des matrices initialement réfractaires. L'ensemble de ces données suggère un couplage fonctionnel entre l'intégration du VIH-1 et les mécanismes de remodelage de la chromatine. Cependant, il a été montré que le remodelage de la chromatine par la sous-unité SNF2H du complexe ACF/SNF2H

appartenant à la famille ISWI ne permet pas la restauration de l'intégration concertée sur matrice polynucléosomale circulaire. Il est possible qu'une interaction directe entre l'IN et SWI/SNF soit nécessaire pour une intégration optimale (Lesbats *et al*, 2011). L'ensemble de ces données est en accord avec la nécessité d'un couplage physique entre le complexe de remodelage de la chromatine et le complexe d'intégration.

5. Objectifs de l'étude

L'ensemble des travaux du laboratoire s'est focalisé ces dernières années sur l'analyse de l'influence de la chromatine sur l'intégration du VIH-1 ainsi que sur sa régulation par des facteurs de remodelage de la chromatine ATP-dépendants. L'ensemble de ces travaux a pu montrer que la chromatine joue un rôle important dans l'intégration du VIH-1. En effet, l'utilisation de substrat de type polynucléosome circulaire permettant le positionnement des nucléosome le long de l'ADN grâce à une répétition de séquence de type 5S-G5E4 reproduisant ainsi une chromatine compactée, affecte négativement l'intégration du VIH-1. De plus, ces travaux ont pu montrer pour la première fois que le couplage de l'intégration avec des complexes de remodelage ATP-dépendants de la famille SWI/SNF va permettre de restaurer une intégration sur ces substrats initialement réfractaires à l'intégration. Le laboratoire a aussi montré que l'utilisation de complexe de remodelage ATP-dépendant de la famille ISWI ne permet pas de restaurer l'intégration concertée sur ce type de substrat.

Afin de déterminer les facteurs régissant la modulation de l'intégration par la chromatine, nous nous sommes préalablement intéressés à l'étude de l'influence des structures nucléosomales sur différents systèmes rétroviraux de type alpharétrovirus (ASV), gammarétrovirus (MLV), lentivirus (VIH-1) et spumavirus (PFV) connus notamment pour leurs différences de préférence de site d'intégration dans la cellule afin de réaliser une étude comparative de ces différentes structures intasomales sur l'intégration dans la chromatine.

De plus, nous avons investigué le rôle d'autres facteurs associés à la chromatine. En effet, les études du complexe SWI/SNF montrent la nécessité d'un couplage fonctionnel et physique entre l'intasome du VIH-1 et ces complexes de remodelage. C'est pourquoi nous avons développé une stratégie visant à sélectionner et identifier des partenaires cellulaires interagissant avec l'intasome du VIH-1. Nous avons étudié le mécanisme d'action de

ces facteurs cellulaires associés à la chromatine tel que le complexe appartenant à la famille des histones chaperonnes, en particulier, le complexe FACT.

Finalement, nous avons développé un troisième axe de travail se focalisant sur le substrat final de l'intégration qu'est le nucléosome. C'est pourquoi, nous avons étudié l'interaction fonctionnelle entre une protéine virale et une modification post-traductionnelle d'histone retrouvée enrichie au niveau des sites d'intégrations du VIH-1 et associée à une transcription active. En effet, nous avons pour la première fois démontré l'importance d'un domaine de type Tudor dans le domaine C-terminale de l'intégrase permettant la reconnaissance spécifique de la lysine 21 mono-méthylée de l'histone H4.

Matériels et méthodes

1. Matériels

1. 1. Culture cellulaire

a) Cellules de mammifères

- HeLa P4

Les cellules HeLa P4 sont une lignée de cellules adhérentes et cancéreuses. Elles expriment de manière constitutive le récepteur CD4 au niveau de la membrane cytoplasmique et sont transformées avec le gène LacZ sous contrôle du promoteur LTR.

- HEK 293T

Les cellules HEK 293T sont une lignée de cellules épithéliales adhérentes issues de cellules embryonnaires de rein humain 293. Ces cellules possèdent l'antigène T du SV40.

- K562

Les cellules K562 sont une lignée cancéreuse établie à partir de cellules d'une leucémie myéloïde chronique prélevée chez une patiente de 53 ans (Lozzio & Lozzio, 1975). Ces cellules indifférenciées sont non adhérentes. De plus, les K562 se caractérisent par le chromosome de Philadelphie qui résulte de la translocation réciproque entre les chromosomes 9 et 22. Au cours de cette translocation le proto-oncogène c-Abl du chromosome 9 fusionne avec le gène BCR situé sur le chromosome 22. Le gène hybride qui résulte de cette translocation conduit à l'expression d'une protéine fusion ABL-BCR possédant une activité tyrosine kinase. Les cellules K562 sont un type cellulaire choisi par le projet ENCODE ayant pour objectif d'identifier l'ensemble des éléments fonctionnels du génome humain.

1. 2. Vecteurs lentiviraux pseudotypés VSV-G

Les vecteurs lentiviraux pseudo-typés VSV-G sont des virus intégratifs, mais non réplicatifs. Ces vecteurs possèdent les mêmes caractéristiques du VIH-1 au niveau de la séquence gag-pol, mais l'ARN génomique est remplacé par un ARN portant un gène rapporteur codant pour l'EGFP sous contrôle d'un promoteur interne CMV. L'enveloppe de ce virus est constituée de la protéine G du VSV (virus de la stomatite vésiculaire). La production de ces vecteurs

lentiviraux ainsi que l'ensemble des mutants utilisés dans ce travail sont assurés par la plateforme de vectorologie de Bordeaux 2.

1. 3. Plasmides

a) Plasmide d'expression bactérien

- Intégrase du VIH-1

La région du gène *pol* codant pour l'intégrase du VIH-1 se trouve sous le contrôle du promoteur T7 ARN polymérase dans le plasmide pET-21b-IN ayant pour marqueur de sélection l'ampicilline.

b) Plasmide d'expression pour la production des vecteurs lentiviraux pseudotypés VSV-G

La production des vecteurs lentiviraux pseudotypés VSV-G nécessite trois plasmides afin de permettre la production de particules infectieuses non répliquatives.

- Le plasmide de transfert permet la transcription d'un ARN mimant l'ARN génomique viral. Cet ARN contient les éléments essentiels pour l'encapsidation (Ψ), pour la transcription inverse (cPPT) et pour l'intégration (LTR). Une délétion au niveau de la région U3 conduit à la production de LTR mutés après la transcription inverse. De plus, cet ARN code pour la protéine EGFP sous la dépendance du promoteur CMV.
- Le plasmide d'empaquetage (pCMV Δ 8.91) permet l'expression des protéines structurales et de Tat et Rev, mais sans les protéines ENV et les protéines accessoires
- Le plasmide d'enveloppe codant pour la protéine G du VSV. Les virus ainsi produits sont alors pseudotypés VSV-G et possèdent un tropisme non spécifique afin d'infecter un large choix de cellules de mammifères.

1. 4. Protéines commerciales

Les protéines recombinantes SSRP1 et Spt16 sont exprimées par baculovirus dans des cellules d'insectes Sf9. Ces protéines contiennent une étiquette FLAG en N-terminal. Ces protéines de 80 kDa et de 140 kDa sont utilisées pour la reconstitution du complexe FACT *in vitro*.

Nom	Fournisseurs	Position tag	Tag
SSRP1	Active Motif	N-ter	FLAG
Spt16	Active Motif	N-ter	FLAG

1. 5 Les anticorps

Noms	Clone	Cibles	Fournisseurs	Espèces	Diilution
IN	Polyclonal	IN du VIH-1	BioProducts	Lapin	WB : 1/2000
Actine	Polyclonal	Actine	Sigma	Lapin	WB : 1/1000
SSRP1	Polyclonal	SSRP1	Abcam	Lapin	WB : 1/1000
Spt16	Polyclonal	Supt16h	Abcam	Lapin	WB : 1/1000
H4K20me1	Monoclonal	H4K20me1	Active Motif	Souris	WB : 1/1000
H4K20me3	Monoclonal	H4K20me3	Active Motif	Souris	WB : 1/1000
His-Tag	Monoclonal	His-Tag	Novagen	Souris	WB : 1/1000

1. 6 Oligonucléotides

Utilisation	Nom	Séquence
Sélection de partenaires	U5-Biot	5'-Biot-CATGTGGAAAATCTCTAGCA-3'
Sélection de partenaires	U3	5'-CAGTGAATTAGCCCCTTCCA-3'
Génération donneur ASV	ASV_U3_SupF5'	5'AATGTAGTCTTATGCAATACTCTTGTAGTCTT GCAATTAACGTTGCCCGGATCCGGTCGCGC 3'
Génération donneur ASV	ASV_U3_SupF3'	5'AATGTAGTCTTATGCAATACTCTTGTAGTCTT GCAAGCGGCGCGTCATTTGATATGATGCG 3'
Génération donneur PFV	PFV-SupF5'	5'ATTGTCATGGAATTTTGTATATTGATTATCCTT TAA CGTTGCCCGGATCCGGTCGCGC 3'
Génération donneur PFV	PFVSupF3'	5' ATTGTCATGGAATTTTGTATATTGATTATCC TGCGGCGCGTCATTTGATATGATGCG 3'
Génération donneur MLV	MLV SupF5'	5'TATGAAAGACCCACCTGTAGTTAACGTTGC CCGGATCCGGTCGCGC 3'
Génération donneur MLV	MLVSupF3'	5'TATGAAAGACCCCGCTGACGCGGCGCGTC ATTTGATATGATGCG 3'
Génération donneur VIH	HIV-U5	5'AGGATAATCATGTGGAAAATCTCTAGCA 3'
Génération donneur VIH	HIV-U3	5'AGGATAATCAGTGAATTAGCCCCTTCCA3'

1. 7. siRNA

Utilisation	Nom	Fournisseurs
siRNA-SSRP1	SSRP1	Santa Cruz Biotech
siRNA Contrôle	Contrôle	Santa Cruz Biotech

1. 8. Peptides

Nom	Fournisseurs	Séquence
H4K20me0	ANASPEC	H-S-G-R-G-K-G-G-K-G-L-G-K-G-G-A-K-R-H-R-K-V-L-R-G-G-K(biotin)
H4K20me1	ANASPEC	H-S-G-R-G-K-G-G-K-G-L-G-K-G-G-A-K-R-H-R-K(me1)-V-L-R-G-G-K(biotin)
H4K20me2	ANASPEC	H-S-G-R-G-K-G-G-K-G-L-G-K-G-G-A-K-R-H-R-K(me2)-V-L-R-G-G-K(biotin)
H4K20me3	ANASPEC	H-S-G-R-G-K-G-G-K-G-L-G-K-G-G-A-K-R-H-R-K(me3)-V-L-R-G-G-K(biotin)
H3K36me3	ANASPEC	H-A-T-K-A-A-R-K-S-A-P-A-T-G-G-V-K(me3)-K-P-H-R-Y-R-P-G-G-K(biotin)
H2A	ANASPEC	H-S-G-R-G-K-Q-G-G-K-A-R-A-K-A-K-T-R-S-S-R-G-G-K(biotin)
H2BK5me1	ANASPEC	H-P-E-P-A-K(me1)-S-A-P-A-P-K-K-G-S-K-K-A-V-T-K-A-Q-G-G-K(biotin)
H4K20me1-FITC	Plateforme IGBMC Strasbourg	FITC-K-G-G-R-H-R-K(me1)-V-L-R
H4K20me2-FITC	Plateforme IGBMC Strasbourg	FITC-K-G-G-R-H-R-K(me2)-V-L-R
H4K20me3-FITC	Plateforme IGBMC Strasbourg	FITC-K-G-G-R-H-R-K(me3)-V-L-R
H4K20me0-FITC	Plateforme IGBMC Strasbourg	FITC-K-G-G-R-H-R-K-V-L-R
H4K20me0	Plateforme IGBMC Strasbourg	K-G-G-R-H-R-K-V-L-R

1. 9. Nucléosome

Nom	Fournisseurs	Remarque
Mononucléosomes humain biotinylés	EPICYHER	Mononucléosomes assemblés à partir d'histones recombinantes humaines exprimées chez <i>E.coli</i> entourées par un ADN de 147 pb de séquence 601. Un groupement biotine-TEG est présent en 5' de l'ADN
Séquence d'assemblage 601-Biotinylée	EPICYHER	Séquence d'ADN 601 de 147 pb biotinylé en 5'
Mononucléosomes humains natifs	Active motif	Mononucléosomes natifs comportant un ensemble de modifications des histones

2. Méthodes

2. 1. Cultures bactériennes

a) Culture

Les cultures bactériennes des souches DH5 et Rosetta sont cultivées à 37 °C sous agitation pour les cultures en milieu liquide (220 rpm) dans un milieu LB (Luria-Bertani : 1 % bacto tryptone, 0,5 % yeast extract, 0,5 % NaCl et 1,5 % agar pour le milieu solide). Pour les milieux de sélection : 50 µg/ml ampicilline et 25 µg/ml chloramphénicol.

b) Transformation

Les transformations bactériennes sont réalisées par l'utilisation de bactéries chimio-compétentes. Cette transformation est basée sur un choc thermique de ces bactéries en présence de l'ADN plasmidique. 200 µl de bactéries sont incubées à 4 °C pendant 30 min avec 100 ng d'ADN plasmidique. Puis, le mix est incubé à 42 °C pendant 1 min 30 s. Finalement, 800 µL de milieu LB sans antibiotique est ajouté afin de permettre l'expression phénotypique et le mélange est incubé 1 h à 37 °C sous agitation (220 rpm). La sélection se

fait ensuite sur milieu solide LB agar supplémenté avec l'antibiotique correspondant à 37 °C durant la nuit.

c) Extraction et purification de l'ADN plasmidique

Pour l'extraction et la purification de l'ADN plasmidique, différentes quantités de culture de bactéries sont utilisées : 10 ml pour les minipreps et 500 ml pour les maxi-preps. Les extractions et purification sont ensuite réalisées par les kit QIAprep Spin Miniprep (QIAGEN) et Nucleobon Xtra Maxi Plus EF (Machery-Nagel).

2. 2. Cultures de cellules mammifères

Les cellules HeLa-P4, HEK 293T et K562 sont cultivées dans du milieu Dulbecco's modified minimal essential medium (DMEM, Invitrogen) supplémenté avec 10 % de sérum de veau fœtal (SVF, Invitrogen) et 50 mg/ml de gentamicine. De plus, 1 mg/ml de G418 (Invitrogen) est ajouté au milieu de culture des cellules HeLa-P4 afin de maintenir l'expression du récepteur CD4 des cellules HeLa-P4. L'ensemble des cultures cellulaires est réalisé à 37 °C en présence de 5 % de CO₂. A confluence des cellules, une solution de trypsine-EDTA (Invitrogen) est ajoutée afin de décoller les cellules de leur surface puis les cellules sont diluées au 1/10 dans un milieu DMEM complet.

2. 3. Infection des cellules par les vecteurs lentiviraux

a) Production des vecteurs lentiviraux pseudotypés VSV-G

L'ensemble des vecteurs lentiviraux pseudotypés VSV-G est réalisé par la plateforme de vectorologie de Bordeaux. Pour cela, les cellules 293T sont mises en culture dans une boîte de pétri de 10 cm (6 à 7 10⁶ cellules) dans 10 ml de DMEM complet. Après 24 heures, les cellules sont transfectées avec les plasmides d'empaquetage, d'enveloppe et de transfert décrits précédemment. Pour cela l'ADN plasmidique (10 µg d'ADN plasmidique dans 450 µl d'eau) et du CaCl₂ (50 µl de solution à 2,5 M) sont filtrés (0,22 µm) puis versés sur 500 µl de HBS 2x (280 mM NaCl, 1,5 mM Na₂HPO₄-7H₂O, 100 mM HEPES pH 7,2). Puis, après trois minutes, le mélange est mis en contact des cellules. Quatre heures post-transfection, le milieu est changé par du DMEM complet et mis à incubation. 24 à 48 h après, le surnageant de culture est récupéré et centrifugé 5 min à 2500 rpm puis filtré (0,22 µm) puis ultra-centrifugé à 20 000 rpm pendant 2 heures. Le culot est ensuite repris dans 1 ml de DMEM.

b) Titration et dosage

La titration des vecteurs lentiviraux ainsi que l'ensemble des mutants est réalisée par la plateforme de Vectorologie de Bordeaux. Afin de déterminer le titre infectieux, des cellules 293T sont infectées avec des dilutions sériées de vecteurs lentiviraux. Puis, la fluorescence de la protéine EGFP est mesurée par cytométrie de flux. Le titre est ensuite calculé en réalisant une moyenne du pourcentage de fluorescence obtenu pour chaque dilution rapportée au nombre de cellules infectées pour une transduction inférieure à 20 %. La quantité de p24 présente dans le surnageant viral est dosée par méthode ELISA (InGen).

c) Traitement DNase

Afin d'empêcher la contamination du surnageant contenant les particules infectieuses par les plasmides utilisés pour la production notamment le plasmide de transfert codant pour l'EGFP, les surnageants sont traités 1 heure à 37 °C avec 40 U/ml de RNase free DNase (Promega) puis stockés à -80 °C.

d) Transduction

Les transductions des cellules 293T et HeLa-P4 sont réalisées en laboratoire de confinement de type P2. Pour cela, les cellules sont mises en culture à 50 000 cellules/puits dans des plaques 48 puits dans un milieu DMEM complet. 48 heures après, les cellules sont transduites avec différentes MOI de vecteurs lentiviraux pseudotypés VSV-G dans un volume final de 200 µL de DMEM par puits. Enfin, après 2 heures 200 µL de DMEM est ajouté dans chaque puit. 48 après, les cellules sont lavées en PBS et les cellules sont ensuite décollées dans un mélange PBS/Trypsine (50 %/50 %) et ensuite centrifugées et repercent dans un milieu PBS 0,2 mM EDTA et 0,5 % SVF pour être analysées par cytométrie de flux (BD FACS Canto).

2. 4. Double transfection par siRNA-SSRP1

Les cellules 293T et HeLa-P4 sont mises en culture dans des plaques 48 puits à 50 000 cellules/ml. A 60 % de confluence, les cellules sont transfectées par une gamme croissante de siRNA-SSRP1 ou du siRNA contrôle (de 1 nM à 40 nM) à l'aide de l'agent de transfection INTERFERin (Polyplus). Pour cela, les siRNA sont dilués dans 50 µl final de DMEM. Puis, l'agent de transfection est dilué dans le mélange à raison de 1,5 µl dans les 50 µL final. Le

mélange est ensuite vortexé 10 sec, centrifugé et incubé 10 min à température ambiante afin de permettre la formation de complexe siRNA-agent de transfection. Puis, après incubation, 250 µl de DMEM complet est ajouté au 50 µl du mélange siRNA-agent de transfection. Les 300 µl sont ensuite disposés sur les cellules précédemment rincées au PBS et les cellules sont ensuite incubées à 37 °C en présence de 5 % de CO₂. 24 heures après, une nouvelle transfection est réalisée en utilisant le même protocole.

a) Extraction des protéines cellulaires

Les cellules sont dans un premier temps lavées en PBS puis sont incubées dans 50 µl de tampon de lyse : ½ tablette inhibiteur de protéase, 1 mM PMSF, RIPA (Sigma) à 4 °C pendant 30 minutes. Les cellules sont ensuite décollées puis centrifugées 10 min à 8000 g à 4 °C. Le surnageant protéique est ensuite conservé à -20 °C.

2. 5. Manipulation des acides nucléiques

a) Extraction des acides nucléiques de cellules infectées

Suite à la transduction, l'ADN total contenu dans les cellules de mammifères est extrait à l'aide du kit QIamp DNA Blood Mini kit (QIAGEN) selon le protocole du fournisseur.

b) Mutagenèses dirigées

La mutagenèse dirigée et le séquençage des plasmides pET21b-IN et PCMVΔ8.91 ont été effectués par BioInnovatise Inc.

2. 6. Préparation des cellules pour le séquençage haut débit des sites d'intégration

Le protocole de préparation des échantillons pour le séquençage haut débit a été mis au point à partir de l'article du Dr Angela Ciuffi (Ciuffi & Barr, 2011). L'ADN génomique de cellules K562 infectées est extrait à l'aide du kit commercial QIamp DNA Blood Mini kit et la qualité de l'extraction est vérifiée sur gel d'agarose.

2. 7. Manipulation des protéines

a) Production de l'IN recombinante

Afin de réaliser la purification d'IN recombinante WT ainsi que l'ensemble des IN mutantes, une préculture est réalisée par ensemencement d'une colonie de Rosetta transformée avec le pET-21b-IN dans 20 ml de LB + ampicilline (100 µg/ml) + chloramphénicol (34 µg/ml) durant la nuit à 37 °C sous agitation (220 rpm). Puis, la préculture est ensemencée dans 500 ml de LB complémenté avec de l'ampicilline et du chloramphénicol jusqu'à l'obtention d'une DO à 600nm de 0,6. L'induction de la production intégrase se fait par l'ajout d'IPTG (500 µM). La culture est ensuite incubée 3 heures à 37 °C sous agitation (220 rpm). Les bactéries sont ensuite centrifugées 10 min à 4000 rpm à 20 °C et les culots protéiques sont congelés à -80 °C.

Tampon de lyse : 50 mM HEPES pH 7,5, 5 mM EDTA pH 8, 1 mM DTT, 1 mM PMSF, inhibiteurs de protéases (Roche)

Tampon de solubilisation : 50 mM HEPES pH 7,5, 1 M NaCl, 1mM EDTA, 1mM DTT, 7 mM CHAPS

Tampon LCS (Low salt concentration) : 50 mM HEPES pH 7,5, 0,2 M NaCl, 0,1 mM EDTA, 1mM DTT, 7 mM CHAPS, 10 % glycérol.

Tampon HCS-Butyl (High salt concentration) : 50 mM HEPES pH 7,5, 0,2 M NaCl, 0,1 mM EDTA, 1 mM DTT, 7 mM CHAPS, 1M Sulfate d'ammonium, 10 % glycérol.

Tampon HCS-Héparine : 50 mM HEPES pH 7,5, 1 M NaCl, 0,1 mM EDTA, 1 mM DTT, 7 mM CHAPS, 10 % glycérol.

Tampon de dilution : 50 mM HEPES pH 7,5, 0,1 mM EDTA, 1mM DTT, 7 mM CHAPS, 10 % glycérol

Les culots protéiques sont repris dans 20 ml de tampon de lyse complémenté avec du lysozyme (20 mg/ml) et de la DNase (15 µl/ml) (RQI Rnase free Promega). Après sonication à 60 W pendant 20 secondes, une centrifugation à 10 000 rpm pendant 10 minutes est réalisée. Le surnageant constituant la fraction protéique soluble et le culot est ensuite lavé dans 20 ml de tampon de lyse puis est centrifugé. Le culot est ensuite repris dans 13 ml de tampon de solubilisation pendant 1 heure à 4 °C sous agitation. Puis, le culot solubilisé est

centrifugé 1 h à 33 000 rpm. Le surnageant est ensuite supplémenté avec du sulfate d'ammonium à une concentration final de 1M et agité 5 min à 4 °C. L'ensemble est ensuite centrifugé 20 min à 10 000 rpm à 4 °C et le surnageant obtenu est alors chargé sur colonne HiTrap Butyl Sépharose 4B (1 ml, GE Healthcare) préalablement lavée 5 volumes de LSC puis équilibrée avec 5 volumes de HSC. L'élution se fait par un step de NaCl (1 à 0 M). La densité optique des fractions éluées est lue à 280 nm. Les fractions contenant l'IN sont regroupées et diluées au 1/3,3 dans le tampon de dilution. La deuxième colonne est une HiTrap Héparine Sépharose CL-4B (1 ml, GE Healthcare). Cette colonne est premièrement lavée avec 5 volumes de tampon HS puis équilibrée avec 5 volumes de tampon LCS. L'ensemble des fractions diluées est ensuite chargé sur la colonne et les protéines sont éluées par choc ionique de NaCl. Les fractions contenant l'IN sont diluées 4 fois dans du tampon de dilution puis rechargées sur une seconde colonne Héparine Sépharose CL-4B. L'élution de cette dernière colonne se fait par gradient de NaCl (0,2 à 1 M). L'ensemble des fractions sont analysées par SDS-PAGE 12 % suivie d'une coloration au bleu de Coomassie et la concentration est déterminée au Nanodrop.

2. 8. Tests *in vitro*

a) Test d'intégration concertée

Le test d'intégration concertée nécessite la pré-incubation pendant 30 minutes sur glace de 100 nM d'IN recombinante en présence de 10 ng d'ADN donneur mimant l'ADN viral, 50 g d'ADN receveur mimant l'ADN cellulaire. Puis la réaction est lancée par l'ajout du mix réactionnel (20mM HEPES pH 7, 10 mM DTT, 10 mM MgCl₂, 15 % DMSO, 8 %PEG 6000, 20 µM ZnCl₂, 100 mM NaCl) pendant 2 heures à 37 °C. La réaction est ensuite arrêtée par l'ajout d'une solution de protéinase K (0,8 mg/ml) pendant 1 h à 55 °C. Les produits d'intégration sont ensuite extraits par traitement phénol-chloroforme-alcool isoamylique (24/25/1 v/v/v) et déposés sur gel d'agarose 1 % pour une migration 200V. Le gel est ensuite traité au TCA 5 % pendant 15 minutes et séché 1 h à 80 °C sous vide avant autoradiographie. L'ensemble des produits d'intégration sont ensuite quantifiés par ImageJ. Les tests d'intégration sur mononucléosomes recombinants 601 ou sur ADN 601 ont été effectués en utilisant la même procédure cependant, la concentration d'IN utilisée est de 400 nM final et la quantité de mononucléosome est de 250 ng par test et 125 ng d'ADN receveur. Les produits d'intégration sont ensuite précipités à l'éthanol et chargés sur un gel d'acrylamide 6 %. Le gel est ensuite séché 1 h à 80 °C sous vide avant autoradiographie. L'ensemble des produits est ensuite quantifié par ImageJ.

b) Histone peptide array

L'expérience de histone peptide array (Active Motif) a été effectuée selon le protocole du fournisseur à une concentration d'IN de 40 nM dans un tampon d'interaction : 20mM HEPES pH 7.5, 20 mM DTT, 15 % DMSO, 16 % PEG, et 10 mM MgCl₂ pendant 2 heures à 37 °C. La détection de l'IN se fait par l'ajout d'anticorps anti-IN du VIH-1 (1/3000) dilués dans une solution de PBS BSA 0,3 % pendant la nuit à 4 °C. Après 3 lavages au PBS, l'anticorps secondaire anti- lapin couplé à la peroxydase (1/3000) dilué dans une solution de PBS BSA 0,3 % est ajouté sur la membrane pendant 1 heure. L'histone peptide array est ensuite lavé 3 fois dans du PBS 1X puis une solution ECL est ajoutée et la visualisation se fait grâce au système LAS4000. La quantification est réalisée via le logiciel Array Analyze (Active Motif).

c) FAIRE “Formaldehyde-Assisted Isolation of Regulatory Elements”

In cellulo, des cellules 293T ou HeLa P4 traitées par la molécule CBL0137 ou par siRNA SSRP1 sont mises en cultures jusqu'à 90 % de confluence. Puis, une solution de formaldéhyde est ajoutée à une concentration finale de 1 % à température ambiante et incubée 1, 2, 4 ou 7 minutes avec les cellules. Puis, une solution stop (125 mM glycine) est ajoutée pendant 5 min à température ambiante. Les cellules sont ensuite lavées 3 fois dans une solution PBS+PMSF. Les cellules sont ensuite décollées et centrifugées à 2000 rpm durant 4 minutes et congelées rapidement dans l'azote liquide. Les cellules sont ensuite reprises dans un 1ml de tampon de lyse (2 % Triton X-100, 1 % SDS, 100 mM NaCl, 10 mM Tris-HCl at pH 8.0, 1 mM EDTA) pour 0,4g de cellules et sont ensuite lysées par l'utilisation de billes de verre durant 5 fois 1 min à 4°C. Puis, après sonication des échantillons (5 sessions de 6 pulsations de 1 sec on/1s ec off), les échantillons sont centrifugés à 15000 rcf pendant 5 min à 4°. L'ADN est ensuite extrait des débris cellulaires par un traitement phénol/chloroforme/alcool isoamylique 25/24/1 et centrifugation à 15000 rpm pendant 5 min à 4°C. La phase aqueuse est ensuite prélevée et stockée. Finalement, 500 µl de solution TE est ajoutée à la phase organique et le mélange est vortexé et centrifugé à 15000 rpm 5 min à 4°C. La phase aqueuse est isolée et ajoutée à la première et une extraction phénol/chloroforme est effectuée sur la fraction contenant le mélange des phases aqueuses afin d'éliminer les protéines. L'ADN est ensuite précipité par l'ajout d'une solution d'acétate de sodium 0.3M, glycogène 20 µg/mL et 2 volumes d'éthanol 95 % et le mélange est incubé à -20°C sur la nuit puis centrifugé à 15000 rpm pendant 10 min à 4°C et le culot est ensuite lavé avec de l'éthanol 70 % et séché au Speed-Vac. Le culot est ensuite repris dans une solution

d'eau distillée traitée à la Rnase A (100 µg/mL) et incubé à 37°C pendant 2 heures. La concentration de l'ADN est ensuite évaluée au nanodrop.

In vitro, 100 ng de matrices polynucléosomales sont traitées avec le complexe FACT et incubées 30 minutes dans une solution 2 % Formaldéhyde. Après sonication de l'ADN 3x30sec, une extraction par l'ajout de 50 µl de solution phénol/chloroforme est effectuée et le mélange est centrifugé à 13000 g 10 minutes. La concentration de l'ADN est ensuite évaluée au nanodrop.

d) Expériences de coprécipitation

12,5 µl de billes magnétiques Dynabeads MyOne Streptavidin (Invitrogen) sont lavées dans 287,5 µl de solution de coprécipitation (50 mM HEPES, 0,1 mg/ml BSA, 1 mM DTT, 0,1 % Tween 20, 10 % glycérol, 50 mM NaCl) pendant 10 min à température ambiante sous agitation. Les billes sont ensuite magnétisées et l'opération de lavage est répétée. Puis, 10 pmoles d'IN sont incubées avec 1 pmole de nucléosomes humains biotinylés (Epiccypher) dans un volume final de 10 µl de solution de co-précipitation. Le mélange est ensuite incubé 20 minutes à 4 °C puis 20 minutes à température ambiante. Les billes magnétiques préalablement rincées sont reprises dans 287,5 µl de solution de coprécipitation. Puis, le mélange IN/Nucléosome est incubé 1 heure à température ambiante sous agitation avec les billes. Le mélange est ensuite magnétisé et le culot est rincé trois fois dans 300 µl de solution de coprécipitation. Finalement, le culot est repris dans une solution de LAEMLI et chargé sur gel SDS-PAGE 12 %.

e) Microscale thermophoresis

Des dilutions en série du domaine CTD de l'IN du VIH-1 de 185 µM à 5 nM sont effectuées. Puis, le peptide d'intérêt est ajouté au CTD de l'IN du VIH-1 à une concentration final de 5 nM. Après mélange, l'échantillon est incubé 15 min à température ambiante. L'ensemble des mesures est effectué par le Nanotemper Monolith NT.015 avec un temps de « laser-on » de 30 sec et un temps de « laser-off » de 5 sec dans les conditions de 20 % LED et 40 % MST IR-Laser power. Pour les expériences de compétition, 0,5 µM de peptide (H4K20me0) non marqué est ajouté à l'échantillon. Puis, 0,5 µM de peptide fluorescent monométhylé est ajouté et la solution est incubée 15 min à température ambiante.

f) Far dot blot

Des dilutions en série de l'intégrase du VIH-1 de 1 pmol à 20 pmoles sont effectuées dans un tampon d'interaction (20 mM HEPES, pH 7,5, 100 mM NaCl, 1 mM EDTA, 1 mM DTT). Une membrane de nitrocellulose est ensuite incubée 10 min dans le tampon d'interaction. Puis, trois morceaux de papier Whatman sont humidifiés avec le tampon d'interaction et placés dans l'appareil à dot. La membrane de nitrocellulose est ensuite posée au-dessus des papiers Whatman. L'appareil est ensuite fermé et les puits sont rincés trois fois avec 300 µl de tampon d'interaction grâce à l'aspiration sous vide. Puis, 200 µL de chaque échantillon est disposé dans les puits et aspiré. Les puits sont ensuite rincés trois fois avec 300 µl de tampon d'interaction et la membrane est incubée dans 4 ml de solution de saturation (tampon d'interaction + BSA 1 %) sous agitation à 4 °C sur la nuit. La membrane est ensuite rincée deux fois durant 5 min avec le tampon d'interaction. Le peptide est ensuite ajouté à une concentration de 1 µM pendant 1 heure. La membrane est ensuite lavée trois fois 10 min avec le tampon d'interaction. Suite à cela, la membrane est incubée une heure avec une solution d'interaction + BSA 0,3 % + ExtrAvidin-Peroxydase (Sigma Aldrich) au 1/4000^e et la membrane est lavée trois fois 10 min avec 4 ml de tampon d'interaction. Puis une solution ECL est ajoutée et la visualisation se fait grâce au système LAS4000.

g) Sélection des cofacteurs cellulaires du complexe IN-ADNv

Des fractions enrichies en complexes nucléoprotéiques biotinylés IN-ADNv sont générées par l'incubation de l'IN recombinante avec un ODN, comportant les 21 derniers nucléotides de l'extrémité virale U5, couplé à la biotine dans des conditions mises au point au préalable (Lesbats *et al*, 2008). Les complexes sont ensuite incubés avec les extraits protéiques cellulaires issus de cellules HeLa P4 après contre sélection de ces protéines sur des billes magnétiques couplées à l'ADN seul. Les protéines cellulaires sont ensuite sélectionnées après 1h d'incubation à 37°C avec le complexe IN-ADNv fixé aux billes magnétiques couplées à la streptavidine dans un tampon d'interaction (50 mM HEPES, pH7.5 ; 1µg/ml BSA;1 mM DTT; 0.1 % Tween 20 ; 10 % glycérol ; 100 mM NaCl). Après trois lavages avec le tampon d'interaction, les protéines sont éluées par l'ajout du tampon de charge LAEMLI et chauffées à 95°C. Les protéines sont ensuite chargées sur gel SDS-PAGE et les bandes correspondantes aux protéines sélectionnées sont analysées par spectrométrie de masse MS/MS.

2. 8. Tests *in silico*

a) Docking moléculaire

Dans l'ensemble des expériences de docking moléculaire, le domaine CTD de l'IN du VIH-1 correspondant au domaine 210-270 issu de la structure cristallographique (PDB 1EX4) a été utilisé comme récepteur. Le pentapeptide H4K20me1 dérivé de la structure du complexe MSL3/H4K20me1 (PDB 3OA6) a été comme ligand. Puis, le docking du récepteur et du ligand a été analysé avec AutoDockTools 1.5.6. Dans les premières expériences de docking utilisant AutoDock 4.2.6 et AutoDock Vina 1.1.2, la surface du récepteur a été traitée comme rigide.

Résultats

Virus	Taille de la duplication	Sélectivité	Cofacteur cellulaire de ciblage
MLV	4pb	Sites d'initiation de la transcription	Protéines BET
PFV			
HIV	5pb	Gènes	LEDGF/p75
ASLV	6pb	Pas de préférence marquée	?
MMTV			
HLTV			

Figure 49 : Caractéristiques génomiques des sites d'intégration rétroviraux

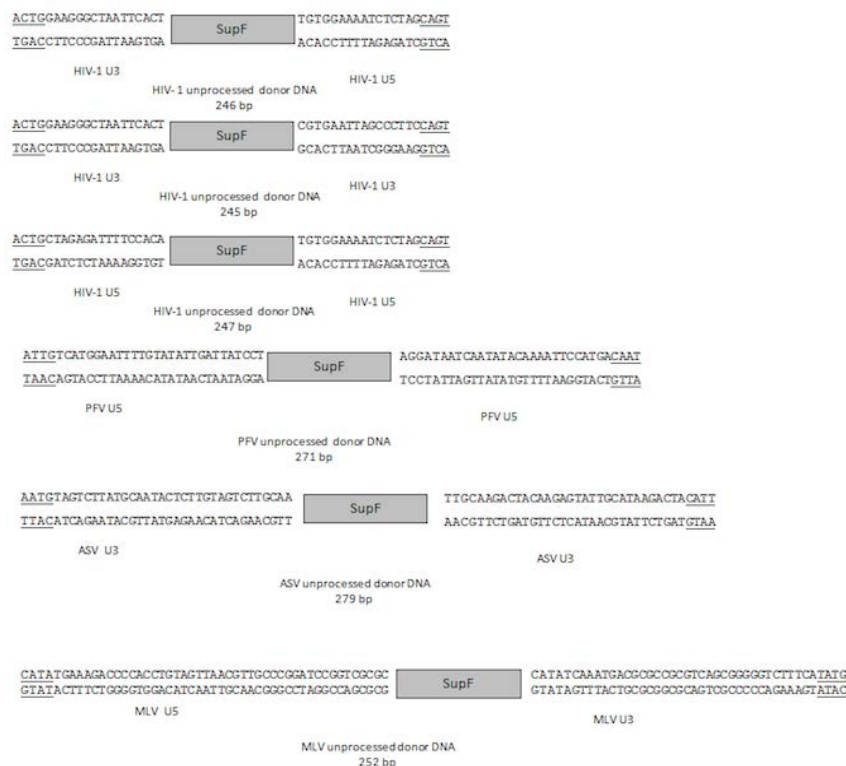


Figure 50 : Structure des ADN donneurs viraux

Les ADN donneurs contiennent le gène supF nécessaire au clonage des intégrants et contiennent les extrémités U3 et U5 spécifiques de VIH-1, PFV, ASV et MLV

Partie 1 : Impact de la structure chromatinienne sur l'intégration rétrovirale

Ce travail a mené à une publication incluant une partie de mon travail de thèse : “ *Intasome architecture and chromatin density modulate retroviral integration into nucleosome*”

Mohamed Salah Benleulmi, Julien Matysiak *et al.* *Retrovirology*

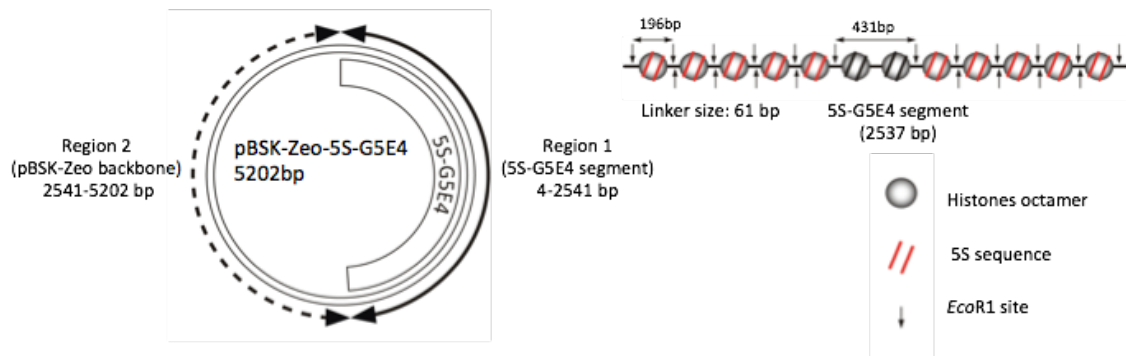
2015 12 : 13

1. Objectifs

À la différence de la situation *in vitro*, dans la cellule infectée, l'ADN cible n'est pas nu, mais compacté et protégé par son enroulement autour de protéines histones. Ce premier degré de compaction forme la chromatine qui est constituée d'une succession de nucléosomes structurés en nucléofilament ou « collier de perles » plus ou moins compact selon les régions du génome (voir chapitre 3.). Très tôt, il a été montré que la structure de la chromatine, la courbure de l'ADN et la présence de mononucléosomes pouvaient influencer l'intégration rétrovirale *in vitro* (Pruss *et al*, 1994 ; Katz *et al*, 1998 ; Pryciak *et al*, 1992). Plus récemment, des études biochimiques menées au laboratoire utilisant des modèles reproduisant la chromatine ont montré que la compaction des nucléosomes pouvait influencer négativement l'intégration du VIH-1 (Lesbats *et al*, 2011). Ainsi l'ensemble de ces résultats suggère que le nucléosome s'avère être une structure favorable à l'intégration rétrovirale et que l'environnement chromatinien permet sa régulation. Par ailleurs, dans la cellule infectée, les différents rétrovirus montrent des préférences distinctes pour certaines régions de la chromatine *via* l'interaction entre les intasomes et des facteurs de ciblage cellulaire (**Figure 49**). Ainsi, l'intégration rétrovirale apparaît être un processus multifactoriel comportant des composants à la fois viraux et cellulaires. Afin de comprendre ce processus complexe de régulation de l'intégration rétrovirale par la chromatine, le laboratoire a choisi d'entreprendre une étude comparative *in vitro* et *in vivo* de l'intégration de différents rétrovirus.

Dans cet objectif, plusieurs rétrovirus ayant une sélectivité différente dans la cellule infectée ont été sélectionnés, car ils sont susceptibles de présenter des processus de régulation distincts. Pour cela, nous avons sélectionné des systèmes rétroviraux de type alpharétroviral (ASV) caractérisé par une faible préférence d'insertion dans les unités de

a



b

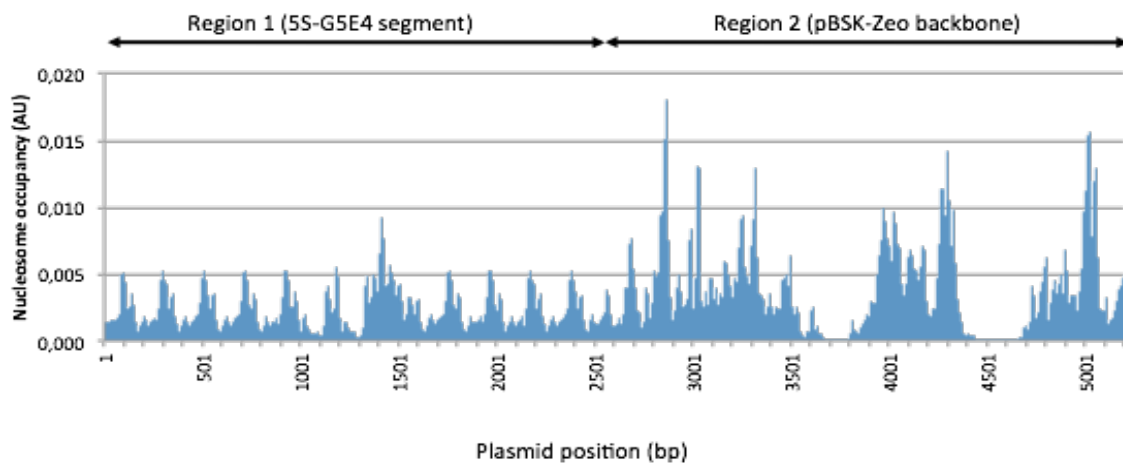


Figure 51 : Représentation du plasmide pBSK-Zeo-G5E4 (a) Détail de la séquence positionnant les nucléosomes 5S-G5E4 clonée dans le plasmide pBSK-zeo. Les flèches représentent les sites de restriction EcoRI. (b) Positionnement des nucléosomes sur le vecteur pBSK-Zeo-G5E4 suivant un algorithme de prédiction. Deux régions sont décrites, la région I contenant les nucléosomes stables et régulièrement associés ainsi que la région II aux nucléosomes plus labiles et dynamiques.

transcription, gammarétroviral (MLV) ayant une préférence pour les sites d'initiation de la transcription des gènes actifs, lentiviral (VIH-1) s'intégrant au sein des unités de transcription active et spumaviral (PFV) caractérisé par une intégration dans les sites d'initiation de la transcription de gènes peu transcrits afin de déterminer la sensibilité de chacun de ces intasomes rétroviraux pour un ADN nu et nucléosomal. De manière intéressante, la comparaison préalable de ces différents modèles dans la cellule infectée montre une corrélation entre le décalage de paire de bases de l'ADN cible au site d'intégration et la sélectivité de ces rétrovirus. Étant donné que ce décalage de paire de bases est directement lié à la structure de l'intasome, cette étude comparative permettrait de déterminer l'influence de ces différentes structures intasomales sur la sensibilité de l'intégration dans la chromatine.

2. Effet de la chromatine sur l'intégration de différents modèles rétroviraux

2.1. Intégration *in vitro* de différents systèmes rétroviraux

Dans l'objectif de pouvoir comparer l'influence de la structure chromatinienne des quatre modèles rétroviraux sélectionnés ASV, VIH-1, MLV et PFV en intégration concertée, des conditions d'intégration *in vitro* spécifiques ont été mises au point afin de permettre une étude comparative de ces différents systèmes viraux en conditions optimales. Cette mise au point préalable à mon travail a constitué une partie du travail de Mohamed Benleulmi, chercheur Postdoctoral ANR au laboratoire. En résumé, des ADN donneurs mimant les ADN de chacun de ces rétrovirus ont été générés par PCR avec des amorces contenant les 21 derniers nucléotides spécifiques des extrémités U5 et U3 à partir d'une matrice codante pour le gène *supF* (**Figure 50**). L'ADN receveur utilisé dans ce test est le pBSK-Zeo-5S-G5E4 contenant la séquence 5S-G5E4 insérée dans le plasmide pBSK-Zeo (**Figure 51a**). Cette séquence 5S-G5E4 contient des répétitions de 2x5 séquences « 5S » connues pour permettre le positionnement stable des nucléosomes. Ces séquences permettront ainsi l'assemblage *in vitro* d'un ADN polynucléosomal circulaire (PN) plus ou moins compact selon le ratio Histone/ADN utilisé.

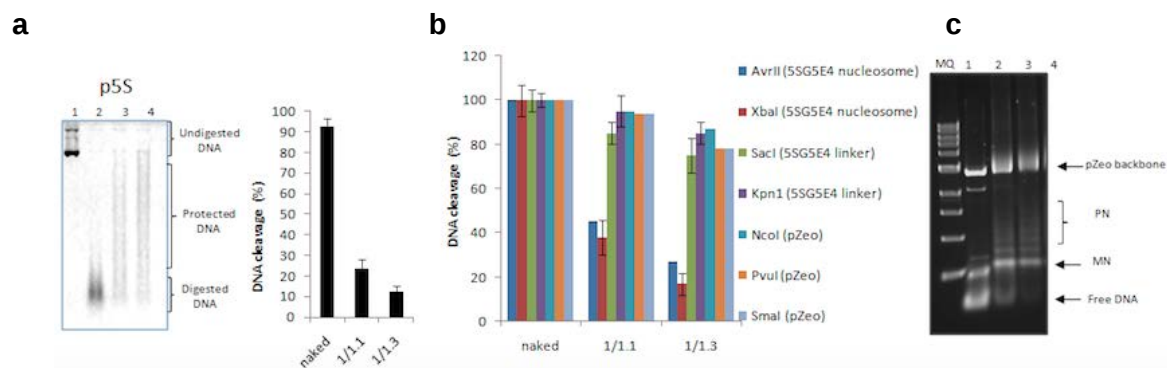


Figure 52 : Analyse biochimique du plasmide pBSK-Zeo-G5E4

(a) Analyse du pBSK-Zeo-G5E4 par DNase1. Piste 1 : p5S non digéré, piste 2 : p5S nu, piste 3 : p5S chromatinisé 1/1.1, piste 4 : p5S 1/1.3. (b) Analyse du pBSK-Zeo-G5E4 par technique REA et représentation en histogramme du pourcentage d'ADN clivé (c) Analyse du pBSK-Zeo-G5E4 par *EcoR1* et gel retard. Piste 1 : *EcoR1* + p5S nu, piste 2 : *EcoR1* + p5S chromatinisé 1/1.1, piste 3 : *EcoR1* + p5S 1/1.3.

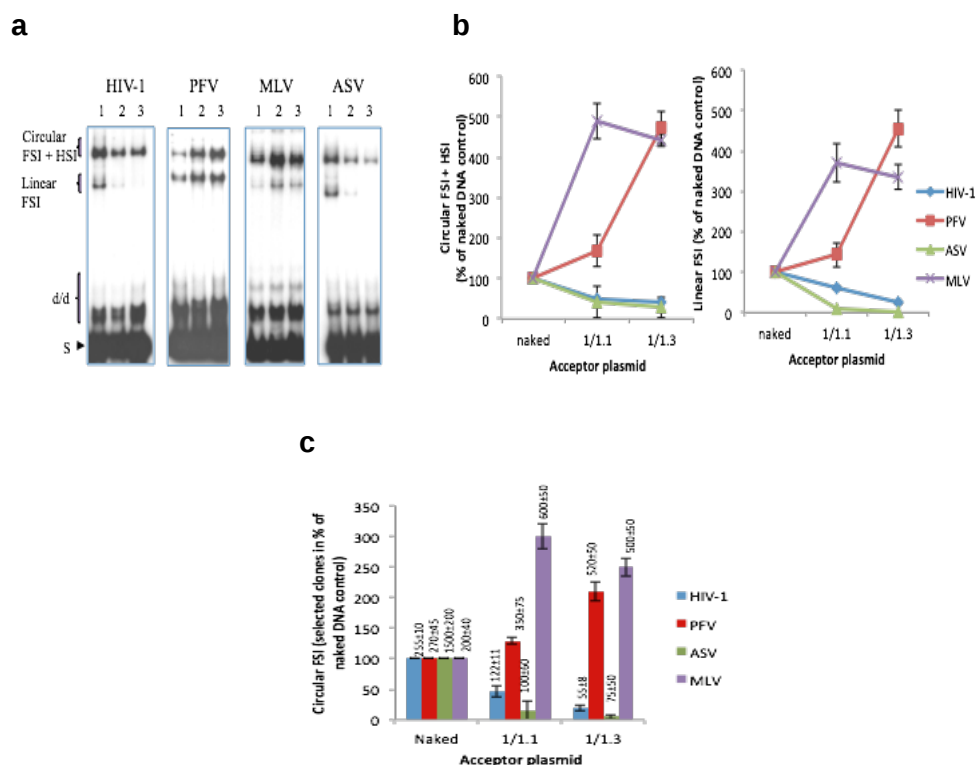


Figure 53 : Intégration concertée de VIH-1, PFV, MLV et ASV dans le receveur pBSK-Zeo 5S-G5E4 chromatinisé.

(a) Tests d'intégration concertée sur plasmide nu (ligne1) et sur plasmide polynucélosomale avec deux ratios massiques ADN/histones ($\mu\text{g}/\mu\text{g}$) (1/1.1 ligne 2 et 1/1.3 ligne 3). Les produits d'intégration sont ensuite déposés sur gel d'agarose 1 %. (b) Quantification des formes d'intégration FSI + HSI ainsi que les produits FSI linéaires. (c) Quantification des formes d'intégration FSI circulaire par transformation bactérienne puis sélection.

a) Analyse *in silico* du pBSK-Zeo 5S-G5E4 chromatinisé

L'analyse de cet ADN polynucléosomal par un algorithme de prédiction de positionnement des nucléosomes développé par C.Vaillant (ENS Lyon) montre que cet ADN peut permettre le positionnement de nucléosomes régulièrement espacés. L'ADN receveur ainsi assemblé contiendrait alors deux régions bien distinctes de structures chromatinienne différentes. En effet, le profil énergétique qui traduit la stabilité des nucléosomes sur une séquence précise montre la présence d'une région I contenant la séquence 5S-G5E4 constituée d'une chromatine dense où les nucléosomes sont stablement et régulièrement positionnés. En revanche, la région II correspondant à la séquence du plasmide pBSK-Zeo permet la formation de nucléosomes, mais de façon moins régulière et moins organisée mimant une structure chromatinienne dynamique proche des régions transcrites.

b) Analyse biochimique du pBSK-Zeo 5S-G5E4 chromatinisé

La structure des ADN receveurs a pu être validée par analyse biochimique de l'ADN polynucléosomal circulaire obtenu après assemblage selon la méthode de dialyse en sel. Pour cela nous avons entrepris des expériences de protection au clivage par la DNase1 ou par différentes enzymes de restriction coupant dans la région nucléosomale grâce à une approche de REA (Restriction Enzymatic Assay). La présence des nucléosomes sur les sites de coupure des enzymes de restriction va inhiber leur activité endonucléasique se traduisant par une protection au clivage visible en gel d'agarose. Ces analyses montrées en **Figure 52** viennent confirmer les études *in silico* quant à la présence d'une région I peu accessible aux enzymes de restriction marqueur d'une région riche en nucléosomes et d'une région II qui reste accessible aux clivages de l'ADN témoignant d'une plus faible densité en nucléosomes. De plus, le positionnement des nucléosomes dans la séquence 5S-G5E4 a pu être déterminé plus précisément par une digestion par *EcoRI* clivant entre chaque répétition 5S. Ainsi la digestion permet d'isoler chaque mononucléosome qui est ensuite visualisable sur un gel retard confirmant pleinement le bon assemblage du plasmide polynucléosomal (**Figure 52C**).

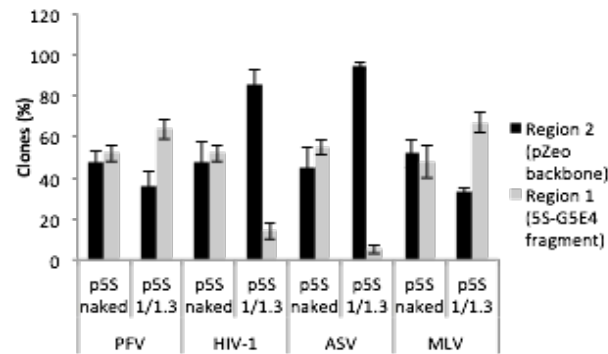


Figure 54 : Localisation des sites d'intégration de VIH-1, PFV, MLV et ASV dans les deux régions du pBSK-Zeo 5S-G5E4. Pour chaque virus, 100 intégrants issus de l'intégration sur plasmide nu ou plasmide polynucléosomal ont été sélectionnés après transformation bactérienne. Le rapport 1/1.3 a été choisi. Le positionnement des sites d'intégration dans la région I ou II du plasmide PBSK-Zeo-G5E4 a été identifié par séquençage. La localisation du site d'intégration est représentée en pourcentage en fonction de la région I ou II.

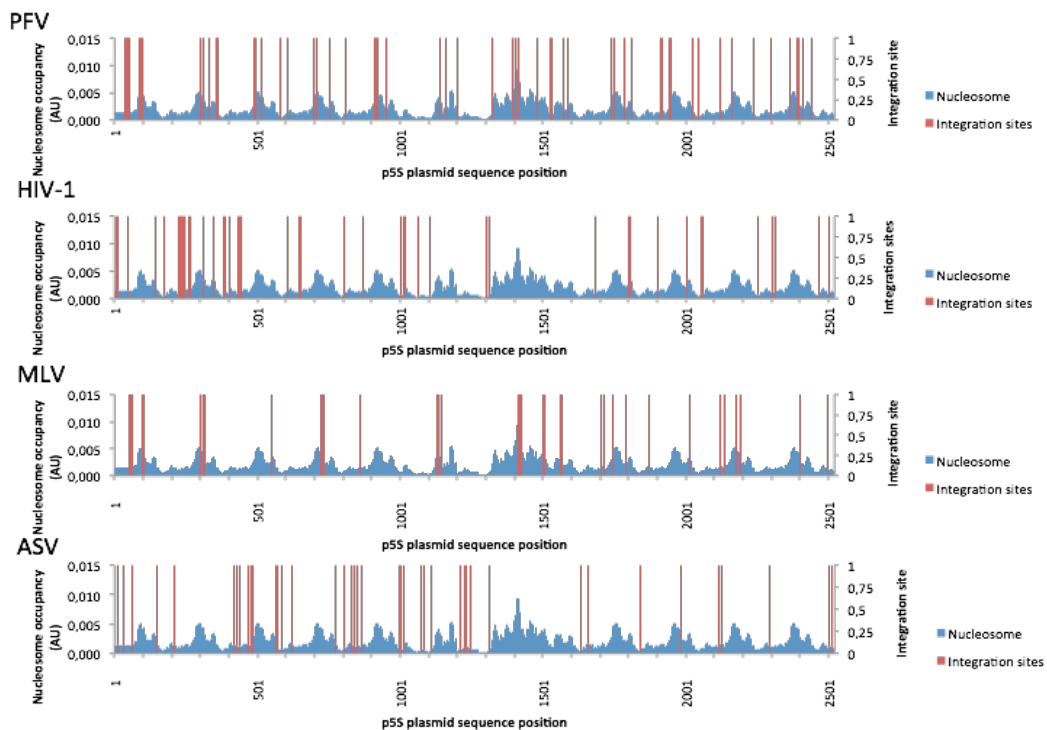


Figure 55 : Comparaison des sites d'intégration de VIH-1, PFV, MLV et ASV avec la prédiction de positionnement des nucléosomes sur le plasmide pBSK-Zeo 5S-G5E4. 100 sites d'intégrations de VIH-1, PFV, MLV et ASV issu de la réaction d'intégration concertée sur plasmide polynucléomale 1/1.3 ont été positionnés sur la séquence d'ADN et comparés à la prédiction de positionnement de nucléosome.

c) Impact de la structure chromatinienne sur l'intégration de différents modèles rétroviraux

Les ADN receveurs ainsi validés ont ensuite été utilisés comme substrat dans les analyses d'intégration des différents systèmes rétroviraux. Les enzymes virales ont été testées à une concentration identique de 100nM finale dans des conditions optimales mises au point préalablement au laboratoire pour chaque enzyme. Dans ces conditions optimisées, l'ensemble de ces IN sont retrouvées actives (**Figure 53a**). De plus, les produits issus de la réaction d'intégration concertée ont été transformés dans des bactéries. Seules les formes circulaires d'intégration concertées des deux extrémités virales (Full Site Integration ou FSI) sont répliquative dans les bactéries. L'isolement et la quantification par clonage et transformation bactérienne confirment que ces enzymes sont également efficaces pour la catalyse de cette activité physiologique. Au contraire, les tests d'intégrations concertées utilisant les substrats polynucléosomaux montrent de fortes différences d'activités entre les systèmes viraux étudiés. En effet, la quantification des formes d'intégrations FSI et HSI (**Figure 53 b**) ainsi que les résultats de clonage (**Figure 53c**) montrent que, contrairement à VIH-1 les activités d'intégration de PFV et MLV sont fortement stimulées d'un facteur 4 à 5 sur un substrat chromatinisé. La quantification de l'activité d'intégration du système ASV montre une forte inhibition d'activité sur une matrice nucléosomale comme observée pour VIH-1. Ces effets sont proportionnels au ratio histone/ADN. De manière intéressante les formes d'autointégration donneur/donneur sont peu ou pas touchées par la présence de nucléosomes sur l'ADN receveur démontrant que l'effet observé est spécifique de la chromatinisation de l'ADN cible.

d) Impact de la structure chromatinienne sur la sélectivité de différents systèmes rétroviraux

Afin d'étudier l'impact de la structure chromatinienne sur le choix du site d'intégration de différents systèmes rétroviraux, les loci d'intégration obtenus sur les plasmides polynucléosomaux ont été séquencés et comparés aux sites d'intégration sur plasmide nu après transformation bactérienne des produits d'intégration circulaires et séquençage. Ce séquençage a permis de montrer qu'en fonction du rétrovirus, différentes régions d'intégration sont privilégiées (**Figure 54**). En effet, les sites d'intégration du VIH-1 sont essentiellement retrouvés dans la région II (86 % des sites d'intégration) correspondant à la séquence du plasmide pBSK-Zeo qui apparaît comme plus dynamique que la région I

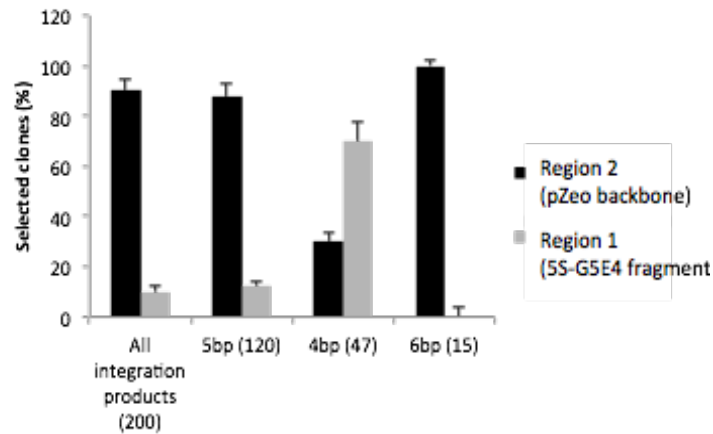


Figure 56 : Localisation des sites d'intégration du VIH-1 en fonction de la duplication de 4, 5 et 6 paires de bases. 100 intégrants issus de l'intégration concertée VIH-1 sur plasmide polynucléosomal ont été sélectionnés après transformation bactérienne. Puis, le positionnement des sites d'intégration dans la région I ou II du plasmide PBSK-Zeo-G5E4 a été réalisé par séquençage. La duplication de paire de bases retrouvée au site d'intégration est représentée en pourcentage en fonction de la région I ou II.

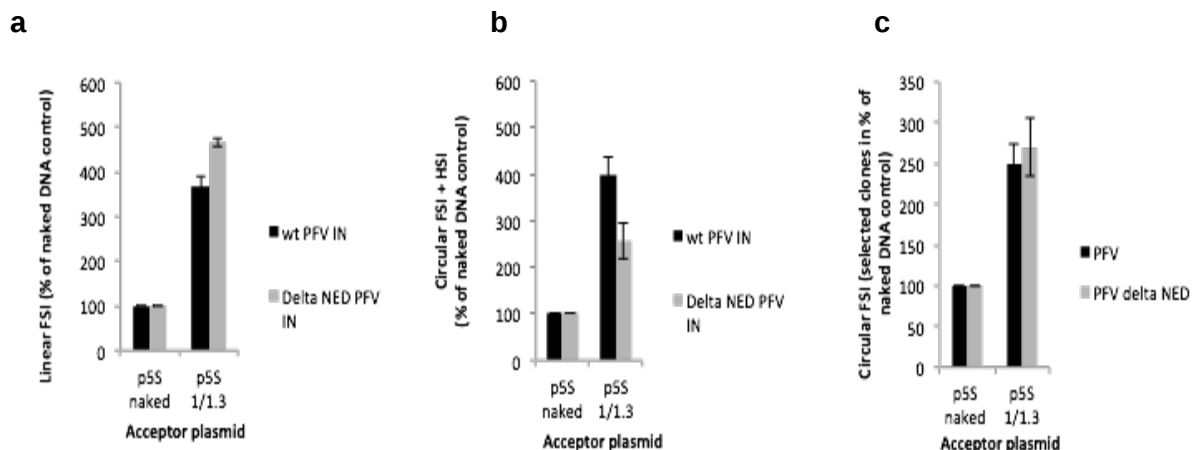


Figure 57 : Comparaison de l'intégration concertée de PFV et PFVΔNED dans le receveur pBSK-Zeo 5S-G5E. (a) Quantification des formes d'intégration FSI linéaire ainsi que (b) des produits d'intégration FSI + HSI. (c) Quantification des formes d'intégration FSI circulaire par transformation bactérienne puis sélection. L'ensemble des clones résistants sont représentés en pourcentage par rapport à la réaction contrôle sur ADN nu.

confirmant les données préalables du laboratoire (Lesbats *et al*, 2011). Ce profil de sélectivité est aussi observé pour ASV où la redistribution des sites d'intégration vers la région II est encore plus importante. Au contraire, les IN de PFV et de MLV montrent une distribution plus homogène avec une préférence significative pour la région stablement positionnée (64 % des sites d'intégration pour PFV et 74 % pour MLV). Ces deux enzymes semblent s'accommoder plus facilement aux deux types de chromatine avec cependant une préférence pour la région dense et stable en nucléosomes. De plus, lorsque le receveur est non chromatinisé, l'intégration dans le cas de tous les virus est sensiblement répartie sur les deux régions I et II confirmant l'absence de biais de ciblage lié à la séquence du plasmide.

Afin de corrélérer plus finement la position des nucléosomes avec les sites d'intégration, nous avons comparé l'ensemble des loci d'intégration obtenus dans la région I du plasmide (nucléosomes stablement positionnés) avec les sites de positionnement des nucléosomes sur le plasmide pBSK-Zeo 5S-G5E4 pouvant être précisément localisés par l'algorithme de prédiction (**Figure 55**). L'analyse de 100 sites d'intégrations pour chaque système rétroviral a permis de montrer que les sites d'intégration de MLV et de PFV étaient retrouvés préférentiellement enrichis dans des régions de forte probabilité de présence de nucléosome. Le calcul de la moyenne de présence de nucléosome à chaque site d'intégration pour MLV et PFV confirme une forte préférence d'intégration sur le nucléosome par rapport à la moyenne de présence de nucléosome de sites d'intégration générés aléatoirement dans la séquence 5S-G5E4. L'ensemble de ces résultats démontre que l'intégration de MLV et PFV *in vitro* est favorisée sur un nucléosome localisé dans une région chromatinienne stable et dense en nucléosomes à la différence de VIH-1 et ASV préférant un nucléosome dans une région dynamique de la chromatine.

e) Impact des structures intasomales sur la sélectivité d'intégration du VIH-1

In vivo, l'intégrase du VIH-1 catalyse l'intégration avec un décalage majoritaire de 5 paires de bases entre les extrémités d'ADN virales. Cependant, nous avons montré qu'*in vitro* l'IN catalyse également une grande proportion d'événements de 4 et 6 paires de bases (**Figure 56**). Au contraire des sites d'intégration portant un décalage de 5 paires de bases retrouvées dans la région II, l'analyse des sites d'intégration du VIH-1 possédant un décalage de 4 paires de bases montre que ces événements sont favorisés dans la région I dense en nucléosome comme observé pour les sites d'intégration de PFV (4pb). Étant donné que la taille de la duplication dépend de la distance entre les deux sites catalytiques au sein des intasomes, ces résultats suggèrent que la différence de sensibilité à l'égard du

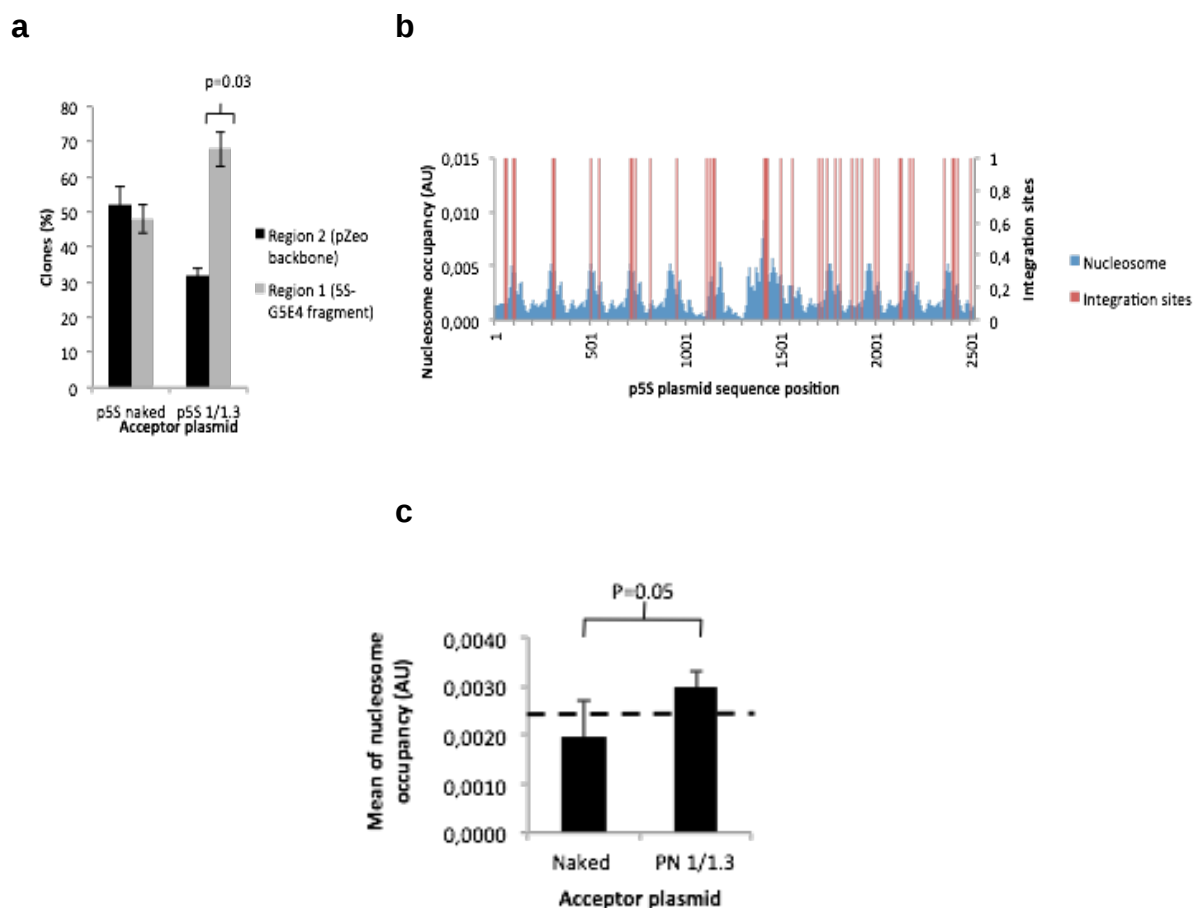


Figure 58 : Localisation et comparaison des sites d'intégration de PFVΔNED. (a) 100 intégrants issus de l'intégration de PFVΔNED sur matrice nue ou chromatinisée ont été sélectionnés chez la bactérie. Le positionnement des sites d'intégration dans la région I ou II du plasmide pBSK-Zeo-G5E4 a ensuite pu être effectué. (b) Les sites d'intégration sont localisés sur la séquence 5S-G5E4 et comparés à la prédiction de positionnement de nucléosomes. (c) Moyenne de prédiction de positionnement de nucléosome à chaque site d'intégration sur plasmide nu et polynucléosomal 1/1.3.

nucléosome de ces deux intasomes viraux pourrait être liée à ce décalage de paire de bases. La structure de l'intasome de PFV publiée durant nos travaux confirme que des contraintes physiques entre l'intasome et l'ADN cible impliquent une structure fortement courbée de l'ADN cible comme retrouvé à la surface du nucléosome et compatible avec une intégration finale sur le nucléosome. Dans le cas du VIH-1 la structure de l'intasome pourrait induire des contraintes différentes requérant des courbures moindres et ainsi des structures chromatinienne optimales différentes.

2. 2. Rôle du « N-terminal extension domain » (NED) de PFV sur l'intégration dans la chromatine

La comparaison de l'alignement des séquences primaires des différentes IN montre que PFV et MLV possèdent un domaine additionnel appelé « N-terminal extension domain ». Ainsi, nous avons investigué le rôle de ce domaine NED afin de savoir si ce domaine pouvait être responsable de la différence de sensibilité de l'intégration *in vitro* vis-à-vis de la chromatine.

a) Impact du domaine NED sur l'intégration *in vitro*

Pour cela, nous avons exprimé et purifié une IN de PFV tronquée pour son domaine NED et nous avons analysé son activité d'intégration dans la chromatine (**Figure 57a**). Les tests d'intégration *in vitro* montrent que l'IN PFV et l'IN PFV Δ NED sont retrouvées aussi actives sur le plasmide nu suggérant ainsi que ces deux enzymes partagent une même structure fonctionnelle. Ces données sont renforcées par l'analyse comparative de la duplication de paire de bases montrant que ces deux enzymes possèdent strictement la même fréquence de duplication de 4 paires de bases au niveau des sites d'intégration confirmant ainsi le partage d'une même structure catalytique fonctionnelle. L'analyse de ces deux enzymes sur l'intégration dans une structure chromatinienne (**Figure 57b et 57c**) montre une forte stimulation des produits d'hétérointégration d'un facteur 3 à 4 pour les produits linéaires FSI et les produits circulaires HSI+FSI. De plus, le clonage et la transformation des produits circulaires FSI confirment la forte stimulation d'activité de ces enzymes sur cette structure.

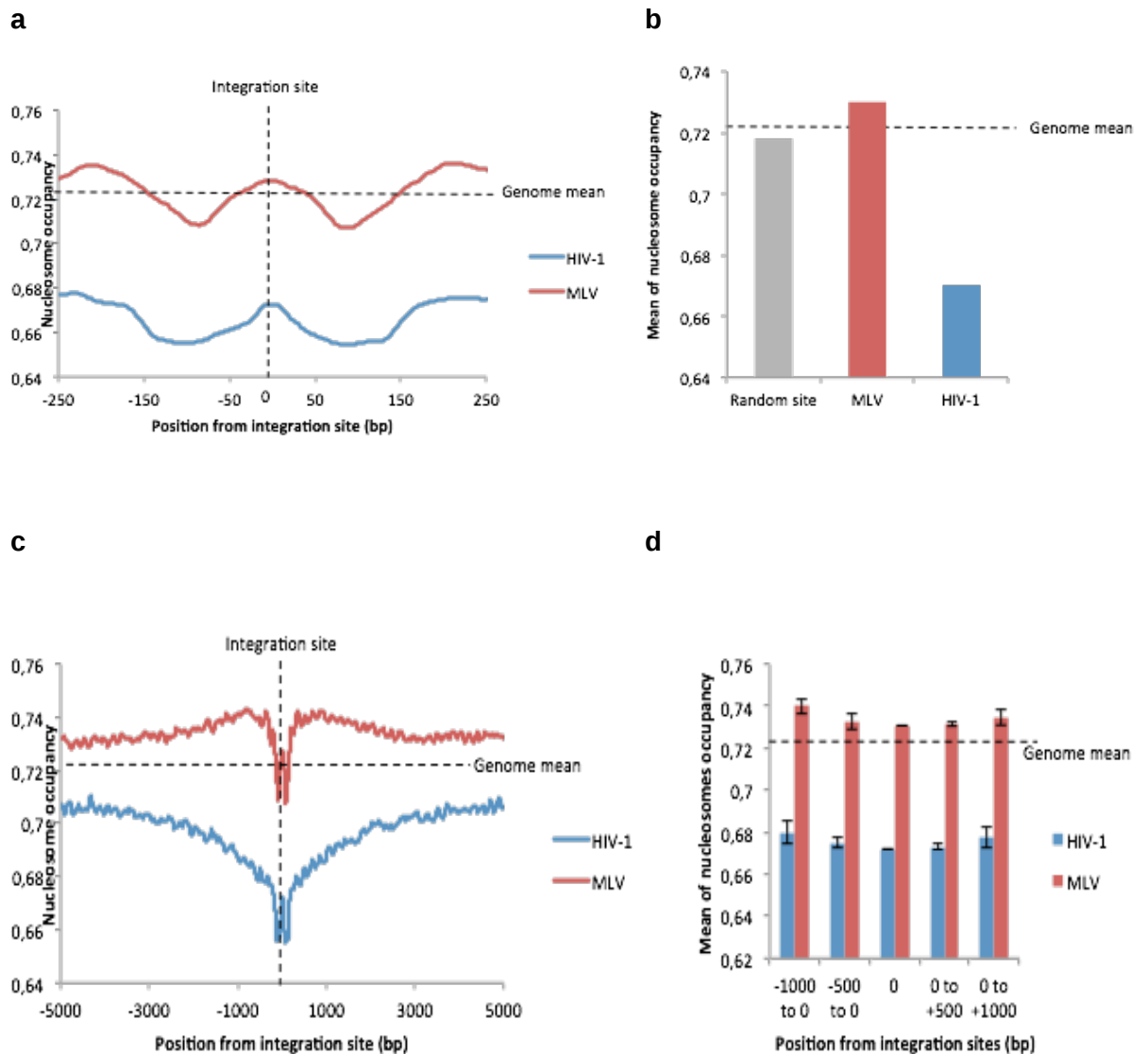


Figure 59 : Corrélation entre localisations des sites d'intégration MLV et VIH-1 avec la prédiction de positionnement des nucléosomes. 41 435 sites d'intégration du VIH-1 et 32 631 de MLV ont été soumis à une analyse de prédiction d'occupation des nucléosomes. (a) Données de prédiction de positionnement de nucléosome entourant le site d'intégration +/- 250 pb pour VIH-1 et MLV. (b) Moyenne de prédiction de positionnement de nucléosome au site d'intégration. (c) Données de prédiction de positionnement de nucléosome entourant le site d'intégration +/- 5000 pb pour VIH-1 et MLV. (d) Moyenne de prédiction de positionnement de nucléosome en fonction de la distance du site d'intégration pour VIH-1 et MLV divisé en quatre régions (-1000 à 0, -500 à 0, 0 à 500, 0 à 1000).

b) Impact du domaine NED sur la sélectivité de l'intégration *in vitro*

L'étude de la sélectivité des sites d'intégration pour l'IN PFVΔNED réalisée par séquençage de plus de 50 sites d'intégration montre une distribution similaire des loci d'intégration vers la région I dense en nucléosome (**Figure 58a**). De plus, l'étude de la sélectivité par corrélation des sites d'intégration avec la position prédite des nucléosomes ne montre pas de différences majeures entre les deux enzymes avec une intégration importante sur les nucléosomes (**Figure 58b et 58c**). L'ensemble de ces données démontre que le domaine NED de l'IN PFV n'est ni responsable de la stimulation de l'intégration dans la chromatine *in vitro*, ni impliqué dans la sélectivité de l'intégration *in vitro*. Ceci conforte l'hypothèse du rôle des sites catalytiques engagés dans l'intégration et de leur distance dans la structure de l'intasome afin de moduler l'intégration à la courbure de l'ADN cible.

2. 3. Étude comparative des sites d'intégration rétroviraux et du positionnement des nucléosomes dans la cellule infectée

Nous avons préalablement montré que les intasomes de différents systèmes rétroviraux possèdent une sensibilité inégale vis-à-vis d'une structure chromatinienne *in vitro*. Afin d'évaluer si cette sensibilité pouvait s'exprimer dans un contexte plus physiologique, nous avons analysé la localisation des sites d'intégration dans la cellule infectée de deux rétrovirus possédant un phénotype d'intégration distinct et opposé par rapport à la prédiction de positionnement des nucléosomes dans le génome. Nous avons ainsi analysé 41 435 et 32 631 sites d'intégration non redondants de VIH-1 et de MLV respectivement publiés dans la littérature (Krishnan & Engelman, 2012; Cattoglio *et al*, 2010). Ces séquences ont été soumises à l'algorithme de prédiction de positionnement des nucléosomes en collaboration avec C.Vaillant (ENS Lyon) comme précédemment utilisé (Lesbats *et al*, 2011). Le positionnement est prédit par la capacité intrinsèque de séquences d'ADN à exclure les nucléosomes. L'ensemble des probabilités d'occupation des nucléosomes a été calculé pour chacune des 41 435 et 32 631 séquences. La **Figure 59a** montre la moyenne de prédiction de positionnement de nucléosome à 250 pb de part et d'autre du site d'intégration. Les résultats montrent clairement que dans les deux cas VIH-1 et MLV, une forte probabilité de présence d'un nucléosome est trouvée sur le site d'intégration (0,73 et 0,67 respectivement). Ceci confirme la préférence d'intégration sur un nucléosome dans un contexte cellulaire. Cependant, le calcul de la moyenne d'occupation

en nucléosome au site d'intégration apparaît comme plus importante pour MLV indiquant une préférence de MLV pour les régions denses en nucléosomes alors que VIH-1 possède une préférence pour les régions de plus faible densité en nucléosomes confirmant ainsi nos résultats *in vitro* (**Figure 59b**). Une analyse plus globale à +/- 5000 pb du site d'intégration montre une augmentation de l'occupation en nucléosome pour MLV alors qu'une baisse est observée pour les loci d'intégration du VIH-1 (**Figure 59c**). Le même type de résultat est observé à 1000 pb du site d'intégration avec une densité en nucléosome de 0,739 et 0,679 pour MLV et VIH-1 respectivement (**Figure 59d**). L'ensemble de ces résultats confirme donc que MLV et VIH-1 présentent tous deux une préférence d'intégration au sein du nucléosome, mais positionnée dans des environnements chromatinien différents dans la cellule infectée. Cette analyse a ensuite pu être étendue récemment à d'autres modèles rétroviraux (PFV, ASV) et les données obtenues (non montré) ont pu confirmer la corrélation entre la capacité des intasomes à intégrer dans des matrices chromatinien *in vitro* et leur sélectivité dans les cellules infectées montrant le rôle important joué par l'intégrase dans le choix du site d'insertion. Ces résultats supportent pour la première fois le concept de modulation de l'intégration rétrovirale par la densité nucléosomale autour du nucléosome ciblé par l'intégration

3. Conclusion et discussion

Les premières données de la littérature ont montré que la structure de la chromatine pouvait influencer l'intégration des rétrovirus suggérant que les propriétés intrinsèques de l'intasome, mais aussi la structure, la densité et la dynamique de la chromatine environnant le nucléosome cible jouent un rôle fondamental dans la sélectivité de l'intégration. Afin de mieux comprendre l'importance de cette régulation de l'intégration par la chromatine nous avons choisi différentes intégrases appartenant à des genres rétroviraux distincts et présentant une sélectivité et un nombre de paires de bases dupliquées différents au site d'intégration. Les intégrases de VIH-1, PFV, ASV et MLV ont ainsi été testées *in vitro* afin d'évaluer leur activité d'intégration sur substrat chromatinien. Les différents résultats ont montré que ces enzymes n'étaient pas affectées de la même façon. En effet, dans nos conditions d'intégration *in vitro*, il apparaît que l'intégration de MLV et de PFV est stimulée dans un plasmide chromatinisé alors que l'intégration de VIH-1 et ASV est fortement inhibée par ces structures chromatinien. De plus, la localisation des sites d'intégrations montre que PFV et MLV s'intègrent dans la région I correspondant à une chromatine dense et stable. En revanche, les sites d'intégrations

de VIH-1 et ASV se trouvent préférentiellement dans la région II formée par une chromatine moins dense et dynamique en nucléosome. Nous avons ensuite pu montrer que cette différence d'intégration dans la chromatine n'était pas dépendante de la présence du domaine NED de PFV suggérant alors une différence dans l'architecture même du domaine catalytique au sein de ces intasomes et pas de la structure globale des différents intasomes. Ces données sont supportées par l'analyse de la duplication de 4, 5 et 6 paires de bases des sites d'intégration du VIH-1. En effet, les duplications de 5 et 6 paires de bases sont exclues de la région dense en nucléosome et les duplications de 4 paires de bases sont plus dispersées avec une préférence pour la région dense en nucléosome suggérant que la distance entre les sites catalytiques ne permet pas l'intégration dans une chromatine compacte. Par conséquent, l'ensemble de ces données indique que les intasomes de PFV et MLV sont capables de par l'architecture de leur site catalytique, de s'adapter aux contraintes imposées par l'ADN nucléosomal dans ces régions denses contrairement à VIH-1 et ASV. Ces résultats rapportent pour la première fois des comportements distincts de différents intasomes et renforcent la possibilité qu'ils puissent présenter des structures actives distinctes. Ceci a pu être confirmé depuis par l'arrivée de nouvelles données structurales d'intasomes rétroviraux montrant des variations importantes dans l'agencement et le nombre de monomères constituant la structure active (Ballandras-Colas *et al*, 2016 ; Yin *et al*, 2016). Afin de savoir si ces résultats *in vitro* étaient concordants avec les mécanismes d'intégration *in vivo*, nous avons comparé la relation entre les sites d'intégration de deux rétrovirus VIH-1 et MLV avec la présence de nucléosome dans la cellule infectée. Cette étude a permis de montrer que les sites d'intégration de VIH-1 et MLV sont localisés dans des environnements chromatinien différents. En effet dans le cas de MLV, les sites d'intégrations sont retrouvés enrichis dans les régions plus denses en nucléosome par rapport à VIH-1 alors que dans les deux cas un nucléosome est trouvé préférentiellement sur le site. Les données cellulaires semblent en accord avec les résultats *in vitro* et sont compatibles avec l'hypothèse d'une sensibilité différentielle des intasomes rétroviraux vis-à-vis de l'ADN nucléosomal et de leur capacité à s'adapter à une chromatine plus ou moins dense. Par conséquent, l'intégration rétrovirale dans la chromatine apparaît être un processus en plusieurs étapes ; tout d'abord le ciblage de l'intasome vers les régions appropriées de la chromatine dû à l'interaction spécifique des intasomes rétroviraux avec leur facteur cellulaire de ciblage comme LEDGF/p75 pour VIH-1 et les protéines BET pour MLV ; puis, une étape de ciblage très local, gouvernée par la capacité intrinsèque des intasomes rétroviraux à cibler l'ADN nucléosomal courbé et la chromatine dense environnant le loci d'intégration. Les dernières étapes sont dépendantes des contraintes physiques imposées par la structure de l'intasome qui modulent l'interaction fonctionnelle avec le

nucléosome. L'incapacité de l'intasome de VIH-1 à intégrer dans la chromatine dense peut être liée à un manque d'accès à l'ADN ou à un manque d'interaction avec certains composants du nucléosome. En effet, des processus de remodelage de la chromatine pourraient compenser ces propriétés réfractaires des structures nucléosomales denses. Afin de mieux comprendre ces processus de régulation, la suite de mon travail s'est focalisée vers l'identification de facteurs cellulaires potentiellement impliqués dans la régulation de l'intégration du VIH-1 par la modulation de la structure chromatinienne de l'hôte.

Partie 2 : Identification de nouveaux cofacteurs cellulaires de l'intégration : rôle du complexe protéique FACT « Facilitates Chromatin Transcription »

Ce travail a donné lieu à une publication en premier auteur, en cours de soumission « *The FACT chromatin remodeling complex associated to RNA polymerase II machinery generates partially dissociated nucleosomes structures that are favored substrates for HIV-1 integration* » Ce travail a permis le développement d'une stratégie nouvelle d'identification de cofacteurs de complexe nucléoprotéique et a permis d'identifier le complexe FACT comme potentiel cofacteur de l'intégration du VIH-1.

1. Objectifs

Bien que l'intégrase virale soit capable d'effectuer seule la réaction d'intégration *in vitro*, il est aujourd'hui établi que plusieurs protéines cellulaires participent à cette étape dans la cellule infectée. Ces dernières années, une attention toute particulière s'est portée sur l'identification de cofacteurs cellulaires de l'intégrase qui ont pu être identifiés indépendamment par plusieurs laboratoires, dont le notre, tels que LEDGF/p75, INI1, BAF, HMGA1, RAD51 et HSP60 (Kalpana *et al*, 1994; Lee & Craigie, 1994; Farnet & Bushman, 1997; Cherepanov *et al*, 2003; Lin & Engelman, 2003; Cosnefroy *et al*, 2012). Ces cofacteurs de l'intégrase du VIH-1 ont été décrits pour jouer un rôle sur la catalyse, l'import nucléaire, la sélectivité et la réparation du site d'intégration (voir Chapitre 2. 4).

Les interactions de l'IN avec ces facteurs cellulaires sont des cibles potentielles pour le développement d'inhibiteurs antirétroviraux telles que des petites molécules inhibitrices de l'interaction protéine-protéine ou SMIPPIs « small-molecule inhibitor of protein-protein interaction ». C'est pourquoi notre laboratoire s'est également focalisé sur la recherche des cofacteurs cellulaires de l'intégrase du VIH-1 et sur l'identification de leur mécanisme d'action, notamment dans le cadre de l'association de l'intasome avec la chromatine de l'hôte. C'est donc dans ce contexte que les mécanismes d'action de la protéine RAD51 et du complexe SWI/SNF ont pu être étudiés au laboratoire (Lesbats *et al*, 2011; Thierry *et al*, 2015). Ces résultats ont notamment pu montrer pour la première fois un couplage fonctionnel entre le processus d'intégration rétrovirale et le remodelage de la chromatine.

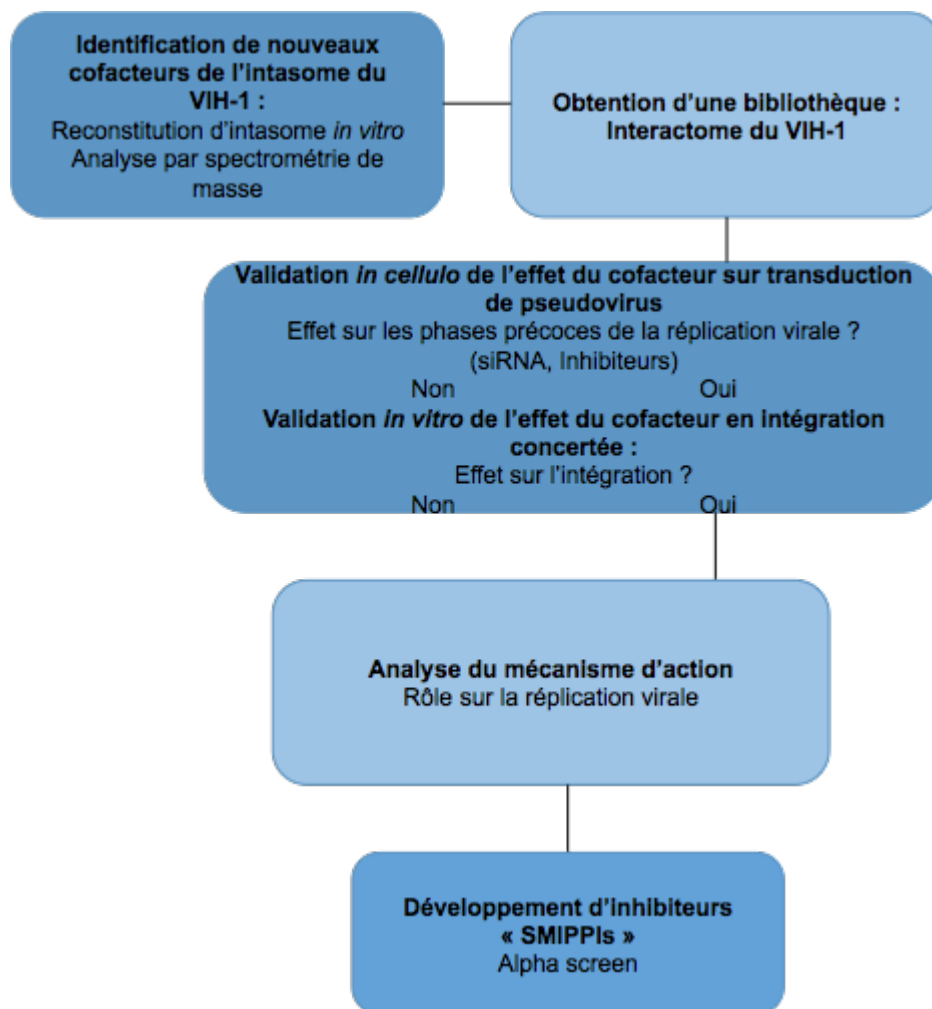


Figure 60 : Algorithme de validation des cofacteurs cellulaires de l'intasome du VIH-1.

Afin d'identifier de nouveaux cofacteurs cellulaires de l'intasome du VIH-1, nous avons dans un premier temps reconstitué *in vitro* un complexe nucléoprotéique actif puis nous avons ensuite identifié les protéines cellulaires présentant une interaction directe ou indirecte avec ce complexe par spectrométrie de masse MS/MS afin d'obtenir une bibliothèque de cofacteurs cellulaires. Dans un second temps, nous avons analysé l'effet *in vivo* et *in vitro* de ces cofacteurs potentiels sur l'intégration du VIH-1. Dans un troisième temps, les facteurs cellulaires ayant un effet *in vivo* et *in vitro* ont pu être analysés plus finement afin de comprendre le rôle de ces protéines sur l'intégration. Finalement, le développement d'inhibiteurs de type SMIPPIs visant à inhiber l'interaction de ces facteurs avec le complexe IN-ADN pourra être entrepris.

Comme reporté dans le chapitre précédent, de tels facteurs associés à la chromatine sont susceptibles de réguler l'accès du site d'intégration aux complexes d'intégration. Dans la continuité de ces études, nous avons poursuivi la caractérisation de ces facteurs.

La majorité des cofacteurs cellulaires de l'intégrase ont été sélectionnés sur la base d'une interaction avec l'IN seule. Cependant le complexe actif dans la cellule est l'intasome constitué de l'IN associée à son ADN viral. Ainsi nous avons développé une stratégie alternative et novatrice permettant l'identification de cofacteurs cellulaires spécifiques de ce complexe nucléoprotéique. Compte tenu des précédents résultats, nous nous sommes focalisés sur l'identification de facteurs impliqués dans le remodelage de la chromatine. Il est à noter que cette approche développée au laboratoire ne se limite pas aux complexes rétroviraux et pourra être appliquée dans le cadre de la recherche d'interactions protéine-protéine d'autres complexes nucléoprotéiques impliqués dans la mobilité de l'ADN, mécanismes étudiés depuis de longues dates au laboratoire.

2. Identification et validation des cofacteurs cellulaires de l'intasome du VIH-1

2.1. Stratégie d'identification de cofacteurs cellulaires

La stratégie développée dans cette étude se base sur l'identification de l'interactome d'un complexe nucléoprotéique intasomal (**Figure 60**). Pour cela, l'approche nécessite dans un premier temps la reconstitution *in vitro* de cet intasome puis la mise en contact du complexe avec un extrait protéique cellulaire. Les protéines cellulaires interagissant directement et indirectement avec l'intasome sont ensuite identifiées par spectrométrie de masse MS/MS. Les cofacteurs putatifs sont ensuite soumis à un algorithme de validation afin d'analyser leur implication dans l'étape d'intégration du VIH-1. Premièrement *in vitro* dans des tests d'intégration concertée afin d'analyser l'effet de ces cofacteurs sur l'intégration, puis *in vivo* par une analyse de l'impact de ces facteurs cellulaires sur les phases précoces de réplication de pseudovirus par des techniques d'extinction de leurs gènes (siRNA). Dans un second temps, le rôle de ces cofacteurs sur la réplication du VIH-1 sera analysé plus finement. Finalement, le développement d'inhibiteurs antirétroviraux de type « SMIPPIs » visant les différents cofacteurs cellulaires pourra être entrepris.

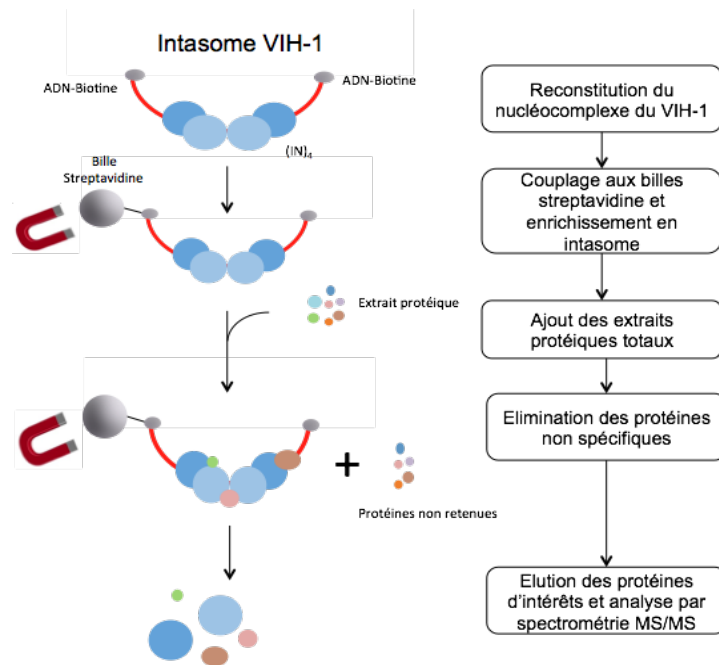


Figure 61 : Représentation schématique de la stratégie d'identification des cofacteurs cellulaires de l'intasome du VIH-1. La stratégie mise au point consiste en la reconstitution d'un complexe nucléoprotéique actif mimant l'intasome du VIH-1. Ce complexe est ensuite fixé à des billes magnétiques couplées à la streptavidine. Puis, un extrait protéique est ajouté et les protéines ayant une interaction directe ou indirecte avec le complexe sont ensuite éluées et identifiées par spectrométrie de masse MS/MS.

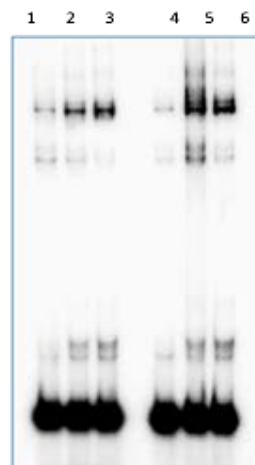


Figure 62 : Analyse fonctionnelle du complexe nucléoprotéique IN-ADN viral. (a) Les conditions optimales à la formation d'un complexe IN-ADN ont été déterminées par un test d'intégration concertée dans des conditions précédemment optimisées (Lesbats et al NAR 2008). Les pistes 1 à 3 correspondent à un test réalisé en présence d'un ADN donneur viral non traité et les lignes 4 à 6 en présence d'un ADN pré-traité. Les concentrations d'IN utilisées sont de 100 nM pour les pistes 1 et 4, 200 nM pour les pistes 2 et 5, 400 nM pour les pistes 3 et 6.

2.2. Sélection des cofacteurs putatifs de l'intasome du VIH-1

Dans l'objectif d'identifier de potentiels nouveaux cofacteurs de l'intasome du VIH-1, nous avons développé une stratégie illustrée en **Figure 61**. Cette stratégie consiste dans un premier temps à une reconstitution *in vitro* d'un complexe nucléoprotéique mimant l'intasome du VIH-1. Pour cela, l'IN du VIH-1 est incubé avec un ODN (oligodésoxyribonucléotide) contenant les 21 derniers nucléotides de l'extrémité U5 virale et marquée à la biotine au niveau de l'extrémité 5' en complexe avec son brin complémentaire. Les conditions de reconstitution du complexe ont pu être déterminées préalablement au laboratoire sur la base des études structures fonctions menées par l'équipe depuis de longues années (Faure *et al* 2004, Lesbats *et al*, 2008). Le nucléocomplexe reconstitué a ensuite pu être testé *in vitro* afin d'analyser son activité dans des tests d'intégration concertée (**Figure 62**). Puis, nous avons effectué un couplage non covalent de l'intasome précédemment reconstitué *in vitro* sur des billes magnétiques couplées à la streptavidine. Puis, après lavage, un extrait protéique total issu de cellules 293T ou HeLa-P4 est incubé avec les intasomes reconstitués comme décrits dans la partie matériels et méthodes. Pour des études plus spécifiques, un extrait protéique issu de cellules de lignées différentes ainsi que des extraits protéiques nucléaires ou cytoplasmiques peuvent se substituer aux extraits totaux. Après plusieurs lavages à force ionique élevée afin d'éliminer les protéines non spécifiques, l'ensemble des protéines ayant une affinité avec l'intasome est élué et analysé en spectrométrie de masse MS/MS. De plus, en vue d'éliminer les protéines de liaison à l'ADN non spécifique de l'intasome du VIH-1 reconstitué *in vitro*, une expérience semblable est effectuée en parallèle avec des billes couplées uniquement avec l'ADN viral seul. Ainsi, les protéines de liaison à l'ADN et aux billes pourront être identifiées et soustraites des protéines d'intérêt.

2.3. Obtention de l'interactome de l'intasome du VIH-1

Les résultats de spectrométrie de masse MS/MS ont permis l'obtention d'un interactome de l'intasome du VIH-1 constitué de près de 50 protéines interagissant directement ou indirectement de manière reproductible avec le nucléocomplexe. De façon intéressante, plusieurs protéines déjà décrites dans la littérature comme cofacteurs de l'IN ou jouant un rôle dans l'intégration ont pu être identifiées par cette stratégie telles que LEDGF/p75, HMGA1, RAD51, FEN1 et INI1 validant notre stratégie (**Tableau 63**).

Trafic intracellulaire et import nucléaire	Voies de dégradation cellulaires	Réparation de l'ADN	Dynamique et remodelage de la chromatine
Importine 7	VBP1	FEN-1	LEDGF/p75
TNPO1	HSP105	RAD51	INI1, SMARCB1
TNPO2	HSP110	Ku70/Ku80	HMGA1
MAP4	ClpX	BRCA1	SETD2
NUP93	SUMO	MRE11	HDAC1, HDAC2
NUP160		APEX1	NAP1L1
		DNA-PK	FACT(SSRP1, SUPT16H)
			PARP1

Tableau 63 : Protéines cellulaires identifiées. En rouge sont représentés les cofacteurs cellulaires décrits dans la littérature. En noir, les protéines candidates identifiées par la stratégie d'identification de cofacteurs et classées par famille

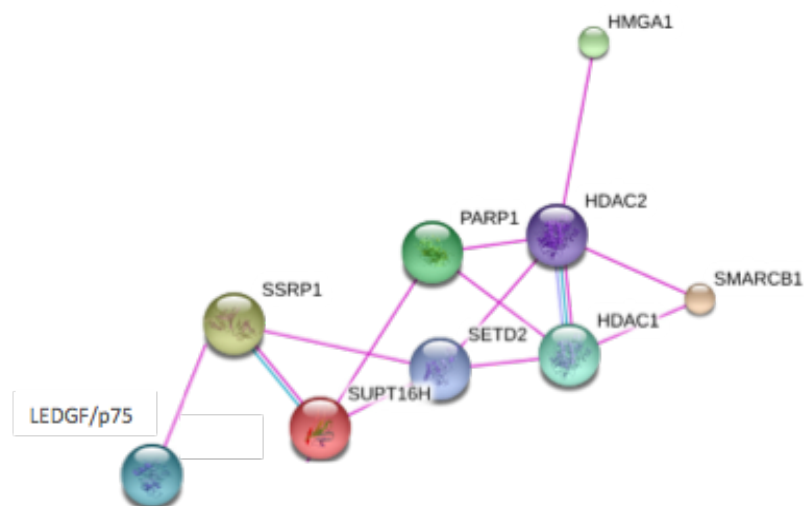


Figure 64 : Réseau d'interaction entre les protéines identifiées. En rose sont représentées les interactions déterminées expérimentalement dans la littérature, en bleu sont représentées les interactions existantes dans la base de données et en violet sont représentés les liens de forte homologie protéique. Adapté de (Szklarczyk *et al*, 2015).

Nous nous sommes ensuite focalisés sur les protéines possédant une fonction liée à l'intégration telles que l'import nucléaire, la réparation de l'ADN, la régulation de la chromatine et les voies de dégradation protéique. Au vu des problématiques présentées dans les chapitres précédents, nous nous sommes focalisés sur les protéines ayant un lien direct avec la régulation et la dynamique de la chromatine. Ces protéines ont dans un premier temps pu être analysées par l'outil string-db.org afin d'établir l'appartenance de ces protéines à des réseaux d'interactions protéine-protéine (**Figure 64**). Puis, afin de valider de nouveaux cofacteurs cellulaires de l'intégration, ces cofacteurs putatifs ont pu être testés premièrement *in vitro* dans des tests d'intégration concertée, et *in vivo* par une analyse de l'impact de ces facteurs cellulaires sur la réplication de pseudovirus par des techniques de siRNA.

3. Rôle du complexe protéique FACT sur l'intégration du VIH-1

3. 1. Introduction

La dynamique de la chromatine est un processus déterminant afin de permettre l'ensemble des mécanismes ADN dépendant tels que l'expression des gènes, la réparation de l'ADN et la division cellulaire. Afin de rendre l'ADN accessible à l'ensemble de ces mécanismes cellulaires, l'action orchestrée de facteurs de remodelage de la chromatine tels que les histones chaperonnes, les complexes de remodelage ATP-dépendant et les enzymes responsables de modifications post-traductionnelles des histones permettent le réarrangement de l'architecture de la chromatine au niveau du nucléosome (voir Chapitre 3. 3.). De plus, le laboratoire a déjà pu démontrer l'importance du remodelage pour réguler l'intégration du VIH-1 notamment dans le cas du complexe SWI/SNF. En effet, les différentes données démontrent que le complexe SWI/SNF grâce à son activité de remodelage de la chromatine ATP-dépendant va permettre une intégration optimale du VIH-1 *in vitro* sur ADN polynucléosomal.

La stratégie d'identification de nouveaux cofacteurs de l'intasome du VIH-1 nous a permis d'identifier un nouveau complexe protéique appartenant à la famille des histones chaperonnes. Ce complexe FACT (Facilitates chromatin transcription) est un complexe protéique composé des protéines SSRP1 (structure-specific recognition protein 1) et Spt16 (suppressor of Ty16) et associé à l'ARN polymérase de type II. Étant donné que l'intégration

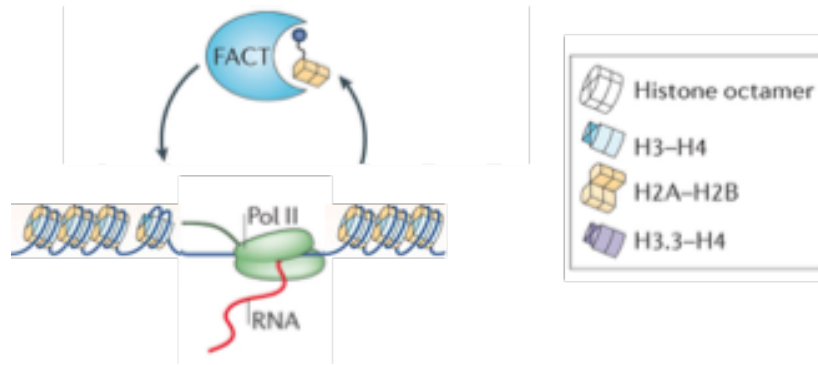


Figure 65 : Représentation schématique de l'activité du complexe de remodelage FACT.

Le complexe FACT (Facilitates chromatin transcription) est un complexe protéique composé des protéines SSRP1 (structure-specific recognition protein 1) et Spt16 (suppressor of Ty16). Ce complexe appartenant à la famille des histones chaperonnes et intimement lié à l'ARN polymérase de type II permet l'éjection et le réassemblage du dimère d'histones H2A/H2B suite à la transcription. Adapté de (Venkatesh & Workman, 2015).

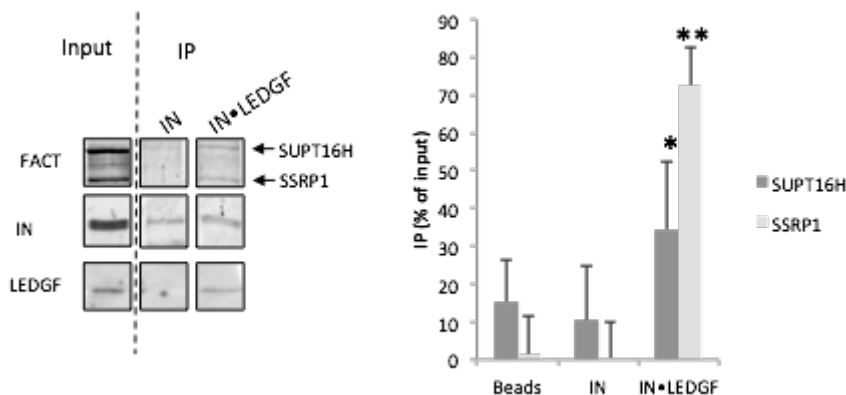


Figure 66 : Interaction entre le complexe IN-LEDGF/p75 et le complexe de remodelage FACT.

Le complexe FACT est composé des protéines SSRP1 et Spt16. Grâce à des expériences de co-immunoprécipitation dans lesquelles des billes magnétiques couplées à la protéine G sont incubées avec des anticorps anti-IN, puis avec l'IN ou le complexe IN-LEDGF/p75, nous avons montré que le complexe FACT est capable d'interagir significativement avec le complexe IN-LEDGF/p75, mais ne semble pas interagir avec l'intégrase seule. La protéine LEDGF/p75 est révélée par western blot, L'IN, SSRP1 et Spt16 par coloration au nitrate d'argent. Les résultats sont représentés en pourcentage de la charge initiale (gauche) et sont issus de la moyenne d'au moins trois expériences effectuées $\pm$ écart type. Les valeurs p ont été calculées par le test « t » de Student et sont représentées par * $p < 0,05$ et ** $p < 0,005$ pour indiquer la probabilité de différences significatives par rapport aux témoins négatifs obtenus avec les billes seules.

du VIH-1 se fait préférentiellement au niveau des régions transcrites et reste intimement liée à l'ARN polymérase de type II, nous nous sommes particulièrement intéressés à ce complexe protéique. Contrairement aux complexes de remodelages, les histones chaperonnes ne sont pas ATP-dépendantes et utilisent le mouvement de l'ADN autour de l'axe dyade pour déstabiliser le nucléosome localement. Une telle activité permettrait à l'intasome d'accéder à une structure nucléosomale adaptée pour l'intégration. Ainsi, nous avons donc investigué le rôle du complexe FACT sur l'intégration du VIH-1. Ce complexe est décrit dans la littérature comme jouant un rôle dans la réplication, la réparation et la transcription de l'ADN grâce à son activité d'assemblage et de désassemblage du nucléosome. En effet, il est décrit que le complexe FACT agit en permettant l'éviction d'un dimère H2A-H2B afin d'obtenir une structure hexamérique du nucléosome (**Figure 65**). Ce complexe va ensuite permettre le réassemblage du nucléosome après la transcription via l'ARN polymérase de type II. Afin de valider le complexe FACT comme cofacteur de l'intégration, nous avons dans un premier temps identifié son mode d'interaction avec l'intasome du VIH-1. Puis, nous avons analysé l'impact d'une déplétion du complexe FACT par siRNA *in cellulo* sur la réplication de vecteurs lentiviraux pseudotypés VSV-G dans des expériences de « single round infection assay » et finalement, nous avons analysé le rôle du complexe FACT sur l'intégration concertée *in vitro*

3. 2. Mode d'interaction du FACT avec l'intasome du VIH-1

Dans le but d'étudier le mode d'interaction direct ou indirect du complexe FACT avec l'intasome du VIH-1, nous avons réalisé des expériences de co- immunoprécipitation. Ainsi, nous avons pu tester l'interaction directe des protéines recombinantes SSRP1 et Spt16 avec l'IN recombinante du VIH-1. Les résultats n'ont pas pu mettre en évidence d'interaction directe entre ces partenaires. De manière intéressante, une interaction entre la protéine SSRP1 et LEDGF/p75 a pu être récemment démontrée (Lopez *et al*, 2016). Ces résultats pouvant suggérer que le FACT interagit avec l'IN via LEDGF/p75, nous avons testé cette interaction. Comme décrit dans la **Figure 66** nous avons pu démontrer que le complexe tripartite FACT-LEDGF/p75-IN pouvait s'effectuer. Ainsi l'ensemble de ces résultats est en faveur d'une interaction indirecte entre l'intasome du VIH-1 et le complexe FACT via le cofacteur LEDGF/p75. Ces résultats montrent aussi que l'interaction de l'IN et de SSRP1 à LEDGF/p75 n'est pas exclusive et que l'interaction IN-SSRP1 peut se réaliser au sein du complexe physiologique IN-LEDGF/p75. Ces données suggèrent également que ces trois

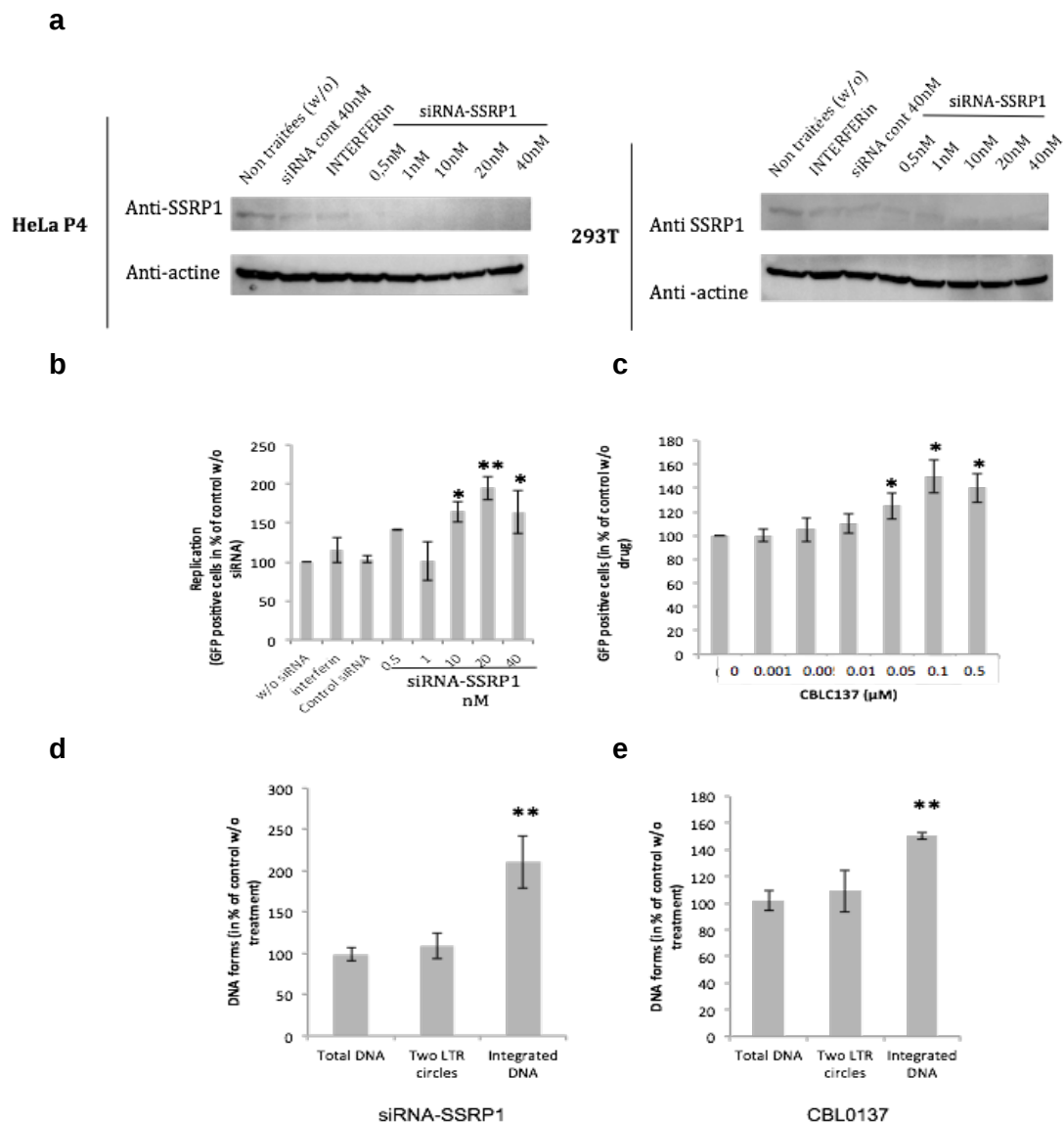


Figure 67 : Effet du FACT sur les phases précoces de la réplication virale (a) Mise au point de l'extinction du gène codant pour SSRP1 par siRNA. Des extraits protéiques de cellules HeLa P4 (gauche) et 293T (droite) obtenus après transfection de quantités croissantes de siRNA SSRP1 sont séparés par SDS-PAGE puis transférés sur membrane PVDF. La protéine SSRP1 et l'actine sont détectées par western blot. b) Quantification des cellules GFP positives. Les cellules sont transfectées par une gamme croissante de siRNA-SSRP1 puis transduites 48 h après par des vecteurs lentiviraux. Le pourcentage de cellules GFP positives est exprimé en pourcentage des cellules non traitées (c) Quantification des cellules GFP positives suite au traitement des cellules par une gamme croissante de CBL0137 et transduction de vecteurs lentiviraux. Le pourcentage de cellules GFP positives est exprimé en pourcentage des cellules non traitées (d) et (e) Quantification de l'ADN total, des cercles à 2 LTR et de l'ADN intégré suite au traitement des cellules par le siRNA-SSRP1 et par la molécule CBL0137 respectivement. Les résultats sont représentés en pourcentage par rapport à la condition non traitée et sont issus de la moyenne d'au moins trois expériences effectuées en double $\pm$ écart type. Les valeurs p ont été calculées par le test « t » de Student et sont représentées par * $p < 0,05$ et ** $p < 0,005$ pour indiquer la probabilité de différences significatives par rapport aux témoins négatifs.

partenaires pourraient se trouver à proximité sur le site d'intégration confortant l'idée d'une action simultanée au moment de l'intégration.

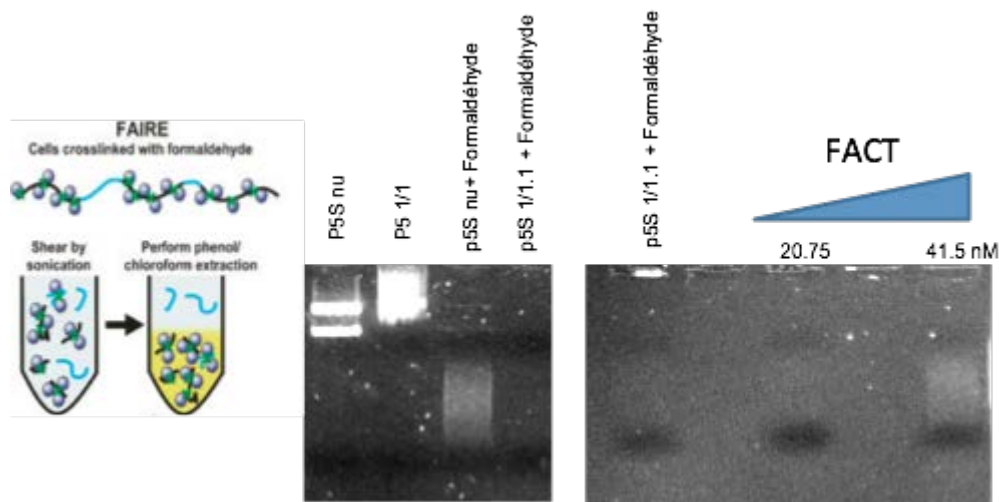
3. 3. Rôle du complexe FACT sur les phases précoces de la réplication *in cellulo*

Dans l'objectif d'analyser le rôle potentiel du complexe FACT *in cellulo*, nous avons choisi d'inhiber l'expression de la protéine SSRP1 par des techniques de siRNA. Pour cela, une double transfection de siRNA-SSRP1 dans les lignées cellulaires 293 T et HeLa P4 a été entreprise afin d'obtenir une inhibition d'expression de la protéine SSRP1 concomitante avec le pic d'activité d'intégration déterminé autour de 18 h post infection (Mohammadi *et al*, 2013). L'analyse de l'expression de la protéine à 18 heures post-infection montre une baisse d'expression de la protéine SSRP1 dans les deux lignées cellulaires utilisées (**Figure 67a**). Nous avons ensuite pu analyser l'infektivité de vecteurs rétroviraux lors de la baisse d'expression de SSRP1. Les résultats montrent une augmentation des cellules GFP positives lors de l'inhibition du complexe FACT de près de deux fois par rapport aux cellules non traitées (**Figure 67b**).

Ces données ont ensuite pu être confirmées par une deuxième approche visant à inhiber l'activité du FACT *in cellulo* par l'utilisation des drogues inhibitrices telles que les curaxines. Ces composés sont décrits comme permettant l'inhibition du complexe par séquestration du FACT sur la chromatine cellulaire (Gasparian *et al*, 2011). Ainsi, les résultats montrent de façon similaire une augmentation des cellules GFP positives lors de l'inhibition du complexe FACT par la molécule CBL0137 (**Figure 67c**).

Afin de savoir quelle étape des phases précoces de la réplication virale était affectée lors de l'inhibition du complexe FACT, les populations d'ADNc, d'ADN intégré et des cercles à 2 LTR ont pu être quantifiées (collaboration avec le Dr Olivier Delelis, ENS Cachan) (**Figure 67d et 67e**). Cette analyse montre que la production d'ADNc et des cercles à 2 LTR n'est pas affectée ne suggérant aucune altération dans l'étape de rétrotranscription de l'ARN viral en ADN et dans l'étape de l'import nucléaire du complexe intasomal. Néanmoins, une augmentation significative de l'ADN intégré est observée lors de la déplétion du complexe FACT. Ces résultats confirment que le FACT impacte l'étape d'intégration du virus. Nous avons donc voulu déterminer le mode d'action de ce complexe sur cette étape de la réplication virale.

a



b

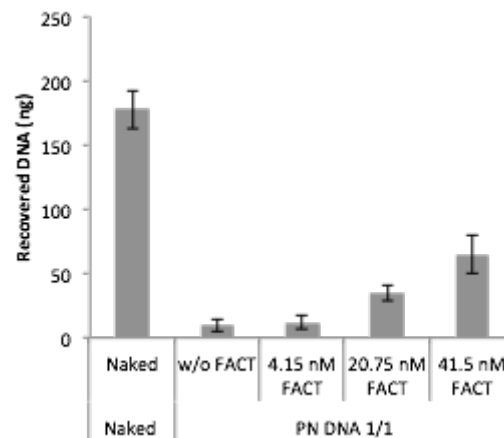


Figure 68 : Mise en évidence de l'activité du FACT (a) Gel représentatif d'une expérience de FAIRE « Formaldehyde-Assisted Isolation of Regulatory Elements ». Les matrices polynucléomales traitées ou non par le FACT puis fixées au formaldéhyde comme décrit dans la partie matériels et méthodes sont ensuite soumises à une extraction phénol/chloroforme et la phase aqueuse est chargée sur gel d'agarose. La présence d'ADN sur gel montre la présence d'une activité de désassemblage des nucléosomes visualisable par la présence d'un smear d'ADN comme observé sur la piste p5S nu + Formaldéhyde (b) Quantification de l'ADN correspondant à l'effet de désassemblage des nucléosomes. Les résultats sont représentés en pourcentage par rapport à la condition non traitée et sont issus de la moyenne d'au moins trois expériences $\pm$ écart type.

3. 4. Rôle du complexe FACT sur l'intégration concertée *in vitro*

Afin d'analyser et de confirmer le rôle du complexe de remodelage de la chromatine FACT sur l'intégration du VIH-1, nous avons couplé ces deux réactions ensemble. Pour cela, le complexe FACT a dans un premier temps été purifié et reconstitué *in vitro* en collaboration avec A.Hamiche (IGBMC Strasbourg) et E.Campos (NYU New York) et son activité de remodelage du nucléosome a pu être vérifiée par des expériences de FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) (**Figure 68**). Ayant vérifié l'activité du FACT *in vitro*, nous avons ensuite regardé son effet dans un test d'intégration concertée (**Figure 69a**). Les résultats montrent que le remodelage de la chromatine par le complexe FACT entraîne une stimulation de l'intégration d'un facteur 1,2 par rapport à l'activité de base sur matrice nue par réduction de la réassociation des nucléosomes (non montré). En revanche, les résultats montrent une baisse de l'activité d'intégration sur ADN nu de manière proportionnelle à la quantité de FACT qui peut être expliquée par une fixation non spécifique du FACT à l'ADN cible inhibant ainsi l'intégration.

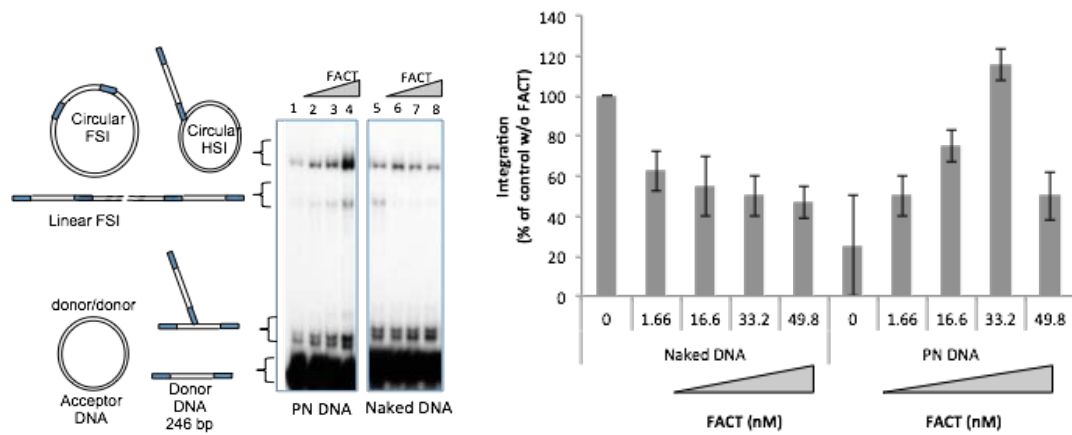
Nous avons ensuite quantifié les formes circulaires d'intégration après transformation chez la bactérie. Des résultats semblables sont obtenus et confirment ces résultats préliminaires (**Figure 69b**). Ces données sont donc en accord avec un rôle du complexe FACT dans l'intégration VIH-1 *in vitro* et *in cellulo* et confirment l'identification du FACT comme cofacteur de l'intégration du VIH-1.

4. Étude du mécanisme d'action impliqué dans la restauration de l'intégration médié par le complexe FACT

4. 1. Étude de l'activité de remodelage du complexe FACT

Afin de savoir si l'activité de remodelage du complexe FACT est responsable de la restauration de l'intégration du VIH-1 *in vitro*, nous avons effectué un pontage covalent des matrices polynucléosomales aux UV afin de permettre la formation d'un complexe stable non remodelable. Ainsi, nous avons pu dans un second temps coupler l'intégration concertée du VIH-1 avec l'activité de remodelage du FACT sur ces matrices polynucléosomales pontées. Les résultats obtenus montrent que le pontage des matrices polynucléosomales bloque la restauration de l'activité de l'intégration par le complexe FACT (**Figure 70**). Ces résultats confirment que la restauration de l'activité de l'intégration est liée à l'activité de remodelage du nucléosome par le FACT.

a



b

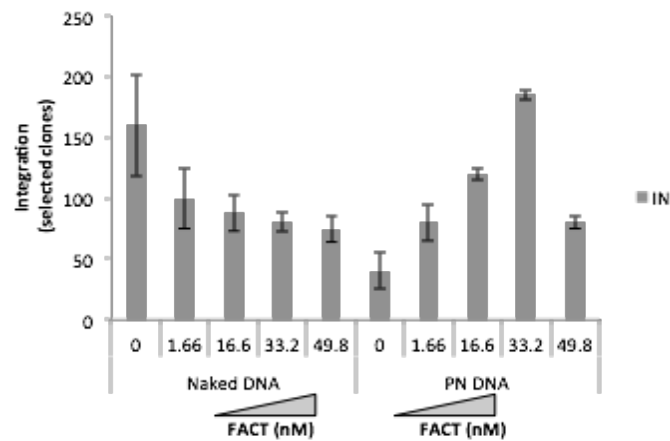


Figure 69 : Effet du complexe FACT sur l'intégration concertée du VIH-1 (a) Gel représentatif et quantification de l'effet du FACT recombinant en intégration concertée *in vitro*. Les résultats sont exprimés en pourcentage par rapport à l'intégration sans traitement par le FACT. Piste 1 et 5 : sans traitement, piste 2 et 6 : 1,66 nM FACT, piste 3 et 7 : 16,6 nM FACT, piste 4 et 8 : 33,2 nM FACT (b) Quantification des formes d'intégration FSI circulaire par transformation bactérienne puis sélection. Les résultats sont représentés en pourcentage par rapport à la condition sans traitement et sont issus de la moyenne d'au moins trois expériences $\pm$ écart type.

4. 2. Étude de l'intégration sur nucléosome tétramérique assemblé *in vitro*

L'activité de remodelage du complexe FACT est décrite comme permettant l'éviction d'un dimère H2A/H2B afin d'obtenir une structure hexamérique du nucléosome. Étant donné les propriétés réfractaires d'une chromatine compacte sur l'intégration du VIH-1 et de l'effet du complexe FACT sur la restauration de l'intégration concertée, nous nous sommes demandé si l'activation de l'intégration médiée par le FACT sur chromatine pourrait être due à une augmentation d'accessibilité locale du site due à l'éviction du dimère H2A/H2B. Pour répondre à cette question, nous avons préalablement généré au laboratoire des matrices polynucléosomales formées par une succession de nucléosomes constitués des histones H3/H4. Ces structures tétramériques du nucléosome partagent une structure proche du nucléosome dissocié (Hamiche *et al*, 1996) et constituent donc un modèle d'étude de chromatine remodelée par le FACT.

Ces matrices polynucléosomales assemblées avec les histones H3 et H4 (PN H4) ont été testées sur l'intégration du VIH-1 et comparées à l'intégration sur l'ADN nu ou sur chromatine assemblée (PN H8). Pour cela, le plasmide pBSK-601-Zeo contenant la séquence 601 permettant l'assemblage stable a été utilisé. Comme reporté en **Figure 71** l'intégration est favorisée sur les matrices nucléosomales tétramériques (PN H4) par rapport au substrat correspondant à l'ADN nu ou au substrat chromatinien octamérique (PN H8). Afin de confirmer ces résultats préliminaires de préférence de l'intégration pour les nucléosomes dissociés, nous avons effectué des expériences de sélectivité préalablement mise au point (Lesbats *et al*, 2011). Ces expériences de sélectivité consistent à réaliser l'intégration concertée sur un mélange de plasmides p5S et p601 de taille différente. De cette façon les produits d'intégration dans le p5S et p601 seront identifiés sur gel par leur taille différente. Ces plasmides seront nus ou chromatinisés avec soit des nucléosomes tétramériques (p601 PN H4), soit octamériques (p601PN H8). Les résultats illustrés en **Figure 72a** montrent que l'intégration est retrouvée préférentiellement dans l'ADN nu quand le mélange comprend la matrice polynucléosomale p601 PN H8 et l'ADN p5S nu. Au contraire, lors du mélange de l'ADN nu avec la matrice polynucléosomale formée de nucléosomes tétramériques (p601 PN H4), l'intégration est trouvée en préférence dans ce dernier (**Figure 72 b**). Ces résultats viennent donc confirmer les données obtenues

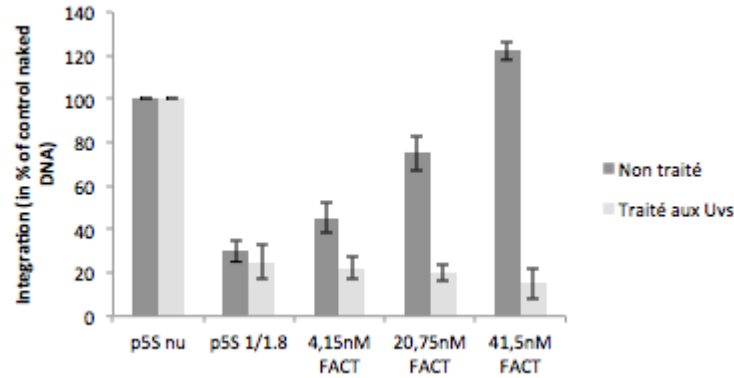


Figure 70 : Effet du complexe FACT sur l'intégration concertée utilisant des matrices pontées aux UV. Quantification de l'effet du complexe FACT en intégration concertée *in vitro* sur matrices nues ou chromatinisées pontées aux UV. Les produits d'intégration sont détectés sur gel et quantifiés par Image J. Les résultats sont représentés en pourcentage par rapport à l'intégration dans l'ADN nu et sont issus de la moyenne d'au moins trois expériences $\pm$ écart type.

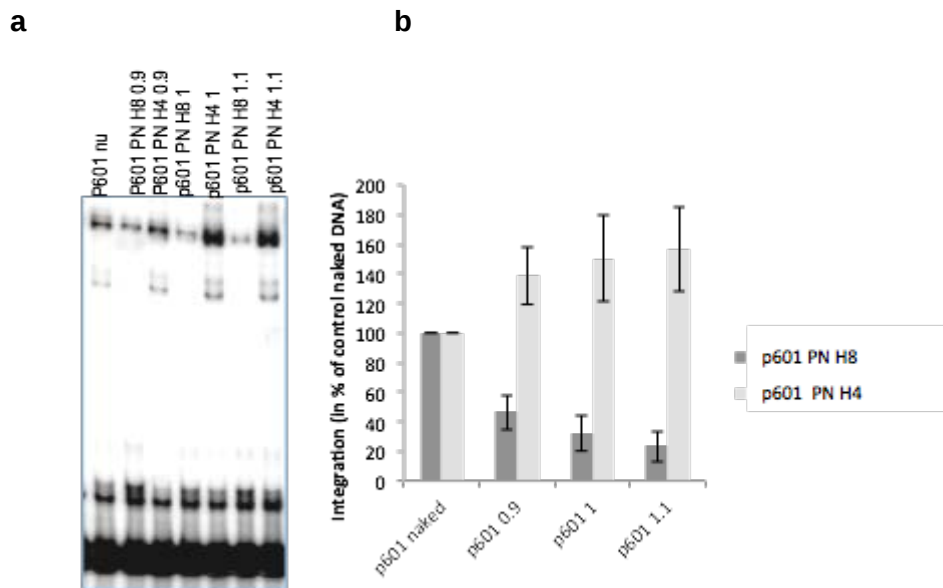


Figure 71 : Intégration concertée *in vitro* du VIH-1 sur matrices tétramériques et octamériques. (a) Gel représentatif d'un test d'intégration concertée utilisant comme receveur une matrice chromatinisée comportant des nucléosomes tétramériques (PN H4) ou octamériques (PN H8) à un ratio histone/ADN croissant. (b) Quantification des produits d'intégration obtenus en test d'intégration concertée. Les résultats sont exprimés en pourcentage par rapport à l'intégration contrôle sur l'ADN nu. Les résultats sont représentés en pourcentage par rapport à l'intégration sur ADN nu et sont issus de la moyenne d'au moins trois expériences $\pm$ écart type.

préalablement et suggèrent que l'intégration efficace du VIH-1 dans la chromatine nécessite des structures partiellement dissociées du nucléosome qui pourraient être générées par des complexes de remodelage comme le complexe FACT.

5. Conclusion et discussion

Nous avons pu grâce au développement au laboratoire d'une stratégie originale d'identification de cofacteurs cellulaires, sélectionner un ensemble de protéines possédant une interaction directe et indirecte avec le complexe IN-ADN. Étant donné les problématiques exposées précédemment, nous nous sommes focalisés sur le complexe FACT. Ce complexe appartenant à la famille des histones chaperonnes est étroitement lié à l'ARN polymérase de type II. Ce complexe est décrit comme permettant le remodelage de la chromatine via l'éjection des protéines histones H2A-H2B afin de permettre la génération d'un nucléosome hexamérique. Les données obtenues au laboratoire démontrent pour la première fois, une interaction indirecte entre l'intasome du VIH-1 et le FACT via l'association entre LEDGF/p75 et SSRP1. Dans l'objectif de valider le FACT comme cofacteur de l'intégration, nous avons suivi un algorithme de validation visant à mettre en évidence un effet *in cellulo* et *in vitro* du cofacteur identifié. C'est pourquoi, grâce à des techniques de siRNA *in cellulo* ainsi que par l'utilisation de molécules inhibitrices du FACT, les curaxines, nous avons pu démontrer un rôle du FACT dans les phases précoces de l'intégration. En effet, l'inhibition du complexe FACT *in cellulo* entraîne une augmentation globale de la quantité de chromatine ouverte dans les régions Pol II par réduction de la réassociation des nucléosomes après le passage de l'ARN polymérase de type II (non montré). Ainsi, nous avons pu montrer que l'inhibition d'expression du FACT entraîne une meilleure efficacité d'intégration.

Nous avons pu valider dans un second temps un rôle du complexe FACT sur l'intégration par des tests d'intégration concertée *in vitro*. Dans ce cadre, le complexe FACT est capable de restaurer l'intégration sur matrices chromatinisées grâce à son activité de remodelage du nucléosome. De plus, des expériences menées *in vitro* utilisant des substrats nucléosomales tétramériques ont finalement permis de montrer que l'intégration du VIH-1 est favorisée sur une structure partiellement dissociée du nucléosome.

L'ensemble de ces résultats est en faveur d'un rôle du complexe du FACT dans la phase d'approche du nucléosome par l'intasome. En effet, grâce à son action de remodelage locale de la chromatine, le complexe FACT va permettre de générer des structures nucléosomales favorables à l'intégration.

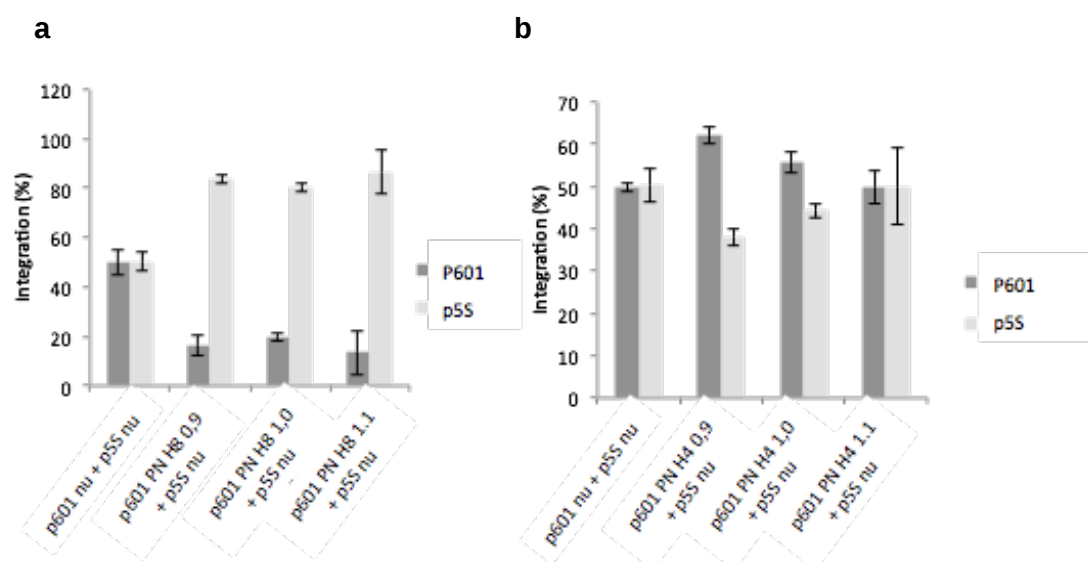


Figure 72 : Test de sélectivité d'intégration *in vitro*. (a) Quantification des produits d'intégrations issus des tests de sélectivité d'intégration dans la matrice nue (p5S nu) ou chromatinisée octamérique (p601 PN H8) à un ratio histone/ADN croissant (0,9, 1,0 et 1,1) (b) Quantification des produits d'intégrations dans la matrice nue (p5S nu) ou chromatinisée tétramérique (p601 PN H4) à un ratio histone/ADN croissant (0,9, 1,0 et 1,1). Les résultats sont représentés en pourcentage par rapport à l'intégration sur ADN nu et sont issus de la moyenne d'au moins trois expériences $\pm$ écart type.

Étant donné que ce remodelage de la chromatine peut potentiellement permettre de rendre l'ADN accessible à d'autres partenaires protéiques de l'intasome du VIH-1, nous nous sommes donc ensuite orientés vers l'étude de ces partenaires lors du contact entre l'intasome du VIH-1 avec le nucléosome.

Partie 3 : Interaction fonctionnelle entre l'IN du VIH-1 et le peptide d'histone H4K20me1

Ce travail a donné lieu à une publication en copremier auteur, en cours de révision à ***PLOS Pathogens*** « *HIV-1 integrase hosts an ancestral chromatin reader domain that binds the human histone H4 tail and is required for functional association with nucleosome and efficient integration* ». Ce manuscrit apporte la démonstration d'une nouvelle interaction hôte-pathogène et montre pour la première fois la présence d'un motif de lecture de la chromatine dans une intégrase rétrovirale participant au choix du site d'intégration.

1. Objectifs

Comme indiqué dans les chapitres précédents, le choix du site d'intégration rétrovirale dans la chromatine n'est pas un processus aléatoire et fait intervenir de nombreux facteurs (voir Chapitre 2. 5.). La sélectivité de l'intégration est un processus déterminant pour l'expression des gènes viraux. En effet, l'expression des gènes va varier en fonction de la localisation du site d'intégration dans la chromatine. Dans le cas du VIH-1 l'intégration est favorisée dans les unités de transcription active. Ce ciblage vers les régions actives repose entre autres sur l'interaction du domaine PWWP de LEDGF/p75 avec la modification d'histone H3K36me3 permettant la mise à proximité de l'intasome du VIH-1 avec les nucléosomes portant cette modification. Cependant, les données de la littérature montrent qu'une déplétion des facteurs de ciblage tels que LEDGF/p75 dans la cellule ne peut à elle seule expliquer ce ciblage vers les régions transcrites du génome suggérant alors d'autres mécanismes de ciblage de l'intasome du VIH-1 vers ces régions. Par ailleurs nos études précédentes montrent que dans le cas du VIH-1 l'intégration dans la chromatine peut nécessiter des interactions supplémentaires entre l'intasome et le nucléosome. Étant donné qu'une interaction entre l'IN du VIH-1 et le nucléosome est nécessaire à une intégration efficace dans la cellule nous nous sommes demandés si l'IN possédait des propriétés intrinsèques de liaison aux protéines formant le nucléosome. Pour cela nous avons analysé la capacité de l'IN du VIH-1 à se lier spécifiquement aux queues d'histones, candidats privilégiés à des interactions avec l'intasome approchant la chromatine.

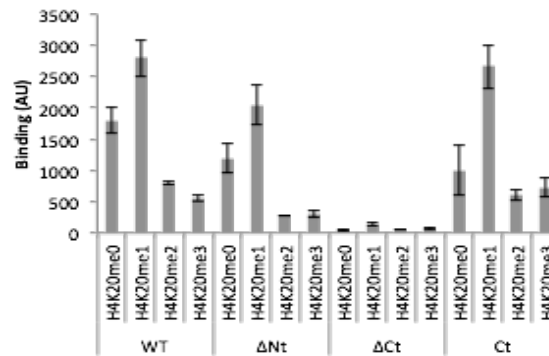
2. Interaction entre l'IN du VIH-1 et la chromatine

2. 1. Interaction entre l'IN du VIH-1 et les queues d'histones

Afin d'analyser l'interaction potentielle entre l'IN du VIH-1 et les queues d'histones, nous avons criblé un ensemble de 384 combinaisons de modifications d'histones issues des queues d'histones H2A, H2B, H3 et H4 avec l'IN du VIH-1 (**Figure 73a**). Grâce à cette approche globale, nous avons pu identifier une interaction de l'IN du VIH-1 avec la queue d'histone H4 (résidus 11 à 30). L'analyse quantitative de l'interaction entre l'IN et la queue H4 montre que cette interaction peut être modulée par des modifications post-traductionnelles. Le calcul du facteur de spécificité (Rapport entre l'intensité moyenne de l'ensemble des points contenant la marque et l'intensité moyenne des autres points ne la contenant pas) de l'interaction de l'IN du VIH-1 avec le peptide H4 monométhylé sur la lysine 20 (H4K20me1) est de 137,25 alors que les modifications H4R19me2a (2^{eme}) et H4K16 (3^{eme}) possèdent un facteur de spécificité de 19,13 et 13,44 respectivement (**Figure 73 b**). De façon intéressante, la quantification de l'interaction de l'IN avec les peptides H4K20me0, H4K20me1, H4K20me2 et H4K20me3 suggère une modulation de l'interaction en fonction de l'état de méthylation de la queue H4 sur ce résidu (**figure 73c**). De plus, ce criblage suggère que des modifications secondaires associées à la modification primaire comme l'acétylation sur la lysine 16 (H4K16acK20me1) et la diméthylation de l'arginine 19 (H4K19me2AK20me1) permettent d'augmenter l'affinité de l'IN du VIH-1 à la queue d'histone H4.

Étant donné que l'IN est complexée *in vivo* avec son cofacteur de ciblage LEDGF/p75, nous avons également testé la capacité du complexe IN-LEDGF/p75 à reconnaître ces différentes modifications d'histones en utilisant la même technique de criblage global. L'ensemble des résultats montre que LEDGF/p75 n'affecte ni fixation de l'IN à la queue d'histone H4 ni sa spécificité pour la queue H4K20me1 (non montré). Ces résultats suggèrent donc que les interactions IN-H4 et IN-LEDGF/p75 ne sont pas exclusives et que l'interaction entre l'IN et la queue d'histone H4 peut avoir lieu dans le contexte du complexe physiologique IN-LEDGF/p75.

a



b

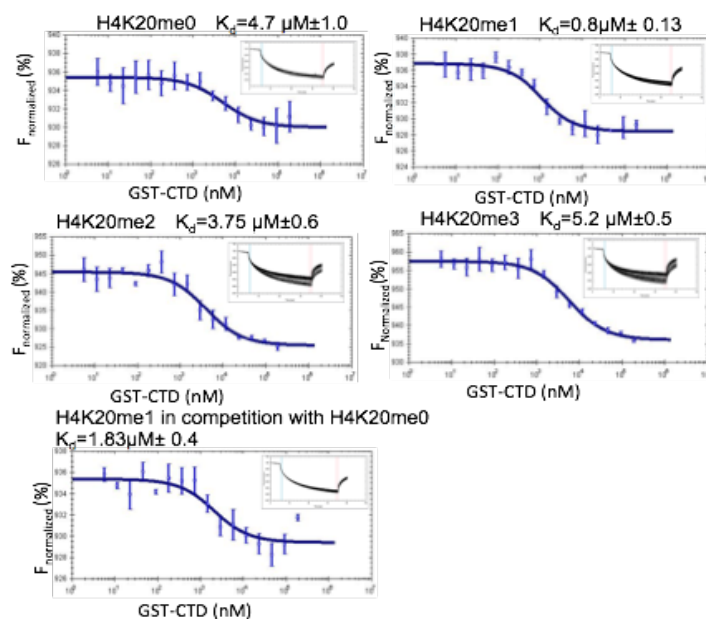


Figure 74 : Identification du domaine de l'IN responsable de l'interaction avec le peptide H4K20me1. (a) Des expériences de far dot blot ont été effectuées afin d'identifier le domaine de l'IN du VIH-1 responsable de la reconnaissance de la queue d'histone H4. Pour cela, des quantités croissantes de protéines entières (WT), dépourvues du domaine NTD (ΔNt) ou CTD (ΔCt) ou du CTD isolé (Ct) ont été immobilisées et incubées avec les peptides H4K20me0, H4K20me1, K4K20me2 et H4K20me3. L'interaction a été quantifiée par l'ajout de streptavidine couplée à la peroxydase. Les résultats sont représentés comme la moyenne de trois à six expériences indépendantes ± écart-type (b) L'interaction entre le domaine CTD et les peptides fluorescents H4K20me1, H4K20me1, H4K20me2 et H4K20me3 a été analysée par « microscale thermophoresis ».

Les expériences de compétition ont été effectuées en utilisant le peptide H4K20me0 non marqué et le peptide H4K20me1 fluorescent. Les données représentent la fluorescence normalisée en fonction de la concentration de CTD et ont été analysées par le logiciel d'analyse NT. Les valeurs indiquées correspondent aux moyennes ± écart-type calculés pour chaque condition de trois répétitions d'expériences indépendantes en utilisant le signal de thermophorèse + T-jump. Les valeurs de K_D sont rapportées pour chaque interaction.

2. 2. Identification du domaine de l'IN du VIH-1 impliqué dans l'interaction.

Dans l'objectif d'identifier le domaine de l'IN responsable de l'interaction avec la queue d'histone H4, nous avons testé l'interaction de plusieurs constructions d'IN tronquées pour leur domaine NTD ou leur domaine CTD avec les queues d'histones H4K20me0, H4K20me1, H4K20me2 et H4K20me3 en far dot blot (**Figure 74a**). Cette approche a permis de montrer que la construction dépourvue du domaine NTD garde une interaction proche de l'enzyme WT indiquant que ce domaine ne semble pas impliqué dans l'interaction. Cependant, la construction dépourvue du domaine CTD perd la fixation aux différentes queues d'histones indiquant un rôle important du domaine CTD dans l'interaction. Ces données ont pu être confirmées par l'analyse de l'interaction du domaine CTD isolé qui garde une interaction avec les queues d'histones de façon similaire à l'enzyme WT.

Afin de quantifier la force d'interaction du domaine CTD avec les différentes queues d'histones, nous avons, en collaboration avec le laboratoire de M.Ruff (IGMC Strasbourg), effectué des analyses de « microscale thermophoresis » (MST) (**Figure 74b**). Cette technologie permet d'analyser quantitativement l'interaction de molécules grâce à la mesure de leur mouvement de molécules dans un gradient de température qui dépend leur l'enveloppe d'hydratation, de leur charge et leur taille. Pour cela, les molécules sont dans un premier temps mises en solution et diffusent passivement. Puis, lors de l'activation du laser, les molécules en complexes avec un partenaire vont connaître un changement de mouvement dans le gradient thermique visualisable par fluorescence. Cette étude a permis de montrer une forte interaction entre le CTD et le peptide H4K20me1 avec un $K_D = 0,8 \pm 0,13 \mu M$ et une interaction de plus faible affinité est montrée avec les peptides H4K20me0, H4K20me2 et H4K20me3 ($K_D = 4,7 \pm 1,0 \mu M$, $3,75 \pm 0,6 \mu M$, $5,2 \pm 0,5 \mu M$ respectivement). De plus, des expériences de compétition par addition du peptide H4K20me0 ne montrent pas d'influence sur l'interaction CTD/H4K20me1 confirmant la spécificité de l'interaction IN/H4K20me1.

2. 3. Analyse structurale et modélisation *in silico* de l'interaction

Le domaine C-terminal (CTD) est le domaine qui est le moins conservé des trois domaines de l'intégrase. Ce domaine possède un repliement de type SH3 connu pour être impliqué dans des interactions protéines/protéines. L'analyse de la structure du CTD de l'IN

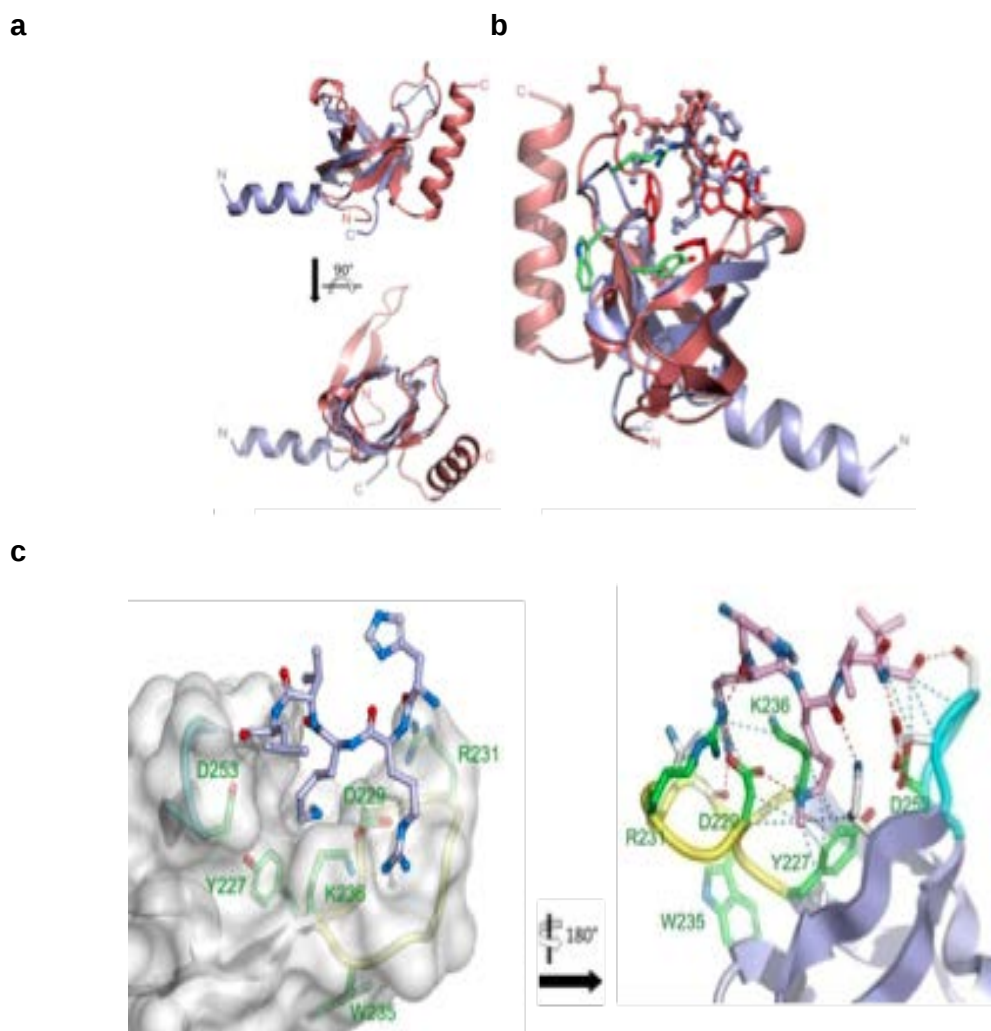


Figure 75 : Représentation de l'interaction entre l'IN du VIH-1 avec le peptide H4K20me1 par docking moléculaire. (a) Représentation en cartoon de la structure RMN du domaine CTD de l'IN du VIH-1 en bleu avec la structure cristallographique de la protéine hMSL3 en rouge dans deux orientations (b) Superposition du modèle de « docking » CTD-H4K20me1 avec la structure cristallographique de la protéine hMSL3 en complexe avec le peptide H4K20me1. Dans les deux structures, le peptide H4K20me1 est représenté en boule et bâton (c) Modèle de docking moléculaire représentant le pentapeptide H4K20me1 (boule et bâton) « docké » à la surface du CTD de l'IN du VIH-1 (surface grise). Les boucles 228-235 et 253-257 sont représentées en jaune et cyan respectivement. Les résidus Y227, R231 et W235 sont en représentation bâton et de couleur verte. Dans ce modèle, la chaîne latérale de la lysine 20 mono-méthylée pointe vers le bas dans la cavité en V définie par les boucles 228-235 et 253-257. Les liaisons hydrogènes prédites et des contacts hydrophobes sont représentés par des lignes en pointillé rouge et bleu, respectivement.

de VIH-1 (PDB : 1EX4) résidus 220 à 270 a pu être comparée avec la structure de la protéine hMSL3 (hMSL3) (PDB : 3OA6) possédant un chromo barrel domaine connu pour interagir avec le peptide H4K20me1 (voir Chapitre 3. 2. b). La superposition des structures du CTD de l'IN et de hMSL3 montre une forte homologie structurale avec un RMSD (Root-mean-square deviation) de 1,9 Å suggérant une similitude d'interaction (**Figure 75a et 75b**). Ces résultats montrent pour la première fois une homologie structurale entre l'IN du VIH-1 et le chromodomaine d'une protéine cellulaire suggérant ainsi une fonction de reconnaissance des modifications post-traductionnelles d'histones conservée entre ces protéines.

Afin de comprendre les bases moléculaires de l'interaction entre CTD et la queue d'histone H4, nous avons en collaboration avec le laboratoire de P.Gouet (MMSB Lyon), effectué une modélisation moléculaire *in silico* de l'interaction (**Figure 75c**). Cette modélisation se base sur un fragment (résidus 210 à 270) de la structure du double domaine CCD-CTD de l'IN du VIH-1 et un pentapeptide correspondant au peptide H4K20me1 de séquence primaire H₁₈RK_{me1}VL. Afin de réaliser cette analyse, deux programmes AutoDock et AutoDock Vina ont été utilisés en parallèle pour déterminer une région potentielle d'interaction dans le fragment du CTD de l'IN considéré comme rigide dans cette analyse. Cette modélisation moléculaire a pu mettre en relief deux sites d'interaction potentielle. Le premier site d'interaction présente une cavité en forme de V délimité par les boucles 228-235 et 235-257. De manière marquante, dans cette modélisation le peptide H4K20me1 possède la même localisation et orientation que dans la structure du complexe hMSL3/H4K20me1 avec un RMSD de 3,4 Å confirmant une forte similitude d'interaction entre ces complexes. Basé sur cette modélisation, nous avons déterminé les résidus de l'IN impliqués dans l'interaction avec la queue d'histone H4. Pour cela, les acides aminés formant la cavité en V ont été traités comme flexible dans le programme AutoDock (Y227, D229, S230, R231, N232, L234, W235, K236, D253, N254, S255, D256, K258 et K264). Cette analyse a permis d'identifier 56 groupes conformationnels dans lesquels le peptide H4K20me1 s'associe à la cavité en V du CTD de l'IN. Étant donné l'importance de la monométhylation et du calcul de l'énergie d'interaction, une solution a pu être sélectionnée. Dans cette modélisation, le peptide H4K20me1 engage 7 liaisons hydrogènes avec les résidus D229, R231, S255, D256, K258 et 15 liaisons hydrophobes avec les résidus Y227, D229, N232, K236, D256, K258 et V260. De plus, dans ce modèle la monométhylation de la lysine 20 du peptide pointe dans la cavité en V et engage à elle seule 7 des 15 liaisons hydrophobes et 1 des 7 liaisons hydrogènes démontrant l'importance de cette modification dans la régulation de l'interaction.

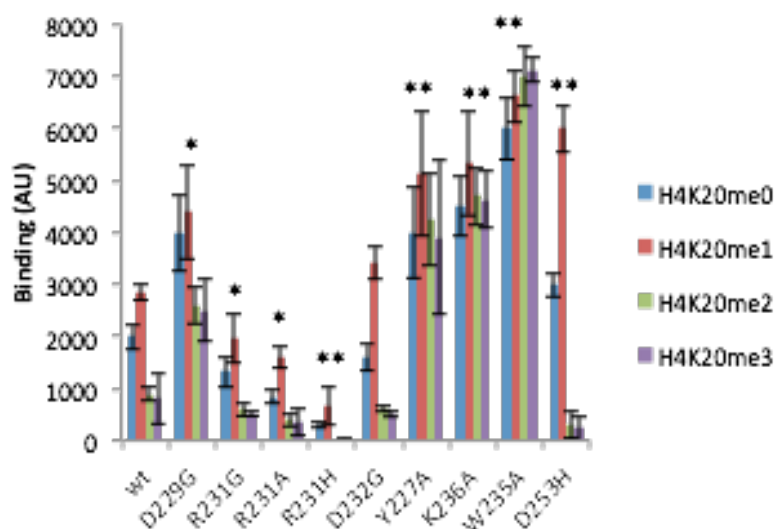


Figure 76 : Détermination des acides aminés du domaine CTD de l'IN du VIH-1 impliqués dans l'interaction avec la queue H4. Des mutations ponctuelles ont été introduites dans l'IN du VIH-1 au niveau des résidus potentiellement impliqués dans l'interaction avec le peptide H4K20me1. Puis, l'interaction des mutants avec la queue H4 et ses variants de méthylation a été analysée par des expériences de far dot blot. Les résultats sont représentés comme la moyenne de trois à dix expériences indépendantes $\pm$ écart-type. Les valeurs p ont été calculées par le test t de Student et sont représentées par * $p < 0,05$ et ** $p < 0,005$ pour indiquer la probabilité d'obtenir des différences significatives par rapport aux résultats obtenus avec l'enzyme WT.

2. 4. Identification des acides aminés impliqués dans l'interaction avec la queue H4

Afin de pouvoir valider la modélisation *in silico* et ainsi confirmer l'implication des acides aminés impliqués dans l'interaction avec la queue d'histone H4, nous avons développé une approche de mutagenèse dirigée. La mutagenèse dirigée se base principalement sur l'homologie structurale entre le CTD de l'IN du VIH-1 et hMSL3 et sur les acides aminés responsables des différentes liaisons hydrogènes et hydrophobes modélisés *in silico*. Ainsi, les résidus Y227, D229, R231, W235, K236 et D253 ont pu être mutés puisque, le Y227 forme le plancher de la cavité en V et pourrait mimer la phénylalanine retrouvée dans la cage aromatique des domaines de la « Royal family », le D229 et le K236, car ces résidus engagent des liaisons avec les résidus K20 et R19 du pentapeptide et le R231, car il engage des liaisons hydrogènes avec le résidu H18 du peptide. Les résidus W235 et D253 ont été mutés, car ils sont localisés dans les boucles 228-236 et 253-257 respectivement et peuvent participer à l'architecture de la cavité en V. De plus, le résidu D253 superpose avec le résidu F56 chez hMSL3 et participe à la structuration de la boucle 253-257 de la cavité. Ainsi, les mutations sur l'ensemble de ces six positions devraient, si ces résidus sont impliqués dans l'interaction, affecter l'affinité en augmentant ou, en inhibant l'interaction avec la queue d'histone H4.

Les mutants ont été purifiés et leur structure fonctionnelle a été vérifiée par quantification de leur activité dans un test d'intégration *in vitro*. La quantification de l'hétérointégration montre que les mutants Y227A et W235A sont affectés de manière importante (70/90 % pertes d'activité) et les mutants K236A et D229G montrent une faible perte d'activité (20/40 %) alors que les mutants R231A/H/G et D253H sont retrouvés aussi actifs que l'IN WT (non montré). Dans le but d'identifier les acides aminés impliqués dans l'interaction avec la queue d'histone, nous avons analysé l'interaction de ces mutants avec la queue d'histone H4 et ses variants méthylés dans des tests de « far dot blot » (**Figure 76**). Les mutants R231A et R231H montrent une baisse d'affinité pour la queue H4 avec une inhibition de près de 80 % pour le R231H. Le mutant R231G reste moins affecté que les autres mutants avec un phénotype intermédiaire par rapport au sauvage (baisse de 10-20 % de l'interaction) ce qui constitue un résultat logique étant donné qu'il s'agit d'un polymorphe naturel. Le second polymorphe naturel D232G conserve une bonne affinité pour la queue d'histone. Le mutant D229G montre quant à lui une augmentation d'affinité à la queue H4 par rapport à l'enzyme WT. Les mutants Y227A, W235A et K236A montrent aussi

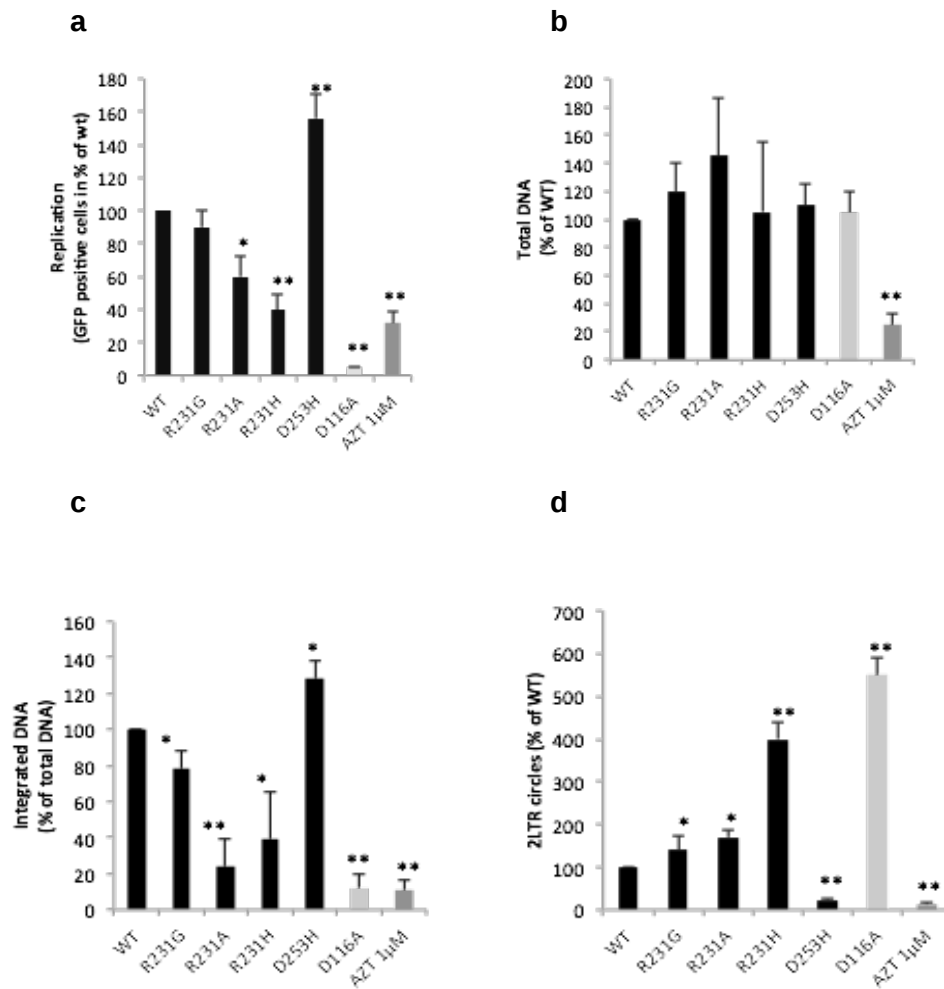


Figure 77 : Effet des mutations impliquées dans l'interaction avec le peptide H4K20me1 sur les phases précoces de la réplication virale. (a) Les cellules 293 T sont infectées avec des vecteurs lentiviraux pseudotypés VSV-G codant pour une IN sauvage ou portant les différentes mutations R231A/H/G, D253H ou le mutant de classe I catalytiquement inactif D116A. La réplication virale a été quantifiée sur le pourcentage de cellule eGFP positive mesuré par cytométrie en flux à une MOI de 4. Les étapes de réplication affectées par les mutations ont été déterminées en mesurant les quantités des différentes espèces d'ADN viral produites par qPCR. Ainsi la quantité d'ADN total (b), d'ADN intégré (c) et de cercles à 2 LTR (d) ont été contrôlés pour vérifier les défauts potentiels aux étapes de transcription inverse, d'intégration et d'import nucléaire du complexe de pré-intégration, respectivement. La quantité de chaque espèce d'ADN viral produite par les pseudovirus WT a été fixée à 100 %. Les données sont représentées comme la moyenne d'au moins trois expériences indépendantes $\pm$ écart-type. Les résultats obtenus dans des conditions de traitement des cellules avec l'inhibiteur de la transcription inverse AZT (1 μ M) est également signalé. Les valeurs p ont été calculées par le test t de Student et sont représentés par * p < 0,05 et ** p < 0,005.

une augmentation d'affinité pour la queue H4, mais ces mutants sont aussi accompagnés d'une perte de spécificité vis-à-vis de la monométhylation de la queue H4. Finalement, le D253H montre une augmentation de l'affinité pour la monométhylation de la queue H4 cependant, son affinité pour les autres variants n'est pas modifiée. Ces résultats suggèrent que l'ensemble des mutants à l'exception du polymorphe naturel D232G est affecté dans la fixation au peptide H4K20me1 et peut être aussi accompagné dans certains cas d'une perte de spécificité permettant de valider pour la première fois la présence d'un site de fixation du peptide H4K20me1 dans le domaine CTD de l'IN du VIH-1.

2. 5. Impact des mutations impliquées dans l'interaction sur la réplication et l'intégration *in vivo*

Dans le but d'évaluer l'importance de l'interaction IN/H4 dans le cycle réplcatif viral, nous avons comparé l'efficacité de réplication de vecteurs lentiviraux portant les mutations impliquées dans l'interaction avec la queue H4 *in vitro* sans affecter l'activité catalytique de l'IN et sa fixation à l'ADN *in vitro*. Pour cela, nous avons sélectionné les mutations R231A/H/G et D253H qui ont été introduites dans la séquence de l'IN du vecteur lentiviral et leur infectivité a été comparé au WT dans un « single round infection assay » dans plusieurs types cellulaires (HEK-293T, HeLa P4 et K562) (**Figure 77a**). Cette analyse a permis de montrer que les mutants R231G, R231A et R231H présentent une diminution de l'infectivité virale allant de 10-20 % pour le R231G à 40-60 % pour les mutants R231A/H. Au contraire, une augmentation de l'infectivité virale est observée pour le mutant D253H suggérant une corrélation directe entre capacité de liaison à la queue H4 et efficacité de réplication.

Afin de savoir quelle(s) étape(s) des phases précoces de la réplication virale était (en) t affectée(s) par ces mutations, les populations d'ADNc, d'ADN intégré et des cercles à 2 LTR pour chaque mutant ont pu être quantifiées. Cette analyse montre que la production d'ADNc n'est pas affectée ne suggérant aucune altération significative dans l'étape de transcription inverse de l'ARN viral en ADN pour l'ensemble des mutants d'intérêt (**Figure 77b**). Cependant, la quantification de l'ADN intégré et des cercles à 2 LTR montrent des différences majeures (**Figure 77c et 77d**). En effet, on observe chez le mutant R231A et R231H une baisse de l'ADN intégré de 60 et 80 % respectivement avec une accumulation

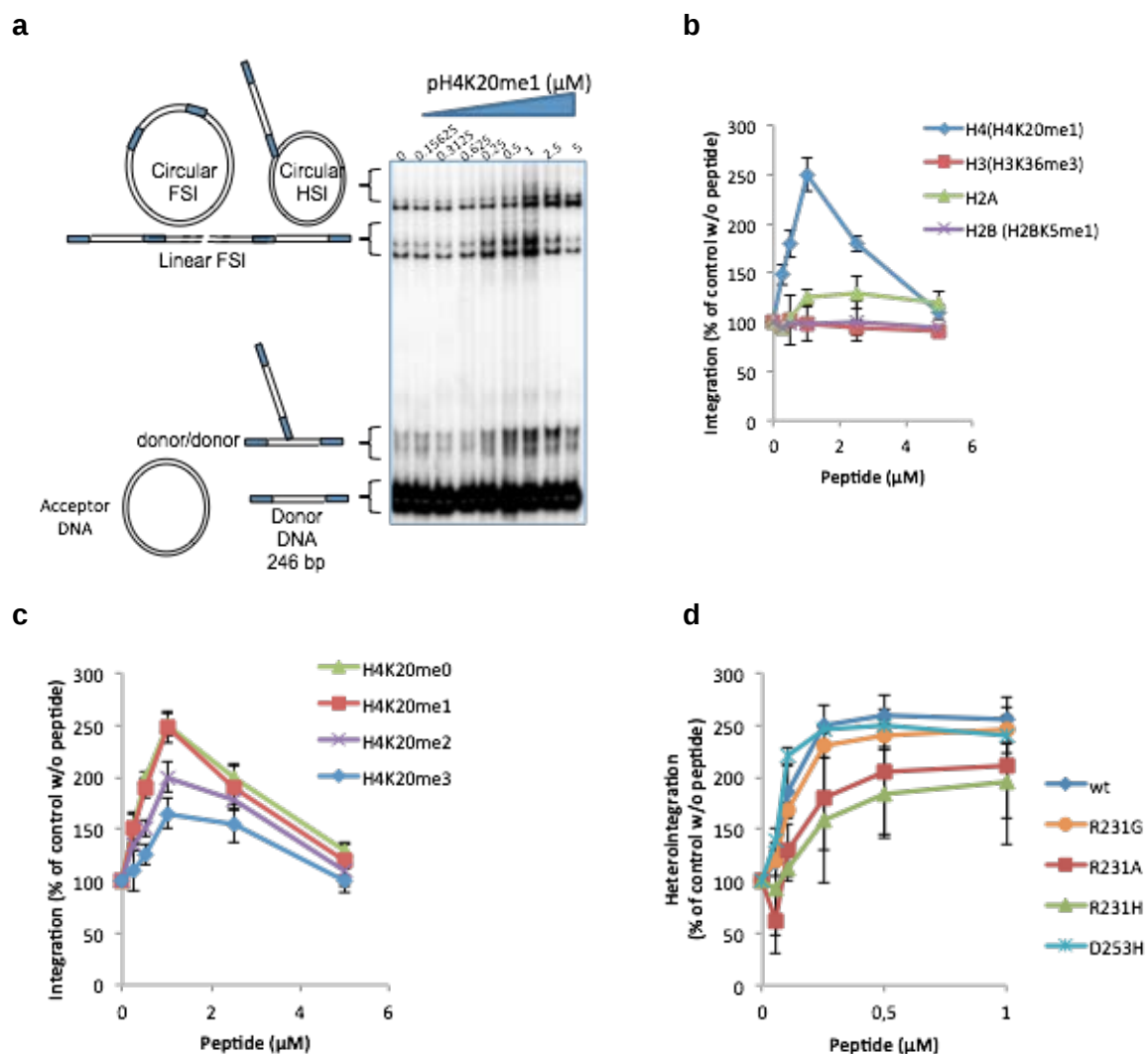


Figure 78 : Effet des IN mutantes et des peptides issus de l'histone H4 sur l'intégration du VIH-1 *in vitro*. (a) Gel représentatif d'un test d'intégration concertée *in vitro* utilisé afin d'analyser l'effet d'une gamme croissante de peptide dérivé des histones H4, H3, H2A et H2B sur l'enzyme WT (b). En (c) effet des peptides H4 non méthylés, mono di et tri méthylés sur l'intégration. (d) Effet d'une gamme croissante de peptide H4K20me1 sur l'intégration des IN WT et mutantes. La quantification des produits d'intégration (HSI, FSI et de l'auto-intégration d/d) a été réalisée. Les résultats sont représentés en pourcentage par rapport à l'activité de l'enzyme WT sans peptide. Toutes les valeurs sont montrées comme la moyenne $\pm$ écart-type (barres d'erreur) d'au moins trois séries d'expériences indépendantes.

des cercles à 2 LTR. Un résultat similaire est obtenu pour le mutant R231G restant cependant le plus proche du WT. Le mutant D253H montre quant à lui un phénotype opposé avec une augmentation de 20 à 30 % de l'ADN intégré associé avec une diminution des cercles à 2 LTR suggérant une meilleure efficacité de l'étape d'intégration pour ce vecteur.

L'ensemble des résultats obtenus *in vitro* sur la capacité des différents mutants à interagir avec la queue H4 et *in vivo* sur l'analyse de la réplication suggèrent une relation directe entre la capacité de ces mutants à interagir avec la queue H4 et leur intégration efficace dans la chromatine. Ceci suggère un rôle de cette interaction dans le processus d'intégration.

3. Étude fonctionnelle de l'interaction IN-queue H4

3. 1. Impact de l'interaction avec la queue d'histone H4 sur l'intégration *in vitro*

Afin de comprendre le rôle de l'interaction avec la queue H4 sur l'étape d'intégration, nous avons analysé préalablement l'effet du peptide H4 et ses variants sur l'activité d'intégration dans un test d'intégration concertée *in vitro* (**Figure 78a et 78d**). Nous avons tout d'abord évalué l'effet du peptide sur un substrat de type plasmide nu. La quantification montre que l'ajout du peptide H4K20me1 (résidus 1 à 23) induit une forte stimulation de l'intégration catalysée par l'IN WT d'un facteur 2 à 2,5. Au contraire, les mutants R231A/H/G possédant une affinité diminuée pour le peptide H4K20me1 sont aussi moins sensibles à la stimulation par le peptide H4K20me1 suggérant que la stimulation de l'activité de l'IN induit par le peptide requiert une interaction entre l'IN et le peptide. Les protéines gardant une affinité forte pour le peptide H4K20me1 comme le mutant D253H sont aussi stimulées. Des expériences complémentaires utilisant les variants de méthylation du peptide H4 (H4K20me0, H4K20me2 et H4K20me3) en intégration concertée confirment la modulation de la stimulation par l'état de méthylation du peptide H4 (**Figure 78c**). De plus, les tests d'intégration utilisant des peptides issus d'autres histones (H2A, H2B et H3) ne montrent pas d'effet significatif sur l'intégration concertée suggérant une spécificité du peptide méthylé de la queue H4 (**Figure 78b**).

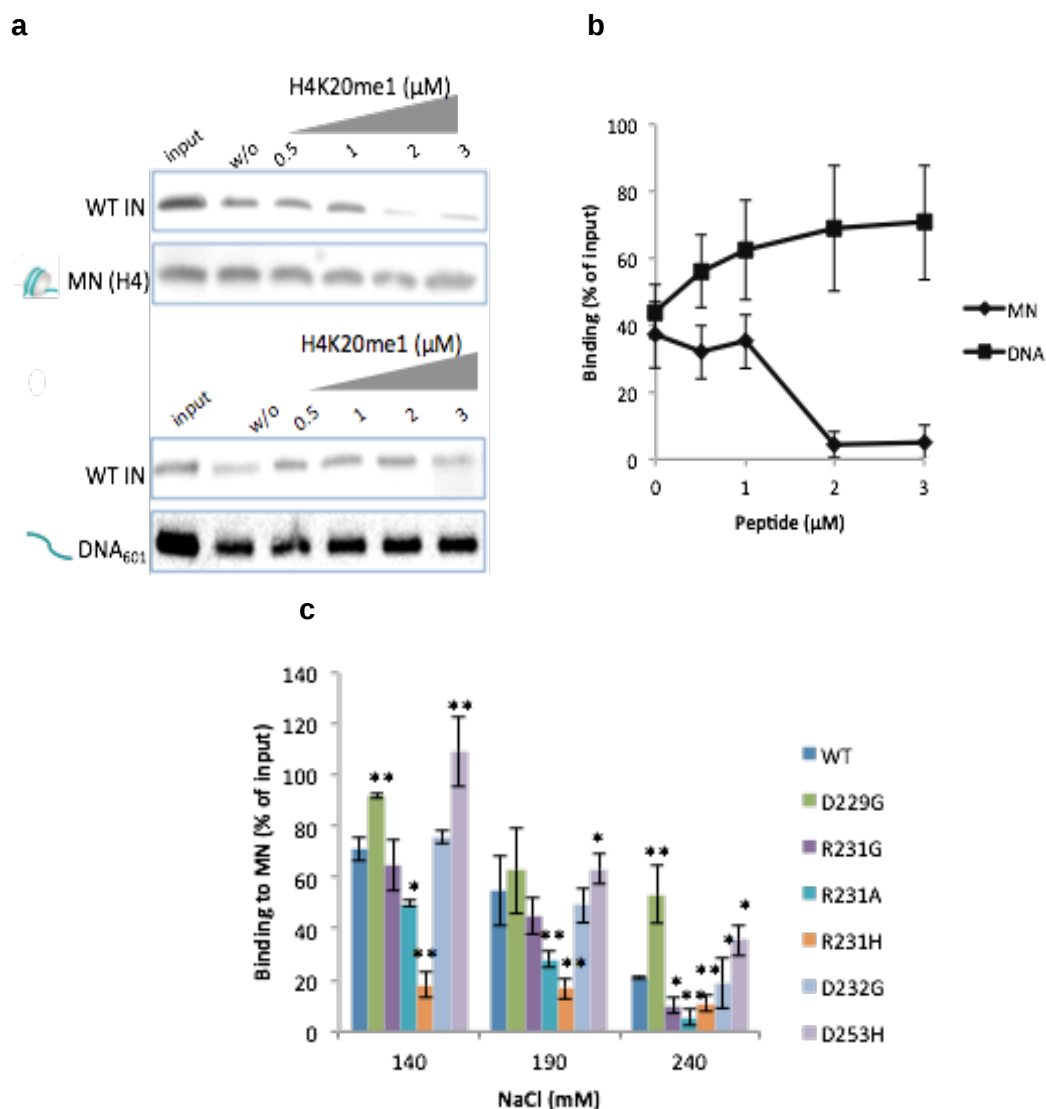


Figure 79 : Effet du peptide H4K20me1 sur la fixation de l'IN au mononucléosome. (a) Des expériences de co-précipitations utilisant l'IN WT et des mononucléosomes recombinants sur séquence 601 ou sur ADN 601 nu en présence d'une gamme croissante de peptide H4K20me1 ont été effectuées. La détection de l'IN et de l'histone H4 se fait par Western-blot. L'ADN nucléosomal est détecté par coloration au SYBR®Safe. (b) La quantification de la quantité d'IN est représentée en pourcentage par rapport à la charge initiale d'IN co-précipitée sans ajout de peptide. (c) Expériences de co-précipitations utilisant l'IN WT et les mutants d'interaction ainsi que des mononucléosomes recombinants sur séquence 601 à 140, 190 et 240 mM NaCl. Toutes les valeurs sont montrées comme la moyenne $\pm$ écart-type (barres d'erreur) de trois à six séries d'expériences indépendantes. Les valeurs p ont été calculées par le test t de Student et sont représentés par * p < 0,05 et ** p < 0,005.

3. 2. Impact des mutants d'interaction pour la queue H4 sur l'interaction avec le mononucléosome

Afin de savoir si l'interaction entre l'IN du VIH-1 et le peptide H4 pouvait jouer un rôle dans l'association fonctionnelle de l'IN avec le mononucléosome, nous avons testé la capacité du peptide d'histone H4 à entrer en compétition avec le mononucléosome. Pour cela, nous avons réalisé des expériences de co-précipitation en utilisant l'IN du VIH-1 et des mononucléosomes assemblés sur la séquence W601 de 147 paires de bases biotinylée en 5'. La séquence d'ADN correspondante nue sert de contrôle pour l'interaction à l'ADN. Les expériences de compétition réalisées par l'ajout du peptide H4K20me1 entraînent une forte diminution de l'affinité de l'IN pour le mononucléosome (**Figure 79a et 79b**). Au contraire, l'ajout du peptide n'inhibe pas la fixation de l'IN pour la séquence d'ADN contrôle W601 suggérant que l'effet compétitif observé n'est pas lié à un effet du peptide sur l'interaction entre l'IN et l'ADN, mais bien avec la queue d'histone.

Dans l'objectif de confirmer ces résultats, nous avons analysé la capacité de fixation aux mononucléosomes des mutants d'IN précédemment décrits (**Figure 79c**). La figure montre que les mutants Y227A, W235A et K236A perdent leur l'affinité pour l'ADN qui peut être corrélée avec leur faible activité en intégration *in vitro*. Ils seront donc exclus de l'étude pour éviter tout biais dans l'analyse d'affinité de l'IN pour le mononucléosome. Les mutants de l'IN ne montrant pas de défaut d'affinité pour l'ADN ont pu être analysés pour leur affinité aux mononucléosomes par co-immunoprécipitation. Ainsi, les mutants R231A et R231H décrits préalablement pour leur défaut d'affinité avec le peptide de la queue H4 montrent une baisse importante d'affinité pour le mononucléosome. Le mutant R231G est aussi affecté, mais dans une moindre mesure, ce qui est également corrélable avec sa baisse d'affinité moindre pour la queue d'histone. À l'opposé, les mutants D229G et D253H qui possèdent une fixation plus importante au peptide H4 par rapport à l'IN WT montrent une augmentation d'affinité pour le nucléosome. Le mutant D232G ne montre pas quant à lui de changement significatif d'affinité pour le nucléosome. Ces résultats sont compatibles avec un rôle de l'interaction IN avec le peptide H4 dans l'association final avec le mononucléosome.

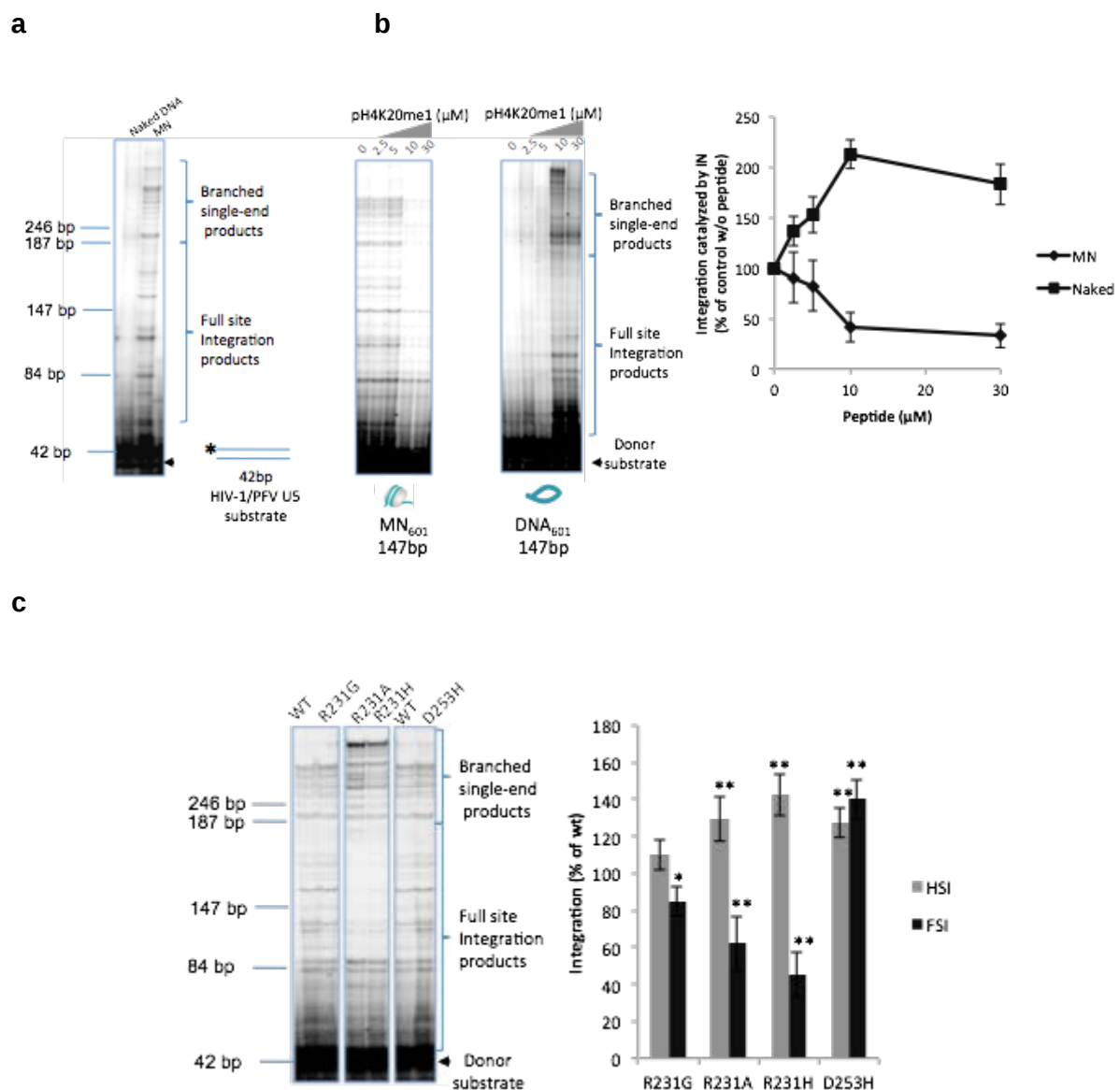


Figure 80 : Impact du peptide H4K20me1 sur l'intégration sur mononucléosome du VIH-1. (a) Intégration du VIH-1 sur ADN 601 nu de 147 pb ou sur mononucléosome recombinant comportant la séquence W601 et incubé avec un ADN donneur radiomarqué en 5' de 42 pb (b) Ajout du peptide H4K20me1 sur l'intégration sur mononucléosome (MN601) et sur l'ADN nu (ADN601) (c) Quantification de l'intégration sur mononucléosome de l'IN WT sur MN601 et ADN601 (d) Intégration sur mononucléosome des mutants de l'IN impliqués dans l'interaction avec le peptide H4K20me1 (e) Quantification de l'intégration HSI et FSI sur mononucléosome des mutants R231G, R231A, R231H et D253H par ImageJ. Toutes les valeurs sont montrées comme la moyenne $\pm$ écart-type (barres d'erreur) de trois séries d'expériences indépendantes. Les valeurs p ont été calculées par le test t de Student et sont représentées par * $p < 0,05$ et ** $p < 0,005$.

3. 3. Impact des mutants d'interaction pour la queue H4 sur l'intégration dans le mononucléosome

Étant donné que l'intégration sur un substrat polynucléosomal est une structure réfractaire à l'intégration, nous avons analysé le rôle de l'interaction IN/H4 sur l'intégration catalysée par l'IN du VIH-1 sur mononucléosome. L'intégration sur MN *in vitro* a pu être mise au point au laboratoire en suivant les méthodes décrites dans la littérature (Maskell *et al*, 2015). Dans ce test d'intégration sur mononucléosome, les produits d'intégration half-site (HSI) correspondent à des produits d'intégration d'une seule extrémité virale dans l'ADN nucléosomale et sont visualisables sur gel à une taille supérieure à 147 + 42 pb (taille de l'ADN 601 nucléosomal et de l'ADN donneur viral). Les produits d'intégration full-site (FSI) correspondent à l'intégration de deux extrémités de deux molécules différentes dans l'ADN de 147 pb et aboutissent après coupure à deux produits, l'un de X+42 pb et l'autre de ([147-X]-5) +42 pb (X correspondant au site de coupure dans l'ADN cible et 5 à la taille entre les deux sites de coupure). La quantification des produits d'intégration montre que l'intégration du VIH-1 est stimulée sur un substrat mononucléosomal par rapport à l'ADN nu confirmant que le MN isolé est un substrat préférentiel pour l'intégration (**Figure 80a**). Nous avons ensuite pu ajouter le peptide H4K20me1 afin d'effectuer des expériences de compétition dans ces tests d'intégration sur mononucléosome (**Figure 80b**). Ainsi, comme démontré préalablement en test d'intégration utilisant des substrats plasmidiques, l'ajout du peptide H4K20me1 montre un effet stimulateur de l'intégration sur l'ADN W601 linéaire nu. Au contraire, l'ajout du peptide H4K20me1 montre un effet inhibiteur de l'intégration sur mononucléosome. Nous avons ensuite pu tester l'activité des mutants de l'IN afin d'analyser l'impact des mutations affectant l'interaction avec la queue H4 (**Figure 80c**). Les mutants R231A et R231H montrent un profil d'intégration différent de l'enzyme WT. En effet, ces mutants catalysent les produits HSI plus efficacement que les produits d'intégration FSI en comparaison avec l'enzyme WT. Le mutant R231G reste de nouveau moins affecté avec cependant une baisse de 10 à 20 % de l'activité d'intégration sur MN. Au contraire, le mutant D253H montre une activité de 20 à 30 % plus importante que l'enzyme WT que l'on peut corrélérer avec son augmentation de liaison à la queue d'histone et au nucléosome. L'ensemble de ces données suggèrent donc que la fixation des IN à la queue H4 permet la modulation de l'intégration sur un substrat physiologique nucléosomal confirmant ainsi l'importance de cette association pour une intégration optimale.

4. Conclusion et discussion

L'utilisation de multiples approches complémentaires nous ont permis de démontrer pour la première fois que l'IN du VIH-1 est capable de se lier spécifiquement à la queue de l'histone humaine H4. En effet, les études de docking moléculaire, de mutagenèse dirigée et l'analyse biochimique cette interaction nous ont permis de confirmer la présence d'un site de fixation à la queue d'histone H4 dans le domaine CTD de l'IN du VIH-1. De plus, nous avons pu déterminer plusieurs acides aminés impliqués dans la reconnaissance et la fixation de la queue d'histone H4. Ainsi, les acides aminés Y227, D229, R231, K236 et D253 structurent le site de fixation responsable de l'interaction avec le peptide issu de la queue H4 dans les expériences de docking moléculaire. Ce site est localisé dans une cavité en V qui mime le site de fixation de ligand méthylé retrouvé dans le domaine formant la « Royal family » incluant les chromo, Tudor, MBT, chromo barrel et PWWP domaines. De façon marquante, parmi les protéines possédant un domaine de reconnaissance des ligands méthylés, la protéine hMSL3 a été décrite préalablement pour reconnaître le peptide H4K20me1 (Kim *et al*, 2010). Nous avons pu montrer que le chromodomaine de hMSL3 possède un fort degré d'homologie structurale avec le domaine CTD de l'IN du VIH-1 qui présente de la même façon une préférence marquée pour le peptide H4K20me1. De plus, le modèle de docking permet d'expliquer l'influence de l'état de méthylation de la lysine en position 20 du peptide sur l'interaction avec l'IN du VIH-1. En effet, un changement d'affinité peut être attendu lors du retrait ou du remplacement du groupement méthyle par un groupement di ou triméthyle dû à la perte de liaisons de stabilisation ou à un réarrangement structural énergétiquement défavorable au sein de la cavité en V mise en évidence.

Étant donné les effets des mutations introduites dans le domaine CTD de l'IN du VIH-1, le rôle des acides aminés Y227, R231, K236 et W235 dans la reconnaissance du peptide H4K20me1 a pu être identifié grâce à la comparaison avec la structure de hMSL3 en complexe avec le peptide H4K20me1. En effet, le résidu K26 de hMSL3 occupe une position similaire au résidu R231 du VIH-1. Le résidu R231 est responsable dans notre modèle d'une liaison hydrogène avec le peptide H4K20me1 et par analogie, il pourrait être impliqué dans la stabilisation de la boucle 228-235 afin de la maintenir dans une conformation favorable à la fixation du peptide H4K20me1. Les mutations sur les résidus Y227 formant le plancher de la cavité et W235 de la boucle 228-235 pourraient affecter sa position et élargir la cavité en V. En effet, les mutations introduites Y227A et W235A mènent à une perte de spécificité vis-à-vis du peptide H4K20me1 et à un défaut catalytique majeur. Le résidu 253 analogue du résidu F56 de hMSL3 participe à la structuration de la boucle 253-257. La mutation 253H

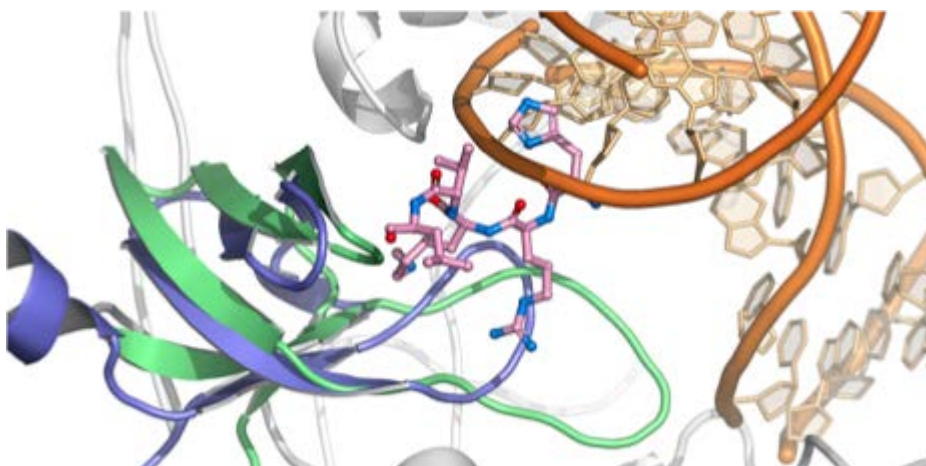


Figure 81 : Superposition du modèle de docking moléculaire H4K20me1/CTD avec le complexe de transfert de brin de PFV. Représentation en cartoon du domaine CTD de VIH-1 et PFV représentés en bleu et en vert respectivement. La partie en gris correspond à la partie restante des protomères du complexe de transfert de brin de PFV (PDB : 3OS0). L'ADN nucléosomal est en orange et le peptide H4K20me1 comme en figure 75.

introduite montre une augmentation d'affinité pour le peptide H4K20me1 suggérant que ce résidu permet de réguler l'accès au site de fixation. Ainsi, le modèle de docking moléculaire et l'ensemble des données biochimiques suggèrent que les boucles 228-235 et 253-257 permettent de réguler l'accès à la cavité en V alors que le plancher de la cavité formée par le résidu Y227 permet la stabilisation de l'interaction.

L'analyse fonctionnelle a montré que les mutations empêchant l'association IN-queue H4, altèrent également la réplication virale au niveau de l'étape d'intégration. Ainsi, le phénotype observé est caractéristique de mutations de classe II reflétant une déficience d'interaction de l'IN avec un cofacteur cellulaire. Étant donné que les mutants de l'IN sont toujours capables d'interagir avec le cofacteur cellulaire LEDGF/p75, nous proposons que la perte d'interaction entre l'IN et la queue d'histone H4 seraient responsables du défaut d'intégration observé.

Ces données sont confirmées par l'analyse mécanistique de l'interaction IN/queue H4 sur l'intégration *in vitro* par l'étude des propriétés de fixation des mutants d'interaction et de leur capacité à intégrer dans l'ADN nucléosomal. En effet, la forte corrélation entre la capacité de l'IN à interagir avec la queue d'histone H4 et au mononucléosome et leur capacité à s'intégrer dans l'ADN nucléosomal démontrent le rôle crucial de cette interaction dans l'intégration. De plus, la stimulation de l'intégration *in vitro* induite par l'ajout du peptide dérivé de la queue d'histone H4 supporte l'implication de l'interaction dans l'optimisation du processus d'intégration. L'effet stimulateur du peptide dérivé de la queue d'histone H4 en intégration concertée indique que le peptide favorise la stabilisation du complexe actif IN/ADN. Ces résultats nous laissent donc penser que suite à l'approche du nucléosome par l'intasome du VIH-1, une étape d'activation par l'interaction de l'IN du VIH-1 avec le peptide issu de l'histone H4 est nécessaire. Cette activation permettrait la réaction d'intégration des deux extrémités virales de manière efficace et éviterait l'intégration partielle d'une seule extrémité virale, délétère pour le virus. Cette hypothèse est supportée par la superposition du modèle de docking moléculaire avec le complexe de transfert de brin de PFV (**Figure 81**). En effet, l'analyse suggère que la fixation du domaine CTD de l'IN du VIH-1 à l'ADN cible et au peptide H4K20me1 n'est pas exclusive et que le positionnement et l'orientation du peptide sont compatibles pour une interaction dans la cellule au sein du nucléosome. Ce phénotype d'intégration du VIH-1 dans le nucléosome est observé de la même façon chez PFV qui est décrit comme interagissant directement via l'interaction de l'IN à la surface des histones H2A/H2B.

De façon intéressante, chez VIH-1 et PFV les mutants déficients pour leur liaison au nucléosome présentent une altération de l'intégration et une préférence d'intégration

diminuée pour les nucléosomes dans la cellule infectée (Maskell *et al*, 2015). Ces données suggèrent que le docking de l'IN du VIH-1 au nucléosome est nécessaire pour une intégration efficace et que PFV et VIH-1 ont pu mettre en place au cours de l'évolution différents systèmes de fixation des histones afin d'interagir avec de telles structures. De plus, les données montrent une modulation de l'interaction entre l'IN du VIH-1 et l'état de méthylation de la queue H4 suggérant que le domaine CTD de l'IN du VIH-1 pourrait posséder un domaine de lecture de la chromatine. Les chromodomaines retrouvés chez les transposons jouent un rôle crucial dans le ciblage de l'intégration vers les régions hétérochromatiniennes via des interactions spécifiques avec des modifications post-traductionnelles d'histones. La question du rôle de cette interaction IN/H4 et de la reconnaissance spécifique de la modification H4K20me1 dans le choix du site d'intégration reste entière et constitue les perspectives de ce travail.

Partie 4 : Conclusion générale et perspectives

Afin de s'intégrer de façon pérenne dans le génome de la cellule cible, le CPI dont la composition reste en question, mais semble contenir les protéines virales RT, MA, Vpr et CA accompagné des protéines cellulaires LEDGF/p75, BAF et HMGA1, est nécessaire à l'import nucléaire et à l'insertion stable de l'ADN viral dans l'ADN cellulaire. Le choix du site d'intégration dans l'ADN cellulaire est une étape cruciale puisqu'elle permet d'établir une persistance virale de longue durée dans la cellule infectée. Les études préalables à cette thèse ont pu montrer que ce processus n'était pas aléatoire et que des structures de la chromatine pouvaient influencer la sélectivité de façon spécifique du rétrovirus. En effet, les résultats suggèrent la participation de facteurs cellulaires spécifiques jouant un rôle dans l'efficacité et le choix du site d'intégration. Ces données reposent dans un premier temps sur les données de séquençage des sites d'intégrations lentiviraux (VIH-1) et gammarétroviraux (MLV) et plus récemment pour d'autres genres de rétrovirus. Ainsi, il est aujourd'hui bien établi que les lentivirus favorisent les unités de transcription actives du génome enrichies en marques épigénétiques H3K4me, H3K4me2, H3K9me, H3K27me, H3K36me3, H4K20me1. Les gammarétrovirus possèdent quant à eux une préférence pour les extrémités 5' des gènes ainsi que pour les promoteurs enrichis en modifications d'histones H3K4me1 et H3K27ac. Les autres genres de rétrovirus possèdent une sélectivité plus modérée pour les unités de transcription ou les sites d'initiation de la transcription. Seuls les betarétrovirus ne montrent pas de préférence d'intégration marquée.

La sélectivité de l'intégration peut être expliquée en partie par un processus de ciblage grâce à l'interaction du CPI avec LEDGF/p75 dans le cas des lentivirus et avec les protéines BET incluant BRD2, BRD3 et BRD4 dans le cas des gammarétrovirus permettant ainsi un ciblage du CPI vers les régions d'intérêt. En effet, l'introduction de mutations dans le domaine de fixation à la chromatine de ces protéines résulte à un changement de préférence d'intégration. Cependant, d'autres études montrent que l'intasome rétroviral est capable en absence de cofacteurs cellulaires de participer au choix du site d'intégration dans l'ADN nucléosomal. Le premier axe de travail de ma thèse s'est alors focalisé sur l'analyse des propriétés intrinsèques des intasomes rétroviraux à accommoder des structures chromatinienne en l'absence de cofacteurs cellulaires.

1. 1. Contrôle de l'intégration par la structure de la chromatine et de l'intasome

Dans la cellule infectée, l'ADN cible est compacté et protégé par son enroulement autour de protéines histones afin de former la chromatine cellulaire. Il a été montré que la structure, la densité et la dynamique de la chromatine pouvaient influencer l'intégration rétrovirale *in vitro*. Dans le premier axe de travail de ma thèse, nous avons pu montrer que l'intégration de MLV, VIH-1, ASV et PFV était influencée de manière différentielle par des structures mimant la chromatine *in vitro*. En effet, en nous appuyant sur le système *in vitro* d'intégration concertée développé au laboratoire et utilisant un ADN chromatinien contenant deux régions de structure et de dynamique différentes, nous avons montré qu'une chromatine dense et compacte représente une structure inhibitrice à l'intégration concertée du VIH-1 et de ASV alors que le nucléosome isolé apparaît un bon substrat d'intégration. Au contraire, cette même structure apparaît comme favorable à l'intégration de PFV et MLV. Ces résultats obtenus *in vitro* concernant l'intégration du VIH-1 et MLV sont supportées par l'analyse des sites d'intégration issus de cellules infectées qui montrent qu'*in vivo* l'intégration du VIH-1 se fait également dans des régions de faible densité en nucléosome et qu'au contraire, l'intégration de MLV s'effectue dans des régions de forte densité en nucléosome. **L'ensemble de ces résultats représente la première démonstration que la structure de la chromatine entourant le nucléosome ciblé module l'intégration rétrovirale.**

Nous avons aussi pu montrer que la différence d'intégration observée entre VIH-1/ASV et MLV/PFV n'est pas dû à la présence d'un domaine additionnel « NED » indiquant alors que l'effet observé est majoritairement dû à l'architecture de la poche catalytique et pas à la structure globale de ces intasomes. Ces données sont supportées par l'analyse des sites d'intégrations du VIH-1 portant une duplication de 4, 5 et 6 paires de bases. En effet, nous avons montré que les sites d'intégrations portant une duplication de 5 et 6 paires de bases sont défavorisés dans les régions stables de la chromatine alors que les sites d'intégration portant une duplication de 4 paires de bases sont trouvés enrichis dans la chromatine dense. **Ces résultats nous ont permis de montrer que l'intasome comporte des déterminants structuraux qui participent à sa préférence pour les structures chromatiniennes ciblées.**

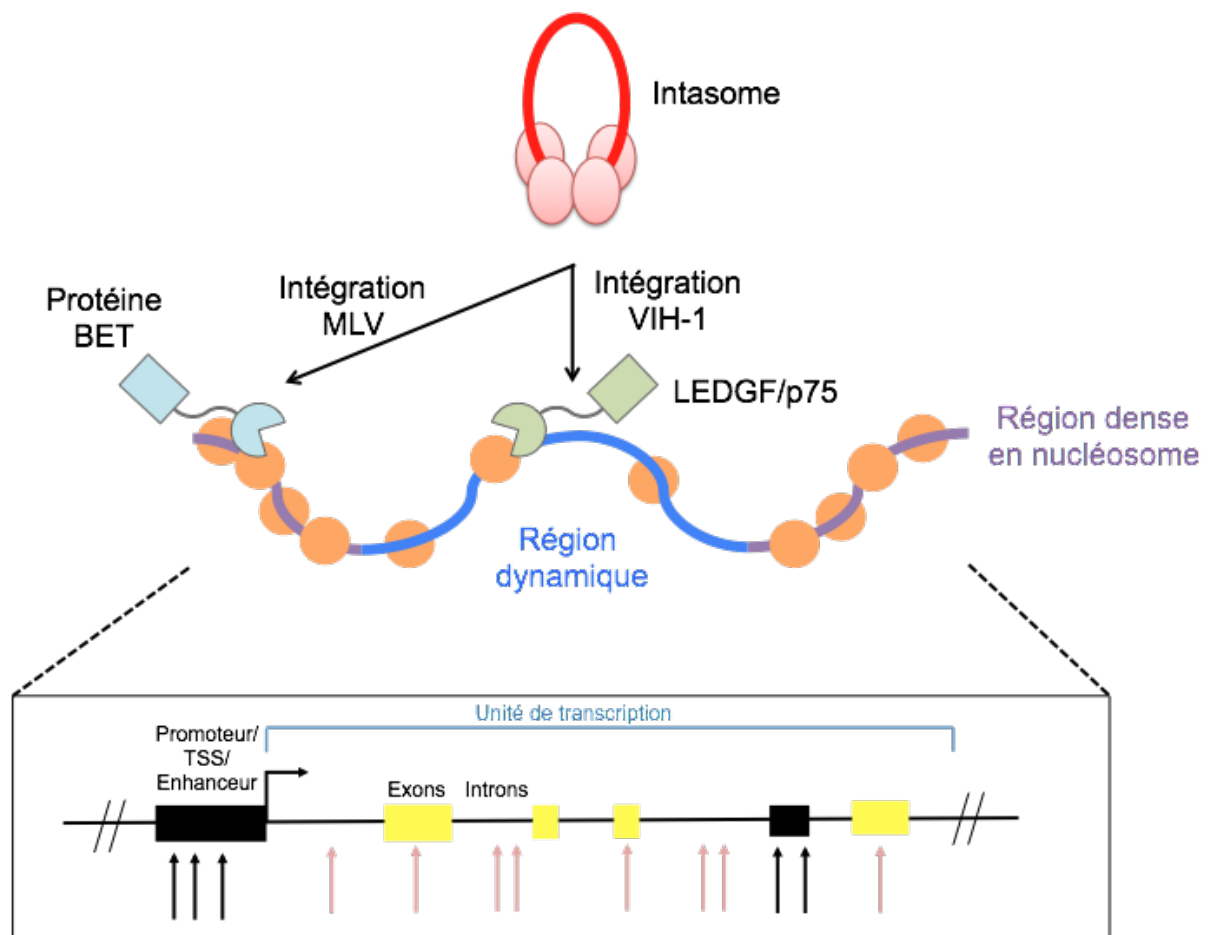


Figure 82 : Modèle d'intégration du VIH-1 et MLV dans la chromatine. Dans ce modèle, l'intasome du VIH-1 est ciblé via son cofacteur LEDGF/p75 vers les régions dynamiques de la chromatine grâce à l'interaction de LEDGF/p75 avec la modification d'histone H3K36me3 retrouvée enrichie dans cette région de la chromatine. Ainsi, l'intasome du VIH-1 est ensuite capable de s'adapter à l'ADN nucléosomal afin de s'intégrer dans les unités de transcription actives (flèches rouges). En revanche, l'intasome du MLV est ciblé par les protéines BET vers les régions denses en nucléosomes de la chromatine enrichies en modification d'histone acétylée permettant l'insertion du virus dans les promoteurs et dans les régions enhanceurs (flèches noires).

De cette analyse, un modèle d'une sélectivité de l'intégration rétrovirale en deux temps a pu être imaginé (**Figure 82**). Dans ce modèle, l'intégration du VIH-1 et ASV a lieu vers les régions nucléosomales dynamiques et l'intégration de MLV et PFV vers les régions denses de la chromatine grâce à un ciblage préalable via LEDGF/p75 ou les protéines BET. Puis est mis en jeu un contrôle local de la préférence pour les structures chromatiniques plus ou moins denses lié aux contraintes physiques au sein de l'intasomes. Ces structures pouvant être modulées par les mécanismes de remodelage de la chromatine, nous nous sommes orientés vers leurs études.

1. 2. Le FACT un nouveau cofacteur de l'intégration du VIH-1

Afin d'investiguer le rôle des facteurs de remodelage sur l'intégration du VIH-1, nous avons développé une stratégie visant à identifier les cofacteurs cellulaires du complexe nucléoprotéique intasomal du VIH-1 en se focalisant sur les protéines ayant un lien étroit avec le remodelage de la chromatine. Grâce au développement de ce système, nous avons pu identifier pour la première fois le complexe FACT comme interagissant potentiel de l'intasome du VIH-1. Ce complexe constitué des protéines SSRP1 et Spt16 est décrit dans la littérature comme permettant la maintenance de la chromatine grâce à une activité histone chaperone exercée localement sur le nucléosome. Dans l'objectif de valider le complexe FACT comme cofacteur cellulaire de l'intégration du VIH-1 nous avons suivi un algorithme de validation permettant premièrement d'étudier le mode d'interaction directe ou indirecte du FACT avec l'intasome du VIH-1 et deuxièmement de mettre en évidence un effet *in vivo* et *in vitro* du complexe FACT sur l'intégration du VIH-1. Des expériences de co-précipitation ont permis de déterminer que le FACT interagit avec l'IN de VIH-1 *via* l'association entre la protéine SSRP1 et LEDGF/p75. Cette interaction a notamment pu être confirmée très récemment (Lopez *et al*, 2016).

Dans l'objectif de caractériser le rôle potentiel du complexe FACT sur l'intégration du VIH-1, nous avons pu montrer qu'une baisse d'expression de la protéine SSRP1 *in cellulo* par siRNA lors de l'étape d'intégration du VIH-1 mène à une augmentation des événements d'intégrations dans les cellules 293T et HeLa-P4 associée à une augmentation de l'accessibilité globale de la chromatine. Des résultats similaires ont pu être obtenus lors d'approches pharmacologiques visant à inhiber le complexe chimiquement. En effet, l'ajout du complexe FACT dans le test d'intégration concertée *in vitro* permet la restauration de l'intégration sur matrice nucléosomale compacte et ceci de manière dépendante de l'activité

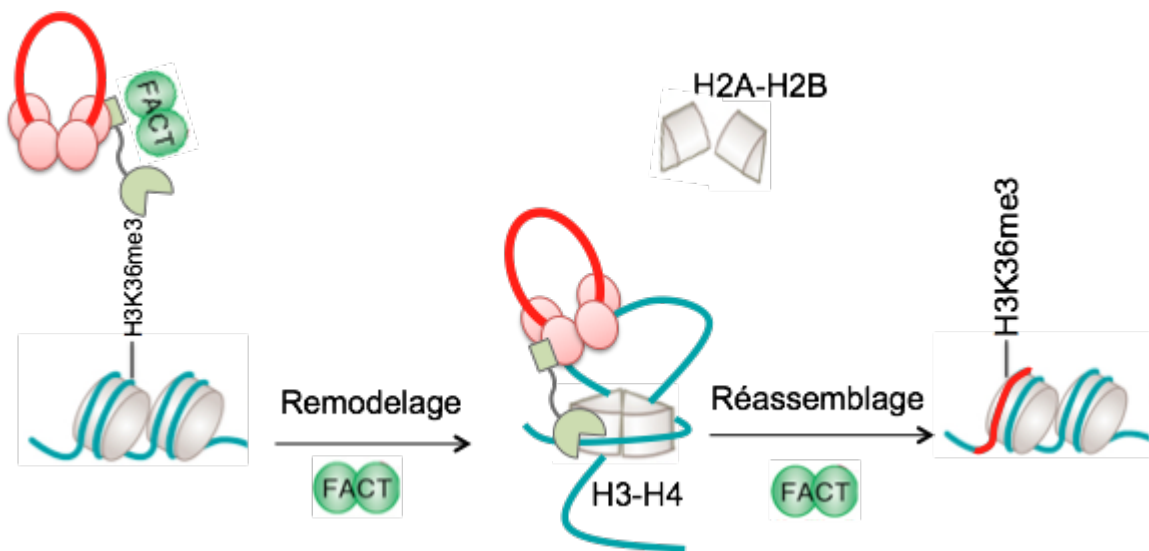


Figure 83 : Modèle d'intégration du VIH-1 en association avec le complexe protéique FACT. Dans ce modèle, l'intasome du VIH-1 est ciblé vers les régions de transcription actives via l'interaction entre son facteur LEDGF/p75 et H3K36me3. Dans ces régions l'interaction IN-LEDGF/p75-FACT peut s'effectuer couplant le remodelage du nucléosome et l'intégration. L'éjection des protéines histones H2A et H2B par le FACT permet de générer un nucléosome hexamérique favorable à l'intégration. Finalement, suite au passage de l'ARN polymérase de type II, le complexe FACT va ensuite permettre le réassemblage du nucléosome afin d'obtenir un produit final d'intégration sur un nucléosome dans les régions dynamiques de la chromatine.

de remodelage du FACT. **Ces résultats suggèrent que le FACT génère des structures nucléosomales partiellement dissociées favorisant l'intégration du VIH-1.**

Ceci a pu être confirmé par l'utilisation de structures chromatiniques tétramériques mimant les structures générées par ce complexe. En effet, ces structures apparaissent comme des substrats d'intégration optimale *in vitro*. De cette analyse, un modèle de couplage entre intégration rétrovirale et remodelage peut être imaginé (**Figure 83**). Dans ce modèle l'intégration, le CPI est ciblé vers les régions dynamiques de la chromatine via son cofacteur de ciblage LEDGF/p75. Le couplage avec le FACT peut alors se faire via l'interaction LEDGF/SSRP1 et l'activité de remodelage du complexe FACT permet de générer les structures nucléosomales favorables à l'intégration.

Le rôle exact de ce remodelage local dans l'intégration reste obscur. Cependant, la dissociation du nucléosome pourrait permettre de rendre l'ADN accessible. Ce mécanisme reste à être élucidé et constitue un axe de recherche développé par le laboratoire.

1. 3. Un chromodomaine dans l'IN du VIH-1

Étant donné que les queues d'histone sont des candidats potentiels pour l'intasome approchant le nucléosome, nous avons analysé l'existence d'interactions IN/histones. La reconnaissance des queues d'histones dans la cellule requiert la présence d'une structure bien définie appelée bromodomaine pour la reconnaissance des marques acétylées, les marques méthylées sont quant à elles reconnues par de nombreux domaines formant la « Royal family » incluant les chromo, Tudor, MBT (Malignant brain tumor), chromo barrel et PWWP domaines. Ainsi, nous avons analysé la capacité de l'IN du VIH-1 à reconnaître spécifiquement les marques épigénétiques *in vitro*. Les expériences par histone peptide array ont permis de montrer que **l'IN du VIH-1 est capable de reconnaître spécifiquement la queue d'histone H4 mono-méthylée sur la lysine 20 (H4K20me1)**. Grâce à l'utilisation de plusieurs constructions tronquées pour leur domaine NTD ou CTD, nous avons montré que le domaine CTD de l'IN était responsable de cette interaction. Ces résultats ont ensuite pu être confirmés par des expériences de « microscale thermophoresis » (MST) permettant de quantifier l'interaction avec un K_D de $0,8 \pm 0,13 \mu M$ pour l'interaction CTD-H4K20me1.

Dans un second temps, afin d'identifier le mode d'interaction du CTD avec le peptide H4K20me1 nous avons comparé la structure de la protéine MSL3 (hMSL3) possédant un chromo barrel domaine connu pour interagir avec le peptide H4K20me1 avec la structure du CTD de l'IN du VIH-1. **Ainsi la superposition des structures du CTD de l'IN et de la protéine hMSL3 montre une forte homologie structurale avec un RMSD de 1,9 Å suggérant une similitude d'interaction non reportée jusqu'à ce jour.**

Puis, afin d'identifier les acides aminés responsables de l'interaction, nous avons développé une approche structurale *in silico* visant à modéliser l'interaction du peptide H4K20me1 dans le domaine CTD de l'IN. De cette analyse, et grâce au calcul de l'énergie d'interaction, une modélisation a été retenue. Dans cette modélisation le peptide H4K20me1 engage 7 liaisons hydrogènes avec les résidus D229, R231, S255, D256, K258 et 15 liaisons hydrophobes avec les résidus Y227, D229, N232, K236, D256, K258 et V260. De plus, dans ce modèle la monométhylation de la lysine 20 du peptide pointe dans la cavité en V et engage à elle seule 7 des 15 liaisons hydrophobes et 1 des 7 liaisons hydrogènes. Puis, afin de valider notre modèle et confirmer l'implication des acides aminés identifiés, nous avons introduit les mutations sur les acides aminés responsables de l'architecture de la cavité en V ou des interactions hydrogènes ou hydrophobes avec le peptide H4K20me1. Nous avons ensuite pu tester l'ensemble de ces protéines mutantes pour leur interaction avec le peptide H4 et ses variants méthylés. Nous avons pu montrer grâce à ces expériences de « far dot blot » que les mutants R231A et R231H montrent une baisse d'affinité pour la queue H4 avec une inhibition de près de 80 % pour le R231H. Le mutant D229G montre quant à lui une augmentation d'affinité à la queue H4 par rapport à l'enzyme WT. Les mutants Y227A, W235A et K236A montrent aussi une augmentation d'affinité pour la queue H4, mais ces mutants sont aussi accompagnés d'une perte de spécificité vis-à-vis de la monométhylation de la queue H4. Le D253H montre une augmentation de l'affinité pour la monométhylation de la queue H4 cependant, son affinité pour les autres variants n'est pas modifiée. Ainsi, nous avons pu montrer que la majorité des mutants est affectée dans la fixation au peptide H4K20me1 et que dans certains cas une perte de spécificité est observée. **Ces résultats permettent de valider pour la première fois la présence d'un site de fixation du peptide H4K20me1 dans le domaine CTD de l'IN du VIH-1.**

Dans un troisième temps, nous pu tester le rôle de l'interaction IN/H4 dans le cycle réplcatif viral. Pour cela, nous avons comparé l'efficacité de réplcation de vecteurs rétroviraux portant les mutations impliquées dans l'interaction avec la queue H4 *in vitro* dans

des expériences de « single round infection assay » dans plusieurs types cellulaires (HEK-293T, HeLa P4 et K562). Nous avons ainsi pu montrer que les mutants R231A et R231H présentent une diminution de la réplication virale (40-60 %) au contraire, une augmentation de la réplication virale est observée pour le mutant D253H suggérant une corrélation directe entre capacité de liaison à la queue H4 et efficacité de réplication. Nous avons ensuite pu montrer que les mutants R231A et R231H présentent une baisse de l'ADN intégré de 60 et 80 % respectivement avec une accumulation des cercles à 2LTR. Le mutant D253H montre quant à lui un phénotype opposé avec une augmentation de 20 à 30 % de l'ADN intégré associé avec une diminution des cercles à 2LTR suggérant une meilleure efficacité de l'étape d'intégration pour ce mutant.

Dans une seconde partie, afin de comprendre le rôle de l'interaction avec la queue H4 sur l'étape d'intégration, nous avons analysé l'effet du peptide H4 et ses variants sur l'activité d'intégration dans un test d'intégration concerté *in vitro*. Ainsi, nous avons montré que l'ajout du peptide H4K20me1 induit une forte stimulation de l'intégration catalysée par l'IN WT d'un facteur 2 à 2,5. Au contraire, les mutants R231A et R231H possédant une affinité diminuée pour le peptide H4K20me1 sont aussi moins sensibles à la stimulation par le peptide H4K20me1 suggérant que la stimulation de l'activité de l'IN induit par le peptide requiert une interaction entre l'IN et le peptide. De plus, les tests d'intégration utilisant des peptides issus d'autres histones (H2A, H2B et H3) ainsi que des peptides contrôles (reverse et scramble) ne montrent pas d'effet significatif sur l'intégration concertée suggérant une spécificité du peptide méthylé sur la position 20 de la queue H4 (non montré). Puis, nous avons montré que l'ajout du peptide H4K20me1 entraîne une forte diminution de l'affinité de l'IN pour le mononucléosome. Au contraire, l'ajout du peptide n'inhibe pas l'interaction de l'IN pour la séquence d'ADN contrôle W601 suggérant donc que l'interaction de l'IN avec la queue H4 est cruciale pour son interaction avec le nucléosome. Ces résultats ont ensuite pu être confirmés par l'analyse de la capacité de fixation des mutants d'IN précédemment décrits aux mononucléosomes dans des tests classiques de coprécipitation. Ainsi, les mutants R231A et R231H décrits préalablement pour leur défaut d'affinité avec la queue H4 montrent une baisse d'affinité significative pour le mononucléosome. À l'opposé, les mutants D229G et D253H qui possèdent une fixation plus importante au peptide H4 par rapport à l'IN WT montrent une augmentation d'affinité par le nucléosome.. **Ainsi, nous avons pu montrer l'importance de l'interaction de l'IN avec le peptide H4 dans l'association finale avec le mononucléosome.** Finalement, nous avons expérimenté l'impact de cette interaction sur l'intégration catalysée par l'IN du VIH-1 sur mononucléosome. Dans les tests

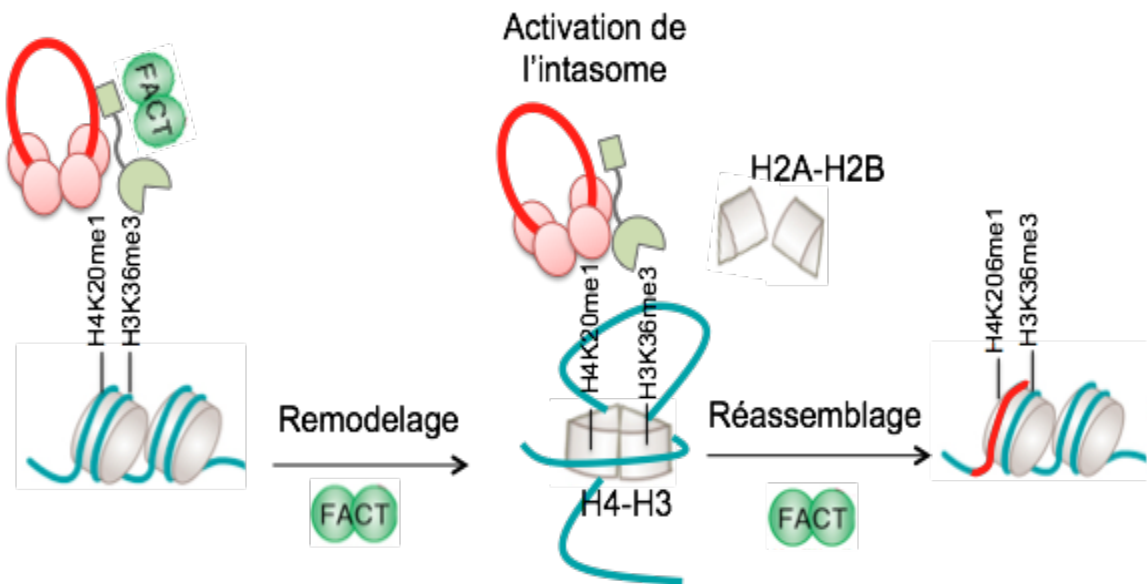


Figure 84 : Modèle d'intégration du VIH-1 en association avec l'interaction IN/H4K20me1 et le complexe de remodelage FACT. Dans ce modèle, l'intasome du VIH-1 est ciblé vers les régions de transcription actives via l'interaction entre son cofacteur LEDGF/p75 et H3K36me3. Ce ciblage va ensuite permettre l'activation de l'intasome du VIH-1 grâce à son interaction avec le peptide d'histone H4K20me1 afin de s'intégrer efficacement dans une structure partiellement dissociée du nucléosome générée par le complexe FACT. Suite au passage de l'ARN polymérase de type II, le complexe FACT va ensuite permettre le réassemblage du nucléosome afin d'obtenir un produit final d'intégration sur un nucléosome dans une région enrichie en modification post-traductionnelles H3K36me3 et H4K20me1.

d'intégration sur mononucléosome, nous avons montré que les mutants R231A et R231H possèdent un profil d'intégration différent de l'enzyme WT. En effet, ces mutants catalysent les produits HSI plus efficacement que les produits d'intégration FSI en comparaison avec l'enzyme WT. Au contraire, le mutant D253H montre une activité de 20 à 30 % plus importante que l'enzyme WT. Grâce à l'ensemble de ces données obtenues *in vitro*, nous avons mis en relief l'importance de la fixation des IN à la queue H4 afin de permettre l'intégration sur un substrat physiologique nucléosomal. Ces résultats confirmant ainsi, l'importance de cette association pour une intégration optimale.

De ces résultats, un modèle d'intégration dans lequel suite à l'approche du nucléosome par l'intasome du VIH-1, une étape d'activation par l'interaction de l'IN du VIH-1 avec le peptide issu de l'histone H4 peut être envisagée (**Figure 84**). De manière intéressante la modification H4K20me1 est enrichie dans les régions ciblées par le virus. La question du rôle de la reconnaissance de cette modification dans le choix du site d'intégration est donc soulevée. L'étude de la sélectivité des virus portant les mutations d'interaction IN/H4 permettra de répondre à cette question.

Par ailleurs, les régions ciblées par le virus sont enrichies en complexes de transcription et de remodelage de la chromatine dont le FACT précédemment identifié comme un partenaire fonctionnel possible de l'intégration du virus. De manière intéressante la dissociation partielle du nucléosome par le FACT pourrait libérer la queue d'histone H4 pour une interaction optimale et une activation de l'intasome. En effet dans une structure chromatinienne stable la queue d'histone H4 est engagée dans des interactions avec le nucléosome voisin (Luger *et al*, 1997). L'action de remodelage a donc pour conséquence directe de libérer cette queue permettant l'interaction avec l'intasome. C'est donc cette hypothèse qui est privilégiée et étudiée par le laboratoire.

En plus d'apporter des éclaircissements quant aux mécanismes d'intégration rétroviraux, notre travail a pu également identifier de nouveaux leviers afin d'affecter la réplication virale constituant ainsi des cibles thérapeutiques intéressantes. En effet, l'interaction IN/H4 semble cruciale pour le cycle réplcatif et l'intégration du virus permettant de valider l'interaction comme candidat de choix dans de nouvelles approches antivirales. C'est donc dans cet objectif que la recherche de composés dissociant l'association IN/H4 est en cours d'identification au laboratoire par des méthodes d'« alphascreen ».

De plus, si la reconnaissance de la marque H4K20me1 joue un rôle dans la sélectivité de l'intégration, l'association IN/H4 constituerait également un outil d'intérêt dans les approches de thérapie génique utilisant des vecteurs lentiviraux. Cet axe est actuellement étudié au laboratoire par l'analyse des sites d'intégration de vecteurs portant

les mutations altérant l'interaction. De plus, les composés sélectionnés par « alphascreen » pourraient également servir de modulateurs chimiques dans ces approches.

D'un point de vue plus fondamental les similitudes structurales et fonctionnelles retrouvées entre l'IN du VIH-1 et le facteur cellulaire hMSL3 ainsi que les analogies entre le domaine de lecture de la chromatine identifié dans le CTD de l'IN du VIH-1 et les chromodomaines de rétrotransposons soulèvent la question du lien évolutif existant entre ces différents modèles. Même si ce pan du travail ouvre de vastes projets d'étude, nos résultats suggèrent que la reconnaissance de marques épigénétiques par les acteurs de la mobilité des génomes comme les intégrases rétrovirales, pourrait être un processus plus général dans le monde du vivant.

Bibliographie

- Albanese A, Arosio D, Terreni M & Cereseto A (2008) HIV-1 pre-integration complexes selectively target decondensed chromatin in the nuclear periphery. *PLoS One* **3**: e2413
- Arents G, Burlingame RW, Wang BC, Love WE & Moudrianakis EN (1991) The nucleosomal core histone octamer at 3.1 Å resolution: a tripartite protein assembly and a left-handed superhelix. *Proc. Natl. Acad. Sci. U. S. A.* **88**: 10148–10152
- Ariumi Y, Serhan F, Turelli P, Telenti A & Trono D (2006) The integrase interactor 1 (INI1) proteins facilitate Tat-mediated human immunodeficiency virus type 1 transcription. *Retrovirology* **3**: 47
- Ballandras-Colas A, Brown M, Cook NJ, Dewdney TG, Demeler B, Cherepanov P, Lyumkis D & Engelman AN (2016) Cryo-EM reveals a novel octameric integrase structure for betaretroviral intasome function. *Nature* **530**: 358–361
- Ballaré C, Lange M, Lapinaite A, Martin GM, Morey L, Pascual G, Liefke R, Simon B, Shi Y, Gozani O, Carlomagno T, Benitah SA & Di Croce L (2012) Phf19 links methylated Lys36 of histone H3 to regulation of Polycomb activity. *Nat. Struct. Mol. Biol.* **19**: 1257–1265
- Ball LJ, Murzina NV, Broadhurst RW, Raine AR, Archer SJ, Stott FJ, Murzin AG, Singh PB, Domaille PJ & Laue ED (1997) Structure of the chromatin binding (chromo) domain from mouse modifier protein 1. *EMBO J.* **16**: 2473–2481
- Bannister AJ, Zegerman P, Partridge JF, Miska EA, Thomas JO, Allshire RC & Kouzarides T (2001) Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **401**: 120–124
- Barré-Sinoussi F, Chermann JC, Rey F, Nugeyre MT, Chamaret S, Gruest J, Dautet C, Axler-Blin C, Vézinet-Brun F, Rouzioux C, Rozenbaum W & Montagnier L (1983) Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* **220**: 868–871
- Becker PB & Workman JL (2013) Nucleosome remodeling and epigenetics. *Cold Spring Harb. Perspect. Biol.* **5**: Available at: <http://dx.doi.org/10.1101/cshperspect.a017905>
- Benleulmi MS, Matysiak J, Henriquez DR, Vaillant C, Lesbats P, Calmels C, Naughtin M, Leon O, Skalka AM, Ruff M, Lavigne M, Andreola M-L & Parissi V (2015) Intasome architecture and chromatin density modulate retroviral integration into nucleosome. *Retrovirology* **12**: 13
- Bernstein E & Hake SB (2006) The nucleosome: a little variation goes a long way. *Biochem. Cell Biol.* **84**: 505–517
- de Béthune M-P (2010) Non-nucleoside reverse transcriptase inhibitors (NNRTIs), their discovery, development, and use in the treatment of HIV-1 infection: a review of the last 20 years (1989-2009). *Antiviral Res.* **85**: 75–90
- Blumenthal R, Durell S & Viard M (2012) HIV entry and envelope glycoprotein-mediated fusion. *J. Biol. Chem.* **287**: 40841–40849
- Bondarenko MT, Maluchenko NV, Valieva ME, Gerasimova NS, Kulaeva OI, Georgiev PG & Studitsky VM (2015) Structure and function of histone chaperone FACT. *Mol. Biol.* **49**: 796–809

Botbol Y, Raghavendra NK, Rahman S, Engelman A & Lavigne M (2008) Chromatinized templates reveal the requirement for the LEDGF/p75 PWWP domain during HIV-1 integration in vitro. *Nucleic Acids Res.* **36**: 1237–1246

Briggs JAG, Riches JD, Glass B, Bartonova V, Zanetti G & Kräusslich H-G (2009) Structure and assembly of immature HIV. *Proc. Natl. Acad. Sci. U. S. A.* **106**: 11090–11095

Brik A & Wong C-H (2003) HIV-1 protease: mechanism and drug discovery. *Org. Biomol. Chem.* **1**: 5–14

Brin E, Yi J, Skalka AM & Leis J (2000) Modeling the late steps in HIV-1 retroviral integrase-catalyzed DNA integration. *J. Biol. Chem.* **275**: 39287–39295

Brown DT, Izard T & Misteli T (2006) Mapping the interaction surface of linker histone H10 with the nucleosome of native chromatin in vivo. *Nat. Struct. Mol. Biol.* **13**: 250–255

Bukrinskaya A (2007) HIV-1 matrix protein: a mysterious regulator of the viral life cycle. *Virus Res.* **124**: 1–11

Burgess RJ & Zhang Z (2013) Histone chaperones in nucleosome assembly and human disease. *Nat. Struct. Mol. Biol.* **20**: 14–22

Bushman FD & 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

Buzon V, Natrajan G, Schibli D, Campelo F, Kozlov MM & Weissenhorn W (2010) Crystal structure of HIV-1 gp41 including both fusion peptide and membrane proximal external regions. *PLoS Pathog.* **6**: e1000880

Cai L, Rothbart SB, Lu R, Xu B, Chen W-Y, Tripathy A, Rockowitz S, Zheng D, Patel DJ, Allis CD, Strahl BD, Song J & Wang GG (2013) An H3K36 methylation-engaging Tudor motif of polycomb-like proteins mediates PRC2 complex targeting. *Mol. Cell* **49**: 571–582

Cai M, Huang Y, Zheng R, Wei SQ, Ghirlando R, Lee MS, Craigie R, Gronenborn AM & Clore GM (1998) Solution structure of the cellular factor BAF responsible for protecting retroviral DNA from autointegration. *Nat. Struct. Biol.* **5**: 903–909

Cai M, Zheng R, Caffrey M, Craigie R, Clore GM & Gronenborn AM (1997) Solution structure of the N-terminal zinc binding domain of HIV-1 integrase. *Nat. Struct. Biol.* **4**: 567–577

Cai Y, Jin J, Yao T, Gottschalk AJ, Swanson SK, Wu S, Shi Y, Washburn MP, Florens L, Conaway RC & Conaway JW (2007) YY1 functions with INO80 to activate transcription. *Nat. Struct. Mol. Biol.* **14**: 872–874

Campbell EM & Hope TJ (2015) HIV-1 capsid: the multifaceted key player in HIV-1 infection. *Nat. Rev. Microbiol.* **13**: 471–483

Cattoglio C, Pellin D, Rizzi E, Maruggi G, Corti G, Miselli F, Sartori D, Guffanti A, Di Serio C, Ambrosi A, De Bellis G & Mavilio F (2010) High-definition mapping of retroviral integration sites identifies active regulatory elements in human multipotent hematopoietic progenitors. *Blood* **116**: 5507–5517

- Chakravarthy S, Park Y-J, Chodaparambil J, Edayathumangalam RS & Luger K (2005) Structure and dynamic properties of nucleosome core particles. *FEBS Lett.* **579**: 895–898
- Chen H & Engelman A (1998) The barrier-to-autointegration protein is a host factor for HIV type 1 integration. *Proc. Natl. Acad. Sci. U. S. A.* **95**: 15270–15274
- Chen JC, Krucinski J, Miercke LJ, Finer-Moore JS, Tang AH, Leavitt AD & Stroud RM (2000) Crystal structure of the HIV-1 integrase catalytic core and C-terminal domains: a model for viral DNA binding. *Proc. Natl. Acad. Sci. U. S. A.* **97**: 8233–8238
- Cherepanov P (2007) LEDGF/p75 interacts with divergent lentiviral integrases and modulates their enzymatic activity in vitro. *Nucleic Acids Res.* **35**: 113–124
- Cherepanov P, Ambrosio ALB, Rahman S, Ellenberger T & Engelman A (2005) Structural basis for the recognition between HIV-1 integrase and transcriptional coactivator p75. *Proc. Natl. Acad. Sci. U. S. A.* **102**: 17308–17313
- Cherepanov P, Devroe E, Silver PA & Engelman A (2004) Identification of an Evolutionarily Conserved Domain in Human Lens Epithelium-derived Growth Factor/Transcriptional Co-activator p75 (LEDGF/p75) That Binds HIV-1 Integrase. *J. Biol. Chem.* **279**: 48883–48892
- Cherepanov P, Maertens G, Proost P, Devreese B, Van Beeumen J, Engelborghs Y, De Clercq E & Debyser Z (2003) HIV-1 integrase forms stable tetramers and associates with LEDGF/p75 protein in human cells. *J. Biol. Chem.* **278**: 372–381
- Chow SA, Vincent KA, Ellison V & Brown PO (1992) Reversal of integration and DNA splicing mediated by integrase of human immunodeficiency virus. *Science* **255**: 723–726
- Christ F, Shaw S, Demeulemeester J, Desimmie BA, Marchand A, Butler S, Smets W, Chaltin P, Westby M, Debyser Z & Pickford C (2012) Small-molecule inhibitors of the LEDGF/p75 binding site of integrase block HIV replication and modulate integrase multimerization. *Antimicrob. Agents Chemother.* **56**: 4365–4374
- Christ F, Thys W, De Rijck J, Gijsbers R, Albanese A, Arosio D, Emiliani S, Rain J-C, Benarous R, Cereseto A & Debyser Z (2008) Transportin-SR2 imports HIV into the nucleus. *Curr. Biol.* **18**: 1192–1202
- Ciuffi A & Barr SD (2011) Identification of HIV integration sites in infected host genomic DNA. *Methods* **53**: 39–46
- Ciuffi A, Llano M, Poeschla E, Hoffmann C, Leipzig J, Shinn P, Ecker JR & Bushman F (2005) A role for LEDGF/p75 in targeting HIV DNA integration. *Nat. Med.* **11**: 1287–1289
- Clapier CR & Cairns BR (2009) The biology of chromatin remodeling complexes. *Annu. Rev. Biochem.* **78**: 273–304
- Coffin J, Haase A, Levy JA, Montagnier L, Oroszlan S, Teich N, Temin H, Toyoshima K, Varmus H & Vogt P (1986) What to call the AIDS virus? *Nature* **321**: 10
- Connor RI, Chen BK, Choe S & Landau NR (1995) Vpr is required for efficient replication of human immunodeficiency virus type-1 in mononuclear phagocytes. *Virology* **206**: 935–944

- Cosnefroy O, Tocco A, Lesbats P, Thierry S, Calmels C, Wiktorowicz T, Reigadas S, Kwon Y, De Cian A, Desfarges S, Bonot P, San Filippo J, Litvak S, Cam EL, Rethwilm A, Fleury H, Connell PP, Sung P, Delelis O, Andréola ML, et al (2012) Stimulation of the human RAD51 nucleofilament restricts HIV-1 integration in vitro and in infected cells. *J. Virol.* **86**: 513–526
- Costanzi C & Pehrson JR (1998) Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals. *Nature* **393**: 599–601
- Czarnota GJ, Bazett-Jones DP, Mendez E, Allfrey VG & Ottensmeyer FP (1997) High resolution microanalysis and three-dimensional nucleosome structure associated with transcribing chromatin. *Micron* **28**: 419–431
- Daniel R, Kao G, Taganov K, Greger JG, Favorova O, Merkel G, Yen TJ, Katz RA & Skalka AM (2003) Evidence that the retroviral DNA integration process triggers an ATR-dependent DNA damage response. *Proc. Natl. Acad. Sci. U. S. A.* **100**: 4778–4783
- Dantzer F, Amé J-C, Schreiber V, Nakamura J, Ménissier-de Murcia J & de Murcia G (2006) Poly(ADP-ribose) polymerase-1 activation during DNA damage and repair. *Methods Enzymol.* **4** :**9**: 493–510
- Darlix J-L, Chang YU, Berthoux L, Ottmann M, Jullian N & Roques B (1995) La nucléocapside du VIH :1: un paradigme pour la recherche et ses applications médicales. *MS. Médecine sciences* :**1**: 420–429
- De Guzman RN, Wu ZR, Stalling CC, Pappalardo L, Borer PN & Summers MF (1998) Structure of the HIV-1 nucleocapsid protein bound to the SL3 psi-RNA recognition element. *Science* **279**: 384–388
- Delelis O, Carayon K, Saïb A, Deprez E & Mouscadet J-F (2008) Integrase and integration: biochemical activities of HIV-1 integrase. *Retrovirology* **5**: 114
- Demeulemeester J, De Rijck J, Gijsbers R & Debyser Z (2015) Retroviral integration: Site matters: Mechanisms and consequences of retroviral integration site selection. *Bioessays* **37**: 1202–1214
- De Rijck J, de Kogel C, Demeulemeester J, Vets S, El Ashkar S, Malani N, Bushman FD, Landuyt B, Husson SJ, Busschots K, Gijsbers R & Debyser Z (2013) The BET family of proteins targets moloney murine leukemia virus integration near transcription start sites. *Cell Rep.* **5**: 886–894
- Desfarges S & Ciuffi A (2010) Retroviral integration site selection. *Viruses* **2**: 111–130
- Desfarges S, San Filippo J, Fournier M, Calmels C, Caumont-Sarcos A, Litvak S, Sung P & Parissi V (2006) Chromosomal integration of LTR-flanked DNA in yeast expressing HIV-1 integrase: down regulation by RAD51. *Nucleic Acids Res.* **34**: 6215–6224
- Di Nunzio F, Fricke T, Miccio A, Valle-Casuso JC, Perez P, Souque P, Rizzi E, Severgnini M, Mavilio F, Charneau P & Diaz-Griffero F (2013) Nup153 and Nup98 bind the HIV-1 core and contribute to the early steps of HIV-1 replication. *Virology* **440**: 8–18
- Di Primio C, Quercioli V, Allouch A, Gijsbers R, Christ F, Debyser Z, Arosio D & Cereseto A (2013) Single-cell imaging of HIV-1 provirus (SCIP). *Proc. Natl. Acad. Sci. U. S. A.* **110**: 5636–5641

- Drosopoulos WC, Rezende LF, Wainberg MA & Prasad VR (1998) Virtues of being faithful: can we limit the genetic variation in human immunodeficiency virus? *J. Mol. Med.* **76**: 604–612
- Dyda F, Hickman AB, Jenkins TM, Engelman A, Craigie R & Davies DR (1994) Crystal structure of the catalytic domain of HIV-1 integrase: similarity to other polynucleotidyl transferases. *Science* **266**: 1981–1986
- Eidahl JO, Crowe BL, North JA, McKee CJ, Shkriabai N, Feng L, Plumb M, Graham RL, Gorelick RJ, Hess S, Poirier MG, Foster MP & Kvaratskhelia M (2013) Structural basis for high-affinity binding of LEDGF PWWP to mononucleosomes. *Nucleic Acids Res.* **41**: 3924–3936
- Eijkelenboom AP, van den Ent FM, Vos A, Doreleijers JF, Hård K, Tullius TD, Plasterk RH, Kaptein R & Boelens R (1997) The solution structure of the amino-terminal HHCC domain of HIV-2 integrase: a three-helix bundle stabilized by zinc. *Curr. Biol.* **7**: 739–746
- Eijkelenboom AP, Lutzke RA, Boelens R, Plasterk RH, Kaptein R & Hård K (1995) The DNA-binding domain of HIV-1 integrase has an SH3-like fold. *Nat. Struct. Biol.* **2**: 807–810
- Emiliani S, Mousnier A, Busschots K, Maroun M, Van Maele B, Tempé D, Vandekerckhove L, Moisan F, Ben-Slama L, Witvrouw M, Christ F, Rain J-C, Dargemont C, Debyser Z & Benarous R (2005) Integrase Mutants Defective for Interaction with LEDGF/p75 Are Impaired in Chromosome Tethering and HIV-1 Replication. *J. Biol. Chem.* **280**: 25517–25523
- Engelman A, Bushman FD & Craigie R (1993) Identification of discrete functional domains of HIV-1 integrase and their organization within an active multimeric complex. *EMBO J.* **12**: 3269–3275
- Engelman A & Craigie R (1992) Identification of conserved amino acid residues critical for human immunodeficiency virus type 1 integrase function in vitro. *J. Virol.* **66**: 6361–6369
- van den Ent FMI, Vos A & Plasterk RHA (1999) Dissecting the Role of the N-Terminal Domain of Human Immunodeficiency Virus Integrase bytrans-Complementation Analysis. *J. Virol.* **73**: 3176–3183
- Falkenberg KJ & Johnstone RW (2014) Histone deacetylases and their inhibitors in cancer, neurological diseases and immune disorders. *Nat. Rev. Drug Discov.* **13**: 673–691
- Fan Y, Nikitina T, Zhao J, Fleury TJ, Bhattacharyya R, Bouhassira EE, Stein A, Woodcock CL & Skoultchi AI (2005) Histone H1 depletion in mammals alters global chromatin structure but causes specific changes in gene regulation. *Cell* **123**: 1199–1212
- Farnet CM & Bushman FD (1997) HIV-1 cDNA integration: requirement of HMG I(Y) protein for function of preintegration complexes in vitro. *Cell* **88**: 483–492
- Fassati A (2012) Multiple roles of the capsid protein in the early steps of HIV-1 infection. *Virus Res.* **170**: 15–24

Ferris AL, Wu X, Hughes CM, Stewart C, Smith SJ, Milne TA, Wang GG, Shun M-C, Allis CD, Engelman A & Hughes SH (2010) Lens epithelium-derived growth factor fusion proteins redirect HIV-1 DNA integration. *Proc. Natl. Acad. Sci. U. S. A.* **107**: 3135–3140

Fink AL (1999) Chaperone-mediated protein folding. *Physiol. Rev.* **79**: 425–449

Freed EO (2015) HIV-1 assembly, release and maturation. *Nat. Rev. Microbiol.* **13**: 484–496

Friedrich M, Setz C, Hahn F, Matthaei A, Fraedrich K, Rauch P, Henklein P, Traxdorf M, Fossen T & Schubert U (2016) Glutamic Acid Residues in HIV-1 p6 Regulate Virus Budding and Membrane Association of Gag. *Viruses* **8**: Available at: <http://dx.doi.org/10.3390/v8040117>

Gasparian AV, Burkhart CA, Purmal AA, Brodsky L, Pal M, Saranadasa M, Bosykh DA, Commane M, Guryanova OA, Pal S, Safina A, Sviridov S, Koman IE, Veith J, Komar AA, Gudkov AV & Gurova KV (2011) Curaxins: anticancer compounds that simultaneously suppress NF- κ B and activate p53 by targeting FACT. *Sci. Transl. Med.* **3**: 95ra74

Van Gent DC, Vink C, Groeneger AA & Plasterk RH (1993) Complementation between HIV integrase proteins mutated in different domains. *EMBO J.* **12**: 3261–3267

Georgopoulos K (2002) Haematopoietic cell-fate decisions, chromatin regulation and ikaros. *Nat. Rev. Immunol.* **2**: 162–174

Gijsbers R, Ronen K, Vets S, Malani N, De Rijck J, McNeely M, Bushman FD & Debyser Z (2010) LEDGF hybrids efficiently retarget lentiviral integration into heterochromatin. *Mol. Ther.* **18**: 552–560

Godde JS & Ura K (2008) Cracking the enigmatic linker histone code. *J. Biochem.* **143**: 287–293

Groom HCT, Anderson EC, Dangerfield JA & Lever AML (2009) Rev regulates translation of human immunodeficiency virus type 1 RNAs. *J. Gen. Virol.* **90**: 1141–1147

Guideline on When to Start Antiretroviral Therapy and on Pre-Exposure Prophylaxis for HIV (2015) Geneva: World Health Organization

Hacein-Bey-Abina S, Von Kalle C, Schmidt M, McCormack MP, Wulffraat N, Leboulch P, Lim A, Osborne CS, Pawliuk R, Morillon E, Sorensen R, Forster A, Fraser P, Cohen JL, de Saint Basile G, Alexander I, Wintergerst U, Frebourg T, Aurias A, Stoppa-Lyonnet D, et al (2003) LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science* **302**: 415–419

Hake SB & Allis CD (2006) Histone H3 variants and their potential role in indexing mammalian genomes: the 'H3 barcode hypothesis'. *Proc. Natl. Acad. Sci. U. S. A.* **103**: 6428–6435

Hall JA & Georgel PT (2007) CHD proteins: a diverse family with strong ties. *Biochem. Cell Biol.* **85**: 463–476

- Hamiche A, Carot V, Alilat M, De Lucia F, O'Donohue M-F, Revet B & Prunell A (1996) Interaction of the histone (H3-H4) 2 tetramer of the nucleosome with positively supercoiled DNA minicircles: Potential flipping of the protein from a left- to a right-handed superhelical form. *Proceedings of the National Academy of Sciences* **93**: 7588–7593
- Hare S, Gupta SS, Valkov E, Engelman A & Cherepanov P (2010) Retroviral intasome assembly and inhibition of DNA strand transfer. *Nature* **464**: 232–236
- Hare S, Shun M-C, Gupta SS, Valkov E, Engelman A & Cherepanov P (2009a) A novel co-crystal structure affords the design of gain-of-function lentiviral integrase mutants in the presence of modified PSIP1/LEDGF/p75. *PLoS Pathog.* **5**: e1000259
- Hare S, Shun M-C, Gupta SS, Valkov E, Engelman A & Cherepanov P (2009b) A novel co-crystal structure affords the design of gain-of-function lentiviral integrase mutants in the presence of modified PSIP1/LEDGF/p75. *PLoS Pathog.* **5**: e1000259
- Hemelaar J (2012) The origin and diversity of the HIV-1 pandemic. *Trends Mol. Med.* **18**: 182–192
- Henderson A, Bunce M, Siddon N, Reeves R & Tremethick DJ (2000) High-mobility-group protein I can modulate binding of transcription factors to the U5 region of the human immunodeficiency virus type 1 proviral promoter. *J. Virol.* **74**: 10523–10534
- Henderson A, Holloway A, Reeves R & Tremethick DJ (2004) Recruitment of SWI/SNF to the human immunodeficiency virus type 1 promoter. *Mol. Cell. Biol.* **24**: 389–397
- Hill CP, Worthylake D, Bancroft DP, Christensen AM & Sundquist WI (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
- Hindmarsh P, Ridky T, Reeves R, Andrade M, Skalka AM & Leis J (1999) HMG protein family members stimulate human immunodeficiency virus type 1 and avian sarcoma virus concerted DNA integration in vitro. *J. Virol.* **73**: 2994–3003
- Hock R, Furusawa T, Ueda T & Bustin M (2007) HMG chromosomal proteins in development and disease. *Trends Cell Biol.* **17**: 72–79
- Hombrouck A, De Rijck J, Hendrix J, Vandekerckhove L, Voet A, De Maeyer M, Witvrouw M, Engelborghs Y, Christ F, Gijssbers R & Debyser Z (2007) Virus evolution reveals an exclusive role for LEDGF/p75 in chromosomal tethering of HIV. *PLoS Pathog.* **3**: e47
- Hütter G, Nowak D, Mossner M, Ganepola S, Müssig A, Allers K, Schneider T, Hofmann J, Kücherer C, Blau O, Blau IW, Hofmann WK & Thiel E (2009) Long-term control of HIV by CCR5 Delta32/Delta32 stem-cell transplantation. *N. Engl. J. Med.* **360**: 692–698
- Hu W-S & Hughes SH (2012) HIV-1 reverse transcription. *Cold Spring Harb. Perspect. Med.* **2**: Available at: <http://dx.doi.org/10.1101/cshperspect.a006882>
- Hu WS & Temin HM (1990) Retroviral recombination and reverse transcription. *Science* **250**: 1227–1233

Jacobo-Molina A, Ding J, Nanni RG, Clark AD Jr, Lu X, Tantillo C, Williams RL, Kamer G, Ferris AL & Clark P (1993) Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 Å resolution shows bent DNA. *Proc. Natl. Acad. Sci. U. S. A.* **90**: 6320–6324

Jacobs SA & Khorasanizadeh S (2002) Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail. *Science* **2**:5: 2080–2083

Jacquot G & Benichou S (2006) Import nucléaire du matériel génétique du virus de l'immunodéficience humaine de type 1. *Virologie* **10**: 43–53

Jeanmougin F, Wurtz JM, Le Douarin B, Chambon P & Losson R (1997) The bromodomain revisited. *Trends Biochem. Sci.* **22**: 151–153

Jenkins TM, Engelman A, Ghirlando R & Craigie R (1996) A soluble active mutant of HIV-1 integrase: involvement of both the core and carboxyl-terminal domains in multimerization. *J. Biol. Chem.* **271**: 7712–7718

Jin J, Cai Y, Li B, Conaway RC, Workman JL, Conaway JW & Kusch T (2005) In and out: histone variant exchange in chromatin. *Trends Biochem. Sci.* **30**: 680–687

Kalpana GV, Marmon S, Wang W, Crabtree GR & Goff SP (1994) Binding and stimulation of HIV-1 integrase by a human homolog of yeast transcription factor SNF5. *Science* **266**: 2002–2006

Katz RA, Gravuer K & Skalka AM (1998) A preferred target DNA structure for retroviral integrase in vitro. *J. Biol. Chem.* **273**: 24190–24195

Kim D, Blus BJ, Chandra V, Huang P, Rastinejad F & Khorasanizadeh S (2010) Corecognition of DNA and a methylated histone tail by the MSL3 chromodomain. *Nat. Struct. Mol. Biol.* **17**: 1027–1029

Kohlstaedt LA, Wang J, Friedman JM, Rice PA & Steitz TA (1992) Crystal structure at 3.5 Å resolution of HIV-1 reverse transcriptase complexed with an inhibitor. *Science* **256**: 1783–1790

Kotova S, Li M, Dimitriadis EK & Craigie R (2010) Nucleoprotein intermediates in HIV-1 DNA integration visualized by atomic force microscopy. *J. Mol. Biol.* **399**: 491–500

Kouzarides T (2007) Chromatin modifications and their function. *Cell* **128**: 693–705

Krishnan L & Engelman A (2012) Retroviral integrase proteins and HIV-1 DNA integration. *J. Biol. Chem.* **287**: 40858–40866

Krishnan L, Li X, Naraharisetty HL, Hare S, Cherepanov P & Engelman A (2010) Structure-based modeling of the functional HIV-1 intasome and its inhibition. *Proc. Natl. Acad. Sci. U. S. A.* **107**: 15910–15915

Kulkosky J, Jones KS, Katz RA & Mack JP (1992) Residues critical for retroviral integrative recombination in a region that is highly conserved among retroviral/retrotransposon integrases and bacterial insertion *Molecular and cellular* Available at: <http://mcb.asm.org/content/12/5/2331.short>

Kvaratskhelia M, Sharma A, Larue RC, Serrao E & Engelman A (2014) Molecular mechanisms of retroviral integration site selection. *Nucleic Acids Res.* **42**: 10209–10225

- Lachner M, O'Carroll D, Rea S, Mechtler K & Jenuwein T (2001) Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410**: 116–120
- Latham JA & Dent SYR (2007) Cross-regulation of histone modifications. *Nat. Struct. Mol. Biol.* **14**: 1017–1024
- Lee MS & Craigie R (1994) Protection of retroviral DNA from autointegration: involvement of a cellular factor. *Proc. Natl. Acad. Sci. U. S. A.* **91**: 9823–9827
- Leh H, Brodin P, Bischerour J, Deprez E, Tauc P, Brochon JC, LeCam E, Coulaud D, Auclair C & Mouscadet JF (2000) Determinants of Mg²⁺-dependent activities of recombinant human immunodeficiency virus type 1 integrase. *Biochemistry* **39**: 9285–9294
- Lelek M, Casartelli N, Pellin D, Rizzi E, Souque P, Severgnini M, Di Serio C, Fricke T, Diaz-Griffero F, Zimmer C, Charneau P & Di Nunzio F (2015) Chromatin organization at the nuclear pore favours HIV replication. *Nat. Commun.* **6**: 6483
- Lesbats P, Botbol Y, Chevereau G, Vaillant C, Calmels C, Arneodo A, Andreola M-L, Lavigne M & Parissi V (2011) Functional coupling between HIV-1 integrase and the SWI/SNF chromatin remodeling complex for efficient in vitro integration into stable nucleosomes. *PLoS Pathog.* **7**: e1001280
- Lesbats P, Métifiot M, Calmels C, Baranova S, Nevinsky G, Andreola ML & Parissi V (2008) In vitro initial attachment of HIV-1 integrase to viral ends: control of the DNA specific interaction by the oligomerization state. *Nucleic Acids Res.* **36**: 7043–7058
- Leschziner AE, Saha A, Wittmeyer J, Zhang Y, Bustamante C, Cairns BR & Nogales E (2007) Conformational flexibility in the chromatin remodeler RSC observed by electron microscopy and the orthogonal tilt reconstruction method. *Proc. Natl. Acad. Sci. U. S. A.* **104**: 4913–4918
- Lewinski MK, Bisgrove D, Shinn P, Chen H, Hoffmann C, Hannenhalli S, Verdin E, Berry CC, Ecker JR & Bushman FD (2005) Genome-Wide Analysis of Chromosomal Features Repressing Human Immunodeficiency Virus Transcription. *J. Virol.* **79**: 6610–6619
- Li H, Fischle W, Wang W, Duncan EM, Liang L, Murakami-Ishibe S, Allis CD & Patel DJ (2007) Structural basis for lower lysine methylation state-specific readout by MBT repeats of L3MBTL1 and an engineered PHD finger. *Mol. Cell* **28**: 677–691
- Li L, Li HS, Pauza CD, Bukrinsky M & Zhao RY (2005) Roles of HIV-1 auxiliary proteins in viral pathogenesis and host-pathogen interactions. *Cell Res.* **15**: 923–934
- Lin C-W & Engelman A (2003) The barrier-to-autointegration factor is a component of functional human immunodeficiency virus type 1 preintegration complexes. *J. Virol.* **77**: 5030–5036
- Lodi PJ, Ernst JA, Kuszewski J, Hickman AB, Engelman A, Craigie R, Clore GM & Gronenborn AM (1995) Solution structure of the DNA binding domain of HIV-1 integrase. *Biochemistry* **34**: 9826–9833

- Lopez AP, Kugelman JR, Garcia-Rivera J, Urias E, Salinas SA, Fernandez-Zapico ME & Llano M (2016) The Structure-Specific Recognition Protein 1 Associates with Lens Epithelium-Derived Growth Factor Proteins and Modulates HIV-1 Replication. *J. Mol. Biol.* Available at: <http://dx.doi.org/10.1016/j.jmb.2016.05.013>
- Loyola A & Almouzni G (2007) Marking histone H3 variants: how, when and why? *Trends Biochem. Sci.* **32**: 425–433
- Lozzio CB & Lozzio BB (1975) Human chronic myelogenous leukemia cell-line with positive Philadelphia chromosome. *Blood* **45**: 321–334
- Luger K & Hansen JC (2005) Nucleosome and chromatin fiber dynamics. *Curr. Opin. Struct. Biol.* **15**: 188–196
- Luger K, Mäder AW, Richmond RK, Sargent DF & Richmond TJ (1997) Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**: 251–260
- Lutzke RA & Plasterk RH (1998) Structure-based mutational analysis of the C-terminal DNA-binding domain of human immunodeficiency virus type 1 integrase: critical residues for protein oligomerization and DNA binding. *J. Virol.* **72**: 4841–4848
- Lv Z, Chu Y & Wang Y (2015) HIV protease inhibitors: a review of molecular selectivity and toxicity. *HIV AIDS* **7**: 95–104
- Maartens G, Celum C & Lewin SR (2014) HIV infection: epidemiology, pathogenesis, treatment, and prevention. *Lancet* **384**: 258–271
- Maddon PJ, Dalgleish AG, McDougal JS, Clapham PR, Weiss RA & Axel R (1986) The T4 gene encodes the AIDS virus receptor and is expressed in the immune system and the brain. *Cell* **47**: 333–348
- Maertens G, Cherepanov P, Debyser Z, Engelborghs Y & Engelman A (2004) Identification and Characterization of a Functional Nuclear Localization Signal in the HIV-1 Integrase Interactor LEDGF/p75. *J. Biol. Chem.* **279**: 33421–33429
- Maertens G, Cherepanov P, Pluymers W, Busschots K, De Clercq E, Debyser Z & Engelborghs Y (2003) LEDGF/p75 is essential for nuclear and chromosomal targeting of HIV-1 integrase in human cells. *J. Biol. Chem.* **278**: 33528–33539
- Maertens GN, Hare S & Cherepanov P (2010) The mechanism of retroviral integration from X-ray structures of its key intermediates. *Nature* **468**: 326–329
- Mangeat B, Turelli P, Caron G, Friedli M, Perrin L & Trono D (2003) Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. *Nature* **424**: 99–103
- Margottin F, Bour SP, Durand H, Selig L, Benichou S, Richard V, Thomas D, Strebel K & Benarous R (1998/3) A Novel Human WD Protein, h-βTrCP, that Interacts with HIV-1 Vpu Connects CD4 to the ER Degradation Pathway through an F-Box Motif. *Mol. Cell* **1**: 565–574
- Maroun M, Delelis O, Coadou G, Bader T, Segéral E, Mbemba G, Petit C, Sonigo P, Rain J-C, Mouscadet J-F, Benarous R & Emiliani S (2006) Inhibition of Early Steps of HIV-1 Replication by SNF5/Ini1. *J. Biol. Chem.* **281**: 22736–22743

Marshall HM, Ronen K, Berry C, Llano M, Sutherland H, Saenz D, Bickmore W, Poeschla E & Bushman FD (2007) Role of PSIP1/LEDGF/p75 in lentiviral infectivity and integration targeting. *PLoS One* **2**: e1340

Martin C & Zhang Y (2005) The diverse functions of histone lysine methylation. *Nat. Rev. Mol. Cell Biol.* **6**: 838–849

Maskell DP, Renault L, Serrao E, Lesbats P, Matadeen R, Hare S, Lindemann D, Engelman AN, Costa A & Cherepanov P (2015) Structural basis for retroviral integration into nucleosomes. *Nature* **523**: 366–369

Massiah MA, Starich MR, Paschall C, Summers MF, Christensen AM & Sundquist WI (1994) Three-dimensional structure of the human immunodeficiency virus type 1 matrix protein. *J. Mol. Biol.* **244**: 198–223

Massiah MA, Worthylake D, Christensen AM, Sundquist WI, Hill CP & Summers MF (1996) Comparison of the NMR and X-ray structures of the HIV-1 matrix protein: evidence for conformational changes during viral assembly. *Protein Sci.* **5**: 2391–2398

Maurer-Stroh S, Dickens NJ, Hughes-Davies L, Kouzarides T, Eisenhaber F & Ponting CP (2003) The Tudor domain 'Royal Family': Tudor, plant Agenet, Chromo, PWWP and MBT domains. *Trends Biochem. Sci.* **28**: 69–74

Mellor J & Morillon A (2004) ISWI complexes in *Saccharomyces cerevisiae*. *Biochim. Biophys. Acta* **1677**: 100–112

Michel F, Crucifix C, Granger F, Eiler S, Mouscadet J-F, Korolev S, Agapkina J, Ziganshin R, Gottikh M, Nazabal A, Emiliani S, Benarous R, Moras D, Schultz P & Ruff M (2009) Structural basis for HIV-1 DNA integration in the human genome, role of the LEDGF/P75 cofactor. *EMBO J.* **28**: 980–991

Min J, Allali-Hassani A, Nady N, Qi C, Ouyang H, Liu Y, MacKenzie F, Vedadi M & Arrowsmith CH (2007) L3MBTL1 recognition of mono- and dimethylated histones. *Nat. Struct. Mol. Biol.* **14**: 1229–1230

Mirambeau G, Lonnais S & Gorelick RJ (2010) Features, processing states, and heterologous protein interactions in the modulation of the retroviral nucleocapsid protein function. *RNA Biol.* **7**: 724–734

Mitchell RS, Beitzel BF, Schroder ARW, Shinn P, Chen H, Berry CC, Ecker JR & Bushman FD (2004) Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. *PLoS Biol.* **2**: E234

Mohammadi P, Desfarges S, Bartha I, Joos B, Zangger N, Muñoz M, Günthard HF, Beerenwinkel N, Telenti A & Ciuffi A (2013) 24 hours in the life of HIV-1 in a T cell line. *PLoS Pathog.* **9**: e1003161

Mohrmann L & Verrijzer CP (2005) Composition and functional specificity of SWI2/SNF2 class chromatin remodeling complexes. *Biochim. Biophys. Acta* **1681**: 59–73

Molina J-M, Capitant C, Spire B, Pialoux G, Cotte L, Charreau I, Tremblay C, Le Gall J-M, Cua E, Pasquet A, Raffi F, Pintado C, Chidiac C, Chas J, Charbonneau P, Delaugerre C, Suzan-Monti M, Loze B, Fonsart J, Peytavin G, et al (2015) On-Demand Preexposure Prophylaxis in Men at High Risk for HIV-1 Infection. *N. Engl. J. Med.* **373**: 2237–2246

- Montes de Oca R, Lee KK & Wilson KL (2005) Binding of barrier to autointegration factor (BAF) to histone H3 and selected linker histones including H1.1. *J. Biol. Chem.* **280**: 42252–42262
- Moore SA, Ferhatoglu Y, Jia Y, Al-Jiab RA & Scott MJ (2010) Structural and biochemical studies on the chromo-barrel domain of male specific lethal 3 (MSL3) reveal a binding preference for mono- or dimethyllysine 20 on histone H4. *J. Biol. Chem.* **285**: 40879–40890
- Morozov A, Yung E & Kalpana GV (1998) Structure-function analysis of integrase interactor 1/hSNF5L1 reveals differential properties of two repeat motifs present in the highly conserved region. *Proc. Natl. Acad. Sci. U. S. A.* **95**: 1120–1125
- Mueller-Planitz F, Klinker H & Becker PB (2013) Nucleosome sliding mechanisms: new twists in a looped history. *Nat. Struct. Mol. Biol.* **20**: 1026–1032
- Mujtaba S, Zeng L & Zhou M-M (2007) Structure and acetyl-lysine recognition of the bromodomain. *Oncogene* **26**: 5521–5527
- Müller HP & Varmus HE (1994) DNA bending creates favored sites for retroviral integration: an explanation for preferred insertion sites in nucleosomes. *EMBO J.* **13**: 4704–4714
- Musinova YR, Sheval EV, Dib C, Germini D & Vassetzky YS (2016) Functional roles of HIV-1 Tat protein in the nucleus. *Cell. Mol. Life Sci.* **73**: 589–601
- Musselman CA, Avvakumov N, Watanabe R, Abraham CG, Lalonde M-E, Hong Z, Allen C, Roy S, Nuñez JK, Nickoloff J, Kulesza CA, Yasui A, Côté J & Kutateladze TG (2012a) Molecular basis for H3K36me3 recognition by the Tudor domain of PHF1. *Nat. Struct. Mol. Biol.* **19**: 1266–1272
- Musselman CA, Lalonde M-E, Côté J & Kutateladze TG (2012b) Perceiving the epigenetic landscape through histone readers. *Nat. Struct. Mol. Biol.* **19**: 1218–1227
- Naughtin M, Haftek-Terreau Z, Xavier J, Meyer S, Silvain M, Jaszczyszyn Y, Levy N, Miele V, Benleulmi MS, Ruff M, Parissi V, Vaillant C & Lavigne M (2015) DNA Physical Properties and Nucleosome Positions Are Major Determinants of HIV-1 Integrase Selectivity. *PLoS One* **10**: e0129427
- Navia MA, Fitzgerald PM, McKeever BM, Leu CT, Heimbach JC, Herber WK, Sigal IS, Darke PL & Springer JP (1989) Three-dimensional structure of aspartyl protease from human immunodeficiency virus HIV-1. *Nature* **337**: 615–620
- Neamati N & Wang B (2011) HIV-1 Integrase: Mechanism and Inhibitor Design Wiley
- Nielsen PR, Nietlispach D, Mott HR, Callaghan J, Bannister A, Kouzarides T, Murzin AG, Murzina NV & Laue ED (2002) Structure of the HP1 chromodomain bound to histone H3 methylated at lysine 9. *Nature* **416**: 103–107
- Ocwieja KE, Brady TL, Ronen K, Huegel A, Roth SL, Schaller T, James LC, Towers GJ, Young JAT, Chanda SK, König R, Malani N, Berry CC & Bushman FD (2011) HIV integration targeting: a pathway involving Transportin-3 and the nuclear pore protein RanBP2. *PLoS Pathog.* **7**: e1001313

Parissi V, Calmels C, De Soultrait VR, Caumont A, Fournier M, Chaignepain S & Litvak S (2001) Functional interactions of human immunodeficiency virus type 1 integrase with human and yeast HSP60. *J. Virol.* **75**: 11344–11353

Peterlin BM & Trono D (2003) Hide, shield and strike back: how HIV-infected cells avoid immune eradication. *Nat. Rev. Immunol.* **3**: 97–107

Peterson CL & Laniel M-A (2004) Histones and histone modifications. *Curr. Biol.* **14**: R546–51

Plantier J-C, Leoz M, Dickerson JE, De Oliveira F, Cordonnier F, Lemée V, Damond F, Robertson DL & Simon F (2009) A new human immunodeficiency virus derived from gorillas. *Nat. Med.* **15**: 871–872

Pluta K & Kacprzak MM (2009) Use of HIV as a gene transfer vector. *Acta Biochim. Pol.* **56**: 531–595

Podtelezhnikov AA, Gao K, Bushman FD & McCammon JA (2003) Modeling HIV-1 integrase complexes based on their hydrodynamic properties. *Biopolymers* **68**: 110–120

Pommier Y, Johnson AA & Marchand C (2005) Integrase inhibitors to treat HIV/AIDS. *Nat. Rev. Drug Discov.* **4**: 236–248

Popovic M, Sarngadharan MG, Read E & Gallo RC (1984) Detection, isolation, and continuous production of cytopathic retroviruses (HTLV-III) from patients with AIDS and pre-AIDS. *Science* **224**: 497–500

Pornillos O, Ganser-Pornillos BK & Yeager M (2011) Atomic-level modelling of the HIV capsid. *Nature* **469**: 424–427

Pruss D, Bushman FD & Wolffe AP (1994) Human immunodeficiency virus integrase directs integration to sites of severe DNA distortion within the nucleosome core. *Proc. Natl. Acad. Sci. U. S. A.* **91**: 5913–5917

Pryciak PM, Sil A & Varmus HE (1992) Retroviral integration into minichromosomes in vitro. *EMBO J.* **11**: 291–303

Pryciak PM & Varmus HE (1992) Nucleosomes, DNA-binding proteins, and DNA sequence modulate retroviral integration target site selection. *Cell* **69**: 769–780

Rain J-C, Cribier A, Gérard A, Emiliani S & Benarous R (2009/4) Yeast two-hybrid detection of integrase–host factor interactions. *Methods* **47**: 291–297

Ranjan A, Wang F, Mizuguchi G, Wei D, Huang Y & Wu C (2015) H2A histone-fold and DNA elements in nucleosome activate SWR1-mediated H2A.Z replacement in budding yeast. Available at: <http://elifesciences.org/lookup/doi/10.7554/eLife.06845>

Reeves R (2001) Molecular biology of HMGA proteins: hubs of nuclear function. *Gene* **277**: 63–81

Rerks-Ngarm S, Pitisuttithum P, Nitayaphan S, Kaewkungwal J, Chiu J, Paris R, Prem Sri N, Namwat C, de Souza M, Adams E, Benenson M, Gurunathan S, Tartaglia J, McNeil JG, Francis DP, Stablein D, Birx DL, Chunsuttiwat S, Khamboonruang C, Thongcharoen P, et al (2009) Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand. *N. Engl. J. Med.* **361**: 2209–2220

- Romani B, Engelbrecht S & Glashoff RH (2010) Functions of Tat: the versatile protein of human immunodeficiency virus type 1. *J. Gen. Virol.* **91**: 1–12
- Saha A, Wittmeyer J & Cairns BR (2006) Mechanisms for nucleosome movement by ATP-dependent chromatin remodeling complexes. *Results Probl. Cell Differ.* **41**: 127–148
- Sarafianos SG, Marchand B, Das K, Himmel DM, Parniak MA, Hughes SH & Arnold E (2009) Structure and function of HIV-1 reverse transcriptase: molecular mechanisms of polymerization and inhibition. *J. Mol. Biol.* **385**: 693–713
- Schalch T, Duda S, Sargent DF & Richmond TJ (2005) X-ray structure of a tetranucleosome and its implications for the chromatin fibre. *Nature* **436**: 138–141
- Schaller T, Ocwieja KE, Rasaiyaah J, Price AJ, Brady TL, Roth SL, Hué S, Fletcher AJ, Lee K, KewalRamani VN, Noursadeghi M, Jenner RG, James LC, Bushman FD & Towers GJ (2011) HIV-1 capsid-cyclophilin interactions determine nuclear import pathway, integration targeting and replication efficiency. *PLoS Pathog.* **7**: e1002439
- Schrijvers R, Vets S, De Rijck J, Malani N, Bushman FD, Debyser Z & Gijsbers R (2012) HRP-2 determines HIV-1 integration site selection in LEDGF/p75 depleted cells. *Retrovirology* **9**: 84
- Schröder ARW, Shinn P, Chen H, Berry C, Ecker JR & Bushman F (2002) HIV-1 integration in the human genome favors active genes and local hotspots. *Cell* **110**: 521–529
- Sekaly R-P (2008) The failed HIV Merck vaccine study: a step back or a launching point for future vaccine development? *J. Exp. Med.* **205**: 7–12
- Serrao E, Ballandras-Colas A, Cherepanov P, Maertens GN & Engelman AN (2015) Key determinants of target DNA recognition by retroviral intasomes. *Retrovirology* **12**: 39
- Sherman PA & Fyfe JA (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
- Shilatifard A (2006) Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. *Annu. Rev. Biochem.* **75**: 243–269
- Shun M-C, Raghavendra NK, Vandegraaff N, Daigle JE, Hughes S, Kellam P, Cherepanov P & Engelman A (2007) LEDGF/p75 functions downstream from preintegration complex formation to effect gene-specific HIV-1 integration. *Genes Dev.* **21**: 1767–1778
- Silvers RM, Smith JA, Schowalter M, Litwin S, Liang Z, Geary K & Daniel R (2010) Modification of integration site preferences of an HIV-1-based vector by expression of a novel synthetic protein. *Hum. Gene Ther.* **21**: 337–349
- Skala AM (2014) Retroviral DNA Transposition: Themes and Variations. *Microbiol Spectr* **2**: Available at: <http://dx.doi.org/10.1128/microbiolspec.MDNA3-0005-2014>
- Sloan RD & Wainberg MA (2011) The role of unintegrated DNA in HIV infection. *Retrovirology* **8**: 52

- Smith CL, Horowitz-Scherer R, Flanagan JF, Woodcock CL & Peterson CL (2003) Structural analysis of the yeast SWI/SNF chromatin remodeling complex. *Nat. Struct. Biol.* **10**: 141–145
- Speranzini V, Pilotto S, Sixma TK & Mattevi A (2016) Touch, act and go: landing and operating on nucleosomes. *EMBO J.* **35**: 376–388
- Stec I, Nagl SB, van Ommen GJ & den Dunnen JT (2000) The PWWP domain: a potential protein-protein interaction domain in nuclear proteins influencing differentiation? *FEBS Lett.* **473**: 1–5
- Stucki M, Clapperton JA, Mohammad D, Yaffe MB, Smerdon SJ & Jackson SP (2005) MDC1 directly binds phosphorylated histone H2AX to regulate cellular responses to DNA double-strand breaks. *Cell* **123**: 1213–1226
- Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou KP, Kuhn M, Bork P, Jensen LJ & von Mering C (2015) STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* **43**: D447–52
- Taganov KD, Cuesta I, Daniel R, Cirillo LA, Katz RA, Zaret KS & Skalka AM (2004) Integrase-specific enhancement and suppression of retroviral DNA integration by compacted chromatin structure in vitro. *J. Virol.* **78**: 5848–5855
- Taverna SD, Li H, Ruthenburg AJ, Allis CD & Patel DJ (2007) How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat. Struct. Mol. Biol.* **14**: 1025–1040
- Thierry S, Benleulmi MS, Sinzelle L, Thierry E, Calmels C, Chaignepain S, Waffo-Teguo P, Merillon J-M, Budke B, Pasquet J-M, Litvak S, Ciuffi A, Sung P, Connell P, Hauber I, Hauber J, Andreola M-L, Delelis O & Parissi V (2015) Dual and Opposite Effects of hRAD51 Chemical Modulation on HIV-1 Integration. *Chem. Biol.* **22**: 712–723
- Thomas JA & Gorelick RJ (2008/6) Nucleocapsid protein function in early infection processes. *Virus Res.* **134**: 39–63
- Tressler R & Godfrey C (2012) NRTI Backbone in HIV Treatment. *Drugs* **72**: 2051–2062
- Turelli P, Doucas V, Craig E, Mangeat B, Klages N, Evans R, Kalpana G & Trono D (2001) Cytoplasmic recruitment of INI1 and PML on incoming HIV preintegration complexes: interference with early steps of viral replication. *Mol. Cell* **7**: 1245–1254
- Vandekerckhove L, Christ F, Van Maele B, De Rijck J, Gijsbers R, Van den Haute C, Witvrouw M & Debyser Z (2006) Transient and stable knockdown of the integrase cofactor LEDGF/p75 reveals its role in the replication cycle of human immunodeficiency virus. *J. Virol.* **80**: 1886–1896
- Venkatesh S & Workman JL (2015) Histone exchange, chromatin structure and the regulation of transcription. *Nat. Rev. Mol. Cell Biol.* **16**: 178–189
- Wang GP, Ciuffi A, Leipzig J, Berry CC & Bushman FD (2007) HIV integration site selection: analysis by massively parallel pyrosequencing reveals association with epigenetic modifications. *Genome Res.* **17**: 1186–1194

- Wang H, Jurado KA, Wu X, Shun M-C, Li X, Ferris AL, Smith SJ, Patel PA, Fuchs JR, Cherepanov P, Kvaratskhelia M, Hughes SH & Engelman A (2012) HRP2 determines the efficiency and specificity of HIV-1 integration in LEDGF/p75 knockout cells but does not contribute to the antiviral activity of a potent LEDGF/p75-binding site integrase inhibitor. *Nucleic Acids Res.* **40**: 11518–11530
- Wang JY, Ling H, Yang W & Craigie R (2001) Structure of a two-domain fragment of HIV-1 integrase: implications for domain organization in the intact protein. *EMBO J.* **20**: 7333–7343
- Wang W, Côté J, Xue Y, Zhou S, Khavari PA, Biggar SR, Muchardt C, Kalpana GV, Goff SP, Yaniv M, Workman JL & Crabtree GR (1996) Purification and biochemical heterogeneity of the mammalian SWI-SNF complex. *EMBO J.* **15**: 5370–5382
- Wang WK, Tereshko V, Boccuni P, MacGrogan D, Nimer SD & Patel DJ (2003) Malignant brain tumor repeats: a three-leaved propeller architecture with ligand/peptide binding pockets. *Structure* **11**: 775–789
- Whitlock G, Nwokolo N & McOwan A (2015) Is PEP prescribed appropriately? *HIV Med.* **16**: 519–520
- Wielens J, Crosby IT & Chalmers DK (2005) A three-dimensional model of the human immunodeficiency virus type 1 integration complex. *J. Comput. Aided Mol. Des.* **19**: 301–317
- Wilen CB, Tilton JC & Doms RW (2012) HIV: cell binding and entry. *Cold Spring Harb. Perspect. Med.* **2**: Available at: <http://dx.doi.org/10.1101/cshperspect.a006866>
- Woodcock CL (2006) Chromatin architecture. *Curr. Opin. Struct. Biol.* **16**: 213–220
- Wu C, Bassett A & Travers A (2007) A variable topology for the 30-nm chromatin fibre. *EMBO Rep.* **8**: 1129–1134
- Yin Z, Shi K, Banerjee S, Pandey KK, Bera S, Grandgenett DP & Aihara H (2016) Crystal structure of the Rous sarcoma virus intasome. *Nature* **530**: 362–366
- Yukl SA, Boritz E, Busch M, Bentsen C, Chun T-W, Douek D, Eisele E, Haase A, Ho Y-C, Hütter G, Justement JS, Keating S, Lee T-H, Li P, Murray D, Palmer S, Pilcher C, Pillai S, Price RW, Rothenberger M, et al (2013) Challenges in detecting HIV persistence during potentially curative interventions: a study of the Berlin patient. *PLoS Pathog.* **9**: e1003347
- Yung E, Sorin M, Pal A, Craig E, Morozov A, Delattre O, Kappes J, Ott D & Kalpana GV (2001) Inhibition of HIV-1 virion production by a transdominant mutant of integrase interactor 1. *Nat. Med.* **7**: 920–926
- Zhang D, Li W & Jiang S (2015) Peptide fusion inhibitors targeting the HIV-1 gp41: a patent review (2009 - 2014). *Expert Opin. Ther. Pat.* **25**: 159–173
- Zhao Z, McKee CJ, Kessl JJ, Santos WL, Daigle JE, Engelman A, Verdine G & Kvaratskhelia M (2008) Subunit-specific protein footprinting reveals significant structural rearrangements and a role for N-terminal Lys-14 of HIV-1 Integrase during viral DNA binding. *J. Biol. Chem.* **283**: 5632–5641

Intégration du VIH-1 : Contrôle et régulation de l'interaction fonctionnelle entre l'intégrase et la chromatine

Résumé

L'intégrase (IN) du VIH-1 est une enzyme clé du cycle viral catalysant l'insertion stable de l'ADN viral dans le génome de la cellule infectée. L'IN participe également à de nombreuses étapes du cycle viral telles que la transcription inverse ou la maturation virale. Ainsi, la compréhension des mécanismes régulant l'intégration du VIH-1 représente un enjeu majeur dans le cadre notamment d'approches thérapeutiques. En effet, les études montrent que ces mécanismes sont finement régulés dans la cellule par des facteurs viraux et cellulaires agissant à différentes étapes du cycle viral. C'est donc dans ce contexte que nous avons étudié les facteurs à la fois viraux et cellulaires régulant ce processus. Dans un premier temps, les déterminants viraux modulant l'intégration dans la chromatine ont été analysés dans le cas de plusieurs modèles rétroviraux. Puis, dans un second temps, nous avons étudié l'impact de facteurs cellulaires, identifiés au laboratoire, sur les mécanismes d'insertion de l'ADN viral dans le génome cellulaire.

Mon travail de thèse s'est ainsi articulé en trois axes majeurs aboutissant à :

- La démonstration de la régulation de l'intégration rétrovirale par la structure chromatinienne de l'hôte
- L'identification de nouveaux cofacteurs cellulaires participant à la régulation de l'intégration dans la chromatine dont le complexe de remodelage FACT « Facilitates Chromatin Transcription »
- L'identification d'une nouvelle interaction fonctionnelle entre l'IN du VIH-1 et la queue amino-terminale de l'histone humaine H4 et de son rôle dans la sélectivité de l'intégration